

# Topology

---

*CAUTION: This book is still under construction.*

- some sections and subsections are unfinished or have gaps;
- many figures are currently shown as empty boxes;
- some figures show old, hand-drawn images that will be replaced by new, computer-generated images; and
- typos and other errors are likely.



# TOPOLOGY

---

*Second edition*

**Murray Eisenberg** 

*University of Massachusetts Amherst*

Murray Eisenberg  
Rockville, MD USA  
<http://people.math.umass.edu/~murray/>  
<mailto://murray@math.umass.edu>  
[orcid.org/0000-0003-0989-4066](https://orcid.org/0000-0003-0989-4066)

This book was composed using the  $\text{\LaTeX} 2_{\epsilon}$  typesetting system, based upon the  $\text{\LaTeX}$  system created by Leslie Lamport and the  $\text{\TeX}$  system created by Donald Knuth, and the memoir document class. It is typeset with newtx text and math fonts created by Michael Sharpe.

First edition 1974, published by Holt, Rinehart and Winston, Inc.  
Second edition 2022, published by the author.  
This is version 1.9.1 created August 19, 2023.

Copyright © 1974, 2023 Murray Eisenberg  
All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the author.

***To the memory of my father***



# Detailed Contents

<b>Detailed Contents</b>	<b>vii</b>
<b>Short Contents</b>	<b>xiii</b>
<b>List of Figures</b>	<b>xv</b>
<b>List of Tables</b>	<b>xix</b>
<b>Preface</b>	<b>xxi</b>
Aims . . . . .	xxiii
Prerequisites . . . . .	xxiii
Contents . . . . .	xxiii
Exercises . . . . .	xxiii
Changes from the first edition . . . . .	xxiii
Numbering scheme . . . . .	xxiv
References . . . . .	xxiv
Russian names . . . . .	xxiv
Suggested course outlines . . . . .	xxiv
Acknowledgements . . . . .	xxvi
<b>0 Sets and Maps</b>	<b>1</b>
<b>0.1 Logic</b> . . . . .	<b>2</b>
Predicate logic . . . . .	4
Exercises for Section 0.1 . . . . .	7
<b>0.2 Sets</b> . . . . .	<b>7</b>
Building sets . . . . .	8
Special sets . . . . .	9
Subsets . . . . .	10
Mathematical induction . . . . .	11
Power sets . . . . .	13
Union and intersection of two sets . . . . .	13
Complements . . . . .	14
Pairs and products . . . . .	14
Relations . . . . .	15
Exercises for Section 0.2 . . . . .	17
<b>0.3 Maps</b> . . . . .	<b>19</b>
Maps . . . . .	19
Operations . . . . .	22
Restriction and extension . . . . .	23
Composition of maps . . . . .	25

Images and inverse images . . . . .	27
Injections, surjections, and bijections . . . . .	28
Inverse of a bijection . . . . .	30
Exercises for Section 0.3 . . . . .	31
<b>0.4 Families and Products . . . . .</b>	<b>34</b>
$n$ -tuples and sequences . . . . .	35
$n$ th power of a set . . . . .	35
Euclidean spaces . . . . .	36
Line segments and lines . . . . .	37
Arbitrary power of a set . . . . .	37
Union and intersection of a family of sets . . . . .	38
Union and intersection of a collection of sets . . . . .	41
Product of a family of sets . . . . .	42
Axiom of Choice . . . . .	44
Exercises for Section 0.4 . . . . .	45
<b>0.5 Recursion . . . . .</b>	<b>47</b>
Ordinary recursion . . . . .	47
Primitive and complete recursion . . . . .	50
Exercises for Section 0.5 . . . . .	52
<b>0.6 Countability . . . . .</b>	<b>53</b>
Finite and infinite sets . . . . .	53
Denumerable, countable, and uncountable sets . . . . .	57
Exercises for Section 0.6 . . . . .	61
<b>0.7 Ordering relations . . . . .</b>	<b>63</b>
Preorderings . . . . .	63
Partial orderings . . . . .	65
Intervals . . . . .	70
Bounds . . . . .	71
Total orderings . . . . .	72
Exercises for Section 0.7 . . . . .	75
<b>0.8 Order-Completeness of the Real Numbers . . . . .</b>	<b>78</b>
Order-completeness . . . . .	79
Base expansion . . . . .	79
Archimedean Ordering and Nested Interval Properties . . . . .	81
Properties of intervals in the real line . . . . .	85
Natural numbers, integers, and rationals within the reals . . . . .	87
Complex numbers . . . . .	89
Exercises for Section 0.8 . . . . .	91
<b>0.9 Equivalence Relations . . . . .</b>	<b>93</b>
Equivalence relations . . . . .	93
Quotient sets . . . . .	96
Partitions . . . . .	98
Saturated sets . . . . .	99
Passing to the quotient . . . . .	100
Exercises for Section 0.9 . . . . .	103
<b>0.10 Ordinals, Cardinality, and Zorn's Lemma . . . . .</b>	<b>109</b>
Well-ordered sets . . . . .	109
The first uncountable ordinal . . . . .	110
Maximal and minimal elements . . . . .	112
Zorn's Lemma . . . . .	113
The Well-ordering Theorem . . . . .	116
Cardinality . . . . .	118



Exercises for Section 0.10 . . . . .	123
<b>1 Metric Spaces</b> . . . . .	<b>127</b>
<b>1.1 Metrics</b> . . . . .	128
The Euclidean metric . . . . .	128
Other metrics on $\mathbb{R}^n$ . . . . .	132
Metrics on function spaces . . . . .	135
Metrics and metric spaces . . . . .	139
Exercises for Section 1.1 . . . . .	142
<b>1.2 Open Sets and Closed Sets</b> . . . . .	146
Balls, disks, and spheres . . . . .	147
Open sets in a metric space . . . . .	152
Exercises for Section 1.2 . . . . .	161
<b>1.3 Equivalent Metrics</b> . . . . .	164
Equivalent metrics on a set . . . . .	164
Bounded metric spaces . . . . .	167
The extended real line . . . . .	168
Isometries . . . . .	172
Exercises for Section 1.3 . . . . .	176
<b>1.4 Continuity and Convergence</b> . . . . .	179
Continuous maps between metric spaces . . . . .	180
Convergent sequences in metric spaces . . . . .	185
Relationship between convergence and continuity . . . . .	194
Exercises for Section 1.4 . . . . .	195
<b>1.5 Completeness</b> . . . . .	200
Cauchy sequences . . . . .	200
Complete metric spaces . . . . .	202
Completeness of sequence and functions spaces . . . . .	206
The Nested Set and Baire Category Theorems . . . . .	212
Exercises for Section 1.5 . . . . .	217
<b>2 Topological Spaces</b> . . . . .	<b>225</b>
<b>2.1 Topologies</b> . . . . .	225
Open sets . . . . .	226
Closed sets . . . . .	229
Subspaces . . . . .	230
Exercises for Section 2.1 . . . . .	233
<b>2.2 Neighborhoods</b> . . . . .	235
Properties of neighborhoods . . . . .	237
Defining a topology by its neighborhoods . . . . .	238
Isolated points and limit points . . . . .	240
Hausdorff spaces . . . . .	241
Exercises for Section 2.2 . . . . .	243
<b>2.3 Boundary, Interior, and Closure</b> . . . . .	246
Boundary of a set . . . . .	246
Interior of a set . . . . .	250
Closure of a set . . . . .	252
Nowhere dense sets . . . . .	256
Boundary, interior, and closure in a subspace . . . . .	258
Exercises for Section 2.3 . . . . .	259
<b>2.4 Bases and Local Bases</b> . . . . .	263
Local bases . . . . .	263

Zero-dimensional spaces . . . . .	266
First-countable spaces . . . . .	268
Bases . . . . .	269
Second-countable spaces . . . . .	277
Separable spaces . . . . .	280
Subbases . . . . .	282
Exercises for Section 2.4 . . . . .	285
<b>2.5 Separation Properties . . . . .</b>	<b>292</b>
The separation properties hierarchy . . . . .	295
Separation properties of new spaces constructed from old . . . . .	297
Exercises for Section 2.5 . . . . .	298
<b>3 Continuity and Convergence . . . . .</b>	<b>301</b>
<b>3.1 Continuous Maps . . . . .</b>	<b>302</b>
Continuous maps of topological spaces . . . . .	302
Extending and gluing maps . . . . .	307
Open and closed maps . . . . .	312
Continuous images of spaces . . . . .	314
Exercises for Section 3.1 . . . . .	316
<b>3.2 Homeomorphisms . . . . .</b>	<b>319</b>
Homeomorphisms—definition and examples . . . . .	320
Topological properties . . . . .	333
Local homeomorphism . . . . .	335
Locally Euclidean spaces . . . . .	336
Topological manifolds . . . . .	343
Embedding . . . . .	346
Preservation of topological properties . . . . .	349
Exercises for Section 3.2 . . . . .	350
<b>3.3 Product Spaces . . . . .</b>	<b>359</b>
Topology of the product of countably many metric spaces . . . . .	360
The product topology . . . . .	363
Products and maps . . . . .	367
Further properties of product spaces . . . . .	377
Exercises for Section 3.3 . . . . .	379
<b>3.4 Quotient Spaces . . . . .</b>	<b>385</b>
The unit circle as a quotient space . . . . .	385
The quotient topology . . . . .	388
Quotient spaces and continuous maps . . . . .	390
Quotient maps . . . . .	392
Concrete realizations of quotient spaces . . . . .	393
Countability and separation properties of quotient spaces . . . . .	408
Exercises for Section 3.4 . . . . .	410
<b>3.5 Convergence . . . . .</b>	<b>420</b>
Sequential convergence in topological spaces . . . . .	420
Directed sets and nets . . . . .	423
Convergence of nets . . . . .	425
Subnets . . . . .	435
Limit of a map . . . . .	437
Filters . . . . .	440
Convergence of filters . . . . .	443
Comparison of filters . . . . .	448
Filters vs. nets . . . . .	449

Exercises for Section 3.5 . . . . .	451
<b>4 Compactness</b> . . . . .	<b>457</b>
<b>4.1 Compact Spaces</b> . . . . .	458
Compactness of the closed unit interval . . . . .	458
Open covers and compact spaces . . . . .	459
Closed sets and compact sets . . . . .	462
The Cantor set . . . . .	466
Separation properties in compact spaces . . . . .	473
Continuous maps on compact spaces . . . . .	475
Products of compact spaces . . . . .	477
Separation in quotients of compact spaces . . . . .	480
Compactness and convergence . . . . .	482
Exercises for Section 4.1 . . . . .	484
<b>4.2 Compact Metric Spaces</b> . . . . .	493
Countability properties in compact metrizable spaces . . . . .	493
Totally bounded metric spaces . . . . .	494
Convergence in compact metrizable spaces . . . . .	495
More about totally bounded metric spaces . . . . .	498
More metric consequences of compactness . . . . .	501
Exercises for Section 4.2 . . . . .	504
<b>4.3 Locally Compact Spaces</b> . . . . .	508
Locally compact spaces . . . . .	508
Local compactness and continuity . . . . .	511
One-point compactification . . . . .	513
Local compactness and countability . . . . .	517
Exercises for Section 4.3 . . . . .	518
<b>5 Connectedness</b> . . . . .	<b>523</b>
<b>5.1 Connected Spaces</b> . . . . .	524
Connected spaces . . . . .	524
Connected subsets of topological spaces . . . . .	526
Continuous maps on connected spaces . . . . .	527
Covering by connected sets . . . . .	532
Products of connected spaces . . . . .	536
Exercises for Section 5.1 . . . . .	541
<b>5.2 Components and Locally Connected Spaces</b> . . . . .	545
Components . . . . .	545
Totally disconnected spaces . . . . .	549
Locally connected spaces . . . . .	552
Exercises for Section 5.2 . . . . .	555
<b>5.3 Path-Connected Spaces</b> . . . . .	558
Paths . . . . .	558
Path-connectedness . . . . .	561
Curves . . . . .	564
Exercises for Section 5.3 . . . . .	570
<b>5.4 Homotopy</b> . . . . .	571
Contractible spaces . . . . .	572
Retracts and deformation retracts . . . . .	574
Homotopies . . . . .	576
Path-homotopies . . . . .	577
Exercises for Section 5.4 . . . . .	585

<b>5.5 Simply connected spaces</b>	587
Simple connectedness	587
The circle and simple connectedness	593
Exercises for Section 5.5	602
<b>6 Embedding</b>	<b>605</b>
<b>6.1 The Tychonoff Product Theorem</b>	605
Proof of the Tychonoff Theorem using subbases	606
Proof of the Tychonoff Theorem using ultrafilters	607
Exercises for Section 6.1	610
<b>6.2 Complete regularity and normality</b>	612
Completely regular spaces	612
Normal spaces	616
Urysohn's Lemma	618
The Tietze Extension Theorem	623
Exercises for Section 6.2	627
<b>6.3 Embedding theorems</b>	631
Embedding in Tychonoff cubes	631
Exercises for Section 6.3	637
<b>6.4 Urysohn Metrization Theorem</b>	638
Exercises for Section 6.4	638
<b>6.5 Compactification</b>	638
Exercises for Section 6.5	638
<b>6.6 Dimension</b>	638
Exercises for Section 6.6	642
<b>7 Fundamental Group and Covering Spaces</b>	<b>643</b>
<b>7.1 The fundamental group</b>	643
Definition of the fundamental group	643
Exercises for Section 7.1	643
<b>7.2 Covering spaces</b>	643
Definition of a covering space	643
Exercises for Section 7.2	643
<b>Guide to the Exercises</b>	<b>645</b>
<b>Bibliography</b>	<b>647</b>
<b>Additional Readings</b>	<b>653</b>
<b>List of Named Results</b>	<b>657</b>
<b>Index of Notation</b>	<b>661</b>
<b>Index of Names</b>	<b>667</b>
<b>Index</b>	<b>669</b>

# Short Contents

<b>Detailed Contents</b>	<b>vii</b>
<b>Short Contents</b>	<b>xiii</b>
<b>List of Figures</b>	<b>xv</b>
<b>List of Tables</b>	<b>xix</b>
<b>Preface</b>	<b>xxi</b>
<b>0 Sets and Maps</b>	<b>1</b>
0.1 Logic . . . . .	2
0.2 Sets . . . . .	7
0.3 Maps . . . . .	19
0.4 Families and Products . . . . .	34
0.5 Recursion . . . . .	47
0.6 Countability . . . . .	53
0.7 Ordering relations . . . . .	63
0.8 Order-Completeness of the Real Numbers . . . . .	78
0.9 Equivalence Relations . . . . .	93
0.10 Ordinals, Cardinality, and Zorn's Lemma . . . . .	109
<b>1 Metric Spaces</b>	<b>127</b>
1.1 Metrics . . . . .	128
1.2 Open Sets and Closed Sets . . . . .	146
1.3 Equivalent Metrics . . . . .	164
1.4 Continuity and Convergence . . . . .	179
1.5 Completeness . . . . .	200
<b>2 Topological Spaces</b>	<b>225</b>
2.1 Topologies . . . . .	225
2.2 Neighborhoods . . . . .	235
2.3 Boundary, Interior, and Closure . . . . .	246
2.4 Bases and Local Bases . . . . .	263
2.5 Separation Properties . . . . .	292
<b>3 Continuity and Convergence</b>	<b>301</b>
3.1 Continuous Maps . . . . .	302
3.2 Homeomorphisms . . . . .	319
3.3 Product Spaces . . . . .	359
3.4 Quotient Spaces . . . . .	385

3.5	Convergence . . . . .	420
<b>4</b>	<b>Compactness</b>	<b>457</b>
4.1	Compact Spaces . . . . .	458
4.2	Compact Metric Spaces . . . . .	493
4.3	Locally Compact Spaces . . . . .	508
<b>5</b>	<b>Connectedness</b>	<b>523</b>
5.1	Connected Spaces . . . . .	524
5.2	Components and Locally Connected Spaces . . . . .	545
5.3	Path-Connected Spaces . . . . .	558
5.4	Homotopy . . . . .	571
5.5	Simply connected spaces . . . . .	587
<b>6</b>	<b>Embedding</b>	<b>605</b>
6.1	The Tychonoff Product Theorem . . . . .	605
6.2	Complete regularity and normality . . . . .	612
6.3	Embedding theorems . . . . .	631
6.4	Urysohn Metrization Theorem . . . . .	638
6.5	Compactification . . . . .	638
6.6	Dimension . . . . .	638
<b>7</b>	<b>Fundamental Group and Covering Spaces</b>	<b>643</b>
7.1	The fundamental group . . . . .	643
7.2	Covering spaces . . . . .	643
	<b>Guide to the Exercises</b>	<b>645</b>
	<b>Bibliography</b>	<b>647</b>
	<b>Additional Readings</b>	<b>653</b>
	<b>List of Named Results</b>	<b>657</b>
	<b>Index of Notation</b>	<b>661</b>
	<b>Index of Names</b>	<b>667</b>
	<b>Index</b>	<b>669</b>

# List of Figures

0.1	The product $\mathbb{N} \times Y$ as a “picket fence.”	39
0.2	In the lexicographic ordering of the product set $X \times Y$ , point $p \leq q$ and $p \leq r$ .	74
0.3	Visualizing the product $\mathbb{N} \times [0, 1[$ with its lexicographic ordering.	74
0.4	Geometric meaning of the Archimedean ordering property	82
0.5	Passing to the quotient.	102
1.1	The triangle inequality.	132
1.2	Taxicab distance.	133
1.3	Max metric distance between two functions.	135
1.4	The $d_1$ -distance between functions as area.	136
1.5	Ball of radius $\varepsilon$ at $x$ in $\mathbb{R}^1$ .	148
1.6	Ball of radius $\varepsilon$ at $x$ in $\mathbb{R}^2$ .	148
1.7	Ball of radius $\varepsilon$ at $x$ in $\mathbb{R}^3$ .	149
1.8	A ball for the max metric is a cube.	150
1.9	The $\varepsilon$ -ball for the sup metric in $C([0, 1])$ .	151
1.10	Ball contains smaller ball centered at one of its points.	152
1.11	Ball and open set for metric induced on subset.	156
1.12	Ball disjoint from a disk	158
1.13	Separating a point from a $d$ -closed set in a metric space.	159
1.14	Distances between two functions for the sup metric $d_\infty$ and the $L_1$ -metric $d_1$ on $C([0, 1])$ .	166
1.15	Function in a $d_1$ -ball but not in a $d_\infty$ -ball.	167
1.16	Graph of $y = \varphi(x) = x/(1 +  x )$ .	169
1.17	Sets in proof of (1) $\implies$ (2) in Theorem 1.53.	183
1.18	Sets in proof of (2) $\implies$ (3) in Theorem 1.53.	183
1.19	Coordinate-wise convergence of a sequence in $\mathbb{R}^2$ .	191
1.20	Coordinate-wise convergence of a sequence in $\mathbb{R}^k$ .	191
1.21	Proving that a Cauchy sequence in the real line converges in the case that it has infinite range.	204
1.22	Proving continuity of the uniform limit of a sequence of continuous real-valued functions.	211
1.23	Start of the construction for proving the Baire Category Theorem.	214
1.24	Constructing a sawtooth function from a piecewise linear function.	217
2.1	Open set in a subspace.	231
2.2	Relative topology induced by a metric.	231
2.3	Neighborhood of $0'$ in line with two origins.	240
2.4	Alternative depiction of the line with two origins.	240
2.5	Neighborhood of $t \in L$ in half-disk space.	242
2.6	Why the half-disk space is not metrizable.	243

2.7	Boundary of an annulus. . . . .	247
2.8	Points not in boundary of a ball in $\mathbb{R}^n$ . . . . .	248
2.9	Point on boundary of ball. . . . .	249
2.10	Typical neighborhoods in the Thomas plank $P$ . . . . .	267
2.11	A basic open set in the Sorgenfrey plane. . . . .	275
2.12	Some functions in a basic open set for the topology of pointwise convergence. . . . .	275
2.13	Constructing a rational-sided, rational-center cube inside a neighborhood of a point in $\mathbb{R}^n$ . . . . .	277
2.14	A typical neighborhood of a point in the Bing triangle space. . . . .	291
2.15	Typical subbasic open set $U_\epsilon(x, y)$ in the Bing triangle space. . . . .	292
3.1	Oscillating function $\sin(1/x)$ . . . . .	308
3.2	Projections of an open square and a hyperbolic arc in the plane. . . . .	314
3.3	Constructing a homeomorphism from the centered $n$ -cube to the $n$ -disk. . . . .	326
3.4	Constructing a homeomorphism from the punctured $n$ -sphere to the equatorial hyperplane in $\mathbb{R}^{n+1}$ . . . . .	328
3.5	Generating a surface $T$ in $\mathbb{R}^3$ that is homeomorphic to the torus $S_1 \times S_1$ . . . . .	331
3.6	The surface $T$ in $\mathbb{R}^3$ obtained by rotating a circle. . . . .	332
3.7	Steps in proving that each point in an $n$ -dimensional locally Euclidean space has arbitrarily small neighborhoods that are Euclidean. . . . .	338
3.8	Local coordinates near a point in a 2-dimensional locally Euclidean space. . . . .	340
3.9	Transition map for two charts $\langle U, \varphi \rangle$ and $\langle V, \psi \rangle$ in a 2-dimensional locally Euclidean space $M$ . . . . .	341
3.10	Trefoil knot (left) and unknotted circle (right) embedded in 3-space. . . . .	349
3.11	The figure-eight. . . . .	352
3.12	The standard 2-simplex in $\mathbb{R}^3$ (left) and its homeomorph in $\mathbb{R}^2$ (right). . . . .	354
3.13	Whiskers attached to the 2-sphere outward (left) and inward (right). . . . .	354
3.14	Neighborhood of a point in the product of two spaces. . . . .	366
3.15	The real line embedded as a subspace $L$ of $\mathbb{R}^3$ . . . . .	376
3.16	Point $x \in V \subset S_1$ contains an open arc. Case (i): $x \neq \langle 1, 0 \rangle$ . . . . .	387
3.17	Point $x \in V \subset S_1$ contains an open arc. Case (ii): $x = \langle 1, 0 \rangle$ . . . . .	388
3.18	An open set in a quotient space and its inverse image under the quotient map. . . . .	389
3.19	Identifying ends of the unit interval $I$ to obtain the unit circle $S_1$ . . . . .	394
3.20	Graphs of period 1 $\cos$ and $\sin$ functions. . . . .	395
3.21	Identify opposite vertical edges of a square to obtain a cylinder. . . . .	396
3.22	Identify opposite edges of a square to obtain a torus. . . . .	397
3.23	Rotate a circle in the $yz$ -plane about the $z$ -axis to obtain a torus. . . . .	398
3.24	Constructing a homeomorphism of the torus $S_1 \times S_1$ with the surface $T$ in $\mathbb{R}^3$ . . . . .	399
3.25	Identify edges of a square to obtain the Möbius strip. . . . .	400
3.26	How edges of the square are identified in order to the Klein bottle. . . . .	400
3.27	Trying to identify the end circles of a cylinder inside $\mathbb{R}^3$ to obtain the Klein bottle. . . . .	401
3.28	Antipodal points and great circles on the 2-sphere $S_2$ . . . . .	401
3.29	Antipodal points on the lower hemisphere are diametrically opposite points on the equatorial circle. . . . .	402
3.30	Constructing the embedding $r$ of the plane $H$ into the lower hemisphere $S_2^-$ . . . . .	403
3.31	Collapsing the bounding circle $S_1$ of the disk $D_2$ to a point to obtain the sphere $S_2$ . . . . .	405
3.32	Constructing a continuous closed surjection $f$ from the disk $D_2$ to the sphere $S_2$ that maps the inside of the disk injectively onto the complement of the north pole. . . . .	405
3.33	Attaching the unit interval $[0, 1]$ to the disk $D_2$ by the map $f$ that sends the endpoints of $[0, 1]$ to the origin. . . . .	407
3.34	Lower and upper Riemann sums $L_P$ and $U_P$ for a partition $P$ of the closed interval $[a, b]$ . . . . .	427



3.35	A sequence in $\mathbb{R}$ that clusters but does not converge. . . . .	433
3.36	Relationship between a net with index set $I$ in a set $X$ and a subnet with index set $J$ . . . . .	436
4.1	A cover of a closed subset $K$ of a topological space $X$ by open subsets of $X$ . . .	464
4.2	Proving that a compact subset of a Hausdorff space is closed. . . . .	465
4.3	The first three steps in forming the Cantor set $K$ . . . . .	467
4.4	The first three approximations to Cantor's function $\kappa$ . . . . .	472
4.5	Cantor's function $\kappa$ . . . . .	473
4.6	A tube around a slice in $\mathbb{R} \times [0, 1]$ . . . . .	477
4.7	The first and second stages in constructing the Sierpinski carpet. . . . .	491
4.8	The fourth stage $S_4$ in constructing the Sierpinski carpet. . . . .	491
4.9	The first and third stages in constructing the Menger sponge. . . . .	492
4.10	Constructing a piecewise linear function $g$ to uniformly approximate a continuous function $f: [a, b] \rightarrow \mathbb{R}$ . . . . .	503
4.11	Collapsing the "ends" of the extended real line to a single point. . . . .	514
5.1	Geometric meaning of the Brouwer Fixed-point Theorem in dimension 1. . . .	531
5.2	Antipodal points on the circle $S_1$ . . . . .	531
5.3	Configuration of line segments for proving that the complement of a countable set in $\mathbb{R}^n$ is connected. . . . .	535
5.4	Intersecting slices in a product of connected spaces. . . . .	537
5.5	Constructing a point $y$ at the inductive step in proving that the product of a sequence of connected spaces is connected. . . . .	539
5.6	Kuratowski's example of a continuous bijection $f$ between non-homeomorphic unions of sequences of intervals. . . . .	548
5.7	Kuratowski's example of a continuous bijection $g$ between non-homeomorphic unions of sequences of intervals. . . . .	548
5.8	The Knaster-Kuratowski example of an uncountable connected space having an explosion point. . . . .	550
5.9	The topologist's comb: a connected space that is not locally connected. . . .	554
5.10	The product of two paths. . . . .	560
5.11	The topologist's sine curve is connected but not locally connected. . . . .	562
5.12	The $\varepsilon$ -ball at a point $\langle 0, y \rangle$ on the limiting line $L$ of the topologist's sine curve. . .	562
5.13	Configurations of opposite corners of the 8 subsquares of a given square. . . .	567
5.14	The first 9 squares and their designated opposite corners. . . . .	568
5.15	The first 81 squares and their designated corners. . . . .	569
5.16	The 2-disk $D_2$ and the annulus $A$ do not seem to be homeomorphic. . . . .	572
5.17	A deformation retraction of the annulus onto the circle. . . . .	574
5.18	A path-homotopy as a map on the unit square. . . . .	578
5.19	Dividing the unit square into two rectangle to make $\sigma * \tau \sim \sigma' * \tau'$ . . . . .	579
5.20	Example of non-associativity of path-products. . . . .	580
5.21	Constructing a path-homotopy $h: \sigma * (\tau * \eta) \sim (\sigma * \tau) * \eta$ . . . . .	582
5.22	Constructing a path-homotopy $h: \sigma * \overleftarrow{\sigma} \sim \nu_x$ . . . . .	584
5.23	A nullhomotopic loop. . . . .	588
5.24	. . . . .	589
5.25	Constructing the loop $\sigma$ in the proof of Theorem 5.97. . . . .	590
5.26	Constructing the path-homotopy $F: \sigma * \omega * \overleftarrow{\sigma} \sim \nu_x$ in the proof of Theorem 5.97. . .	591
5.27	Simple and non-simple closed curves. . . . .	592
5.28	Which points are inside, and which are outside, this simple closed curve? . . .	592
5.29	Evenly covering an open arc on the unit circle. . . . .	594
5.30	Constructing the map $G_j$ to prove the homotopy-lifting property for the circle. . .	596

- 5.31 The domains of  $G_j$  and  $G_{j+1}^i$  in the proof of the homotopy-lifting property. . . 598
- 5.32 Construction of  $r$  in the proof of the Brouwer Fixed-point Theorem in Dimension 2. 599
- 6.1 A basic neighborhood  $U_n(z)$  at a point  $z \in C_1$  in the Alexandroff double circle. 636

# List of Tables

0.1	Meanings of logical symbols. . . . .	3
0.2	Some common logical abbreviations. . . . .	7
2.1	Basic relationships among separation properties. . . . .	297
2.2	Relationships among complete regularity and other separation properties. . .	297
3.1	Preservation of topological properties: separation and countability properties— version 1 . . . . .	351
3.2	Preservation of topological properties: separation and countability properties— version 2 . . . . .	411



# Preface

chap:preface

This book introduces the basic concepts, facts, and techniques of general topology and the fundamental group at a level appropriate to a student's first exposure to the subject. It is suitable as a text for a variety of undergraduate and graduate courses of differing lengths and emphases, and for classes having varying backgrounds; some possible course outlines are suggested below. It may be used in a conventional course environment or for self-study.

The only mathematical prerequisite for reading this book is calculus, although some exposure to a first course in linear algebra is desirable. No knowledge of the topology of Euclidean spaces or metric spaces is assumed. Some prior experience with "epsilonotics" is desirable, but not indispensable; the greater the prior experience, the more material can be covered.

The first edition avoided Zorn's Lemma and ordinal numbers entirely. In this second edition, both those topics are introduced and exploited, but the development has been structured so that they may still be avoided if desired.

One of our aims has been to assist the student's mathematical maturation. Hence careful attention has been paid to motivating new notions. For example, eleven pages of examples of metrics precede the actual definition of a metric, and a proof of the compactness of the unit interval (together with the corollary that continuous functions on it are bounded) precede the definition of compactness. There are examples galore of everything. Special pains have been taken to explain the significance of theorems and to write enough proofs in enough detail to provide models for the student's own proof-making. In using a preliminary version of the first edition, the author found that students can read much of the text themselves with only minimal guidance by the instructor, so that classroom time can be devoted mainly to filling in details or missing parts of proofs, solving the exercises, and discussing the more difficult points.

Chapter 0 concisely presents the necessary preliminaries on sets and maps, induction and recursion, families and products, countability, ordering relations, order-completeness of the real numbers, and equivalence relations. It introduces as well the Axiom of Choice and related matters—including Zorn's Lemma, ordinals, and cardinality—for those choosing to study the topics and examples involving these topics.

. The time spent on this material will, of course, depend on the student's background, but it is suggested that all students at least read the chapter rapidly for review and to fix terminology. It is a good idea to review the section on equivalence relations in conjunction with the study of quotient spaces in Chapter 3.

Chapter 1 introduces open sets, closed sets, neighborhoods, continuous maps, and convergent sequences in metric spaces. The terminology used here— $d$ -closed set,  $(d, d')$ -continuous map, and so on—calls attention to the particular metrics involved. At the same time the entire thrust of the chapter is to justify the later definition of a topological space by demonstrating that notions of continuity and convergence remain unchanged when the metrics are replaced by equivalent ones. The section on completeness includes the Baire category theorem and its application to the existence of nowhere differentiable, continuous

functions. Owing to its greater technical difficulty as well as its treatment of uniform, as distinct from topological, ideas, this section may be postponed or even omitted; except for occasional mention, completeness does not appear again until Section 2 of Chapter 4, where it is used to characterize compact metric spaces and to prove the Tychonoff theorem for a sequence of compact metric spaces.

The study of topology proper is begun in Chapter 2, where topologies, neighborhoods, Hausdorff spaces, bases and local bases, and countability properties are discussed. Of the whole hierarchy of separation properties, which in its entirety is liable to confuse the beginner, the property  $T_2$  is dealt with at length separately, the others being postponed to a subsequent separate section and, in the case of the circle of ideas relating complete regularity and normality, to a later chapter.

Continuity is the theme of Chapter 3. Here product and quotient spaces are constructed and their mapping properties are emphasized. Here, too, the theory of convergence of nets is introduced at a level kept elementary by treating subnets only briefly and by avoiding universal nets entirely. This theory deserves to be included in a first topology course: It places sequential convergence in proper perspective, facilitates a later study of filter convergence, and reveals the diverse kinds of limits the student has previously encountered as instances of a single unifying concept. Nevertheless, the treatment of nets can be omitted without substantial loss: the only places nets are used again are the net characterization of compactness (4.27), which can itself be omitted, and the sequence characterization of compactness of a metric space (4.36), which requires only the equivalence of sequential clustering with subsequential convergence (3.69). The elementary facts about compactness are developed in Chapter 4. To accommodate those who do not wish to include Zorn's Lemma, in this chapter the Tychonoff theorem is proved here only for the product of finitely many spaces and separately for the product of a sequence of metrizable spaces; proof of the general case is deferred until Chapter xx. Sequential compactness and other variants of compactness are not studied in their own right, only as equivalents of compactness in the metrizable case. Also considered for metric spaces is the relationship of compactness to uniform continuity and to completeness. The discussion of locally compact spaces includes the one-point compactification.

The first three sections of Chapter 5 present the standard facts about connected sets, components, locally connected spaces, and path-connected spaces. This material, which is technically if not conceptually simpler than that on compactness, can be read before Chapter 4. The final two sections, on homotopy, culminate in proofs of the Brouwer fixed-point theorem in dimension 2 and the fundamental theorem of algebra. The only thing about compactness needed in these two sections is the existence of a Lebesgue number (4.41) for an open cover of the unit interval or of the unit square. In order to keep the algebraic machinery to a minimum and make things geometrically more transparent, our initial treatment of homotopy avoids explicit mention of the fundamental group; a full treatment of the fundamental group, such as is common in introductory graduate courses, appears separately.

The exercises, found at the end of each section, number x83 in all. Ranging in difficulty from the routine to the challenging, they are meant both to test comprehension of the ideas presented in the text and to provide applications, additional examples, and extensions of these ideas. Many call not just for proofs, but for answers to such questions as "Is it true that ... ?" or "What can be said about ... ?" or "Is there an analog . . . ?" Included in the exercises are a number of topics this author did not deem so essential to a first course in topology to have included them in the text proper, but which are interesting and important in their own right. These topics are as diverse as completion of a metric space, Cartesian sum topologies of general families (the case of two being included in the body of the text), manifolds with boundary, topological groups, the closed-graph theorem, and cut points.

There are surely more exercises than an instructor would want to assign to any one class. We have therefore appended a Guide to the Exercises in which we cite each exercise needed for exercises in subsequent sections.

All definitions, theorems, and examples within a single chapter are numbered consecutively, so that 3.15 refers to the fifteenth item in Chapter 3. The exercises in each chapter are separately numbered consecutively; a reference to the fifth exercise of Chapter 3 would be "Exercise 15" if made within that chapter, but "Exercise 3.15" if made in another chapter.

The Bibliography includes only those books and articles referred to in the text or suggested for further reading. References to bibliographic entries are made by numbers enclosed in brackets. Appended to the Bibliography is a list of suggested readings on special topics about which individual students might report to the class.

## Aims

MORE

## Prerequisites

MORE

## Contents

MORE

## Exercises

The exercises, found at the end of each section, number xxx in all. Ranging in difficulty from the routine to the challenging, they are meant both to test comprehension of the ideas presented in the text and to provide applications, additional examples, and extensions of these ideas. Many call not just for proofs, but for answers to such questions as "Is it true that . . . ?" or "What can be said about . . . ?" or "Is there an analog . . . ?"

Included in the exercises are a number of topics this author did not deem so essential to a first course in topology to have included them in the text proper, but that are interesting and important in their own right. These topics are as diverse as completion of a metric space, Cartesian sum topologies, manifolds with boundary, topological groups, the closed-graph theorem, and cut points. Some topics relegated to exercises in the earlier chapters are taken up again later within the body of the text, especially in sections and chapters added for this new edition.

There are surely more exercises than an instructor would want to assign to any one class. We have therefore appended a [Guide to the Exercises \(page 645\)](#) in which we cite each exercise that is needed for exercises in subsequent sections.

## Changes from the first edition

In general, now that is free of the commercial constraint of a limit on the number of pages, this edition includes additional topics and expanded coverage of topics from the first addition.

- [Chapter 0](#) has been considerably enlarged so as to include, among other things, explicit treatment of recursion; a brief introduction to groups (needed for XXX); and an entire new section on well-ordering, the first uncountable ordinal, and Zorn's lemma. The last provides background needed to understand the long line and the Tychonoff plank,

which provide important counterexamples, as well as the foundational machinery for proving the Tychonoff Product Theorem in full generality.

Note that it is entirely possible to skip the new material on well-ordering and ordinals and then just ignore examples involving the long line and the Tychonoff plank.

- Product spaces are no longer restricted to the situation of countably many factors, and accordingly the full Tychonoff Product Theorem is proved.
- The fundamental group is now explicitly defined and studied. (Nearly all the machinery to define it already appeared in the first edition.)
- Completely regular and normal spaces, treated only briefly in the first edition, are now covered much more fully. While normal spaces appear first in This coverage includes now Urysohn's Lemma and the Tietze Extension Theorem; the Urysohn Metrization Theorem; and Tychonoff's Embedding Theorem.
- (*MAYBE*) The Stone-Čech compactification is defined.

MORE

## Numbering scheme

Each chapter is divided into numbered sections, so that, for example, Section 2.4 designates the fourth section of Chapter 2. Each section in turn is subdivided into unnumbered subsections.

*All definitions, theorems, propositions, lemmas, corollaries, and examples within a single chapter are numbered consecutively*, so that, for example, 3.15 refers to the fifteenth such item in Chapter 3. The exercises within each chapter are separately numbered consecutively; a reference to the fifth exercise of Chapter 3 would be "Exercise 5" if made within that chapter, but "Exercise 3.5" if made in another chapter.

## References

The main [Bibliography](#) includes only those books and articles referred to in the text or suggested for additional background or further reading. References to bibliographic entries are made by numbers enclosed in brackets. A separate [Additional Readings](#) provides a guide to literature on some special topics that could be used for individual or group investigations, reports, and projects.

## Russian names

As with any topology text, this one necessarily includes work by the Russian/Soviet mathematicians with the Cyrillic surnames Александров, Тихонов, and Урысон. Often today these names are transliterated into English as Aleksandrov, Tikhonov, and Uryson, respectively. However, the work in question became known primarily through papers that appeared in German or French journals. Accordingly, we use the same, older, transliterations as did those journals, namely: Alexandroff, Tychonoff, and Urysohn.

## Suggested course outlines

The list below is not meant to be exhaustive, but only' suggestive of possible courses that can be based on this text. The portion of the text covered by any given class will, of course, depend on the students' preparation; it will also depend heavily on the number and difficulty of problems assigned from among the many we have provided.



*A minimal course*  
(1 quarter or 1 semester)

Sections 1 through 5 Sections 1 through 4 omit 2.41, 2.42, and 2.56(5) and (6) Section 1 except 3.11 (2) ; Section 2 except 3.22 (3) and 3.23; Section 3 through 3.40, except 3.35(5); Section 6 of Chapter 0; Section 4 except 3.49(7) through (9); Section 5 through 3.54 Section 1 except 4.27 Section 1 except 5.25 and 5.26; Section 2 through 5.33 or 5.35(4); Section 3 through 5.51-optional

*A second course in topology*  
(1 quarter or 1 semester)

Chapter 1 Section 5 Chapter 2 Chapter 3 Chapter 4 2.41 and 2.42 Section 3 from 3.41; Sections 4 and/ or 5 4.27 (if Section 5 of Chapter 3 is included); Sections 2 and 3, or 4.41 through 4.44 and Section 3 through 4.56 Section 2 from 5.34; Sections 3 through 5 Chapter 5 Additional readings or individual projects (see the Bibliography)

*A complete course*  
(2 semesters or 3 quarters)

Chapters 0 through 5 Additional readings or individual projects

*A standard course—emphasis on geometry*  
(1 semester or 2 quarters)

Sections 1 through 5 Sections 1 through 4 omit 2.41, 2.42, and 2.56(5) and (6) Section 1 except 3.11(2); Section 2 except 3.23; Sectionf 3 through 3.40; Section 6 of Chapter 0; Section 4; Section 5 through 3.53 Section 1; Theorem 4.41 omit 5.35(5); include Exercises 5.97, 5.98, 5.107

*A standard course—emphasis on analysis*  
(1 semester or 2 quarters)

Sections 1 through 5 include Exercises 1.85 and 1.86 Sections 1 through 3; Section 6 of Chapter 0; Section 4 except 3.49(7) through (9); Section 5 Section 1; Section 2 through 5.33; Section 3 through 5.51-optional

*A brief course in set theory*  
(3 to 5 weeks)

Selections from Chapter 0

*A course in special topics*  
(variable time)

1.71 through 1.73; 1.69, 1.70, and Exercises 1.85 and 1.86 3.23 (include Example 1.9 and Exercise 1.23) 4.28 and 4.29; 4.41 Examples 5.35 (5) and/ or 5.52; Sections 3 through 5

## Acknowledgements

My students at the University of Massachusetts Amherst spotted numerous errors and detected defects in exposition in a preliminary version of the first edition. Students there also provided feedback on additional materials, now incorporated into this second edition, that were used for several years in the graduate topology course. For the first edition, Professor Victor Klee and several anonymous reviewers corrected infelicities of style and suggested many improvements. For the second edition, several contributors to [math.stackexchange.com](https://math.stackexchange.com) provided useful examples.

Then-state-of-the-art technology was used for the first edition, namely manuscript was typed on a typewriter from my hand-written scrawl, Ms. Margo Vidrine and Mrs. Rita Warner. This second edition was prepared directly on a computer using the  $\text{\LaTeX}$  document preparation system. To help me bend  $\text{\LaTeX}$  and various packages for it to my will, I am greatly indebted to TeXperts (egreg, marmot, etc. in [tex.stackexchange](https://tex.stackexchange.com); Richard Koch and Herbert Schulz for TeXShop); etc.

MORE

MURRAY EISENBERG  
*Rockville, Maryland*  
August 19, 2023

## CHAPTER

# 0

## Sets and Maps

chap:setsmaps

### Chapter Contents

0.1 Logic . . . . .	2	Exercises for Section 0.5 . . . . .	52
Predicate logic . . . . .	4	0.6 Countability . . . . .	53
Exercises for Section 0.1 . . . . .	7	Finite and infinite sets . . . . .	53
0.2 Sets . . . . .	7	Denumerable, countable, and un-	
Building sets . . . . .	8	countable sets . . . . .	57
Special sets . . . . .	9	Exercises for Section 0.6 . . . . .	61
Subsets . . . . .	10	0.7 Ordering relations . . . . .	63
Mathematical induction . . . . .	11	Preorderings . . . . .	63
Power sets . . . . .	13	Partial orderings . . . . .	65
Union and intersection of two sets . . . . .	13	Intervals . . . . .	70
Complements . . . . .	14	Bounds . . . . .	71
Pairs and products . . . . .	14	Total orderings . . . . .	72
Relations . . . . .	15	Exercises for Section 0.7 . . . . .	75
Exercises for Section 0.2 . . . . .	17	0.8 Order-Completeness of the Real Numbers . . . . .	78
0.3 Maps . . . . .	19	Order-completeness . . . . .	79
Maps . . . . .	19	Base expansion . . . . .	79
Operations . . . . .	22	Archimedean Ordering and Nested	
Restriction and extension . . . . .	23	Interval Properties . . . . .	81
Composition of maps . . . . .	25	Properties of intervals in the real line . . . . .	85
Images and inverse images . . . . .	27	Natural numbers, integers, and ra-	
Injections, surjections, and bijections . . . . .	28	tionals within the reals . . . . .	87
Inverse of a bijection . . . . .	30	Complex numbers . . . . .	89
Exercises for Section 0.3 . . . . .	31	Exercises for Section 0.8 . . . . .	91
0.4 Families and Products . . . . .	34	0.9 Equivalence Relations . . . . .	93
$n$ -tuples and sequences . . . . .	35	Equivalence relations . . . . .	93
$n$ th power of a set . . . . .	35	Quotient sets . . . . .	96
Euclidean spaces . . . . .	36	Partitions . . . . .	98
Line segments and lines . . . . .	37	Saturated sets . . . . .	99
Arbitrary power of a set . . . . .	37	Passing to the quotient . . . . .	100
Union and intersection of a family		Exercises for Section 0.9 . . . . .	103
of sets . . . . .	38	0.10 Ordinals, Cardinality, and Zorn's Lemma . . . . .	109
Union and intersection of a collec-		Well-ordered sets . . . . .	109
tion of sets . . . . .	41	The first uncountable ordinal . . . . .	110
Product of a family of sets . . . . .	42	Maximal and minimal elements . . . . .	112
Axiom of Choice . . . . .	44	Zorn's Lemma . . . . .	113
Exercises for Section 0.4 . . . . .	45	The Well-ordering Theorem . . . . .	116
0.5 Recursion . . . . .	47	Cardinality . . . . .	118
Ordinary recursion . . . . .	47	Exercises for Section 0.10 . . . . .	123
Primitive and complete recursion . . . . .	50		

## Introduction

In this preliminary chapter we collect some essential information about sets and maps that are used throughout the text. Much of this material will doubtless not be new to you and is therefore covered rapidly, with few examples or proofs, so as to remind you of what you already know and to fix the particular terminology and notation that we shall use. Those topics that are likely to be less familiar—countability, order-completeness of the real numbers, the Axiom of Choice, Zorn’s Lemma, well-ordered sets, and ordinal numbers—are covered in somewhat greater detail.

For a fuller treatment of these preliminaries at an elementary level see, for example, Eisenberg [22], Lakins [42], or Pinter [53]. For an advanced axiomatic treatment see Bourbaki [9], Rubin [56], or Bernays [3] (or, with a different axiomatic foundation, Suppes [63]).

### 0.1 Logic

sec:logic

Logic makes explicit and crisp those patterns of deductive reasoning used by careful speakers and writers. Only in a few instances—for example, the more generous meanings of ‘or’ and ‘there exists’—does logical language differ a bit from everyday usage.

Any systematic or detailed treatment of logic would take us far afield of even the set-theoretic preliminaries presented in the remaining sections of this chapter, and of topology itself, presented in the rest of the book. For informal treatments of logic and, especially, methods of proof, see, for example, Lakins [42, chapters 1 and 2], Sundstrom [62, chapters 1–3], or Velleman [65]. For a more thorough and formal treatment, see, for example, Bourbaki [9, chapter I].

In principle, logic ought to precede everything else in mathematics that uses it, yet without any actual mathematics to treat, logic is wholly abstract. Accordingly, to present notions from logic we shall have to talk about mathematical objects. Subsequent sections of this chapter do present basics about such mathematical objects.

The logic to be discussed here concerns mathematical *statements* (in the formal vocabulary of logic, **propositions**—assertions which may be true or false—about mathematical objects such as numbers, sets, functions, and geometric structures. For example:

$$e < \pi.$$

$$n \text{ is a multiple of } 4.$$

$$|x^2 - 3^2| < \varepsilon.$$

The following logical symbols are used to combine or qualify such statements:

- the logical **connectives**  $\neg$  (negation),  $\vee$  (disjunction),  $\wedge$  (conjunction),  $\implies$  (implication), and  $\iff$  (equivalence);
- the logical **quantifiers**  $\forall$  (universal quantifier),  $\exists$  (existential quantifier), and  $\iota$  (descriptor);
- the set-theoretic signs  $=$  (equality) and  $\in$  (elementhood)—the latter not to be confused with the Greek epsilon,  $\varepsilon$ ; and
- parentheses, as punctuation marks, when needed to avoid ambiguity.

For example:

$$5\pi/15 = \pi/3 \text{ and } e < \pi.$$

Every integer that is a multiple of 4 is even.

There exists a real number  $c$  such that  $c^2 = -1$ .

For every  $\varepsilon > 0$ , there exists some  $\delta > 0$  such that  $|x^2 - 3^2| < \varepsilon$  whenever  $|x - 3| < \delta$ .

The positive real number  $s$  for which  $s^2 = 2$  is not rational.

Table 0.1: Meanings of logical symbols.

tab:logic-words

<i>Symbolism</i>	<i>Meanings</i>
$\neg(P)$	not $P$ it is not the case that $P$
$(P) \& (Q)$	$P$ and $Q$ both $P$ and $Q$ $P$ but $Q$ $P$ yet $Q$ $P$ whereas $Q$
$(P) \vee (Q)$	$P$ or $Q$ $P$ and/or $Q$ either $P$ or $Q$ or both $P$ and $Q$
$(P) \implies (Q)$	$P$ implies $Q$ if $P$ , then $Q$ $P$ only if $Q$ $Q$ if $P$ $Q$ provided that $P$ $P$ is sufficient for $Q$ $Q$ is necessary for $P$ $Q$ because $P$ $Q$ since $P$ $P$ whence $Q$
$(P) \iff (Q)$	$P$ if and only if $Q$ $P$ iff $Q$ $P$ is (logically) equivalent to $Q$ $P$ is necessary and sufficient for $Q$ $P$ precisely when $Q$
$(\forall x)(P)$	for all $x$ , $P$ for every $x$ , $P$ for each $x$ , $P$ for arbitrary $x$ , $P$ for any $x$ , $P$
$(\exists x)(P)$	there exists $x$ such that $P$ there is an $x$ such that $P$ for some $x$ , $P$
$(\iota x)(P)$	the $x$ such that $P$ the unique $x$ such that $P$
$x = y$	$x$ equals $y$
$x \in Y$	$x$ is an element of $Y$ $x$ is a member of $Y$ $x$ belongs to $Y$ $x$ is in $Y$

predicate

subsec: predicate-logic

**Predicate logic****A predicate**

A letter  $x$  is said to be **bound** in a string  $S$  in case any one of the strings  $(\forall x)$ ,  $(\exists x)$ ,  $(\iota x)$  occurs as part of  $S$ . Otherwise, the letter is said to be **free** in the string. In particular, the letter is free in the string if it does not occur there at all. Here are two examples, which use the notations  $\mathbb{N}$  for the set  $\{0, 1, 2, \dots\}$  of all natural numbers and  $\mathbb{Z}$  for the set  $\{\dots, -2, -1, 0, 1, 2, 3, \dots\}$  of all integers. First,  $n$  is bound in

$$(\forall n)(n \in \mathbb{N} \implies 2n > n)$$

(which means that the double of each natural number is greater than the number itself). Second,  $n$  is bound in

$$(\iota n)(n \in \mathbb{Z} \ \& \ 2n = n)$$

(which represents the one and only natural number whose double is itself, namely, 0). However,  $n$  is free in each of the three strings:

$$2n > n, \quad n \in \mathbb{N} \implies 2n > n, \quad k \in \mathbb{N} \implies 2k > k$$

When a particular occurrence of a letter  $x$  is bound in a string  $S$  because that occurrence is within part of  $S$  that has one of the three form  $(\forall x)(P)$ ,  $(\exists x)(P)$ ,  $(\iota x)(P)$ , then that occurrence of  $x$  is said to be **within the scope** of that quantifier  $\forall$ ,  $\exists$ , or  $\iota$ .

For example,

$$\{\text{eqn: xNOTEQx}\} \ (*) \quad (\forall x)(\neg(x = x))$$

and

$$\{\text{eqn: xINyIFFxNOTEQx}\} \ (**) \quad (\iota A)((\forall x)((x \in A) \iff \neg(x = x)))$$

are formulas. (The first of these makes the *false* statement that each object is unequal to itself. The second represents the object characterized by the property that no object belongs to it—the *empty set*).

Parentheses are used to group

The order of precedence in logic is similar to the one in algebra where multiplication takes precedence over addition, so that  $x \cdot y + z$  means  $(x \cdot y) + z$  rather than  $x \cdot (y + z)$ . In logic, if we consider the connectives and set-theoretic signs in the order

$$\longleftarrow \text{weaker} \quad = \in \neg \ \& \ \vee \implies \iff \longrightarrow \text{stronger}$$

from “weakest” to “strongest”, then the rule is: **when parentheses are missing, the stronger sign reaches further.** For example,  $x \in A \vee x \in B$  means  $(x \in A) \vee (x \in B)$ , and  $x = y \implies x \in A$  means  $(x = y) \implies (x \in A)$ . The formula  $(*)$  on page 4 may now be written in the shorter form

$$(\forall x)(\neg x = x)$$

and the term  $(**)$  on page 4 may now be written much more readably as:

$$(\iota A)(\forall x)(x \in A \iff \neg x = x)$$

Some quantified formulas, such as  $(\forall x)(x = x)$ , are intended to make “universal” statements about all possible objects. Many quantified formulas, however, are intended to

make statements about only those objects that are elements of some particular set. Here are two, expressed in typically informal mathematical language:

Every integer  $n$  that is a multiple of 4 is even.  
Some real number has square 2.

These statements are, in other words:

For every integer  $n$ , if  $n$  is a multiple of 4, then  $n$  is even.  
There is a real number  $x$  such that  $x^2 = 2$ .

The intended meanings of these statements are:

For every  $n \in \mathbb{Z}$ , if  $n$  is a multiple of 4, then  $n$  is even.  
There exists  $x \in \mathbb{R}$  such that  $x^2 = 2$ .

And the latter two statements, in turn, are understood to mean:

For every  $n$ , if  $n \in \mathbb{Z}$ , then: if  $n$  is a multiple of 4, then  $n$  is even.  
There exists  $x$  such that  $x \in \mathbb{R}$  and  $x^2 = 2$ .

Using the final two abbreviations in Table 0.2, we may express these statements more formally, using quantifiers, as:

$$(\forall n \in \mathbb{Z})(n \text{ is a multiple of } 4 \implies n \text{ is even}).$$

$$(\exists x \in \mathbb{R})(x^2 = 2).$$

Suppose you were asked to prove the first of these two statements. To do so, you would, of course, have to know the precise meaning of ‘ $n$  is a multiple of 4’ and ‘ $n$  is even’. By definition,  $n$  is even when  $n = 2k$  for some integer  $k$ , that is:

$$(\exists k \in \mathbb{Z})(n = 2k)$$

According to the rules for forming terms and formulas, terms are of two types—those of the form  $(\iota x)(P)$  and those that are single letters. A term of the form  $(\iota x)(P)$  represents a specific object; such a term is like a proper noun (‘Neil Armstrong’) or a definite description (‘the first human to walk on the Moon’) in ordinary English. A term that is a single letter—also called a **variable**—ambiguously represents an unspecified object; a variable is like a common noun (‘a person’, ‘a thing’) or a pronoun (‘she’, ‘it’).

When a formula or term is, or includes, a string of the form  $(\forall x)(P)$ ,  $(\exists x)(P)$ , or  $(\iota x)(P)$ , then the letter  $x$  is sometimes called a *dummy variable*. This terminology is used to indicate that, without the meaning being altered, the variable could be replaced by any other letter not already occurring in that formula or term. Thus, the two formulas

$$\text{For every } x \in \mathbb{N}, x^2 \geq 0$$

$$\text{For every } n \in \mathbb{N}, n^2 \geq 0$$

although different, say in effect the same thing (namely, that the square of each natural number is nonnegative). We stipulate that in such a case, one is true exactly when the other is true. Likewise, the terms

$$(\iota y)(\forall x)(x \in y \iff x \neq x)$$

$$(\iota S)(\forall y)(y \in S \iff y \neq y)$$

$$(\exists y)(\forall x)(x \in y \iff x \neq x)$$

$$(\forall y)(\forall x)(x \in y \iff x \neq x)$$

are different but represent the same object, namely, the empty set  $\emptyset$ . We stipulate that in such a case, the two objects are equal.

The situation with dummy variables in logic is akin to the one in calculus, where the letter  $x$  in  $\int_0^1 e^{-x^2} dx$  is a dummy variable that could be replaced by any other letter (except  $e$  and  $d$ ):  $\int_0^1 e^{-x^2} dx = \int_0^1 e^{-t^2} dt$ . But it is different from the situation with limits, where there is a distinction between limit of a function of a real variable  $x$ , on the one hand, and the limit of a sequence defined in terms of a discrete, integer-valued variable  $n$ , on the other hand. The convention in calculus is that letters such as  $x$  and  $y$  and  $t$  refer to real variables, whereas letters such as  $m$  and  $m$  and  $k$  refer to integers. Thus the statement  $\lim_{x \rightarrow \infty} 2x/(1+x) = 2$  refers to the limiting behavior of the function  $f(x) = 2x/(1+x)$ , defined for all real  $x \neq -1$ , whereas the statement  $\lim_{n \rightarrow \infty} 2n/(1+n) = 2$  refers to the limiting behavior of the sequence  $2/2, 4/3, 6/4, 8/5, 10/6, \dots$  (Of course the formula for the limit of the function of the real variable  $x$  implies the formula for the limit of the sequence involving  $n$ .)

In various realms of mathematical discourse, it is not unusual to adopt conventions about the domains of different sets of letters just like those used in calculus. In this book, however, we shall generally avoid such conventions and, instead, explicitly state the domain of the variables we use. For example, we would *not* write  $(\forall n)(n^2 \geq 0)$  to suggest, “by agreement”, that the  $n$  is restricted to integers. Instead, we write  $(\forall n \in \mathbb{Z})(n^2 \geq 0)$ .

To deal with formulas involving dummy variables, some notation helps. If  $x$  is a letter that occurs in a string  $P$  and that  $S$  is a string, then

$$P[x \rightarrow S]$$

will mean the string obtained by replacing each occurrence of  $x$  in  $P$  by  $S$ . For example, if  $P$  is the formula  $(\forall x \in \mathbb{N})(x^2 \geq 0)$  and  $S$  is the letter  $y$ , then  $P[x \rightarrow y]$  is the formula  $(\forall y \in \mathbb{N})(y^2 \geq 0)$ .

A formula in our sense is like a declarative sentence in ordinary language. Thus, ‘ $0 < 1$ ’ and ‘The empty set is a subset of  $X$ ’ are (informal renderings of) formulas, whereas the command ‘Solve  $x^2 = 1$ ’ and the question ‘Is  $\pi^e > e^\pi$ ?’ are not.

Formulas, like terms, are of two types:

- A *closed* formula is one that contains no free letters. A closed formula may make a definite assertion about the specific objects named in it; for example, ‘The empty set  $\emptyset$  is an element of the one-element set  $\{\emptyset\}$ ’. Or, a closed formula may make a definite assertion about all objects; for example, ‘For every  $X$ , the empty set is a subset of  $X$ ’.
- An *open* formula is one that contains at least one free letter. For example, ‘The empty set is a subset of  $X$ ’. An open formula makes an actual statement only when the free letters in it are replaced by terms that are not letters. For example, replacing  $X$  by  $\emptyset$  in ‘The empty set is a subset of  $X$ ’ yields the formula ‘The empty set is a subset of  $\emptyset$ ’, which is no longer open and so is a statement.

A closed formula is sometimes called a **statement**, or **sentence**. It makes sense to ask whether a statement is true. However, we shall soon become sloppy about this usage and refer even to open formulas such as  $x = x$  as “statements”.

To emphasize that a formula  $P$  is open because the variable  $x$  occurring in it is free, we sometimes write  $P(x)$  and then call the formula a **predicate in  $x$** . For example,  $(\exists y)(x \in y)$  is a predicate in  $x$ ; but it is *not* a predicate in  $y$  (because  $y$  is bound in the whole formula). Similarly, we can have a predicate  $P(x, y)$  in two variables, such as  $x \in y \vee y \in x$ .

It does not seem to make sense to ask whether an open formula is true, because of the ambiguity of its free variables. For example, if  $P(x)$  is the predicate  $x \in \mathbb{R} \wedge x > 0$ , then  $P(x)$  is true when the variable  $x$  is replaced by the term  $\sqrt{2}$  but false when  $x$  is replaced by  $-1$ .



Such examples suggest two of the commonly used **definitions**—abbreviations, in effect—shown in Table 0.2. These definitions provide another way to shorten formulas and terms.

Table 0.2: Some common logical abbreviations.

table:four-logic-abbrevs

Abbreviation	Stands for	Meaning
$x \neq y$	$\neg(x = y)$	$x$ is not equal to $y$
$x \notin A$	$\neg(x \in A)$	$x$ is not an element of $A$
$(\forall x \in X)(P)$	$(\forall x)(x \in X \implies P)$	for all $x$ in $X$ , $P$
$(\exists x \in X)(P)$	$(\exists x)((x \in X) \ \& \ (P))$	there exists $x$ in $X$ such that $P$

Ways to express “such that”: such that, provided that, when, where, satisfying, subject to the condition, . . .

EXERCISES FOR SECTION 0.1

1.
2. (a)  
(b)
3.

0.2 Sets

sec:sets

Intuitively, a **set** is a collection of mathematical objects. If  $x$  is one of the objects comprising a set  $X$ , then we write

$$x \in X$$

and say that  $x$  is an **element**, **member**, or **point of**  $X$ , and that  $x$  **belongs to**  $X$ ; in the contrary case we write

$$x \notin X.$$

For example, if  $\mathbb{Z}$  is the set of all integers, then  $-5 \in \mathbb{Z}$  but  $1/2 \notin \mathbb{Z}$ .

Two sets  $X$  and  $Y$  are **equal** to one another—in symbols

$$X = Y$$

—precisely when they have the same elements, that is, when

{eq:def-equal-sets} (\*) 
$$(\forall x)(x \in X \iff x \in Y).$$

When it is not the case that  $X = Y$ , we write

$$X \neq Y$$

(Similar use of a slash mark to negate a relational symbol will be made in the future without further explanation.) For an illustration of this notation, consider the set  $\mathbb{Q}$  of all rational numbers along with the set  $\mathbb{Z}$  of all integers. Then  $\mathbb{Q} \neq \mathbb{Z}$  because, for example,  $1/2 \notin \mathbb{Z}$  whereas  $1/2 \in \mathbb{Q}$ .

set!list of elements@anti of elements

set-builder notation

subsec.build-sets

set!set-builder notation@and set-builder notation

predicate

## Building sets

Two notational devices are used to specifying particular sets (or particular forms of sets). The first simply lists the elements of the set between braces (that is, “curly brackets”). For example,

$$\{-1, 1\}$$

is the set having just the two elements  $-1$  and  $1$ . Thus  $-1 \in \{-1, 1\}$  and  $1 \in \{-1, 1\}$ , but  $0 \notin \{-1, 1\}$ . And

$$\{2, 4, 6, \dots\}$$

is the set of all even positive integers. Since the latter set is infinite (see [Definition 0.36](#)), its elements cannot all be listed explicitly, but their identity is supposed to be implicit in the few actually listed in conjunction with the context. In a moment we shall see a notational device for making explicit just what the elements of this set are. In an example like this, however, we may indicate the form of the “general” element of the set:

$$\{2, 4, 6, \dots, 2n, \dots\}$$

The second notational device uses the **set-builder** (or “set-descriptor”) **notation**

$$\{x : P(x)\},$$

read as “the set of all  $x$  **such that**  $P(x)$ ,” to specify the set consists of those objects  $x$  that have a given property  $P(x)$ . [Strictly speaking,  $P(x)$  is a *predicate* in  $x$  in the sense of [page 4](#).]

For example, if  $\mathbb{R}$  denotes the set of all real numbers, then

$$\{x : x \in \mathbb{R}, x^2 = 1\} = \{-1, 1\}.$$

(On the left-hand side of that equation, the comma is a substitute for the connective ‘and’.) This set, which consists of those elements of the set  $\mathbb{R}$  that satisfy a certain condition, may also be specified by the modified notation

$$\{x \in \mathbb{R} : x^2 = 1\},$$

which is read as “the set of all  $x$  in  $\mathbb{R}$  such that  $x^2 = 1$ .” More generally, if  $A$  is some set and  $P(x)$  is some property, then

$$\{x \in A : P(x)\}$$

is the set consisting of those  $x$  that are elements of  $A$  and have the property  $P(x)$ .

A variant form of set-builder notation specifies the form of a typical element to the left of the “such that” colon. For example, if  $\mathbb{N}^* = 1, 2, 3, \dots$  denotes the set of all positive integers, then

$$\{2n : n \in \mathbb{N}^*\},$$

is the set of all even positive integers—the same set denoted more informally earlier by  $\{2, 4, 6, \dots, 2n, \dots\}$ .

Some mathematicians prefer a vertical bar ( $|$ ) instead of a colon ( $:$ ) to mean ‘such that’ in set-builder notation, as in,

$$\{x \in \mathbb{R} \mid x^2 = 1\}.$$

This alternative notation can be confusing, though, when the property  $P$  involves absolute values, as in

$$\{x \in \mathbb{R} \mid |x| < 1\}.$$

(The use of a colon for ‘such that’ can itself be problematic when the objects constituting the set are functions (see [Section 0.3](#)), as in

$$\{f: \mathbb{R} \rightarrow \mathbb{R} : f(1) = 0\},$$

and in such a case the vertical bar may indeed be preferable:

$$\{f: \mathbb{R} \rightarrow \mathbb{R} \mid f(1) = 0\}$$

Some mathematicians use (square) brackets instead of braces to form sets, as in  $[-1, 1]$  and  $[x \in \mathbb{R} : x^2 = 1]$ ; and others use angle-brackets, as in  $\langle -1, 1 \rangle$  and  $\langle x \in \mathbb{R} : x^2 = 1 \rangle$ . Unfortunately, brackets in set-descriptor notation are too easily confused with brackets denoting a closed interval ([subsection “Intervals”](#)), while angle-brackets are often used for generalizations of the dot-product in Euclidean space.

**Caution!** The set-builder notation  $\{x : P(x)\}$  must not be used wantonly. See [Exercise 25](#) for the sort of difficulty that may arise if it is.

## Special sets

subsec:special-sets

If  $x$  is an object, then the **singleton**

$$\{x\}$$

is the set having the lone member  $x$ ; the set  $\{x\}$  is just as different from the object  $x$  as a caged lion is from a loose lion!

If  $x$  and  $y$  are objects with  $x \neq y$ , then the **doubleton**  $\{x, y\}$  is the set having  $x$  and  $y$  as its sole members. (Notice that  $\{x, y\}$  is actually a singleton in case  $x = y$ .) A set is said to be **nondegenerate** when it has at least two different members.

The **empty set** is the set  $\emptyset$  having no members at all. Thus

$$\{x : x \neq x\} = \emptyset = \{x \in \mathbb{R} :: x < x\},$$

etc. Any other set is **nonempty**.

Some sets of numbers for which we reserve special notation are:

$\mathbb{N}$  = the set of all natural numbers =  $\{0, 1, 2, 3, \dots\}$ ,

$\mathbb{Z}$  = the set of all integers =  $\{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$ ,

$\mathbb{Q}$  = the set of all rational numbers =  $\{m/n : m, n \in \mathbb{Q}, n \neq 0\}$ ,

$\mathbb{R}$  = the set of all real numbers,

$\mathbb{C}$  = the set of all complex numbers =  $\{a + bi : a, b \in \mathbb{R}\}$ ,

$I$  = the (closed) unit interval =  $[0, 1] = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$ ,

$J$  = the (closed) *centered* interval  $[-1, 1] = \{x \in \mathbb{R} : -1 \leq x \leq 1\}$

Notice the different typefaces: double-struck (“blackboard bold”) letters  $\mathbb{N}$ ,  $\mathbb{Z}$ , etc., for the various number systems; boldface letters  $I$  and  $J$  for the special subsets of  $\mathbb{R}$ .

singleton  
doubleton  
nondegenerate set  
empty set  
natural numbers  
real numbers  
integers  
rational numbers  
complex numbers

nonnegative  
positive  
containment  
subset  
superset

To omit the number 0 from one of those special sets, we adorn its name with an asterisk ‘\*’. For example,

$\mathbb{N}^*$  = the set of all positive natural numbers =  $\{1, 2, 3, \dots\}$ ,

$\mathbb{R}^*$  = the set of all nonzero real numbers.

And in the case of the “ordered” sets  $\mathbb{Q}$  and  $\mathbb{R}$ , to form the set of just those elements that are *nonnegative*, we adorn its name with a superscript plus-sign ‘+’. Thus

$\mathbb{Q}^+$  = the set of all *nonnegative* rational numbers =  $\{q \in \mathbb{Q} : q \geq 0\}$ ,

$\mathbb{R}^+$  = the set of all *nonnegative* real numbers =  $\{x \in \mathbb{R} : x \geq 0\}$ .

In terms of standard interval notation, then  $\mathbb{R}^+ = [0, \infty[$ : see the subsection “Intervals” (page 70). (We have no need for the name  $\mathbb{Z}^+$  since it would be the same set as  $\mathbb{N}$ .)

Notice the distinction between a real number  $x$  being **nonnegative**, that is,  $x \geq 0$ , on the one hand, and being (strictly) **positive**, that is,  $x > 0$ , on the other hand.

In printed mathematics, it used to be common to see boldface letters ( $\mathbf{N}$ ,  $\mathbf{R}$ , etc.) to designate these special sets. But how can one write boldface letters on paper with pen or pencil, or on a blackboard with chalk? At some point, mathematicians just wrote ordinary upper-case letters ( $N$ ,  $R$ , etc.) and then added extra vertical strokes to them. The “blackboard bold” letters ( $\mathbb{N}$ ,  $\mathbb{R}$ , etc.) used here and commonly elsewhere in print, are stylized forms of those extra-stroke letters.

## Subsets

subsec:subsets

We say that a set  $Y$  is **contained in** a set  $X$ , call  $Y$  a **subset of**  $X$ , and write

$$Y \subset X$$

to mean that each element of  $Y$  is an element of  $X$ , that is,

$$(**) \quad (\forall x) (x \in Y \implies x \in X).$$

We also write

$$X \supset Y$$

to mean the same thing and then say that  $X$  **contains**  $Y$  and call  $X$  a **superset of**  $Y$ .

Examples of subset containment are the following relations among the special sets introduced above:

$$\begin{aligned} \mathbb{N} &\subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}, \\ \mathbb{N}^* &\subset \mathbb{N}, \quad \mathbb{Z}^* \subset \mathbb{Z}, \quad \text{etc.}, \\ \mathbb{N}^* &\subset \mathbb{Z}^* \subset \mathbb{Q}^*, \quad \text{etc.}, \\ \mathbb{I} &\subset \mathbb{R}. \end{aligned}$$

Moreover:

$$\mathbb{R} \supset \mathbb{I},$$

but

$$\mathbb{I} \not\subset \mathbb{Q}.$$

If  $x$  is an object, then

$$\{x\} \subset X \iff x \in X.$$

**The empty set is a subset of every set  $X$ :**

$$\emptyset \subset X.$$

(*Proof:* Since  $\emptyset$  has no elements at all, it does not have any element that fails to be an element of  $X$ .)

A set is nondegenerate if and only if it has some doubleton as a subset.

From properties (\*), page 7, and (\*\*), page 10, evidently

$$X = Y \iff X \subset Y \text{ and } Y \subset X.$$

Thus the inclusion  $Y \subset X$  does not preclude the possibility that  $X = Y$ . When  $Y \subset X$  but  $Y \neq X$ , we call  $Y$  a **proper** subset of  $X$  and write  $Y \subsetneq X$ .

**Caution!** Some authors denote by  $Y \subseteq X$  what we denote by  $Y \subset X$ , and then they reserve the notation  $Y \subset X$  to mean what we denote by  $Y \subsetneq X$ .

The relation  $\subset$  between sets may be called “subset inclusion.” And when  $X \subset Y$ , we may say that  $Y$  **includes**  $X$ , and that  $X$  is **included in**  $Y$ . Unfortunately, it is common to use the word ‘include’ in the sense of element membership, as in “1 is included in  $\mathbb{N}$ .” Note that set  $X$  may simultaneously be an element of another set  $Y$  and a subset of  $Y$ —see Exercise 9.

## Mathematical induction

subsec:induction

The language of subsets may be used to state the principle underlying “proof by induction.”

axiom:induction **0.1 Axiom of Induction.** Let  $E$  be a subset of the set  $\mathbb{N}$  of natural numbers. Suppose:

- $0 \in E$ ; and
- $n + 1 \in E$  whenever  $n \in E$ .

Then  $E = \mathbb{N}$ .

The axiom of induction, stated above in terms of a subsets of  $\mathbb{N}$ , may be reformulated in the following equivalent form, in terms of properties of natural numbers.

thm:principle-induction **0.2 Principle of Mathematical Induction.** Let  $P(n)$  be a predicate in  $n$ . If

- $P(0)$ ; and
- $(\forall n \in \mathbb{N}) (P(n) \implies P(n + 1))$ ,

then

$$(\forall n \in \mathbb{N}) (P(n)).$$

To deduce this reformulation from the axiom, take  $E = \{n \in \mathbb{N} : P(n)\}$ .

In the sequel we shall say simply that we are providing a “proof by induction” to mean that we are using the Principle of Mathematical Induction. In such a proof, verifying  $P(0)$  is the **base step**, and deducing that  $P(n) \implies P(n + 1)$  is the **inductive step**.

To illustrate proof by induction, let us prove:

$$(*) \qquad 2^n > n \qquad (n \in \mathbb{N})$$

Base step. First,  $2^0 = 1 > 0$ .

proper subset  
subset!proper  
inclusion  
mathematical induction  
proof by induction  
base step  
inductive step

**Inductive step.** We wish to prove that, for every  $n \in \mathbb{N}$ , if  $2^n > n$ , then  $2^{n+1} > n + 1$ .  
 Let  $n \in \mathbb{N}$  be arbitrary. Assume that  $2^n > n$ . Then  
 $2^{n+1} = 2 \cdot 2^n > 2 \cdot n = n + n \geq n + 1$ ,  
 and so  $2^{n+1} > n + 1$ .  $\square$

The symbol  $\square$  in the line above indicates the end of the proof.

Like most proofs by induction, the preceding one is written informally, without explicit mention of the predicate  $P(n)$ . The predicate implicitly involved there is just “ $2^n > n$ .”

The fundamental properties of the set  $\mathbb{N}$  of natural numbers include that the natural number 0 is not strictly less than any other natural number, in other words, that 0 is the least element of  $\mathbb{N}$ . That property generalizes from  $\mathbb{N}$  itself to arbitrary nonempty subsets of  $\mathbb{N}$ ; the generalization is a consequence of the Axiom of Induction (0.1).

thm:wo-principle

**0.3 Well-ordering Principle.** Each nonempty subset of  $\mathbb{N}$  has a least element.

The preceding theorem is so-named because, more generally, a set having an ordering relation is said to be “well-ordered” when each of its nonempty subsets has a least element. Well-ordering in general is discussed in Section 0.10.

**Proof.** Let  $A$  be a nonempty subset of  $\mathbb{N}$ . Then there is some natural number  $n \in A$ , and so a fortiori there is some natural number  $\leq n$  in  $A$ . Accordingly, it suffices to prove that, for each natural number  $n$ :

Every subset of  $\mathbb{N}$  that includes some natural number  $\leq n$  has a least element.

And to prove that, we use induction on  $n$ .

**Base step:**  $n = 0$ . A subset of  $\mathbb{N}$  that includes some natural number  $\leq 0$  in fact includes 0 and therefore has 0 as its least element.

**Inductive step.** Now let  $n \geq 0$  and assume that every subset of  $\mathbb{N}$  that includes some natural number  $\leq n$  has a least element. Let  $A$  be a subset of  $\mathbb{N}$  that includes some natural number  $\leq n + 1$ . If in fact  $A$  includes some natural number  $\leq n$ , then by the inductive assumption  $A$  has a least element.

So suppose that  $A$  does not include any natural number  $\leq n$ . Then  $n + 1 \in A$ , and  $n + 1$  is the least element of  $A$ .  $\square$

**0.4 Example.** One nice application of Well-ordering Principle is a proof that  $\sqrt{2}$  is irrational. Recall that  $\sqrt{2}$  is the unique positive real number whose square is 2. (For a proof that such a real number exists, see Exercise 131.) Note that  $1 < \sqrt{2} < 2$  (because  $1^2 < 2 < 2^2$ ).

Just suppose that  $\sqrt{2}$  is rational, so that there is a positive integer  $k$  for which  $k\sqrt{2}$  is also a positive integer. Then there is a least such  $k$ ; denote it by  $n$ . Now the number  $m = n\sqrt{2} - n$  is also a positive integer, and  $m\sqrt{2} = 2n - n\sqrt{2}$  is a positive integer. But this is impossible since  $m < n$ .  $\diamond$

It so happens that the Well-ordering Principle is in fact logically equivalent to the Axiom of Induction (0.1). The proof that the Well-ordering Principle implies the statement in the Axiom of Induction is even easier than the preceding proof. (See Exercise 24.)

subsec:power-set

**Power sets**

The **power set** of a given set  $X$  is the collection

$$\mathcal{P}(X) = \{A : A \subset X\}$$

consisting of all the subsets of  $X$ . For example,

$$\mathcal{P}(\{0, 1\}) = \{\emptyset, \{0\}, \{0, 1\}\}.$$

In general,

$$\emptyset \in \mathcal{P}(X), \quad X \in \mathcal{P}(X)$$

for each set  $X$ .

A **collection** of sets is just a set whose elements are themselves sets. In particular, given a set  $X$ , a set  $\mathcal{A}$  consisting of subsets of  $X$  is often referred to as a **collection of subsets of  $X$** . Thus  $\mathcal{A}$  is a collection of subsets of  $X$  exactly when  $\mathcal{A} \subset \mathcal{P}(X)$ .

For example, let  $E = \{\dots, -6, -4, -2, 0, 2, 4, 6, \dots\}$  be the set of all even integers, let  $F = \{\dots, -12, -8, -4, 0, 4, 8, 12, \dots\}$  be the set of all integers that are multiples of 4, and let  $Q = \{\dots, -15, -10, -5, 0, 5, 10, 15, \dots\}$  be the set of all integers that are multiples of 5. Then  $\mathcal{S} = \{E, F, Q\}$  is a collection of subsets of  $\mathbb{Z}$ .

Typically, we use lower-case letters  $a, b, c, \dots, t, s, x, y, z$  to denote elements of a set; upper-case letters  $A, B, C, \dots, X, Y, Z$  to denote sets consisting of such elements; and script letters  $\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots, \mathcal{T}, \mathcal{U}, \mathcal{V}, \mathcal{W}$  to denote collections of subsets of such sets.

The term *class* is sometimes used to refer to a totality of collections of subsets of one or more sets, as in “the class of all power sets of subsets of  $\mathbb{N}$ .” However, the same term **class** has a more general meaning, referring to *any* totality of sets. Technically speaking, a **set** is defined to be a class that is a member of some class; and a class that is not a set—for example, the class of all sets or the class of all groups—is called a **proper class**. Thus a proper class is a totality of so many sets that it is “too big” to be itself a set. In a careful development of set theory from axioms, we would need to be scrupulous about the distinction between classes in general and sets in particular. In such a development, we would never describe a class  $\{x : P(x)\}$  as a set unless we knew it was actually a member of some class (compare the footnote on page 15).

power set  
collection  
class  
class  
set  
proper class  
class!proper  
set-builder notation  
union!two sets@of two sets  
union!two sets@of two sets  
intersection!two sets@of two sets  
intersection!two sets@of two sets  
disjoint sets  
intersecting sets

**Union and intersection of two sets**

subsec:union-intersect-two

Set  $A$  and  $B$  be sets. The **union of  $A$  and  $B$**  is the set

$$A \cup B = \{x : x \in A \text{ or } x \in B\}$$

consisting of all those objects that belong to *at least one* of the sets  $A$  and  $B$ . And the **intersection of  $A$  and  $B$**  is the set

$$A \cap B = \{x : x \in A \text{ and } x \in B\}$$

consisting of all those objects that belong to *both* of the sets  $A$  and  $B$ . For example,  $\{2, 3, 4\} \cup \{2, 4, 6, 8\} = \{2, 3, 4, 6, 8\}$  and  $\{2, 3, 4\} \cap \{2, 4, 6, 8, 10\} = \{2, 4\}$ .

The set  $A$  is **disjoint from  $B$**  when

$$A \cap B = \emptyset,$$

that is, when  $A$  and  $B$  have no elements in common whatsoever; but  $A$  **intersects  $B$**  in the contrary case, namely, when at least one object is an element of both  $A$  and  $B$ . For example, the set  $\{x \in \mathbb{R} : 0 \leq x \leq 1\}$  is disjoint from  $\{x \in \mathbb{R} : x > 1\}$  but intersects  $\{x \in \mathbb{R} : x > 1/2\}$ .

complement

De Morgan, Augustus

De Morgan's laws

complement!union@of union

complement!intersection@of intersection

union!two sets@of two sets

intersection!two sets@of two sets

ordered pair

Here are some handy general formulas concerning union and intersection:

$$A \cap B \subset A \subset A \cup B,$$

$$A \cap A = A = A \cup A,$$

$$A \cap \emptyset = \emptyset,$$

$$A \cup \emptyset = A,$$

$$A \cap B = B \cap A,$$

$$A \cup B = B \cup A,$$

$$A \cup (B \cup C) = (A \cup B) \cup C,$$

$$A \cap (B \cap C) = (A \cap B) \cap C,$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C),$$

$$A \cap B = A \iff A \subset B \iff A \cup B = B.$$

## Complements

subsec:complement

For sets  $A$  and  $X$ , the **complement of  $A$  in  $X$**  is the set

$$X \setminus A = \{x \in X : x \notin A\}$$

consisting of those elements of  $X$  that do *not* belong to  $A$ .

For any sets  $A$  and  $X$ :

$$X \setminus \emptyset = X,$$

$$X \setminus X = \emptyset,$$

$$A \cup (X \setminus A) = X,$$

$$A \cap X \setminus A = \emptyset,$$

$$X \setminus (X \setminus A) = A.$$

If  $A$  and  $B$  are subsets of  $X$ , then

$$A \subset B \iff X \setminus B \subset X \setminus A.$$

eq:DeMorgans-laws

For any sets  $X$ ,  $A$ , and  $B$ ,

$$X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B),$$

$$X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B).$$

These two **De Morgan's Laws**, named after Augustus De Morgan, will be generalized in [Section 0.4](#).

## Pairs and products

subsec:ordered-pairs

The **ordered pair**  $\langle x, y \rangle$  formed from objects  $x$  and  $y$  is a new object in which  $x$  is the **first coordinate** and  $y$  is the **second coordinate**. Equality of ordered pairs is governed by the rule

$$\langle x, y \rangle = \langle a, b \rangle \iff x = a \quad \text{and} \quad y = b.$$

[It is interesting, but unnecessary for our needs, to know that  $\langle x, y \rangle$  may be defined as the set  $\{\{x\}, \{x, y\}\}$ . Then the preceding rule about equality of ordered pairs may be deduced from this definition.]

Observe that the ordered pair  $\langle x, y \rangle$  is *not* the same thing as the set  $\{x, y\}$ . Indeed, whereas

$$\{x, y\} = \{y, x\}$$

for every  $x$  and  $y$ ,

$$x \neq y \implies \langle x, y \rangle \neq \langle y, x \rangle.$$

The **(Cartesian) product** of sets  $X$  and  $Y$  is the set

$$X \times Y = \{\langle x, y \rangle : x \in X, y \in Y\}$$



of all ordered pairs whose first coordinate belongs to  $X$  and whose second coordinate belongs to  $Y$ . For example, if  $X = \{1/3, 2/7\}$  and  $Y = \{\pi, \sqrt{2}, e\}$ , then

$$X \times Y = \{\langle 1/3, \pi \rangle, \langle 1/3, \sqrt{2} \rangle, \langle 1/3, e \rangle, \langle 2/7, \pi \rangle, \langle 2/7, \sqrt{2} \rangle, \langle 2/7, e \rangle\};$$

and if instead  $Y = \mathbb{R}$ , then  $X \times Y$  is the subset of the plane consisting of the two vertical lines with equations  $x = 1/3$  and  $x = 2/7$ , respectively.

For arbitrary sets  $X$  and  $Y$ , clearly

$$X \times Y \neq \emptyset \iff X \neq \emptyset \text{ and } Y \neq \emptyset.$$

The notion of an **ordered triple**  $\langle x, y, z \rangle$  may be defined in terms of ordered pairs:  $\langle x, y, z \rangle = \langle \langle x, y \rangle, z \rangle$ . Then  $\langle x, y, z \rangle = \langle a, b, c \rangle$  if and only if  $x = a$ ,  $y = b$ , and  $z = c$ . And the **(Cartesian) product** of sets  $X$ ,  $Y$ , and  $Z$  is the set

$$X \times Y \times Z = \{\langle x, y, z \rangle : x \in X, y \in Y, z \in Z\}$$

To deal with ordered quadruples, quintuples, etc., along with the corresponding Cartesian products of four, five, etc., sets, it is easier to frame definitions in terms of functions—see the subsection “ $n$ -tuples and sequences” (page 35).

## Relations

subsec:relations

A set  $\alpha$  each of whose elements is an ordered pair is called a **(binary) relation**.<sup>1</sup> Given such a relation: the set of all the first coordinates of its pairs is called the **domain** of the relation; and the set of all the second coordinates of its pairs is called the **range** of the relation. A relation whose domain is a set  $X$  and whose range is a *subset* of a set  $Y$  is said to be a **relation in  $X$  to  $Y$** . In other words, a relation in  $X$  to  $Y$  is just a subset of the Cartesian product  $X \times Y$ .

When  $Y = X$ , we call a relation in  $X$  to  $Y$  more simply a **relation in  $X$** .

A given relation  $\alpha$  gives rise to a corresponding two-variable predicate “ $x \alpha y$ ,” where

$$x \alpha y$$

means the very same thing as

$$\langle x, y \rangle \in \alpha.$$

And then we write  $x \not\alpha y$  (with a slash through the  $\alpha$ ) to mean it is *not* the case that  $x \alpha y$ , in other words, it *is* the case that  $\langle x, y \rangle \notin \alpha$ .

Ordinarily we regard a relation  $\alpha$  as being the corresponding two-variable predicate “ $x \alpha y$ .” But when we want to regard a relation  $\alpha$  as a set of ordered pairs, we may refer to it as the **graph of  $\alpha$** .<sup>2</sup>

**0.5 Examples.** (1) For a given set  $X$ , the subset  $\Delta_X$  of the product  $X \times X$  is a relation in  $X$ , called the **diagonal of  $X \times X$** . Thus:

$$x \Delta_X y \iff x = y \quad (x, y \in X)$$

Thus  $\Delta_X$  realizes equality within a given set  $X$  as a relation. The graph of this relation is the set  $\Delta_X$  and the corresponding two-variable predicate is “ $x = y$ .”

fn:set-vs-class

<sup>1</sup>Sometimes it is convenient to relax the requirement that a relation be a set. (Strictly speaking, this is possible only in a system of set theory that allows such mathematical objects as “classes” that, while having sets as their elements, need not be sets themselves.) Later, for example, we shall consider the relation of metric spaces being isometric and the relation of topological spaces being homeomorphic, even though neither the class of all metric spaces nor the class of all topological spaces is a set (just as the class of all sets is not itself a set: see Exercise 25).

<sup>2</sup>Some formulations of set theory—for example, that in Bourbaki [9]—clearly distinguish between a (binary) relation and a graph: namely, a *relation* is a two-variable predicate, whereas a *graph* is a collection of ordered pairs.

product!two sets@of two sets  
relation  
domain!relation@of a relation  
range!relation@of a relation  
graph!relation@of a relation  
relation!graph@and graph  
diagonal

empty relation  
empty relation  
relation!empty  
trivial relation  
relation!trivial  
functional relation  
relation!functional  
ordering relation  
equivalence relation  
functional relation  
usual ordering

(2) For any sets  $X$  and  $Y$ , the empty set  $\emptyset$  is a subset of  $X \times Y$  and hence a relation in  $X$  to  $Y$ . It “relates no elements of  $X$  to any elements of  $Y$ ” and may thus be called the **trivial relation in  $X$  to  $Y$** .  $\diamond$

In the sequel we shall be especially interested in three special types of relations:

- *functional relations*—those relations that are the graphs of functions (see [Definition 0.6](#), below, and [Section 0.3](#));
- *ordering relations*—those relations in a set to itself that allow us to say that some elements of the set “precede” other elements of the set (see [Section 0.7](#)); and
- *equivalence relations*—generalizations of equality that allow us to classify pairs of elements of a set as being either “alike” one another or else not alike one another (see [Section 0.9](#)).

The first of these three special types of relations is defined as follows.

def:functional-relation

**0.6 Definition.** A relation  $\alpha$  is said to be **functional** when no two of its elements have the same *first* coordinate, in other words when, for all  $x$ ,  $y$ , and  $y'$ ,

$$x \alpha y \text{ and } x \alpha y' \implies y = y'.$$

Then for each element  $x$  of the domain of  $\alpha$ , there is a unique element  $y$  in the range of  $\alpha$  for which  $x \alpha y$ , and this  $y$  is called the **value of  $\alpha$  at  $x$**  and is denoted by  $\alpha(x)$ .

Here is an example of each of these types of relations.

exs:types-of-relations

**0.7 Examples.** (1) The graph

$$\alpha = \{\langle x, y \rangle \in \mathbb{R} \times \mathbb{R} : x^2 = y\}$$

of the squaring function for real numbers, given by  $f(x) = x^2$ , is a functional relation. The value  $\alpha(x)$  of this relation at an  $x \in \mathbb{R}$  is precisely the value of the function  $f$  at  $x$  in the usual sense.

Observe that two elements of this functional relation can have the same *first* coordinate: in fact, the ordered pairs  $\langle 3, 9 \rangle$  and  $\langle -3, 9 \rangle$  are both members of  $\alpha$ , that is,  $3 \alpha 9$  and  $-3 \alpha 9$ , even though  $3 \neq -3$ .

By contrast, the relation

$$\beta = \{\langle x, y \rangle \in \mathbb{R} \times \mathbb{R} : x = y^2\}$$

is *not* functional because the ordered pairs  $\langle 9, 3 \rangle$  and  $\langle 9, -3 \rangle$  are both members of  $\beta$ , that is,  $9 \beta 3$  and  $9 \beta (-3)$ .

(2) The **usual ordering** of the set  $\mathbb{Z}$  of all integers, which is the relation  $\alpha$  in  $\mathbb{Z}$  given by

$$\alpha = \{\langle m, n \rangle \in \mathbb{Z} \times \mathbb{Z} : m \leq n\},$$

so that

$$m \alpha n \iff m \leq n \quad (m, n \in \mathbb{Z}).$$

(Here the parenthetical expression on the right means “for every  $m \in \mathbb{Z}$  and for every  $n \in \mathbb{Z}$ ” and qualifies the statement to its left.) Of course we ordinarily denote this relation simply by  $\leq$ .

Similarly we have the **usual ordering**  $\leq$  of the set  $\mathbb{N}$  of natural numbers; the usual ordering of the set  $\mathbb{Q}$  of rational numbers; and the usual ordering of the set  $\mathbb{R}$  of real numbers.

usual ordering  
equivalence relation  
relation!reverse

(3) The relation

$$\alpha = \{\langle m, n \rangle \in \mathbb{Z} \times \mathbb{Z} : m - n = 2k \text{ for some } k \in \mathbb{Z}\}$$

is an equivalence relation in  $\mathbb{Z}$  that classifies two integers as being “alike” when they are either both even or else both odd. The usual notation used for  $m \alpha n$  in number theory is  $m \equiv n \pmod{2}$ .  $\diamond$

Any relation may be “turned around” to form a new one, its *reverse*.

def:reverse-relation

**0.8 Definition.** Given a relation  $\alpha$  in a set  $X$  to a set  $Y$ , the **reverse of  $\alpha$**  is the relation in  $Y$  to  $X$  denoted by  $\alpha^{-1}$  and defined by

$$y \alpha^{-1} x \iff x \alpha y \quad (y \in Y, x \in X).$$

When a relation  $\alpha$  is regarded as a set of ordered pairs, then its reverse  $\alpha^{-1}$  consists of the same ordered pairs but with the coordinates of each pair in the opposite order.

In [Examples 0.7](#), the corresponding reverse relations are as follows:

(1) For real numbers  $y$  and  $x$ , we have  $y \alpha^{-1} x \iff y = x^2$ . The reverse relation  $\alpha^{-1}$  is *not* functional because the set

$$\{\langle y, x \rangle \in \mathbb{R} \times \mathbb{R} : y = x^2\}$$

is not the graph of a function.

(2) For integers  $n$  and  $m$ , we have  $n \alpha^{-1} m \iff m \leq n$ . This reverse relation is a new “order relation.”

(3) The reverse relation  $\alpha^{-1} = \alpha$ , because if  $m - n = 2k$  for an integer  $k$ , then  $n - m = 2(-k)$ .

## EXERCISES FOR SECTION 0.2

prob:sets

4. Write the set  $\{\dots, -5, -3, -1, 1, 3, 5, \dots\}$  explicitly in set-builder form  $\{n \in \_\_ : \_\_\_\_\_\_ \}$ . Then rewrite it in set-builder form  $\{\_\_\_\_\_\_ : k \in \mathbb{Z}\}$ .

5. Given a real number  $\varepsilon > 0$ , find a real number  $\delta > 0$  such that

$$\{x \in \mathbb{R} : |x - 1| < \delta\} \subset \{x \in \mathbb{R} : |(3x - 1) - 2| < \varepsilon\}.$$

6. Let  $A = \{x \in \mathbb{R} : x^2 < 2\}$  and  $B = \{x \in \mathbb{R} : x^3 = 3\}$ .

(a) Compute and draw pictures of the sets  $A \cup B$ ,  $A \cap B$ ,  $\mathbb{R} \setminus A$ , and  $\mathbb{R} \setminus B$ .

(b) Compute  $B \cap \mathbb{Q}$  and  $A \cap \mathbb{Z}$ .

prob:proper-subsets

7. (a) Do there exist two sets each of which is a proper subset of the other?

(b) If  $X$  is a proper subset of  $Y$  and  $Y$  is a proper subset of  $Z$ , must  $X$  be a proper subset of  $Z$ ?

8. Let  $A = \{\emptyset\}$ ,  $B = \{\emptyset, A\}$ , and  $C = \{\emptyset, A, B\}$ .

prob-part:union-intersect-and-induction

prob-part:power-sets-and-principle-of-strong-induction

composite of relations  
diagonal

prob:both-element-and-subset

- (a) Compute the union and intersection of each two of these three sets.
- (b) Compute  $\mathcal{P}(A)$ ,  $\mathcal{P}(B)$ , and  $\mathcal{P}(C)$ .
- (c) Considering  $A$ ,  $B$ , and  $C$  together with all the sets you computed in (a) and (b), determine which are elements of others, which are subsets of others, and which are equal to others.

9. Find two sets  $X$  and  $Y$  such that  $X \in Y$  and  $X \subset Y$ .

10. Establish the *absorption laws*:  $(A \cup B) \cap B = B$  and  $(A \cap B) \cup B = B$ .

11. (a) Exhibit subsets  $A$  and  $B$  of  $\mathbb{R}$  for which  $\mathbb{R} \setminus (A \cup B) \neq (\mathbb{R} \setminus A) \cup (\mathbb{R} \setminus B)$ .

- (b) Exhibit subsets  $A$  and  $B$  of  $\mathbb{R}$  for which  $\mathbb{R} \setminus (A \cap B) \neq (\mathbb{R} \setminus A) \cap (\mathbb{R} \setminus B)$ .

12. For sets  $A$  and  $B$ , prove that  $A = B$  if and only if  $A \cup B = A \cap B$ .

13. Let  $A$  and  $B$  be subsets of a set  $X$ . Prove:

(a)  $A = B \iff X \setminus A = X \setminus B$ .

(b)  $A \subset B \iff A \cap (X \setminus B) = \emptyset$ .

14. (a) Express the intersection  $(A \times B) \cap (C \times D)$  of products as the product of two sets.

- (b) Show by example that a union  $(A \times B) \cup (C \times D)$  of products is not necessarily a product of two sets.

15. For which sets  $X$  and  $Y$ , if any, does  $X \times Y = Y \times X$ ?

prob:integer-gaps

16. (a) Use induction to prove that there is no integer strictly between 0 and 1.

- (b) Prove that, for each  $m \in \mathbb{N}$ , there is no integer strictly between  $m$  and  $m + 1$ .

prob:strong-induction

17. The **Principle of Strong Induction** is:

Let  $E$  be a subset of  $\mathbb{N}$  such that, for each  $n \in \mathbb{N}$ ,

$$\text{if } \{k \in \mathbb{N} : k \leq n\} \subset E, \text{ then } n + 1 \in E.$$

Then  $E = \mathbb{N}$ .

- (a) Deduce the Principle of Strong Induction from the Axiom of Induction (0.1) or from one of its other consequences.

- (b) Formulate a version of this principle in terms of a predicate  $P(n)$ , just as Principle of Mathematical Induction (0.2) is a predicate version of the Axiom of Induction (0.1).

prob:compose-relations

18. For a relation  $\alpha$  in a set  $X$  to a set  $Y$  and a relation  $\beta$  in  $Y$  to a set  $Z$ , the **composite of  $\alpha$  and  $\beta$**  is the relation in  $X$  to  $Z$  denoted by  $\beta \circ \alpha$  and defined by:

$$x (\beta \circ \alpha) z \iff x \alpha y \text{ and } y \beta z \text{ for some } y \in Y$$

- (a) Take  $X = \{0, 1, 2, 3\}$  and let  $\alpha$  and  $\beta$  be the relations in  $X$  given by  $\alpha = \{\langle 0, 0 \rangle, \langle 0, 1 \rangle, \langle 2, 1 \rangle, \langle 3, 3 \rangle\}$  and  $\beta = \{\langle 0, 0 \rangle, \langle 1, 2 \rangle\}$ . Calculate the composites  $\beta \circ \alpha$  and  $\alpha \circ \beta$ . Are they equal?

- (b) Let  $\alpha$  be a relation in a set  $X$  to a set  $Y$ . Prove that  $\alpha \circ \Delta_X = \alpha$  and  $\Delta_Y \circ \alpha = \alpha$ . Here  $\Delta_X$  and  $\Delta_Y$  are the diagonals of  $X \times X$  and  $Y \times Y$ , respectively.

- (c) Prove that composition of relations is associative.

(d) Express the reverse  $(\beta \circ \alpha)^{-1}$  in terms of  $\alpha^{-1}$  and  $\beta^{-1}$ .

19. A subset  $A$  of  $\mathbb{N}$  is said to be “bounded above” when there is some natural number  $b$  such that  $a \leq b$  for every  $a \in A$ .

(a) Give a couple of examples of nonempty subsets of  $\mathbb{N}$  that are bounded above.

(b) The empty set  $\emptyset$  is a subset of  $\mathbb{N}$ . Is it bounded above?

(c) Let  $A$  be a *nonempty* subset of  $\mathbb{N}$  that is bounded above. Prove: There is a least natural number  $b$  making  $A$  bounded above; that is, there is a natural number  $b$  such that  $a \leq b$  for every  $a \in A$  and, if  $b'$  is another natural number such that  $a \leq b'$  for every  $a \in A$ , then  $b < b'$ .

20. Prove: for each positive integer  $k$  and each positive integer  $m$ , there is a positive integer  $n$  for which  $nk > m$ .

21. Let  $n$  be a positive integer and  $m$  an arbitrary integer. Prove that there exist unique integers  $q$  and  $r$  such that  $m = qn + r$  and  $0 \leq r < n$ .

[Hint: For existence, to obtain  $r$  apply the Well-ordering Principle (0.3) to the set  $\{m - kn : k \in \mathbb{Z} \text{ and } m - kn \geq 0\}$ .]

22. Let  $b$  be an integer with  $b > 1$ . Prove:

(a) For each natural number  $a$ , there is some  $m \in \mathbb{N}$  such that  $a < b^m$ .

(b) For each positive integer  $a$  there is a unique  $n \in \mathbb{N}$  such that  $b^n \leq a < b^{n+1}$ .

23. (a) Prove that for each positive integer  $n$  there exists a unique natural number  $n$  and a unique  $(n+1)$ -tuple  $\langle d_0, d_1, d_2, \dots, d_n \rangle$  of integers in  $\{0, 1, 2, \dots, 9\}$  such that  $d_n > 0$  and  $a = \sum_{j=0}^n d_j 10^j$ . [Hint: For existence, apply the Principle of Strong Induction (Exercise 17) on  $n$ . Use Exercises 21 and 22.]

(b) Generalize (a) by replacing 10 with an arbitrary “base” integer  $b > 1$ .

24. Assuming the statement of the Well-ordering Principle (0.3), deduce the statement of the Axiom of Induction (0.1) from it.

25. Let  $X = \{x : x \notin x\}$ . Is  $X \in X$ ? If not, is  $X \notin X$ ?

base expansion!  
positive integers@  
base for expansion  
Axiom of Induction!  
Well-ordering Pri  
Well-ordering Principle!  
Axiom of Ind  
map  
map  
function  
domain of a map  
codomain  
graph  
graph  
map!functional relation@and functi  
functional relation!map@and map  
rule of a map

## 0.3 Maps

### Maps

A **map** (or **function**)

$$f: X \rightarrow Y$$

**from** (or **on**)  $X$  **to** (or **into**)  $Y$  consists of sets  $X$  and  $Y$  together with a “rule”  $f$  that assigns to each  $x \in X$  exactly one element  $f(x) \in Y$  called the **value of  $f$  at  $x$** . The set  $X$  is the **domain**, the set  $Y$  is the **codomain**, and the rule  $f$  is the **graph** of the map. Sometimes when denoting a map we put the name of the rule above the arrow, as in:

$$X \xrightarrow{f} Y$$

In more technical terms, a map  $f: X \rightarrow Y$  is an ordered triple  $\langle X, f, Y \rangle$  in which  $X$  and  $Y$  are sets and the “rule”  $f$  is a *functional relation* in  $X$  to  $Y$  (Definition 0.6)—that is,  $f$  is a

identity map  
graph  
maps (verb)  
sends

carries  
characteristic function  
absolute-value function

relation having the property that for each  $x \in X$  there is exactly one  $y \in Y$  with  $\langle x, y \rangle \in f$ , and then that unique  $y$  is the value  $f(x)$ .

According to that technical definition, the graph  $f$  of a map  $f: X \rightarrow Y$  is the set

$$f = \{\langle x, y \rangle : x \in X \text{ and } y = f(x)\} = \{\langle x, f(x) \rangle : x \in X\}.$$

Thus the term ‘graph’ for the rule of a function is the very same notion as used in elementary mathematics.

For example, given a set  $X$ , the **identity map of  $X$** —the map  $\iota_X: X \rightarrow X$  given by

$$\iota_X(x) = x \quad (x \in X),$$

which sends each element of  $X$  to itself—has as its graph the diagonal  $\Delta_X = \{(x, x) : x \in X\}$  of  $X \times X$ .

Ordinarily, the technical meaning of the “rule” of a map as a relation will not be of concern. In a few situations it will be—for example, in [Exercise 3.124](#) and [Exercise 4.44](#).

Let  $f: X \rightarrow Y$  be a map. For  $x \in X$  and  $y = f(x)$ , we write

$$x \mapsto y$$

—with a vertical bar at the left end of the arrow—and say that  $f$  **sends** or **maps** or **carries  $x$  to  $y$** . When  $f(x)$  is specified by a single formula involving  $x$  for arbitrary  $x \in X$ , we write

$$\begin{aligned} f: X &\rightarrow Y \\ x &\mapsto f(x) \end{aligned}$$

or more compactly,

$$f: X \rightarrow Y: x \mapsto f(x).$$

For example,

$$\begin{aligned} f: \mathbb{R} &\rightarrow \mathbb{R}^+ \\ x &\mapsto x^2 \end{aligned}$$

—written more compactly as

$$f: \mathbb{R} \rightarrow \mathbb{R}^+: x \mapsto x^2$$

—is the map from  $\mathbb{R}$  to  $\mathbb{R}^+$  such that

$$f(x) = x^2 \quad (x \in \mathbb{R}).$$

(Recall that  $\mathbb{R}^+$  denotes the set of all *nonnegative* real numbers.)

Sometimes several formulas specifying  $f(x)$  are needed for various parts of the domain to which  $x$  might belong. For example, if  $A \subset X$ , then the **characteristic function of  $A$  in  $X$**  is the map

$$\chi_A: X \rightarrow \{0, 1\}$$

such that

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \in X \setminus A. \end{cases}$$

In the notation  $f: X \rightarrow Y$ , at times the rule  $f$  is not named explicitly but instead is indicated by a form with a placeholder dot, as for the **absolute-value function**

$$|\cdot|: \mathbb{R} \rightarrow \mathbb{R}: x \mapsto |x|$$

where

$$|x| = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0. \end{cases}$$

And at times the rule is not named at all, as for the “successor function”

$$\mathbb{N} \rightarrow \mathbb{N}: n \mapsto n + 1.$$

According to both the technical definition of ‘map’ in terms of ordered triples and the initial, informal, meaning of the term, two maps

$$f: X \rightarrow Y, \quad g: A \rightarrow B$$

are equal precisely when:

- they have the same domain, that is,  $X = A$ ;
- they have the same rule, that is,  $f(x) = g(x)$  for all  $x \in X$ ; *and*
- **they have the same codomain**, that is,  $Y = B$ .

Thus the two maps

$$\begin{array}{ccc} f: \mathbb{R} \rightarrow \mathbb{R} & \text{and} & g: \mathbb{R} \rightarrow \mathbb{R}^+, \\ x \mapsto x^2 & & x \mapsto x^2 \end{array}$$

which have the very same domain and are defined by the very same rule, are different because their codomains are different. In general, **a domain and a “rule” together do not determine a map: a specific codomain is required as well.**

At times we become a bit sloppy about precise usage of the terminology and refer to a map  $f: X \rightarrow Y$  as just  $f$ . Thus we denote the domain, codomain, and graph of a map  $f: X \rightarrow Y$  simply by  $\text{dom}(f)$ ,  $\text{codom}(f)$ , and  $\text{graph}(f)$ , respectively.

The **range** of a map  $f: X \rightarrow Y$  is the set

$$\{f(x) : x \in X\}$$

of all its values. Thus the range is a subset of the codomain and, as the preceding example indicates, can be a proper subset. A map is **real-valued** when its range is contained in  $\mathbb{R}$ —that is, when each of its values is a real number. A map is **constant** when its range is a singleton  $\{c\}$ , that is, when there is some  $c \in Y$  such that

$$f(x) = c \quad (x \in X).$$

When a map  $f$  has as its domain the product  $X \times Y$  of two sets, it may be called a **function of two variables** (or **of two arguments**). In such a case the value of  $f$  at an element  $\langle x, y \rangle$  of  $X \times Y$  ought to be denoted by

$$f(\langle x, y \rangle).$$

Customarily however, we write more simply:

$$f(x, y).$$

For example, we have the real-valued function  $f$  of two real variables defined by  $f(x, y) = x^2/16 - y^2/9$ . And similarly for a map whose domain is a product of three or more sets.

For a function  $f: X \rightarrow \mathbb{R}$ , a **zero** (or **root**) of  $f$  is an element  $a$  in the domain  $X$  of  $f$  for which  $f(a) = 0$ . For example,  $-\sqrt{2}$  is a zero of the function  $f: \mathbb{R} \rightarrow \mathbb{R}$  defined by  $f(x) = x^2 - 2$ ; and  $1 + i$  is a zero of the function  $f: \mathbb{C} \rightarrow \mathbb{R}$  given by  $f(x) = x^2 - 2x + 2$ .

For a map  $f: X \rightarrow X$  from a set to itself, a **fixed-point** of  $f$  is an element  $p \in X$  for which  $f(p) = p$ . For example,  $-1$  is a fixed-point of the function  $f: \mathbb{R} \rightarrow \mathbb{R}$  defined by  $f(x) = x^2 + 3x + 1$ .

equal maps  
equality!maps@of maps  
graph of map  
domain  
codomain  
range  
real-valued map  
constant map  
zero of a function  
root!function@of a function  
fixed-point

## Operations

Sometimes a function  $f: X \times X \rightarrow X$  of two variables is called an **operation in** (or **on**)  $X$ —more precisely, a **binary operation on  $X$** —and then typically the value of  $f$  at an input pair  $\langle x, y \rangle$  is written in “infix form” as  $x f y$ , with the function’s name placed between its two arguments, rather than in the “prefix form”  $f(x, y)$ . For example, on the set  $\mathbb{N}$  of natural numbers we have the operations

$$+: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}: \langle m, n \rangle \mapsto m + n, \quad \cdot: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}: \langle m, n \rangle \mapsto m \cdot n$$

of **addition** and **multiplication**, respectively. Similarly we have such operations on  $\mathbb{Z}$ ,  $\mathbb{Q}$ ,  $\mathbb{R}$ , and  $\mathbb{C}$ . [For details about  $\mathbb{C}$ , see the subsection “Complex numbers” (page 89).] For multiplication, even the function’s name, ‘ $\cdot$ ’ is usually elided and the multiplication denoted just by juxtaposing the elements, as in  $xy$ ; for clarity in formulation fundamental properties of multiplication below, however, we leave the ‘ $\cdot$ ’ in.

Addition in each of  $X = \mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$ ,  $\mathbb{R}$ , and  $\mathbb{C}$  is associative and commutative and has an identity element:

(A1) Associativity: *For all  $x, y \in X$ ,*

$$(x + y) + z = x + (y + z).$$

(A2) Commutativity: *For all  $x, y \in X$ ,*

$$(x + y) + z = x + (y + z).$$

(A3) Identity element: *There is a unique element  $0 \in X$  such that, for all  $x \in X$ ,*

$$x + 0 = x = 0 + x.$$

Multiplication in each of  $X = \mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$ ,  $\mathbb{R}$ , and  $\mathbb{C}$  is associative and commutative and has an identity element:

(M1) Associativity: *For all  $x, y \in X$ ,*

$$(x \cdot y) \cdot z = x \cdot (y \cdot z).$$

(M2) Commutativity: *For all  $x, y \in X$ ,*

$$(x \cdot y) \cdot z = x \cdot (y \cdot z).$$

(M3) Identity element: *There is a unique element  $1 \in X$  such that  $1 \neq 0$  and, for all  $x \in X$ ,*

$$x \cdot 1 = x = 1 \cdot x.$$

Multiplication in each of  $X = \mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$ ,  $\mathbb{R}$ , and  $\mathbb{C}$  distributes over addition:

(D1) Distributivity: *For all  $x, y, z \in X$ ,*

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z).$$

With respect to addition in each of  $X = \mathbb{Z}$ ,  $\mathbb{Q}$ ,  $\mathbb{R}$ , and  $\mathbb{C}$ —but *not* in  $\mathbb{N}$ —all elements have inverses, that is, negatives:

(A4) Additive inverses: *For each  $x \in X$ , there exists a unique element  $-x \in X$  such that*

$$x + (-x) = 0 = (-x) + x.$$



With respect to multiplication in each of  $X = \mathbb{Q}$ ,  $\mathbb{R}$ , and  $\mathbb{C}$ —but *not* in  $X = \mathbb{N}$  or  $X = \mathbb{Z}$ —nonzero elements have inverses:

mult-recipe (M4) Multiplicative inverses: *For each  $x \in X$  with  $x \neq 0$ , there exists a unique element  $x^{-1} \in X$  such that*

$$x \cdot x^{-1} = 1 = x^{-1} \cdot x.$$

pg-ref:field-end Together, properties (A1)–(A4), (M1)–(M4), and (D1) say that the operations of addition and multiplication in each of  $X = \mathbb{Q}$ ,  $\mathbb{R}$ , and  $\mathbb{C}$  make it a **field**.

In each of  $X = \mathbb{N}$  and  $\mathbb{Z}$ , which are *not* fields, addition and multiplication allow “cancellation”:

add-cancel (A5) Cancellation: *For all  $x, y, z \in X$ ,*

$$x + z = y + z \implies x = y.$$

mult-cancel (M5) Cancellation of nonzero elements: *For all  $x, y, z \in X$ ,*

$$x \cdot z = y \cdot z \quad \text{and} \quad z \neq 0 \implies x = y.$$

The same two cancellation properties (A5) and (M5) hold also in each of the fields  $X = \mathbb{Q}$ ,  $\mathbb{R}$ , and  $\mathbb{C}$ —but as consequences of the field properties (A1)–(A4), (M1)–(M4), and (D1) rather than as separate fundamental properties.

### Restriction and extension

subsec:restrict-extend

The **restriction** of a map  $f: X \rightarrow Y$  to a set  $E \subset X$  is the map

$$f|_E: E \rightarrow Y \\ x \mapsto f(x)$$

obtained by applying the rule  $f$  only to elements of  $E$ ; an alternative notation is  $f|_E$ . (Occasionally we shall also restrict the codomain of a map to a subset of its codomain that contains its range.)

For example, if  $E \subset X$ , then the **inclusion map**

$$E \rightarrow X: x \mapsto x$$

of  $E$  into  $X$ , also denoted by

$$E \hookrightarrow X$$

(with a hooked arrow), is the restriction

$$(\iota_X)|_E: E \rightarrow X: x \mapsto x$$

of the identity map  $\iota_X: X \rightarrow X$ .

multiplicative inverse

operation!multiplicative inverse@and field

operation!cancellation@and cancellation

cancellation

cancellation

operation!cancellation@and cancellation

operation

restriction of map

restriction of map

inclusion map

identity map

extension of map  
 extension of map  
 extends a map  
 restriction of map!domain-codomain  
 domain-codomain restriction  
 restriction of map!codomain  
 codomain restriction  
 restriction of map  
 map!piecewise@defined piecewise  
 piecewise

A map

$$g: Z \rightarrow Y$$

$$f: X \rightarrow Y$$

is called an **extension** of a map  
 to  $Z$  and is said to **extend**  $f$  to  $Z$  when

$$X \subset Z \quad \text{and} \quad g|_X = f,$$

that is, when

$$X \subset Z \quad \text{and} \quad g(x) = f(x) \quad (x \in X).$$

For example, the two maps

$$g_1: \mathbb{R} \rightarrow \mathbb{R}: x \mapsto x^2 \quad \text{and} \quad g_2: \mathbb{R} \rightarrow \mathbb{R}: x \mapsto x^3$$

both extend to  $\mathbb{R}$  the inclusion map  $f: \{0, 1\} \rightarrow \mathbb{R}$  of the two-element set  $\{0, 1\}$  into  $\mathbb{R}$ .

The restriction  $f|_E$  of a map  $f: X \rightarrow Y$  to a subset  $E$  of its domain  $X$ , as defined above, leaves the codomain of the restriction to be the same as that of the original map  $f$ . Sometimes we shall want to cut down the codomain of the restriction so as to be as small as possible, that is, to be the range  $f(E) = \{f(x) : x \in E\}$  of  $f|_E$ ; in this situation we denote the new map by  $f|_{E, f(E)}: E \rightarrow f(E)$  and refer to it as the **domain-codomain restriction of  $f$  to  $E$** . In the special case that  $E = X$ , the original domain, we may refer to the map  $f|_{X, f(X)}$  as the **codomain restriction of  $f$  to its range**.

Often you will see the absolute-value function defined “piecewise” by

$$|x| = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x \leq 0, \end{cases}$$

and it will be noted that “the two formulas agree at the overlap  $x = 0$ .” Actually, we are starting with two separate domains

$$A = \{x \in \mathbb{R} : x \geq 0\}, \quad B = \{x \in \mathbb{R} : x \leq 0\}$$

and two separate functions

$$\begin{array}{ll} f: A \rightarrow \mathbb{R}, & g: B \rightarrow \mathbb{R} \\ x \mapsto x & x \mapsto -x \end{array}$$

that agree on the overlap  $A \cap B = \{0\}$  of their individual domains, that is, such that  $f|_{A \cap B} = g|_{A \cap B}$ . And then the absolute value function allegedly defined by the “piecewise” formulas is the function  $h: A \cup B = \mathbb{R} \rightarrow \mathbb{R}$  that agrees with  $f$  on  $A$  and  $g$  on  $B$ , that is,  $h|_A = f$  and  $h|_B = g$ . That this is a legitimate way to define a function is guaranteed by the following proposition.

prop:extend-two-pieces

**0.9 Proposition (construction of a map from two pieces).** Let

$$f: A \rightarrow Y, \quad g: B \rightarrow Y$$

be maps with common codomain  $Y$  for which

$$f|_{A \cap B} = g|_{A \cap B},$$

that is, for which

$$f(x) = g(x) \text{ for all } x \in A \cap B.$$

Then there exists a unique map

$$h: A \cup B \rightarrow Y$$

such that

$$h|_A = f, \quad h|_B = g,$$

that is,

$$h(x) = \begin{cases} f(x) & \text{if } x \in A, \\ g(x) & \text{if } x \in B. \end{cases}$$

map!piecewise@defined piecewise  
piecewise  
extension of map  
composite of maps  
composite of maps

**Proof.** Existence. Let  $H = \text{graph } f \cup \text{graph } g$ , so that  $H$  is a subset of  $(A \cup B) \times Y$ , that is, a relation in  $A \cup B$  to  $Y$ .

We show that  $H$  is the graph of a map. Suppose  $\langle x, y \rangle$  and  $\langle x, y' \rangle$  are elements of  $H$  having the same first coordinate. We want to show that  $y = y'$ . If  $\langle x, y \rangle \in \text{graph } f$  and  $\langle x, y' \rangle \in \text{graph } f$ , then  $y = y'$  because  $f$  is a map. Similarly,  $y = y'$  if  $\langle x, y \rangle \in \text{graph } g$  and  $\langle x, y' \rangle \in \text{graph } g$ . If  $\langle x, y \rangle \in \text{graph } f$  and  $\langle x, y' \rangle \in \text{graph } g$ , then  $f(x) = y$  and  $g(x) = y'$  with  $x \in A \cap B$  whence  $y = y'$  because  $f|_{A \cap B} = g|_{A \cap B}$ . Similarly,  $y = y'$  if  $\langle x, y \rangle \in \text{graph } g$  and  $\langle x, y' \rangle \in \text{graph } f$ .

Let  $h: A \cup B \rightarrow Y$  be the map whose graph is the relation  $H$ . Then  $h|_A = f$  and  $h|_B = g$ . In fact, if  $x \in A$ , then  $(x, f(x)) \in \text{graph } f \subset H = \text{graph } h$  and so  $h(x) = f(x)$ . Similarly, if  $x \in B$ , then  $h(x) = g(x)$ .

Uniqueness. Suppose  $h, h': A \cup B \rightarrow Y$  are maps both of whose restrictions to  $A$  equal  $f$  and both of whose restrictions to  $B$  equal  $g$ . Let  $x \in A \cup B$ . If  $x \in A$ , then  $h(x) = f(x)$  and  $h'(x) = f(x)$  whence  $h(x) = h'(x)$ . Similarly,  $h(x) = h'(x)$  if  $x \in B$ .  $\square$

The preceding proposition will be generalized to the case of any number of “pieces”: see [Proposition 0.21](#).

### Composition of maps

subsec:composite

For maps

$$f: X \rightarrow Y \quad \text{and} \quad g: Y \rightarrow Z,$$

their **composite** is the map

$$\begin{aligned} g \circ f: X &\rightarrow Z \\ x &\mapsto g(f(x)) \end{aligned}$$

that assigns to each  $x \in X$  the value of  $g$  at the element  $f(x)$  of  $Y$ . *Note that the composite  $g \circ f$  is defined only when the codomain of  $f$  is the same as the domain of  $g$ .*

diagram  
commutative  
commutative diagram

For example, the composite  $g \circ f$  of

$$\begin{array}{ll} f: \mathbb{R} \rightarrow \mathbb{R}^+ & \text{and} \quad g: \mathbb{R}^+ \rightarrow \mathbb{R} \\ x \mapsto e^x & y \mapsto 2y \end{array}$$

is the map

$$\begin{array}{l} g \circ f: \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto 2e^x. \end{array}$$

The composite  $f \circ g$  in the reverse order is

$$\begin{array}{l} f \circ g: \mathbb{R}^+ \rightarrow \mathbb{R}^+ \\ y \mapsto e^{2y}. \end{array}$$

In the preceding example  $f \circ g \neq g \circ f$ ; thus composition of maps is *not* commutative. Composition is, however, associative:

$$h \circ (g \circ f) = (h \circ g) \circ f$$

for maps  $f: X \rightarrow Y$ ,  $g: Y \rightarrow Z$ ,  $h: Z \rightarrow W$ . In fact,

$$\begin{aligned} (h \circ (g \circ f))(x) &= h((g \circ f)(x)) = h(g(f(x))) \\ &= (h \circ g)(f(x)) = ((h \circ g) \circ f)(x) \end{aligned}$$

for all  $x \in X$ .

Three maps  $f: X \rightarrow Y$ ,  $g: Y \rightarrow Z$ ,  $h: X \rightarrow Z$  may be displayed together in a single triangular diagram showing “where the maps are from and where they are to”:

$$\begin{array}{ccc} & Y & \\ f \nearrow & & \searrow g \\ X & \xrightarrow{h} & Z \end{array}$$

The same relationship may be shown by reorienting the triangle, for example:

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ & \searrow h & \downarrow g \\ & & Z \end{array}$$

The triangle is said to **commute** (or to be **commutative**) when

$$h = g \circ f,$$

in other words, when the result  $g(f(x))$  of applying to an arbitrary  $x \in X$  first the map  $f$  and then the map  $g$  on the “path” from  $X$  to  $Z$  passing through  $Y$  is the same as the result  $h(x)$  of applying to  $x$  the map  $h$  on the direct path from  $X$  to  $Z$ .

Similarly, four maps  $f: X \rightarrow Y$ ,  $g: Y \rightarrow W$ ,  $h: X \rightarrow Z$ ,  $k: Z \rightarrow W$  may be displayed on a square diagram:

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ h \downarrow & & \downarrow k \\ Z & \xrightarrow{g} & W \end{array}$$

The square is said to **commute** when

$$g \circ h = k \circ f.$$

A more complicated diagram of maps is said to **commute** when each of its constituent triangles and squares commutes.

For example, consider the product  $X \times Y$  of two sets. The **first** and **second projections** of  $X \times Y$  are the respective maps

$$\begin{aligned} p: X \times Y &\rightarrow X, & q: X \times Y &\rightarrow Y. \\ (x, y) &\mapsto x & (x, y) &\mapsto y \end{aligned}$$

Suppose further that  $f: Z \rightarrow X$  and  $g: Z \rightarrow Y$  are maps having the same domain  $Z$ . Then there is a unique map  $h: Z \rightarrow X \times Y$  such that

$$p \circ h = f, \quad q \circ h = g,$$

in other words, such that the following diagram commutes:

$$\begin{array}{ccccc} & & Z & & \\ & f \swarrow & \downarrow h & \searrow g & \\ X & \xleftarrow{p} & X \times Y & \xrightarrow{q} & Y \end{array}$$

(Note: There the vertical arrow is dashed, not solid, so as to indicate that the existence of  $h$  is at issue.) This map  $h$  is defined by

$$h(x) = (f(x), g(x)).$$

### Images and inverse images

subsec:images

Let  $f: X \rightarrow Y$  be a map. For  $A \subset X$ , the **(direct) image of  $A$  under  $f$**  is the set

$$f(A) = \{f(x) : x \in A\}$$

consisting of all the values that  $f$  takes at the elements of  $A$ . In particular,

$$f(\text{dom}(f)) = \text{range}(f).$$

For  $D \subset Y$ , the **inverse image of  $D$  under  $f$**  is the set

$$f^{-1}(D) = \{x \in X : f(x) \in D\}$$

consisting of all those elements of  $X$  that  $f$  maps to elements of  $D$ . In particular, for  $y \in Y$ , the inverse image  $f^{-1}(\{y\})$  is the set of all those  $x \in X$  at which  $f$  takes the particular value  $y$ . Note that  $f^{-1}(\{y\}) = \emptyset$  unless  $y \in \text{range}(f)$ .

For a point  $y$  in the codomain of  $f: X \rightarrow Y$ , the inverse image  $f^{-1}(\{y\})$  of the singleton  $\{y\}$  is often denoted more simply by  $f^{-1}(y)$ . Thus for  $y \in Y$ ,

$$f^{-1}(y) = \{x \in X : f(x) = y\}.$$

page:def-fiber

This subset of the domain is called the **fiber (of  $X$ ) over  $y$  (for  $f$ )**.

angle-bracket-direct-inverse-image

The angle-bracket notations  $f\langle A \rangle$  and  $f^{-1}\langle D \rangle$  or the square-bracket notations  $f[A]$  and  $f^{-1}[D]$  are sometimes used for direct image  $f(A)$  and inverse image  $f^{-1}(D)$ , respectively. Then the notation  $f^{-1}\langle y \rangle$  or  $f^{-1}[y]$  may be used as an abbreviation for the fiber  $f^{-1}\langle \{y\} \rangle$  over a point  $y$  in the codomain of  $f$ ; this alternative notation disambiguates the meaning of  $f^{-1}(y)$  when the map  $f$  is “bijective”—see [page 28](#) and the cautionary note on [page 30](#).

projection  
diagram  
commutative  
commutative diagram  
composite of maps  
image of a set  
image of a set  
range  
inverse image  
inverse image  
fiber

image of a set!union@and union  
image of a set!intersection@and intersection  
inverse image of a set!union@and union  
inverse image of a set!intersection@and intersection  
the formulas:

$$\begin{aligned} f(A \cup B) &= f(A) \cup f(B), & f(A \cap B) &\subset f(A) \cap f(B), \\ f^{-1}(D \cup E) &= f^{-1}(D) \cup f^{-1}(E), & f^{-1}(D \cap E) &= f^{-1}(D) \cap f^{-1}(E). \end{aligned}$$

image of a set  
inverse image of a set  
injection  
injective map  
injection  
surjection  
surjective map  
surjection  
bijection  
bijection  
bijective map  
one-to-one correspondence.

Finally,

$$f^{-1}(Y \setminus D) = X \setminus f^{-1}(D)$$

for any  $D \subset Y$ . Thus:

- inverse images preserve unions, intersections, and set differences;
- direct images preserve unions but *not* necessarily intersections [see [Exercise 36 \(a\)](#)].

### Injections, surjections, and bijections

subsec:inj-surj-bij

A map  $f: X \rightarrow Y$  is said to be **injective** (or **one-to-one** and is called an **injection** when it assigns distinct values to distinct members of its domain, that is:

$$x_1 \neq x_2 \implies f(x_1) \neq f(x_2) \quad (x_1, x_2 \in X).$$

For example, of the two maps

$$f: \mathbb{R} \rightarrow \mathbb{R}: x \mapsto x^3, \quad g: \mathbb{R} \rightarrow \mathbb{R}: x \mapsto x^2,$$

the first is injective, but the second is not [because, for example,  $g(-1) = g(1)$ ].

A map  $f: X \rightarrow Y$  is said to be **surjective**, is said to map its domain  $X$  **onto** its codomain, and is called a **surjection** when its range equals its codomain, that is, when

$$f(X) = Y,$$

or in more detailed terms, when

$$y \in Y \implies y = f(x) \quad \text{for some } x \in X.$$

In still other words,  $f: X \rightarrow Y$  is surjective when  $f^{-1}(y) \neq \emptyset$  for each  $y \in Y$ . For example, of the two maps

$$f: \mathbb{R} \rightarrow \mathbb{R}^+: x \mapsto x^2, \quad g: \mathbb{R} \rightarrow \mathbb{R}: x \mapsto x^2$$

the first is surjective, but the second is not (because  $-1 \neq x^2$  for any  $x \in \mathbb{R}$ ). Notice that whether a map is surjective depends crucially on its domain and not just on its codomain and the rule applied to elements of that domain.

page:bijection  
def:one-one-corr

A map  $f: X \rightarrow Y$  is said to be **bijective** and is called a **bijection**, and is said to be a **one-to-one correspondence between  $X$  and  $Y$** , when it is both injective and surjective—in other words, when for each  $y \in Y$  there is *one and only one*  $x \in X$  such that  $f(x) = y$ . Thus a bijection effects a pairing, or “matching up”, of all the elements of its domain one by one with all the elements of its codomain.

For example, the doubling map

$$\begin{array}{ccc} \{1, 2, 3, \dots\} & \rightarrow & \{2, 4, 6, \dots\} \\ n & \mapsto & 2n \end{array}$$

is a one-to-one correspondence between the set of all positive integers and the set of all even positive integers. The maps

$$\begin{array}{ccc} f: \mathbb{R} \rightarrow \mathbb{R}, & g: \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto x^3 & y \mapsto \sqrt[3]{y} \end{array}$$

are both bijections from  $\mathbb{R}$  onto  $\mathbb{R}$ . In general, a bijection from a set onto itself is often called a **permutation** of that set and may be regarded as a way of rearranging the elements of the set. For example, the identity map  $\iota_X$  of any set  $X$  is such a permutation; and the rule  $0 \mapsto 3, 1 \mapsto 2, 2 \mapsto 0, 3 \mapsto 1$  defines a permutation of  $\{0, 1, 2, 3\}$ .

permutation

prop:inj-surj-bij-composite

**0.10 Proposition (compositions of injective and surjective maps).** Let  $f: X \rightarrow Y$  and  $g: Y \rightarrow Z$  be maps.

- (1) If  $f$  and  $g$  are both injective, then their composite  $g \circ f$  is also injective.
- (2) If  $f$  and  $g$  are both surjective, then their composite  $g \circ f$  is also surjective.
- (3) If  $f$  and  $g$  are both bijective, then their composite  $g \circ f$  is also bijective.

prop-part:composite-of-bij

**Proof.** (1) Assume that  $f$  and  $g$  are both injective. Let  $x_1, x_2 \in X$  with  $x_1 \neq x_2$ . Then

$$f(x_1) \neq f(x_2)$$

because  $f$  is injective. Hence

$$g(f(x_1)) \neq g(f(x_2))$$

because  $g$  is also injective.

(2) Assume that  $f$  and  $g$  are both surjective. Then

$$(g \circ f)(X) = g(f(X)) = g(Y) = Z.$$

(3) This follows from (1) and (2).  $\square$

Statements (1) and (2) in the preceding proposition have partial converses:

prop:composite-inj-surj-implies-one-is

**0.11 Proposition (injective and surjective compositions).** Let  $f: X \rightarrow Y$  and  $g: Y \rightarrow Z$  be maps.

- (1) If the composite  $g \circ f$  is injective, then  $f$  is injective.
- (2) If the composite  $g \circ f$  is surjective, then  $g$  is surjective.

**Proof.** This is [Exercise 42](#).  $\square$

Those partial converses are often applied in case  $g \circ f$  is an identity map.

**0.12 Corollary.** Let  $f: X \rightarrow Y$  be a map.

- (1) If there exists a map  $r: Y \rightarrow X$  such that  $r \circ f = \iota_X$ , then  $f$  is injective.
- (2) If there exists a map  $s: Y \rightarrow X$  such that  $f \circ s = \iota_Y$ , then  $f$  is surjective.

Both parts of the preceding corollary have converses. See [Exercise 43](#) for the converse of (1); see [Proposition 0.30](#) for the converse of (2).

### Inverse of a bijection

Suppose

$$f: X \rightarrow Y$$

is a bijection. Then a new map

$$f^{-1}: Y \rightarrow X,$$

called the **inverse of  $f$** , is obtained by defining,

$$f^{-1}(y) = \text{the unique } x \in X \text{ such that } f(x) = y$$

for each  $y \in Y$ . Thus

$$f(x) = y \iff f^{-1}(y) = x \quad (x \in X, y \in Y).$$

Hence

$$f^{-1}(f(x)) = x \quad (x \in X), \quad f(f^{-1}(y)) = y \quad (y \in Y),$$

or more succinctly,

$$f^{-1} \circ f = \iota_X, \quad f \circ f^{-1} = \iota_Y,$$

where  $\iota_X$  and  $\iota_Y$  are the identity maps of  $X$  and  $Y$ , respectively.

Notice that, as a relation, the graph of the inverse  $f^{-1}$  of a bijection  $f$  is just the reverse ([Definition 0.8](#)) of the graph of  $f$ .

**Caution!** When a map  $f: X \rightarrow Y$  is a bijection, the notation  $f^{-1}(y)$  may be ambiguous: it could mean either the fiber

$$f^{-1}(\{y\}) = \{x \in X : f(x) = y\}$$

over  $y$  for  $f$ , which is a *subset* of the domain  $X$ , or else

the unique *element*  $x$  of the domain  $X$  such that  $f(x) = y$ .

The context often indicates which meaning is intended. To avoid any ambiguity, we may use the alternative angle-bracket notation  $f^{-1}\langle y \rangle$  ([page 27](#)) for inverse image to indicate the former, whose relationship to the latter is:

$$f^{-1}\langle y \rangle = f^{-1}(\{y\}) = \{f^{-1}(y)\}$$



prop:both-composites-identities

**0.13 Proposition (inverse of map with two-sided identities).** Let  $f: X \rightarrow Y$  be a map. Suppose there is a map  $g: Y \rightarrow X$  whose domain is the codomain of  $f$  and whose codomain is the domain of  $f$  such that the composites  $g \circ f$  and  $f \circ g$  are identity maps, namely:

$$g \circ f = \iota_X, \quad f \circ g = \iota_Y.$$

Then  $f$  is bijective and

$$g = f^{-1}.$$

inverse  
bijection

**Proof.** Apply successively the two parts of [Proposition 0.11](#): Since  $\iota_X$  is injective and  $g \circ f = \iota_X$ , then  $f$  is injective too; since  $\iota_Y$  is surjective and  $f \circ g = \iota_Y$ , then  $f$  is surjective too. .

To see that  $g = f^{-1}$ , let  $y \in Y$ . Then

$$f(g(y)) = \iota_Y(y) = y = (f \circ f^{-1})(y) = f(f^{-1}(y)),$$

whence  $g(y) = f^{-1}(y)$  because  $f$  is injective.  $\square$

In examples (\*) of the subsection “Injections, surjections, and bijections” ([page 28](#)), the map  $g$  is the inverse of the bijection  $f$ . Moreover,  $f$  is the inverse of  $g$ . The same thing is true in general, as part (2) of the following proposition asserts.

prop:inverse-map-properties

**0.14 Proposition.** (1) For any set  $X$ , the identity map  $\iota_X: X \rightarrow X$  is a bijection and

$$(\iota_X)^{-1} = \iota_X.$$

prop-part:inverse-of-inverse

(2) If  $f: X \rightarrow Y$  is a bijection, then  $f^{-1}: Y \rightarrow X$  is also a bijection and

$$(f^{-1})^{-1} = f.$$

prop-part:inverse-of-composite

(3) If  $f: X \rightarrow Y$  and  $g: Y \rightarrow Z$  are both bijections, then their composite is also a bijection and

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

**Proof.** This is [Exercise 45](#).  $\square$

### EXERCISES FOR SECTION 0.3

- 26.** For a given set  $Y$ , how many maps are there from  $\emptyset$  to  $Y$ , and what are they?
- 27.** Show that the following identities concerning the characteristic functions of subsets  $A$  and  $B$  of a set  $X$  hold for all  $x \in X$ :
- (a)  $\chi_{X \setminus A}(x) = 1 - \chi_A(x).$
  - (b)  $\chi_{A \cap B}(x) = \chi_A(x) \cdot \chi_B(x).$
  - (c)  $\chi_{A \cup B}(x) = \chi_A(x) + \chi_B(x) - \chi_{A \cap B}(x).$
- 28.** Verify the following properties of the absolute-value function:
- (a)  $|x| \geq 0.$
  - (b)  $|x| = 0 \iff x = 0.$

liftable map  
lift of a map  
map!liftable

$$(c) |\lambda x| = |\lambda| \cdot |x|.$$

$$(d) |x + y| \leq |x| + |y|.$$

29. Let  $\varepsilon > 0$ . Prove:  $|x| < \varepsilon \iff -\varepsilon < x < \varepsilon$ .

30. Let  $f: \mathbb{R} \rightarrow \mathbb{R}$  and  $g: \mathbb{R} \rightarrow \mathbb{R}$  be defined by  $f(x) = x + x^2$  and  $g(x) = |x|$ . Determine the ranges of  $f$ ,  $g$ ,  $g \circ f$ , and  $f \circ g$ .

prob:fixed-pts-R

31. (a) Find the fixed-points, if any, of the function  $f: \mathbb{R} \rightarrow \mathbb{R}$  defined by  $f(x) = x/2 + 1/x$ .

prob-part:fixed-pt-as-zero

(b) Show that a fixed-point of an arbitrary function  $f: \mathbb{R} \rightarrow \mathbb{R}$  is a zero of another function  $g: \mathbb{R} \rightarrow \mathbb{R}$ .

32. For an arbitrary map  $f: X \rightarrow Y$ , show that

$$f \circ \iota_X = f, \quad \iota_Y \circ f = f.$$

33. Let  $f: X \rightarrow Y$  be a map. Consider the maps

$$\begin{aligned} h: X &\rightarrow \mathcal{P}(X), & k: Y &\rightarrow \mathcal{P}(Y). \\ x &\mapsto \{x\} & y &\mapsto \{y\} \end{aligned}$$

Construct a map  $F: \mathcal{P}(X) \rightarrow \mathcal{P}(Y)$  making the following diagram commutative:

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ h \downarrow & & \downarrow k \\ \mathcal{P}(X) & \xrightarrow{F} & \mathcal{P}(Y) \end{array}$$

prob:liftable

34. Let  $p: E \rightarrow B$  be a given map. Given a map  $f: X \rightarrow B$  be a map from another set to the codomain of  $p$ , we say that  $f$  is **liftable to  $E$  through  $p$**  when there is a map  $\tilde{f}: X \rightarrow E$  such that  $p \circ \tilde{f} = f$ . Such an  $\tilde{f}$  is called a **lift** (or **lifting**) of  $f$  to  $E$  through  $p$ .

(a) Draw an appropriate commutative diagram illustrating that such an  $f$  is liftable to  $E$  through  $p$ .

prob-part:liftable-1st-ex

(b) Now let  $p: \mathbb{R} \rightarrow S_1$  be the map from the line to the unit circle given by  $p(\theta) = \langle \cos 2\pi\theta, \sin 2\pi\theta \rangle$ . Consider the “path”  $\sigma: [0, 1] \rightarrow S_1$  defined by  $\sigma(t) = \langle \cos \pi t/4, \sin \pi t/4 \rangle$ . Show that  $\sigma$  is liftable to  $\mathbb{R}$  through  $p$  by constructing an explicit lift.

*Note:* This same example is posed in terms of complex numbers in [Exercise 141](#).

35. Let  $p: X \times Y \rightarrow X$  and  $q: X \times Y \rightarrow Y$  be the first and second projections. Is it necessarily true that  $E = p(E) \times q(E)$  for each nonempty subset  $E$  of  $X \times Y$ ?

prob-part:image-and-inv-image-composites

36. (a) Exhibit a map  $f: \mathbb{R} \rightarrow \mathbb{R}$  and subsets  $A$  and  $B$  of  $\mathbb{R}$  for which

$$f(A \cap B) \neq f(A) \cap f(B),$$

thereby showing that the inclusion  $f(A \cap B) \subset f(A) \cap f(B)$ , [page 28](#), may be strict.

(b) If  $A \subset X$  and  $f: X \rightarrow Y$  is a map, prove that

$$f^{-1}(Y \setminus f(X \setminus A)).$$

prob-part:image-and-inv-image-composites

37. For a map  $f: X \rightarrow Y$ , show that:

(a)  $A \subset f^{-1}(f(A))$  for every subset  $A$  of  $X$ , but equality need not hold.

saturated set!map@and map

(b)  $f(f^{-1}(B)) \subset B$  for every subset  $B$  of  $Y$ , but equality need not hold.

*Note:* For situations in which equalities necessarily hold above, see [Exercise 46](#).

prob:set-saturated-by-map

**38.** Let  $f: X \rightarrow Y$  be a map. We say that a subset  $A$  of  $X$  is **saturated by  $f$**  when  $A$  contains the inverse image  $f^{-1}(y)$  of each point  $y$  of  $Y$  for which  $A$  intersects  $f^{-1}(y)$ . In other words,  $A$  is saturated by  $f$  when each point of  $X$  that is mapped to the value of  $f$  at some point of  $A$  is already a member of  $A$ .

(a) Show that a subset  $A$  of  $X$  is saturated by  $f$  if and only if  $A = f^{-1}(f(A))$ .

(b) Show that a subset  $A$  of  $X$  is saturated by  $f$  if and only if it has the form  $A = f^{-1}(B)$  for some subset  $B$  of  $Y$ .

(c) Suppose  $A$  is a saturated subset of  $X$ . Deduce that  $f(A \cap E) = f(A) \cap f(E)$  for every subset  $E$  of  $X$ .

**39.** If  $f: X \rightarrow Y$  and if  $B \subset A \subset X$ , must  $f(A \setminus B) = f(A) \setminus f(B)$ ?

**40.** For real numbers  $\alpha$  and  $\beta$ , let  $f: \mathbb{R} \rightarrow \mathbb{R}$  be the affine linear function given by  $x \mapsto \alpha x + \beta$ .

(a) For which  $\alpha$  and  $\beta$  is  $f$  constant?

(b) For which  $\alpha$  and  $\beta$  is  $f$  injective?

(c) For which  $\alpha$  and  $\beta$  is  $f$  surjective?

(d) For which  $\alpha$  and  $\beta$  is  $f$  bijective? For such  $\alpha$  and  $\beta$ , compute the inverse function  $f^{-1}$ .

**41.** Let  $f: \mathbb{Z} \rightarrow \mathbb{Z}$  be the doubling map given by  $n \mapsto 2n$ , so that  $f$  is a bijection. Explain the difference between  $f^{-1}(\{6\})$ , on the one hand, and  $f^{-1}(6)$ , on the other hand.

ve-composite-inj-surj-implies-one-is

**42.** Prove [Proposition 0.11](#).

prob:inj-implies-exists-retraction

**43.** Prove the following converse of [Corollary 0.12 \(1\)](#): If  $f: X \rightarrow Y$  is an injection with *nonempty* domain, then there exists a map  $r: Y \rightarrow X$  with  $r \circ f = \iota_X$ .

*Note:* For the corresponding result for a surjection, see [Proposition 0.30](#).

**44.** If  $f: X \rightarrow Y$  is an injection and if  $g_1, g_2: Z \rightarrow X$  are maps from the same set into the domain of  $f$  with  $f \circ g_1 = f \circ g_2$ , deduce that  $g_1 = g_2$ .

prob:prove-inverse-map-properties

**45.** Prove [Proposition 0.14](#).

e-and-inv-image-composites-inj-surj

**46.** (Continuation of [Exercise 37](#).)

For a map  $f: X \rightarrow Y$ , show that:

(a) If  $f$  is injective, then  $A = f^{-1}(f(A))$  for every subset  $A$  of  $X$ .

(b) If  $f$  is surjective, then  $f(f^{-1}(B)) = B$  for every subset  $B$  of  $Y$ .

**47.** Let  $C$  be the set of all continuous real-valued functions having domain  $\mathbb{R}$ . A map  $I: C \rightarrow C$  is defined by

$$I(f)(x) = \int_0^x f(t) dt \quad (f \in C, x \in \mathbb{R}).$$

Is  $I$  injective? Is  $I$  surjective? (*Hint:* The Fundamental Theorem of Calculus is useful here.)

prob:knaster-kuratowski-fixed-point-theorem-48: Prove the **Knaster-Kuratowski Fixed-point Theorem**: Let  $\varphi: \mathcal{P}(X) \rightarrow \mathcal{P}(X)$  be a map from the power set of a set  $X$  to itself that is increasing with respect to set inclusion, that is, for which  $\varphi(A) \subset \varphi(B)$  whenever  $A \subset B$ . Then  $\varphi$  has some fixed-point. [Hint: Let  $S = \bigcup \mathcal{A}$  where  $\mathcal{A} = \{A \in \mathcal{P}(X) : A \subset \varphi(A)\}$ . Show that  $S \in \mathcal{A}$ ,  $\varphi(S) \in \mathcal{A}$ , and finally  $\varphi(S) = S$ .]

*Note:* The theorem was originally proved by Bronisław Knaster and later generalized by Alfred Tarski beyond the situation of power sets involved here. It may be used to furnish an especially slick proof of the Cantor-Bernstein Theorem (0.126): see Exercise 185.

fixed-point  
Knaster, Bronisław  
Tarski, Alfred  
family  
functional relation  
family!functional relation  
index  
index set  
coordinate  
entry  
subscript  
doubly-indexed family  
family!doubly-indexed  
sec:families

## 0.4 Families and Products

A sequence such as

$$1/2, -1/2^2, 1/2^3, -1/2^4, \dots$$

is, roughly speaking, a list that has a first entry, a second entry, a third entry, and so forth, and “goes on and on forever.” Such a list  $x_1, x_2, x_3, \dots$  is actually a *functional relation*  $\{(n, x_n) : n \in \mathbb{N}^*\}$  (see Definition 0.6) in which each positive integer  $n$  is related to some number  $x_n$ , and that number  $x_n$  may be regarded as labeled by its position  $n$  in the list.

More generally, when an arbitrary functional relation is regarded as labeling the various values of its range by the members of its domain at which it takes these values, the special language and notation of families is employed. Specifically, when a functional relation  $x$  with domain  $I$  is referred to as a **family**, each element  $i \in I$  is called an **index** of the family, and the set  $I$  is called the **index set** of the family. For each index  $i \in I$ , the value  $x(i)$  of  $x$  at an index  $i$  is called the  **$i$ th coordinate** (or  **$i$ th entry**) of the family  $x$  and is denoted by means of a subscript as

$$x_i,$$

and then the family  $x$  is denoted by

$$\langle x_i \rangle_{i \in I}$$

or by

$$\langle x_i : i \in I \rangle.$$

When the range of such a functional relation is a subset of a set  $X$ , so that  $x_i \in X$  for all indices  $i \in I$ , we say that  $\langle x_i \rangle_{i \in I}$  is a family **in**  $X$ . Thus a family  $x = \langle x_i \rangle_{i \in I}$  in a set  $X$  is, in effect, a map  $x: I \rightarrow X$ .

Notice that  $\langle x_i : i \in I \rangle$  and  $\{x_i : i \in I\}$  denote quite different things: the former is a map, whereas the latter is the range of that map. The notation  $\langle x_i \rangle_{i \in I}$  not only is more concise than  $\langle x_i : i \in I \rangle$  but also helps distinguish visually the family from the set  $\{x_i : i \in I\}$ .

In a 2-by-3 matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix},$$

the location of each entry  $a_{ij}$  is indicated by its row-number  $i$  and column-number  $j$ , so the matrix is a family whose index set is the product  $\{1, 2, 3\} \times \{1, 2\}$ . Likewise, the family  $((-1)^i / 2^j)_{i=0,1,2,\dots; j=1,2,3,\dots}$ , where each entry is determined by the power  $i$  of  $-1$  in its numerator and the positive integer  $j$  that is squared in its denominator, the index set is the product  $\mathbb{N} \times \mathbb{N}^*$ . Both examples are instances of a **double-indexed family**  $\langle a_{i,j} \rangle_{(i,j) \in I \times J}$ , that is, a family whose index set is the product  $I \times J$  of two sets.

A matrix  $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$  may be regarded as either a pair of rows  $\langle \langle a_{ij} \rangle_{j \in \{1,2,3\}} \rangle_{i \in \{1,2\}}$  or else a triple of columns  $\langle \langle a_{ij} \rangle_{i \in \{1,2\}} \rangle_{j \in \{1,2,3\}}$ . More generally, a doubly-indexed family

$(a_{i,j})_{(i,j) \in I \times J}$  may be regarded also as: a family  $\langle \langle a_{i,j} \rangle_{i \in I} \rangle_{j \in J}$ , indexed by the set  $J$ , of families, each entry in the latter being indexed by  $I$ ; or a family  $\langle \langle a_{i,j} \rangle_{j \in J} \rangle_{i \in I}$ , indexed by the set  $I$ , of families, each entry in the latter being indexed by  $J$ .

n-tuple@n-tuple  
n-tuple@n-tuple  
ordered pair!2-tuple@as 2-tuple  
n-tuple@n-tuple  
sequence  
sequence!index origin 0  
sequence!index origin 1  
index origin  
sequence  
power of set

subsec:tuples-sequences

### ***n*-tuples and sequences**

The families we shall encounter most often are indexed by subsets of the set  $\mathbb{N} = \{0, 1, 2, \dots\}$  of natural numbers. When for some natural number  $n \geq 1$  a family  $\langle x_i \rangle_{i \in I}$  of elements of a set  $X$  is indexed by the set  $I = \{1, 2, \dots, n\}$ , then  $\langle x_i \rangle_{i \in I}$  is called an ***n*-tuple** (in  $X$ ) and is denoted by

$$\langle x_i \rangle_{i=1,2,\dots,n},$$

or at times by

$$\langle x_i \rangle_{i=1}^n,$$

or more simply by

$$\langle x_1, x_2, \dots, x_n \rangle.$$

[For a specific small  $n$ , the ellipsis dots in the family notation are not needed; for example, if  $n = 4$  then we may write  $\langle x_i \rangle_{i=1,2,3,4}$  or simply  $\langle x_1, x_2, x_3, x_4 \rangle$ .]

A 2-tuple is a family  $\langle x_i \rangle_{i=1,2}$  assigning an object  $x_1$  to index 1 and an object  $x_2$  to index 2 and hence is conceptually the same thing as an ordered pair  $\langle x_1, x_2 \rangle$ . In fact, the rule

$$\langle x_1, x_2 \rangle \mapsto \langle x_i \rangle_{i=1,2}$$

is a one-to-one correspondence between the set  $X \times X$  of all ordered pairs of elements of  $X$  and the set of all 2-tuples in  $X$ . Hence we make no notational distinction between a 2-tuple and an ordered pair when we write  $\langle x_1, x_2 \rangle$ .

Similarly, a 1-tuple  $\langle x_i \rangle_{i=1}$  assigns an object  $x_1$  to the lone index 1 and hence is essentially the same thing as the object  $x_1$  itself.

A family  $\langle x_i \rangle_{i \in I}$  indexed by the set  $I = \mathbb{N}$  of all natural numbers is called a **sequence** and is denoted by

$$\langle x_i \rangle_{i=0,1,2,\dots},$$

or at times by

$$\langle x_i \rangle_{i=0}^\infty$$

(even though the latter notation is *not* meant to suggest that  $\infty$  is one of the indices!), or more simply by

$$\langle x_0, x_1, x_2, \dots \rangle.$$

Families indexed by the set  $I = \mathbb{N}^* = \{1, 2, 3, \dots\}$  of all positive integers are called sequences, too, and then the obvious notation is employed. (To distinguish these two meanings of “sequence”, we may refer them as using *index origin 0* and *index origin 1*, respectively.) More generally, a sequence may be indexed by all the integers starting at any particular one.

### ***n*th power of a set**

subsec:nth-power-of-set

Suppose  $X$  is a set. For a positive integer  $n$ , the ***n*-th power of  $X$**  is defined to be the set of all  $n$ -tuples of elements of  $X$  and is denoted by  $X^n$ ; in other terms,

$$X^n = \{ \langle x_1, x_2, \dots, x_n \rangle : x_1 \in X, x_2 \in X, \dots, x_n \in X \}.$$

As previously indicated, since a 2-tuple is essentially the same thing as an ordered pair, we shall make no distinction between the power  $X^2$  and the Cartesian product  $X \times X$ . Likewise, we shall make no distinction between  $X^1$  and  $X$ .

Euclidean  $n$ -space  
 Euclidean  $n$ -space  
 origin  
 Euclidean  $0$ -space  
 negative of vector

An important instance of the preceding definition occurs when  $X = \mathbb{R}$ , which we examine in the next subsection.

### Euclidean spaces

subsec:euclidean-spaces

Let  $n$  be a positive integer. Then the set

$$\mathbb{R}^n = \{\langle x_1, x_2, \dots, x_n \rangle : x_1 \in \mathbb{R}, x_2 \in \mathbb{R}, \dots, x_n \in \mathbb{R}\}$$

of all  $n$ -tuples of real numbers is  **$n$ -dimensional Euclidean space**. The  $n$ -tuple

$$\mathbf{0} = \langle 0, 0, \dots, 0 \rangle$$

all of whose coordinates are the real number 0 is the **origin** of  $\mathbb{R}^n$  and at times, to distinguish it from the number 0, may be indicated by vector notation **0** (in boldface) or  $\vec{0}$  (*with an arrow*). In particular:

$$\mathbb{R}^1 = \mathbb{R}$$

is the **real line**;

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$$

is the **(Euclidean) plane**; and

$$\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$$

is **(Euclidean) 3-space**. For some purposes it is convenient to define **(Euclidean) 0-space**

$$\mathbb{R}^0 = \{0\},$$

a set with just the single element 0, its origin.

On  $\mathbb{R}^n$  two algebraic operations of vector addition and multiplication by scalars are defined “coordinatewise” as follows. If

$$x = \langle x_1, x_2, \dots, x_n \rangle \in \mathbb{R}^n, \quad y = \langle y_1, y_2, \dots, y_n \rangle \in \mathbb{R}^n,$$

then their vector sum

$$x + y = \langle x_1 + y_1, x_2 + y_2, \dots, x_n + y_n \rangle.$$

And if

$$x = \langle x_1, x_2, \dots, x_n \rangle \in \mathbb{R}^n, \quad \alpha \in \mathbb{R},$$

then their product

$$\alpha x = \langle \alpha x_1, \alpha x_2, \dots, \alpha x_n \rangle$$

(with no operation symbol between the scalar ‘ $\alpha$ ’ and the vector ‘ $x$ ’). The negative  $-x$  of an element  $x \in \mathbb{R}^n$  is

$$-x = \langle -x_1, -x_2, \dots, -x_n \rangle.$$

The following properties say that these two operations make  $\mathbb{R}^n$  into a (real) **vector space**:

$$\begin{aligned} (x + y) + z &= x + (y + z), \\ \mathbf{0} + x &= x = x + \mathbf{0}, \\ x + (-x) &= \mathbf{0} = (-x) + x, \\ x + y &= y + x, \\ (\alpha\beta)x &= \alpha(\beta x), \\ (\alpha + \beta)x &= \alpha x + \beta x, \\ \alpha(x + y) &= \alpha x + \alpha y, \\ 1x &= x. \end{aligned}$$

Here  $x, y, z$  are arbitrary elements of  $\mathbb{R}^n$  and  $\alpha, \beta$  are arbitrary real numbers. When these algebraic operations are involved, we often refer to an element of  $\mathbb{R}^n$  as a **vector**—more precisely, an  **$n$ -vector**—and an element of  $\mathbb{R}$  as a **scalar**.

pg-ref:C-as-R2

The set  $\mathbb{C}$  of complex numbers is the same as the set  $\mathbb{R}^2$ : an arbitrary complex number  $z = x + yi$  with real part  $x = \operatorname{Re}(z)$  and imaginary part  $y = \operatorname{Im}(z)$  is the ordered pair  $\langle x, y \rangle$  of real numbers. Interpreted as operations on  $\mathbb{C}$ , the operations of addition and multiplication by real scalars make  $\mathbb{C}$  into a vector space, too. What distinguishes  $\mathbb{C}$  from  $\mathbb{R}^2$  is that it has an additional operation of multiplication making it a field (page 23). For details about  $\mathbb{C}$ , see the subsection “Complex numbers” (page 89).

vector sum  
vector product by scalar  
vector  
n-vector@ $\mathbb{R}^n$ -vector  
complex numbers  
line segment  
line segment  
line  
convex set  
line segment  
Euclidean  $n$ -space@Euclidean  $\mathbb{R}^n$ -space  
n-tuple@ $\mathbb{R}^n$ -tuple

### Line segments and lines

subsec:segments-lines

Let  $x, y \in \mathbb{R}^n$ . The **line segment joining  $x$  to  $y$**  is the subset

$$\{(1-t)x + ty : 0 \leq t \leq 1\}$$

of  $\mathbb{R}^n$ . (Note that this line segment reduces to the singleton  $\{x\}$  in case  $y = x$ .) If  $x \neq y$ , then the **line (passing) through  $x$  and  $y$**  is the subset

$$\{(1-t)x + ty : t \in \mathbb{R}\}$$

of  $\mathbb{R}^n$  containing that line segment. In  $\mathbb{R}^3$ , for example, the line passing through two points  $x = (x_1, x_2, x_3)$  and  $y = (y_1, y_2, y_3)$  consists of all points  $z = (z_1, z_2, z_3)$  whose coordinates satisfy the equations

$$\begin{cases} z_1 = (1-t)x_1 + ty_1, \\ z_2 = (1-t)x_2 + ty_2, \\ z_3 = (1-t)x_3 + ty_3 \end{cases}$$

for some  $t \in \mathbb{R}$ , and these are the usual parametric equations for the line.

inlinedef:convex

A subset  $K$  of  $\mathbb{R}^n$  is said to be **convex** when it contains the entire line segment joining any two of its points—that is, when

$$x, y \in K \implies (1-t)x + ty \in K \text{ for all } 0 \leq t \leq 1.$$

For example, the line segment joining two points of  $\mathbb{R}^n$  is itself a convex set.

### Arbitrary power of a set

subsec:power-of-set

def:power-of-set

**0.15 Definition.** Let  $X$  is any set. For any index set  $I$ , the  **$I$ th power of  $X$**  is the set

$$X^I = \{\langle x_i \rangle_{i \in I} : x_i \in X \text{ for all } i \in I\},$$

consisting of all families in  $X$  that are indexed by  $I$ . An alternative notation for this set is  $\mathcal{F}(I, X)$ .

The reason for the alternative notation  $\mathcal{F}(I, X)$  is that a family in  $X$  indexed by  $I$  is just a map from  $I$  to  $X$ . Thus  $\mathcal{F}(I, X)$  is the set of all maps from  $I$  to  $X$ .

Taking again  $I = \{1, 2, 3, \dots, n\}$  for a positive integer  $n$ , we obtain as a special case the set  $X^n$  of all  $n$ -tuples of elements of  $X$ . [The notations  $X^n$  and  $X^I$  for  $I = \{1, 2, 3, \dots, n\}$  are almost consistent: according to the foundations of set theory, a positive integer  $n$  is the set  $\{0, 1, 2, \dots, n-1\}$ . The two notations would be perfectly consistent if we used “origin 0” indexing, that is if we indexed  $n$ -tuples by  $\{0, 1, 2, \dots, n-1\}$  rather than by  $\{1, 2, 3, \dots, n\}$ .]

sequence  
power of set

Taking instead  $I = \mathbb{N}^* = \{1, 2, 3, \dots\}$ , we obtain the set

$$X^{\mathbb{N}^*} = \{\langle x_1, x_2, x_3, \dots \rangle : x_1 \in X, x_2 \in X, x_3 \in X, \dots\}$$

of all sequences of elements of  $X$ .

For another instance of a power, take  $X$  to be an arbitrary set. Then  $\{0, 1\}^I$  is the set of all maps from  $I$  to  $\{0, 1\}$ , in other words, the set of all characteristic functions of subsets of  $X$ . This set is often denoted by  $2^X$ . (Writing  $2^X$  instead of  $\{0, 1\}^X$  is consistent with the equality  $2 = \{0, 1\}$  of ordinals as discussed in [Remark 0.110](#).) In particular,  $2^{\mathbb{N}}$  is just the set of all binary sequences—the set of all sequences in  $\{0, 1\}$ .

A subset of a given set is uniquely determined by its characteristic function. More precisely, we have the following result.

prop:2-power-01-powerset

**0.16 Proposition (correspondence between subsets and characteristic functions).** *If  $X$  is any set, then the map*

$$\begin{aligned} \mathcal{P}(X) &\rightarrow 2^X \\ A &\mapsto \chi_A \end{aligned}$$

*sending each subset of  $X$  to its characteristic function is a bijection.*

**Proof.** This is [Exercise 64](#).  $\square$

### Union and intersection of a family of sets

subsec:union-intersect-family

def:union-intersection-family

**0.17 Definition.** Let  $\langle X_i \rangle_{i \in I}$  be a family of sets. Then its **union** is the set

$$\bigcup_{i \in I} X_i = \{x : x \in X_i \text{ for some } i \in I\}$$

consisting of those objects that belong to at least one of the sets in the family. And, when  $I \neq \emptyset$ , its **intersection** is the set

$$\bigcap_{i \in I} X_i = \{x : x \in X_i \text{ for every } i \in I\}$$

consisting of those objects that belong to every one of the sets in the family.

(The case  $I = \emptyset$  is excluded when forming intersections because it is vacuously true that every set  $x$  belongs to  $X_i$  for each  $i \in \emptyset$ , and to allow a “set of all sets” would lead to certain logical paradoxes.)

When  $I = \{1, 2, \dots, n\}$ , the notations

$$\bigcup_{i=1}^n X_i, \quad X_1 \cup X_2 \cup \dots \cup X_n$$

are also used for  $\bigcup_{i \in I} X_i$ , and when  $I = \{1, 2, 3, \dots\}$ , the notations

$$\bigcup_{i=1}^{\infty} X_i, \quad X_1 \cup X_2 \cup X_3 \cup \dots$$

are also used for  $\bigcup_{i \in I} X_i$ ; similarly for intersections. Other notations such as

$$\bigcup_{i=2}^5 X_i, \quad X_0 \cup X_1 \cup X_2 \cup \dots$$

should require no further explanation.



The union and intersection of two sets in the sense of the subsection “Union and intersection of two sets” (page 13) are special cases of the union and intersection of a family. Indeed, if  $X_1$  and  $X_2$  are sets, then

$$X_1 \cup X_2 = \bigcup_{i=1}^2 X_i, \quad X_1 \cap X_2 = \bigcap_{i=1}^2 X_i.$$

Moreover, the formulas involving union and intersection in the subsection “Union and intersection of two sets” generalize to the case of arbitrary families. For example, we have the distributive laws

$$Y \cap \bigcup_{i \in I} X_i = \bigcup_{i \in I} (Y \cap X_i), \quad Y \cup \bigcap_{i \in I} X_i = \bigcap_{i \in I} (Y \cup X_i).$$

product-of-2-sets-as-stacked-copies **0.18 Example.** Let  $X$  and  $Y$  be sets. Then their Cartesian product  $X \times Y$  may be represented as

$$X \times Y = \bigcup_{x \in X} Y_x \text{ where } Y_x = \{x\} \times Y \text{ for each } x \in X.$$

For each fixed  $x \in X$  the map

$$\begin{aligned} Y_x &\rightarrow Y \\ (x, y) &\mapsto y \end{aligned}$$

is a one-to-one correspondence that we may use to regard  $Y_x$  as a “copy” of  $Y$ . Hence we may think of the product  $X \times Y$  as a stack of copies of  $Y$ , one copy for each  $x \in X$ .

When  $X$  has an “ordering” in the sense of Section 0.7, we may use it to arrange the copies of  $Y$  into the corresponding order. For example, when  $X = \mathbb{N}$ , we may arrange the copies  $Y_n$  of  $Y$  into the order  $Y_0, Y_1, Y_2, \dots$ . Then we may picture  $\mathbb{N} \times Y$  as an infinitely long picket fence, each of whose pickets is just like  $Y$ , as shown in Figure 0.1.  $\diamond$

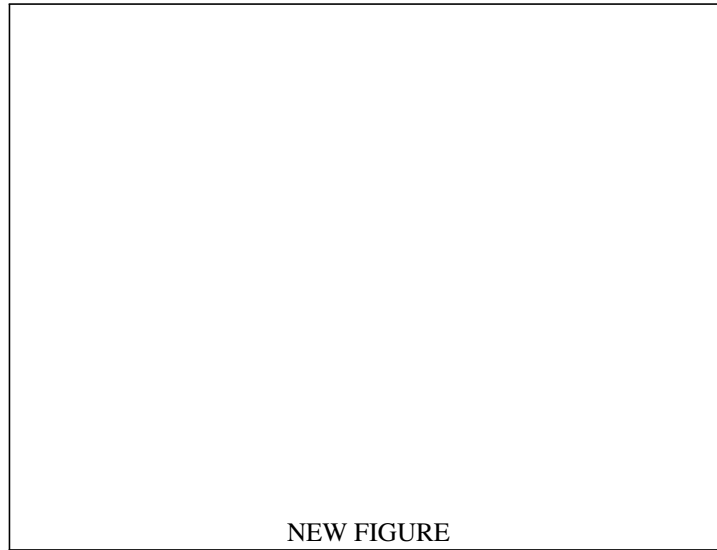


Figure 0.1: The product  $\mathbb{N} \times Y$  as a “picket fence.”

fig:product-of-2-sets-as-stacked-cop

disjoint family  
disjoint sets  
disjoint collection  
disjoint sets

A family  $\langle X_i \rangle_{i \in I}$  of sets is said to be **(pairwise) disjoint** when

$$i \neq j \implies X_i \cap X_j = \emptyset \quad (I, j \in I).$$

inverse image!union@and union  
inverse image!intersection@and intersection  
union!inverse image@and inverse image  
intersection!inverse image@and inverse image  
image!union@and union  
image!intersection@and intersection  
union!image@and image  
intersection!image@and image

Similarly, a collection  $\mathcal{A}$  of sets is said to be **(pairwise) disjoint** when

$$A \neq B \implies A \cap B = \emptyset \quad (A, B \in \mathcal{A}).$$

Observe that **a family  $\langle X_i \rangle_{i \in I}$  need not be disjoint even if  $\bigcap_{i \in I} X_i = \emptyset$** . For example, the sequence  $\langle \{n, n+1\} \rangle_{n \in \mathbb{N}}$  of 2-element subsets of  $\mathbb{N}$  has empty intersection, yet  $\{0, 1\} \cap \{1, 2\} \neq \emptyset$ . Observe also that when a family  $\langle X_i \rangle_{i \in I}$  of *nonempty* sets is disjoint, necessarily  $X_i \neq X_j$  whenever  $i \neq j$ .

The following proposition generalizes, to an arbitrary number of sets, De Morgan's Laws formulated previously (page 14) for just two sets.

prop:DeMorgans-laws

**0.19 De Morgan's Laws.** Let  $\langle A_i \rangle_{i \in I}$  be a family of sets with  $I \neq \emptyset$  and let  $X$  be a set. Then

$$X \setminus \bigcup_{i \in I} A_i = \bigcap_{i \in I} (X \setminus A_i), \quad X \setminus \bigcap_{i \in I} A_i = \bigcup_{i \in I} (X \setminus A_i).$$

inverse-image-union-and-intersection The formation of inverse images under a map  $f: X \rightarrow Y$  “preserves” unions and intersections according to the formulas

$$f^{-1}\left(\bigcup_{i \in I} B_i\right) = \bigcup_{i \in I} f^{-1}(B_i), \quad f^{-1}\left(\bigcap_{i \in I} B_i\right) = \bigcap_{i \in I} f^{-1}(B_i),$$

where  $\langle B_i \rangle_{i \in I}$  is a family of subsets of the codomain  $Y$  of  $f$ . For direct images, although the equality

$$f\left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} f(A_i)$$

holds for an arbitrary family  $\langle A_i \rangle_{i \in I}$  of subsets of the domain  $X$  of  $f$ , in general only the inclusion

$$f\left(\bigcap_{i \in I} A_i\right) \subset \bigcap_{i \in I} f(A_i)$$

holds—even when the index set  $I$  is finite. Those relationships generalize those stated earlier for two sets (see page 28).

In order to state the next result, we shall want the following definition.

def:cover-family

**0.20 Definition.** Let  $A$  be a subset of a set  $X$ . We say that a family  $\langle C_i \rangle_{i \in I}$  of subsets of  $X$  **covers  $A$  in  $X$** , and that  $A$  is **covered by  $\langle C_i \rangle_{i \in I}$  in  $X$** , when  $A \subset \bigcup_{i \in I} C_i$ , that is, when for each  $x \in X$ , there is at least one  $i \in I$  with  $x \in C_i$ . When  $A = X$  and  $A \subset \bigcup_{i \in I} C_i$ , we say simply that  $\langle C_i \rangle_{i \in I}$  **covers  $X$** , and that  $X$  is **covered by  $\langle C_i \rangle_{i \in I}$** .

For example, the family  $\langle C_i \rangle_{i \in \{1,2\}}$  given by  $C_1 = \{x \in \mathbb{R} : x \leq 0\}$  and  $C_2 = \{x \in \mathbb{R} : x \geq 0\}$  covers  $\mathbb{R}$ , and it also covers any subset of  $\mathbb{R}$  in  $\mathbb{R}$ . However, if the inequalities  $\leq$  and  $\geq$  are changed to strict inequalities  $<$  and  $>$ , then  $\langle C_i \rangle_{i \in \{1,2\}}$  no longer covers  $\mathbb{R}$  (because then 0 is “uncovered”—it belongs to neither  $C_1$  nor  $C_2$ ); however, with the strict inequalities,  $\langle C_i \rangle_{i \in \{1,2\}}$  does cover the interval  $[1, 2]$  in  $\mathbb{R}$ .

When  $\langle C_i \rangle_{i \in I}$  is a family of subsets of a given set  $X$ , then to say that  $\langle C_i \rangle_{i \in I}$  covers  $X$  means simply that  $\bigcup_{i \in I} C_i = X$ .

The following result, generalizing [Proposition 0.9](#), justifies defining a map “piecewise” using any number “pieces.”

map!defined piecewise  
piecewise  
cover  
restriction of map  
extension of map  
map!defined piecewise  
piecewise  
union!collection of sets@of a collect  
intersection!collection of sets@of a

prop:extend-many-pieces

**0.21 Proposition (piecewise construction of a map).** Let  $\langle C_i \rangle_{i \in I}$  be a family of subsets of a set  $X$  that covers  $X$  and let  $\langle f_i : C_i \rightarrow Y \rangle_{i \in I}$  be a family of maps all having the same codomain  $Y$ . Suppose each pair of these maps agree on the overlap of their domains, in other words,

$$f_i \upharpoonright_{C_i \cap C_j} = f_j \upharpoonright_{C_i \cap C_j} \quad (i, j \in I).$$

Then there is a unique map

$$h : X = \bigcup_{i \in I} C_i \rightarrow Y$$

that extends each of the individual maps  $f_i$ , that is, such that

$$h \upharpoonright_{C_i} = f_i \quad (i \in I).$$

The proof of the preceding proposition is similar to that of [Proposition 0.9](#) and is left as an exercise.

### Union and intersection of a collection of sets

It is sometimes convenient to form the union and intersection of a number of sets even when these sets have not been explicitly indexed as the values of a family.

subsec:union-intersect-of-collection

def:union-intersect-collection

**0.22 Definition.** Let  $\mathcal{A}$  be a collection of sets. Then the **union of  $\mathcal{A}$**  is the set

$$\bigcup \mathcal{A} = \{x : x \in A \text{ for some } A \in \mathcal{A}\},$$

and when the collection  $\mathcal{A}$  is nonempty the **intersection of  $\mathcal{A}$**  is the set

$$\bigcap \mathcal{A} = \{x : x \in A \text{ for every } A \in \mathcal{A}\}.$$

If we do want to index the sets belonging to a collection  $\mathcal{A}$ , the simplest way is to index each set  $A \in \mathcal{A}$  by itself, in other words, to form the family  $\langle A : A \in \mathcal{A} \rangle$ . And then:

$$\bigcup_{A \in \mathcal{A}} A = \bigcup \mathcal{A}, \quad \bigcap_{A \in \mathcal{A}} A = \bigcap \mathcal{A},$$

(the latter equality holding when the collection  $\mathcal{A}$  is nonempty).

As an example, take  $\mathcal{A}$  to be the collection

$$\mathcal{A} = \{ ]-1/n, 1 - 1/n[ : n = 2, 3, 4, \dots \},$$

of open intervals  $]-1/n, 1 - 1/n[ = \{x \in \mathbb{R} : -1/n < x < 1 - 1/n\}$  for  $n = 2, 3, 4, \dots$ . Then

$$\begin{aligned} \bigcup \mathcal{A} &= \bigcup_{n=2}^{\infty} ]-1/n, 1 - 1/n[ = ]-1/2, 1[ = \{x \in \mathbb{R} : -1/2 < x < 1\}, \\ \bigcap \mathcal{A} &= \bigcap_{n=2}^{\infty} ]-1/n, 1 - 1/n[ = [0, 1/2[ = \{x \in \mathbb{R} : 0 \leq x < 1/2\}. \end{aligned}$$

As another example, if  $\mathcal{A} = \{A, B\}$  for sets  $A$  and  $B$ , then

$$\bigcup \mathcal{A} = \bigcup \{A, B\} = A \cup B, \quad \bigcap \mathcal{A} = \bigcap \{A, B\} = A \cap B.$$

union!collection of sets@of a collection of sets  
 intersection!collection of sets@of a collection of sets  
 subcover  
 refinement!refinement of cover  
 cover  
 product!family of sets

**0.23 Definition.** Let  $S$  be a subset of a set  $X$ . A **cover of  $S$  in  $X$**  is a collection  $C$  of subsets of the entire set  $X$  for which  $S \subset \bigcup C$ , that is, for which each point of  $S$  belongs to at least one member of  $C$ . When  $C$  is a cover of  $S$  in  $X$ , we say that  $C$  **covers  $S$  (in  $X$ )** and that  $S$  is **covered by  $C$  (in  $X$ )**.  
 When  $S = X$ , a cover of  $S$  in  $X$  is called simply a **cover of  $X$** .

In particular, if  $C$  is a collection of subsets of  $X$ , then to say that  $C$  covers  $X$  means simply that  $\bigcup C = X$ . For example, the collection  $\mathcal{A} = \{\{n, n+1, n+2\} : n \in \mathbb{N}\}$  is a cover of the set  $\mathbb{N}$  of natural numbers.

According to the preceding definition, a collection  $C$  is a cover of  $S$  in  $X$  when the associated family  $\langle C \rangle_{C \in C}$  covers  $S$  in  $X$  in the sense of [Definition 0.20](#).

**Usage note.** Some mathematicians use the term *covering* for what we are calling a cover. However, we eschew that alternative because it clashes with the same word used in the term *covering space*—a topological notion to be examined in [Section 7.2](#).

We distinguish two ways of obtaining a new cover of a set from a given cover.

**0.24 Definition.** Let  $\mathcal{A}$  be a cover of a set  $X$  (in  $X$ ).  
 A **subcover of  $\mathcal{A}$**  is a subcollection of  $\mathcal{A}$  that is still a cover of  $X$ .  
 A **refinement** of  $\mathcal{A}$  is a cover  $\mathcal{C}$  of  $X$  such that, for each  $C \in \mathcal{C}$ , there is some  $A \in \mathcal{A}$  with  $C \subset A$ .

Thus a subcover  $\mathcal{S}$  of a cover  $\mathcal{A}$  of  $X$  just has fewer sets than the given cover: after discarding some (possibly none!) of the members of  $\mathcal{A}$ , we retain enough sets for  $\mathcal{S}$  still to cover  $X$ .

By contrast, a refinement  $\mathcal{C}$  of a cover  $\mathcal{A}$  of  $X$  has “smaller” members: after discarding elements (possibly none!) from each member of  $\mathcal{A}$ , we retain sets that are still big enough that together they all cover  $X$ .

For example, the collection  $\mathcal{S} = \{\{m, m+1, m+2\} : m \in \mathbb{N} \text{ and } m \text{ is even}\}$  is a subcover of the cover  $\mathcal{A} = \{\{n, n+1, n+2\} : n \in \mathbb{N}\}$  of  $\mathbb{N}$ .

Obviously a subcover of a given cover  $\mathcal{A}$  of a set  $X$  is a refinement of  $\mathcal{A}$ . However, a refinement of  $\mathcal{A}$  need not be a subcover of  $\mathcal{A}$ . For example, the collection  $\mathcal{C} = \{\{n, n+1\} : n \in \mathbb{N}\}$  is a refinement of that same  $\mathcal{A}$  but is *not* a subcover of  $\mathcal{A}$ .

## Product of a family of sets

subsec:family-product

Instead of families all of whose coordinates belong to the same set, we may consider families whose coordinates belong to (possibly) different sets.

Let  $X_1, X_2, \dots, X_n$  be sets, where  $n$  is a positive integer. Then the **product of  $\langle X_1, X_2, \dots, X_n \rangle$**  is the set

$$\begin{aligned} \bigtimes_{i=1}^n X_i &= X_1 \times X_2 \times \cdots \times X_n \\ &= \{\langle x_1, x_2, \dots, x_n \rangle : x_1 \in X_1, x_2 \in X_2, \dots, x_n \in X_n\} \end{aligned}$$

of all those  $n$ -tuples in  $\bigcup_{i=1}^n X_i$  whose first coordinate belongs to  $X_1$ , whose second coordinate belongs to  $X_2$ , and so on. Because we make no conceptual distinction between 2-tuples and

ordered pairs, when  $n = 2$  we may write

$$\bigtimes_{i=1}^2 X_i = X_1 \times X_2,$$

so that products in the present sense generalize products in the sense of the [subsection “Pairs and products”](#) (page 14). For  $j = 1, 2, \dots, n$ , the  **$j$ th projection** is the map

$$p_j: X_1 \times X_2 \times \cdots \times X_n \rightarrow X_j \\ \langle x_1, x_2, \dots, x_n \rangle \mapsto x_j$$

that sends each element of the product to its  $j$ th coordinate. In particular, when  $n = 2$  these projections

$$p_1: X_1 \times X_2 \rightarrow X_1 \quad p_2: X_1 \times X_2 \rightarrow X_2 \\ \langle x_1, x_2 \rangle \mapsto x_1, \quad \langle x_1, x_2 \rangle \mapsto x_2$$

are the first and second projections considered in the [subsection “Composition of maps”](#) (page 25).

Next, let  $\langle X_1, X_2, X_3, \dots \rangle$  be a sequence of sets. Then the **product** of this sequence is the set

$$\bigtimes_{i=1}^{\infty} X_i = X_1 \times X_2 \times X_3 \times \cdots \\ = \{ \langle x_1, x_2, x_3, \dots \rangle : x_1 \in X_1, x_2 \in X_2, x_3 \in X_3, \dots \}$$

of all those sequences in  $\bigcup_{i=1}^{\infty} X_i$  whose  $i$ th coordinate belongs to the  $i$ th set  $X_i$  for each  $i \in I$ . As before, for each  $j = 1, 2, 3, \dots$  we have the  **$j$ th projection**

$$p_j: X_1 \times X_2 \times X_3 \times \cdots \rightarrow X_j \\ \langle x_1, x_2, x_3, \dots \rangle \mapsto x_j$$

Both kinds of products just considered can be denoted uniformly by

$$\bigtimes_{i \in I} X_i$$

where either  $I = \{1, 2, \dots, n\}$  for some  $n$  or else  $I = \mathbb{N}^* = \{1, 2, 3, \dots\}$ . The same notation is used, with the obvious meaning, when the index set  $I$  is an subset of  $\mathbb{N}$ .

The index set for a product can, even more generally, be an arbitrary set.

def:product-of-family

**0.25 Definition.** Let  $\langle X_i \rangle_{i \in I}$  be a family of sets indexed by a set  $I$ . Then the **product of  $\langle X_i \rangle_{i \in I}$** , denoted by  $\bigtimes_{i \in I} X_i$ , is the set

$$\{ \langle x_i \rangle_{i \in I} : \langle x_i \rangle_{i \in I} \text{ is a family and } x_i \in X_i \text{ for each } i \in I \}$$

consisting of all those families in  $\bigcup_{i \in I} X_i$  that are indexed by  $I$  and whose  $i$ th entry belongs to the  $i$ th set  $X_i$  for each index  $i \in I$ .

For each  $j \in I$ , the  **$j$ th projection** is the map

$$p_j: \bigtimes_{i \in I} X_i \rightarrow X_j \\ \langle x_i \rangle_{i \in I} \mapsto x_j$$

whose value at an element of the product is its  $j$ th coordinate.

Some mathematicians prefer the notation  $\prod_i X_i$  for the Cartesian product.

product!two sets@of two sets  
projection  
product!sequence of sets@of sequen  
projection

power of set  
product! family of sets  
Axiom of Choice

In practice, we may more informally write the definition as

$$\prod_{i \in I} X_i = \{ \langle x_i \rangle_{i \in I} : x_i \in X_i \text{ for each } i \in I \}.$$

Notice that each element  $\langle x_i \rangle_{i \in I}$  of the product  $\prod_{i \in I} X_i$  is necessarily a family in the set  $\bigcup_{i \in I} X_i$ . Accordingly, it is sometimes useful to think of the product  $\prod_{i \in I} X_i$  as being a set of all maps from  $I$  to  $\bigcup_{i \in I} X_i$ , namely, the set of those maps from  $I$  to  $\bigcup_{i \in I} X_i$  for which the value at  $i$  belongs to  $X_i$  for each  $i$ .

All the earlier types of products of families are special cases of the preceding general notion of product.

Suppose there is a set  $X$  with

$$X_i = X \quad (i \in I).$$

Then

$$\prod_{i \in I} X_i = X^I,$$

the  $I$ th power of  $X$  (Definition 0.15). Thus in this case the product is just the set of all maps  $I \rightarrow X$ . In particular, when  $I = \{1, 2, \dots, n\}$  for a positive integer  $n$ , then

$$\prod_{i \in \{1, 2, \dots, n\}} X_i = \prod_{i=1}^n X_i = X^n,$$

the  $n$ th power of  $X$  as defined in the subsection “ $n$ th power of a set” (page 35). For example,

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ copies}}.$$

Families whose index sets are neither finite like  $\{1, 2, \dots, n\}$  nor “denumerable” like  $\{1, 2, 3, \dots\}$  arise, for example, in the study of differential equations. For example, the first-order linear differential equation  $dx/dt = t + 3$  has, for each real number  $c$ , the particular solution  $x_c: \mathbb{R} \rightarrow \mathbb{R}$  given by  $x_c(t) = t^2/2 + 3t + c$ ; then the family  $\langle x_c \rangle_{c \in \mathbb{R}}$  of all these solutions belongs to the product set  $\mathbb{R}^{\mathbb{R}}$ .

### Axiom of Choice

subsec:AC

If  $X_1$  and  $X_2$  are nonempty sets, there exist elements  $c_1 \in X_1$  and  $c_2 \in X_2$ , and so the ordered pair  $\langle c_1, c_2 \rangle \in X_1 \times X_2$ . In other words, we may “simultaneously” an element from each of the two sets. The same thing is true for “finitely many” nonempty sets  $X_1, X_2, \dots, X_n$ —see Exercise 66.

The corresponding property for a family of nonempty sets whose index set is not finite may seem intuitively “obvious.” However, it cannot be proved without a separate axiom.

ax:AC **0.26 Axiom of Choice.** Let  $\langle X_i \rangle_{i \in I}$  be a family of nonempty sets. Then there exists a map  $c: I \rightarrow \bigcup_{i \in I} X_i$  such that, for each  $i \in I$ , the value  $c(i) \in X_i$ .

Because such a map  $c$  is just a family in  $\bigcup_{i \in I} X_i$ , another formulation of the Axiom of Choice is the following:

cor:AC-product **0.27 Corollary (Axiom of Choice—product version).** *The product  $\prod_{i \in I} X_i$  of a family  $\langle X_i \rangle_{i \in I}$  of nonempty sets is itself nonempty.*

section of a surjection  
surjection!section@and section  
Axiom of Choice

The Axiom of Choice has equivalent formulations in terms of collections of nonempty sets.

cor:AC-choice-function **0.28 Corollary (Axiom of Choice—choice function version).** *Let  $\mathcal{A}$  be a collection of nonempty sets. Then there exists a map  $c: \mathcal{A} \rightarrow \bigcup \mathcal{A}$  such that  $c(A) \in A$  for each  $A \in \mathcal{A}$ .*

page:choice-fn In the notation of [Corollary 0.28](#), such a map  $c$  is called a **choice function** of  $\mathcal{A}$ .

cor:AC-choice-set **0.29 Corollary (Axiom of Choice—choice set version).** *Let  $\mathcal{A}$  be a collection of nonempty sets. Then there exists a set  $C$  such that, for each  $A \in \mathcal{A}$  the intersection  $C \cap A$  is a singleton.*

page:choice-set In the notation of [Corollary 0.29](#), such a set  $C$  is called a **choice set** of  $\mathcal{A}$ .  
All the preceding versions of the Axiom of Choice are, in fact, logically equivalent—see [Exercise 69](#). Accordingly, from now on, we shall indiscriminately refer to any of them simply as “the Axiom of Choice.”

The Axiom of Choice will be needed later to show that an infinite set necessarily has a subset that is “denumerable” in the sense that it can be put into one-to-one correspondence with the set  $\mathbb{N}$  of natural numbers. (See the proof of [Proposition 0.47](#)).

One consequence of the Axiom of Choice is the following converse of [Corollary 0.12 \(2\)](#); its proof is requested in [Exercise 68](#).

prop:surj-implies-exists-section **0.30 Proposition.** *If  $f: X \rightarrow Y$  is a surjection, then there exists a map  $s: Y \rightarrow X$  such that  $f \circ s = \iota_Y$ .*

In the notation of the preceding proposition, such a map  $s$  is referred to as a **section** of the surjection  $f$ . Necessarily, such a section is injective.

## EXERCISES FOR SECTION 0.4

49. (a) Find a sequence  $\langle X_n : n = 0, 1, 2, \dots \rangle$  of sets such that  $X_{n+1}$  is a proper subset of  $X_n$  for each  $n$ . Then compute  $\bigcup_{n=0}^{\infty} X_n$  and  $\bigcap_{n=0}^{\infty} X_n$ .  
(b) Find a sequence  $\langle Y_n \rangle_{n \in \mathbb{N}}$  of sets such that  $Y_n \in Y_{n+1}$  for each  $n$ . Then compute  $\bigcup_{n=0}^{\infty} Y_n$  and  $\bigcap_{n=0}^{\infty} Y_n$ .
50. Let  $x = (1, 2, 3, 4)$ ,  $y = (4, 3, 2, 1) \in \mathbb{R}^4$  and let  $L$  be the line in  $\mathbb{R}^4$  passing through  $x$  and  $y$ .  
(a) Determine two points other than  $x$  and  $y$  that also belong to  $L$ .  
(b) Does the origin belong to  $L$ ?
51. Let  $K = \{x \in \mathbb{R}^n : |x_1| + |x_2| + \dots + |x_n| \leq 1\}$ .  
(a) Draw pictures of  $K$  for the cases  $n = 1, 2$ , and  $3$ .  
(b) Show that  $K$  is convex for arbitrary  $n$ .
52. Let  $f$  be a continuous real-valued function with domain the set of real numbers.

- (a) Describe, as an element of an appropriate product set, a family that arises from the various definite integrals  $\int_0^x f(t) dt$ .
- (b) Repeat (a) but for a family that arises from the various definite integrals  $\int_a^b f(t) dt$ .

53. Prove or disprove:

- (a) If  $\langle K_i \rangle_{i \in I}$  is any family of convex subsets of  $\mathbb{R}^n$ , then  $\bigcap_{i \in I} K_i$  is convex.
- (b) If  $\langle K_i \rangle_{i \in I}$  is any family of convex subsets of  $\mathbb{R}^n$ , then  $\bigcup_{i \in I} K_i$  is convex.

54. Let  $X_1 = \{0, 1\}$ ,  $X_2 = \{1, 2\}$ ,  $X_3 = \{2, 3\}$ , and  $X_n = \{3\}$  for  $n \geq 4$ . Compute:

- (a)  $\bigcup_{n=1}^3 X_n$ . (d)  $\bigcup_{n=1}^{\infty} X_n$ .
- (b)  $\bigcap_{n=1}^{\infty} X_n$ . (e)  $\bigcap_{n=1}^3 X_n$ .
- (c)  $\bigtimes_{n=1}^3 X_n$ . (f)  $\bigtimes_{n=1}^{\infty} X_n$ .

55. Represent the product  $X \times Y$  of two sets as a bundle of copies of  $X$ . Then draw a picture of the representation in the case  $Y = \mathbb{N}$ .

56. (a) Prove: If  $\langle X_i : i \in I \rangle$  and  $\langle Y_i : i \in I \rangle$  are families of sets indexed by the same set  $I$ , then

$$\left( \bigcap_{i \in I} X_i \right) \times \left( \bigcap_{i \in I} Y_i \right) = \bigcap_{i \in I} (X_i \times Y_i).$$

(b) Does the analog of (a) for union instead of intersection also hold?

57. Prove [Proposition 0.21](#).

58. If  $X$  is a set, what are  $\bigcup \mathcal{P}(X)$  and  $\bigcap \mathcal{P}(X)$ ?

59. If  $\mathcal{A}$  is a collection of sets and  $X \in \mathcal{A}$ , show that  $\bigcap \mathcal{A} \subset X \subset \bigcup \mathcal{A}$ .

60. Find collections  $\mathcal{A}$  and  $\mathcal{B}$  of sets for which  $(\bigcap \mathcal{A}) \cap (\bigcap \mathcal{B}) \neq \bigcap (\mathcal{A} \cap \mathcal{B})$ .

61. Let  $\mathcal{A}$  and  $\mathcal{B}$  be collections of sets with  $\mathcal{A} \subset \mathcal{B}$ . Prove:

$$\bigcup \mathcal{A} \subset \bigcup \mathcal{B}, \quad \bigcup \mathcal{B} \subset \bigcup \mathcal{A}.$$

62. (a) What is  $\bigcup \emptyset$ ?

(b) If  $X$  is any set, what is the power set  $X^\emptyset$ ?

prob:no-intersect-empty-collection

63. Why in [Definition 0.22](#) was the intersection  $\bigcap \mathcal{A}$  formed only when the collection  $\mathcal{A}$  is nonempty? In other words, why did we *not* include the definition

$$\bigcap \emptyset = \{x : x \in \emptyset \text{ for every } A \in \emptyset\}?$$

prob:pf-1-1-powerst-char-fns

64. Prove [Proposition 0.16](#).

prob:pf-extend-many-pieces

65. Prove [Proposition 0.21](#).

prob:finite-product-nonempty

66. Let  $n$  be a positive integer and let  $(X_i)_{i=1,2,\dots,n}$  be a family of *nonempty* sets.

(a) Prove that the product  $\bigtimes_{i=1}^n X_i$  is nonempty. Do this *without* using any version of the Axiom of Choice.

(b) Deduce from (a) that each projection  $p_j : \bigtimes_{i=1}^n X_i \rightarrow X_j$  is surjective.



67. *Without* using any version of the Axiom of Choice:

- (a) Find a choice function for the collection of all nonempty subsets of  $\mathbb{N}$ .
- (b) Find a choice function for the collection of all nonempty subsets of the set of nonnegative rational numbers.

recursion  
recursion!ordinary  
initial condition  
recurrence relation

68. Prove the converse [Proposition 0.30](#) of [Corollary 0.12](#) (2).

69. The purpose of this exercise is to show that all the stated versions of the Axiom of Choice—[0.26](#), [0.27](#), [0.28](#), and [0.29](#)—are in fact logically equivalent.

- (a) *Without* using the Axiom of Choice ([0.26](#)), show that the statements of its choice function version ([Corollary 0.28](#)) and its choice set version ([Corollary 0.29](#)) are equivalent.
- (b) Deduce the choice function version ([Corollary 0.28](#)) from the Axiom of Choice ([0.26](#)).
- (c) *Without* assuming the Axiom of Choice ([0.26](#)), deduce its statement from its choice function version ([Corollary 0.28](#)).

## 0.5 Recursion

sec:recurse

Some proofs require constructing a sequence each of whose entries is defined in terms of the immediately preceding entry, or even in terms of all the immediately preceding entries, of that sequence. A sequence obtained in such a way is said to be defined **recursively**, and we say and that we are using **recursion**.

### Ordinary recursion

subsec:ordinary-recurse

Here is a simple example of recursion.

**0.31 Example.** Form a sequence  $\langle s_n \rangle_{n=0,1,2,\dots}$  of integers as follows: begin with some particular integer  $c$  and then repeatedly add the same increment  $b$  to obtain successive entries. To be specific, take  $c = 2$  and  $b = 5$ . In other words:

$$\{eq:arith-progress-recurse\} \quad (*) \quad \begin{cases} s_0 = 2, \\ s_{n+1} = s_n + 5 \end{cases} \quad (n = 0, 1, 2, \dots).$$

From that “initial condition”  $s_0 = 2$  and “recurrence relation”  $s_{n+1} = s_n + 5$ , we may readily determine all the entries up to any particular index  $n$ . For example:

$$\begin{aligned} s_1 &= s_0 + 5 = 2 + 5 = 7, \\ s_2 &= s_1 + 5 = 7 + 5 = 12, \\ s_3 &= s_2 + 5 = 12 + 5 = 17. \end{aligned}$$

But why does such an entire sequence—whose domain is the entire set of natural numbers—actually exist? As intuitively compelling as recursion may seem, it nonetheless requires justification. The next theorem provides that justification.

Incidentally, the sequence in this example may be obtained explicitly, by a single formula, rather than implicitly using recursion, namely,  $s_n = 2 + 5n$  for every  $n \in \mathbb{N}$ . In the example following the next theorem, in general there is no such explicit formula.  $\diamond$

In the preceding example, express the recurrence relation  $s_{n+1} = s_n + 5$  as  $s_{n+1} = G(s_n)$  where  $G: \mathbb{N} \rightarrow \mathbb{N}$  is the map defined by  $G(k) = k + 5$ . Then the existence of the sequence

**initial condition**  $\langle s_n \rangle_{n=0,1,2,\dots}$  may be cast in terms of the initial value  $c = 2$  and the map  $G$ . Now a sequence  
**recurrence relation**  $\langle s_n \rangle_{n=0,1,2,\dots}$  in the set  $\mathbb{N}$  is just a map  $h: \mathbb{N} \rightarrow \mathbb{N}$  with  $h_n = s(n)$  for each  $n \in \mathbb{N}$ . Then the following theorem justifies using recursion such as in the preceding example.

thm-ordinary-recursion **0.32 Principle of Ordinary Recursion.** Let  $c$  be a given element of a set  $X$  and let

$$G: X \rightarrow X$$

be a given map. Then there exists a unique map

$$h: \mathbb{N} \rightarrow X$$

such that

$$\{ \text{prop-eq:IC-and-recurrence} \} (*) \quad \begin{cases} h(0) = c, \\ h(n+1) = G(h(n)) \end{cases} \quad (n \in \mathbb{N}).$$

The equality  $h(0) = c$  is called the **initial condition**, and the relation  $h(n+1) = G(h(n))$  is called the **recurrence relation**.

**Proof.** We prove uniqueness first, then existence.

Uniqueness. Suppose each of two maps

$$h: \mathbb{N} \rightarrow X, \quad h': \mathbb{N} \rightarrow X$$

satisfies the initial condition and the recurrence relation (\*). We use induction on  $n$  to prove that  $h(n) = h'(n)$  for every  $n \in \mathbb{N}$ .

First,  $h(0) = c = h'(0)$  from the initial condition.

Next, let  $n \in \mathbb{N}$  and assume that  $h(n) = h'(n)$ . Then

$$h(n+1) = G(h(n)) = G(h'(n)) = h'(n+1)$$

by the recurrence relation.

Existence. We are going to construct a sequence of maps  $\langle h_n \rangle_{n \in \mathbb{N}}$  with

$$h_n: \{0, 1, 2, \dots, n\} \rightarrow X \quad (n \in \mathbb{N})$$

such that, for each  $n$ :

$$\{ \text{pf-eq:recurse-1} \} (1) \quad \begin{cases} h_n(0) = c, \\ h_n(k+1) = G(h_n(k)) \end{cases} \quad (0 \leq k < n)$$

and

$$\{ \text{pf-eq:recurse-2} \} (2) \quad h_n \upharpoonright_{\{0,1,2,\dots,m\}} = h_m \quad (0 \leq m < n).$$

For  $0 \leq m < n$ , we will then have  $(\text{dom } h_m) \cap (\text{dom } h_n) = \text{dom } h_m$ , and by property (2),

$$h_n \upharpoonright_{\text{dom } h_m} = h_m.$$

Since  $\bigcup_{n \in \mathbb{N}} \text{dom } h_n = \bigcup_{n \in \mathbb{N}} \{0, 1, 2, \dots, n\} = \mathbb{N}$ , from [Proposition 0.21](#) it will follow that there exists a map

$$h: \mathbb{N} \rightarrow X$$

such that

$$h \upharpoonright_{\{0,1,2,\dots,n\}} = h_n \quad (n \in \mathbb{N}).$$

From (1) the map  $h$  will satisfy the desired initial condition and recurrence relation (\*).

In order to obtain such maps  $h_n: \{0, 1, 2, \dots, n\} \rightarrow X$ , define  $M_n$  to be the set of *all* iterate of map maps  $h: \{0, 1, 2, \dots, n\} \rightarrow X$  having the properties

$$\{\text{pf-eq:ordinary-recurse-3}\} \quad (3) \quad \begin{cases} h(0) = c, \\ h(k+1) = G(h(k)) \end{cases} \quad (k \in \mathbb{N}, 0 < k < n).$$

Clearly, for each  $n$ ,

$$\{\text{pf-eq:recurse-4}\} \quad (4) \quad h \in M_{n+1} \implies h|_{\{0,1,2,\dots,n\}} \in M_n.$$

An inductive argument based upon that implication shows:

$$h \in M_n \implies h|_{\{0,1,2,\dots,m\}} \in M_m \quad (n, m \in \mathbb{N}; m < n).$$

To complete the proof, we use induction on  $n$  to show that for each  $n \in \mathbb{N}$ , the set  $M_n$  contains a *unique* map

$$h_n: \{0, 1, 2, \dots, n\} \rightarrow X.$$

For the base step, observe that  $M_0$  contains exactly one function, namely, the constant map  $h_0: \{0\} \rightarrow X$ .

For the inductive step, let  $n \in \mathbb{N}$  and assume that  $M_n$  contains a unique member  $h_n$ . Then the map

$$h_{n+1}: \{0, 1, 2, \dots, n, n+1\} \rightarrow X$$

having as its domain the set  $\{0, 1, 2, \dots, n, n+1\} = \{0, 1, 2, \dots, n\} \cup \{n+1\}$  and defined piecewise by

$$h_{n+1}|_{\{0,1,2,\dots,n\}} = h_n, \quad h_{n+1}(n+1) = G(h_n(n))$$

is a member of  $M_{n+1}$ . And it is the only member of  $M_{n+1}$ . In fact, let  $h \in M_{n+1}$ . Then  $h|_{\{0,1,2,\dots,n\}} \in M_n$  by (4), so that  $h|_{\{0,1,2,\dots,n\}} = h_n$  by the inductive hypothesis. Also

$$h(n+1) = G(h(n)) = G(h_n(n)) = h_{n+1}(n+1).$$

Hence indeed  $h = h_{n+1}$ . This completes the induction that  $M_n$  contains a unique member for each  $n$ .  $\square$

A minor modification of the preceding theorem provides an analogous principle when the initial index  $n$  is 1, or any other particular natural number, rather than 0.

0.33 Example. Let  $f: A \rightarrow A$  be any map having the same codomain as domain. Recursively define the sequence  $\langle f^n \rangle_{n=1,2,3,\dots}$  of maps from  $A \rightarrow A$  by

$$\begin{cases} f^0 = \iota_A & (\text{the identity map of } A), \\ f^{n+1} = f \circ f^n & (n = 0, 1, 2, \dots). \end{cases}$$

For example

$$f^1 = f \circ \iota_A = f, \quad f^2 = f \circ f, \quad f^3 = f \circ (f^2) = f \circ (f \circ f) = f \circ f \circ f.$$

For each  $n \in \mathbb{N}$ , the map  $f^n$  so defined is called the  *$n$ th iterate of  $f$* .

To make explicit how Theorem 0.32 applies here: take  $X = A^A$ , the set of all maps from  $A$  to  $A$ ; take  $c = X$ ; and take  $G: X \rightarrow X$  be the map given by

$$G(\varphi) = f \circ \varphi \quad (\varphi \in X). \quad \diamond$$

### recursion!ordinary Primitive and complete recursion

subsec:primitive-complete-recurse

The factorial function  $n \mapsto n!$  is typically defined recursively by

$$\begin{cases} 0! = 1, \\ (n+1)! = (n+1) \cdot (n!). \end{cases}$$

Unfortunately, the existence of this function is *not* directly guaranteed by [Principle of Ordinary Recursion \(0.32\)](#)—because the value of the function at  $n+1$  depends not just on the value at  $n$ , but on  $n$  itself as well.

The right-hand-side of the recurrence relation for the factorial function has the form  $(n+1)! = G(n, n!)$  where  $G: \mathbb{N} \times \mathbb{N}$  is the function of two variables given by  $G(n, x) = (n+1)x$ . Then the following generalization of [Principle of Ordinary Recursion \(0.32\)](#), a consequence of it, does justify the recursive definition of the factorial function.

cor:primitive-recursion

**0.34 Principle of Primitive Recursion.** Let  $c$  be a given element of a set  $X$  and let

$$G: \mathbb{N} \times X \rightarrow X$$

be a given map. Then there exists a unique map

$$h: \mathbb{N} \rightarrow X$$

such that

{cor-eq-primitive-recurse} (\*\*)

$$\begin{cases} h(0) = c, \\ h(n+1) = G(n, h(n)) \quad (n \in \mathbb{N}). \end{cases}$$

**Proof.** Existence. Define the map

$$K: \mathbb{N} \times X \rightarrow \mathbb{N} \times X$$

by

$$K(n, x) = \langle n+1, G(n, x) \rangle \quad (n \in \mathbb{N}, x \in X).$$

By Principle of Ordinary Recursion (0.32), there is a map

$$f: \mathbb{N} \rightarrow \mathbb{N} \times X$$

such that

{pf-eq-primitive-recurse-1} (1)

$$\begin{cases} f(0) = \langle 0, c \rangle, \\ f(n+1) = K(f(n)) \quad (n \in \mathbb{N}). \end{cases}$$

Let

$$\begin{aligned} p: \mathbb{N} \times X &\rightarrow \mathbb{N}, & q: \mathbb{N} \times X &\rightarrow X \\ \langle n, x \rangle &\mapsto n & \langle n, x \rangle &\mapsto x \end{aligned}$$

be the projection maps, so that

$$f(n) = \langle p \circ f(n), q \circ f(n) \rangle \quad (n \in \mathbb{N}).$$

Define

$$h = q \circ f: \mathbb{N} \rightarrow X.$$

We are going to show that  $h$  satisfies the desired initial condition and recurrence relation (\*\*).

fix: converting  
equal to aligned  
and inserting  
shortintertext  
here...

Let

$$g = p \circ f: \mathbb{N} \rightarrow \mathbb{N}$$

so that

$$\{ \text{pf-eq-primitive-recurse-2} \} \quad (2) \quad f(n) = \langle g(n), h(n) \rangle \quad (n \in \mathbb{N}).$$

Then

$$\{ \text{pf-eq-primitive-recurse-3} \} \quad (3) \quad \begin{aligned} K(f(n)) &= K(g(n), h(n)) \\ &= \langle g(n) + 1, G(g(n), h(n)) \rangle \end{aligned} \quad (n \in \mathbb{N}).$$

Next, we shall show that  $g = \iota_{\mathbb{N}}$ , the identity map of  $\mathbb{N}$ . From (3),

$$\begin{cases} g(0) = p(f(0)) = p(0, c) = 0, \\ g(n+1) = p(f(n+1)) = p(K(f(n))) \\ \quad = p(g(n) + 1, G(g(n), h(n))) \\ \quad = g(n) + 1 \end{cases} \quad (n \in \mathbb{N}),$$

and so

$$\{ \text{pf-eq-primitive-recurse-4} \} \quad (4) \quad \begin{cases} g(0) = 0, \\ g(n+1) = g(n) + 1 \end{cases} \quad (n \in \mathbb{N})$$

Now the identity map  $\mathbb{N}$  satisfies  $\mathbb{N}(0) = 0$  and  $\mathbb{N}(n+1) = \mathbb{N}(n) + 1$  for all  $n \in \mathbb{N}$ . From (4) and the uniqueness part of the theorem on Principle of Ordinary Recursion (0.32) we obtain

$$g = \iota_{\mathbb{N}}.$$

Since  $g = \iota_{\mathbb{N}}$ , equation (2) simplifies to

$$\{ \text{pf-eq-primitive-recurse-5} \} \quad (5) \quad f(n) = \langle n, h(n) \rangle \quad (n \in \mathbb{N})$$

and (3) simplifies to

$$\{ \text{pf-eq-primitive-recurse-6} \} \quad (6) \quad K(f(n)) = \langle n+1, G(n, h(n)) \rangle \quad (n \in \mathbb{N}).$$

Hence from (1), (5), (3), and (6):

$$\begin{cases} h(0) = q(f(0)) = q(0, c) = c, \\ h(n+1) = q(f(n+1)) = q(K(f(n))) = q(n+1, G(n, h(n))) \\ \quad = G(n, h(n)) \end{cases} \quad (n \in \mathbb{N}),$$

as desired.

**Uniqueness.** If  $h, h': \mathbb{N} \rightarrow X$  are maps each of which satisfies (\*\*), then an induction on  $n$  shows that, for each  $n \in \mathbb{N}$ , the equality  $h(n, x) = h'(n, x)$  holds for all  $x \in X$ .  $\square$

Sometimes when defining a sequence recursively, the  $(n+1)$ st entry depends not just on  $n$  and the  $n$ th entry, but on all the entries for indices  $0, 1, 2, \dots, n$  up to the  $n$ th. The following theorem, which generalizes both Principle of Ordinary Recursion and Principle of Primitive Recursion, justifies this kind of recursion. And now, the domain of the given map  $G$  has to allow  $n$ -tuples of elements of the set  $X$ , that is, each set of the form  $\mathbb{N} \times X^n$ , for every for every  $n = 1, 2, 3, \dots$

recursion!ordinary  
recursion!primitive  
Principle of Primitive Recursion

recursion  
thm:complete-recursion

**0.35 Principle of Complete Recursion.** Let  $c$  be a given element of a set  $X$  and let

$$G: \mathbb{N} \times \bigcup_{n=1}^{\infty} X^n \rightarrow X$$

be a given map. Then there exists a unique map

$$f: \mathbb{N} \rightarrow X$$

such that

$$\begin{aligned} f(0) &= c, \\ f(n+1) &= G(n, f(0), f(1), \dots, f(n)) \quad (n \in \mathbb{N}). \end{aligned}$$

For an example of Principle of Complete Recursion, see the proof of [Proposition 0.46](#).

### EXERCISES FOR SECTION 0.5

- 70.** Give a recursive definition of the  $n$ th powers  $a^n$  of a real number (or, if you prefer, an integer)  $a$ , where  $n = 0, 1, 2, \dots$ . What particular form of recursion—ordinary, primitive, or complete—does your definition use, and what are the initial condition and the relevant map  $G$ ?
- 71.** Let  $f: \mathbb{R} \rightarrow \mathbb{R}$  be a differentiable function. Assume, for simplicity, that  $f'(x) \neq 0$  for all  $x$ . The Newton-Raphson method for finding a solution of the equation  $f(x) = 0$  is to start with some initial approximation  $x_0$  to a solution and then, once the  $n$ th approximation  $x_n$  has been obtained, to define the next approximation  $x_{n+1}$  by

$$x_{n+1} = x_n - f(x_n)/f'(x_n).$$

For the given  $x_0$ , the method thus provides a sequence  $\langle x_n \rangle_{n=0}^{\infty}$  of approximations.

- (a) What form of recursion—ordinary, primitive, or complete—is being used in the Newton-Raphson method, and what are the relevant initial condition and the map  $G$ ?
- (b) Define  $g: \mathbb{R} \rightarrow \mathbb{R}$  by  $g(x) = x - f(x)/f'(x)$ . Express the sequence  $\langle x_n \rangle_{n=0}^{\infty}$  obtained by the Newton-Raphson method in terms of iterates of the function  $g$ .
- 72.** Let  $X$  be a given set.  $\langle A_i \rangle_{i=1}^{\infty}$  be a sequence of sets.
- (a) Give a recursive definition of the union  $\bigcup_{i=1}^n A_i$  of a family of  $n$  subsets  $A_i$  of  $X$ , where  $n = 1, 2, 3, \dots$
- (b) Make explicit which form of recursion—ordinary, primitive, or complete—you are using and what the relevant map  $G$  is.
- (c) Prove by induction that each such union, so defined, is the same as the set obtained from [Definition 0.17](#) for the case of index set  $I = \{1, 2, \dots, n\}$ .

prob:powers-of-rel

- 73.** Let  $\alpha$  be a relation in a set  $X$  to itself.
- (a) Give a recursive definition of the powers  $\alpha^n$  for  $n = 0, 1, 2, \dots$  of  $\alpha$  in such a way that, in particular,  $\alpha^2 = \alpha \circ \alpha$ , the composite of relations in the sense of [Exercise 18](#). *Note:* This generalizes [Example 0.33](#).
- (b) Prove that  $\alpha^m \circ \alpha^n = \alpha^{m+n}$  for all  $m, n \in \mathbb{N}$ .
- (c) If negative powers of  $\alpha$  are defined by  $\alpha^{-n} = (\alpha^{-1})^n$  for  $n = 1, 2, 3, \dots$ , is it true that  $\alpha^m \circ \alpha^n = \alpha^{m+n}$  for all  $m, n \in \mathbb{Z}$ ?

**74.** The Fibonacci numbers  $F_0, F_1, F_2, \dots$  are defined recursively by:

$$\begin{cases} F_0 = F_1 = 1, \\ F_{n+1} = F_n + F_{n-1} \end{cases} \quad (n = 1, 2, 3, \dots).$$

infinite set  
infinite set  
number of elements

Notice that there are two initial conditions!

- (a) Use the initial conditions and recurrence relation for the Fibonacci numbers to calculate  $F_2, F_3, F_4$ , and  $F_5$ .
  - (b) Formulate an existence and uniqueness theorem about this kind of recursion, where there are two initial conditions and each entry in the sequence depends upon the two immediately preceding two entries.
  - (c) Deduce from [Theorem 0.32](#), [Corollary 0.34](#), or [Theorem 0.35](#) the result you formulated in (b).
- 75.** Let  $a$  and  $b$  be given elements of sets  $A$  and  $B$ , respectively, and let  $H: A \times B \rightarrow A$  and  $K: A \times B \rightarrow B$  be given maps. Prove that there exist unique maps  $h: \mathbb{N} \rightarrow A$  and  $k: \mathbb{N} \rightarrow B$  such that:

$$\begin{aligned} h(0) &= a, & k(0) &= b, \\ h(n+1) &= H(h(n), k(n)), & k(n+1) &= K(h(n), k(n)) \quad (n \in \mathbb{N}). \end{aligned}$$

**76.** Prove the theorem on Principle of Complete Recursion ([0.35](#)).

## 0.6 Countability

sec:countable

Some sets have so small a number of elements that their elements can be “counted.” The underlying idea is to regard two sets as having the same “number” of elements when the elements of one of the sets can be paired up with the elements of the other—in other words, when there is a one-to-one correspondence ([page 28](#)) between the two sets.

### Finite and infinite sets

subsec:finite

def:finite **0.36 Definition.** A set  $X$  is said to be **finite** when either  $X = \emptyset$  or else there exists, for some positive integer  $n$ , a bijection from the set  $\{1, 2, \dots, n\}$  onto  $X$ . And  $X$  is said to be **infinite** when it is not finite.

In particular,  $\{1, 2, \dots, n\}$  is finite for each positive integer  $n$ . Obviously a set will be finite (respectively, infinite) when there is a bijection from it onto some finite (respectively, infinite) set.

The following lemma will allow us to specify the “number of elements” in a finite set.

lem:num-elts-finite-unique

**0.37 Lemma.** If  $m$  and  $n$  are distinct positive integers, then there is no bijection from  $\{1, 2, \dots, m\}$  to  $\{1, 2, \dots, n\}$ .

**Proof.** We use induction on  $n$  to prove that, for each integer  $m > n$ , there is no bijection from  $\{1, 2, \dots, m\}$  to  $\{1, 2, \dots, n\}$ .

Base step:  $n = 1$ . If  $m > 1$ , there cannot exist even an injection  $\{1, 2, \dots, m\} \rightarrow \{1\}$ .

number of elements  
cardinality of finite set

**Inductive step:** Now let  $n \geq 1$ . Assume that, for all  $m > n$ , there exists *no* bijection  $\{1, 2, \dots, m\} \rightarrow \{1, 2, \dots, n\}$ . Let  $m > n + 1$  be arbitrary and just suppose there exists some bijection

$$h: \{1, 2, \dots, n, n+1, \dots, m\} \rightarrow \{1, 2, \dots, n, n+1\}.$$

If  $h(m) = n + 1$ , then the domain-codomain restriction  $\{1, 2, \dots, m-1\} \rightarrow \{1, 2, \dots, n\}$  of  $h$  would be a bijection, contrary to the inductive assumption. Hence  $h(j) = n + 1$  for some  $j \in \{1, 2, \dots, m-1\}$ . Then the modified map

$$h': \{1, 2, \dots, n, n+1, n+2\} \rightarrow \{1, 2, \dots, n, n+1\}$$

given by

$$h'(i) = \begin{cases} h(i) & \text{if } i \neq n+2 \text{ and } h(i) \neq n+1, \\ h(n+2) & \text{if } h(i) = n+1, \\ n+1 & \text{if } i = n+2 \end{cases}$$

would also be a bijection, but with

$$h'(n+2) = n+1.$$

Hence the map obtained from  $h'$  by restricting its domain to  $\{1, 2, \dots, n, n+1\}$  and its codomain to  $\{1, 2, \dots, n\}$  would be a bijection, contrary to the inductive assumption that no such bijection exists.  $\square$

In particular, for a positive integer  $n$  there is *no* bijection from  $\{1, 2, \dots, n, n+1\}$  to  $\{1, 2, \dots, n\}$ .

In view of [Lemma 0.37](#), if a nonempty set  $X$  is finite, there exists a *unique* positive integer  $n$  for which  $X$  is in one-to-one correspondence with  $\{1, 2, \dots, n\}$ ; this unique  $n$  is called the **number of elements in  $X$**  and may be denoted by  $\#(X)$  or simply  $\#X$ . Since the empty set has no elements whatsoever, we also define  $\#(\emptyset) = 0$ .

It is easy to exhibit finite sets besides  $\emptyset$  and  $\{1, 2, \dots, n\}$ : any singleton  $\{x\}$  is finite, as is any doubleton  $\{x, y\}$ . It is somewhat harder to establish the existence of infinite sets.

**0.38 Example.** The set  $\mathbb{N}$  of all natural numbers is infinite. In fact, just suppose that  $\mathbb{N}$  were finite. Since  $\mathbb{N}$  is nonempty, there exists a bijection

$$f: \{1, 2, \dots, n\} \rightarrow \mathbb{N}$$

for some positive integer  $n$ . Define

$$g: \{1, 2, \dots, n, n+1\} \rightarrow \mathbb{N}$$

that shifts each value of  $f$  by 1 to the right—thereby “uncovering”  $0 \in \mathbb{N}$  as no longer a value—and that sends  $n+1$  to 0. In other words:

$$g(i) = \begin{cases} f(i) + 1 & \text{if } i < n+1, \\ 0 & \text{if } i = n+1. \end{cases}$$

Then  $g$  is bijective, so the composite

$$f^{-1} \circ g: \{1, 2, \dots, n, n+1\} \rightarrow \{1, 2, \dots, n\}$$

is a bijection. But this is impossible according to [Lemma 0.37](#).  $\diamond$



The set  $\mathbb{N}^*$  of all positive integers is also infinite, since the map

$$\begin{aligned}\mathbb{N} &\rightarrow \mathbb{N}^* \\ n &\mapsto n + 1\end{aligned}$$

is a bijection.

The next several results give methods for obtaining finite sets from other finite sets. The first is a technical result that will be subsumed in [Proposition 0.41](#).

family!finite  
finite family  
family!infinite  
infinite family

lem:disj-union-finite-sets

**0.39 Lemma (finiteness of union of two disjoint finite sets).** *Let  $A$  and  $B$  be finite sets with  $A$  disjoint from  $B$ . Then  $A \cup B$  is finite.*

**Proof.** There is nothing to prove if  $A = \emptyset$  or  $B = \emptyset$ , so we assume that neither  $A$  nor  $B$  is empty. Since they are finite, there exist positive integers  $m$  and  $n$  and bijections

$$f: \{1, 2, \dots, m\} \rightarrow A, \quad g: \{1, 2, \dots, n\} \rightarrow B.$$

Then the map

$$h: \{1, 2, \dots, m, m+1, m+2, \dots, m+n\} \rightarrow A \cup B$$

given by

$$h(i) = \begin{cases} f(i) & \text{if } 1 \leq i \leq m, \\ g(i-m) & \text{if } m+1 \leq i \leq m+n \end{cases}$$

is a bijection.  $\square$

prop:subset-of-finite

**0.40 Proposition (finiteness of subset of finite set).** *A subset  $X$  of a finite set  $Y$  is finite, and  $\#X \leq \#Y$ .*

**Proof.** It suffices to show that for each  $n \geq 1$  any subset of  $\{1, 2, \dots, n\}$  is finite. We do this by induction on  $n$ .

Base step:  $n = 1$ . The only subsets of  $\{1\}$  are  $\emptyset$  and  $\{1\}$ , both of which are finite.

Inductive step: Now let  $n \geq 1$  and assume that each subset of  $\{1, 2, \dots, n\}$  is finite. Let  $A$  be a subset of  $\{1, 2, \dots, n, n+1\}$ . If  $A \subset \{1, 2, \dots, n\}$ , then already  $A$  is finite by the inductive assumption.

Suppose now  $A \not\subset \{1, 2, \dots, n\}$ . Then  $n+1 \in A$ . By the inductive assumption the subset  $A \setminus \{n+1\}$  of  $\{1, 2, \dots, n\}$  is finite. Now the singleton  $\{n+1\}$  is also finite. From [Lemma 0.39](#) it follows that the set

$$A = (A \setminus \{n+1\}) \cup \{n+1\}$$

is finite as well.  $\square$

According to this proposition, a set is infinite if it contains an infinite set. In view of [Example 0.38](#), then, each of the sets  $\mathbb{Z}$ ,  $\mathbb{Q}$ , and  $\mathbb{R}$  is infinite.

Below we shall speak of a **finite family**, meaning a family indexed by a finite set. (Later we shall need to form the intersection of a finite family of sets, and then we shall understand that the index set is not only finite, but nonempty as well.) Similarly, one may speak of an **infinite family**, meaning a family indexed by an infinite set.

prop:union-finite-family-of-finite-sets

**0.41 Proposition (finiteness of union of finitely many finite sets).** *The union of a finite family of finite sets is itself finite.*

**Proof.** First note that the union of any two (not necessarily disjoint) finite sets  $A$  and  $B$  is finite, because:

$$A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B);$$

the pairwise disjoint sets  $A \setminus B$ ,  $B \setminus A$ , and  $A \cap B$  are finite by [Proposition 0.40](#); and so  $A \cup B$  is finite by two applications of [Lemma 0.39](#).

Now let  $\langle X_i \rangle_{i \in I}$  be a finite family of finite sets. If  $I = \emptyset$ , then  $\bigcup_{i \in I} X_i = \emptyset$  too. So suppose  $I \neq \emptyset$ . Then there is a bijection

$$\sigma\{1, 2, \dots, n\} \rightarrow I$$

for some  $n \geq 1$ . The family  $\langle A_j \rangle_{j=1,2,\dots,n}$  defined by

$$A_j = X_{\sigma(j)} \quad (j = 1, 2, \dots, n)$$

is indexed by  $\{1, 2, \dots, n\}$  and has union

$$\bigcup_{j=1}^n A_j = \bigcup_{i \in I} X_i.$$

Thus it suffices to prove that for each  $n = 1, 2, \dots$ , each family  $\langle A_i \rangle_{i=1,2,\dots,n}$  of  $n$  finite sets has a finite union. To do that, we use induction on  $n$ . The case  $n = 1$  is obvious, and the case  $n = 2$  follows from the very first paragraph of this proof. The inductive step follows from the formula

$$\bigcup_{j=1}^{n+1} A_j = \left( \bigcup_{j=1}^n A_j \right) \cup A_{n+1}$$

with the aid of the first paragraph.  $\square$

prop:finite-product-of-finite-sets

**0.42 Proposition (finiteness of product of finitely many finite sets).** *The product of a nonempty finite family of finite sets is itself finite.*

**Proof.** First we show that the product of two finite sets  $A$  and  $B$  is finite. We have

$$A \times B = \bigcup_{a \in A} (\{a\} \times B).$$

For each  $a \in A$  the map

$$\begin{aligned} \{a\} \times B &\rightarrow B \\ \langle a, b \rangle &\mapsto b \end{aligned}$$

is bijective, and so  $\{a\} \times B$  is finite because  $B$  is. It follows from [Proposition 0.41](#) that  $A \times B$  is finite.

As with [Proposition 0.41](#), to prove the present proposition it suffices to show that for each  $n = 1, 2, \dots$ , each family  $\langle X_i \rangle_{i=1,2,\dots,n}$  of  $n$  finite sets has a finite product. The case  $n = 1$  is obvious, and the case  $n = 2$  follows from what we just proved. To execute the inductive step, use the fact that the map

$$\begin{aligned} \times_{i=1}^{n+1} X_i &\rightarrow \left( \times_{i=1}^n X_i \right) \times X_{n+1} \\ \langle x_1, x_2, \dots, x_n, x_{n+1} \rangle &\mapsto \langle \langle x_1, x_2, \dots, x_n \rangle, x_{n+1} \rangle \end{aligned}$$

is a bijection.  $\square$

In particular, if  $S$  and  $I$  are finite sets, then the power set  $S^I$  is finite.

For our final result concerning finite sets we shall need the Well-ordering Principle for  $\mathbb{N}$ (0.3): Each nonempty subset of  $\mathbb{N}$  has a least member.

prop:image-of-finite-set

**0.43 Proposition (finiteness of range of surjection from ffinite set).** *Let  $X$  be a finite set and let  $f: X \rightarrow Y$  be a surjection. Then  $Y$  is also finite and  $\#Y \leq \#X$ .*

**Proof.** If  $X = \emptyset$ , then  $Y = \emptyset$ , too. So suppose now  $X \neq \emptyset$ . Without of generality we may assume that

$$X = \{1, 2, \dots, n\}$$

for some positive integer  $n$ . Define a map

$$g: Y \rightarrow X$$

as follows. For each  $y \in Y$ , its inverse image  $f^{-1}(y) \neq \emptyset$ , and take  $g(y)$  to be the least element of  $f^{-1}(y)$ . Then  $g$  is an injection; the range  $g(Y)$  is finite by Proposition 0.40; and hence  $Y$  is finite.  $\square$

prop-power-set-finite-set

**0.44 Proposition (finiteness of power set of finite set).** *If  $X$  is a finite set, then its power set  $\mathcal{P}(X)$  is finite and*

$$\#(\mathcal{P}(X)) = 2^{\#(X)}.$$

**Proof.** See Exercise 84.  $\square$

### Denumerable, countable, and uncountable sets

subsec:denumerable

We now distinguish between two kinds of infinite sets—those that are as “small” as  $\mathbb{N}$  and those that are not.

def:denumerable

**0.45 Definition.** A set  $X$  is said to be **denumerable** if there exists some bijection

$$f: \mathbb{N} \rightarrow X,$$

**countable** if it is finite or denumerable, and **uncountable** if it is not countable.

**Caution!** Some authors use ‘countable’ or ‘countably infinite’ to mean what we are calling ‘denumerable’.

Because  $\mathbb{N}$  is infinite, each denumerable set is infinite. Hence **a set is infinite if and only if it is either denumerable or else uncountable**. Of course, each set is either countable or uncountable.

The set  $\mathbb{N}^*$  of all positive integers is denumerable because the map  $n \mapsto n + 1$  from  $\mathbb{N}$  to  $\mathbb{N}^*$  is a bijection.<sup>3</sup> Similarly, the set  $\{2, 4, 6, \dots\}$  of all even positive integers is denumerable.

A set will obviously be denumerable (respectively, countable, uncountable) if there is a bijection from it to some denumerable (respectively, countable, uncountable) set.

<sup>3</sup>This bijection may be described fancifully in terms of “Hilbert’s hotel”—more properly, “the paradox of the Grand Hotel,” which was presented in a popular lecture by David Hilbert. Imagine a hotel with infinitely many rooms, consecutively numbered  $0, 1, 2, \dots$ , and with every room occupied. To accommodate a newly arrived guest, the desk clerk has each occupant move to the next room, thereby leaving vacant room 0, into which the new guest is placed. (See also Footnote 6, page 119.)

fn:hilbert-hotel

denumerable set!sequence@and sequence  
sequence!denumerable set@and denumerable set  
infinite set!denumerable subset@and denumerable subset  
A bijection  $f: \mathbb{N} \rightarrow X$  is nothing but a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  of elements of  $X$  such that  $X = \{x_n : n \in \mathbb{N}\}$  and  $x_i \neq x_j$  whenever  $i \neq j$ .

Hence a set is denumerable if and only if it is the set of values of some sequence of *distinct* elements.

prop:inf-subset-of-countable

**0.46 Proposition (denumerability of infinite subset of countable set).** *An infinite subset of a countable set is denumerable.*

**Proof.** Let  $A$  be an infinite subset of a countable set  $X$ . By [Proposition 0.40](#) the set  $X$  must be denumerable. Without loss of generality we may assume that  $X = \mathbb{N}$ .

Let  $x_0$  the least element of  $A$ .

Since the subset  $A \setminus \{x_0\}$  of  $\mathbb{N}$  is still infinite, it is nonempty; let  $x_1$  be the least element of  $A \setminus \{x_0\}$ .

Next, since  $A \setminus \{x_0, x_1\}$  is also infinite, it is nonempty; let  $x_2$  be the least element of  $A \setminus \{x_0, x_1\}$ .

In general, once  $x_0, x_1, \dots, x_n$  have already been constructed, let  $x_{n+1}$  be the least element of  $A \setminus \{x_0, x_1, \dots, x_n\}$ .

By construction,  $x_i \neq x_j$  whenever  $i \neq j$ . Induction on  $m$  establishes that

$$m \in A \implies m = x_n \text{ for some } n \in \mathbb{N}.$$

Hence  $A = \{x_n : n \in \mathbb{N}\}$ .  $\square$

In particular, each subset of  $\mathbb{N}$  is either finite or else denumerable.

An obvious corollary of [Proposition 0.46](#) is that each subset of a countable set is countable.

By [Example 0.38](#) and [Proposition 0.40](#), a set is infinite if it contains some denumerable subset. Conversely, we have the following proposition.

prop:inf-set-has-den-subset

**0.47 Proposition (infinite set has a denumerable subset).** *Let  $X$  an infinite set. Then there exists a denumerable subset of  $X$ .*

**Proof.** Arbitrarily choose an element  $x_0 \in X$ . The set  $X \setminus \{x_0\}$  is still infinite and therefore nonempty; arbitrarily choose an  $x_1 \in X \setminus \{x_0\}$ . Next, the set  $X \setminus \{x_0, x_1\}$  is still infinite and therefore nonempty; choose an  $x_2 \in X \setminus \{x_0, x_1\}$ . Continuing in this way we obtain a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in  $X$  with  $x_i \neq x_j$  whenever  $i \neq j$ . Then  $D = \{x_n : n \in \mathbb{N}\}$  is the desired denumerable subset of  $X$ .  $\square$

Evidently the preceding proof uses the Principle of Complete Recursion ([0.35](#)). Still, the argument is not quite rigorous: it really requires a guarantee that *all* the necessary choices be made “simultaneously.” And that guarantee is provided by the Axiom of Choice, specifically, by its choice function version ([Corollary 0.28](#)).

prop:image-of-countable-set

**0.48 Proposition (countability of range of a map on a countable set).** *Let  $X$  be a countable set and let  $f: X \rightarrow Y$  be a surjection. Then  $Y$  is also countable.*

**Proof.** If  $X$  is finite, then by [Proposition 0.43](#) the set  $Y$  is finite and hence countable. Suppose now that  $X$  is denumerable. Without loss of generality we may suppose that  $X = \mathbb{N}$ . For each  $y \in Y$ , the subset  $f^{-1}(y)$  of  $\mathbb{N}$  is nonempty, and we define  $g(y)$  to be the least member of  $f^{-1}(y)$ . The map  $g: Y \rightarrow \mathbb{N}$  so obtained is injective. Then the range  $g(Y)$  of  $g$  is countable, and so  $Y$ , too, is countable.  $\square$

Another way of stating the preceding proposition is that the image of a countable set under a map is itself countable.

From [Proposition 0.48](#) it follows that a nonempty set  $Y$  is countable if and only if its elements can be arranged in a sequence—that is, there is a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in  $Y$  such that  $Y = \{x_n : n \in \mathbb{N}\}$ .

countable set!sequence@and seque  
sequence!countable set@and count  
countable family

lem:N-times-N-denumerable

**0.49 Lemma.** *The product set  $\mathbb{N} \times \mathbb{N}$  is denumerable.*

**Proof.** Because it contains the infinite set  $\mathbb{N} \times \{0\}$ , the set  $\mathbb{N} \times \mathbb{N}$  is already infinite. To show that it is countable as well, we need only construct an injection  $\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ .

We claim that

$$\begin{aligned} \mathbb{N} \times \mathbb{N} &\rightarrow \mathbb{N} \\ \langle m, n \rangle &\mapsto 2^m 3^n \end{aligned}$$

is an injection. In fact, suppose  $\langle m, n \rangle, \langle i, j \rangle \in \mathbb{N} \times \mathbb{N}$  with

$$2^m 3^n = 2^i 3^j.$$

We shall deduce that  $m = i$  and  $n = j$ . If  $m \neq i$ , say  $m > i$ , then

{pf-eq:two-three-powers} (\*)

$$2^{m-i} 3^n = 3^j,$$

which is impossible since the left-hand side  $2^{m-i} 3^n$  is even whereas the right-hand side  $3^j$  is odd. Hence  $m = i$ . Then equation (\*) reduces to  $3^n = 3^j$  whence also  $n = j$ .  $\square$

When we refer to a **countable family** below, we shall mean a family that is indexed by a countable set.

m:countable-union-of-countable-sets

**0.50 Theorem (countability of union of countably many countable sets).** *The union of a countable family of countable sets is itself countable.*

**Proof.** Let  $\langle X_i \rangle_{i \in I}$  be a countable family of countable sets. If  $I = \emptyset$ , there is nothing to prove, so we assume that  $I \neq \emptyset$ . Without loss of generality we may assume that  $I \subset \mathbb{N}$ . In fact, we may even assume that  $I = \mathbb{N}$ , for if we choose some  $j \in I$  and define

$$X_n = X_j \quad (n \in \mathbb{N} \setminus I),$$

then

$$\bigcup_{n \in \mathbb{N}} X_n = \bigcup_{i \in I} X_i.$$

For each  $n \in \mathbb{N}$  there is a surjection

$$f_n : \mathbb{N} \rightarrow X_n.$$

Define

$$f : \mathbb{N} \times \mathbb{N} \rightarrow \bigcup_{n \in \mathbb{N}} X_n$$

by

$$f(m, n) = f_n(m).$$

Clearly  $f$  is surjective. That the union of  $\langle X_n \rangle_{n \in \mathbb{N}}$  is countable now follows from [Proposition 0.48](#) and [Lemma 0.49](#).  $\square$

Cantor, Georg  
cor:union-denumerable

**0.51 Corollary.** *The union of a nonempty countable family of denumerable sets is itself denumerable.*

A particular consequence of [Theorem 0.50](#) is that *the set  $\mathbb{Z}$  of all integers is denumerable*. In fact, the set  $\mathbb{Z}$  is infinite because it contains the infinite set  $\mathbb{N}$ , and it is countable because

$$\mathbb{Z} = \mathbb{N} \cup \{-n : n \in \mathbb{N}^*\}.$$

Actually, it is easy to construct an explicit bijection from  $\mathbb{N}$  to  $\mathbb{Z}$  (see [Exercise 88](#)).

r:product-finite-family-countable-sets

**0.52 Corollary (countability of product of finitely many countable sets).** *The product of a finite family of countable sets is countable.*

**Proof.** The product of two countable sets  $A$  and  $B$  is countable because

$$A \times B = \bigcup_{a \in A} (\{a\} \times B).$$

The general result follows by induction.  $\square$

In particular, if  $X$  is a countable set, then  $X^n$  is countable for each  $n = 1, 2, 3, \dots$

From [Corollary 0.52](#) we can deduce now that *the set  $\mathbb{Q}$  of all rational numbers is denumerable*. In fact, the set  $\mathbb{Q}$  is infinite because it contains  $\mathbb{N}$ , and it countable because the map

$$\begin{array}{ccc} \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) & \rightarrow & \mathbb{Q} \\ \langle m, n \rangle & \mapsto & m/n \end{array}$$

is surjective with domain the product of the countable sets  $\mathbb{Z}$  and  $\mathbb{Z} \setminus \{0\}$ .

**Caution!** The product of an infinite family of countable sets need not be countable!

uncountable-product-of-doubletons

**0.53 Example.** The set  $\{0, 1\}^{\mathbb{N}^*}$  consisting of all sequences of 0s and 1s (indexed by  $\{1, 2, 3, \dots\}$ ) is uncountable. In fact, just suppose this set is countable. Then

$$\{0, 1\}^{\mathbb{N}^*} = \{x_n : n = 1, 2, 3, \dots\}$$

for some sequence  $\langle x_n \rangle_{n=1,2,3,\dots}$ . For each  $n$ , the member  $x_n$  of  $\{0, 1\}^{\mathbb{N}^*}$  is itself a sequence

$$x_n = \langle x_{n,i} \rangle_{i=1,2,3,\dots}$$

Here the doubly-subscripted notation  $x_{n,i}$  is shorthand for  $(x_n)_i$ , that is, the  $i$ th coordinate of the  $n$ th sequence  $x_n$ .

We are now going to use Georg Cantor's original **diagonal argument** to obtain a sequence  $y$  of 0s and 1s that differs from  $x_n$  for each  $n$ , thereby deriving a contradiction.

Consider the array shown below, in which the values (0 or 1) of the sequence  $x_n$  are written on the  $n$ th line.

$$\begin{array}{ccccccc} x_{1,1} & x_{1,2} & x_{1,3} & \cdots & x_{1,n} & \cdots \\ x_{2,1} & x_{2,2} & x_{2,3} & \cdots & x_{2,n} & \cdots \\ x_{3,1} & x_{3,2} & x_{3,3} & \cdots & x_{3,n} & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \\ x_{n,1} & x_{n,2} & x_{n,3} & \cdots & x_{n,n} & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \ddots \end{array}$$

Form the sequence  $\langle x_{1,1}, x_{2,2}, x_{3,3}, \dots \rangle$  by taking the elements on the main diagonal of this binary expansion array. Define the sequence

$$y = \langle y_1, y_2, y_3, \dots \rangle$$

of 0s and 1s by

$$y_n = \begin{cases} 1 & \text{if } x_{n,n} = 0, \\ 0 & \text{if } x_{n,n} = 1. \end{cases}$$

For each  $n \geq 1$ , we have  $y \neq x_n$  because the  $n$ th coordinate  $y_n$  is not the same as the  $n$ th coordinate  $x_{n,n}$  of  $x_n$ .  $\diamond$

Likewise, the set  $\{0, 1\}^{\mathbb{N}}$  is uncountable. More generally, a power  $\{0, 1\}^I$  is uncountable whenever  $I$  is denumerable.

By using binary expansion of real numbers—see the subsection “Base expansion” (page 79), below—it is possible to deduce from [Example 0.53](#) that the set  $\mathbb{R}$  of all real numbers is uncountable. We shall give a different proof of this fact ([Corollary 0.85](#)) in [Section 0.8](#).

### EXERCISES FOR SECTION 0.6

- |                                   |  |
|-----------------------------------|--|
| prob:num-elts-union-2-finite      | <p><b>77. (a)</b> Improve <a href="#">Lemma 0.39</a> by showing that if <math>A</math> and <math>B</math> are disjoint finite sets, then <math>\#(A \cup B) = (\#A) + (\#B)</math>.</p> <p><b>(b)</b> Deduce that if <math>A</math> and <math>B</math> are any finite sets, then <math>\#(A \cup B) \leq (\#A) + (\#B)</math>.</p>   |
| prob:num-elts-subset-finite       | <p><b>78.</b> Improve <a href="#">Proposition 0.40</a> by showing that if <math>Y</math> is a proper subset of a finite set <math>X</math>, then <math>\#Y &lt; \#X</math>.</p>  |
| prob:num-elts-finite-union-finite | <p><b>79. (a)</b> Improve <a href="#">Proposition 0.41</a> by showing that if <math>\langle X_i \rangle_{i \in I}</math> is a finite family of finite sets, then <math>\#(\bigcup_{i \in I} X_i) \leq \sum_{i \in I} \#X_i</math>.</p> <p><b>(b)</b> Show, further, that if <math>\langle X_i \rangle_{i \in I}</math> is a finite family of <i>disjoint</i> finite sets, then <math>\#(\bigcup_{i \in I} X_i) = \sum_{i \in I} \#X_i</math>.</p>  |
| prob:num-elts-finite-prod-finite  | <p><b>80.</b> Improve <a href="#">Proposition 0.42</a> by showing that if <math>\langle X_i \rangle_{i \in I}</math> is a nonempty finite family of finite sets, then <math>\#(\prod_{i \in I} X_i) = \prod_{i \in I} \#X_i</math>.</p> <p><i>Note:</i> This result shows, in particular, that if <math>S</math> and <math>I</math> are finite sets, then <math>\#(S^I) = (\#S)^{\#I}</math>.</p>  |
| prob:num-elts-image-finite        | <p><b>81.</b> Complete the proof of <a href="#">Proposition 0.43</a> by showing that if <math>X</math> is a finite set and <math>f: X \rightarrow Y</math> is a surjection, then <math>\#Y \leq \#X</math>.</p> <p><b>82.</b> A certain collection <math>\mathcal{A}</math> of sets has the property that <math>A \cap B \in \mathcal{A}</math> whenever <math>A \in \mathcal{A}</math> and <math>B \in \mathcal{A}</math>. Show that the intersection of any finite family of members of <math>\mathcal{A}</math> is again a member of <math>\mathcal{A}</math>.</p> <p><b>83.</b> Let <math>\mathcal{T} = \{A : A \subset \mathbb{R}, \mathbb{R} \setminus A \text{ is finite}\}</math>. Prove:</p> <p><b>(a)</b> If <math>\langle A_i \rangle_{i \in I}</math> is a finite family with <math>A_i \in \mathcal{T}</math> for each <math>i \in I</math>, then also <math>\bigcap_{i \in I} A_i \in \mathcal{T}</math>.</p> <p><b>(b)</b> If <math>\langle A_i \rangle_{i \in I}</math> is any family with <math>A_i \in \mathcal{T}</math> for each <math>i \in I</math>, then also <math>\bigcup_{i \in I} A_i \in \mathcal{T}</math>.</p> |
| prob:pf-power-set-finite-set      | <p><b>84.</b> Prove <a href="#">Proposition 0.44</a> in each of these two ways:</p> <p><b>(a)</b> Directly, using induction.</p> <p><b>(b)</b> By exploiting <a href="#">Proposition 0.16</a>.</p> <p><b>85.</b> Let <math>f: X \rightarrow X</math> be a map from a finite set <math>X</math> into <math>X</math>. Prove that <math>f</math> is bijective if and only if <math>f</math> is injective, or equivalently, if and only if <math>f</math> is surjective.</p>   |

**algebraic number** 86. Verify that each of the following sets is denumerable:

**transcendental number**

- (a) The set  $\mathbb{N} \setminus \{n\}$ , where  $n \in \mathbb{N}$ .
- (b) The set  $\mathbb{N} \setminus \{m, n\}$ , where  $m, n \in \mathbb{N}$  with  $m \neq n$ .
- (c) The set  $\mathbb{N} \setminus F$ , where  $F$  is any finite subset of  $\mathbb{N}$ .

87. Prove that a set  $X$  is countable if and only if for some subset  $A$  of  $\mathbb{N}$  there is some bijection from  $A$  to  $X$ .

prob:Z-is-denumerable 88. Construct an explicit bijection from  $\mathbb{N}$  to  $\mathbb{Z}$ . (*Hint:* Write  $\mathbb{Z}$  as the union of  $\{n \in \mathbb{Z} : n < 0\}$  and  $\{n \in \mathbb{Z} : n \geq 0\}$ .)

89. Let  $X$  be a *nonempty* set. Prove that the following conditions are equivalent:

- (i) The set  $X$  is countable.
- (ii) There exists some surjection  $\mathbb{N} \rightarrow X$ .
- (iii) There exists some injection  $X \rightarrow \mathbb{N}$ .

90. Show that each of the following sets is infinite and determine whether it is denumerable or else uncountable:

- (a) The set of all nonvertical lines in  $\mathbb{R}^2$  having rational slopes.
- (b) The set  $\{x \in \mathbb{R}^2 : x^2 \in \mathbb{Q}\}$ .
- (c) The set  $\{x \in \mathbb{R} : x = m/2^n \text{ for some } m \in \mathbb{Z}, n \in \mathbb{N}\}$ .

91. Must the union of a denumerable family of countable sets be denumerable?

92. For an arbitrary set  $S$ , construct a bijection between  $(S^{\mathbb{N}})^{\mathbb{N}}$  and  $S^{(\mathbb{N} \times \mathbb{N})}$ .

ob:algebraic-numbers-denumerable 93. A real number is said to be **algebraic** if it is a root of some polynomial

$$\{\text{eq:poly-for-algebraic}\} (*) \quad p = a_0x^n + a_1x^{n-1} + \cdots + a_{n-2}x^2 + a_{n-1}x + a_n$$

having integers as its coefficients  $a_0, a_1, \dots, a_{n-2}, a_{n-1}, a_n$ .

- (a) Show that each rational number is algebraic. (Thus algebraic numbers are generalizations of rational numbers.)
- (b) Exhibit three algebraic numbers that are not rational.
- (c) Prove that the set  $A$  of all algebraic numbers is denumerable. [*Hint:* You may assume that a polynomial of degree  $n$  has at most  $n$  real roots. If  $p$  is a polynomial given by  $(*)$ , define its “height”  $h(p)$  to be the positive integer

$$h(p) = n + |a_0| + |a_1| + \cdots + |a_n|.$$

How many polynomials  $p$  are there with height  $h(p) = m$  for a given positive integer  $m$ ?

- (d) Deduce that the set  $\mathbb{R} \setminus A$  of all **transcendental numbers** is uncountable.  
*Note:* This establishes, in particular, the existence of transcendental numbers, but it does not provide so much as a single specific example of such a number. By completely different methods it can be proved that both  $\pi$  and  $e$  are transcendental.

94. Prove:

- (a) The collection of all finite subsets of  $\mathbb{N}$  is denumerable.
- (b) The collection of all denumerable subsets of  $\mathbb{N}$  is uncountable.

95. The proof that an infinite set contains a denumerable subset ([Proposition 0.47](#)) in reality uses the Axiom of Choice—choice function version ([0.28](#)). What, in the notation of [Corollary 0.28](#), is the relevant collection  $\mathcal{A}$ ?



## 0.7 Ordering relations

sec:order-relations

Each of the sets  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$ , and  $\mathbb{R}$  has, besides the operations of addition and multiplication that provide it with its algebraic structure, an “ordering relation”  $\leq$  that allows comparing its members with other members. Before stating the fundamental properties of ordering of these number systems, we take a look at such relations in general.

precedes  
predecessor  
succeeds  
successor  
divisibility relation  
divides

### Preorderings

subsec:preorder

To be deemed an “ordering relation”, at the very least a relation ought to have the two properties listed in the following definition.

**0.54 Definition.** A relation  $\leq$  in a set  $X$  is said to **preorder**  $X$  and is called a **preordering of  $X$**  (or, more briefly, a **preorder of  $X$** ) if it has the following two properties:

(I1) Reflexitivity: For each  $x \in X$ ,

$$x \leq x.$$

(I2) Transitivity: For all  $x, y \in X$ ,

$$x \leq y \quad \text{and} \quad y \leq z \implies x \leq z.$$

A set together with a specific preordering of it is called a **preordered set**.

Each of the sets  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$ , and  $\mathbb{R}$  is a preordered set—and more!—with respect to its usual ordering  $\leq$ .

For a general preordering relation we use the symbol ‘ $\preceq$ ’ (with the “curly” less-than  $\prec$ ) instead of the more familiar ‘ $\leq$ ’ (with the ordinary ‘ $<$ ’), lest one ascribe additional properties to the relation that are typically connoted by  $\leq$  (such as “antisymmetry” or “comparability”—see (I3), page 66, and (I4), page 73).

When  $x \preceq y$  we say that  $x$  **precedes**  $y$ , and we call  $x$  a **predecessor of  $y$** ; moreover, we say that  $y$  **succeeds**  $x$ , and we call  $y$  a **successor of  $x$** .

**0.55 Examples.** (1) Take  $X = \mathbb{Z}$ , the set of all integers. Define the relation  $\leq$  in  $\mathbb{Z}$  by

$$m \leq n \iff m^2 \leq n^2$$

where the relation  $\leq$  on the right-hand side above is the usual ordering of  $\mathbb{Z}$ . Then  $\leq$  preorders  $\mathbb{Z}$ .

For this preordering, the inequalities  $m \leq n$  and  $n \leq m$  can both hold yet it does *not* necessarily follow that  $m = n$ . For example, take  $m = -1$  and  $n = 1$ . (In the terminology of the next subsection, the preordering here does not *not* “partially order”  $\mathbb{Z}$ .)

(2) On the set  $\mathbb{Z}$  of integers, define the **divisibility relation**  $|$  by

$$m | n \iff n = k m \text{ for some } k \in \mathbb{Z}.$$

For example,  $2 | 6$ ; also,  $m | 0$  for every  $m \in \mathbb{Z}$ , and in particular  $0 | 0$ ; but  $4 \nmid 6$ . When  $m | n$ , we say that  $m$  **divides**  $n$  and that  $n$  is **divisible by  $m$** .

(If we imposed for divisibility the stronger condition that  $n = k m$  for some *nonzero*  $k \in \mathbb{Z}$ , then the relation  $|$  would no longer be reflexive.)

ex:preordering

ex:divisibility

preordering!induced  
preordering!inherited

relation!irreflexive  
irreflexive relation  
relation!transitive  
transitive relation  
strict preordering  
preordering!strict  
weak preordering  
preordering!weak

For this preordering, the inequalities  $m \leq n$  and  $n \leq m$  together *do* imply that  $m = n$ . (In the terminology of the next subsection, the relation  $|$  does “partially order”  $\mathbb{N}$ .)

- (3) For a positive integer  $k$ , denote by  $P(k)$  the number of primes that divide  $k$ . For example,  $P(4840) = P(2^3 \cdot 5 \cdot 11^2) = 3$ . Then the relation  $\leq'$  on the set  $\mathbb{N}^*$  of positive integers defined by

$$m \leq' n \iff P(m) \leq P(n) \quad (m, n \in \mathbb{N}^*).$$

is a preordering of  $\mathbb{N}^*$ .

For this preordering, the inequalities  $m \leq' n$  and  $n \leq' m$  can both hold and yet  $m \neq n$ : for example,  $P(2) = 1 = P(3)$ .

This preordering has the additional property that for each  $m, n \in \mathbb{N}^*$ , there is some  $k \in \mathbb{N}^*$  with  $m \leq k$  and  $n \leq k$ . It is an example of a kind of ordering—a “direction”—that will be used in [Section 3.5](#) to study convergence.

ex:preorder-on-subset

- (4) Any subset  $Y$  of a preordered set  $X$  becomes a preordered set in its own right if we restrict the preordering of  $X$  just to elements of  $Y$ .

In more detail, let  $\leq$  preorder a set  $X$  and let  $Y \subset X$ . Then the relation  $\leq_Y$  in  $Y$  defined by

$$x \leq_Y y \iff x \leq y \quad (x, y \in Y)$$

is a preordering of  $Y$ , said to be **induced by**, or **inherited from**  $\leq$ . As a set, the induced preordering  $\leq_Y$  is just the subset

$$\leq \cap (Y \times Y)$$

of  $Y \times Y$ . Ordinarily we use the same symbol  $\leq$  for an induced preordering as for the preordering that induces it.

- (5) The trivial (empty) relation in a *nonempty* set  $X$  is transitive but not reflexive, hence does *not* preorder  $X$ .  $\diamond$

For real numbers  $x$  and  $y$ , the “strict inequality”  $x < y$  means that  $x \leq y$  but  $x \neq y$ . The meaning of strict inequality for a general preordering has to be formulated a bit differently.

strict ordering

**0.56 Definition.** Given a preorder  $\leq$  on  $X$ , the associated **strict preordering** is the relation  $<$  in  $X$  defined by

$$x < y \iff x \leq y \text{ and } y \not\leq x.$$

The strict preordering  $<$  associated with a preordering  $\leq$  of a set  $X$  has the following properties:

def:irreflexive

- (SI1) Irreflexivity: *For each*  $x \in X$ ,

$$x \not< x.$$

- (SI2) Transitivity: *For all*  $x, y \in X$ ,

$$x < y \text{ and } y < z \implies x < z.$$

Conversely, given a relation  $<$  in a set  $X$  that is irreflexive and transitive, there is an associated preordering  $\leq$  on  $X$  defined by

$$x \leq y \iff x < y \text{ or } x = y,$$

called the associated **weak ordering** of  $X$ .

When  $x < y$  we say that  $x$  **strictly precedes**  $y$  and call  $x$  a **strict predecessor** of  $y$ ; moreover, we say that  $y$  **strictly succeeds**  $x$  and call  $y$  a **strict successor** of  $x$ .

For a preordering  $\leq$  of  $X$ , its reverse  $\geq$ , given by

$$x \geq y \iff y \leq x \quad (x, y \in X),$$

is again a preordering; the reverse  $>$  of the associated strict preordering  $<$ , given by

$$x > y \iff y < x \quad (x, y \in X),$$

is the strict preordering associated with  $\geq$ .

Given a preordering  $\leq$  of  $X$  and elements  $x, y \in X$ , we say that  $x$  **immediately precedes**  $y$ , and call  $x$  an **immediate predecessor** of  $y$ , when  $x < y$  and there is *no*  $z \in X$  such that  $x < z < y$ . With the obvious definition, we speak of **immediately succeeds** and refer to an **immediate successor**. For example, in  $\mathbb{N}$  with its usual ordering, 5 succeeds 2 and 3 is the immediate successor of 2; in  $\mathbb{R}$  with its usual ordering, no element has an immediate predecessor or an immediate successor.

def:order-preserving

**0.57 Definition.** Let  $X$  and  $Y$  be sets sets preordered by relations  $\leq_X$  and  $\leq_Y$ , respectively. A map  $f: X \rightarrow Y$  is said to **preserve order**, and to be **order-preserving**, when for all  $x_1, x_2 \in X$ ,

$$x_1 \leq_X x_2 \implies f(x_1) \leq_Y f(x_2).$$

Similarly,  $f$  is said to **reverse order**, and to be **order-reversing**, when for all  $x_1, x_2 \in X$ ,

$$x_1 \leq_X x_2 \implies f(x_1) \geq_Y f(x_2).$$

ler-preserving-map-preordered-sets

**0.58 Examples.** (1) Provide  $\mathbb{R}$  with its usual ordering. Then the maps  $x \mapsto x^3$ ,  $x \mapsto -x$ , and  $x \mapsto x^2$  of  $\mathbb{R} \rightarrow \mathbb{R}$  are, respectively, order-preserving, order-reversing, and neither order-preserving nor order-reversing.

-divisible-to-number-prime-divisors

(2) Let  $\leq$  be the preordering induced on  $\mathbb{N}^*$  by the divisibility preordering  $|$  of  $\mathbb{Z}$  [Examples 0.55 (2)], so that

$$m \leq n \iff m \text{ is a divisor of } n \quad (m, n \in \mathbb{N}^*),$$

and let  $\leq'$  be the preordering of  $\mathbb{N}^*$  given by

$$m \leq' n \iff P(m) \leq P(n) \quad (m, n \in \mathbb{N}^*),$$

where  $P(k)$  is the number of primes that divide a positive integer  $k$  [see Examples 0.55 (3)]. Then the identity function of  $\mathbb{N}^*$ , considered as a map  $\langle \mathbb{N}^*, \leq \rangle \rightarrow \langle \mathbb{N}^*, \leq' \rangle$  of preordered sets, is order-preserving.  $\diamond$

## Partial orderings

subsec:partial-order

To the two conditions (I1)–(I2) required for a relation to be a preordering we now add one more.

precedes!strictly  
predecessor!strict  
succeeds!strictly  
successor!strict  
reverse relation!preordering@and p  
precedes!immediately  
predecessor!immediarte  
succeeds!immediately  
successor!immediate  
order-preserving map

def:partial-order

**0.59 Definition.** A relation  $\leq$  in a set  $X$  is said to **partially order**  $X$  and is called a **partial ordering of  $X$**  (or sometimes, more briefly, a **partial order of  $X$** ) if it has all three of the following properties:

(I1) Reflexivity: For each  $x \in X$ ,

$$x \leq x.$$

(I2) Transitivity: For all  $x, y \in X$ ,

$$x \leq y \quad \text{and} \quad y \leq z \implies x \leq z.$$

def:antisymmetric

(I3) Antisymmetry: For each  $x \in X$  and  $y \in X$ ,

$$x \leq y \quad \text{and} \quad y \leq x \implies x = y.$$

In other words,  $\leq$  partially orders  $X$  if and only if it preorders  $X$  and if, in addition, it is antisymmetric. A set together with a specific preordering on it is called a **partially ordered set** (or, sometimes, more briefly, a **poset**).

Each of the sets  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$ , and  $\mathbb{R}$  is a partially ordered set—and more!— with respect to its usual ordering.

m-sys-preordered by partial order

**0.60 Examples.** (1) The relation  $\leq$  in  $\mathbb{R}$  defined by

$$x \leq y \iff |x| \leq |y|$$

is a preordering of  $\mathbb{R}$ . However, this relation is *not* a partial ordering because antisymmetry fails for it: for example,  $-1 \leq 1$  and  $1 \leq -1$  yet  $-1 \neq 1$ .

Similarly, we obtain preorderings of  $\mathbb{Z}$  and of  $\mathbb{Q}$  that are not partial orderings.

ex:divisibility-po

(2) The divisibility relation [Examples 0.55 (2)] partially orders  $\mathbb{Z}$ .

ex:inclusion-po

(3) Let  $S$  be a given set. On the power set  $\mathcal{P}(S)$  of  $S$ , define the relation  $\leq$  by

$$A \leq B \iff A \subset B.$$

(Recall that  $A \subset B$  is the case even when  $A = B$ ; that is why some mathematicians write  $A \subseteq B$  for what we denote by  $A \subset B$ .) Then the relation  $\leq$  partially orders  $\mathcal{P}(S)$ .

ex:induced-po

(4) Given a partially ordered set  $X$  with partial ordering  $\leq$ , the preordering on  $Y$  induced by  $\leq$  [Examples 0.55 (4)] is again a partial ordering. Thus any subset  $Y$  of a partially ordered set  $X$  becomes a partially ordered set in its own right by restricting the partial ordering of  $X$  just to elements of  $Y$ .

For example, the usual ordering of  $\mathbb{R}$  induces on  $\mathbb{Q}$  the usual ordering of  $\mathbb{Q}$ ; the usual ordering of  $\mathbb{Q}$  induces on  $\mathbb{Z}$  the usual ordering of  $\mathbb{Z}$ ; and the usual ordering of  $\mathbb{Z}$  induces on  $\mathbb{N}$  the usual ordering of  $\mathbb{N}$ .

As a different example, the divisibility relation of Examples 0.55 (2) induces on the set  $\mathbb{N}^*$  of positive integers the partial ordering  $|$  given by

$$m | n \iff n = k m \text{ for some } k \in \mathbb{N}^* \quad (m, n \in \mathbb{N}^*).$$

(5) The reverse of a partial ordering in a set partially orders the same set.

In particular, for a set  $S$ , its power set  $\mathcal{P}(S)$  is partially ordered by *reverse* set inclusion  $\supset$ , that is:

$$A \leq B \iff A \supset B \quad (A, B \in \mathcal{P}(S))$$

partial ordering!product@and product!  
product!partial ordering@and partial  
partial ordering!coordinatewise  
relation!irreflexive  
irreflexive relation  
relation!transitive  
transitive relation  
relation!asymmetric  
asymmetric relation

ex:product-ordering

- (6) Let  $\langle X, \leq_X \rangle$  and  $\langle Y, \leq_Y \rangle$  be partially ordered sets. On the product set  $X \times Y$  define a relation  $\leq$  as follows. For all  $\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle \in X \times Y$ ,

$$\langle x_1, y_1 \rangle \leq_X \langle x_2, y_2 \rangle \iff x_1 \leq_X x_2 \quad \text{and} \quad y_1 \leq_Y y_2.$$

Then  $\leq$  is a partial ordering of  $X \times Y$ , called the **product ordering of  $X \times Y$  induced by  $\leq_X$  and  $\leq_Y$** . We say that this partial ordering is defined **coordinatewise**.

More generally, on the product  $\times_{i \in I} X_i$  of any family of partially ordered sets, there is a product partial ordering, defined coordinatewise.

A quite different ordering of a product—the *lexicographic ordering*, which is useful for examples in topology—is discussed in [Example 0.70](#).

ex:fn-extension-as-po

- (7) Let  $X$  and  $Y$  be sets. In the collection

$$\mathcal{F} = \{f: S \rightarrow Y : S \subset X\}$$

of all “partial maps” from  $X$  to  $Y$ , define a relation  $\leq$  by

$$f \leq g \iff g \text{ extends } f.$$

In other terms,  $f \leq g$  if and only if, as graphs of maps,  $f$  and  $g$  satisfy the set inclusion  $f \subset g$ . Then  $\leq$  is a partial ordering of  $\mathcal{F}$ .  $\diamond$

Because a partial ordering  $\leq$  of  $X$  is antisymmetric, its associated strict ordering  $<$ , defined by

$$x < y \iff x \leq y \text{ and } y \not\leq x \quad (x, y \in X),$$

actually satisfies the condition

$$x < y \iff x \leq y \text{ and } x \neq y \quad (x, y \in X).$$

This is the reason for the switch from the symbol ‘ $\leq$ ’ for a general preordering to the more familiar inequality symbol ‘ $\leq$ ’ for a partial ordering. Moreover, it is why  $x \leq y$  is legitimately read as “ $x$  is less than or equal to  $y$ ” while  $x < y$  is read as “ $x$  is less than  $y$ ” or, for emphasis, as “ $x$  is *strictly* less than  $y$ .”

The strict preordering  $<$  associated with a partial ordering  $\leq$  of a set  $X$  has the following properties:

(SI1) Irreflexivity: For each  $x \in X$ ,

$$x \not< x.$$

(SI2) Transitivity: For all  $x, y \in X$ ,

$$x < y \text{ and } y < z \implies x < z.$$

(SI3) Asymmetry: For all  $x, y \in X$ ,

$$x < y \implies y \not< x.$$

reverse relation! partial ordering and partial ordering  
 maximum  
 minimum  
 increasing map  
 increasing map  
 decreasing map  
 decreasing map  
 strictly increasing map  
 strictly decreasing map  
 def: greatest-least

The notations  $y \geq x$  and  $y > x$  are used for the reverses of a partial ordering  $\leq$  and of its associated strict partial ordering  $<$ , respectively. As usual, ‘ $y \geq x$ ’ is read as “ $y$  is greater than or equal to  $x$ ”, and ‘ $y > x$ ’ is read as “ $y$  is greater than  $x$ ” or, for emphasis, as “ $y$  is *strictly* greater than  $x$ .” The reverse  $\geq$  of a partial ordering  $\leq$  is again a partial ordering, and the reverse  $>$  of the associated strict partial ordering  $<$  is again a strict partial ordering.

**0.61 Definition.** Let  $A$  be a subset of a partially ordered set. A **greatest** (respectively, **least**) **element of  $A$**  is an  $x \in A$  such that

$$y \in A \implies y \leq x \quad (\text{respectively, } y \in A \implies x \leq y).$$

Synonyms for ‘greatest’ are ‘largest’ and ‘last’; synonyms for ‘least’ are ‘smallest’ and ‘first’.

For example, partially order the power set  $\mathcal{P}(S)$  of a set  $S$  by subset inclusion [Examples 0.60 (3)]; then the empty set  $\emptyset$  is a least element, and the entire set  $S$  a greatest element, of  $\mathcal{P}(S)$ .

A nonempty set  $A$  need not have a greatest or a least element: consider, for example, the subset  $A = \{x \in \mathbb{R} : 0 < x < 1\}$  of  $\mathbb{R}$ , where  $\mathbb{R}$  has its usual ordering. If, however, a set  $A$  does have a greatest (respectively, least) element, antisymmetry guarantees that it has exactly one such, which is then also called the **maximum** (respectively, **minimum**) of  $A$  and is denoted by

$$\max A \quad (\text{respectively, } \min A).$$

Notations such as

$$\max_{1 \leq i \leq n} x_i, \quad \min_{1 \leq i \leq n} x_i$$

used when the elements of  $A$  are indexed, should be self-explanatory.

Observe that, for a given partial ordering  $\leq$  of a set  $X$ , a maximum (respectively, minimum) of a subset  $A$  of  $X$  for  $\leq$  is a minimum (respectively, maximum) of  $A$  for the reverse partial ordering  $\geq$ .

The notions of a map between partially ordered sets being *order-preserving* or *order-reversing* are the ones introduced in Definition 0.57 for the more general case of preordered sets. In the case of partially ordered sets, we also use the synonyms **increasing** and **decreasing** for ‘order-preserving’ and ‘order-reversing’, respectively.

A map  $f: X \rightarrow Y$  between partially ordered sets is said to be **strictly increasing** when for all  $x_1, x_2 \in X$ ,

$$x_1 < x_2 \implies f(x_1) < f(x_2),$$

and **strictly decreasing** when for all  $x_1, x_2 \in X$ ,

$$x_1 < x_2 \implies f(x_1) > f(x_2)$$

**Usage note.** Some mathematicians use “nondecreasing” and “nonincreasing,” respectively, for what we call “increasing” and “decreasing”; then they reserve “increasing” and “decreasing” for what we call “strictly increasing” and “strictly decreasing,” respectively.

**0.62 Example.** Provide  $\mathbb{R}$  with its usual ordering. Then the maps  $x \mapsto x^3$  and  $x \mapsto -x$  of  $\mathbb{R} \rightarrow \mathbb{R}$  are strictly increasing and strictly decreasing, respectively.

The map  $f: \mathbb{R} \rightarrow \mathbb{R}$  given by

$$f(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } x \geq 0 \end{cases}$$

is increasing but not strictly increasing; the map  $g: \mathbb{R} \rightarrow \mathbb{R}$  given by  $g(x) = 1 - f(x)$  is decreasing but not strictly decreasing.  $\diamond$

A sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in a set  $X$  is a map from the ordered set  $\mathbb{N}$  to  $X$ . When the set  $X$  is also partially ordered, then the preceding terminology is applicable. Thus a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in a partially ordered set  $X$  is: **increasing** when  $x_m \leq x_n$  for all  $m, n \in \mathbb{N}$  with  $m \leq n$ ; and **decreasing** when  $x_m \geq x_n$  for all  $m, n \in \mathbb{N}$  with  $m \leq n$ ; **strictly increasing** when  $x_m < x_n$  for all  $m, n \in \mathbb{N}$  with  $m < n$ ; and **strictly decreasing** when  $x_m > x_n$  for all  $m, n \in \mathbb{N}$  with  $m < n$ . Similar language applies for sequences indexed by  $\mathbb{N}^* = \{1, 2, 3, \dots\}$  or other subsets of  $\mathbb{Z}$ .

It is easy to see that a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  is increasing, decreasing, strictly increasing, or strictly decreasing when

$$\begin{aligned} x_n &\leq x_{n+1} && \text{for all } n \in \mathbb{N}, \\ x_n &\geq x_{n+1} && \text{for all } n \in \mathbb{N}, \\ x_n &< x_{n+1} && \text{for all } n \in \mathbb{N}, \end{aligned}$$

or

$$x_n > x_{n+1} \quad \text{for all } n \in \mathbb{N},$$

respectively.

That a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  is increasing often is indicated by writing

$$x_0 \leq x_1 \leq x_2 \leq \dots \leq x_n \leq x_{n+1} \leq \dots,$$

and that it is decreasing by writing

$$x_0 \geq x_1 \geq x_2 \geq \dots \geq x_n \geq x_{n+1} \geq \dots,$$

and similarly for a strictly increasing or strictly decreasing sequence.

Our usage of ‘increasing’ and ‘decreasing’ is consistent with our use of the subset inclusion relation ‘ $\subset$ ’ to include the possibility of equality. Thus a sequence  $\langle A_n \rangle_{n \in \mathbb{N}}$  of subsets of a given set  $S$  **increasing** or **decreasing** when it is increasing or decreasing, respectively, as a sequence in the partially ordered set  $\langle \mathcal{P}(S), \subset \rangle$ .

def:order-iso

**0.63 Definition.** A map  $f: X \rightarrow Y$  is called an **order-isomorphism from  $\langle X, \leq_X \rangle$  to  $\langle Y, \leq_Y \rangle$**  if  $f$  is a bijection and both  $f$  and  $f^{-1}$  are order-preserving. The partially ordered set  $\langle X, \leq_X \rangle$  is said to be **order isomorphic to  $\langle Y, \leq_Y \rangle$**  when there is some order-isomorphism from the one to the other.

An order-isomorphism from one partially-ordered set to another is thus a one-to-one correspondence between the elements of the sets for which pairs of corresponding elements bear the same order relationship. Thus two partially-ordered sets that are order-isomorphic may be regarded as “essentially the same” with regard to their orderings.

sequence!increasing  
sequence!decreasing  
increasing sequence  
decreasing sequence  
sequence!increasing  
sequence!decreasing  
increasing sequence  
decreasing sequence  
sequence!strictly increasing  
sequence!strictly decreasing  
sequence of sets!increasing  
sequence of sets!decreasing  
sequence!increasing  
sequence!decreasing  
increasing sequence  
decreasing sequence  
increasing map  
decreasing map

**0.64 Examples.** (1) The identity map of any partially-ordered set is an order-isomorphism. (2) Give the set  $\mathbb{Z}$  of integers its usual ordering relation  $\leq$ . Then the map  $n \mapsto -n$  is an order-isomorphism from  $\langle \mathbb{Z}, \leq \rangle$  to  $\langle \mathbb{Z}, \geq \rangle$ . (3) Let  $X = \mathbb{N}$  and  $Y = \{n/(n+1) : n \in \mathbb{N}\}$ . Give both sets their partial orders induced by the usual ordering of  $\mathbb{Q}$ . Then the map  $n \mapsto n/(n+1)$  is an order-isomorphism.  $\diamond$

### Intervals

**0.65 Definition.** An **interval** in a partially ordered set  $X$  is a subset  $J$  of  $X$  with the property

$$x, y \in J \text{ and } z \in X \text{ and } x < z < y \implies z \in J,$$

in other words, any element of  $X$  that lies “between” two elements of  $J$  must itself belong to  $J$ .

An interval is said to be **nondegenerate** when it consists of more than one element.

According to this definition, both  $\emptyset$  and  $X$  are intervals in any partially ordered set  $X$ . For each  $a \in X$ , the **open rays**

$$]a, \rightarrow[ = \{x \in X : x > a\}, \quad ]\leftarrow, a[ = \{x \in X : x < a\}$$

and the **closed rays**

$$[a, \rightarrow[ = \{x \in X : x \geq a\}, \quad ]\leftarrow, a] = \{x \in X : x \leq a\}$$

are intervals in  $X$ . Also, for  $a, b \in X$  with  $a \leq b$ , the **open interval**

$$]a, b[ = \{x \in X : a < x < b\},$$

the **closed interval**

$$[a, b] = \{x \in X : a \leq x \leq b\},$$

and the **half-open intervals**

$$]a, b] = \{x \in X : a < x \leq b\}, \quad [a, b[ = \{x \in X : a \leq x < b\}$$

are intervals in  $X$ . A half-open interval of the form  $]a, b]$  is **left-open, right-closed**; a half-open interval of the form  $[a, b[$  is **left-closed, right-open**.

Later we shall explain the use of the adjectives ‘open’ and ‘closed’ above.

Notice that for an arbitrary partially ordered set we use  $\leftarrow$  and  $\rightarrow$  where, for the set  $\mathbb{R}$  with its usual ordering, we customarily use the symbols  $-\infty$  and  $+\infty$ . Thus for  $\mathbb{R}$  with its usual ordering:

$$\begin{aligned} ]a, +\infty[ &= ]a, \rightarrow[ = \{x \in \mathbb{R} : x > a\}, \\ ]-\infty, b[ &= ]\leftarrow, b[ = \{x \in \mathbb{R} : x < b\}, \\ [a, +\infty[ &= [a, \rightarrow[ = \{x \in \mathbb{R} : x \geq a\}, \\ ]-\infty, b] &= ]\leftarrow, b] = \{x \in \mathbb{R} : x \leq b\}. \end{aligned}$$

An interval in an arbitrary partially ordered set need not have one of the forms listed! In  $\mathbb{Q}$ , for example, the set

$$\{x \in \mathbb{Q} : x > 0, x^2 < 2\}$$

is an interval that is not of any of those forms. We shall shortly prove, however, that any interval in  $\mathbb{R}$  does have one of those forms (see [Theorem 0.87](#)).



An open interval  $]a, b[$  or a closed interval  $[a, b]$  in  $\mathbb{R}$  or in  $\mathbb{Q}$  is nondegenerate if and only if  $a < b$ . nondegenerate interval  
interval!nondegenerate

### Bounds

subsec:bounds

We now generalize the notion of least and greatest elements.

def:upper-lower-bounds

**0.66 Definition.** Let  $X$  be a partially ordered set and let  $A \subset X$ . An **upper bound of  $A$  (in  $X$ )** is an element  $b$  of  $X$  such that

$$x \in A \implies x \leq b.$$

A **lower bound of  $A$  (in  $X$ )** is an element  $b$  of  $X$  such that

$$x \in A \implies b \leq x.$$

The set  $A$  is said to be **bounded above in  $X$**  if it has some upper bound in  $X$ , and **bounded below in  $X$**  if it has some lower bound in  $X$ .

The empty subset  $\emptyset$  of a set  $X$  is bounded above in  $X$  if and only if  $X$  is nonempty, and then each element of  $X$  is an upper bound of  $\emptyset$  in  $X$ .

If  $A$  has a greatest element, then necessarily that element will be an upper bound of  $A$  in  $X$ ; likewise, if  $A$  has a least element, then necessarily that element will be a lower bound of  $A$  in  $X$ . However, an upper (respectively, a lower) bound of  $A$  need not belong to  $A$  and hence need not be the greatest (respectively, least) element of  $A$ .

A subset  $A$  of  $X$  may have neither an upper nor a lower bound: for example,  $A = \mathbb{Z}$  itself for the usual ordering of  $\mathbb{Z}$ . Moreover, if  $A$  does have an upper bound  $b$ , it need not be unique, because each element  $y$  of  $X$  with  $y \geq b$ , if there are any such, will also be an upper bound of  $A$ ; and similarly, if  $A$  does have a lower bound, it need not be unique.

def:sup-inf

**0.67 Definition.** Let  $X$  be a partially ordered set  $X$  and let  $A \subset X$ . An element of  $X$  is called a **supremum, or least upper bound, of  $A$  (in  $X$ )** if it is a least element of the set of all upper bounds of  $A$  in  $X$ .

An element of  $X$  is called an **infimum, or greatest lower bound, of  $A$**  if it is a greatest element of the set of all lower bounds of  $A$ .

In order that a subset  $A$  of  $X$  have a supremum (respectively, an infimum) it is evidently necessary that  $A$  have at least one upper (respectively, lower) bound in  $X$ . If, however,  $A$  does have a supremum (respectively, an infimum), it will be unique and is then denoted by

$$\sup A \quad (\text{respectively, } \inf A)$$

According to this definition, for  $b \in X$ ,

$$b = \sup A$$

precisely when

$$x \in A \implies x \leq b, \quad b' < b \implies b' < y \quad \text{for some } y \in A.$$

Similarly,

$$b = \inf A$$

endpoint  
interval!endpoints@and endpoints  
Dedekind, Richard

precisely when

$$x \in A \implies b \leq x, \quad b < b' \implies y < b' \quad \text{for some } y \in A \quad .$$

If a subset  $A$  of  $X$  has a supremum (respectively, infimum)  $x$  in  $X$  that belongs to  $A$ , then necessarily  $x$  will be the greatest (respectively, the least) element of  $A$ .

The infimum of an interval  $J$  in a partially ordered set  $X$ , if it exists, is called the **left endpoint of  $J$** , and the supremum of  $J$ , if it exists, is called the **right endpoint of  $J$** . Thus an element  $a$  of  $X$  is a left endpoint of each interval of the form  $]a, \rightarrow[$ ,  $[a, \rightarrow[$ ,  $]a, b[$ ,  $]a, b]$ ,  $[a, b]$ , or  $[a, b[$ ; and similarly for a right endpoint. In particular, a real number  $a$  is a left endpoint of each interval of the form  $]a, +\infty[$ ,  $[a, +\infty[$ ,  $]a, b[$ ,  $]a, b]$ ,  $[a, b]$ , or  $[a, b[$ ; and similarly for a right endpoint.

An interval is said to be **closed on the left** when it has a left endpoint that belongs to the interval, but **open on the left** otherwise. Similarly, an interval is said to be **closed on the right** when it has a right endpoint that belongs to the interval, and **open on the right** otherwise. An interval that has both a left endpoint and a right endpoint is said to be **closed** when it is both closed on the left and closed on the right, and **open** when it is both open on the left and open on the right. (This terminology explains the use of ‘open’ and ‘closed’ in the earlier examples of various kinds of intervals.)

**Caution!** An interval in a partially ordered set—even an interval in  $\mathbb{R}$ —need not have a left endpoint or a right endpoint; when it does, the endpoint need not belong to the interval. Thus *the supremum or infimum of a set need not belong to the set*.

When a set  $A$  does have a greatest (respectively, least) element  $b$ , then  $b = \sup A$  (respectively,  $b = \inf A$ ). Hence suprema and infima are generalizations of greatest and least elements, respectively.

The set  $\mathbb{N}$  of natural numbers, with its usual ordering, has the property that each nonempty subset of  $\mathbb{N}$  bounded above in  $\mathbb{N}$  has a supremum (namely, its greatest element—see [Exercise 103](#)). Let us name this property.

def:order-complete

**0.68 Definition.** A partially ordered set  $X$  and its partial ordering are said to be **order-complete** when each nonempty subset of  $X$  that is bounded above in  $X$  has a supremum in  $X$ .

“Dedekind complete” is commonly used instead of “order-complete.”

The set  $\mathbb{R}$  of real numbers, with its usual ordering, is order-complete—see the subsection “Order-completeness” ([page 79](#)). But the set  $\mathbb{Q}$  of rational numbers, with its usual ordering, is *not* order-complete—see [Exercise 132](#). Aside from  $\mathbb{R}$ , the main order-complete set that will be of interest to us is the set  $\Omega^+ = [0, \Omega]$  consisting of all of all “ordinals” up to and including the first uncountable ordinal—see the subsection “The first uncountable ordinal” ([page 110](#)).

## Total orderings

subsec:total-order

To the three conditions (I1)–(I3) required for a relation to be a partial ordering we add yet another.

def:total-ordering

**0.69 Definition.** A relation  $\leq$  in a set  $X$  is said to **totally order**  $X$  and is called a **total ordering of  $X$**  (or sometimes, more briefly, a **total order of  $X$** ) if it has all four of the following properties:

(I1) Reflexivity: *For each  $x \in X$ ,*

$$x \leq x.$$

(I2) Transitivity: *For all  $x, y, z \in X$ ,*

$$x \leq y \text{ and } y \leq z \implies x \leq z.$$

(I3) Antisymmetry: *For all  $x, y \in X$ ,*

$$x \leq y \text{ and } y \leq x \implies x = y.$$

def:comparable

(I4) Comparability: *If  $x, y \in X$ , then*

$$x \leq y \text{ or } y \leq x.$$

In other words,  $\leq$  totally orders  $X$  if it partially orders  $X$  and if, in addition, any two elements of  $X$  are comparable for  $\leq$ . A set together with a specific total ordering of it is called a **totally ordered set**.

pg-ref-total-ordering-end

The usual ordering of  $\mathbb{R}$  is a total ordering. The partial ordering induced on a subset of a totally ordered set is itself a total ordering. Thus any subset  $Y$  of a totally ordered set  $X$  becomes a totally ordered set in its own right by restriction of the total ordering of  $X$  just to elements of  $Y$ . In this way,  $\mathbb{N}$ ,  $\mathbb{Q}$ , and  $\mathbb{Z}$  are totally ordered sets.

ex:lexicographic-order

**0.70 Example.** Let  $X$  and  $Y$  be sets that are totally ordered by the relations  $\leq_X$  and  $\leq_Y$ , respectively. On the product set  $X \times Y$ , define a relation  $\leq$  as follows: For  $\langle x, y \rangle, \langle u, v \rangle \in X \times Y$ ,

$$\langle x, y \rangle \leq \langle u, v \rangle \iff (x <_X u) \text{ or } (x = u \text{ and } y \leq_Y v).$$

Then the relation  $\leq$  in  $X \times Y$  is a total ordering of  $X \times Y$  that, for obvious reasons, is called the **dictionary ordering**—in fancier terminology, the **lexicographic ordering**—of  $X \times Y$  (induced by  $\leq_X$  and  $\leq_Y$ ). When provided with its lexicographic ordering, we speak of the product of two totally ordered sets as being **lexicographically ordered**.

It can be helpful to visualize the lexicographic ordering as follows. Represent the product  $X \times Y$  as if it were a subset of the plane, with  $X$  a horizontal interval and  $Y$  a vertical interval. Think of each “stalk”  $\{x\} \times Y$  of the product as the vertical segment through the point  $x$  on the horizontal axis (as in [Example 0.18](#)). Then for points  $p$  and  $q$  in  $X \times Y$ ,

$$p \leq q$$

for the lexicographic ordering exactly when *either*:

- (i)  $p$  is on a stalk strictly to the left of the stalk containing  $q$ ; *or else*
- (ii) both  $p$  and  $q$  are on the same stalk and  $p$  is at or below  $q$  there.

(See [Figure 0.2](#).)

Take, in particular,  $X = \mathbb{N}$  and  $Y = [0, 1[$  with their usual orderings. Give the product  $\mathbb{N} \times [0, 1[$  its lexicographic ordering. Visualize the set  $\mathbb{N} \times [0, 1[$ , as shown in [Figure 0.3\(a\)](#), as a sequence of vertical intervals  $Y_0, Y_1, Y_2, \dots$  in the plane, each a copy of  $[0, 1[$ , with the  $n$ th copy standing upon the  $n$ th natural number on the  $x$ -axis. In short, attach a copy of  $[0, 1[$  to each element of  $\mathbb{N}$ . Now rotate each of the vertical intervals clockwise a quarter-turn

lexicographic ordering  
lexicographically ordered product

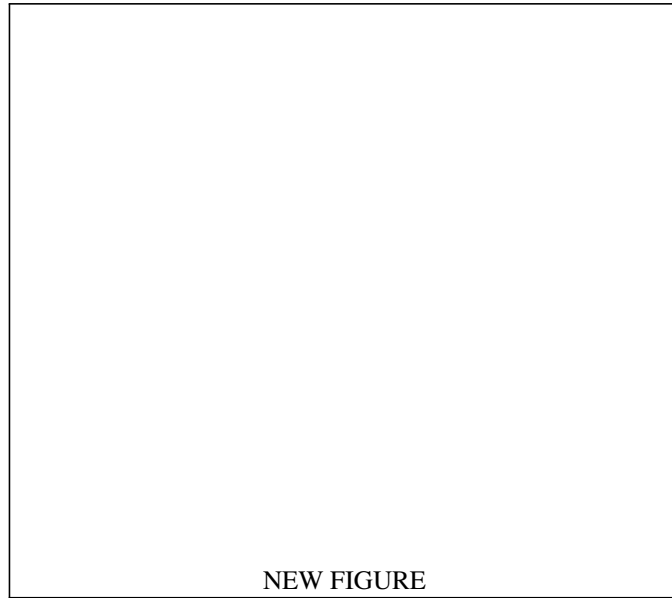
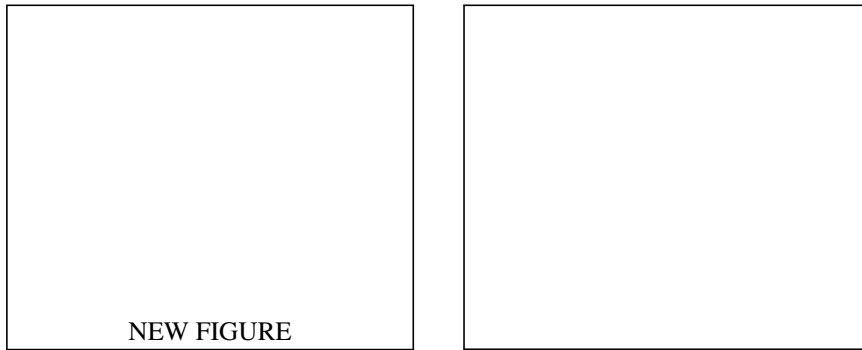


Figure 0.2: In the lexicographic ordering of the product set  $X \times Y$ , point  $p \leq q$  and  $p \leq r$ . fig:lexicographic-order



subfig:seq-vertical-intervals (a) The sequence of copies of the interval  $[0, 1]$  subfig:rotated-intervals (b) Rotate the intervals to fill out the ray  $[0, +\infty[$ .

Figure 0.3: Visualizing the product  $\mathbb{N} \times [0, 1[$  with its lexicographic ordering.

fig:lexicographic-order-N-times-inter

around its bottom end until it lies flat on the  $x$ -axis. The result is to fill out the entire ray  $[0, +\infty[$ , as shown in [Figure 0.3\(b\)](#). This seeming resemblance of  $\mathbb{N} \times [0, +\infty[$  to  $[0, +\infty[$  is genuine: the two totally ordered sets are order-isomorphic—see [Exercise 114](#).  $\diamond$

Recall that a partial ordering of a set  $X$  induces a partial ordering on each of its subsets.

def:chain

**0.71 Definition.** Let  $X$  be a partially ordered set. A **chain in  $X$**  is a subset of  $X$  that is *totally* ordered by its induced partial ordering. Equivalently, a chain in  $X$  is a subset each two of whose elements are comparable.

The notion of a chain is meaningful even when  $X$  is preordered but not partially ordered.

exs:chains **0.72 Examples.** (1) Partially order  $\mathcal{P}(\mathbb{N})$  by set inclusion. Then the collection

$$\{\{1, 3\}, \{1, 3, 7\}, \{1, 3, 6, 7, 10\}\}$$

is a chain in  $\mathcal{P}(\mathbb{N})$ .

- (2) Partially order the set  $\mathbb{Z}$  of all integers by its divisibility relation [Examples 0.55 (2)]. Then the subsets  $\{1, 2, 4, 12\}$  and  $\{3, 6, 12\}$  are chains in  $\mathbb{Z}$  whereas the subset  $\{2, 5, 6\}$  is not.  $\diamond$

relation!irreflexive  
irreflexive relation  
relation!transitive  
transitive relation  
relation!asymmetric  
asymmetric relation  
trichotomy

The strict preordering  $<$  associated with a total ordering  $\leq$  of a set  $X$  has the following properties:

(SI1) Irreflexivity: For each  $x \in X$ ,

$$x \not< x.$$

(SI2) Transitivity: For all  $x, y \in X$ ,

$$x < y \quad \text{and} \quad y < z \implies x < z.$$

(SI3) Asymmetry: For all  $x, y \in X$ ,

$$x < y \implies y \not< x.$$

(SI4) Trichotomy: For all  $x, y \in X$ ,

$$x < y, x = y, \text{ or } y < x.$$

Because of asymmetry and irreflexivity, the trichotomy property actually hold in the stronger form that for all  $x, y \in X$ , *exactly one* of the three conditions

$$x < y, \quad x = y, \quad y < x$$

holds.

The partial ordering  $\geq$  that is the reverse of a total ordering  $\leq$  is again a total ordering.

We distinguish those totally ordered sets in which between each two elements there is another element, or an element of a specified subset.

def:order-dense **0.73 Definition.** Let  $X$  be a totally ordered set. A subset  $D$  of  $X$  is said to be **order-dense in  $X$**  when each open interval in  $X$  contains some member of  $D$ ; in other words when, for all  $x, y \in X$  with  $x < y$ , there is some  $z \in D$  such that  $x < z < y$ . The totally ordered set  $X$  is said to be **order-dense** when  $X$  is order-dense in itself; in other words, when for all  $x, y \in X$  with  $x < y$ , there is some  $z \in X$  such that  $x < z < y$ .

pg:Q-order-dense-in-R For example, the set  $\mathbb{Q}$  of rational numbers is order-dense in the ordered set  $\mathbb{R}$  of real numbers—see Corollary 0.81 in the subsection “Order-completeness” (page 79), below. However, the set  $\mathbb{N}$  of natural numbers is *not* order-dense in either  $\mathbb{Q}$  or in  $\mathbb{R}$ . With their usual orderings,  $\mathbb{Q}$  and  $\mathbb{R}$  are order-dense (in themselves)—see Exercise 116. However  $\mathbb{N}$ , with its usual ordering, is *not* order-dense—see that same exercise.

## EXERCISES FOR SECTION 0.7

prob:ordering

- 96.** If  $\leq$  preorders a set  $X$ , prove that the corresponding strict ordering  $<$  is in fact irreflexive and transitive.

lexicographic ordering 97. Which if any, of the following relations  $\leq$  on  $\mathbb{R} \times \mathbb{R}$  are preorderings? which are order-complete set!lexicographic ordering and lexicographic ordering partial orderings? which are total orderings?

sequence!increasing  
sequence!decreasing  
increasing sequence  
decreasing sequence

- (a) The relation defined by  $\langle a, b \rangle \leq \langle c, d \rangle \iff a \leq c$  and  $b \leq d$ .
- (b) The relation defined by  $\langle a, b \rangle \leq \langle c, d \rangle \iff a \leq c$  or  $b \leq d$ .
- (c) The relation defined by  $\langle a, b \rangle \leq \langle c, d \rangle \iff a + b \leq c + d$ .
- (d) The relation defined by  $\langle a, b \rangle \leq \langle c, d \rangle \iff a < c$  or  $(a = c \text{ and } b \leq d)$ .
- (e) The relation defined by  $\langle a, b \rangle \leq \langle c, d \rangle \iff \sqrt{a^2 + b^2} \leq \sqrt{c^2 + d^2}$ .

98. Prove the assertion made on page 71 that an open interval  $]a, b[$  or a closed interval  $[a, b]$  in  $\mathbb{R}$  or in  $\mathbb{Q}$  is nondegenerate if and only if  $a < b$ .

99. Show that the preordering  $\leq$  defined in Examples 0.55 (3) does “direct”  $\mathbb{N}^*$  in the sense that for each  $m, n \in \mathbb{N}^*$ , there is some  $k \in \mathbb{N}^*$  such that  $m \leq k$  and  $n \leq k$ .

100. Construct an order isomorphism from the set  $\{1, 2, 3, 6\}$ , partially ordered by the divisibility relation, to the power set  $\mathcal{P}(\{0, 1\})$ , partially ordered by subset inclusion.

101. Let  $X$  be a totally ordered set and let  $x, y \in X$ . Prove:

- (a) Either  $] \leftarrow, x[ \subset ] \leftarrow, y[$  or  $] \leftarrow, y[ \subset ] \leftarrow, x[$ .
- (b) If  $] \leftarrow, x[ \subset ] \leftarrow, y[$ , then  $x \leq y$ .

102. Denote by  $\leq$  the lexicographic ordering (Example 0.70) of the plane  $\mathbb{R} \times \mathbb{R}$  induced by the usual ordering of the real line  $\mathbb{R}$ , and denote by  $<$  the corresponding strict ordering. Describe geometrically an interval in  $\mathbb{R} \times \mathbb{R}$  having the form  $[x, z[$  where  $x < z$ .

prob:N-order-complete 103. Show that  $\mathbb{N}$ , with its usual ordering, is order-complete.

er-complete-to-set-has-glb-property 104. Let  $X$  be a partially ordered set. Show that if  $X$  is order-complete, then each nonempty subset of  $X$  bounded below in  $X$  has a greatest lower bound. Is the converse true as well?

o:lex-prod-order-complete-min-max 105. Show that, when given its lexicographic ordering (Example 0.70), the product of two order-complete totally ordered sets, each having a greatest element, is order-complete.

prob-part:incr-seq-general-rel 106. (a) Prove: If  $\langle x_n \rangle_{n \in \mathbb{N}}$  is an increasing sequence in a partially ordered set  $X$ , then  $a_n \leq a_m$  for all  $n, m \in \mathbb{N}$  with  $n \leq m$ .

(b) State and prove the analog of (a) for decreasing sequences.

strictly-incr-seq-ints-not-bdd-above (c) Prove: A strictly increasing sequence in  $\mathbb{N}$  is never bounded above in  $\mathbb{N}$ .

b-part:incr-seq-sets-union-intersect (d) If  $\langle A_n \rangle_{n \in \mathbb{N}}$  is an increasing sequence of subsets of a set  $S$ , what are  $\bigcap_{n=0}^m A_n$  and  $\bigcup_{n=0}^m A_n$ ?

(e) Repeat (d) but for a decreasing sequence of sets.

107. (Continuation of Exercise 106.)

(a) Give an example of an increasing sequence in  $\mathbb{R}$  that is not bounded above and a decreasing sequence in  $\mathbb{R}$  that is not bounded below.

(b) Give an example of an increasing sequence in  $\mathbb{Q}$  that is bounded above but has no supremum in  $\mathbb{Q}$ .

(c) For  $n = 1, 2, 3, \dots$ , let

$$a_n = (1 + 1/n)^n, \quad b_n = (1 + 1/n)^{n+1}.$$

Prove: the sequence  $\langle a_n \rangle_{n \in \mathbb{N}^*}$  is increasing and bounded above in  $\mathbb{R}$ ; and the sequence  $\langle b_n \rangle_{n \in \mathbb{N}^*}$  is decreasing and bounded below in  $\mathbb{R}$ .

Cantor's theorem  
Cantor, Georg

**108.** Show that the partial ordering defined coordinatewise on the product of two partially ordered sets—see [Examples 0.60 \(6\)](#)—need not be a total ordering.

preorder-product-coordinatewise **109.** On the product  $X \times Y$  of two *preordered* sets  $X$  and  $Y$  with preorderings  $\leq_X$  and  $\leq_Y$ , respectively, let  $\leq$  be the relation defined “coordinatewise”, as in [Examples 0.60 \(6\)](#). Show that  $\leq$  preorders  $X \times Y$ .

**110.** Find all chains in  $\mathcal{P}(\{0, 1, 2\})$  when this power set is partially ordered by subset inclusion.

**111.** Is the open interval  $]0, 1[$ , with its usual ordering, order-isomorphic to the closed interval  $[0, 1]$ , with its usual ordering?

**112.** Let  $(X, \leq_X)$  and  $(Y, \leq_Y)$  be partially ordered sets and let  $f: X \rightarrow Y$  be a map such that

$$x_1 \leq_X x_2 \iff f(x_1) \leq_Y f(x_2) \quad (x_1, x_2 \in X).$$

Prove that  $f$  must be injective.

**113.** Prove that an order-preserving bijection between *totally* ordered sets must be an order-isomorphism.

totally-ordered-order-isomorphic-to-ray **114.** Give the product  $\mathbb{N} \times [0, 1[$  its lexicographic ordering induced by the usual orderings of  $\mathbb{N}$  and  $[0, 1[$ . Show that this totally ordered set is order-isomorphic to the ray  $[0, +\infty[$  when the latter is given its usual ordering.

**115.** Partially ordering coordinatewise the product of a family of partially ordering sets applies in particular to the  $I$ th power  $X^I$  of a partially ordered set  $X$ ; explicitly:

$$\langle x_i \rangle_{i \in I} \leq \langle y_i \rangle_{i \in I} \iff x_i \leq_X y_i \text{ for all } i \in I,$$

where  $\leq_X$  is the given partial ordering on  $X$ .

Take  $X = \{0, 1\}$  with its usual ordering. Prove that  $\{0, 1\}^I$ , with its coordinatewise partial ordering, is order isomorphic to the power set  $\mathcal{P}(X)$ , with set inclusion as the partial ordering.

$\mathbb{Q}$ -and- $\mathbb{R}$ -order-dense-in-themselves **116.** (a) Show that the set  $\mathbb{Q}$  of rational numbers and the set of real numbers, each with its usual ordering, is order-dense (in itself).

(b) Prove that the set  $\mathbb{N}$  of natural numbers is *not* order-dense (in itself).

Cantor-thm-countable-order-dense **117.** (a) Prove Cantor's theorem: any two countable, order-dense, totally ordered sets, each having neither an upper bound nor a lower bound, are order-isomorphic.

(b) Deduce that an order-dense, totally ordered set having neither an upper bound nor a lower bound is countable if and only if it is order-isomorphic to  $\mathbb{Q}$  (with its usual ordering).

prob:generate-preorder-from-surj  
field/ordered  
ordered field

**118.** Let  $\leq$  be a partial order on a set  $Y$  and let  $f: X \rightarrow Y$  be a surjection. Show that the relation  $\leq$  in  $X$  defined by

$$x \leq y \iff f(x) \leq f(y) \quad (x, y \in X)$$

preorders  $X$ . *Note:* In [Exercise 119](#) you will show that every preorder can be obtained in this way.

prob:preorder-is-generated-like-that

**119.** (*Continuation of Exercise 118.* This exercise requires familiarity with equivalence relations and the associated quotient maps, as described in [Section 0.9](#).) Let  $\leq$  be a preorder of a set  $X$ . Define a relation  $\sim$  in  $X$  by

$$x \sim y \iff x \leq y \text{ and } y \leq x \quad (x, y \in X).$$

(a) Check that  $\sim$  is an equivalence relation on  $X$ .

Let  $Y = X/\sim$  be the quotient set and let  $q: X \rightarrow Y$  be the corresponding quotient map, so that  $q$  is surjective.

(b) Show that the formula

$$[x] \leq [y] \iff x \leq y$$

gives a well-defined relation  $\leq$  in  $Y$  and that this relation partially orders  $Y$ .

(c) Show that, for  $x, y \in X$ , we have  $x \leq y$  in  $X$  if and only if  $q(x) \leq q(y)$  in  $Y$ .

*Note:* Thus every preorder of a set may be obtained by the construction described in [Exercise 118](#).

## 0.8 Order-Completeness of the Real Numbers

sec:ordercomplete

The algebraic and ordering properties of the set  $\mathbb{R}$  of real numbers that we have stipulated so far are as follows:

prop-R:field (R1) *The operations of addition and multiplication make  $\mathbb{R}$  a field.* (See [pages 22–23](#)).

prop-R:ordered (R2) *The relation  $\leq$  makes  $\mathbb{R}$  a totally ordered set.* (See [pages 73–73](#)).

We stipulate also the following compatibility conditions between the field operations on  $\mathbb{R}$  and the total ordering relation of  $\mathbb{R}$ :

prop-R:ordered-field (R3) *For all  $x, y, z \in \mathbb{R}$ ,*

$$\begin{aligned} x \leq y &\implies x + z \leq y + z, \\ x \leq y &\implies x \cdot z \leq y \cdot z \text{ provided that } z \geq 0. \end{aligned}$$

Together, properties (R1)–(R3) say that  $\mathbb{R}$  is an **ordered field**

Operations of addition and multiplication along with an ordering relation make the rational numbers  $\mathbb{Q}$ , like  $\mathbb{R}$ , an ordered field. What distinguishes  $\mathbb{R}$  from  $\mathbb{Q}$  (and from other ordered fields) is one more property of  $\mathbb{R}$ —*order-completeness*—introduced in the following subsection. That  $\mathbb{R}$ , unlike  $\mathbb{Q}$ , is order-complete has, as we shall see, significant consequences.



## Order-completeness

subsec:order-completeness

A subset  $A$  of  $\mathbb{R}$  may have neither an upper nor a lower bound: for example,  $A = \mathbb{R}$  itself. Moreover, if  $A$  does have an upper bound  $b$ , it will never be unique, because every real number  $y$  with  $y \geq b$  will also be an upper bound of  $A$ ; and similarly, if  $A$  does have a lower bound, it will never be unique.

We shall accept as an “article of faith”—a fundamental property of the set of real numbers—that  $\mathbb{R}$  is order-complete (Definition 0.68).

infimum  
supremum  
base expansion  
base for expansion  
base- $b$  digits@base- $b$  digits

ax:R-order-complete

**0.74 Axiom (order-completeness).** Each nonempty subset of  $\mathbb{R}$  that has an upper bound has a supremum.

This axiom, which distinguishes  $\mathbb{R}$  from other ordered fields such as  $\mathbb{Q}$ , asserts that there are no “holes” in the real line. It guarantees, for example, the existence of a square-root of each real number  $a \geq 0$ : in fact, if  $b = \sup\{x \in \mathbb{R} : x \geq 0, x^2 \leq a\}$ , then  $b^2 = a$ . We proceed to deduce several of the axiom’s more general, important, consequences.

prop:infs-in-R

**0.75 Proposition (existence of infimum of set of reals bounded below).** Each nonempty subset of  $\mathbb{R}$  that has a lower bound has an infimum.

**Proof.** This follows from Exercise 104. We give a different proof here. The idea is to reflect points in, and subsets of, the real line around the origin, thereby converting upper bounds to lower bounds, and vice versa.

Let  $\emptyset \neq B \subset \mathbb{R}$  and let  $B$  have a lower bound  $m \in \mathbb{R}$ . Reflect both the set  $B$  and the point  $m$  around the origin in the real line, that is, define

$$A = \{-x : x \in B\}$$

and form  $-m$ . Then  $A \neq \emptyset$  and  $-m$  is an upper bound of  $A$ . (Supply the details!) By Axiom 0.74,  $A$  has a supremum in  $\mathbb{R}$ . Let

$$b = \sup A.$$

We claim that

$$-b = \inf B.$$

Clearly  $-b$  is a lower bound of  $B$ . Suppose that  $b'$  is a lower bound of  $B$ . Then  $-b'$  is an upper bound of  $A$ , so that  $b \leq -b'$  because  $b$  is the least upper bound of  $A$ , and hence  $b' \leq -b$ .  $\square$

## Base expansion

subsec:base-expand

Every integer  $m$  has a decimal, that is, base-10, representation  $m = m_n m_{n-1} \dots m_2 m_1 m_0$  in the sense that  $m = \sum_{j=0}^n m_j / 10^j$  with  $m_j \in \{0, 1, 2, \dots, 9\}$  for each  $j$ . [See Exercise 23 (a).]

The first consequence of order-completeness we obtain is the counterpart of such representations for real numbers between 0 and 1—and for arbitrary base  $b > 1$ , not just 10. In this context, we refer to the members of  $\{0, 1, 2, \dots, b-1\}$  as **base- $b$  digits**.

Let  $\langle d_n \rangle_{n \in \mathbb{N}^*}$  be an arbitrary sequences of base- $b$  digits. For each  $n = 1, 2, 3, \dots$ , the value of the finite sum  $\sum_{j=1}^n d_j / b^j$  is nonnegative and no greater than the value of the geometric sum  $\sum_{j=1}^n (b-1)/b^j$  and thus satisfies  $0 \leq \sum_{j=1}^n d_j / b^j \leq 1$ . Moreover, the

sequence  $\langle \sum_{j=1}^n d_j/b^j \rangle_{n \in \mathbb{N}^*}$  of nonnegative real numbers is *increasing*. From that and the order-completeness of  $\mathbb{R}$ , we obtain the following result.

**0.76 Lemma.** Let  $\langle d_n \rangle_{n \in \mathbb{N}^*}$  be an arbitrary sequence of base- $b$  digits for a base  $b > 1$ . Then the set  $\{ \sum_{j=1}^n d_j/b^j : n \in \mathbb{N}^* \}$  has a supremum. Moreover, the equation

$$(1) \quad x = \sup \left\{ \sum_{j=1}^n \frac{d_j}{b^j} : n \in \mathbb{N}^* \right\}$$

is equivalent to the condition:

$$(2) \quad \text{for each } \varepsilon > 0, \text{ there exists } k \in \mathbb{N}^* \text{ such that } 0 \leq x - \sum_{j=1}^n \frac{d_j}{b^j} < \varepsilon \text{ for all } n \geq k.$$

We express (1), or equivalently (2), by saying that  $(0.d_1 d_2 d_3 \dots d_n \dots)_b$  is a **base- $b$  expansion** of  $x$ , by writing

$$(*) \quad x = (0.d_1 d_2 d_3 \dots d_n \dots)_b.$$

and by calling  $d_1, d_2, d_3, \dots$  the **digits** of this base- $b$  expansion.

Later, in the subsection “Convergent sequences in metric spaces” (page 185), we shall formally introduce limits of sequences and series in the context of metric spaces. Anticipating that, we also express equation (\*) by writing

$$x = \lim_{n \rightarrow \infty} \sum_{j=1}^n \frac{d_j}{b^j} = \sum_{j=1}^{\infty} \frac{d_j}{b^j}.$$

When there is some nonnegative integer  $k$  with  $d_n = 0$  for all  $n > k$ , the expansion (\*) is said to be **terminating**; otherwise, it is said to be **nonterminating**. When an expansion (\*) is nonterminating, then the first inequality in condition (2), above, will be strict.

In the familiar case  $b = 10$ , that is, **decimal expansion**, the subscript in equation (\*) is ordinarily dropped and we write simply

$$x = 0.d_1 d_2 d_3 \dots d_n \dots$$

In the sequel, the most important cases of base expansions will be not for  $b = 10$  but rather for  $b = 2$  and  $b = 3$ , which we refer to as **binary** and **ternary** expansions, respectively.

For arbitrary  $b$ , we referred to (\*) as *a* base- $b$  expansion and *not* as *the* base- $b$  expansion. Indeed, certain real numbers  $x$  with  $0 < x \leq 1$  will have *two* base- $b$  expansions, one terminating and the other non-terminating. For example, we have the two decimal expansions

$$\begin{aligned} 1/2 &= (0.500 \dots 0 \dots)_{10}, \\ 1/2 &= (0.499 \dots 9 \dots)_{10} \end{aligned}$$

and the two binary expansions

$$\begin{aligned} 1/2 &= (0.100 \dots 0 \dots)_2, \\ 1/2 &= (0.011 \dots 1 \dots)_2 \end{aligned}$$

We saw above how, from a sequence of base- $b$  digits, we may obtain an  $x$  in  $[0, 1]$  whose base- $b$  expansion is given by that sequence. Now we consider the reverse process, namely: how, from an  $x$  in  $[0, 1]$ , to obtain a base- $b$  expansion of it. Already,  $x = 0$  has the *unique* base- $b$  expansion  $0 = (0.000 \dots 0 \dots)_b$  no matter what the base  $b$  is.

base expansion

prop:base-b-expand-x

**0.77 Proposition (base expansion of a real number).** Let  $b$  be an integer with  $b > 1$ . Let  $x$  be a real number with  $0 < x \leq 1$ . Then:

(1) There is a unique sequence  $\langle d_n \rangle_{n \in \mathbb{N}^*}$  of base- $b$  digits for which

{eq:base-digits-construct-inequal} (3) 
$$0 < x - \sum_{j=1}^n \frac{d_j}{b^j} \leq \frac{1}{b^n} \text{ for each } n = 1, 2, 3, \dots$$

(2) The real number  $x$  has the unique nonterminating base- $b$  expansion  $(0.d_1 d_2 d_3 \dots d_n \dots)_b$ .

**Proof.** Any base- $b$  expansion  $x = (0.d_1 d_2 d_3 \dots d_n \dots)_b$  satisfying (3) is necessarily nonterminating. In fact, if it is terminating, then  $x = \sum_{j=1}^k d_j / b^j$  for some  $k$ , contrary to (3).

Existence. Construct recursively the sequence  $\langle d_n \rangle_{n \in \mathbb{N}}$  of base- $b$  digits as follows. First, take  $d_1$  to be the unique base- $b$  digit  $k$  such that

$$\frac{k}{b^1} < x \leq \frac{k+1}{b^1},$$

so that

$$\frac{d_1}{b^1} < x \leq \frac{d_1+1}{b^1}$$

and hence

$$0 < x - \frac{d_1}{b^1} \leq \frac{1}{b^1}.$$

Once  $d_1, d_2, \dots, d_{n-1}$  have been constructed, take  $d_n$  to be the unique base- $b$  digit  $k$  such that

$$\frac{k}{b^n} < x - \sum_{j=1}^{n-1} \frac{d_j}{b^j} \leq \frac{k+1}{b^n},$$

so that

$$\frac{d_n}{b^n} < x - \sum_{j=1}^{n-1} \frac{d_j}{b^j} \leq \frac{d_n+1}{b^n}$$

and hence

$$0 < x - \sum_{j=1}^n \frac{d_j}{b^j} \leq \frac{1}{b^n}.$$

Thus the sequence  $\langle d_n \rangle_{n \in \mathbb{N}}$  satisfies condition (3) on page 81. In view of the Archimedean Ordering Property (Theorem 0.78), it also satisfies condition (2) there. From Lemma 0.76 we conclude that  $(0.d_1 d_2 d_3 \dots d_n \dots)_b$  is a base- $b$  expansion of  $x$ .

Uniqueness. The proof is left as an exercise.  $\square$

## Archimedean Ordering and Nested Interval Properties

subsec:AOP-and-NIP

We continue with more consequences of order-completeness.

thm:R-archimedean

**0.78 Theorem (Archimedean Ordering Property).** Let  $x$  and  $\varepsilon$  be real numbers with  $\varepsilon > 0$ . Then there exists a positive integer  $n$  such that

$$n\varepsilon > x.$$

**Proof.** Just suppose, to the contrary, that

$$n\varepsilon \leq x \quad (n = 1, 2, 3, \dots).$$

Then

$$n \leq \frac{x}{\varepsilon} \quad (n = 1, 2, 3, \dots),$$

so that  $x/\varepsilon$  is an upper bound in  $\mathbb{R}$  of the set

$$\mathbb{N}^* = \{1, 2, 3, \dots\}$$

of all positive integers. By order-completeness we may form the real number

$$b = \sup \mathbb{N}^*$$

Then

$$n \leq b \quad (n \in \mathbb{N}^*)$$

so also

$$n + 1 \leq b \quad (n \in \mathbb{N}^*).$$

Hence

$$n \leq b - 1 \quad (n \in \mathbb{N}^*).$$

Thus  $b - 1$  is an upper bound of  $\mathbb{N}^*$  in  $\mathbb{R}$ . This is impossible because  $b$  is the *least* upper bound of  $\mathbb{N}^*$ .  $\square$

In geometric language, the Archimedean Ordering Property says that by starting at the origin and laying end-to-end sufficiently many line segments all of fixed length  $\varepsilon$ , no matter how small  $\varepsilon$  may be, we can get beyond any given point on the real line (see [Figure 0.4](#)).



Figure 0.4: Geometric meaning of the Archimedean ordering property

fig:AOP-geometry

Taking  $\varepsilon = 1$  in the Archimedean Ordering Property ([Theorem 0.78](#)) yields the following special case.

cor:N-unbdd-above-in-R

**0.79 Corollary.** The set of positive integers is not bounded above in  $\mathbb{R}$ : for each real number  $x$ , there is some positive integer  $n$  with  $n > x$ .

We shall need the following result later when later we discuss limits of sequences. In the language of limits, the result says that

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0.$$

cor:limit-1-over-n-is-0

**0.80 Corollary.** Let  $y \in \mathbb{R}$  with

$$0 \leq y \leq \frac{1}{n} \quad (n = 1, 2, 3, \dots).$$

Then  $y = 0$ .dyadic rational  
order-dense set

**Proof.** By hypothesis,  $ny \leq 1$  for each  $n = 1, 2, 3, \dots$ . If  $y > 0$ , then by taking  $\varepsilon = y$  and  $x = 1$  in [Theorem 0.78](#) we see, however, that  $ny > 1$  for some positive integer  $n$ .  $\square$

Because  $2^n > n$ , it follows from [Corollary 0.80](#) that also

$$\text{if } 0 \leq y \leq \frac{1}{2^n} \text{ for each } n = 1, 2, 3, \dots, \text{ then } y = 0.$$

Another consequence of order-completeness of the real numbers is that the set  $\mathbb{Q}$  of rational numbers is order-dense in  $\mathbb{R}$ , which was noted earlier on [page 75](#), in connection with [Definition 0.73](#).

cor:rational-order-dense-in-reals

**0.81 Corollary (order-density of the rationals in the reals).** Each open interval in  $\mathbb{R}$  contains a rational number.

**Proof.** An arbitrary open interval  $]a, b[$  in  $\mathbb{R}$  can be put into the form  $]x - \varepsilon, x + \varepsilon[$  by taking  $x = (a + b)/2$  and  $\varepsilon = (b - a)/2$ . Hence what we must show is equivalent to:

$$(\ast) \quad x \in \mathbb{R} \text{ and } \varepsilon > 0 \implies |x - q| < \varepsilon \quad \text{for some } q \in \mathbb{Q}.$$

It is enough to consider the case  $x \geq 0$ ; indeed, if  $x < 0$  and if  $q \in \mathbb{Q}$  with  $|(-x) - q| < \varepsilon$ , then  $-q \in \mathbb{Q}$  and  $|x - (-q)| < \varepsilon$ .

Accordingly, let  $x \geq 0$  and  $\varepsilon > 0$ . By the Archimedean Ordering Property ([Theorem 0.78](#)), there exists a positive integer  $n$  such that

$$\frac{1}{n} < \varepsilon.$$

Also by the Archimedean Ordering Property ([Theorem 0.78](#)), there exists a positive integer  $k$  such that  $k > nx$ ; let  $m$  be the least such  $k$ . Then  $m - 1$  is not such a  $k$ , that is,

$$m - 1 \leq nx,$$

and so

$$\frac{m}{n} - \frac{1}{n} \leq x < \frac{m}{n}.$$

Thus the rational number  $q = m/n$  satisfies  $|x - q| < \varepsilon$ .  $\square$

Property (\*) above says that any real number can be approximated as closely as we wish by rational numbers.

Actually, there are countable subsets of real numbers that are in a sense “sparser” than the set of all rationals yet are still order-dense.

ex:dyadic-rationals

**0.82 Example.** A **dyadic rational** is a number of the form  $m/2^n$ , where  $m$  is an integer and  $n$  is a nonnegative integer. For example,  $0 = 0/2$ ,  $1 = 1/2^0$ ,  $5/8 = 5/2^3$ ,  $49/32 = 49/2^5$ , and  $-9/16 = (-9)/2^4$  are dyadic rationals.

The set of all dyadic rationals is order-dense in  $\mathbb{R}$ . In fact, let  $a, b \in \mathbb{R}$  be arbitrary with  $a < b$ . By [Corollary 0.79](#), there is some natural number  $n$  with  $n > 1/(b - a)$ . Then a fortiori  $2^n > 1/(b - a)$ , and so  $2^n b - 2^n a > 1$ . Since the collection  $\{[k, k + 1) : k \in \mathbb{N}\}$

increasing sequence of sets  
 decreasing sequence of sets  
 sequence!increasing  
 sequence!decreasing

is a partition of  $[0, \infty[$ , there is some integer  $m$  with  $2^n a < m < 2^n b$ . This means that  $a < m/2^n < b$ .

From what was just proved, it follows that the set

$$\mathcal{A} = \left\{ \frac{m}{2^n} : n \in \mathbb{N}, m = 0, 1, 2, \dots, 2^n \right\}$$

is dense in the unit interval  $[0, 1]$ . (Notice that  $\mathcal{A}$  is just the set of those numbers in  $[0, 1]$  having a terminating binary expansion.) And the set

$$\mathcal{A}^* = \left\{ \frac{m}{2^n} : n \in \mathbb{N}^*, m = 1, 2, \dots, 2^n - 1 \right\}$$

of dyadic rationals in  $]0, 1[$  is order-dense in the open unit interval  $]0, 1[$  as well as in the closed unit interval  $[0, 1]$ .  $\diamond$

The next theorem refers to a “decreasing sequence” of closed intervals. In general, a sequence  $\langle X_n \rangle_{n \in \mathbb{N}}$  of sets is said to be **decreasing** when

$$X_0 \supset X_1 \supset X_2 \supset \dots \supset X_n \supset X_{n+1} \supset \dots$$

and **increasing** when

$$X_0 \subset X_1 \subset X_2 \subset \dots \subset X_n \subset X_{n+1} \subset \dots$$

(See [Exercise 106](#).) Note that a sequence  $\langle X_n \rangle_{n \in \mathbb{N}}$  of sets will be both increasing and decreasing exactly in case  $X_0 = X_1 = X_2 = \dots = X_n = \dots$ .

The next theorem also uses also the condition  $\lim_{n \rightarrow \infty} (b_n - a_n) = 0$ . This has the usual meaning of the limit of a sequence from calculus:

For each  $\varepsilon > 0$ , there is some  $k$  such that  $|b_n - a_n| < \varepsilon$  for all  $n \geq k$ .

[In the subsection “Convergent sequences in metric spaces” ([page 185](#)) of [Chapter 1](#) (Metric Spaces) we shall study sequential convergence in a more general setting.]

thm:NIP **0.83 Theorem (Nested Interval Property).** Let  $\langle [a_n, b_n] : n \in \mathbb{N} \rangle$  be a decreasing sequence of closed intervals in  $\mathbb{R}$  with

$$\lim_{n \rightarrow \infty} (b_n - a_n) = 0.$$

Then there exists a unique point  $x$  such that

$$x \in \bigcap_{n=0}^{\infty} [a_n, b_n].$$

**Proof.** Since  $[a_n, b_n] \subset [a_0, b_0]$  for each  $n$ , then

$$a_n \leq b_n \quad (n \in \mathbb{N}).$$

Thus  $b_0$  is an upper bound of the nonempty set  $\{a_n : n \in \mathbb{N}\}$ . By order-completeness ([Axiom 0.74](#)), we may form

$$x = \sup\{a_n : n \in \mathbb{N}\}.$$

We show that  $x \in [a_n, b_n]$  for every  $n$ . Since  $x$  is an upper bound of  $\{a_n : n \in \mathbb{N}\}$ , already  $a_n \leq x$  for each  $n$ . Just suppose that  $b_m < x$  for some  $m$ . Now

$$n \leq m \implies [a_n, b_n] \supset [a_m, b_m] \implies a_n \leq a_m \leq b_m,$$

and

$$n > m \implies [a_n, b_n] \subset [a_m, b_m] \implies a_n \leq b_n \leq b_m.$$

Hence  $b_m$  is also an upper bound of  $\{a_n : n \in \mathbb{N}\}$ . This is impossible because  $b_m < x$  and  $x$  is the *least* upper bound of this set.

Thus there is some number  $x \in \bigcap_{n=0}^{\infty} [a_n, b_n]$ . To complete the proof, we show that there is only one such  $x$ . Just suppose also  $y \in [a_n, b_n]$  for every  $n$  but  $y \neq x$ . Since  $\lim_{n \rightarrow \infty} (b_n - a_n) = 0$ , there exists some  $n$  with

$$b_n - a_n < |x - y|.$$

However,

$$|x - y| \leq b_n - a_n$$

because  $x \in [a_n, b_n]$  and  $y \in [a_n, b_n]$ .  $\square$

We have taken order-completeness of  $\mathbb{R}$  as an axiom and from it have deduced the intuitively appealing Archimedean Ordering Property and Nested Interval Property. It is possible, conversely, to deduce order-completeness from these two properties together.

### Properties of intervals in the real line

Let us return now to the subject of “sizes” of sets, as introduced in [Section 0.6](#). Recall that a set is said to be *uncountable* when it is neither finite nor denumerable.

**0.84 Theorem (uncountability of closed interval in  $\mathbb{R}$ ).** Each closed interval  $[a, b]$  in  $\mathbb{R}$  with  $a < b$  is uncountable.

**Proof.** Just suppose, to the contrary, that  $[a, b]$  is countable. Then

$$[a, b] = \{x_n : n \in \mathbb{N}\}$$

for some sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$ .

Construct a sequence of real numbers and a sequence of closed intervals recursively as follows. First, choose a closed interval  $[a_0, b_0] \subset [a, b]$  with

$$x_0 \notin [a_0, b_0], \quad b_0 - a_0 < \frac{1}{2} (b - a).$$

Next, choose a closed interval  $[a_1, b_1] \subset [a_0, b_0]$  with

$$x_1 \notin [a_1, b_1], \quad b_1 - a_1 < \frac{1}{2} (b_0 - a_0).$$

In general, once  $x_{n-1}$  and  $[a_{n-1}, b_{n-1}]$  have been obtained, choose a closed interval  $[a_n, b_n] \subset [a_{n-1}, b_{n-1}]$  with

$$x_n \notin [a_n, b_n], \quad b_n - a_n < \frac{1}{2} (b_{n-1} - a_{n-1}).$$

We obtain a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  of real numbers and a decreasing sequence  $\langle [a_n, b_n] \rangle_{n \in \mathbb{N}}$  of closed intervals with

$$x_n \notin [a_n, b_n], \quad b_n - a_n < \frac{1}{2^{n+1}} (b - a)$$

for every  $n \in \mathbb{N}$ .

Clearly  $\lim_{n \rightarrow \infty} (b_n - a_n) = 0$ . By the Nested Interval Property ([Theorem 0.83](#)), there is some

$$x \in \bigcap_{n=0}^{\infty} [a_n, b_n].$$

In particular,  $x \in [a_0, b_0] \subset [a, b]$ , and so  $x = x_n$  for some  $n$ . Then  $x_n = x \in [a_n, b_n]$ , contrary to the choice of  $x_n$ .  $\square$

cor:reals-uncountable

**0.85 Corollary (uncountability of the set of real numbers).** *The set  $\mathbb{R}$  of all real numbers is uncountable.*

We already saw, as a consequence of order-completeness, that every nonempty open interval in  $\mathbb{R}$  contains rational numbers. Since such an interval, as we shall see, must be uncountable, it will contain irrational numbers as well.

cor:open-interval-contains-irrationals

**0.86 Corollary (order-density of irrational numbers in the reals).** *Each nonempty open interval in  $\mathbb{R}$  contains some irrational number.*

**Proof.** Consider an open interval  $]a, b[$ , where  $a < b$ . Choose  $c < d$  with  $[c, d] \subset ]a, b[$ . The closed interval  $[c, d]$  is uncountable by [Theorem 0.84](#), and so the open interval  $]a, b[$  is uncountable as well. Now

$$]a, b[ = (\mathbb{Q} \cap ]a, b[) \cup ((\mathbb{R} \setminus \mathbb{Q}) \cap ]a, b[)$$

with the subset  $\mathbb{Q} \cap ]a, b[$  of  $]a, b[$  countable, and so  $(\mathbb{R} \setminus \mathbb{Q}) \cap ]a, b[$  is uncountable. In particular,  $(\mathbb{R} \setminus \mathbb{Q}) \cap ]a, b[$  is nonempty.  $\square$

To conclude our treatment of order-completeness we establish the form of all subsets of  $\mathbb{R}$  that are intervals in the sense of [Definition 0.65](#).

thm:intervals-in-R

**0.87 Theorem (characterization of intervals in  $\mathbb{R}$ ).** *A subset of  $\mathbb{R}$  is an interval in  $\mathbb{R}$  if and only if it is the empty set, an open ray, a closed ray, an open interval, a closed interval, a half-open interval, or  $\mathbb{R}$  itself.*

**Proof.** We have already noted ([page 70](#)) that the empty set, open rays, and so on, are intervals in any totally ordered set. Conversely, let  $J$  be an interval in  $\mathbb{R}$ . Suppose  $J \neq \emptyset$ . We distinguish several cases.

Case (i):  $J$  has both a lower bound and an upper bound. In this case, order completeness allows us to form the numbers

$$a = \inf J, \quad b = \sup J.$$

Clearly

$$x < a \text{ or } x > b \implies x \notin J,$$

so that

$$J \subset [a, b].$$

Also,

$$]a, b[ \subset J,$$

for if  $a < x < b$ , then  $x$  is neither a lower nor an upper bound of  $J$ , which implies there exist  $c, d \in J$  with  $c < x < d$ , and hence  $x \in J$  because  $J$  is an interval. In this case, we conclude



that:  $J = [a, b]$  if  $a \in J$  and  $b \in J$ ;  $J = [a, b[$  if  $a \in J$  but  $b \notin J$ ;  $J = ]a, b]$  if  $b \in J$  but  $a \notin J$ ; and  $J = ]a, b[$  if  $a \notin J$  and  $b \notin J$ .

all-intervals-lower-not-upper-bound

Case (ii):  $J$  has a lower bound but not an upper bound. In this case, let

$$a = \inf J.$$

Clearly

$$x < a \implies x \notin J$$

so that

$$J \subset ]a, +\infty[.$$

Also,

$$]a, +\infty[ \subset J,$$

for if  $x > a$ , then  $x$  is not a lower bound of  $J$ , which implies there exists  $c \in J$  with  $c < x$ , there exists  $d \in J$  with  $x < d$  because  $x$  is not an upper bound of  $J$ , and finally  $x \in J$  because  $J$  is an interval. In this case,  $J = [a, +\infty[$  if  $a \in J$ , and  $J = ]a, +\infty[$  if  $a \notin J$ .

Case (iii):  $J$  has an upper bound but not a lower bound. In this case, let

$$b = \sup J.$$

As in Case (ii),  $J = ]-\infty, b]$  if  $b \in J$ , and  $J = ]-\infty, b[$  if  $b \notin J$ .

Case (iv):  $J$  has neither an upper nor a lower bound. Let  $x$  be an arbitrary real number. Then  $x$  is neither a lower nor an upper bound of  $J$ , which implies there exist  $c, d \in J$  with  $c < x < d$ , and so  $x \in J$  because  $J$  is an interval. Hence  $J = \mathbb{R}$  in this case.  $\square$

### Natural numbers, integers, and rationals within the reals

subsec:N-Z-Q-within-R

Earlier in this chapter we dealt with the sets  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$ , and  $\mathbb{R}$  of natural numbers, integers, rational numbers, and real numbers, respectively, somewhat informally, trusting to your prior familiarity with them and their properties. Then at the beginning of this section we formalized the status of  $\mathbb{R}$  by making the fundamental assumption that it is an order-complete ordered field [(R1)–(R3) and (0.74), pages 78–79].

We now proceed to identify within  $\mathbb{R}$  the sets  $\mathbb{N}$ ,  $\mathbb{Z}$ , and  $\mathbb{Q}$  of natural numbers, integers, and rational numbers, respectively.

Step 1: Obtain  $\mathbb{N}$  within  $\mathbb{R}$ .

Call a subset  $E$  of  $\mathbb{R}$  *inductive* when:

- $0 \in E$ ; and
- $n + 1 \in E$  whenever  $n \in E$ .

The collection  $\mathcal{I}$  of all inductive subsets of  $\mathbb{R}$  is nonempty because the entire set  $\mathbb{R}$  is inductive. Then we may form the intersection  $\bigcap \mathcal{I}$  of this collection which, with respect to subset inclusion, is the smallest inductive subset of  $\mathbb{R}$ . And that is the subset we identify as the set  $\mathbb{N}$  of all natural numbers.

From this very definition of  $\mathbb{N}$ , it follows that if  $E \subset \mathbb{N}$  with  $0 \in E$  and  $n + 1 \in E$  whenever  $n \in \mathbb{N}$ , then  $E = \mathbb{N}$ . In other words, the Axiom of Induction (0.1) is satisfied, and hence the Principle of Mathematical Induction (0.2) holds.

Inductive proofs may now be used to establish that if  $m, n \in \mathbb{N}$ , then  $m + n \in \mathbb{N}$  and  $m \cdot n \in \mathbb{N}$ . Thus it is meaningful to restrict the operations of addition and multiplication on  $\mathbb{R}$  to  $\mathbb{N}$ . Owing to the assumption that  $\mathbb{R}$  is a field, it follows immediately that the resulting

operations of addition and multiplication on  $\mathbb{N}$  satisfy all the usual algebraic properties (A1)–(A3) and (A5), (M1)–(M3) and (M5), and (D1), as listed on pages 22–23.

The total ordering on  $\mathbb{R}$  induces a total ordering of  $\mathbb{N}$ . Owing to the assumption that  $\mathbb{R}$  is an ordered field, it follows immediately that the resulting total ordering of  $\mathbb{N}$  satisfies the compatibility properties (R3), as formulated on page 78.

Now that we have identified the integers  $\mathbb{N}$  within  $\mathbb{R}$  and seen why it has all the usual properties, we turn to  $\mathbb{Z}$  and  $\mathbb{Q}$ .

Step 2: Obtain the integers  $\mathbb{Z}$  within the reals  $\mathbb{R}$ .

We identify  $\mathbb{Z}$  as the subset

$$\mathbb{Z} = \mathbb{N} \cup \{-n : n \in \mathbb{N}\}$$

of the set  $\mathbb{R}$  of all real numbers.

From the fundamental assumption about  $\mathbb{R}$ , it follows that addition and multiplication on  $\mathbb{R}$  restrict to operations on  $\mathbb{Z}$  and that these operations satisfy algebraic properties (A1)–(A5), (M1)–(M3) and (M5), and (D1). Moreover, the total ordering of  $\mathbb{R}$  induces a total ordering of  $\mathbb{Z}$  that is compatible with the algebraic operations in the sense of (R3).

Step 3: Obtain  $\mathbb{Q}$  inside  $\mathbb{R}$ . We identify  $\mathbb{Q}$  as the subset

$$\mathbb{Q} = \{m \cdot n^{-1} : m \in \mathbb{Z}, n \in \mathbb{Z}, n \neq 0\}$$

of the set  $\mathbb{R}$  of all real numbers.

Addition and multiplication on  $\mathbb{R}$  restrict to operations on  $\mathbb{Q}$ , the total ordering of  $\mathbb{R}$  induces a total ordering of  $\mathbb{Q}$ , and then  $\mathbb{Q}$  becomes an ordered field.

Having made the fundamental assumption that  $\mathbb{R}$  is a order-complete ordered field, we can thus successively recover  $\mathbb{N}$ ,  $\mathbb{Z}$ , and  $\mathbb{Q}$  as subsets of  $\mathbb{R}$ .

It is possible, conversely, to start with fundamental assumptions about  $\mathbb{N}$ —the **Peano Postulates**—and then successively *construct* the sets  $\mathbb{Z}$ ,  $\mathbb{Q}$ , and  $\mathbb{R}$  and their algebraic and ordering relations. (In fact, it is even possible to retreat a step back from  $\mathbb{N}$ , by using the axioms of set theory to define the set  $\mathbb{N}$  and to verify that it satisfies the Peano Postulates—see Remark 0.110.)

The constructive path consists of two stages:

Stage 1: Construct  $\mathbb{Q}$ . This stage is typically carried out in two steps:

Step (i): Construct  $\mathbb{Z}$  from  $\mathbb{N}$ .

Step (ii): Construct  $\mathbb{Q}$  from  $\mathbb{Z}$ .

We hint at how to effect those two steps in Exercise 161 and Exercise 162, respectively, from Section 0.9; both exercises use the notions of “equivalence relation” and “quotient set.”

Another way to construct  $\mathbb{Z}$  from  $\mathbb{N}$  is described in Eisenberg [22, Section D.4] and carried out in detail in Garling [28, Section 2.5]. For details of constructing  $\mathbb{Q}$  from  $\mathbb{Z}$ , as the “field of quotients” of  $\mathbb{Z}$ , see, for example, Birkhoff and Mac Lane [5, Section 2.2].

Stage 2: Construct  $\mathbb{R}$  from  $\mathbb{Q}$ . There are two principal methods for carrying out this step, both of which are outlined in Wikipedia [69]:

- Construct  $\mathbb{R}$  from “Dedekind cuts” of  $\mathbb{Q}$ . This method is carried out in Rudin [57, Appendix to Chapter 1]. The notion of a Dedekind cut is defined in our Exercise 133.
- Construct  $\mathbb{R}$  by “completing” the metric space  $\mathbb{Q}$ . This method is indicated in our Exercises 1.130–1.131 from Section 1.5 and carried out in detail in Taylor [64, Section 6].

Taking the constructive path to the real number system is beyond the scope of this book. But if you wish to pursue the matter, see the references above.

subsec:complex-numbers

**Complex numbers**complex numbers  
complex numbers  
real part  
imaginary part

The set  $\mathbb{C}$  of complex numbers was introduced in the [subsection “Special sets”](#) (page 9) as  $\{x + yi : x, y \in \mathbb{R}\}$ . While complex numbers are indeed sometimes described as being creatures of the form  $x + yi$  with  $x$  and  $y$  real, it is useful to view complex numbers as being ordered pairs  $\langle x, y \rangle$  of real numbers, that is, as points in the plane  $\mathbb{R} \times \mathbb{R}$ . Thus

$$\mathbb{C} = \mathbb{R} \times \mathbb{R} = \{\langle x, y \rangle : x \in \mathbb{R}, y \in \mathbb{R}\}$$

as sets. The first coordinate  $x$  of a complex number  $z = \langle x, y \rangle$  is its **real part** and the second coordinate  $y$  is its **imaginary part**.

Operations of addition and multiplication on the set  $\mathbb{C}$  of complex numbers are defined by the formulas

$$\begin{aligned}\langle x, y \rangle + \langle u, v \rangle &= \langle x + u, y + v \rangle, \\ \langle x, y \rangle \cdot \langle u, v \rangle &= \langle xu - yv, xv + yu \rangle.\end{aligned}$$

Direct calculation establishes that these operations make  $\mathbb{C}$  a field, as already asserted in the [subsection “Operations”](#) (page 22). This field has additive identity  $\langle 0, 0 \rangle$  and multiplicative identity  $\langle 1, 0 \rangle$ . The special complex number  $i$  is defined as

$$i = \langle 0, 1 \rangle,$$

and then  $i^2 = \langle 0, -1 \rangle$ , where as usual the square of a number is its product with itself.

The injection

$$\begin{aligned}\iota: \mathbb{R} &\rightarrow \mathbb{C} \\ x &\mapsto \langle x, 0 \rangle\end{aligned}$$

—that symbol is a Greek iota, not  $i$  or  $i$ —creates a one-to-one correspondence between the set of real numbers, on the one hand, and the set of complex numbers having imaginary part 0, on the other hand. Addition and multiplication of real numbers are performed the same way as are addition and multiplication of the complex numbers corresponding to them under  $\iota$ ; in other words,  $\iota(x + u) = \iota(x) + \iota(u)$ , and similarly for multiplication. For this reason, we make no distinction between a real number  $x$  and the corresponding complex number  $\iota(x) = \langle x, 0 \rangle$ : we say that we *identify* each real number  $x$  with the corresponding complex number  $\langle x, 0 \rangle$ . This identification makes  $\mathbb{R}$  a subset of  $\mathbb{C}$  (and, in fact, a subfield).

After that identification, each complex number  $z = \langle x, y \rangle$  can be written uniquely in the form

$$z = x + yi,$$

for real numbers  $x$  and  $y$ , because that just means

$$\langle x, y \rangle = \langle x, 0 \rangle + \langle y, 0 \rangle \langle 0, 1 \rangle.$$

It follows that  $i^2 = -1$ . Moreover, the operations of addition and multiplication in  $\mathbb{C}$  are now given by

$$\begin{aligned}(x + yi) + (u + vi) &= (x + u) + (y + v)i, \\ (x + yi) \cdot (u + vi) &= xu - yv + (xv + yu)i,\end{aligned}$$

Each nonzero  $z = x + yi \in \mathbb{C}$  has the multiplicative inverse

$$z^{-1} = \frac{1}{z} = \frac{x}{x^2 + y^2} - \frac{y}{x^2 + y^2} i.$$

conjugate  
modulus  
complex-valued function  
zero of a function  
root!function@of a function  
complex exponential function

As usual in algebra, for a complex number  $z$ , its nonnegative powers are defined recursively by

$$\begin{aligned} z^0 &= 1, \\ z^{n+1} &= z^n z \end{aligned} \quad (n = 0, 1, 2, \dots)$$

[and its negative powers by  $z^{-k} = (z^{-1})^k$  for  $k = 1, 2, 3, \dots$ ]

The **conjugate**  $\bar{z}$  of a complex  $z = x + yi$  is defined by  $\bar{z} = x - yi$ . The **modulus**  $|z|$  of a complex  $z = x + yi$  is defined by  $|z| = \sqrt{x^2 + y^2}$ , so that  $|z| = \sqrt{z\bar{z}}$ .

Complex modulus has the following two properties, which hold for all  $z, w \in \mathbb{C}$ :

$$\begin{aligned} |zw| &= |z||w| \\ |z + w| &\leq |z| + |w| \end{aligned}$$

{eq:complex-modulus-properties} (\*)

The first of those properties implies that  $|z^n| = |z|^n$  for all nonnegative integers  $n$ ; the second extends to a sum of finitely many complex numbers.

A **complex-valued function** on a set  $X$  is a function with domain  $X$  and codomain  $\mathbb{C}$ . As in the case for a real-valued function, so for a complex-valued function  $f: X \rightarrow \mathbb{C}$ , a **zero** (or **root**) of  $f$  is an element  $a$  in the domain  $X$  of  $f$  for which  $f(a) = 0$ . For example, both  $i$  and  $-i$  are roots of the polynomial function  $f: \mathbb{C} \rightarrow \mathbb{C}$  defined by  $f(z) = z^2 + 1$ ; and both  $1 + \sqrt{3}i$  and  $1 - \sqrt{3}i$  are roots of the polynomial function  $f: \mathbb{C} \rightarrow \mathbb{C}$  defined by  $f(z) = z^2 - 2z + 4$ .

Our use of complex numbers will be when describing points on the unit circle

$$S_1 = \{(x, y) : x^2 + y^2 = 1\}$$

in the  $xy$ -plane  $\mathbb{R}^2$ . In terms of complex numbers, the unit circle is the subset

$$S_1 = \{z \in \mathbb{C} : |z| = 1\}$$

of the “ $z$ -plane”  $\mathbb{C}$  or, in terms of polar coordinates,

$$S_1 = \{\cos \theta + i \sin \theta : 0 \leq \theta \leq 2\pi\} = \{\cos 2\pi s + i \sin 2\pi s : 0 \leq s \leq 1\}$$

It is more succinct to represent such a point on  $S_1$  in the complex exponential form

$$z = \exp(\theta i) = e^{\theta i}.$$

For our purposes, for a real number  $\theta$  we shall *define* both the value  $\exp(\theta i)$  of the complex exponential function and the power  $e^{\theta i}$  by:

$$\exp(\theta i) = e^{\theta i} = \cos \theta + i \sin \theta,$$

so that

$$\exp(2\pi s i) = e^{2\pi s i} = \cos 2\pi s + i \sin 2\pi s$$

The addition properties, periodicity, and special values of the real sine and cosine functions yield the basic identities

$$\begin{aligned} e^{(\alpha+\beta)i} &= e^{\alpha i} e^{\beta i}, \\ e^{-\alpha i} &= 1/e^{\alpha i}, \\ e^{2\pi i} &= 1, \quad e^{\pi i} = -1 \end{aligned}$$

for the complex exponential.<sup>4</sup>

An induction on  $n$  establishes the following identity.

<sup>4</sup>The complex exponential function  $z \mapsto \exp z$  can, in fact, be defined as the sum of the everywhere convergent power series  $\sum_{n=0}^{\infty} z^n/n!$ ; and the preceding identities can be obtained by applying properties of power series. Then the addition formulas and other properties of the real sine and cosine may be derived from those.

prop:deMoivre

**0.88 Proposition (De Moivre's Formula).** For all  $\theta \in \mathbb{R}$ ,

$$(e^{i\theta})^n = e^{n i \theta} \quad (n = 0, 1, 2, \dots)$$

standard map from  $\mathbb{R}$  to  $S_1$  @ standard  
complex exponential function  
complex numbers

In terms of real and complex parts, De Moivre's Formula asserts that, for all real  $\theta$ ,

$$(\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta \quad (n = 0, 1, 2, \dots)$$

The standard map

$$\begin{aligned} \mathbb{R} &\rightarrow S_1 \\ s &\mapsto e^{2\pi s i} \end{aligned}$$

from the real line onto the circle  $S_1$  and its restriction

$$\begin{aligned} I &\rightarrow S_1 \\ s &\mapsto e^{2\pi s i} \end{aligned}$$

from the unit interval  $I = [0, 1]$  onto the circle  $S_1$  will be of particular interest later, when we study paths in topological space, beginning in [Section 5.3](#).

### EXERCISES FOR SECTION 0.8

**120.** Let  $A$  and  $B$  be nonempty sets of positive real numbers. Denote by  $A + B$  the set  $\{a + b : a \in A, b \in B\}$ .

prob-part:max-sum-two-sets

(a) Suppose each of  $A$  and  $B$  has a maximum. Show that  $A + B$  has one also and that  $\max(A + B) \leq \max A + \max B$ .

prob-part:sup-sum-two-sets

(b) Suppose each of  $A$  and  $B$  is bounded above. Show that  $A + B$  has a supremum and that  $\sup(A + B) \leq \sup A + \sup B$ .

(c) Can the inequalities in (a) and (b) be strengthened to equalities?

**121.** Reprove [Proposition 0.75](#) as follows: Let  $\emptyset \neq B \subset \mathbb{R}$  have a lower bound. Define  $L$  to be the set of all lower bounds of  $B$ . Show that  $L$  has a supremum and then that  $\sup L = \inf B$ .

**122.** Let  $E$  be a nonempty set of integers.

(a) Suppose  $E$  has an upper bound in  $\mathbb{R}$ . Use order-completeness to show that  $E$  has a greatest element.

(b) Suppose  $E$  has an upper bound  $b$  that is an integer. *Without* using order-completeness, show that  $E$  has a greatest element. [Hint: Apply the Well-ordering Principle (0.3) for  $\mathbb{N}$  to the set  $\{b - x : x \in E\}$ .]

prob:x-between-n-nplus1

**123.** Let  $x \in \mathbb{R}$ . Establish the existence of some integer  $n$  such that  $n \leq x < n + 1$ . Then prove that such an  $n$  is unique.

prob:Archimedean-analog-rationals

**124.** *Without* using the Archimedean Ordering Property, prove the following analog of that property for the rational numbers: Let  $r$  and  $\varepsilon$  be rational numbers with  $\varepsilon > 0$ . Then there exists a positive integer  $n$  such that  $n\varepsilon > r$ . (Hint: Use [Exercise 20](#).)

**125.** Let  $x \in \mathbb{R}$  and  $\varepsilon > 0$ .

(a) Show that there exist infinitely many rational numbers  $q$  with  $|x - q| < \varepsilon$ .

(b) Must there be infinitely many irrational numbers  $y$  with  $|x - y| < \varepsilon$ ?

prob:inject-R-into-powerset-Q

**126.** Show that the map  $\mathbb{R} \rightarrow \mathcal{P}(\mathbb{Q})$  given by  $x \mapsto \{q \in \mathbb{Q} : q \leq x\}$  is injective.

prob:disj-collection-open-intervals-in- $\mathbb{R}$  127. (a) Exhibit a denumerable collection  $\mathcal{D}$  of pairwise disjoint open intervals in  $\mathbb{R}$  such that  $\bigcup \mathcal{D} = \mathbb{R}$ .

repeating expansion  
open-intervals-in- $\mathbb{R}$  (b) Prove that there does not exist any uncountable collection of pairwise disjoint open intervals in  $\mathbb{R}$ . (*Hint*: Use the order-density of  $\mathbb{Q}$  in  $\mathbb{R}$ .)

terminating expansion  
uncountable set!real numbers@and real numbers  
real numbers!uncountable set@as uncountable set  
128. Prove the uniqueness assertions of Proposition 0.77.

prob:repeating-expansion-iff-rational 129. For a base  $b > 1$ , a base- $b$  expansion  $(0.d_1d_2d_3 \dots d_n \dots)_b$  of a number  $x \in [0, 1]$  is said to be **repeating** when it has the form

binary expansion  
Dedekind cut  
real numbers!Dedekind cuts@and Dedekind cuts  
Dedekind, Richard

for some positive integers  $m$  and  $k$ ; such an expansion may be abbreviated by the form

$$(0.d_1d_2 \dots d_{m-1} \overline{d_md_{m+1}d_{m+2}d_{m+k} \dots})_b.$$

In particular, a terminating expansion is repeating. For example,

$$3/40 = (0.075\overline{0} \dots)_{10}, \quad 1/10 = (0.000\overline{1100} \dots)_2.$$

Show that a real number in  $[0, 1]$  is rational if and only if it has *some* repeating base- $b$  expansion.

130. Use binary expansion [see the subsection “Base expansion” (page 79)] to give another proof that  $[0, 1]$ , and hence  $\mathbb{R}$ , is uncountable.

[*Hint*: The map  $\{0, 1\}^{\mathbb{N}^*} \rightarrow [0, 1]$  sending a sequence  $\langle d_1, d_2, d_3, \dots, d_n, \dots \rangle$  to the real number with binary expansion  $(0.d_1d_2d_3 \dots d_n \dots)_2$  is surjective but not injective. Restrict its domain so that it remains surjective and becomes injective.]

prob:sqrt2-exists 131. From order-completeness of  $\mathbb{R}$  deduce the existence of  $\sqrt{2}$ , that is, of a real number  $s > 0$  for which  $s^2 = 2$ .

[*Hint*: First show that the set  $A = \{a \in \mathbb{R} : a > 0 \text{ and } a^2 < 2\}$  has a supremum  $s$ . If  $s^2 < 2$ , show that there is a positive integer  $m$  with  $(s + 1/m)^2 \leq s^2 + (2s + 1)/n < 2$ ; if  $s^2 > 2$ , show that there is a positive integer  $n$  with  $(s - 1/n)^2 > s^2 - (2s)/n > 2$ .]

prob:Q-not-order-complete 132. Prove that the rationals  $\mathbb{Q}$ , with their usual order, are *not* order-complete.

[*Hint*: Show that the subset  $A = \{a \in \mathbb{Q} : a^2 < 2\}$  is bounded above in  $\mathbb{Q}$  but does not have a least upper bound in  $\mathbb{Q}$ .]

prob:dedekind-cut 133. A **Dedekind cut** is a pair  $\langle A, B \rangle$  of subsets of  $\mathbb{Q}$  such that  $A \neq \emptyset$ ,  $B \neq \emptyset$ ,  $A \cap B = \emptyset$ ,  $A \cup B = \mathbb{Q}$ , and  $a < b$  for each  $a \in A$  and each  $b \in B$ . (The term ‘Dedekind cut’ is named after Richard Dedekind, who based a construction of the set of real numbers upon such “cuts” of the set of rational numbers.)

For example, the pair  $\langle A, B \rangle$  given by

$$A = \{a \in \mathbb{Q} : a < 0 \text{ or } a^2 \leq 2\}, \quad B = \{b \in \mathbb{Q} : b > 0 \text{ and } 2 \leq b^2\}$$

is a Dedekind cut. (These sets  $A$  and  $B$  were deliberately described strictly in terms of rational numbers. However, they may be more succinctly described in terms of real numbers as

$$A = \{a \in \mathbb{Q} : a < \sqrt{2}\}, \quad B = \{b \in \mathbb{Q} : \sqrt{2} < b\},$$

respectively.)

Prove that if  $\langle A, B \rangle$  is any Dedekind cut, then there exists a *unique* real number  $x$  with either

$$A = \{a \in \mathbb{Q} : a \leq x\}, \quad B = \{b \in \mathbb{Q} : x < b\}$$

or

$$A = \{a \in \mathbb{Q} : a < x\}, \quad B = \{b \in \mathbb{Q} : x \leq b\}.$$

*Note:* The number  $x$  need not be rational; when  $x$  is irrational, the two possibilities for  $(A, B)$  indicated above are really one and the same.

**134.** Prove that the set of irrational numbers is order-dense in the set of real numbers.

**135.** Let  $\langle [c_n, d_n] \rangle_{n \in \mathbb{N}}$  be a decreasing sequence of closed intervals in  $\mathbb{R}$ ; we do *not* assume that  $\lim_{n \rightarrow \infty} (d_n - c_n) = 0$ . Show that  $\bigcap_{n=0}^{\infty} [c_n, d_n] \neq \emptyset$ . (*Hint:* Apply the Nested Interval Property to an appropriate sequence of subintervals of the given intervals.)

**136.** Does the Nested Interval Property remain true if the intervals are open instead of closed?

**137.** Use the definition of  $\mathbb{N}$  from the subsection “Natural numbers, integers, and rationals within the reals” (page 87) to prove that the operations of addition and multiplication on  $\mathbb{R}$  do, in fact, restrict to operations on  $\mathbb{N}$ ; that is, show that  $m + n \in \mathbb{N}$  and  $m \cdot n \in \mathbb{N}$  whenever  $m, n \in \mathbb{N}$ .

**138.** Use the definition of  $\mathbb{Q}$  from the subsection “Natural numbers, integers, and rationals within the reals” (page 87) to prove that  $\mathbb{Q}$  is an ordered field.

**139. (a)** Verify that the definitions of addition and multiplication of complex numbers, as given in the subsection “Complex numbers” (page 89), do make  $\mathbb{C}$  a field.

**(b)** Establish an explicit formula for the multiplicative inverse  $z^{-1}$  of a complex number  $z = x + yi$  in terms of its real and imaginary parts.

**140.** Prove the properties (\*) of complex modulus given on page 90.

**141.** Let  $p: \mathbb{R} \rightarrow S_1$  be the map from the real line to the unit circle (the latter set considered as a subset of  $\mathbb{C}$ ) given by  $p(\theta) = e^{2\pi i \theta}$ . Consider the “path”  $\sigma: [0, 1] \rightarrow S_1$  defined by  $\sigma(t) = e^{\pi i t/4}$ . Show that  $\sigma$  is liftable to  $\mathbb{R}$  through  $p$  in the sense of Exercise 34. *Note:* This restates Exercise 34 (b) in terms of complex numbers.

## 0.9 Equivalence Relations

sec:equiv

Previously we examined two important special types of relations—functional relations (Section 0.3) and ordering relations (Section 0.7). Now we study a third special type—equivalence relations.

### Equivalence relations

subsec:equiv-rel

Suppose that the objects in a given collection are classified by some rule that declares certain of the objects to be alike and others unlike. We may represent the given collection by a set  $X$  and the classifying rule by a relation  $\sim$  in  $X$  to  $X$ , with ‘ $x \sim y$ ’ being interpreted to mean that  $x$  is like  $y$  under the rule. Now the connotations of ‘like’ suggest that each  $x$  is like itself, that  $y$  is like  $x$  if  $x$  is like  $y$ , and that  $x$  is like  $z$  if  $x$  is like  $y$  and  $y$  is like  $z$ . Hence

congruence modulo  $m$  the relation representing the classifying rule should be an “equivalence relation” in the following sense.

**0.89 Definition.** Let  $X$  be a set. An **equivalence relation** on  $X$  is a relation  $\sim$  in  $X$  to  $X$  that has all three of the properties:

**(EqR1) Reflexivity:** For each  $x \in X$ ,

$$x \sim x.$$

**(EqR2) Symmetry:** For all  $x, y \in X$ ,

$$x \sim y \implies y \sim x.$$

**(EqR3) Transitivity:** For all  $x, y, z \in X$ ,

$$x \sim y \text{ and } y \sim z \implies x \sim z.$$

**0.90 Examples.** (1) Let  $\sim$  be the relation in  $\mathbb{R}$  to  $\mathbb{R}$  given by

$$x \sim y \iff |x| = |y|.$$

Then  $\sim$  is an equivalence relation on  $\mathbb{R}$ .

(2) Fix a positive integer  $m$ . Let  $\sim$  be the relation in  $\mathbb{Z}$  to  $\mathbb{Z}$  given by

$$n \sim k \iff n - k = dm \text{ for some } d \in \mathbb{Z}.$$

Equivalently (pardon the pun),

$$n \sim k \iff m \mid (n - k),$$

where  $\mid$  is the divisibility relation [see [Examples 0.55 \(2\)](#)]. Or, in still other terms,  $n \sim k$  if and only if  $n$  and  $k$  have the same remainder when divided by  $m$ . Then  $\sim$ , so defined, is an equivalence relation on  $\mathbb{Z}$ , which is called **congruence modulo  $m$** . Instead of  $n \sim k$  we usually write

$$n \equiv k \pmod{m},$$

which is read as “ $n$  is congruent to  $k$  modulo  $m$ ” or, more tersely, “ $n$  is congruent to  $k$  mod  $m$ .” We call  $m$  the **modulus** of this equivalence relation.

Take  $m = 2$  above. Then

$$\begin{aligned} n \sim k &\iff n \equiv k \pmod{2} \\ &\iff n - k \text{ is even} \\ &\iff n \text{ and } k \text{ are both even or both odd.} \end{aligned}$$

(3) Let  $\sim$  be the relation in  $\mathbb{R}$  to  $\mathbb{R}$  given by

$$x \sim y \iff x - y \in \mathbb{Z}.$$

Then  $\sim$  is an equivalence relation on  $\mathbb{R}$ .

(4) Let  $X$  be the set of all triangles in the plane. Define  $\sim$  by

$$x \sim y \iff x \text{ is congruent to } y$$

(where ‘congruent’ has its usual geometric meaning.) Then  $\sim$  is an equivalence relation on  $X$ .



- (5) Let  $\mathcal{A}$  be a collection of sets. For  $A, B \in \mathcal{A}$ , define

equivalence kernel

$$A \sim B \iff \text{there exists a bijection } f: A \rightarrow B.$$

Then  $\sim$  is an equivalence relation on  $\mathcal{A}$  that classifies two sets belonging to  $\mathcal{A}$  as alike when they have the same “number” of elements—technically, when they have the same “cardinality”; see the [subsection “Cardinality”](#) of [Section 0.10](#).

If  $\mathbb{N} \in \mathcal{A}$ , then for  $A \in \mathcal{A}$  we have

$$A \sim \mathbb{N} \iff A \text{ is denumerable.}$$

ex:equality-as-eq-rel

- (6) Let  $X$  be any set. Then the relation  $\sim$  in  $X$  to  $X$  defined by

$$x \sim y \iff x = y$$

is an equivalence relation on  $X$  that classifies no object as like another distinct from itself.

ex:equiv-ker

- (7) Let  $f: X \rightarrow Y$  be a map. Then the relation  $\sim$  in  $X$  to  $X$  given by

$$x \sim y \iff f(x) = f(y) \quad (x, y \in X)$$

is an equivalence relation on  $X$ , called the **equivalence kernel of  $f$** . For example, the equivalence kernel of an identity map  $\iota_X: X \rightarrow X$  is the diagonal  $\Delta_X$  of  $X \times X$ . More generally, when  $f: X \rightarrow Y$  is injective, the equivalence kernel is just the identity relation; the larger the equivalence kernel is, the further  $f$  is from being injective.

from-non-reflexive-not-transitivity-rel

- (8) Define a relation  $\sim$  in the set  $X = \{0, 1, 2, 3\}$  to  $X$  by

$$0 \sim 2, \quad 2 \sim 3.$$

This relation is not reflexive, not symmetric, and not transitive. Extend this relation by taking also

$$0 \sim 0, \quad 1 \sim 1, \quad 2 \sim 2, \quad 3 \sim 3,$$

so that now this relation is also reflexive. But it is still neither symmetric nor transitive. Extend the extended relation by taking also

$$2 \sim 0, \quad 3 \sim 2,$$

so that now it is symmetric as well as reflexive. But still the relation is not transitive. However, extending it further by taking also

$$0 \sim 3, \quad 3 \sim 0$$

makes this relation now reflexive, symmetric, and transitive, in other words, makes it an equivalence relation.  $\diamond$

The preceding example illustrates a situation we shall encounter in constructing certain topological “quotient spaces”: we may explicitly give only the fundamental identifications a relation should create and leave implicit all the additional identifications that symmetry, reflexivity, and transitivity would demand so as to provide an actual equivalence relation.

In general, let  $\alpha$  be a relation in a set  $X$  to itself, not necessarily reflexive, symmetric, or transitive. Then we can, as in the preceding example, form the least equivalence relation  $\sim$  on  $X$  for which  $x \alpha y$  implies  $x \sim y$ , namely, the intersection of all equivalence relations on  $X$  containing  $\alpha$ . (The collection of equivalence relations on  $X$  containing  $\alpha$  is nonempty because  $X \times X$  is such an equivalence relation!)

reflexive, symmetric,  
transitive closure  
quotient set

**0.91 Definition.** Let  $\alpha$  be a relation in a set  $X$  to itself. The least equivalence relation  $\sim$  on  $X$  for which  $x \alpha y$  implies  $x \sim y$  is called the **reflexive, symmetric, transitive closure of  $\alpha$** .

reflexive-symmetric-transitive-closure

**0.92 Examples.** (1) The reflexive, symmetric, transitive closure of the relation  $\{\langle 0, 2 \rangle, \langle 2, 3 \rangle\}$  in  $\{0, 1, 2, 3\}$  to itself is the equivalence relation  $\sim$  of [Examples 0.90 \(8\)](#).

(2) Fix a positive integer  $m$ . Let  $\alpha$  be the relation in  $\mathbb{Z}$  to itself given by

$$n \alpha (n + m) \quad (n \in \mathbb{Z})$$

Then the reflexive, symmetric, transitive closure of  $\alpha$  is the equivalence relation of [Examples 0.90 \(2\)](#), namely, the relation  $\sim$  given by

$$n \sim k \iff n - k = dm \text{ for some } d \in \mathbb{Z}.$$

(3) Let  $\sim$  be the relation in  $\mathbb{Z} \times \mathbb{Z}^*$  to itself that identifies  $\langle m, n \rangle$  with  $\langle mk, nk \rangle$  for all  $m \in \mathbb{Z}$  and all  $n, k \in \mathbb{Z}^*$ . Then the reflexive, symmetric, transitive closure of this relation is the equivalence relation  $\sim$  on  $\mathbb{Z} \times \mathbb{Z}^*$  given by

$$\langle m, n \rangle \sim \langle s, t \rangle \iff mt = ns. \quad \diamond$$

### Quotient sets

subsec:quot-set

Given an equivalence relation on a set  $X$ , we may group together with each  $x \in X$  all those elements of  $X$  (including  $x$  itself) that are “like”  $x$ .

def:quotient-set

**0.93 Definition.** Let  $\sim$  be an equivalence relation on a set  $X$ . For each  $x \in X$ , the **equivalence class of  $x$  under  $\sim$**  is the subset

$$\llbracket x \rrbracket_{\sim} = \{y \in X : x \sim y\}$$

of  $X$ , and each  $y \in \llbracket x \rrbracket_{\sim}$  is a **representative of** this equivalence class. The **quotient set of  $X$  under  $\sim$**  is the collection

$$X/\sim = \{\llbracket x \rrbracket_{\sim} : x \in X\}$$

consisting of all these equivalence classes, and the **quotient map (induced by  $\sim$ )** is the surjection

$$\begin{aligned} q: X &\rightarrow X/\sim \\ x &\mapsto \llbracket x \rrbracket_{\sim} \end{aligned}$$

that sends each element of  $X$  to its equivalence class under  $\sim$ .

When a particular equivalence relation  $\sim$  on  $X$  is understood, we denote the equivalence class  $\llbracket x \rrbracket_{\sim}$  of an element  $x$  of  $X$  simply by  $\llbracket x \rrbracket$ . Then the corresponding quotient map is given by

$$\begin{aligned} q: X &\rightarrow X/\sim \\ x &\mapsto [x] \end{aligned}$$

**0.94 Examples.** (1) On the set  $\mathbb{Z}$  of integers, let  $\sim$  be the relation of congruence modulo  $m = 2$ —see [Examples 0.90 \(2\)](#). Then the preceding definition gives quotient set

$$\llbracket n \rrbracket = \begin{cases} E & \text{if } n \text{ is even,} \\ O & \text{if } n \text{ is odd,} \end{cases}$$

where

$$\begin{aligned} E &= \{\dots, -6, -4, -2, 0, 2, 4, 6, \dots\}, \\ O &= \{\dots, -5, -3, -1, 1, 3, 5, \dots\}, \end{aligned}$$

the set of all even integers and the set of all odd integers, respectively. Thus the quotient set is

$$\mathbb{Z}/\sim = \{E, O\}.$$

(2) On the set  $\mathbb{R}$  of real numbers let  $\sim$  be the equivalence relation of [Examples 0.90 \(1\)](#), namely,

$$x \sim y \iff |x| = |y|.$$

For each  $x \in \mathbb{R}$ , its equivalence class is

$$\llbracket x \rrbracket = \{x, -x\}.$$

Thus the quotient set is

$$X/\sim = \{\{x, -x\} : x \in \mathbb{R} \text{ and } x \geq 0\}.$$

(3) For the equivalence relation of equality in any set  $X$ —see [Examples 0.90 \(6\)](#)—the definition gives

$$\llbracket x \rrbracket = \{x\}$$

for each  $x \in X$ , and so

$$X/\sim = \{\{x\} : x \in X\},$$

the collection of all singletons formed by elements of  $X$ .  $\diamond$

In the terms of [Definition 0.89](#) we can say in several different ways when two objects are alike under an equivalence relation.

**0.95 Lemma (characterization of equivalent elements).** *Let  $\sim$  be an equivalence relation on a set  $X$  let and let  $x, y \in X$ . Then the following conditions are equivalent:*

(i)  $x \sim y$ .

(ii)  $\llbracket x \rrbracket = \llbracket y \rrbracket$ .

(iii)  $q(x) = q(y)$ , where  $q: X \rightarrow X/\sim$  is the quotient map.

(iv)  $\llbracket x \rrbracket$  intersects  $\llbracket y \rrbracket$ .

**Proof.** (i)  $\implies$  (ii): Assume (i), namely,  $x \sim y$ . If  $z \in \llbracket x \rrbracket$ , then  $x \sim z$ , by symmetry  $z \sim x$ , by transitivity  $z \sim y$ , by symmetry again  $y \sim z$ , and so  $z \in \llbracket y \rrbracket$ ; thus  $\llbracket x \rrbracket \subset \llbracket y \rrbracket$ . Similarly,  $\llbracket y \rrbracket \subset \llbracket x \rrbracket$

partition  
equivalence relation! partition induced by partition

(ii)  $\implies$  (i): Assume (ii), namely,  $\llbracket x \rrbracket = \llbracket y \rrbracket$ . Then  $y \in \llbracket y \rrbracket$  because  $y \sim y$  by reflexivity, and so  $y \in \llbracket x \rrbracket$ , that is,  $x \sim y$ .

(ii)  $\iff$  (iii) by definition of  $q$ , and (ii)  $\implies$  (iv) because  $x \in \llbracket x \rrbracket$ .

(iv)  $\implies$  (i): Assume (iv). Choose some  $z \in \llbracket x \rrbracket \cap \llbracket y \rrbracket$ . Then  $x \sim z$  and  $y \sim z$ , and so by symmetry and transitivity  $x \sim y$ .  $\square$

## Partitions

subsec:partitions

The preceding lemma tells us, most importantly, that any two equivalence classes are either identical or else disjoint. And that suggests the following definition.

def:partition

**0.96 Definition.** A collection  $\mathcal{A}$  of *nonempty* subsets of a set  $X$  is called a **partition** of  $X$ , and is said to **partition**  $X$ , when each point of  $X$  belongs to exactly one member of  $\mathcal{A}$ . In other terms, a collection  $\mathcal{A}$  of subsets of  $X$  is a partition of  $X$  when:

- each member  $A$  of  $\mathcal{A}$  is nonempty;
- each element of  $X$  belongs to some member of  $\mathcal{A}$  (that is,  $X = \bigcup \mathcal{A}$ ); and
- no two members of  $\mathcal{A}$  intersect (that is,  $\mathcal{A}$  is pairwise disjoint).

A fine point: In the preceding definition, we do *not* require that the collection  $\mathcal{A}$  itself be nonempty, only that each of its members—if there are any!—be nonempty. If the set  $X$  is empty, then necessarily any partition of  $X$  will have no members, that is, the only partition of  $X$  will be the empty collection. And conversely, if a partition of  $X$  is empty, then the set  $X$  must be empty.

thm:equiv-rel-vs-partition

**0.97 Theorem (characterization of quotient sets as partitions).** Let  $X$  be a set.

thm-part:quotient-set-is-partition

(1) If  $\sim$  is an equivalence relation on  $X$ , then the quotient set  $X/\sim$  is a partition of  $X$ .

thm:part:partition-induces-equiv-rel

(2) If  $\mathcal{A}$  is a partition of  $X$ , then the relation  $\sim$  in  $X$  to  $X$  given by

$$x \sim y \iff x \in A \text{ and } y \in A \text{ for some } A \in \mathcal{A}$$

is an equivalence relation on  $X$ .

**Proof.** (1) follows from [Lemma 0.95](#) and the fact that  $x \in \llbracket x \rrbracket \in X/\sim$  for each  $x \in X$ .

(2) Let  $\mathcal{A}$  be a partition of  $X$  and define  $\sim$  as indicated. We show that  $\sim$  is transitive. Suppose  $x, y, z \in X$  with  $x \sim y$  and  $y \sim z$ . There exist members  $A, B \in \mathcal{A}$  with

$$x, y \in A, \quad y, z \in B.$$

Since  $y \in A \cap B$ , we have  $A = B$ . Then  $x, z \in A$ , and so  $x \sim z$ . Thus  $\sim$  is transitive. The proofs that  $\sim$  is reflexive and symmetric are even easier.  $\square$

Given a partition  $\mathcal{A}$  of a set  $X$ , the equivalence relation defined in [Theorem 0.97 \(2\)](#) is said to be **induced by  $\mathcal{A}$** .

A partition  $\mathcal{A}$  of a set  $X$  may be thought of as a way to classify the objects of  $X$  by dividing them into different classes. Then the equivalence relation induced by  $\mathcal{A}$  considers two objects alike precisely when they belong to the same class. On the intuitive level it is clear that equivalence relations and partitions are just two different mathematical constructs

expressing a single concept of classification. On the formal level the same thing is true: the map

$$\sim \mapsto X/\sim$$

assigning to each equivalence relation on  $X$  the quotient set under  $\sim$  is a one-to-one correspondence between the collection of all equivalence relations on  $X$ , on the one hand, and the collection of all partitions of  $X$ , on the other hand (see [Exercise 155](#)).

partition

congruence modulo  $m$ @congruence

### Saturated sets

By definition, the equivalence class  $E = [z]$  of an element  $z$  in a set  $X$  under an equivalence relation  $\sim$  has the property that if  $y \in X$  and if  $y \sim x$  for some  $x \in E$ , then necessarily  $y \in E$ , too. We generalize this property to other subsets of  $X$ .

**0.98 Definition.** Let  $\sim$  be an equivalence relation on a set  $X$ . A subset  $S$  of  $X$  is said to be **saturated by  $\sim$**  when

$$y \in X \text{ and } y \sim x \text{ for some } x \in S \implies y \in S.$$

When the equivalence relation  $\sim$  is understood, we say simply that such a set  $S$  is **saturated**.

If  $\mathcal{A}$  is a partition of a set  $X$ , then a subset  $A$  of  $X$  is said to be **saturated by  $\mathcal{A}$**  when it is saturated by the equivalence relation induced by  $\mathcal{A}$ .

Evidently, a subset  $S$  of  $X$  is saturated by an equivalence relation  $\sim$  on  $X$  precisely when  $S$  contains each equivalence class that it intersects. In other words,  $S$  is saturated if and only if it is the union of some collection of equivalence classes under  $\sim$ .

**0.99 Examples.** (1) On the set  $\mathbb{Z}$  of integers, let  $\sim$  be the relation of congruence modulo 4—see [Examples 0.90 \(2\)](#). Then each of the sets  $\{4k : k \in \mathbb{Z}\}$  and  $\{a + 4k : k \in \mathbb{Z}, a = 1 \text{ or } 2\}$  are saturated by  $\sim$ . However, neither of the sets  $\{2k : k \in \mathbb{Z}\}$  nor  $\{3k : k \in \mathbb{Z}\}$  is saturated.

(2) On the set  $\mathbb{R}$  of real numbers let  $\sim$  be the equivalence relation of [Examples 0.90 \(1\)](#), namely,

$$x \sim y \iff |x| = |y|.$$

Then the subset  $\{\sqrt{2}, -\sqrt{2}, \pi, -\pi, 3/4, 3/4\}$  of  $\mathbb{R}$  is saturated by  $\sim$ , but the subset  $[0, 1]$  is not.

(3) Let  $X$  be an arbitrary set. Then every subset of  $X$  is saturated by the equality relation on  $X$ .  $\diamond$

**0.100 Definition.** Let  $\sim$  be an equivalence relation on a set  $X$  and let  $S$  be a subset of  $X$ . Then the **saturation of  $S$  (by  $\sim$ )** is the subset

$$\{y \in X : \text{there exists some } x \in S \text{ such that } y \sim x\}$$

of the set  $X$ .

If  $\mathcal{A}$  is a partition of a set  $X$ , then the **saturation of  $S$  (by  $\mathcal{A}$ )** is the saturation of  $S$  by the equivalence relation induced by  $\mathcal{A}$ .

Thus the saturation of a subset  $S$  of  $X$  by an equivalence relation  $\sim$  is the union of all equivalence classes of elements of  $S$ ; in other terms, the saturation of  $S$  by  $\sim$  is the set

passing to the quotient  $q^{-1}(q(S))$ , where  $q: X \rightarrow X/\sim$  is the quotient map. And the saturation of  $S$  by a partition  $\mathcal{A}$  of  $X$  is the union of all members of  $\mathcal{A}$  that intersect  $S$ .  
 complex numbers  
 complex exponential function  
 standard map from  $\mathbb{R}$  to  $S_1$  @standard map from  $\mathbb{R}$  to  $S_1$

ex:saturation **0.101 Examples.** (1) On the set  $\mathbb{Z}$  of integers let  $\sim$  be the relation of congruence modulo 4. Then the saturation of the set  $\{1, 3\}$  is the set of all odd integers.

(2) On the set  $\mathbb{R}$  of real numbers let  $\sim$  be the equivalence relation of Examples 0.90 (1), namely,

$$x \sim y \iff |x| = |y|.$$

Then the saturation by  $\sim$  of an arbitrary subset  $A$  of  $\mathbb{R}$  is  $A \cup \{-a : a \in A\}$ .  $\diamond$

prop:saturation **0.102 Proposition (saturation of subset by equivalence relation).** Let  $\sim$  be an equivalence relation on a set  $X$  and let  $q: X \rightarrow X/\sim$  be the quotient map. Let  $A$  be a subset of  $X$ . Then:

- op-part:saturation-least-saturated-set (1) The saturation of  $A$  by  $\sim$  is the least saturated subset of  $X$  that contains  $A$ .
- prop-part:saturation-via-quotient-map (2) The saturation of  $A$  by  $\sim$  is the subset  $q^{-1}(q(A))$  of  $X$ .
- prop-part:saturation-criterion (3) The set  $A$  is saturated by  $\sim$  if and only if  $A = q^{-1}(q(A))$ .
- prop-part:inverse-image-is-saturated (4) The inverse image  $q^{-1}(B)$  of each subset  $B$  of  $X/\sim$  is a saturated subset of  $X$ .

**Proof.** Part (3) is an immediate consequence of part (2); and (4) of (3). The proofs of (1) and (2) are left as exercises.  $\square$

### Passing to the quotient

subsec:pass-to-quotient

From certain “nice” maps with domain a given set  $X$  we can obtain corresponding maps with domain the quotient set  $X/\sim$  of  $X$  under an equivalence relation.

ex: $\mathbb{R}$ -to- $S_1$ -well-defined **0.103 Example.** Let  $\sim$  be the relation in  $\mathbb{R}$  to itself defined by

$$t \sim s \iff t - s \in \mathbb{Z}.$$

As you may readily check,  $\sim$  is an equivalence relation on  $\mathbb{R}$ .

Let  $X = \mathbb{R}$  and

$$Y = S_1 = \{\langle x, y \rangle \in \mathbb{R}^2 : x^2 + y^2 = 1\},$$

the unit circle in the euclidean plane, and define the map

$$f: X \rightarrow Y$$

by

$$f(t) = e^{2\pi i t} = \langle \cos 2\pi t, \sin 2\pi t \rangle \quad (t \in \mathbb{R})$$

(every value  $f$  does lie in  $S_1$  because  $\cos^2 u + \sin^2 u = 1$  for all real  $u$ ). We shall refer to this map  $\mathbb{R} \rightarrow S_1$  as the **standard map**.

We would like to define a map

$$f^*: X/\sim \rightarrow Y$$

by the formula

{eq:fstar-from-f-via-elements} (\*)

$$f^*([t]) = f(t)$$

However, an equivalence class  $[t]$  is identical to  $[s]$  for each  $s$  satisfying  $t \sim s$ . Thus in order that the formula (\*) allegedly defining  $f^*$  does do so, we must check that

$$t \sim s \implies f(t) = f(s).$$

That implication does hold (as does in fact the corresponding logical equivalence), and so the formula defining  $f^*$  is legitimate.

What we have just done—justifying that a value of  $f^*$  at an equivalence class in its domain  $X/\sim$  is *independent of the choice of representative* of that equivalence class—is referred to as showing that the map  $f^*$  is **well-defined**.  $\diamond$

Observe in the preceding example that:

- the map  $f$  is constant on each equivalence class under  $\sim$ ; and
- the map  $f^*$  is related to the original  $f$  by

$$f^* \circ q = f,$$

where  $q: X \rightarrow X/\sim$  is the quotient map.

And these observations suggest a general formulation for constructing the map  $f^*: X/\sim \rightarrow Y$  from a given map  $f: X \rightarrow Y$ .

thm:pass-to-quotients-sets

**0.104 Theorem (passing to the quotient).** Let  $q: X \rightarrow X/\sim$  be the quotient map induced by an equivalence relation on a set  $X$ . Let  $f: X \rightarrow Y$  be a map from  $X$  to a set  $Y$  that is constant on each equivalence class, that is,

$$x \sim t \implies f(x) = f(t) \quad (x, t \in X).$$

Then there is a unique map

$$f^*: X/\sim \rightarrow Y$$

such that

$$f^* \circ q = f.$$

Moreover:

thm-part:fstar-surj-iff-f-is-sets

(1) The map  $f^*$  is surjective if and only if  $f$  is surjective.

inct-vals-on-distinct-equiv-classes-sets

(2) The map  $f^*$  is injective if and only if  $f$  takes distinct values at representatives of different equivalence classes under  $\sim$ , that is, if the equality  $f(x) = f(t)$  implies the equivalence  $x \sim t$ .

The proof of the theorem appears at the end of this section.

well-defined map  
map!well-defined

def:map-pass-to-quot

**0.105 Definition.** In the notation of [Theorem 0.104](#), we call  $f^*$  the **map obtained from  $f$  by passing to the quotient (under  $\sim$ )**, and we say that  $f^*$  is **induced by  $f$**  and that we **pass to the quotient** to obtain  $f^*$  from  $f$ .

In terms of elements, the equation  $f^* \circ q = f$  means that, for each  $x \in X$ ,

$$f^*([y]) = f(x) \text{ for every } y \sim x,$$

and so in particular,

$$f^*([x]) = f(x).$$

The relationship among the original map  $f$ , the map  $f^*$ , and the quotient map  $q$  is shown in the following commutative diagram.

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ q \downarrow & \nearrow f^* & \\ X/\sim & & \end{array}$$

This relationship is indicated schematically in [Figure 0.5](#), where equivalence classes under  $\sim$  are represented by vertical lines in the region representing  $X$ .

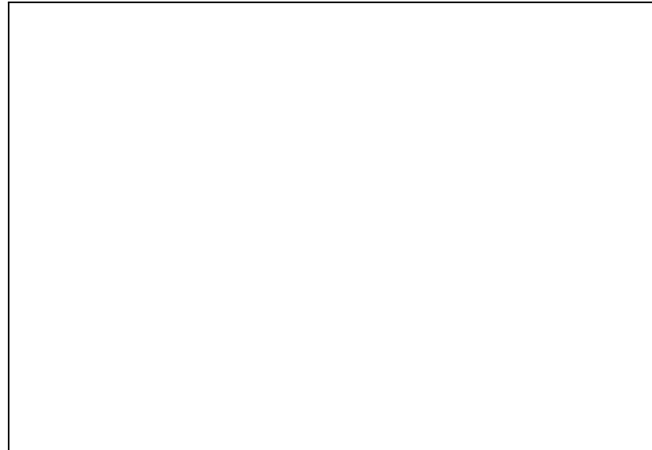


Figure 0.5: Passing to the quotient.

fig:map-pass-to-quotient

Let us see how the theorem is applied in practice for the situation of the preceding example.

ex:R-to-S1-pass-to-quotient-sets

**0.106 Example.** As in [Example 0.103](#), let  $f$  be the standard map

$$\begin{aligned} f: \mathbb{R} &\rightarrow S_1 \\ t &\mapsto e^{2\pi i t} = \langle \cos 2\pi t, \sin 2\pi t \rangle \end{aligned}$$

and let  $\sim$  be the equivalence relation on  $\mathbb{R}$  given by

$$t \sim s \iff t - s \in \mathbb{Z}$$

Then  $f$  is constant on each equivalence class, that is,  $t - s \in \mathbb{Z}$  implies  $f(t) = f(s)$  for all  $t, s \in \mathbb{R}$ . Hence we are justified in passing to the quotient so as to obtain a unique map

$$f^*: \mathbb{R}/\sim \rightarrow S_1$$



such that  $f^* \circ q = f$ .

Moreover, the map  $f^*$  is surjective because  $f$  is. And  $f^*$  is also injective because  $f$  does take distinct values at representative of different equivalence classes under  $\sim$ . Thus  $f^*$  is actually a bijection between the quotient set  $\mathbb{R}/\sim$  and the unit circle  $S_1$ .  $\diamond$

passing to the quotient  
equivalence relation

**Proof of Theorem 0.104.** For convenience, set  $Z = X/\sim$ .

**Existence.** We construct  $f^*$  as follows. If  $z \in Z$ , then  $z$  is an equivalence class under  $\sim$ , and we define  $f^*(z)$  to be  $f(x)$  for any representative of this equivalence class; the value  $f^*(z)$  is determined by  $z$  alone, independently of the choice of  $x$ , because if  $x$  and  $t$  are both representatives of  $z$ , then  $f(x) = f(t)$  by hypothesis.

From this definition of  $f^*$ , it follows that  $f^* \circ q = f$ . In fact, if  $x \in X$  is arbitrary, then  $x$  is a representative of the equivalence class  $q(x)$  and so  $f^*(q(x)) = f(x)$  by the very definition of  $f^*$ .

**Uniqueness.** Let  $f^*$  be the map just constructed. Suppose  $g: Z \rightarrow Y$  is a map such that  $g \circ q = f$ . If  $z$  is an arbitrary element of  $Z$ , then  $z = q(x)$  for some  $x \in X$ , and so

$$g(z) = g(q(x)) = f(x) = f^*(q(x)) = f^*(z).$$

Hence  $g = f^*$ .

- (1) Assume first that  $f$  is surjective. Let  $y \in Y$  be arbitrary. Then  $y = f(x)$  for some  $x \in X$ . Hence  $z = q(x) \in Z$  with  $f^*(z) = f^*(q(x)) = f(x) = y$ .

Conversely, assume that  $f^*$  is surjective. Let  $y \in Y$  be arbitrary. Then  $y = f^*(z)$  for some  $z \in Z$ . Choose a representative  $x$  of  $z$ , so that  $z = [x] = q(x)$  for some  $x \in X$ . Then  $x \in X$  with  $f(x) = f^*(q(x)) = f^*(z) = y$ .

- (2) Assume first that  $f$  takes distinct values at representatives of different equivalence classes. Suppose  $z_1, z_2 \in Z$  with  $f^*(z_1) = f^*(z_2)$ . We shall deduce that  $z_1 = z_2$ . Choose representatives  $x_1$  and  $x_2$  of  $z_1$  and  $z_2$ , respectively. Then  $z_1 = q(x_1)$ ,  $z_2 = q(x_2)$ , and so

$$f(x_1) = f^*(q(x_1)) = f^*(z_1) = f^*(z_2) = f^*(q(x_2)) = f(x_2).$$

By assumption,  $x_1$  and  $x_2$  must be representatives of the same equivalence class. Hence  $z_1 = z_2$ , as desired.

Conversely, assume that  $f^*$  is injective. Suppose that  $x_1, x_2$  are representatives of different equivalence classes. This means that  $[x_1] \neq [x_2]$ . By assumption,  $f^*([x_1]) \neq f^*([x_2])$ . But  $f^*([x_1]) = f(x_1)$  and  $f^*([x_2]) = f(x_2)$ . Hence  $f(x_1) \neq f(x_2)$ , as desired.  $\square$

Notions of “passing to the quotient” for a relation in a set and for a binary operation on a set are defined similarly. See [Exercises 159](#) and [160](#) for details.

## EXERCISES FOR SECTION 0.9

- 142.** In each case below, decide whether the specified relation  $\sim$  is an equivalence relation on the given set  $X$ .

- (a)  $X = \mathbb{R}$ ; for  $x, y \in X$ ,  $x \sim y \iff y - x \in \mathbb{N}$ .
- (b)  $X = \mathbb{Z}^*$ ; for  $x, y \in X$ ,  $x \sim y \iff y = kx$  for some  $k \in X$ .
- (c)  $X = \mathcal{P}(\mathbb{R})$ ; for  $A, B \in X$ ,  $A \sim B \iff A \cap B \neq \emptyset$ .

prob:R-mod-Z-or-Q-equiv-rel **143.** Show that the specified relation  $\sim$  is an equivalence relation on  $\mathbb{R}$ .

prob-part:R-mod-Z-equiv-rel (a) The relation  $\sim$  in  $\mathbb{R}$  given by  $x \sim y \iff y - x \in \mathbb{Z}$ .

The quotient set  $\mathbb{R}/\sim$  is called  $\mathbb{R} \bmod \mathbb{Z}$  and is denoted by  $\mathbb{R}/\mathbb{Z}$ .

prob-part:R-mod-Q-equiv-rel (b) The relation  $\sim$  in  $\mathbb{R}$  given by  $x \sim y \iff y - x \in \mathbb{Q}$ .

The quotient set  $\mathbb{R}/\sim$  is called  $\mathbb{R} \bmod \mathbb{Q}$  and is denoted by  $\mathbb{R}/\mathbb{Q}$ .

prob:examples-equiv-rel **144.** In each case below, verify that the specified relation  $\sim$  is an equivalence relation on the given set  $X$  and determine the equivalence class of each  $x \in X$ .

prob-part:equiv-rel-on-NxN-for-Z (a)  $X = \mathbb{N} \times \mathbb{N}$ ; for  $\langle m, n \rangle \in X$  and  $\langle i, j \rangle \in X$ ,

$$\langle m, n \rangle \sim \langle i, j \rangle \iff m + j = n + i.$$

prob-part:equiv-rel-on-ZxZstar-for-Q (b)  $X = \mathbb{Z} \times \mathbb{Z}^*$ ; for  $\langle m, n \rangle \in X$  and  $\langle i, j \rangle \in X$ ,

$$\langle m, n \rangle \sim \langle i, j \rangle \iff m j = n i.$$

(c)  $X = [0, 1] \times [0, 1]$ ; for  $\langle t, s \rangle \in X$  and  $\langle u, v \rangle \in X$ ,

$$\langle t, s \rangle \sim \langle u, v \rangle \iff \langle t, s \rangle = \langle u, v \rangle \text{ or } \langle t, s \rangle = \langle u, 1 - v \rangle.$$

(d)  $X$  is the set of all differentiable functions  $f: \mathbb{R} \rightarrow \mathbb{R}$ ; for  $f \in X$  and  $g \in X$ ,  $f \sim g$  if and only if  $f'(t) = g'(t)$  for all  $t \in \mathbb{R}$ , where the prime denotes the derivative.

**145.** Let  $\alpha$  be the relation in  $\mathbb{R}$  to itself given by  $x \alpha y \iff y = x + 1$ . What equivalence relation on  $\mathbb{R}$  is the reflexive, symmetric, transitive closure of  $\alpha$ ?

**146.** Describe in the form  $z \sim w \iff \dots$  the reflexive, symmetric, transitive closure of the relation  $\alpha$  in  $\mathbb{R} \times \{0, 1\}$  to itself given by  $(x, 0) \alpha (x, 1)$  for all  $x \in \mathbb{R}^*$ .

**147.** Find the least equivalence relation  $\sim$  on  $\mathbb{Z} \times \{0, 1\}$  for which  $\langle x, 0 \rangle \sim \langle -x, 0 \rangle$  for all  $x > 0$  and  $\langle x, 0 \rangle \sim \langle -x, 1 \rangle$  for all  $x < 0$ .

prob:rel-closure-explicit **148.** Let  $\alpha$  be a relation in a set  $X$  to itself and let  $\sim$  be the reflexive, symmetric, transitive closure of  $\alpha$  (Definition 0.91). Show that, as a subset of  $X \times X$ ,

$$\sim = \Delta X \cup \bigcup_{n=1}^{\infty} \beta^n,$$

where

$$\beta = \alpha \cup \alpha^{-1},$$

$\alpha^{-1}$  is the reverse of  $\alpha$  (Definition 0.8), and the  $\beta^n$  are the  $n$ th powers of  $\beta$  as defined in Exercise 73.

**149.** Let  $A$  be a nonempty proper subset of a set  $X$ . For  $x, y \in X$ , let

$$x \sim y \iff x = y \text{ or } (x \in A \text{ and } y \in A).$$

In other words,  $\sim$  classifies two elements of  $X$  as alike exactly when they are equal or they both belong to  $A$ .

(a) Verify that  $\sim$  is an equivalence relation on  $X$ .

(b) Show that the quotient set  $X/\sim$  is  $\{A\} \cup (X \setminus A)$ .

(c) Let  $q: X \rightarrow X/\sim$  be the quotient map. Show that its restriction

$$\begin{aligned} f: X \setminus A &\rightarrow (X/\sim) \setminus \{A\} \\ x &\mapsto q(x) \end{aligned}$$

is a bijection.

equivalence kernel  
saturated set!map@and map  
equivalence kernel  
punctured plane

prob:equiv-ker-of-map **150.** (a) Determine explicitly the equivalence kernel [Examples 0.90 (7)] of the map  $f: \mathbb{R} \rightarrow \mathbb{R}$  defined by the formula  $f(x) = \cos 2\pi x$ .

(b) Determine explicitly the equivalence kernel of the map  $f: \mathbb{R} \rightarrow S_1$  defined by the formula  $f(t) = \langle \cos 2\pi t, \sin 2\pi t \rangle$  that was considered in Examples 0.103 and 0.106.

(c) Given a map  $f: X \rightarrow Y$ , show that a subset  $S$  of  $X$  is saturated by the equivalence kernel of  $f$  if and only if it is saturated by  $f$  in the sense of Exercise 38.

(d) What is the partition induced by the equivalence kernel of a map  $f: X \rightarrow Y$ ?

(e) If  $\sim$  is an equivalence relation on a set  $X$ , what is the equivalence kernel of the quotient map  $q: X \rightarrow X/\sim$ ?

**151.** In each case below, either construct a partition  $\mathcal{A}$  of the plane  $\mathbb{R}^2$  having the stated properties or else tell why no such partition exists.

(a) The partition  $\mathcal{A}$  is denumerable and each  $A \in \mathcal{A}$  is denumerable.

(b) The partition  $\mathcal{A}$  is denumerable and each  $A \in \mathcal{A}$  is uncountable.

(c) The partition  $\mathcal{A}$  is uncountable and each  $A \in \mathcal{A}$  is denumerable.

prob:partition-by-punctured-lines **152.** For each line  $L$  in the plane  $\mathbb{R}^2$  that passes through the origin  $\mathbf{0} = \langle 0, 0 \rangle$ , call the set  $L \setminus \{\mathbf{0}\}$  a “deleted line.” Verify that the collection of all deleted lines is a partition of the “punctured plane”  $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ . Then describe algebraically the equivalence relation  $\sim$  on the punctured plane that is induced by this partition.

**153.** Determine the saturation (Definition 0.100) of the specified subset of the given set by the indicated equivalence relation or partition.

(a) The saturation of the 2-disk  $D_2$  in the plane  $\mathbb{R}^2$  by the equivalence relation  $\sim$  given by  $\langle x, y \rangle \sim \langle s, t \rangle$  if and only if  $y = t$ .

(b) The saturation of the circle  $S_1$  in the “punctured plane”  $X = \mathbb{R}^2 \setminus \{\mathbf{0}\}$  by the partition described in Exercise 152.

(c) The saturation of the closed unit interval  $[0, 1]$  in the real line  $\mathbb{R}$  by the partition  $\mathbb{R}/\mathbb{Z}$  of  $\mathbb{R}$ —see Exercise 143 (a).

(d) The saturation of the closed unit interval  $[0, 1]$  in the real line  $\mathbb{R}$  by the partition  $\mathbb{R}/\mathbb{Q}$  of  $\mathbb{R}$ —see Exercise 143 (b).

**154.** Let  $\sim$  and  $\simeq$  be two equivalence relations on the same set  $X$  such that  $x \sim y$  implies  $x \simeq y$  for all  $x, y \in X$ . Denote by  $\mathcal{A}$  and  $\mathcal{B}$  the quotient sets  $X/\sim$  and  $X/\simeq$ , respectively. How is  $\mathcal{A}$  related to  $\mathcal{B}$ ?

prob:1-1-corr-equiv-rels-and-partitions **155.** Let  $X$  be a set. For each partition  $\mathcal{A}$  of  $X$ , denote by  $X/\mathcal{A}$  the equivalence relation on  $X$  induced by  $\mathcal{A}$ , so that for  $x \in X$  and  $y \in X$ , we have

$$x (X/\mathcal{A}) y \iff x \in A \text{ and } y \in A \text{ for some } A \in \mathcal{A}.$$

(a) Let  $\mathcal{A}$  be any partition of  $X$ . Then  $X/\mathcal{A}$  is an equivalence relation  $\sim$  on  $X$  whose quotient set  $X/\sim = X/(X/\mathcal{A})$  is therefore a partition of  $X$ . Prove that  $X/\sim = \mathcal{A}$ , in other words, that

$$X/(X/\mathcal{A}) = \mathcal{A}.$$

**(b)** Let  $\sim$  be any equivalence relation on  $X$ . Then the quotient set  $X/\sim$  is a partition  $\mathcal{A}$  of  $X$  which therefore induces the equivalence relation  $X/\mathcal{A} = X/(X/\sim)$  on  $X$ . Prove that  $X/\mathcal{A} = \sim$ , in other words, that

$$X/(X/\sim) = \sim.$$

**(c)** Use (a) and (b) to show that the rule

$$\sim \mapsto X/\sim$$

defines a one-to-one correspondence between the collection  $\eta$  of all equivalence relations on  $X$  and the collection  $\pi$  of all partitions of  $X$ .

**156.** Comment on the following, possibly apocryphal, claim by a first-year college student: “I’m the only one in the dorm with a roommate.”<sup>5</sup>

**157.** Let  $f: X \rightarrow Y$  be a map that is constant on equivalence classes under an equivalence relation  $\sim$ . Must  $\sim$  be the equivalence kernel ([Exercise 150](#)) of  $f$ ? If so, prove it; if not, construct a counterexample.

**158.** Let  $\sim$  be the equivalence relation on the closed unit interval  $[0, 1]$  given by

$$s \sim t \iff s = t \text{ or } \{s, t\} = \{0, 1\}.$$

Construct a map  $[0, 1]/\sim \rightarrow S_1$  by starting with a suitable map  $[0, 1] \rightarrow S_1$  and then passing to the quotient.

**159.** Let  $\sim$  be an equivalence relation on a set  $X$  and let  $q: X \rightarrow X/\sim$  be the quotient map. A relation  $\alpha$  in  $X$  is said to be **compatible with**  $\sim$  when elements equivalent to  $\alpha$ -related elements are themselves  $\alpha$ -related, that is,

$$s \sim x \text{ and } t \sim y \text{ and } x \alpha y \implies s \alpha t$$

for all  $x, y, s, t \in X$ . (Examples of this situation appear in [Exercises 161](#) and [162](#).)

**(a)** Prove that there exists a unique relation  $\tilde{\alpha}$  on the quotient set  $X/\sim$  such that

$$x \alpha y \iff \llbracket x \rrbracket \tilde{\alpha} \llbracket y \rrbracket$$

for all  $x, y \in X$ . We say that the relation  $\tilde{\alpha}$  is **obtained from  $\alpha$  by passing to the quotient (under  $\sim$ )**.

**(b)** Show that if  $\alpha$  is actually a partial ordering (respectively, a total ordering) of  $X$ , then  $\tilde{\alpha}$  is a partial ordering (respectively, a total ordering) of  $X/\sim$ .

**160.** Let  $\sim$  be an equivalence relation on a set  $X$  and let  $p: X \rightarrow X/\sim$  be the quotient map. A binary operation  $*$  on  $X$  is said to be **compatible with**  $\sim$  when

$$s \sim x \text{ and } t \sim y \implies s * t = x * t$$

for all  $x, y, s, t \in X$ . (Examples of this situation appear in [Exercises 161](#) and [162](#).)

**(a)** Prove that there exists a unique binary operation  $\tilde{*}$  on the quotient set  $X/\sim$  such that

$$\llbracket x * y \rrbracket = \llbracket x \rrbracket \tilde{*} \llbracket y \rrbracket$$

for all  $x, y \in X$ .

We say that the operation  $\tilde{*}$  is **obtained from  $*$  by passing to the quotient (under  $\sim$ )**.

<sup>5</sup>Suggested by an item in “The Back Page”, Notices Amer. Math. Soc. **64** (2017), 808.

- (b) Show that if  $*$  is an associative (respectively, a commutative) operation on  $X$ , then  $\widetilde{*}$  is an associative (respectively, a commutative) operation on the quotient set  $X/\sim$ .
- (c) Suppose now there is an identity element  $e$  in  $X$  for the operation  $*$ . Show that  $\llbracket e \rrbracket$  is an identity element in  $X/\sim$  for the operation  $\widetilde{*}$ .  
Show, further, that if an  $x \in X$  has an inverse  $x'$  in  $X$  for the operation  $*$  that is,  $x * x' = e = x' * x$ , then  $\llbracket x' \rrbracket$  is an inverse of  $\llbracket x \rrbracket$  in  $X/\sim$  for the operation  $\widetilde{*}$ , that is,  $\llbracket x \rrbracket \widetilde{*} \llbracket x' \rrbracket = \llbracket e \rrbracket = \llbracket x' \rrbracket \widetilde{*} \llbracket x \rrbracket$ .

prob:Z-from-N **161.** *Note:* The purpose of this exercise is to construct the integers from the natural numbers. The idea behind the construction is that every integer, positive or negative or zero, should be the difference  $m - n$  of a pair  $\langle m, n \rangle$  of natural numbers  $m$  and  $n$ . But a given integer would be the difference of many such pairs. Indeed, for such pairs  $\langle m, n \rangle$  and  $\langle i, j \rangle$ , we would have  $m - n = i - j$  exactly when  $m + j = n + i$ . Hence we want to form a quotient set of  $\mathbb{N} \times \mathbb{N}$ , and forming this involves only natural numbers. Accordingly, temporarily imagine you are completely unfamiliar with the set  $\mathbb{Z}$  of integers.

Form the quotient set

$$Z' = (\mathbb{N} \times \mathbb{N})/\sim$$

of  $\mathbb{N} \times \mathbb{N}$  under the equivalence relation  $\sim$  defined in [Exercise 144 \(a\)](#), namely,

$$\langle m, n \rangle \sim \langle i, j \rangle \iff m + j = n + i.$$

Let

$$Z' = (\mathbb{N} \times \mathbb{N})/\sim$$

be the corresponding quotient set.

prob-part:Z-from-N-def-leq

- (a) Verify that the relation  $\leq$  in  $\mathbb{N} \times \mathbb{N}$  defined by

$$\langle m, n \rangle \leq \langle k, l \rangle \iff m + l \leq k + n \quad (m, n, k, l \in \mathbb{N})$$

is compatible with  $\sim$  in the sense of [Exercise 159](#).

Show that the relation  $\widetilde{\leq}$  obtained by passing to the quotient is a total ordering of  $Z'$ .

prob-part:Z-from-N-def-add

- (b) Verify that the binary operation

$$\begin{aligned} + : (\mathbb{N} \times \mathbb{N}) \times (\mathbb{N} \times \mathbb{N}) &\rightarrow \mathbb{N} \times \mathbb{N} \\ \langle \langle m, n \rangle, \langle k, l \rangle \rangle &\mapsto \langle m + k, n + l \rangle \end{aligned}$$

on  $\mathbb{N} \times \mathbb{N}$  is compatible with  $\sim$  in the sense of [Exercise 160](#).

Show that the binary operation  $\widetilde{+}$  on  $Z'$  obtained by passing to the quotient is associative and commutative and has  $0' = \llbracket (0, 0) \rrbracket$  as an identity element.

Show further that each element of  $Z'$  has an inverse with respect to the operation  $\widetilde{+}$ . (*Caution:* The inverse of an element  $\llbracket \langle m, n \rangle \rrbracket$  of  $Z'$  is *not*  $\llbracket \langle -m, -n \rangle \rrbracket$ ; we are working here only with natural numbers and not also with negative integers!)

prob-part:Z-from-N-def-mult

- (c) Verify that the binary operation

$$\begin{aligned} \cdot : (\mathbb{N} \times \mathbb{N}) \times (\mathbb{N} \times \mathbb{N}) &\rightarrow \mathbb{N} \times \mathbb{N} \\ \langle \langle m, n \rangle, \langle k, l \rangle \rangle &\mapsto \langle ml + nk, nl \rangle \end{aligned}$$

on  $\mathbb{N} \times \mathbb{N}$  is compatible with  $\sim$ .

Show that the binary operation  $\widetilde{\cdot}$  on  $Z'$  obtained by passing to the quotient is associative and commutative and has  $1' = \llbracket (1, 0) \rrbracket$  as an identity element for  $\widetilde{\cdot}$ .

Show also that each element of  $Z'$  other than the additive identity  $0'$  has an inverse with respect to the operation  $\sim$ .

Show, further, that  $\sim$  distributes over  $\bar{+}$ .

Show, finally, that the cancellation law holds for  $\sim$ : if  $x', y', z' \in Z'$  with  $x' \sim y' = x' \sim z'$  and if  $x' \neq 0'$ , then  $y' = z'$ .

- (d) Show that the operations  $\bar{+}$  and  $\sim$  on  $Z'$  are compatible with the total ordering  $\preceq$  in the sense that for all  $x', y', z' \in Z'$ ,

$$\begin{aligned} x' \preceq y' &\implies x' \bar{+} z' \preceq y' \bar{+} z', \\ x' \preceq y' &\implies x' \cdot z' \preceq y' \cdot z' \text{ provided that } 0' \preceq z'. \end{aligned}$$

- (e) Show that the map

$$\begin{aligned} \varphi: \mathbb{N} &\rightarrow Z' \\ k &\mapsto \llbracket \langle k, 0 \rangle \rrbracket \end{aligned}$$

is an injection.

Show, further, that the injection  $\varphi$  “preserves addition and multiplication” in the sense that

$$\begin{aligned} \varphi(k + l) &= \varphi(k) \bar{+} \varphi(l), \\ \varphi(k \cdot l) &= \varphi(k) \cdot \varphi(l) \end{aligned}$$

for all  $j, k \in \mathbb{N}$ .

Show, finally, that  $\varphi$  is order-preserving.

*Note:* Thus we can use  $\varphi$  to identify each natural number  $k$  with the corresponding element  $\varphi(k)$  of  $Z'$ , and natural numbers add and multiply and compare the same way the corresponding elements of  $Z'$  do.

For the following part, you may once again assume that the number system  $\mathbb{Z}$  of all integers is known. Let  $q: \mathbb{N} \times \mathbb{N} \rightarrow Z'$  be the quotient map.

- (f) Define  $\psi: \mathbb{Z} \rightarrow \mathbb{N} \times \mathbb{N}$  by

$$\psi(k) = \begin{cases} \langle k, 0 \rangle & \text{if } k \geq 0, \\ \langle 0, k \rangle & \text{if } k < 0. \end{cases}$$

Prove that the composite  $\psi' = q \circ \psi: \mathbb{Z} \rightarrow Z'$  is a bijection that preserves addition and multiplication and is order-preserving. Prove, further, that composite  $\psi^{-1} \circ \varphi: \mathbb{N} \rightarrow Z'$  is just the inclusion map  $j: \mathbb{N} \rightarrow \mathbb{Z}$ .

*Note:* Thus if we knew about  $\mathbb{N}$  but not  $\mathbb{Z}$ , then the set  $Z'$ , when equipped with its operations of  $\bar{+}$  and  $\sim$  and its total ordering  $\preceq$ , for all intents and purposes is the system of integers, with the subset  $\varphi(\mathbb{N})$  of  $Z'$  being essentially the same as  $\mathbb{N}$ . It is in this sense that we have constructed  $\mathbb{Z}$  from  $\mathbb{N}$ .

prob:Q-from-Z **162.** *Note:* The purpose of this exercise is to construct the rational numbers from the integers. The idea behind the construction is that every rational number should be the quotient  $m/n$  of a pair  $\langle m, n \rangle$  of integers  $m$  and  $n$  with  $n \neq 0$ , that is, for an element  $\langle m, n \rangle$  of  $\mathbb{Z} \times \mathbb{Z}^*$ . But a given rational number would be the quotient of many such pairs. Indeed, for such pairs  $\langle m, n \rangle$  and  $\langle i, j \rangle$ , we would have  $m/n = i/j$  exactly when  $mj = ni$ . Hence we shall want to form a quotient set of  $\mathbb{Z} \times \mathbb{Z}^*$ , and forming this involves only integers.

Accordingly, temporarily imagine you are completely unfamiliar with the set  $\mathbb{Q}$  of rational numbers but do know about the set  $\mathbb{Z}$  of integers.

Form the quotient set

$$Q' = (\mathbb{Z} \times \mathbb{Z}^*) / \sim,$$

where  $\sim$  is the equivalence relation on  $\mathbb{Z} \times \mathbb{Z}^*$  defined in [Exercise 144 \(b\)](#), namely,

$$\langle m, n \rangle \sim \langle i, j \rangle \iff mj = ni.$$

Let  $Q' = (\mathbb{Z} \times \mathbb{Z}^*) / \sim$  be the quotient set. Carry out the analogs of the steps Exercises (a)–(e) from [Exercise 161](#) for  $Q'$ . Show, further, that  $Q'$  is actually an ordered field for the addition, multiplication, and total ordering so obtained. Finally, assuming once again that the ordered field  $\mathbb{Q}$  is known, carry out the analog of step (f) from [Exercise 161](#) for  $Q$  and  $Q'$ .

## 0.10 Ordinals, Cardinality, and Zorn's Lemma

sec:ordcard

subsec:wo-sets

### Well-ordered sets

In a totally ordered set, any two-element subset  $\{x, y\}$  has a least element; that is an immediate consequence of the comparability property (I4). By induction, each nonempty *finite* subset of a totally ordered set has a least element. According to the Well-ordering Principle (0.3), each nonempty subset—finite or infinite—of the set  $\mathbb{N}$  of natural numbers has a least element. In general, however, a nonempty subset of a totally ordered set need not have a least element: for example, with its usual ordering the set  $\mathbb{Z}$  of all integers has no least element.

We generalize those examples.

**0.107 Definition.** A relation  $\leq$  in a set  $X$  is said to **well-order**  $X$  and is called a **well-ordering of  $X$**  if it partially orders  $X$  and if, in addition, it has the property:

(WO)<sub>p</sub> *Every nonempty subset of  $X$  has a least element.*

When a set is well-ordered by a given relation, we refer to it simply as a **well-ordered set**.

A *well-ordered set* is *totally ordered* because each of its two-element subsets has a least element. (Which is why some authors include being totally ordered as part of the definition of being well-ordered.)

As in  $\mathbb{N}$  (with its usual ordering), so **in any well-ordered set each element has an immediate successor**—see [Exercise 163 \(a\)](#).

Aside from  $\mathbb{N}$ , the main example of a well-ordered set for our study of topology will be set  $[0, \Omega]$  of “ordinals”, described in the following subsection “The first uncountable ordinal” ([page 110](#)).

Later we shall use the following property of well-ordered sets.

prop:bded-above-subset-wo-has-sup

**0.108 Proposition (order-completeness of well-ordered set).** *Let  $(X, \leq)$  be a well-ordered set. Then  $(X, \leq)$  is order-complete, that is, each nonempty subset of  $X$  that is bounded above in  $X$  has a supremum.*

**Proof.** Let  $A \subset X$  and assume that  $A$  is bounded above in  $X$ . Then the subset  $\{b \in X : b \text{ is an upper bound of } A\}$  of  $X$  is nonempty and hence has a least element.  $\square$

In view of the Well-ordering Principle (0.3), the set  $\mathbb{N}$  of natural numbers, with its usual ordering, is order-complete.

ordinal  
 subsec:1st-uncountable-ordinal  
 ordinal!first uncountable  
 ordinal!finite  
 ordinal!first infinite  
 ordinal

### The first uncountable ordinal

The principal aim of this section is to establish existence of a certain uncountable well-ordered set that will be used in later chapters for certain specialized counterexamples. The result we seek is the following.

thm:OmegaPlus-exists **0.109 Theorem (the first uncountable ordinal).** *There exists a totally-ordered set  $(\Omega^+, \leq)$  with the following properties:*

property:OmegaPlus1 (Ω1)  $\Omega^+$  is well-ordered by  $\leq$ . In particular,  $\Omega^+$  has a least element 0.

(Ω2)  $\Omega^+$  has a greatest element  $\Omega$ .

(Ω3) The set  $\Omega^+$  is uncountable.

property:OmegaPlus4 (Ω4) For each  $\alpha \in \Omega^+$  with  $\alpha < \Omega$ , the interval  $]\leftarrow, \alpha]$  is countable.

The proof of this theorem appears at the end of this subsection. That proof will depend on certain results—Zorn’s Lemma (0.115) and the Well-ordering Theorem (0.117)—that are established only in subsequent subsections, at the end of this section.

We denoted the least and greatest elements of  $\Omega^+$  by 0 and  $\Omega$ , respectively. Thus

$$\Omega^+ = [0, \Omega],$$

and then property (Ω4) may be restated as: for each  $\alpha \in [0, \Omega[$ , the interval  $[0, \alpha]$  is countable.

There is essentially just one totally ordered set having the properties ((Ω1))–((Ω4)) just listed. More precisely, it can be proved that two such sets are order-isomorphic (the proof is beyond the scope of this book).

The elements of  $[0, \Omega[$  are known as the **countable ordinals**. The greatest element of  $[0, \Omega]$  is the **first uncountable ordinal**.

Those elements  $\alpha$  of  $[0, \Omega]$  for which the initial interval  $[0, \alpha]$  is finite are called **finite ordinals**. Because the entire set  $[0, \Omega]$  is uncountable, hence infinite, there is an element of  $[0, \Omega]$  that is not a finite ordinal; the first such element is called the **first infinite ordinal** and is denoted by  $\omega$ . Thus the finite ordinals are those elements  $\alpha \in \Omega^+$  with  $\alpha < \omega$ . Just as we had  $\Omega^+ = [0, \Omega]$ , so we write

$$\omega^+ = [0, \omega]$$

for the set of all finite ordinals together with the first infinite ordinal  $\omega$ .

rem:omega-within-Omega **0.110 Remark.** It is no accident that we denoted the first element of the set  $\Omega^+$  by 0. In an axiomatic development of set theory, one defines an **ordinal** (or **ordinal number**) to be a set  $\alpha$  with the properties:

(Ord1) *The set  $\alpha$  is well-ordered by the weak ordering  $\leq$  associated with the relation  $\in$ , that is,  $x \leq y$  if and only if  $x \in y$  or  $x = y$ .*

(Ord2) *If  $x \in \alpha$ , then  $x \subset \alpha$ .*

Then each ordinal is the set consisting of all ordinals strictly less than itself.



For any set  $S$ , let  $S^+$  denote the set  $S \cup \{S\}$ . Then one defines

$$\begin{aligned} 0 &= \emptyset, \\ 1 &= 0^+ = 0 \cup \{0\} = \emptyset \cup \{\emptyset\} = \{\emptyset\} = \{0\}, \\ 2 &= 1^+ = 1 \cup \{1\} = \{\emptyset\} \cup \{1\} = \{0, 1\}, \\ 3 &= 2^+ = \{0, 1, 2\}, \end{aligned}$$

and so forth. According to these definitions,  $0 \in 1 \in 2 \in 3 \in \dots$ . The ordinals  $0, 1, 2, \dots$  so obtained are the finite ordinals, introduced just before this remark; each is in thus a finite set. Further, the first infinite ordinal  $\Omega$ , also introduced just before this remark, is in fact the set  $\{0, 1, 2, \dots\}$  consisting of all the finite ordinals, well-ordered by the weak ordering associated with  $\in$ . (An axiom of set theory justifies forming this infinite set.)

Thus when they are considered strictly as sets, there is no distinction between natural numbers, on the one hand, and finite ordinals, on the other hand; hence  $\mathbb{N} = \Omega$ , too. Moreover, in this axiomatic approach, the definition of addition of natural numbers includes the special case

$$n + 1 = n^+,$$

so that  $1 = 0 + 1, 2 = 1 + 1, 3 = 2 + 1$ , and so forth.

Already 0 denotes the least element of  $\Omega^+$ . In view of [Remark 0.110](#), we denote the immediate successor of each  $\alpha < \Omega$  by  $\alpha + 1$ . Then still the immediate successor  $0 + 1$  of 0 is denoted by 1, the immediate successor  $1 + 1$  of 1 by 2, and so forth.

Further, denote by  $\alpha + 2$  the immediate successor of the immediate successor  $\alpha + 1$  of  $\alpha$ , and so forth. Then the ordinals begin, in order, with the finite ordinals

$$0, 1, 2, \dots, n, n + 1, n + 2, n + 3, \dots$$

def:limit-ordinal **0.111 Definition.** An ordinal in  $\Omega^+$  other than 0 is said to be a **limit ordinal** if it is *not* the immediate successor of another ordinal.

Thus  $\Omega$  is the least limit ordinal, and the well-ordered set  $\Omega^+$  begins, in order:

$$0, 1, 2, \dots, n, n + 1, n + 2, \dots, \Omega, \Omega + 1, \Omega + 2, \dots$$

The least limit ordinal after all these is denoted, of course, by  $\Omega + \Omega$  or by  $\Omega \cdot 2$ —and *not* by  $2 \cdot \Omega$ —(these notations are consistent with the definitions of ordinal sum and ordinal product given in [Exercises 168](#) and [169](#)). Then with the obvious meanings for such notations as  $\Omega \cdot 3$  and  $\Omega^2$ , the set  $\Omega^+$  begins, in order:

$$\begin{aligned} &0, 1, 2, \dots, n, n + 1, \dots, \Omega, \\ &\Omega + 1, \Omega + 2, \dots, \Omega \cdot 2, \\ &\Omega \cdot 2 + 1, \Omega \cdot 2 + 2, \dots, \Omega \cdot 2 + \Omega, \\ &\Omega \cdot 2 + \Omega + 1, \Omega \cdot 2 + \Omega + 2, \dots, \Omega \cdot 3, \dots, \\ &\Omega \cdot 4, \dots, \Omega \cdot 5, \dots, \Omega^2, \\ &\Omega^2 + 1, \Omega^2 + 2, \dots, \Omega^2 + \Omega, \\ &\Omega^2 + \Omega + 1, \Omega^2 + \Omega + 2, \dots, \Omega^2 + \Omega \cdot 2, \\ &\Omega^2 + \Omega \cdot 2 + 1, \Omega^2 + \Omega \cdot 2 + 2, \dots, \Omega^2 + \Omega \cdot 3, \dots, \\ &\Omega^2 \cdot 2, \dots, \Omega^2 \cdot 3, \dots, \Omega^3, \dots, \\ &\Omega^4, \dots, \Omega^5, \dots \end{aligned}$$

ordinal

We shall not need these explicit names for the countable ordinals, but the list should give an idea of the richness of the supply of these—all before we reach the first uncountable ordinal  $\Omega$ .

The following is a consequence of [Proposition 0.108](#) and the fact that  $[0, \Omega]$  has  $\Omega$  as a greatest element.

prop:sups-exist-in-OmegaPlus

**0.112 Proposition.** *Each nonempty subset of  $\Omega^+$  has a supremum.*

[Proposition 0.112](#) has an analog in the half-open interval  $[0, \Omega[$ —but just for *countable* subsets.

pp:countable-sups-exist-in-OmegaPlus

**0.113 Proposition.** *Each countable subset of  $[0, \Omega[$  has a supremum in that open interval.*

**Proof.** In view of [Proposition 0.108](#), it suffices to show that each countable subset of  $[0, \Omega[$  is bounded above in that open interval.

Let  $A$  be a countable subset of  $[0, \Omega[$ . From property [\(Ω4\)](#), for each  $\alpha \in A$ , the interval  $[0, \alpha[$  is countable. Hence the union  $B = \bigcup_{\alpha \in A} [0, \alpha[$  is a countable subset of  $[0, \Omega[$ . But  $[0, \Omega[$  is uncountable. Hence there exists some  $\beta \in [0, \Omega[$  with  $\beta \notin B$ .

We show that  $\beta$  is in fact an upper bound of  $A$  in  $[0, \Omega[$ . Let  $\alpha \in A$ . If  $\beta < \alpha$ , then  $\beta \in [0, \alpha[ \subset B$ , which is impossible because  $\beta \notin B$ .  $\square$

Proving existence of the [the first uncountable ordinal](#) ([Theorem 0.109](#)) depends on the existence of *some* uncountable well-ordered set. That such a set exists will be an immediate consequence of the Well-ordering Theorem ([0.117](#)). Assuming that, we can now prove the theorem about  $\Omega^+$ .

pf:OmegaPlus-exists

**Proof of Theorem 0.109.** Let  $Y$  be some uncountable well-ordered set. Well-order the two-element set  $\{0, 1\}$  by its usual ordering:  $0 \leq 0$ ,  $0 \leq 1$ , and  $1 \leq 1$ . Give the product set  $\{0, 1\} \times Y$  its lexicographic ordering ([Example 0.70](#)). As you are asked to prove in [Exercise 166](#), this lexicographic ordering is in fact a well-ordering.

For each  $y \in Y$ , the interval  $] \leftarrow, (1, y)[$  in  $\{0, 1\} \times Y$  is uncountable. In fact,

$$\begin{aligned} ] \leftarrow, (1, y)[ &= \{ \langle 0, u \rangle : u \in Y \} \cup \{ \langle 1, v \rangle : v \in Y, v < z \} \\ &= (\{0\} \times Y) \cup (\{1\} \times ] \leftarrow, z[), \end{aligned}$$

and its subset  $\{0\} \times Y$  is uncountable (because  $Y$  is). Thus there exist elements  $\beta$  of  $\{0, 1\} \times Y$  for which the interval  $] \leftarrow, \beta[$  is uncountable.

Let  $\Omega$  be the *least* such element  $\beta$ . Define

$$\Omega^+ = ] \leftarrow, \Omega] = ] \leftarrow, \Omega[ \cup \{\Omega\}.$$

Give  $\Omega^+$  its well-ordering induced by that of  $\{0, 1\} \times Y$ . Properties [\(\(Ω1\)\)](#)–[\(\(Ω4\)\)](#) are now immediate.  $\square$

There are ordinals beyond the members of the set  $\Omega^+$ . And as with  $\Omega$ , so the class of all ordinals is itself well-ordered by the weak ordering associated with the relation *in*, that is by the relation  $\leq$  defined by  $\alpha \leq \beta$  if and only if  $\alpha \in \beta$  or  $\alpha = \beta$ .

## Maximal and minimal elements

First we generalize the notion of a greatest element (*maximum*) of an ordered set to that of a maximal element.

def:maximal element

**0.114 Definition.** Let  $X$  be a partially ordered set. An element  $m$  of  $X$  is said to be **maximal (in  $X$ )** if no element of  $X$  is strictly greater than  $m$ . Similarly, an element  $m$  of  $X$  is said to be **minimal (in  $X$ )** if no element of  $X$  is strictly less than  $m$ .

maximal principle  
Tukey, John W.  
Hausdorff, Felix  
Kuratowski, Kazimierz  
Kneser, Hellmuth

Evidently, if a partially ordered set  $X$  has a greatest element  $m$ , then  $m$  is a maximal element.

A totally ordered set need not have a greatest element—for example,  $\mathbb{N}$  with its usual ordering—and hence need not have a maximal element. However, a maximal element need not be a greatest element. For example, partially order the collection  $\{\emptyset, \{0\}, \{1\}\}$  by subset inclusion; then both  $\{0\}$  and  $\{1\}$  are maximal elements, but neither is a greatest element. That same example shows that when a maximal element exists, it need not be unique.

In case the partially ordered set  $X$  is actually totally ordered, then  $m \in X$  is maximal if and only if  $m$  is a greatest element—hence *the* maximum element—of  $X$ .

The following subsection gives conditions under which a partially ordered set is guaranteed to have a maximal element.

subsec:zorn

### Zorn's Lemma

For subsequent application—including the proof that every set has some well-ordering—the key result is the following “maximal principle.” In the form stated, it is actually due to J. W. Tukey

ZornLemma

**0.115 Zorn's Lemma.** Let  $X$  be a partially ordered set in which each chain is bounded above. Then  $X$  has a maximal element.

Notice that  $\emptyset$  is trivially a chain in  $X$ ; to say that  $\emptyset$  has an upper bound in  $X$  is equivalent to saying that  $X$  is nonempty. Thus the hypothesis of Zorn's Lemma implicitly requires that  $X$  be nonempty. Consequently, when applying Zorn's Lemma, typically one first verifies explicitly that  $X$  is nonempty and then establishes that each *nonempty* chain in  $X$  is bounded above.

The original proofs of maximal principles such as Zorn's Lemma, by Hausdorff and Kuratowski, were based upon the Well-ordering Theorem (0.117). The proof of Zorn's Lemma here, by contrast, directly uses the Axiom of Choice (0.26) and avoids any use of the Well-ordering Theorem. The argument is due to H. Kneser [41]; its presentation below is based upon versions by Grayson [30] and Gaillard [26].

Given a subset  $A$  of a partially ordered set  $B$ , we shall call  $A$  an *initial segment* of  $B$  when

$$B \cap ]\leftarrow, a[ \subset A \text{ for each } a \in A,$$

or equivalently,

$$\{eq:star-zorn-pf\} \quad (*) \quad B \cap ]\leftarrow, a[ = A \cap ]\leftarrow, a[ \text{ for each } a \in A,$$

lem:zorn-prelim-lemma

**0.116 Lemma.** Let  $\mathcal{V}$  be a collection of well-ordered subsets of a partially ordered set  $X$  such that for all  $V, W \in \mathcal{V}$ , one of  $V$  and  $W$  is an initial segment of the other. Then  $\bigcup \mathcal{V}$  is well-ordered, and each  $V \in \mathcal{V}$  is an initial segment of  $\bigcup \mathcal{V}$ .

**Proof.** Let  $U = \bigcup \mathcal{V}$ . First we show that  $U$  is totally ordered. Certainly  $U$  is partially ordered. Now let  $x, y \in U$ . There exist  $V, W \in \mathcal{V}$  with  $x \in V, y \in W$ . By hypothesis, one of  $V$  and  $W$  is an initial segment of the other. Then both  $x$  and  $y$  belong to the same well-ordered—hence totally ordered—set and so are comparable.

Second, we show that each  $V \in \mathcal{V}$  is an initial segment of  $U$ . Let  $V \in \mathcal{V}$ . Certainly  $V$  is a subset of  $U$ . Let  $v \in V$ . To see that  $U \cap ]\leftarrow, v[ \subset V$ , let  $x \in U \cap ]\leftarrow, v[$  and just suppose that  $x \notin V$ . Then  $x \in W$  for some  $W \in \mathcal{V}$ . Since  $W \not\subset V$ , by hypothesis  $V$  must be an initial segment of  $W$ . Then  $W \cap ]\leftarrow, v[ \subset V$ . But this is impossible because  $x \in W \cap ]\leftarrow, v[$  but  $x \notin V$ .

Third, we show that the totally ordered set  $U$  is well-ordered. Let  $E$  be a nonempty subset of  $U$ . Then  $E \cap W \neq \emptyset$  for some  $W \in \mathcal{V}$ . Because  $W$  is well-ordered, we may let  $m$  be the least element of  $E \cap W$ . We claim that  $m \leq x$  for all  $x \in E$ . In fact, just suppose there is some  $x \in E$  with  $m \not\leq x$ . Since  $U$  is totally ordered, then  $x < m$ . But this is impossible because  $W$  is an initial segment of  $U$ .  $\square$

**Proof of Zorn's lemma.** Just suppose that  $X$  has *no* maximal element. We shall derive a contradiction. From the hypothesis, each *well-ordered* subset of  $X$  has an upper bound in  $X$ .

Denote by  $\mathcal{W}$  the collection of all well-ordered subsets of  $X$ . For  $W \in \mathcal{W}$  and  $y \in X$ , write  $W < y$  to mean that  $w < y$  for all  $w \in W$ .

pf:zorn-step-1 Step 1: There is a map  $\gamma: \mathcal{W} \rightarrow X$  with  $W < \gamma(W)$  for each  $W \in \mathcal{W}$ .

To prove this, note first that for each  $W \in \mathcal{W}$ , the set

$$A_W = \{x \in X : W < x\}$$

is nonempty. In fact, let  $W \in \mathcal{W}$ . By hypothesis,  $W$  has some upper bound  $b$  in  $X$ . If  $b \notin W$ , then already  $W < b$ . Suppose now that  $b \in W$ . Since  $b$  is *not* maximal in  $X$ , there is an  $x \in X$  with  $b < x$ , and so  $W < x$ .

By the Axiom of Choice (0.26), there is a map  $\gamma': \mathcal{W} \rightarrow \bigcup_{W \in \mathcal{W}} A_W$  such that  $\gamma'(W) \in A_W$  for each  $W$ . Replace the codomain of  $\gamma'$  with  $X$  to obtain  $\gamma$ .

Observe that, since each subset of a well-ordered set is itself well-ordered, the domain of  $\gamma$  includes each subset of a member of  $\mathcal{W}$ .

(Note: Step 1 is the *only* place in the proof where we invoke the Axiom of Choice, and the only place we apply the supposition that  $X$  has no maximal element!)

pf:zorn-step-2 Step 2: Form the collection

$$\mathcal{V} = \{W \in \mathcal{W} : \gamma(W \cap ]\leftarrow, w[) = w : w \in W\}$$

and define

$$U = \bigcup \mathcal{V}.$$

In the remaining steps, we are going to obtain a contradiction by showing:

- $U \in \mathcal{V} \implies U \cup \{\gamma(U)\} \in \mathcal{V}$ ;
- $U \cup \{\gamma(U)\} \notin \mathcal{V}$ ; and
- $U \in \mathcal{V}$ .

pf:zorn-step-3 Step 3:  $U \in \mathcal{V} \implies U \cup \{\gamma(U)\} \in \mathcal{V}$ .

In fact, assume  $U \in \mathcal{V}$ . Let  $v \in U \cup \{\gamma(U)\}$ , so that  $v \in U$  or  $v = \gamma(U)$ .

Suppose first that  $v \in U$ . Since  $U < \gamma(U)$ , we have  $\gamma(U) \notin ]\leftarrow, v[$ . Then

$$\gamma((U \cup \{\gamma(U)\}) \cap ]\leftarrow, v[) = \gamma(U \cap ]\leftarrow, v[) = v$$

because by assumption  $U \in \mathcal{V}$ .

Suppose now that  $v = \gamma(U)$ . Since  $U < \gamma(U)$ , we have  $U \cap ]\leftarrow, \gamma(U)[ = U$ . Hence

$$\gamma((U \cup \{\gamma(U)\}) \cap ]\leftarrow, \gamma(U)[) = \gamma(U \cap ]\leftarrow, \gamma(U)[) = \gamma(U).$$

pf:zorn-step-4 Step 4:  $U \cup \{\gamma(U)\} \notin \mathcal{V}$ .

In fact,  $U \cup \{\gamma(U)\} \not\subseteq U$  because  $U < \gamma(U)$ .

pf:zorn-step-5 Step 5: An initial segment of a member of  $\mathcal{V}$  is a member of  $\mathcal{V}$ .

In fact, let  $W \in \mathcal{V}$  and let  $S$  be an initial segment of  $W$ . Being a subset of a well-ordered set,  $S$  is itself well-ordered. Now let  $s \in S$ . From (\*),

$$\gamma(S \cap ]\leftarrow, s[) = \gamma(W \cap ]\leftarrow, s[).$$

But  $\gamma(W \cap ]\leftarrow, s[) = s$  because  $s \in W \in \mathcal{V}$ . Hence  $\gamma(S \cap ]\leftarrow, s[) = s$ .

pf:zorn-step-6 Step 6: If  $V, W \in \mathcal{V}$ , then  $V$  is an initial segment of  $W$  or  $W$  is an initial segment of  $V$ .

In fact, let  $V, W \in \mathcal{V}$ . Define  $I$  to be the set of those  $x \in X$  that belong to some initial segment of  $V$  that is also an initial segment of  $W$ . We establish several claims.

pf-zorn-claim-a Claim (a):  $I$  is itself an initial segment of both  $V$  and  $W$ , namely, the largest such initial segment.

In fact, already  $I \subset V$  and  $I \subset W$ . Let  $x \in I$ . This means that  $x \in S$  for some set  $S$  that is an initial segment of both  $V$  and  $W$ . Then  $S \subset I$ , and so

$$V \cap ]\leftarrow, x[ \subset S \subset I, \quad W \cap ]\leftarrow, x[ \subset S \subset I.$$

pf-zorn-claim-b Claim (b):  $I$  is a member of  $\mathcal{V}$ .

This is a consequence of Step 5 and (a).

pf-zorn-claim-c Claim (c):  $I = V$  or  $I = W$ .

In fact, just suppose  $I \neq V$  and  $I \neq W$ . We are going to show that then  $I \cup \{\gamma(I)\}$  is also an initial segment of both  $V$  and  $W$ , which will be impossible because  $I$  is the largest such initial segment.

Since  $V$  and  $W$  are well ordered, their subsets  $V \setminus I$  and  $W \setminus I$  have least elements  $v_0$  and  $w_0$ , respectively.

Let

$$I' = I \cup \{\gamma(I)\}.$$

We shall now show that  $I'$  is also an initial segment of both  $V$  and  $W$ , which will contradict the fact already established that  $I$  is the largest such common initial segment.

Notice first that

$$V \cap ]\leftarrow, v_0[ = I = W \cap ]\leftarrow, w_0[.$$

Since  $V, W \in \mathcal{V}$ ,

$$\gamma(V \cap ]\leftarrow, v_0[) = v_0, \quad \gamma(W \cap ]\leftarrow, w_0[) = w_0,$$

and so

$$v_0 = \gamma(I) = w_0.$$

Then  $\gamma(I)$  belongs to both  $V$  and  $W$ . Hence  $I' \subset V$  and  $I' \subset W$ .

Zorn's Lemma  
Well-ordering Theorem  
Cantor, Georg  
Zermelo, Ernst

It remains to show that  $V \cap ]\leftarrow, j[ \subset I'$  for all  $j \in I'$ ; and similarly for  $W$ . If  $j \in I$ , then  $V \cap ]\leftarrow, j[ \subset I$  because  $I$  is an initial segment of  $V$ . And from

$$\gamma(I) = \gamma(V \cap ]\leftarrow, v_0[) = v_0,$$

we obtain

$$V \cap ]\leftarrow, \gamma(I)[ = V \cap ]\leftarrow, v_0[ \subset I$$

because  $v_0 \in V$  and  $I$  is an initial segment of  $V$ .

pf:zorn-step-7 Step 7:  $U \in \mathcal{V}$ .

By Step 6 and Lemma 0.116, the set  $U$  is well-ordered, that is,  $U \in \mathcal{W}$ .

Finally, we show that  $\gamma(U \cap ]\leftarrow, u[) = u$  for all  $u \in U$ . Let  $u \in U$ . Then  $u \in W$  for some  $W \in \mathcal{V}$ , so that

$$\gamma(W \cap ]\leftarrow, u[) = u.$$

By Lemma 0.116,  $W$  is an initial segment of  $U$ , and so from (\*) on page 113,

$$\gamma(U \cap ]\leftarrow, u[) = \gamma(W \cap ]\leftarrow, u[) = u. \quad \square$$

## The Well-ordering Theorem

The theorem asserts that *any set can be well-ordered*. It is essential to realize that:

- it does *not* say that any given total ordering is a well-ordering; and
- it does *not* provide a recipe for how to construct any specific well-ordering: it is purely an existence statement.

Cantor, the creator of set theory, had simply accepted the statement of the Well-ordering Theorem as not requiring any proof. In 1904, Zermelo deduced the theorem from the [Axiom of Choice](#).

thm:wo **0.117 Well-ordering Theorem.** *On each set there exists some well-ordering.*

**Proof.** Let  $X$  be a set.

pfstep-wothm-collection Step 1: Form a partially ordered collection of sets.

Define

$$\mathcal{E} = \{ \langle E, \alpha \rangle : E \subset F \text{ and } \alpha \text{ is a well-ordering of } E \}.$$

We want to show that  $\langle X, \alpha \rangle \in \mathcal{E}$  for some  $\alpha$ . Define a relation  $\leq$  in  $\mathcal{E}$  as follows: For  $\langle E, \alpha \rangle, \langle F, \beta \rangle \in \mathcal{E}$ ,

$$\begin{array}{ll} \{\text{pf-wo-cond-i}\} & \text{(i)} \\ \{\text{pf-wo-cond-ii}\} & \text{(ii)} \\ \{\text{pf-wo-cond-iii}\} & \text{(iii)} \end{array} \quad \langle E, \alpha \rangle \leq \langle F, \beta \rangle \iff \begin{cases} E \subset F, \\ \alpha \subset \beta, \text{ that is, } \beta \text{ induces } \alpha \text{ on } E, \\ x \in E \text{ and } y \in F \setminus E \implies x \beta y. \end{cases}$$

In other words,  $\langle E, \alpha \rangle \leq \langle F, \beta \rangle$  means that  $E$  is a subset of  $F$  and that  $\beta$  extends  $\alpha$  from  $E$  to  $F$ .

It is easy to check that  $\leq$  partially orders  $\mathcal{E}$ . We are going to apply Zorn's Lemma to the partially ordered set  $(\mathcal{E}, \leq)$ .

pfstep-wothm-chains-bded

Step 2: Each chain in  $\mathcal{E}$  is bounded above.

In fact, the collection  $\mathcal{E}$  is nonempty because trivially  $\langle \emptyset, \emptyset \rangle \in \mathcal{E}$ , and so the empty subcollection of  $\mathcal{E}$  is bounded above in  $\mathcal{E}$ . Now let  $\mathcal{A}$  be a nonempty chain in  $\mathcal{E}$ . We want to find an upper bound of  $\mathcal{A}$  in  $\mathcal{E}$ .

Write

$$\mathcal{A} = \{ \langle A_i, \alpha_i \rangle : i \in I \}$$

for some family  $\langle A_i \rangle_{i \in I}$  of subsets of  $X$  and some family  $\langle \alpha_i \rangle_{i \in I}$  of relations. The collection  $\{A_i : i \in I\}$  is totally ordered by set inclusion according to (i), and  $\{\alpha_i : i \in I\}$  is totally ordered by set inclusion according to (ii). Define

$$B = \bigcup_{i \in I} A_i, \quad \beta = \bigcup_{i \in I} \alpha_i.$$

Then  $B \subset X$  and  $\beta$  is a relation in  $B$ . We are going to show that  $(B, \beta)$  is an upper bound of  $\mathcal{A}$  in  $\mathcal{E}$ .

Observe that:

{pf-wothm-same-i} (\*) If  $x, y \in B$ , there is some  $i \in I$  with  $x, y \in A_i$ .

{pf-wothm-beta-and-alpha} (\*\*) If  $i \in I$  and  $x, y \in A_i$ , then  $x \beta y \iff x \alpha_i y$ .

In fact, (\*) holds because  $\{A_i : i \in I\}$  is totally ordered by inclusion. We show that (\*\*) holds. First, if  $x \alpha_i y$ , then  $x \beta y$  because  $\alpha_i \subset \beta$ . Conversely, suppose  $x \beta y$ . There is some  $j \in I$  with  $x \alpha_j y$ . From (ii) and the comparability of  $\langle A_i, \alpha_i \rangle$  with  $\langle A_j, \alpha_j \rangle$ , we have  $x \alpha_i y$ .

To show that  $(B, \beta)$  is an upper bound of  $\mathcal{A}$  in  $\mathcal{E}$ , we must verify that  $\beta$  actually well-orders  $B$ . Certainly  $\beta$  is reflexive because each  $\alpha_i$  is. From (\*) and (\*\*),  $\beta$  is antisymmetric and transitive because each  $\alpha_i$  is.

Let  $E$  be a nonempty subset of  $B$ . We show that  $E$  has a least element for  $\beta$ . There is some  $i \in I$  with  $E \cap A_i \neq \emptyset$ , and then  $E \cap A_i$  has a least element  $x_0$  for the well-ordering  $\alpha_i$  of  $A_i$ . Now let  $x \in E$ . There is some  $j \in I$  with  $x \in A_j$ . If also  $x \in A_i$ , then  $x_0 \alpha_i x$  and so  $x_0 \beta x$  by (\*\*). If, however,  $x \notin A_i$ , so that  $A_j \not\subset A_i$ , then  $\langle A_i, \alpha_i \rangle < \langle A_j, \alpha_j \rangle$  by (i) and the comparability property of  $\leq$ ; in this case,  $x_0 \alpha_j x$  by (ii), and so again  $x_0 \beta x$  by (\*\*).

pfstep-wothm-get-maximal

Step 3:  $\mathcal{E}$  has a maximal member  $M$ .

This follows from Zorn's Lemma (0.115).

pfsteps-wothm-M-equals-X

Step 4:  $M = X$ .

In fact, just suppose  $M \neq X$ . Choose some  $x \in X \setminus M$ , let  $E = M \cup \{x\}$ , and let  $\beta$  be the extension of  $\alpha$  to  $E$  that makes  $x$  the greatest element. Then  $\langle E, \beta \rangle \in \mathcal{E}$  but  $\langle M, \alpha \rangle < \langle E, \beta \rangle$ . This contradicts the maximality of  $M$  in  $\mathcal{E}$ .  $\square$

The preceding proof is a typical procedure for applying Zorn's Lemma. To show that a set  $X$  has a certain property, carry out the following steps:

Step 1: Form the collection  $\mathcal{E}$  of all subsets of  $X$  having this property and suitably partially order  $\mathcal{E}$ .

Step 2: Verify that each chain of  $\mathcal{E}$  has an upper bound in  $X$ .

Step 3: Apply Zorn's Lemma to obtain a maximal element  $M$  of  $\mathcal{E}$ .

Step 4: Conclude that  $M = X$  by showing that otherwise  $M \cup A \in \mathcal{E}$  for some  $A \neq \emptyset$ .

cardinality  
cardinality/finite set@of finite set  
subsec:card  
finite set@cardinality@and cardinality  
cardinality!denumerable set@of denumerable set  
denumerable set@cardinality@and cardinality

## Cardinality

For finite sets  $X$  and  $Y$ , existence of some bijection from  $X$  to  $Y$  means that the two sets have the same number of elements:  $\#(X) = \#(Y)$ . The following definition generalizes this idea to arbitrary sets.

**0.118 Definition.** Let  $X$  and  $Y$  be sets. When some bijection from a  $X$  to  $Y$  exists, we say that  $X$  **has the same cardinality as**  $Y$ , or that  $X$  and  $Y$  **have the same cardinality**, and we write

$$\text{card } X = \text{card } Y.$$

in the contrary case, when no such bijection exists, we write  $\text{card } X \neq \text{card } Y$ .

Note that, for a set  $S$ , we are not giving any meaning here to  $\text{card } S$  as a mathematical object! (How to give such meaning is discussed at the end of this subsection; see [page 122](#).) And so, at least for now, the expression “ $\text{card } X = \text{card } Y$ ” is *not* asserting the equality of two mathematical objects: it is just shorthand for “ $X$  has the same cardinality as  $Y$ .” Fortunately, the notation is harmless. In fact, Propositions [0.14 \(1\)](#), [0.14 \(2\)](#), and [0.10 \(3\)](#), respectively, tell us:

- it:card reflexive • each set  $X$  has the same cardinality as itself;
- it:card symmetric • if  $X$  has the same cardinality as  $Y$ , then  $Y$  has the same cardinality as  $X$ ; and
- it:card-transitive • if  $X$  has the same cardinality as  $Y$  and if  $Y$  has the same cardinality as  $Z$ , then  $X$  has the same cardinality as  $Z$ .

In the card notation, that means:

- $\text{card } X = \text{card } X$ ;
- if  $\text{card } X = \text{card } Y$ , then  $\text{card } Y = \text{card } X$ ; and
- if  $\text{card } X = \text{card } Y$  and  $\text{card } Y = \text{card } Z$ , then  $\text{card } X = \text{card } Z$ .

In this regard, “equality” of cardinalities behaves just like as ordinary equality of mathematical objects—even though, so far, we do not have any actual mathematical objects being equal.

For a *finite* set  $X$ , we do define  $\text{card } X$  as an object, namely:

$$\text{card } X = \#X,$$

so that  $\text{card } X$  is a natural number. And in the case of a *denumerable* set  $X$ , we write

$$\text{card } X = \aleph_0,$$

where “ $\aleph_0$ ” is read as “aleph zero” or “aleph null.” ( $\aleph$  is the first letter of the Hebrew alphabet.) For now, and for practical purposes, the expression “ $\text{card } X = \aleph_0$ ” is nothing more than a notation meaning “ $X$  is denumerable.” Later, in the subsection “The first uncountable ordinal” ([page 110](#)), we shall identify  $\aleph_0$  as an actual mathematical object.

**0.119 Examples.** (1) From results in the preceding subsection:

$$\text{card } \mathbb{N}^* = \text{card}(\mathbb{N} \times \mathbb{N}) = \text{card } \mathbb{Z} = \text{card } \mathbb{Q} = \aleph_0.$$

(2) Adjoining a finite set to an infinite set does not change its cardinality. In other words, if  $S$  is an *infinite* set and  $F$  is a *finite* set disjoint from  $S$ , then

$$\text{card}(S \cup F) = \text{card } S.$$

In fact, if  $F$  is empty, we are done, so suppose now that  $F$  is nonempty. Consider first the particular case that  $F = \{y\}$ , a singleton. By [Proposition 0.47](#), the infinite set  $S$



has a denumerable subset  $D$ ; there is a sequence  $\langle d_n \rangle_{n \in \mathbb{N}}$  of *distinct* elements of  $D$  with  $D = \{d_n : n \in \mathbb{N}\}$ . Informally, we now apply a “Hilbert’s hotel” argument (see [Footnote 3, page 57](#)) to shift all the elements in this sequence to the right, thereby uncovering  $d_0$ , and to put in its place the new element  $y$ . Formally, we define a map  $f: S \cup \{y\} \rightarrow S$  by

$$f(x) = \begin{cases} d_0 & \text{if } x = y, \\ d_{n+1} & \text{if } x = d_n \text{ for some } n \in \mathbb{N}, \\ x & \text{if } x \in S \setminus D. \end{cases}$$

This map is bijective.

The general case of nonempty finite  $F$  now follows by induction on the number of elements of  $F$ .

- ex:card-interval (3) Every nondegenerate interval in  $\mathbb{R}$  has the same cardinality as  $\mathbb{R}$ ; in particular,

$$\text{card } ]0, 1[ = \text{card } ]0, 1] = \text{card } [0, 1[ = \text{card } [0, 1] = \text{card } \mathbb{R}.$$

In fact, for  $a < b$  we have  $\text{card } ]a, b] = \text{card } (]a, b[ \cup \{b\}) = \text{card } ]a, b[$  from (2); similarly,  $\text{card } [a, b[ = \text{card } ]a, b[$  and  $\text{card } [a, b] = \text{card } ]a, b[$ . For  $a < b$ , the linear function  $[0, 1] \rightarrow [a, b]$  given by  $t \mapsto a + (b - 1)t$  is bijective, as are its restrictions  $[0, 1[ \rightarrow [a, b[, ]0, 1] \rightarrow ]a, b]$ , and  $]0, 1[ \rightarrow ]a, b[$ .

Similarly, each open or closed ray has the same cardinality as the open ray  $]0, +\infty[$ .

Finally, the function  $] -1, 1[ \rightarrow \mathbb{R}$  given by  $t \mapsto \arctan t$  is bijective, and so  $\text{card } \mathbb{R} = \text{card } ] -1, 1[ = \text{card } ]0, 1[$ .

- ex:card-2-power-01-powerset (4) According to [Proposition 0.16](#), for any set  $X$ ,

$$\text{card } \mathcal{P}(X) = \text{card}(2^X).$$

This common cardinality is denoted by  $2^{\text{card } X}$ , so that  $\text{card } \mathcal{P}(X) = 2^{\text{card } X}$ . In particular, the power set of a denumerable set has cardinality  $2^{\aleph_0}$ .  $\diamond$

The following result improves [Examples 0.119 \(2\)](#).

prop:card-adjoin-count-to-infinite

**0.120 Proposition.** *Let  $S$  be an infinite set and  $C$  be a countable set. Then*

$$\text{card}(S \cup C) = \text{card } S.$$

**Proof.** Since  $S \cup C = S \cup (C \setminus S)$  and  $C \setminus S$  is still countable, we may assume without loss of generality that  $S$  and  $C$  are disjoint. If  $C$  is finite, the result follows already from [Examples 0.119 \(2\)](#). So suppose now that  $C$  is infinite, hence denumerable. The infinite set  $S$  has a denumerable subset  $D$ . Hence there are sequences  $\langle d_n \rangle_{n \in \mathbb{N}}$  and  $\langle c_n \rangle_{n \in \mathbb{N}}$ , each of distinct elements, with  $D = \{d_n : n \in \mathbb{N}\}$  and  $C = \{c_n : n \in \mathbb{N}\}$ . Construct a bijection  $S \cup C \rightarrow S$  by shifting each  $d_n$  to  $d_{2n+1}$ , thereby uncovering all the even-index members of the sequence, and replacing these even-index members with the successive  $c_n$ .<sup>6</sup>  $\square$

Here is an important application of the preceding proposition.

<sup>6</sup>This variation of “Hilbert’s hotel” (see [Footnote 3, page 57](#)) allows the previously full hotel to accommodate denumerably many new guests!

**0.121 Example (cardinality of  $\mathbb{R}$ ).** As a consequence of order-completeness of the real numbers,  $\text{card } \mathbb{R} = 2^{\aleph_0}$ .

$$\text{card } \mathbb{R} = 2^{\aleph_0}.$$

To prove this, it suffices to show that  $\text{card } 2^{\aleph_0} = \text{card } ]0, 1[$ , since  $\text{card } \mathbb{R} = \text{card } ]0, 1[$  [Examples 0.119 (3)].

Let

$$C = \{b \in 2^{\aleph_0} : \text{there is some } j \geq 0 \text{ with } b_i = 0 \text{ for all } i \geq j\}.$$

be the set of binary sequences  $\langle b_n \rangle_{n \in \mathbb{N}^*}$  that are eventually constant and let  $B = 2^{\aleph_0} \setminus C$ , the set of those that are *not*. Since the subset  $B$  of  $2^{\aleph_0}$  is denumerable while the interval  $]0, 1[$  is uncountable, by Proposition 0.120,  $\text{card}(]0, 1[ \cup B) = \text{card } ]0, 1[$ . Hence it suffices to show that  $\text{card } 2^{\aleph_0} = \text{card}(]0, 1[ \cup B)$ .

According to Proposition 0.77 applied to binary expansion, for each  $x \in ]0, 1[$ , there is a *unique*  $\langle x_n \rangle_{n \in \mathbb{N}^*} \in B$  for which  $x = \sum_{n=1}^{\infty} x_n / 2^n$ . Define the map  $f: 2^{\aleph_0} \rightarrow ]0, 1[ \cup B$  as follows:

$$f(b) = \begin{cases} \sum_{i=1}^{\infty} b_i / 2^i & \text{if } b \notin C, \\ b & \text{if } b \in B, \end{cases}$$

Then  $f$  is bijective.

For a somewhat different proof, see Exercise 182.  $\diamond$

Beyond distinguishing between finite sets and infinite sets, and between denumerable and uncountable sets, we may compare the sizes of any two sets.

**0.122 Definition.** Let  $X$  and  $Y$  be sets. When some injection  $X \rightarrow Y$  exists, we write  $\text{card } X \leq \text{card } Y$ . When  $\text{card } X \leq \text{card } Y$  but  $\text{card } X \neq \text{card } Y$ , we write  $\text{card } X < \text{card } Y$ .

Equivalently,  $\text{card } X \leq \text{card } Y$  exactly when  $\text{card } X = \text{card } S$  for some subset  $S$  of  $Y$ .

The definition is consistent with both the relation  $\#X \leq \#Y$  that holds for a subset  $X$  of a finite set  $Y$  (Proposition 0.40) and the fact that a subset of a countable set is countable (see Proposition 0.46).

In terms of the definition,  $\text{card } A \leq \text{card } B$  whenever  $A \subset B$ ; however, it need not be the case that  $\text{card } A < \text{card } B$  even when  $A \neq B$  (for example,  $A = \mathbb{N}$  and  $B = \mathbb{Z}$ ). For any finite set  $F$ , denumerable set  $D$ , and uncountable set  $K$ , we have  $\text{card } F < \text{card } D < \text{card } K$ .

The following result generalizes both Proposition 0.43 and Proposition 0.48. It is a consequence of existence of a section for a surjection (Proposition 0.30).

**0.123 Proposition (cardinality of range of a map).** If  $X$  and  $Y$  are sets for which there exists a surjection  $X \rightarrow Y$ , then  $\text{card } Y \leq \text{card } X$ .

The  $\leq$  notation for comparing cardinalities suggests that a partial ordering or even total ordering of sets is at issue. And this in fact turns out to be true. Reflexivity and transitivity are immediate.

**0.124 Proposition (cardinality transitivity).** (1) For a set  $X$ , we have  $\text{card } X \leq \text{card } X$ .

(2) For sets  $X, Y$ , and  $Z$ , if  $\text{card } X \leq \text{card } Y$  and  $\text{card } Y \leq \text{card } Z$ , then  $\text{card } X \leq \text{card } Z$ .

Antisymmetry and comparability for cardinality are dealt with below, in Theorem 0.126 and Theorem 0.127, respectively.) The following theorem establishes existence of sets of arbitrarily large cardinality.

thm:cantor

**0.125 Cantor's Theorem.** *Let  $X$  be a set. Then*

$$\text{card } X < \text{card } \mathcal{P}(X).$$

Cantor, Georg  
Bernstein, Felix  
Schröder, Ernst

**Proof.** First,  $\text{card } X \leq \text{card } \mathcal{P}(X)$  because the map  $x \mapsto \{x\}$  is injective. To see that  $\text{card } X \neq \text{card } \mathcal{P}(X)$ , just suppose there exists a surjection  $f: X \rightarrow \mathcal{P}(X)$ . Form the element  $A = \{x \in X : x \notin f(x)\}$  of  $\mathcal{P}(X)$ . Then  $A = f(x)$  for some  $x \in X$ . But this is manifestly impossible.  $\square$

In view of [Examples 0.119 \(4\)](#), the conclusion of Cantor's Theorem may be expressed as

$$\text{card } X < 2^{\text{card } X}.$$

Thus we have the infinite sequence of successively larger cardinalities

$$\aleph_0 < 2^{\aleph_0} < 2^{(2^{\aleph_0})} < 2^{(2^{(2^{\aleph_0})})} < \dots$$

The notation  $\leq$  used in comparisons  $\text{card } X \leq \text{card } Y$  suggests that if  $\text{card } X \leq \text{card } Y$  and  $\text{card } Y \leq \text{card } X$ , then  $\text{card } X = \text{card } Y$ . But this requires proof! The implication was originally asserted, without proof, by Cantor and was subsequently proved by Cantor's student Felix Bernstein, and others. ("Schröder" is often included in the theorem's name, referring to Ernst Schröder, who proffered a proof before Bernstein's that turned out to be flawed.)

thm:CB

**0.126 Cantor-Bernstein Theorem.** *Let  $X$  and  $Y$  be sets. If  $\text{card } X \leq \text{card } Y$  and  $\text{card } Y \leq \text{card } X$ , then  $\text{card } X = \text{card } Y$ .*

**Proof.** Assume there are injections  $f: X \rightarrow Y$  and  $g: Y \rightarrow X$ . Define sequences  $\langle A_n \rangle_{n \in \mathbb{N}}$  and  $\langle B_n \rangle_{n \in \mathbb{N}}$  of subsets of  $X$  and  $Y$ , respectively, by

$$\begin{aligned} A_n &= \begin{cases} X \setminus g(Y) & \text{if } n = 0, \\ g(f(A_{n-1})) & \text{if } n > 0. \end{cases} \\ B_n &= f(A_n) \text{ for } n = 0, 1, 2, \dots, \end{aligned}$$

and set

$$A = \bigcup_{n=0}^{\infty} A_n, \quad B = \bigcup_{n=0}^{\infty} B_n.$$

Notice that  $A_n = g(B_{n-1})$  for each  $n$ .

To simplify the notation, write  $g^{-1}$  to mean the map  $g(Y) \rightarrow X$  that is the inverse of the codomain restriction  $Y \rightarrow g(Y)$  of  $g$ . In terms of that notation, define  $h: X \rightarrow Y$  by

$$h(x) = \begin{cases} f(x) & \text{if } x \in A, \\ g^{-1}(x) & \text{if } x \in X \setminus A. \end{cases}$$

The definition of  $h(x)$  in the latter case makes sense since in this case  $x \notin A_0$  and so  $x \in g(Y)$ . We shall show that  $h: X \rightarrow Y$  is bijective.

The map  $h$  is injective: Let  $x, u \in X$  with  $x \neq u$ . If either  $x$  and  $u$  both belong to  $A$  or neither belongs to  $A$ , then already  $h(x) \neq h(u)$  because both  $f$  and  $g^{-1}$  are injective. Now let  $x \in A$  but  $u \notin A$  (the reverse situation is handled similarly). Choose  $n$  with  $x \in A_n$ . Then  $h(x) \in B_n$ . On the other hand,  $h(u) = g^{-1}(u) \notin B_n$ , since otherwise  $u \in A_{n+1}$ . Hence  $h(x) \neq h(u)$ .

Cantor, Georg  
Zermelo, Ernst  
trichotomy!  
cardinal number!  
ordinal!  
cardinal number  
ordinal!  
cardinality!

The map  $h$  is surjective: Already  $B \subset h(X)$ , so suppose  $y \in Y \setminus B$ . Then  $g(y) \notin X \setminus g(Y) = A$ . Moreover, for each  $n$  we have  $y \notin B_n$ , and so  $g(y) \notin g(B_n) = A_{n+1}$ . Hence  $g(y) \notin \bigcup_{n \in \mathbb{N}} A_{n+1} = h(g(Y))$ .  $\square$

For a different proof, see [Exercise 185](#).

One application of the Cantor-Bernstein Theorem is to a proof of  $\text{card } \mathbb{R} = 2^{\aleph_0}$  that, unlike the one in [Example 0.121](#), avoids use of [Examples 0.119 \(3\)](#): see [Exercise 182](#).

Cantor stated the following result but never proved it. The first proof, by Zermelo, applied the Well-ordering Theorem ([0.117](#)); the one outlined here uses just Zorn's Lemma ([0.115](#)).

thm:cards-comparable

**0.127 Theorem (cardinal comparability).** For sets  $X$  and  $Y$ , either  $\text{card } X \leq \text{card } Y$  or  $\text{card } Y \leq \text{card } X$ .

**Proof.** Just suppose it is *not* the case that  $\text{card } Y \leq \text{card } X$ , that is, there is *no* injection  $Y \rightarrow X$ . The aim is to deduce that existence of some injection  $X \rightarrow Y$ . In the collection

$$\mathcal{F} = \{f: S \rightarrow Y : S \subset X \text{ and } f \text{ is injective}\}$$

of “partial injections,” define the relation  $\leq$  by  $f \leq g$  if and only if  $g$  extends  $f$ . This relation is a partial ordering of  $\mathcal{F}$  [compare [Examples 0.60 \(7\)](#)].

It can be verified that each chain in  $\mathcal{F}$  has an upper bound in  $\mathcal{F}$ , and so according Zorn's Lemma ([0.115](#)),  $\mathcal{F}$  has a maximal member  $m$ . If the domain of  $m$  is a proper subset  $S$  of  $X$ , then it can be extended to a member of  $\mathcal{F}$  with domain having one additional element. But this contradicts the maximality of  $m$  in  $\mathcal{F}$ .  $\square$

In view of cardinal comparability ([Theorem 0.127](#)) and the Cantor-Bernstein Theorem ([0.126](#)), we have the following **trichotomy law** for cardinalities: For two sets  $X$  and  $Y$ , exactly one of the three alternatives  $\text{card } X < \text{card } Y$ ,  $\text{card } X = \text{card } Y$ ,  $\text{card } Y < \text{card } X$  holds.

So far we have used such expressions as “ $\text{card } X = \text{card } Y$ ” and “ $\text{card } X < \text{card } Y$ ” without identifying  $\text{card } X$  or  $\text{card } Y$  as mathematical objects, and that suffices for our purposes in developing topology. But cardinal numbers may be defined as actual mathematical objects, namely, certain ordinal numbers.

pg:card-object-def

To see how to define cardinal numbers, recall that the class of all ordinals is well-ordered by the relation  $\leq$  defined by  $\beta \leq \alpha$  if and only if  $\beta \in \alpha$  or  $\beta = \alpha$ . By definition a **cardinal number** is an *initial ordinal*, that is, an ordinal  $\alpha$  having the property that there is *no* ordinal  $\beta < \alpha$  (that is,  $\beta \in \alpha$ ) for which there is a bijection between  $\beta$  and  $\alpha$ .

According to this definition, each natural number is a cardinal number; in fact, a set is a finite cardinal number exactly when it is a natural number. The set  $\omega$  is the first infinite cardinal number, and in this role is precisely what we previously denoted by  $\aleph_0$ . By contrast, the ordinal  $\omega + 1 = \omega \cup \{\omega\}$  is *not* a cardinal number.

pg:omega-successor-not-cardinal

It can be proved that, for each set  $X$  there is a unique cardinal number  $\mathfrak{a}$  for which there exists a bijection  $X \rightarrow \mathfrak{a}$ , and then this cardinal number is called the **cardinality of**  $X$  and is denoted by  $\text{card } X$ .

It turns out that for cardinal numbers  $\mathfrak{a}$  and  $\mathfrak{b}$ :

- $\mathfrak{a} = \mathfrak{b}$  if and only if there exists a bijection between the sets  $\mathfrak{a}$  and  $\mathfrak{b}$ , in other words,  $\text{card } \mathfrak{a} = \text{card } \mathfrak{b}$  in the sense of [Definition 0.118](#); and
- $\mathfrak{a} < \mathfrak{b}$  if and only if there is an injection, but no bijection, from  $\mathfrak{a}$  to  $\mathfrak{b}$ , in other words,  $\text{card } \mathfrak{a} < \text{card } \mathfrak{b}$  in the sense of [Definition 0.122](#).

There is an arithmetic of cardinal numbers that extends the arithmetic of natural numbers.. The sum  $m + n$  of two cardinal numbers is defined as the cardinality of the Cartesian sum of the two sets; the product  $m \cdot n$  is defined as the cardinality of the Cartesian product of the two sets; and the power  $m^n$  is defined as the cardinality of the corresponding power set consisting of all maps  $n \rightarrow m$ . The usual “laws of arithmetic” for natural numbers also hold for cardinal numbers. For example, cardinal numbers  $s$ ,  $i$ , and  $j$  satisfy the power law  $(a^i)^j = a^{i \cdot j}$ ; this is a consequence of the following result.

cardinal number!ordinal@and ordin  
cardinality  
choice function!well-ordered set@ar

prop:power-law-cardinalities

**0.128 Proposition (power law of cardinality).** Let  $A$ ,  $I$ , and  $J$  be sets. Then  $\text{card}(A^I)^J = \text{card } A^{I \times J}$ .

**Proof.** Define maps  $\varphi: (A^I)^J \rightarrow A^{I \times J}$  and  $\psi: A^{I \times J} \rightarrow (A^I)^J$  as follows. For arbitrary  $x \in (A^I)^J$ ,

$$\varphi(x)\langle i, j \rangle = (x(j))(i) \quad (\langle i, j \rangle \in I \times J),$$

and for arbitrary  $y \in A^{I \times J}$ ,

$$(\psi(y)(j))(i) = y(i, j) \quad (j \in J, i \in I).$$

Then  $\psi \circ \varphi$  and  $\varphi \circ \psi$  are the identity maps of their respective domains.  $\square$

The only bit of cardinal arithmetic we shall need in the sequel is the following consequence of the preceding proposition.

cor:card-R-power-N

**0.129 Corollary (cardinality of  $\mathbb{R}^{\mathbb{N}}$ ).**  $\text{card } \mathbb{R}^{\mathbb{N}} = 2^{\aleph_0}$ .

**Proof.** We have the string of equalities  $\text{card } \mathbb{R}^{\mathbb{N}} = \text{card}(2^{\mathbb{N}})^{\mathbb{N}} = \text{card } 2^{\mathbb{N} \times \mathbb{N}} = \text{card } 2^{\mathbb{N}} = 2^{\aleph_0}$ , where the penultimate equality holds because  $\mathbb{N} \times \mathbb{N}$  is denumerable.  $\square$

For a complete and rigorous development of about cardinal numbers, including the facts presented above, see, for example, Rubin [56].

## EXERCISES FOR SECTION 0.10

prob:two immediate neighbors

- 163.** (a) Show that each element of a well-ordered set has an immediate successor.  
(b) Must each element of a well-ordered set have an immediate predecessor?

- 164.** Let  $\mathcal{A}$  be a collection of nonempty subsets of a well-ordered set  $X$ . Without using any form of the Axiom of Choice, show that there exists a choice function (page 45) for  $\mathcal{A}$ .

- 165.** Let  $R$  be a relation in a set  $X$  for which each nonempty subset of  $X$  has a *unique* least element. Prove that  $R$  in fact totally orders  $X$  and hence well-orders  $X$  in the sense of Definition 0.107.

rob:lexicographic-order-product-w

- 166.** Prove that the product of two well-ordered sets is well-ordered by the lexicographic ordering (Example 0.70).

prob:extend-well-ordering-greatest-elt **167.** Let  $\langle X, \leq_X \rangle$  be a well-ordered set, let  $z$  be a set with  $z \notin X$ —take  $z = X$ , for example—and let  $X^+ = X \cup \{z\}$ . Extend the ordering of  $X$  to a relation  $\leq$  in  $X^+$  by adjoining  $z$  as a greatest element, that is:

ordinal product

reverselexicographic ordering

Principle of Transfinite Induction

transfinite induction

induction!transfinite

$$\begin{aligned} u \leq v &\iff u, v \in X \text{ and } u \leq_X v, \text{ or} \\ &u \in X \text{ and } v = z, \text{ or} \\ &u = v = z. \end{aligned}$$

- (a) Write the relation  $\leq$  as a subset of  $X^+ \cup X^+$ .  
 (b) Show that  $\leq$  well-orders  $X^+$ .

prob:ordinal-sum-2-wos **168.** (*Generalization of Exercise 167.*) Let  $\langle X, \leq_X \rangle, \langle Y, \leq_Y \rangle$  be two well-ordered sets with  $X$  and  $Y$  disjoint. Define the relation  $\leq$  in  $X \cup Y$  by

$$\begin{aligned} u \leq v &\iff u, v \in X \text{ and } u \leq_X v, \text{ or} \\ &u, v \in Y \text{ and } u \leq_Y v, \text{ or} \\ &u \in X \text{ and } v \in Y. \end{aligned}$$

In other words,  $\leq$  extends the orderings of  $X$  and  $Y$  to  $X \cup Y$  and places all elements of  $Y$  after all elements of  $X$ . Show that  $\leq$  well-orders  $X \cup Y$ . The well-ordered set  $\langle X \cup Y, \leq \rangle$  is called the **ordinal sum of  $\langle X, \leq_X \rangle$  and  $\langle Y, \leq_Y \rangle$** .

prob:ordinal-product-2-wos **169.** Let  $\langle X, \leq_X \rangle, \langle Y, \leq_Y \rangle$  be two well-ordered sets. Define the relation  $\leq$  in  $X \times Y$  as follows: for all  $\langle x, y \rangle, \langle s, t \rangle \in X \times Y$ ,

$$\begin{aligned} \langle x, y \rangle \leq \langle s, t \rangle &\iff y \leq t, \text{ or else} \\ &y = t \text{ and } x \leq s. \end{aligned}$$

(This is known as the **reverse lexicographic ordering** of the product—it is *not* the same as of the lexicographic ordering relation!)

- (a) Show that  $\leq$  well-orders  $X \times Y$ .  
 (b) Explain why  $\Omega \cdot 2$  does not mean the same thing as  $2 \cdot \Omega$ .

prob:transfinite-induction **170.** Prove the **principle of transfinite induction**:

Let  $E$  be a subset of a well-ordered set  $X$ . If for each  $x \in E$ ,

$$]\leftarrow, x[ \subset E \implies x \in E,$$

then  $E = X$ .

[Note that this generalizes to arbitrary well-ordered sets the principle of strong induction for  $\mathbb{N}$  (Exercise 17).]

**171.** Prove the following converse of the principle of transfinite induction (Exercise 170): Let  $X$  be a totally ordered set. Assume that for *every* subset  $E$  of  $X$ , the condition

$$]\leftarrow, x[ \subset E \implies x \in E \quad (x \in X)$$

holds. Then  $X$  is well-ordered.

- 172.** (a) Find all maximal elements, if any, of the set  $\mathbb{N}^*$  of positive integers when it is partially ordered by the divisibility relation, as in Examples 0.55 (2).  
 (b) Find all maximal members of  $\mathcal{P}(\mathbb{N}) \setminus \{\emptyset, \mathbb{N}\}$  when this collection is ordered by subset inclusion.

- (c) Find all maximal members of the collection of all *infinite* proper subsets of  $\mathbb{N}$ . maximal chain
- 173.** A **maximal chain** a partially ordered set  $X$  is a chain in  $X$  that is not properly contained in any other chain. Hausdorff Maximal Principle  
Hausdorff, Felix  
Zorn's Lemma  
maximal chain  
of finite character  
Teichmüller-Tukey Lemma@Teichmüller, Oswald  
Tukey, John W.  
of finite character
- (a) *Without* using the Axiom of Choice, Zorn's Lemma, or any other maximal principle, prove: If  $C$  is a maximal chain in a *finite* partially ordered set, then  $C$  contains a maximal element of  $X$ .
- (b) Use Zorn's Lemma to prove **Hausdorff's Maximal Principle**: A partially ordered set contains a maximal chain.
- prob:of-finite-character **174.** A collection  $\mathcal{F}$  of sets is said to be **of finite character** when it has the following property: A set  $E$  is a member of  $\mathcal{F}$  if and only if each finite subset of  $E$  is a member of  $\mathcal{F}$ .
- (a) Let  $X$  be a set and let  $x_0 \in X$ . Show that  $\{E : E \subset X, x_0 \notin E\}$  is of finite character.
- (b) Let  $X$  be a preordered set. Show that the collection of all chains in  $X$  is of finite character.
- (c) Give an example of a collection of sets that is not of finite character.
- (d) Let  $\mathcal{F}$  be a collection of finite character and let  $E \in \mathcal{F}$ . Show that  $A \in \mathcal{F}$  for each  $A \subset E$ . *Note*: In particular, then, if  $\mathcal{F}$  is a *nonempty* collection that is of finite character, then  $\emptyset \in \mathcal{F}$ .
- prob:Tukey-Lemma **175.** (*Continuation of Exercise 174.*) Prove the **Teichmüller–Tukey Lemma**: A nonempty collection of sets that is of finite character has a maximal member (with respect to subset inclusion).
- 176.** Deduce the axiom of choice (0.28) from Zorn's Lemma. [*Hint*: Let  $\mathcal{A}$  be a collection of nonempty sets. Begin by partially ordering the set  $\mathcal{E}$  consisting of all those ordered pairs  $(\mathcal{B}, f)$  for which  $\mathcal{B} \subset \mathcal{A}$  and  $f: \mathcal{B} \rightarrow \bigcup \mathcal{B}$  is a map such that  $f(A) \in A$  for each  $A \in \mathcal{B}$ .]
- 177.** Show that if  $A$  is uncountable but  $B$  is countable, then  $A \setminus B$  has the same cardinality as  $A$ .
- 178.** In the proof of Theorem 0.125, why it is impossible that  $A = f(x)$  for some  $x$ .
- 179.** Show that, for a set  $X$ , there does *not* exist any injection  $\mathcal{P}(X) \rightarrow X$ .
- sets-implies-injection-of-powersets **180.** Prove or disprove: If  $\text{card } X < \text{card } Y$ , then  $\text{card } \mathcal{P}(X) < \text{card } \mathcal{P}(Y)$ .
- prob:pf-cards-comparable **181.** Fill in the missing details in the proof of Theorem 0.127. Specifically, show that:
- (a) The relation  $\leq$  is a partial ordering of  $\mathcal{F}$ .
- (b) Each chain in  $\mathcal{F}$  has an upper bound there.
- (c) If the domain of a maximal element  $m$  is a proper subset  $S$  of  $X$ , then it can be extended to a member of  $\mathcal{F}$  with domain having one additional element.
- prob:pf-card-R-use-CB **182.** Prove anew that  $\text{card } \mathbb{R} = 2^{\aleph_0}$  by applying the Cantor-Bernstein Theorem (0.126) after constructing injections  $]0, 1[ \rightarrow 2^{\aleph_0}$  and  $2^{\aleph_0} \rightarrow ]0, 1[$  as follows:
- (i) Map each  $b \in ]0, 1[$  to the sequence  $\langle b_n \rangle_{n \in \mathbb{N}^*}$  in  $\{0, 1\}$  that is *not* eventually constant and for which  $x$  has binary representation  $x = \sum_{n=1}^{\infty} b_n / 2^n$ .

real numbers! Dedekind cuts (ii) Compose a suitable injection  $2^{\mathbb{N}^*} \rightarrow 3^{\mathbb{N}^*}$  of binary sequences to ternary sequences with the map  $3^{\mathbb{N}^*} \rightarrow ]0, 1[$  that sends each ternary sequence  $\langle t_n \rangle_{n \in \mathbb{N}^*}$  to  $\sum_{n \in \mathbb{N}^*} t_n / 3^n$ .  
Dedekind cut  
real numbers! cardinality@and cardinality

Cantor-Bernstein Theorem! Knaster-Tarski Fixed-point Theorem@and Knaster-Tarski Fixed-point Theorem  
183. Show anew that  $\text{card } \mathbb{R} \leq 2^{\aleph_0}$  by exhibiting a bijection between  $\mathbb{R}$  and the collection of all Dedekind cuts (Exercise 133).  
Knaster-Tarski Fixed-point Theorem! Cantor-Bernstein Theorem@and Cantor-Bernstein Theorem

184. In the proof of the power law in Proposition 0.128, express the bijection between  $(A^I)^J$  and  $A^{I \times J}$  in terms of families and families of families.

prob:pf-CB-via-knaster-tarski 185. Apply the Knaster-Kuratowski Fixed-point Theorem (Exercise 48) to prove anew the Cantor-Bernstein Theorem (0.126).

[Hint: Assume there exist injections  $f: X \rightarrow Y$  and  $g: Y \rightarrow X$ . Form the map  $\varphi: \mathcal{P}(X) \rightarrow \mathcal{P}(X)$  given by  $A \mapsto X \setminus g(Y \setminus f(A))$ . Establish that  $\varphi$  has a fixed-point  $S$ . Show  $X \setminus g(Y \setminus f(S)) = S$  and then  $g(Y \setminus f(S)) = X \setminus S$ , so that it is meaningful to form the domain-codomain restrictions  $f|_{S, f(S)}: S \rightarrow f(S)$  and  $g|_{Y \setminus f(S), X \setminus S}: Y \setminus f(S) \rightarrow X \setminus S$  of  $f$  and  $g$ , respectively. Verify that these restrictions are bijections. Finally, glue together  $f|_{S, f(S)}$  and  $(g|_{Y \setminus f(S), X \setminus S})^{-1}$  to obtain the desired bijection  $X \rightarrow Y$ .]

186. (a) Justify the assertion on page 122 that the ordinal number  $\omega + 1$  is not a cardinal number.

(b) Which, if any, of the ordinals  $\Omega$ ,  $\Omega + 1$ , and  $\Omega^2$  are cardinal numbers?



## CHAPTER

# 1

## Metric Spaces

chap:metricspace

### Chapter Contents

1.1 Metrics . . . . .	128	1.4 Continuity and Convergence . . . . .	179
The Euclidean metric . . . . .	128	Continuous maps between metric	
Other metrics on $\mathbb{R}^n$ . . . . .	132	spaces . . . . .	180
Metrics on function spaces . . . . .	135	Convergent sequences in metric	
Metrics and metric spaces . . . . .	139	spaces . . . . .	185
Exercises for Section 1.1 . . . . .	142	Relationship between convergence	
1.2 Open Sets and Closed Sets . . . . .	146	and continuity . . . . .	194
Balls, disks, and spheres . . . . .	147	Exercises for Section 1.4 . . . . .	195
Open sets in a metric space . . . . .	152	1.5 Completeness . . . . .	200
Exercises for Section 1.2 . . . . .	161	Cauchy sequences . . . . .	200
1.3 Equivalent Metrics . . . . .	164	Complete metric spaces . . . . .	202
Equivalent metrics on a set . . . . .	164	Completeness of sequence and	
Bounded metric spaces . . . . .	167	functions spaces . . . . .	206
The extended real line . . . . .	168	The Nested Set and Baire Category	
Isometries . . . . .	172	Theorems . . . . .	212
Exercises for Section 1.3 . . . . .	176	Exercises for Section 1.5 . . . . .	217

### Introduction

As a prelude to our study of topology proper, in this chapter we discuss “metric spaces”—sets of points on which there are given numerical measures of distances between these points. Here we have three aims:

- to provide a number of important examples that will recur through the entire text;
- to abstract from the familiar setting of one-, two-, and three-dimensional Euclidean spaces notions of continuous functions and convergent sequences that are meaningful in any metric space; and
- to demonstrate that many concepts associated with metric spaces, and in particular continuity, depend not on the particular measures of distance involved, but only on the “open” sets they define. Thus we prepare the way for the more abstract concept of a “topological space” in which there need be no notion of distance at all.

Euclidean plane  
continuous map!  
Euclidean metric

The final section of the chapter concerns completeness, a notion not amenable to generalization to arbitrary topological spaces. Indeed, this metric property will not reappear until Section 4.2 of Chapter 4 (Compactness), where it is related to the topological property of compactness. Yet completeness is hardly insignificant: the fact that a certain metric space is complete whose “points” are functions will allow us to demonstrate that “most” continuous real-valued functions on an interval have derivatives at no point whatsoever.

## 1.1 Metrics

sec:metrics

Consider  $n$ -dimensional Euclidean space  $\mathbb{R}^n$ , which consists of all  $n$ -tuples  $x = \langle x_1, x_2, \dots, x_n \rangle$  of real numbers, and a real-valued function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  defined on  $\mathbb{R}^n$ . For example, when  $n = 1$ , the set  $\mathbb{R}^n$  is the real line, and  $f$  might be the function of a single variable given by  $f(x) = x \sin x$ . When  $n = 2$ , the set  $\mathbb{R}^n$  is the Euclidean plane, and  $f$  might be the function of two variables given by

$$f(x_1, x_2) = (x_1^2 + x_2^2) / (1 + x_1^2 + x_2^3).$$

When  $n = 3$ , the set  $\mathbb{R}^n$  is Euclidean 3-space, and  $f$  might be the function of three variables given by

$$f(x_1, x_2, x_3) = x_1 x_2 \sin(|1 + x_3|)$$

From calculus you should be familiar (at least for  $n = 1, 2$ , or  $3$ ) with the following informal definition of continuity of such a function  $f$  at a point  $x \in \mathbb{R}^n$ :

The value  $f(u)$  will be as close to the value  $f(x)$  as one wishes provided that  $u \in \mathbb{R}^n$  is close enough to  $x$ .

More precisely,  $f$  will be continuous at  $x$  when:

Given an arbitrary  $\varepsilon > 0$ , there is a corresponding  $\delta > 0$  such that  $f(u)$  is at a distance less than  $\varepsilon$  from  $f(x)$  whenever  $u \in \mathbb{R}^n$  is at a distance less than  $\delta$  from  $x$ .

As it stands, the preceding statement is not quite an adequate definition of continuity, for it does not prescribe how to measure distances between points in  $\mathbb{R}$  and between points in  $\mathbb{R}^n$ . It is to numerical measures of distance in Euclidean and other “spaces” that this section is devoted.

### The Euclidean metric

We begin with measures of distance in  $\mathbb{R}^n$ .

**1.1 Example.** Formulas for distance in  $\mathbb{R}^n$  are familiar enough for dimensions  $n = 1, 2$ , and 3. The distance from a point  $x$  to a point  $y$  in  $\mathbb{R}^1 = \mathbb{R}$  is

$$d(x, y) = |x - y|.$$

The distance from a point  $x = \langle x_1, x_2 \rangle$  to a point  $y = \langle y_1, y_2 \rangle$  in the plane  $\mathbb{R}^2$  is

$$d(x, y) = [(x_1 - y_1)^2 + (x_2 - y_2)^2]^{1/2} = \left[ \sum_{i=1}^2 (x_i - y_i)^2 \right]^{1/2}.$$

The distance from a point  $x = \langle x_1, x_2, x_3 \rangle$  to a point  $y = \langle y_1, y_2, y_3 \rangle$  in  $\mathbb{R}^3$  is

$$d(x, y) = [(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2]^{1/2} = \left[ \sum_{i=1}^3 (x_i - y_i)^2 \right]^{1/2}.$$

The distance formula for dimension 1 may be made to resemble the formulas for dimensions 2 and 3: recalling that  $[a^2]^{1/2} = |a|$  for each real number  $a$ , then for points  $x = x_1$  and  $y = y_1$  in  $\mathbb{R}^1$  we may write

$$d(x, y) = \left[ \sum_{i=1}^1 (x_i - y_i)^2 \right]^{1/2}. \quad \diamond$$

Distance in  $\mathbb{R}^n$  for an arbitrary dimension  $n$  may now be defined so as to fit the same “square-root of the sum of the squares of the differences of the coordinates” pattern.

**1.2 Definition.** Let  $n$  be a positive integer. For points  $x = \langle x_1, x_2, \dots, x_n \rangle$  and  $y = \langle y_1, y_2, \dots, y_n \rangle$  in  $\mathbb{R}^n$ , the **Euclidean distance from  $x$  to  $y$**  is defined to be the real number

$$[(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2]^{1/2} = \left[ \sum_{i=1}^n (x_i - y_i)^2 \right]^{1/2}.$$

The assignment of the Euclidean distance from  $x$  to  $y$  for each ordered pair  $\langle x, y \rangle$  of points in  $\mathbb{R}^n$  is a function

$$d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

called the **Euclidean metric on  $\mathbb{R}^n$** .

This definition does make sense, for the quantity in square brackets there is nonnegative and hence has a unique (nonnegative) square-root.

We must emphasize that the formula

$$d(x, y) = \left[ \sum_{i=1}^n (x_i - y_i)^2 \right]^{1/2}$$

for the Euclidean distance from  $x$  to  $y$  is our *definition*—not something proved by geometrical reasoning. Of course, for dimensions  $n = 2$  and 3 this formula is often proved in elementary mathematics courses under certain assumptions about distance (compare [Exercise 4](#))

In [Definition 1.2](#), the integer  $n$  is positive. Recall that  $\mathbb{R}^0$  denotes the single-element set  $\{0\}$ . Then to expand the definition so as to allow  $n$  to be any *nonnegative* integer, we also define the **Euclidean metric on  $\mathbb{R}^0$**  trivially by  $d(0, 0) = 0$ .

In order to derive from the definition several essential properties of Euclidean distance, we shall exploit the algebraic structure of  $\mathbb{R}^n$  as a vector space, as described in subsection “Euclidean spaces” ([page 36](#)). We begin by observing that the distance—that is, the

Euclidean metric on  $\mathbb{R}^0$  is the zero metric on the Euclidean 0-space

norm  
norm!Euclidean  
norm!Euclidean  
Euclidean norm

Euclidean distance—from one point to another in  $\mathbb{R}^n$  can always be expressed as the distance from a point to the origin  $\mathbf{0} = \langle 0, 0, \dots, 0 \rangle \in \mathbb{R}^n$ . Indeed, given  $x = \langle x_1, x_2, \dots, x_n \rangle$  and  $y = \langle y_1, y_2, \dots, y_n \rangle$  in  $\mathbb{R}^n$ , we have  $x - y = \langle x_1 - y_1, x_2 - y_2, \dots, x_n - y_n \rangle$ , so that

$$d(x, y) = d(x - y, \mathbf{0}).$$

In view of that observation it is natural to consider for each point

$$x = \langle x_1, x_2, \dots, x_n \rangle \in \mathbb{R}^n$$

the **(Euclidean) norm of  $x$** , namely, the real number  $\|x\|$  defined by

$$\|x\| = d(x, \mathbf{0}),$$

or in terms of coordinates,

$$\|x\| = \left[ \sum_{i=1}^n x_i^2 \right]^{1/2}.$$

In dimensions  $n = 2$  and  $n = 3$ , the norm  $\|x\|$  is the usual “length” of the vector  $x$ . And in dimension  $n = 1$ , the norm  $\|x\| = |x|$ , the absolute value of  $x$ . Then the following lemma generalizes familiar properties of absolute value.

lem:norm-properties **1.3 Lemma (properties of Euclidean norm).** For all  $x \in \mathbb{R}^n$ ,  $y \in \mathbb{R}^n$ , and  $\lambda \in \mathbb{R}$ :

- lem-part:norm-nonneg (1)  $\|x\| \geq 0$ .
- lem-part:norm-strict-pos (2)  $\|x\| = 0$  if and only if  $x = \mathbf{0}$ .
- lem-part:norm-homog (3)  $\|\lambda x\| = |\lambda| \|x\|$ .
- lem-part:norm-triangle (4)  $\|x + y\| \leq \|x\| + \|y\|$ .

**Proof.** Property 1 is obvious, as is the “if” part of property 2. The “only if” part of 2 is an immediate consequence of the fact that if a sum of nonnegative numbers is 0, then each of the numbers must be 0.

(3) Write  $x = \langle x_1, x_2, \dots, x_n \rangle$ . Then  $\lambda x = \langle \lambda x_1, \lambda x_2, \dots, \lambda x_n \rangle$ , so that

$$\|\lambda x\| = \left[ \sum_{i=1}^n (\lambda x_i)^2 \right]^{1/2} = \left[ \lambda^2 \sum_{i=1}^n x_i^2 \right]^{1/2} = |\lambda| \left[ \sum_{i=1}^n x_i^2 \right]^{1/2} = |\lambda| \|x\|.$$

(4) By (1), both  $\|x + y\|$  and  $\|x\| + \|y\|$  are nonnegative. hence it will suffice to show that  $\|x + y\|^2 \leq (\|x\| + \|y\|)^2$ . Now

$$(\|x\| + \|y\|)^2 = \|x\|^2 + 2\|x\| \|y\| + \|y\|^2$$

and

$$\begin{aligned} \|x + y\|^2 &= \sum_{i=1}^n (x_i + y_i)^2 = \sum_{i=1}^n (x_i^2 + 2x_i y_i + y_i^2) \\ &= \sum_{i=1}^n x_i^2 + 2 \sum_{i=1}^n x_i y_i + \sum_{i=1}^n y_i^2 = \|x\|^2 + 2 \sum_{i=1}^n x_i y_i + \|y\|^2. \end{aligned}$$

Hence it remains only to show that  $\sum_{i=1}^n x_i y_i \leq \|x\| \|y\|$ . That will follow from the next lemma and hence complete the proof.

lem:CS-inequality

**1.4 Lemma (Cauchy–Schwarz Inequality).** Let  $x = \langle x_1, x_2, \dots, x_n \rangle$  and  $y = \langle y_1, y_2, \dots, y_n \rangle$  in  $\mathbb{R}^n$ . then

$$\left| \sum_{i=1}^n x_i y_i \right| \leq \|x\| \|y\|.$$

**Proof.** Set

$$\alpha = \sum_{i=1}^n x_i y_i.$$

We wish to show that  $|\alpha| \leq \|x\| \|y\|$ . Form the quadratic polynomial

$$p(\lambda) = (\|x\|^2) \lambda^2 + (2\alpha)\lambda + \|y\|^2$$

in  $\lambda$ . It will suffice to show that its discriminant

$$\delta = (2\alpha)^2 - 4\|x\|^2\|y\|^2$$

satisfies  $\delta \leq 0$ , for then  $\alpha^2 \leq \|x\|^2\|y\|^2$  and so  $|\alpha| \leq \|x\| \|y\|$ , as desired.

For each real number  $\lambda$ , the value  $p(\lambda)$  of  $p$  at  $\lambda$  satisfies

$$p(\lambda) = \lambda^2 \sum_{i=1}^n x_i^2 + 2\lambda \sum_{i=1}^n x_i y_i + \sum_{i=1}^n y_i^2 = \sum_{i=1}^n (\lambda x_i + y_i)^2 \geq 0.$$

Now if the discriminant  $\delta$  satisfied  $\delta > 0$ , then the quadratic  $p(\lambda)$  would have exactly two distinct real roots and consequently would assume negative as well as nonnegative values, contrary to the inequality  $p(\lambda) \geq 0$  just established. Hence  $\delta \leq 0$ , as required.  $\square$

The Cauchy–Schwarz Inequality (Lemma 1.4) just established completes the proof of Lemma 1.3.

**Remark.** Creating the quadratic polynomial  $p(\lambda)$  to prove the Cauchy–Schwarz inequality may seem a “trick”—like a magician’s “pulling a rabbit out of a hat.” However, once you realize that the three quantities  $\sum_{i=1}^n x_i y_i$ ,  $\|x\|^2 = \sum_{i=1}^n x_i^2$ , and  $\|y\|^2 = \sum_{i=1}^n y_i^2$  are themselves quadratic polynomials in the  $x_i$ ’s and  $y_i$ ’s, conjuring up a quadratic polynomial  $p(\lambda)$  is not at all unreasonable—especially if you realize that those three quantities appear in the discriminant of  $p(\lambda)$ .

Since  $d(x, y) = d(x - y, \mathbf{0})$ , then

$$d(x, y) = \|x - y\|$$

for all  $x, y \in \mathbb{R}^n$ .

Using this expression for the Euclidean distance in terms of the Euclidean norm, from Lemma 1.3 we can at last derive the properties of the Euclidean metric that will be important in the sequel.

prop:Euclidean-d-properties

**1.5 Proposition (properties of Euclidean metric).** Let  $d$  be the Euclidean metric on  $\mathbb{R}^n$ . Then for all points  $x, y, z \in \mathbb{R}^n$ :

prop-part:Euclidean-d-nonneg

$$(1) \quad d(x, y) \geq 0.$$

prop-part:Euclidean-d-strict-pos

$$(2) \quad d(x, y) = 0 \text{ if and only if } x = y.$$

prop-part:Euclidean-d-symmetric

$$(3) \quad d(x, y) = d(y, x).$$

prop-part:Euclidean-d-triangle

$$(4) \quad d(x, y) \leq d(x, z) + d(z, y).$$

norm!Euclidean

triangle inequality  
metric examples of

**Proof.** (2) Since  $d(x, y) = \|x - y\|$ , then from Lemma 1.3 2 we have  $d(x, y) = 0$  if and only if  $x - y = \mathbf{0}$ , that is  $x = y$ .

(3) From Lemma 1.3 3,

$$d(x, y) = \|x - y\| = \|(-1)(y - x)\| = |-1| \|y - x\| = d(y, x).$$

(4) From Lemma 1.3 4,

$$\begin{aligned} d(x, z) &= \|x - z\| = \|(x - y) + (y - z)\| \\ &\leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z). \quad \square \end{aligned}$$



Figure 1.1: The triangle inequality.

fig:triangle-inequality

Property 1.5 2 says that the distance from one point to another is zero precisely when the two points are actually the same. Property 1.5 3 says that the distance from one point to a second is the same as the distance from the second to the first. Property 1.5 4, called the **triangle inequality**, says that the length of one side of a triangle is at most the sum of the lengths of the other two sides (see Figure 1.1). These properties may seem geometrically obvious, so was all the work to prove them really necessary? Yes, because our definition of Euclidean distance was purely algebraic and did not invoke any prior geometric meaning of distance.

### Other metrics on $\mathbb{R}^n$

Let us return for a moment to the plane  $\mathbb{R}^2$ . A crow flying in a straight line from one point to another would want to use the Euclidean metric to compute how far he travels. But the driver of a taxicab cruising a city's streets laid out in a rectangular grid (such as downtown Philadelphia, PA) would not. Instead, she would total the distance she travels on north-south streets and the distance she travels on east-west streets. (Just as we implicitly assumed that the portion of the earth's surface the crow flies over is nearly planar, so we now assume that the streets the taxicab cruises are not one-way.) By superimposing a rectangular coordinate system on the street grid with axes parallel to the streets (and origin at City Hall, say, as in the case of Philadelphia), we see that the taxicab driver would compute the distance from  $x = \langle x_1, y_1 \rangle$  to  $y = \langle y_1, y_2 \rangle$  to be

$$|x_1 - y_1| + |x_2 - y_2|$$

(see Figure 1.2).

This discussion suggests a different definition of distance in not just  $\mathbb{R}^2$  but more generally in  $\mathbb{R}^n$ .

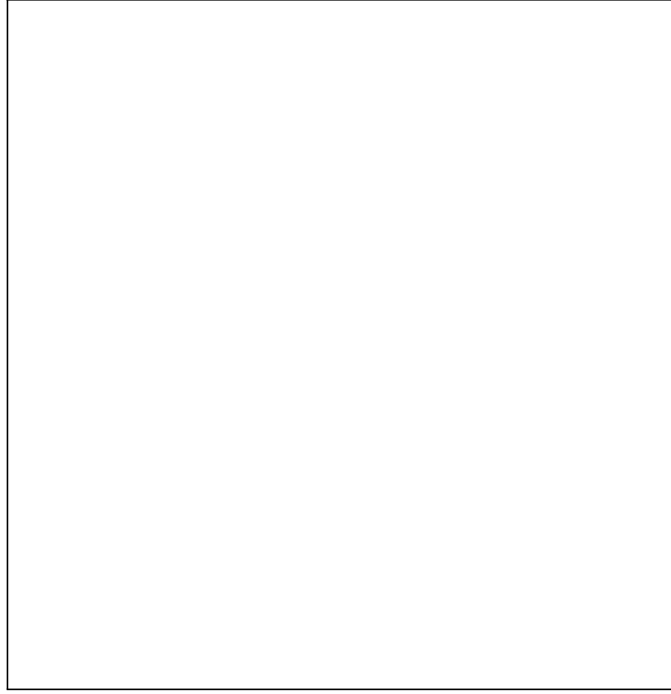


Figure 1.2: Taxicab distance.

fig:taxicab-dist

def:taxicab-metric

**1.6 Definition.** For points  $x = \langle x_1, x_2, \dots, x_n \rangle$  and  $y = \langle y_1, y_2, \dots, y_n \rangle$  in  $\mathbb{R}^n$ , the **taxicab distance from  $x$  to  $y$**  is defined to be the real number

$$\sum_{i=1}^n |x_i - y_i|.$$

The function

$$d_1 : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

assigning to each ordered pair  $(x, y)$  of points in  $\mathbb{R}^n$  the taxicab distance from  $x$  to  $y$  is called the **taxicab metric on  $\mathbb{R}^n$** .

The taxicab metric  $d_1$  has the very same properties listed in [Proposition 1.5](#), namely, for all  $x, y, z \in \mathbb{R}^n$ :

- |                         |  |
|-------------------------|--|
| prop-part:d1-nonneg     | (1) $d_1(x, y) \geq 0$ .                     |
| prop-part:d1-strict-pos | (2) $d_1(x, y) = 0$ if and only if $x = y$ . |
| prop-part:d1-symmetric  | (3) $d_1(x, y) = d_1(y, x)$ .                |
| prop-part:d1-triangle   | (4) $d_1(x, y) \leq d_1(x, z) + d_1(z, y)$ . |

The proofs of (1)–(4) here are all quite easy. For example, here is the proof of (4). Since for each  $i = 1, 2, \dots, n$  we have

$$|x_i - z_i| = |(x_i - y_i) + (y_i - z_i)| \leq |x_i - y_i| + |y_i - z_i|,$$

then

$$\begin{aligned} d_1(x, z) &= \sum_{i=1}^n |x_i - z_i| \leq \sum_{i=1}^n (|x_i - y_i| + |y_i - z_i|) \\ &= \sum_{i=1}^n |x_i - y_i| + \sum_{i=1}^n |y_i - z_i| = d_1(x, y) + d_1(y, z). \quad \square \end{aligned}$$

Already in dimension  $n = 2$  the Euclidean metric  $d$  and the taxicab metric  $d_1$  differ—for example:

$$d(\langle 1, 1 \rangle, \langle 0, 0 \rangle) = \sqrt{2} \neq 2 = d_1(\langle 1, 1 \rangle, \langle 0, 0 \rangle)$$

In dimension  $n = 1$ , of course, they coincide:

$$d(x, y) = d_1(x, y) \quad (x, y \in \mathbb{R})$$

Thus both Euclidean distance and taxicab distance generalize to arbitrary dimension  $n$  the usual distance in dimension 1. Next comes yet a third generalization to dimension  $n$ .

def:max-metric

**1.7 Definition.** For points  $x = \langle x_1, x_2, \dots, x_n \rangle$  and  $y = \langle y_1, y_2, \dots, y_n \rangle$  in  $\mathbb{R}^n$ , the **max distance from  $x$  to  $y$**  is defined to be the real number

$$\max_{1 \leq i \leq n} |x_i - y_i|.$$

The function

$$d_\infty: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

assigning to each ordered pair  $(x, y)$  of points in  $\mathbb{R}^n$  the max distance from  $x$  to  $y$  is called the **max metric on  $\mathbb{R}^n$** .

The max distance in  $\mathbb{R}^n$  does indeed generalize the usual distance in  $\mathbb{R}$ , for when  $n = 1$ , each point in  $\mathbb{R}^n$  has a single coordinate, and then  $d_\infty(x, y) = |x - y|$ .

The max metric has the same four properties that the Euclidean and taxicab metrics share:

- |                           |   |
|---------------------------|---|
| prop-part:dmax-nonneg     | (1) $d_\infty(x, y) \geq 0$ .                               |
| prop-part:dmax-strict-pos | (2) $d_\infty(x, y) = 0$ if and only if $x = y$ .           |
| prop-part:dmax-symmetric  | (3) $d_\infty(x, y) = d_\infty(y, x)$ .                     |
| prop-part:dmax-triangle   | (4) $d_\infty(x, y) \leq d_\infty(x, y) + d_\infty(y, z)$ . |

This time let us write out the proofs of both (2) and (4).

If  $d_\infty(x, y) = 0$ , then  $0 \leq |x_i - y_i| \leq 0$ , that is,  $x_i = y_i$ , for each  $i = 1, 2, \dots, n$ , and hence  $x = y$ . Conversely, if  $x = y$ , then  $|x_i - y_i| = 0$  for each  $i$ , and so  $d_\infty(x, y) = 0$ . This proves (2).

To prove (4), from  $|x_i - z_i| \leq |x_i - y_i| + |y_i - z_i|$  for each  $i$  we deduce

$$\begin{aligned} d_\infty(x, z) &= \max_{1 \leq i \leq n} |x_i - z_i| \\ &\leq \max_{1 \leq i \leq n} (|x_i - y_i| + |y_i - z_i|) \\ &\leq \max_{1 \leq i \leq n} |x_i - y_i| + \max_{1 \leq i \leq n} |y_i - z_i| \\ &= d_\infty(x, y) + d_\infty(y, z). \quad \square \end{aligned}$$

In his short story *A Day's Wait*, Ernest Hemingway tells of the ailing boy Schatz waiting all day to die. The doctor has reported the boy's temperature to be  $102^\circ$ , and Schatz has remembered his school companions in France telling him that one cannot live with



a temperature over  $44^\circ$ . He is relieved only after his father explains that just as there are kilometers and miles, so there are different temperature scales: the doctor had used a Fahrenheit thermometer, not a Celsius thermometer!

sup metric  
function space

If you have grown up computing distance always by the “square-root of the sum of the squares of the differences of the coordinates” formula, you may well believe that only the Euclidean metric gives the “true” distance. To be sure, there are contexts in which the Euclidean metric is the only appropriate measure of distance in  $\mathbb{R}^n$  (see [Exercise 4](#)). Nonetheless, as we shall demonstrate later in this chapter, for purposes of defining continuity the taxicab and max metrics are just as good as the Euclidean metric.

*Note:* If you are impatient to get to the definition of ‘metric’, you may wish to jump ahead to the [subsection “Metrics and metric spaces”](#) (page 139), but then be sure to return to the following subsection.

### Metrics on function spaces

Distances can also be measured in certain “spaces” whose “points” are in fact functions.

subsec:metrics-fn-and-seq-spaces

ex:sup-metric-on-functions

**1.8 Example (sup metric on  $C(X)$ ).** Let  $X = C([0, 1])$  be the set of all continuous real-valued functions on the closed interval  $[0, 1]$ ; the ordinary calculus sense of ‘continuous’ is meant here. Contrary to the usual practice in calculus, typical functions belonging to  $X$  will be denoted here by the letters  $x$ ,  $y$ , and  $z$ , whereas typical members of their domain  $[0, 1]$  will be denoted by  $t$  and  $s$ .

We shall assume here, without proving it yet, the extreme value theorem from calculus: a continuous real-valued function on  $[0, 1]$  must attain a maximum value. (We shall prove the extreme value theorem later, in a more general setting: see [Corollary 4.27](#).)

Given functions

$$x: [0, 1] \rightarrow \mathbb{R}, \quad y: [0, 1] \rightarrow \mathbb{R}$$

belonging to  $X$ , the function  $t \mapsto |x(t) - y(t)|$  is continuous and consequently attains a maximum value on  $[0, 1]$ ; we define  $d_\infty(x, y)$  to be this maximum value, so that

$$d_\infty(x, y) = \max_{0 \leq t \leq 1} |x(t) - y(t)|.$$

Thus  $d_\infty(x, y)$  is the greatest vertical distance between the graphs of  $x$  and  $y$  (see [Figure 1.3](#)).



Figure 1.3: Max metric distance between two functions.

fig:max-metric-dist-fns

max metric!mainref  
 sup metric  
 sup metric  
 L1 metric@\$L\_1\$-metric  
 function space

The map  $d_\infty : X \times X \rightarrow \mathbb{R}$  is called the **max metric on  $C([0, 1])$** . Note that

$$d_\infty(x, y) = \sup\{|x(t) - y(t)| : 0 \leq t \leq 1\} \quad (x, y \in C([0, 1])).$$

For that reason we also refer to  $d_\infty$  as the **sup metric**—even though the supremum is actually attained as a maximum. [But the notation  $d_\infty$  is the same as that for the max metric on  $\mathbb{R}^n$  (Definition 1.7).]

The sup metric  $d_\infty$  has the properties:

- prop-part:dmax-fns-nonneg (1)  $d_\infty(x, y) \geq 0$ .  
 prop-part:dmax-fns-strict-pos (2)  $d_\infty(x, y) = 0$  if and only if  $x = y$ .  
 prop-part:dmax-fns-symmetric (3)  $d_\infty(x, y) = d_\infty(y, x)$ .  
 prop-part:dmax-fns-triangle (4)  $d_\infty(x, y) \leq d_\infty(x, z) + d_\infty(z, y)$ .

We prove only property (4). For each  $t \in [0, 1]$ ,

$$\begin{aligned} |x(t) - z(t)| &\leq |x(t) - y(t)| + |y(t) - z(t)| \\ &\leq \max_{0 \leq s \leq 1} |x(s) - y(s)| + \max_{0 \leq s \leq 1} |y(s) - z(s)| \\ &= d_\infty(x, y) + d_\infty(y, z), \end{aligned}$$

and so

$$d_\infty(x, z) = \max_{0 \leq t \leq 1} |x(t) - z(t)| \leq d_\infty(x, y) + d_\infty(y, z) \quad \diamond$$

ex:L1-metric-on-functions **1.9 Example ( $L^1$ -metric).** Again let  $X$  be the set of all continuous functions  $x : [0, 1] \rightarrow \mathbb{R}$ . For  $x, y \in X$  define:

$$d_1(x, y) = \int_0^1 |x(t) - y(t)| dt$$

That is meaningful because the function  $t \mapsto |x(t) - y(t)|$ , being continuous on  $[0, 1]$ , is integrable there. Thus  $d_1(x, y)$  is the area of the plane region enclosed by the graphs of  $x$  and  $y$  (see Figure 1.4), which is a reasonable measure of how far apart  $x$  and  $y$  are.

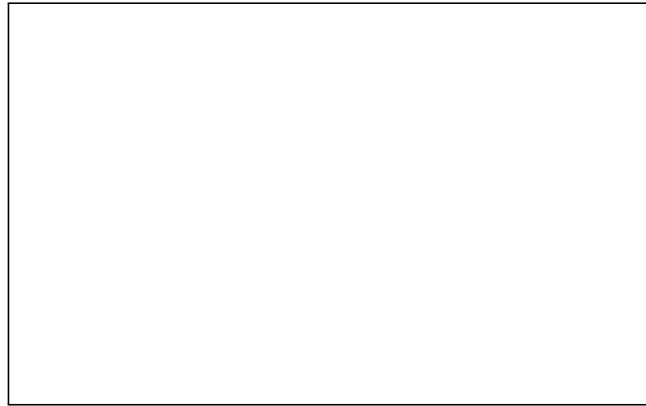


Figure 1.4: The  $d_1$ -distance between functions as area.

fig:d1-dist-between-functions

Note that  $d_1 \neq d_\infty$ . For example, if  $x(t) = t$  and  $y(t) = 0$  for all  $t \in [0, 1]$ , then

$$d_1(x, y) = \int_0^1 t dt = 1/2 \neq 1 = \max_{0 \leq t \leq 1} t = d_\infty(x, y).$$

By now it should come as no surprise that for all  $x, y, z \in X$ :

prop-part:d1-fns-nonneg	(1) $d_1(x, y) \geq 0$ .
prop-part:d1-fns-strict-pos	(2) $d_1(x, y) = 0$ if and only if $x = y$ .
prop-part:d1-fns-symmetric	(3) $d_1(x, y) = d_1(y, x)$ .
prop-part:d1-fns-triangle	(4) $d_1(x, y) \leq d_1(x, z) + d_1(z, y)$ .

square-summable sequence

This time we give a detailed proof just of the “only if” part of (2). Suppose  $d_1(x, y) = 0$ . Then

$$\int_0^1 f(t) dt = 0,$$

where  $f$  is the continuous real-valued function on  $[0, 1]$  given by

$$f(t) = |x(t) - y(t)|.$$

To show that  $x = y$  we shall show, equivalently, that  $f(t) = 0$  for all  $t \in [0, 1]$ . Now on the interval  $[0, 1]$  the function  $f$  has an antiderivative  $g$ , and the Fundamental Theorem of Calculus tells us that

$$\int_0^t f(s) ds = g(t) - g(0) \quad (0 \leq t \leq 1).$$

Since  $f$  is nonnegative, for each  $0 \leq t \leq 1$  we have

$$0 \leq \int_0^t f(s) ds \leq \int_0^t f(s) ds + \int_t^1 f(s) ds = \int_0^1 f(s) ds = 0,$$

so that

$$0 = \int_0^t f(s) ds$$

and hence

$$0 = g(t) - g(0).$$

Thus  $g$  is constant on the interval  $[0, 1]$ . But then  $f$ , the derivative of  $g$ , is identically zero on this interval.  $\diamond$

The next example will be an “infinite-dimensional” analog of the finite-dimensional Euclidean spaces  $\mathbb{R}^n$ . It will be a space consisting of certain sequences of real numbers in which, by analogy with Euclidean distance, the distance from a sequence  $x = \langle x_i \rangle_{i=1,2,3,\dots}$  to a sequence  $y = \langle y_i \rangle_{i=1,2,3,\dots}$  will be defined by

$$d_2(x, y) = \left( \sum_{i=1}^{\infty} (x_i - y_i)^2 \right)^{1/2}.$$

Of course this formula does not have meaning for arbitrary  $x$  and  $y$  because the infinite series on its right-hand side need not converge. So we must restrict attention to just some sequences. Now our space will include the special sequence  $\mathbf{0}$  all of whose entries are the number 0. Then we must impose the restriction that any other point  $x = \langle x_i \rangle_{i=1,2,3,\dots}$  of our space be one for which

$$d_2(x, \mathbf{0}) = \left( \sum_{i=1}^{\infty} x_i^2 \right)^{1/2}$$

has meaning—as one says, that the sequence  $x = \langle x_i \rangle_{i=1,2,3,\dots}$  be “square-summable”. It will turn out that no further restrictions are needed: the formula for  $d_2(x, y)$  will be meaningful for any two sequences  $x$  and  $y$  that are square-summable in this sense.

ex: hilbert sequence space  
 Hilbert sequence space  
 Hilbert, David  
 square-summable sequence  
 Hilbert norm  
 norm  
 Hilbert distance  
 Hilbert metric

**1.10 Example (Hilbert metric).** The **Hilbert sequence space**, denoted by  $\ell^2$  (read “little ell two”), is the set of all sequences  $x = \langle x_i \rangle_{i=1,2,3,\dots}$  that are *square-summable* in the sense that the series  $\sum_{i=1}^{\infty} x_i^2$  converges.

For example, the sequence  $\langle 1/i \rangle_{i=1,2,3,\dots} = \langle 1, 1/2, 1/3, \dots \rangle \in \ell^2$ , but the alternating sequence  $\langle (-1)^i / \sqrt{i} \rangle_{i=1,2,3,\dots} = \langle -1, +1/\sqrt{2}, -1/\sqrt{3}, \dots \rangle \notin \ell^2$ . Trivially, the zero sequence  $\mathbf{0} = \langle 0, 0, \dots, 0, \dots \rangle \in \ell^2$ .

By analogy with the Euclidean distance  $(\sum_{i=1}^n (x_i - y_i)^2)^{1/2}$  in  $\mathbb{R}^n$ , we would like to define the distance from a sequence  $x = \langle x_i \rangle_{i=1,2,3,\dots}$  to a sequence  $y = \langle y_i \rangle_{i=1,2,3,\dots}$  in  $\ell^2$  to be  $[\sum_{i=1}^{\infty} (x_i - y_i)^2]^{1/2}$ . To do that, however, we first need to know that the difference sequence  $\langle x_i - y_i \rangle_{i=1,2,3,\dots}$  actually belongs to  $\ell^2$ . Moreover, as in the case with Euclidean distance, in order to establish properties of distance in  $\ell^2$ , we shall want to reduce matters to considering distances just from the zero element, in other words, to introduce a notion of ‘norm’ in  $\ell^2$ . Hence we shall begin with the norm and only then return to distance in  $\ell^2$ .

By analogy with the Euclidean norm, the **Hilbert norm** of an  $x = \langle x_i \rangle_{i=1,2,3,\dots} \in \ell^2$  is defined to be the real number

$$\|x\|_2 = \left[ \sum_{i=1}^{\infty} x_i^2 \right]^{1/2}.$$

(The definition makes sense because, by definition of  $\ell^2$ , the sequence is square-summable.)

Clearly

$$\|x\|_2 \geq 0, \quad \|x\|_2 = 0 \iff x = \mathbf{0}$$

for each  $x \in \ell^2$ . Moreover, if  $x = \langle x_i \rangle_{i=1,2,3,\dots} \in \ell^2$  and  $\lambda \in \mathbb{R}$ , then the sequence

$$\lambda x = \langle \lambda x_i \rangle_{i=1,2,3,\dots}$$

also belongs to  $\ell^2$  and satisfies

$$\|\lambda x\|_2 = |\lambda| \|x\|_2.$$

Let  $x = \langle x_i \rangle_{i=1,2,3,\dots} \in \ell^2$  and  $y = \langle y_i \rangle_{i=1,2,3,\dots} \in \ell^2$ . For each  $n \geq 1$ , the inequality 1.3 (4) concerning norms in  $\mathbb{R}^n$  says that

$$\left[ \sum_{i=1}^n (x_i + y_i)^2 \right]^{1/2} \leq \left[ \sum_{i=1}^n x_i^2 \right]^{1/2} + \left[ \sum_{i=1}^n y_i^2 \right]^{1/2} \leq \left[ \sum_{i=1}^{\infty} x_i^2 \right]^{1/2} + \left[ \sum_{i=1}^{\infty} y_i^2 \right]^{1/2} = \|x\|_2 + \|y\|_2.$$

This proves two things: First, since all the partial sums of the series  $\sum_{i=1}^{\infty} (x_i + y_i)^2$  are bounded above by the number  $(\|x\|_2 + \|y\|_2)^2$ , this series converges [see Examples 1.61 (2)]; hence the sum sequence

$$x + y = \langle x_i + y_i \rangle_{i=1,2,3,\dots}$$

belongs to  $\ell^2$ . Second, that sum sequence satisfies

$$\|x + y\|_2 \leq \|x\|_2 + \|y\|_2.$$

Given  $x = \langle x_i \rangle_{i=1,2,3,\dots} \in \ell^2$  and  $y = \langle y_i \rangle_{i=1,2,3,\dots} \in \ell^2$ , the sequence

$$x - y = x + (-1)y = \langle x_i - y_i \rangle_{i=1,2,3,\dots}$$

also belongs to  $\ell^2$ , and so we may define the **Hilbert distance** from  $x$  to  $y$  to be the real number

$$d_2(x, y) = \|x - y\|_2 = \left[ \sum_{i=1}^{\infty} (x_i - y_i)^2 \right]^{1/2}.$$

The map  $d_2: \ell^2 \times \ell^2 \rightarrow \mathbb{R}$  so defined is the **Hilbert metric** on  $\ell^2$ .

We know now that the four properties of the Euclidean norm on  $\mathbb{R}^n$  listed in Lemma 1.3 hold as well for the Hilbert norm on  $\ell^2$ . Hence the same proof used to deduce Proposition 1.5 establishes that for  $x, y, z \in \ell^2$ :

- (1)  $d_2(x, y) \geq 0$ .
- (2)  $d_2(x, y) = 0$  if and only if  $x = y$ .
- (3)  $d_2(x, y) = d_2(y, x)$ .
- (4)  $d_2(x, y) \leq d_2(x, z) + d_2(y, z)$ .

Hilbert cube  
Hilbert sequence space  
metric!examples of  
Fréchet, Maurice  
triangle inequality  
metric!examples of

Of particular interest later will be the subset

$$\ell^\infty = \{x \in \ell^2 : |x_i| \leq 1/i \text{ for each } i = 1, 2, \dots\},$$

of  $\ell^2$ , known as the **Hilbert cube**. (See [Exercise 36](#).)  $\diamond$

### Metrics and metric spaces

It is time to let the “secret” out. All the preceding examples—including the Euclidean metric  $d$  on  $\mathbb{R}^n$ , the max metric  $d_\infty$  on  $\mathbb{R}^n$ , the sup metric  $d_\infty$  on  $C([0, 1])$ , and the Hilbert metric  $d_2$  on the Hilbert sequence space  $\ell^2$ —are instances of a general notion first defined by Maurice Fréchet in 1906.

**1.11 Definition.** A **metric** on a set  $X$  is a function

$$d: X \times X \rightarrow \mathbb{R}$$

such that for all  $x, y, z \in X$ :

- property:metric-nonnegative (M1) *Nonnegativity:*  $d(x, y) \geq 0$ .
- property:metric-positive (M2) *Positivity:*  $d(x, y) = 0$  if and only if  $x = y$ .
- property:metric-symmetric (M3) *Symmetry:*  $d(x, y) = d(y, x)$ .
- property:metric-triangle-inequality (M4) *Triangularity:*  $d(x, z) \leq d(x, y) + d(y, z)$ .

If  $x, y \in X$ , then  $d(x, y)$  is called the  **$d$ -distance from  $x$  to  $y$** .

A set  $X$  together with a metric  $d$  on  $X$  comprise a pair  $\langle X, d \rangle$ , called a **metric space**.

The inequality in (M4) is known as **the triangle inequality**.

Property (M2) asserts two things: first, the equality

$$d(x, x) = 0$$

for each  $x$ , and second, in view of (M1), the strict inequality

$$x \neq y \implies d(x, y) > 0.$$

When we are showing that a particular  $d$  is actually a metric, it is ordinarily this strict inequality and triangularity (M4) that cause any trouble; usually we do not even bother to mention that the remaining properties hold.

Different metrics  $d$  and  $d'$  on the same set  $X$  give rise to different metric spaces  $\langle X, d \rangle$  and  $\langle X, d' \rangle$ . For example, when  $n > 1$ , the Euclidean metric  $d$ , the taxicab metric  $d_1$ , and the max metric  $d_\infty$  on  $\mathbb{R}^n$  yield three distinct metric spaces  $\langle \mathbb{R}^n, d \rangle$ ,  $\langle \mathbb{R}^n, d_1 \rangle$ , and  $\langle \mathbb{R}^n, d_\infty \rangle$ .

We close this section with additional examples of metric spaces.

metric!discrete-  
ex:discrete-metric  
discrete metric  
metric!discrete  
discrete metric  
Euclidean 0-space  
Euclidean metric  
metric!Euclidean  
metric!on subset  
metric!induced  
induced metric  
Euclidean metric  
max metric  
metric!on product

**1.12 Example (discrete metric).** Let  $X$  be any set. Then the function  $\delta: X \times X \rightarrow \mathbb{R}$  defined by

$$\delta(x, y) = \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x \neq y \end{cases}$$

is a metric on  $X$ , called the **discrete metric**. The triangle inequality

$$\delta(x, z) \leq \delta(x, y) + \delta(y, z)$$

certainly holds in case  $x = y$  and  $y = z$ , for then both sides are 0; it holds as well in case  $x \neq y$  or  $y \neq z$ , for then  $\delta(x, y) + \delta(y, z) \geq 1$  whereas  $\delta(x, z) \leq 1$ .

The adjective ‘discrete’ is applied to  $\delta$  so as to suggest that this metric isolates distinct points. Indeed, for  $0 < \varepsilon \leq 1$ , the only point of  $X$  whose  $\delta$ -distance from an  $x \in X$  is less than  $\varepsilon$  is  $x$  itself.

This example shows that any set, no matter what the nature of its elements, can be made into a metric space by endowing it with its discrete metric. Of course, a particular set— $\mathbb{R}^n$ , for example, when  $n > 0$ —may well have metrics much more interesting than the discrete one.

On a 1-element set, the discrete metric is the only metric. In particular, on Euclidean 0-space  $\mathbb{R}^0 = \{0\}$  the unique metric is the Euclidean metric  $d$  given by  $d(0, 0) = 0$ .  $\diamond$

ex:induced-d **1.13 Example (induced metric on subset).** Given a metric space  $\langle X, d \rangle$  and a subset  $Y$  of  $X$ , we may use  $d$  to measure distances just between points of  $Y$  by forming the restriction

$$d' = d|_{Y \times Y}: Y \times Y \rightarrow \mathbb{R}$$

of  $d: X \times X \rightarrow \mathbb{R}$  to the new domain  $Y \times Y$ . Since  $d'(x, y) = d(x, y)$  for all  $x, y \in Y$ , the function  $d'$  is clearly a metric on  $Y$ . This metric  $d'$  on  $Y$  is said to be **induced by  $d$** .

When  $X = \mathbb{R}^n$  and  $d$  is the Euclidean metric on  $\mathbb{R}^n$ , then the induced metric  $d'$  on  $Y \subset X$  is called the **Euclidean metric on  $Y$** , and  $Y$  becomes a metric space in its own right when provided with  $d'$ .  $\diamond$

ex:max-metric-on-product **1.14 Example (max metric on finite product).** Let  $\langle X_1, d_1 \rangle, \langle X_2, d_2 \rangle, \dots, \langle X_n, d_n \rangle$  be finitely many metric spaces. On the product set

$$X = X_1 \times X_2 \times \cdots \times X_n$$

the **max metric  $d_\infty$  induced by  $\langle d_1, d_2, \dots, d_n \rangle$**  is obtained by setting

$$d_\infty(x, y) = \max_{1 \leq i \leq n} d_i(x_i, y_i)$$

for any two points  $x = \langle x_1, x_2, \dots, x_n \rangle$  and  $y = \langle y_1, y_2, \dots, y_n \rangle$  of  $X$ .

When  $X_1 = X_2 = \cdots = X_n = \mathbb{R}$  and  $d_1 = d_2 = \cdots = d_n =$  the Euclidean metric on  $\mathbb{R}$ , then  $d_\infty$  is the very same max metric on  $\mathbb{R}^n$  defined earlier (Definition 1.7). Of course the Euclidean and taxicab metrics on  $\mathbb{R}^n$  likewise suggest constructing metrics on a product of metric spaces (Exercise 12), but these metrics are less convenient to work with than the one above.  $\diamond$

ex:sequence-space-metric **1.15 Example.** Let  $X$  be the set of all sequences  $x = \langle x_i \rangle_{i \in \mathbb{N}}$  in a given nonempty set  $S$ . We want to regard two points  $x = \langle x_i \rangle_{i \in \mathbb{N}}$  and  $y = \langle y_i \rangle_{i \in \mathbb{N}}$  of  $X$  to be close together if for some large  $n$  their first  $n$  coordinates  $x_0, x_1, \dots, x_{n-1}$  and  $y_0, y_1, \dots, y_{n-1}$  agree; the larger

the value of  $n$ , the closer together the points are to be. To accomplish this we define a distance function

$$\{eq:metric-on-sequences\} (*) \quad d(x, y) = \begin{cases} 0 & \text{if } x = y, \\ \frac{1}{1 + \min\{i : x_i \neq y_i\}} & \text{if } x \neq y. \end{cases}$$

binary sequence  
bounded function  
function!bounded  
sup norm  
sup metric  
metric!examples of

The triangle inequality

$$d(x, z) \leq d(x, y) + d(y, z)$$

certainly holds if  $x = z$  or  $x = y$  or  $y = z$ . To verify it when  $x \neq z$ ,  $x \neq y$ , and  $y \neq z$ , set

$$j = \min\{i : x_i \neq y_i\}, \quad k = \min\{i : y_i \neq z_i\}, \quad m = \min\{i : x_i \neq z_i\}.$$

For  $i \in \mathbb{N}$ , of course

$$x_i = y_i \quad \text{and} \quad y_i = z_i \implies x_i = z_i.$$

Then  $m \geq \min\{j, k\}$ , so that

$$\frac{1}{1+m} \leq \frac{1}{1+j} + \frac{1}{1+k}.$$

The sequences just considered have index origin 0, that is, their indices are  $0, 1, 2, \dots$ . For sequences having index origin 1, that is, whose indices are  $1, 2, 3, \dots$ , the same formula (\*) still defines a metric on the set of such sequences.

A notable special case of this example occurs when  $S$  is the two-element set  $\{0, 1\}$ , for then  $X$  is the set of all *binary sequences*, that is, all sequences whose entries are the bits 0 or 1.  $\diamond$

The preceding example of a metric on a set of sequences is different from the metric  $d_2$  on the Hilbert sequence space  $\ell^2$  in [Example 1.10](#). [If you previously skipped the subsection “Metrics on function spaces” ([page 135](#)), you should read it now!]

**1.16 Example (sup metric on  $\mathcal{B}(X)$ ).** Let  $X$  be a nonempty set. A real-valued function  $f: X \rightarrow \mathbb{R}$  is said to be **bounded** when there is some constant  $c$  such that  $|f(x)| \leq c$  for all  $x \in X$ . The collection of all bounded real-valued functions on  $X$  is denoted by  $\mathcal{B}(X)$ .

For  $f \in \mathcal{B}(X)$ , the set  $\{|f(x)| : x \in X\}$  is nonempty and bounded in  $\mathbb{R}$  (for the usual metric of  $\mathbb{R}$ ), and so we may define

$$\|f\|_\infty = \sup\{|f(x)| : x \in X\}.$$

The map  $f \mapsto \|f\|_\infty$  of  $\mathcal{B}(X) \rightarrow \mathbb{R}$  is called the **sup norm on  $X$** . It shares the characteristic properties of the norms previously encountered:

$$\begin{aligned} \|f\|_\infty &\geq 0, \\ \|f\|_\infty &= 0 \text{ if and only if } f = 0, \\ \|f + g\|_\infty &\leq \|f\|_\infty + \|g\|_\infty, \\ \|kf\|_\infty &= |k| \|f\|_\infty, \end{aligned}$$

for all  $f, g \in \mathcal{B}(X)$  and all constants  $k \in \mathbb{R}$ .

Just as with the norms previously encountered, so with the sup norm the formula

$$d_\infty(f, g) = \|(\|_\infty f - g) = \sup\{|f(x) - g(x)| : x \in X\}$$

defines a metric—the **sup metric**—on  $\mathcal{B}(X)$ .  $\diamond$

translation

translation vector

rotation

angle of rotation

translation-invariant metric

max norm

taxicab norm

**EXERCISES FOR SECTION 1.1**

1. Verify that for all  $x, y \in \mathbb{R}^n$  the identity

$$\|x\|^2 \|y\|^2 - \left(\sum_{i=1}^n x_i y_i\right)^2 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (x_i y_j - x_j y_i)^2$$

holds and from it derive anew the the Cauchy–Schwarz Inequality.

2. (a) For which pairs of points  $x$  and  $y$  of  $\mathbb{R}^n$  does  $\|x + y\| = \|x\| + \|y\|$ ? [Hint: Examine the proofs of Lemma 1.3 and the Cauchy–Schwarz Inequality.]  
 (b) Use (a) to determine when equality actually holds in the triangle inequality for the Euclidean metric on  $\mathbb{R}^n$ .

3. Verify the inequalities

$$d(x, y) \leq d_1(x, y) \leq \sqrt{n} d(x, y)$$

relating the Euclidean metric  $d$  and the taxicab metric  $d_1$  on  $\mathbb{R}^n$ .

4. A **translation** of  $\mathbb{R}^2$  is a map  $T_c: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  of the form

$$T_c(x) = x + c$$

for some fixed  $c \in \mathbb{R}^2$ —the *translation vector*.

A **rotation** of  $\mathbb{R}^2$  (around the origin) is a map  $R_\theta: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  of the form

$$R(x_1, x_2) = \langle x_1 \cos \theta - x_2 \sin \theta, x_1 \sin \theta + x_2 \cos \theta \rangle$$

for some fixed real number  $\theta$ —the *angle of rotation*. A rotation around any other point  $p$  may be obtained as a composition  $T_p \circ R_\theta \circ T_{-p}$ .

Note: Each such rotation  $R_\theta$  of  $\mathbb{R}^2$  around the origin is uniquely determined by a  $2 \times 2$  matrix  $M_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ ; the corresponding rotation map  $R$  is given by  $R_\theta\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = M_\theta \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ , where the dot denotes matrix multiplication.

- (a) Show that the Euclidean metric  $d$  on  $\mathbb{R}^2$  is **translation-invariant** in the sense that

$$d(T(x), T(y)) = d(x, y) \quad (x, y \in \mathbb{R}^2)$$

for every translation  $T$  of  $\mathbb{R}^2$ , and that  $d$  is **rotation-invariant** in the sense that

$$d(R(x), R(y)) = d(x, y) \quad (x, y \in \mathbb{R}^2)$$

for every rotation  $R$  of  $\mathbb{R}^2$ .

- (b) Are the taxicab and max metrics on  $\mathbb{R}^2$  also translation-invariant? rotation-invariant?

5. (a) The **max norm** on  $\mathbb{R}^n$  is the map  $x \mapsto \|x\|_\infty$  of  $\mathbb{R}^n$  to  $\mathbb{R}$  defined by

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Which of the properties in Lemma 1.3 hold when the max norm is used instead of the Euclidean norm?

- (b) The **taxicab norm** on  $\mathbb{R}^n$  is the map  $x \mapsto \|x\|_1$  of  $\mathbb{R}^n$  to  $\mathbb{R}$  defined by

$$\|x\|_1 = \sum_{1 \leq i \leq n} |x_i|.$$

Which of the properties in Lemma 1.3 hold when the the taxicab norm is used instead of the Euclidean norm?



defines a metric on  $J^\infty$ . *Note:* This example is generalized in [Exercise 51](#); that generalization is studied further in [Proposition 3.51](#).

prob:sup-metric-sequences

pseudometric

pseudometric

pseudometric!and metric

indiscrete pseudometric

17. Denote by  $\ell^\infty$  the set of all *bounded* sequences of real numbers. For  $x = \langle x_n \rangle_{n=1}^\infty$  and  $y = \langle y_n \rangle_{n=1}^\infty$  such sequences, let

$$d_\infty(x, y) = \sup_{n=1,2,3,\dots} |x_n - y_n|$$

Show that the function  $d_\infty$  is a metric on  $\ell^\infty$ .

(Hint: You may wish first to define the “sup norm”

$$\|x\|_\infty = \sup_{n=1,2,3,\dots} |x_n|$$

and to verify that it has the corresponding properties to those for the Euclidean norm that are listed in [Lemma 1.3](#); and then express  $d_\infty$  in terms of the sup norm.)

Instead of real-valued bounded sequences, we may consider complex-valued bounded sequences, with absolute value replaced by modulus.

prob:TGS-max-2-metrics

18. If  $d_1$  and  $d_2$  are metrics on a set  $X$ , show that the formula

$$d(x, y) = \max\{d_1(x, y), d_2(x, y)\}$$

defines a metric on  $X$ .

prob:TGS-metric-on-tangent-disk

19. Fix  $r > 0$  and  $x \in \mathbb{R}$ , and let  $N = \{\langle x, 0 \rangle\} \cup B_r(\langle x, 0 \rangle; d)$  where  $d$  is the Euclidean metric.

prob-part:prep-def-rho

- (a) Let  $z = \langle u, v \rangle \in N$  with  $z \neq \langle x, 0 \rangle$ . Construct a unique  $\varepsilon$  with  $0 < \varepsilon < r$  for which  $z \in S_\varepsilon(\langle x, \varepsilon \rangle; d)$ .
- (b) For  $z = \langle u, v \rangle \in N$  with  $z \neq \langle x, 0 \rangle$ , let  $\rho(z)$  be the  $\varepsilon$  obtained in ((a)); let  $\rho(\langle x, 0 \rangle) = 0$ . Show that the formula

$$D(z, w) = \max\{d(z, w), |\rho(z) - \rho(w)|\}$$

defines a metric on  $N$ .

- (c) Sketch the  $D$ -ball  $B_\delta(z; D)$  for  $z = \langle x, 0 \rangle$  and for some  $z \in N$  with  $z \neq \langle x, 0 \rangle$ .

prob:pseudometric

20. A **pseudometric** on a set  $X$  is a function  $d: X \times X \rightarrow \mathbb{R}$  that satisfies properties (M1), (M3), and (M4) for a metric together with the following weakened form of (M2):

$$(M2^*) \quad d(x, x) = 0 \quad \text{for all } x \in X.$$

A trivial example of a pseudometric is the “indiscrete pseudometric” defined, for any set  $X$ , by  $d(x, y) = 0$  for all  $x, y \in X$ .

In each case below, verify that the given function  $d$  is a pseudometric on the given set  $X$  and then determine whether it is actually a metric.

- (a)  $X$  is the set of all continuous functions  $x: [0, 1] \rightarrow \mathbb{R}$ , and  $d(x, y) = |x(0) - y(0)|$ .
- (b)  $X$  is any set, and for some function  $f: X \rightarrow \mathbb{R}$ , the function  $d$  is defined by  $d(x, y) = |f(x) - f(y)|$ .
- (c)  $X$  is the set of all convergent sequences  $x = \langle x_n \rangle_{n \in \mathbb{N}}$  of real numbers, and  $d(x, y) = \lim_{n \rightarrow \infty} |x_n - y_n|$ .

prob:pseudometric-identify-distance-0

21. Let  $d$  be a pseudometric ([Exercise 20](#)) on a nonempty set  $X$ . Show that a metric space  $\langle X^*, d^* \rangle$  is obtained by “identifying points of  $X$  that are zero distance apart,” as follows.

- (a) Show that the relation  $\sim$  on  $X$  given by

$$x \sim y \iff d(x, y) = 0$$

is an equivalence relation on  $X$ .

- (b) For each  $x \in X$ , let

$$x^* = \{y \in X : x \sim y\},$$

that is,  $x^*$  is the equivalence class of  $x$  under  $\sim$ . And let  $X^*$  denote the quotient set  $X/\sim = \{x^* : x \in X\}$  consisting of all these equivalence classes. Show that there is a unique metric  $d^*$  on  $X^*$  such that

$$d^*(x^*, y^*) = d(x, y) \quad (x, y \in X).$$

- (c) Describe  $X^*$  and  $d^*$  in the case that  $d$  is already a metric.

pseudometric  
ultrametric  
metric! and ultrametric  
ultrametric inequality  
ultrametric space  
p-adic norm @ \$p\$-adic norm

- prob:ultrametric 22. An **ultrametric** on a set  $X$  is a function  $d: X \times X \rightarrow \mathbb{R}$  that satisfies properties (M1)–(M3) of a metric and in addition, instead of the triangle inequality, the stronger **ultrametric inequality**:

$$(UM) \text{ (property UM)} \quad d(x, z) \leq \max\{d(x, y), d(y, z)\}$$

for all  $x, y, z \in X$ . A metric space  $\langle X, d \rangle$  for which  $d$  is an ultrametric is called an **ultrametric space**.

- (a) Show that equality holds in (UM) whenever  $d(x, y) \neq d(y, z)$ .  
 (b) If  $x, y, z$  are points in an ultrametric space  $\langle X, d \rangle$ , show that at least two of the distances  $d(x, y)$ ,  $d(y, z)$ , and  $d(x, z)$  are the same. (In geometric terms, this says that every triangle in an ultrametric space is isosceles!)  
 (c) Verify that the metric of Example 1.15 is actually an ultrametric.  
 (d) As in (c), let  $S$  be a nonempty set and let  $X$  be the set of all sequences in  $S$ , but now for points  $x = \langle x_i \rangle_{i \in \mathbb{N}}$  and  $y = \langle y_i \rangle_{i \in \mathbb{N}}$  in  $X$ , define

$$d(x, y) = \begin{cases} 0 & \text{if } x = y, \\ 1/2^n & \text{if } x \neq y \text{ and } n \text{ is the least } i \in \mathbb{N} \text{ with } x_i \neq y_i. \end{cases}$$

Show that this  $d$ , too, is an ultrametric on  $X$ .

(Another example of an ultrametric appears in Exercise 24, for which Exercise 23 is preparation. For additional properties of ultrametrics, see Exercise 45.)

- prob:p-adic-norm 23. Let  $p$  be a fixed prime number (that is,  $p$  is a positive integer whose only positive integer divisors are 1 and  $p$  itself).

The purpose of this problem and the next is to obtain on the set  $\mathbb{Q}$  of rational numbers a new metric  $d_p$ , depending on the particular prime  $p$ , that is actually an ultrametric (Exercise 22).

On the set  $\mathbb{Q}$  of all rational numbers the  **$p$ -adic norm**  $x \mapsto |x|_p$  is defined as follows. First, set

$$|0|_p = 0.$$

Next, if  $0 \neq x \in \mathbb{Q}$ , write  $x$  in the form

$$x = p^k \frac{m}{n}$$

for integers  $k, m, n$  with  $p$  dividing neither  $m$  nor  $n$ , and set

$$|x|_p = p^{-k}.$$

**p-adic metric** **(a)** Calculate  $|1|_p$  and  $|-40/81|_p$  for each of the primes  $p = 2, 3, 5, 7$ , and  $11$ .

**metric!examples of ultrametric**

**(b)** Establish the following properties of the  $p$ -adic norm:

(i)  $|x|_p \geq 0$ .

(ii)  $|x|_p = 0$  if and only if  $x = 0$ .

(iii)  $|xy|_p = |x|_p |y|_p$

(iv)  $|x + y|_p \leq |x|_p + |y|_p$  (triangle inequality).

In order to establish the triangle inequality (iv), first prove the following stronger inequality (v).

(v)  $|x + y|_p \leq \max\{|x|_p, |y|_p\}$  (ultrametric inequality).

[*Hint:* To prove the ultrametric inequality (v), you could proceed as follows. For  $0 \neq x \in \mathbb{Q}$  write  $x$  in the form  $x = p^k m/n$  for integers  $k, m$ , and  $n$  with  $p$  dividing neither  $m$  nor  $n$ , and set  $v(x) = k$ . Show that  $v(a+b) \geq \min\{v(a), v(b)\}$  whenever  $a, b \in \mathbb{Q}$  with  $a \neq 0 \neq b$  and  $a + b \neq 0$ .]

**prob;p-adic-metric**

**24.** (Continuation of [Exercise 23](#).) On the set  $\mathbb{Q}$  of all rational numbers the  **$p$ -adic metric**  $d_p: \mathbb{Q} \times \mathbb{Q} \rightarrow \mathbb{R}$  is defined in terms of the  $p$ -adic norm by

$$d_p(x, y) = |x - y|_p.$$

Clearly  $d_p$  has the properties (M1)–(M3) of a metric. Show that  $d_p$  also satisfies the triangle inequality (M4) by proving that it is, in fact, an ultrametric ([Exercise 22](#)).

## 1.2 Open Sets and Closed Sets

**sec:openclosed**

The notion of continuity of a function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  suggested in the preceding section will become a precise definition once we introduce metrics to measure distances. On  $\mathbb{R}^n$  we have defined several different metrics, each well-behaved in its own way. But for any prejudice bred of familiarity with the Euclidean metric, there is no good reason to prefer one of these metrics over the others. Fortunately, it will turn out that the Euclidean, taxicab, and max metrics all yield the same continuous functions, because points that are close to one another for one of these metrics are close for the other two as well.

The extent to which continuity of a function is independent of the particular metrics used will be clarified in [Section 1.4](#), where we begin to study continuity in a systematic way. As preparation, we first study sets of points that are close to a given point in a metric space.

### Balls, disks, and spheres

subsec:balls-disks-spheres

def:ball-disk-sphere

**1.17 Definition (balls, disks, and spheres).** Let  $\langle X, d \rangle$  be a metric space. For a point  $x \in X$  and a real number  $\varepsilon > 0$ , the  $d$ -ball of radius  $\varepsilon$  at  $x$  (or **with center**  $x$ , or **centered at**  $x$ ) is the set

$$B_\varepsilon(x; d) = \{x \in X : d(x, y) < \varepsilon\},$$

the  $d$ -disk of radius  $\varepsilon$  at  $x$  is the set

$$D_\varepsilon(x; d) = \{x \in X : d(x, y) \leq \varepsilon\},$$

and the  $d$ -sphere of radius  $\varepsilon$  at  $x$  is the set

$$S_\varepsilon(x; d) = \{x \in X : d(x, y) = \varepsilon\}.$$

When the particular metric  $d$  is understood, we may refer to  $B_\varepsilon(x; d)$ ,  $D_\varepsilon(x; d)$ , and  $S_\varepsilon(x; d)$  simply as the  $\varepsilon$ -ball at  $x$ , the  $\varepsilon$ -disk at  $x$ , and the  $\varepsilon$ -sphere at  $x$ , respectively. (The terminology about balls and disks is far from standardized. Some authors use ‘open ball’ or, just to really confuse things, ‘open sphere’ instead of ‘ball’, and they use ‘closed ball’ or ‘closed sphere’ instead of ‘disk’. Some, alas, even use ‘sphere’ for ‘ball’.)

Observe that

$$\begin{aligned} x &\in B_\varepsilon(x; d), \quad x \notin S_\varepsilon(x; d), \\ B_\varepsilon(x; d) \cup S_\varepsilon(x; d) &= D_\varepsilon(x; d), \quad B_\varepsilon(x; d) \cap S_\varepsilon(x; d) = \emptyset. \end{aligned}$$

exs:balls-spheres-disks **1.18 Examples.** (1) Take  $d$  to be the Euclidean metric on  $\mathbb{R}$ . Let  $x \in \mathbb{R}$  and  $\varepsilon > 0$ . Then

$$\begin{aligned} B_\varepsilon(x; d) &= \{y \in \mathbb{R} : d(x, y) < \varepsilon\} = \{y \in \mathbb{R} : |x - y| < \varepsilon\} \\ &= \{y \in \mathbb{R} : x - \varepsilon < y < x + \varepsilon\} = ]x - \varepsilon, x + \varepsilon[, \end{aligned}$$

the open interval of length  $2\varepsilon$  centered at  $x$ .

ex:ball-in-Rn (2) More generally, take  $d$  to be the Euclidean metric on  $\mathbb{R}^n$ . Let  $x \in \mathbb{R}^n$  and  $\varepsilon > 0$ . Then the  $d$ -sphere  $S_\varepsilon(x; d)$  is the set of all  $y \in \mathbb{R}^n$  satisfying the quadratic equation

$$\sum_{i=1}^n (x_i - y_i)^2 = \varepsilon^2.$$

The  $d$ -ball  $B_\varepsilon(x; d)$  is the set of all  $y \in \mathbb{R}^n$  satisfying the quadratic strict inequality

$$\sum_{i=1}^n (x_i - y_i)^2 < \varepsilon^2,$$

that is,  $B_\varepsilon(x; d)$  is the set of all  $y \in \mathbb{R}^n$  lying “inside” the sphere  $S_\varepsilon(x; d)$  (see [Exercise 26](#)). And the  $d$ -disk  $D_\varepsilon(x; d)$  is the set of all  $y \in \mathbb{R}^n$  satisfying the quadratic inequality

$$\sum_{i=1}^n (x_i - y_i)^2 \leq \varepsilon^2,$$

that is,  $D_\varepsilon(x; d)$  is the set of all  $y \in \mathbb{R}^n$  “lying inside or on” that sphere.

In dimension  $n = 1$ : the  $d$ -ball  $B_\varepsilon(x; d)$  is the interval  $]x - \varepsilon, x + \varepsilon[$ ; the  $d$ -sphere  $S_\varepsilon(x; d)$  is just the set  $\{x - \varepsilon, x + \varepsilon\}$  of the two endpoints of that interval (see

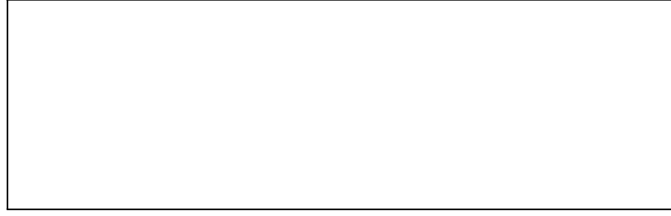
Figure 1.5: Ball of radius  $\varepsilon$  at  $x$  in  $\mathbb{R}^1$ .

fig:ball-in-R1

Figure 1.5); and the  $d$ -disk  $D_\varepsilon(x; d)$  is the interval  $[x - \varepsilon, x + \varepsilon]$ , which includes those endpoints.

In dimension  $n = 2$ : the  $d$ -ball  $B_\varepsilon(x; d)$  is the “circular region” of radius  $\varepsilon$  centered at  $x$ ; the  $d$ -sphere  $S_\varepsilon(x; d)$  is the circle surrounding that region (see Figure 1.6); and the  $d$ -disk  $D_\varepsilon(x; d)$  consists of the circular region together with the surrounding circle.

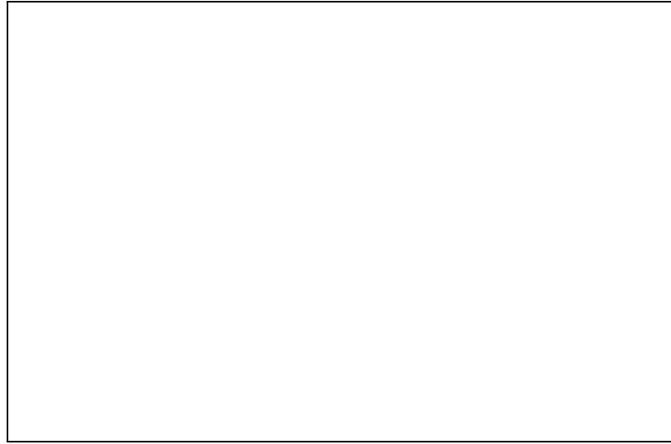
Figure 1.6: Ball of radius  $\varepsilon$  at  $x$  in  $\mathbb{R}^2$ .

fig:ball-in-R2

In dimension  $n = 3$ : the  $d$ -ball  $B_\varepsilon(x; d)$  is the solid spherical region of radius  $\varepsilon$  centered at  $x$ ; the  $d$ -sphere  $S_\varepsilon(x; d)$  is the spherical surface surrounding that region (see Figure 1.7); and the  $d$ -disk  $D_\varepsilon(x; d)$  consists of solid spherical region together with the surrounding spherical surface.

The  $d$ -ball  $B_\varepsilon(x; d)$  is convex (see the subsection “Line segments and lines”). In fact, let  $p$  be a point on the line segment joining two points  $z, w \in B_\varepsilon(x; d)$ , so that for some  $t$ ,

$$y = (1 - t)z + tw, \quad 0 \leq t \leq 1.$$

From properties Lemma 1.3 of the Euclidean norm,

$$\begin{aligned} d(x, y) &= \|x - y\| = \|[ (1 - t)x - (1 - t)z ] + [ tx - tw ]\| \\ &\leq (1 - t)\|x - z\| + t\|x - w\| = (1 - t)d(x, z) + td(x, w) \\ &< (1 - t)\varepsilon + t\varepsilon = \varepsilon. \end{aligned}$$

Hence  $y \in B_\varepsilon(x; d)$  also.

In the present context we are mainly interested in relatively “small” values of the radius  $\varepsilon$  so that we can capture the idea of points being “close to” a given center.



n-ball@n\$-ball  
 n-disk@n\$-disk  
 n-sphere@\$(n-1)\$-sphere  
 unit circle

Figure 1.7: Ball of radius  $\varepsilon$  at  $x$  in  $\mathbb{R}^3$ .

fig:ball-in-R3

Later we shall be especially interested in the particular case where the center is the origin  $\mathbf{0} = \langle 0, 0, \dots, 0 \rangle$  and the radius  $\varepsilon = 1$ . In this case we call the ball, disk, or sphere the **(unit)  $n$ -ball**, the **(unit)  $n$ -disk**, and **(unit)  $(n - 1)$ -sphere in  $\mathbb{R}^n$**  and denote them by  $B_n$ ,  $D_n$ , or  $S_{n-1}$ , respectively. Thus:

$$\begin{aligned} B_n &= \left\{ y \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 < 1 \right\} && [n\text{-ball}], \\ D_n &= \left\{ y \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 \leq 1 \right\} && [n\text{-disk}], \\ S_{n-1} &= \left\{ y \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 = 1 \right\} && [(n-1)\text{-sphere}]. \end{aligned}$$

In particular, the 0-sphere  $S_0$  is the set

$$S_0 = \{x \in \mathbb{R}^1 : x^2 = 1\} = \{-1, 1\},$$

and the 1-sphere, also called simply the **unit circle**, is the set

$$S_1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

**Caution!** Yes, that notation is what we intended:  $S_{n-1}$  for the sphere in  $\mathbb{R}^n$  with  $n - 1$  and *not*  $n$ . The reason concerns the “dimension” of the sphere itself rather than the dimension  $n$  of the surrounding Euclidean space  $\mathbb{R}^n$ . (The notion of dimension of a metric space—or of a more general kind of space—will be treated later.) If in the definition of the  $(n - 1)$ -sphere we replace  $n$  by  $n + 1$ , then the  **$n$ -sphere**

$$S_n = \left\{ y \in \mathbb{R}^{n+1} : \sum_{i=1}^{n+1} y_i^2 = 1 \right\},$$

so obtained is a subset of  $\mathbb{R}^{n+1}$ —and *not* of  $\mathbb{R}^n$ .

max metric  
metric!max  
discrete metric  
metric!discrete



Figure 1.8: A ball for the max metric is a cube.

fig:ball-for-max-metric

(Note on notation: Some authors denote by  $B^n$ ,  $D^n$ , and  $S^{n-1}$  what we have denoted by  $B_n$ ,  $D_n$ , and  $S_{n-1}$ . We prefer our notation so as to avoid any connotation that the ball, disk, or sphere is itself a product set.)

- ex:ball-for-dmax (3) Again let  $x \in \mathbb{R}^n$  and  $\varepsilon > 0$ , but now consider the max metric  $d_\infty$  on  $\mathbb{R}^n$ , so that for each  $y \in \mathbb{R}^n$ ,

$$d_\infty(x, y) = \max_{1 \leq i \leq n} |x_i - y_i|$$

(see [Definition 1.7](#)). Since

$$d_\infty(x, y) < \varepsilon$$

if and only if

$$|x_i - y_i| < \varepsilon \quad (i = 1, 2, \dots, n),$$

that is,

$$x_i - \varepsilon < y_i < x_i + \varepsilon \quad (i = 1, 2, \dots, n),$$

then the ball  $B_\varepsilon(x; d_\infty)$  is the product

$$]x_1 - \varepsilon, x_1 + \varepsilon[ \times ]x_2 - \varepsilon, x_2 + \varepsilon[ \times \cdots \times ]x_n - \varepsilon, x_n + \varepsilon[$$

of open intervals each of length  $2\varepsilon$ , with the  $i$ th interval being symmetric about  $x_i$  (see [Figure 1.8](#) for the case  $n = 3$ ). Thus the ball  $B_\varepsilon(x; d_\infty)$  is the  $\varepsilon$ -cube centered at  $x$ . In the case of dimension  $n = 2$ , this set is the “**square region**” with sides of length  $2\varepsilon$  and centered at  $x$ .

- (4) Let  $\delta$  be the discrete metric on a set  $X$  ([Example 1.12](#)). For  $x \in X$  and  $\varepsilon > 0$ ,

$$B_\varepsilon(x; \delta) = \begin{cases} \{x\} & \text{if } \varepsilon \leq 1, \\ X & \text{if } \varepsilon > 1. \end{cases}$$



- (5) Let  $\langle X, d \rangle$  be any metric space, and let  $d'$  be the metric on a subset  $Y$  of  $X$  that is induced by  $d$  (Example 1.13). If  $x \in Y$ , then

$$B_\varepsilon(x; d') = B_\varepsilon(x; d) \cap Y$$

for each  $\varepsilon > 0$ .

- (6) Let  $X = C([0, 1])$  be the set of all continuous real-valued functions on the closed interval  $[0, 1]$ , and let  $d_\infty$  be the sup metric on  $X$  as defined in Example 1.8. Let  $x: [0, 1] \rightarrow \mathbb{R}$  belong to  $X$  and let  $\varepsilon > 0$ . Recall that

$$d_\infty(x, y) = \max_{0 \leq t \leq 1} |x(t) - y(t)|.$$

Then for every  $y \in X$ ,

$$d_\infty(x, y) < \varepsilon$$

if and only if

$$|x(t) - y(t)| < \varepsilon \quad (0 \leq t \leq 1).$$

Hence the ball  $B_\varepsilon(x; d_\infty)$  consists of all those continuous functions  $y: [0, 1] \rightarrow \mathbb{R}$  whose graphs lie strictly between the graphs of the vertically translated pair of functions

$$\begin{array}{ll} x_{-\varepsilon}: [0, 1] \rightarrow \mathbb{R}, & x_\varepsilon: [0, 1] \rightarrow \mathbb{R}, \\ t \mapsto x(t) - \varepsilon & t \mapsto x(t) + \varepsilon \end{array}$$

(see Figure 1.9).  $\diamond$

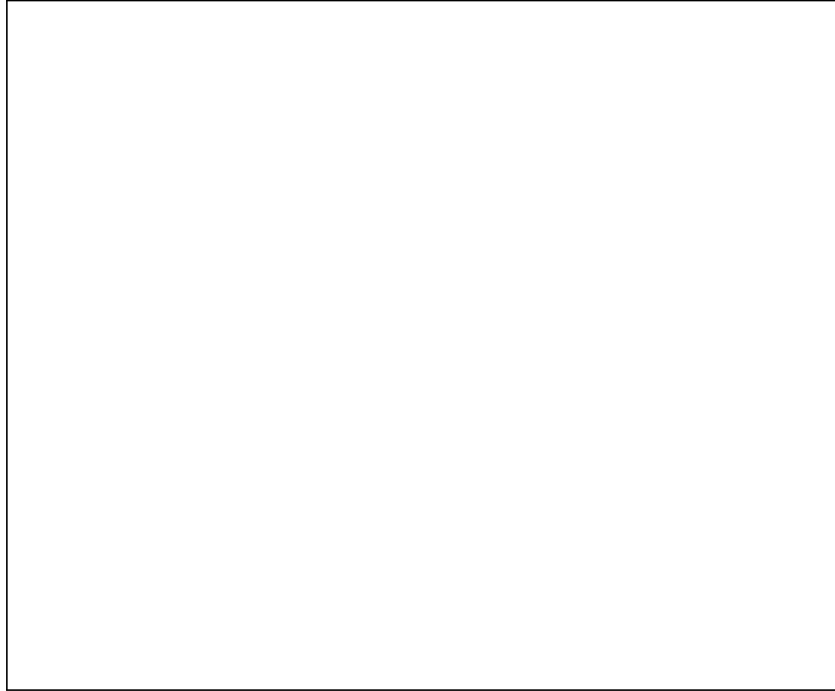


Figure 1.9: The  $\varepsilon$ -ball for the sup metric in  $C([0, 1])$ .

fig:ball-in-C-unit-interval

If  $B$  is the region in the plane enclosed by (but not containing any of the points on) a circle, then it is geometrically evident that each point of  $B$  is the center of another such circular region sufficiently small so as to be completely contained in  $B$  (see Figure 1.10). A similar result is true of a ball in any metric space.

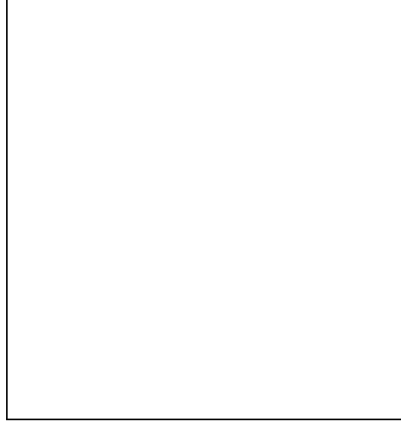


Figure 1.10: Ball contains smaller ball centered at one of its points.

fig:smaller-ball-inside-ball

prop:small-ball-inside-ball

**1.19 Proposition.** Let  $x$  be a point of a metric space  $\langle X, d \rangle$  and let  $\varepsilon > 0$ . Then for each  $y \in B_\varepsilon(x; d)$  there is some  $\eta > 0$  for which

$$B_\eta(y; d) \subset B_\varepsilon(x; d)$$

—in fact, any  $\eta$  with  $0 < \eta \leq \varepsilon - d(x, y)$  will do.

**Proof.** Let  $y \in B_\varepsilon(x; d)$ . The  $\varepsilon - d(x, y) > 0$ . Choose any  $\eta$  such that  $0 < \eta \leq \varepsilon - d(x, y)$ . If  $z \in B_\eta(y; d)$ , then  $d(y, z) < \eta$ , by the triangle inequality

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + (\varepsilon - d(x, y)) = \varepsilon,$$

and so  $z \in B_\varepsilon(x; d)$ .  $\square$

### Open sets in a metric space

Proposition 1.19 says that a  $d$ -ball  $U$  at a point has the property that all points “sufficiently close” to any given point of  $U$  also belong to  $U$ . Let us give a name this property and then look at other sets that have it.

def:d-open

**1.20 Definition.** Let  $\langle X, d \rangle$  be a metric space. A subset  $U$  of  $X$  is said to be  **$d$ -open** if for each  $x \in U$  there is some  $\varepsilon > 0$  such that  $B_\varepsilon(x; d) \subset U$ .

Thus **any  $d$ -ball in a metric space is  $d$ -open**. In order to construct other examples of  $d$ -open sets, it will be convenient to establish the following properties of  $d$ -open sets.

**1.21 Theorem (properties of  $d$ -open sets).** Let  $\langle X, d \rangle$  be a metric space. Then:

- (1) The empty set  $\emptyset$  and the entire set  $X$  are both  $d$ -open.
- (2) The union of any number of  $d$ -open sets is itself  $d$ -open.
- (3) The intersection of finitely many  $d$ -open sets is itself  $d$ -open.

**Proof.** (1) Since there is no  $x \in \emptyset$  whatsoever, it is vacuously true that at each  $x \in \emptyset$  there is some  $d$ -ball contained in  $\emptyset$ . Hence  $\emptyset$  is  $d$ -open.

The whole set  $X$  is  $d$ -open since  $B_\varepsilon(x; d) \subset X$  for every  $x \in X$  and every  $\varepsilon > 0$ .

- (2) Let  $\langle U_i \rangle_{i \in I}$  be any family of  $d$ -open sets, and let  $U = \bigcup_{i \in I} U_i$  be their union. Let  $z \in U$ . Then  $z \in U_j$  for some  $j \in J$ . Since  $U_j$  is  $d$ -open, for some  $\varepsilon > 0$  the ball  $B_\varepsilon(z; d) \subset U_j$ . Then also  $B_\varepsilon(z; d) \subset U$ . Thus the union  $U$  is  $d$ -open.
- (3) Let  $U_1, U_2, \dots, U_n$  be  $d$ -open sets, and let  $U = \bigcap_{i=1}^\infty U_i$ . Let  $x \in U$ ; we must show that  $B_\varepsilon(x; d) \subset U$  for some  $\varepsilon > 0$ . For each  $i = 1, 2, \dots, n$  we have  $x \in U_i$ , a  $d$ -open set, and so there is some  $\varepsilon_i > 0$  with

$$B_{\varepsilon_i}(x; d) \subset U_i.$$

Set

$$\varepsilon = \min\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}.$$

Then  $\varepsilon > 0$ , too, and

$$B_\varepsilon(x; d) \subset B_{\varepsilon_i}(x; d) \subset U_i \quad (i = 1, 2, \dots, n),$$

so that  $B_\varepsilon(x; d) \subset U$ . Thus the intersection  $U$  is  $d$ -open.  $\square$

The preceding theorem will be crucial for motivating the definition of a topological space as given in [Chapter 2](#) (Topological Spaces).

**1.22 Examples.** (1) Let  $\delta$  be the discrete metric on a set  $X$  ([Example 1.12](#)). Then for each  $x \in X$ , the singleton  $\{x\}$  is  $\delta$ -open, because  $B_1(x; \delta) = \{x\}$ . It follows from [Theorem 1.21](#) that every subset of  $X$  is  $\delta$ -open, for if  $A \subset X$ , then

$$A = \bigcup_{x \in A} \{x\}.$$

- (2) Any open interval  $]a, b[$  in the real line is a  $d$ -open set for the Euclidean metric  $d$  on  $\mathbb{R}$ . In fact, take  $c$  to be the midpoint of  $]a, b[$ , that is,

$$c = \frac{1}{2}(a + b),$$

and let  $\varepsilon$  be distance between that midpoint and each of the ends of  $]a, b[$ , that is,

$$\varepsilon = \frac{1}{2}(b - a) = b - c = c - a.$$

Then  $]a, b[ = B_\varepsilon(c; d)$ , a  $d$ -ball, so that the interval  $]a, b[$  is indeed  $d$ -open.

On the other hand, a closed interval  $[a, b]$  or a half-closed interval  $[a, b[$  or  $]a, b]$  is *not*  $d$ -open. (Why not?)

open set!in metric space  
metric space!and open sets  
discrete metric!and open sets  
open set!in real line  
real line!and open sets  
interval!and open sets  
open interval!and open sets  
closed interval

open-sets-properties-in-metric-space

thm-part:nil-and-whole-are-d-open

n-part:union-of-d-open-sets-is-d-open

f-finitely-many-d-open-sets-is-d-open

exs:d-open-in-R-etc

ex:open-interval-ray-d-open

real line and open sets  
 open ray as open set  
 ray and open sets  
 closed ray  
 real line and open sets

Likewise, any open ray  $]a, +\infty[$  or  $]-\infty, b[$  in  $\mathbb{R}$  is a  $d$ -open set. In fact, a ray of the form  $]a, +\infty[$  is  $d$ -open because if  $x \in ]a, +\infty[$  and if we take  $\varepsilon = x - a$ , then

$$B_\varepsilon(x; d) = ]x - \varepsilon, x + \varepsilon[ \subset ]a, +\infty[.$$

Similarly, a ray of the form  $]-\infty, b[$  is  $d$ -open.

On the other hand, a closed ray  $[a, +\infty[$  or  $]-\infty, b]$  is *not*  $d$ -open. (Why not?)

In view of the preceding facts, from [Theorem 0.87](#) we conclude that *an interval  $J$  in  $\mathbb{R}$  is  $d$ -open for the Euclidean metric  $d$  if and only if  $J$  is  $\emptyset$ ,  $\mathbb{R}$ , an open interval  $]a, b[$ , or an open ray  $]a, +\infty[$  or  $]-\infty, b[$ .*

Thus the meanings of ‘open’ in the terms ‘open interval’ and ‘open ray’, on the one hand, and in the term ‘ $d$ -open set’, on the other hand—are consistent. Moreover, as the following example indicates,  $d$ -open sets are just generalizations of open intervals and open rays.

ex: all  $d$ -open sets in  $\mathbb{R}$

- (3) Again let  $d$  be the Euclidean metric on  $\mathbb{R}$ . we shall determine all the  $d$ -open subsets of  $\mathbb{R}$ .

Already from [Theorem 1.21](#) 2 and 2 preceding, we know that the union of any countable collection of pairwise disjoint open intervals and open rays is  $d$ -open. We shall show, conversely, that *each proper  $d$ -open subset of  $\mathbb{R}$  is the union of a countable collection of pairwise disjoint open intervals and open rays.*

Let  $U$  be a proper  $d$ -open subset of  $\mathbb{R}$ . For each  $x \in U$ , define  $J_x$  to be the set of those  $y \in \mathbb{R}$  such that there exist real numbers  $a$  and  $b$  with  $a < b$  and

$$x \in ]a, b[, \quad y \in ]a, b[, \quad ]a, b[ \subset U.$$

We are going to show that collection  $\{J_x : x \in U\}$  is the desired countable collection of pairwise disjoint open intervals and open rays whose union is  $U$ .

Note that the sets  $J_x$  need not be distinct for different values of  $x$ . For example, if  $U$  were itself an open interval  $]a, b[$ , then  $J_x = ]a, b[$  for every  $x \in U$ .

Step 1: the set  $U$  has the form

$$U = \bigcup_{x \in U} J_x.$$

In fact, if  $x \in U$ , then certainly  $J_x \subset U$ ; thus

$$\bigcup_{x \in U} J_x \subset U.$$

To prove the opposite inclusion, let  $x \in U$ . Since  $U$  is  $d$ -open, then  $x \in ]x - \varepsilon, x + \varepsilon[ \subset U$  for some  $\varepsilon > 0$ . Hence  $x \in J_x$ .

Step 2: if  $x, z \in U$ , then

$$(*) \quad z \in J_x \implies J_x = J_z.$$

In fact, let  $z \in J_x$ , so that

$$x, z \in ]a, b[ \subset U$$

for some  $a < b$ . If  $y \in J_z$ , then also

$$y, z \in ]c, d[ \subset U$$

for some  $c < d$ , and then  $]a, b[ \cup ]c, d[$  is an open interval  $]u, v[$  with

$$x, y \in ]u, v[ \subset U,$$

and so  $y \in J_x$ . Hence  $J_x \subset J_z$ . Similarly,  $J_z \subset J_x$ .

Step 3: the collection  $\{J_x : x \in U\}$  is pairwise disjoint. In fact, let  $x, y \in U$  and suppose there is some  $z \in J_x \cap J_y$ . From (\*),

$$J_x = J_z, \quad J_y = J_z.$$

Hence  $J_x = J_y$ .

So far, we have a pairwise disjoint collection  $\{J_x : x \in U\}$  of subsets of  $U$  whose union is  $U$ .

Step 4: for each  $x \in U$ , the set  $J_x$  is an interval in the sense of Definition 0.65. In fact, fix  $x \in U$ . Let  $u, v \in J_x$  with  $u < v$ . There are open intervals  $]a, b[$  and  $]c, d[$  with

$$u, x \in ]a, b[ \subset U, \quad u, v \in ]c, d[ \subset U.$$

Then  $]a, b[ \cup ]c, d[$  is an open interval with

$$]u, v[ \subset ]a, b[ \cup ]c, d[ \subset U,$$

and so  $]u, v[ \subset J_x$ . Hence  $J_x$  is an interval.

Step 5: for each  $x \in U$ , the interval  $J_x$  is an open interval or an open ray. In fact, fix  $x \in U$ . The interval  $J_x \neq \mathbb{R}$  because  $J_x \subset U \neq \mathbb{R}$ . To see that  $J_x$  is an open interval or an open ray, we show that it contains no endpoint. Just suppose it contained a left endpoint  $y$ . There exist  $a, b$  with

$$a < y < x < b, \quad ]a, b[ \subset U.$$

Choose  $z$  with  $a < z < y$ . Since  $z \notin J_x$  whereas  $z \in U$ , then  $J_z \neq J_x$ . However,  $y \in J_x \cap J_z$ , which is impossible. Similarly,  $J_x$  cannot contain a right endpoint.

Step 6: the collection  $\{J_x : x \in U\}$  is countable. In fact, in each open interval  $J_x$  we may choose some rational number  $r(J_x)$ . Since  $\{J_x : x \in U\}$  is pairwise disjoint,  $r(J_x) \neq r(J_y)$  if  $J_x \neq J_y$ . But there are only countably many rational numbers.

ex:open-sets-subset-metric

- (4) Let  $(X, d)$  be a metric space, let  $Y \subset X$ , and let  $d'$  be the metric on  $Y$  induced by  $d$  (Example 1.13). What is the relationship between the  $d'$ -open subsets of  $Y$  and the  $d$ -open subsets of  $X$ ?

To answer that question, first let  $U$  be a  $d$ -open subset of  $X$ . Then the subset  $V$  of  $Y$  defined by

{eq:open-intersect-subset-metric} (\*)

$$V = U \cap Y$$

is a  $d'$ -open subset of  $Y$ . To see this, let  $y \in V$ . Since  $y \in U$  and  $U$  is  $d$ -open, then  $B_\varepsilon(y; d) \subset U$  for some  $\varepsilon > 0$  (see Figure 1.11). Then

$$B_\varepsilon(y; d') = B_\varepsilon(y; d) \cap Y \subset U \cap Y = V.$$

Hence  $V$  is  $d'$ -open.

Now let  $V$  be a  $d'$ -open subset of  $Y$ . We claim that  $V$  has the form (\*) for some  $d$ -open subset  $U$  of  $X$ . In fact, for each  $y \in V$ , there exists some  $\varepsilon(y) > 0$  with

$$B_{\varepsilon(y)}(y; d') \subset V.$$

Define

$$U = \bigcup_{y \in V} B_{\varepsilon(y)}(y; d').$$

From Theorem 1.21 2 and the fact that any  $d$ -ball is  $d$ -open, it follows that  $U$  is  $d$ -open. An easy computation shows that  $U \cap Y = V$ .

Thus the  $d'$ -open subsets of  $Y$  are precisely the intersections with  $Y$  of the  $d$ -open subsets of  $X$ .

open set!  
in subset  
subspace!  
and open sets

intersection!of open sets  
open set!and intersection

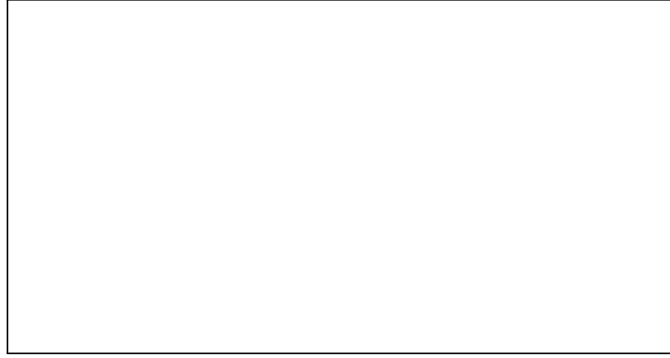


Figure 1.11: Ball and open set for metric induced on subset.

fig:d-prime-open-in-subset

ex:open-sets-in-closed-interval

- (5) Let  $d'$  be the Euclidean metric on the closed interval  $[0, 1]$ , so that  $d'$  is induced by the Euclidean metric  $d$  on  $\mathbb{R}$ . Together, 3 and 4 will allow us to determine all the  $d'$ -open subsets of  $[0, 1]$ .

The intersection of an open interval or an open ray in  $\mathbb{R}$  with  $[0, 1]$  is one of: the empty set  $\emptyset$ ; the entire closed interval  $[0, 1]$ ; an interval of the form  $[0, b[$  for some  $b$  with  $0 < b < 1$ ; an interval of the form  $]a, 1]$  for some  $a$  with  $0 < a < 1$ ; or an interval of the form  $]a, b[$  for some  $a, b$  with  $0 < a < b < 1$ . Hence a subset  $V$  of  $[0, 1]$  is  $d'$ -open if and only if it has one of the forms

$$\begin{aligned} V &= \emptyset, \\ V &= [0, 1], \\ V &= [0, b[ \cup W, \\ V &= ]a, 1] \cup W, \\ V &= [0, b[ \cup ]a, 1] \cup W, \\ V &= W, \end{aligned}$$

where  $0 < b < 1$ ,  $0 < a < 1$ , and  $W$  is a (possibly empty) union of a countable collection of disjoint open intervals contained in  $]0, 1[$ .

Notice that  $[0, 1]$  is  $d'$ -open, but it is *not*  $d$ -open.

ex:open-rectangle-open-set

- (6) Let  $d_\infty$  be the max metric on the plane  $\mathbb{R}^2$ . By Examples 1.18 (3), each “square region” of the form  $]x - \varepsilon, x + \varepsilon[ \times ]y - \varepsilon, y + \varepsilon[$  is  $d_\infty$ -open. Then each square region of the form  $]s, t[ \times ]u, v[$  with sides of the same length  $v - u = t - s$  (and parallel to the axes) is  $d_\infty$ -open, because it is of that form for  $\langle x, y \rangle$  being its center and  $\varepsilon$  being half the length of its sides. Then each rectangular region  $R$  with sides parallel to the axes—that is, each set of the form  $R = ]a, b[ \times ]c, d[$ —is also  $d_\infty$ -open; in fact, such an  $R$  is the union of (overlapping) square regions having sides parallel to the axes, and so Theorem 1.21 (2) is applicable.

This example generalizes to  $\mathbb{R}^n$ : see Exercise 33.

- (7) Although in a metric space  $\langle X, d \rangle$  the intersection of finitely many  $d$ -open sets will always be  $d$ -open, the intersection of arbitrarily many  $d$ -open sets need not be  $d$ -open. In fact, *the intersection of denumerably many  $d$ -open sets need not be  $d$ -open*.

For example, let  $d$  be the Euclidean metric on  $\mathbb{R}$ . Then for each positive integer  $n$  the interval  $] -1/n, 1/n[$  is  $d$ -open, but the intersection

$$\bigcap_{n=1}^{\infty} ] -1/n, 1/n[$$

is *not*  $d$ -open.  $\diamond$

It is natural to single out, along with the  $d$ -open sets in a metric space, the complements of those sets.

**1.23 Definition.** Let  $\langle X, d \rangle$  be a metric space. A subset  $E$  of  $X$  is said to be  **$d$ -closed** if its complement  $X \setminus E$  in  $X$  is  $d$ -open.

Recalling what it means for  $X \setminus E$  to be  $d$ -open, we see that:

- a subset  $E$  of  $X$  is  $d$ -closed exactly when at each point of  $X$  not belonging to  $E$  there is some  $d$ -ball that is disjoint from  $E$ .

Since  $X \setminus (X \setminus A) = A$  for each subset  $A$  of  $X$ , we also see that:

- a subset  $A$  of  $X$  is  $d$ -open if and only if its complement  $X \setminus A$  in  $X$  is  $d$ -closed.

In view of [Examples 1.22 \(3\)](#), for the Euclidean metric  $d$  on the real line, each closed interval  $[a, b]$ , closed ray  $[a, +\infty[$ , or closed ray  $] -\infty, b]$  is a  $d$ -closed subset of  $\mathbb{R}$ . There are many other  $d$ -closed subsets of  $\mathbb{R}$ , too!

The analog of [Theorem 1.21](#) for  $d$ -closed sets is the following theorem.

**1.24 Theorem (properties of  $d$ -closed sets).** Let  $\langle X, d \rangle$  be a metric space. Then:

- (1) The empty set  $\emptyset$  and the entire set  $X$  are both  $d$ -closed.
- (2) The intersection of any number of  $d$ -closed sets is itself  $d$ -closed.
- (3) The union of finitely many  $d$ -closed sets is itself  $d$ -closed.

**Proof.** (1) According to [Theorem 1.21 \(1\)](#), both  $\emptyset$  and  $X$  are  $d$ -open. Hence  $X = X \setminus \emptyset$  and  $\emptyset = X \setminus X$  are  $d$ -closed

- (2) Let  $\langle E_i \rangle_{i \in I}$  be any family of  $d$ -closed sets and let

$$E = \bigcap_{i \in I} E_i.$$

By one of [De Morgan's Laws \(0.19\)](#),

$$X \setminus E = \bigcup_{i \in I} (X \setminus E_i).$$

Now each set  $X \setminus E_i$  is  $d$ -open, so by [Theorem 1.21 \(2\)](#) the above union is also  $d$ -open. Hence  $E$  is closed.

- (3) This proof of [3](#) is similar to the preceding except that it uses part (3) of [Theorem 1.21](#).  $\square$

**1.25 Examples.** (1) In any metric space  $\langle X, d \rangle$ , both the sets  $\emptyset$  and  $X$  are simultaneously  $d$ -open and  $d$ -closed. (When these two are the *only* subsets of  $X$  having this property,  $X$  is said to be “connected.” Connected spaces are studied in detail in [Chapter 5](#).)

- (2) Let  $\delta$  be the discrete metric on a set  $X$ . Since every subset of  $X$  is  $\delta$ -open, then every subset of  $X$  is  $\delta$ -closed as well.
- (3) In the metric space  $\langle \mathbb{R}, d \rangle$ , where  $d$  is the Euclidean metric, the set  $]0, 1[$  is  $d$ -open but not  $d$ -closed.
- (4) In the metric space  $\langle \mathbb{R}, d \rangle$  just considered, the set  $[0, 1[$  is neither  $d$ -open nor  $d$ -closed.

**Caution!** This example should correct any mistaken idea that a set must be  $d$ -open if it is not  $d$ -closed, or  $d$ -closed if it is not  $d$ -open. Even a door may be neither open nor closed, but instead ajar [Of course, a door cannot be both open and closed at the same time, but a set in a metric space  $\langle X, d \rangle$  may be both  $d$ -open and  $d$ -closed.]

Sets in a metric space that are neither open nor closed are in fact sometimes said to be **ajar**—realliy!

- (5) Any  $d$ -disk  $D_r(x; d)$  at a point  $x$  in a metric space  $\langle X, d \rangle$  is  $d$ -closed. In fact, let  $y \in X$  with  $y \notin D_r(x; d)$ . Then  $d(x, y) > r$ . Define

$$\varepsilon = d(x, y) - r,$$

so that  $\varepsilon > 0$ . An easy use of the triangle inequality shows that  $B_\varepsilon(y; d)$  is disjoint from  $D_r(x; d)$ —see [Figure 1.12](#).

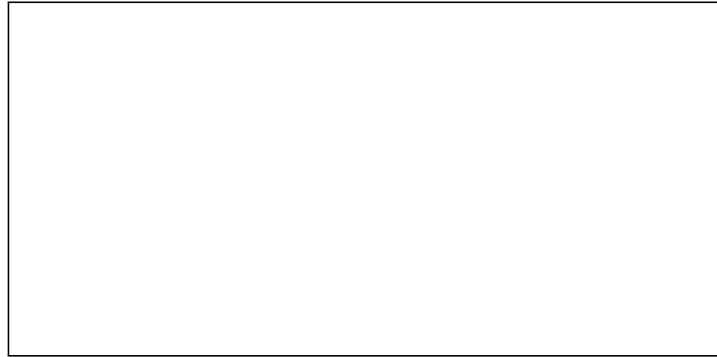


Figure 1.12: Ball disjoint from a disk

fig:ball-disk-disjoint

A similar argument shows that any  $d$ -sphere  $S_\varepsilon(x; d)$  is  $d$ -closed.

- (6) Let  $\langle X, d \rangle$  be a metric space. Then each singleton  $\{x\}$  is  $d$ -closed; in fact, if  $y \neq x$ , then  $B_\varepsilon(y; d)$  is disjoint from  $\{x\}$  whenever  $0 < \varepsilon \leq d(x, y)$ . Said in another way, if  $x$  and  $y$  are distinct points in a metric space  $\langle X, d \rangle$ , then each of them lies in some  $d$ -open set not containing the other. Later, we shall express this situation by saying that  $X$  is a  $T_1$ -space; see [Definition 2.94](#) (1).

From [Theorem 1.24](#) (3) it follows that any finite subset of  $X$  is  $d$ -closed.  $\diamond$



It was noted above that a point  $x$  not belonging to a  $d$ -closed set  $A$  in a metric space  $\langle X, d \rangle$  is contained in some  $d$ -open set—in fact, in a  $d$ -ball—that is disjoint from  $A$ . Actually, something more is true: in a metric space, *a  $d$ -closed set and a point not in that set can be separated by  $d$ -open sets*. Later, we shall express this situation by saying that  $X$  is *regular*; see [Definition 2.94 \(3\)](#).

**1.26 Proposition (separating a point from a  $d$ -closed set).** *Let  $\langle X, d \rangle$  be a metric space. If  $A$  is a  $d$ -closed subset of  $X$  and if  $x$  is a point of  $X$  with  $x \notin A$ , then there are disjoint  $d$ -open sets  $U$  and  $V$  with*

$$x \in U, \quad A \subset V.$$

**Proof.** Since  $A$  is  $d$ -closed, there exists  $\delta > 0$  such that  $B_\delta(x; d)$  is disjoint from  $A$ . Choose  $\varepsilon$  and  $\eta$  with

$$0 < \varepsilon < \eta < \delta.$$

Define

$$U = B_\varepsilon(x; d), \quad V = X \setminus D_\eta(x; d)$$

(see [Figure 1.13](#)). Then  $U$  is  $d$ -open by [Proposition 1.19](#), and  $V$  is  $d$ -open by [Examples 1.25 \(5\)](#). Since

$$B_\varepsilon(x; d) \subset D_\eta(x; d) \subset B_\delta(x; d),$$

we have  $A \subset V$  and  $U \cap V = \emptyset$ .  $\square$

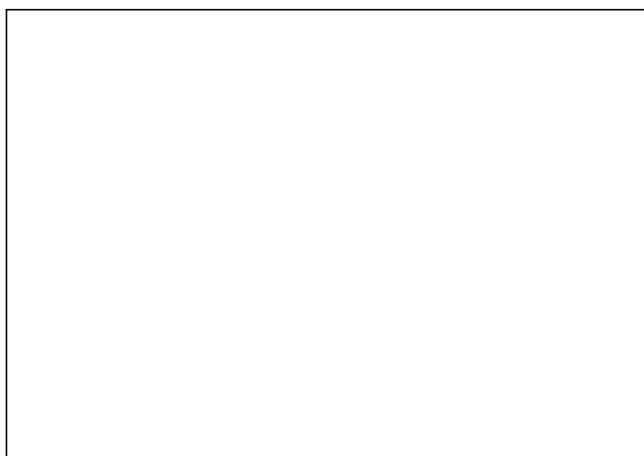


Figure 1.13: Separating a point from a  $d$ -closed set in a metric space.

open set!in metric space  
metric space!and open sets  
closed set!in metric space  
metric space!and closed sets  
regular space!metric space@and me  
metric space!regular space@as regul  
regular space!metric space@and me  
metric space!regular space@as regul  
Hausdorff space!metric space@and  
T2-space@Ttwo-space!metric space  
metric space!T2-space@as Ttwo-sp

fig:separate-pt-from-closed-in-metri

In view of [Examples 1.25 \(6\)](#), we could take  $A$  to be a singleton  $\{y\}$  in [Proposition 1.26](#) to obtain the following corollary, which says that, in a metric space, distinct points can be separated by disjoint  $d$ -open sets. Later, we shall express this situation by saying that  $X$  is a  *$T_2$ -space*, and that it is a *Hausdorff space*; see [Definition 2.24](#).

Hausdorff space!  
cor:separate-pts-in-metric  
space@Vtwo-sp  
metric space!  
normal space!  
metric space!  
normal space!

**1.27 Corollary.** If  $x$  and  $y$  are distinct points in a metric space  $\langle X, d \rangle$ , then there are disjoint  $d$ -open subsets  $U$  and  $V$  of  $X$  with  $x \in U$  and  $y \in V$ .

[separating two points in a metric space]

Exercise 30 asks you to prove 1.27 directly, without any use of Proposition 1.26.

def:dist-pt-to-set

**1.28 Definition (distance from point to set).** Let  $x$  be a point and  $A$  be a nonempty subset of a metric space  $\langle X, d \rangle$ . Then the **distance from  $x$  to  $A$**  is defined to be the nonnegative real number

$$d(x, A) = \inf\{d(x, y) : y \in A\}$$

This definition does make sense, because the set  $\{d(x, y) : y \in A\}$  is nonempty,  $A$  being nonempty, and has 0 as a lower bound.

When  $A$  reduces to a singleton  $\{y\}$ , then  $d(x, A) = d(x, y)$ . If there happens to be a point  $y \in A$  for which  $d(x, y) = d(x, A)$ —so that the infimum in Definition 1.28 is an actual minimum—then we may call  $y$  a **nearest point of  $A$  to  $x$** .

However, such a nearest point need not exist! Indeed, it can happen that  $d(x, A) = 0$  yet  $x \notin A$ . For example, take  $d$  to be the Euclidean metric on  $\mathbb{R}$  and  $A = \{1/n : n = 1, 2, 3, \dots\}$ . Then

$$0 \leq d(0, A) = \inf\left\{\frac{1}{n} : n = 1, 2, 3, \dots\right\} \leq \frac{1}{k} \quad (k = 1, 2, 3, \dots),$$

so that  $d(0, A) = 0$  but  $0 \notin A$ . This example is explained by the following criterion alluded to above.

prop:d-closed-via-dist-from-pts

**1.29 Proposition (distance of points from  $d$ -closed set).** Let  $\langle X, d \rangle$  be a metric space and let  $A$  be a nonempty subset of  $X$ . Then a necessary and sufficient condition for  $A$  to be  $d$ -closed is that  $d(x, A) > 0$  for each  $x \in X$  not belonging to  $A$ .

**Proof.** Necessity. Assume that  $A$  is  $d$ -closed. Let  $x \in X \setminus A$ . There exists  $\varepsilon > 0$  with

$$B_\varepsilon(x; d) \subset X \setminus A.$$

If  $y \in A$ , then  $y \notin B_\varepsilon(x; d)$  and so  $d(x, y) \geq \varepsilon$ . Hence

$$d(x, A) = \inf\{d(x, y) : y \in A\} \geq \varepsilon > 0.$$

Sufficiency. Assume that the condition holds. Let  $x \in X \setminus A$ . Then  $d(x, A) > 0$ , and we may form the  $d$ -ball  $B_\varepsilon(x; d)$  of radius  $\varepsilon = d(x, A)$ . Since  $d(x, y) \geq d(x, A)$  for each  $y \in A$ , then  $B_\varepsilon(x; d)$  is disjoint from  $A$ . Hence  $A$  is  $d$ -closed.  $\square$

Reasoning about distances from points to sets may be used (see Exercise 42) to strengthen Proposition 1.26 so as to show that *any two disjoint  $d$ -closed sets in a metric space can be separated from each other by  $d$ -open sets*. Later, we shall express this situation by saying that  $X$  is a *normal space*; see Definition 2.94 (5).

prop:metric-normal

**1.30 Proposition (separating  $d$ -closed sets).** If  $A$  and  $B$  are disjoint  $d$ -closed subsets of a metric space  $\langle X, d \rangle$ , then there are disjoint  $d$ -open subsets  $U$  and  $V$  such that  $A \subset U$  and  $B \subset V$ .

In generalizing from  $d$ -balls to  $d$ -open sets we neglected the “local” aspect of  $d$ -balls—the fact that a  $d$ -ball consists of points sufficiently close to a *given* point. Next we introduce a generalization of  $d$ -balls that does retain this aspect.

normal space! metrizable space! as norm  
metric space! normal space! as norm

neighborhood! in metric space!

**1.31 Definition ( $d$ -neighborhood).** Let  $x$  be a point in a metric space  $\langle X, d \rangle$ . A subset  $V$  of  $X$  is called a  $d$ -neighborhood of  $x$  if there exists some  $d$ -open set  $U$  with  $x \in U \subset V$ .

Clearly:

- Subset  $V$  of  $X$  is a  $d$ -neighborhood of  $x$  if there exists some  $d$ -ball  $B_\varepsilon(x; d)$  at  $x$  with  $B_\varepsilon(x; d) \subset V$ .
- Any  $d$ -open set containing  $x$ —in particular any  $d$ -ball at  $x$ —is a  $d$ -neighborhood of  $x$ .
- Any  $d$ -disk  $D_\varepsilon(x; d)$  at  $x$  is a  $d$ -neighborhood of  $x$ . Thus:

**Caution!** According to our definition, a  $d$ -neighborhood of a point need not be  $d$ -open. (Some authors include as part of the definition of a  $d$ -neighborhood the requirement that it be  $d$ -open!)

By their very definition, the  $d$ -neighborhoods of a point are determined by the  $d$ -open sets. Conversely, as the following proposition states, the  $d$ -open sets are determined by the  $d$ -neighborhoods of points.

prop:  $d$ -open iff nbd each pt

**1.32 Proposition (characterization of  $d$ -open sets by neighborhoods).** Let  $\langle X, d \rangle$  be a metric space. Then a subset  $V$  of  $X$  is  $d$ -open if and only if  $V$  is a  $d$ -neighborhood of each of its points.

**Proof.** Clearly a  $d$ -open set is a  $d$ -neighborhood of each of its points.

Conversely, let  $V$  be a subset of  $X$  that is a  $d$ -neighborhood of each of its points. Then for each  $x \in V$  there is some  $d$ -open set  $U_x$  such that

$$x \in U_x \subset V.$$

Since then

$$V = \bigcup_{x \in V} U_x$$

the set  $V$  is  $d$ -open by Theorem 1.21 (2).  $\square$ .

cor:  $d$ -closed via nbds

**1.33 Corollary.** Let  $\langle X, d \rangle$  be a metric space. Then a subset  $E$  of  $X$  is  $d$ -closed if and only if every point of  $X$ , each neighborhood of which intersects  $E$ , belongs to  $E$ .

## EXERCISES FOR SECTION 1.2

**25.** Given  $x \in \mathbb{R}^2$  and  $\varepsilon > 0$ , describe and draw pictures of  $B_\varepsilon(x; d_1)$ ,  $D_\varepsilon(x; d_1)$ , and  $S_\varepsilon(x; d_1)$  where  $d_1$  is the taxicab metric on  $\mathbb{R}^2$ .

**26.** Let  $d$  be the Euclidean metric on  $\mathbb{R}^n$ , let  $x \in \mathbb{R}^n$ , and let  $\varepsilon > 0$ . Show that a point  $y \in \mathbb{R}^n$  not belonging to  $S_\varepsilon(x; d)$  belongs to  $B_\varepsilon(x; d)$  if and only if it lies on some line segment joining  $x$  to a point of  $S_\varepsilon(x; d)$ .

prob: pt inside sphere on segment

**Hilbert sequence space** **27.** If  $d_1$  is the taxicab metric on  $\mathbb{R}^n$ , is every  $d_1$ -ball convex? Answer the corresponding question for the max metric  $d_\infty$ .

**Hilbert cube**

**disjunct-open-sets-in-metric-space**

**28. (a)** Show that any pairwise disjoint collection of  $d$ -open subsets of  $\mathbb{R}$ , where  $d$  is the Euclidean metric, must be countable.

**(b)** Is the corresponding result true for  $\mathbb{R}^2$ ? for  $\mathbb{R}^n$ ?

**29.** Determine all  $d'$ -open subsets of  $[0, 1[$ , where  $d'$  is the Euclidean metric on  $[0, 1[$ . Do the same for  $]0, 1]$ .

**direct-proof-separate-pts-in-metric**

**30.** Prove [Corollary 1.27](#) directly, without any use of [Proposition 1.26](#).

**31.** Let  $X$  be the set of all continuous functions  $x: [0, 1] \rightarrow \mathbb{R}$  and let  $d_\infty$  be the sup metric on  $X$ , as defined in [Example 1.8](#). Suppose  $x, y \in X$  with  $x(t) < y(t)$  for all  $0 \leq t \leq 1$ . Is the subset

$$\{x \in X : x(t) < z(t) < y(t) \text{ for all } 0 \leq t \leq 1\}$$

of  $X$  a  $d_\infty$ -ball?

**prob:d-open-iff-union-d-balls**

**32.** Prove that a subset of a metric space  $\langle X, d \rangle$  is  $d$ -open if and only if it is a union of  $d$ -balls.

**prob:open-hypercube-is-dmax-open**

**33.** Generalize [Examples 1.22 \(6\)](#) to  $\mathbb{R}^n$  for  $n \geq 2$ .

**34.** Let  $d$  be the Euclidean metric on  $\mathbb{R}^n$ , let  $x \in \mathbb{R}^n$ , and let  $\varepsilon > 0$ . Suppose  $A$  is a subset of  $\mathbb{R}^n$  with  $B_\varepsilon(x; d) \subset A \subset D_\varepsilon(x; d)$ . When is the set  $A$   $d$ -open?  $d$ -closed?

**35.** Under what circumstances will the intersection  $\bigcap_{\varepsilon > 0} B_x(\varepsilon; d)$  of the collection of all open balls at a point  $x$  be open in a metric space  $\langle X, d \rangle$ ?

**prob:hilbert-space-and-cube**

**36.** Let  $(\ell^2, d_2)$  be the Hilbert sequence space of [Example 1.10](#).

**(a)** Show that each sequence  $x = \langle x_i \rangle_{i=1,2,3,\dots}$  of real numbers with  $|x_i| \leq 1/i$  for each  $i$  is square-summable, thereby verifying that the Hilbert cube  $l^\infty$  is, in fact, a subset of  $\ell^2$ .

**prob-part:hilbert-cube**

**(b)** Show that  $l^\infty$  is  $d_2$ -closed in  $\ell^2$ .

**(c)** Is the set

$$\{x \in \ell^2 : |x_i| < 1/i \text{ for each } i = 1, 2, \dots\}$$

$d_2$ -open in  $\ell^2$ ?

**37.** Let  $d'$  be the metric induced on a subset  $Y$  of a set  $X$  by a metric  $d$  on  $X$ .

**(a)** Is every subset of  $Y$  of the form  $E \cap Y$ , with  $E$  a  $d$ -closed subset of  $X$ , a  $d'$ -closed set? Is every  $d'$ -closed subset of  $Y$  of this form?

**(b)** Let  $y \in Y$ . Is every subset of  $Y$  of the form  $V \cap Y$ , with  $V$  a  $d$ -neighborhood of  $y$ , a  $d'$ -neighborhood of  $y$ . Is every  $d'$ -neighborhood of  $y$  of this form?

**38.** Let  $A$  be a nonempty subset of a metric space  $\langle X, d \rangle$ .

**(a)** Show that

$$|d(x, A) - d(y, A)| \leq d(x, y)$$

for all  $x, y \in A$ .

**(b)** Prove that the set

$$\{x \in X : d(x, A) < \varepsilon\}$$

is  $d$ -open for every  $\varepsilon > 0$ .

be  $d$ -closed for every  $\varepsilon > 0$ ?

distance!between sets

e-intersects-with-disj-open-in-space

- 40.** Let  $A$  and  $B$  be subsets of a metric space  $\langle X, d \rangle$  with  $A \subset B$ . Must  $d(x, \text{bdy } A) \leq d(x, \text{bdy } B)$  for all  $x \in A$  if:

- (a)**  $X = \mathbb{R}$  with the Euclidean metric?      **(d)**  $X$  is a subset of  $\mathbb{R}^2$  with the Euclidean metric?  
**(b)**  $X$  is a subset of  $\mathbb{R}$  with the Euclidean metric?      **(e)**  $X = \mathbb{R}^n$  for arbitrary  $n$ ?  
**(c)**  $X = \mathbb{R}^2$  with the Euclidean metric?      **(f)**  $X$  is a subset of  $\mathbb{R}^n$ , for arbitrary  $n$ , with the Euclidean metric?

prob:Gdelta-Fsigma-metric

- 41.** A subset  $A$  of a metric space  $\langle X, d \rangle$  is called a  $G_\delta$ -**set** if it is the intersection of some sequence of  $d$ -open sets, and  $A$  is called an  $F_\sigma$ -**set** if it is the union of some sequence of  $d$ -closed sets.

*Note:* In  $G_\delta$ , the  $G$  is from the German *Gebeit*, meaning “domain”; and the subscript  $\delta$  is from the German *Durchschnitt*, meaning “intersection” mathematically (and “section” or “average” more generally). In  $F_\sigma$ , the  $F$  is from the French *fermé*, meaning “closed”; and the subscript  $\sigma$  is from the French *somme*, meaning “union” mathematically (and “sum” more generally).

prob-part:d-closed-is-G-delta

- (a) Prove that every  $d$ -closed set  $A$  is a  $G_\delta$ -set. [Hint: . If  $A \neq \emptyset$ , consider the sets  $\{x \in X : d(x, A) < 1/n\}$  for  $n = 1, 2, 3, \dots$ .]
- (b) Deduce from (a) that every  $d$ -open set is an  $F_\sigma$ -set.
- (c) For the Euclidean metric  $d$  on  $\mathbb{R}$ , show that  $\mathbb{Q}$  is an  $F_\sigma$ -set that is not  $d$ -open.
- (d) For the Euclidean metric  $d$  on  $\mathbb{R}$ , give a  $G_\delta$ -set that is not  $d$ -closed.

part:dist-to-Ables dist-to-Bisite pets

- 42. (a)** If  $A$  and  $B$  are nonempty subsets of a metric space  $\langle X, d \rangle$ , show that the set

$$\{x \in X : d(x, A) < d(x, B)\}$$

is  $d$ -open.

prob-part:metric-is-normal

- (b) Use (a) to prove Proposition 1.30: If  $A$  and  $B$  are disjoint  $d$ -closed sets in a metric space  $\langle X, d \rangle$ , then there are disjoint  $d$ -open subsets  $U$  and  $V$  of  $X$  such that  $A \subset U$  and  $B \subset V$ .

prob:dist-between-sets

- 43.** Given two nonempty sets  $A$  and  $B$  in a metric space  $\langle X, d \rangle$ , the **distance from  $A$  to  $B$**  is defined to be the number

$$d(A, B) = \inf\{d(x, y) : x \in A, y \in B\}.$$

- (a) Construct disjoint nonempty  $d$ -open subsets  $A$  and  $B$  of  $\mathbb{R}^2$ , where  $d$  is the Euclidean metric, such that  $d(A, B) = 0$  yet  $d(x, y) > 0$  for all  $x \in A$  and all  $y \in B$ .

- (b)** Show that, in general,

$$d(A, B) = \inf\{d(x, B) : x \in A\}.$$

pseudometric  
ultrametric

*Note:* : When  $A$  and  $B$  are “compact” in the sense defined in [Chapter 4](#) (Compactness), the distance between them is actually the distance between some point of  $A$  and some point of  $B$ : see [Exercise 4.29 \(b\)](#).

prob:open-etc-for-pseudometric

**44.** If  $d$  is a pseudometric on a set  $X$  ([Exercise 20](#)), then the notions of ‘ $d$ -ball of radius  $\varepsilon$  at  $x$ ’, ‘ $d$ -open set’, ‘ $d$ -closed set’, and ‘ $d$ -neighborhood of  $x$ ’ may be defined exactly as for a metric  $d$ .

(a) Verify that all results of this section concerning metric spaces, except [Corollary 1.27](#), hold for pseudometrics as well.

(b) Given a pseudometric  $d$  on a set  $X$ , let  $\langle X^*, d^* \rangle$  be the metric space constructed in [Exercise 21](#) and let  $p: X \rightarrow X^*$  be the quotient map defined by  $p(x) = x^*$  for all  $x \in X$ . Show that a subset  $V$  of  $X^*$  is  $d^*$ -open if and only if the subset  $p^{-1}(V)$  of  $X$  is  $d$ -open.

If  $U$  is a  $d$ -open subset of  $X$ , must  $p(U)$  be  $d^*$ -open? Must  $U$  be  $d$ -open if  $p(U)$  is  $d^*$ -open?

prob:ultrametric-bis

**45.** Let  $d$  be an ultrametric on a set  $X$  in the sense of [Exercise 22](#). [For example,  $d$  could be a  $p$ -adic metric on  $\mathbb{Q}$  ([Exercise 24](#)) or the metric of [Example 1.15](#).] Prove:

(a) Each  $d$ -ball is  $d$ -closed (as well as  $d$ -open).

(b) Given  $x \in X$  and  $\varepsilon > 0$ , we have  $B_\varepsilon(y; d) = B_\varepsilon(x; d)$  for every  $y \in B_\varepsilon(x; d)$ . (In other words, every point of a  $d$ -ball is a “center” of that  $d$ -ball!)

(c) If two  $d$ -balls intersect, then one of them contains the other.

### 1.3 Equivalent Metrics

sec:equivalent

A definition of continuity of a function from one metric space to another will be given in the next section. It will turn out that the continuity depends not on the specific metrics themselves, but only on the open sets they define. For this reason we now investigate the question of when two metrics on the same set define the same open sets.

#### Equivalent metrics on a set

subsec:equiv-metrics

def:equivalent-metrics

**1.34 Definition (equivalence of metrics).** Metrics  $d$  and  $d'$  on the same set  $X$  are said to be **(topologically) equivalent** when each  $d$ -open subset of  $X$  is  $d'$ -open and each  $d'$ -open subset of  $X$  is  $d$ -open. Of course  $d$  and  $d'$  are said to be **inequivalent** when they are not equivalent.

Recall that a subset  $U$  of a metric space  $\langle X, d \rangle$  was defined to be  $d$ -open if at each point of  $U$  there is some  $d$ -ball contained in  $U$ . Hence to decide whether two metrics on the same set are equivalent, we need only look at  $d$ -balls and  $d'$ -balls.

prop:equiv-metrics-via-balls

**1.35 Proposition (characterization of equivalent metrics by balls).** Let  $d$  and  $d'$  be metrics on the same set  $X$ . Then a necessary and sufficient condition for  $d$  to be equivalent to  $d'$  is that, given an arbitrary point  $x \in X$ , each  $d$ -ball at  $x$  contains some  $d'$ -ball at  $x$ , and each  $d'$ -ball at  $x$  contains some  $d$ -ball at  $x$ .

**Proof.** Necessity. Assume that  $d$  is equivalent to  $d'$ . Let  $x \in X$ . Consider an arbitrary  $d$ -ball  $B_\varepsilon(x; d)$  at  $x$ . Since  $B_\varepsilon(x; d)$  is  $d$ -open, by our assumption it must be  $d'$ -open as well. Hence at the point  $x \in B_\varepsilon(x; d)$  there is some  $d'$ -ball  $B_\eta(x; d')$  with

$B_\eta(x; d') \subset B_\varepsilon(x; d)$ . Similarly, for each  $d'$ -ball  $B_\varepsilon(x; d')$  there is some  $d$ -ball  $B_\eta(x; d)$  with  $B_\eta(x; d) \subset B_\varepsilon(x; d')$ .

Sufficiency. Assume the condition about balls holds. We show that each  $d$ -open set is  $d'$ -open; the proof that each  $d'$ -open set is  $d$ -open is similar. Let  $U$  be a  $d$ -open subset of  $X$ . Let  $x \in U$ . Since  $U$  is  $d$ -open, there is some  $\varepsilon > 0$  with

$$B_\varepsilon(x; d) \subset U.$$

By assumption there is some  $\eta > 0$  with

$$B_\eta(x; d') \subset B_\varepsilon(x; d).$$

Then  $B_\eta(x; d') \subset U$ . Since  $x$  was an arbitrary point of  $U$ , we conclude that  $U$  is indeed  $d'$ -open.  $\square$

It is geometrically evident that each square region centered at a point of  $\mathbb{R}^2$  contains some circular region centered at the same point, and vice versa. The analogous statement for cubical regions and spherical regions in  $\mathbb{R}^3$  is also evident. Hence in dimensions  $n = 2$  and  $n = 3$  the Euclidean metric  $d$  and the max metric  $d_\infty$  are equivalent. The same thing is true for arbitrary  $n$ .

prop:d-and-dmax-equiv

**1.36 Proposition (equivalence of Euclidean and max metrics).** *The Euclidean metric  $d$  and the max metric  $d_\infty$  on  $\mathbb{R}^n$  are equivalent.*

**Proof.** In view of Proposition 1.35, it suffices to show that

$$B_\varepsilon(x; d) \subset B_\varepsilon(x; d_\infty)$$

and

$$B_{\varepsilon/\sqrt{n}}(x; d_\infty) \subset B_\varepsilon(x; d)$$

for each  $x \in \mathbb{R}^n$  and each  $\varepsilon > 0$ . Hence it is enough to show that

$$d_\infty(x, y) \leq d(x, y) \leq \sqrt{n} d_\infty(x, y)$$

for all  $x, y \in \mathbb{R}^n$ . To establish these inequalities, it is enough to prove that

$$\max_{1 \leq i \leq n} a_i \leq \sqrt{\sum_{i=1}^n a_i^2} \leq \sqrt{n} \max_{1 \leq i \leq n} a_i$$

for every  $n$ -tuple  $\langle a_1, a_2, \dots, a_n \rangle$  of nonnegative real numbers. Given such an  $n$ -tuple, let  $1 \leq j \leq n$  with

$$a_j = \max_{1 \leq i \leq n} a_i.$$

Then

$$a_j = \sqrt{a_j^2} \leq \sqrt{\sum_{i=1}^n a_i^2},$$

and since  $a_i \leq a_j$  for each  $i$ , then

$$\sum_{i=1}^n a_i^2 \leq n \cdot a_j^2. \quad \square$$

equivalent metrics  
inequivalent metrics  
sup metric  
inequivalent metrics

The taxicab metric is also equivalent to the Euclidean metric ([Exercise 47](#)). However, not every metric on  $\mathbb{R}^n$  is equivalent to the Euclidean metric. For example, each singleton  $\{x\}$  is  $\delta$ -open for the discrete metric  $\delta$  on  $\mathbb{R}^n$ , but no singleton is  $d$ -open for the Euclidean metric.

Here is a more substantial example of two inequivalent metrics on the same set.

ex:inequiv-metrics-on-fns

**1.37 Example (non-equivalence of sup and  $L^1$ -metrics).** Let  $X = C([0, 1])$ , the set of all continuous functions  $x: [0, 1] \rightarrow \mathbb{R}$ . Consider the sup metric  $d_\infty$  and the  $L_1$ -metric  $d_1$  of Examples 1.8 and 1.9 given respectively by

$$d_\infty(x, y) = \max_{0 \leq t \leq 1} |x(t) - y(t)|, \quad d_1(x, y) = \int_0^1 |x(t) - y(t)| dt.$$

Note that if  $m = d_\infty(x, y)$ , then (see [Figure 1.14](#))



Figure 1.14: Distances between two functions for the sup metric  $d_\infty$  and the  $L_1$ -metric  $d_1$  on  $C([0, 1])$ .

fig:dmax-vs-done

$$d_1(x, y) \leq \int_0^1 m dt = m = d_\infty(x, y),$$

and hence

$$B_\varepsilon(x; d_\infty) \subset B_\varepsilon(x; d_1)$$

for each  $x \in X$  and each positive  $\varepsilon$ . Nonetheless,  $d_1$  is *not* equivalent to  $d_\infty$ , as we are about to demonstrate.

Define  $x: [0, 1] \rightarrow \mathbb{R}$  to be the function that is constantly zero. We claim that

$$B_\eta(x; d_1) \not\subset B_1(x; d_\infty) \quad (\eta > 0).$$

In fact, let  $\eta > 0$ . Define  $y: [0, 1] \rightarrow \mathbb{R}$  to be the function that takes the values 2 at  $t = 0$  and 0 at  $t = \eta/2$ , is linear on  $[0, \eta/2]$ , and is constantly zero on  $[\eta/2, 1]$  (see [Figure 1.15](#)). Then  $d_1(x, y)$  is just the area of a right triangle of height 2 and base  $\eta/2$ , so that

$$d_1(x, y) = \frac{1}{2} (2) \left( \frac{\eta}{2} \right) = \frac{\eta}{2} < \eta.$$

Thus  $y \in B_\eta(x; d_1)$ . However,

$$d_\infty(x, y) = 2 > 1,$$

so that  $y \notin B_1(x; d_\infty)$ .  $\diamond$



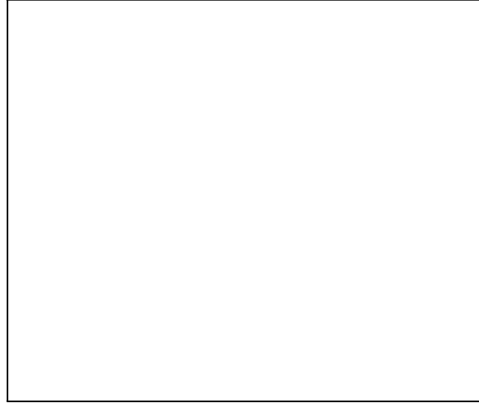
Figure 1.15: Function in a  $d_1$ -ball but not in a  $d_\infty$ -ball.

fig:func-not-in-ball

### Bounded metric spaces

The next example of equivalent metrics, provided by [Proposition 1.40](#) below, will be a general one concerning metrics for which points cannot be arbitrarily far apart.

def:bounded

**1.38 Definition (bounded sets).** Let  $\langle X, d \rangle$  be a metric space. A subset  $A$  of  $X$  is said to be  **$d$ -bounded** if there is some constant  $m$  for which  $d(x, y) \leq m$  for all  $x, y \in A$ . If  $A$  is  $d$ -bounded and nonempty, then the real number

$$\text{diam}_d A = \sup\{d(x, y) : x \in A, y \in A\}$$

(which exists thanks to order-completeness of the reals—see [Axiom 0.74](#)) is called the  **$d$ -diameter of  $A$** . When the metric  $d$  is understood, the diameter  $\text{diam}_d A$  may be denoted simply by  $\text{diam } A$ .

When the set  $X$  itself is  $d$ -bounded, then both the metric space  $\langle X, d \rangle$  and the metric  $d$  on  $X$  are said to be **bounded**.

When a set or metric or metric space is not bounded, it is said to be **unbounded**.

exs:bounded-sets

**1.39 Examples.** (1) The discrete metric on any set is bounded.

(2) Any finite set in a metric space  $\langle X, d \rangle$  is  $d$ -bounded.

(3) If  $A$  is a bounded subset of a metric space, then its diameter need not be an actual maximum distance between two particular points; in other words, the sup in [Definition 1.38](#) need not be a max. For example,  $]0, 1[$  is a bounded subset of  $\mathbb{R}$  for the Euclidean metric  $d$ , and  $\text{diam}_d ]0, 1[ = 1$  even though  $d(x, y) < 1$  for all  $x, y \in ]0, 1[$ .

(4) If  $B$  is a bounded set in a metric space and if  $A \subset B$ , then  $A$  is also bounded and  $\text{diam } A \leq \text{diam } B$ ,

(5) Any  $d$ -disk, and *a fortiori* any  $d$ -ball, in a metric space  $\langle X, d \rangle$  is  $d$ -bounded, because  $y, z \in D_r(x; d)$  implies

$$d(y, z) \leq d(y, x) + d(x, z) \leq r + r = 2r.$$

bounded set  
extended realline  
extended realline

- (6) A subset of  $\mathbb{R}$  is  $d$ -bounded for the Euclidean metric  $d$  if and only if it has both an upper bound and a lower bound in  $\mathbb{R}$ .
- (7) The Euclidean metric  $d$  on  $\mathbb{R}^n$  is *unbounded*, since for each  $c > 0$  the  $d$ -distance from  $\langle c, 0, \dots, 0 \rangle$  to  $\langle 0, 0, \dots, 0 \rangle$  is  $c$ .  $\diamond$

We show now that from any metric we can construct a bounded metric that is equivalent to it.

prop:bded-metric-equiv-to-given

**1.40 Proposition (equivalence of metric with bounded metric).** *Let  $\langle X, d \rangle$  be any metric space. There there is a bounded metric on  $X$  that is equivalent to  $d$  and for which  $\text{diam } X \leq 1$ . One such metric  $d^*$  is given by*

$$d^*(x, y) = \min\{1, d(x, y)\} \quad (x, y \in X).$$

**Proof.** First we show that  $d^*$  given by the above equation is a metric. Because the other properties of a metric are evidently satisfied by  $d^*$ , we verify only the triangle inequality

$$d^*(x, z) \leq d^*(x, y) + d^*(y, z).$$

If  $d(x, y) \leq 1$  and  $d(y, z) \leq 1$ , then

$$d^*(x, z) \leq d(x, z) \leq d(x, y) + d(y, z) = d^*(x, y) + d^*(y, z).$$

Now suppose  $d(x, y) > 1$  or  $d(y, z) > 1$ . If  $d(x, y) > 1$ , then

$$d^*(x, z) \leq 1 \leq 1 + d^*(y, z) = d^*(x, y) + d^*(y, z).$$

The case  $d(y, z) > 1$  is treated similarly. Thus  $d^*$  is a metric.

The metric  $d^*$  is bounded and  $\text{diam } X \leq 1$  for  $d^*$  because  $d^*(x, y) \leq 1$  for all  $x, y \in X$ .

Since  $d^*(x, y) = d(x, y)$  whenever  $d(x, y) < 1$ , then

$$B_\varepsilon(x; d^*) = B_\varepsilon(x; d)$$

for all  $x \in X$  and every  $\varepsilon$  with  $0 < \varepsilon \leq 1$ . Hence  $d^*$  is equivalent to  $d$ .  $\square$

### The extended real line

subsec:extended-reals

The preceding [Proposition 1.40](#) provides, in particular, a bounded metric equivalent to the Euclidean metric  $d$  on the real line  $\mathbb{R}$ . In the example that follows, we shall construct an entirely different metric that is equivalent to  $d$ . This example will be important later in showing that infinite limits and limits-at-infinity, familiar from calculus, are just special cases of a more general concept of limit [see [Examples 3.126 \(3\) and 4](#)].

ex:extended-reals

**1.41 Example (extended real line).** The **extended real line** is the set

$$\widehat{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$$

consisting of all real numbers together with two distinct objects  $-\infty$  and  $+\infty$  not belonging to  $\mathbb{R}$ . (What the particular objects  $-\infty$  and  $+\infty$  really are does not matter, only that they are different from one another and are not themselves real numbers.)

We extend the usual total ordering of  $\mathbb{R}$  to a total ordering of  $\widehat{\mathbb{R}}$  by setting

$$\begin{aligned} -\infty &< +\infty \\ -\infty &< x < +\infty \quad (x \in \mathbb{R}). \end{aligned}$$

Thus we have adjoined a largest element  $+\infty$  and a smallest element  $-\infty$  to  $\mathbb{R}$ . Then

$$\widehat{\mathbb{R}} = [-\infty, +\infty], \quad \mathbb{R} = ]-\infty, +\infty[$$

as intervals in  $\widehat{\mathbb{R}}$ . For each  $a \in \mathbb{R}$ , the intervals  $] \leftarrow, a[$  and  $]a, \rightarrow[$  in  $\mathbb{R}$  are just the intervals  $] -\infty, a[$  and  $]a, +\infty[$  in  $\widehat{\mathbb{R}}$ . Thus each open ray in  $\mathbb{R}$  becomes an actual interval in  $\widehat{\mathbb{R}}$  having both a left and a right endpoint, and similarly for closed rays. For this reason, we may henceforth dispense with the arrows  $\leftarrow$  and  $\rightarrow$  when denoting and use instead  $-\infty$  and  $+\infty$ , respectively.

We are going to make  $\widehat{\mathbb{R}}$  into a bounded metric space that “looks like” the closed interval  $[-1, 1]$  with its Euclidean metric. More precisely, we shall construct a certain one-to-one correspondence  $\widehat{\varphi}$  between  $\widehat{\mathbb{R}}$  and  $[-1, 1]$  that maps intervals to intervals, and then define the distance between a pair of points in  $\widehat{\mathbb{R}}$  to be the same as the distance between the pair of points in  $[-1, 1]$  that correspond to them under  $\widehat{\varphi}$ .

Step 1: construct a certain one-to-one correspondence  $\varphi$  between  $\mathbb{R}$  and  $] -1, 1[$ . Since

$$\left| \frac{x}{1 + |x|} \right| = \frac{|x|}{1 + |x|} < 1 \quad (x \in \mathbb{R}),$$

we may define a map

$$\varphi: \mathbb{R} \rightarrow ] -1, 1[$$

by setting

$$\varphi(x) = \frac{x}{1 + |x|} \quad (x \in \mathbb{R}).$$

The graph of  $\varphi$ , shown in [Figure 1.16](#), is easily sketched using the usual methods of elementary calculus. Direct computation shows that  $\varphi$  has the inverse

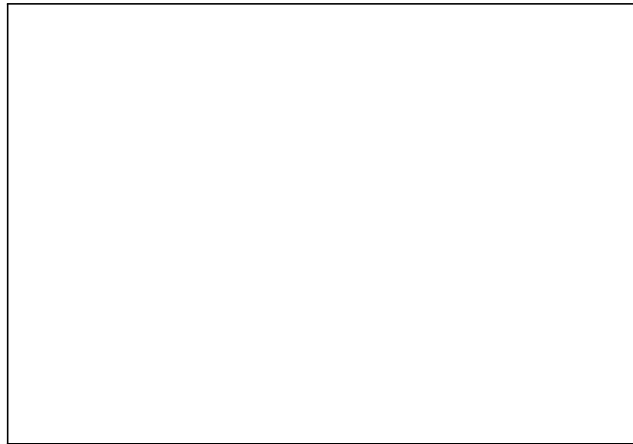


Figure 1.16: Graph of  $y = \varphi(x) = x/(1 + |x|)$ .

fig:map-R-to-interval

$$\psi: ]-1, 1[ \rightarrow \mathbb{R}$$

given by

$$\psi(y) = \frac{y}{1 - |y|} \quad (-1 < y < 1).$$

Hence  $\varphi: \mathbb{R} \rightarrow ]-1, 1[$  is a bijection.

Step 2: show that  $\varphi$  maps intervals in  $\mathbb{R}$  to intervals in  $\widehat{\mathbb{R}}$ . Some manipulation of inequalities (or a calculation showing that the derivatives of  $\varphi$  and  $\psi$  are positive) shows that both  $\varphi$  and its inverse  $\psi$  are strictly increasing functions. Then

$$\varphi([a, b[) = ]\varphi(a), \varphi(b)[$$

whenever  $a < b$ . In fact,  $a < x < b$  implies  $\varphi(a) < \varphi(x) < \varphi(b)$  since  $\varphi$  is strictly increasing. Conversely,  $\varphi(a) < y < \varphi(b)$  implies  $y = \varphi(x)$  for some  $x \in \mathbb{R}$ , so that

$$\psi(\varphi(a)) < \psi(\varphi(x)) < \psi(\varphi(b))$$

since  $\psi$  is strictly increasing, and hence  $a < x < b$  since  $\psi \circ \varphi$  is the identity map on  $\mathbb{R}$ . In the same way one proves that

$$\varphi([a, +\infty[) = ]\varphi(a), 1[, \quad \varphi(]-\infty, b]) = ]-1, \varphi(b)[$$

for each  $a \in \mathbb{R}$  and each  $b \in \mathbb{R}$ .

Step 3: extend  $\varphi: \mathbb{R} \rightarrow ]-1, 1[$  to a map  $\widehat{\mathbb{R}} \rightarrow [-1, 1]$ . The graph of  $\varphi$  has the lines  $y = 1$  and  $y = -1$  as horizontal asymptotes, for

$$\begin{aligned} \lim_{x \rightarrow +\infty} \varphi(x) &= \lim_{x \rightarrow +\infty} \frac{x}{1+x} = 1, \\ \lim_{x \rightarrow -\infty} \varphi(x) &= \lim_{x \rightarrow -\infty} \frac{x}{1-x} = -1. \end{aligned}$$

Hence it is natural to extend  $\varphi: \mathbb{R} \rightarrow ]-1, 1[$  to a map

$$\widehat{\varphi}: \widehat{\mathbb{R}} \rightarrow [-1, 1]$$

by setting

$$\widehat{\varphi}(-\infty) = -1, \quad \widehat{\varphi}(+\infty) = 1.$$

Clearly  $\widehat{\varphi}$  is a bijection, and

$$\widehat{\varphi}([a, +\infty]) = ]\varphi(a), 1], \quad \widehat{\varphi}([-\infty, a]) = [-1, \varphi(a)[$$

for each  $a \in \mathbb{R}$ .

Step 4: define a metric  $\widehat{d}$  on the extended real line  $\widehat{\mathbb{R}}$ . To do this, we make the distance between two points of  $\widehat{\mathbb{R}}$  be the same as the Euclidean distance between the points in  $[-1, 1]$  to which they correspond under  $\widehat{\varphi}$ . In other words, we define

$$\widehat{d}(x, y) = |\widehat{\varphi}(x) - \widehat{\varphi}(y)| \quad (x, y \in \widehat{\mathbb{R}}).$$

The proof that  $\widehat{d}$  is actually a metric on  $\widehat{\mathbb{R}}$  uses no property of  $\widehat{\varphi}$  except that it is injective. First,  $\widehat{d}(x, y) \geq 0$  for all  $x, y \in \widehat{\mathbb{R}}$ . Since  $\widehat{\varphi}$  is an injective function,

$$\widehat{d}(x, y) = 0 \iff \widehat{\varphi}(x) = \widehat{\varphi}(y) \iff x = y.$$

The symmetry of  $\widehat{d}$  is obvious. Finally,

$$\begin{aligned} \widehat{d}(x, z) &= |\widehat{\varphi}(x) - \widehat{\varphi}(z)| \\ &= |[\widehat{\varphi}(x) - \widehat{\varphi}(y)] + [\widehat{\varphi}(y) - \widehat{\varphi}(z)]| \\ &\leq |\widehat{\varphi}(x) - \widehat{\varphi}(y)| + |\widehat{\varphi}(y) - \widehat{\varphi}(z)| \\ &= \widehat{d}(x, y) + \widehat{d}(y, z) \end{aligned}$$

for all  $x, y, z \in \widehat{\mathbb{R}}$ , which establishes the triangle inequality.

Step 5: the metric space  $(\widehat{\mathbb{R}}, \widehat{d})$  is bounded. In fact,

$$\widehat{d}(x, y) = |\widehat{\varphi}(x) - \widehat{\varphi}(y)| \leq |\widehat{\varphi}(x)| + |\widehat{\varphi}(y)| \leq 1 + 1 = 2$$

for all  $x, y \in \widehat{\mathbb{R}}$ . Moreover, since  $\widehat{d}(-\infty, +\infty) = 2$ , we have

$$\text{diam } \widehat{\mathbb{R}} = 2. \quad \diamond$$

bounded metric!  
equivalent metric@  
metric!bounded

**1.42 Proposition.** *The metric  $d^*$  induced on  $\mathbb{R}$  by the metric  $\widehat{d}$  on  $\widehat{\mathbb{R}}$  is a bounded metric equivalent to the Euclidean metric  $d$  on  $\mathbb{R}$ .*

**Proof.** The metric  $d^*$  is bounded because  $\widehat{d}$  is bounded.

Let  $x \in \mathbb{R}$ . First we show that each  $d^*$ -ball at  $x$  of sufficiently small radius contains some  $d$ -ball at  $x$ . Since  $B_\varepsilon(x; d^*) = B_\varepsilon(x; \widehat{d}) \cap \mathbb{R}$ , we consider only  $d^*$ -balls at  $x$  of radius  $\varepsilon$  for which  $-\infty, +\infty \notin B_\varepsilon(x; \widehat{d})$ , that is, for which

$$B_\varepsilon(x; d^*) = B_\varepsilon(x; \widehat{d}).$$

Let

$$0 < \varepsilon < \min\{\varphi(x) - (-1), 1 - \varphi(x)\}.$$

Then

$$-1 < \varphi(x) - \varepsilon < \varphi(x) + \varepsilon < 1,$$

and

$$B_\varepsilon(x; \widehat{d}) = ]\psi(\varphi(x) - \varepsilon), \psi(\varphi(x) + \varepsilon)[$$

is an open interval in  $\mathbb{R}$  that certainly contains a  $d$ -ball at  $x$ .

Next we show that each  $d$ -ball at  $x$  contains some  $d^*$ -ball at  $x$ . More generally, consider a  $d$ -open set containing  $x$  having the form  $]a, b[$  with

$$-\infty < a < x < b < +\infty.$$

Choose  $\varepsilon$  with

$$0 < \varepsilon < \min\{\varphi(x) - \varphi(a), \varphi(b) - \varphi(x)\}$$

Then

$$-1 < \varphi(a) < \varphi(x) - \varepsilon < \varphi(x) + \varepsilon < \varphi(b) < 1,$$

so that

$$B_\varepsilon(x; d^*) = B_\varepsilon(x; \widehat{d}) = ]\psi(\varphi(x) - \varepsilon), \psi(\varphi(x) + \varepsilon)[ \subset ]a, b[. \quad \square$$

For later use we also establish a relationship between intervals in  $\widehat{\mathbb{R}}$  and  $\widehat{d}$ -balls at the points  $+\infty$  and  $-\infty$ .

extended real line  
translation  
rotation  
translation-invariant  
rotation-invariant

**1.43 Lemma (balls at infinity as rays in extended real line).** If  $0 < \varepsilon < 1$ , then

$$B_\varepsilon(+\infty; \widehat{d}) = ]1/\varepsilon - 1, +\infty],$$

and if  $u > 0$ , then

$$]u, +\infty] = B_{1/(1+u)}(+\infty; \widehat{d}).$$

If  $0 < \varepsilon < 1$ , then

$$B_\varepsilon(-\infty; \widehat{d}) = [-\infty, 1 - 1/\varepsilon[,$$

and if  $u < 0$ , then

$$[-\infty, u[ = B_{1/(1-u)}(-\infty; \widehat{d}).$$

**Proof.** We prove only the assertions concerning  $+\infty$ , leaving proof of the others to the reader.

Let  $0 < \varepsilon < 1$ . Set

$$B = B_\varepsilon(+\infty; \widehat{d}), \quad J = ]1/\varepsilon - 1, +\infty].$$

We must show

$$(*) \quad x \in B \iff x \in J$$

Since  $+\infty$  belongs to both  $B$  and  $J$  and the point  $-\infty$  belongs to neither, we need only consider  $x \in \mathbb{R}$ . Now each  $x \in J$  is positive since  $(1/\varepsilon) - 1 > 0$ , and each  $x \in B$  is positive since  $\varepsilon < 1$ . Hence we need only prove  $(*)$  for  $x > 0$ . But  $x > 0$  implies

$$\widehat{d}(x, +\infty) = \left| \frac{x}{1+x} - 1 \right| = \frac{1}{1+x},$$

and  $(*)$  follows at once.

Now let  $u > 0$ . Set

$$\varepsilon = \frac{1}{1+u}.$$

Then  $0 < \varepsilon < 1$  and by what we just proved,

$$]u, +\infty] = B_\varepsilon(+\infty; \widehat{d}). \quad \square$$

## Isometries

subsec:isometry

The metric  $\widehat{d}$  on the extended real line  $\widehat{\mathbb{R}}$  was so constructed that the  $\widehat{d}$ -distance between any two points  $x_1, x_2$  of  $\mathbb{R}$  is the same as the  $d$ -distance between the points  $\widehat{\varphi}(x_1), \widehat{\varphi}(x_2)$  of  $[-1, 1]$  corresponding to them under the bijection  $\widehat{\varphi}: \widehat{\mathbb{R}} \rightarrow [-1, 1]$ . This relationship between  $(\widehat{\mathbb{R}}, \widehat{d})$  and  $([-1, 1], d)$  suggests the following definition.

def:distance-preserving map

**1.44 Definition (distance-preserving map).** Let  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$  be metric spaces. A map  $f: X \rightarrow Y$  is said to be  $\langle d, d' \rangle$ -**distance-preserving**, or more simply to be **distance-preserving** and to **preserve distance**, when the metrics are understood, if

$$d(x_1, x_2) = d'(f(x_1), f(x_2)) \quad (x_1, x_2 \in X).$$

Translations of the Euclidean plane  $\mathbb{R}^2$  are distance-preserving; so are rotations of  $\mathbb{R}^2$ . That is just another way of saying, in the language of [Exercise 4](#), that the Euclidean metric on  $\mathbb{R}^2$  is translation-invariant and rotation invariant, respectively.

A distance-preserving map between metric spaces is necessarily injective. However, such a map need not be surjective; for example, the map  $\mathbb{R} \rightarrow \mathbb{R}^2$  defined by  $x \mapsto \langle x, 0 \rangle$  is distance-preserving for the Euclidean metrics but is not surjective. When a distance-preserving map is, in addition, surjective we refer to it as an “isometry.”

distance-preserving map  
translation  
reflection

def:isometry

**1.45 Definition (isometry).** Let  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$  be metric spaces. A map  $f: X \rightarrow Y$  is called an **isometry from  $\langle X, d \rangle$  to  $\langle Y, d' \rangle$**  if it is bijective and distance-preserving. The metric space  $\langle X, d \rangle$  is said to be **isometric to** the metric space  $\langle Y, d' \rangle$  when there exists some isometry from  $\langle X, d \rangle$  to  $\langle Y, d' \rangle$ .

Thus an isometry from a metric space  $\langle X, d \rangle$  to a metric space  $\langle Y, d' \rangle$  is a *distance-preserving* one-to-one correspondence between their points.

**Intuitive idea—*isometric metric spaces*.** We may regard metric spaces  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$  to be essentially indistinguishable, in so far as they are metric spaces, when they are isometric.

Observe that if two metric spaces are isometric and if one of them is bounded, then the other is bounded, too, and the two must have the same diameter.

translation-and-reflections:isometries

**1.46 Examples.** (1) Let  $d$  be the Euclidean metric on  $\mathbb{R}^n$ . Then for each fixed  $c \in \mathbb{R}^n$ , the **translation**

$$\begin{aligned} T: \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ x &\mapsto x + c \end{aligned}$$

by  $c$  is an isometry from  $\langle \mathbb{R}^n, d \rangle$  to itself, because

$$\begin{aligned} d(T(x), T(y)) &= d(x + c, y + c) \\ &= \|(x + c) - (y + c)\| \\ &= \|x - y\| \\ &= d(x, y) \end{aligned}$$

for all  $x, y \in \mathbb{R}^n$ . (This generalizes [Exercise 4](#).)

Also, for each fixed  $1 \leq j \leq n$ , the **reflection**

$$\begin{aligned} \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ \langle x_1, x_2, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_n \rangle &\mapsto \langle x_1, x_2, \dots, x_{j-1}, -x_j, x_{j+1}, \dots, x_n \rangle \end{aligned}$$

in the  $j$ th coordinate is an isometry from  $\langle \mathbb{R}^n, d \rangle$  to itself.

There are many isometries from  $\langle \mathbb{R}^n, d \rangle$  to itself besides translations and such reflections. For example, in  $\mathbb{R}^2$ , a rotation around the origin is such an isometry (see [Exercise 4](#)); and in  $\mathbb{R}^3$ , a rotation around one of the coordinate axes is such an isometry (see [Exercise 56](#)). With the aid of linear algebra it is possible to determine all isometries from  $\langle \mathbb{R}^n, d \rangle$  to itself. (For the cases  $n = 2$  and  $n = 3$ , see Jacob and Bailey [38, Theorems 8.1.3 and 8.2.2] or Cooper [16] and Cooper [17]. For the general case see Fleming [25, pages 125–126] or Conrad [14]. See also [Exercise 64](#).)

line-segment-isometric-closed-interval

(2) For  $a, b \in \mathbb{R}^n$  with  $a \neq b$ , let  $d'$  be the Euclidean metric on the line segment

$$Y = \{(1 - t)a + tb : 0 \leq t \leq 1\}$$

joining  $a$  to  $b$ . Let  $d$  be the Euclidean metric on the closed interval

$$X = [0, c]$$

rigid motion

metrically equivalent metric spaces

metric spaces!metrically equivalent

isometry

in  $\mathbb{R}$ , where  $c > 0$ . If  $c = d'(a, b)$ , that is, if  $c = \|a - b\|$ , then the map

$$\begin{aligned} X &\rightarrow Y \\ x &\mapsto \left(1 - \frac{x}{c}\right)a + \frac{x}{c}b \end{aligned}$$

is an isometry from  $\langle X, d \rangle$  to  $\langle Y, d' \rangle$ . However,  $\langle X, d \rangle$  is *not* isometric to  $\langle Y, d' \rangle$  if  $c \neq d'(a, b)$ , since in that case  $\text{diam } X = c$  whereas  $\text{diam } Y = d'(a, b)$ .

ex:isometric-embed- $\mathbb{R}^n$ -in- $\mathbb{R}^{n+1}$ 

(3) Let  $d$  be the Euclidean metric on  $\mathbb{R}^n$  and let  $d'$  be the Euclidean metric on the subset

$$Y = \{ \langle x_1, x_2, \dots, x_n, x_{n+1} \rangle \in \mathbb{R}^{n+1} : x_{n+1} = 0 \}$$

of  $\mathbb{R}^{n+1}$ . Then

$$\begin{aligned} \mathbb{R}^n &\rightarrow Y \\ \langle x_1, x_2, \dots, x_n \rangle &\mapsto \langle x_1, x_2, \dots, x_n, 0 \rangle \end{aligned}$$

is an isometry from  $\langle \mathbb{R}^n, d \rangle$  to  $\langle Y, d' \rangle$ .  $\diamond$

An isometry from  $\mathbb{R}^n$ , with its Euclidean metric, to itself is often called a *rigid motion*.

It so happens that any distance-preserving map from  $\langle \mathbb{R}^n, d \rangle$  to itself is necessarily surjective and hence is an isometry (We shall not use this fact and hence offer no proof of it; for a proof, see Fleming [25, pages 125–126] or Conrad [14].)

Fundamental properties about isometries are as follows.

prop:isometric-embed-is-isometry

**1.47 Proposition (properties of isometries).** (1) If  $\langle X, d \rangle$  is any metric space, then the identity map  $x: X \rightarrow X$  is an isometry from  $\langle X, d \rangle$  to itself.

o-part:inverse-of-isometry-is-isometry

(2) Let  $f$  be an isometry from a metric space  $\langle X, d \rangle$  to a metric space  $\langle Y, d' \rangle$ . Then  $f^{-1}: Y \rightarrow X$  is an isometry from  $\langle Y, d' \rangle$  to  $\langle X, d \rangle$ .

t:composite-of-isometries-is-isometry

(3) Let  $f: X \rightarrow Y$  be an isometry from  $\langle X, d \rangle$  to  $\langle Y, d' \rangle$  and let  $g: Y \rightarrow Z$  be an isometry from  $\langle Y, d' \rangle$  to  $\langle Z, d'' \rangle$ . Then the composite  $g \circ f: X \rightarrow Z$  is an isometry from  $\langle X, d \rangle$  to  $\langle Z, d'' \rangle$ .

**Proof.** (3) If  $y_1, y_2 \in Y$ , let  $x_1 = f^{-1}(y_1)$ ,  $x_2 = f^{-1}(y_2)$ ; then

$$d'(y_1, y_2) = d'(f(x_1), f(x_2)) = d(x_1, x_2) = d(f^{-1}(y_1), f^{-1}(y_2)). \quad \square$$

The three parts of Proposition 1.47 say that the relation ‘is isometric to’ is reflexive, symmetric, and transitive, respectively. Thus the relation ‘is isometric to’ is an equivalence relation among metric spaces. Accordingly, metric spaces that are isometric to one another are often said to be **metrically equivalent**.

There is a more general sense in which metric spaces may be called equivalent to one another.

def:top-equiv-metric-spaces

**1.48 Definition (topological equivalence of metric spaces).** Let  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$  be metric spaces. A map  $f: X \rightarrow Y$  is called a **topological equivalence from  $\langle X, d \rangle$  to  $\langle Y, d' \rangle$**  if  $f$  is a bijection such that  $f(U)$  is a  $d'$ -open subset of  $Y$  for each  $d$ -open subset  $U$  of  $X$  and  $f^{-1}(V)$  is a  $d$ -open subset of  $X$  for each  $d'$ -open subset  $V$  of  $Y$ . The metric space  $\langle X, d \rangle$  is said to be **topologically equivalent to** the metric space  $\langle Y, d' \rangle$  if exists some topological equivalence from  $\langle X, d \rangle$  to  $\langle Y, d' \rangle$

Thus a topological equivalence from  $\langle X, d \rangle$  to  $\langle Y, d' \rangle$  is a one-to-one correspondence between their points under which the  $d$ -open subsets of  $X$  correspond to the  $d'$ -open subsets of  $Y$ .



ex:equivalent-metric-spaces

**1.49 Examples.** (1) If  $d$  and  $d'$  are equivalent metrics on a set  $X$ , in the sense of [Definition 1.34](#), then  $\langle X, d \rangle$  is topologically equivalent to  $\langle X, d' \rangle$ . In fact, the identity map of  $X$  will serve as the needed bijection  $f: X \rightarrow X$ .

- (2) More generally than (1), metrically equivalent metric spaces are topologically equivalent. In fact, let  $f: X \rightarrow Y$  be an isometry from  $\langle X, d \rangle$  to  $\langle Y, d' \rangle$ . Then

$$f(B_\varepsilon(x; d)) = B_\varepsilon(f(x); d')$$

for all  $x \in X$  and all  $\varepsilon > 0$ , and

$$f^{-1}(B_\eta(y; d')) = B_\eta(f^{-1}(y); d)$$

for all  $y \in Y$  and all  $\eta > 0$ .

Now let  $U$  be a  $d$ -open subset of  $X$ . We show that the image  $f(U)$  of  $U$  under  $f$  is  $d'$ -open. Let  $y \in f(U)$ . Then  $x = f^{-1}(y) \in U$ , and so

$$B_\varepsilon(x; d) \subset U$$

for some  $\varepsilon > 0$ . Hence

$$B_\varepsilon(y; d') = f(B_\varepsilon(x; d)) \subset f(U).$$

A similar argument shows that  $f^{-1}(V)$  is  $d$ -open for each  $d'$ -open subset  $V$  of  $Y$ .

ex:intervals-top-equiv-metric-context

- (3) For  $c > 0$  and  $a < b$ , consider the closed intervals  $[0, c]$  and  $[a, b]$  in  $\mathbb{R}$  with their Euclidean metrics  $d$  and  $d'$ , respectively.

- (a) The metric space  $([0, c], d)$  is topologically equivalent to  $([a, b], d')$ . To see this, observe that the map

$$\begin{aligned} f: [0, c] &\rightarrow [a, b] \\ x &\mapsto a + \frac{x}{c} (b - a) \end{aligned}$$

—the composite of a “dilation” or “contraction” followed by a translation—is a bijection. Moreover,  $f$  has the property

$$d'(f(x_1), f(x_2)) = \frac{b-a}{c} d(x_1, x_2),$$

and its inverse

$$\begin{aligned} f^{-1}: [a, b] &\rightarrow [0, c] \\ y &\mapsto \frac{c}{b-a} (y - a), \end{aligned}$$

which is the composite of a translation followed by a dilation or contraction, has the property

$$d(f^{-1}(y_1), f^{-1}(y_2)) = \frac{c}{b-a} d'(y_1, y_2).$$

- (b) However, if  $c \neq b - a$ , then the metric space  $([0, c], d)$  is *not* metrically equivalent to  $([a, b], d')$ .  $\diamond$

The following fundamental properties of topological equivalences among metric spaces are analogs of the corresponding properties of isometries ([Proposition 1.47](#)).

**1.50 Proposition (properties of topological equivalence).** (1) If  $\langle X, d \rangle$  is any metric space, then the identity map  $\iota_X: X \rightarrow X$  is a topological equivalence from  $\langle X, d \rangle$  to itself.

(2) Let  $f$  be a topological equivalence from a metric space  $\langle X, d \rangle$  to a metric space  $\langle Y, d' \rangle$ . Then  $f^{-1}: Y \rightarrow X$  is a topological equivalence from  $\langle Y, d' \rangle$  to  $\langle X, d \rangle$ .

(3) Let  $f: X \rightarrow Y$  be a topological equivalence from  $\langle X, d \rangle$  to  $\langle Y, d' \rangle$  and let  $g: Y \rightarrow Z$  be a topological equivalence from  $\langle Y, d' \rangle$  to  $\langle Z, d'' \rangle$ . Then the composite  $g \circ f: X \rightarrow Z$  is a topological equivalence from  $\langle X, d \rangle$  to  $\langle Z, d'' \rangle$ .

Part (1) of the preceding proposition just restates [Examples 1.49 \(1\)](#). The other two parts are left as exercises.

We shall return to the topic of topological equivalence in a more general setting in [Chapter 2](#) (Topological Spaces).

### EXERCISES FOR SECTION 1.3

**46.** Given a metric  $d$  on a set  $X$ , show that, for each real number  $\alpha > 0$ , the multiple  $\alpha d$  of  $d$  is a metric on  $X$  that is equivalent to  $d$ .

**47.** Prove that the taxicab metric  $d_1$  is equivalent to the Euclidean metric  $d$  on  $\mathbb{R}^n$ . [*Hint:* Use [Exercise 3](#).]

**48.** Given a metric  $d$  on a set  $X$ , show that the formula

$$d'(x, y) = \frac{d(x, y)}{1 + d(x, y)}$$

defines a metric on  $X$  that is bounded and equivalent to  $d$ .

**49.** Let  $d$  be a metric on a set  $X$ .

(a) Given a real number  $r > 0$ , show that the formula  $d'(x, y) = r d(x, y)$  defines a metric on  $X$  that is equivalent to  $d$ .

(b) Deduce that for each  $s > 0$  there is a metric  $d'$  on  $X$  equivalent to  $d$  for which  $d(x, y) < s$  for all  $x, y \in X$ .

**50.** Let  $X$  be the set of all index-origin 1 sequences  $\langle x_i \rangle_{i \in \mathbb{N}^*}$  in a nonempty set  $S$ .

(a) Show that the formula

$$(**) \quad d'(x, y) = \begin{cases} 0 & \text{if } x = y, \\ 1/\min\{i : x_i \neq y_i\} & \text{if } x \neq y. \end{cases}$$

defines a metric on  $X$ .

(b) Is this metric  $d'$  equivalent to the metric  $d$  of [Example 1.15](#)?

**51.** (a) Let  $\langle X_1, d_1 \rangle, \langle X_2, d_2 \rangle, \langle X_3, d_3 \rangle, \dots$  be a sequence of bounded metric spaces, each of diameter at most 1. Show that the formula

$$d'(x, y) = \sum_{i=1}^{\infty} d_i(x_i, y_i)/2^i$$

defines a bounded metric  $d'$  on the product set

$$X = \prod_{i=1}^{\infty} X_i.$$

- (b) If  $\langle X_i, d_i \rangle = \langle Y, \delta \rangle$  for every  $i = 1, 2, 3, \dots$ , where  $Y$  is some set and  $\delta$  is its discrete metric, show that the metric  $d'$  in (a) is equivalent to the metric  $d$  of Example 1.15. Hilbert cube  
Hilbert sequence space  
Hausdorff metric
- (c) Take  $X_i = [-1/i, 1/i]$  and  $d_i$  = the Euclidean metric on  $X_i$  for each  $i$ . Then  $X$  is the Hilbert cube [Exercise 36 (b)]. Is  $d'$  equivalent to the metric induced on  $X$  by the metric  $d_2$  on the Hilbert sequence space (Example 1.10)? rotation  
angle of rotation

52. Let  $d$  and  $d'$  be metrics on the same set  $X$ .

- (a) Prove that a necessary condition for  $d$  and  $d'$  to be equivalent is that each  $d$ -closed set be  $d'$ -closed and each  $d'$ -closed set be  $d$ -closed. Is this condition also sufficient?
- (b) Formulate a criterion for  $d$  and  $d'$  to be equivalent in terms of  $d$ -neighborhoods and  $d'$ -neighborhoods.

-part:diam-ball-equal-diam-disk-Rn

53. (a) For the Euclidean metric  $d$  on  $\mathbb{R}^n$ , show that

$$\text{diam } B_\varepsilon(x; d) = \text{diam } D_\varepsilon(x; d) = \text{diam } S_\varepsilon(x; d) = 2\varepsilon$$

for each  $x \in \mathbb{R}^n$  and each  $\varepsilon > 0$ .

(b) Does (a) generalize to arbitrary metric spaces?

prob:Hausdorff-metric

54. Let  $\mathcal{F}$  be the collection of all nonempty  $d$ -closed subsets of a bounded metric space  $\langle X, d \rangle$ . For  $A, B \in \mathcal{F}$ , let

$$d^*(A, B) = \max \left\{ \sup_{x \in A} d(x, B), \sup_{y \in B} d(y, A) \right\}.$$

- (a) Show that  $d^*$  is a metric on  $\mathcal{F}$ . This metric is called the **Hausdorff metric** induced by  $d$ .
- (b) Compare  $d^*(A, B)$  with  $d(A, B)$  [Exercise 43] and with  $\text{diam}(A \cup B)$ .
- (c) Take  $d$  to be the Euclidean metric on  $[0, 1]$ . Compute  $d^*(A, B)$  if  $A = [u, v]$  and  $B = [t, s]$  are closed subintervals of  $[0, 1]$ .
- (d) Take  $d$  now to be the Euclidean metric on  $\mathbb{R}^2$ . Compute  $d^*(A, B)$  if  $A = I^2 = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y \leq 1\}$  and  $B = D_2 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ .
- (e) In general, would  $d^*$  still be a metric if  $\mathcal{F}$  were instead the collection of all nonempty subsets of  $X$ .

prob:rotation-R2-isometry

55. (a) Provide the plane  $\mathbb{R}^2$  with its Euclidean metric. Show that each rotation  $R_\theta$  of  $\mathbb{R}^2$  (Exercise 4) is an isometry.
- (b) Show that, for an arbitrary angle  $\theta$ , the rotation  $R_\theta$  maps the circle (the 1-sphere)  $S_1$  onto itself and hence that its domain-codomain restriction  $R_\theta|_{S_1, S_1}$  is an isometry of  $S_1$ .  
We refer to such a restriction more simply as a “rotation of  $S_1$ .”
- (c) Show that, for any two points  $z$  and  $w$  of the 1-sphere  $S_1$ , there is a rotation of  $S_1$  that maps  $z$  to  $w$ .

prob:rotations-R3-isometries

56. A rotation of  $\mathbb{R}^3$  around the  $z$ -axis is a map  $R_{z, \gamma}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  of the form

$$R_{z, \gamma}(p) = \begin{pmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot p$$

for some fixed real number  $\gamma$ —the *angle of rotation*. Here the dot denotes matrix multiplication and the point  $p \in \mathbb{R}^3$  is written as a column vector.

**isometric embedding** Similarly one has rotations  $\mathbb{R}_{x,\alpha}$  and  $R_{y,\beta}$  around the  $x$ - and  $y$ -axes, respectively.

**Euclidean inner product** Show that each such rotation  $\mathbb{R}_{x,\alpha}$ ,  $R_{y,\beta}$ , and  $R_{z,\gamma}$  of  $\mathbb{R}^3$  around one of the coordinate axes is an isometry for the Euclidean metric.

**isometric-to-2-sphere-less-northpole** 57. The purpose of this exercise is to show that the complement  $S_2 \setminus \{q\}$  of an arbitrary point on the 2-sphere  $S_2$  is isometric to the complement  $S_2 \setminus \{\mathbf{p}\}$  of the north pole  $\mathbf{p} = \langle 0, 0, 1 \rangle$ .

(a) Show that the map  $\langle x, y, z \rangle \mapsto \langle x, -y, -z \rangle$  is an isometry of  $\mathbb{R}^3$  that maps the 2-sphere  $S_2$  onto itself. Conclude that the domain-codomain restriction of this map to  $S_2$  is an isometry of  $S_2$  that maps the north pole  $\langle 0, 0, 1 \rangle$  to the south pole  $\langle 0, 0, -1 \rangle$ .

(b) Given an arbitrary point  $\langle s, t, u \rangle \in S_2$  other than the south pole  $\langle 0, 0, -1 \rangle$ , show that the formula

$$R(x, y, z) = \begin{bmatrix} \frac{t^2+u+u^2}{1+u} & -\frac{st}{1+z} & x \\ -\frac{st}{1+u} & \frac{1-t^2+u}{1+u} & t \\ -s & -t & u \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

—where the dot denotes matrix multiplication—defines a surjection  $S_2 \rightarrow S_2$  that is an isometry and carries the north pole  $\langle 0, 0, 1 \rangle$  of  $S_2$  to the point  $\langle s, t, u \rangle$ .

**prob:Sn-isometrically-homogeneous** 58. Prove that, for any two points  $a$  and  $b$  of the  $n$ -sphere  $S_n$ , with its usual metric, there is an isometry of  $S_n$  mapping  $a$  to  $b$ .

**prob:isometric-embedding** 59. An **isometric embedding** of a metric space  $\langle X, d \rangle$  into a metric space  $\langle Y, D' \rangle$  is a map  $f: X \rightarrow Y$  whose codomain restriction  $X \rightarrow Y' = f(X)$  is an isometry from  $\langle X, d \rangle$  to  $\langle Y, D' \rangle$ , where  $D'$  is the metric on  $Y'$  induced by  $D$ .

Construct an isometric embedding of  $\langle \mathbb{R}^n, d \rangle$ , where  $d$  is the Euclidean metric, into the Hilbert sequence space  $\langle \ell^2, d_2 \rangle$  of [Example 1.10](#).

60. Let  $X = \{x, y, z, w\}$  be a set consisting of four distinct points.

(a) Show that there is a unique metric  $d$  on  $X$  for which

$$\begin{aligned} d(x, y) &= d(y, z) = d(z, x) = 2, \\ d(x, w) &= d(y, w) = d(z, w) = 1. \end{aligned}$$

(b) Show that there is no isometric embedding of  $\langle X, d \rangle$  into the Hilbert sequence space  $\langle \ell^2, d_2 \rangle$  of [Example 1.10](#).

**prob:isometric-embedding-vs-isometric** 61. Must two metric spaces be isometric to one another if there are isometric embeddings of each into the other?

62. If  $d$  is the Euclidean metric and  $d_\infty$  is the max metric on  $\mathbb{R}^2$ , is  $\langle \mathbb{R}^2, d \rangle$  isometric to  $\langle \mathbb{R}^2, d_\infty \rangle$ ?

**prob:euclidean-ip** 63. The **Euclidean inner product** on  $\mathbb{R}^n$  is the map

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R}^n &\rightarrow \mathbb{R} \\ \langle x, y \rangle &\mapsto \langle x | y \rangle \end{aligned}$$

defined by

$$\langle x | y \rangle = \sum_{i=1}^n x_i y_i.$$

(This generalizes to dimensions  $n$  the “dot product” of  $\mathbb{R}^2$  and  $\mathbb{R}^3$ .)

The Euclidean inner product is related to the Euclidean norm (see the discussion preceding [Lemma 1.3](#)) by the formula

$$\langle x|x \rangle = \|x\|^2.$$

Consequently, the inner product has the property or *positive definiteness*:

$$\begin{aligned} \langle x|x \rangle &\geq 0, \\ \langle x|x \rangle = 0 &\iff x = 0 \quad (x \in \mathbb{R}^n). \end{aligned}$$

Moreover, the Cauchy–Schwarz Inequality ([Lemma 1.4](#)) says

$$|\langle x|y \rangle| \leq \|x\| \cdot \|y\|.$$

Verify the following additional properties of the Euclidean inner product:

- (a) *Symmetry*:  $\langle x|y \rangle = \langle y|x \rangle$  for all  $x, y \in \mathbb{R}^n$ .
- (b) *Additivity*:  $\langle x + y|z \rangle = \langle x|z \rangle + \langle y|z \rangle$  for all  $x, y, z \in \mathbb{R}^n$ .
- (c) *Homogeneity*:  $\langle \lambda x|y \rangle = \lambda \langle x|y \rangle$  for all  $\lambda \in \mathbb{R}$  and all  $x, y \in \mathbb{R}^n$ .
- (d) *Polarization identity*:  $\langle x|y \rangle = \frac{1}{2} (\|x\|^2 - \|x - y\|^2 + \|y\|^2)$  for all  $x, y \in \mathbb{R}^n$ .

prob:ip-and-isometry

- 64.** (*Continuation of Exercise 63.*) A map  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  is said to “preserve inner products” if

$$\langle f(x)|f(y) \rangle = \langle x|y \rangle \quad (x, y \in \mathbb{R}^n).$$

prob-part:rotation-R2-preserves-ip

- (a) Show that each rotation of  $\mathbb{R}^2$  ([Exercise 4](#)) preserves inner products.

part:central-isometry-preserves-ip

- (b) Let  $d$  be the Euclidean metric on  $\mathbb{R}^n$ . Prove that every isometry  $f$  from  $\langle \mathbb{R}^n, d \rangle$  to itself for which  $f(0) = 0$  preserves inner products.

part:isometry-R2-maps-lines-to-lines

- (c) Let  $f$  be an isometry from the Euclidean plane to itself. Show that the image under  $f$  of each line is a line. [*Hint*: If  $f(0) = z$ , then the translate  $g$  of  $f$  given by  $g(x) = f(x) - z$  is also an isometry, but now with  $g(0) = 0$ . The equation of a line may be put in the form  $\langle \langle a, b \rangle | \langle x_1, x_2 \rangle \rangle = c$ .]

- 65.** Let  $B$  and  $b'$  be two balls in  $\mathbb{R}^n$  (for the Euclidean metric on  $\mathbb{R}^n$ ), and let  $d$  and  $d'$  be their Euclidean metrics.

- (a) Prove that  $\langle B, d \rangle$  is topologically equivalent to  $\langle B', d' \rangle$ .

- (b) When will  $\langle B, d \rangle$  actually be metrically equivalent to  $\langle B', d' \rangle$ ?

- 66.** Let  $d$  be the Euclidean metric on  $\mathbb{R}$ . Construct a metric  $d'$  on  $\mathbb{R}$  such that  $\langle \mathbb{R}, d \rangle$  is topologically equivalent to  $\langle \mathbb{R}, d' \rangle$  but  $d$  is *not* equivalent to  $d'$ . (*Moral*: The word ‘equivalent’ has several meanings.)

- 67.** Prove parts (2) and (3) of [Proposition 1.50](#).

## 1.4 Continuity and Convergence

sec:contconv

At long last we are ready to define continuity of a function from one metric space to another and to show that continuity depends not on the particular metrics involved but only on the open sets they determine. We also introduce the notion of sequential convergence and relate it to continuity.

### Continuous maps between metric spaces

subsec:cont-metric

The definition that follows is merely a generalization of the one suggested at the beginning of [Section 1.1](#) for continuity of a function from  $\mathbb{R}^n$  to  $\mathbb{R}$ .

def:cont-metric

**1.51 Definition (continuity of map of metric spaces).** Let  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$  be metric spaces and let  $f: X \rightarrow Y$  be a map. Given  $x \in X$ , the map  $f$  is said to be  $\langle d, d' \rangle$ -**continuous at  $x$**  when

for each  $\varepsilon > 0$  there exists some  $\delta > 0$  such that for all  $u \in X$

$$d(x, u) < \delta \implies d'(f(x), f(u)) < \varepsilon,$$

or purely in symbols,

$$(\forall \varepsilon > 0) (\exists \delta > 0) \left( (\forall u \in X) \left( d(x, u) < \delta \implies d'(f(x), f(u)) < \varepsilon \right) \right).$$

The map  $f$  is said to be  $\langle d, d' \rangle$ -**continuous** when it is  $\langle d, d' \rangle$ -continuous at every  $x \in X$ . The negation of “continuous” is **discontinuous**.

**Intuitive idea—continuity at a point.** Loosely speaking,  $f$  is  $\langle d, d' \rangle$ -continuous at  $x$  if  $f(u)$  can be made as  $d'$ -close to  $f(x)$  as we wish by taking  $u$  sufficiently  $d$ -close to  $x$ .

exs:cont-metric

**1.52 Examples.** (1) Let  $d$  be the Euclidean metric on  $\mathbb{R}$ . Then a function  $f: \mathbb{R} \rightarrow \mathbb{R}$  is  $\langle d, d \rangle$ -continuous at a point  $x \in \mathbb{R}$  precisely when for each  $\varepsilon > 0$  there exists some  $\delta > 0$  such that  $|f(x) - f(u)| < \varepsilon$  whenever  $u \in \mathbb{R}$  with  $|x - u| < \delta$ . Hence [Definition 1.51](#) generalizes the familiar calculus definition of continuity of a real-valued function of a real value.

We stipulate that *all the basic elementary functions encountered in calculus—including polynomial functions, rational functions, exponential and logarithmic functions, and trigonometric and inverse trigonometric functions—are continuous at all points of their domains.*

(2) For this example it is best to write out explicitly what it means for  $f: X \rightarrow Y$  to be  $\langle d, d' \rangle$ -discontinuous at  $x \in X$ :

There is some  $\varepsilon > 0$  such that for each  $\delta > 0$ ,  
there is at least one  $u \in X$  for which  
 $d(x, u) < \delta$  but  $d'(f(x), f(u)) \geq \varepsilon$ .

Now let  $d$  be the Euclidean metric on  $\mathbb{R}$ , and define  $f: \mathbb{R} \rightarrow \mathbb{R}$  by

$$f(x) = \begin{cases} u & \text{if } u \neq 0, \\ 1 & \text{if } u = 0. \end{cases}$$

Then  $f$  is  $\langle d, d \rangle$ -discontinuous at 0. In fact, let  $\varepsilon = 1$ . Take an arbitrary  $\delta > 0$ . Then the point  $u = -\delta/2$  is one for which  $d(0, u) < \delta$  but nonetheless  $d(f(0), f(u)) = 1 + \delta/2 \geq \varepsilon$ .

(3) Let  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$  be any metric spaces and let

$$\begin{aligned} f: X &\rightarrow Y \\ x &\mapsto c \end{aligned}$$

a constant map. Then  $f$  is  $\langle d, d' \rangle$ -continuous because

$$d'(f(x), f(u)) = d'(c, c) = 0$$

for all  $x, u \in X$ .

- (4) Let  $d'$  be the metric induced on a subset  $Y$  of a metric space  $\langle X, d \rangle$ . Then the inclusion map

$$\begin{aligned} j: Y &\rightarrow X \\ x &\mapsto x \end{aligned}$$

is  $\langle d', d \rangle$ -continuous, because

$$d'(x, u) < \varepsilon \implies d(j(x), j(u)) = d(x, u) = d'(x, u) < \varepsilon$$

for all  $x, u \in Y$ .

Specialize the above by taking  $Y = X$ , so that  $d' = d$  and  $j$  is the identity map

$$\begin{aligned} \iota_X: X &\rightarrow X \\ x &\mapsto x \end{aligned}$$

of  $X$ . Then  $\iota_X$  is  $\langle d, d \rangle$ -continuous.

ex:composite-cont-is-cont-metric

- (5) Let  $\langle X, d \rangle$ ,  $\langle Y, d' \rangle$ , and  $\langle Z, d'' \rangle$  be metric spaces, let

$$f: X \rightarrow Y, \quad g: Y \rightarrow Z$$

be given maps, and form the composite map

$$\begin{aligned} g \circ f: X &\rightarrow Z \\ x &\mapsto g(f(x)) \end{aligned}.$$

Suppose  $f$  is  $\langle d, d' \rangle$ -continuous at a point  $x \in X$  and  $g$  is  $\langle d', d'' \rangle$ -continuous at the point  $f(x) \in Y$ . Then the composite  $g \circ f$  is  $\langle d, d'' \rangle$ -continuous at  $x$ . In fact, let  $\varepsilon > 0$ . There exists  $\eta > 0$  such that

$$d'(f(x), y) < \eta \implies d''(g(f(x)), g(y)) < \varepsilon$$

for all  $y \in Y$ . Next, there exists  $\delta > 0$  such that

$$d(x, u) < \delta \implies (f(x), f(u)) < \eta.$$

for all  $u \in X$ . Hence

$$d(x, u) < \delta \implies d''((g \circ f)(x), (g \circ f)(u)) < \varepsilon$$

for all  $u \in X$ .

By what we just proved,  $g \circ f$  will be  $\langle d, d'' \rangle$ -continuous whenever  $f$  is  $\langle d, d' \rangle$ -continuous and  $g$  is  $\langle d', d'' \rangle$ -continuous.

- (6) Let  $\delta$  be the discrete metric on a set  $Z$ , and let  $\langle X, d \rangle$  be an arbitrary metric space. Then *every* map  $f: Z \rightarrow X$  is  $\langle \delta, d \rangle$ -continuous. In fact, since  $\delta(x, u) < 1$  precisely when  $x = u$ , we have

$$\delta(x, u) < 1 \implies d(f(x), f(u)) = 0 < \varepsilon$$

for all  $\varepsilon > 0$  and all  $x, u \in Z$ .

ex:isometry-metrically-cont (7) Prove: An isometry  $f: \text{opair } X, d \rightarrow \langle Y, d' \rangle$  of metric spaces is  $\langle d, d' \rangle$ -continuous, and its inverse  $f^{-1}: \langle Y, d' \rangle \rightarrow \text{opair } X, d$  is  $\langle d', d \rangle$ -continuous.

ex:metric-cont-fn (8) Let  $\langle X, d \rangle$  be any metric space. The metric  $d$  is a map

$$d: X \times X \rightarrow \mathbb{R}$$

about whose continuity it makes sense to inquire. Here we provide  $\mathbb{R}$  with its Euclidean metric  $d'$  and  $X \times X$  with its max metric  $d_\infty$  given by

$$d_\infty((x_1, x_2), (u_1, u_2)) = \max\{d(x_1, u_1), d(x_2, u_2)\}$$

(see [Example 1.14](#)). We claim that the map  $d$  is  $\langle d_\infty, d' \rangle$ -continuous.

Let  $(x_1, x_2) \in X \times X$  and let  $\varepsilon > 0$  be given. We seek some  $\delta > 0$  for which

$$d_\infty((x_1, x_2), (u_1, u_2)) < \delta \implies |d(x_1, x_2) - d(u_1, u_2)| < \varepsilon.$$

Now

$$\begin{aligned} |d(x_1, x_2) - d(u_1, u_2)| &\leq |d(x_1, x_2) - d(u_1, x_2)| + |d(u_1, x_2) - d(u_1, u_2)| \\ &\leq d(x_1, u_1) + d(x_2, u_2) \\ &\leq 2 d_\infty((x_1, x_2), (u_1, u_2)). \end{aligned}$$

Hence  $\delta = \varepsilon/2$  will do.  $\diamond$

The implication

$$d(x, u) < \delta \implies d'(f(x), f(u)) < \varepsilon$$

appearing in [Definition 1.51](#) may be restated in the form

$$u \in B_\delta(x; d) \implies f(u) \in B_\varepsilon(f(x); d')$$

or even more concisely in the form

$$f(B_\delta(x; d)) \subset B_\varepsilon(f(x); d').$$

Hence  $f: X \rightarrow Y$  is  $\langle d, d' \rangle$ -continuous at  $x \in X$  if and only if each  $d'$ -ball at  $f(x)$  contains the image of some  $d$ -ball at  $x$ .

That observation allows us to characterize continuity at a point in terms of arbitrary neighborhoods.

thm:metric-cont-at-pt-via-nbds **1.53 Theorem (characterization of continuity in metric spaces in terms of neighborhoods).** Let  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$  be metric spaces, let  $f: X \rightarrow Y$  be a map, and let  $x \in X$ . Then the following statements are equivalent:

- property:metric-cont-at-pt-via-nbds-cont (i) The map  $f$  is  $\langle d, d' \rangle$ -continuous at  $x$ .
- metric-cont-at-pt-via-nbds-exists-nbd (ii) For each  $d'$ -neighborhood  $N$  of  $f(x)$  in  $Y$  there exists some  $d$ -neighborhood  $M$  of  $x$  in  $X$  such that  $f(M) \subset N$ .
- metric-cont-at-pt-via-nbds-inverse-image (iii) The inverse image  $f^{-1}(N)$  of each  $d'$ -neighborhood  $N$  of  $f(x)$  in  $Y$  is a  $d$ -neighborhood of  $x$  in  $X$ .

**Proof.** We give a “circular proof” which consists in proving each of the implications (i)  $\implies$  (ii), (ii)  $\implies$  (iii), (iii)  $\implies$  (i). It will then follow that all of the equivalences (1)  $\iff$  (2), (2)  $\iff$  (3), (1)  $\iff$  (3) are valid.



(i)  $\implies$  (ii). Assume (i). Let  $N$  be a  $d'$ -neighborhood of  $f(x)$ . Choose  $\varepsilon > 0$  with

$$B_\varepsilon(f(x); d') \subset N$$

(see Figure 1.17). By (i), there exists  $\delta > 0$  with

$$f(B_\delta(x; d)) \subset B_\varepsilon(f(x); d').$$

Then  $M = B_\delta(x; d)$  is a  $d$ -neighborhood of  $x$  for which  $f(M) \subset N$ .

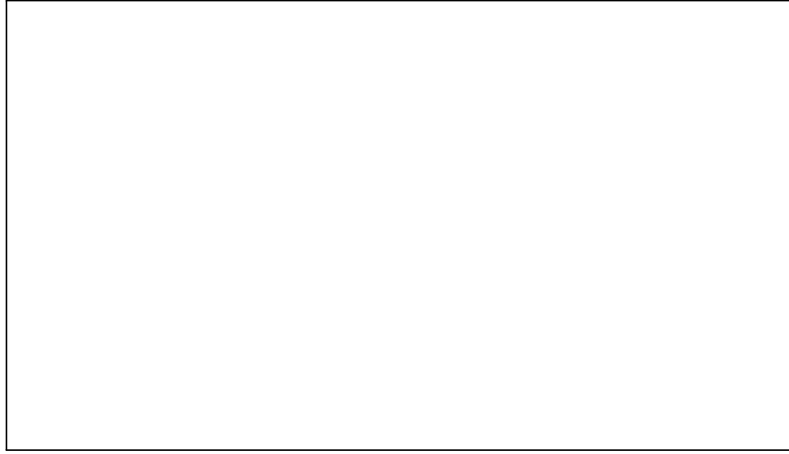


Figure 1.17: Sets in proof of (1)  $\implies$  (2) in Theorem 1.53.

fig:pf-1then2-metric-cont-at-pt

(ii)  $\implies$  (iii). Assume (ii). Let  $N$  be a  $d'$ -neighborhood of  $f(x)$ . By (ii), there exists a  $d$ -neighborhood  $M$  of  $x$  with

$$f(M) \subset N.$$

By the definition of ' $d$ -neighborhood at  $x$ ', there is a  $d$ -open set  $U$  with

$$x \in U \subset M$$

(see Figure 1.18). Since  $M \subset f^{-1}(N)$ ,

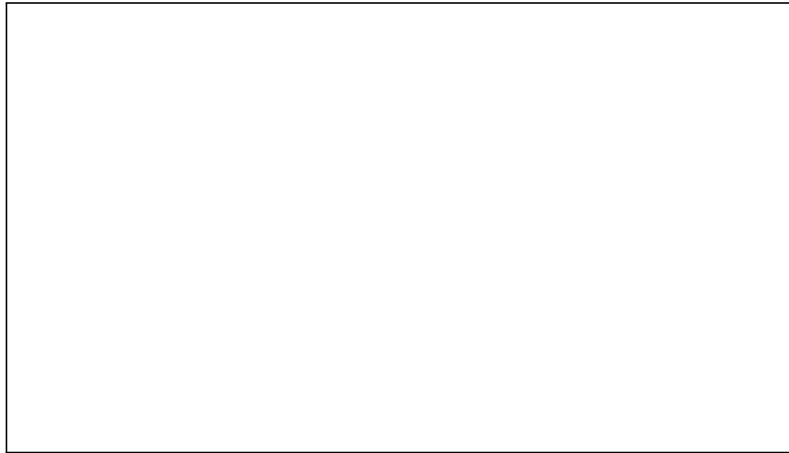


Figure 1.18: Sets in proof of (2)  $\implies$  (3) in Theorem 1.53.

fig:pf-2then3-metric-cont-at-pt

$$x \in U \subset f^{-1}(N).$$

Hence  $f^{-1}(N)$  is a  $d$ -neighborhood of  $x$ .

(iii)  $\implies$  (i). Assume (iii). Let  $\varepsilon > 0$ . Since  $B_\varepsilon(f(x); d')$  is a  $d'$ -neighborhood of  $f(x)$ , by (iii) the set

$$M = f^{-1}(B_\varepsilon(f(x); d'))$$

is a  $d$ -neighborhood of  $x$ . Then

$$B_\delta(x; d) \subset M$$

for some  $\delta > 0$ . Hence

$$f(B_\delta(x, d)) \subset B_\varepsilon(f(x); d') \quad \square$$

Statement (iii) above definitely does *not* say that the image of each  $d$ -neighborhood of  $x$  is a  $d'$ -neighborhood of  $f(x)$ . And it should not!

ex:metric-map-cont-not-open **1.54 Example.** Let  $\langle X, d \rangle$  be any metric space, let  $d'$  be the Euclidean metric on  $\mathbb{R}$ , and let  $f: X \rightarrow \mathbb{R}$  be the constant function with value 0. Then  $f$  is  $\langle d, d' \rangle$ -continuous at every  $x \in \mathbb{R}$ . Now  $f(M) = \{0\}$  for each nonempty set  $M \subset X$ , yet the image  $\{0\}$  of  $M$  under  $f$  is not a  $d'$ -neighborhood of 0. (See also [Exercise 74](#).)  $\diamond$

Because equivalent metrics determine the same open sets and hence the same neighborhoods of a point, [Theorem 1.53](#) has the following immediate corollary.

cor:cont-same-for-equiv-metrics **1.55 Corollary (continuity of equivalent metrics).** In the notation of [Theorem 1.53](#), let  $D$  be a metric on  $X$  that is equivalent to  $d$  and let  $D'$  be a metric on  $Y$  that is equivalent to  $d'$ . Then the map  $f$  is  $\langle d, d' \rangle$ -continuous at  $x$  if and only if it is  $\langle D, D' \rangle$ -continuous at  $x$ .

From [Theorem 1.53](#) we may now deduce an especially simple criterion for continuity—one that involves only open sets.

thm:continuity-in-metric-via-open-sets **1.56 Theorem (continuity in terms of open sets).** Let  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$  be metric spaces and let  $f: X \rightarrow Y$  be a map. Then a necessary and sufficient condition for  $f$  to be  $\langle d, d' \rangle$ -continuous is that the inverse image  $f^{-1}(V)$  of each  $d'$ -open subset  $V$  of  $Y$  be a  $d$ -open subset of  $X$ .

**Proof.** Necessity. Assume first that  $f$  is  $\langle d, d' \rangle$ -continuous. Let  $V \subset Y$  be an arbitrary  $d'$ -open set. To show that  $f^{-1}(V)$  is  $d$ -open, it suffices by [Proposition 1.32](#) to show that  $f^{-1}(V)$  is a  $d$ -neighborhood of each of its points. Let  $x \in f^{-1}(V)$ . then  $f(x) \in V$ , so that  $V$  is a  $d'$ -neighborhood of  $f(x)$ , and hence by condition (iii) of [Theorem 1.53](#) the set  $f^{-1}(V)$  is a  $d$ -neighborhood of  $x$ .

Sufficiency. Conversely, assume the condition holds. Let  $x \in X$ . To show that  $f$  is  $\langle d, d' \rangle$ -continuous at  $x$  we verify condition (ii) of [Theorem 1.53](#). Accordingly, Let  $N$  be a  $d'$ -neighborhood of  $f(x)$  in  $Y$ . Choose a  $d'$ -open subset  $V$  of  $Y$  with

$$f(x) \in V \subset N.$$

By assumption, the set  $M = f^{-1}V$  is  $d$ -open. Hence  $M$  is a  $d$ -neighborhood of  $x$  with

$$f(M) = V \subset N. \quad \square$$

Suppose in the preceding [Theorem 1.56](#) that the map  $f: X \rightarrow Y$  is a bijection. Then we may apply the theorem to the map  $f^{-1}: Y \rightarrow X$  to conclude that  $f^{-1}$  is  $\langle d', d \rangle$ -continuous precisely when for each  $d$ -open subset  $U$  of  $X$ , the subset

$$f(U) = (f^{-1})^{-1}(U)$$

is  $d'$ -open. Hence **a metric space  $\langle X, d \rangle$  is topologically equivalent to a metric space  $\langle Y, d' \rangle$  if and only if there is some bijection  $f: X \rightarrow Y$  such that  $f$  is  $\langle d, d' \rangle$ -continuous and  $f^{-1}$  is  $\langle d', d \rangle$ -continuous.**

The study of continuity will be resumed in [Chapter 3](#) (Continuity and Convergence), in the more general setting of “topological spaces.”

### Convergent sequences in metric spaces

subsec:conv-seq-metric

We turn now to the topic of sequential convergence in metric spaces. Recall from calculus that a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  of real number is said to converge to the real number  $x$  if:

for each  $\varepsilon > 0$  there is some index  $m \in \mathbb{N}$  such that  $|x - x_n| < \varepsilon$  for all  $n \geq m$ .

Now  $|x - x_n|$  is just the distance  $d(x, x_n)$  for the Euclidean metric  $d$  on  $\mathbb{R}$ , so it is straightforward to generalize the notion of sequential convergence to arbitrary metric spaces.

def:seq-conv-metric

**1.57 Definition (sequential convergence in metric space).** Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence of points in a metric space  $\langle X, d \rangle$ . If  $x \in X$ , then we say that  $\langle x_n \rangle_{n \in \mathbb{N}}$  **converges to  $x$  in  $\langle X, d \rangle$** , or that it  **$d$ -converges to  $x$** , and write  $\langle x_n \rangle_{n \in \mathbb{N}} \xrightarrow{d} x$  to mean:

for each  $\varepsilon > 0$  there exists some index  $m \in \mathbb{N}$  such that  $d(x, x_n) < \varepsilon$  for all  $n \in \mathbb{N}$  with  $n \geq m$ .

When  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to some point in  $\langle X, d \rangle$ , then we say that  $\langle x_n \rangle_{n \in \mathbb{N}}$  **converges in  $\langle X, d \rangle$** ; otherwise, we say that it **diverges in  $\langle X, d \rangle$** .

**Intuitive idea—convergence of a sequence.** To say that a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x$  in  $\langle X, d \rangle$  means that we can make  $x_n$  as  $d$ -close to  $x$  as we wish by taking  $n$  sufficiently large.

Because the preceding definition generalizes the definition from calculus, you already know examples such as the following.

exs:seq-convergence-R **1.58 Examples.** Let  $d$  be the Euclidean metric on  $\mathbb{R}$ .

ex:1-over-n-converges-to-0

- (1) The sequence  $\langle 1/(n+1) \rangle_{n \in \mathbb{N}}$  converges to 0 in  $\langle \mathbb{R}, d \rangle$ . This is an immediate consequence of the Archimedean ordering property of  $\mathbb{R}$  (see [Theorem 0.78](#)).
- (2) The “oscillating” sequence  $\langle (-1)^n \rangle_{n \in \mathbb{N}}$ , consisting of entries that are alternately 1 and  $-1$ , diverges in  $\langle \mathbb{R}, d \rangle$ .
- (3) For a given real number  $r$ , the sequence  $\langle r^n \rangle_{n \in \mathbb{N}}$  converges in  $\langle \mathbb{R}, d \rangle$  if and only if  $-1 < r \leq 1$ . Moreover, if  $-1 < r < 1$ , the sequence converges to 0; if  $r = 1$ , it converges to 1.

continuous map!metric spaces@bet  
convergent sequence

**increasing sequence** (4) For a real number  $r$ , let  $\langle s_n \rangle_{n \in \mathbb{N}}$  be the sequence of “geometric sums” given by **sequence!increasing**

$$s_n = \sum_{j=0}^n r^j \quad (n \in \mathbb{N}).$$

Then  $\langle s_n \rangle_{n \in \mathbb{N}}$  converges in  $\langle \mathbb{R}, d \rangle$  if and only if  $|r| < 1$ , because

$$s_n = \frac{1}{1-r} - \frac{r^{n+1}}{1-r} \quad (n \in \mathbb{N}).$$

Moreover, the sequence converges to  $1/(1-r)$  when  $|r| < 1$ .  $\diamond$

Before providing additional examples, we name and establish a necessary condition for convergence.

**def:seq-bded** **1.59 Definition (bounded sequence).** A sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in a metric space  $\langle X, d \rangle$  is said to be  **$d$ -bounded** when its range  $\{x_n : n \in \mathbb{N}\}$  is a  $d$ -bounded subset of  $X$ .

**prop:conv-seq-bded** **1.60 Proposition.** If a sequence converges in a metric space  $\langle X, d \rangle$ , then it is  $d$ -bounded.

**Proof.** Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence that converges in  $\langle X, d \rangle$ . Corresponding to  $\varepsilon = 1$  there is an  $m \in \mathbb{N}$  such that

$$\{ineq:dxnx-less-than-1\} \quad (*) \quad d(x_n, x) < 1 \quad (n \geq m).$$

All we need do is find a constant  $c$  for which

$$d(x_n, x_m) \leq c \quad (n \in \mathbb{N}),$$

because then

$$d(x_j, x_k) \leq d(x_j, x_m) + d(x_m, x_k) \leq 2c \quad (j, k \in \mathbb{N}).$$

Now from  $(*)$  we have

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < 2 \quad (n \geq m).$$

Set

$$b = \max_{1 \leq n \leq m} d(x_n, x_m).$$

Then  $c = \max\{2, b\}$  will do.  $\square$

**1.61 Examples.** (1) **An increasing sequence of real numbers converges if and only if it is bounded above.**

In fact, let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence in  $\mathbb{R}$  that is **increasing** in the sense that

$$x_n \leq x_{n+1} \quad (n \in \mathbb{N})$$

(compare [Exercise 0.106](#).) We shall prove that  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges in  $\langle \mathbb{R}, d \rangle$ , where  $d$  is the Euclidean metric, if and only if the range  $\{x_n : n \in \mathbb{N}\}$  of the sequence

is bounded above in  $\mathbb{R}$  (Definition 0.66). Moreover, in this situation the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges in  $\langle \mathbb{R}, d \rangle$  to

$$x = \sup\{x_n : n \in \mathbb{N}\}.$$

In fact, in view of Proposition 1.60, the range  $\{x_n : n \in \mathbb{N}\}$  will certainly be bounded above in  $\mathbb{R}$  if the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges in  $\langle \mathbb{R}, d \rangle$ .

Conversely, assume that  $\{x_n : n \in \mathbb{N}\}$  is bounded above in  $\mathbb{R}$ . We deduce that  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x = \sup\{x_n : n \in \mathbb{N}\}$  in  $\langle \mathbb{R}, d \rangle$ . Let  $\varepsilon > 0$  be arbitrary. Since  $x - \varepsilon < x$ , then  $x - \varepsilon$  is not an upper bound of  $\{x_n : n \in \mathbb{N}\}$ , that is,

$$x - \varepsilon < x_m$$

for some  $m \in \mathbb{N}$ . Since the sequence is increasing and its range has  $x$  as an upper bound,

$$x - \varepsilon < x_m \leq x_n \leq x < x + \varepsilon \quad (n \geq m).$$

Hence  $d(x, x_n) < \varepsilon$  for all  $n$  with  $n \geq m$ . Since the sequence is increasing and

g-conv-iff-partial-sums-bdd-above

- (2) The result just established in (1) has an application to convergence of infinite series that was already used in our discussion of the Hilbert sequence space  $\ell^2$  (Example 1.10, page 138).

Recall that a series  $\sum_{j=1}^{\infty} a_j$  of real numbers is said to **converge** when its sequence  $\langle s_n \rangle_{n=1,2,3,\dots}$  of partial sums converges, where the  $n$ th partial sum  $s_n = \sum_{j=1}^n a_j$ .

Suppose  $\sum_{j=1}^{\infty} a_j$  is a series of *nonnegative* terms  $a_j$ . Then the sequence  $\langle s_n \rangle_{n=1,2,3,\dots}$  of its partial sums

$$s_n = \sum_{j=1}^n a_j$$

is increasing because

$$s_{n+1} = s_n + a_{n+1} \geq s_n \quad (n = 1, 2, 3, \dots).$$

Hence the infinite series  $\sum_{j=1}^{\infty} a_j$  converges if and only if the set  $\{s_n : n = 1, 2, 3, \dots\}$  is bounded above in  $\mathbb{R}$ , and in this event  $s_n \leq \sum_{j=1}^{\infty} a_j$  for each  $n$ .

ex:comparison-test

- (3) A corollary of (2) is the comparison test from calculus: Suppose  $\sum_{j=1}^{\infty} a_j$  and  $\sum_{j=1}^{\infty} b_j$  are two series with the first “dominated by” the second in the sense that

$$a_j \leq b_j \quad (j = 1, 2, 3, \dots),$$

and suppose the “larger” series  $\sum_{j=1}^{\infty} b_j$  converges. Then the “smaller” series  $\sum_{j=1}^{\infty} a_j$  also converges. The reason is that for each  $n$ ,

$$s_n = \sum_{j=1}^n a_j \leq \sum_{j=1}^n b_j \leq \sum_{j=1}^{\infty} b_j,$$

so that the increasing sequence  $\langle s_n \rangle_{n=1,2,3,\dots}$  is bounded above by the real number  $\sum_{j=1}^{\infty} b_j$ .

ex:conv-power-series-base

- (4) In calculus, you learn that from the closed-form formula

$$\sum_{j=0}^n a r^j = \frac{a}{1-r} - \frac{a r^{n+1}}{1-r}$$

for the sum of a geometric progression it follows that the series  $\sum_{j=0}^{\infty} a r^j$  converges to  $a/(1-r)$  when  $|r| < 1$ . Then the series  $\sum_{j=1}^{\infty} a/b^j$  converges to  $a/(b-1)$  for each  $b > 1$ .

series  
convergent series  
comparison test

base expansion  
binary base  
ternary base  
base!binary  
base!ternary  
series

In particular, let  $b$  be an integer with  $b > 1$ . Then the series  $\sum_{j=1}^{\infty} (b-1)/b^j$  converges—in fact, converges to 1.

From the comparison test—see Example (3)—it follows that for each sequence  $\langle d_j \rangle_{j=1,2,3,\dots}$  of “digits”  $d_j \in \{0, 1, 2, \dots, b-1\}$  for the base  $b$ , the series

$$\sum_{j=1}^{\infty} \frac{d_j}{b^j} = \frac{d_1}{b} + \frac{d_2}{b^2} + \dots + \frac{d_j}{b^j} + \dots$$

converges—in fact, converges to a real number in  $[0, 1]$ .

For base  $b = 10$ , this is the reason that any (possibly non-terminating) decimal  $0.d_1 d_2 d_3 \dots d_j \dots$  represents a real number. Similarly for other bases  $b$ , including the important *binary* base  $b = 2$  and the *ternary* base  $b = 3$ .

Conversely, as was shown in the subsection “Base expansion” (page 79), given an integer  $b > 1$ , each real number  $x$  in  $[0, 1]$  is representable by such an infinite series; that is, there exists a sequence  $\langle d_j \rangle_{j=1,2,3,\dots}$  in  $\{0, 1, 2, \dots, b-1\}$  such that  $x = \sum_{j=1}^{\infty} d_j/b^j$ , and so  $x = (0.d_1 d_2 d_3 \dots d_j \dots)_b$ .

Binary expansion was already exploited in Example 0.121 to show that  $\text{card } \mathbb{R} = 2^{\aleph_0}$ ; ternary expansion will be applied when we analyze the “Cantor set” (Example 4.16).

ex:seq-bded-not-conv

- (5) If  $d$  is the Euclidean metric on  $\mathbb{R}$ , then the sequence  $\langle (-1)^n \rangle_{n \in \mathbb{N}}$  is  $d$  bounded but does not converge in  $\langle \mathbb{R}, d \rangle$ . Thus the converse of Proposition 1.60 fails.

conv-in-subsp-then-in-whole-metric

- (6) Let  $\langle y_n \rangle_{n \in \mathbb{N}}$  be a sequence in a subset  $Y$  of a metric space  $\langle X, d \rangle$ , and let  $d'$  be the metric on  $Y$  induced by  $d$ . For  $y \in Y$ , then clearly  $\langle y_n \rangle_{n \in \mathbb{N}}$  converges to  $y$  in  $\langle Y, d' \rangle$  if and only if it converges to  $y$  in  $\langle X, d \rangle$ . However, the converse of this last statement need not hold—see (7).

conv-in-whole-metric-not-in-subsp

- (7) Let  $d$  and  $d'$  be the Euclidean metrics on the reals  $\mathbb{R}$  and rationals  $\mathbb{Q}$ , respectively. We shall construct a sequence  $\langle y_n \rangle_{n \in \mathbb{N}}$  in  $\mathbb{Q}$  that is strictly increasing, is bounded above in  $\mathbb{Q}$ , does *not* converge in  $\langle \mathbb{Q}, d' \rangle$ , and yet converges in  $\langle \mathbb{R}, d \rangle$ . This sequence will consist of successively better approximations to  $\sqrt{2}$ .

Choose  $y_1 \in \mathbb{Q}$  with

$$\sqrt{2} - 1 < y_1 < \sqrt{2},$$

and set  $y_0 = y_1$ . Next, choose  $y_2 \in \mathbb{Q}$  with

$$\sqrt{2} - \frac{1}{2} < y_2 < \sqrt{2}, \quad y_1 < y_2.$$

In general, once  $y_1, y_2, \dots, y_{n-1}$  have been selected, choose  $y_n \in \mathbb{Q}$  with

$$\sqrt{2} - \frac{1}{n} < y_n < \sqrt{2}, \quad y_{n-1} < y_n.$$

At each stage the choice is possible because the rational numbers are order-dense in  $\mathbb{R}$ , that is, each open interval in  $\mathbb{R}$  contains a rational number (see Corollary 0.81).

Clearly  $\langle y_n \rangle_{n \in \mathbb{N}}$  is strictly increasing, and any rational number greater than  $\sqrt{2}$  (for example, 2) is an upper bound of  $\langle y_n \rangle_{n \in \mathbb{N}}$  in  $\mathbb{Q}$ . The sequence converges to  $\sqrt{2}$  in  $\langle \mathbb{R}, d \rangle$  because

$$d(\sqrt{2}, y_n) < \frac{1}{n} \quad (n \geq 1).$$

eventually constant sequence  
sequence!eventually constant  
uniform convergence  
converge!uniformly

Just suppose  $\langle y_n \rangle_{n \in \mathbb{N}}$  converges in  $\langle \mathbb{Q}, d' \rangle$  to some point  $y \in \mathbb{Q}$ . Since  $\sqrt{2}$  is irrational,  $y \neq \sqrt{2}$ . Let

$$\varepsilon = \varepsilon = d(y, \sqrt{2}) > 0.$$

There is an  $m \in \mathbb{N}$  for which

$$d(y, y_n) = d'(y, y_n) < \frac{\varepsilon}{2} \quad (n \geq m)$$

because  $\langle y_n \rangle_{n \in \mathbb{N}}$  converges to  $y$  in  $\langle \mathbb{Q}, d' \rangle$ ; and there is a  $k \in \mathbb{N}$  for which

$$d(y_n, \sqrt{2}) < \frac{\varepsilon}{2} \quad (n \geq k)$$

because  $\langle y_n \rangle_{n \in \mathbb{N}}$  converges to  $\sqrt{2}$  in  $\langle \mathbb{R}, d \rangle$ . Then for  $n = \max\{m, k\}$ ,

$$d(y, \sqrt{2}) \leq d(y, y_n) + d(y_n, \sqrt{2}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

which contradicts the definition of  $\varepsilon$ .

eventually-cst-seq-converges-metric

- (8) Let  $\langle X, d \rangle$  be a metric space and let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence in  $X$ . Suppose the sequence is constant, that is, there exists some  $p \in X$  such that  $x_n = p$  for all  $n$ . Or more generally, suppose the sequence is **eventually constant** in the sense that there exist some  $p \in X$  and some  $m \in \mathbb{N}$  such that  $x_n = p$  for all  $n \geq m$ . Then  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $p$  in  $\langle X, d \rangle$ .

ex:conv-seq-discrete-metric

- (9) Let  $\delta$  be the discrete metric on a set  $X$ . For  $x \in X$  and  $\varepsilon \leq q$ , the only point  $y \in X$  satisfying  $\delta(x, y) < \varepsilon$  is  $x$  itself. Hence a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges in  $(X, \delta)$  if and only if it is eventually constant.
- (10) Let  $d_\infty$  be the sup metric, as constructed in [Example 1.8](#), on the set  $X = C([0, 1])$  of all continuous functions  $x: [0, 1] \rightarrow \mathbb{R}$ . Then to say that a sequence of functions  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to a function  $x$  in  $(X, d_\infty)$  means that for each  $\varepsilon > 0$  there is an index  $m \in \mathbb{N}$  such that

$$|x(t) - x_n(t)| < \varepsilon \quad (n \geq m, 0 \leq t \leq 1).$$

Roughly speaking, this says that we can simultaneously make all the vertical distances  $|x(t) - x_n(t)|$  between the graphics of  $x$  and  $x_n$  as small as we wish by taking  $n$  sufficiently large. For this reason, convergence of a sequence in this metric space is said to be **uniform**.

- (11) Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence in an arbitrary metric space  $\langle X, d \rangle$ , and let  $d'$  be the Euclidean metric on  $\mathbb{R}$ . Then  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to a point  $x$  in  $\langle X, d \rangle$  precisely when the sequence  $\langle d(x, x_n) \rangle_{n \in \mathbb{N}}$  of real numbers converges to 0 in  $(\mathbb{R}, d')$ . In fact, if we let

$$y_n = d(x, x_n)$$

for each  $n \in \mathbb{N}$ , then

$$d'(0, y_n) = d(x, x_n)$$

for every  $n \in \mathbb{N}$ .  $\diamond$

The inequality

$$d(x, x_n) < \varepsilon,$$

appearing in [Definition 1.57](#) may be restated as

$$x_n \in B_\varepsilon(x; d).$$

Then in the language of [Definition 3.98](#), the definition of sequential convergence may be expressed as:

A sequence converges to a point  $x$  in  $\langle X, d \rangle$  if and only if, for each  $\varepsilon > 0$ , the sequence is eventually in the  $d$ -ball  $B_\varepsilon(x; d)$ .

Even more succinctly:

A sequence converges to a point  $x$  in  $\langle X, d \rangle$  if and only if it is eventually in each  $d$ -ball at  $x$ .

Each  $d$ -ball at a point  $x$  in a metric space  $\langle X, d \rangle$  is a  $d$ -neighborhood of  $x$ , and each  $d$ -neighborhood of  $x$  contains some  $d$ -ball at  $x$ . Hence we have the following criterion for convergence.

thm:conv-via-nbds-metric

**1.62 Theorem (sequential convergence in terms of neighborhoods).** *A sequence converges to a point  $x$  in a metric space  $\langle X, d \rangle$  if and only if it is eventually in each  $d$ -neighborhood of  $x$ .*

Be sure to interpret correctly the use of ‘each’ above! The convergence condition means: “for each  $d$ -neighborhood  $V$  of  $x$ , the sequence is in  $V$ .” In detail, where  $\langle x_n \rangle_{n \in \mathbb{N}}$  is the sequence at issue: “for each  $d$ -neighborhood  $V$  of  $x$ , there exists an index  $m_V$  (depending on  $V$ ) such that  $x_n \in V$  for all  $n \geq m_V$ .” (It definitely does *not* mean that there is some  $m$  such that, for all  $n \geq m$ , the  $n$ th entry of the sequence is in every neighborhood of  $x$ .)

r:conv-seq-invariant-of-equiv-metrics

**1.63 Corollary (preservation of convergence for equivalent metrics).** *Let  $d$  and  $d'$  be equivalent metrics on a set  $X$ . Then a sequence converges to a point in  $\langle X, d \rangle$  if and only if it converges to the same point in  $\langle X, d' \rangle$ .*

This independence of sequential convergence from the particular metric is exploited in the following example.

eq-conv-Rn-iff-conv-coordinatewise

**1.64 Example (convergence in  $\mathbb{R}^n$  and coordinatewise convergence).** Denote by  $d_k$  the Euclidean metric on  $\mathbb{R}^k$ . Let  $x \in \mathbb{R}^k$  and let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence of points in  $\mathbb{R}^k$ . Denote the  $k$  coordinates of  $x$  by

$$x^1, x^2, \dots, x^k,$$

and for each  $n \in \mathbb{N}$  denote the  $k$  coordinates of  $x_n$  by

$$x_n^1, x_n^2, \dots, x_n^k$$

(our use of subscripts to denote the entries in a sequence necessitates this departure from our usual practice of using subscripts for coordinates).

*Claim:* The sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x$  in  $(\mathbb{R}^k, d_k)$  if and only if for each coordinate  $j = 1, 2, \dots, k$  the sequence  $\langle x_n^j \rangle_{n \in \mathbb{N}}$  of real numbers converges to the number  $x^j$  in  $(\mathbb{R}, d)$ , where  $d$  is the Euclidean metric. See [Figure 1.19](#) for the case of dimension  $k = 2$ , and examine [Figure 1.20](#) to keep track of what is going on in the general case.



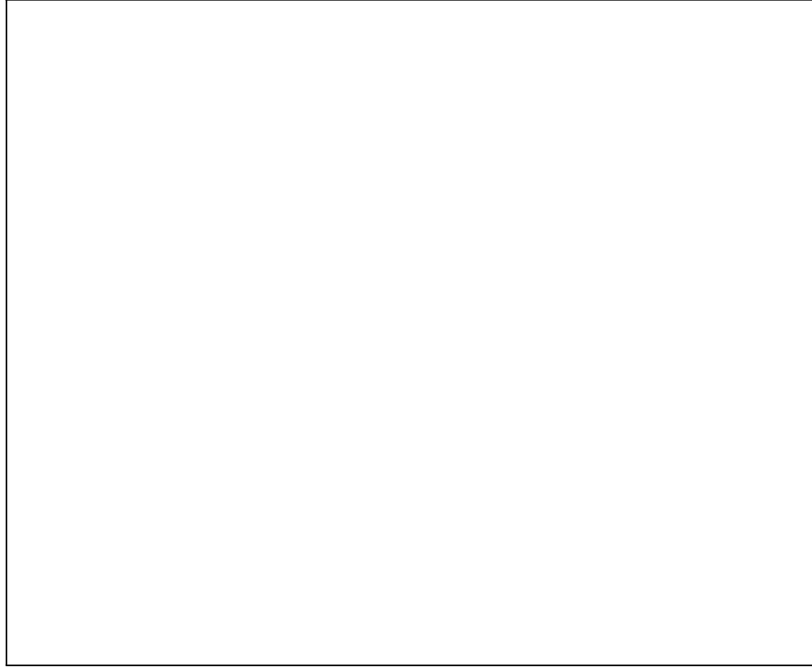
Figure 1.19: Coordinate-wise convergence of a sequence in  $\mathbb{R}^2$ .

fig:coordinatewise-conv-seq-R2

$$\begin{array}{rcl}
 \mathbb{R}^k & = & \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} \times \cdots \times \mathbb{R} \\
 x_1 & = & (x_1^1, x_1^2, \dots, x_1^j, \dots, x_1^k) \\
 x_2 & = & (x_2^1, x_2^2, \dots, x_2^j, \dots, x_2^k) \\
 \vdots & & \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\
 x_n & = & (x_n^1, x_n^2, \dots, x_n^j, \dots, x_n^k) \\
 \vdots & & \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\
 \downarrow & & \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\
 x & = & (x^1, x^2, \dots, x^j, \dots, x^k)
 \end{array}$$

Figure 1.20: Coordinate-wise convergence of a sequence in  $\mathbb{R}^k$ .

fig:coordinatewise-conv-seq-Rk

We know that the max metric  $d_\infty$  on  $\mathbb{R}^k$  is equivalent to  $d_k$  ([Proposition 1.36](#)). In view of [Corollary 1.63](#), to prove our claim we may consider convergence in  $(\mathbb{R}, d_\infty)$  instead of in  $(\mathbb{R}, d_k)$ .

Assume first that  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x$  in  $(\mathbb{R}^k, d_\infty)$ . Fix  $j$  with  $1 \leq j \leq k$ . Let  $\varepsilon > 0$  be given. There exists  $m \in \mathbb{N}$  with

$$d_\infty(x, x_n) < \varepsilon \quad (n \geq m).$$

Since

$$d(x^j, x_n^j) \leq d_\infty(x, x_n)$$

for all  $n$ , then

$$d(x^j, x_n^j) < \varepsilon \quad (n \geq m).$$

Hence  $\langle x_n^j \rangle_{n \in \mathbb{N}}$  converges to  $x^j$  in  $\langle \mathbb{R}, d \rangle$ .

Conversely, assume that  $\langle x_n^j \rangle_{n \in \mathbb{N}}$  converges to  $x^j$  in  $\langle \mathbb{R}, d \rangle$  for each  $j = 1, 2, \dots, k$ . Let  $\varepsilon > 0$  be given. For each  $j$  with  $1 \leq j \leq k$ , there exists some  $m(j) \in \mathbb{N}$  with

$$d(x^j, x_n^j) < \varepsilon \quad (n \geq m(j)).$$

Set

$$m = \max\{m(1), m(2), \dots, m(k)\}.$$

Then

$$d(x^j, x_n^j) < \varepsilon \quad (1 \leq j \leq k, n \geq m),$$

and so

$$d_\infty(x, x_n) < \varepsilon \quad (n \geq m).$$

Hence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x$  in  $(\mathbb{R}^k, d_\infty)$ .  $\diamond$

[Example 1.64](#) has an obvious generalization concerning the max metric on the product of any finite number of metric spaces ([Example 1.14](#)).

The expression ‘ $x$  is a limit of  $\langle x_n \rangle_{n \in \mathbb{N}}$ ’ is frequently used to mean ‘ $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x$ ’. Note that we did *not* say ‘ $x$  is *the* limit of  $\langle x_n \rangle_{n \in \mathbb{N}}$ ’, because we have no right to do so unless and until we know that a sequence cannot converge to two distinct points. Fortunately, it is both true and easy to prove that limits of sequences in a metric space are unique.

thm:unique-seq-lim-metric

**1.65 Theorem (uniqueness of sequential limits).** Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence in a metric space  $\langle X, d \rangle$  which converges in  $\langle X, d \rangle$  both to a point  $x$  and to a point  $y$ . Then  $x = y$ .

**Proof.** Just suppose that  $x \neq y$ . By [Corollary 1.27](#) there are disjoint  $d$ -neighborhoods  $U$  of  $x$  and  $V$  of  $y$ . Since  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x$ , there is some  $m \in \mathbb{N}$  with

$$x_n \in U \quad (n \geq m).$$

Since the same sequence converges to  $y$ , there is some  $k \in \mathbb{N}$  with

$$x_n \in V \quad (n \geq k)$$

Then for  $n = \max\{m, k\}$  we have

$$x_n \in U \cap V,$$

which contradicts the disjointness of  $U$  from  $V$ .  $\square$

[The alert reader will have noticed that essentially the same argument was already used in the fourth paragraph of [Examples 1.61 \(7\)](#).]

[Theorem 1.65](#) justifies the following definition.

def:limit-seq-metric

**1.66 Definition (limit of sequence).** If  $\langle x_n \rangle_{n \in \mathbb{N}}$  is a convergent sequence in a metric space  $\langle X, d \rangle$ , then the unique point of  $X$  to which it converges is called **the  $d$ -limit of  $\langle x_n \rangle_{n \in \mathbb{N}}$** , or **the limit of  $\langle x_n \rangle_{n \in \mathbb{N}}$  in  $\langle X, d \rangle$** , and is denoted by  $d\text{-}\lim_{n \rightarrow \infty} x_n$ . When the metric  $d$  is understood—for example, when  $d$  is the Euclidean metric on  $\mathbb{R}$ , that unique point may be called simply **the limit of  $\langle x_n \rangle_{n \in \mathbb{N}}$**  and may be denoted by  $\lim_{n \rightarrow \infty} x_n$ .

If a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in  $\langle X, d \rangle$  is constant—that is, if there is some  $p \in X$  with  $x_n = p$  for all  $n$ —then clearly  $\lim_{n \rightarrow \infty} x_n = p$ . This is a result you

From calculus, if  $\langle x_n \rangle_{n \in \mathbb{N}}$  is a sequence of real numbers that converges, then so does the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  of their multiples by a constant  $c$ , and then

$$\{\text{eq:seq-lim-cst-multiple}\} \quad (1) \quad \lim_{n \rightarrow \infty} c x_n = c \lim_{n \rightarrow \infty} x_n.$$

Further, if  $\langle x_n \rangle_{n \in \mathbb{N}}$  and  $\langle y_n \rangle_{n \in \mathbb{N}}$  are sequences of real numbers that converge, then so do their sum, difference, product, and (if all  $y_n \neq 0$  and  $\lim_{n \rightarrow \infty} y_n \neq 0$ ) quotient sequences, and then

$$\{\text{eq:seq-lim-sum}\} \quad (2) \quad \lim_{n \rightarrow \infty} (x_n + y_n) = \left( \lim_{n \rightarrow \infty} x_n \right) + \left( \lim_{n \rightarrow \infty} y_n \right),$$

$$\{\text{eq:seq-lim-diff}\} \quad (3) \quad \lim_{n \rightarrow \infty} (x_n - y_n) = \left( \lim_{n \rightarrow \infty} x_n \right) - \left( \lim_{n \rightarrow \infty} y_n \right),$$

$$\{\text{eq:seq-lim-product}\} \quad (4) \quad \lim_{n \rightarrow \infty} (x_n y_n) = \left( \lim_{n \rightarrow \infty} x_n \right) \left( \lim_{n \rightarrow \infty} y_n \right),$$

$$\{\text{eq:seq-lim-quotient}\} \quad (5) \quad \lim_{n \rightarrow \infty} \frac{x_n}{y_n} = \frac{\lim_{n \rightarrow \infty} x_n}{\lim_{n \rightarrow \infty} y_n}.$$

Properties (1)–(3) hold as well in  $\mathbb{R}^n$ —see Exercises 89–90.

If we know which sets in a metric space  $\langle X, d \rangle$  are  $d$ -open, then we know which sequences converge to which points in the space. Surprisingly enough, the converse is true as well; this is a consequence of the next theorem and the fact that the  $d$ -open sets are just the complements of the  $d$ -closed sets.

thm:d-closed-iff-seq-limit-in-set

**1.67 Theorem (closedness in terms of sequential convergence).** *Let  $E$  be a subset of a metric space  $\langle X, d \rangle$ . Then the following statements are equivalent:*

cond:set-d-closed (i) *The subset  $E$  of  $X$  is  $d$ -closed.*

cond:seq-limit-in-set (ii) *Each point  $x \in X$  to which some sequence of points of  $E$  converges in  $\langle X, d \rangle$  itself belongs to the subset  $E$ .*

**Proof.** (i)  $\implies$  (ii). Assume (i). Let  $x \in X$  and let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence in the subset  $E$  that converges to a point  $x$  in  $\langle X, d \rangle$ . Just suppose that  $x \notin E$ . Then the set  $V = X \setminus E$  is a  $d$ -neighborhood of  $x$  such that  $x_n \notin V$  for each  $n$ . This contradicts the convergence of the sequence to  $x$ .

(ii)  $\implies$  (i). Assume (ii). Just suppose that  $E$  is not  $d$ -closed. Then  $X \setminus E$  is not  $d$ -open. This means there is some point  $x \in X \setminus E$  with the property that no  $d$ -ball at  $x$  is contained in  $X \setminus E$ . Hence for each  $n \in \mathbb{N}$  the  $d$ -ball of radius  $1/n$  at  $x$  meets  $E$ , and so we may choose some point

$$x_n \in B_{1/n}(x; d).$$

Then the sequence  $\text{seq} x_n$  so obtained is a sequence of points of  $E$ , and this sequence converges to  $x$  in  $\langle X, d \rangle$  because

$$d(x, x_n) < \frac{1}{n} \quad (n \in \mathbb{N}).$$

But  $x \notin E$ . This contradicts (ii).  $\square$

Because sequential convergence in a metric space determines which sets are closed there, it also determines which sets are open there.

or:  $d$ -open-via-sequential-convergence

**1.68 Corollary (openness in terms of sequential convergence).** *Let  $U$  be a subset of a metric space  $\langle X, d \rangle$ . Then  $U$  is  $d$ -open if and only if for each point  $x \in U$ , each sequence  $\text{seq} x_n$  in  $X$  that  $d$ -converges to  $x$  is eventually in  $U$ .*

subsec:metric-conv-vs-cont

### Relationship between convergence and continuity

The open sets in two metric spaces determine which functions from the one space to the other are continuous. And sequential convergence determines which sets in a metric space are open there. Hence sequential convergence in two metric spaces should determine which functions from one to the other are continuous. This is indeed true.

thm:cont-via-seq-conv-metric

**1.69 Theorem (characterization of continuity in terms of sequential convergence).** *Let  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$  be metric spaces, let  $f: X \rightarrow Y$  be a map, and let  $x \in X$ . Then a necessary and sufficient condition for  $f$  to be  $\langle d, d' \rangle$ -continuous at  $x$  is that for each sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converging to  $x$  in  $\langle X, d \rangle$ , the sequence  $\langle f(x_n) \rangle_{n \in \mathbb{N}}$  converge to  $f(x)$  in  $\langle Y, d' \rangle$ .*

**Proof.** Necessity. Assume that  $f$  is  $\langle d, d' \rangle$ -continuous at  $x$ . Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence that converges to  $x$  in  $\langle X, d \rangle$ . To show that  $\langle f(x_n) \rangle_{n \in \mathbb{N}}$  converges to  $f(x)$  in  $\langle Y, d' \rangle$ , let  $V$  be an arbitrary  $d'$ -neighborhood of  $f(x)$ . By [Theorem 1.53](#), there is a  $d$ -neighborhood  $U$  of  $x$  with

$$f(U) \subset V.$$

Choose  $m \in \mathbb{N}$  with

$$x_n \in U \quad (n \geq m).$$

Then

$$f(x_n) \in V \quad (n \geq m).$$

Sufficiency. Assume the condition holds. Let  $V$  be an arbitrary  $d'$ -neighborhood of  $f(x)$ . According to [Theorem 1.53](#), it suffices to show that  $V$  contains the image  $f(U)$  of some  $d$ -neighborhood  $U$  of  $x$ . Just suppose this is not the case. In particular, then, for each  $n \in \mathbb{N}$ ,

$$f(B_{1/n}(x; d)) \not\subset V,$$

and so we may choose a point  $x_n$  such that

$$x_n \in B_{1/n}(x; d), \quad f(x_n) \notin V.$$

Then the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  so obtained converges to  $x$  in  $\langle X, d \rangle$ , but certainly the image sequence  $\langle f(x_n) \rangle_{n \in \mathbb{N}}$  does not converge to  $f(x)$  in  $\langle Y, d' \rangle$ . This contradicts our original assumption.  $\square$

Thus when  $\langle x_n \rangle_{n \in \mathbb{N}} \xrightarrow{d} x$  and  $f$  is  $\langle d, d' \rangle$ -continuous at  $x$ , we have

$$f\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} f(x_n).$$

[Theorem 1.69](#) tells us that sequential convergence determines continuity. We close this section by showing that the reverse is true. More precisely, we shall demonstrate that any given sequence in a metric space converges precisely when a certain map taking values in that metric space is continuous. This fact is not at all essential to topology (and is

therefore infrequently stated), but it does provide some additional insight into the meaning of sequential convergence.

For motivation, recall that a sequence in a set  $X$  is actually a map  $n \mapsto x_n$  of  $\mathbb{N} \rightarrow X$ . If  $d$  is a metric on  $X$ , then to say that  $\text{seq} x_n$  converges in  $\langle X, d \rangle$  is to say something about the behavior of this map for “large” numbers in its domain  $\mathbb{N}$ . Now  $\mathbb{N} \subset \mathbb{R} \cup \{+\infty\}$ , and the “closer” an  $n \in \mathbb{N}$  is to  $+\infty$  in the extended real line  $\widehat{\mathbb{R}}$ , the “larger”  $n$  will be. Hence it is natural to extend the map  $n \mapsto x_n$  to all of  $\mathbb{N} \cup \{+\infty\}$  by appropriately assigning a value to  $+\infty$ .

extended realline  
convergent sequence!continuous ma  
convergent sequence!continuous ma  
extended realline  
convergent sequence

rem:seq-conv-as-cont-fn-metric **1.70 Remark.** Denote by  $d'$  the metric induced on the subset

$$\mathbb{N}' = \mathbb{N} \cup \{+\infty\}$$

of the extended real line

$$\widehat{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$$

by the metric  $\widehat{d}$  on  $\widehat{\mathbb{R}}$ , where  $d'$  is as defined in [Example 1.41](#). For a given sequence  $\langle x_n : n \in \mathbb{N} \rangle$  in a metric space  $\langle X, d \rangle$  and a given point  $x \in X$ , define a map

$$f: \mathbb{N}' \rightarrow X$$

by

$$\begin{aligned} f(n) &= x_n & (n \in \mathbb{N}), \\ f(+\infty) &= x. \end{aligned}$$

We claim that the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x$  in  $\langle X, d \rangle$  if and only if  $f$  is  $\langle d', d \rangle$ -continuous at  $+\infty$ .

To prove this assertion, observe first that each  $d'$ -neighborhood of  $+\infty$  in  $\mathbb{N}'$  contains a set of the form

$$U_m = \{n \in \mathbb{N} : n \geq m\} = (m-1, +\infty] \cap \mathbb{N}'$$

for some integer  $m \geq 1$ , and each set of this form is a  $d'$ -neighborhood of  $+\infty$  (see [Lemma 1.43](#)). For an arbitrary  $d$ -neighborhood  $V$  of  $x$  certainly  $f(+\infty) = x \in V$ , and so

$$f(u) \in V \text{ for all } u \in U_m \iff x_n \in V \text{ for all integers } n \geq m.$$

Now apply [Theorem 1.53](#) and [Theorem 1.62](#).  $\square$

## EXERCISES FOR SECTION 1.4

**68.** Suppose that in the condition

$$d(x, u) < \delta \implies d'(f(x), f(u))$$

of  $\langle d, d' \rangle$ -continuity at a point  $x$  ([Definition 1.51](#)) one or the other or both of the strict inequalities  $<$  were changed to the weak inequality  $\leq$ . Would this change which maps  $f: X \rightarrow Y$  are  $\langle d, d' \rangle$ -continuous at  $x$ ?

**69.** If  $d$  is the Euclidean metric on  $\mathbb{R}$ , determine whether the given function  $f: \mathbb{R} \rightarrow \mathbb{R}$  is  $\langle d, d \rangle$ -continuous, and if it is not find all points at which it is  $\langle d, d \rangle$ -discontinuous:

- (a)  $f(x) = x \sin(1/x)$  if  $x \neq 0$ , and  $f(0) = 0$ ,
- (b)  $f(x) = x/|x|$  if  $x \neq 0$ , and  $f(0) = 0$ .
- (c)  $f(x) = 0$  if  $x$  is irrational, and  $f(x) = 1$  if  $x$  is rational.

- (d)  $f(x) = 0$  if  $x$  is irrational, and  $f(x) = 1/q$  if  $x$  is rational with  $q$  being that positive integer for which  $x = p/q$  where  $p$  is an integer that is relatively prime to  $q$  (that is, 1 is the only positive integer that divides both  $p$  and  $q$ ).

70. (a) If  $d$  is the Euclidean metric on  $\mathbb{R}$  and  $d_2$  is the Euclidean metric on  $\mathbb{R} \times \mathbb{R}$ , show that the operations

$$\begin{aligned} \mathbb{R} \times \mathbb{R} &\rightarrow \mathbb{R} & \mathbb{R} \times \mathbb{R} &\rightarrow \mathbb{R} \\ (x, y) &\mapsto x + y & (x, y) &\mapsto xy \end{aligned}$$

of addition and multiplication, respectively, are  $\langle d_2, d \rangle$ -continuous.

- (b) If  $d$  is the Euclidean metric on  $\mathbb{R}$ , show that the map

$$\begin{aligned} \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto -x \end{aligned}$$

is  $\langle d, d \rangle$ -continuous.

- (c) If  $d$  is the Euclidean metric on  $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ , show that the map

$$\begin{aligned} \mathbb{R}^* &\rightarrow \mathbb{R}^* \\ x &\mapsto 1/x \end{aligned}$$

is  $\langle d, d \rangle$ -continuous.

- (d) If  $d$  is the Euclidean metric on  $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ , show that the map

$$\begin{aligned} \mathbb{R}^* &\rightarrow \mathbb{R}^* \\ x &\mapsto 1/x \end{aligned}$$

is  $\langle d, d \rangle$ -continuous.

- (e) Show that the operations

$$\begin{aligned} \mathbb{R} \times \mathbb{R} &\rightarrow \mathbb{R} & \mathbb{R} \times \mathbb{R}^* &\rightarrow \mathbb{R} \\ (x, y) &\mapsto x - y & (x, y) &\mapsto x/y \end{aligned}$$

of subtraction and division, respectively, are continuous with respect to appropriate metrics.

sum-and-scalar-multiplication-cont

71. Show that the vector space operations

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R}^n &\rightarrow \mathbb{R}^n & \mathbb{R} \times \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ (x, y) &\mapsto x + y & (\alpha, x) &\mapsto \alpha x \end{aligned}$$

of addition and scalar multiplication, respectively, are continuous with respect to appropriate metrics.

or-diff-and-mult-by-fixed-scalar-cont

72. (Continuation of [Exercise 71](#).)

Deduce that the operations

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R}^n &\rightarrow \mathbb{R}^n & \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ (x, y) &\mapsto x - y & x &\mapsto \lambda x \end{aligned}$$

are continuous with respect to appropriate metrics. For the second operation,  $\lambda$  is a fixed scalar.

73. Show that the Euclidean inner product operation

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R}^n &\rightarrow \mathbb{R} \\ (x, y) &\mapsto x \cdot y \end{aligned}$$

is continuous with respect to appropriate metrics.

ob:metric-open-map-nowhere-cont

- 74.** Construct metric spaces  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$  and a map  $f: X \rightarrow Y$  such that  $f(U)$  is  $d'$ -open for each  $d$ -open subset  $U$  of  $X$ , yet  $f$  is  $\langle d, d' \rangle$ -continuous at *no* point of  $X$ . projection Hilbert sequence space

-part:proj-finite-prod-metric-is-cont

- 75. (a)** Given metric spaces  $(X_1, d_1), (X_2, d_2), \dots, (X_n, d_n)$ , let  $d_\infty$  be the max metric induced by  $(d_1, d_2, \dots, d_n)$  on the product set  $X = \times_{i=1}^n X_i$  (see [Example 1.14](#)). Prove that for each  $j = 1, 2, \dots, n$ , the  $j$ th projection

$$\begin{aligned} p_j: X &\rightarrow X_j \\ x &\mapsto x_j \end{aligned}$$

is  $\langle d_\infty, d_j \rangle$ -continuous.

- (b)** Give an analog of **(a)** for the product of a sequence of bounded metric spaces (see [Exercise 51](#)).

aps-from-seq-spaces-metric-cont-Q

- 76.** If  $d'$  is the Euclidean metric on  $\mathbb{R}$ , is it true that for all  $n \in \mathbb{N}$  the maps

$$\begin{aligned} X &\rightarrow \mathbb{R} \\ x &\mapsto x_n \end{aligned}$$

are  $\langle d, d' \rangle$ -continuous when:

- (a)** the set  $X = \ell^2$ , the Hilbert sequence space, and  $d$  is the metric  $d_2$  on  $X$  constructed in [Example 1.10](#)?
- (b)** the set  $X$  is the set of all sequences in  $\mathbb{R}$  and  $d$  is the metric on  $X$  constructed in [Example 1.15](#)?

uation-map-and-def-integral-cont-Q

- 77.** Let  $X = C([0, 1])$ , the set of all continuous functions  $x: [0, 1] \rightarrow \mathbb{R}$ , and let  $d'$  be the Euclidean metric on  $\mathbb{R}$ .

- (a)** Is it true that for all  $t \in [0, 1]$  the “evaluation maps”

$$\begin{aligned} X &\rightarrow \mathbb{R} \\ x &\mapsto x(t) \end{aligned}$$

are  $\langle d, d' \rangle$ -continuous if  $d = d_\infty$ , the sup metric on  $X$  ([Example 1.8](#))? if  $d = d_1$ , the metric constructed in [Example 1.9](#)?

- (b)** Is the map

$$\begin{aligned} X &\rightarrow \mathbb{R} \\ x &\mapsto \int_0^1 x(t) \, dt \end{aligned}$$

$\langle d_\infty, d' \rangle$ -continuous? Is it  $\langle d_1, d' \rangle$ -continuous?

prob:inv-im-closed-is-closed-metric

- 78.** Let  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$  be metric spaces and let  $f: X \rightarrow Y$  be a map. Prove that if  $f$  is  $\langle d, d' \rangle$ -continuous, then the inverse image  $f^{-1}(E)$  of each  $d'$ -closed subset  $E$  of  $Y$  is a  $d$ -closed subset of  $X$ .

Does the converse hold?

prob:ip-cont-metric

- 79.** Let  $d$  be the Euclidean metric on  $\mathbb{R}^n$  and let  $d_\infty$  be the max metric that it induces on  $\mathbb{R}^n \times \mathbb{R}^n$ . Prove that the Euclidean inner product on  $\mathbb{R}^n$  ([Exercise 63](#)) is  $\langle d_\infty, d' \rangle$ -continuous, where  $d'$  is the Euclidean metric on  $\mathbb{R}$ .

hyperhalfspaces

Euclidean inner product

open-halfspace

closed-halfspace

polyhedron

convex set

distance!point to set@from point to set

Urysohn's lemma!metrizable space@and metrizable space

completely regular space!metrizable space@and metrizable space

metrizable space!completely regular space@as completely regular space

**80. A hyperplane in  $\mathbb{R}^n$  is a set of the form**

$$\{x \in \mathbb{R}^n : \langle x|a \rangle = c\}$$

for some point  $a \in \mathbb{R}^n$  and some real number  $c$ , where  $\langle x|a \rangle$  is the Euclidean inner product as defined in [Exercise 63](#); an **open-halfspace in  $\mathbb{R}^n$**  is a set of the form

$$\{x \in \mathbb{R}^n : \langle x|a \rangle < c\}$$

for some  $a \in \mathbb{R}^n$  and some  $c \in \mathbb{R}$ ; and a **closed-halfspace in  $\mathbb{R}^n$**  is a set of the form

$$\{x \in \mathbb{R}^n : \langle x|a \rangle \leq c\}$$

for some  $a \in \mathbb{R}^n$  and some  $c \in \mathbb{R}$ .

**(a)** Show that every hyperplane, open-halfspace, closed-halfspace, or polyhedron in  $\mathbb{R}^n$  is convex ([page 37](#)).

**(b)** Prove that each open-halfspace in  $\mathbb{R}^n$  is  $d$ -open, where  $d$  is the Euclidean metric on  $\mathbb{R}^n$ . Prove also that each hyperplane, each closed-halfspace, and each polyhedron in  $\mathbb{R}^n$  is  $d$ -closed. (*Hint:* Use [Exercise 78](#) and [Exercise 79](#).)

prob:dist-pt-to-set-continuous

**81.** Let  $A$  be a subset of a metric space  $\langle X, d \rangle$ . Let  $d'$  be the Euclidean metric on  $\mathbb{R}$ . Prove;

prob-part:dist-pt-to-set-continuous

**(a)** The map

$$\begin{aligned} X &\rightarrow \mathbb{R} \\ x &\mapsto d(x, A) \end{aligned}$$

is  $\langle d, d' \rangle$ -continuous.

**(b)** Given  $x \in X$ , there is a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in  $A$  such that the sequence  $\langle d(x, x_n) \rangle_{n \in \mathbb{N}}$  converges to  $d(x, A)$  in  $(\mathbb{R}, d')$ .

prob:Urysohn-for-metrizable

**82.** Prove the following special case of Urysohn's Lemma ([6.26](#)): If  $A$  and  $B$  are disjoint closed sets in a metrizable space  $X$ , then there is a continuous function  $f: X \rightarrow [0, 1]$  with  $f(a) = 0$  for all  $a \in A$  and  $f(b) = 1$  for all  $b \in B$ . [*Hint:* If  $d$  induces the topology of  $X$ , consider the function defined by  $f(x) = d(x, A) / (d(x, A) + d(x, B))$ .]  
*Note:* In the language of the “separation properties” introduced later ([Definition 2.94](#)), it says that a metrizable space is *completely regular*.

**83.** Express the definition ([1.57](#)) that a sequence converges to a point  $x$  in a metric space  $\langle X, d \rangle$  purely in symbols, with all needed quantifiers included.

prob:seq-equal-eventually-converge

**84.** Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  and  $\langle y_n \rangle_{n \in \mathbb{N}}$  be two sequences in a metric space  $\langle X, d \rangle$  that are “eventually equal” in the sense that there is an index  $k \in \mathbb{N}$  such that  $x_n = y_n$  for all  $n \geq k$ . Show that one of these sequences will converge in  $\langle X, d \rangle$  if and only if the other does, and that in this case they will converge to the same point.

**85.** Establish the analog of [Example 1 \(1\)](#) for decreasing sequences of real numbers.

**86.** Define  $d: \mathbb{N}^* \times \mathbb{N}^* \rightarrow \mathbb{R}$  by  $d(m, n) = |1/m - 1/n|$ .

**(a)** Verify that  $d$  is a metric on  $\mathbb{N}^*$ .

**(b)** Which sequences in  $\langle \mathbb{N}^*, d \rangle$  converge there?

**87.** Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a convergent sequence in a metric space  $\langle X, d \rangle$ . Prove: If  $\lim_{n \rightarrow \infty} x_n = p$ , then also  $\lim_{n \rightarrow \infty} x_{n+1} = p$ .

prob:real-ops-preserve-limits

**88.** Prove properties (1)–(5) from [page 193](#).



ob:vector-space-ops-preserve-limits

- 89.** Prove that the vector space operations of addition and scalar multiplication in  $\mathbb{R}^n$  preserve limits of sequences. More precisely, prove:

Cantor-Bernstein Theorem  
p-adic metric@Sp\$-adic metric

- (a) Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  and  $\langle y_n \rangle_{n \in \mathbb{N}}$  be sequences in  $\mathbb{R}^n$  that converge in  $\langle \mathbb{R}^n, d_n \rangle$ , where  $d_n$  is the Euclidean metric. Then the sequence  $\langle x_n + y_n \rangle_{n \in \mathbb{N}}$  of sums also converges there and, in fact,

$$\lim_{n \rightarrow \infty} (x_n + y_n) = \left( \lim_{n \rightarrow \infty} x_n \right) + \left( \lim_{n \rightarrow \infty} y_n \right).$$

- (b) Let  $\langle \alpha_n \rangle_{n \in \mathbb{N}}$  be a sequence of scalars that converges in  $\langle \mathbb{R}, d \rangle$ , where  $d$  is the Euclidean metric; and let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence in  $\mathbb{R}^n$  that converges in  $\langle \mathbb{R}^n, d_n \rangle$ , where  $d_n$  is the Euclidean metric. Then the sequence  $\langle \alpha x_n \rangle_{n \in \mathbb{N}}$  of scalar multiples also converges there, and, in fact,

$$\lim_{n \rightarrow \infty} (\alpha_n x_n) = \left( \lim_{n \rightarrow \infty} \alpha_n \right) \left( \lim_{n \rightarrow \infty} x_n \right).$$

mult-by-fixed-scalar-preserve limits

- 90.** (Continuation of [Exercise 89](#).) Deduce that the vector space operation of subtraction in  $\mathbb{R}^n$  preserves limits of sequences, as does the operation of multiplication of vectors by a fixed scalar.

prob:find-seq-lim-assuming-exists

- 91.** Define the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  of reals recursively by

$$\begin{aligned} x_0 &= 1, \\ x_{n+1} &= \sqrt{2 + x_n} \quad (n = 0, 1, 2, \dots). \end{aligned}$$

- (a) Assuming that  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges in  $\mathbb{R}$ , determine the value of  $L = \lim_{n \rightarrow \infty} x_n$ . [Hint: If  $\lim_{n \rightarrow \infty} x_n = L$ , then also  $\lim_{n \rightarrow \infty} x_{n+1} = L$ . From this and the recurrence relation obtain an algebraic equation that  $L$  must satisfy.]

- (b) Prove that  $\langle x_n \rangle_{n \in \mathbb{N}}$  does actually converge to that  $L$ . [Hint: Define  $\varepsilon_n = L - x_n$  for each  $n$ . Use the recurrence relation to express  $\varepsilon_{n+1}$  first in terms of  $x_n$  and then in terms of  $\varepsilon_n$ . Change the form of the expression you obtain for  $\varepsilon_{n+1}$  in terms of  $\varepsilon_n$  so as to establish that  $\varepsilon_{n+1} < \varepsilon_n/2$ . Conclude by showing that then  $\varepsilon_n \leq 1/2^n$  for every  $n$ .]

- 92.** The sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  of reals is defined recursively by  $x_0 = 1/2$  and  $x_{n+1} = (1 + x_n^2)/2$  for  $n = 0, 1, 2, \dots$ .

part:recursive-monotonic-bded-seq

- (a) Show that this sequence converges by verifying that it is either increasing or decreasing) and bounded. Then find its limit.  
(b) Redo (a) but with the initial condition changed to  $x_0 = 2$ .

prob:inject-powerset-Q-into-R

- 93.** (a) Show that the map  $\{0, 2\}^{\mathbb{N}^*} \rightarrow \mathbb{R}$  given by  $\langle d_j \rangle_{j \in \mathbb{N}^*} \mapsto \sum_{j=1}^{\infty} d_j/3^j$  is injective.  
(b) Deduce the existence of an injection  $\mathcal{P}(\mathbb{Q}) \rightarrow \mathbb{R}$ ; in other words, show that  $\text{card } \mathcal{P}(\mathbb{Q}) \leq \text{card } \mathbb{R}$ .

prob-part-exists-inj-powerset-Q-to-R

Note: According to [Exercise 0.126](#), also  $\text{card } \mathbb{R} \leq \text{card } \mathcal{P}(\mathbb{Q})$ . In fact, from either the Cantor-Bernstein Theorem ([0.126](#)) or [Example 0.121](#), it is true that  $\text{card } \mathcal{P}(\mathbb{Q}) = \text{card } \mathbb{R}$ .

- 94.** Let  $p$  be a prime number. Show that the sequence  $\langle p^n \rangle_{n \in \mathbb{N}}$  converge in the metric space  $(\mathbb{Q}, d_p)$ , where  $d_p$  is the  $p$ -adic metric ([Exercise 24](#)), but not in  $\langle \mathbb{R}, d \rangle$ , where  $d$  is the Euclidean metric.

- 95.** (a) Prove an analog of [Example 1.64](#) for the product of a sequence of metric spaces (see [Example 1.15](#)).

infinite limit! sequence of reals  
 limit of a sequence! infinite  
 Cauchy sequence  
 Cauchy, Augustin-Louis

(b) Does the analog of [Example 1.64](#) hold for the Hilbert sequence space  $(\ell^2, d_2)$  of [Example 1.10](#).

96. Let  $X = C([0, 1])$ , the set of all continuous functions  $x: [0, 1] \rightarrow \mathbb{R}$ . Denote the Euclidean metric on  $\mathbb{R}$  by  $d$ .

- (a) Construct a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in  $X$  with the property that for each  $t \in [0, 1]$  the sequence  $\langle x_n(t) \rangle_{n \in \mathbb{N}}$  of real numbers converges in  $(\mathbb{R}, d)$ , yet the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  of functions does *not* converge in  $(X, d_\infty)$ , where  $d_\infty$  is the sup metric on  $X$  ([Example 1.8](#)).
- (b) Does the sequence you constructed in (a) converge in  $(X, d_1)$ , where  $d_1$  is the metric defined in [Example 1.9](#) in terms of integrals?

97. (a) Write down the calculus definition of

$$\lim_{n \rightarrow \infty} x_n = +\infty$$

for a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  of real numbers. Explain what this statement means in the context of sequential convergence in a suitable metric space.

(b) Do the same for the statement

$$\lim_{n \rightarrow \infty} x_n = -\infty.$$

98. The definitions of continuity and sequential convergence given in this section still make sense if we deal with pseudometrics ([Exercise 20](#)) instead of metrics. Check that all general results of this section except the uniqueness of sequential limits ([Theorem 1.65](#)) extend to the pseudometric case.

## 1.5 Completeness

sec:complete

To prove that a given sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in a metric space converges, sometimes we can proceed by first determining the point  $x$  to which it ought to converge—assuming that the sequence converges—and second by checking that  $\langle x_n \rangle_{n \in \mathbb{N}}$  does actually converge to  $x$ . (See [Exercise 91](#) for such an example.) Many times, though, such a procedure is not applicable. And often it is of greater importance to know *that* a sequence converges rather than to know to which particular point it converges.

What is needed, then, is a “convergence criterion” for testing whether a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges by looking only at the values  $x_n$  of the sequence itself, not at some other point  $x$  as well. We have already seen one such convergence criterion in the case of sequences of real numbers: an increasing sequence of reals that is bounded above in  $\mathbb{R}$  converges there—see [Examples 1.61 \(1\)](#). In this section we shall introduce a different convergence criterion that is not restricted to sequences in  $\mathbb{R}$ .

### Cauchy sequences

Roughly speaking, a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges when its values  $x_n$  get closer and closer to some fixed point  $x$  for larger and larger values of  $n$ . Then a convergent sequence should have the property that its values  $x_n$  get closer and closer to *one another* for larger and larger values of  $n$ . It was precisely this property that Cauchy introduced in the early nineteenth century as a criterion for convergence of sequences of real numbers. In the remainder section we are going to study metric spaces in which Cauchy’s condition always guarantees convergence.

def:cauchy-seq-metric

**1.71 Definition (Cauchy sequence).** A sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in a metric space  $\langle X, d \rangle$  is said to be a **Cauchy sequence in  $\langle X, d \rangle$** , or to be a  **$d$ -Cauchy sequence**, when:

For each  $\varepsilon > 0$ , there exists some  $m \in \mathbb{N}$  such that  
 $d(x_n, x_k) < \varepsilon$  for all  $n \geq m$  and all  $k \geq m$ .

convergent sequence!Cauchy sequence!

Cauchy sequence!bounded sequence!

Many examples of Cauchy sequences are furnished by the first part of the following result, which also formalizes the idea that values of a sequence get close to one another if they get close to some fixed point.

prop:convergent-and-cauchy-and-bded

**1.72 Proposition (sequential convergence and Cauchy sequence).** Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence in a metric space  $\langle X, d \rangle$ .

prop-part:conv-implies-cauchy

(1) If  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges in  $\langle X, d \rangle$ , then it is a Cauchy sequence in  $\langle X, d \rangle$ .

prop-part:cauchy-implies-bded

(2) If  $\langle x_n \rangle_{n \in \mathbb{N}}$  is a Cauchy sequence in  $\langle X, d \rangle$ , then it is  $d$ -bounded, that is, its range  $\{x_n : n \in \mathbb{N}\}$  is  $d$ -bounded.

More tersely: (1) every convergent sequence is a Cauchy sequence; and (2) every Cauchy sequence is bounded.

**Proof.** (1) Assume that  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges in  $\langle X, d \rangle$  to a point  $x$ . Let  $\varepsilon > 0$  be arbitrary. There exists  $m \in \mathbb{N}$  with

$$d(x_n, x) < \frac{\varepsilon}{2} \quad (n \geq m).$$

Then  $n \geq m$  and  $k \geq m$  implies

$$d(x_n, x_k) \leq d(x_n, x) + d(x, x_k) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

(2) Assume that  $\langle x_n \rangle_{n \in \mathbb{N}}$  is a Cauchy sequence in  $\langle X, d \rangle$ . Corresponding to  $\varepsilon = 1$  there is an  $m \in \mathbb{N}$  with

$$d(x_n, x_k) < 1 \quad (n \geq m, k \geq m).$$

Let

$$b = \max\{d(x_n, x_k) : 1 \leq n \leq m, 1 \leq k \leq m\}.$$

Then

$$d(x_n, x_k) \leq 1 + b$$

for all  $n, k \in \mathbb{N}$ .  $\square$

Neither of the implications above is reversible, as the following examples demonstrate.

ex:convergent-but-not-cauchy-sequence

**1.73 Examples.** (1) Let  $\langle Y, d' \rangle$  be a metric space and let  $d$  be the metric induced by  $d'$  on a subset  $X$  of  $Y$ . Suppose  $\langle x_n \rangle_{n \in \mathbb{N}}$  is a sequence in  $X$  that converges to a point  $y \in Y \setminus X$  in  $\langle Y, d' \rangle$ . By uniqueness of sequential limits ([Theorem 1.65](#)), the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  does *not* converge in  $\langle X, d \rangle$ . However,  $\text{seq } x_n$  is a Cauchy sequence in  $\langle X, d \rangle$  because by [Proposition 1.72 \(1\)](#) it is a Cauchy sequence in  $\langle Y, d' \rangle$ .

ex:cauchy-seq-in-Q-not-conv-in-Q

(2) As a specific instance of (1), take  $Y = \mathbb{R}$  and  $X = \mathbb{Q}$ , let  $d'$  and  $d$  be the Euclidean metrics, and let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be the sequence constructed in [Examples 1.61 \(7\)](#).

ex:bded-non-cauchy-sequence (3) Let  $d$  be the Euclidean metric on  $\mathbb{R}$ . Then the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  given by  
 Cauchy complete metric space @ Cauchy complete metric space  
 rational numbers! incomplete space @ as incomplete space  
 open interval! incomplete space @ as incomplete space

$$x_n = (-1)^n \quad (n \in \mathbb{N})$$

is *not* a Cauchy sequence in  $\langle \mathbb{R}, d \rangle$  because  $d(x_n, x_{n+1}) = 2$  for all  $n$ . However, this sequence is  $d$ -bounded because  $\{x_n : n \in \mathbb{N}\} = \{-1, 1\}$ .  $\diamond$

### Complete metric spaces

subsec:complete-metric-spaces

Let us name those metric spaces that, unlike those in Examples 1.73 (1)–(2), *do* contain all the points that “ought to be there” from the viewpoint of convergence.

def:complete **1.74 Definition (completeness).** Both a metric space  $\langle X, d \rangle$  and its metric  $d$  are said to be **complete** when each Cauchy sequence in  $\langle X, d \rangle$  converges in  $\langle X, d \rangle$ . A metric space that is not complete is said to be **incomplete**.

To distinguish this concept from other kinds of “completeness”, a metric space that is complete in the preceding sense is sometimes said to be **Cauchy complete**.

complete metric spaces **1.75 Examples.** (1) The simplest example of a complete metric is the discrete metric  $\delta$  on any set  $X$ .

In fact, suppose  $\langle x_n \rangle_{n \in \mathbb{N}}$  is a Cauchy sequence in  $\langle X, \delta \rangle$ . Corresponding to  $\varepsilon = 1$ , there is some  $m \in \mathbb{N}$  with  $\delta(x_n, x_m) < 1$  for all  $n \geq m$ , that is,  $x_n = x_m$  for all  $n \geq m$ . According to Examples 1.61 (9), this is precisely the condition needed for  $\langle x_n \rangle_{n \in \mathbb{N}}$  to converge in  $\langle X, \delta \rangle$ .

ex:rational-not-complete (2) In view of Examples 1.73 (2), *the set of rational numbers with its Euclidean metric is not complete*.

ex:open-interval-not-complete (3) In (1), take  $Y = \mathbb{R}$  and  $X = ]0, 1[$ , and let  $d'$  and  $d$  be the Euclidean metrics. Then the sequence  $\langle 1/(n+1) \rangle_{n \in \mathbb{N}}$  is a Cauchy sequence in  $\langle \mathbb{Q}, d \rangle$  that does not converge in  $\langle \mathbb{Q}, d \rangle$ . More generally, *The Euclidean metric on an open interval in  $\mathbb{R}$  is not complete*.  $\diamond$

The single most important example of a complete metric space is the real line with its Euclidean metric.

thm:reals-complete **1.76 Theorem (completeness of the real line).** The Euclidean metric on  $\mathbb{R}$  is complete.

**Proof.** Denote the Euclidean metric on  $\mathbb{R}$  by  $d$ . Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be an arbitrary Cauchy sequence in  $\langle \mathbb{R}, d \rangle$ . Let

$$R = \{x_n : n \in \mathbb{N}\},$$

the range of the sequence.

Case (i): the range  $R$  is finite. In this case, there is some  $x \in \mathbb{R}$  with  $x_n = x$  for infinitely many values of  $n$ . We show that  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x$  in  $\langle \mathbb{R}, d \rangle$ . Let  $\varepsilon > 0$  be arbitrary. There is some  $i \in \mathbb{N}$  such that

$$|x_n - x_m| < \varepsilon \quad (n \geq i, m \geq i).$$

Since  $\{m \in \mathbb{N} : x_m = x\}$  is infinite whereas  $\{m \in \mathbb{N} : m < i\}$  is finite, there is an  $m \geq i$  with  $x_m = x$ . Then  $n \geq m$  implies  $|x_n - x| < \varepsilon$ .

Case (ii): the range  $R$  is infinite. Since  $R$  is  $d$ -bounded (Proposition 1.60),

$$R \subset [a, b]$$

for some reals  $a$  and  $b$  with  $a < b$ .

We apply a bisection argument to the interval

$$[a_0, b_0] = [a, b].$$

Next, since  $R$  is infinite, bisecting  $[a_0, b_0]$  produces two subintervals at least one of which still contains infinitely many points of  $R$ ; call such an interval  $[a_1, b_1]$ . Since  $R \cap [a_1, b_1]$  is infinite, bisecting  $[a_1, b_1]$  produces two subintervals at least one of which still contains infinitely many points of  $R$ ; call such an interval  $[a_2, b_2]$ . Continuing in this way, we obtain a nested sequence

$$[a_0, b_0] \supset [a_1, b_1] \supset \cdots [a_n, b_n] \supset [a_{n+1}, b_{n+1}] \supset \cdots$$

of closed intervals with the  $n$ th interval having length

$$b_n - a_n = \frac{b_0 - a_0}{2^n}$$

and containing infinitely many points of  $R$ . By the Nested Interval Property (Theorem 0.83), there exists a unique point

$$x \in \bigcap_{n=0}^{\infty} [a_n, b_n].$$

To complete the proof in this case we show that  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to this  $x$ . Let  $\varepsilon > 0$  be arbitrary. Because we have a Cauchy sequence, there is some  $i \in \mathbb{N}$  such that

$$|x_n - x_k| < \frac{\varepsilon}{2} \quad (n \geq i, k \geq i).$$

By the Archimedean Ordering Property (Theorem 0.78), there is some  $m \in \mathbb{N}$  such that

$$m \geq i, \quad b_m - a_m < \frac{\varepsilon}{2}.$$

Since  $R \cap [a_m, b_m]$  is infinite, there is some  $k \in \mathbb{N}$  such that

$$k \geq i, \quad x_k \in [a_m, b_m].$$

Then  $n \geq m$  implies

$$|x_n - x| \leq |x_n - x_k| + |x_k - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

(see Figure 1.21).  $\square$

The preceding proof used in a crucial way both the Nested Interval Property and the Archimedean Ordering Property of the real numbers. Hence Cauchy completeness of the Euclidean metric on  $\mathbb{R}$  is a consequence of the order-completeness of  $\mathbb{R}$ .

According to Examples 1.73 (1), the set  $\mathbb{Q}$  of all rational numbers with its Euclidean metric  $d$  is not complete. We have just proved that the larger set  $\mathbb{R}$  of all real numbers with its Euclidean metric  $d'$  is complete. Thus each Cauchy sequence in the smaller space  $\langle \mathbb{Q}, d \rangle$  converges in the larger space  $\langle \mathbb{R}, d' \rangle$ . Moreover, it is clear from the order-density of  $\mathbb{Q}$  in  $\mathbb{R}$  (Corollary 0.81) that  $\mathbb{Q}$  is “ $d'$ -dense” in  $\mathbb{R}$  in the sense that for each point  $x$  of the larger set  $\mathbb{R}$  there is some sequence in the smaller set  $\mathbb{Q}$  which converges to  $x$  in  $\langle \mathbb{R}, d' \rangle$ . Hence by passing from  $\mathbb{Q}$  to  $\mathbb{R}$  we supply all the points that are “missing” from  $\mathbb{Q}$ , and so we say that  $\langle \mathbb{R}, d' \rangle$  completes  $\langle \mathbb{Q}, d \rangle$ .



Figure 1.21: Proving that a Cauchy sequence in the real line converges in the case that it has infinite range.

fig:pf-R-complete-case-ii

More generally, let  $\langle X, d \rangle$  be any metric space. Then we can construct a **completion of**  $\langle X, d \rangle$ , that is, a complete metric space  $\langle X', d' \rangle$  such that  $X \subset X'$ , the metric  $d'$  induces  $d$  on  $X$ , and  $X$  is  $d'$ -dense in the same sense as above. The construction of such a completion is outlined in [Exercise 130](#). (For technical reasons, the definition of a completion there is slightly different from the one just described.)

From completeness of the real line ([Theorem 1.76](#)) we are going to deduce that the Euclidean metric on  $\mathbb{R}^k$  is complete for every  $k \geq 1$ . For that, we establish first a more general result.

**1.77 Proposition (completeness of max metric).** *Let  $\langle X_1, d_1 \rangle, \langle X_1, d_1 \rangle, \dots, \langle X_k, d_k \rangle$  be  $k$  complete metric spaces. Then the max metric  $d_\infty$  induced by  $\langle d_1, d_2, \dots, d_k \rangle$  on the product*

$$X = \prod_{i=1}^k X_i$$

*is also complete.*

**Proof.** Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a Cauchy sequence in  $\langle X, d_\infty \rangle$ . As in [Example 1.64](#), for each  $n$  denote the  $i$ th coordinate of  $x_n$  by  $x_n^i$ , so that

$$x_n = \langle x_n^1, x_n^2, \dots, x_n^i, \dots, x_n^k \rangle.$$

We show that for each  $i = 1, 2, \dots, k$ , the sequence  $\langle x_n^i \rangle_{n \in \mathbb{N}}$  is a Cauchy sequence in  $\langle X_i, d_i \rangle$ . Let  $\varepsilon > 0$  be arbitrary. There exists  $m \in \mathbb{N}$  such that

$$d_\infty(x_n, x_j) < \varepsilon \quad (n \geq m, j \geq m).$$

By definition of  $d_\infty$ , for each  $i = 1, 2, \dots, k$  we have

$$d_i(x_n^i, x_j^i) < \varepsilon \quad (n \geq m, j \geq m).$$

For each  $i = 1, 2, \dots, k$  the metric space  $\langle X_i, d_i \rangle$  is complete and so the Cauchy sequence  $\langle x_n^i \rangle_{n \in \mathbb{N}}$  converges to some point  $x^i$  in  $\langle X_i, d_i \rangle$ . Take

$$x = \langle x^1, x^2, \dots, x^i, \dots, x^k \rangle \in X.$$

Then exactly as in [Example 1.64](#) we see that  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x$  in  $\langle X, d_\infty \rangle$ .  $\square$

cor:Rn-complete

**1.78 Corollary (completeness of Euclidean metric).** *The Euclidean metric on Euclidean  $k$ -space  $\mathbb{R}^k$  is complete.*

**Proof.** Denote the Euclidean metric on  $\mathbb{R}^k$  by  $d$ . Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a Cauchy sequence in  $\langle \mathbb{R}^k, d \rangle$ . From the inequality

$$d(x, y) \leq \sqrt{k} d_\infty(x, y) \quad (x, y \in \mathbb{R}^k)$$

established in the course of proving [Proposition 1.36](#), it follows that  $\langle x_n \rangle_{n \in \mathbb{N}}$  is also a Cauchy sequence in  $\langle \mathbb{R}^k, d_\infty \rangle$ , where  $d_\infty$  is the max metric. By completeness of the real line ([Theorem 1.76](#)) and [Proposition 1.77](#), the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  also converges in  $\langle \mathbb{R}^k, d_\infty \rangle$ . Now by [Proposition 1.36](#) the metrics  $d$  and  $d_\infty$  are equivalent. From [Corollary 1.63](#) we conclude that  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x$  in  $\langle \mathbb{R}^k, d \rangle$ .  $\square$

**Caution!** Do not be misled by the preceding proof into believing that it is always true that if one of two equivalent metrics is complete, then the other is as well. The following example demonstrates that this is *not* true.

ex:non-complete-metric-on-R

**1.79 Example.** Denote by  $d^*$  the bounded metric induced on  $\mathbb{R}$  by the bounded metric  $\widehat{d}$  on the extended real line  $\widehat{\mathbb{R}}$  that was constructed in [Example 1.41](#). By [Proposition 1.42](#), this metric  $d^*$  is equivalent to the Euclidean metric  $d$  on  $\mathbb{R}$ . We already saw in [Theorem 1.76](#) that  $\langle \mathbb{R}, d \rangle$  is complete. We claim, however, that  $\langle \mathbb{R}, d^* \rangle$  is *not* complete.

To justify our claim, consider the bijection

$$\varphi: \mathbb{R} \rightarrow ]-1, 1[$$

that was also constructed in [Example 1.41](#). Denote the Euclidean metric on  $]-1, 1[$  by  $d'$ . The bijection  $\varphi$  is an isometry from  $\langle \mathbb{R}, d^* \rangle$  to  $\langle ]-1, 1[, d' \rangle$ , because  $x, y \in \mathbb{R}$  implies

$$\begin{aligned} d'(\varphi(x), \varphi(y)) &= |\varphi(x) - \varphi(y)| \\ &= \widehat{d}(x, y) \\ &= d^*(x, y). \end{aligned}$$

Then  $\langle \mathbb{R}, d^* \rangle$  will be complete if and only if  $\langle ]-1, 1[, d' \rangle$  is complete. However, according to [Examples 1.73 \(3\)](#), the latter metric space is *not* complete.  $\diamond$

Despite the preceding example, it is nonetheless true that any complete metric is equivalent to a bounded complete metric.

metric-has-equiv-bded-complete-metric

**1.80 Proposition.** *Let  $\langle X, d \rangle$  be a complete metric space. then the bounded metric  $d^*$  on  $X$  given by*

$$d^*(x, y) = \min\{1, d(x, y)\} \quad (x, y \in X)$$

*is also complete.*

**Proof.** That  $d^*$  is a metric on  $X$  that is bounded and equivalent to  $d$  was established in [Proposition 1.40](#).

By [Corollary 1.63](#), the convergent sequences in  $\langle X, d \rangle$  are the same as those in  $\langle X, d^* \rangle$ . Hence it remains only to show that the Cauchy sequences in  $\langle X, d \rangle$  are the same as those in  $\langle X, d^* \rangle$ . But this follows from the fact that for arbitrary  $\varepsilon$  with  $0 < \varepsilon < 1$ ,

$$d(x, y) < \varepsilon \iff d^*(x, y) < \varepsilon. \quad \square$$

From the complete metric spaces we have already considered, the first part of the next theorem supplies us with many more.

**1.81 Theorem (completeness and closed sets).** *Let  $\langle X, d \rangle$  be a metric space, let  $Y \subset X$ , and let  $d'$  be the metric on  $Y$  induced by  $d$ .*

- (1) *If  $\langle X, d \rangle$  is complete and  $Y$  is  $d$ -closed, then  $\langle Y, d' \rangle$  is also complete.*
- (2) *If  $\langle Y, d' \rangle$  is complete, then  $Y$  is  $d$ -closed.*

**Proof.** (1) Assume that  $\langle X, d \rangle$  is complete and  $Y$  is  $d$ -closed. Let  $\langle y_n \rangle_{n \in \mathbb{N}}$  be a Cauchy sequence in  $\langle Y, d' \rangle$ . Then  $\langle y_n \rangle_{n \in \mathbb{N}}$  is also a Cauchy sequence in  $\langle X, d \rangle$ , and so by assumption it converges to some  $y$  in  $\langle X, d \rangle$ . Since  $Y$  is  $d$ -closed, by [Theorem 1.67](#), the point  $y \in Y$ . Hence  $\langle y_n \rangle_{n \in \mathbb{N}}$  converges to  $y$  in  $\langle Y, d' \rangle$ .

- (2) Assume that  $\langle Y, d' \rangle$  is complete. Let  $\langle y_n \rangle_{n \in \mathbb{N}}$  be a sequence in  $Y$  that converges to a point  $x$  in  $\langle X, d \rangle$ . In view of [Theorem 1.67](#), it suffices to show that  $x \in Y$ . Now  $\langle y_n \rangle_{n \in \mathbb{N}}$ , being a Cauchy sequence in  $\langle X, d \rangle$ , is a Cauchy sequence in  $\langle Y, d' \rangle$ . By assumption, this sequence must converge to some  $y$  in  $\langle Y, d' \rangle$ , and hence in  $\langle X, d \rangle$  as well. Then  $x = y$  by uniqueness of sequential limits ([Theorem 1.65](#)), and so  $x \in Y$ .  $\square$

An important application of [Theorem 1.81](#) will be the proof, in a series of steps in the next subsection, that the sup metric  $d_\infty$  on the set  $C([a, b])$  of continuous functions from a closed interval  $[a, b]$  to  $\mathbb{R}$  is complete. The proof is completed in [Corollary 1.87](#).

### Completeness of sequence and functions spaces

The infinite-dimensional analog of finite-dimensional Euclidean spaces  $\mathbb{R}^k$  is another complete metric space.

**1.82 Theorem (completeness of  $\ell^2$ ).** *The Hilbert sequence space  $\langle \ell^2, d_2 \rangle$  is complete.*

**Proof.** Each point of  $\ell^2$  is itself a sequence of real numbers. Then in order to distinguish a point in  $\ell^2$  from a sequence of points in  $\ell^2$ , we use superscripts to denote the coordinates of a point of  $\ell^2$  (just as we did when dealing with sequences in  $\mathbb{R}^k$ ). The strategy we use is similar to that in the proof of [Proposition 1.77](#). Denote the Euclidean metric on  $\mathbb{R}$  by  $d$ .

Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a Cauchy sequence in  $\langle \ell^2, d_2 \rangle$ . For each  $n \in \mathbb{N}$ , write

$$x_n = \langle x_n^i \rangle_{i=1,2,3,\dots}.$$



Step 1: define a certain sequence  $x$  of real numbers. For all  $n, k$ , and  $i$ , we have

$$|x_n^i - x_k^i| \leq \left[ \sum_{j=1}^{\infty} (x_n^j - x_k^j)^2 \right]^{1/2} = d_2(x_n, x_k).$$

Then for each  $i = 1, 2, 3, \dots$ , the sequence  $\langle x_n^i \rangle_{n \in \mathbb{N}}$  of real numbers is a Cauchy sequence in  $\langle \mathbb{R}, d \rangle$  and therefore converges to some real number  $x^i$  in  $\langle \mathbb{R}, d \rangle$ . Define

$$x = \langle x^i \rangle_{i=1,2,3,\dots}.$$

We shall complete the proof by showing, first, that  $x \in \ell^2$  and, second, that  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x$  in  $\langle \ell^2, d_2 \rangle$ .

Step 2:  $x \in \ell^2$ . It suffices to find an  $m$  such that  $x_m - x \in \ell^2$ , for then

$$x = x_m - (x_m - x)$$

will belong to  $\ell^2$  as well. Corresponding to  $\varepsilon = 1$ , there is an  $m \in \mathbb{N}$  such that

$$\{eq:xn-xm-dtwo\} \quad (*) \quad d_2(x_n, x_k) < 1 \quad (n \geq m, k \geq m).$$

We claim that the sequence

$$x_m - x = \langle x_m^i - x^i \rangle_{i=1,2,3,\dots}$$

does indeed belong to  $\ell^2$ , that is, the series

$$\sum_{i=1}^{\infty} (x_m^i - x^i)^2$$

converges.

To show that this series of nonnegative real numbers converges, it is enough to show that all its partial sums are bounded above, say by 1, that is:

$$\sum_{i=1}^p (x_m^i - x^i)^2 \leq 1 \quad (p = 1, 2, 3, \dots).$$

Fix  $p \geq 1$ . From  $(*)$  we obtain

$$\sum_{i=1}^p (x_m^i - x_n^i)^2 < 1 \quad (n \geq m).$$

Now for each  $i = 1, 2, \dots, p$ , the sequence  $\langle x_n^i \rangle_{n \in \mathbb{N}}$  converges to  $x^i$ , and from the elementary properties (2)–(4) of limits of sequences of real numbers enumerated on [page 193](#), we conclude

$$\sum_{i=1}^p (x_m^i - x^i)^2 \leq 1,$$

as desired.

bounded function  
sup metric

Step 3: the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x$  in  $(\ell^2, d_2)$ . Let  $\varepsilon > 0$  be arbitrary. There exists  $m \in \mathbb{N}$  such that

$$d_2(x_n, x_k) < \varepsilon \quad (n \geq m, k \geq m).$$

Then for each  $p = 1, 2, 3, \dots$ ,

$$\sum_{i=1}^p (x_n^i - x_k^i)^2 < \varepsilon^2 \quad (n \geq m, k \geq m).$$

For each  $i \geq 1$ , the sequence  $\langle x_k^i \rangle_{k \in \mathbb{N}}$  converges to  $x^i$ , and so for each  $p = 1, 2, 3, \dots$ , we have

$$\sum_{i=1}^p (x_n^i - x^i)^2 \leq \varepsilon^2 \quad (n \geq m).$$

Hence

$$d_2(x_n, x) = \left( \sum_{i=1}^{\infty} (x_n^i - x^i)^2 \right)^{1/2} \leq \varepsilon \quad (n \geq m),$$

as needed.  $\square$

As stated earlier, [Theorem 1.81](#) will allow us to prove completeness of the sup metric  $d_\infty$  on the set  $C([a, b])$  of all continuous functions from a closed interval  $[a, b]$  to  $\mathbb{R}$  ([Example 1.8](#)). The larger complete space we need to apply the theorem will consist of all bounded functions from  $[a, b]$  to  $\mathbb{R}$ .

We shall actually consider, more generally, bounded real-valued functions on an arbitrary nonempty set  $X$ . Recall from [Example 1.16](#) that a function  $f: X \rightarrow \mathbb{R}$  is said to be *bounded* when there is some constant  $c$  such that  $|f(x)| \leq c$  for all  $x \in X$ . Recall also that the formula

$$\|f\|_\infty = \sup\{|f(x)| : x \in X\}.$$

defines a norm on the set  $\mathcal{B}(X)$  of all bounded functions from  $X$  to  $\mathbb{R}$ , the *sup norm*, and that then the formula

$$d_\infty(f, g) = \|f - g\|_\infty = \sup\{|f(x) - g(x)| : x \in X\}$$

defines a metric on the set  $\mathcal{B}(X)$ , the *sup metric*.

**1.83 Theorem (completeness of sup metric on bounded functions).** Let  $\mathcal{B}(X)$  be the set of all bounded real-valued functions on a nonempty set  $X$ . Then the sup metric  $d_\infty$  on  $\mathcal{B}(X)$  is complete.

**Proof.** (The strategy is similar to that used to prove [Theorem 1.82](#).)

Let  $\langle f_n \rangle_{n \in \mathbb{N}}$  be a Cauchy sequence in  $(\mathcal{B}(X), d_\infty)$ .

Step 1: obtain a certain function  $f: X \rightarrow \mathbb{R}$ . Given  $\varepsilon > 0$ , there is some  $m \in \mathbb{N}$  such that

$$d_\infty(f_n, f_k) < \varepsilon \quad (n \geq m, k \geq m)$$

and hence

$$|f_n(x) - f_k(x)| < \varepsilon \quad (n \geq m, k \geq m)$$

for each  $x \in X$ . Thus for each  $x \in X$ , the sequence  $\langle f_n(x) \rangle_{n \in \mathbb{N}}$  of real numbers is a Cauchy sequence in  $(\mathbb{R}, d)$ , where  $d$  is the Euclidean metric. Then for each  $x \in X$  the sequence

$\langle f_n(x) \rangle_{n \in \mathbb{N}}$  converges in  $\langle \mathbb{R}, d \rangle$  to some number, which we denote by  $f(x)$ . In this way we obtain a function  $f: X \rightarrow \mathbb{R}$ . It remains to show that  $f \in \mathcal{B}(X)$  and that  $\langle f_n \rangle_{n \in \mathbb{N}}$  converges to  $f$  in  $\langle \mathcal{B}(X), d_\infty \rangle$ .

pointwise convergent sequence of fu  
pointwise convergent sequence of fu

Step 2: the function  $f$  is bounded. Corresponding to  $\varepsilon = 1$ , there is some  $m \in \mathbb{N}$  such that

$$d_\infty(f_n, f_k) < 1 \quad (n \geq m, k \geq m).$$

Since  $f_m \in \mathcal{B}(X)$ , there is some  $c \in \mathbb{R}$  with  $|f_m(x)| \leq c$  for all  $x \in X$ . Fix  $x \in X$ . Choose  $n \geq m$  such that

$$|f(x) - f_n(x)| < 1.$$

Then

$$|f(x)| \leq |f(x) - f_n(x)| + |f_n(x) - f_m(x)| + |f_m(x)| < 2 + c.$$

Since  $x$  was arbitrary and  $c$  did not depend on  $x$ , the function  $f$  is indeed bounded.

Step 3: the sequence  $\langle f_n \rangle_{n \in \mathbb{N}}$  converges to  $f$  in  $(\mathcal{B}(X), d_\infty)$ . Let  $\varepsilon > 0$  be arbitrary. Choose  $m \in \mathbb{N}$  such that

$$\{eq:dmax-fn-fk-eps-over-2\} \quad (*) \quad d_\infty(f_n, f_k) < \frac{\varepsilon}{2} \quad (n \geq m, k \geq m).$$

We claim that

$$d_\infty(f, f_n) \leq \varepsilon \quad (n \geq m).$$

To prove this, we show that

$$|f(x) - f_n(x)| < \varepsilon \quad (n \geq m)$$

for each  $x \in X$ . Fix  $x \in X$ . Since  $\langle f_n(x) \rangle_{n \in \mathbb{N}}$  converges to  $f(x)$ , there exists some  $k \geq m$  such that

$$|f(x) - f_k(x)| < \frac{\varepsilon}{2}.$$

From this and (\*),  $n \geq m$  implies

$$|f(x) - f_n(x)| \leq |f(x) - f_k(x)| + |f_k(x) - f_n(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \quad \square$$

Suppose now that  $\langle f_n \rangle_{n \in \mathbb{N}}$  is a sequence of functions from a nonempty set  $X$  to the real line  $\mathbb{R}$ . Let  $d'$  be the Euclidean metric on  $\mathbb{R}$ . If for each  $x \in X$  the sequence  $\langle f_n \rangle_{n \in \mathbb{N}}$  converges in  $\langle \mathbb{R}, d' \rangle$ , then we obtain a function  $f: X \rightarrow \mathbb{R}$  given by:

$$\{eq:ptwise-conv-seq-fns-to-R-lim-form\} \quad (*) \quad f(x) = \lim_{n \rightarrow \infty} f_n(x) \quad (x \in X).$$

In this case we say that the sequence  $\langle f_n \rangle_{n \in \mathbb{N}}$  of functions **converges pointwise to**  $f$ . When expanded, condition (\*) is:

$$\{eq:ptwise-conv-seq-fns-to-R\} \quad (\text{Pointwise}) \quad \text{For each } x \in X \text{ and each } \varepsilon > 0, \text{ there exists } m \in \mathbb{N} \text{ such that } |f_n(x) - f(x)| < \varepsilon \text{ for all } n \geq m.$$

Suppose, further, that  $f$  and all the  $f_n$  are all bounded. Then it is meaningful to form the sup-distances

$$d_\infty(f_n, f) = \sup_{x \in X} |f_n(x) - f(x)|.$$

uniformly convergent sequence of functions:  
 uniformly convergent sequence of functions  
 pointwise convergent sequence of functions

Notice that the condition

$$\lim_{n \rightarrow \infty} d_{\infty}(f_n, f) = 0$$

—which may or may not hold for the given sequence—is equivalent to the condition:

{eq:uniform-conv-seq-fns-to-R-lim-form} (Uniform)

For every  $\varepsilon > 0$ , there exists an  $m \in \mathbb{N}$  such that  
 $|f_n(x) - f(x)| < \varepsilon$  for all  $n \geq m$  and all  $x \in X$ .

When the preceding condition holds, we say that the sequence  $\langle f_n \rangle_{n \in \mathbb{N}}$  of functions **converges uniformly to  $f$** .

The distinction between pointwise and uniform convergence is crucial. On the one hand, the index  $m$  in the pointwise convergence condition depends on the particular point  $x \in X$  as well as the given  $\varepsilon$ . On the other hand, the index  $m$  in the uniform convergence condition depends only on the given  $\varepsilon$  and not on  $x$ ; as we might say, “the same  $m$  ‘works’ for all  $x \in X$ .”

Evidently, *if a sequence of functions converges uniformly to a function  $f$ , then it converges pointwise to  $f$* . However, the converse fails.

ex:ptwise-not-unif-conv **1.84 Example (non-uniformly convergence of convergent sequence of functions).** For each  $n = 0, 1, 2, \dots$ , define the function  $f_n: [0, 1] \rightarrow \mathbb{R}$  by

$$f_n(x) = \frac{nx}{1 + n^2x^2} \quad (x \in [0, 1]).$$

Then

$$\lim_{n \rightarrow \infty} f_n(x) = 0$$

for each  $x \in [0, 1]$ , and so  $\langle f_n \rangle_{n \in \mathbb{N}}$  converges pointwise to the zero function  $f$  given by

$$f(x) = 0 \quad (x \in [0, 1]).$$

However,  $\langle f_n \rangle_{n \in \mathbb{N}}$  does *not* converge uniformly to  $f$ , because

$$\left| f_n\left(\frac{1}{n}\right) - f\left(\frac{1}{n}\right) \right| = \frac{1}{2}$$

for each  $n = 1, 2, 3, \dots$ . This implies that for, say,  $\varepsilon = 1/4$ , there is no single positive integer  $m$  such that  $|f_n(x) - f(x)| < \varepsilon$  for all  $n \geq m$  and all  $x \in [0, 1]$ .  $\diamond$

unif-limit-bded-cont-fns-is-bded-cont

**1.85 Lemma (continuity of uniform limit of uniformly convergent sequence).** Let  $\langle X, d \rangle$  be a nonempty metric space, let  $d'$  be the Euclidean metric on  $\mathbb{R}$ , let  $f: X \rightarrow \mathbb{R}$  be a function and let  $\langle f_n \rangle_{n \in \mathbb{N}}$  be a sequence of bounded,  $\langle d, d' \rangle$ -continuous functions that converges uniformly to  $f$ . Then  $f$  is bounded and  $\langle d, d' \rangle$ -continuous.

**Proof.** The proof that  $f$  must be bounded does not require continuity of the functions  $f_n$  and is left an exercise [Exercise 113 (b)].

We prove that  $f$  must be  $\langle d, d' \rangle$ -continuous. Fix  $x \in X$ . We show that  $f$  is  $\langle d, d' \rangle$ -continuous at  $x$ . Let  $\varepsilon > 0$  be arbitrary. There exists  $m \in \mathbb{N}$  such that

$$|f_m(t) - f(t)| < \frac{\varepsilon}{3} \quad (t \in X)$$

and, in particular,

$$|f_m(x) - f(x)| < \frac{\varepsilon}{3}.$$

By continuity of  $f_m$  at  $x$ , there is some  $d$ -neighborhood  $U$  of  $x$  such that

$$|f_m(t) - f_m(x)| < \frac{\varepsilon}{3} \quad (t \in U)$$

(see [Figure 1.22](#)). Then  $t \in U$  implies

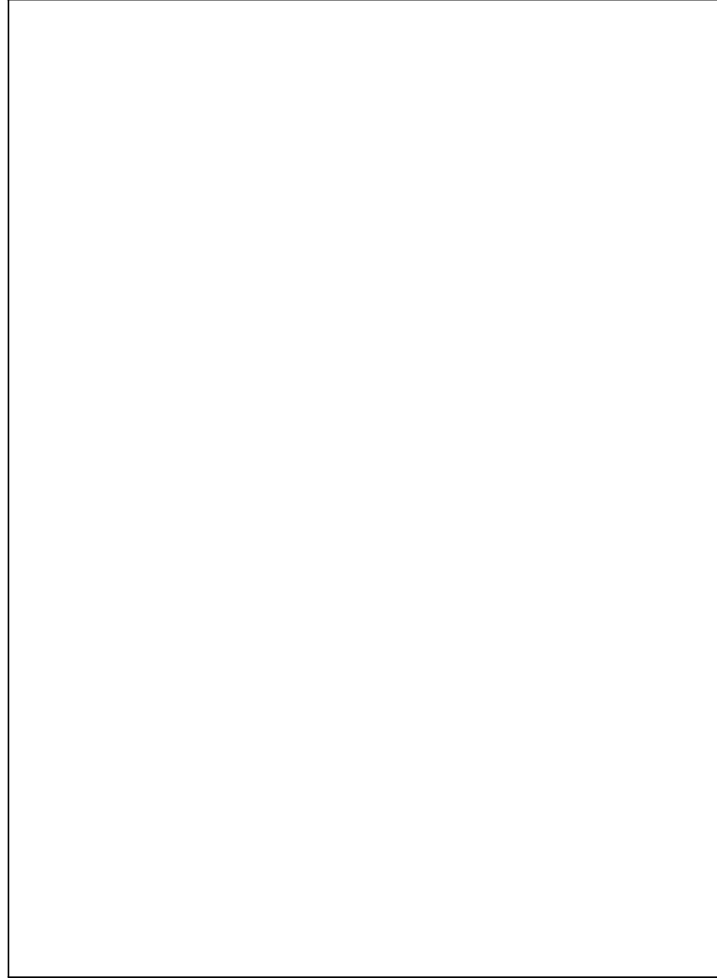


Figure 1.22: Proving continuity of the uniform limit of a sequence of continuous real-valued functions.

fig:pf-unif-lim-cont-is-cont

$$\begin{aligned} |f(t) - f(x)| &\leq |f(t) - f_m(t)| + |f_m(t) - f_m(x)| + |f_m(x) - f(x)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \quad \square \end{aligned}$$

Given a metric space  $\langle X, d \rangle$ , denote by  $C^*(X, d)$ , or simply  $C^*(X)$ , the set of all *bounded*  $\langle d, d' \rangle$ -continuous functions from  $X$  to  $\mathbb{R}$ . The sup metric  $d_\infty$  on the set  $\mathcal{B}(X)$  of all bounded real-valued functions on  $X$  (see [Exercise 16](#)) induces on  $C^*(X)$  a metric that we again denote by  $d_\infty$ .

uniformly converge  
thm:sup-metric-bdd-cont-complete  
Hausdorff, Lem  
Nested Interval Prop

**1.86 Theorem (completeness of sup metric on continuous bounded functions).** Let  $\langle X, d \rangle$  be a metric space, let  $d'$  be the Euclidean metric on  $\mathbb{R}$ , and let  $C^*(X)$  be the set of all bounded  $\langle d, d' \rangle$ -continuous functions from  $X$  to  $\mathbb{R}$ . Then the sup metric on  $C^*(X)$  is complete.

**Proof.** By Theorem 1.81 (1), we need only show that the set  $C^*(X)$  is  $d_\infty$ -closed in  $\mathcal{B}(X)$ , where  $\langle \mathcal{B}(X), d_\infty \rangle$  is the metric space of Theorem 1.83. According to the criterion Theorem 1.67 (ii), it suffices to prove that if a sequence in  $C^*(X)$  converges in  $\langle \mathcal{B}(X), d_\infty \rangle$  to a function  $f \in \mathcal{B}(X)$ , then  $f \in C^*(X)$  as well. But Lemma 1.85 says precisely that.  $\square$

Recall now from calculus two facts:

- a continuous real-valued function on a closed interval  $[a, b]$  is bounded; and
- a continuous real-valued function on a closed interval  $[a, b]$  assumes a maximum value at some point of the interval.

[These facts will be proved in Chapter 4 (Compactness)—specifically, in Corollaries 4.2 and 4.27, respectively.] Then for continuous functions  $f, g: [a, b] \rightarrow \mathbb{R}$ , the distance  $d_\infty(f, g)$  given by the sup metric is actually a maximum:

$$d_\infty(f, g) = \max\{|f(x) - g(x)| : x \in [a, b]\}.$$

cor:cont-fnsl-to-R-complete

**1.87 Corollary.** The set  $C([a, b])$  of all continuous real-valued functions on a closed interval  $[a, b]$  in  $\mathbb{R}$  is complete with respect to the sup metric  $d_\infty$ .

### The Nested Set and Baire Category Theorems

Knowing now so many complete metric spaces, we deduce two of the most significant consequences of completeness. The first says in a new way that points which “ought to be there” really are there. First proved for arbitrary complete metric spaces by Hausdorff, this consequence of completeness generalizes the Nested Interval Property (Theorem 0.83).

thm:nested-set-thm

**1.88 Nested Set Theorem.** Let  $\langle E_n \rangle_{n \in \mathbb{N}}$  be a decreasing sequence of nonempty,  $d$ -closed,  $d$ -bounded subsets of a complete metric space  $\langle X, d \rangle$  such that

$$\lim_{n \rightarrow \infty} \text{diam } E_n = 0.$$

Then there is a unique point

$$x \in \bigcap_{n=0}^{\infty} E_n.$$

**Proof.** Uniqueness. Suppose  $x, y \in \bigcap_{n=0}^{\infty} E_n$  with  $x \neq y$ . Set

$$\varepsilon = d(x, y),$$

so that  $\varepsilon > 0$ . Choose  $n \in \mathbb{N}$  with  $\text{diam } E_n < \varepsilon$ . Since  $x \in E_n$  and  $y \in E_n$ , we have  $d(x, y) < \varepsilon$ , which gives a contradiction.

Existence. Arbitrarily choose a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  with

$$x_n \in E_n \quad (n \in \mathbb{N}).$$

We claim that  $\langle x_n \rangle_{n \in \mathbb{N}}$  is a Cauchy sequence in  $\langle X, d \rangle$ . In fact, let  $\varepsilon > 0$  be arbitrary. Choose some  $m \in \mathbb{N}$  such that

$$\text{diam } E_m < \varepsilon.$$

Then  $n \geq m$  and  $k \geq m$  implies  $x_n \in E_n \subset E_m$  and  $x_k \in E_k \subset E_m$ , so that

$$d(x_n, x_k) < \varepsilon.$$

Since  $\langle X, d \rangle$  is complete, the Cauchy sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to some point  $x \in X$ .

We conclude by showing that  $x \in E_n$  for each  $n \in \mathbb{N}$ . Fix  $n \in \mathbb{N}$ . The sequence  $\langle x_{n+j} \rangle_{j \in \mathbb{N}}$  also converges to  $x$  in  $\langle X, d \rangle$ , and each of its values  $x_{n+j}$  belongs to  $E_n$ . Because  $E_n$  is  $d$ -closed, it follows from [Theorem 1.67](#) that  $x \in E_n$ , too.  $\square$

For closed intervals in the real line and certain other complete metric spaces, the hypotheses on the sets  $E_n$  can be considerably relaxed. In a “compact” metric space  $\langle X, d \rangle$ , any sequence of  $d$ -closed sets will have a nonempty intersection provided that each finite collection of these sets has nonempty intersection (see [Theorem 4.9](#)).

The second significant consequence of completeness concerns “dense” sets.

def:d-dense

**1.89 Definition ( $d$ -dense set).** A subset  $D$  of a metric space  $\langle X, d \rangle$  is said to be  **$d$ -dense in  $X$**  when it intersects every nonempty  $d$ -open set. Equivalently,  $D$  is  $d$ -dense in  $X$  when each  $d$ -neighborhood of each point of  $X$  intersects  $D$ .

**Intuitive idea— $d$ -dense set.** Roughly speaking, then, a set  $D$  is  $d$ -dense in  $X$  when it contains points arbitrarily  $d$ -close to any given point of  $X$ .

examples:d-dense-sets

**1.90 Examples.** (1) The set of rational numbers is  $d$ -dense in  $\mathbb{R}$ , where  $d$  is the Euclidean metric. This is an immediate consequence of the order-density of  $\mathbb{Q}$  in  $\mathbb{R}$  ([Corollary 0.81](#)).

(2) The complement in  $\mathbb{R}$  of any finite set is a  $d$ -open set that is  $d$ -dense in  $\mathbb{R}$ .  $\diamond$

thm:Baire-category-thm

**1.91 Baire Category Theorem.** In a complete metric space  $\langle X, d \rangle$ , the intersection of any sequence of  $d$ -open subsets of  $X$  that are  $d$ -dense in  $X$  is itself  $d$ -dense in  $X$  (and hence nonempty if  $X$  is nonempty).

**Proof.** Let

$$A = \bigcap_{n=0}^{\infty} A_n$$

where each  $A_n$  is both  $d$ -open and  $d$ -dense in  $X$ . Let  $U$  be an arbitrary nonempty  $d$ -open subset of  $X$ . To show that  $A$  intersects  $U$ , we are going to construct a sequence  $\langle E_n \rangle_{n \in \mathbb{N}}$  of sets satisfying the hypotheses of the Nested Set Theorem ([1.88](#)) such that

$$E_n \subset A_n \cap U \quad (n \in \mathbb{N}).$$

Then the point belonging to  $\bigcap_{n=0}^{\infty} E_n$  will belong to  $A \cap U$ .

Baire, Ren\`e-Louis



Figure 1.23: Start of the construction for proving the Baire Category Theorem.

fig:Baire-pf

Since  $A_0$  is  $d$ -dense in  $X$ , there is some point

$$x_0 \in A_0 \cap U,$$

and then since  $A_0 \cap U$  is  $d$ -open, we may choose some number  $\varepsilon_0$  with

$$0 < \varepsilon_0 < \frac{1}{2}, \quad D_{\varepsilon_0}(x_0; d) \subset A_0 \cap U$$

(see Figure 1.23).

Nest, since  $A_1$  is  $d$ -dense in  $X$ , there is some point

$$x_1 \in A_1 \cap B_{\varepsilon_0}(x_0; d),$$

and then we may choose some  $\varepsilon_1$  with

$$0 < \varepsilon_1 < \frac{1}{2^2}, \quad D_{\varepsilon_1}(x_1; d) \subset A_1 \cap B_{\varepsilon_0}(x_0; d).$$

Continuing in this way, we obtain a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  of points together with a sequence  $\langle \varepsilon_n \rangle_{n \in \mathbb{N}}$  of positive numbers such that

$$D_{\varepsilon_{n+1}}(x_{n+1}; d) \subset A_{n+1} \cap B_{\varepsilon_n}(x_n; d),$$

$$2\varepsilon_n < \frac{1}{2^n}$$

for each  $n$ .

The decreasing sequence  $\langle E_n \rangle_{n \in \mathbb{N}}$  of sets to which we apply the Nested Set Theorem (1.88) is given by

$$E_n = D_{\varepsilon_n}(x_n; d) \quad (n \in \mathbb{N}).$$

The requirement that  $\lim_{n \rightarrow \infty} \text{diam } E_n = 0$  is satisfied because  $\text{diam } E_n \leq 2\varepsilon_n < 1/2^n$  for each  $n$ .  $\square$

This theorem just proved is named in honor of René-Louis Baire, who proved it in 1899; the term ‘category’ appears for reasons explained on page 257. Observe that the hypothesis concerns completeness, a property of the particular metric  $d$  used, whereas the conclusion



concerns only the  $d$ -open sets. (Essentially the same “topological” conclusion can be deduced from a completely different hypothesis not involving metrics—see [Exercise 4.128](#).)

piecewise linear function  
Banach, Stefan

A  $d$ -dense set in a metric space contains “many” points of  $X$ . For this reason, the Baire Category Theorem provides a method, and a powerful one at that, for proving that many points in a complete metric space have a particular property. Unfortunately, most significant applications of this method are anything but elementary and would take us far afield of the subject of this book. One application, to topology, is made in [Examples 5.38 \(5\)](#). Others may be found in Boas, Jr. and Boas [7] and Goldberg [29].

We shall rest content with presenting here just one striking application of the Baire Category Theorem: the rarity of continuous real-valued functions of a real variable that have derivatives. The argument is complicated and involves more analysis than topology, so you may wish to postpone studying it. And it uses a result not proved until later ([Application 4.66](#)), namely:

Given any continuous function  $f: [0, 1] \rightarrow \mathbb{R}$  and any  $\varepsilon > 0$ , there is a continuous *piecewise linear* function  $p: [0, 1] \rightarrow \mathbb{R}$  such that  $|f(x) - p(x)| < \varepsilon$  for all  $x \in [0, 1]$ .

By a continuous function  $p: [0, 1] \rightarrow \mathbb{R}$  that is **piecewise linear** we mean one whose graph consists of finitely many line segments.

app:non-diff-fns

**1.92 Application (uniform approximation of continuous function by nowhere differentiable function).** In

calculus you doubtless spent a considerable time learning how to compute derivatives. Occasionally, though, you encountered continuous functions whose graphs have “corners” and which therefore fail to be differentiable at a few points: any continuous piecewise linear, but nonlinear, function—such as the absolute-value function—is of this type. We are going to establish the surprising fact that among all continuous functions, those having a derivative at even a single point are rare! The argument is due to Banach.

Provide the set  $C([0, 1])$  of all continuous functions  $f: [0, 1] \rightarrow \mathbb{R}$  with its sup metric  $d_\infty$  ([Example 1.8](#)). Let

$$\mathcal{A} = \{f: [0, 1] \rightarrow \mathbb{R} \mid f \text{ has a derivative at no point in } [0, 1[ \}.$$

(Conceivably the set  $\mathcal{A}$  could be empty, but the argument will show, in particular, that it is not.) We shall show that *the set  $\mathcal{A}$  is  $d_\infty$ -dense in  $C([0, 1])$* , in other words:

*each continuous function  $f: [0, 1] \rightarrow \mathbb{R}$  can be approximated, uniformly over  $[0, 1]$ , to within any prescribed error  $\varepsilon$ , by a continuous function  $g: [0, 1] \rightarrow \mathbb{R}$  that is nowhere differentiable on the interval  $[0, 1[$ .*

If  $f \in C([0, 1])$  does have a derivative at a particular point  $x \in [0, 1[$ , then the right-hand difference quotients

$$\frac{f(x+h) - f(x)}{h} \quad (0 < h < 1 - x)$$

are bounded; hence for a sufficiently large positive integer  $n$ , the function  $f$  belongs to the complement of the set

$$\mathcal{A}_n = \{f \in C([0, 1]) : \text{for each } x \in [0, 1 - 1/n] \text{ there exists } h \text{ with } 0 < h < 1 - x \text{ and } |f(x+h) - f(x)| > nh\}.$$

Thus

$$\mathcal{A} \supset \bigcap_{n=1}^{\infty} \mathcal{A}_n.$$

To show that  $\mathcal{A}$  is  $d_\infty$ -dense in  $C([0, 1])$ , it suffices to show that the smaller set  $\bigcap_{n=1}^{\infty} \mathcal{A}_n$  is  $d_\infty$ -dense. Now  $d_\infty$  is a complete metric ([Corollary 1.87](#)). According to the Baire Category

Theorem (1.91), then, it suffices to show that each  $\mathcal{A}_n$  is both  $d_\infty$ -open and  $d_\infty$ -dense in  $C([0, 1])$ .

For the remainder of the argument, the positive integer  $n$  is fixed.

Proof that  $\mathcal{A}_n$  is  $d_\infty$ -open in  $C([0, 1])$ . We shall show that the complement of  $\mathcal{A}_n$  is  $d_\infty$ -closed, by using criterion Theorem 1.67. Let  $\langle f_i \rangle_{i \in \mathbb{N}}$  be a sequence in  $C([0, 1]) \setminus \mathcal{A}_n$  that converges to  $f \in C([0, 1])$  with respect to  $d_\infty$ . We must show that  $f \notin \mathcal{A}_n$ .

For each  $i \in \mathbb{N}$  there is some point  $x_i$  with

$$0 \leq x_i \leq 1 - \frac{1}{n},$$

$$|f_i(x_i + h) - f_i(x_i)| \leq nh \quad (0 < h < 1 - x_i).$$

A bisection argument produces a point  $x \in [0, 1 - 1/n]$  such that each  $d$ -neighborhood of  $x$  contains  $x_i$  for infinitely many values of  $i$ , where  $d$  is the Euclidean metric on  $[0, 1]$ .

We are about to show that

$$|f(x + h) - f(x)| < nh + \varepsilon \quad (\varepsilon > 0, 0 < h < 1 - x).$$

It will then follow that

$$|f(x + h) - f(x)| \leq nh \quad (0 < h < 1 - x),$$

which will mean that  $f \notin \mathcal{A}_n$ .

Let  $\varepsilon > 0$  be arbitrary, and let  $0 < h < 1 - x$ . By continuity of  $f$  at  $x$  and at  $x + h$ , there are  $d$ -neighborhoods  $U$  of  $x$  and  $V$  of  $x + h$  such that

$$u \in U \implies |f(u) - f(x)| < \frac{\varepsilon}{4}, \quad v \in V \implies |f(x + h) - f(v)| < \frac{\varepsilon}{4}.$$

Next, there is an  $i \in \mathbb{N}$  such that

$$x_i \in U, \quad x_i + h \in V, \quad d_\infty(f, f_i) < \frac{\varepsilon}{4}.$$

Then

$$\begin{aligned} |f(x + h) - f(x)| &\leq |f(x + h) - f(x_i + h)| + |f(x_i + h) - f_i(x_i + h)| \\ &\quad + |f_i(x_i + h) - f_i(x_i)| + |f_i(x_i) - f(x_i)| \\ &\quad + |f(x_i) - f(x)| \\ &< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + nh + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = nh + \varepsilon. \end{aligned}$$

Proof that  $\mathcal{A}_n$  is  $d_\infty$ -dense in  $C([0, 1])$ . Let  $f \in C([0, 1])$  and let  $\varepsilon > 0$ . By a comment above, there is a piecewise linear function  $p \in C([0, 1])$  with  $d_\infty(f, p) < \varepsilon/2$ . Hence we need only find some  $g \in \mathcal{A}_n$  with  $d_\infty(p, g) \leq \varepsilon/2$ .

The finitely many line segments comprising the graph of  $p$  have slopes  $m_1, m_2, \dots, m_k$ . Choose a number  $m$  with

$$m > n + \max_{1 \leq i \leq k} |m_i|,$$

and let  $s$  be the “sawtooth” piecewise linear function whose line segments have slopes  $\pm m$  and for which

$$0 \leq s(x) \leq \frac{\varepsilon}{2} \quad (0 \leq x \leq 1)$$

(see Figure 1.24). Set

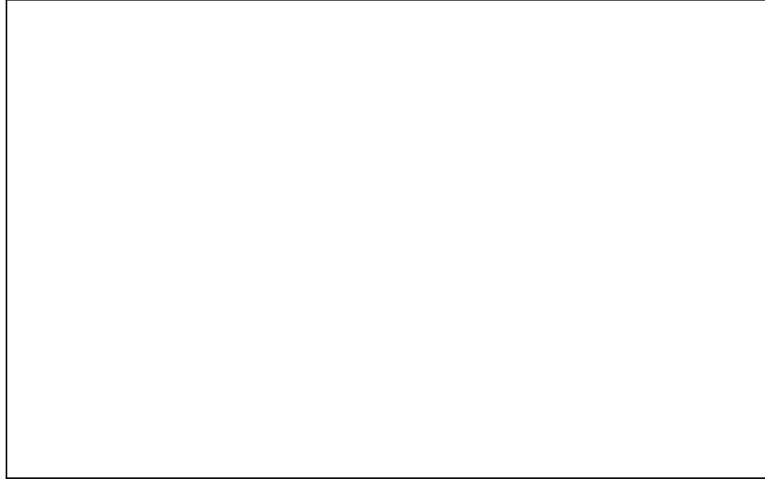


Figure 1.24: Constructing a sawtooth function from a piecewise linear function.

Baire Category Theorem  
dense set!metric space@in metric sp  
metric space!dense subset@and den  
complete metric space

cluster point

clustering sequence

fig:sawtooth-from-piecewise-linear

$$g = p + s.$$

Then  $g \in C([0, 1])$  with

$$d_\infty(p, g) = \max_{0 \leq x \leq 1} s(x) = \frac{\varepsilon}{2}.$$

We conclude the proof by showing that  $g \in \mathcal{A}_n$ . Let  $0 \leq x < 1 - 1/n$ . Choose  $h$  with  $0 < h < 1 - x$  so small that the points  $\langle x, p(x) \rangle$  and  $\langle x + h, p(x + h) \rangle$  both lie on the same line segment of the graph of  $p$ , of slope  $m_i$ , and at the same time the points  $\langle x, s(x) \rangle$  and  $\langle x + h, s(x + h) \rangle$  both lie on the same line segment of the graph of  $s$ . Then

$$\begin{aligned} |p(x + h) - p(x)| &= |m_i|h, \\ |s(x + h) - s(x)| &= mh > nh + |m_i|h. \end{aligned}$$

Hence

$$|g(x + h) - g(x)| > nh.$$

This means that  $g \in \mathcal{A}_n$ .  $\diamond$

It should be noted that the Baire Category Theorem is not really needed in order to demonstrate the existence of everywhere continuous, nowhere differentiable functions. Such functions can be exhibited explicitly—see, for example, Goldberg [29, Section 9.7], McCarthy [47], or Hildebrandt [35].

### EXERCISES FOR SECTION 1.5

**99.** What can be said about a Cauchy sequence in an arbitrary metric space if its range is finite?

**100.** Let  $\langle X, d \rangle$  be a metric space. A sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in  $X$  is said to **cluster at** a point  $x$  in  $\langle X, d \rangle$ , and  $x$  is called a **cluster point of**  $\langle x_n \rangle_{n \in \mathbb{N}}$  in  $\langle X, d \rangle$ , when  $\langle x_n \rangle_{n \in \mathbb{N}}$  is frequently in each  $d$ -ball at  $x$ .

Prove or disprove:

prob:seq-cluster-in-metric-space

- uniformly continuous map (a) A Cauchy sequence in  $\langle X, d \rangle$  cannot cluster at two different points.
- uniformly continuous map (b) If a Cauchy sequence clusters at a point  $x$  in  $\langle X, d \rangle$  and converges to a point  $y$  in  $\langle X, d \rangle$ , then  $x = y$ .
- continuous map!uniformly (c) If a Cauchy sequence clusters at a point in  $\langle X, d \rangle$ , then it converges there.
- map!uniformly continuous (d) If a Cauchy sequence converges to a point in  $\langle X, d \rangle$ , then it clusters there.
- extension!uniformly continuous map@of uniformly continuous map (e) If a Cauchy sequence clusters at a point in  $\langle X, d \rangle$  and if  $\langle X, d \rangle$  is complete, then it converges to that point.
- uniformly continuous map (f) If a Cauchy sequence in a  $d$ -closed subset of  $\langle X, d \rangle$  clusters at some point of  $X$ , then the point belongs to that subset.
- extended realline

prob:unif-cont **101.** Given metric spaces  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$ , a map  $f: X \rightarrow Y$  is said to be  $(d, d')$ -**uniformly continuous** when for each  $\varepsilon > 0$  there exists some  $\delta > 0$  such that for all  $x, t \in X$ ,

$$d(x, t) < \delta \implies d'(f(x), f(t)) < \varepsilon.$$

For example, any isometry from  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$  is  $\langle d, d' \rangle$ -uniformly continuous.

prob-part:unif-cont-implies-cont

- (a) Prove: If  $f$  is  $\langle d, d' \rangle$ -uniformly continuous, then it is  $(d, d')$  continuous.
- (b) By considering the function  $x \mapsto x^2$  from  $\mathbb{R}$  to  $\mathbb{R}$ , or the function  $x \mapsto 1/x$  from  $\{x \in \mathbb{R} : x > 0\}$ , show that the converse of (a) fails.
- Note:* The converse of (a) does hold for certain spaces among which are all closed bounded intervals in  $\mathbb{R}$ —see [Theorem 4.65](#).
- (c) Let  $d$  be a metric on a set  $X$ , let  $d_\infty$  be the corresponding max-metric on  $X \times X$ , and let  $d'$  be the Euclidean metric on  $\mathbb{R}$ . Show that  $d: X \times X \rightarrow \mathbb{R}$  is  $\langle d_\infty, d' \rangle$ -uniformly continuous. *Note:* Compare [Examples 1.52 \(8\)](#).
- (d) Give an analog of [Examples 1.52 \(5\)](#) for uniform continuity.
- (e) Which of the maps from  $X$  to  $\mathbb{R}$  considered in Exercises [76](#) and [77](#) are  $\langle d, d' \rangle$ -uniformly continuous?

prob:unif-cont-bis **102.** (Continuation of [Exercise 101](#).)

Suppose  $f$  is  $\langle d, d' \rangle$ -uniformly continuous.

- (a) If  $\langle x_n \rangle_{n \in \mathbb{N}}$  is a Cauchy sequence in  $\langle X, d \rangle$ , show that then  $\langle f(x_n) \rangle_{n \in \mathbb{N}}$  is a Cauchy sequence in  $\langle Y, d' \rangle$ . Is the same true if  $f$  is only  $\langle d, d' \rangle$ -continuous?
- (b) If  $D$  and  $D'$  are metrics on  $X$  and  $Y$  that are equivalent to  $d$  and  $d'$ , respectively, must  $f$  also be  $\langle D, D' \rangle$ -uniformly continuous?

ob:extend-unif-cont-from-dense-set **103.** (Continuation of [Exercise 102](#).) Let  $f: D \rightarrow Y$  be a  $\langle d, d' \rangle$ -uniformly continuous function from a  $d$ -dense subset  $D$  of a metric space  $\langle X, d \rangle$  to a *complete* metric space  $\langle Y, d' \rangle$ . Prove:

- (a) The function  $f$  has a unique  $\langle d, d' \rangle$ -continuous extension  $F: X \rightarrow Y$  to  $X$ .
- (b) The extension  $F$  is also  $\langle d, d' \rangle$ -uniformly continuous.

ces-between-cauchy-seq-converges **104.** Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  and  $\langle y_n \rangle_{n \in \mathbb{N}}$  be Cauchy sequences in a metric space  $\langle X, d \rangle$ . Prove that  $\langle d(x_n, y_n) \rangle_{n \in \mathbb{N}}$  converges in  $(\mathbb{R}, d')$ , where  $d'$  is the Euclidean metric.

**105.** If  $d$  is a complete metric on  $X$ , is the bounded metric  $d'$  on  $X$  given by  $d'(x, y) = d(x, y)/(1 + d(x, y))$  necessarily complete? (See [Exercise 48](#).)

**106.** Is the metric  $\widehat{d}$  on the extended real line  $\widehat{\mathbb{R}}$  ([Example 1.41](#)) complete?

- prob:p-adic-metric-complete-Q **107.** For a prime number  $p$ , is the  $p$ -adic metric  $d_p$  on the rationals  $\mathbb{Q}$  complete?
- 108.** Construct a metric space  $\langle X, d \rangle$  in which every subset is  $d$ -open, yet  $\langle X, d \rangle$  is not complete (and hence  $d$  is not the discrete metric).
- prob:complete-metric-on-finite-product **109.** In the notation of [Exercise 51 \(a\)](#), suppose the metrics  $d_1, d_2, d_3, \dots$  on  $X_1, X_2, X_3, \dots$ , respectively, are all complete. Prove that then the metric  $d'$  on the product set  $X = \times_{i=1}^{\infty} X_i$  is complete.
- 110.** Given a nonempty set  $X$  and a complete, bounded metric  $d$  on a set  $Y$ , show that the formula
$$d_{\infty}(f, g) = \sup\{d(f(x), g(x)) : x \in X\}$$
defines a complete metric on the set of all maps from  $X$  to  $Y$ .
- 111.** Let  $V$  be a  $d$ -open set in a complete metric space  $\langle X, d \rangle$  with  $V \neq X$ . Let  $d'$  be the metric induced on  $V$  by  $d$ . Prove that there is a complete metric  $d^*$  on  $V$  that is equivalent to  $d'$ .  
You may want to carry out the following steps.
- (a) The function  $f: V \rightarrow \mathbb{R}$  defined by  $f(x) = 1/d(x, X \setminus V)$  is  $\langle d', d_1 \rangle$ -continuous, where  $d_1$  is the Euclidean metric.
  - (b) The formula  $d^*(x, y) = d(x, y) + |f(x) - f(y)|$  defines a metric on  $V$ .
  - (c) The metric  $d^*$  is equivalent to  $d'$ .
  - (d) The metric space  $\langle V, d^* \rangle$  is complete.
- prob:ptwise-vs-unif-conv **112.** Determine whether the sequence  $\langle f_n \rangle_{n \in \mathbb{N}}$  of real-valued functions on  $[0, 1]$  is pointwise convergent and, if so, whether it is also uniformly convergent.
- (a)  $f_n(x) = n/(1 + n)$ .
  - (b)  $f_n(x) = x^n$ .
  - (c)  $f_n(x) = n/(1 + nx)$ .
- prob:unif-vs-ptwise-lim-bded-fns **113.** Let  $\langle f_n \rangle_{n \in \mathbb{N}}$  be a sequence of *bounded* real-valued functions on a nonempty set  $X$  that converges pointwise to a function  $f: X \rightarrow \mathbb{R}$ .
- (a) Show by example that the pointwise-limit function  $f$  need not be bounded.
  - (b) Prove, however, that if  $\langle f_n \rangle_{n \in \mathbb{N}}$  converges uniformly to  $f$ , then  $f$  must also be bounded.
- prob-part:unif-lim-bded-fns-is-bded **114.** Show by example that the Nested Set Theorem ([1.88](#)) no longer remains true if any of the hypotheses that  $\langle X, d \rangle$  be complete, that each  $E_n$  be  $d$ -closed, or that  $\lim_{n \rightarrow \infty} \text{diam } E_n = 0$  is omitted.
- 115.** Does the Nested Set Theorem ([1.88](#)) remain true if the hypothesis that  $\langle E_n \rangle_{n \in \mathbb{N}}$  be decreasing is replaced by the weaker hypothesis that each finite subcollection of  $\{E_n : n \in \mathbb{N}\}$  have nonempty intersection?
- 116.** Generalize the Nested Interval Property ([Theorem 0.83](#)) to  $\mathbb{R}^k$ . *Note:* The analog of a closed interval  $[a, b]$  is, of course, a cube (the product of  $k$  closed intervals) in  $\mathbb{R}^k$ , so you should consider a decreasing sequence of such cubes. No assumption is to be made, however, that the diameters of these cubes tend to zero, and so the Nested Set Theorem ([1.88](#)) cannot be applied directly to the cubes.

**117.** Prove the following converse to the Nested Set Theorem (1.88): A metric space  $\langle X, d \rangle$  is complete if each decreasing sequence  $\langle E_n \rangle_{n \in \mathbb{N}}$  of nonempty,  $d$ -closed,  $d$ -bounded sets with  $\lim_{n \rightarrow \infty} \text{diam } E_n = 0$  has nonempty intersection.

**Banach, Stefan**  
**fixed-point**

**118. (a)** Let  $\langle E_n \rangle_{n \in \mathbb{N}}$  be a sequence of  $d$ -closed subsets of a complete metric space  $\langle X, d \rangle$  such that no  $E_n$  contains any nonempty  $d$ -open set. Show that  $\bigcup_{n=0}^{\infty} E_n$  contains no nonempty  $d$ -open set.

**(b)** Apply (a) so as to show that the plane is not the union of countably many lines.

**119.** The Euclidean metric  $d$  on the open interval  $]0, 1[$  is not complete. Show that, nonetheless, the intersection of any sequence of  $d$ -open,  $d$ -dense sets in  $]0, 1[$  is itself  $d$ -dense in  $]0, 1[$ .

**120.** Does there exist any complete metric on  $\mathbb{Q}$  that is equivalent to the Euclidean metric?

**121.** In the proof of Baire Category Theorem (1.91), make precise the informal use of the recursive definitions of the families  $\langle x_n \rangle_{n \in \mathbb{N}}$  and  $\langle \varepsilon_n \rangle_{n \in \mathbb{N}}$ —including the implicit use of the Axiom of Choice.

**122.** The argument in Application 1.92 actually shows that the set of functions  $f \in C([0, 1])$  for which the right-hand derivative

$$\lim_{h \rightarrow 0^+} \frac{f(x+h) - f(x)}{h}$$

exists at no  $x \in [0, 1[$  is  $d_{\infty}$ -dense in  $C([0, 1])$ . Prove the analogous result concerning left-hand derivatives.

**123.** Given a metric space  $\langle X, d \rangle$ , a map  $T: X \rightarrow X$  is called a **contraction** if there is some constant  $c < 1$  such that

$$d(T(x), T(y)) \leq c d(x, y) \quad (x, y \in X).$$

Verify that  $T$  is a contraction when:

**(a)**  $\langle X, d \rangle = \langle \ell^2, d_2 \rangle$ , the Hilbert sequence space (Example 1.10) and  $T$  is defined by  $T(x) = \langle x_i/2 \rangle_{i=1,2,3,\dots}$  for each  $x = \langle x_i \rangle_{i=1,2,3,\dots} \in \ell^2$ .

**(b)**  $X = [0, 1]$  with its Euclidean metric  $d$ , and  $T: X \rightarrow X$  is a continuous functions that is differentiable on  $]0, 1[$  such that  $|T'(x)| \leq c$  for some  $c < 1$  and all  $x \in ]0, 1[$ . (*Hint:* Use the Mean-Value Theorem from calculus.)

**124.** (*Continuation of Exercise 123.*) Prove Banach's **Contraction Mapping Principle**:

A contraction  $T$  of a complete metric space  $\langle X, d \rangle$  has a unique fixed-point  $x \in X$  to which, for each  $y \in X$ , the sequence  $\langle T^n(y) : n \in \mathbb{N} \rangle$  converges in  $\langle X, d \rangle$ .

Recall that to say  $x$  is a *fixed-point* of  $T$  (see page 21) means that  $T(x) = x$ . And here  $T^n$  denotes the  $n$ th iterate of  $T$ , so that  $T^1 = T$ ,  $T^2 = T \circ T$ ,  $T^3 = T \circ T^2$ , etc.).

*Note:* Some applications of this theorem are given in the following two exercises. Another application is the proof of the Inverse Function Theorem, which says that a continuously differentiable function from  $\mathbb{R}^n$  to  $\mathbb{R}^n$  with a nonsingular Jacobian has at each point a “local inverse”—see, for example, Rudin [57, pages 221–223]. Some other fixed-point theorems are considered in Chapter 5 (Connectedness).

**125.** (Continuation of [Exercise 124](#).)

Use Banach's Contraction Mapping Principle to show that  $\cos x = x$  for some  $x \in [0, 1]$  and to approximate such  $x$  to 3 decimal places.

Picard Existence Theorem

Picard, \Emile

Lipschitz condition

Lipschitz, Rudolf

Contraction Mapping Principle

contraction of metric space

pseudometric

pseudometric!completeness@and co

prob:Picard-existence-thm-ODEs

**126.** (Continuation of [Exercise 124](#).)

Prove the **Picard Existence Theorem** for differential equations: Let  $f$  be a real-valued continuous function on a closed rectangle

$$D = [x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b]$$

in the plane that satisfies a "Lipschitz condition"

$$|f(x, y) - f(x, z)| \leq K |y - z| \quad (\langle x, y \rangle, \langle x, z \rangle \in D,)$$

on  $D$  for some constant  $K > 0$ . Then on some interval  $[x_0 - \delta, x_0 + \delta]$  containing  $x_0$  there exists a solution  $g$  of the differential equation

$$\{\text{eq:ODE}\} \quad (\text{ODE}) \quad g'(x) = f(x, g(x))$$

satisfying the initial condition

$$\{\text{eq:IC}\} \quad (\text{IC}) \quad g(x_0) = y_0.$$

[*Suggestion:* Show first that an arbitrary continuous function  $g: [x_0 - \delta, x_0 + \delta] \rightarrow \mathbb{R}$  satisfies both (ODE) and (IC) if and only if

$$g(x) = y_0 + \int_{x_0}^x f(t, g(t)) dt \quad (|x - x_0| < \delta).$$

Next, choose a constant  $M > 0$  with

$$|f(x, y)| \leq M \quad ((x, y) \in D)$$

and then a  $\delta > 0$  such that

$$K\delta < 1, \quad [x_0 - \delta, x_0 + \delta] \times [y_0 - M\delta, y_0 + M\delta] \subset D.$$

Define

$$\mathcal{E} = \{g \in C([x_0 - \delta, x_0 + \delta]) : |g(x) - y_0| \leq M\delta \text{ for } |x - x_0| < \delta\}.$$

Show that  $\mathcal{E}$  is  $d_\infty$ -closed in  $C([x_0 - \delta, x_0 + \delta])$  and hence that the sup metric on  $\mathcal{E}$  is complete. Finally, show that the formula

$$(T(g))(x) = y_0 + \int_{x_0}^x f(t, g(t)) dt$$

defines a map  $T: \mathcal{E} \rightarrow \mathcal{E}$  that is a contraction.]

(For more about differential equations and the Picard Existence Theorem, see, for example, Boyce, DiPrima, and Meade [[11](#), Section 2.8].)

prob:pseudometric-Cauchy-complete

**127.** Let  $d$  be a pseudometric on a set  $X$  ([Exercise 20](#)). Cauchy sequences and completeness for  $\langle X, d \rangle$  may be defined exactly for the case that  $d$  is a metric. Show that  $\langle X, d \rangle$  is complete if and only if the associated metric space  $\langle X^*, d^* \rangle$  constructed in [Exercise 21](#) is complete.

prob:complete-if-dense-set-etc

**128.** Let  $d$  be a metric (or even a pseudometric) on a set  $X$ . Suppose there is a  $d$ -dense set  $Z$  in  $X$  such that each Cauchy sequence in  $\langle X, d \rangle$  consisting of points of  $Z$  converges in  $\langle X, d \rangle$ . Prove that  $\langle X, d \rangle$  is complete.

y:seq-distances-conv-pseudometric completion of a metric space **129.** Given a pseudometric  $d$  on  $X$  and let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a Cauchy sequence in  $\langle X, d \rangle$  (see [Exercise 127](#)).

rt:seq-distances-conv-pseudometric completion of a metric space **(a)** Show that the sequence  $\langle d(y, x_k) \rangle_{k \in \mathbb{N}}$  of real numbers converges for each  $y \in X$ .  
**(b)** Show that  $\lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} d(x_n, x_k) = 0$ . *Note:* In view of **(a)**, for each  $n \in \mathbb{N}$  the limit  $\lim_{k \rightarrow \infty} d(x_n, x_k)$  exists.

Hausdorff, Felix  
Cantor, Georg

prob:completion **130.** Let  $\langle X, d \rangle$  be any metric space. A **completion** of  $\langle X, d \rangle$  is an isometric embedding  $f$  of  $\langle X, d \rangle$  into a complete metric space  $\langle X', d' \rangle$  in which the image  $f(X)$  is  $d'$ -dense. Carry out the following steps in Hausdorff's construction of a completion of  $\langle X, d \rangle$ , based upon Cantor's construction of the real numbers from the rational numbers.

**(a)** Let  $Y$  be the set whose points are the Cauchy sequences in  $\langle X, d \rangle$ . [Exercise 104](#) justifies forming

$$D(x, y) = \lim_{n \rightarrow \infty} d(x_n, y_n)$$

for any “points”  $x = \langle x_n \rangle_{n \in \mathbb{N}}$  and  $y = \langle y_n \rangle_{n \in \mathbb{N}}$  of  $Y$ . Verify that  $D$  is a pseudometric on  $Y$ .

**(b)** Show that the map  $g: X \rightarrow Y$  defined by the rule that  $g(x) = \langle x_n \rangle_{n \in \mathbb{N}}$  with  $x_n = x$  for all  $n \in \mathbb{N}$  is an isometric embedding of  $\langle X, d \rangle$  into  $\langle Y, D \rangle$ .

**(c)** Show that  $g(X)$  is  $D$ -dense in  $Y$ .

**(d)** Prove that  $\langle Y, D \rangle$  is complete. (*Hint:* Use [Exercise 128](#). You may also want to use [Exercise 129](#).)

**(e)** Let  $\langle Y^*, D^* \rangle$  be the metric space constructed from  $\langle Y, D \rangle$  as in [Exercise 21](#), so that  $\langle Y^*, D^* \rangle$  is complete by [Exercise 127](#). In the notation of [Exercise 21](#), let  $p: Y \rightarrow Y^*$  be the map sending each  $y \in Y$  to its equivalence class  $y^*$  in  $Y$ . Show that  $f = p \circ g: X \rightarrow Y^*$  is an isometric embedding of  $\langle X, d \rangle$  into  $\langle Y^*, D^* \rangle$  with  $f(X)$  being  $D^*$ -dense in  $Y^*$ .

prob:completion-bis **131.** (*Continuation of [Exercise 130](#).*) This exercise establishes the essential uniqueness—or as one says, “uniqueness up to isometry”—of a completion of an arbitrary metric space.

Suppose  $f$  and  $f'$  are isometric embeddings of  $\langle X, d \rangle$  into metric spaces  $\langle Y, D \rangle$  and  $\langle Y', D' \rangle$ , respectively, that define completions of  $\langle X, d \rangle$ . Prove that then there is a unique isometry  $h$  from  $\langle Y, D \rangle$  to  $\langle Y', D' \rangle$  such that  $h \circ f = f'$ .

[*Hint:* To define  $h(y)$  for a point  $y \in Y$ , first use the fact that  $f(X)$  is  $D$ -dense in  $Y$  so as to find a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in  $X$  such that  $\langle f(x_n) \rangle_{n \in \mathbb{N}}$  converges to  $y$  in  $\langle Y, D \rangle$ . Next, show that  $\langle f'(x_n) \rangle_{n \in \mathbb{N}}$  converges to some point  $y' \in Y'$ , and finally set  $h(y) = y'$ . You will need to show that the  $y' \in Y'$  so obtained is independent of the choice of the particular sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$ ; in other words, if  $\langle z_n \rangle_{n \in \mathbb{N}}$  is another sequence in  $X$  such that  $\langle f(z_n) \rangle_{n \in \mathbb{N}}$  converges to  $y$  in  $\langle Y, D \rangle$ , then  $\langle f'(z_n) \rangle_{n \in \mathbb{N}}$  also converges to  $y'$ . ]

prob:completion-via-Cstar **132.** (*Continuation of [Exercise 130](#).*) This exercise provides an alternative construction of a completion of a metric space  $\langle X, d \rangle$ . Arbitrarily choose a point  $z \in X$ .

**(a)** For each  $x \in X$  let  $f_x: X \rightarrow \mathbb{R}$  be the function defined by

$$f_x(y) = d(x, y) - d(z, y) \quad (y \in X).$$

Verify that each  $f_x$  is bounded and  $\langle d, d' \rangle$ -continuous, where  $d'$  is the Euclidean metric on  $\mathbb{R}$ . In other words, in the notation of [Theorem 1.86](#), verify that  $f_x \in C^*(X)$  for each  $x \in X$ .



- (b) Let  $F: X \rightarrow C^*(X)$  be the map  $x \mapsto f_x$ . Show that  $F$  is an isometric embedding of  $\langle X, d \rangle$  into  $(C^*(X), d_\infty)$ , where  $d_\infty$  is the sup metric. completion of a metric space

*Note:* Although  $(C^*(X), d_\infty)$  is complete, we do not yet necessarily have a completion of  $\langle X, d \rangle$ , because  $F(X)$  need not be  $d_\infty$ -dense in  $C^*(X)$ .

- (c) Define  $Y$  to be the set of those  $g \in C^*(X)$  to which some sequence  $\langle g_n \rangle_{n \in \mathbb{N}}$  in  $F(X)$  converges in  $(C^*(X), d_\infty)$ . Let  $D$  be the metric on  $Y$  induced by  $d_\infty$ . Prove that  $F(X)$  is  $D$ -dense in  $Y$  and that  $Y$  is  $d_\infty$ -closed in  $C^*(X)$ . Conclude that  $\langle Y, D \rangle$  is complete. [Thus  $F$ , regarded as a map from  $X$  into  $Y$ , defines a completion of  $\langle X, d \rangle$ .]



## CHAPTER

# 2

## Topological Spaces

chap:spaces

### Chapter Contents

2.1	Topologies . . . . .	225	Boundary, interior, and closure in	
	Open sets . . . . .	226	a subspace . . . . .	258
	Closed sets . . . . .	229	Exercises for Section 2.3 . . . . .	259
	Subspaces . . . . .	230	2.4 Bases and Local Bases . . . . .	263
	Exercises for Section 2.1 . . . . .	233	Local bases . . . . .	263
2.2	Neighborhoods . . . . .	235	Zero-dimensional spaces . . . . .	266
	Properties of neighborhoods . . . . .	237	First-countable spaces . . . . .	268
	Defining a topology by its neighborhoods	238	Bases . . . . .	269
	Isolated points and limit points . . . . .	240	Second-countable spaces . . . . .	277
	Hausdorff spaces . . . . .	241	Separable spaces . . . . .	280
	Exercises for Section 2.2 . . . . .	243	Subbases . . . . .	282
2.3	Boundary, Interior, and Closure . . . . .	246	Exercises for Section 2.4 . . . . .	285
	Boundary of a set . . . . .	246	2.5 Separation Properties . . . . .	292
	Interior of a set . . . . .	250	The separation properties hierarchy . . . . .	295
	Closure of a set . . . . .	252	Separation properties of new spa-	
	Nowhere dense sets . . . . .	256	ces constructed from old . . . . .	297
			Exercises for Section 2.5 . . . . .	298

### Introduction

In the preceding chapter the notion of a metric space was abstracted from Euclidean spaces. In this chapter the abstraction is carried a step further to obtain the notion of a topological space. In a topological space there is no longer any quantitative measure of distance provided by a metric. Nevertheless, for such a space we can still talk about open and closed sets, neighborhoods, subspaces, and (as we shall see in the next chapter) limits and continuity.

### 2.1 Topologies

sec:topologies

We have seen that a given set (even  $\mathbb{R}^n$ ) can have a number of different metrics that yield the same open sets. Our definition of continuity of a map from one metric space to another was phrased in terms of the particular metrics on the two spaces. According to

provide set with a topology  
endow set with a topology

**Theorem 1.56**, however, continuity of such a map depends not on the metrics themselves, but solely on the open sets those metrics define. Similarly, for convergence of sequences it is the open sets and not the metrics that matter. All this suggests that we ought to be able to discuss continuity and convergence without any recourse to metrics at all, just so long as we can still talk about open sets. Hence we are about to discard what is not really essential for understanding continuity and convergence—namely, the metrics—and shall study open sets in a general setting devoid of any metrics.

### Open sets

subsec:open-sets

For a metric space, the notion ‘open set’ was defined by means of the metric, but now we want to talk about open sets in a context where there need be no distances and no  $\varepsilon$ -balls. Hence we no longer specify which sets are open by means of some prior notion such as that of a metric. Rather, we suppose already given in a set certain subsets said to be “open”; all we presume to know about these subsets is specified by a few properties—taken as axioms—that we *assume* they have. The properties we use as axioms are, of course, the very ones that are true of the open sets obtained from a metric when there is one.

**2.1 Definition.** Let  $X$  be a set. A **topology on  $X$**  is a collection  $\mathcal{T}$  of subsets of  $X$  such that:

(O1) *The empty set  $\emptyset \in \mathcal{T}$ , and the whole set  $X \in \mathcal{T}$ .*

(O2) *The union of any number of sets belonging to  $\mathcal{T}$  also belongs to  $\mathcal{T}$ .*

(O3) *The intersection of finitely many sets belonging to  $\mathcal{T}$  also belongs to  $\mathcal{T}$ .*

The subsets of  $X$  belonging to a topology  $\mathcal{T}$  are variously said to be  **$\mathcal{T}$ -open**, **open for  $\mathcal{T}$** , **open in  $X$**  (when the particular topology  $\mathcal{T}$  is understood), or simply **open** (when the particular set  $X$  and the particular topology  $\mathcal{T}$  are understood).

To verify that a collection  $\mathcal{T}$  of subsets of  $X$  is a topology on  $X$ , it is enough to check instead of (O3) the simpler condition:

(O3') *The intersection of any two sets belonging to  $\mathcal{T}$  also belongs to  $\mathcal{T}$ .*

In fact, an easy application of mathematical induction shows that (O3') implies (O3).

At last we come to the abstraction from metric spaces that is the principal object of study in general topology.

**2.2 Definition.** A **topological space** is an ordered pair  $(X, \mathcal{T})$  where  $X$  is a set, called the **underlying set of  $(X, \mathcal{T})$** , and  $\mathcal{T}$  is a topology on  $X$ , called the **topology of  $(X, \mathcal{T})$** .

Informally, then, a topological space consists of a set  $X$  provided with a specified collection of “open” subsets, subject only to the conditions that the empty set and the whole set  $X$  are open, that any union of open sets is open, and that any intersection of finitely many open sets is open.

When we specify a particular topology  $\mathcal{T}$  on a set  $X$ , we may say that we **provide** or **endow  $X$  with** that topology; and in the passive usage of those terms, we may say that the set  $X$  is **provided with** or **endowed with** that topology.

Ordinarily, when only one specific topology  $\mathcal{T}$  on a given set  $X$  is at issue, we need not explicitly name or mention  $\mathcal{T}$ . Accordingly, then, we may refer to  $X$  itself as the topological

space and speak simply of an “open subset of  $X$ .” And when no other underlying set of a topological space is at issue, we may speak simply of “an open set.”

**2.3 Examples.** (1) Let  $\langle X, d \rangle$  be a metric space. Take  $\mathcal{T}$  to be the collection of all  $d$ -open subsets of  $X$  in the sense of Definition 1.20; thus a subset  $U$  of  $X$  belongs to  $\mathcal{T}$  precisely when at each point of  $U$  there is some  $d$ -ball contained in  $U$ . In view of the motivation for Definition 2.1 it should come as no surprise that  $\mathcal{T}$  is a topology on  $X$ : Theorem 1.21 says just that.

The topology  $\mathcal{T}$  is said to be **induced by**  $d$  and is sometimes denoted by  $\mathcal{T}(d)$  to indicate its dependence on the particular metric  $d$ . The topological space  $\langle X, \mathcal{T} \rangle$  is said to be **associated with** the metric space  $\langle X, d \rangle$ .

**Convention.** Unless otherwise indicated, *any future reference to a topology on the set underlying a metric space shall be to the one induced by the metric.* Accordingly, when a specific metric space  $\langle X, d \rangle$  is at issue, we may speak of “an open subset of  $X$ ” or even “an open set” to mean a subset of  $X$  that is open for the topology induced by the metric, in other words, to mean a  $d$ -open subset of  $X$ .

Owing to the many examples of metric spaces presented in Chapter 1, numerous examples of topological spaces are now at hand.

(2) Specialize (1) by taking  $X$  to be a subset of some Euclidean space  $\mathbb{R}^n$  and  $d$  to be the Euclidean metric on  $X$ . Then the topology  $\mathcal{T}(d)$  induced by  $d$  is called the **usual topology** on  $X$ . In particular,  $n$ -dimensional Euclidean space  $\mathbb{R}^n$  becomes a topological space when provided with its usual topology.

Take  $n = 0$ : the usual topology on Euclidean 0-space  $\mathbb{R}^0 = \{0\}$  is just the collection  $\{\emptyset, \{0\}\}$ .

**Convention.** Unless otherwise indicated, *any future reference to a topology on a subset of a Euclidean space is to its usual topology.*

(3) Let  $X = \mathbb{R}^n$ , let  $d$  be the Euclidean metric on  $X$ , and let  $d_\infty$  be the max metric on  $X$ . By Proposition 1.36, a subset of  $X$  is  $d$ -open if and only if it is  $d_\infty$ -open. In other words, the topology  $\mathcal{T}(d)$  induced on  $X$  by  $d$  is the same as the topology  $\mathcal{T}(d_\infty)$  induced by  $d_\infty$ . However,  $d \neq d_\infty$  if  $n > 1$ . Thus the metric spaces  $\langle X, d \rangle$  and  $\langle X, d_\infty \rangle$  are different if  $n > 1$ , but their associated topological spaces  $\langle X, \mathcal{T}(d) \rangle$  and  $\langle X, \mathcal{T}(d_\infty) \rangle$  are the same.

(4) This example will show that any set  $X$  can be made into the underlying set of a topological space. Let  $X$  be a set. Then the collection  $\mathcal{P}(X)$  of all subsets of  $X$  is obviously a topology on  $X$ , called its **discrete topology**. Every subset of  $X$  is open for this topology. Now the discrete metric  $\delta$  on  $X$  has the property that every subset of  $X$  is  $\delta$ -open [see Examples 1.25 (2)]. Hence the discrete topology is induced by the discrete metric,

A topological space  $\langle X, \mathcal{T} \rangle$  is called a **discrete space** if its topology  $\mathcal{T}$  is the discrete topology on  $X$ .

For example, Euclidean 0-space  $\mathbb{R}^0 = \{0\}$  is a discrete space.

(5) Let  $X$  be any set. Then the collection  $\{\emptyset, X\}$  is trivially a topology on  $X$ , called the **indiscrete topology**. (The pun is intentional!)

topology!induced by metric  
topological space!associated with m  
Euclidean  $\mathbb{R}^n$ -space!usual topology  
usual topology!subset of Euclidean s  
Euclidean 0-space  
max metric!topology induced by  
topology!discrete  
discrete topology  
discrete space  
Euclidean 0-space  
indiscrete topology  
topology!indiscrete

ex-topology-induced-topologies

ex:euclidean-topology

convention:usual-top-on-subset-of- $\mathbb{R}^n$

ex:d-infinity-topology-on- $\mathbb{R}^n$

ex:discrete-topology

ex:indiscrete-topology

Sierpinski space

Sierpiński, Wacław

topology!finite-complement

finite-complement topology

topological space

topology

Of all topologies on a given set  $X$ , the indiscrete topology is the “smallest” and the discrete topology is the “largest” in the sense that

$$\{\emptyset, X\} \subset \mathcal{T} \subset \mathcal{P}(X)$$

for every topology  $\mathcal{T}$ . [See [Examples 3.9 \(10\)](#) for further discussion of the comparison of topologies on the same underlying set.]

On a set of two points, (4) and (5) already give two different topologies. The next example is a third one.

ex:Sierpinski-space

- (6) Let  $X = \{0, 1\}$ . Then the collection  $\mathcal{T} = \{\emptyset, \{0\}, X\}$  is a topology on  $X$  different from both the discrete and indiscrete topologies on  $X$ . The topological space  $\langle X, \mathcal{T} \rangle$  is called the **Sierpinski space**.

ex:finite-compl-topology

- (7) Let  $X$  be any set. Define  $\mathcal{T}$  to be the collection consisting of  $\emptyset$  together with all those subsets of  $X$  whose complements in  $X$  are finite. Then  $\mathcal{T}$  is a topology on  $X$ , called the **finite-complement topology** (or, more simply, the **co-finite topology**).

To see that  $\mathcal{T}$  is actually a topology on  $X$ , we must verify that axioms (O1)–(O3) are satisfied. We verify, for example (O3). Let  $\langle U_j \rangle_{j \in J}$  be a nonempty finite family of sets belonging to  $\mathcal{T}$ ; thus the index set  $J$  is a nonempty finite set. We must show that  $\bigcap_{j \in J} U_j \in \mathcal{T}$ . By definition of  $\mathcal{T}$ , for each  $j \in J$ , either  $U_j = \emptyset$  or  $X \setminus U_j$  is finite. Then there are two cases to consider.

Case (i): For some  $j \in J$ , the set  $U_j = \emptyset$ . In this case,  $\bigcap_{j \in J} U_j = \emptyset$  also, and so this intersection belongs to  $\mathcal{T}$ .

Case (ii): For each  $j \in J$ , the set  $U_j \neq \emptyset$ . In this case, the set  $X \setminus U_j$  is finite for each  $j$ . Now

$$X \setminus \bigcap_{j \in J} U_j = \bigcup_{j \in J} (X \setminus U_j)$$

by one of De Morgan’s Laws (0.19), and the union on the right is finite. Hence  $\bigcap_{j \in J} U_j$  belongs to  $\mathcal{T}$  in this case too.

However “unnatural” the finite-complement topology might seem, it does satisfy the axioms for a topology and hence is just as genuine a topology as any other. In fact, it is interesting precisely because of its pathology.  $\diamond$

For some later purposes we shall require the following notion.

def:pted-space

**2.4 Definition.** A **pointed (topological) space** is an ordered pair  $(X, x_0)$  consisting of a topological space  $X$  and some particular point  $x_0 \in X$ , referred to as the **distinguished point**.

Among all the topological spaces, those associated with metric spaces are important enough to deserve a special name.

def:metrizable

**2.5 Definition.** A topological space  $\langle X, \mathcal{T} \rangle$  and its topology  $\mathcal{T}$  are both said to be **metrizable** if  $\mathcal{T}$  is the topology induced by *some* metric on  $X$ , and then any such metric is said to be **compatible** with the topology. A topological space that is *not* metrizable is said to be **nonmetrizable**.

**Caution!** Note the distinction between ‘metric space’ and ‘metrizable space.’ A *metric* space  $\langle X, d \rangle$  consists of a set  $X$  together with one particular metric  $d$  on  $X$ ; it gives rise to the metrizable topological space  $\langle X, \mathcal{T}(d) \rangle$  but is not itself a topological space. By way of contrast, a *metrizable* space  $\langle X, \mathcal{T} \rangle$  is a topological space; its topology is induced by some metric (and usually by many metrics), but no particular metric is specified.

metrizable space  
ajar subset  
clopen subset  
closed set

Metrizable spaces are the “concrete” objects that topological spaces generalize (just as metric spaces are the objects that Euclidean spaces generalize). Hence they will be important for motivating definitions, suggesting theorems, and serving as test cases for conjectures about topological spaces in general.

There would be ample justification for studying topological spaces even were there not a single example of a nonmetrizable space (see the first paragraph of this section). The fact is, though, that nonmetrizable spaces not only do exist but can behave in peculiar ways that metrizable spaces cannot. Nonetheless, it is all too easy for the novice to fall into the psychological trap of believing that the only kind of topological space is a metrizable one. Machinery for dispelling this belief will very soon be at hand.

### Closed sets

subsec:closed-in-top-sp

The definition of ‘closed set’ in a metric space made no direct reference to the metric, only to the open sets the metric determined. By using the very same definition, we may then talk about closed sets in an arbitrary topological space.

def:closed-set

**2.6 Definition.** Given a topological space  $\langle X, \mathcal{T} \rangle$ , a subset  $E$  of  $X$  is variously said to be  $\mathcal{T}$ -**closed**, **closed for**  $\mathcal{T}$ , **closed in**  $X$ , or simply **closed** when its complement  $X \setminus E$  in  $X$  is  $\mathcal{T}$ -open.

The basic properties ([Theorem 1.24](#)) of  $d$ -closed sets in a metric space hold almost verbatim for closed sets in any topological space.

thm:closed-sets-properties

**2.7 Theorem.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space. Then:

thm-part:nil-and-whole-are-closed

(1) The empty set  $\emptyset$  and the entire set  $X$  are both closed in  $X$ .

part:intersect-of-closed-sets-is-closed

(2) The intersection of any number of closed subsets of  $X$  is itself closed in  $X$ .

of-finitely-many-closed-sets-is-closed

(3) The union of finitely many closed subsets of  $X$  is itself closed in  $X$ .

**Proof.** The proof is essentially the same as that for [Theorem 1.24](#). Indeed, aside from De Morgan’s Laws, the only thing used to prove [Theorem 1.24](#) was [Theorem 1.21](#); and the properties of open sets listed in [Theorem 1.21](#) hold by assumption in an arbitrary topological space.  $\square$

As in the special situation for metric spaces, so in any topological space a given subset may be: open but not closed; closed but not open; neither open nor closed (“*ajar*”); or both closed and open (“*clopen*,” as some mathematicians say—but we shall not use that ugly term!).

## Subspaces

subsec:subspaces

We are now going to give a method for making each and every subset of a given topological space into a topological space in its own right. Let  $\langle X, \mathcal{T} \rangle$  be a topological space and let  $Y$  be a subset of  $X$ . Of course there are topologies on  $Y$ —the discrete topology, for example—but we want to use the topology  $\mathcal{T}$  on  $X$  to construct a topology on  $Y$ .

metric!induced on subset

For motivation, consider the special case when  $\langle X, \mathcal{T} \rangle$  is actually metrizable. Choose a metric  $d$  on  $X$  that induces  $\mathcal{T}$ . Restrict the metric  $d: X \times X \rightarrow \mathbb{R}$  to the domain  $Y \times Y$  so as to obtain the metric  $d'$  on  $Y$  induced by  $d$ . The metric  $d'$  in turn induces a topology  $\mathcal{S}$  on  $Y$ , yielding a new topological space  $\langle Y, \mathcal{S} \rangle$ . By [Examples 1.22 \(4\)](#), the  $d'$ -open subsets of  $Y$  are just the intersections with  $Y$  of the  $d$ -open subsets of  $X$ , that is,

$$\mathcal{S} = \{U \cap Y : U \in \mathcal{T}\}.$$

Observe that metrics do not appear in any way in the preceding expression for  $\mathcal{S}$ . Hence this expression could be used to define a topology  $\mathcal{S}$  on  $Y$  even when  $\langle X, \mathcal{T} \rangle$  is not metrizable—provided we could be assured that  $\mathcal{S}$  really *is* a topology even when there are no metrics involved.

lem:subspace-topology

**2.8 Lemma.** *Let  $\langle X, \mathcal{T} \rangle$  be an arbitrary topological space and let  $Y \subset X$ . Then the collection*

$$\mathcal{S} = \{U \cap Y : U \in \mathcal{T}\}.$$

*of subsets of  $Y$  is a topology on  $Y$ .*

**Proof.** We verify axioms (O1)–(O3) for  $\mathcal{S}$ .

(O1) Since  $\emptyset = \emptyset \cap Y$  and  $Y = X \cap Y$ , both  $\emptyset$  and  $Y$  belong to  $\mathcal{S}$ .

(O2) Let  $\langle V_i \rangle_{i \in I}$  be any family of sets belonging to  $\mathcal{S}$ . For each  $i \in I$  we may write

$$V_i = U_i \cap Y$$

for some  $U_i \in \mathcal{T}$ . Since  $\bigcup_{i \in I} V_i \in \mathcal{S}$ , then the set

$$\bigcup_{i \in I} V_i = \bigcup_{i \in I} (U_i \cap Y) = \left( \bigcup_{i \in I} U_i \right) \cap Y$$

belongs to  $\mathcal{S}$ .

(O3) In the notation of the preceding proof of (O2), suppose in addition that  $I$  is finite. Since  $\bigcap_{i \in I} V_i \in \mathcal{S}$ , then the set

$$\bigcap_{i \in I} V_i = \bigcap_{i \in I} (U_i \cap Y) = \left( \bigcap_{i \in I} U_i \right) \cap Y$$

belongs to  $\mathcal{S}$ .  $\square$

That lemma justifies the following definition.

def:subspace

**2.9 Definition.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and let  $Y$  be a subset of  $X$ . The topology  $\mathcal{S} = \{U \cap Y : U \in \mathcal{T}\}$  on  $Y$  is called the **relative topology on  $Y$**  (induced by  $\mathcal{T}$ , or **inherited from  $\mathcal{T}$** ), and the topological space  $\langle Y, \mathcal{S} \rangle$  is said to be a **subspace of  $\langle X, \mathcal{T} \rangle$** . Moreover, we may refer to  $\langle X, \mathcal{T} \rangle$  as the **ambient space** of  $\langle Y, \mathcal{S} \rangle$ .



**Convention!** Unless otherwise explicitly indicated, *any future reference to a topology on a subset  $Y$  of a given topological space  $\langle X, \mathcal{T} \rangle$  is to the relative topology on  $Y$  induced by the given topology  $\mathcal{T}$  on  $X$ .*

Thus for a set  $A$  in a subspace  $\langle Y, \mathcal{S} \rangle$  of a topological space  $\langle X, \mathcal{T} \rangle$ ,

$$A \text{ is open in } Y \iff A = U \cap Y \text{ for some set } U \text{ open in } X$$

(see Figure 2.1).



Figure 2.1: Open set in a subspace.

fig:open-in-subspace

**2.10 Examples.** (1) To recapitulate the discussion preceding Lemma 2.8 in the language of Definition 2.9, let  $\langle X, \mathcal{T} \rangle$  be a metrizable topological space, let  $d$  be a metric on  $X$  inducing  $\mathcal{T}$ , and let  $Y \subset X$ . If  $d'$  is the metric on  $Y$  induced by  $d$ , then the topology  $\mathcal{S}$  on  $Y$  induced by  $d'$  is the same as the relative topology on  $Y$  induced by the topology  $\mathcal{T}$ . This fact is indicated schematically in Figure 2.2, where each arrow stands for ‘induces’.

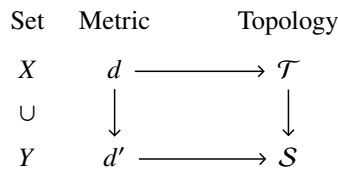


Figure 2.2: Relative topology induced by a metric.

fig:rel-topology-induced-by-metric

Metrizability is thus a “hereditary” property of topological spaces in the sense that *every subspace of a metrizable space is metrizable*.

(2) Earlier (page 227) we adopted the convention that a subset  $Y$  of  $\mathbb{R}^n$  is to be considered a topological space by using its “usual” topology, that is, the topology induced by the Euclidean metric on  $Y$ . Hence *the usual topology on a subset of  $\mathbb{R}^n$  is just the relative topology induced by the usual topology on  $\mathbb{R}^n$ .*

ex:R-subspace-of-Rextended (3) Specialize (1) by taking  $X = \widehat{\mathbb{R}}$  (the extended real line),  $d = \widehat{d}$  (the metric on  $\widehat{\mathbb{R}}$  defined in [Example 1.41](#)), and  $Y = \mathbb{R}$ . Let  $d'$  be the metric on  $\mathbb{R}$  induced by  $\widehat{d}$ . According to [Proposition 1.42](#), both  $d'$  and the Euclidean metric on  $\mathbb{R}$  induce the very same topology on  $\mathbb{R}$ . Hence *the real line  $\mathbb{R}$  (with its usual topology) is a subspace of the extended real line  $\widehat{\mathbb{R}}$  (with its topology induced by  $\widehat{d}$ )*.  $\diamond$

It is easy to tell which sets are closed in a subspace.

prop:closed-in-subspace **2.11 Proposition.** *Let  $\langle Y, \mathcal{S} \rangle$  be a subspace of a topological space  $\langle X, \mathcal{T} \rangle$ , and let  $A \subset Y$ . Then  $A$  is closed in  $Y$  if and only if  $A = E \cap Y$  for some closed subset  $E$  of  $X$ .*

**Proof.** First, assume that  $A$  is closed in  $Y$ . Then  $Y \setminus A$  is open in  $Y$ , so by definition of the relative topology on  $Y$  there is some open subset  $U$  of  $X$  with

$$Y \setminus A = U \cap Y.$$

But then

$$A = Y \setminus (U \cap Y) = Y \cap (X \setminus U)$$

with  $X \setminus U$  being closed in  $X$ .

Conversely, assume that  $A = E \cap Y$  with  $E$  closed in  $X$ . Then the set  $U = X \setminus E$  is open in  $X$ , the set

$$Y \setminus A = Y \setminus (E \cap Y) = Y \setminus ((X \setminus U) \cap Y) = U \cap Y$$

is open in  $Y$ , and hence  $A$  is closed in  $Y$ .  $\square$

Together, [Definition 2.9](#) and [Proposition 2.11](#) tell us that if  $A \subset Y$  and if  $A$  is open or closed in the entire space  $X$ , then  $A = A \cap Y$  is open or closed, respectively, in the subspace  $Y$ . The converse is *not* true in general: for example,  $A = [0, 1[$  is both open and closed in  $Y = [0, 1[$ , but  $A$  is neither open nor closed in  $X = \mathbb{R}$ . The following partial converse is true, however.

**2.12 Proposition.** *Let  $\langle Y, \mathcal{S} \rangle$  be a subspace of a topological space  $\langle X, \mathcal{T} \rangle$ , and let  $A \subset Y$ . Then:*

- (1) *If  $A$  is open in  $Y$  and if  $Y$  is open in  $X$ , then  $A$  is open in  $X$ .*
- (2) *If  $A$  is closed in  $Y$  and if  $Y$  is closed in  $X$ , then  $A$  is closed in  $X$ .*

**Proof.** (1) If  $A$  is open in  $Y$ , then  $A = U \cap Y$  for some open subset  $U$  of  $X$ . If in addition  $Y$  is open in  $X$ , then the intersection  $A = U \cap Y$  of two open subsets of  $X$  is open in  $X$ .

(2) Replace ‘open’ by ‘closed’ above.  $\square$

We have by now two good reasons for using at times such expressions as ‘open for  $\mathcal{S}$ ’ and ‘closed in  $X$ ’ instead of the simpler ‘open’ and ‘closed’. First, even so innocent a set as  $X = \{0, 1\}$  has several different topologies, so that a reference to an open or a closed subset of  $X$  will be ambiguous until a particular topology on  $X$  is singled out. Second, a set  $A$  in a topological space  $\langle X, \mathcal{T} \rangle$  may be a subset of a subspace  $\langle Y, \mathcal{S} \rangle$  of  $\langle X, \mathcal{T} \rangle$ , and then being open or closed in  $Y$  (that is, for  $\mathcal{S}$ ) need not be the same as being open or closed, respectively, in  $X$  (that is, for  $\mathcal{T}$ ).

Ordinarily, our discussion of open or closed sets will concern either an arbitrary topology or else some particular topology whose identity is clear from the context. In that case, when there is no danger of genuine ambiguity, we may suppress explicit mention of the topology ‘ $\mathcal{T}$ ’ and indulge in the following “abuse of language”:

relative topology  
subspace  
included point topology  
topology!included point  
excluded point topology  
topology!excluded point

**Convention!** Given a topological space  $\langle X, \mathcal{T} \rangle$ , we may refer to  $X$  itself as a topological space.

As an example of this convention, if a set  $X$  underlies a topological space  $\langle X, \mathcal{T} \rangle$ , we may call a subset  $Y$  of  $X$  simply a “subspace of  $X$ .” Of course we must still distinguish between being open in  $Y$  from being open in  $X$  (unless it so happens that  $Y$  is itself already open in  $X$ ).

### EXERCISES FOR SECTION 2.1

prob:topologies  
prob:4-pt-regular-not-T0-space

1. (a) Verify that  $\mathcal{T} = \{\emptyset, \{0, 1\}, \{2, 3\}, X\}$  is a topology on the 4-element set  $X = \{0, 1, 2, 3\}$ .  
(b) List all the closed sets in the topological space  $\langle X, \mathcal{T} \rangle$ .

prob:finite-space-topologies

2. (a) Show that a set of one element has a unique topology. Does any set having more than one element have a unique topology?  
(b) Find a fourth topology on  $\{0, 1\}$  different from the three obtained earlier. Are there more than four?  
(c) Find all topologies on the three-element set  $\{0, 1, 2\}$ . Are some of these like others of these in some sense?  
(d) Show that there are at most  $2^{(2^n-2)}$  topologies on a finite set that consists of  $n \geq 1$  elements.  
*Note:* A better upper bound than  $2^{(2^n-2)}$  can be established, namely,  $2^{n(n-1)}$ . However, it is still an unsolved problem to find a closed-form formula telling exactly how many topologies there are on a set of  $n$  elements, or even to calculate the actual number for all but fairly small values of  $n$ .

3. Is a discrete space the only kind of topological space having the property that each one-point set  $\{x\}$  is open?

included-and-excluded-pt-topologies

4. The Sierpinski space [Examples 2.3 (6)] can be generalized in three ways as follows. Let  $X$  be a set.  
(a) Let  $A$  be a given nonempty proper subset of  $X$ . Verify that  $\{\emptyset, A, X\}$  is a topology on  $X$ .

Now let  $x_0$  be a given element of  $X$ .

prob-part:included-pt-topology

- (b) Verify that  $\{U \subset X : U = \emptyset \text{ or } x_0 \in U\}$  is a topology on  $X$ . This is called an **included point topology**.

prob-part:excluded-pt-topology

- (c) Verify that  $\{U : U \subset X, x_0 \notin U\}$  is a topology on  $X$ . This is called an **excluded point topology**.

5. The condition  $x_0 \in U$  used in Exercise 4 (b) to define an included point topology may be written  $\{x_0\} \subset U$ . This suggests a generalization. Let  $A$  be a given subset of a set  $X$ .

countable-complement topology (a) Verify that  $\{U : U = \emptyset \text{ or } A \subset U \subset X\}$  is a topology on  $X$ . This is called an “included set topology.”

linearly open set (b) Can excluded point topologies be similarly generalized to “excluded set topologies”?

pseudometrizable space

Sierpinski space

6. Are there only countably many or uncountably many topologies on the set  $\mathbb{N}$  of all natural numbers?

prob:countable-complement-topology

7. Let  $X$  be any set. Call a subset  $U$  of  $X$  “open” whenever  $U = \emptyset$  or  $X \setminus U$  is countable. Show that this collection of “open” subsets of  $X$  really does constitute a topology on  $X$ . This topology is called the **countable-complement topology on  $X$** .

prob:cf-3-topologies-on-R

8. Consider on the set  $\mathbb{R}$  the usual topology, the finite-complement topology, and the countable-complement topology. Are any of these contained in others of these?

9. Is the intersection of any two topologies on a set necessarily a topology of the set? the union of two topologies?

prob:linearly-open-in-plane

10. Call a subset  $U$  of the plane  $\mathbb{R}^2$  “linearly open” if its intersection with each line  $L$  in  $\mathbb{R}^2$  is open in  $L$  (for the usual topology of  $L$ ). Let  $\mathcal{S}$  be the collection of all linearly open subsets of  $\mathbb{R}^2$ .

(a) Verify that  $\mathcal{S}$  is a topology on  $\mathbb{R}^2$ .

(b) If  $K$  is a convex subset of  $\mathbb{R}^2$ , show that  $K$  is open for the topology  $\mathcal{S}$  if and only if it is open for the usual topology, and closed for  $\mathcal{S}$  if and only if it is closed for the usual topology.

(c) It is obvious from the definition that a subset of the plane will be linearly open if it is open for the usual topology. Is the converse true?

prob:pseudometrizable

11. Let  $d$  be a pseudometric on a set  $X$  (see [Exercise 1.20](#)). According to [Exercise 1.44](#), the collection  $\mathcal{T}(d)$  of all  $d$ -open sets forms a topology on  $X$ . Just as in the special case when  $d$  is a metric, we say that the topology  $\mathcal{T}(d)$  is **induced by** the pseudometric  $d$ . A **pseudometrizable space** is a topological space whose topology is induced by some pseudometric.

(a) Give an example of a pseudometrizable space that is not metrizable.

(b) Is the Sierpinski space [\[Examples 2.3 \(6\)\]](#) pseudometrizable?

12. (a) Are there topological spaces in which the intersection of *every* collection of open sets is open?

(b) Answer the analog of (a) for unions of closed sets.

13. Let  $X$  be a set and let  $\mathcal{E}$  be a given collection of subsets of  $X$  such that: (i)  $\emptyset \in \mathcal{E}$  and  $X \in \mathcal{E}$ ; (ii) the intersection of any number of members of  $\mathcal{E}$  is again a member of  $\mathcal{E}$ ; and (iii) the union of finitely many members of  $\mathcal{E}$  is again a member of  $\mathcal{E}$ . Prove that there is then a unique topology  $\mathcal{T}$  on  $X$  for which  $\mathcal{E}$  is the collection of all  $\mathcal{T}$ -closed sets.

14. Let  $Y$  be a subset of a topological space  $X$ .

(a) Show by example that the collection

$$\mathcal{S} = \{U : U \subset Y, U \text{ is open in } X\}$$

need not be a topology on  $Y$ .

(b) Is  $\mathcal{S}$  ever a topology on  $Y$ ?

15. Let  $\mathcal{T}$  be the usual topology on  $\mathbb{R}$  and let  $f: \mathbb{R} \rightarrow \mathbb{R}$  be a function.

(a) Is the collection  $\{f(U) : U \in \mathcal{T}\}$  necessarily a topology on  $\mathbb{R}$ ? If so, prove it. If not, is it ever a topology on  $\mathbb{R}$  for particular functions  $f$ ?

(b) Repeat (a) but for the collection  $\{f^{-1}(V) : V \in \mathcal{T}\}$ .

16. Prove: A subspace  $Z$  of a subspace  $Y$  of a topological space  $X$  is itself a subspace of  $X$ . In other words: if  $\langle X, \mathcal{T} \rangle$  is a topological space, if  $Z \subset Y \subset X$ , and if  $\mathcal{T}_Y$  is the relative topology on  $Y$  induced by  $\mathcal{T}$ , then the relative topology on  $Z$  induced by  $\mathcal{T}_Y$  is the same as the relative topology on  $Z$  induced by  $\mathcal{T}$ . *Note:* This removes an apparent ambiguity in our convention that a subset of a space is to be provided with its relative topology.

17. Can a subspace consisting of more than one point of a topological space  $X$  be discrete when the large space  $X$  is not discrete? even when the subspace is infinite? when the subspace is uncountable?

prob:cartesian-sum-disjoint

18. Let  $\langle X_i \rangle_{i \in I}$  be a family of topological spaces with  $X_i$  disjoint from  $X_j$  whenever  $i \neq j$ , and let  $X = \bigcup_{i \in I} X_i$ .

(a) Show that there is a unique topology  $\mathcal{T}$  on  $X$  for which each  $X_i$ , with its given topology, is an open subspace.

The resulting topology on  $X$  is called the **(Cartesian) sum topology on  $X$** , and the topological space with underlying set  $X$  and this topology is called the **(Cartesian) sum of** the pairwise disjoint family  $\langle X_i \rangle_{i \in I}$  of topological spaces. *Note:* This construction will be generalized in [Exercise 3.106](#) to the situation where the family of spaces is not necessarily disjoint.

(b) Prove that each  $X_i$  is also closed in the sum space  $X$ .

(c) Suppose  $I = \mathbb{N}$  and  $X_i = \mathbb{R} \times \{i\}$  for each  $i \in I$ , so that  $X \subset \mathbb{R} \times \mathbb{R}$ . Verify that the sum topology on  $X$  is the same as the usual topology on  $X$ .

(d) Construct an example where each  $X_i$  is a subspace of  $\mathbb{R} \times \mathbb{R}$  yet the sum space  $X$  is *not* a subspace of  $\mathbb{R} \times \mathbb{R}$ .

prob:Gdelta-Fsigma-top

19. The notions of  $G_\delta$ -sets and  $F_\sigma$ -sets, defined in [Exercise 1.41](#) for metric spaces, generalize to arbitrary topological spaces: A subset  $A$  of a topological space  $X$  is called a  **$G_\delta$ -set** if it is the intersection of some sequence of open subsets of  $X$ , and  $A$  is called an  **$F_\sigma$ -set** if it is the union of some sequence of closed subsets of  $X$ . Evidently the complement in  $X$  of a  $G_\delta$ -set is an  $F_\sigma$ -set, and vice versa.

(a) Show that the intersection of a countable collection of  $G_\delta$ -sets is again a  $G_\delta$ -set.

(b) Give an example of a denumerable collection of  $G_\delta$ -sets in some space  $X$  whose union is not a  $G_\delta$ -set.

(c) Show that the union of finitely many  $G_\delta$ -sets is again a  $G_\delta$ -set.

(d) Do the counterparts of (a)–(c) for  $F_\sigma$ -sets.

o-part:intersect-count-Gdeltas-is-Gd

union-count-Gdeltas-not-nec-Gdelta

o-part:union-finite-Gdeltas-is-Gdelta

## 2.2 Neighborhoods

sec:nbds

In a metric space, points  $\varepsilon$ -close to a given point  $x$  were those belonging to the  $d$ -ball of radius  $\varepsilon$  at  $x$ . More generally, points belonging to any neighborhood  $V$  of  $x$  could be thought of as being close to  $x$ , although  $V$  no longer had to provide a numerical measure of

Cartesian sum!family of spaces@of a  
Gdelta-set@\$G\_\delta\$-set  
Fsigma-set@\$F\_\sigma\$-set

discrete space closeness to  $x$ . Now our definition of neighborhood in a metric space referred only to open sets, not to the metric. Hence simply by extending that definition we can introduce into any topological space the idea of points close to a given point.

indiscrete space

finite-complement topology

def:nbd **2.13 Definition.** Let  $x$  be a point in a topological space  $X$ . Then a subset  $V$  of  $X$  is called a **neighborhood of  $x$  in  $X$**  when  $x \in U \subset V$  for some open subset  $U$  of  $X$ . The collection of all neighborhoods of  $x$  in  $X$  is called the **neighborhood system at the point  $x$**  and is denoted by  $\mathcal{N}_x(X)$  or, when the space  $X$  is understood, simply by  $\mathcal{N}_x$ .

As usual, we may refer to a “ $\mathcal{T}$ -neighborhood of  $x$ ” or a “neighborhood of  $x$  for  $\mathcal{T}$ ” in order to indicate the particular topology  $\mathcal{T}$  on  $X$  under consideration.

**Usage note.** According to the preceding definition, *a neighborhood of a point need not be an open set*. Some mathematicians, especially in the older literature, restrict the term ‘neighborhood’ to mean what for us is an open neighborhood.

exs:nbds **2.14 Examples.** (1) Let  $\langle X, d \rangle$  be a metric space and let  $\langle X, \mathcal{T} \rangle$  be the associated topological space. Given  $x \in X$ , a subset  $V$  of  $X$  is a  $d$ -neighborhood of  $x$  in the metric space sense exactly when it is a  $\mathcal{T}$ -neighborhood of  $x$ . (After all, [Definition 1.31](#) was generalized to the preceding definition so as to ensure precisely that!)

Hence the neighborhoods of a point in a metrizable space are exactly the  $d$ -neighborhoods obtained from any metric  $d$  that induces the topology.

- (2) In a discrete space  $X$ , every subset of  $X$  containing a given point  $x \in X$  is a neighborhood of  $x$ . At the opposite extreme, in an indiscrete space  $X$ , the only neighborhood of a point is the entire space  $X$ .
- (3) Let  $X$  be a set provided with its finite-complement topology [[Examples 2.3 \(7\)](#)] and let  $x \in X$ . Then a subset  $V$  of  $X$  containing  $x$  is a neighborhood of  $x$  if and only if  $X \setminus V$  is finite. In fact, for  $x \in U \subset V$  we have:  $X \setminus V \subset X \setminus U$ ; the set  $U$  is open if and only if  $X \setminus U$  is finite; and in that event  $X \setminus V$  is also finite. Note that then *every neighborhood of a point in this space is open!*  $\diamond$

Recall that the open sets in a subspace  $Y$  of a space  $X$  are just the intersections with  $Y$  of the open sets in the entire space  $X$ . Hence a similar result holds for neighborhoods.

prop:nbds-in-subspace **2.15 Proposition.** Let  $y$  be a point in a subspace  $Y$  of a topological space  $X$ . Then the neighborhood system at  $y$  in  $Y$  is the collection

$$\{V \cap Y : V \text{ is a neighborhood of } y \text{ in } X\}.$$

**Proof.** Let  $W$  be a neighborhood of  $y$  in the subspace  $Y$ . There is an open set  $U_1$  in  $Y$  with

$$y \in U_1 \subset W.$$

Since the topology on  $Y$  is the relative one,

$$U_1 = U \cap Y$$

for some open set  $U$  in  $X$ . Because  $y \in U$  and  $U$  is open, the set

$$V = W \cup U$$

is a neighborhood of  $y$  in  $X$ . Moreover,

$$V \cap Y = (W \cup U) \cap Y = (W \cap Y) \cup (U \cap Y) = W \cup U_1 = W.$$

Conversely, let  $V$  be a neighborhood of  $y$  in  $X$ . Then  $y \in U \subset V$  for some open set  $U$  in  $X$ . Now  $y \in U \cap Y$ , the set  $U \cap Y$  is open in  $Y$ , and  $U \cap Y$  is contained in  $V \cap Y$ . Hence  $V \cap Y$  is a neighborhood of  $x$  in  $Y$ .  $\square$

In the notation of [Definition 2.13](#), [Proposition 2.15](#) asserts that

$$\mathcal{N}_y(Y) = \{V \cap Y : V \in \mathcal{N}_y(X)\}.$$

Among the neighborhoods of a point  $x$  in a topological space  $X$  are all those open subsets of  $X$  that contain  $x$ . Of course, a neighborhood of  $x$  may be closed in  $X$ ; indeed, if  $V$  is a closed subset of  $X$  and if  $x \in U \subset V$  for some open subset  $U$  of  $X$ , then  $V$  is a neighborhood of  $x$ .

The proof of [Proposition 1.32](#) used only properties of open sets in a metric space that hold more generally in any topological space. Hence the very same proof establishes the following proposition.

prop:open-iff-nbd-each-pt

**2.16 Proposition.** *A subset of a topological space is open in the space if and only if it is a neighborhood of each of its points.*

The concept of neighborhood of a point has a natural generalization to that of a neighborhood of a set.

def:nbd-of-set

**2.17 Definition.** Let  $A$  be a subset of a topological space  $X$ . A **neighborhood of  $A$  in  $X$**  is a subset  $V$  of  $X$  such that  $A \subset U \subset V$  for some open set  $U$  in  $X$ .

Thus a neighborhood of a singleton  $\{x\}$  is just a neighborhood of the point  $x$  in the original sense of the term ‘neighborhood’. In the Euclidean plane  $\mathbb{R}^2$ , the horizontal strip set  $\{(x, y) \in \mathbb{R}^2 : |y| < \varepsilon\}$  is, for each  $\varepsilon > 0$ , a neighborhood of the  $x$ -axis  $\{(x, 0) : x \in \mathbb{R}\}$ .

### Properties of neighborhoods

subsec:nbd-properties

The fundamental properties of neighborhoods are listed in the next theorem

thm:nbd-axioms

**2.18 Theorem (properties of neighborhood systems).** *Let  $X$  be a topological space. For each  $x \in X$  the neighborhood system  $\mathcal{N}_x$  at  $x$  has the properties:*

property:N1 (N1) *There exists at least one member of  $\mathcal{N}_x$ .*

property:N2 (N2) *The point  $x$  belongs to each member of  $\mathcal{N}_x$ .*

property:N3 (N3) *The intersection of any two members of  $\mathcal{N}_x$  also belongs to  $\mathcal{N}_x$ .*

property:N4 (N4) *Each subset of  $X$  containing a member of  $\mathcal{N}_x$  also belongs to  $\mathcal{N}_x$ . (That is, each “superset” of a member of  $\mathcal{N}_x$  belongs to  $\mathcal{N}_x$ .)*

property:N5 (N5) *Each member  $V$  of  $\mathcal{N}_x$  contains some member  $U$  of  $\mathcal{N}_x$  such that  $U$  belongs to  $\mathcal{N}_y$  for each  $y \in U$ .*

**Proof.** (N1) Let  $x \in X$ . Then  $X \in \mathcal{N}_x$ . In fact,  $X$  is an open set with  $x \in X \subset X$ , so that  $X$  is a neighborhood of  $x$ .

neighborhood

(N3) Let  $x \in X$ , let  $V \in \mathcal{N}_x$ , and let  $W \in \mathcal{N}_x$ . Since  $V$  and  $W$  are neighborhoods of  $x$ , there exist open sets  $U_1$  and  $U_2$  such that

$$x \in U_1 \subset V, \quad x \in U_2 \subset W.$$

Then  $x \in U_1 \cap U_2 \subset V \cap W$  with  $U_1 \cap U_2$  being open in  $X$ . Hence  $V \cap W$  is a neighborhood of  $x$ , that is,  $V \cap W \in \mathcal{N}_x$ .

(N5) Let  $x \in X$  and  $V \in \mathcal{N}_x$ . There exists an open set  $U$  with  $x \in U \subset V$ . By [Proposition 2.16](#), for each  $y \in U$  the set  $U$  is a neighborhood of  $y$ , that is,  $U \in \mathcal{N}_y$ .

Verification of (N2) and (N4) is left to the reader.  $\square$

### Defining a topology by its neighborhoods

subsec:def-top-by-nbds

The preceding theorem describes the behavior of neighborhoods of points in a set on which a topology is already given. Suppose now we are given just a set  $X$ , without any topology. Suppose, further, that along with each point  $x$  of  $X$  we are given certain subsets of  $X$  that behave as they would were they really neighborhoods of  $x$  obtained from a topology on  $X$ . The next theorem says that we can then uniquely reconstruct a topology for which the neighborhoods of each  $x \in X$  are precisely the given subsets associated with  $x$ .

thm:topology-from-nbds

**2.19 Theorem (constructing a topology from neighborhood systems).** *Let  $X$  be a set. For each  $x \in X$  let  $\mathcal{M}_x$  be a given collection of subsets of  $X$ . Assume that these collections have properties ((N1))–((N5)), as listed in 2.18 [but for the collections  $\mathcal{M}_x$  rather than  $\mathcal{N}_x$ ]. Then there exists a unique topology  $\mathcal{T}$  on  $X$  such that for each  $x \in X$ , the collection  $\mathcal{M}_x$  is the neighborhood system  $\mathcal{N}_x$  at  $x$ . Moreover, a subset  $V$  of  $X$  is  $\mathcal{T}$ -open if and only if  $V \in \mathcal{M}_x$  for each  $x \in V$ .*

**Proof.** First we shall prove uniqueness and, in doing so, show that any such topology  $\mathcal{T}$  must have a specific form. This will tell us actually how to construct  $\mathcal{T}$ .

Uniqueness. Let  $\mathcal{T}$  be a topology on  $X$  such that for each  $x \in X$ , the collection  $\mathcal{M}_x = \mathcal{N}_x$ , the neighborhood system at  $x$ . Apply [Proposition 2.16](#) to the topological space  $\langle X, \mathcal{T} \rangle$ : A subset  $V$  of  $X$  is  $\mathcal{T}$ -open if and only if  $V$  is a neighborhood of each  $x \in V$ . Hence  $V$  is  $\mathcal{T}$ -open if and only if  $V \in \mathcal{M}_x$  for each  $x \in V$ . Thus

$$\{eq:top-via-nbds\} \quad (*) \quad \mathcal{T} = \{V : V \subset X, V \in \mathcal{M}_x \text{ for each } x \in V\}.$$

This establishes uniqueness of  $\mathcal{T}$ .

Existence. Define  $\mathcal{T}$  by condition (\*). We must show, first, that  $\mathcal{T}$  is a topology on  $X$  and, second, that for each  $x \in X$ , the collection  $\mathcal{M}_x$  is the  $\mathcal{T}$ -neighborhood system  $\mathcal{N}_x$  at  $x$ . To show these two things, we must be careful to use only the definition (\*) of  $\mathcal{T}$  and the assumed properties ((N1))–((N5)) for the collections  $\mathcal{M}_x$ .

We show that  $\mathcal{T}$  is a topology on  $X$  by verifying in turn each of the axioms ((O1))–((O3)) for a topology.

(O1) Since no  $x \in \emptyset$ , it is vacuously true that  $\emptyset \in \mathcal{M}_x$  for each  $x \in \emptyset$ . Hence  $\emptyset \in \mathcal{T}$ . To see that  $X \in \mathcal{T}$ , let  $x \in X$  be arbitrary. By (N1) there exists some  $U \in \mathcal{M}_x$ . Since  $U \subset X$ , then  $X \in \mathcal{M}_x$  by (N4).



(O2) Let  $\langle V_i \rangle_{i \in I}$  be a family of sets belonging to  $\mathcal{T}$  and let  $V = \bigcup_{i \in I} V_i$ . To see that  $V \in \mathcal{T}$ , let  $x \in V$  be arbitrary; we must show that  $V \in \mathcal{M}_x$ . Choose  $i \in I$  with  $x \in V_i$ . Since  $V_i \in \mathcal{T}$ , then  $V_i \in \mathcal{M}_x$ . Since  $V_i \subset V$ , then  $V \in \mathcal{M}_x$  by (N4).

(O3) is left to the reader to verify.

To complete the proof, we show that for each  $x \in X$  the collection  $\mathcal{M}_x = \mathcal{N}_x$ , the  $\mathcal{T}$ -neighborhood system  $\mathcal{N}_x$  at  $x$ . Let  $x \in X$  be arbitrary.

First we show that  $V \in \mathcal{M}_x \implies V \in \mathcal{N}_x$ . Let  $V \in \mathcal{M}_x$ . By (N5) there is some  $U \in \mathcal{M}_x$  such that  $U \subset V$  and  $U \in \mathcal{M}_y$  for each  $y \in U$ . Then  $U \in \mathcal{T}$  by definition of  $\mathcal{T}$ , and  $x \in U$  by (N2). Thus  $x \in U \subset V$  with  $U$  being  $\mathcal{T}$ -open. By definitions of neighborhoods for a topology, the set  $V \in \mathcal{N}_x$ .

Conversely, we show that  $V \in \mathcal{N}_x \implies V \in \mathcal{M}_x$ . Let  $V \in \mathcal{N}_x$ . Then  $x \in U \subset V$  for some  $\mathcal{T}$ -open set  $U$ . By definition of  $\mathcal{T}$ , the set  $U \in \mathcal{M}_x$ . Hence  $V \in \mathcal{M}_x$  by (N4).  $\square$

To define ‘topological space’ we used ‘open set’ as the primitive notion, subject to axioms ((O1))–((O3)) listed in 2.1. Then we defined ‘neighborhood’ in terms of ‘open sets’ and deduced from these axioms the properties ((N1))–((N5)) of neighborhoods, as listed in Theorem 2.18. However, Theorem 2.19 indicates we would have arrived at the very same notion of a topological space had we used ‘neighborhood’ as the primitive terms and assumed ((N1))–((N5)) as axioms.

Historically it was just this latter approach to topological spaces that was followed. In 1914 Felix Hausdorff abstracted from various “spaces” the concept of neighborhood, took ‘neighborhood’ as an undefined term, and set out as axioms for neighborhoods essentially properties ((N1))–((N5)). (Hausdorff included also a “separation axiom” that is discussed below.)

In addition to this theoretical interest, Theorem 2.19 provides a method for constructing examples of topological spaces.

exs:topology!right-interval

**2.20 Examples.** (1) Take  $X = \mathbb{R}$  as a set. For each  $x \in X$ , define  $\mathcal{M}_x$  as follows:  $V \in \mathcal{M}_x$  if and only if  $V$  contains the left-closed, right-open interval  $[x, z[$  with left endpoint  $x$  for some  $z > x$ .

Let  $x \in X$  be arbitrary. Clearly properties (N1), (N2), and (N4) hold at  $x$ .

To verify (N3), let  $V_1, V_2 \in \mathcal{M}_x$ . For some  $z_1 > x_1$  and  $z_2 > x_2$ , we have  $[x, z_1[ \subset V_1$  and  $[x, z_2[ \subset V_2$ . Let  $z$  be the minimum of  $z_1$  and  $z_2$ . Then  $[x, z[ \subset V_1 \cap V_2$ , and so  $V_1 \cap V_2 \in \mathcal{M}_x$ .

To verify (N5), let  $V \in \mathcal{M}_x$ . Choose  $z > x$  with  $[x, z) \subset V$ . Then  $[x, z[ \in \mathcal{M}_y$  for each  $y \in [x, z[$ . In fact, if  $y \in [x, z[$ , then  $[y, z[ \subset [x, z[$  whence  $[x, z[ \in \mathcal{M}_y$ .

By Theorem 2.19 there is a unique topology on  $\mathbb{R}$  for which the neighborhood system of an  $x \in \mathbb{R}$  is exactly  $\mathcal{M}_x$ . This topology is variously called the **right-interval topology** or the **lower-limit topology** on  $\mathbb{R}$ . When the set  $\mathbb{R}$  is provided with this topology, the resulting topological space is known as the **Sorgenfrey line**, after R. Sorgenfrey, and is denoted by  $\mathbb{R}_l$ . [The reason for the designation ‘lower-limit’ and the subscript  $l$  in the notation will be explained later: see Examples 3.101 (4).]

Later we shall see that the Sorgenfrey line provides a rich source of counterexamples.

ex:right-interval-topology

(2) More generally, let  $X$  be any totally ordered set having no greatest element. For each  $x \in X$ , define  $\mathcal{M}_x$  to consist of all those subsets of  $X$  that contain intervals of the form  $[x, z[$  with  $z > x$ . Just as in (1), the collections  $\mathcal{M}_x$  satisfy ((N1))–((N5)). The resulting topology given by Theorem 2.19 is called the **right-interval topology on  $X$** .

Hausdorff, Felix  
topology!right-interval  
right-interval topology  
Sorgenfreyline  
Sorgenfrey, Robert  
topology!right-interval  
right-interval topology

ex:line-with-two-origins (3) Let  $X = \mathbb{R} \cup \{0'\}$ , where  $0' \notin \mathbb{R}$ . It does not matter what the object  $0'$  is, just that it topology!neighborhood@ does not belong to  $\mathbb{R}$ . For brevity call any neighborhood of an  $x \in \mathbb{R}$  for the usual topology of  $\mathbb{R}$  simply an “ $\mathbb{R}$ -neighborhood of  $x$ .”

For each  $x \in \mathbb{R}$ , let  $\mathcal{N}_x$  consist of all subsets of  $X$  that contain some  $\mathbb{R}$ -neighborhood of  $x$ . In other words, each member of  $\mathcal{N}_x$  is either already an  $\mathbb{R}$ -neighborhood of  $x$  or else has the form  $V \cup \{0'\}$  for some  $\mathbb{R}$ -neighborhood  $V$  of  $x$ . Further, let  $\mathcal{N}_{0'}$  consist of all subsets of  $X$  that contain  $(V \setminus \{0\}) \cup \{0'\}$  for some  $\mathbb{R}$ -neighborhood  $V$  of 0.

If in the  $\langle x, y \rangle$ -plane the set  $\mathbb{R}$  is represented as usual by the  $x$ -axis and if  $0'$  is represented by the point  $\langle 0, 1 \rangle$ , then a typical member of  $\mathcal{N}_{0'}$  may be represented as in Figure 2.3. The topological space obtained by applying Theorem 2.19 is called the



Figure 2.3: Neighborhood of  $0'$  in line with two origins.

fig:line-with-two-origins

**line with two origins.** This space has the interesting property that each neighborhood  $U$  of  $0'$  intersects each neighborhood  $V$  of 0, because  $U \cap V$  will contain a set of the form  $]-\varepsilon, 0[ \cup ]0, \varepsilon[$  for some  $\varepsilon > 0$ .

It is sometimes useful to represent this space a bit differently, as suggested by Figure 2.4.



Figure 2.4: Alternative depiction of the line with two origins.

fig:alt-line-with-two-origins

Namely, as a set  $X = \mathbb{R}^* \cup \{0', 0''\}$ , where  $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$  and  $0', 0'' \notin \mathbb{R}$ . In this space a neighborhood of a point  $x \in \mathbb{R}^*$  is any subset containing some  $\mathbb{R}$ -neighborhood  $V$  of  $x$  with  $0 \notin V$ ; a neighborhood of  $0'$  is any subset containing  $(V \setminus \{0\}) \cup \{0'\}$  for some  $\mathbb{R}$ -neighborhood  $V$  of 0; and a neighborhood of  $0''$  is any subset containing  $(V \setminus \{0\}) \cup \{0''\}$  for some  $\mathbb{R}$ -neighborhood  $V$  of 0.

The preceding two representations of the line with two origins—the first using  $\mathbb{R} \cup \{0'\}$  and the second  $\mathbb{R}^* \cup \{0', 0''\}$ —give essentially the same topological space: see Exercise 3.50.  $\diamond$

### Isolated points and limit points

subsec:isolated-pts-and-limit-points

In a discrete space, every singleton (single-element subset) is open. It is also possible to have a topological space in which some, but not all, singletons are open.

def:isolated-pt

**2.21 Definition.** Let  $X$  be a topological space. Given a subset  $A$  of  $X$ , a point  $x$  of  $X$  is said to be **isolated in  $A$**  and is called an **isolated point of  $A$**  when the singleton  $\{x\}$  is open in  $X$ . In particular, a point  $x$  of  $X$  is said simply to be **isolated** when it is isolated in  $X$ , in other words, when  $\{x\}$  is open in  $X$ .

isolated point  
Riesz, Frigyes  
Fréchet, Maurice  
accumulation point  
separation properties

For example, in the subspace  $X = \{0\} \cup \{1/n : n = 1, 2, 3, \dots\}$  of  $\mathbb{R}$ , each of the points  $1/n$  is isolated whereas  $0$  is not; but in  $\mathbb{R}$  itself, none of those points is isolated. The space  $\mathbb{R}$  has no isolated points whatsoever.

The relevance of isolated points to the current topic of neighborhoods is the following observation.

**2.22 Proposition.** A point  $x$  of a subset  $A$  of a topological space  $X$  is isolated in  $A$  if and only if there is some neighborhood  $U$  of  $x$  in  $X$ —and hence some open neighborhood  $U$  of  $x$ —such that  $U \cap A = \{x\}$ .

Let  $x$  be a point  $x$  in a subset  $A$  of a space  $X$ . Then  $x$  is not isolated in  $A$  if and only if each neighborhood of  $x$  in  $X$  contains some point of  $A$  other than  $x$ . This suggests the following definition.

def:limit-pt

**2.23 Definition.** Let  $X$  be a topological space. Given a subset  $A$  of  $X$ , a point  $x$  of  $X$  is said to be a **limit point of  $A$**  (**in  $X$** ) when each neighborhood of  $x$  in  $X$  contains some point of  $A$  different from  $x$ .

For example, in the subspace  $X = \{0\} \cup \{1/n : n = 1, 2, 3, \dots\}$  of  $\mathbb{R}$ , the element  $0$  is a limit point of the subset  $A = \{1/n : n = 1, 2, 3, \dots\}$ .

In general, when  $x \in A$ , then either  $x$  is a limit point of  $A$  or else  $x$  is an isolated point of  $A$ . On the other hand, when  $x \notin A$ , then  $x$  is a limit point of  $A$  if and only if each neighborhood of  $x$  in  $X$  intersects  $A$ .

The first attempt at an axiomatic treatment of spaces more general than metric spaces, made by Frigyes Riesz and Maurice Fréchet in the period 1904–1906, was based on limit points. Today limit points are out of fashion as a foundation for topology but still have their uses.

**Usage note.** Some mathematicians use the term “**accumulation point**” as a synonym for “limit point.” This usage can lead to confusion with a stronger notion, namely, “ **$\omega$ -accumulation point**” (see [Exercise 27](#)).

### Hausdorff spaces

subsec:hausdorff

We now introduce the separation property that Hausdorff took as an additional axiom for neighborhoods.

def:hausdorff-space

**2.24 Definition.** A topological space  $X$  is called a **Hausdorff space** or a  **$T_2$ -space** when each two distinct points of  $X$  have some disjoint neighborhoods.

**Usage note.** Some mathematicians regard it as a barbarism to write “ $X$  is Hausdorff” because it deploys the proper name ‘Hausdorff’ as a predicate adjective, and they insist on writing instead “ $X$  is a Hausdorff space.” (Do they allow writing “ $X$  is  $T_2$ ”?) We shall not hesitate to write both “ $X$  is Hausdorff” and “ $X$  is  $T_2$ .”

**Intuitive idea—Hausdorff space.** Loosely speaking,  $X$  is a Hausdorff space when, given any  $x, y \in X$  with  $x \neq y$ , it is impossible to find points arbitrarily close both to  $x$  and to  $y$ . Note that the condition that each two distinct points have disjoint neighborhoods is equivalent to the condition that each two distinct points have disjoint *open* neighborhoods (because each neighborhood of a point contains an open set containing that point).

**2.25 Examples.** (1) The line with two origins [Examples 2.20 \(3\)](#) is *not* a Hausdorff space, for 0 and  $z$  fail to have any disjoint neighborhoods.

(2) Every metrizable space  $X$  is a Hausdorff space. In fact, let  $d$  be a metric inducing the topology of  $X$  and let  $x, y \in X$  with  $x \neq y$ . Let  $\varepsilon = d(x, y)/2$ . Then  $B_\varepsilon(x; d)$  and  $B_\varepsilon(y; d)$  are neighborhoods of  $x$  and  $y$ , respectively, that are disjoint.

(3) Let  $L$  be the  $x$ -axis and  $H$  be the open upper half-plane in the plane  $\mathbb{R} \times \mathbb{R}$ , that is:

$$L = \mathbb{R} \times \{0\} = \{\langle x, 0 \rangle : x \in \mathbb{R}\}, \quad H = \{\langle x, y \rangle \in \mathbb{R} \times \mathbb{R} : y > 0\}.$$

Let

$$X = H \cup L = \{\langle x, y \rangle \in \mathbb{R} \times \mathbb{R} : y \geq 0\},$$

the closed upper half-plane. We shall use [Theorem 2.19](#) to construct a certain topology  $\mathcal{T}$  on  $X$ . The resulting topological space will be called the **half-disk space**.

Denote by  $d$  the usual metric on  $\mathbb{R} \times \mathbb{R}$ .

Let  $z \in H$ . Define  $\mathcal{N}_z$  to consist of all neighborhoods of  $z$  for the usual topology (induced by  $d$ ) on  $X$ . Since  $z$  is at a positive  $d$ -distance above  $L$ , each member of  $\mathcal{N}_z$  contains a  $d$ -ball  $B_\varepsilon(z; d)$ .

Now let  $z = \langle x, 0 \rangle \in L$ . For each real number  $r > 0$ , define

$$H_r(z) = \{z\} \cup (B_r(z; d) \cap H)$$

(see [Figure 2.5](#)). Define  $\mathcal{N}_z$  to consist of all subsets of  $X$  that contain the set  $H_r(z)$

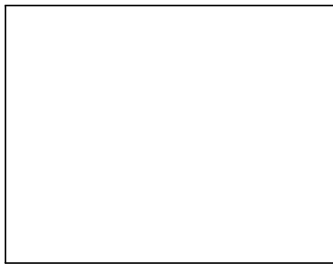


Figure 2.5: Neighborhood of  $t \in L$  in half-disk space.

fig:nbd-x-half-disk-space

for some  $r > 0$ .

The verification of ((N1))–((N5)) for these collections  $\mathcal{N}_x$  uses the corresponding properties of neighborhoods of points in the Euclidean plane. Then  $\mathcal{T}$  is the topology on  $X$  given by [Theorem 2.19](#).

We shall show that *the half-disk space is a Hausdorff space that is not metrizable*.

That  $\langle X, \mathcal{T} \rangle$  is a Hausdorff space follows from two facts. First, each  $d$ -ball at a point  $z \in X$  contains some  $\mathcal{T}$ -neighborhood of  $z$ . Second, any two distinct points of  $X$  have some disjoint  $d$ -balls about them.

half-disk space  
topological property!hereditary  
topological property!inherited by su  
Hausdorff space

Just suppose  $\langle X, \mathcal{T} \rangle$  is metrizable. Because each  $d$ -ball at a point of the plane contains a  $d$ -disk at that point, the  $\mathcal{T}$ -neighborhood  $H_1(\langle 0, 0 \rangle)$  contains some  $\mathcal{T}$ -neighborhood  $W$  of  $\langle 0, 0 \rangle$  that is  $\mathcal{T}$ -closed (see Figure 2.6). Since  $W \in \mathcal{N}_{\langle 0, 0 \rangle}$ , then  $H_r(\langle 0, 0 \rangle) \subset W$

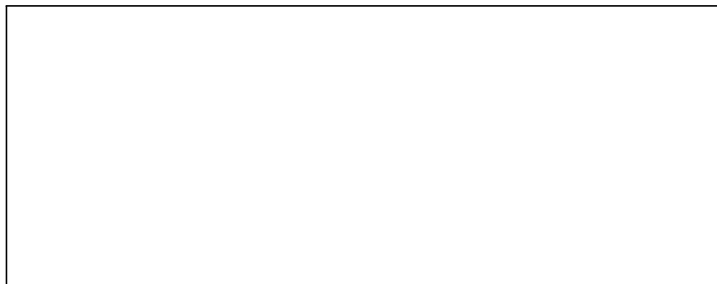


Figure 2.6: Why the half-disk space is not metrizable.

fig:half-disk-space-non-metrizable

for some  $r > 0$ . Take any  $x \in L$  with  $x \neq 0$  and  $d(x, 0) < r$ . Clearly  $x \notin H_r(\langle 0, 0 \rangle)$ . We shall obtain a contradiction by showing that  $x \in W$ .

Just suppose  $x \notin W$ . Since  $W$  is closed in  $X$ , the point  $x$  has some  $\mathcal{T}$ -neighborhood  $V$  that is disjoint from  $W$ . Choose  $s > 0$  with  $H_s(x) \subset V$ . Now  $H_s(x)$  intersects  $H_r(\langle 0, 0 \rangle)$  at points of  $H$ . Thus  $V$  intersects  $W$  after all.  $\diamond$

def:hereditary

Being a Hausdorff space is the first “topological property” encountered so far. (The precise meaning of the term ‘topological property’ is given in Definition 3.27.) One of the things we shall want to know about any topological property is whether it is **hereditary** (or as we may also say, “inherited by subspaces”)—that is, whether every subspace of a space having the property also has the same property. Being a Hausdorff space *is* a hereditary property.

prop:hausdorff-hereditary

**2.26 Proposition.** Every subspace of a Hausdorff space is itself a Hausdorff space.

**Proof.** Let  $X$  be a Hausdorff space, let  $Y \subset X$ , and let  $x, y \in Y$  with  $x \neq y$ . There are neighborhoods  $U$  and  $V$  of  $x$  and  $y$ , respectively, in  $X$  that are disjoint. Then  $U \cap Y$  and  $V \cap Y$  are disjoint, and by Proposition 2.15 these sets are neighborhoods of  $x$  and  $y$ , respectively, in  $Y$ .  $\square$

## EXERCISES FOR SECTION 2.2

**20.** Describe explicitly the neighborhood system of each point in:

- (a) an indiscrete space;
- (b) the Sierpinski space [Examples 2.3 (6)];
- (c) a set provided with its countable-complement topology (Exercise 7); and
- (d) a set provided with an included point topology (Exercise 4).

**21.** Show that the following is a necessary and sufficient condition for a set  $E$  to be closed in a topological space  $X$ : Given an arbitrary  $x \in X$ , if each neighborhood of  $x$  intersects  $E$ , then  $x \in E$ .

neighborhood!set@o2-set (a) Must the intersection of finite many neighborhoods of a point necessarily be a neighborhood of the point? the intersection of arbitrarily many neighborhoods?

isolated point (b) Must the union of finitely many neighborhoods of a point necessarily be a neighborhood of the point? the union of arbitrarily many neighborhoods?

omega-accumulation point@omega-accumulation point (c) Do analogs of the properties ((N1))–((N5)) from Theorem 2.18 hold for neighborhoods of sets?

limit point!vs.\ omega-accumulation point@vs.\ \$\omega\$-accumulation point

limit point!T1-space@in \Tone-space (d) Let  $A$  be a subset of a topological space  $X$ . Suppose  $U$  is a neighborhood of  $A$  in  $X$  and  $V$  is a neighborhood of  $A$  in  $U$ . Show that  $V$  is then a neighborhood of  $A$  in  $X$ .

derived set (e) Let  $Y$  be a topological space and let  $E$  be any set disjoint from  $Y$ . Construct a topology on  $X = Y \cup E$  that makes  $Y$ , with its given topology, a subspace of  $X$  and makes each  $x \in E$  an isolated point of  $X$ .

prob:adjoin-isolated-pts

(Thus any number of isolated points may be “adjoined” to a topological space.)

prob:limit-pt (a) Give an example of a subset  $A$  of  $\mathbb{R}$  having exactly two limit points neither of which belongs to  $A$ .

(b) For  $X = \mathbb{R}^n$ , show that  $x$  is a limit point of  $A$  if and only if each neighborhood of  $x$  contains infinitely many points of  $A$ . *Note:* See also the following Exercise 27.

(c) For an arbitrary space  $X$ , prove that  $A$  is closed in  $X$  if and only if every limit point of  $A$  belongs to  $A$ .

prob:omega-accumulation-pt (d) For a subset  $A$  of a space  $X$ , a point  $x$  of  $X$  is said to be an  $\omega$ -accumulation point of  $A$  (in  $X$ ) when each neighborhood of  $x$  in  $X$  contains infinitely many points of  $A$ .

Establish the following:

(a) In any space  $X$ , an  $\omega$ -accumulation point of a subset  $A$  is necessarily a limit point of  $A$ .

in-T1-then-omega-accumulation-pt

(b) In a  $T_1$ -space  $X$ , a limit point of a subset  $A$  of  $X$  is necessarily an  $\omega$ -accumulation point of  $A$ .

(c) In an infinite  $T_0$ -space  $X$ , a limit point of a subset  $A$  of  $X$  need not be an  $\omega$ -accumulation point of  $X$ .

*Note:* Thus in a  $T_1$ -space, a limit point is the same thing as an  $\omega$ -accumulation point; it is for this reason that some mathematicians use “accumulation point” synonymously with “limit point.” For the stronger notion of a *condensation point*, see Exercise 120.

28. Let  $E$  and  $F$  be closed subsets in a topological space  $X$  that cover  $X$ . Show that  $X$  must be a Hausdorff space if each of  $E$  and  $F$ , with its relative topology, is a Hausdorff space.

prob:derived-set-ex (e) For a subset  $A$  of a topological space  $X$ , the **derived set of  $A$  (in  $X$ )** is defined to be the set  $A'$  of all limit points of  $A$  in  $X$ .

*Note:* The notation  $A'$  for the derived set is fairly standard, even though it fails to indicate the space  $X$  to which the limit points are supposed to belong. *Caution:* This meaning of  $A'$  conflicts with that used temporarily, in the discussion on page 255, for the complement of  $A$ !

Find the derived set  $A'$  of the given subset  $A$  of the given topological space  $X$ :

(a)  $X = \mathbb{R}$  and  $A$  is the set of all irrational numbers.

(b)  $X = \mathbb{R}$ , the real line, and  $A = \{1 - 1/n : n = 2, 3, 4, \dots\}$ .

(c)  $X = \mathbb{R}^2$  and  $A$  is the union of the circles centered at the origin and of radii  $1 - 1/n$  for  $n = 2, 3, 4, \dots$

derived set

self-dense space

(d)  $X = \mathbb{C}$ , the complex plane, and

self-dense space

$$A = \{(1 - 1/n) \exp(2k\pi i/n) : n = 1, 2, 3, \dots \text{ and } k = 0, 1, 2, \dots, n-1\}.$$

dense-in-itself space

self-dense space

(e)  $A = \{1/m + 1/n : m = 1, 2, 3, \dots, n = 1, 2, 3, \dots\}$  and  $X = \mathbb{R}$ .

scattered space

prob:derived-set-properties

30. (Continuation of [Exercise 29](#).)

Let  $A$  be a subset of a space  $X$ .

(a) Show that  $A'$  is closed in  $X$ .

(b) Show that  $(A')' \subset A'$ .

Now  $B$  be a second subset of  $X$ .

(c) For subsets  $A$  and  $B$  of a given space, show that  $A' \subset B'$  if  $A \subset B$ .

(d) For subsets  $A$  and  $B$  of a given space, what is the relationship of  $(A \cup B)'$  to  $A' \cup B'$ ? of  $(A \cap B)'$  to  $A' \cap B'$ ?

prob:self-dense

31. A topological space is said to be **self-dense** (or **dense-in-itself**) when none of its points are isolated. A subset of a topological space is said to be self-dense when, with its relative topology, it is a self-dense space.

(a) Explain why the space  $\mathbb{Q}$  of rationals and the space  $\mathbb{R} \setminus \mathbb{Q}$  of irrationals are self-dense.

*Note:* A remarkable example of a self-dense and closed subspace of the real line is discussed in the subsection “The Cantor set” ([page 466](#)) of Section 4.1.

(b) Show that the union of a collection of self-dense subspaces of a topological space is again self-dense.

(c) Show that the intersection of a self-dense subspace of a space with an open subset of that space is again self-dense. Must the intersection of a self-dense subspace of a space with a closed subset of that space be self-dense?

(d) Must the closure of a self-dense subspace itself be self-dense?

(e) Express the property of being self-dense in terms of limit points.

prob:self-dense-in-complete-metric

32. (Continuation of [Exercise 31](#).) Prove that a self-dense topological space that is metrizable with a complete metric must be uncountable.

prob:scattered

33. A topological space is said to be **scattered** when it contains no nonempty self-dense subspace or, equivalently, when each nonempty subspace has an isolated point. (For example, a discrete space is necessarily scattered.)

(a) Show that a scattered space is necessarily a  $T_0$ -space but need not be a  $T_1$ -space. scattered if and only if each nonempty subspace has an isolated point.

(b) Give an example of a denumerable scattered subspace of the real line whose closure there is not scattered.

*Note:* For examples of uncountable nondiscrete scattered spaces, see [Exercise 103](#).

prob:1-pt-cptn-discrete-space

34. Let  $X$  be any set, let  $z$  be any object such that  $z \notin X$  (for example,  $z = \{X\}$ ), and let  $Y = X \cup \{z\}$ . For each  $x \in X$ , let  $\mathcal{N}_x$  be the collection of all subsets of  $Y$  that contain  $x$ ; and let  $\mathcal{N}_z$  be the collection of all subsets of  $Y$  that contain  $z$  and whose complements in  $Y$  are finite.

tangent disk space

Moore, Robert L.

Niemytzki, Viktor Vladimirovich

tangent disk space! separation properties@and separation properties

Cartesian sum! family of spaces@or a family of spaces

(a) Verify that properties (N1)–(N5) hold for these collections.

Provide  $Y$  with the topology having these collections as its neighborhood systems.(b) What is the relative topology on  $X$  in this space?(c) Show that  $Y$  is a Hausdorff space.

prob:open-closed-in-half-disk-space

35. Which of the following subsets of the half-disk space [Examples 2.25 (3)] are open? which are closed?

(a) The  $x$ -axis  $\mathbb{R} \times \{0\}$ .(b) The subset  $\{0\} \times \{y \in \mathbb{R} : y \geq 0\}$  of the  $y$ -axis.(c) The subset  $\{0\} \times \{y \in \mathbb{R} : y > 0\}$  of the  $y$ -axis.(d) The set  $[0, 1] \times \{0\}$  along the  $x$ -axis.(e) The set  $]0, 1[ \times \{0\}$  along the  $x$ -axis.36. Must a discrete subspace of a Hausdorff space  $X$  be closed in  $X$ ?

prob:tangent-disk-space

37. Let  $L$  be the  $x$ -axis, let  $H$  be the open upper half-plane, let  $d$  be the usual metric on the plane, and let  $X = H \cup L$ , the same set as in Examples 2.25 (3). For each  $z \in H$ , define  $\mathcal{N}_z$  as in that example. For each  $z = \langle x, 0 \rangle \in L$ , however, redefine  $\mathcal{N}_z$  to consist of all subsets of  $X$  that contain a set of the form

$$T_r(z) = \{z\} \cup B_r(z; d)$$

for some  $r > 0$  (draw a picture!).

(a) Verify that ((N1))–((N5)) from Theorem 2.18 hold.

The topological space  $\Gamma$  obtained with the use of Theorem 2.19 is called the **tangent disk space**. It is also known variously as the **Moore plane**, after R. L. Moore, and as the **Niemytzki plane**, after V. V. Niemytzki.(b) What is the topology induced on the subset  $L$  of  $\Gamma$ ?(c) Show that this space is a  $T_2$ -space.

prob-part:tangent-disk-sp-T2

38. If each of a family  $\langle X_i \rangle_{i \in I}$  of pairwise disjoint spaces is a Hausdorff space, then must their Cartesian sum (Exercise 18) be a Hausdorff space?

## 2.3 Boundary, Interior, and Closure

sec:bdy

With the aid of neighborhoods we are going to describe three operations that, when applied to the subsets of a topological space, yield new subsets of the space. The first two—forming the boundary and the interior of a set—are natural extensions of familiar geometric ideas to topological spaces. The third—forming the closure of a set—will be of interest in connection with limits.

### Boundary of a set

txt:annular-set

Consider the annular subset

$$A = \{x \in \mathbb{R}^2 : 1/2 \leq d(x, c) < 1\}$$

of the plane  $\mathbb{R}^2$ , where  $c$  is some point of  $\mathbb{R}^2$  and  $d$  is the Euclidean metric (see Figure 2.7).





Figure 2.7: Boundary of an annulus.

fig:annulus-pts

The set  $A$  includes all points on the circle of radius 1 centered at  $c$  but none of the points on the circle of radius 2. It is perfectly natural to refer to the set

$$B = \{x \in \mathbb{R}^2 : d(x, c) = 1/2\} \cup \{x \in \mathbb{R}^2 : d(x, c) = 1\}$$

of points on these two concentric circles as the “boundary” of  $A$ . Clearly  $B$  consists of just those points  $x \in \mathbb{R}^2$  each of whose neighborhoods intersects both the set  $A$  and its complement  $\mathbb{R}^2 \setminus A$ . Hence this example suggests a purely topological definition of “boundary” meaningful in any topological space.

**2.27 Definition.** Let  $A$  be a subset of a topological space  $X$ . Then the **boundary of  $A$**  (in  $X$ ), denoted by  $\text{bdy } A$ , is defined to be the set of all  $x \in X$  such that each neighborhood of  $x$  in  $X$  intersects both  $A$  and  $X \setminus A$ .

In symbols, the definition states that, for  $x \in X$ :

$$x \in \text{bdy } A \iff (\forall V \in \mathcal{N}_x)(V \cap A \neq \emptyset \text{ and } V \cap (X \setminus A) \neq \emptyset)$$

**Caution!** Some authors use the notation  $\partial A$  for the boundary of a subset  $A$  of a topological space  $X$ . Unfortunately, this conflicts with a quite different meaning commonly given to  $\partial A$  in the context of manifolds-with-boundary. (See [Exercise 3.97](#).)

intuit:bdy **Intuitive idea—boundary.** Intuitively, we may think of the boundary of a set  $A$  as its “edge” or “periphery.” And then we may regard points of  $\text{bdy } A$  as lying arbitrarily close both to  $A$  and to the complement of  $A$ .

To check that a point  $x$  in a metric space  $\langle X, d \rangle$  belongs to  $\text{bdy } A$ , it is enough to check that each  $d$ -ball at  $x$  intersects both  $A$  and  $X \setminus A$ . (This fact will be generalized in [Proposition 2.57](#).)

exs:bdy **2.28 Examples.**

- (1) Let  $X = \mathbb{R}$  and  $A = (0, 1]$ . Then  $\text{bdy } A = \{0, 1\}$ . In fact, for each  $\varepsilon > 0$ , the neighborhood  $(-\varepsilon, \varepsilon)$  intersects both  $A$  and  $X \setminus A = (-\infty, 0] \cup (1, \infty)$ , as does the

boundary!ball@of ball neighborhood  $(1 - \varepsilon, 1 + \varepsilon)$  of 1, so  $0 \in \text{bdy } A$  and  $1 \in \text{bdy } A$ . However, if  $x \neq 0, 1$ ,  
 sphere!boundary@as boundary then for

$$\varepsilon = \min\{|x|, |x - 1|\}$$

the neighborhood  $(x - \varepsilon, x + \varepsilon)$  of  $x$  cannot intersect both  $A$  and  $X \setminus A$ .

This example shows that a point belonging to the boundary of a set  $A$  may, but need not, belong to  $A$  itself.

(2) In any topological space  $X$ , both  $\emptyset$  and  $X$  have empty boundaries.

ex:bdy-line-in-plane (3) Let  $A = \mathbb{R} \times \{0\}$ , the  $x$ -axis in the  $opair$   $x, y$ -plane  $\mathbb{R} \times \mathbb{R}$ . Now  $A$  is a topological space in its own right (under its usual topology induced by its Euclidean metric), and by (2) the boundary of  $A$  in the space  $A$  is empty. (This example will be elucidated by [Proposition 2.54](#) and the discussion immediately following it.)

**Caution!** The boundary of a given set  $A$  depends on the particular ambient topological space under consideration that contains  $A$ .

ex:sphere-bdy-of-ball-in-Rn (4) Let  $d$  be the Euclidean metric on  $\mathbb{R}^n$ , let  $x \in \mathbb{R}^n$ , and let  $r > 0$ . Then in  $\mathbb{R}^n$ :

$$\text{bdy } B_r(x; d) = S_r(x; d)$$

[Recall that  $S_r(x; d) = \{y \in \mathbb{R}^n : d(x, y) = r\}$ , the  $d$ -sphere of radius  $r$  at  $x$ .]

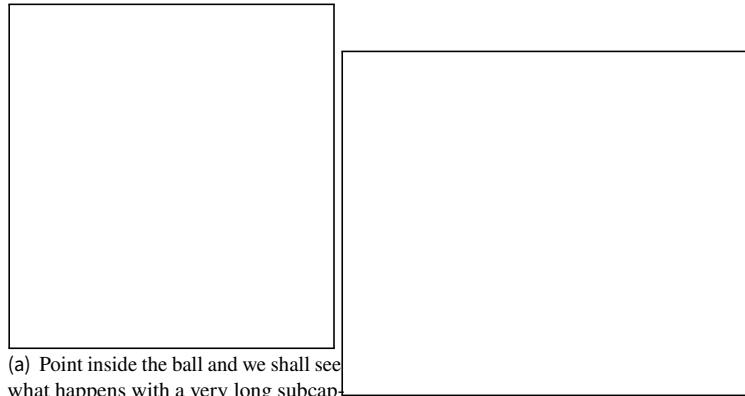
In particular, *the boundary of the  $n$ -ball  $B_n$  in Euclidean  $n$ -space  $\mathbb{R}^n$  is the  $(n - 1)$ -sphere  $S_{n-1}$ .*

To prove that  $\text{bdy } B_r(x; d) = S_r(x; d)$ , let  $y \in \mathbb{R}^n$ .

Case (i):  $d(x, y) < r$ , that is,  $y$  is in the ball  $B_r(x; d)$ . Then

$$B_\varepsilon(y; d) \subset B_r(x; d) \text{ for } \varepsilon = r - d(x, y)$$

[see [Figure 2.8\(a\)](#)].



subfig:pt-inside-ball (a) Point inside the ball and we shall see what happens with a very long subcaption. (b) Point outside the disk. subfig:pt-outside-disk

Figure 2.8: Points not in boundary of a ball in  $\mathbb{R}^n$ .

fig:pts-outside-ball-bdy

In fact,  $z \in B_\varepsilon(y; d)$  implies

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + \varepsilon = r.$$

Hence the neighborhood  $B_\varepsilon(y; d)$  of  $y$  does not intersect the complement  $\mathbb{R}^n \setminus B_r(x; d)$ , and so  $y \notin B_r(x; d)$ .

Case (ii):  $d(x, y) > r$ , that is,  $y$  is not in the disk  $D_r(x; d)$ . Then

$$B_\varepsilon(y; d) \subset \mathbb{R}^n \setminus B_r(x; d) \text{ for } \varepsilon = d(x, y) - r$$

[see Figure 2.8(b)], and so again  $y \notin B_r(x; d)$ .

Case (iii):  $d(x, y) = r$ , that is,  $y$  is on the sphere  $S_r(x; d)$ . We shall show that each  $d$ -ball at  $y$  intersects both the ball  $B_r(x; d)$  and its complement in  $\mathbb{R}^n$ , so that each neighborhood of  $y$  does meet both those sets.

Consider a ball  $B_\varepsilon(y; d)$  centered at  $y$ ; without loss of generality we may suppose that  $0 < \varepsilon < r$ . Already  $y$  is a point in  $B_\varepsilon(y; d)$  that does not belong to  $B_r(x; d)$ . It remains to find a point  $z$  in  $B_\varepsilon(y; d)$  that does belong to  $B_r(x; d)$ .

We shall, in fact, construct such a point  $z$  that lies on the line segment joining  $x$  to  $y$ . (See Figure 2.9.) Each point  $z$  on this line segment has, in vector notation, the form

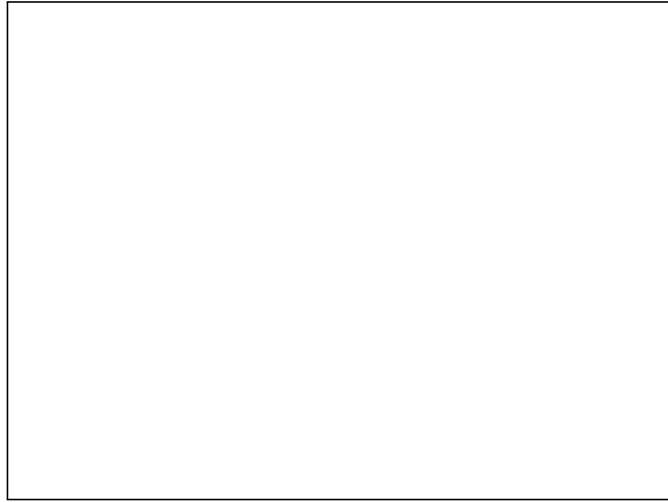


Figure 2.9: Point on boundary of ball.

fig:pt-on-ball-bdy

$$z = (1 - t)x + ty, \quad 0 \leq t \leq 1.$$

For  $z$  of this form,

$$d(x, z) = tr, \quad d(z, y) = (1 - t)r,$$

and we require a  $z$  with  $d(x, z) < r$  and  $d(z, y) < \varepsilon$ . Hence we take any  $t$  satisfying  $1 - \varepsilon/r < t < 1$ .

ex:sphere-bdy-of-disk-in-Rn

- (5) An argument similar to the one used in Example (4) proves that, for  $x \in \mathbb{R}^n$  and  $r > 0$ :

$$\text{bdy } D_r(x; d) = S_r(x; d)$$

where again  $d$  is the Euclidean metric on  $\mathbb{R}^n$ .

In particular, **the boundary of the  $n$ -disk  $D_n$  in Euclidean  $n$ -space  $\mathbb{R}^n$  is the  $(n - 1)$ -sphere  $S_{n-1}$ .**  $\diamond$

## boundary

In view of the intuitive interpretation (page 247) of “boundary” as “edge” or “periphery”, the following elementary properties should seem reasonable.

**2.29 Proposition.** *Let  $A$  be any subset of a topological space  $X$ . Then:*

- prop-part:bdy-of-complement (1)  $\text{bdy}(X \setminus A) = \text{bdy } A$ .
- prop-part:bdy-is-closed (2)  $\text{bdy } A$  is closed in  $X$ .
- prop-part:open-in-terms-of-bdy (3)  $A$  is open in  $X$  if and only if  $\text{bdy } A \subset X \setminus A$ .
- prop-part:closed-in-terms-of-bdy (4)  $A$  is closed in  $X$  if and only if  $\text{bdy } A \subset A$ .

**Proof.** (2) We show that  $X \setminus \text{bdy } A$  is open in  $X$ . Let  $x \in X \setminus \text{bdy } A$ . Then some open neighborhood  $V$  of  $x$  does not intersect both  $A$  and  $X \setminus A$ . If  $y \in V$ , then  $V$  is a neighborhood of  $y$  not intersecting both  $A$  and  $X \setminus A$ , so that  $y \notin \text{bdy } A$ . Hence  $V \subset X \setminus \text{bdy } A$ .

op-part-proof:open-in-terms-of-bdy (3) Assume first that  $A$  is open in  $X$ . We show that  $\text{bdy } A \subset X \setminus A$  by showing  $A \subset X \setminus \text{bdy } A$ . Let  $x \in A$ . Then some neighborhood  $V$  of  $x$  is contained in  $A$ , so that  $V$  does *not* intersect  $X \setminus A$ . Hence  $x \notin \text{bdy } A$ .

Conversely, assume that  $\text{bdy } A \subset X \setminus A$ . We show that  $A$  is open in  $X$ . Let  $x \in A$ . Then  $x \notin X \setminus A$ , by assumption  $x \notin \text{bdy } A$ , and so  $x$  has a neighborhood  $V$  that does not intersect both  $A$  and  $X \setminus A$ . But  $x \in V \cap A$ , so that  $V$  cannot intersect  $X \setminus A$ , that is,  $V \subset A$ .

(4) Since statement (3) is true of any subset  $A$  of  $X$ , it is true of  $X \setminus A$ ; in other words,  $X \setminus A$  is open in  $X$  if and only if  $\text{bdy}(X \setminus A) \subset A$ . Now apply (1).

[Statement (4) may also be proved directly without making any use of (3).]  $\square$

In the space  $\mathbb{Q}$  of rational numbers, the set  $A = ]\sqrt{2}, \sqrt{3}[ \cap \mathbb{Q}$  has empty boundary; and  $A$  is both open and closed in  $\mathbb{Q}$ . That is no accident.

cor:clopen-iff-empty-bdy **2.30 Corollary.** *A subset of a topological space is both open and closed if and only if its boundary is empty.*

## Interior of a set

subsec:interior

intuit:int **Intuitive idea—interior.** Just as we regard points belonging to the boundary of a set  $A$  as being arbitrarily close both to  $A$  and the complement of  $A$ , so we should think of points of  $A \setminus \text{bdy } A$  as lying entirely “inside”  $A$ , away from all points not in  $A$ ; such points constitute the “interior” of  $A$ .

def:interior

**2.31 Definition.** Let  $A$  be subset of a topological space  $X$ . Then the **interior of  $A$**  (in  $X$ ), denoted by  $\text{int } A$ , is defined to be the subset  $A \setminus \text{bdy } A$  of  $A$ . An  $x \in \text{int } A$  is said to be **interior to  $A$**  and is called an **interior point of  $A$** .

Other notations for the interior of  $A$  are  $A^\circ$  and  $A_\circ$ . The former of these two is particularly useful when combined with other topological operations on sets, such as taking the “closure” (see Definition 2.37, page 252, and the “closure-interior duality” formulas on page 255 ).

We shall often use the following characterization of interior points:

prop:interior-pt-via-nbds

**2.32 Proposition.** *Let  $A$  be subset of a topological space  $X$  and let  $x \in X$ . Then:*

$$x \in \text{int } A \iff \text{some neighborhood } x \text{ is contained in } A.$$

In symbols, the proposition states that, for  $x \in X$ :

$$x \in \text{int } A \iff (\exists V \in \mathcal{N}_x)(V \subset A)$$

**Proof.** First, assume that  $x \in \text{int } A$ . Then  $x \notin \text{bdy } A$ , and so some neighborhood  $V$  of  $x$  does not intersect both  $A$  and  $X \setminus A$ . But  $V$  does intersect  $A$  because  $x \in V \cap A$ . Hence  $V$  does not intersect  $X \setminus A$ , that is,  $V \subset A$ .

Conversely, assume there is some neighborhood  $V$  of  $x$  such that  $V \subset A$ . Then  $x \notin \text{bdy } A$  because  $V$  does not intersect  $X \setminus A$ ; also,  $x \in V \subset A$ . Hence  $x \in A \setminus \text{bdy } A = \text{int } A$ .  $\square$

exs:interiors

**2.33 Examples.** (1) For the annular subset

$$A = \{x \in \mathbb{R}^2 : 1 \leq d(x, c) < 2\}$$

of  $\mathbb{R}^2$  considered at the beginning of this section (see [page 246](#)),

$$\text{int } A = \{x \in \mathbb{R}^2 : 1 < d(x, c) < 2\}.$$

(2) In any topological space  $X$ ,

$$\text{int } \emptyset = \emptyset \quad \text{int } X = X.$$

ex:rational-interior

(3) The set  $\mathbb{Q}$  of all rational numbers has empty interior in  $\mathbb{R}$ . In fact, each neighborhood of a rational number contains irrational numbers because  $\mathbb{Q}$  is order-dense in  $\mathbb{R}$  (see [Corollary 0.81](#)). Likewise, the set  $\mathbb{R} \setminus \mathbb{Q}$  of all irrational numbers has empty interior in  $\mathbb{R}$ . (Why?)

(4) Since an open set is a neighborhood of each of its points, it follows from [Proposition 2.32](#) that the interior of an open set  $A$  of a topological space is just  $A$  itself. (See also [Corollary 2.35](#).)

In particular,

$$\text{int } B_r(x; d) = B_r(x; d)$$

for every  $r > 0$  and every point  $x$  of a metric space  $\langle X, d \rangle$ .  $\diamond$

Although the boundary of a set can be disjoint from the set, the interior of a set is always contained in that set. The next theorem makes the relationship of a set to its interior more precise.

thm:int-largest-open

**2.34 Theorem.** *Let  $A$  be a subset of a topological space  $X$ . Then  $\text{int } A$  is the largest open subset of  $X$  that is contained in  $A$ , that is:*

thm-part:int-open

(1) *The set  $\text{int } A$  is open in  $X$ , and  $\text{int } A \subset A$ .*

thm-part:int-largest-open-means

(2) *If  $U$  is open in  $X$  and  $U \subset A$ , then  $U \subset \text{int } A$ .*

**Proof.** (2) Let  $U$  be an open subset of  $X$  with  $U \subset A$ . If  $x \in U$ , then  $U$  is a neighborhood of  $x$  contained in  $A$ , so that  $x \in \text{int } A$  according to Proposition 2.32. Thus  $U \subset \text{int } A$ .

(1) Since  $\text{int } A = A \setminus \text{bdy } A$ , we have  $\text{int } A \subset A$ . To see that  $\text{int } A$  is open in  $X$ , let  $x \in \text{int } A$ . By Proposition 2.32 there is an open neighborhood  $U$  of  $x$  with  $U \subset A$ . Then  $U \subset \text{int } A$  according to part (2).  $\square$

**2.35 Corollary.** A subset  $A$  of a topological space  $X$  is open in  $X$  if and only if  $A = \text{int } A$ .

**Proof.** If  $A = \text{int } A$ , then  $A$  is open in  $X$  because  $\text{int } A$  is open in  $X$ . Conversely, if  $A$  is open in  $X$ , then  $A$  is itself the largest subset of  $A$  that is open in  $X$ , and so  $A = \text{int } A$ .  $\square$

**2.36 Corollary.** For all subsets  $A$  and  $B$  of a topological space  $X$ :

(1) If  $A \subset B$ , then  $\text{int } A \subset \text{int } B$ .

(2)  $\text{int}(A \cap B) = (\text{int } A) \cap (\text{int } B)$ .

**Proof.** (1) Assume  $A \subset B$ . Then  $\text{int } A$  is open in  $X$  and  $\text{int } A \subset A \subset B$ . Since  $\text{int } B$  is the largest open subset of  $X$  contained in  $B$ , it follows that  $\text{int } A \subset \text{int } B$ .

(2) Since  $A \cap B \subset A$ , part (1) implies that  $\text{int}(A \cap B) \subset \text{int } A$ ; similarly,  $\text{int}(A \cap B) \subset \text{int } B$ . Thus  $\text{int}(A \cap B) \subset (\text{int } A) \cap (\text{int } B)$ . The reverse inclusion is true because  $(\text{int } A) \cap (\text{int } B)$  is an open set that is contained in  $A \cap B$ .  $\square$

**Caution!** The analog of Corollary 2.36 (2) for unions is *not* true in general. For example if  $X = \mathbb{R}$ ,  $A = \mathbb{Q}$ , and  $B = \mathbb{R} \setminus \mathbb{Q}$ , then

$$\text{int}(A \cup B) = \text{int } \mathbb{R} = \mathbb{R} \neq \emptyset = (\text{int } A) \cup (\text{int } B)$$

according to Examples 2.33 (3). What is true in general is only that

$$(\text{int } A) \cup (\text{int } B) \subset \text{int}(A \cup B)$$

because  $A \subset A \cup B$  and  $B \subset A \cup B$ .

### Closure of a set

subsec:closure

We now combine the interior and boundary of a set.

**2.37 Definition.** Let  $A$  be a subset of a topological space  $X$ . Then the **closure of  $A$**  (in  $X$ ), denoted by  $\text{cls } A$ , is the subset  $(\text{int } A) \cup (\text{bdy } A)$  of  $X$ .

Other notations for the closure of  $A$  are  $A^-$  and  $\bar{A}$ . The former of these two is particularly useful when combined with other topological operations on sets, such as taking the interior. (See, for example, the “closure-interior duality” formulas on page 255.)

Examples 2.28 and Examples 2.33 allow us to compute at once the closures of some sets.

exs:closure **2.38 Examples.** (1) For the annular subset

$$A = \{x \in \mathbb{R}^2 : 1 \leq d(c, c) < 2\}$$

of  $\mathbb{R}^2$  with its Euclidean metric we have

$$\text{cls } A = \{x \in \mathbb{R}^2 : 1 \leq d(x, c) \leq 2\}.$$

(2) In any topological space  $X$ ,

$$\text{cls } \emptyset = \emptyset, \quad \text{cls } X = X.$$

ex:cls-of-Q (3) In the real line  $\mathbb{R}$ ,

$$\text{cls } \mathbb{Q} = \mathbb{R} = \text{cls}(\mathbb{R} \setminus \mathbb{Q}).$$

(4) Let  $d$  be the Euclidean metric on  $\mathbb{R}^n$  let  $x \in \mathbb{R}^n$ , and let  $r > 0$ . Then in  $\mathbb{R}^n$ ,

$$\text{cls } B_r(x; d) = D_r(x; d),$$

the  $d$ -disk of radius  $r$  at  $x$ .

(5) If  $A$  is a subset of a topological space  $X$ , then a point  $x$  is a limit point of  $A$  in  $X$  exactly when  $x \in \text{cls}(A \setminus \{x\})$ .  $\diamond$

intuit:cls **Intuitive idea—closure.** Intuitively, the boundary of a set  $A$  consists of all points that are arbitrarily close both to  $A$  and to the complement of  $A$  (page 247); and the interior of  $A$  consists of all points in  $A$  that are *not* arbitrarily close to the complement of  $A$  (page 250). Hence it is intuitively reasonable that *the closure of  $A$  consists of all points that are arbitrarily close to  $A$ .*

According to the following proposition, the preceding intuition about the closure is correct.

prop:closure-in-terms-of-nbhds **2.39 Proposition.** Let  $A$  be a subset of a topological space  $X$  and let  $x$  be a point of  $X$ . Then  $x \in \text{cls } A$  if and only if each neighborhood of  $x$  intersects  $A$ .

In symbols, the proposition states that, for  $x \in X$ :

$$x \in \text{cls } A \iff (\forall V \in \mathcal{N}_x)(V \cap A \neq \emptyset)$$

**Proof.** Assume first that  $x \in \text{cls } A$ . If  $x \in \text{int } A$ , then  $x \in A$ , so that each neighborhood of  $x$  intersects  $A$  at least at the point  $x$ . If, on the other hand,  $x \in \text{bdy } A$ , then each neighborhood of  $x$  intersects  $A$  (as well as  $X \setminus A$ ).

Conversely, assume that each neighborhood of  $x$  intersects  $A$ . If some neighborhood of  $x$  is actually contained in  $A$ , then  $x \in \text{int } A \subset \text{cls } A$ . Otherwise, each neighborhood of  $x$  must intersect  $X \setminus A$  as well as  $A$ , so that  $x \in \text{bdy } A \subset \text{cls } A$ .  $\square$

ex:infy-in-cls **2.40 Example.** Let  $A \subset \mathbb{R}$ . Denote by  $\text{cls } A$  the closure of  $A$  not in  $\mathbb{R}$  but rather in the extended real line  $\widehat{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$  (see Example 1.41).

It readily follows from the preceding Proposition 2.39 and Lemma 1.43 that:

$$+\infty \in \text{cls } A \iff A \text{ is not bounded above in } \mathbb{R},$$

$$-\infty \in \text{cls } A \iff A \text{ is not bounded below in } \mathbb{R}.$$

These facts will be used in Chapter 3 (Continuity and Convergence) in the discussion of “limits at infinity.”  $\diamond$

For a metric space, the characterization in [Proposition 2.39](#) of the closure of  $A$  can be stated more concretely in terms of the distance of points from  $A$  (see [Definition 1.28](#)).

discrete space in discrete space

**2.41 Corollary.** Let  $A$  be a subset of the metric space  $\langle X, d \rangle$ . Then

$$\text{cls } A = \{x \in X : d(x, A) = 0\}$$

Let  $x \in X$  and assume  $d(x, A) = \delta > 0$ . Then  $d(x, a) \geq \delta$  for all  $a \in A$ , so that

the  $d$ -ball  $B_{\delta/2}(x; d)$  does not intersect  $A$ , and hence  $x \notin \text{cls } A$ .

Conversely, let  $x \in X$  and assume  $x \notin \text{cls } A$ . Then some  $d$ -ball  $B_\varepsilon(x; d)$  at  $x$  does not intersect  $A$ , which means that  $d(x, a) \geq \varepsilon$  for all  $a \in A$ , and hence  $d(x, A) \geq \varepsilon > 0$ .  $\square$

For an application of [Proposition 2.39](#) let us generalize the notion of a  $d$ -dense set in a metric space as given in [Definition 1.89](#).

**2.42 Definition.** A subset  $D$  of a topological space  $X$  is said to be **dense in  $X$**  when  $D$  intersects each nonempty open subset of  $X$ .

Evidently  $D$  is dense in  $X$  if and only if each neighborhood of each point of  $X$  intersects  $D$ . Then according to [Proposition 2.39](#),

$$D \text{ is dense in } X \iff \text{cls } D = X.$$

**Intuitive idea—dense set.** Thus, roughly speaking, a subset  $D$  of a space  $X$  is dense in  $X$  precisely when there are points of  $D$  arbitrarily close to each point of  $X$ .

**2.43 Examples.** (1) [Examples 2.38 \(3\)](#) says that the set  $\mathbb{Q}$  of all rational numbers and the set  $\mathbb{R} \setminus \mathbb{Q}$  of all irrational numbers are both dense in the real line  $\mathbb{R}$ .

(2) The only subset of a discrete space  $X$  that is dense in  $X$  is the entire space  $X$ .

(3) In the space  $[0, 1]$  each of the following sets is dense:  $]0, 1[$ ,  $[0, 1[$ ,  $]0, 1/2[ \cup ]1/2, 1[$ ,  $]0, 1[ \setminus \{1/n : n = 1, 2, 3, \dots\}$ .

(4) By [Example 2.40](#), the real line  $\mathbb{R}$  is dense in the extended real line  $\widehat{\mathbb{R}}$ .  $\diamond$

[Proposition 2.39](#) allows us to express the boundary of a set  $A$  in a space  $X$  in terms of closures, namely,

$$\text{bdy } A = (\text{cls } A) \cap \text{cls}(X \setminus A),$$

and consequently,

$$\text{bdy } A = \text{bdy}(X \setminus A)$$

Moreover, it allows us to express closures in terms of interiors, and vice versa.

**2.44 Theorem (closure-interior duality).** For each subset  $A$  of a topological space  $X$ :

(1)  $X \setminus \text{cls } A = \text{int}(X \setminus A)$

(2)  $X \setminus \text{int } A = \text{cls}(X \setminus A)$



**Proof.** (1) Let  $x \in X$ . Then  $x \in X \setminus \text{cls } A$  if and only if some neighborhood  $V$  of  $x$  does not intersect  $A$ , that is,  $V \subset X \setminus A$ . But  $x \in \text{int}(X \setminus A)$  if and only if some neighborhood of  $x$  is contained in  $X \setminus A$ .

(2) First, replace  $A$  by  $X \setminus A$  in (1)—this is legitimate because (1) holds for every subset of  $X$ —to obtain

$$X \setminus \text{cls}(X \setminus A) = \text{int}(X \setminus (X \setminus A)) = \text{int } A.$$

Now take the complements in  $X$  of both sides of the preceding equation.  $\square$

Since  $\text{bdy } A = (\text{cls } A) \cap \text{cls}(X \setminus A)$ , equation (2) above yields the following result.

cor:bdy-as-cls-less-int

**2.45 Corollary.** For every subset  $A$  of a topological space  $X$ ,

$$\text{bdy } A = (\text{cls } A) \setminus (\text{int } A).$$

**Theorem 2.44** can be put into an easily remembered form through the use of superscripts  $^-$ ,  $^\circ$ , and  $'$  to denote closure, interior, and complement in  $X$ , respectively:

$$A^{-'} = A'^\circ, \quad A^{\circ'} = A'^- \quad (\text{closure-interior duality})$$

Thus as working rules of calculation: *interchange  $^-$  and  $'$  and then replace  $^-$  by  $^\circ$ ; and interchange  $^\circ$  and  $'$  and then replace  $^\circ$  by  $^-$ .*

By taking the complements in  $X$  of both sides of equations (1) and (2) in **Theorem 2.44**, we obtain equations (1) and (2), respectively, in the following corollary.

cor:cls-int-duality

**2.46 Corollary.** For every subset  $A$  of a topological space  $X$ :

cor-part:cls-as-int-compl

$$(1) \text{ cls } A = X \setminus \text{int}(X \setminus A)$$

cor-part:int-as-cls-compl

$$(2) \text{ int } A = X \setminus \text{cls}(X \setminus A)$$

In the superscript notation, the equations of the preceding corollary are:

$$A^- = A'^{\circ'}, \quad A^\circ = A'^{-'}$$

**Theorem 2.44** and **Corollary 2.46** allow us to derive from each property of interiors a corresponding “dual” property of closures, and vice versa. For example, if  $A$  is a subset of a space  $X$ , then **Theorem 2.34** says that  $\text{int}(X \setminus A)$  is open in  $X$  and is contained in  $X \setminus A$ ; then by **Corollary 2.46** (1), we have  $\text{cls } A = \text{int}(X \setminus A)$  is closed in  $X$  and contains  $A$ . This establishes the first part of the following dual of **Theorem 2.34**.

thm:cls-smallest-closed

**2.47 Theorem.** Let  $A$  be a subset of a topological space  $X$ . Then  $\text{cls } A$  is the smallest closed subset of  $X$  that contains  $A$ , that is:

(1) The set  $\text{cls } A$  is closed in  $X$ , and  $A \subset \text{cls } A$ .

(2) If  $E$  is closed in  $X$  and  $A \subset E$ , then  $\text{cls } A \subset E$ .

**Proof.** (2) Let  $E$  be a closed subset of  $X$  with  $A \subset E$ . Then  $X \setminus E$  is an open subset of  $X$  with  $X \setminus E \subset X \setminus A$ . By [Theorem 2.34 \(2\)](#),

$$X \setminus E \subset \text{int}(X \setminus A).$$

Then by [Corollary 2.46 \(1\)](#),

$$\text{cls } A = X \setminus \text{int}(X \setminus A) \subset X \setminus (X \setminus E) = E. \quad \square$$

cor:closed-iff-equals-closure

**2.48 Corollary.** A subset  $A$  of a topological space  $X$  is closed in  $X$  if and only if  $A = \text{cls } A$ .

**Proof.** If  $A = \text{cls } A$ , then  $A$  is closed in  $X$  because  $\text{cls } A$  is. Conversely, if  $A$  is closed in  $X$ , then  $A$  is itself the smallest closed subset of  $X$  that contains  $A$ , and so  $A = \text{cls } A$ .  $\square$

Of course [Corollary 2.48](#) can also be deduced from [Corollary 2.35](#) by using closure-interior duality ([Corollary 2.46](#)).

Perhaps you have already anticipated the following dual of [Corollary 2.36](#).

cor:closure-and-inclusion-and-union

**2.49 Corollary.** For all subsets  $A$  and  $B$  of a topological space  $X$ :

cor-part:subset-preserve-closure

(1) If  $B \subset A$ , then  $\text{cls } B \subset \text{cls } A$ .

cor-part:closure-preserve-union

(2)  $\text{cls}(A \cup B) = (\text{cls } A) \cup (\text{cls } B)$ .

**Proof.** (2) From [Corollary 2.46 \(1\)](#), De Morgan's Laws ([page 14](#)), and [Corollary 2.36 \(2\)](#) we obtain

$$\begin{aligned} \text{cls}(A \cup B) &= X \setminus \text{int}(X \setminus (A \cup B)) \\ &= X \setminus \text{int}((X \setminus A) \cap (X \setminus B)) \\ &= X \setminus [(\text{int}(X \setminus A)) \cap (\text{int}(X \setminus B))] \\ &= (X \setminus \text{int}(X \setminus A)) \cup (X \setminus \text{int}(X \setminus B)) \\ &= (\text{cls } A) \cup (\text{cls } B). \quad \square \end{aligned}$$

closure!intersection@and intersection

**Caution!** The analog of (2) for intersection is *not* true in general. For example, if  $A = \mathbb{R} \setminus \mathbb{Q}$  and  $B = \mathbb{Q}$  in the space  $\mathbb{R}$ , then

$$\text{cls}(A \cap B) = \text{cls } \emptyset = \emptyset \neq \mathbb{R} = \mathbb{R} \cap \mathbb{R} = (\text{cls } A) \cap (\text{cls } B).$$

What is true in general is that

$$\text{cls}(A \cap B) \subset (\text{cls } A) \cap (\text{cls } B)$$

for all subsets  $A$  and  $B$  of a topological space. (Why?)

subsec:nowhere-dense

### Nowhere dense sets

A set  $A$  is dense in a space  $X$  when its closure is the largest open subset of  $X$ —namely,  $X$  itself. At the opposite extreme, the closure of  $A$  may contain no open subset of  $X$  whatsoever except the empty one.

def:nowhere-dense

**2.50 Definition.** A subset  $A$  of a topological space  $X$  is said to be **nowhere dense in**  $X$  when  $\text{cls } A$  contains no nonempty open subset of  $X$ ; that is, when  $\text{int}(\text{cls } A) = \emptyset$ . A synonym for ‘nowhere dense’ is **rare**.

Cantor set  
Baire Category Theorem  
Baire, René-Louis

For example, each finite subset of  $\mathbb{R}^n$  is nowhere dense in  $\mathbb{R}^n$ . The infinite set  $\{1/n : n = 1, 2, 3, \dots\}$  is nowhere dense in  $\mathbb{R}$ . In Example 4.16 we shall encounter a remarkable subset of the unit interval  $[0, 1]$  that is nowhere dense in that interval—the *Cantor set*.

By closure-interior duality [Theorem 2.44 (2)], a set  $A$  is nowhere dense in a space  $X$  if and only if

$$X = X \setminus \emptyset = X \setminus \text{int}(\text{cls } A) = \text{cls}(X \setminus \text{cls } A),$$

in other words, if and only if  $X \setminus \text{cls } A$  is dense in  $X$ . In particular,  $A$  is a nowhere dense closed set in  $X$  if and only if its complement  $X \setminus A$  is a dense open set in  $X$ . Dually,  $A$  is a dense open set in  $X$  if and only if  $X \setminus A$  is a nowhere dense closed set in  $X$ .

The preceding observations suggest equivalent statements of the Baire Category Theorem (1.91), whose topological conclusion concerned dense open sets. For these statements, Baire introduced the following classification of subsets  $B$  of a topological space  $X$  into two types:  $B$  is **of first-category in**  $X$  if  $B$  is the union of some sequence of sets that are nowhere dense in  $X$ ; and  $B$  is **of second-category in**  $X$  if it is *not* first-category in  $X$ . Unfortunately, the terms ‘first-category’ and ‘second-category’ are not suggestive of their meanings, and so we prefer the following more modern terminology.

term:category

def:meager-set

**2.51 Definition.** Let  $B$  be a subset of a topological space  $X$ . Then  $B$  is said to be **meager (in**  $X)$  if  $B$  is the union of some sequence of subsets of  $X$  each of which is nowhere dense in  $X$ . On the other hand,  $B$  is said to be **nonmeager (in**  $X)$  if  $B$  is *not* meager in  $X$ .

**Intuitive idea—meager and nonmeager sets.** We should think of a meager subset of  $X$  as having its points sparsely scattered through the space and hence as being “thin” or “negligible.” Then we should think of a nonmeager subset of  $X$  as being “thick” or “nonnegligible.”

prop:meager-characterize

**2.52 Proposition.** For a topological space  $X$  the following conditions are equivalent:

prop-part:intersect-dense-opens

(i) The intersection of each sequence of dense open subsets of  $X$  is dense in  $X$ .

prop-part:meager-has-empty-int

(ii) Each meager subset of  $X$  has empty interior.

prop-part:open-is-nonmeager

(iii) Each nonempty open subset of  $X$  is nonmeager in  $X$ .

**Proof.** Assume (i). We deduce (ii). Let  $B$  be a meager subset of  $X$ , so that

$$B = \bigcup_{n=0}^{\infty} A_n$$

for some sequence  $\langle A_n \rangle_{n \in \mathbb{N}}$  of nowhere dense subsets of  $X$ . For each  $n \in \mathbb{N}$  the set

$$D_n = X \setminus \text{cls } A_n$$

is dense and open in  $X$ . By (i),

$$\text{cls} \left( \bigcap_{n=0}^{\infty} D_n \right) = X.$$

Now

$$D_n \subset X \setminus A_n \quad (n \in \mathbb{N}),$$

so also

$$\text{cls} \left[ \bigcap_{n=0}^{\infty} (X \setminus A_n) \right] = X.$$

Using closure-interior duality [Theorem 2.44 (1)] we conclude that

$$\begin{aligned} \text{int } B &= \text{int} \left( \bigcup_{n=0}^{\infty} A_n \right) = \text{int} \left( X \setminus \bigcap_{n=0}^{\infty} [X \setminus A_n] \right) \\ &= X \setminus \text{cls} \bigcap_{n=0}^{\infty} [X \setminus A_n] = \emptyset. \end{aligned}$$

Similarly, (ii) implies (i). That (ii) is equivalent to (iii) is easy to see.  $\square$

### Boundary, interior, and closure in a subspace

subsec:bdy-int-cls-subspace

Already in Examples 2.28 (3) we observed that a given set may have different boundaries in a space and in a subspace of that space; the same is true of closures and interiors. Consequently, we need to enhance the notations for boundary, closure, and interior so as to specify the space in which these operations are to be taken.

**2.53 Notation.** Let  $A$  be a subset of a subspace  $Y$  of a topological space  $X$ . Then  $\text{bdy}_Y A$ ,  $\text{cls}_Y A$ , and  $\text{int}_Y A$  denote the boundary, closure, and interior, respectively, of  $A$  in the topological space  $Y$ .

In particular,  $\text{bdy}_X A$ ,  $\text{cls}_X A$ , and  $\text{int}_X A$  are what we earlier denoted by  $\text{bdy } A$ ,  $\text{cls } A$ , and  $\text{int } A$ , respectively, when the ambient space  $X$  was understood.

For an arbitrary topological space  $X$ , we have  $\text{cls } X = X$ ,  $\text{int } X = X$ , and  $\text{bdy } X = \emptyset$ .

The next proposition clarifies the relationships between the boundary, closure, and interior of a set in a subspace, on the one hand, and the boundary, closure, and interior of that same set in the larger space, on the other hand.

prop:bdy-relative

**2.54 Proposition.** Let  $A$  be a subset of a subspace  $Y$  of a topological space  $X$ . Then:

prop-part:cls-in-subspace

$$(1) \quad \text{cls}_Y A = Y \cap \text{cls}_X A$$

prop-part:int-in-subspace

$$(2) \quad \text{int}_Y A \supset Y \cap \text{int}_X A$$

prop-part:bdy-in-subspace

$$(3) \quad \text{bdy}_Y A \subset Y \cap \text{bdy}_X A$$

**Proof.** (1) In view of Theorem 2.47 we need only show that  $Y \cap \text{cls}_X A$  is the smallest subset of  $X$  that contains both  $A$  and that is closed in  $Y$ .

Since  $A \subset \text{cls}_X A$  and  $\text{cls}_X A$  is closed in  $X$ , then  $A \subset Y \cap \text{cls}_X A$  and  $Y \cap \text{cls}_X A$  is closed in  $Y$ . Now let  $E$  be any subset of  $Y$  such that  $A \subset E$  and  $E$  closed in  $Y$ . Write  $E = Y \cap F$  with  $F$  closed in  $X$ . Then  $\text{cls}_X A \subset F$  because  $A \subset F$ . Hence  $Y \cap \text{cls}_X A \subset Y \cap F = E$ .

(2) The proof of (2) is similar to the above but employs Theorem 2.34.

(3) Statement (3) follows from (1) and (2):

$$\begin{aligned}
 \text{bdy}_Y A &= (\text{cls}_Y A) \setminus (\text{int}_Y A) \\
 &\subset (\text{cls}_Y A) \setminus (Y \cap \text{int}_X A) \\
 &= (Y \cap \text{cls}_X A) \setminus (Y \cap \text{int}_X A) \\
 &= Y \cap ((\text{cls}_X A) \setminus (\text{int}_X A)) \\
 &= Y \cap \text{bdy}_X A. \quad \square
 \end{aligned}$$

In general, the inclusions (2) and (3) above cannot be strengthened to equalities. For example, take

$$X = \mathbb{R}^2, \quad A = Y = \mathbb{R} \times \{0\}.$$

Then

$$\text{int}_X A = \emptyset, \quad \text{bdy}_X A = A,$$

but

$$\text{int}_Y A = A, \quad \text{bdy}_Y A = \emptyset.$$

Thus if  $Y$  is a subspace of a topological space  $X$ , then the boundary, closure or interior of  $Y$  in itself need not be same as the boundary, closure, or interior, respectively, of  $Y$  in the ambient space  $X$ .

### EXERCISES FOR SECTION 2.3

**39.** Find the boundary, interior, and closure of the given set  $A$  in the given space  $X$ :

- (a)  $A = \{1/n : n = 1, 2, 3, \dots\}$ ,  $X = \mathbb{R}$ .
- (b)  $A = \{0\} \cup \{1/n : n = 1, 2, 3, \dots\}$ ,  $X = \mathbb{R}$ .
- (c)  $A = \mathbb{N}$ ,  $X = \mathbb{R}$ .
- (d)  $A = \mathbb{R} \setminus \mathbb{N}$ ,  $X = \mathbb{R}$ .
- (e)  $A = \{(x, y) \in \mathbb{R}^2 : (x^2/a^2) + (y^2/b^2) \leq 1\}$ , where  $a$  and  $b$  are positive real numbers,  $X = \mathbb{R}^2$ .
- (f)  $A = \{(x, y) \in \mathbb{R}^2 : x > 0, 0 < y \leq 1/x\}$ ,  $X = \mathbb{R}^2$ .
- (g)  $A = \{(x, y) \in \mathbb{R}^2 : 0 < x^2 + y^2 \leq 1\}$ ,  $X = \mathbb{R}^2$ .
- (h)  $A = \{(x, y) \in \mathbb{R}^2 : 0 < x \leq 1, y = \sin(1/x)\}$ ,  $X = \mathbb{R}^2$ .
- (i)  $A = \{(x, y) \in \mathbb{R}^2 : 0 < x \leq 1, y = x \sin(1/x)\}$ ,  $X = \mathbb{R}^2$ .
- (j)  $A = \{0\}$ ,  $X$  = the Sierpinski space [Examples 2.3 (6)].
- (k)  $A = ]0, 1]$ ,  $X = \mathbb{R}_I$ , the Sorgenfrey line [Examples 2.20 (1)].
- (l)  $A = ]0, 1]$ ,  $X$  = the line with two origins [Examples 2.20 (3)].
- (m)  $A = H$  (the open upper half-plane),  $X$  = the half-disk space [Examples 2.25 (3)].
- (n)  $A = H$  (the open upper half-plane),  $X$  = the tangent disk space  $\Gamma$  (Exercise 37).

**40.** Let  $A$  and  $B$  be subsets of a space  $X$  with  $A \subset B$ . We know that  $\text{int } A \subset \text{int } B$  and  $\text{cls } A \subset \text{cls } B$ . But is  $\text{bdy } A \subset \text{bdy } B$ ?

**41.** Prove that

$$\text{bdy } A = (\text{cls } A) \cap \text{cls}(X \setminus A)$$

for an arbitrary subset  $A$  of a topological space  $X$ .

boundary!closure@and closure  
closure!boundary@and boundary

boundary!closure@and closure 42. Prove that the inclusions

closure!boundary@and boundary

$$\text{bdy}(\text{cls } A) \subset \text{bdy } A \text{ and } \text{bdy}(\text{int } A) \subset \text{bdy } A$$

interior!boundary@and boundary

boundary!interior@and interior hold for an arbitrary subset  $A$  of a topological space  $X$ , but that equality need not hold in either.

limit point!closure@and closure

closure!limit point@and limit point

43. What is the relationship between  $\text{bdy}(\text{bdy } A)$  and  $\text{bdy } A$ ?

derived set!closure@and closure

closure!derived set@and derived set

44. For the Euclidean metric  $d$  on  $X = \mathbb{R}^n$  we know that, in  $\mathbb{R}^n$ ,

self-dense space!dense set@and dense set

$$\text{bdy } B_r(x; d) = S_r(x; d), \quad \text{cls } B_r(x; d) = D_r(x; d)$$

dense set!self-dense space@and self-dense space

for each  $x \in X$  and each  $r > 0$ .

(a) Show that these equations do not hold in general for an arbitrary metric space  $\langle X, d \rangle$ .

(b) Do they hold for  $X = \mathbb{R}^n$  and  $d$  an arbitrary metric inducing the usual topology on  $X$ ?

45. Prove [Corollary 2.30](#): A subset of a topological space is both open and closed if and only if its boundary is empty.

46. Let  $A$  be a nonempty subset of  $\mathbb{R}$  having both an upper bound and a lower bound in  $\mathbb{R}$ .

(a) Show that  $\sup A \in \text{bdy } A$  and  $\inf A \in \text{bdy } A$ .

(b) From (a) we obtain:  $\sup A \in A$  and  $\inf A \in A$  if  $A$  is closed in  $\mathbb{R}$ . Does the converse hold?

47. Prove that  $X$  is a Hausdorff space if and only if for each  $x \in X$  and each  $y \in X$  with  $y \neq x$ , there is a neighborhood  $V$  of  $x$  such that  $y \notin \text{cls } V$ .

prob:closure-limit-pts

48. For a subset  $A$  of a topological space  $X$ , show that  $\text{cls } A = A \cup A'$ , where  $A'$  is the derived set of  $A$  in  $X$ , that is, the set of all limit points of  $A$  in  $X$ . (See [Exercise 30](#).)

prob:diameter-and-cls

49. Let  $A$  be a  $d$ -bounded set in a metric space  $\langle X, d \rangle$ .

(a) Prove: The set  $\text{cls } A$  is also  $d$ -bounded, and  $\text{diam}(\text{cls } A) = \text{diam } A$ .

(b) Discuss the analogous statement for  $\text{int } A$ .

50. If  $B$  is a subset of  $\mathbb{R}^n$  and if  $A \subset B$ , must  $d(x, \text{bdy } A) \leq d(x, \text{bdy } B)$  for all  $x \in A$ ?

51. (a) If  $A \subset B \subset \text{cls } A$ , what can be said about  $\text{cls } B$ ?

(b) What can be said about a closed subset  $E$  of  $X$  when  $E$  contains a dense subset of  $X$ .

52. Bob claims that the boundary of the 2-disk  $D_2$  is the enclosing circle  $S_1$ ; Alice claims that the boundary is not that. Who is right, Alice or Bob?

53. Prove or disprove: If  $Y$  is a dense subspace of a topological space  $X$  and if  $D \subset Y$  is dense in  $Y$ , then  $D$  is dense in  $X$ .

54. Let  $D_1$  and  $D_2$  be two dense sets in a topological space  $X$ . Show that  $D_1 \cap D_2$  need not be dense in  $X$  but will be in case  $D_1$  or  $D_2$  is also open in  $X$ .

55. Given an arbitrary topological space  $X$ , then the entire set  $X$  is dense in  $X$  in the sense of [Definition 2.42](#). Show, however, that a self-dense subset of a topological space need not be dense in that space.

prob:perfect

**56.** A subset of a topological space is said to be **perfect** when it is self-dense (Exercise 31) and closed in  $X$ .

- (a) Show that a closed interval  $[a, b]$  in  $\mathbb{R}$  is perfect.
- (b) Show that a subset  $A$  of a space  $X$  is perfect if and only if  $A$  is the set of all limit points of  $A$  in  $X$ .
- (c) Show that the closure of a self-dense subset of a topological space is necessarily perfect.
- (d) Prove that a perfect subset of  $\mathbb{R}$  must be uncountable.
- (e) Prove, more generally than (d), that a perfect subset of a complete metric space must be uncountable.

*Note:* The Cantor set, already mentioned on page 257 as a nowhere dense subset of the unit interval, is also a perfect subset of that interval: see Example 4.16.

prob-part:perfect-in-R-uncountable

prob-partition-space-perfect-scattered

**57.** Prove that an arbitrary topological space has a unique partition into a subspace that is perfect (Exercise 56) and a subspace that is scattered (Exercise 33).

**58.** Let  $K \subset \mathbb{R}^n$  with  $K$  convex (page 37).

- (a) Show that  $\text{cls } A$  must also be convex. Are  $\text{bdy } A$  and  $\text{int } K$  necessarily convex?
- (b) Suppose  $K$  is also closed in  $\mathbb{R}^n$ . An **extreme point** of  $K$  is a point  $x \in K$  that does *not* have the form

$$x = (1 - t)a + tb$$

for any  $0 < t < 1$  and any  $a, b \in K$ . Prove that  $\text{bdy } K$  contains the set of all extreme points of  $K$  but need not equal this set.

prob:limit-pt-and-cls

**59.** Let  $A$  be a subset of a topological space  $X$ .

- (a) How can ‘ $x$  is a limit point of  $A$ ’ (Exercise 26) be phrased in terms of closures?
- (b) What is the relationship between  $\text{cls } A$  and the set  $D$  of all limit points of  $A$  in  $X$ ?

prob:exterior

**60.** The **exterior** of a set  $A$  in a topological space  $X$  is the subset  $\text{ext } A$  of  $X$  defined as  $\text{ext } A = \text{int}(X \setminus A)$ . Show that the three sets  $A$ ,  $\text{bdy } A$ , and  $\text{ext } A$  are always pairwise disjoint and have the entire space  $X$  as their union.

s-is-intersection-of-closed-supersets

**61.** (a) Prove: For a subset  $A$  of a topological space  $X$ ,

$$\text{cls } A = \bigcap \{E : A \subset E \subset X, E \text{ closed in } X\}.$$

- (b) Formulate and prove the dual of (a) for  $\text{int } A$ .

prob:cls-part-with-union-of-intersections

**62.** (a) Prove that

$$\text{cls}(\bigcup_{i \in I} A_i) = \bigcup_{i \in I} \text{cls } A_i$$

for any nonempty *finite* family  $\langle A_i \rangle_{i \in I}$  of subsets of a space  $X$ , thereby generalizing Corollary 2.49 (2). Show, however, that this equation no longer remains true for an infinite—even a denumerable—family of subsets.

part:closure-of-intersection-arbitrary

- (b) Prove that

$$\text{cls}(\bigcap_{i \in I} A_i) \subset \bigcap_{i \in I} \text{cls } A_i$$

for an arbitrary family  $\langle A_i \rangle_{i \in I}$  of subsets of  $X$ .

- (c) Discuss the analogs of (a) and (b) for interiors.

perfect set

self-dense space

Cantor set

perfect space

scattered space

convex set

limit point

closure!limit point@and!limit point

exterior

closure!intersection@and intersection

closure!union@and union

closure!intersection@and intersection

interior!union@and union

interior!intersection@and intersection

prob:locally-finite-family 63. A family  $\langle A_i \rangle_{i \in I}$  of subsets of a topological space  $X$  is said to be **locally finite** when each  $x \in X$  has some neighborhood that intersects  $A_i$  for only finitely many (and possibly no) values of  $i \in I$ .

closure!union@and union  
nowhere dense set  
residual set

closure!axioms@and axioms

Kuratowski closure axioms

Kuratowski, Kazimierz

(a) Verify that if  $A_i = [i, \infty)$  for each  $i \in \mathbb{N}$ , then the family  $\langle A_i \rangle_{i \in \mathbb{N}}$  of subsets of  $\mathbb{R}$  is locally finite even though, for each  $j \in \mathbb{N}$ , the set  $A_j$  intersects  $A_i$  for infinitely many  $i$ .

(b) Construct a family  $\langle A_i \rangle_{i \in I}$  of subsets of  $\mathbb{R}$  that is not locally finite but such that each point of  $\mathbb{R}$  belongs to  $A_i$  for only finitely many  $i \in I$ .

(c) If  $\langle A_i \rangle_{i \in I}$  is a locally finite family of subsets of a space  $X$ , show that the family  $\langle \text{cls } A_i \rangle_{i \in I}$  is also locally finite

(d) Generalize Exercise 62 (a) by proving that

$$\text{cls}(\cup_{i \in I} A_i) = \cup_{i \in I} \text{cls } A_i$$

whenever the family  $\langle A_i \rangle_{i \in I}$  is locally finite.

prob:Rn-open-with-finite-set-closed

64. Let  $A$  be a nonempty open subset of  $\mathbb{R}^n$  such that  $A \cup F$  is closed in  $\mathbb{R}^n$  for some nonempty finite subset  $F$ . Show that  $A$  is dense in  $\mathbb{R}^n$  if  $n \geq 2$  but that it need not be dense if  $n = 1$ .

65. Let  $A$  be a subset of a topological space  $X$ . Prove:

(a) If  $A$  is nowhere dense in  $X$ , then  $X \setminus A$  is dense in  $X$ . However, the converse need not hold unless  $A$  is closed in  $X$ .

(b) If  $A$  is closed in  $X$ , then  $A$  is nowhere dense in  $X$  if and only if  $A = \text{bdy } A$ .

(c) The set  $A$  is nowhere dense in  $X$  exactly when  $\text{cls } A$  is nowhere dense in  $X$ .

(d) The boundary of  $A$  is nowhere dense in  $X$  in case  $A$  is either open or closed in  $X$ , but otherwise  $\text{bdy } A$  need not be nowhere dense in  $X$ .

(e) The union of finitely many nowhere dense subsets of  $X$  is itself nowhere dense in  $X$ .

66. Verify that for each  $1 \leq m < n$ , the subset

$$\{ \langle x_1, x_2, \dots, x_n \rangle \in \mathbb{R}^n : x_{m+1} = x_{m+2} = \dots = x_n = 0 \}$$

of  $\mathbb{R}^n$  is nowhere dense in  $\mathbb{R}^n$ .

67. The complement of a meager set (Definition 2.51) in a topological space  $X$  is said to be **residual in  $X$** . If the topology of  $X$  is induced by some complete metric, show that each residual set in  $X$  is dense in  $X$ .

68. From the intuitive idea of “meager”, it is unlikely that the open unit interval  $]0, 1[$  is meager in  $\mathbb{R}$ . Show that it is actually nonmeager.

69. Can the inclusions (2) and (3) of Proposition 2.54 be strengthened to equalities when  $Y$  is closed in  $X$ ? when  $Y$  is open in  $X$ ?

prob:Kuratowski-closure-axioms

70. Let  $X$  be a set. Suppose  $c: \mathcal{P}(X) \rightarrow \mathcal{P}(X)$  is a map, assigning to each subset  $A$  of  $X$  a subset  $c(A)$  of  $X$ , which satisfies the following four **Kuratowski closure axioms**:

(K1)  $c(\emptyset) = \emptyset$ .

(K2) If  $A \subset X$ , then  $A \subset c(A)$ .

(K3) If  $A \subset X$ , then  $c(c(A)) = A$ .



(K4) If  $A \subset X$  and  $B \subset X$ , then  $c(A \cup B) = c(A) \cup c(B)$ .

Prove that then there is a unique topology  $\mathcal{T}$  on  $X$  such that in the topological space  $\langle X, \mathcal{T} \rangle$ , the equality  $\text{cls } A = c(A)$  holds for each subset  $A$  of  $X$ . (*Hint*: A subset of  $X$  will be open for a topology  $\mathcal{T}$  when its complement is its own closure.)

Kuratowski closure-complement problem

Kuratowski, Kazimierz

local base

neighborhood

base

neighborhood base

discrete space

metrizable space

owski-closure-complement-problem

**71.** (*The Kuratowski closure-complement problem.*) For a given subset  $A$  of a topological space  $X$ , let  $\mathcal{K}$  be the collection of all subsets of  $X$  obtained by successive steps, starting with  $A$ , that consist of taking the closure or complement in  $X$  of a set obtained in the preceding set. Thus  $\mathcal{K}$  has as members the sets  $A$ ,  $\text{cls } A$ ,  $X \setminus \text{cls } A$ ,  $\text{cls}(X \setminus \text{cls } A)$ , etc.

(a) Prove that  $\mathcal{K}$  can never contain more than 14 distinct subsets of  $X$ . (*Suggestion*: Use the superscript notation  $\bar{\phantom{x}}$  for closure and superscript prime  $'$  to denote taking the complement in  $X$ .)

(b) Show by example (say in  $\mathbb{R}$ ) that the bound of 14 can actually be attained, that is, that  $\mathcal{K}$  can consist of exactly 14 distinct sets.

## 2.4 Bases and Local Bases

sec:bases

In a metric space  $\langle X, d \rangle$ , the collection of all  $d$ -balls at a point  $x$  determines all the neighborhoods of  $x$ , and the collection of all  $d$ -balls determines all the open subsets of  $X$ . Likewise, in an arbitrary topological space, a “local base” at a point will determine all the neighborhoods of the point, and a “base” will determine all the open sets. By examining a local base or a base, it is often possible to reduce questions concerning all neighborhoods of a point or all open sets to a simpler one concerning just certain “nice” neighborhoods or open sets.

### Local bases

subsec:local-bases

def:local-base

**2.55 Definition.** Let  $x$  be a point of a topological space  $X$ . A collection  $\mathcal{M}$  of neighborhoods of  $x$  such that each neighborhood of  $x$  contains a member of  $\mathcal{M}$  is called a **local base at  $x$** .

When a particular local base  $\mathcal{M}$  at a point  $x$  is specified (or is understood from the context), a member of  $\mathcal{M}$  is sometimes referred to as a **typical neighborhood of  $x$** .

Synonyms for ‘local base at  $x$ ’ are ‘**neighborhood base at  $x$** ’ and ‘**fundamental system of neighborhoods of  $x$** ’. As usual, we may refer to a “ $\mathcal{T}$ -local base at  $x$ ” in order to indicate the particular topology  $\mathcal{T}$  under consideration.

ex:nbd-sys:local-base

**2.56 Examples.** (1) The entire neighborhood system at a point  $x$  of a topological space is itself a local base at  $x$ .

ex:open-nbds-local-base

(2) The collection of all open neighborhoods of a point  $x$  of a topological space is a local base at  $x$ .

ex:singleton-local-base-discrete

(3) For a point  $x$  of a discrete space, the collection  $\{\{x\}\}$  having the single member  $\{x\}$  is a local base at  $x$ .

ex:metric-local-bases

(4) Let  $X$  be a metrizable space and let  $x \in X$ . Take any metric  $d$  that induces the topology of  $X$ . Then each of the following collections is a local base at  $x$ :

- $\{B_\varepsilon(x; d) : \varepsilon > 0\}$ ;

realtime!local base@andlocal base •  $\{D_\varepsilon(x; d) : \varepsilon > 0\};$

Euclidean n-space@Euclidean n-space!local base@andlocal base •  $\{B_r(x; d) : r \in \mathbb{Q}, r > 0\},$

Sorgenfreyline!local base

half-disk space

half-disk space

- $\{B_{1/n}(x; d) : n = 1, 2, \dots\};$  and
- $\{B_\varepsilon(x; d) : 0 < \varepsilon < \eta\},$  where  $\eta$  is some fixed positive number.

ex:intervals-local-base-R (5) Specialize (4) by taking  $X$  to be the real line  $\mathbb{R}$ . Then the four local bases at  $x$  given by (4) are, respectively:

- the collection of all open intervals  $]x - \varepsilon, x + \varepsilon[$  symmetric about  $x$ ;
- the collection of all closed intervals  $[x - \varepsilon, x + \varepsilon]$  of positive length and symmetric about  $x$ ;
- the collection of all open intervals  $]x - r, x + r[$  of positive rational length and symmetric about  $x$ ; and
- the collection of all open intervals of the form  $]x - 1/n, x + 1/n[$  for  $n = 1, 2, 3, \dots$

Two more local bases at  $x$  are the collection

$$\{]a, b[ : a \in \mathbb{R}, b \in \mathbb{R}, a < x < b\}$$

of all open intervals containing  $x$  and the collection

$$\{[a, b] : a \in \mathbb{R}, b \in \mathbb{R}, a < x < b\}$$

of all closed intervals of positive length and containing  $x$  in their interior.

ex:open-cubes-local-base-Rn (6) Specialize (4) by taking  $X = \mathbb{R}^n$ . Since the max metric  $d_\infty$  induces the topology on  $\mathbb{R}^n$  [see Examples 2.3 (3) and Proposition 1.36], the collection

$$\{B_\varepsilon(x; d_\infty) : \varepsilon > 0\}$$

of all open *cubes* centered at  $x$  is a local base at  $x \in \mathbb{R}^n$ . Similarly as in (4), instead of taking as radii all positive real numbers as radii, we may take just positive rational numbers or the particular positive rational numbers  $1, 1/2, 1/3, \dots$

ex:Sorgenfrey-line-local-base (7) At each given real number  $x$ , the Sorgenfrey line [Examples 2.20 (1)] has as a local base the collection  $\{[x, y[ : y \in \mathbb{R}, x < y\}$  of half-open intervals with  $x$  as left endpoint.

(8) Consider the half-disk space  $H \cup L$  of Examples 2.20 (3). At each  $x \in L$ , the collection  $\{H_\varepsilon(x) : \varepsilon > 0\}$  is a local base at  $x$ . And at each  $z = \langle z_1, z_2 \rangle \in H$ , the collection

$$\{B_\varepsilon(z; d) : 0 < \varepsilon < z_2\}$$

is a local base at  $z$ .  $\diamond$

In some situations involving the neighborhood system at a point, it suffices to consider just a local base at that point. The following is such a situation.

prop:ops-and-local-base

**2.57 Proposition.** *Let  $A$  be a subset of a topological space  $X$ , and let  $x \in X$ . Then  $x \in \text{bdy } A$  if and only if there is some local base  $\mathcal{V}$  at  $x$  in  $X$  such that each member of  $\mathcal{V}$  intersects both  $A$  and  $X \setminus A$ .*

**Proof.** First assume  $x \in \text{bdy } A$ . Then the neighborhood system  $\mathcal{N}_x$  at  $x$  is such a local base. (Moreover, if  $\mathcal{V}$  is any local base at  $x$ , then each member of  $\mathcal{V}$  must intersect both  $A$  and its complement in  $X$ .)

Conversely, assume there exists some local base  $\mathcal{V}$  at  $x$  in  $X$  such that each member of  $\mathcal{V}$  intersects both  $A$  and  $X \setminus A$ . To show that  $x \in \text{bdy } A$ , let  $V$  be an arbitrary neighborhood of  $x$ . There is some  $W \in \mathcal{V}$  with  $x \in W \subset V$ . By assumption,  $W$  intersects both  $A$  and  $X \setminus A$ , and so the superset  $V$  of  $W$  intersects both as well.  $\square$

For local bases there are analogs of the properties listed in [Theorem 2.18](#) for neighborhood systems. Their proofs are left as an exercise: see [Exercise 84](#).

prop:properties-of-local-bases

**2.58 Proposition (properties of local bases).** For each point  $x$  in a topological space  $X$ , let  $\mathcal{B}_x$  be a local base at  $x$  consisting solely of open sets. Then for all  $x \in X$ :

property:LB1 (LB1) There exists at least one member of  $\mathcal{B}_x$ .

property:LB2 (LB2) The point  $x$  belongs to each member of  $\mathcal{B}_x$ .

property:LB3 (LB3) The intersection of any two members of  $\mathcal{B}_x$  contains some member of  $\mathcal{B}_x$ .

property:LB4 (LB4) If  $V \in \mathcal{B}_x$ , then for each  $y \in V$  there is some  $W \in \mathcal{B}_y$  such that  $W \subset V$ .

[Theorem 2.19](#) described how a topology may be constructed from neighborhood systems. A topology may also be constructed from local bases.

prop:topology-from-local-bases

**2.59 Proposition (constructing a topology from local bases).** Let  $X$  be a set and let  $\langle \mathcal{B}_x \rangle_{x \in X}$  be a family of collections of subsets of  $X$  having properties ((LB1))–((LB4)) above. Then there exists a unique topology  $\mathcal{T}$  on  $X$  such that, for each  $x \in X$ , the collection  $\mathcal{B}_x$  is a local base at  $x$  consisting of  $\mathcal{T}$ -open sets.

**Proof.** Existence. For each  $x \in X$ , define  $\mathcal{M}_x$  to be the collection of all those subsets  $V$  of  $X$  for which there is some  $B \in \mathcal{B}_x$  with  $B \subset V$ . The family  $\langle \mathcal{M}_x \rangle_{x \in X}$  has properties ((N1))–((N5)) as listed in [2.18](#) but for the collections  $\mathcal{M}_x$  rather than  $\mathcal{N}_x$ . According to [Theorem 2.19](#), there is a unique topology  $\mathcal{T}$  on  $X$  such that, for each  $x \in X$ , the collection  $\mathcal{M}_x$  is the  $\mathcal{T}$ -neighborhood system at  $x$ . Then for each  $x \in X$ , the given collection  $\mathcal{B}_x$  is a local base at  $x$  for  $\mathcal{T}$ .

Uniqueness. Let  $\mathcal{S}$  be a topology on  $X$  such that, for each  $x \in X$ , the given collection  $\mathcal{B}_x$  is a local base at  $x$ . Then at each  $x \in X$  the collection  $\mathcal{M}_x$  defined above is the  $\mathcal{S}$ -neighborhood system at  $x$ . That  $\mathcal{S} = \mathcal{T}$  now follows from the uniqueness assertion of [Theorem 2.19](#).  $\square$

It is especially when applying the preceding proposition to construct a topology that we refer to “typical” neighborhoods of a point.

exs:top-from-local-bases

**2.60 Examples.** (1) The discrete topology on a set  $X$  is obtained by taking as a local base at an arbitrary point  $x$  the collection  $\{\{x\}\}$ .

(2) The topology induced by a metric  $d$  on a set  $X$  is obtained by taking as a local base at an arbitrary point  $x$  the collection  $\mathcal{M}_x = \{B_x(\varepsilon; d) : \varepsilon > 0\}$  of all open balls at  $x$ .

(3) The topology of the line with two origins may be obtained as follows. Again the underlying set is  $X = \mathbb{R} \cup \{0'\}$  with  $0' \notin \mathbb{R}$ . At an  $x \in \mathbb{R}$ , a typical neighborhood is

Thomas plank

Thomas, John

topology!local base@andlocal base

ex:Thomas plank

local base topology@and topology

dimension

zero-dimensional space

an  $\mathbb{R}$ -neighborhoods of  $x$  in  $\mathbb{R}$ ; at the point  $0'$  a typical neighborhood has the form  $(V \setminus \{0\}) \cup \{0'\}$  where  $V$  is an arbitrary  $\mathbb{R}$ -neighborhood of  $0$  in  $\mathbb{R}$ .

(4) This example, unlike the preceding ones, involves a space not previously introduced, the **Thomas plank**. (It is named after John Thomas, who gave a similar example. The one described here is the modification of Thomas's example constructed by Steen and Seebach, Jr. [61, pages 113–114].)

Let

$$L_n = \begin{cases} [0, 1[ \times \{1/n\} & \text{if } n = 1, 2, 3, \dots, \\ ]0, 1[ \times \{0\} & \text{if } n = 0. \end{cases}$$

Then the Thomas plank is the set

$$P = \bigcup_{n=0}^{\infty} L_n = ([0, 1[ \times \mathbb{N}) \setminus \{(0, 0)\}$$

provided with the topology for which:

- for  $n = 1, 2, 3, \dots$ , at each point  $\langle x, 1/n \rangle$  of  $L_n$  with  $0 < x < 1$ , a typical neighborhood is the singleton  $\{\langle x, 1/n \rangle\}$ ;
- for  $n = 1, 2, 3, \dots$ , at the point  $\langle 0, 1/n \rangle$  of  $L_n$ , a typical neighborhood is a subset of  $L_n$  that contains  $\langle 0, 1/n \rangle$  and whose complement in  $L_n$  is finite; and
- at each point  $\langle x, 0 \rangle$  of  $L_0$ , a typical neighborhood is a set having the form  $\{\langle x, 0 \rangle\} \cup \{\langle x, 1/n \rangle : n \geq k\}$  where  $k \geq 1$ —a vertical stack of points starting at height  $k$  together with the point at height 0 directly below the stack.

The set  $P$  and a typical neighborhood at each of these three kinds of points are depicted in [Figure 2.10](#).

Notice that each typical neighborhood of each point as described above is closed in  $P$ .

The Thomas plank  $P$  is a Hausdorff space, as can readily be checked. However it is *not* metrizable: see [Exercise 82](#) or [Exercise 149](#).  $\diamond$

### Zero-dimensional spaces

subsec:0-dim

In a discrete space, each point  $x$  has the smallest possible neighborhood, namely, the singleton  $\{x\}$ , and this neighborhood has empty boundary. Then discreteness has the following generalization, which involves local bases.

def:0-dim

**2.61 Definition.** A topological space is said to be **zero-dimensional** when it is nonempty but at each of its points there is a local base consisting of sets whose boundaries are empty.

The following alternative characterization of zero-dimensionality is an immediate consequence of [Corollary 2.30](#).

prop:0-dim-iff-clopen-base

**2.62 Proposition.** A nonempty topological space is zero-dimensional if and only if at each point there is a local base consisting of sets that are both open and closed.

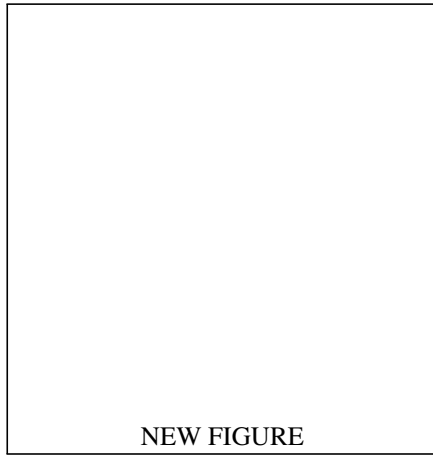
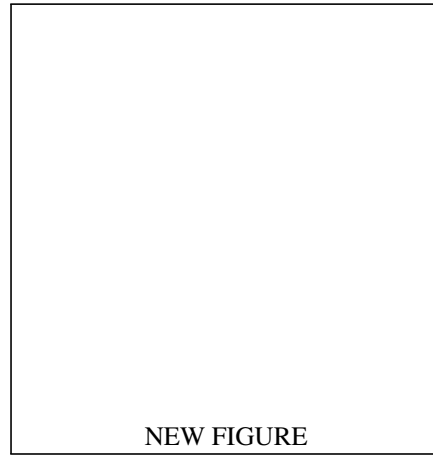
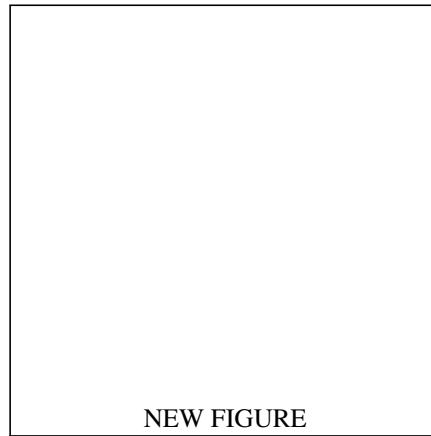
(a) Typical neighborhood at  $\langle x, 1/n \rangle$  for  $0 < x < 1$ .(b) Typical neighborhood at  $\langle x, 1/n \rangle$ .(c) Typical neighborhood at  $\langle x, 1/n \rangle$  for  $0 < x < 1$ .Figure 2.10: Typical neighborhoods in the Thomas plank  $P$ .

fig:Thomas-plank-nbds

exs:0-dim **2.63 Examples.** (1) Any discrete space, finite or infinite, is zero-dimensional.

ex:Q-0-dim (2) The space  $\mathbb{Q}$  of all rational numbers is zero-dimensional. In fact, let  $q \in \mathbb{Q}$  and let  $U$  be a neighborhood of  $q$  in  $\mathbb{Q}$ . Then  $U \supset ]a, b[ \cap \mathbb{Q}$  for some  $a, b \in \mathbb{R}$  with  $a < q < b$ . Choose *irrational* numbers  $u, v$  with  $a < u < q < v < b$  and let  $V = ]u, v[ \cap \mathbb{Q}$ . Then  $V$  is also a neighborhood of  $q$  in  $\mathbb{Q}$  with  $V \subset U$ , and although  $V$  has boundary  $\{u, v\}$  in  $\mathbb{R}$ , it has empty boundary in  $\mathbb{Q}$ .

ex:irrationals-0-dim (3) The space  $\mathbb{R} \setminus \mathbb{Q}$  of all irrational numbers is also zero-dimensional. (See [Exercise 76](#).)

ex:R-is-0-dim (4) The real line  $\mathbb{R}$  is *not* zero-dimensional. To see this, just suppose, to the contrary, that  $\mathbb{R}$  is zero-dimensional. Of course,  $\mathbb{R}$  is nonempty. Consider, say, the point 0 of  $\mathbb{R}$ . Its neighborhood  $] -1, 1[$  contains a neighborhood  $V$  that is both open and closed in  $\mathbb{R}$ . Since the set  $V$  is nonempty and bounded above in  $\mathbb{R}$ , by order-completeness ([Axiom 0.74](#)) it has a supremum  $s$ .

Cantor set  
countability properties

metrizable space!first-countable space@as first-countable space:

half-disk space!first-countable space@and first-countable space

half-disk space

Fort topology

Fort space

Fort, Marion K., Jr.

Now  $s \notin V$ , because otherwise the open set  $V$  would contain some open interval  $]s - \varepsilon, s + \varepsilon[$  around  $s$ ; and then  $s + \varepsilon/2$  would be an element of  $V$ , which is impossible since  $s$  is an upper bound of  $V$  in  $\mathbb{R}$ .

Thus  $s \in \mathbb{R} \setminus V$ . The open set  $\mathbb{R} \setminus V$  contains some open interval  $]s - \delta, s + \delta[$ . Then  $s - \delta/2$  is an upper bound of  $V$ , which is impossible since  $s$  is the *least* upper bound of  $V$  in  $\mathbb{R}$ .

[*Note:* Essentially the same argument will be used later to establish that the closed unit interval in  $\mathbb{R}$  is “compact” (Proposition 4.1 and that any interval in  $\mathbb{R}$ , including  $\mathbb{R}$  itself, is “connected” (Theorem 5.4).]

In Section 6.6 of Chapter 6 (Embedding) we shall give meaning to a space being  $n$ -dimensional and show there that the real line is, in fact, 1-dimensional.

ex:interval-not-0-dim (5) Similarly, no nondegenerate interval in  $\mathbb{R}$  is zero-dimensional. In particular, the unit interval  $[0, 1]$  is not zero-dimensional.  $\diamond$

The most famous zero-dimensional space, the *Cantor set*, which is a subspace of  $[0, 1]$ , will be constructed in the Section 4.1 (page 466).

prop:subsp-0-dim **2.64 Proposition.** A nonempty subspace of a zero-dimensional space is zero-dimensional.

### First-countable spaces

subsec:1st-countable

Using local bases we now define the first of several **countability properties** which a topological space may possess. (The additional countability properties ‘second-countable’ and ‘separable’ are introduced later in this section—see Definitions 2.75 and 2.85; a further countability property, being a ‘Lindelöf space,’ is discussed in Exercise 115.)

def:1st-countable **2.65 Definition.** A topological space  $X$  is said to be **first-countable** when at each point of  $X$  there is some countable local base.

ex:metrizableex:1st-countable **2.66 Examples.** (1) **Every metrizable space is first-countable.** In fact, if  $d$  is a metric that induces the topology of a topological space  $X$  and if  $x \in X$ , then each of the collections  $\{B_r(x; d) : r \in \mathbb{Q}\}$  and  $\{B_{1/n}(x; d) : n = 1, 2, 3, \dots\}$  is a countable local base at  $x$ .

ex:half-disk-space-1st-countable (2) The half-disk space  $H \cup L$  of Examples 2.20 (3), which we already know is Hausdorff but not metrizable, is nonetheless first-countable. In fact, at each  $x \in L$ , the countable collection  $\{H_r(x) : 0 < r \in \mathbb{Q}\}$  is a local base. And at each  $z = \langle z_1, z_2 \rangle \in H$ , the countable collection

$$\{B_r(z; d) : r \in \mathbb{Q}, 0 < r < z_2\}$$

is a local base at  $z$ .

ex:Fort-topology (3) This is an example of a Hausdorff space that is not first-countable.

Let  $X$  be an infinite set and let  $p \in X$  be some specific point. Define  $\mathcal{T}$  to be the collection of all those subsets  $U$  of  $X$  for which  $X \setminus U$  is finite or  $p \notin U$ . Then  $\mathcal{T}$  is a topology on  $X$ , called the **Fort topology** (named after M. K. Fort, Jr.).

We show that  $\langle X, \mathcal{T} \rangle$  is a Hausdorff space. Let  $x, y \in X$  with  $x \neq y$ . One of the two points  $x$  and  $y$ , say  $x$ , is distinct from  $p$ . Then  $X \setminus \{x\}$  and  $\{x\}$  are disjoint open neighborhoods of  $x$  and  $y$ , respectively.

Assume now that  $X$  is *uncountable*. In this case we refer to  $\langle X, \mathcal{T} \rangle$  as a **Fort space**. We show that there is no countable local base at the point  $p$  in  $\langle X, \mathcal{T} \rangle$ . Just suppose, to the contrary, that  $\{V_n : n \in \mathbb{N}\}$  is a local base at  $p$  for some sequence  $\langle V_n \rangle_{n \in \mathbb{N}}$  of (not necessarily distinct) open sets. Since  $X \setminus V_n$  is finite for each  $n \in \mathbb{N}$ , then

first-countable space  
local base

$$X \setminus \bigcap_{n=0}^{\infty} V_n = \bigcup_{n=0}^{\infty} (X \setminus V_n)$$

is countable. Since  $X$  is itself uncountable, there is some  $x \in \bigcap_{n=0}^{\infty} V_n$  with  $x \neq p$ . This is impossible since  $p \in V_n$ .

Note that  $\langle X, \mathcal{T} \rangle$  does have a countable local base at each  $x \neq p$ , namely, the collection consisting of the single open neighborhood  $\{x\}$  of  $x$ .  $\diamond$

As an example of a Hausdorff space that fails to have a countable local base at any of its points, we shall construct below [Examples 2.72 (8)] a certain space whose “points” are functions. That example will be geometrically more natural than the Fort space, but more difficult to describe.

## Bases

subsec:bases

The collection of all  $d$ -balls at a point of a metric space  $\langle X, d \rangle$  has been generalized to the concept of a local base at a point of a topological space. Now we introduce a generalization of the collection of all  $d$ -balls (at all the points) in a metric space.

**2.67 Definition.** Let  $X$  be a topological space with topology  $\mathcal{T}$ . A collection  $\mathcal{B}$  of subsets of  $X$  is called a **base of  $X$**  (and **of  $\mathcal{T}$** ) if:

property:base-member-open

(i) each member of  $\mathcal{B}$  is open in  $X$ ; and

property:open-union-of-base-members

(ii) each open subset of  $X$  is the union of some collection of sets each of which belongs to  $\mathcal{B}$ .

When a particular base  $\mathcal{B}$  of a topology is under consideration, we refer to each of its members as a **basic open set**.

To check condition (ii), it suffices to check that each *nonempty* open subset of  $X$  is the union of some collection of sets belonging to  $\mathcal{B}$ . Notice that since  $X$  itself is open in its given topology, condition (ii) implies that each point of  $X$  belongs to at least one member of  $\mathcal{B}$ , that is, the collection  $\mathcal{B}$  *covers*  $X$  (Definition 0.23).

**2.68 Examples.** (1) The entire topology of a topological space  $X$  is itself a base of  $X$ .

ex:singletons-base-of-discrete

(2) The collection  $\{\{x\} : x \in X\}$  of all singletons in a discrete space  $X$  is a base of  $X$ . For this base, each singleton is a basic open set.

ex:all-balls-base-of-metric

(3) Let  $X$  be a metrizable space. Take any metric  $d$  that induces the topology of  $X$ . By Proposition 1.19 and Exercise 1.32, the collection  $\{B_\varepsilon(x; d) : x \in X, \varepsilon > 0\}$  is a base of  $X$ .

ex:open-ints-base-of-R

(4) Specialize (3) by taking  $X$  to be the real line  $\mathbb{R}$ . Then the collection

$$\{]a, b[ : a \in \mathbb{R}, b \in \mathbb{R}, a < b\}$$

of all open intervals is a base of  $\mathbb{R}$ .

ex:open-cubes-base-of- $\mathbb{R}^n$  (5) Specialize (3) by taking  $X = \mathbb{R}^n$ . Since the maxi metric  $d_\infty$  induces the usual topology on  $\mathbb{R}^n$ , the collection

$$\{B_\varepsilon(x; d_\infty) : x \in \mathbb{R}^n, \varepsilon > 0\}$$

of all open cubes is a base of  $\mathbb{R}^n$ .

Another (larger) base of  $\mathbb{R}^n$  is the collection

$$\{]a_1, b_1[ \times ]a_1, b_1[ \times \cdots \times ]a_n, b_n[ : a_i < b_i \text{ for each } i = 1, 2, \dots, n\}$$

of all open “boxes”—in technical terms,  $n$ -dimensional “hyperrectangles.” (In dimension  $n = 2$  this is the collection of all open rectangles, and in dimension  $n = 3$  the collection of all open rectangular parallelepipeds).

ex:Sorgenfrey-line-base (6) The Sorgenfrey line  $\mathbb{R}_l$  [Examples 2.20 (1)] has as a base the collection of all left-closed, right-open intervals  $[x, y[$ . Observe that each member of this base is closed as well as open. (See Exercise 88.)

ex:base-of-line-with-2-origins (7) The line with two origins [Examples 2.20 (3)] has the collection

$$\{]a, b[ : a < b\} \cup \{]a, 0[ \cup \{0'\} \cup ]0, b[ : a < 0 < b\}$$

as a base.  $\diamond$

Comparison of Examples 2.68 (1)–(5) with Examples 2.56 (1)–(6) suggests the following characterization of a base in terms of local bases.

prop:base-vs-local-base **2.69 Proposition.** Let  $\mathcal{B}$  be a collection of open sets in a topological space  $X$ . Then the following two conditions are equivalent:

cond:base-of- $X$  (i) The collection  $\mathcal{B}$  is a base of  $X$ .

cond:local-base-at- $x$ -in- $X$  (ii) For each  $x \in X$ , the collection

$$\mathcal{B}_x = \{B : x \in B \in \mathcal{B}\}$$

of all those members of  $\mathcal{B}$  containing  $x$  is a local base at  $x$ .

**Proof.** Assume (i). We show (ii). Let  $x \in X$ . If  $B \in \mathcal{B}_x$ , then  $B$  is an open set that contains  $x$ , so that  $B$  is a neighborhood of  $x$ .

Now let  $V$  be any neighborhood of  $x$ . Choose an open set  $U$  with  $x \in U \subset V$ . Since  $\mathcal{B}$  is a base of  $X$ , then  $U = \bigcup_{i \in I} B_i$  for some family  $\langle B_i \rangle_{i \in I}$  of sets belonging to  $\mathcal{B}$ . Then  $x \in B_j$  for some  $j \in I$ . Hence  $B_j \in \mathcal{B}_x$  and  $B_j \subset V$ .

Conversely, assume (ii). We show (i). Let  $U$  be an arbitrary nonempty open set in  $X$ . For each  $x \in U$  the set  $U$  is a neighborhood of  $x$ , and so by assumption there exists some  $B_x \in \mathcal{B}_x$  with  $B_x \subset U$ . Then  $U = \bigcup_{x \in U} B_x$  with  $B_x \in \mathcal{B}$  for each  $x \in U$ .  $\square$

Condition (ii) above for a collection  $\mathcal{B}$  of open sets to be a base of  $X$  has the equivalent formulation:

cond:base-exists-B-at- $x$  (2') For each open set  $U$  in  $X$  and each  $x \in U$ , there exists  $B \in \mathcal{B}$  with  $x \in B \subset U$ .

This condition implies, in particular, that each  $x \in X$  belongs to at least one member of  $\mathcal{B}$ , that is  $\bigcup \mathcal{B} = X$ . In other words, the condition implies that  $\mathcal{B}$  covers  $X$  (Definition 0.23).

Proposition 2.69 furnishes additional examples. Thus if  $d$  is a metric inducing the topology of a space  $X$ , then 2.69 together with Examples 2.66 (1) says that the collection  $\{B_r(x; d) : x \in X, 0 < r \in \mathbb{Q}\}$  is a base of  $X$ .



The next result states how bases and local bases in a subspace are obtained from the corresponding objects in the entire space.

topology!associated with a base

prop:base-local-base-subspace

**2.70 Proposition.** *Let  $Y$  be a subspace of a topological space  $X$ . then:*

(1) *If  $y \in Y$  and if  $\mathcal{M}$  is a local base at  $y$  in  $X$ , then the collection*

$$\{V \cap Y : V \in \mathcal{M}\}$$

*is a local base at  $y$  in the subspace  $Y$ .*

(2) *If  $\mathcal{B}$  is a base of  $X$ , then the collection*

$$\{B \cap Y : B \in \mathcal{B}\}$$

*is a base of the subspace  $Y$ .*

**Proof.** (1) Let  $y \in Y$  and let  $\mathcal{M}$  be a local base at  $y$  in  $X$ . Recall (Proposition 2.15) that the neighborhoods of  $y$  in the subspace  $Y$  are precisely the intersections  $V \cap Y$  of neighborhoods  $V$  of  $y$  in  $X$ . In particular,  $V \cap Y$  is a neighborhood of  $y$  in  $Y$  for each  $V \in \mathcal{M}$ .

Now let  $W$  be an arbitrary neighborhood of  $y$  in  $Y$ . Then  $W = U \cap Y$  for some neighborhood  $U$  of  $y$  in  $X$ . Since  $\mathcal{M}$  is a local base at  $y$  in  $X$ , there is some  $V \in \mathcal{M}$  with  $V \subset U$ . Then  $V \cap Y \subset U \cap Y = W$ .

(2) Use (1) and Proposition 2.69.  $\square$

The next theorem gives us a new way of constructing topological spaces.

thm:top-from-base

**2.71 Theorem (constructing a topology from a base).** *Let  $X$  be a set. Let  $\mathcal{B}$  be a collection of subsets of  $X$  having the two properties:*

property:B1 (B1) *For each  $x \in X$ , there exists  $B \in \mathcal{B}$  such that  $x \in B$ .*

property:B2 (B2) *If  $B_1, B_2 \in \mathcal{B}$  and if  $x \in B_1 \cap B_2$ , then there is some  $B \in \mathcal{B}$  such that  $x \in B \subset B_1 \cap B_2$ .*

*Then there exists a unique topology  $\mathcal{T}$  on  $X$  of which  $\mathcal{B}$  is a base.*

Condition (B1) just says that the collection  $\mathcal{B}$  is a **cover of  $X$**  (Definition 0.23).

In the notation of the preceding theorem, the unique topology  $\mathcal{T}$  of which the given collection  $\mathcal{B}$  is a base is said to be **associated with  $\mathcal{B}$** .

**Proof.** In view of Proposition 2.69, we shall be interested in the two collections

$$\mathcal{B}_x = \{B : x \in B \in \mathcal{B}\},$$

$$\mathcal{N}_x = \{V : x \in B \subset V \text{ for some } B \in \mathcal{B}_x\},$$

defined for each  $x \in X$ .

Uniqueness. Suppose  $\mathcal{T}$  is a topology on  $X$  of which  $\mathcal{B}$  is a base. By Proposition 2.69, for each  $x \in X$  the collection  $\mathcal{B}_x$  is a local base at  $x$  for  $\mathcal{T}$ , and so  $\mathcal{N}_x$  is the  $\mathcal{T}$ -neighborhood system at  $x$ . According to Theorem 2.19, the topology  $\mathcal{T}$  is uniquely determined by the  $\mathcal{T}$ -neighborhood systems at all  $x \in X$ .

Existence. We shall construct the desired topology  $\mathcal{T}$  by applying Theorem 2.19 to the collections  $\mathcal{N}_x$  defined above. Conditions (N2) and (N4) need to apply 2.19 are immediate

order topology consequences of the definition of  $\mathcal{N}_x$ ; properties (N1) and (N3) follow from (B1) and (B2), respectively. To verify (N5), let  $x \in X$  and  $V \in \mathcal{N}_x$ . Choose  $B \in \mathcal{B}$  with  $x \in B \subset V$ . Then  $B \in \mathcal{N}_y$  for each  $y \in B$ .

According to Theorem 2.19, there is a topology  $\mathcal{T}$  on  $X$  such that for each  $x \in X$  the collection  $\mathcal{N}_x$  is the  $\mathcal{T}$ -neighborhood system at  $x$ . It remains to show that  $\mathcal{B}$  is a base of this  $\mathcal{T}$ . Of  $B \in \mathcal{B}$ , then  $B \in \mathcal{N}_x$  for each  $x \in B$ , and so  $B$  is  $\mathcal{T}$ -open. For each  $x \in X$  the collection  $\mathcal{B}_x$  is a local base at  $x$  by definition of  $\mathcal{N}_x$ . It follows from Proposition 2.69 that  $\mathcal{B}$  is a base of  $\mathcal{T}$ .  $\square$

Note that condition (B2) above is automatically satisfied in case the intersection of any two members of  $\mathcal{B}$  again belongs to  $\mathcal{B}$ .

Don't confuse conditions (i)–(ii) in Definition 2.67 with conditions ((B1))–((B2)) of Theorem 2.71:

- in Definition 2.67 we start with a given topology  $\mathcal{T}$  on a set  $X$ ; the conditions (i) and (ii) are those that a collection  $\mathcal{B}$  must satisfy in order to form a base of that particular topology; but by contrast,
- in Theorem 2.71 we start with just a set  $X$  and a collection  $\mathcal{B}$  of subsets of that set that satisfies (B1) and (B2); then the topology on  $X$  arises from  $\mathcal{B}$ .

**2.72 Examples.** (1) Let  $X$  be a totally ordered set (Definition 0.69) under a total ordering  $\leq$ , and suppose that  $X$  has at least two elements. Define  $\mathcal{B}$  to be the collection of all open rays and all open intervals in  $X$ , that is,

$$\mathcal{B} = \{]x, \rightarrow[ : x \in X\} \cup \{]\leftarrow, y[ : y \in X\} \cup \{]x, y[ : x, y \in X\}.$$

The collection  $\mathcal{B}$  satisfies (B1). In fact, if  $x \in X$ , there is some  $y \in X$  with  $y \neq x$ ; then  $x \in ]\leftarrow, y[$  in case  $y > x$ , and  $x \in ]y, \rightarrow[$  in case  $y < x$ . We leave the verification of (B2) to the reader, noting here only that

$$]x, \rightarrow[ \cap ]\leftarrow, y[ = ]x, y[$$

for all  $x, y \in X$ .

The topology  $\mathcal{T}$  associated with  $\mathcal{B}$  is called the **order topology** on the totally ordered set  $\langle X, \leq \rangle$  and is said to be **induced by** the partial ordering  $\leq$ .

Suppose now that  $X$  in fact has neither a least element nor a greatest element. Then for each  $x \in X$ ,

$$]x, \rightarrow[ = \bigcup_{y > x} ]x, y[, \quad ]\leftarrow, x[ = \bigcup_{y < x} ]y, x[.$$

Hence in this case the order topology also has as base the collection

$$\{]x, y[ : x, y \in X\}$$

of all open intervals.

In particular, in view of Examples 2.68 (4), the order topology on  $\mathbb{R}$  induced by the usual ordering of  $\mathbb{R}$  is just the usual topology on  $\mathbb{R}$ .

In general, a totally ordered set  $X$  provided with its order topology is a Hausdorff space. (See Exercise 89.)

(2) Again let  $X$  be a totally ordered set, provide  $X$  with its order topology, and let  $Y$  be a subset of  $X$ . Then on  $Y$  we have two topologies:

- the relative topology induced by the order topology of  $X$ ; and
- the order topology induced by the total ordering of  $Y$  that is inherited from the total ordering of  $X$ .

extended realline  
ordinal  
ordinal!first uncountable  
ordinal space  
ordinal!first infinite

*These two topologies on the subset  $Y$  need not be the same!* For example, take  $X = \mathbb{R}$  with its usual topology, that is, the topology induced by the usual ordering of  $\mathbb{R}$ ; and let  $Y = \{0\} \cup ]2, 3[$ . In the relative topology on  $Y$ , the singleton  $\{0\}$  is open because  $\{0\} = ]-1, 1[ \cap Y$ . However,  $\{0\}$  is *not* open in the order topology induced by the total ordering of  $Y$  inherited from the usual ordering of  $\mathbb{R}$ , because each member of the order topology of  $Y$  that contains 0 must contain a basic open set of the form  $\{y \in Y : 0 \leq y < b\}$  for some  $b$  with  $2 < b < 3$ , and such a set cannot equal  $\{0\}$ .

x:order-topology-on-extended-reals

- (3) Consider the extended real line  $\widehat{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$  with its total ordering  $\leq$  and its metric  $\widehat{d}$  as defined in [Example 1.41](#). Then the order topology  $\mathcal{T}$  induced on  $\widehat{\mathbb{R}}$  by  $\leq$  is the same as the topology  $\mathcal{T}(\widehat{d})$  induced on  $\widehat{\mathbb{R}}$  by the metric  $\widehat{d}$ .

To prove our assertion it suffices to show that each  $\mathcal{T}$ -neighborhood of a point  $x \in \widehat{\mathbb{R}}$  is a  $\mathcal{T}(\widehat{d})$ -neighborhood of  $x$ , and vice versa. Let  $x \in \mathbb{R}$ .

Case (i):  $x \in \mathbb{R}$ . Let  $V$  be an arbitrary  $\mathcal{T}(\widehat{d})$ -neighborhood of  $x$ . Let  $\varphi: \mathbb{R} \rightarrow ]-1, 1[$  be the bijection defined in [Example 1.41](#). Choose  $\varepsilon > 0$  with

$$\varepsilon \leq \min\{|\varphi(x) - 1|, |1 + \varphi(x)|\},$$

so that  $B_\varepsilon(x; \widehat{d}) \subset V$ . In our earlier proof ([Proposition 1.42](#)) that the metric on  $\mathbb{R}$  induced by  $\widehat{d}$  is equivalent to the Euclidean metric on  $\mathbb{R}$ , we actually showed that for such an  $\varepsilon$  the  $\widehat{d}$ -ball  $B_\varepsilon(x; \widehat{d})$  is an open interval in  $\mathbb{R}$  that contains  $x$ . Hence  $V$  is a  $\mathcal{T}$ -neighborhood of  $x$  in  $\widehat{\mathbb{R}}$ .

Conversely, let  $V$  be an arbitrary  $\mathcal{T}$ -neighborhood of  $x$ . Choose  $a$  and  $b$  with  $-\infty < a < x < b < +\infty$  and  $]a, b[ \subset V$ . In the proof just cited we showed that  $B_\varepsilon(x; \widehat{d}) \subset ]a, b[$  for a suitable  $\varepsilon > 0$ . Hence  $V$  is a  $\mathcal{T}(\widehat{d})$ -neighborhood of  $x$ .

ex-case:plus-infty

Case (ii):  $x = +\infty$ . Let  $V$  be an arbitrary  $\mathcal{T}(\widehat{d})$ -neighborhood of  $+\infty$ . Then

$$B_\varepsilon(+\infty; \widehat{d}) \subset V$$

for some  $\varepsilon$  with  $0 < \varepsilon < 1$ . Now

$$B_\varepsilon(+\infty; \widehat{d}) = ]a/\varepsilon - 1, +\infty] = ]1/\varepsilon - 1, \rightarrow[$$

by [Lemma 1.43](#). Thus  $V$  is a  $\mathcal{T}$ -neighborhood of  $+\infty$ .

Conversely, let  $V$  be a  $\mathcal{T}$ -neighborhood of  $+\infty$ . then

$$]u, \rightarrow[ = ]u, +\infty] \subset V$$

for some  $u > 0$ . Now

$$]u, +\infty] = B_{1/(u+1)}(+\infty; \widehat{d})$$

by [Lemma 1.43](#). Thus  $V$  is a  $\mathcal{T}(\widehat{d})$ -neighborhood of  $+\infty$ .

Case (iii):  $x = -\infty$ . This case is treated in the same way as Case (ii).

ex:ordinal-spaces

- (4) Specialize (1) to the well-ordered set  $\Omega^+ = [0, \Omega]$  consisting of all ordinals up to and including the first uncountable ordinal  $\Omega$ , as described in the subsection “Well-ordered sets” ([page 109](#)). When provided with the order-topology, that set and intervals in it are referred to as **ordinal spaces**. In particular, one of these ordinal spaces is the subspace  $\omega^+ = [0, \omega]$  consisting of all ordinals up to and including the first infinite ordinal  $\omega$ .

order topology

product space

product topology

product of two topological spaces

ex:prod-2-spaces

function space

In  $\Omega^+$ , at an ordinal  $\gamma$  with  $0 < \gamma < \Omega$ , the collection of all open intervals  $]\alpha, \beta[$  with  $0 \leq \alpha < \beta \leq \Omega$  is a local base. At 0, the collection of all left-closed, right-open intervals  $[0, \beta[$  with  $0 < \beta < \Omega$  is a local base. And at  $\Omega$ , the collection of all left-open, right-closed intervals  $]\alpha, \Omega]$  with  $0 \leq \alpha < \Omega$  is a local base.

- (5) Let  $X$  and  $Y$  be topological spaces. Form the collection

$$\mathcal{B} = \{U \times V : U \text{ is open in } X \text{ and } V \text{ is open in } Y\}$$

Now for each point  $\langle x, y \rangle \in X \times Y$ , the member  $X \times Y$  of  $\mathcal{B}$  contains  $\langle x, y \rangle$ . Further, if  $B_1 = U_1 \times V_1$  and  $B_2 = U_2 \times V_2$  are two members of  $\mathcal{B}$ , then

$$B_1 \cap B_2 = (U_1 \times V_1) \cap (U_2 \times V_2) = (U_1 \cap U_2) \times (V_1 \cap V_2) \in \mathcal{B}.$$

According to [Theorem 2.71](#), there is a unique topology  $\mathcal{T}$  on the product set  $X \times Y$  of which  $\mathcal{B}$  is a base. This topology  $\mathcal{T}$  is called the **product topology**, and  $(X \times Y, \mathcal{T})$  is known as a **product space**—more specifically, as **the product of the topological spaces  $X$  and  $Y$** .

*An open set for the product topology need itself be the product of open sets.* See (6), below.

*Whenever we consider the product of two topological spaces as a topological space itself, it will be with respect to this product topology.*

The preceding construction of a topology on the product of two topological spaces will be generalized in [Section 3.3](#) to the case of finitely many, and even infinitely many, topological spaces.

ex:prod-top-plane

- (6) For a specific example of the product of two spaces take  $X = Y = \mathbb{R}$ , the real line with its usual topology. We claim that **the product topology on  $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$  is the usual topology** of the Euclidean plane.

To see this, note first that by [Examples 1.18 \(3\)](#), each open square of the form  $]x - \varepsilon, x + \varepsilon[ \times ]y - \varepsilon, y + \varepsilon[$  is a  $d_\infty$ -ball and so is  $d_\infty$ -open, where  $d_\infty$  is the max metric; hence each open rectangle  $]t, s[ \times ]u, v[$ , being the union of such open squares, is  $d_\infty$ -open. It follows that the product topology on  $\mathbb{R}^2$  is the topology induced by  $d_\infty$ . But the usual topology on  $\mathbb{R}^2$  is induced by the Euclidean metric, which by [Proposition 1.36](#) is equivalent to the max metric  $d_\infty$ .

Notice that an open set in the product topology on  $\mathbb{R}^2$  need not be the product of open sets. In fact, by what was just established, a ball for the Euclidean metric is open in the product topology yet is not the product of an open set in  $\mathbb{R}$  with an open set in  $\mathbb{R}$ .

ex:Sorgenfrey-plane

- (7) Another specific example of the product two spaces is the **Sorgenfrey plane**—the product  $\mathbb{R}_l \times \mathbb{R}_l$  of the Sorgenfrey line [[Examples 2.20 \(1\)](#)] with itself. This space has as a base the collection of all rectangular sets  $[x, y[ \times [u, v[ = \{\langle s, t \rangle \in \mathbb{R} \times \mathbb{R} : x \leq s < y, u \leq t < v\}$ . A typical basic open set is depicted in [Figure 2.11](#).

y-ptwise-convergence-all-fns-R-to-R

- (8) Let  $\mathcal{F}$  be the set of all functions  $f: \mathbb{R} \rightarrow \mathbb{R}$ . We are going to define a topology on  $\mathcal{F}$  by applying [Theorem 2.71](#).

Given any integer  $k \geq 1$ , any  $k$ -tuple  $\langle x_1, x_2, \dots, x_k \rangle$  of real numbers, and any  $k$ -tuple  $\langle I_1, I_2, \dots, I_k \rangle$  of open intervals in  $\mathbb{R}$ , define

$$B(x_1, x_2, \dots, x_k; I_1, I_2, \dots, I_k) = \{f \in \mathcal{F} : f(x_j) \in I_j \text{ for each } j = 1, 2, \dots, k\}.$$

Thus a function  $f: \mathbb{R} \rightarrow \mathbb{R}$  belongs to  $B(x_1, x_2, \dots, x_k; I_1, I_2, \dots, I_k)$  if and only if the graph of  $f$  passes through each of the vertical segments  $\{x_1\} \times I_1, \{x_2\} \times$

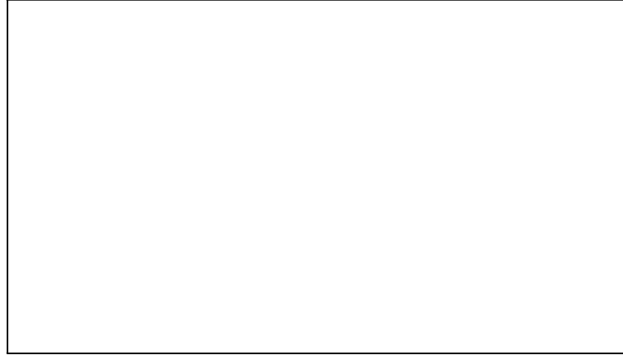


Figure 2.11: A basic open set in the Sorgenfrey plane.

fig:basic-set-Sorgenfrey-plane

$I_2, \dots, \{x_k\} \times I_k$  in the plane  $\mathbb{R} \times \mathbb{R}$ . [For an example, see 2.12, which depicts three members of a set having the form  $B(x_1, x_2, x_3, x_4; I_1, I_2, I_3, I_4)$ .] Define  $\mathcal{B}$  to be the

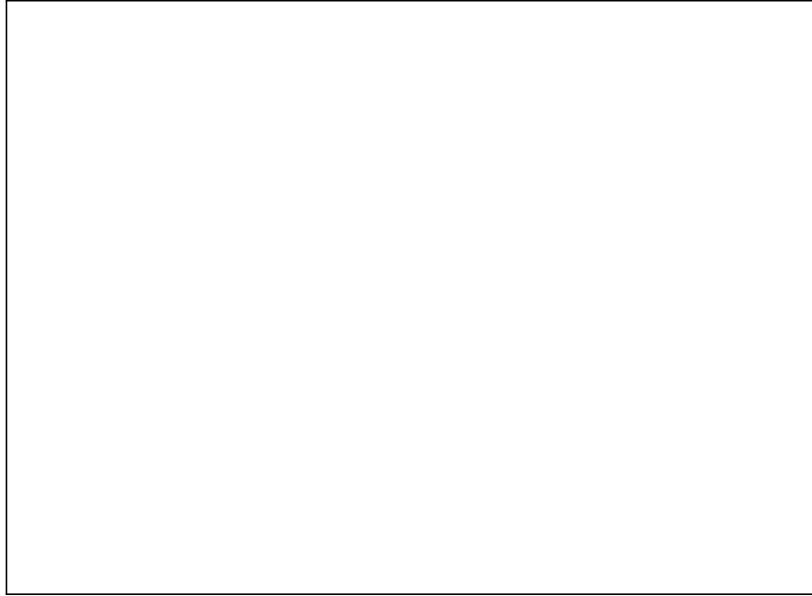


Figure 2.12: Some functions in a basic open set for the topology of pointwise convergence.

fig:3-fns-thru-4-windows

collection of all such subsets  $B(x_1, x_2, \dots, x_k; I_1, I_2, \dots, I_k)$  of  $\mathcal{F}$  for all possible choices of  $k$ , of  $\langle x_1, x_2, \dots, x_k \rangle$ , and of  $\langle I_1, I_2, \dots, I_k \rangle$ . We are going to show that  $\mathcal{B}$  has properties (B1) and (B2).

Property (B1) holds for  $\mathcal{B}$  because if  $f \in \mathcal{F}$ , then  $f \in B(x_1; I_1)$  for any real number  $x_1$  and any open interval  $I_1$  that includes the value  $f(x_1)$ .

That property (B2) holds for  $\mathcal{B}$  follows from the relation

$$\begin{aligned} B(x_1, x_2, \dots, x_k; I_1, I_2, \dots, I_k) \cap B(t_1, t_2, \dots, t_m; J_1, J_2, \dots, J_m) \\ = B(x_1, x_2, \dots, x_k, t_1, t_2, \dots, t_m; I_1, I_2, \dots, I_k, J_1, J_2, \dots, J_m). \end{aligned}$$

pointwise convergence topology!pointwise convergence of pointwise convergence function space

From [Theorem 2.71](#) we obtain a unique topology  $\mathcal{T}$  on  $\mathcal{F}$  having  $\mathcal{B}$  as a base. We call this topology associated with  $\mathcal{B}$  the **topology of pointwise convergence** on the function space  $\mathcal{F}$ . (This terminology will be explained in [Section 3.5](#).)

We show that **the functions space  $(\mathcal{F}, \mathcal{T})$  is a Hausdorff space**. Let  $f, g \in \mathcal{F}$  with  $f \neq g$ . Then  $f(x) \neq g(x)$  for some  $x \in \mathbb{R}$ . Choose disjoint open intervals  $I$  and  $J$  that contain the values  $f(x)$  and  $g(x)$ , respectively. Then  $B(x; I)$  and  $B(x; J)$  are disjoint open neighborhoods of  $f$  and  $g$ , respectively.

We show that **the function space  $(\mathcal{F}, \mathcal{T})$  is not first-countable** by showing that at no  $f \in \mathcal{F}$  is there any countable local base. Just suppose there is a countable local base at some  $f \in \mathcal{F}$ . Since  $\{B : f \in B \in \mathcal{B}\}$  is a local base at  $f$ , there is a sequence  $\text{seq } B_n$  of members of  $\mathcal{B}$  such that  $\{B_n : n \in \mathbb{N}\}$  is also a local base at  $f$ . For each  $n$  the set  $B_n$  has the form

$$B_n = B(x_{1,n}, x_{2,n}, \dots, x_{k_n,n}; I_{1,n}, I_{2,n}, \dots, I_{k_n,n}).$$

The set  $\{x_{j,n} : n \in \mathbb{N}, 1 \leq j \leq k_n\}$  of real numbers used to form these sets  $B_n$  is only countable, whereas  $\mathbb{R}$  itself is uncountable. Hence we may choose some  $x_0 \in \mathbb{R}$  such that  $x_0 \neq x_{j,n}$  for all  $n \in \mathbb{N}$  and all  $1 \leq j \leq k_n$ . Take  $I_0$  to be any open interval that includes the value  $f(x_0)$ , so that  $B(x_0; I_0)$  is a neighborhood of  $f$ . Then

$$B_n \subset B(x_0; I_0)$$

for some  $n$ .

We shall now obtain a contradiction by constructing a  $g \in B_n$  with  $g \notin B(x_0; I_0)$ . Define  $g: \mathbb{R} \rightarrow \mathbb{R}$  by

$$g(x) = \begin{cases} y_{j,n} & \text{if } x = x_{j,n} \text{ for } 1 \leq j \leq k_n, \\ y_0 & \text{if } x = x_0, \\ 0 & \text{otherwise.} \end{cases}$$

In view of what we just proved and [Examples 2.66 \(1\)](#), **the function space  $(\mathcal{F}, \mathcal{T})$  is not metrizable**.  $\diamond$

Although an open set in the product of two spaces [[Examples 2.72 \(5\)](#)] need not itself be a product of open sets, nonetheless the following results about products of open sets do hold. (Proofs are requested in [Exercise 107](#).)

prop:local-base-base-prod-2

**2.73 Proposition.** *Let  $X$  and  $Y$  be topological spaces. Then*

- (1) *If  $\mathcal{M}$  and  $\mathcal{N}$  are local bases at points  $x$  and  $y$  in  $X$  and  $Y$ , respectively, then the collection  $\{U \times V : U \in \mathcal{M}, V \in \mathcal{N}\}$  is a local base at the point  $(x, y)$  in  $X \times Y$ .*
- (2) *If  $\mathcal{U}$  and  $\mathcal{V}$  are bases of  $X$  and  $Y$ , respectively, then the collection  $\{U \times V : U \in \mathcal{U}, V \in \mathcal{V}\}$  is a base of  $X \times Y$ .*

Product spaces provide a setting for the following interesting criterion for being a  $T_2$ -space. (The proof is requested in [Exercise 110](#).)

prop:diag-closed-for-T2

**2.74 Proposition (closed diagonal criterion).** *A topological space  $X$  is a  $T_2$ -space if and only if the diagonal  $\Delta_X = \{ \langle x, x \rangle : x \in X \}$  of  $X \times X$  is closed in the product space  $X \times X$ .*

### Second-countable spaces

Euclidean  $\mathbb{R}^n$ -space!second-countable

subsec:2nd-countable

The next countability property is defined in terms of bases.

def:2nd-countable

**2.75 Definition.** A topological space is said to be **second-countable** if it has some countable base.

ex: $\mathbb{R}^n$  is 2nd-countable

**2.76 Examples.** (1) The metrizable space  $\mathbb{R}^n$  is second-countable.

To construct a countable base, we use the max metric  $d_\infty$  on  $\mathbb{R}^n$  (Definition 1.7). Form the collection

$$\mathcal{B} = \{B_r(z; d_\infty) : 0 < r \in \mathbb{Q}, z_i \in \mathbb{Q} \text{ for each } i = 1, 2, \dots, n\}$$

consisting of all cubes whose sides have rational lengths and whose centers have all coordinates rational. Since both  $\mathbb{Q}^n$  and  $\mathbb{Q}$  are countable, so is  $\mathcal{B}$ . Since  $d_\infty$  induces the usual topology of  $\mathbb{R}^n$  (Proposition 1.36), each member of  $\mathcal{B}$  is open.

To complete the proof that  $\mathcal{B}$  is a base of  $\mathbb{R}^n$ , we verify condition (ii) of Proposition 2.69. Let  $x \in \mathbb{R}^n$  and let  $V$  be any neighborhood of  $x$ . We construct a set  $B \in \mathcal{B}$  such that  $x \in B \subset V$ . Choose  $\varepsilon > 0$  with  $B_\varepsilon(x; d_\infty) \subset V$ . For each  $i = 1, 2, \dots, n$ , there is a rational number  $z_i$  with  $|z_i - x_i| < \varepsilon/2$ . Set  $z = \langle z_1, z_2, \dots, z_n \rangle$ , so that  $d_\infty(z, x) < \varepsilon/2$ . Choose a rational number  $r$  with

$$d_\infty(z, x) < r < \frac{\varepsilon}{2}.$$

(see Figure 2.13). Then  $x \in B_r(z; d_\infty) \in \mathcal{B}$ . Finally,  $B_r(z; d_\infty) \subset B_\varepsilon(x; d_\infty)$ ,

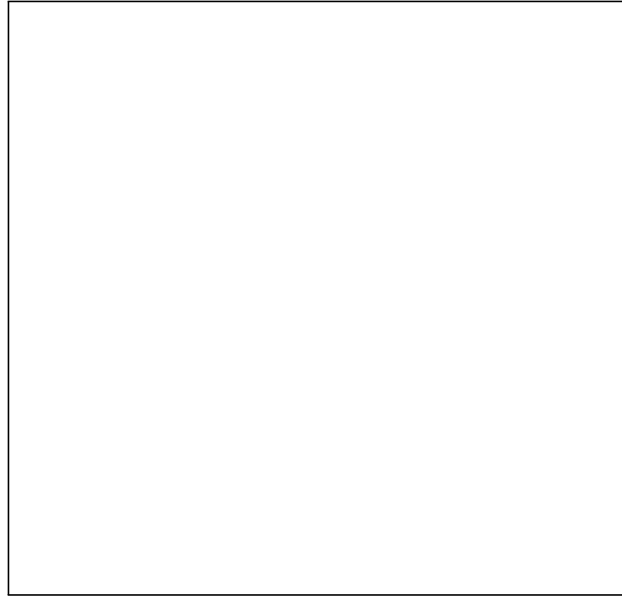


Figure 2.13: Constructing a rational-sided, rational-center cube inside a neighborhood of a point in  $\mathbb{R}^n$ .

fig:rational-cube-inside

because  $y \in B_r(z; d_\infty)$  implies

$$d_\infty(y, x) \leq d_\infty(y, z) + d_\infty(z, x) < r + r < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

ex:2nd-countable-discrete (2) Let  $X$  be a discrete space. Then the topological space  $X$  is second-countable if and only if the underlying set  $X$  is countable. To see this, note first that the set  $X$  is countable if and only if the collection

$$\mathcal{S} = \{\{x\} : x \in X\}$$

of singletons is countable. If  $X$  is countable, then  $\mathcal{S}$  is a countable base of  $X$ . Now suppose that  $X$  is uncountable and let  $\mathcal{B}$  be a base of  $X$ . If  $x \in X$ , then  $x \in B \subset \{x\}$  for some  $B \in \mathcal{B}$ , whence  $\{x\} = B$ . Thus  $\mathcal{S} \subset \mathcal{B}$ , and so  $\mathcal{B}$  is uncountable.

By putting the discrete topology on an uncountable set (for example, on  $\mathbb{R}$ ), we obtain a space that is metrizable—hence first-countable—but *not* second-countable.  $\diamond$

The preceding example shows that *a first-countable space need not be second-countable*. However, as an immediate consequence of Proposition 2.69, the converse is true.

**2.77 Proposition.** Every second-countable space is first-countable.

The following is an immediate consequence of Proposition 2.70.

**2.78 Proposition.** (1) A subspace of a first-countable space is itself first-countable.

(2) A subspace of a second-countable space is itself second-countable.

The next theorem tells us that, by removing enough sets from it, any base in a second-countable space can be “pruned” so as to result in a countable base.

**2.79 Theorem.** In a second-countable space, each base contains some countable base.

**Proof.** Let  $\mathcal{B}$  be a base of the second-countable space  $X$ . Choose some countable base  $\mathcal{C}$  of  $X$ . Define

$$\mathcal{S} = \{\langle C, D \rangle \in \mathcal{C} \times \mathcal{C} : C \subset B \subset D \text{ for some } B \in \mathcal{B}\}.$$

Since  $\mathcal{C}$  is countable, so is  $\mathcal{S}$ . For each  $\langle C, D \rangle \in \mathcal{S}$ , choose some particular set  $B_{C,D} \in \mathcal{B}$  such that

$$C \subset B_{C,D} \subset D.$$

(Some form of the Axiom of Choice was just used!) Define

$$\mathcal{A} = \{B_{C,D} : \langle C, D \rangle \in \mathcal{S}\}.$$

Then  $\mathcal{A} \subset \mathcal{B}$  and, since  $\mathcal{S}$  is countable, so is  $\mathcal{A}$ .

We show that  $\mathcal{A}$  is a base of  $X$ . Let  $x \in X$  and let  $U$  be an open set with  $x \in U$ . Since  $\mathcal{C}$  is a base of  $X$ , there is some  $D \in \mathcal{C}$  with  $x \in D \subset U$ . Next, since  $\mathcal{B}$  is a base of  $X$ , there is some  $B \in \mathcal{B}$  with  $x \in B \subset D$ . Again since  $\mathcal{C}$  is a base of  $X$ , there is some  $C \in \mathcal{C}$  with  $x \in C \subset B$ . Thus

$$x \in C, \quad D \subset U, \quad \text{and} \quad C \subset B \subset D.$$

The last inclusions mean that  $\langle C, D \rangle \in \mathcal{S}$ . Then  $C \subset B_{C,D} \subset D$ . Hence  $B_{C,D} \in \mathcal{A}$  with  $x \in B_{C,D} \subset U$ .  $\square$

The preceding theorem is useful in showing that certain spaces are *not* second-countable.



ex:Rl-1st-not-2nd-countable **2.80 Example.** The Sorgenfrey line  $\mathbb{R}_l$  [Examples 2.20 (1)] is first-countable but *not* second countable. Sorgenfreyline

In fact, at each  $x \in \mathbb{R}_l$  the collection  $\{[x, r[ : r \in \mathbb{Q}, r > x\}$  is a countable local base.

Just suppose that  $\mathbb{R}_l$  were second-countable. Then by Theorem 2.79, the base  $\mathcal{B} = \{[x, y[ : x < y\}$  would have to contain some countable base  $\{[x_n, y_n[ : n \in \mathbb{N}\}$ . Choose any real number  $x$  with  $x \neq x_n$  for all  $n$ , and any  $y > x$ . There must be some  $n \in \mathbb{N}$  such that  $x \in [x_n, y_n[ \subset [x, y[$ ; but this is impossible.  $\diamond$

If  $\mathcal{B}$  is a base of a topological space  $X$ , then it is a collection of open subsets of  $X$  that *covers*  $X$  in the sense of Definition 0.23.

def:open-cover **2.81 Definition.** An **open cover** of a topological space  $X$  is a collection  $\mathcal{U}$  of *open* subsets of  $X$  that covers  $X$ .

The following example illustrates the terminology.

ex:covers-of-unit-interval **2.82 Example.** Let  $\varepsilon_0$  and  $\varepsilon_1$  be fixed real numbers with

$$0 < \varepsilon_0 < 1/2, \quad 0 < \varepsilon_1 < 1/2.$$

And for each  $x \in ]0, 1[$  let  $\varepsilon_x$  be any real number with

$$0 < \varepsilon_x < \min\{x, 1 - x\},$$

in other words, with  $\varepsilon_x$  strictly less than the distances of  $x$  from the endpoints 0 and 1 of the interval  $[0, 1]$ . Then:

- the collection

$$\{[0, \varepsilon_0[, ]1 - \varepsilon_1, 1]\}$$

is a collection of open subsets of  $[0, 1]$  that is *not* a cover of  $[0, 1]$ ;

- the collection

$$\{[0, \varepsilon_0[, ]1 - \varepsilon_1, 1]\} \cup \{]x - \varepsilon_x, x + \varepsilon_x[ : 0 < x < 1\}$$

is a cover of  $[0, 1]$ , but it is *not* an *open* cover of  $[0, 1]$  since its member  $[0, \varepsilon_0]$  is not open in  $[0, 1]$ ; and finally,

- the collection

$$\{[0, \varepsilon_0[, ]1 - \varepsilon_1, 1]\} \cup \{]x - \varepsilon_x, x + \varepsilon_x[ : 0 < x < 1\}$$

is an open cover of  $[0, 1]$ .  $\diamond$

As already noted, a base of a space is necessarily an open cover of the space. However, an open cover need not be a base.

**2.83 Example.** The collection  $\{]n, n + 1[ : n \in \mathbb{N}\}$  is a (countable) open cover of  $\mathbb{R}$ , but it is not a base of  $\mathbb{R}$ .  $\diamond$

The relevance of open covers to countability properties lies in the following theorem, which is an analog of Theorem 2.79 but for open covers instead of bases.

open cover  
 then Lindelöf  
 Lindelöf space@Lin  
 Hilbert sequence space

**2.84 Lindelöf Theorem.** *If a topological space is second-countable, then each open cover of the space contains some countable cover of the space.*

**Proof.** Let  $\mathcal{U}$  be an open cover of a second-countable topological space  $X$ . Choose some countable base  $\mathcal{B}$  of  $X$ . Define

$$\mathcal{A} = \{B \in \mathcal{B} : B \subset U \text{ for some } U \in \mathcal{U}\}.$$

Since  $\mathcal{A} \subset \mathcal{B}$ , the collection  $\mathcal{A}$  is countable. For each  $B \in \mathcal{A}$  choose some one  $U_B \in \mathcal{U}$  with

$$B \subset U_B.$$

Define

$$\mathcal{V} = \{U_B : B \in \mathcal{A}\}.$$

Then  $\mathcal{V} \subset \mathcal{U}$ , and  $\mathcal{V}$  is countable because  $\mathcal{A}$  is.

We show that  $\mathcal{V}$  is a cover of  $X$ . Let  $x \in X$ . By hypothesis,  $x \in U$  for some  $U \in \mathcal{U}$ . Since  $\mathcal{B}$  is a base of  $X$ , then  $x \in B \subset U$  for some  $B \in \mathcal{B}$ . Then  $B \in \mathcal{A}$ ,  $U_B \in \mathcal{V}$ , and  $x \in U_B$ .  $\square$

The preceding theorem is named after Ernst Lindelöf, who in 1903 proved it for Euclidean spaces. In the terminology of [Exercise 115](#) on page 289, the theorem asserts that a second-countable space is a **Lindelöf space**.

### Separable spaces

subsec:separable

Our third countability property uses the notion of a dense set ([Definition 2.42](#)).

def:separable

**2.85 Definition.** A topological space is said to be **separable** when some countable subset is dense in it.

exs:separable

**2.86 Examples.** (1) According to [Examples 2.43 \(1\)](#), the real line  $\mathbb{R}$  is separable.

ex:when-discrete-space-separable

(2) By [Examples 2.43 \(2\)](#), a discrete space  $X$  is separable if and only if the underlying set  $X$  is countable. Then in view of [Examples 2.76 \(2\)](#), a discrete space is separable if and only if it is second-countable.

ex:Rn-separable

(3) While proving in [Examples 2.76 \(1\)](#) that  $\mathbb{R}^n$  is second-countable, we showed that the countable set  $\mathbb{Q}^n$  is dense in  $\mathbb{R}^n$ . Hence Euclidean  $n$ -space  $\mathbb{R}^n$  is separable.

ex:ell2-separable

(4) The Hilbert sequence space  $\ell^2$  ([Example 1.10](#)) is separable.

Indeed, by analogy with the preceding example, we might expect that the set

$$\{y \in \ell^2 : y_i \in \mathbb{Q} \text{ for } i = 1, 2, 3, \dots\}$$

of all “rational points” in  $\ell^2$  is dense in  $\ell^2$ —and it is!—but unfortunately this set is uncountable.

We must consider instead the smaller set  $D$  consisting solely of those rational points in  $\ell^2$  that are eventually zero: function space

$$D = \bigcup_{n=1}^{\infty} D_n$$

where, for each  $n = 1, 2, 3, \dots$ ,

$$D_n = \{y \in \ell^2 : y_1, y_2, \dots, y_n \in \mathbb{Q}, y_{n+1} = y_{n+2} = \dots = 0\}.$$

For each  $n$  there is an obvious one-to-one correspondence between  $D_n$  and the denumerable set  $\mathbb{Q}^n$ , so that  $D_n$  is denumerable. Hence  $D$  is denumerable.

To see that  $D$  is dense in  $\ell^2$ , let  $x = \langle x_i \rangle_{i=1,2,3,\dots} \in \ell^2$  and let  $\varepsilon > 0$ . Since the series  $\sum_{i=1}^{\infty} x_i^2$  converges, there exists some  $n \geq 1$  for which

$$\sum_{i=n+1}^{\infty} x_i^2 < \frac{\varepsilon^2}{2}.$$

For each  $i = 1, 2, \dots, n$  there is a rational number  $y_i$  with

$$|x_i - y_i|^2 < \frac{\varepsilon^2}{2n}.$$

Then the point  $y = \langle y_i \rangle_{i=1,2,3,\dots}$  with  $y_i = 0$  for all  $i > n$  belongs to  $D_n$ , and

$$\begin{aligned} [d_2(x, y)]^2 &= \sum_{i=1}^{\infty} |x_i - y_i|^2 \\ &= \sum_{i=1}^n |x_i - y_i|^2 + \sum_{i=n+1}^{\infty} x_i^2 \\ &< n \frac{\varepsilon^2}{2n} + \frac{\varepsilon^2}{2} = \varepsilon^2. \end{aligned}$$

- ex:C01-separable (5) The space  $C([0, 1])$  of all continuous real-valued functions on  $[0, 1]$ , with the topology induced by the sup metric (Example 1.8), is separable.

One countable dense subset of  $C([0, 1])$  consists of all those functions  $f: [0, 1] \rightarrow \mathbb{R}$  having the following property: for some positive integer  $n$ , the function  $f$  takes rational values at each of the points  $0, 1/n, 2/n, \dots, 1$  and is linear on each of the intervals  $[0, 1/n], [1/n, 2/n], \dots, [(n-1)/n, 1]$ . The proof requires ideas developed in Section 4.2 of Chapter 4 (Compactness): see Application 4.66.

- (6) The space  $\mathcal{B}([0, 1])$  of all bounded real-valued functions on  $[0, 1]$ , with its topology induced by the sup metric  $d_{\infty}$  (see Example 1.16 and the discussion preceding Theorem 1.83), is *not* separable. In fact, consider for each  $x \in [0, 1]$  the characteristic function  $f_x$  of  $\{x\}$  given by

$$f_x(x) = 1, \quad f_x(y) = 0 \quad (y \neq x).$$

Any dense set  $D$  in  $\mathcal{B}([0, 1])$  contains for each  $x \in [0, 1]$  some function  $g_x$  such that

$$d_{\infty}(f_x, g_x) < \frac{1}{2}.$$

Now  $x, y \in [0, 1]$  with  $x \neq y$  implies  $d_{\infty}(f_x, f_y) = 1$  and so  $g_x \neq g_y$ . Thus such a dense set  $D$  contains the uncountable set  $\{g_x : x \in [0, 1]\}$ , and so  $D$  cannot be countable.  $\diamond$

Examples Examples 2.86 (2)–(3) suggest a general fact.

metrizable-separable-iff-2nd-countable  
 half-disk space!sep:  
 separable space!su  
 subspace!separable space@or separable space

**2.87 Theorem.** A metrizable space is separable if and only if it is second-countable.

**Proof.** Let  $X$  be a metrizable space. Choose some metric  $d$  that induces the topology of  $X$ .

Assume first that  $X$  is second-countable. Let  $\langle B_n \rangle_{n \in \mathbb{N}}$  be a sequence of nonempty open subsets of  $X$  such that  $\{B_n : n \in \mathbb{N}\}$  is a base of  $X$ . For each  $n \in \mathbb{N}$ , choose some point  $x_n \in B_n$ . Then  $\{x_n : n \in \mathbb{N}\}$  is a countable set which is easily seen to be dense in  $X$ . Hence  $X$  is separable.

Conversely, assume that  $X$  is separable. Let  $D$  be some countable dense subset of  $X$ . Define the collection  $\mathcal{B}$  of open subsets of  $X$  by

$$\mathcal{B} = \{B_r(z; d) : z \in D, 0 < r \in \mathbb{Q}\}.$$

Since both  $D$  and  $\mathbb{Q}$  are countable, so is  $\mathcal{B}$ .

To complete the proof that  $X$  is second-countable, we shall show that  $\mathcal{B}$  is a base of  $X$  by verifying condition ((ii)) of Proposition 2.69. Let  $x \in X$  and let  $V$  be a neighborhood of  $x$ . We want to show that  $x \in B \subset V$  for some  $B \in \mathcal{B}$ . Choose  $\varepsilon > 0$  with  $B_\varepsilon(x; d) \subset V$ . Since  $D$  is dense in  $X$ , there is some point  $z \in D \cap B_{\varepsilon/2}(x; d)$ . Choose a rational number  $r$  with  $d(z, x) < r < \varepsilon/2$ . Then  $x \in B_r(z; d) \in \mathcal{B}$  and  $B_r(z; d) \subset B_\varepsilon(x; d) \subset V$ , as needed.  $\square$

According to Proposition 2.78, every subspace of a first-countable space is first-countable, and every subspace of a second-countable space is second-countable. By contrast, **a subspace of a separable space need not be separable**, as the following example shows. (A different example appears in Exercise 3.114.)

ex:separability-does-not-inherit

**2.88 Example.** Let  $X = L \cup H$  be the half-disk space of Examples 2.20 (3). The entire space  $X$  is separable because its countable subset

$$\{\langle x, y \rangle \in \mathbb{Q} \times \mathbb{Q} : y > 0\}$$

is dense in  $X$ . Now the relative topology on the subspace  $L = \mathbb{R} \times \{0\}$  is the discrete topology; in fact, for each  $x \in L$  the set

$$H_1(x) = (B_1(x; d) \cap H) \cup \{x\}$$

is an open neighborhood of  $x$  in  $X$ , and so its subset

$$H_1(x) \cap L = \{x\},$$

is an open neighborhood of  $x$  in  $L$ . According to Examples 2.86 (2), an uncountable discrete space cannot be separable. Hence the subspace  $L$  of  $X$  is *not* separable.  $\diamond$

The first two of the following three statements are immediate consequences of Proposition 2.73; for the third, see Exercise 108.

prod-of-2-proper-part-countability-props

**2.89 Proposition.** (1) The product of two first-countable spaces is itself first-countable.

prop-part:prod-2-2nd-count

(2) The product of two second-countable spaces is itself second-countable.

prop-part:prod-2-separable

(3) The product of two separable spaces is itself separable.

## Subbases

subsec:subbases

In general, to obtain the order topology on a totally ordered set  $X$ , we need not only all the open rays  $] \leftarrow, y[$  and  $]x, \rightarrow[$  but also all open intervals  $]x, y[$ . As noted in Examples 2.72 (1),

however, each open interval is the intersection of two rays, so in a sense the open intervals are superfluous for determining the topology; more precisely, the open rays form a “subbase” of the topology.

cover!subbase@and subbase  
subbase!cover@and cover  
cover!subbase@and subbase  
subbase!cover@and cover  
cover!subbase@and subbase  
subbase!cover@and cover

**2.90 Definition.** Let  $X$  be a topological space with topology  $\mathcal{T}$ . A collection  $\mathcal{S}$  of subsets of  $X$  is called a **subbase of  $X$**  (and **of  $\mathcal{T}$** ) if the collection consisting of all intersections of finitely many sets belonging to  $\mathcal{S}$  is a base of  $X$ . This base is said to be **associated with** the subbase  $\mathcal{S}$ .

When the definition speaks of “finitely many” sets belong to  $\mathcal{S}$ , it means as usual a *nonempty* finite subcollection  $\mathcal{E}$  of  $\mathcal{S}$ .

The definition includes the provision that each member of  $\mathcal{S}$  be open in the space, in other words, that  $\mathcal{S} \subset \mathcal{T}$ . Condition (B1) required of a base implies that for  $\mathcal{S}$  to be a subbase of the topology of  $X$ , it is also necessary that each point of  $X$  belong to at least one member of  $\mathcal{S}$ , that is,  $\bigcup \mathcal{S} = X$ —in other words, that  $\mathcal{S}$  is a cover of  $X$  (Definition 0.23).

**2.91 Examples.** (1) Any base of a topology is a subbase of that topology.

(2) Let  $X = \{0, 1, 2, 3, 4, 5\}$ . As you may readily check, the collection

$$\mathcal{T} = \{\emptyset, \{0\}, \{1, 2\}, \{0, 1, 2\}, \{1, 2, 3, 4, 5\}, X\}.$$

is a topology on  $X$ . This topology has the collection

$$\mathcal{S} = \{\{0\}, \{0, 1, 2\}, \{1, 2, 3, 4, 5\}\}$$

as a subbase. However,  $\mathcal{S}$  is *not* a base of  $\mathcal{T}$ , because  $U = \{1, 2\}$  is  $\mathcal{T}$ -open and  $1 \in U$  but there is no  $B \in \mathcal{S}$  with  $1 \in B \subset U$ .

(3) As the discussion preceding Definition 2.90 indicates, the collection

$$\{]-\infty, b[ : b \in \mathbb{R}\} \cup \{]a, +\infty[ : a \in \mathbb{R}\}$$

of all open rays is a subbase of the usual topology of  $\mathbb{R}$ . However, it is *not* a base of that topology.

(4) The collection

$$\{]a, b[ \times \mathbb{R} : a < b\} \cup \{\mathbb{R} \times ]c, d[ : c < d\}$$

of all open horizontal and vertical “infinite strips” is a subbase of the usual topology of the Euclidean plane  $\mathbb{R}^2$ . In fact, the set of open rectangles  $]a, b[ \times ]c, d[$  is a base of that topology, and any such open rectangle is the intersection of a horizontal infinite strip and a vertical infinite strip.  $\diamond$

fix: Add fig for  
strips &  
rectangles as base  
of plane?

An especially simple way to describe a topology on a set  $X$  is to start with *any* collection  $\mathcal{S}$  of subsets of  $X$  that covers  $X$  in the sense of Definition 0.23, that is,  $\bigcup \mathcal{S} = X$ .

**2.92 Proposition.** Let  $\mathcal{S}$  be a collection of subsets of a set  $X$  that covers  $X$ . Then there is a unique topology  $\mathcal{T}$  on  $X$  of which  $\mathcal{S}$  is a subbase. Moreover,  $\mathcal{T}$  is the smallest topology on  $X$  that contains  $\mathcal{S}$ , that is:

- (i)  $\mathcal{S} \subset \mathcal{T}$ ; and
- (ii) if  $\mathcal{T}'$  is a topology on  $X$  such that  $\mathcal{S} \subset \mathcal{T}'$ , then  $\mathcal{T} \subset \mathcal{T}'$ .

**Proof.** Define  $\mathcal{B}$  to be the collection consisting of all intersections of nonempty finite subcollections of  $\mathcal{S}$ . Then already  $\mathcal{S} \subset \mathcal{B}$ .  
 Existence. We show that  $\mathcal{B}$  is a base of some topology on  $X$  by showing that it satisfies the conditions (B1) and (B2) of Theorem 2.71.

Condition (B1) holds because  $\mathcal{S}$  covers  $X$  by hypothesis and  $\mathcal{S} \subset \mathcal{B}$  by construction.

To establish condition (B2), it suffices to show that the intersection of any two members of  $\mathcal{B}$  is itself a member of  $\mathcal{B}$ . Let  $B_1, B_2 \in \mathcal{B}$  and let  $B = B_1 \cap B_2$ . There are nonempty finite subcollections  $\mathcal{F}_1$  and  $\mathcal{F}_2$  of  $\mathcal{S}$  such that  $B_1 = \bigcap \mathcal{F}_1$  and  $B_2 = \bigcap \mathcal{F}_2$ . Then

$$B = \left( \bigcap \mathcal{F}_1 \right) \cap \left( \bigcap \mathcal{F}_2 \right) = \bigcap (\mathcal{F}_1 \cup \mathcal{F}_2)$$

with  $\mathcal{F}_1 \cup \mathcal{F}_2$  being another nonempty finite subcollection of  $\mathcal{S}$ . Thus  $B$  is also a member of  $\mathcal{B}$ .

For the remaining parts of the proof, let  $\mathcal{T}$  be the topology having the collection  $\mathcal{B}$  as a base.

Uniqueness. Let  $\mathcal{T}'$  be an arbitrary topology on  $X$  that has  $\mathcal{S}$  as a subbase. According to Definition 2.90, the collection  $\mathcal{B}$  is a base of  $\mathcal{T}'$ . Since  $\mathcal{B}$  is also a base of  $\mathcal{T}$ , it follows that  $\mathcal{T}' = \mathcal{T}$ .

Characterization. Since  $\mathcal{S}$  is a subbase of  $\mathcal{T}$ , we do have  $\mathcal{S} \subset \mathcal{T}$ .

Now let  $\mathcal{T}'$  be an arbitrary topology on  $X$  such that  $\mathcal{S} \subset \mathcal{T}'$ . By property (O3) of a topology the intersection of each finite subcollection of  $\mathcal{S}$  is also  $\mathcal{T}'$ -open; in other words,  $\mathcal{B} \subset \mathcal{T}'$ . Next, by property (O2) of a topology the union of each subcollection of  $\mathcal{B}$  is  $\mathcal{T}'$ -open. But the collection of such unions is the topology  $\mathcal{T}$ , and hence  $\mathcal{T} \subset \mathcal{T}'$ .  $\square$

In the notation of Proposition 2.92, we call  $\mathcal{T}$  the topology **generated by**  $\mathcal{S}$  and say that the cover  $\mathcal{S}$  of  $X$  **generates**  $\mathcal{T}$ . Because the given cover  $\mathcal{S}$  is a subbase of  $\mathcal{T}$ , then  $\mathcal{T}$  has as a base the collection of all intersections of nonempty finite subcollections of  $\mathcal{S}$ .

**2.93 Examples.** (1) Let  $\mathcal{S}$  be the collection of all vertical lines  $\{a\} \times \mathbb{R}$  in the plane  $\mathbb{R}^2$ . Then  $\mathcal{S}$  covers  $\mathbb{R}^2$  and hence forms a subbase of a topology  $\mathcal{T}$  of  $\mathbb{R}^2$ . This topology is *not* the usual topology because each such vertical line is open for  $\mathcal{T}$  but no such vertical line is open for the usual topology.

(2) Let  $X$  be a nonempty set and let  $Y$  be a topological space. Let  $\mathcal{F}(X, Y)$  be the set of all maps  $f: X \rightarrow Y$ ; in other words,  $\mathcal{F}(X, Y) = Y^X$ , the  $X$ th power of  $Y$  (see Definition 0.15). Define  $\mathcal{S}$  be the collection of all those subsets of  $\mathcal{F}(X, Y)$  that have the form

$$B(x, V) = \{f \in \mathcal{F}(X, Y) : f(x) \in V\},$$

where  $x \in X$  and  $V$  is an open set in  $Y$ . Then  $\mathcal{S}$  covers  $\mathcal{F}(X, Y)$  because if  $f \in \mathcal{F}(X, Y)$  and if we choose some  $x_0 \in X$ , then  $f \in B(x_0, Y)$ . The topology generated by  $\mathcal{S}$  is called the **topology of pointwise convergence on**  $\mathcal{F}(X, Y)$ .

The base  $\mathcal{B}$  of this topology associated with the subbase  $\mathcal{S}$  consists of all sets of the form

$$\begin{aligned} B(x_1, x_2, \dots, x_k; I_1, I_2, \dots, I_k) &= \bigcap_{j=1}^k B(x_j, V_j) \\ &= \{f \in \mathcal{F}(X, Y) : f(x_j) \in V_j \text{ for each } j = 1, 2, \dots, k\}, \end{aligned}$$

where  $k$  is a positive integer,  $\langle x_1, x_2, \dots, x_k \rangle$  is a  $k$ -tuple of elements of  $X$ , and  $\langle V_1, V_2, \dots, V_k \rangle$  is a  $k$ -tuple of open subsets of  $Y$ . Thus the topology of pointwise

convergence generalizes the topology of the same name defined in [Examples 2.72 \(8\)](#) on the set of all functions from  $\mathbb{R}$  to  $\mathbb{R}$ . Accordingly, a basic neighborhood of an element  $f$  of  $\mathcal{F}(X, Y)$  may be visualized, as they were for that special case  $X = Y = \mathbb{R}$ , as in [Figure 2.12](#).

The topology of pointwise convergence on  $\mathcal{F}(X, Y)$  just defined will itself be generalized in the [subsection “The product topology”](#) of [Section 3.3](#) to a topology on the product of sets underlying an arbitrary family of topological spaces.  $\diamond$

Subbases will constitute an essential tool for proving a central result of topology—the Tychonoff Product Theorem ([4.33](#))—that we shall offer in [Section 6.1](#).

### EXERCISES FOR SECTION 2.4

**72.** Give a family of local bases for the space of [Exercise 34](#) different from the family of its neighborhood systems.

**73.** Use [Proposition 2.59](#) to construct anew the topology of the half-disk space [[Examples 2.25 \(3\)](#)] from a family of local bases different from the family of its neighborhood systems.

**74.** Use [Proposition 2.59](#) to construct anew the topology of the tangent disk space ([Exercise 37](#)) from a family of local bases different from the family of its entire neighborhood systems.

**75.** Let  $\mathcal{T}_1$  and  $\mathcal{T}_2$  be two topologies on the same set  $X$  and let  $x \in X$ . Suppose  $\mathcal{M}_1$  and  $\mathcal{M}_2$  are local bases at  $x$  for  $\mathcal{T}_1$  and  $\mathcal{T}_2$ , respectively. In terms of  $\mathcal{M}_1$  and  $\mathcal{M}_2$ , when is the  $\mathcal{T}_1$ -neighborhood system at  $x$  the same as the  $\mathcal{T}_2$ -neighborhood system at  $x$ ?

**76.** Prove that the space  $\mathbb{R} \setminus \mathbb{Q}$  of irrational numbers is zero-dimensional. [This is [Examples 2.63 \(3\)](#).]

**77.** Which of the following spaces are zero-dimensional?

- (a) The space  $\mathbb{Q}^n$  consisting of all points in  $\mathbb{R}^n$  all of whose coordinates are rational.
- (b) The space  $\mathbb{R}^n \setminus \mathbb{Q}^n$  consisting of all points in  $\mathbb{R}^n$  all of whose coordinates are irrational.
- (c) The subspace of  $\mathbb{R}^n$  consisting of all points having at least one coordinate rational.

**78.** Prove [Proposition 2.64](#): A subspace of a zero-dimensional space is zero-dimensional.

**79.** Show that the following spaces are not zero-dimensional:

- (a) A circle in  $\mathbb{R}^2$ .
- (b) A square in  $\mathbb{R}^2$  whose sides are perpendicular to the axes.

**80.** Show that the following nondiscrete spaces are zero-dimensional:

- (a) For a nondegenerate finite discrete space  $Y$ , the collection  $\mathcal{F}$  of all functions  $f: \mathbb{N} \rightarrow Y$ , provided with its topology of pointwise convergence—see [Examples 2.93 \(2\)](#).
- (b) The subspace of the Hilbert sequence space  $\ell^2$  ([Example 1.10](#)) consisting of those square-summable sequences all of whose coordinates are rational.

subbase

half-disk space!local bases@andto

tangent disk space!local bases@and

zero-dimensional space

zero-dimensional space

prob:half-disk-space-via-local-bases

prob:tangent-disk-space-via-local-bases

prob:irrationals-0-dim

prob-part:Erdos-nonzero-dim

- Fort topology  
Thomas plank  
equivalent metrics  
Euclidean  $n$ -space
81. (a) Is the Fort topology [Examples 2.66 (3)] on a countable set necessarily first-countable?  
(b) Is every topological space whose underlying set is countable a first-countable space?
- order topology!  
Hausdorff space  
and Hausdorff space  
prob:Thomas-plank-not-1st-count  
lexicographically ordered square
82. (a) Show that the Thomas plank [Examples 2.60 (4)] is not first-countable (and hence not metrizable).  
(b) Is the Thomas plank separable?
- lexicographic ordering  
subspace!  
ordered space  
of ordered space  
g-seq-at-local-base-in-1st-countable
83. Let  $X$  be a first-countable space and let  $x \in X$ . Prove that there is a sequence  $\langle V_n \rangle_{n \in \mathbb{N}}$  such that  $\{V_n : n \in \mathbb{N}\}$  is a local base at  $x$  and  $V_{n+1} \subset V_n$  for every  $n \in \mathbb{N}$ . Can the sets  $V_n$  all chosen to be open? Can they be chosen such that  $\bigcap_{n=0}^{\infty} V_n = \{x\}$ ?
- prob:pf-properties-of-local-bases
84. Verify the properties of local bases listed in Proposition 2.58.
85. Let  $\mathcal{B}$  be a base of a topological space  $X$ .  
(a) Prove: A subset  $U$  of  $X$  is open in  $X$  if and only if for each  $x \in U$  there is some  $B \in \mathcal{B}$  such that  $x \in B \subset U$ .  
(b) When, in terms of  $\mathcal{B}$ , is a subset  $E$  of  $X$  closed in  $X$ ?  
(c) Do properties ((B1))–((B2)) of Theorem 2.71 necessarily hold for  $\mathcal{B}$ ?
- t:two-bases-two-topologies-contain
86. (a) Let  $\mathcal{B}_1$  and  $\mathcal{B}_2$  be bases of topologies  $\mathcal{T}_1$  and  $\mathcal{T}_2$ , respectively, on the same set  $X$ . Show that  $\mathcal{T}_1 \subset \mathcal{T}_2$  if and only if, for each  $B_1 \in \mathcal{B}_1$  and each  $x \in B_1$ , there exists some  $B_2 \in \mathcal{B}_2$  with  $x \in B_2 \subset B_1$ .  
(b) By applying (a), show that two metrics  $d_1$  and  $d_2$  on a set  $X$  induce the same topology in case there are positive constants  $\alpha$  and  $\beta$  such that  $\alpha d_1(x, y) \leq d_2(x, y) \leq \beta d_1(x, y)$  for all  $x, y \in X$ .
87. Show that each of the following collections is a base of  $\mathbb{R}^n$ :  
(a) The collection of all  $n$ -dimensional “open parallelepipeds”  
$$]a_1, b_1[ \times ]a_2, b_2[ \times \cdots ]a_n, b_n[.$$
  
(b) The collection of all open sets in  $\mathbb{R}^n$  that are convex.
- prob:base-RI-clopen
88. Justify the assertion made in Examples 2.68 (6) that each member of the base  $\{[x, y[ : x < y\}$  of the Sorgenfrey line  $\mathbb{R}_l$  is closed in  $\mathbb{R}_l$ .
- prob:order-top-T2
89. Prove the assertion made in Examples 2.72 (1) that a totally ordered set provided with its order topology is a Hausdorff space.
- prob:lex-square
90. Give the product set  $\mathbb{R} \times \mathbb{R}$  the order topology induced by the lexicographic ordering (Example 0.70), where each factor has its usual ordering; we may call the resulting space  $X$  the lexicographically ordered plane. Give the set  $[0, 1] \times [0, 1]$  its own lexicographic ordering (which is the ordering inherited from the lexicographic ordering of  $\mathbb{R} \times \mathbb{R}$ ); the resulting topological space, denoted by  $l_{\text{lex}}^2$ , is called the **lexicographically ordered square**.  
Is the given topology on  $l_{\text{lex}}^2$  the relative topology inherited from  $X$ ?
91. As in Examples 2.72 (2), let  $X$  be a totally ordered set, provide  $X$  with its order topology, and let  $Y$  be a subset of  $X$ .  
(a) Show that, in general, the relative topology on  $Y$  is finer than the order topology of  $Y$  obtained from the induced ordering of  $Y$ .



- (b) The two topologies on  $Y$  are identical in case  $Y$  has the property that for each  $a, b \in Y$  with  $a < b$ , the interval  $]a, b[$  in  $X$  is contained in  $Y$ .

deleted sequence space  
Sorgenfrey line

prob:deleted-seq-space

92. Let  $K = \{1/n : n = 1, 2, 3, \dots\}$ .

- (a) Show that the collection consisting of all bounded open intervals  $]a, b[$  together with all sets of the form  $]a, b[ \setminus K$  for  $a, b \in \mathbb{R}$  is a topology on the set of real numbers.

lexicographic ordering  
order topology

This topology is variously known as the **deleted sequence topology** or the  **$K$ -topology**; the resulting topological space is known as the **deleted sequence space** or the  **$K$ -space** and is sometimes denoted by  $\mathbb{R}_K$ .

order topology  
finite-complement topology!second-countable  
finite-complement topology!second-countable  
countable-complement topology!second-countable  
half-disk space!second-countable  
extended real line!second-countable  
deleted sequence space!second-countable

- (b) Show that the collection of all sets of the form  $U \setminus A$ , where  $U$  is open in  $\mathbb{R}$  for the usual topology and  $A \subset K$ , is a base of the deleted sequence topology.

- (c) Compare with respect to inclusion the usual, right-interval, and deleted sequence topologies on the set of real numbers.

93. Are properties (N1)–(N5) from Theorem 2.18 still satisfied if for each real number  $x$  we define  $\mathcal{N}_x$  to be the collection of all subsets of  $\mathbb{R}$  that contain a closed interval of the form  $[x, z]$  with  $z > x$ ?

RI-separable-1st-not-2nd-countable

94. Show that the Sorgenfrey line  $\mathbb{R}_l$  [Examples 2.20 (1)] is separable but not metrizable.

prob:lexicographic-topology-plane

95. Give the plane  $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$  the lexicographic ordering  $\leq$  (Example 0.70) induced by the usual ordering of  $\mathbb{R}$ , that is, for  $x = \langle x_1, x_2 \rangle, y = \langle y_1, y_2 \rangle \in \mathbb{R}^2$ ,

$$x \leq y \iff (x_1 < y_1) \text{ or } (x_1 = y_1 \text{ and } x_2 \leq y_2).$$

- (a) Which subsets of  $\mathbb{R}^2$  that are open for the resulting order topology [Examples 2.72 (1)] are open for the right-interval topology [Examples 2.72 (2)] obtained from this same lexicographic ordering of  $\mathbb{R}^2$ , and vice versa?

- (b) When  $\mathbb{R}^2$  is provided with this order topology, what is the relative topology induced on the horizontal line  $\mathbb{R} \times \{0\}$ ? when  $\mathbb{R}^2$  is provided with the right-interval topology?

- (c) When  $\mathbb{R}^2$  is provided with this order topology, what is the relative topology induced on the line  $\{(x, y) \in \mathbb{R} \times \mathbb{R} : x = y\}$ ? when  $\mathbb{R}^2$  is provided with the right-interval topology?

96. Let  $X$  be a totally ordered set and let  $Y$  be a subset of  $X$  having at least two elements. Then on  $Y$  we have the relative topology induced by the order topology [Examples 2.72 (1)] of  $X$ , and the order topology defined by the restriction of the total ordering of  $X$  to  $Y$ . Show that these two topologies may be distinct, even when  $X = \mathbb{R}$  with its usual topology.

prob:examples-countability-properties

97. Which of the following spaces are first-countable? second-countable? separable?

- (a) A denumerable set with its finite-complement topology [Examples 2.3 (7)].

- (b) An uncountable set with its finite-complement topology.

- (c) An uncountable set with its countable-complement topology [Exercise 7].

- (d) The half-disk space [Examples 2.20 (3)].

- (e) The extended real line (see Example 1.41 and Examples 2.72 (3)).

- (f) The deleted sequence space  $\mathbb{R}_K$  (Exercise 92).

finite-complement-top-countability-properties

98. Exhibit a countable base of the tangent disk space  $\Gamma$  (Exercise 37), thereby showing that  $\Gamma$  is second-countable.

99. Show that the tangent disk space  $\Gamma$  (Exercise 37) is not metrizable.

100. Show that the tangent disk space  $\Gamma$  (Exercise 37) is locally metrizable, that is, each of its points has a metrizable neighborhood (and hence each neighborhood of each point contains a metrizable neighborhood).

[Hint: : If  $z \in H$ , the open half-plane, then  $z$  has as a neighborhood in  $\Gamma$  a Euclidean ball, which is metrizable. Now suppose that  $z = \langle a, 0 \rangle \in L$ , the  $x$ -axis. Choose any basic open neighborhood  $T_r(z) = \{z\} \cup B_r(z; d)$  of  $z$  in  $\Gamma$ . For  $w = \langle u, v \rangle \in B_r(z; d)$ , there is a unique  $\varepsilon > 0$  for which  $z \in S_\varepsilon(\langle a, \varepsilon \rangle; d)$ , namely,  $\langle a, \varepsilon \rangle$  is the point where the perpendicular bisector of the line segment from  $w$  to  $z$  intersects the vertical line through  $z$ ; denote that  $\varepsilon$  by  $\rho(w)$ . Further, set  $\rho(z) = 0$ . Define  $d$  by

$$d(w, w') = \sqrt{d(w, w')^2 + (\rho(w) - \rho(w'))^2}.$$

Show that  $d$  is a metric on  $T_r(z)$  inducing the relative topology of  $T_r(z)$  in  $\Gamma$ .]

Note: A quick way to obtain the same result is given in Exercise 6.58.

101. Prove: Each *disjoint* open cover of a second-countable space is countable.

102. This problem concerns countability property of ordinal spaces [Examples 2.72 (4)].

(a) Show that the ordinal space  $\Omega = [0, \Omega[$  is first-countable but neither separable nor second-countable.

(b) Is the ordinal space  $\Omega^+ = [0, \Omega]$  first-countable? separable? second-countable?

(c) Is the ordinal space  $\omega^+ = [0, \omega]$  first-countable? separable? second-countable?

103. Which, if any, of the ordinal spaces  $\omega$ ,  $\omega^+$ ,  $\Omega$ , and  $\Omega^+$  are scattered (Exercise 33).

104. Show that a second-countable scattered space (Exercise 33) must be countable.

105. Let  $A$  and  $B$  be subsets of topological spaces  $X$  and  $Y$ , respectively. Form the product space  $X \times Y$  [Examples 2.72 (5)]. Prove:

(a)  $\text{int}(A \times B) = (\text{int } A) \times (\text{int } B)$ .

(b)  $\text{cls}(A \times B) = (\text{cls } A) \times (\text{cls } B)$ .

(c)  $\text{bdy}(A \times B) = ((\text{bdy } A) \times (\text{cls } B)) \cup ((\text{cls } A) \times (\text{bdy } B))$ .

106. Show by example that for subsets  $A$  and  $B$  of topological spaces  $X$  and  $Y$ , respectively, it need not be the case that  $\text{bdy}(A \times B) = (\text{bdy } A) \times (\text{bdy } B)$  in the product space  $X \times Y$ . [Compare formula (c) in Exercise 105.]

107. Prove Proposition 2.73.

108. Prove the following results about the product of two spaces [Examples 2.72 (5)], the first of which was stated as Proposition 2.89 (3):

(a) The product of two separable spaces is separable.

(b) The product of two metrizable spaces is metrizable.

109. (a) Generalize the product topology of Examples 2.72 (5) to the case of an arbitrary finite number of spaces.

(b) Extend Proposition 2.89 to the product of finitely many spaces.

(c) Similarly extend [Exercise 150](#).

prob:diag-closed-for-T2 **110.** Prove the closed diagonal criterion ([Proposition 2.74](#)): A topological space  $X$  is a  $T_2$ -space if and only if the diagonal  $\Delta_X$  is closed in  $X \times X$ .

prob:Fortissimo-space **111.** Let  $X$  be an infinite set and let  $p \in X$  be some specific point. Define  $\mathcal{T}$  to be the collection of all those subsets  $U$  of  $X$  for which  $X \setminus U$  is countable or  $p \notin U$ .

(a) Verify that  $\mathcal{T}$  is a topology on  $X$ .

We call this topology the **Fortissimo topology**; when the set  $X$  is uncountable, we refer to  $\langle X, \mathcal{T} \rangle$  as a **Fortissimo space**. ('Fortissimo' is a pun on the name 'Fort' in 'Fort topology'.) Just as the Fort topology [[Examples 2.66 \(3\)](#)] is a modification of the finite-complement topology [[Examples 2.3 \(7\)](#)], so the Fortissimo topology is a modification of the countable-complement topology [[Exercise 7](#)].

(b) Assume that  $X$  is uncountable. Is the Fort topology on  $X$  first-countable? separable? second-countable?

**112.** Let  $X = \sum_{i \in I} X_i$  be the Cartesian sum ([Exercise 18](#)) of a family  $\langle X_i \rangle_{i \in I}$  of pairwise disjoint spaces. Discuss the relationship of each of the countability properties for  $X$  to the corresponding property for the spaces  $X_i$ .

**113.** Does the Lindelöf Theorem ([2.84](#)) remain valid if the hypothesis that  $X$  be second-countable is weakened to  $X$  being first-countable? Does it remain valid if  $X$  is metrizable?

**114.** Prove the following converse of the Lindelöf Theorem ([2.84](#)): If each open cover of a metrizable space  $X$  contains a countable cover of  $X$ , then  $X$  is second-countable.

The following Exercises [115–117](#) are included here so as to prepare for [Exercise 121](#); the latter is referenced only for verifying properties of the counterexample [Example 3.88](#).

prob:Lindelof-space **115.** A topological space is called a **Lindelöf space** when each open cover of the space contains a countable cover.

*Note:* According to the Lindelöf Theorem ([2.84](#)), a second-countable space is a Lindelöf space. Thus the property of being a Lindelöf space generalizes second-countability.

prob:Lindelof-via-base (a) Prove: If a topological space  $X$  has some base  $\mathcal{B}$  with the property that each cover of  $X$  by members of  $\mathcal{B}$  contains a countable cover of  $X$ , then  $X$  must be a Lindelöf space. *Note:* This is an analog of [Exercise 4.9](#).

prob-part:RI-Lindelof (b) Show that the Sorgenfrey line [[Examples 2.20 \(1\)](#)], which is not second-countable, is a Lindelöf space. *Note:* For a stronger result, see [Exercise 117](#), below. [*Hint:* Use [Examples 2.68 \(6\)](#).]

prob:Lindelof-subspace **116.** (Continuation of [Exercise 115](#).)

ed-subspace-Lindelof-then-Lindelof (a) Show that a *closed* subspace of a Lindelöf space is necessarily a Lindelöf space. *Note:* Contrast the situation for an open subspace: see [Exercise 4.8 \(b\)](#).

(b) Show that a product of Lindelöf spaces need not be a Lindelöf space by proving that the Sorgenfrey plane [[Examples 2.72 \(7\)](#)] is not a Lindelöf space. [*Hint:* Apply (a) to the reverse diagonal of the Sorgenfrey plane.]

Fortissimo topology

Fortissimo space

Cartesian sum!family of spaces@of a

Lindelof Theorem@Lindelof of Theor

Lindelof Theorem@Lindelof of Theor

Lindelof space@Lindelof of space

Lindelof space@Lindelof of space

Sorgenfreyline!Lindelof space@aski

Lindelof space@Lindelof of space!sub

Sorgenfrey plane!Lindelof@and!ind

Lindelof space@Lindelof of space!pro

117. (Continuation of Exercise 115.)

According to Exercise 4.8, a subspace of a Lindelöf space need not be a Lindelöf space. Show, though, that every subspace of the Sorgenfrey line  $\mathbb{R}_l$  is a Lindelöf space.

(Hint: Show that each collection  $\mathcal{U}$  of open subsets of  $\mathbb{R}_l$  contains a countable subcollection  $\mathcal{V}$  with  $\bigcup \mathcal{U} = \bigcup \mathcal{V}$ .)

118. (Continuation of Exercise 115.)

Show that the ordinal space  $\Omega^+ = [0, \Omega]$  is a Lindelöf space but its open subspace  $\Omega = [0, \Omega[$  is not.

119. For a subset  $A$  of a topological space  $X$ , a **condensation point of  $A$  (in  $X$ )** is a point  $x \in X$  such that each neighborhood of  $x$  contains uncountably many points of  $A$ . For example, each point of  $[0, 1]$  is a condensation point of  $]0, 1[$  in  $\mathbb{R}$ .

- Give an example of a limit point of an infinite subset of  $\mathbb{R}$  that is not a condensation point of that subset.
- A condensation point of a subset  $A$  of a space  $X$  is necessarily an  $\omega$ -accumulation point of that subset in the sense of Exercise 27. Show by example that the converse is not the case.
- Give an example of an uncountable subspace  $A$  of a Hausdorff space  $X$  that has a limit point in  $X$  but no condensation points there.

120. (Continuation of Exercise 119.) Prove that an uncountable subset  $E$  of a second-countable space  $X$  contains a condensation point of itself in  $X$  and that, in fact, only countably many points of  $E$  are *not* condensation points of  $E$ .

(Hint: Choose a countable base  $\mathcal{B}$  of  $X$  and let  $U = \bigcup \{B \in \mathcal{B} : B \cap E \text{ is countable}\}$ . Show that  $X \setminus U$  is the set of condensation points of  $E$ .)

121. A **two-sided condensation point** of a subset  $A$  of  $\mathbb{R}$  is an  $x \in \mathbb{R}$  such that, for each neighborhood  $V$  of  $x$ , both  $\{y \in V \cap A : y < x\}$  and  $\{y \in V \cap A : y > x\}$  are uncountable. (Thus a two-sided condensation point of  $A$  is a condensation point of  $A$  in the sense of Exercise 120.)

Prove that each uncountable subset  $A$  of  $\mathbb{R}$  contains a two-sided condensation point of  $A$ .

(Note: Then each uncountable subset of  $\mathbb{R}$  has a limit point in  $\mathbb{R}$ . Thus the result here strengthens Corollary 4.58.)

[Hint: Use Exercise 117 to show that, for each  $n = 1, 2, 3, \dots$ , both the sets

$$L_n = \{a \in A : ]a - 1/n, a] \cap A \text{ is countable}\},$$

$$R_n = \{a \in A : [a, a + 1/n[ \cap A \text{ is countable}\}$$

are countable. Then show that every point of the set  $B = A \setminus \bigcup_{n=1}^{\infty} (L_n \cup R_n)$ , is a two-sided condensation point of  $A$ .]

122. Provide the set  $\mathcal{F}$  of all continuous functions from  $\mathbb{R}$  to  $\mathbb{R}$  with its topology of pointwise convergence, as in Examples 2.72 (8).

- Show that the topology of  $\mathcal{F}$  is generated by the collection of all sets of the form

$$B(x; I) = \{f \in \mathcal{F} : f(x) \in I\}$$

where  $x$  is a point of  $\mathbb{R}$  and  $I$  an open interval in  $\mathbb{R}$ .

(b) Find the boundary, interior, and closure in  $\mathcal{F}$  of its subset

$$\{f \in \mathcal{F} : 0 < f(x) < 1 \text{ for all } 0 \leq x \leq 1\}.$$

Bing triangle space

- 123.** Must a subspace of a separable metrizable space itself be separable?
- 124. (a)** Let  $\mathcal{T}$  be the topology on a set  $X$  that is generated by a cover  $\mathcal{A}$  of  $X$ . Describe in terms of  $\mathcal{A}$  the  $\mathcal{T}$ -neighborhood system of an  $x \in X$ .
- 125.** Let  $\mathcal{A}$  be the collection of all subsets of  $\mathbb{N}$  having the form  $\{n, n+1, n+2, \dots\}$  for  $n \in \mathbb{N}$ . This collection covers  $\mathbb{N}$  because  $\{0, 0+1, 0+2, \dots\} = \mathbb{N}$ . What topology on  $\mathbb{N}$  does  $\mathcal{A}$  generate?
- 126.** What topology on the plane  $\mathbb{R}^2$  is generated by the collection of all lines in  $\mathbb{R}^2$ ?
- 127.** In [Examples 2.91 \(1\)](#) we noted that a vertical line is open for the topology  $\mathcal{T}$  on  $\mathbb{R}^2$  generated by the set of all vertical lines but is not open for the usual topology. Must a subset of  $\mathbb{R}^2$  that is open for the usual topology be open for  $\mathcal{T}$ ?

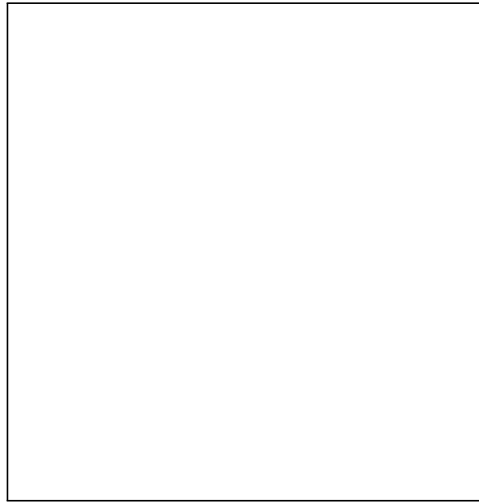


Figure 2.14: A typical neighborhood of a point in the Bing triangle space.

**128.** In the plane  $\mathbb{R} \times \mathbb{R}$ , let  $L$  be the  $x$ -axis  $\mathbb{R} \times \{0\}$ ; denote each point  $\langle x, 0 \rangle \in L$  simply by the corresponding real number  $x$ . Let  $X$  be the countable subset of  $\mathbb{R} \times \mathbb{R}$  consisting of all rational points in the closed upper half-plane, that is,

$$X = \{\langle x, y \rangle \in \mathbb{Q} \times \mathbb{Q} : y \geq 0\}.$$

We shall construct a remarkable topology on  $X$  in terms of a subbase.

Given a number  $\varepsilon > 0$  and a point  $\langle x, y \rangle \in X$  with  $y > 0$ , define a set  $U_\varepsilon(x, y)$  as follows. The three points  $\langle x, y \rangle$ ,  $\langle x - y/\sqrt{3}, 0 \rangle$ , and  $\langle x + y/\sqrt{3}, 0 \rangle$  are the vertices of an equilateral triangle  $T(x, y)$  whose base is on  $L$ , whose other two sides have slopes

Bing triangle space  
 Bing, R. H.  
 separation properties

$\pm\sqrt{3}$ , and which intersects  $L$  at the irrational points  $x - y/\sqrt{3}$  and  $x + y/\sqrt{3}$ . Form the sets

$$B_\varepsilon\left(x - y/\sqrt{3}\right) = \mathbb{Q} \cap ]x - y/\sqrt{3} - \varepsilon, x - y/\sqrt{3} + \varepsilon[,$$

$$B_\varepsilon\left(x + y/\sqrt{3}\right) = \mathbb{Q} \cap ]x + y/\sqrt{3} - \varepsilon, x + y/\sqrt{3} + \varepsilon[$$

of those rational points on  $L$  that belong to the symmetric intervals of length  $2\varepsilon$  about these two points (see Figure 2.15). Let



Figure 2.15: Typical subbasic open set  $U_\varepsilon(x, y)$  in the Bing triangle space.

fig:Bing-triangle-space-subbasic-set

$$U_\varepsilon(x, y) = \{(x, y)\} \cup B_\varepsilon\left(x - y/\sqrt{3}\right) \cup B_\varepsilon\left(x + y/\sqrt{3}\right).$$

(a) Show that the collection  $\{U_\varepsilon(x, y) : \varepsilon > 0, (x, y) \in X\}$  covers  $X$ .

Provide  $X$  with the topology generated by this collection. The resulting topological space is the **Bing triangle space**, named after R. H. Bing. (That is his full name: the letters 'R' and 'H' are not abbreviations!)

Prove:

- (b) Each  $x \in X \cap L$  has a local base consisting of intervals in  $\mathbb{Q}$  of the form  $\mathbb{Q} \cap ]x - \varepsilon, x + \varepsilon[$ .
- (c) For distinct points  $\langle x, y \rangle$  and  $\langle x', y' \rangle$  of  $X$  with  $y > 0$  and  $y' > 0$ , the triangles  $T(x, y)$  and  $T(x', y')$  do not have a vertex in common and do not intersect along an entire line segment not lying on  $L$ .
- (d) The space  $X$  is a Hausdorff space.
- (e) The space  $X$  is second-countable.

## 2.5 Separation Properties

sec:separation

Hausdorff spaces were introduced in the subsection “Hausdorff spaces” of Section 2.2. The alternative designation ‘ $T_2$ ’ for them comes from the German *Trennungssaxiom* meaning

“separation axiom.” The Hausdorff axiom is just one of a family of separation axioms—or as we shall refer to them, **separation properties**.

separation axioms  
Hausdorff space  
T2-space@Ttwo-space  
separation properties

def:separation-axioms **2.94 Definition.** A topological space  $X$  is said to be:

def-item:T0 (0) a  **$T_0$ -space** if, for any two distinct points  $x$  and  $y$  of  $X$ , there is some neighborhood of  $x$  that does not contain  $y$  *or* there is some neighborhood of  $y$  does not contain  $x$ .

def-item:T1 (1) a  **$T_1$ -space** if, for any two distinct points  $x$  and  $y$  of  $X$ , there is some neighborhood of  $x$  that does not contain  $y$  *and* there is some neighborhood of  $y$  that does not contain  $x$ .

def-item:T2 (2) a  **$T_2$ -space** if, for any two distinct points  $x$  and  $y$  of  $X$ , there are disjoint neighborhoods of  $x$  and  $y$ .

def-item:regular (3) **regular** if, for each point  $x \in X$  and each closed subset  $E$  of  $X$  with  $x \notin E$ , there are disjoint neighborhoods of  $x$  and  $E$ .

def-item:completely-regular (4) **completely regular** if, for each point  $x \in X$  and each closed subset  $E$  of  $X$  with  $x \notin E$ , there is a continuous function  $f: X \rightarrow [0, 1]$  such that  $f(x) = 0$  and  $f(y) = 1$  for every  $y \in E$ .

def-item:normal (5) **normal** if, each two disjoint closed subsets have disjoint neighborhoods.

The property of being regular is known as **regularity**; of being completely regular is known as **complete regularity**; and of being normal is known as **normality**.

For completely regularity, to say that a function  $f: X \rightarrow [0, 1]$  is *continuous* means that the inverse image  $f^{-1}(V)$  of each open subset  $V$  of  $[0, 1]$  is open in  $X$ ; as discussed in [Section 3.1](#), continuity in this sense generalizes continuity of maps between metric spaces (see [Theorem 1.56](#)).

As is the case with the condition for being a  $T_2$ -space, for each of the separation properties  $T_0$ ,  $T_1$ , regularity, and normality, the stated condition about neighborhoods is equivalent to the same condition but for *open* neighborhoods.

Separation properties may be described informally, as follows:

- A topological space is a  $T_2$ -space when each two points can be separated (by neighborhoods).
- A topological space is regular when each closed subset and each point not in that subset can be separated (by neighborhoods).
- A topological space is completely regular space when each closed subset and each point not in that subset can be separated by a continuous real-valued function with values in  $[0, 1]$ .
- A topological space is normal when each two closed sets can be separated (by neighborhoods).

Kolmogoroff, Andre  
Fréchet, Maurice  
Riesz, Frigyes  
Hausdorff, Felix  
Vietoris, Leopold  
Urysohn, Pavel  
Tietze, Heinrich

**Caution!** The terminology regarding separation axioms beyond  $T_2$  has never been standardized:

- Some authors include being  $T_0$  as part of the definitions of both regularity and complete regularity, and the property of being  $T_1$  as part of the definition of normality.

**We do not include being  $T_0$  as part of the definition of either regularity or complete regularity! And we do not include being  $T_1$  as part of the definition of normality!**

- Some authors introduce  $T_3$ ,  $T_4$ , and  $T_5$  as synonyms for what we mean by regularity, complete regularity and normality, respectively.
- Others include being  $T_0$  as part of the definitions of both  $T_3$  and  $T_4$ , and of being  $T_1$  as part of the definition of being  $T_5$ .

To prevent possible confusion, **we shall avoid the designations  $T_3$ ,  $T_4$ , and  $T_5$  entirely!**

The separation properties were all introduced in the 1920s:  $T_0$  by Andrey Kolmogoroff;  $T_1$  by Fréchet and Riesz;  $T_2$  by Hausdorff, as mentioned earlier; regularity by Leopold Vietoris; complete regularity by Pavel Urysohn; and normality by Heinrich Tietze.

In a  $T_1$ -space, points are closed. More precisely, we have the following result, whose proof is requested in [Exercise 132](#).

prop:T1-iff-singletons-closed

**2.95 Proposition.** A topological space  $X$  is a  $T_1$ -space if and only if for each point  $x \in X$ , the singleton  $\{x\}$  is closed in  $X$ .

The condition there is equivalent to  $\text{cls}\{x\} = \{x\}$  for each point  $x \in X$ . Regularity and normality also have useful equivalents in terms of closure.

prop:equivalent-regular-and-normal

**2.96 Proposition.** (1) A topological space  $X$  is regular if and only if for each point  $x$  of  $X$  and each neighborhood  $U$  of  $x$ , there is some open neighborhood  $V$  of  $x$  with  $\text{cls } V \subset U$ .

prop-part:equivalent-normal

(2) A topological space  $X$  is normal if and only if for each closed subset  $E$  of  $X$  and each neighborhood  $U$  of  $E$ , there is some open neighborhood  $V$  of  $E$  with  $\text{cls } V \subset U$ .

**Proof.** (1) First assume that  $X$  is a regular space. Let  $x$  be an arbitrary point of  $X$  and let  $U$  be an arbitrary neighborhood of  $x$ . Without loss of generality we may assume that  $U$  is an open set. The set  $E = X \setminus U$  is closed and  $x \notin E$ . By regularity, there are disjoint open neighborhoods  $V$  of  $x$  and  $W$  of  $E$ . Then  $\text{cls } V \subset U$ . In fact, just suppose that  $y \in \text{cls } V$  but  $y \notin U$ . Then  $y \in E$  whence  $y \in W$ . Since  $W$  is a neighborhood of  $y$  and  $y \in \text{cls } V$ , then  $W$  must intersect  $V$ , which is impossible.

Conversely, assume the condition holds for a space  $X$ . Let  $x$  be an arbitrary point and  $E$  be an arbitrary closed subset of  $X$  with  $x \notin E$ . The set  $U = X \setminus E$  is an open neighborhood of  $x$ , and so there exists an open neighborhood  $V$  of  $x$  with  $\text{cls } V \subset U$ . Then  $V$  and  $X \setminus \text{cls } V$  are disjoint open neighborhoods of  $x$  and  $E$ , respectively.

(2) The proof of (2) is similar to that of (1) and is left to the reader ([Exercise 142](#)).  $\square$



Part (1) of the preceding proposition states, equivalently, that a space is regular if and only if each of its points has a local base consisting solely of closed sets.

Part (1) of the preceding proposition also provides a more direct proof than the one for Proposition 1.26 that a metrizable space is regular: Let  $d$  be a metric inducing the topology of a metrizable space  $X$ , let  $x \in X$ , and let  $U$  be a neighborhood of  $x$ . Choose some  $\varepsilon > 0$  with  $B_x(\varepsilon; d) \subset U$ . Take some  $\delta > 0$  with  $\delta < \varepsilon$ . Then  $B_x(\delta; d)$  is an open neighborhood of  $x$  with  $\text{cls } B_x(\delta; d) = D_x(\delta; d) \subset B_x(\varepsilon; d)$ .

Completely regular spaces and deeper properties of normal spaces will be studied in Section 6.2.

ex:Thomas plank

**2.97 Examples.** (1) The Thomas plank [Examples 2.60 (4)] is regular. This follows from Proposition 2.96 (1) because each basic neighborhood of each point is closed as well as open.

ex:RI-regular-T1

(2) The Sorgenfrey line  $\mathbb{R}_l$  [Examples 2.20 (1)] is a regular  $T_1$ -space.

To see that  $\mathbb{R}_l$  is a  $T_1$ -space, let  $x, y \in \mathbb{R}_l$  with  $x < y$ . Then  $[x, y[$  is a neighborhood of  $x$  not containing  $y$ , and  $[y, y + 1[$  is a neighborhood of  $y$  not containing  $x$ .

That  $\mathbb{R}_l$  is a regular space follows from Proposition 2.96 (1) along with the fact that each member of the base  $\{[x, y[ : x < y\}$  of  $\mathbb{R}_l$  is closed in  $\mathbb{R}_l$  [see Examples 2.68 (6) and Exercise 88.]

ex:1st-ex-normal-RI

(3) The Sorgenfrey line  $\mathbb{R}_l$  [Examples 2.20 (1)] is, in fact, a normal  $T_1$ -space.

To see that  $\mathbb{R}_l$  is normal, let  $E$  and  $F$  be disjoint closed subsets of  $\mathbb{R}_l$ . For each  $x \in E$ , choose some  $u_x > x$  for which the basic open neighborhood  $[x, u_x[$  of  $x$  is disjoint from  $F$ ; and for each  $y \in F$ , choose some  $v_y > y$  for which the basic open neighborhood  $[y, v_y[$  of  $y$  is disjoint from  $E$ . Then  $\bigcup_{x \in E} [x, u_x[$  and  $\bigcup_{y \in F} [y, v_y[$  are open neighborhoods of  $E$  and  $F$  respectively.

Let  $x \in E$  and  $y \in F$ , so that  $x \neq y$ ; without loss of generality we may suppose that  $x < y$ . Since the point  $y$  of  $F$  does not belong to  $[x, u_x[$ , necessarily  $u_x \leq y$ . Then no  $z \in \mathbb{R}_l$  can belong to both  $[x, u_x[$  and  $[y, v_y[$ . It follows that  $\bigcup_{x \in E} [x, u_x[$  and  $\bigcup_{y \in F} [y, v_y[$  are disjoint.

*Alternative proof:*  $\mathbb{R}_l$  is a regular “Lindelöf space” [(2) and Exercise 115 (b)]; and a regular Lindelöf space is necessarily normal (Exercise 6.31).  $\diamond$

For non-examples, see the counterexamples in the next subsection.

### The separation properties hierarchy

subsec:separation-hierarchy

The separation properties are arranged in a hierarchy, as indicated in the following proposition.

prop:hierarchy-part1-is-T0

**2.98 Proposition.** (1) Every  $T_1$ -space is a  $T_0$ -space.

prop-part:T2-is-T1

(2) Every  $T_2$ -space is a  $T_1$ -space.

prop-part:regular-T0-is-T2

(3) Every regular  $T_0$ -space is a  $T_2$ -space.

prop-part:normal-T1-is-regular

(4) Every normal  $T_1$ -space is regular.

**Proof.** (3) Let  $x$  and  $y$  be distinct points of a regular  $T_0$ -space  $X$ . Since  $X$  is a  $T_0$ -space, there is an open set  $W$  in  $X$  containing one of the two points  $x$  and  $y$  but not the

Thomas plank!regular space@as reg  
Sorgenfreyline!regular space@as reg  
regular space!Sorgenfreyline@and S  
Sorgenfreyline!normal space@as no  
normal space!Sorgenfreyline@and S  
Lindelof space@Lindelof space

other, without loss of generality, assume that  $x \in W$  but  $y \notin W$ . Let  $E = X \setminus W$ . Then  $E$  is a closed subset of  $X$  with  $x \notin E$ . Since  $X$  is also regular, there are disjoint open neighborhoods  $U$  and  $V$  of  $x$  and  $E$ , respectively. Then  $U$  and  $V$  are disjoint neighborhoods of  $x$  and  $y$ , respectively.  $\square$

Sorgenfrey plane!regular space@and regular space

ex:simple-separation-compact 2.99 Examples. (1) A topological space need not be a  $T_0$ -space. In fact, an indiscrete space with more than one point fails to be a  $T_0$ -space.

normal space!product space@and product space  
ex:T0-not-T1 (2) A  $T_0$ -space need not be a  $T_1$ -space. The Sierpinski space  $X = \{0, 1\}$  with its topology  $\{\emptyset, \{0\}, X\}$  [Examples 2.3 (6)], is such a space. In fact, it is a  $T_0$ -space because its point 0 does have a neighborhood, namely,  $\{0\}$  not include the other point; but it is not a  $T_1$ -space because  $\{0\}$  is not closed.

ex:T1-not-T2 (3) A  $T_1$ -space need not be a  $T_2$ -space. An infinite set  $X$  with its finite-complement topology [Examples 2.3 (7)] is such a space. In fact,  $X$  is a  $T_1$ -space because, by the very definition of the topology, each singleton is closed; but it is not a  $T_2$ -space because if  $U$  and  $V$  are any two open sets, then  $U \cap V$  has finite complement and so is infinite, and *a fortiori*, nonempty.

Another such space is the line with two origins [Examples 2.20 (3)]. In fact, the complement of each singleton in this space is open there, and the points 0 and 0' do not have disjoint neighborhoods. (See Exercise 134).

ex:T2-not-regular (4) A  $T_2$ -space need not be regular. The half-disk space  $X$  [Examples 2.25 (3)] is such a space. Already we know that it is a  $T_2$ -space. However, it is *not* regular, because its subset  $E = ]0, 1[ \times \{0\} = \{\langle x, 0 \rangle : 0 < x < 1\}$  is closed in  $X$  yet cannot be separated from the point  $(1, 0)$  by neighborhoods.

ex:regular-not-T2 (5) A regular space need not be a  $T_0$ -space and hence need not be a  $T_2$ -space. For such a space, see Exercise 138.

ex:regular-not-normal-prelim (6) A regular  $T_0$ -space need not be normal. In fact, the Sorgenfrey plane  $\mathbb{R}_l \times \mathbb{R}_l$  [Examples 2.72 (7)] is such a space: see Exercise 154.

The tangent disk space [Exercise 37] is another such space: see Exercises 152 and 153.

ex:normal-T0-not-T1-not-regular (7) A normal  $T_0$ -space need not be a  $T_1$  space and hence need not be regular. In fact, give the set  $X = \{0, 1, 2\}$  the following topology:

$$\{\emptyset, \{0\}, \{1\}, \{0, 1\}, X\}.$$

Then  $X$  is “vacuously” normal because it has no two disjoint nonempty closed subsets whatsoever. Moreover,  $X$  is a  $T_0$ -space, as can be checked by examining its pairs of elements. However,  $X$  is *not* a  $T_1$ -space because  $\{0\}$  is not closed. Then also  $X$  is *not* regular, for otherwise it would be a  $T_2$ -space and, *a fortiori*, a  $T_1$ -space.  $\diamond$

Table 2.1 summarizes the relationships among the separation properties enumerated in Proposition 2.98 and Examples 2.99. **None of the listed implications are reversible!**

Table 2.2 summarizes additional relationships to be established in Section 6.2. And again, as we shall see there: **none of the listed implications are reversible!**

In view of the following proposition, regular spaces and normal spaces abound because metrizable spaces do. (Another class of regular and normal spaces that abound, but that are not necessarily metrizable, are the “compact” Hausdorff spaces: see Theorem 4.20.)

$\begin{aligned} \text{regular } T_0 &\implies T_2 \implies T_1 \implies T_0, \\ \text{normal } T_1 &\implies \text{regular}, \\ \text{normal} &\not\Rightarrow \text{regular}, \quad \text{regular} \not\Rightarrow T_2, \quad T_2 \not\Rightarrow \text{regular}. \end{aligned}$
--

Table 2.1: Basic relationships among separation properties.

$\begin{aligned} \text{completely regular} &\implies \text{regular}, \\ \text{normal } T_1 &\implies \text{completely regular } T_0 \implies \text{regular } T_1, \\ \text{normal} &\not\Rightarrow \text{completely regular}, \quad \text{completely regular} \not\Rightarrow \text{regular}. \end{aligned}$
---

Table 2.2: Relationships among complete regularity and other separation properties.

**2.100 Proposition.** Every metrizable space is a normal  $T_1$ -space.

**Proof.** This follows from Proposition 1.26 and Corollary 1.27.  $\square$

In particular, every metrizable space is a regular space. However, **a regular  $T_2$ -space need not be metrizable**. The Fort space [Examples 2.66 (3)] is such a space: see Exercise 145. More strongly, **a normal  $T_1$ -space need not be metrizable**.

**2.101 Example.** The Sorgenfrey line [Examples 2.20 (1)] is a normal  $T_1$ -space that is not metrizable: see Examples 2.97 (3) and Exercise 94.  $\diamond$

Separation properties of new spaces constructed from old

Part (3) of the next result just restates Proposition 2.26; exercises request proofs of the other parts.

**2.102 Proposition.** (1) Every subspace of a  $T_0$ -space is itself a  $T_0$ -space.

(2) Every subspace of a  $T_1$ -space is itself a  $T_1$ -space.

(3) Every subspace of a  $T_2$ -space is itself a  $T_2$ -space.

(4) Every subspace of a regular space is itself regular.

(The corresponding result for completely regular spaces appears later, in Proposition 6.18.)

In contrast to the results in Proposition 2.102, **a subspace of a normal space need not be normal**, even when it is open; the Tychonoff plank (Example 6.23) is such a normal space. However, the following is true.

metrizable space!regular space@and  
Fort space  
normal space!metrizable space@and  
Sorgenfreyline!normal space@as no  
Sorgenfreyline!metrizable space@ar  
normal space!metrizable space@and  
Tychonoff plank

table:basic-separation-implications

table:compl-regular-implications

prop:metrizable-is-normal-T1

hfrey-line-normal-T1-not-metrizable

subsec:separation-new-spaces

prop:hereditary-separation-properties

prop-part:T1-inherits

prop-part:T2-inherits

prop-part:regular-inherits

separation axioms

T0-space@\Tzero-sp

T1-space@\Tone-space

T0-space@\Tzero-space

T2-space@\Ttwo-space!T1-space@and \Tone-space

Hausdorff space!T1

Hausdorff space!T1

T1-space@\Tone-sp

T1-space@\Tone-sp

T2-space@\Ttwo-sp

T1-space@\Tone-sp

T1-space@\Tone-sp

**2.103 Proposition.** Every closed subspace of a normal space is itself normal.

Some separation properties are preserved when forming the product of two (or more) spaces having those properties. (Proofs are requested in [Exercise 150](#).)

**2.104 Proposition.** (1) The product of two  $T_0$ -spaces is itself a  $T_0$ -space.

(2) The product of two  $T_1$ -spaces is itself a  $T_1$ -space.

(3) The product of two  $T_2$ -spaces is itself a  $T_2$ -space.

(4) The product of two regular spaces is itself a regular space.

In contrast with those results, *the product of two normal spaces need not be normal*. For a counterexample, see [Exercise 154](#) or [Example 6.32](#).

### EXERCISES FOR SECTION 2.5

**129.** Explain why a topological space  $X$  is a  $T_1$ -space if and only if, given any two distinct points  $x$  and  $y$  of  $X$ , there is an open set  $U$  in  $X$  with  $x \in U$  but  $y \notin U$ . (Thus, the condition originally stated for  $X$  to be a  $T_1$ -space has a redundancy.)

**130.** (a) Prove that every  $T_1$ -space is a  $T_0$ -space. [This is [Proposition 2.98 \(1\)](#).]

(b) Prove that every subspace of a  $T_0$ -space must itself be a  $T_0$ -space. [This is [Proposition 2.102 \(1\)](#).]

**131.** (a) Formulate a condition involving closures that is equivalent to being a  $T_0$ -space.

(b) Formulate a condition involving closures that is equivalent to being a  $T_2$ -space.

**132.** Show that a topological space  $X$  is a  $T_1$ -space if and only if for each  $x \in X$ , the set  $\{x\}$  is the intersection of all neighborhoods of  $x$ .

**133.** Which finite topological spaces, if any, are  $T_1$ -spaces?

**134.** Verify the assertions made in [Examples 2.99 \(3\)](#) about the line with two origins [[Examples 2.20 \(3\)](#)]:

(a) The complement of each singleton is closed, thereby establishing that this space is a  $T_1$ -space.

(b) The points  $0$  and  $0'$  do not have disjoint neighborhoods, thereby establishing that this space is not a  $T_2$ -space.

**135.** (a) Prove that every  $T_2$ -space is a  $T_1$ -space. [This is [Proposition 2.98 \(2\)](#).]

(b) Prove that every subspace of a  $T_1$ -space must itself be a  $T_1$ -space. [This is [Proposition 2.102 \(2\)](#).]

**136.** (a) Give an example of a zero-dimensional space that is not a  $T_1$ -space.

(b) Show that a zero-dimensional  $T_1$ -space is a  $T_2$ -space.

(c) Prove that a zero-dimensional space is regular.

*Note:* A zero-dimensional  $T_1$ -space need not be normal: see [Exercise 6.41](#).

**137.** Prove that a pseudometrizable space ([Exercise 11](#)) is metrizable if and only if it is a  $T_0$ -space.

- prob:4-pt-regular-not-T2-space **138.** Construct a topology on a four-point space making it regular but not a  $T_0$ -space and hence not a  $T_2$ -space. regular spac  
regular space
- prob:equivalents-of-regularity **139.** Prove that a topological space  $X$  is regular if and only if for each point  $x \in X$  and each closed set  $E \subset X$  with  $x \notin E$ , there is some neighborhood  $V$  of  $x$  such that  $E \cap \text{cls } V = \emptyset$ . regular space!closure@and closure  
metrizable space!regular space@and  
regular space!metrizable space@and  
deleted sequence space!separation p  
Fort topology  
Bing triangle space  
regular space
- prob:subspace-of-regular **140.** Prove that every subspace of a regular space must itself be regular. [This is [Proposition 2.102 \(4\)](#).] metrizable space!regular space@and  
regular space!metrizable space@and
- prob:metrizable-is-regular **141.** Prove directly, without using [Proposition 2.100](#), that every metrizable space is regular. *Note:* A regular space, even a regular  $T_1$ -space or a regular  $T_2$ -space, need not be metrizable—see [Exercise 145](#) and [Exercise 146](#). normal space!regular space@and reg  
regular space!normal space@and no  
normal space  
Thomas plank!normal space@and n  
regular space
- prob:normal-characterization-via-cls **142.** Prove the characterization of normal spaces in terms of closure stated in [Proposition 2.96 \(2\)](#). normal space!closure@and closure  
normal space
- prob:dual-criterion-normality **143.** Prove that normality of a space  $X$  is equivalent to the following property: For each cover  $\{G, H\}$  of  $X$  by two open sets, there is a cover  $\{A, B\}$  of  $X$  by closed sets with  $A \subset G$  and  $B \subset H$ . tangent disk space!separation prop  
regular space  
regular space
- 144.** Which separation properties does the deleted sequence space  $\mathbb{R}_K$  ([Exercise 92](#)) have?
- b:Fort-space-regular-not-metrizable **145.** Verify that a Fort topology [[Examples 2.66 \(3\)](#)], which is a Hausdorff space that is not first-countable (hence not metrizable), is regular.
- prob:Bing triangle space-not-regular **146.** Verify that the Bing triangle space ([Exercise 128](#)), which is a second-countable Hausdorff space, is *not* regular and hence *not* metrizable.
- 147.** Give a proof of [Proposition 2.96 \(2\)](#).
- 148.** Prove that a normal  $T_1$ -space must be regular. [This is [Proposition 2.98 \(4\)](#).]
- prob:thomas-plank-not-normal **149.** Show that the Thomas plank [[Examples 2.60 \(4\)](#)], which is a regular  $T_0$ -space [[Examples 2.97 \(1\)](#)], is *not* normal (and hence not metrizable).
- separation-properties-product-of-2 **150.** Prove:
- (a) The product of two  $T_2$ -spaces is a  $T_2$ -space. [This is [Proposition 2.104 \(1\)](#).]
  - (b) The product of two  $T_0$ -spaces is a  $T_0$ -space. [This is [Proposition 2.104 \(1\)](#).]
  - (c) The product of two  $T_1$ -spaces is a  $T_1$ -space. [This is [Proposition 2.104 \(2\)](#).]
  - (d) The product of two regular spaces is a regular space. [This is [Proposition 2.104 \(4\)](#).]
- prob:equivalents-of-normality **151.** Prove that each of the following is a necessary and sufficient condition for a topological space  $X$  to be normal:
- (a) For each two disjoint closed subsets  $E$  and  $F$  of  $X$ , there are open neighborhoods  $U$  and  $V$  of  $E$  and  $F$ , respectively, such that  $\text{cls } U$  is disjoint from  $\text{cls } V$ .
  - (b) For each cover  $\{U, V\}$  of  $X$  by two open sets, there exists a cover  $\{E, F\}$  of  $X$  by closed sets for which  $E \subset U$  and  $V \subset F$ .
- prob:tangent-disk-space-regular **152.** The tangent disk space  $\Gamma$  defined in [Exercise 37](#) is a  $T_2$ space. Show that this space is, in fact, regular.
- (Hint: Let  $x$  be an arbitrary point of  $\Gamma$  and let  $E$  be an arbitrary closed subset of  $\Gamma$  with  $x \notin E$ . Consider the several cases:  $x \in H$  and  $E \subset H$ ;  $x \in L$  and  $E \subset H$ ;  $x \in H$  and  $E \cap L \neq \emptyset$ ; and  $x \in L$  and  $E \cap L \neq \emptyset$ .)*

angent-disk-space-normal-Baire **153.** (*Continuation of Exercise 152.*) Prove that the tangent disk space  $\Gamma$  is *not* normal.  
 tangent disk space!nonnormal space!nonnormal space  
 normal space!tangent disk space!nonnormal space  
 tangent disk space!separation properties!separation properties  
 Sorgenfrey plane!and separation properties!and separation properties  
 product space!normal space!normal space  
 normal space!product space!product space  
 normal space

[*Hint:* Let  $Q$  be the set of rational points in  $L$ , that is,  $Q = \mathbb{Q} \times \{0\}$ , and let  $P = L \setminus Q$ . Just suppose that there are disjoint open sets  $U$  and  $V$  in  $\Gamma$  containing  $Q$  and  $P$ , respectively. Endow  $L$  with its *usual* (Euclidean) topology. For each positive integer  $n$ , define  $A_n$  to be the set of those points  $p \in P$  for which  $B_p(1/n; d) \subset V$ , where  $d$  is the Euclidean metric. Verify that  $\bigcup_{n=1}^{\infty} A_n$  is a nonempty subset of  $L$  with each  $A_n$  open in the usual topology on  $L$ . Apply the Baire Category Theorem (1.91) to obtain an  $n$  for which  $A_n$  is *not* nowhere dense in  $L$ .]

*Note:* A different method of proof uses cardinality: see Exercise 6.30. A more direct proof, avoiding the use of both cardinality and the Baire Category Theorem but applying instead the Nested Set Theorem, is possible: see Vetterlein [66].

Sorgenfrey-plane-not-normal-Baire **154.** Show that the Sorgenfrey plane  $\mathbb{R}_l \times \mathbb{R}_l$  [Examples 2.72 (7)] is regular but *not* normal.

*Note:* In conjunction with Examples 2.97 (3), this exercise shows that the product of two normal  $T_1$ -spaces need not be normal.

[*Hint:* Let  $X = \mathbb{R}_l \times \mathbb{R}_l$ . Form the reverse diagonal  $D = \{\langle x, -x \rangle : x \in \mathbb{R}\}$  of  $X$ . Verify that  $D$  is a closed subspace of  $X$  whose relative topology is discrete. Let  $E$  be the set of those points on  $D$  whose coordinates are rational and let  $F$  be its complement in  $D$ , so that  $E$  and  $F$  are closed in  $D$  and hence in  $X$ . Just suppose that  $E$  and  $F$  have disjoint open neighborhoods  $U$  and  $V$ , respectively, in  $X$ .

For  $x \in \mathbb{R}$  and  $\varepsilon > 0$ , let  $N_\varepsilon(x)$  be the basic open neighborhood  $[x, x + \varepsilon[ \times ]-x, -x + \varepsilon[$  of  $\langle x, -x \rangle$  in  $X$ . For  $n = 1, 2, 3, \dots$ , let  $V_n = \{x \in \mathbb{R} \setminus \mathbb{Q} : N_{1/n}(x) \subset V\}$ . Show that  $\mathbb{R} \setminus \mathbb{Q} = \bigcup_{n=1}^{\infty} V_n$ . Now use the Baire Category Theorem (1.91) to show that there is some  $n$  such that  $V_n$  is *not* nowhere dense in  $\mathbb{R}$  (for the usual topology).]

*Note:* A different method for showing that the Sorgenfrey plane is not normal uses cardinality: see Example 6.32.

## CHAPTER

# 3

## Continuity and Convergence

chap:contconv

### Chapter Contents

3.1 Continuous Maps . . . . .	302	Exercises for Section 3.3 . . . . .	379
Continuous maps of topological spaces . . . . .	302	3.4 Quotient Spaces . . . . .	385
Extending and gluing maps . . . . .	307	The unit circle as a quotient space . . . . .	385
Open and closed maps . . . . .	312	The quotient topology . . . . .	388
Continuous images of spaces . . . . .	314	Quotient spaces and continuous maps . . . . .	390
Exercises for Section 3.1 . . . . .	316	Quotient maps . . . . .	392
3.2 Homeomorphisms . . . . .	319	Concrete realizations of quotient spaces . . . . .	393
Homeomorphisms—definition and examples . . . . .	320	Countability and separation properties of quotient spaces . . . . .	408
Topological properties . . . . .	333	Exercises for Section 3.4 . . . . .	410
Local homeomorphism . . . . .	335	3.5 Convergence . . . . .	420
Locally Euclidean spaces . . . . .	336	Sequential convergence in topological spaces . . . . .	420
Topological manifolds . . . . .	343	Directed sets and nets . . . . .	423
Embedding . . . . .	346	Convergence of nets . . . . .	425
Preservation of topological properties . . . . .	349	Subnets . . . . .	435
Exercises for Section 3.2 . . . . .	350	Limit of a map . . . . .	437
3.3 Product Spaces . . . . .	359	Filters . . . . .	440
Topology of the product of countably many metric spaces . . . . .	360	Convergence of filters . . . . .	443
The product topology . . . . .	363	Comparison of filters . . . . .	448
Products and maps . . . . .	367	Filters vs. nets . . . . .	449
Further properties of product spaces . . . . .	377	Exercises for Section 3.5 . . . . .	451

### Introduction

Our program of abstracting topological ideas from metric spaces continues in this chapter with the study of continuous maps between topological spaces. Homeomorphisms are a special class of continuous maps between topological space that we examine in detail; when there is a homeomorphism from one space to another, the two spaces are topologically the same in the sense that they have precisely the same topological properties. With the aid of continuous maps we define appropriate topologies on product and quotient sets of topological spaces, thereby providing new methods for constructing topological spaces. Finally, after defining sequential convergence in topological spaces and seeing that it is not



adequate for treating topological questions, we present a more general theory of convergence which is adequate.

### 3.1 Continuous Maps

sec:continuous

To define continuity of a map  $f: X \rightarrow Y$  between arbitrary topological spaces we cannot simply mimic the definition (1.51) of the case of metric spaces, for that definition directly employs metrics. According to Theorem 1.53, however, if the topologies  $\mathcal{T}$  and  $\mathcal{S}$  on  $X$  and  $Y$  are induced by metrics  $d$  and  $d'$ , respectively, then  $f$  is  $\langle d, d' \rangle$ -continuous at a point  $x \in X$  precisely when each  $d'$ -neighborhood  $N$  of  $f(x)$  contains the image  $f(M)$  of some  $d$ -neighborhood  $M$  of  $x$ . Now a  $d'$ -neighborhood of  $f(x)$  is just an  $\mathcal{S}$ -neighborhood of  $f(x)$  in the topological space  $Y$ , and a  $d$ -neighborhood of  $x$  is just a  $\mathcal{T}$ -neighborhood of  $x$  in the topological space  $X$ . Hence  $f$  is  $\langle d, d' \rangle$ -continuous at  $x$  precisely when each  $\text{top}\mathcal{S}$ -neighborhood  $N$  of  $f(x)$  in  $Y$  contains the image  $fM$  of some  $\mathcal{T}$ -neighborhood  $M$  of  $x$  in  $X$ . This last condition no longer refers directly to metrics, only to topologies.

#### Continuous maps of topological spaces

To define continuity of a map  $f: X \rightarrow Y$  between arbitrary topological spaces we cannot simply mimic the definition (Definition 1.51) of the case of metric spaces, for that definition directly employs metrics. According to Theorem 1.53, however, if the topologies  $\mathcal{T}$  and  $\mathcal{S}$  on  $X$  and  $Y$  are induced by metrics  $d$  and  $d'$ , respectively, then  $f$  is  $\langle d, d' \rangle$ -continuous at a point  $x \in X$  precisely when each  $d'$ -neighborhood  $N$  of  $f(x)$  contains the image  $f(M)$  of some  $d$ -neighborhood  $M$  of  $x$ . Now a  $d'$ -neighborhood of  $f(x)$  is just an  $\mathcal{S}$ -neighborhood of  $f(x)$  in the topological space  $Y$ , and a  $d$ -neighborhood of  $x$  is just a  $\mathcal{T}$ -neighborhood of  $x$  in the topological space  $X$ . Hence  $f$  is  $\langle d, d' \rangle$ -continuous at  $x$  precisely when each  $\text{top}\mathcal{S}$ -neighborhood  $N$  of  $f(x)$  in  $Y$  contains the image  $fM$  of some  $\mathcal{T}$ -neighborhood  $M$  of  $x$  in  $X$ . This last condition no longer refers directly to metrics, only to topologies.

def:cont **3.1 Definition.** Let  $f: X \rightarrow Y$  be a map from a topological space  $X$  to a topological space  $Y$ . For  $x \in X$ , the map  $f$  is said to be **continuous at  $x$**  if for each neighborhood  $N$  of  $f(x)$  in  $Y$  there exists some neighborhood  $M$  of  $x$  in  $X$  such that  $f(M) \subset N$ . The map  $f$  is said to be **continuous** if it is continuous at each  $x \in X$ . The negation of “continuous” is **discontinuous**.

Sometimes it is convenient to work with the following condition (compare Theorem 1.53) instead of directly with the definition.

lem:cont-via-inverse-image-nbds **3.2 Lemma.** A necessary and sufficient condition for a map  $f: X \rightarrow Y$  between topological spaces to be continuous at a point  $x \in X$  is that the inverse image  $f^{-1}(N)$  be a neighborhood of  $x$  in  $X$  for each neighborhood  $N$  of  $f(x)$  in  $Y$ .

**Proof.** Let  $N$  be an arbitrary neighborhood of  $f(x)$  in  $Y$ . If  $f(M) \subset N$  for a neighborhood  $M$  of  $x$  in  $X$ , then  $f^{-1}(N)$  is a neighborhood of  $x$  in  $X$  because  $M \subset f^{-1}(N)$ . Conversely, if  $f^{-1}(N)$  is a neighborhood of  $x$  in  $X$ , then  $f(M) \subset N$  for  $M = f^{-1}(N)$ .  $\square$

In the same way that we deduced Theorem 1.56 from Theorem 1.53, we may deduce from the preceding Lemma 3.2 a topological criterion for continuity.



thm:cont-via-inverse-images

**3.3 Theorem.** A map  $f: X \rightarrow Y$  from a topological space  $X$  to a topological space  $Y$  is continuous if and only if the inverse image  $f^{-1}(V)$  of each open subset  $V$  of  $Y$  is open in  $X$ .

In terms of the topologies  $\mathcal{T}$  and  $\mathcal{S}$  on  $X$  and  $Y$ , respectively, the criterion in [Theorem 3.3](#) says simply that

$$V \in \mathcal{S} \implies f^{-1}(V) \in \mathcal{T}.$$

When using [Theorem 3.3](#) to establish continuity of a map  $f: X \rightarrow Y$ , we can sometimes reduce the number of open subsets of  $Y$  whose inverse images in  $X$  must be examined. For example, if a metric  $d$  induces the topology of  $Y$ , we need only show that the inverse image of each  $d$ -ball  $B_\varepsilon(y; d)$  is open in  $X$ . This is a special case of part (1) of the following corollary.

cont-via-inverse-images-base-subbase

**3.4 Corollary.** Let  $f: X \rightarrow Y$  be a map from a topological space  $X$  to a topological space  $Y$ . Then:

cor-part:cont-via-inverse-images-base

(1) The map  $f$  is continuous if there is some base  $\mathcal{B}$  of  $Y$  such that  $f^{-1}(B)$  is open in  $X$  for each member  $B$  of  $\mathcal{B}$ .

part:cont-via-inverse-images-subbase

(2) The map  $f$  is continuous if there is some subbase  $\mathcal{S}$  of  $Y$  such that  $f^{-1}(S)$  is open in  $X$  for each member  $S$  of  $\mathcal{S}$ .

Because  $f^{-1}(\emptyset) = \emptyset$  is open in  $X$ , then in each of (1) and (2) we need only consider inverse images of *nonempty* members.

**Proof.** (1) Assume that such a base  $\mathcal{B}$  exists. Let  $V$  be any open subset of  $Y$ . There is a family  $\langle B_i \rangle_{i \in I}$  of members of  $\mathcal{B}$  such that

$$V = \bigcup_{i \in I} B_i.$$

Then the inverse image

$$f^{-1}(V) = f^{-1}\left(\bigcup_{i \in I} B_i\right) = \bigcup_{i \in I} f^{-1}(B_i)$$

is a union of open subsets of  $X$  and hence is open in  $X$ . It follows from [Theorem 3.3](#) that  $f$  is continuous.

(2) See [Exercise 3](#).  $\square$

Since open sets determine closed sets and vice versa, the criterion of [Theorem 3.3](#) may be formulated instead in terms of closed sets.

cor:cont-via-inverse-images-closed

**3.5 Corollary.** The map  $f: X \rightarrow Y$  is continuous if and only if the inverse image  $f^{-1}(E)$  of each closed subset  $E$  of  $Y$  is closed in  $X$ .

**Proof.** Assume first that  $f$  is continuous. Let  $E$  be closed in  $Y$ . Then  $Y \setminus E$  is open in  $Y$ , by [Theorem 3.3](#) the set

$$X \setminus f^{-1}(E) = f^{-1}(Y \setminus E)$$

is open in  $X$ , and hence  $f^{-1}(E)$  is closed in  $X$ .

The proof of the converse is similar.  $\square$

cor:cont-and-closure **3.6 Corollary.** Let  $f: X \rightarrow Y$  be a map between topological spaces. If  $f: X \rightarrow Y$  is continuous, then

$$f(\text{cls } A) \subset \text{cls}(f(A))$$

for each set  $A \subset X$ .

In the preceding corollary, of course,  $\text{cls } A$  is the closure of  $A$  in  $X$ , whereas  $\text{cls } f(A)$  is the closure of  $f(A)$  in  $Y$ . The converse of the implication is also true: see [Exercise 8 \(a\)](#).

**Proof.** Let  $A \subset X$ . The set  $\text{cls } f(A)$  is closed in  $Y$ , so by [Corollary 3.5](#) its inverse image  $f^{-1}(\text{cls } f(A))$  is closed in  $X$ . Now

$$A \subset f^{-1}(f(A)) \subset f^{-1}(\text{cls } f(A))$$

because  $f(A) \subset \text{cls } f(A)$ . Since  $\text{cls } A$  is the *leaset* closed subset of  $X$  containing  $A$ , then

$$\text{cls } A \subset f^{-1}(\text{cls } f(A)).$$

Hence  $f(\text{cls } A) \subset \text{cls}(f(A))$ .  $\square$

Recall that, for a map  $f: X \rightarrow Y$ , the *fiber over* a point  $y$  of  $Y$  is the inverse image  $f^{-1}(y)$  of  $\{y\}$ . (See [page 27](#).)

cor:fibers-closed-for-T1-codomain **3.7 Corollary.** If  $f: X \rightarrow Y$  is a continuous map to a  $T_1$  space  $Y$ , then the fibers of  $f$  over points of  $Y$  are closed in  $X$ .

Despite its trivial proof, the next theorem expresses the single most important property of continuity. It assures us that the most general way of combining continuous maps always leads to continuous maps. Moreover, it is a tool for establishing the continuity of maps built up from simpler maps. For example, from the continuity of the sine and square-root functions it allows us to conclude at once the continuity of the real-valued function  $x \mapsto \sin(\sqrt{x})$  on the set of nonnegative real numbers.

thm:cont-composite **3.8 Theorem (composite of continuous maps).** Let  $f: X \rightarrow Y$  and  $g: Y \rightarrow Z$  be continuous maps between topological spaces. Then the composite map  $g \circ f: X \rightarrow Z$  is continuous.

**Proof.** We use [Theorem 3.3](#) once more. Let  $W$  be any open subset of  $Z$ . By continuity of  $g$ , the set  $g^{-1}(W)$  is open in  $Y$ . Then by continuity of  $f$ , the set

$$(g \circ f)^{-1}(W) = f^{-1}(g^{-1}(W))$$

is open in  $X$ .  $\square$

For a “local” version of the preceding theorem, see [Exercise 1](#).

exs:continuous-maps **3.9 Examples.** (1) Let  $X$  and  $Y$  be metrizable topological spaces. Then the discussion leading to [Definition 3.1](#) shows that a map  $f: X \rightarrow Y$  is continuous precisely when it is  $\langle d, d' \rangle$ -continuous for any metrics  $d$  and  $d'$  that induce the topologies of  $X$  and  $Y$ , respectively.

As a consequence, all the continuous maps between metric spaces discussed in [Chapter 1](#), and in particular all the continuous functions encountered in calculus, are thus examples of continuous maps between topological spaces.

This example demonstrates once again that continuity of a map between metric spaces depends not on the particular metrics, but only on the topologies they induce (compare [Corollary 1.55](#)). Hence it supports our contention that topological spaces provide the proper setting in which to study continuity—even of maps between metric spaces.

- (2) A constant map  $f: X \rightarrow Y$  with constant value  $c$  is continuous, because

$$f^{-1}(V) = \begin{cases} \emptyset & \text{if } c \notin V, \\ X & \text{if } c \in V \end{cases}$$

for each subset  $V$  of  $Y$ .

- (3) Since every subset of a discrete space is open, **a map  $f: X \rightarrow Y$  from a discrete space to an arbitrary topological space  $Y$  is continuous.**

ex:inclusion-cont

- (4) Suppose that  $A$  is a subspace of a topological space  $X$ . Then the inclusion map

$$\begin{aligned} j: A &\rightarrow X \\ x &\mapsto x \end{aligned}$$

is continuous, since by definition of the relative topology on  $A$  the set  $j^{-1}(V) = V \cap A$  is open in  $A$  for each open subset  $V$  of  $X$ .

ex:identity-cont

- (5) In particular—take  $A = X$  in (4)—**the identity map  $\iota_X: X \rightarrow X$  of any topological space is continuous.**

ex:restrict-continuous

- (6) Let  $f: X \rightarrow Y$  be a map from a topological space  $X$  to a topological space  $Y$ . If  $A$  is a subspace of  $X$ , then the restriction

$$\begin{aligned} f|_A: A &\rightarrow Y \\ x &\mapsto f(x) \end{aligned}$$

of  $f$  to  $A$  is just the composite of the inclusion map  $j: A \rightarrow X$  and  $f: X \rightarrow Y$ . Hence by [Theorem 3.8](#), **the restriction of a continuous map between topological spaces to a subspace of its domain is still continuous.** Thus cutting down the domain  $X$  of  $f$  does not affect continuity.

- (7) The map  $x \mapsto 1/x$  with domain  $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$  is continuous whether considered as a map with codomain  $\mathbb{R}$  or else with codomain  $\mathbb{R}^*$ . In general, **cutting down or enlarging the codomain of a map does not affect its continuity:** Let

$$g: X \rightarrow Z, \quad f: X \rightarrow Y$$

be maps on a topological space  $X$ , let  $Z$  be a subspace of the topological space  $Y$ , and suppose

$$f(x) = g(x) \quad (x \in X).$$

Necessarily  $f(X) \subset Z$ ; the two maps differ only when  $Z \neq Y$ . Then  $f$  is continuous if and only if  $g$  is.

In fact,  $f: X \rightarrow Y$  is the composite of  $g: X \rightarrow Z$  and the inclusion map  $j: Z \rightarrow Y$ , so that  $f$  is continuous if  $g$  is. Conversely, if  $f: X \rightarrow Y$  is continuous, then an arbitrary open subset  $V$  of  $Z$  has the form  $W \cap Z$  for some open subset  $W$  of  $Y$ , and  $g^{-1}(V) = f^{-1}(W)$  is open in  $X$ .

If above we take  $Z = f(X)$ , the range of  $f$ , then we obtain a surjection  $g: X \rightarrow f(X)$  having the same graph as  $f: X \rightarrow Y$ .

*Note:* The issue of enlarging the *domain* of a continuous map is taken up in the next subsection “Extending and gluing maps” ([page 307](#)).

inclusion map!continuous map@as  
identity map!continuous map@as co  
continuous map!composite of maps  
restriction of map!continuous map@  
continuous map!restriction@and res  
continuous map!codomain@and co

ex:proj-prod-spaces-cont  
 topology!coarser  
 topology!finer  
 coarser topology  
 finer topology  
 topologies!comparison of

- (8) Let  $X$  and  $Y$  be topological spaces and give  $X \times Y$  the product topology as defined in [Examples 2.72 \(5\)](#). Then the projections

$$\begin{aligned} p: X \times Y &\rightarrow X, & q: X \times Y &\rightarrow Y \\ \langle x, y \rangle &\mapsto x & \langle x, y \rangle &\mapsto y \end{aligned}$$

are continuous. In fact, for an arbitrary open subset  $U$  of  $X$ , its inverse image  $p^{-1}U = U \times Y$ , which is one of the basic open sets used to define the product topology; thus  $p$  is continuous. And similarly  $q$  is continuous.

ex:diagonal-map-cont

- (9) Let  $X$  be a topological space and give the product  $X \times X$  of  $X$  with itself the product topology. Then the **diagonal map**

$$\begin{aligned} \delta_X: X &\rightarrow X \times X \\ x &\mapsto \langle x, x \rangle \end{aligned}$$

is continuous. The proof is requested in [Exercise 4](#).

ex:coarser-finer-topologies

- (10) Let  $\mathcal{T}$  and  $\mathcal{S}$  be two topologies on the same underlying set  $X$ . Consider the identity map  $i: X \rightarrow X$  as a map  $i: \langle X, \mathcal{T} \rangle \rightarrow \langle X, \mathcal{S} \rangle$  from the topological space  $\langle X, \mathcal{T} \rangle$  to the topological space  $\langle X, \mathcal{S} \rangle$ .

[Generally, given a map  $f: X \rightarrow Y$  between sets, and given topologies  $\mathcal{T}$  and  $\mathcal{S}$  on  $X$  and  $Y$ , respectively, we write

$$f: \langle X, \mathcal{T} \rangle \rightarrow \langle Y, \mathcal{S} \rangle$$

when continuity of  $f$  with respect to these topologies is at issue.]

Since  $i^{-1}(V) = V$  for each  $V \subset X$ , then

$$i: \langle X, \mathcal{T} \rangle \rightarrow \langle X, \mathcal{S} \rangle \text{ is continuous} \iff \mathcal{S} \subset \mathcal{T}.$$

We express the relation  $\mathcal{S} \subset \mathcal{T}$  by saying that  $\mathcal{T}$  is **finer** or **stronger** or **larger than**  $\mathcal{S}$ , and that  $\mathcal{S}$  is **coarser** or **weaker** or **smaller than**  $\mathcal{T}$ .

For example, the discrete topology is finer than, and the indiscrete topology weaker than, every topology on  $X$ .

[Example 1.37](#) concerning the set  $X$  of all continuous functions  $x: [0, 1] \rightarrow \mathbb{R}$  shows that each  $d_1$ -ball at a point  $x \in X$  contains a  $d_\infty$ -ball at  $x$ , but not conversely; hence the topology on  $X$  induced by the metric  $d_\infty$  is *strictly* finer than the topology induced by  $d_1$ ; in other words, the  $d_\infty$ -topology is finer than but distinct from the  $d_1$ -topology.

An arbitrary pair of topologies on a set  $X$  need not of course, be comparable with one another; that is, neither need be finer than the other (see [Exercise 2.8](#)).

If the topologies  $\mathcal{S}$  and  $\mathcal{T}$  on  $X$  satisfy  $\mathcal{S} \subset \mathcal{T}$  and if  $\langle Y, \mathcal{U} \rangle$  is any topological space, then each continuous map  $f: \langle X, \mathcal{S} \rangle \rightarrow \langle Y, \mathcal{U} \rangle$  is also continuous as a map  $f: \langle X, \mathcal{T} \rangle \rightarrow \langle Y, \mathcal{U} \rangle$ .  $\diamond$

**Intuitive idea—finer topology.** In the terminology and notation of [Examples 3.9 \(10\)](#):

- the finer the topology on  $X$ , the more open sets there are and hence the more neighborhoods of each point there are; and
- the finer the topology on  $X$ , the more continuous maps on  $X$  there are.

At times we shall want to look at the entire set of continuous maps between two spaces or two pointed spaces (Definition 2.4).

extension!continuous map@of conti  
continuous map!extension@and ext

def:all-cont-maps

**3.10 Definition.** If  $X$  and  $Y$  are topological spaces, then the set of all continuous maps  $f: X \rightarrow Y$  is denoted by  $C(X, Y)$ .  
If  $\langle X, x \rangle$  and  $\langle Y, y \rangle$  are pointed spaces, then the set of all continuous maps  $f: X \rightarrow Y$  for which  $f(x) = y$  is denoted by  $C[(X, x), (Y, y)]$ .

In the special case that  $Y = \mathbb{R}$ , the set  $C(X, Y)$  becomes the set  $C(X)$  of all continuous real-valued functions on  $X$ .

In terms of the notation just defined, a sophisticated version of Theorem 3.8 is that, for topological spaces  $X, Y$ , and  $Z$ , composition yields

$$\begin{aligned} C(X, Y) \times C(Y, Z) &\rightarrow C(X, Z). \\ \langle f, g \rangle &\mapsto g \circ f \end{aligned}$$

And then for pointed topological space  $\langle X, x \rangle$ ,  $\langle Y, y \rangle$ , and  $\langle Z, z \rangle$ , composition yields a map

$$\begin{aligned} C[(X, x), (Y, y)] \times C[(Y, y), (Z, z)] &\rightarrow C[(X, x), (Z, z)]. \\ \langle f, g \rangle &\mapsto g \circ f \end{aligned}$$

### Extending and gluing maps

subsec:extend-glue

Although a restriction of a continuous map to a subspace of its domain is necessarily continuous, what about an extension from a subspace? More precisely, let

$$g: A \rightarrow Y$$

be a continuous map whose domain is a subspace  $A$  of a topological space  $X$ . Then:

- Must a given extension  $f: X \rightarrow Y$  of  $g$  to  $X$  be continuous?
- Must  $g$  have some continuous extension  $f: X \rightarrow Y$  to all of  $X$ ?
- If  $g$  does have a continuous extension  $f: X \rightarrow Y$  to  $X$ , must such an extension be unique?

By examples we shall see that all three questions have the answer “no.” Then we shall consider some conditions under which continuous extensions do exist or are unique.

ex:discont-extension

**3.11 Examples.** (1) *A given extension of a continuous map need not be continuous.* In other words, if  $A \subset X$  and  $g = f|_A: A \rightarrow Y$  is continuous, then the extension  $f$  of  $g$  to  $X$  need not be continuous.

For example, take  $X = Y = \mathbb{R}$ ,  $A = \mathbb{R}^* = \mathbb{R} \setminus \{0\}$ , and  $f: X \rightarrow Y$  the map defined by

$$f(x) = \begin{cases} \sin(1/x) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Then  $g = f|_A$  is continuous, being the composite of the continuous functions  $x \mapsto 1/x$  of  $A \rightarrow \mathbb{R}$  and  $y \mapsto \sin y$  of  $\mathbb{R} \rightarrow \mathbb{R}$ . However,  $f$  cannot be continuous at  $0 \in X$  because its graph oscillates between the lines  $y = -1$  and  $y = 1$  arbitrarily close to the origin (see Figure 3.1).

More precisely, the reason that  $f$  is not continuous at 0 is as follows. Let  $V$  be any neighborhood of  $0 = f(0)$  in  $Y = \mathbb{R}$ . Then  $] -\varepsilon, \varepsilon[ \subset V$  for some  $\varepsilon$  with  $0 < \varepsilon < 1$ .

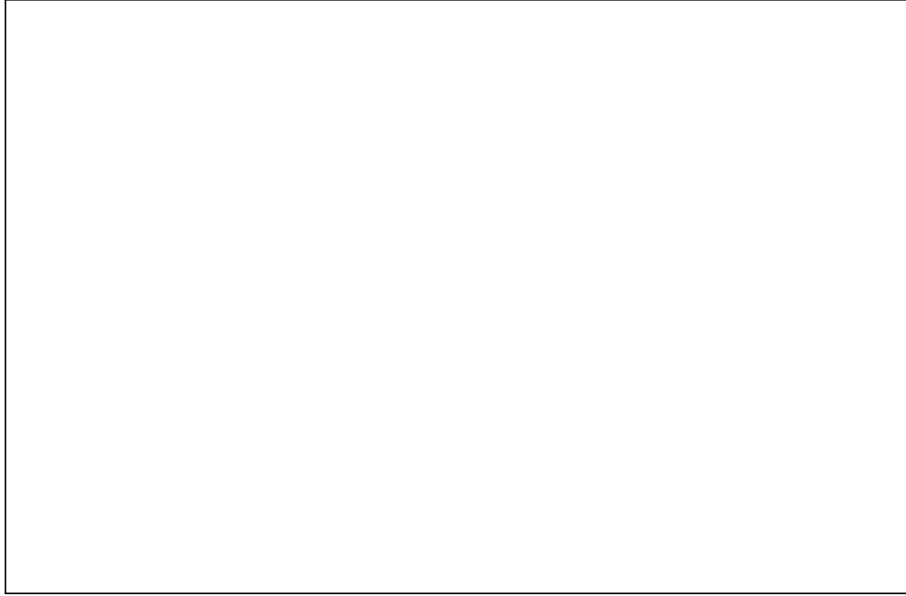
Figure 3.1: Oscillating function  $\sin(1/x)$ .

fig:sin-recip-oscillate-near-0

Each neighborhood  $U$  of 0 in  $X = \mathbb{R}$  contain the number  $x_n = 2/[(4n+1)\pi]$  for some sufficiently large positive integer  $n$ ; then the value  $f(x_n) = \sin[(4n+1)\pi/2] = 1 \notin V$ , and so  $f(U) \not\subset V$ .

ex:no-continuous-extension

- (2) **A continuous map need not have any continuous extension!** For example, consider again the function  $g: A = \mathbb{R}^* \rightarrow Y = \mathbb{R}$  of (1), given by  $x \mapsto \sin(1/x)$ . Any extension  $f: X = \mathbb{R} \rightarrow Y$  of  $g$  will be determined by the single value  $f(0)$ . Even if we take  $f(0)$  to be some number other than 0, the function  $f$  will still not be continuous at 0, as a slight refinement of the preceding argument shows.

In the jargon of advanced calculus,  $f$  has a “nonremovable discontinuity” at  $x = 0$ ; in the language of elementary calculus,  $f$  cannot be continuous at  $x = 0$ , no matter the value  $f(0)$ , because  $\lim_{x \rightarrow 0} f(x)$  does not exist.

x:non-unique-continuous-extension

- (3) **When a continuous map does have a continuous extension, the extension need not be unique.** For example, let  $X$  be a discrete space, let  $A$  be a subspace of  $X$  with  $A \neq X$ , and let  $Y$  contain at least two points. Then any continuous map  $g: A \rightarrow Y$  can be extended to  $X$  in at least two different ways—by assigning one or the other of two distinct points of  $Y$  as a constant value on  $X \setminus A$ .  $\diamond$

There is one important situation when a continuous extension, if it exists at all, must be unique.

thm:extend-identities

**3.12 Theorem (extension of identities).** Let  $f: X \rightarrow Y$  and  $g: X \rightarrow Y$  be continuous maps from the same topological space  $X$  into a Hausdorff space  $Y$ . Suppose  $f|_A = g|_A$  for some dense subset  $A$  of  $X$ . Then  $f = g$ .

**Proof.** Assume that  $f(x) \neq g(x)$  at some  $x \in X$ . Choose disjoint open neighborhoods  $V$  of  $f(x)$  and  $W$  of  $g(x)$  in  $Y$ . The set  $U$  defined by

$$U = f^{-1}(V) \cap g^{-1}(W)$$

is open in  $X$  and, since it contains  $x$ , nonempty. Since  $A$  is dense in  $X$ , there is some  $a \in A \cap U$ . Then  $f(a) \in V$  and  $g(a) \in W$ . This is impossible because  $f(a) = g(a)$  whereas  $V$  and  $W$  are disjoint.  $\square$

Thus under the hypotheses of the theorem, the identity  $f(x) = g(x)$  holds for all  $x$  in the entire domain if it holds for all  $x$  in a dense subset. For example, a continuous function  $f: \mathbb{R} \rightarrow \mathbb{R}$  taking the value 0 at each rational number must take the value 0 at every real number.

We already saw that a continuous map  $g: A \rightarrow Y$  on a subspace  $A$  of a space  $X$  need not have an extension to  $X$ .

When *does* a continuous map on a subspace have a continuous extension to the entire space? This is one of the central questions of topology. A general situation where continuous extensions *do* always exist is given below, in [Examples 3.14 \(2\)](#).

Another instance where *no* continuous extension exists is

$$A = Y = S_1 = \{x \in \mathbb{R}^2 : \|x\| = 1\}, \quad X = D_2 = \{x \in \mathbb{R}^2 : \|x\| \leq 1\},$$

and  $g: A \rightarrow Y$  is the identity map of the unit circle  $A$ . Any continuous extension  $f: X \rightarrow Y$  would be a continuous map of the unit disk  $D_2$  onto its bounding circle  $S_1$  that leaves each point of this circle fixed, that is,  $f(x) = x$  for each  $x \in A = S_1$ . You are hereby challenged to prove no such  $f$  exists—*before* you see the machinery developed in [Sections 5.4–5.5](#) of [Chapter 5](#) (Connectedness)!

What happens if the restrictions of a map  $f: X \rightarrow Y$  to each of a whole family  $\langle A_i : i \in I \rangle$  of subspaces of  $X$  are continuous? To have any hope of inferring continuity of  $f$  on the entire space  $X$ , we should certainly require that each point of  $X$  belong to at least one of the sets  $A_i$ , that is,

$$X = \bigcup_{i \in I} A_i.$$

We express this condition by saying that  $\langle A_i \rangle_{i \in I}$  **covers**  $X$ . However, this condition alone is still not enough to guarantee that  $f$  is continuous—see [Examples 3.9 \(1\)](#) again.

thm:gluing **3.13 Theorem (Gluing Lemma).** Let  $\langle A_i \rangle_{i \in I}$  be a family of subspaces of a topological space  $X$  that covers  $X$ . Suppose either that

- (i) each  $A_i$  is open in  $X$ , or else that
- (ii) the index set  $I$  is finite and each  $A_i$  is closed in  $X$ .

Then a map  $f: X \rightarrow Y$  into a topological space  $Y$  is continuous if its restrictions  $f|_{A_i}: A_i \rightarrow Y$  are continuous for all  $i \in I$ .

**Proof.** Assume that the restriction of  $f$  to each  $A_i$  is continuous.

pf-case:each-Ai-open Case (i): each  $A_i$  is open in  $Y$ . Let  $V$  be an arbitrary open set in  $Y$ . For each  $i \in I$  the set

$$f^{-1}(V) \cap A_i = (f|_{A_i})^{-1}(V)$$

is open in  $A_i$  by [Theorem 3.3](#) and hence is open in  $X$ . Then the union

$$f^{-1}(V) = f^{-1}(V) \cap \bigcup_{i \in I} A_i = \bigcup_{i \in I} (f^{-1}(V) \cap A_i)$$

is open in  $X$ . By [3.3](#) again,  $f$  is continuous.

pf-case:each-Ai-closed Case (ii): the index set  $I$  is finite and each  $A_i$  is closed in  $X$ . The same kind of calculations as in Case (i) show here that the inverse image  $f^{-1}(E)$  of an arbitrary *closed* subset  $E$  of  $Y$  is closed in  $X$ .  $\square$

**absolute-value function** Typically in applications of the Gluing Lemma, we do not start with a map  $f$  already defined on the entire space  $X$ . Rather, we are given a whole family of maps

**Tietze Extension Theorem**  
**normal space**

$$f_i: A_i \rightarrow Y \quad (i \in I),$$

one for each of the subspaces  $A_i$  of  $X$ . Suppose that these maps “match up” on the overlaps of their domains, that is, if

$$f_i|(A_i \cap A_j) = f_j|(A_i \cap A_j) \quad (i, j \in I).$$

Then they may be “glued together”—see [Proposition 0.21](#)—to form a single map  $f: X \rightarrow Y$  defined by the rule that for  $x \in X$ ,

$$f(x) = f_i(x) \text{ where } i \in I \text{ and } x \in A_i.$$

This map is well defined because  $x \in A_i$  and  $x \in A_j$  implies  $f_i(x) = f_j(x)$ . Thus  $f: X \rightarrow Y$  is the unique map such that

$$f|_{A_i} = f_i \quad (i \in I).$$

According to the Gluing Lemma,  $f$  will be continuous if each  $f_i$  is.

**ex:abs-value** **ex:gluing** **3.14 Examples.** (1) Write the definition of the absolute-value function  $f: \mathbb{R} \rightarrow \mathbb{R}$  in the form

$$f(x) = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x \leq 0. \end{cases}$$

Then the Gluing Lemma may be used to establish continuity of  $f$  as follows. Let  $I = \{1, 2\}$  and take

$$A_1 = \{x \in \mathbb{R} : x \geq 0\}, \quad A_2 = \{x \in \mathbb{R} : x \leq 0\},$$

so that the family  $\langle A_i \rangle_{i \in I}$  of two closed sets covers  $\mathbb{R}$ . The restriction  $f|_{A_1}: A_1 \rightarrow \mathbb{R}$  is the inclusion map, which is continuous; the restriction  $f|_{A_2}: A_2 \rightarrow \mathbb{R}$  is also continuous, being a restriction of the continuous function  $x \mapsto -x$  from  $\mathbb{R}$  into  $\mathbb{R}$ .

Of course continuity of  $f$  may also be established directly by an  $\varepsilon$ - $\delta$  argument.

**ex:Tietze-for-metrizable** (2) Here is a nontrivial application of the Gluing Lemma. Any continuous function

$$g: A \rightarrow [0, 1]$$

on a closed subspace  $A$  of a metrizable space  $X$  has a continuous extension  $f: X \rightarrow [0, 1]$  to  $X$ .

This result is a special case of the *Tietze Extension Theorem*, which asserts that the same thing is true more generally for any *normal* space  $X$ . [Normal spaces were defined in [Definition 2.94](#) (5).] That every metrizable space is normal is [Proposition 2.100](#), treated in [Exercise 1.42 \(b\)](#). The general Tietze Extension Theorem (6.34) is proved later.

For technical reasons it is simpler to extend not  $g: A \rightarrow [0, 1]$  but instead the continuous function

$$\begin{aligned} G: A &\rightarrow [1, 2] \\ x &\mapsto g(x) + 1 \end{aligned}$$

given by  $G(x) = g(x) + 1$ . Once a continuous extension  $F: X \rightarrow [1, 2]$  has been constructed, the desired continuous extension of  $g$  will be the function

$$\begin{aligned} f: A &\rightarrow [0, 1] \\ x &\mapsto F(x) - 1 \end{aligned}$$

defined by  $f(x) = F(x) - 1$ .



To construct  $F$ , choose a metric  $d$  that induces the topology of  $X$ . According to [Proposition 1.29](#), the distance

$$d(x, A) = \inf_{a \in A} d(x, a)$$

from a point  $x \in X$  to the closed set  $A$  is positive when  $x \notin A$ . Then it is meaningful to define

$$F(x) = \begin{cases} G(x) & \text{if } x \in A, \\ H(x)/d(x, A) & \text{if } x \notin A, \end{cases}$$

where

$$H(x) = \inf_{a \in A} G(a) d(x, a) \quad (x \in X \setminus A).$$

Since  $1 \leq G(a) \leq 2$  for all  $a \in A$ , then  $1 \leq F(x) \leq 2$  for all  $x \in X \setminus A$ . Thus an extension  $F: X \rightarrow [1, 2]$  of  $G$  has been constructed.

To see that  $F$  is continuous, consider its restrictions to the closed subspaces

$$A_1 = A = \text{cls } A, \quad A_2 = \text{cls}(X \setminus A) = (X \setminus A) \cup \text{bdy } A$$

of  $X$ . Now  $A_1 \cup A_2 = X$ , and the restriction  $F|_{A_1} = G$  is already known to be continuous. So it remains only to show that the restriction  $F|_{A_2}$  is continuous. Let  $x \in A_2$  and consider two cases.

Case (i):  $x \in X \setminus A$ .

According to [Exercise 1.81 \(a\)](#), the real-valued function  $u \mapsto d(u, A)$  is continuous at  $x$ . If we can establish continuity of the function  $H: X \setminus A \rightarrow \mathbb{R}$  at  $x$ , then continuity of  $F|_{A_2}$  will follow by the standard argument used to establish continuity of the quotient of a real-valued function by a positive-valued function on  $\mathbb{R}$  [compare [Exercise 119 \(a\)](#)].

Let  $\varepsilon > 0$  be arbitrary. Let  $u \in X \setminus A$  with

$$d(x, u) < \frac{\varepsilon}{2}.$$

Then  $a \in A$  implies

$$d(x, a) \leq d(x, u) + d(u, a) < \frac{\varepsilon}{2} + d(u, a)$$

and, since  $1 \leq G(a) \leq 2$ ,

$$G(a) d(x, a) < \varepsilon + G(a) d(u, a).$$

Hence

$$H(x) \leq \varepsilon + H(u).$$

Similarly,

$$H(u) \leq \varepsilon + H(x).$$

Thus

$$|H(u) - H(x)| \leq \varepsilon.$$

Case (ii):  $x \in \text{bdy } A$ .

continuous map!gluing@and gluing  
extension!continuous map@of continuous map

Let  $\varepsilon > 0$  be arbitrary. By continuity of  $G$  at  $x$  there is some  $\delta > 0$  such that

$$\{eq:ineq-for-glue-extension-ex\} (*) \quad a \in A \quad \text{and} \quad d(x, a) < \delta \implies |G(a) - G(x)| < \varepsilon.$$

Already, then,  $|F(u) - F(x)| < \varepsilon$  for  $d(x, u) < \delta/4$  in case  $u \in A$ . We shall show that the same thing is true in case  $u \in X \setminus A$ . Let  $u \in X \setminus A$  with

$$d(x, u) < \frac{\delta}{4}.$$

The infimum

$$H(u) = \inf_{a \in A} G(a) d(u, a)$$

will be unchanged if we restrict  $a$  so as to vary over just the subset

$$B = A \cap B_\delta(x; d)$$

of  $A$ , that is,

$$\{eq:restricted-inf-in-glue-extension-ex\} (**) \quad H(u) = \inf_{a \in B} G(a) d(u, a).$$

In fact,  $a \in A \setminus B$  implies

$$d(u, a) \geq d(x, a) - d(x, u) > \delta - \frac{\delta}{4} = 3 \frac{\delta}{4}$$

so that

$$\inf_{a \in B} G(a) d(u, a) \geq 3 \frac{\delta}{4},$$

whereas  $x \in B$  and

$$G(x) d(u, x) \leq 2 d(u, x) < \frac{\delta}{2} < 3 \frac{\delta}{4}.$$

By the same kind of argument just used,

$$\inf_{a \in B} d(u, a) = d(u, A),$$

and from (\*)

$$G(x) - \varepsilon < G(a) < G(x) + \varepsilon \quad (a \in B).$$

Hence

$$(G(x) - \varepsilon) d(u, A) \leq \inf_{a \in B} G(a) d(u, a) \leq (G(x) + \varepsilon) d(u, A).$$

Using (\*\*) we conclude

$$|F(u) - F(x)| = |F(u) - G(x)| \leq \varepsilon. \quad \diamond$$

### Open and closed maps

subsec:open-closed-maps

[Theorem 3.3](#) and [Corollary 3.5](#) say that a map  $f: X \rightarrow Y$  is continuous precisely when the *inverse* image  $f^{-1}(B)$  of each open or closed subset of  $Y$  is open or closed, respectively, in  $X$ . However, they say nothing whatsoever about the *direct* image  $f(A)$  of an open or closed subset  $A$  of  $X$  being open or closed in  $Y$ .

def:open-closed-map

**3.15 Definition.** A map  $f: X \rightarrow Y$  between topological spaces is said to be **open** if the image  $f(U)$  of each open subset  $U$  of  $X$  is open in  $Y$ , and **closed** if the image  $f(E)$  of each closed subset  $E$  of  $X$  is closed in  $Y$ .

projection!closed map@and closed  
closed map!projection@and project  
quotient map

exs:open-closed-maps

**3.16 Examples.** (1) Let  $X$  be  $\mathbb{R}$  with its usual topology. Then the identity map  $\iota_X: X \rightarrow X$  is continuous, open, and closed.

(2) Again let  $X$  be  $\mathbb{R}$  with its usual topology, and let  $Y$  be  $\mathbb{R}$  with its discrete topology. Then the identity map  $X \rightarrow Y$  is *not* continuous but is both open and closed. The identity map  $Y \rightarrow X$ , on the other hand, is continuous but neither open nor closed.

proj-from-R2-cont-open-not-closed

(3) The projection

$$p_1: \mathbb{R}^2 \rightarrow \mathbb{R} \\ \langle x_1, x_2 \rangle \mapsto x_1$$

of the plane onto the horizontal axis is continuous. To see this, we use the max metric  $d_\infty$  on  $\mathbb{R}^2$  and the Euclidean metric  $d$  on  $\mathbb{R}$ . Given  $x = \langle x_1, x_2 \rangle \in \mathbb{R}^2$  and any  $\varepsilon > 0$ , then each  $u = \langle u_1, u_2 \rangle \in \mathbb{R}^2$  that satisfies

$$d_\infty(x, u) = \max\{d(x_1, u_1), d(x_2, u_2)\} < \varepsilon$$

will necessarily satisfy

$$d(p_1(x), p_1(u)) = d(x_1, u_1) < \varepsilon.$$

The projection  $p_1$  is also an open map. In fact, let  $U$  be an open subset of  $\mathbb{R}^2$ . To see that  $p_1(U)$  is open in  $\mathbb{R}$ , let  $x_1 \in p_1(U)$  be arbitrary. Then for some  $x_2$ , the point  $x = (x_1, x_2) \in U$ . For some  $\varepsilon > 0$ , the square region

$$]x_1 - \varepsilon, x_1 + \varepsilon[ \times ]x_2 - \varepsilon, x_2 + \varepsilon[ = B_\varepsilon(x; d_\infty) \subset U.$$

Then  $p_1(U)$  contains the neighborhood  $]x_1 - \varepsilon, x_1 + \varepsilon[$  of  $x_1$  (see the left-hand side of Figure 3.2).

The projection  $p_1$  is, however, *not* closed! In fact, the hyperbolic arc

$$E = \{\langle x_1, x_2 \rangle \in \mathbb{R}^2 : x_1 > 0, x_2 = 1/x_1\}$$

is closed in  $\mathbb{R}^2$ , but its image

$$p_1(E) = \{x_1 \in \mathbb{R} : x_1 > 0\}$$

is certainly not closed in  $\mathbb{R}$  (see the right-hand side of Figure 3.2).

Of course everything just proved about the first projection  $p_1$  holds as well for the second projection  $p_2$ .  $\diamond$

The open or closed maps in the preceding examples are continuous surjections. Both continuous open surjections and continuous closed surjections are instances of a more general class of maps—*quotient maps*—that are introduced in Section 3.4.

Two facts about open maps will be useful later. The first, an analog of Corollary 3.4 (1), generalizes the technique used in Examples 3.16 (3) to show that the projection  $p_1: \mathbb{R}^2 \rightarrow \mathbb{R}$  is open. (For a “local” version of this fact, see Exercise 2.)

continuous image

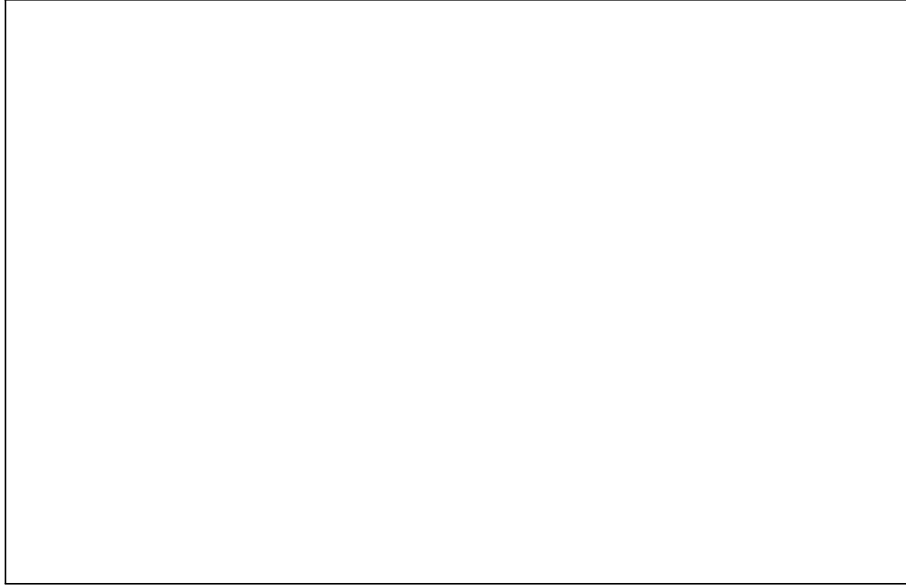


Figure 3.2: Projections of an open square and a hyperbolic arc in the plane.

fig:p1-images-plane-box

prop:open-map-via-base

**3.17 Proposition.** A map  $f: X \rightarrow Y$  between topological spaces is open if there is some base  $\mathcal{B}$  of  $X$  such that  $f(B)$  is open in  $Y$  for each  $B \in \mathcal{B}$ .

prop:open-bij-via-closed-sets

**3.18 Proposition.** A bijection  $f: X \rightarrow Y$  is an open map if and only if it is a closed map.

**Proof.** Use the fact that if  $f$  is a bijection, then

$$Y \setminus f(A) = f(X \setminus A)$$

for all subsets  $A$  of  $X$ . (Compare the proof of [Corollary 3.5](#).)  $\square$

### Continuous images of spaces

subsec:cont-images-spaces

We have met various “topological properties”: the separation properties  $T_0$ ,  $T_1$ ,  $T_2$ , regular, etc.; the countability properties first-countable, second-countable, and separable; and metrizable. (The precise meaning of ‘topological property’ will be explained in [Section 3.2](#).) Earlier it was pointed out that we want to know whether a given topological property is hereditary, that is, every subspace of a space having the property also has the property.

def:cont-image

Another thing we want to know about a topological property is whether it is “preserved under continuous images”—that is, whether every continuous image of a topological space having the property also has the property. By a **continuous image** of a topological space  $X$  we mean a topological space  $Y$  for which there is some continuous *surjection*  $X \rightarrow Y$ . The following result addresses this question for one of the countability properties; its proof is left as an exercise.

prop:cont-image-separable

**3.19 Proposition.** *The continuous image of a separable space is itself separable.*first-countable space!continuous image  
Fort topology

Are first-countability and second-countability also preserved under continuous images?

second-countable space!continuous  
finite-complement topology

image-not-preserve-2nd-countable

**3.20 Examples.** (1) Let  $X$  be an uncountable set with its discrete topology and let  $Y$  be the same set but with its Fort topology [Examples 2.66 (3)], so that  $Y$  is the continuous image of  $X$  under the identity map. The space  $X$  is first-countable but  $Y$  is not.

image-not-preserve-2nd-countable

(2) Let  $X = \mathbb{R}$  with its usual topology and let  $Y = \mathbb{R}$  with its finite-complement topology. The identity map from  $X$  to  $Y$  is continuous because the complement of any finite subset of the real line is open. We already know that  $X$  is second-countable. However,  $Y$  is *not* second-countable, as you may verify [compare Exercise 2.97 (b)].  $\diamond$ 

Thus neither first-countability nor second-countability is preserved under continuous images. However, each *is* preserved if the continuous surjection is further assumed to be an open map.

part:cont-open-image-1st-countable

**3.21 Proposition.** (1) *The image of a first-countable space under a continuous open map is itself first-countable.*

part:cont-open-image-2nd-countable

(2) *The image of a second-countable space under a continuous open map is itself second-countable.*

**Proof.** (1) Let  $f: X \rightarrow Y$  be a continuous open surjection from a first-countable space  $X$  onto a topological space  $Y$ . Let  $y \in Y$  be arbitrary. Since  $f$  is surjective, there is an  $x \in X$  with  $f(x) = y$ . Since  $X$  is first-countable, there is a countable local base  $\mathcal{M}$  at  $x$  in  $X$  that consists of *open* sets. Form the collection

$$\mathcal{N} = \{f(M) : M \in \mathcal{M}\}$$

of images of members of  $\mathcal{M}$ . We shall show that  $\mathcal{N}$  is the desired countable local base at  $y$  in  $Y$ .

First,  $\mathcal{N}$  is also countable. In fact,  $\mathcal{N}$  is the image, under the map

$$\begin{aligned} f^*: \mathcal{P}(X) &\rightarrow \mathcal{P}(Y) \\ A &\mapsto f(A), \end{aligned}$$

of the countable subcollection  $\mathcal{M}$  of the domain of  $f^*$ . By Proposition 0.48,  $\mathcal{N}$  is also countable.

Second,  $\mathcal{N}$  is a local base at  $y$  in  $Y$ . In fact, let  $V$  be an arbitrary neighborhood of  $y$  in  $Y$ . Then  $f^{-1}(V)$  is a neighborhood of  $x$  in  $X$ , and so there exists some  $M \in \mathcal{M}$  such that  $x \in M \subset f^{-1}(V)$ . Then  $f(M)$  is an open set in  $Y$  with  $y \in f(M) \subset V$ .

(2) See Exercise 39.  $\square$

The continuous image of a topological space having even the weakest separation property, namely  $T_0$ , need not be a  $T_0$  space. For example, let  $X$  be any discrete space consisting of more than one point and let  $Y$  be the indiscrete space having the same underlying set as  $X$ . Then  $X$  is a  $T_0$ -space—in fact, a regular  $T_1$ -space (hence a  $T_2$ -space). However, the continuous image  $Y$  of  $X$  has none of those properties.

## EXERCISES FOR SECTION 3.1

- continuous map!point-continuity
- local base!continuity@and continuity
- prob:comp-cont-at-pt
- continuous map!local base@and local base
- line with two origins
- half-disk space
- prob:cont-via-local-base
- only via local base
- continuous map!closure@at a point
- closure
- prob:pf-cont-via-subbase
- prob:diag-map-cont-pf
- b-part:cont-map-on-half-disk-space
- prob:part-cont-on-half-disk-space
- prob:cont-image-of-separable
1. Prove the following “local” version of [Theorem 3.8](#): If  $f: X \rightarrow Y$  is continuous at  $x \in X$  and if  $g: Y \rightarrow Z$  is continuous at  $f(x)$ , show that the composite  $g \circ f$  is continuous at  $x$ .
  2. Prove the following “local” version of [Corollary 3.4 1](#): A map  $f: X \rightarrow Y$  is continuous at a point  $x \in X$  if there is some local base  $\mathcal{M}$  at  $f(x)$  in  $Y$  such that  $f^{-1}(M)$  is a neighborhood of  $x$  for each  $M \in \mathcal{M}$ .
  3. Prove part (2) of [Corollary 3.4](#).
  4. Given an arbitrary topological space  $X$ , prove that the diagonal map  $\delta_X: X \rightarrow X \times X$  [[Examples 3.9 \(9\)](#)] is continuous.
  5. Is the map  $f: \mathbb{Q} \rightarrow \mathbb{R}$  such that  $f(x) = 0$  if  $x^2 < 2$ , and  $f(x) = 0$  otherwise, a continuous map?
  6. Let  $X = \mathbb{R} \cup \{0'\}$  be the line with two origins [[Examples 2.20 \(3\)](#)]. Which of the following maps are continuous?
    - (a) The map  $f: X \rightarrow \mathbb{R}$  with  $f(x) = x$  if  $x \in \mathbb{R}$  and  $f(0') = 0$ .
    - (b) The map  $r: [-1, 1] \rightarrow X$  with  $r(x) = x$  for all  $x$ .
    - (c) The map  $r': [-1, 1] \rightarrow X$  with  $r(x) = x$  for all  $x \neq 0$  and  $r'(0) = 0'$ .
    - (d) The map  $\sigma: [0, 2] \rightarrow X$  with  $\sigma(t) = t$  for  $0 \leq t \leq 4$ ,  $\sigma(t) = 2 - t$  for  $1 \leq t < 2$ , and  $\sigma(2) = 0'$ .
  7. (a) If  $X = H \cup L$  is the half-disk space [[Examples 2.20 \(3\)](#)], is the map  $\langle x, y \rangle \mapsto \langle x, 0 \rangle$  from  $X$  to its subspace  $L$  continuous? Is it open? Do your answers remain the same if  $L$  is given its usual topology instead of its relative topology as a subspace of  $L$ .  
 (b) Repeat (a) if  $X$  is instead the tangent disk space  $\Gamma$  ([Exercise 2.37](#)).
  8. (a) Strengthen [Corollary 3.6](#) by showing that a map  $f: X \rightarrow Y$  is continuous if and only if  $f(\text{cls } A) \subset \text{cls } f(A)$  for each subset  $A$  of  $X$ .  
 (b) Is there an analogous characterization of continuity involving interiors?
  9. Prove: If  $f: X \rightarrow Y$  is a continuous *surjection* from a separable space  $X$  onto a topological space  $Y$ , then  $Y$  must also be separable. (Note: This is [Proposition 3.19](#).)
  10. Does the property that every map  $f: X \rightarrow Y$  on a discrete space  $X$  to an arbitrary topological space be continuous actually characterize discrete spaces?
  11. Determine all topological space  $Y$  such that every map  $f: X \rightarrow Y$  for every topological space  $X$  is continuous.
  12. If  $f: X \rightarrow Y$  is a continuous injection whose codomain  $Y$  is a Hausdorff space  $Y$ , must its domain  $X$  be a Hausdorff space?
  13. Show that a topological space  $X$  must be discrete if every continuous open map with domain  $X$  has as its range a Hausdorff space.
  14. Show that the collection of all subsets of  $\mathbb{R}$  that are both open for the usual topology and bounded for the usual metric is a topology on  $\mathbb{R}$  that is strictly coarser than the usual topology.

15. Which of these topologies on an infinite set  $X$  are finer than which others:

Fort topology

Fortissimo topology

- the finite-complement topology [Examples 2.3 (7)];
- the Fort topology [Examples 2.66 (3)];
- the countable-complement topology [Exercise 2.7]; and
- the Fortissimo topology [Exercise 2.111]?

Consider separately the cases when  $X$  is denumerable and when  $X$  is uncountable.

16. Let  $\mathcal{T}$  and  $\mathcal{S}$  be two topologies on the same set  $X$  such that  $\mathcal{S}$  is finer than  $\mathcal{T}$ .

- Given a subset  $A$  of  $X$ , compare the boundary, interior, and closure of  $A$  for  $\mathcal{S}$  with the boundary, interior, and closure, respectively, for  $\mathcal{T}$ .
- Discuss the relationship between the various countability properties holding for  $\langle X, \mathcal{T} \rangle$  and their holding for  $\langle X, \mathcal{S} \rangle$ .
- Discuss the relationship between the separation properties  $T_0$ ,  $T_1$ , and  $T_2$  holding for  $\langle X, \mathcal{T} \rangle$  and their holding for  $\langle X, \mathcal{S} \rangle$ .

17. Prove that the relative topology on a subset  $A$  of a topological space  $X$  is the coarsest topology making the inclusion map  $j: A \rightarrow X$  continuous.

prob:prod-top-2-coarsest

18. Let  $X$  and  $Y$  be topological spaces. Show that the product topology [Examples 2.72 (5)] is the coarsest topology on the set  $X \times Y$  that makes both the projections  $p: X \times Y \rightarrow X$  and  $q: X \times Y \rightarrow Y$  continuous.

-2-spaces-cont-iff-components-cont

19. Let  $f: Z \rightarrow X \times Y$  be a map from a topological space  $Z$  into the product [Examples 2.72 (5)] of two topological spaces. Let  $p: X \times Y \rightarrow X$  and  $q: X \times Y \rightarrow Y$  be the projections. Show that  $f$  is continuous if and only if its “components”  $p \circ f: Z \rightarrow X$  and  $q \circ f: Z \rightarrow Y$  are continuous.

prob:map-l-to-S1-cont

20. Show that the map

$$\begin{aligned} I &\rightarrow \mathbb{R}^2 \\ t &\mapsto \langle \cos 2\pi t, \sin 2\pi t \rangle \end{aligned}$$

from the unit interval  $I = [0, 1]$  into the Euclidean plane  $\mathbb{R}^2$  is continuous.

*Note:* When  $\mathbb{R}^2$  is identified with the set  $\mathbb{C}$  of all complex numbers (see page 37), the preceding function may be written as

$$\begin{aligned} I &\rightarrow \mathbb{C} \\ t &\mapsto \exp(2\pi i t) \end{aligned}$$

21. Given two topologies  $\mathcal{T}$  and  $\mathcal{S}$  on the same set  $X$ , show that there is a finest topology on  $X$  that is coarser than both  $\mathcal{T}$  and  $\mathcal{S}$ , and there is a coarsest topology on  $X$  that is coarser than both  $\mathcal{T}$  and  $\mathcal{S}$ .

22. Show that a sufficient condition for a map  $f: X \rightarrow Y$  to be continuous is that each  $x \in X$  have some neighborhood  $V_x$  such that  $f|_{V_x}$  is continuous.

23. Given a map  $f: X \rightarrow Y$  from a space  $X$  to a space  $Y$  and a subspace  $A$  of  $X$ , show that continuity of  $f|_A$  is *not* necessarily equivalent to the continuity of  $f$  at each point of  $A$ . Find conditions on  $A$  sufficient to ensure that continuity of  $f|_A$  is equivalent to the continuity of  $f$  at each point of  $A$ .

- 24.** Show directly, *without* using [Examples 3.14 \(2\)](#), that a continuous map  $g: [a, b] \rightarrow [0, 1]$  with domain a closed interval  $[a, b]$  in  $\mathbb{R}$  can always be extended continuously to  $\mathbb{R}$ . Can the codomain  $[0, 1]$  here be replaced by other spaces?
- prob:extension-of-equality **25.** Let  $f, g: X \rightarrow \mathbb{R}^n$  be continuous functions such that  $f(x) = g(x)$  for all  $x$  in some dense subset of  $X$ . Deduce that  $f = g$ .
- prob:Cauchy-eq **26.** Let  $f: \mathbb{R} \rightarrow \mathbb{R}$  be a function that is continuous and satisfies *Cauchy's equation*  $f(a+b) = f(a) + f(b)$  for all real numbers  $a$  and  $b$ . Prove that  $f$  is just multiplication by a constant. [*Hint:* To begin, prove algebraically that  $f(0) = 0$ ; next that  $f(-x) = -f(x)$  for all  $x \in \mathbb{R}$ ; and then that  $f(nx) = f(x)n$  for all  $n \in \mathbb{Z}$  and all  $x \in \mathbb{R}$ .]
- 27.** Show that the conclusion of [Theorem 3.12](#) fails if either the subset  $A$  of  $X$  is not dense or the codomain  $Y$  is not a Hausdorff space.
- 28.** Give the details of the proof of [Theorem 3.13](#) in the case that  $\langle A_i \rangle_{i \in I}$  is a finite family of closed sets.
- 29.** Verify that the conclusion of the Gluing Lemma ([Theorem 3.13](#)) still holds if the family  $\langle A_i \rangle_{i \in I}$  covering  $X$  is only a locally finite family ([Exercise 2.63](#)) of closed sets.
- 30.** In the notation of the paragraph on [page 310](#) that follows the proof of the Gluing Lemma ([Theorem 3.13](#)), show that the graph of  $f: X \rightarrow Y$  is the union of the graphs of the maps  $f_i: A_i \rightarrow Y$  for all  $i \in I$ .
- etze-implies-urysohn-for-metrizable **31.** From [Examples 3.14 \(2\)](#) deduce the following special case of *Urysohn's Lemma* ([6.26](#)): If  $A$  and  $B$  are disjoint *closed* sets in a metrizable space  $X$ , then there is a continuous function  $f: X \rightarrow [0, 1]$  with  $f(a) = 0$  for all  $a \in A$  and  $f(b) = 1$  for all  $b \in B$ . (*Note:* A direct proof of this result is indicated in [Exercise 1.82](#).)
- prob:open-map-via-nbds **32.** Prove: A map  $f: X \rightarrow Y$  is open precisely when for each  $x \in X$ , the image  $f(N)$  of each neighborhood  $N$  of  $x$  is a neighborhood of  $f(x)$  in the codomain  $Y$ .
- 33.** When is the inclusion map  $j: A \rightarrow X$  of a subspace  $A$  of a space  $X$  an open map? When is it a closed map? (In each case, give conditions that are necessary and sufficient.)
- prob:ex-real-valued-fns-open-closed **34.** Which of the following continuous functions are open? Which are closed?
- (a) The function  $f: \mathbb{R} \rightarrow \mathbb{R}$  given by  $f(x) = x^2$ .
- (b) The function  $g: \mathbb{R} \rightarrow [0, \infty[$  given by
- $$g(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ x & \text{if } x \geq 0. \end{cases}$$
- (c) The function  $s: \mathbb{R} \rightarrow [-1, 1]$  given by  $s(x) = \sin x$ .
- (d) The function  $h: [0, 1] \cup [2, 3] \rightarrow [0, 2]$  given by
- $$h(x) = \begin{cases} x & \text{if } 0 \leq x \leq 1, \\ x - 1 & \text{if } 2 \leq x \leq 3. \end{cases}$$
- 35.** Let  $f: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \times [0, \infty[$  be the map given by  $f(x, y) = \langle x, |y| \rangle$ , which folds the plane along the  $x$ -axis onto the upper half-plane. Is  $f$  continuous? Is it open? Is it closed?



**36.** A map  $f: X \rightarrow Y$  can be continuous or not continuous, open or not open, closed or not closed. Show by example that all of the resulting  $8 = 2^3$  possibilities do actually occur. Begin with [Examples 3.16](#) and wherever possible construct surjections as your additional examples.

second-countable space!continuous  
continuous image!second-countable  
continuous map!closure@and closur  
lower semicontinuous map  
upper semicontinuous map  
semicontinuous map

**37.** Does an analog of [Theorem 3.8](#) hold for composition of open maps? for closed maps?

**38. (a)** Let  $f: X \rightarrow Y$  be a closed map. Prove: If  $S$  is a subset of  $X$  that is saturated by  $f$  in the sense of [Exercise 0.38](#), then the domain-codomain restriction  $f|_{S, f(S)}: S \rightarrow f(S)$  is also a closed map.

**(b)** Prove the analog of [\(a\)](#) for open maps.

**39.** Prove: The image of a second-countable space under a continuous *open* map is second-countable. [This is [Proposition 3.21 \(2\)](#).]

**40.** Must the image of a  $T_1$ -space under a continuous open map be a  $T_1$  space?

**41.** Prove that a map  $f: X \rightarrow Y$  is both continuous and closed precisely when  $f(\text{cls } A) = \text{cls } f(A)$  for each subset  $A$  of  $X$ . Give an analogous result for continuous open maps.

**42. (a)** If  $f: X \rightarrow Y$  is a continuous *surjection*, show that then a map  $g: Y \rightarrow Z$  is open if  $g \circ f$  is open,

**(b)** If  $g: Y \rightarrow Z$  is a continuous *injection*, show that then a map  $f: X \rightarrow Y$  is open if  $g \circ f$  is open.

**(c)** Do [\(a\)](#) and [\(b\)](#) remain true if 'open' is replaced by 'closed'?

**43.** A real-valued function  $f$  on a topological space  $X$  is said to be **lower semicontinuous** when the inverse image  $f^{-1}(]c, \infty[)$  is open in  $X$  for each  $c \in \mathbb{R}$ , and **upper semicontinuous** when the inverse image  $f^{-1}(]-\infty, c[)$  is open in  $X$  for each  $c \in \mathbb{R}$ .

**(a)** Construct on a suitable topological space  $X$  a real-valued function that is lower semicontinuous but not upper semicontinuous, and a function that is upper semicontinuous but not lower semicontinuous. [See [\(b\)](#), below.]

Prove:

**(b)** The real-valued function  $f$  on  $X$  is continuous if and only if  $f$  is both lower semicontinuous and upper semicontinuous.

**(c)** The real-valued function  $f$  is lower semicontinuous if and only if its negative  $-f$  is upper semicontinuous.

**(d)** The real-valued function  $f$  is lower semicontinuous if and only if for each  $x \in X$  and each  $c < f(x)$ , there is a neighborhood  $U$  of  $x$  with  $c < f(u)$  for all  $u \in U$ . (This explains the terminology 'lower').

**(e)** The analog of [\(d\)](#) for upper semicontinuity. [Hint: Use [\(c\)](#).]

**(f)** The characteristic function  $\chi_A$  of a subset  $A$  of  $X$ , regarded as a function  $X \rightarrow \mathbb{R}$  with codomain  $\mathbb{R}$ , is lower semicontinuous if  $A$  is open in  $X$  and is upper semicontinuous if  $A$  is closed in  $X$ .

## 3.2 Homeomorphisms

sec:homeomorph

Metric equivalence (isometry) and topological equivalence of metric spaces were introduced earlier, in [Section 1.3](#)—specifically, in the subsections “Equivalent metrics on a set” and “Isometries” ([pages 164](#) and [172](#)). Although metric equivalence depends on the particular metrics involved, topological equivalence depends only on the topologies induced by the metrics. Hence the latter notion may be generalized to a purely topological setting.

### Homeomorphisms—definition and examples

topological transformation  
subsec:homeo

def:homeo

**3.22 Definition.** Let  $X$  and  $Y$  be topological spaces. If  $f: X \rightarrow Y$  is a *bijection* such that

- the image  $f(U)$  of each open subset  $U$  of  $X$  is open in  $Y$ , and
- the inverse image  $f^{-1}(V)$  of each open subset  $V$  of  $Y$  is open in  $X$ ,

then we call  $f$  a **homeomorphism from  $X$  to  $Y$**  and write

$$f: X \cong Y$$

When some homeomorphism  $f: X \cong Y$  from  $X$  to  $Y$  exists, we say that  $X$  is **homeomorphic to  $Y$**  and write

$$X \cong Y.$$

Homeomorphisms are sometimes called *topological transformations*.

Any bijection  $f: X \rightarrow Y$  induces a bijection

$$\begin{aligned} f^*: \mathcal{P}(X) &\rightarrow \mathcal{P}(Y) \\ U &\mapsto f(U) \end{aligned}$$

from the collection of all subsets of  $X$  to the collection of all subsets of  $Y$ . Now the topologies of  $X$  and  $Y$  are subcollections of  $\mathcal{P}(X)$  and  $\mathcal{P}(Y)$ , respectively. Then to say that the bijection  $f$  is a homeomorphism is to say that  $f^*$  maps the topology of  $X$  onto the topology of  $Y$ . Thus: **a homeomorphism  $f: X \cong Y$  is a one-to-one correspondence between the points of  $X$  and the points of  $Y$  that establishes a one-to-one correspondence between the open sets in  $X$  and the open sets in  $Y$ .** For this reason, a topologist regards two homeomorphic spaces as being essentially the same. (More will be said about this shortly.)

In view of the criterion in [Theorem 3.3](#) for a map to be continuous and the meanings from [Definition 3.15](#) of open maps and closed maps, evidently **a homeomorphism  $f: X \cong Y$  is just a continuous open bijection**, or equivalently, a continuous closed bijection. Now a bijection  $f: X \rightarrow Y$  is an open map precisely when its inverse  $f^{-1}: Y \rightarrow X$  is continuous, because

$$(f^{-1})^{-1} = f(U)$$

for each open subset  $U$  of  $X$ . Consequently, **a bijection  $f: X \rightarrow Y$  is a homeomorphism if and only if both  $f$  and its inverse  $f^{-1}$  are continuous.**

This last characterization, coupled with [Theorem 3.8](#) and [Examples 3.9 \(5\)](#), establishes the following theorem.

thm:homeo-basics-properties

**3.23 Theorem.** (1) The identity map  $\iota_X: X \rightarrow X$  of any topological space  $X$  is a homeomorphism

$$\iota_X: X \cong X.$$

the-part:inverse-of-homeo

(2) If

$$f: X \cong Y$$

is a homeomorphism, then its inverse

$$f^{-1}: Y \cong X$$

is also a homeomorphism.

thm-part:composite-of-homeos

(3) If

$$f: X \cong Y, \quad g: Y \cong Z$$

are homeomorphisms, then their composite

$$g \circ f: X \cong Z$$

is also a homeomorphism.

According to this theorem the relation ‘is homeomorphic to’ is an equivalence relation on the class of all topological spaces. That is, for all topological spaces  $X$ ,  $Y$ , and  $Z$ :

$$\begin{aligned} X &\cong X, \\ X &\cong Y \implies Y \cong X, \\ X &\cong Y \quad \text{and} \quad Y \cong Z \implies X \cong Z. \end{aligned}$$

Thus homeomorphism is a means of classifying topological spaces as being alike or unlike one another, with two spaces being classified as alike if and only if they are homeomorphic to one another.

In some of the following examples, and later, we shall at times use the following criterion for a continuous map to be a homeomorphism.

prop:homeo-via-f-and-g

**3.24 Proposition.** A continuous map  $f: X \rightarrow Y$  between topological spaces is a homeomorphism if there is some continuous map  $g: Y \rightarrow X$  such that  $g \circ f$  and  $f \circ g$  are the identity maps of their respective domains  $X$  and  $Y$ .

ex:isometric-implies-homeo

**3.25 Examples.** (1) Isometric metric spaces are, as topological spaces (with the induced topologies), homeomorphic. This is just a restatement of [Examples 1.49 \(1\)](#).

(2) Two discrete spaces  $X$  and  $Y$  are homeomorphic to one another if and only if there is some bijection from  $X$  to  $Y$ , in other words,  $X$  has the same cardinality as  $Y$ . Thus every denumerable discrete space is homeomorphic to  $\mathbb{Z}$ , and in particular  $\mathbb{N} \cong \mathbb{Z}$ .

ex:extended-reals-homeo-interval

(3) The bijection

$$\widehat{\varphi}: \widehat{\mathbb{R}} \rightarrow [-1, 1]$$

used in [Example 1.41](#) to construct the metric  $\widehat{d}$  on the extended real line  $\widehat{\mathbb{R}}$  is a homeomorphism, for by construction it is an isometry with respect to the metric  $\widehat{d}$  on

extended real line!  
homeomorphic to  
closed interval!  
homeomorphism@

real line!homeomorphic to  $\widehat{\mathbb{R}}$  and the Euclidean metric on  $[-1, 1]$ . The restriction

open interval!homeomorphism@and homeomorphism

closed interval!homeomorphism@and homeomorphism

open interval!homeomorphism@and homeomorphism

open interval!homeomorphism@and homeomorphism

closed interval!homeomorphism@and homeomorphism

$$\begin{aligned}\varphi: \mathbb{R} &\rightarrow ]-1, 1[ \\ x &\mapsto \frac{x}{1 + |x|}\end{aligned}$$

of that bijection  $\varphi$  is a homeomorphism, since by Examples 2.10 (3) the real line  $\mathbb{R}$  is a subspace of  $\widehat{\mathbb{R}}$ .

ex:any-two-closed-intervals-homeo (4) According to Examples 1.49 (3), the bijection

$$\begin{aligned}f: [0, 1] &\rightarrow [a, b] \\ x &\mapsto a + (b - a)x\end{aligned}$$

is a homeomorphism when  $a < b$ . Thus the closed unit interval

$$I = [0, 1]$$

is a topological “representative” of all closed intervals  $[a, b]$  in the real line that contain more than one point. Moreover, since  $I \cong [-1, 1]$  and, as observed in (3),  $\widehat{\mathbb{R}} \cong [-1, 1]$ , from Theorem 3.23 we conclude

$$\widehat{\mathbb{R}} \cong I.$$

ex:any-two-open-intervals-homeo (5) In Example (4) we have  $f(0) = a$  and  $f(1) = b$ , so by restricting both the domain and the codomain of  $f$  we obtain a homeomorphism

$$\mathbb{R} \cong ]a, b[$$

whenever  $a < b$ . In particular,  $]0, 1[ \cong ]-1, 1[$ , and so from Example (3) we obtain

$$\mathbb{R} \cong ]0, 1[.$$

For each  $a \in \mathbb{R}$  the translation  $x \mapsto x - a + 1$  gives a homeomorphism

$$]a, +\infty[ \cong ]1, +\infty[,$$

for each  $a \geq 0$  the reflection  $x \mapsto -x$  gives a homeomorphism

$$]a, +\infty[ \cong ]-\infty, -a[,$$

and the inversion  $x \mapsto 1/x$  gives a homeomorphism

$$]1, +\infty[ \cong ]0, 1[.$$

Hence **all nonempty open intervals in  $\mathbb{R}$ —including all open rays and  $\mathbb{R}$  itself—are homeomorphic to one another.**

ex:closed-open-intervals-not-homeo (6) **A closed interval in  $\mathbb{R}$  is not homeomorphic to an open interval.** To prove this it suffices to show that the closed interval  $[0, 1]$  is not homeomorphic to the open interval  $]0, 1[$ . In fact, just suppose to the contrary that there exists some homeomorphism  $h: [0, 1] \cong ]0, 1[$ . Now the inversion  $f: ]0, 1[ \rightarrow \mathbb{R}$  given by  $f(x) = 1/x$  is continuous, so the composite  $f \circ h: [0, 1] \rightarrow \mathbb{R}$  is continuous, too. Then  $f \circ h$  is bounded (once again you are asked to assume that continuous real-valued functions on  $[0, 1]$  are bounded—see Corollary 4.2); in other words, there exists a constant  $c$  such that

$$f(h(t)) \leq c \quad (t \in [0, 1]).$$

Since  $h$  maps  $[0, 1]$  onto  $]0, 1[$ , then

$$f(x) \leq c \quad (x \in ]0, 1[),$$

which is patently absurd.

ex:line-in-Rn-homeo-R

- (7) Given two distinct points
- $x, y \in \mathbb{R}^n$
- , the line

$$L = \{(1-t)x + ty : t \in \mathbb{R}\}$$

in  $\mathbb{R}^n$  passing through  $x$  and  $y$  is homeomorphic to the real line  $\mathbb{R}$ . A suitable homeomorphism is

$$\begin{aligned} h: \mathbb{R} &\rightarrow L \\ t &\mapsto (1-t)x + ty \end{aligned}$$

which satisfies the condition

$$\|h(s) - h(t)\| = |s - t| \|x - y\|.$$

simple closed curve  
unit square

ex:square-homeo-circle

- (8) We are going to show that **the unit circle is homeomorphic to the unit sphere**. Recall that *unit circle* is just another name for the 1-sphere  $S_1$ . By the **unit square** we mean the boundary in  $\mathbb{R}^2$  of its subset  $J \times J$ . Alas, the notation  $S_1$  for the unit circle uses the letter ‘S’ and the word ‘square’ begins with ‘s’; to avoid confusion we temporarily denote the unit circle by  $\bigcirc$  and the unit square by  $\square$ . Thus

$$\begin{aligned} \bigcirc &= \{ \langle x, y \rangle \in \mathbb{R}^2 : x^2 + y^2 = 1 \}, \\ \square &= (\{-1, 1\} \times [-1, 1]) \cup ([-1, 1] \times \{-1, 1\}) \\ &= \{ \langle u, v \rangle \in \mathbb{R}^2 : \max\{|u|, |v|\} = 1 \}. \end{aligned}$$

Notice that

$$\bigcirc = \{ \langle x, y \rangle \in \mathbb{R}^2 : \|\langle x, y \rangle\| = 1 \}$$

and

$$\square = \{ \langle u, v \rangle \in \mathbb{R}^2 : \|\langle u, v \rangle\|_\infty = 1 \},$$

where  $\|\cdot\|$  and  $\|\cdot\|_\infty$  are the Euclidean norm and the max norm, respectively (the latter was defined in [Exercise 1.5](#)).

To prove that  $\bigcirc \cong \square$ , observe first that the circle lies inside the square (and is tangent to it at the centers of each side of the square). Construct maps  $f: \bigcirc \rightarrow \square$  and  $g: \square \rightarrow \bigcirc$  geometrically as follows. Given an arbitrary point  $\langle x, y \rangle \in \bigcirc$ , draw the line segment from the origin to  $\langle x, y \rangle$ , extend it until it intersects the square, and take  $f(\langle x, y \rangle)$  to be the point of intersection; given an arbitrary point  $\langle u, v \rangle \in \square$ , draw the line segment from the origin to  $\langle u, v \rangle$  and take  $g(\langle u, v \rangle)$  to be the point where it intersects the circle. On geometric grounds, it is clear that the maps  $f$  and  $g$  are inverses of each other. To verify this algebraically, use the formulas

$$\begin{aligned} f(\langle x, y \rangle) &= \frac{1}{\|\langle x, y \rangle\|} \langle x, y \rangle, \\ g(\langle u, v \rangle) &= \frac{1}{\max\{|u|, |v|\}} \langle u, v \rangle. \end{aligned}$$

Since the maps  $f$  and  $g$  are continuous and satisfy the criterion of [Proposition 3.24](#), it follows that  $f: \bigcirc \cong \square$ .

Example (8) suggests the following definition.

def:simple-closed-curve

**3.26 Definition.** A **simple closed curve** is a topological space that is homeomorphic to the unit circle.

In terms of this definition, then, the unit square is a simple closed curve.

ex:all-balls-in-Rn-homeo-Bn (9) The homeomorphism from  $\mathbb{B}$  to  $[0, 1]$  of Example (3) has an  $n$ -dimensional analog

ball!homeomorphism@and homeomorphism

$$f: \mathbb{R}^n \cong \mathbb{B}_n$$

ball!homeomorphism@and homeomorphism

Euclidean  $n$ -space@Euclidean  $n$ -space!homeomorphism@and homeomorphism

$n$ -cube@ $n$ -cube!homeomorphism@and homeomorphism

$$\mathbb{B}_n = \{x \in \mathbb{R}^n : \|x\| < 1\},$$

centered  $n$ -cube@centered  $n$ -cube!homeomorphism@and homeomorphism

which is just the Euclidean ball of radius 1 at the origin  $\mathbf{0}$ . To construct  $f$ , note first that for any  $x \in \mathbb{R}^n$ ,

$$\left\| \frac{1}{1 + \|x\|} x \right\| = \frac{1}{1 + \|x\|} \|x\| < 1.$$

Then we have a map

$$\begin{aligned} f: \mathbb{R}^n &\rightarrow \mathbb{B}_n \\ x &\mapsto \frac{1}{1 + \|x\|} x. \end{aligned}$$

Now for  $x \in \mathbb{R}^n$  and  $y \in \mathbb{B}_n$ ,

$$\begin{aligned} y = f(x) &\iff x = (1 + \|x\|) y \quad \text{and} \quad \|x\| = (1 + \|x\|) \|y\| \\ &\iff x = \frac{1}{1 - \|y\|} y, \end{aligned}$$

so that  $f$  is a bijection whose inverse is the map

$$\begin{aligned} g: \mathbb{B}_n &\rightarrow \mathbb{R}^n \\ y &\mapsto \frac{1}{1 - \|y\|} y. \end{aligned}$$

Both  $f$  and  $g$  are continuous, and so

$$f: \mathbb{R}^n \cong \mathbb{B}_n.$$

ex:all-balls-in-Rn-homeo-Bn (10) Let  $d$  be the Euclidean metric on  $\mathbb{R}^n$ . Then **every  $d$ -ball in  $\mathbb{R}^n$  is homeomorphic to the unit  $n$ -ball  $\mathbb{B}_n$ —and hence to  $\mathbb{R}^n$  itself.** In fact, for any point  $y \in \mathbb{R}^n$  and any radius  $r > 0$ , each of the following maps is a homeomorphism: the translation

$$\begin{aligned} B_r(y; d) &\cong B_r(\mathbf{0}; d) \\ x &\mapsto x - y \end{aligned},$$

the “similarity”

$$\begin{aligned} B_r(\mathbf{0}; d) &\cong B_1(\mathbf{0}; d) = \mathbb{B}_n \\ x &\mapsto \frac{1}{r} x \end{aligned},$$

and the map

$$g: \mathbb{B}_n \cong \mathbb{R}^n$$

of the preceding Example (9).

ex:n-cube-homeo-centered-cube (11) Recall the notations

$$I = [0, 1] \quad J = [-1, 1]$$

for the unit interval and the centered interval, respectively. The map

$$\begin{aligned} k: I &\rightarrow J \\ x &\mapsto 2x - 1 \end{aligned}$$

is a bijection—in fact, a homeomorphism—between them.

Form the subspaces

$$I^n = \{x \in \mathbb{R}^n : 0 \leq x_i \leq 1 \text{ for each } i = 1, 2, \dots, n\}$$

and

$$J^n = \{x \in \mathbb{R}^n : -1 \leq x_i \leq 1 \text{ for each } i = 1, 2, \dots, n\}$$

of  $\mathbb{R}^n$ . We call  $I^n$  the ***n*-cube** and  $J^n$  the **centered *n*-cube** (the terminology “centered” is not standard).

The bijection  $k: I \rightarrow J$  induces a bijection

$$h: I^n \rightarrow J^n \\ \langle x_1, x_2, \dots, x_n \rangle \mapsto \langle h(x_1), h(x_2), \dots, h(x_n) \rangle$$

from the *n*-cube to the centered *n*-cube. This bijection is continuous because  $x, u \in I^n$  implies

$$\begin{aligned} d_\infty(h(u), h(x)) &= \max_{1 \leq i \leq n} |k(u_i) - k(x_i)| \\ &= 2 \max_{1 \leq i \leq n} |u_i - x_i| \\ &= 2 d_\infty(u, x), \end{aligned}$$

where  $d_\infty$  is the max metric. Similarly, the inverse

$$h^{-1}: J^n \rightarrow I^n \\ \langle y_1, y_2, \dots, y_n \rangle \mapsto \langle k^{-1}(y_1), k^{-1}(y_2), \dots, k^{-1}(y_n) \rangle$$

of  $h$  is continuous. Hence we have obtained a homeomorphism

$$h: I^n \cong J^n.$$

ex:n-cube-homeo-n-disk (12) We are going to show that ***the n-cube  $I^n$  is homeomorphic to the n-disk***

$$D_n = \{x \in \mathbb{R}^n : \|x\| \leq 1\}.$$

After Example (11), we need only show that  $J^n \cong D_n$ ; working with  $J^n$  instead of  $I^n$  has the advantage that both  $I^n$  and  $D_n$  are centered around the origin  $\mathbf{0}$  of  $\mathbb{R}^n$ . In terms of the Euclidean metric  $d$  and the max metric  $d_\infty$  on  $\mathbb{R}^n$ ,

$$D_n = \{x \in \mathbb{R}^n : d(x, \mathbf{0}) \leq 1\}, \quad J^n = \{x \in \mathbb{R}^n : d_\infty(x, \mathbf{0}) \leq 1\}.$$

Then  $D_n \subset J^n$  because  $d_\infty \leq d$ . Let

$$E = \{x \in \mathbb{R}^n : d_\infty(x, \mathbf{0}) = 1\},$$

which, it so happens, is the boundary of  $J^n$  in  $\mathbb{R}^n$  (see Exercise 67).

Suppose  $x = \langle x_1, x_2, \dots, x_n \rangle \in J^n \setminus \{\mathbf{0}\}$ . Then the ray  $\{\lambda x : \lambda \geq 0\}$  from the origin  $\mathbf{0}$  through  $x$  intersects  $E$  in exactly one point  $e(x)$ , as shown in Figure 3.3. In fact,

centered n-cube@centered  $J^n$ -cube  
n-cube@ $J^n$ -cube!homeomorphism  
n-disk@ $J^n$ -disk!homeomorphism@

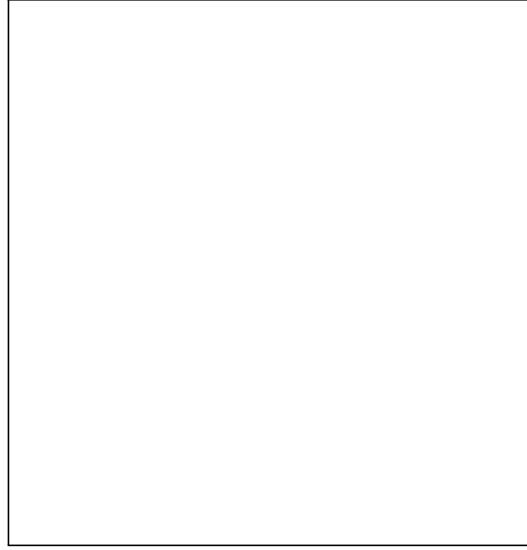


Figure 3.3: Constructing a homeomorphism from the centered  $n$ -cube to the  $n$ -disk.

fig:construct-homeo-Dn-Jn

if  $d_\infty(x, \mathbf{0}) = |x_j|$ , then  $d_\infty(\lambda x, \mathbf{0}) = \lambda |x_j| = \lambda d_\infty(x, \mathbf{0})$  for  $\lambda \geq 0$ , so that  $\lambda x \in E$  if and only if  $\lambda = 1/d_\infty(x, \mathbf{0})$ . Hence the point  $e(x)$  of intersection is given by

$$e(x) = \frac{1}{d_\infty(x, \mathbf{0})} x.$$

Since  $d_\infty$  induces the usual topology of  $\mathbb{R}^n$ , the map  $x \mapsto d_\infty(x, \mathbf{0})$  is continuous. Hence the map

$$e: J^n \setminus \{\mathbf{0}\} \rightarrow E$$

is continuous. For  $x \in J^n \setminus \{\mathbf{0}\}$  we have  $d_\infty(x, \mathbf{0}) \leq 1$  and  $\|x\| = d(x, \mathbf{0}) \geq d_\infty(x, \mathbf{0})$ , so that

$$\|e(x)\| \geq \|x\|, \quad \|e(x)\| \geq 1.$$

We propose to construct a map  $f: J^n \rightarrow D_n$  that linearly contracts each line segment joining the origin  $\mathbf{0}$  to a point of  $E$  onto its segment joining  $\mathbf{0}$  to  $S_{n-1}$ . Here  $S_{n-1}$  is the  $(n-1)$ -sphere

$$S_{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\},$$

the boundary of the  $n$ -disk  $D_n$  in  $\mathbb{R}^n$  [see [Examples 2.28 \(5\)](#)]. The ray from  $\mathbf{0}$  through  $x \in J^n \setminus \{\mathbf{0}\}$  intersects the set  $E$  at the point  $e(x)$  and the sphere  $S_{n-1}$  at the point  $(1/\|x\|)x$ ; if the value of the map  $f$  to be constructed is  $f(x) = \lambda x$  at  $x$ , then we want to make the ratio

$$\frac{d(x, \mathbf{0})}{d(e(x), \mathbf{0})} = \frac{\|x\|}{\|e(x)\|}$$

of distances to be the same as the ratio

$$\frac{d(f(x), \mathbf{0})}{d((1/\|x\|)x, \mathbf{0})} = \frac{\|f(x)\|}{1} = \lambda \|x\|,$$

in other words, we want to take  $\lambda = 1/\|e(x)\|$ . Hence we define

$$f: J^n \rightarrow D_n$$



by

$$f(x) = \begin{cases} \frac{1}{\|e(x)\|} x & \text{if } x \in J^n \setminus \{\mathbf{0}\}, \\ \mathbf{0} & \text{if } x = \mathbf{0}. \end{cases}$$

Continuity of  $f$  at each point of  $J^n \setminus \{\mathbf{0}\}$  follows from continuity of  $e$  there, and continuity of  $f$  at  $x = \mathbf{0}$  follows from the inequality

$$\|f(u) - f(\mathbf{0})\| = \frac{\|u\|}{\|e(u)\|} \leq \|u\| = \|u - \mathbf{0}\| \quad (u \in J^n \setminus \{\mathbf{0}\}).$$

Whenever  $x \in J^n \setminus \{\mathbf{0}\}$  and  $y \in D_n \setminus \{\mathbf{0}\}$  are on the same ray through the origin, we have  $e(x) = e(y)$ , and so

$$y = f(x) \iff x = \|e(x)\| y = \|e(y)\| y.$$

Hence  $f$  is a bijection having the inverse  $g: D_n \rightarrow J^n$  given by

$$g(y) = \begin{cases} \|e(y)\| y & \text{if } y \in D_n \setminus \{\mathbf{0}\}, \\ \mathbf{0} & \text{if } y = \mathbf{0}. \end{cases}$$

The map  $g$  is continuous, too, and we conclude that

$$f: J^n \cong D_n.$$

A topological space that is homeomorphic to the  $n$ -disk  $D_n$  is said to be an  **$n$ -cell** or, for emphasis, a **closed  $n$ -cell**. Thus this example has shown that *the  $n$ -cube  $I^n$  is an  $n$ -cell*; thus the preceding example showed that *the centered  $n$ -cube  $J^n$  is an  $n$ -cell*.

punctured- $n$ -sphere-homeo- $n$ -space (13) Consider the **punctured  $n$ -sphere** obtained by removing from the  $n$ -sphere

$$S_n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$$

its **north pole**

$$\mathbf{p} = \langle 0, 0, \dots, 0, \dots, 0, 1 \rangle.$$

We are going to show that **the punctured  $n$ -sphere is homeomorphic to Euclidean  $n$ -space**:

$$S_n \setminus \{\mathbf{p}\} \cong \mathbb{R}^n$$

To begin, observe that the “equatorial hyperplane”

$$H = \mathbb{R}^n \times \{\mathbf{0}\} = \{x \in \mathbb{R}^{n+1} : x_{n+1} = 0\}$$

in  $\mathbb{R}^{n+1}$  is homeomorphic to  $\mathbb{R}^n$  under the map

$$g: \begin{matrix} H & \rightarrow & \mathbb{R}^n \\ \langle x_1, x_2, \dots, x_n, 0 \rangle & \mapsto & \langle x_1, x_2, \dots, x_n \rangle \end{matrix}$$

Hence we need only show that  $S_n \setminus \{\mathbf{p}\} \cong H$ .

To do that, let  $x \in S_n \setminus \{\mathbf{0}\}$ . Then the line through  $x$  and  $\mathbf{p}$  intersects  $H$  in a unique point  $y$  (see 3.4). In fact, for  $y = \langle y_1, y_2, \dots, y_n, y_{n+1} \rangle \in \mathbb{R}^{n+1}$  and  $\lambda \in \mathbb{R}$ , the

$n$ -cell@ $n$ -cell

$n$ -sphere@ $n$ -sphere!homeomorph

Euclidean  $n$ -space@Euclidean  $n$ -space

punctured sphere

north pole



Figure 3.4: Constructing a homeomorphism from the punctured  $n$ -sphere to the equatorial hyperplane in  $\mathbb{R}^{n+1}$ .

fig:stereographic-proj

conditions

$$\{\text{eq:cond-for-stereographic-proj}\} \quad (*) \quad y_{n+1} = 0, \quad y = \lambda x + (1 - \lambda) \mathbf{p}$$

yield

$$0 = y_{n+1} = \lambda x_{n+1} + (1 - \lambda) \mathbf{p}_{n+1} = \lambda x_{n+1} + (1 - \lambda),$$

that is,

$$\lambda = \frac{1}{1 - x_{n+1}}$$

(since  $x \in S_n$  but  $x \neq \mathbf{p}$ , then  $x_{n+1} \neq 1$ ). Thus equations  $(*)$  have the unique solution

$$y = \frac{1}{1 - x_{n+1}} x + \frac{-x_{n+1}}{1 - x_{n+1}} \mathbf{p}.$$

Call this solution  $f(x)$ . We have thus constructed a continuous map

$$f: S_n \setminus \{\mathbf{p}\} \rightarrow H$$

with

$$\begin{aligned} f(x) &= \frac{1}{1 - x_{n+1}} (x - x_{n+1} \mathbf{p}) \\ &= \left\langle \frac{x_1}{1 - x_{n+1}}, \frac{x_2}{1 - x_{n+1}}, \dots, \frac{x_n}{1 - x_{n+1}}, 0 \right\rangle \end{aligned}$$

for each  $x$ .

To see that  $f$  is bijective and that its inverse is continuous, too, take an arbitrary  $y = \langle y_1, y_2, \dots, y_n, 0 \rangle \in H$ . We want to solve the equations

$$\{\text{eq:to-find-inverse-of-stereographic-proj}\} \quad (**) \quad \begin{cases} y = f(x), \\ \|x\| = 1 \end{cases}$$

for  $x = \langle x_1, x_2, \dots, x_n, x_{n+1} \rangle$  with  $x_{n+1} \neq 1$ . In terms of coordinates equations  $(**)$  say

$$\begin{cases} y_i = \frac{x_i}{1 - x_{n+1}} & (i = 1, 2, \dots, n), \\ \sum_{i=1}^n x_i^2 + x_{n+1}^2 = 1. \end{cases}$$

These last equations have unique solution

$$\begin{cases} x_i = \frac{2}{\|y\|^2 + 1} y_i & (i = 1, 2, \dots, n), \\ x_{n+1} = \frac{\|y\|^2 - 1}{\|y\|^2 + 1}. \end{cases}$$

Thus  $f$  has an inverse given by

$$f^{-1}(y) = \frac{2}{\|y\|^2 + 1} y + \frac{\|y\|^2 - 1}{\|y\|^2 + 1} \mathbf{p} \quad (y \in H).$$

The homeomorphism

$$\pi^+ = g \circ f: S_n \setminus \{\mathbf{p}\} \cong \mathbb{R}^n$$

is known as the **stereographic projection**. This map is of particular interest for dimension  $n = 2$ , because it allows us to represent the complex plane  $\mathbb{C}$ , which as a topological space is just  $\mathbb{R}^2$ , as the complement of a point in the 2-sphere  $S_2$  (see [Example 4.81](#) for a further discussion).

stereographic projection  
antipodal map  
reflection  
locally Euclidean space  
manifold  
n-manifold@n-manifold  
torus

ex:n-sphere-is-n-manifold (14) Let  $-\mathbf{p}$  be the “south pole”  $\langle 0, 0, \dots, 0, -1 \rangle$  of the  $n$ -sphere  $S_n$ . Let

$$\begin{aligned} \alpha: S_n \setminus \{-\mathbf{p}\} &\rightarrow S_n \setminus \{\mathbf{p}\} \\ x &\mapsto -x \end{aligned}$$

be the **antipodal map**, which is a homeomorphism. The antipodal map  $\alpha$  is a domain-codomain restriction of the reflection

$$\begin{aligned} \rho: \mathbb{R}^{n+1} &\rightarrow \mathbb{R}^{n+1} \\ x &\mapsto -x \end{aligned}$$

of  $\mathbb{R}^{n+1}$  through the origin, and  $\rho$  is also a homeomorphism. Notice that  $\rho^{-1} = \rho$ . Finally, form the composite

$$\pi^- = (\rho^{-1} \circ \pi^+ \circ \alpha): S_n \setminus \{-\mathbf{p}\} \rightarrow \mathbb{R}^n,$$

where  $\pi^+$  is the stereographic projection. More explicitly,

$$\pi^-(x) = -\pi^+(-x) \quad (x \in S_n \setminus \{-\mathbf{p}\}).$$

(We could have omitted  $\rho^{-1}$  from the composite, thereby resulting in the seemingly simpler formula  $\pi^-(x) = \pi^+(-x)$ ; however, we do include the extra  $-$  so as simplify certain calculations later.) Thus we have a homeomorphism

$$\pi^-: S_n \setminus \{-\mathbf{p}\} \cong \mathbb{R}^n.$$

The two subsets  $S_n \setminus \{\mathbf{p}\}$  and  $S_n \setminus \{-\mathbf{p}\}$  are open in  $S_n$ , each point of  $S_n$  belongs to at least one of them, and each of them is homeomorphic to  $\mathbb{R}^n$ . Hence, in the language of [Definitions 3.31](#) and [3.40](#), the  $n$ -sphere is *locally Euclidean of dimension  $n$*  and, in fact, is an  *$n$ -dimensional manifold*.

ex:torus-homeo-surface-in-R3 (15) The **torus**, also known as the **2-torus**, is the product  $S_1 \times S_1$  of the unit circle  $S_1$  with itself, where the topology is the product topology in the sense of [Examples 2.72 \(5\)](#).

total surface

Since each  $S_1$  factor is a subspace of  $\mathbb{R}^2$ , then the product  $S_1 \times S_1$  is essentially a subspace of  $\mathbb{R}^4$ ; more precisely, the map

$$\begin{aligned} S_1 \times S_1 &\rightarrow \mathbb{R}^4 \\ \langle \langle x, y \rangle, \langle u, v \rangle \rangle &\mapsto \langle x, y, u, v \rangle \end{aligned}$$

is a continuous open injection and hence defines a natural homeomorphism of  $S_1 \times S_1$  with the subspace

$$\{ \langle x, y, u, v \rangle : x^2 + y^2 = 1, u^2 + v^2 = 1 \}$$

of  $\mathbb{R}^4$ . In the terminology to be introduced in [Definition 3.46](#), this map is an *embedding* of the torus  $S_1 \times S_1$  into Euclidean 4-space  $\mathbb{R}^4$ .

Nonetheless, it does not take 4 dimensions to “accommodate” the torus: we are going to construct a homeomorphism of  $S_1 \times S_1$  with a subspace of Euclidean 3-space  $\mathbb{R}^3$ . And that subspace will, in fact, be a *surface*—a 2-dimensional object.

To motivate the construction, we note that, as a set, the torus has the representation

$$S_1 \times S_1 = \bigcup_{p \in S_1} \{p\} \times S_1.$$

Now for each  $p \in S_1$ , the product  $\{p\} \times S_1$  is homeomorphic to  $S_1$  under the map  $q \mapsto \langle p, q \rangle$ . This means that the torus, at least as a set, may be regarded as comprised of many circles  $K_p$ , one associated with each point  $p$  of the first circle  $S_1$ .

Accordingly, take a circle  $C$  in the  $xy$ -plane of radius  $R > 0$  centered at the origin in  $\mathbb{R}^3$ ; this “large circle” has equation

$$x^2 + y^2 = R^2 \quad (\text{the large circle } C).$$

Next, take a circle  $K$  of radius  $r < R$  centered at the point  $\langle R, 0, 0 \rangle$  on  $C$  and lying in the  $xz$ -plane; this “small circle” has equation

$$(x - R)^2 + z^2 = r^2 \quad (\text{the small circle } K).$$

Both circles are shown in [Figure 3.5\(a\)](#). Now rotate  $K$  around in  $\mathbb{R}^3$  while keeping its center on  $C$  and keeping its plane passing through the  $z$ -axis, as indicated in [Figure 3.5\(b\)](#). The resulting surface  $T = \bigcup_{p \in S_1} K_p$ —a subspace of  $\mathbb{R}^3$ —is the *total surface*; it has equation

$$\{eq:torus-T\} \quad (*) \quad \left( \sqrt{x^2 + y^2} - R \right)^2 + z^2 = r^2.$$

The entire surface  $T$  is shown in [Figure 3.6](#).

We are going to show that the torus  $S_1 \times S_1$  is homeomorphic to the total surface  $T$ .

To do that, first notice that the two copies of the unit circle  $S_1$  in the product  $S_1 \times S_1$  are homeomorphic to the large circle  $C$  and the small circle  $K$ , respectively. In fact, the map

$$\begin{aligned} f_1: S_1 &\rightarrow C, \\ \langle s, t \rangle &\mapsto \langle Rs, Rt, 0 \rangle \end{aligned}$$

which expands the circle  $S_1$  from radius 1 to  $R$  and inserts the result into the  $xy$ -plane in  $\mathbb{R}^3$  with its center at the origin is a homeomorphism; and the map

$$\begin{aligned} f_2: S_1 &\rightarrow K, \\ \langle u, v \rangle &\mapsto \langle R + ru, 0, rv \rangle \end{aligned}$$

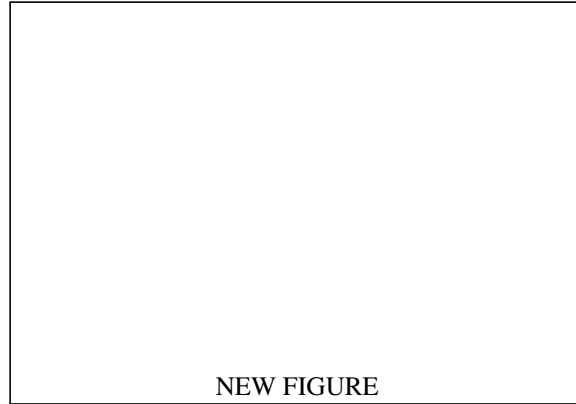
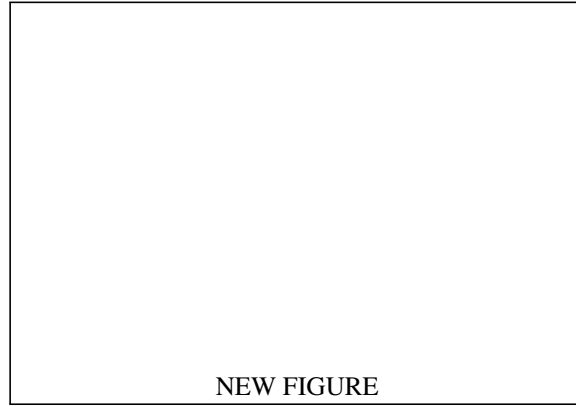
(a) Circles  $C$  and  $K$  in  $\mathbb{R}^3$ .subfig:circle-centered-on-circle(b) Rotate circle  $C$  around circle  $K$ .subfig:rotate-circle-to-get-torus

Figure 3.5: Generating a surface  $T$  in  $\mathbb{R}^3$  that is homeomorphic to the torus  $S_1 \times S_1$ .

fig:generate-torus-rotate-circle-arou

which contracts the circle  $S_1$  from radius 1 to radius  $r$ , inserts the result into the  $xz$ -plane in  $\mathbb{R}^3$  with its center at the origin, and then translates that inserted circle to have new origin  $\langle R, 0, 0 \rangle$  is also a homeomorphism.

We are going to construct a homeomorphism  $h: S_1 \times S_1 \cong T$ . For that we shall use the two-argument version

$$\text{atan2}: \mathbb{R}^2 \setminus \langle 0, 0 \rangle \rightarrow \mathbb{R}$$

of the inverse tangent function, given by

$$\text{atan2}(s, t) = \begin{cases} \arctan(t/s) & \text{if } s > 0, \\ \arctan(t/s) + \pi & \text{if } s < 0 \text{ and } t \geq 0, \\ \arctan(t/s) - \pi & \text{if } s < 0 \text{ and } t < 0, \\ \pi/2 & \text{if } s = 0 \text{ and } t > 0, \\ -\pi/2 & \text{if } s = 0 \text{ and } t < 0, \end{cases}$$

where  $\arctan$  is the usual single-variable inverse tangent having range  $(-\pi/2, \pi/2)$ .

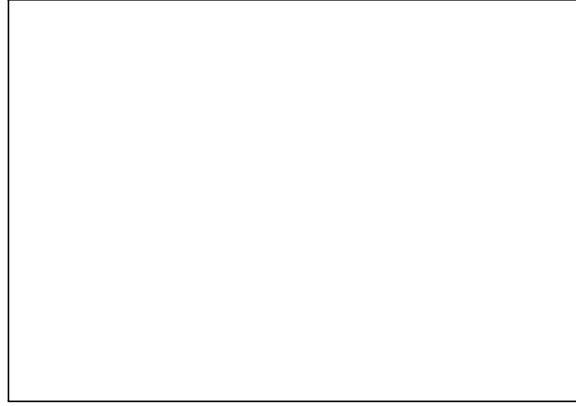


Figure 3.6: The surface  $T$  in  $\mathbb{R}^3$  obtained by rotating a circle.

fig:torus-from-rotated-circle

Transform the unit circle  $S_1$  in  $\mathbb{R}^2$  into the circle  $C$  in  $\mathbb{R}^3$  by means of the map

$$\begin{aligned} h_1: S_1 &\rightarrow C \\ \langle u, v \rangle &\mapsto \langle ru + R, 0, rv \rangle \end{aligned}$$

is a homeomorphism, too.

Define  $h: S_1 \times S_1 \rightarrow T$  as follows. Take an arbitrary point

$$\langle p, q \rangle = \langle \langle s, t \rangle, \langle u, v \rangle \rangle \in S_1 \times S_1.$$

Map the circle  $S_1$  in  $\mathbb{R}^2$  to the circle  $C_1$  in the  $xz$ -plane having radius 1 and centered at the origin, then expand  $C_1$  by a factor of  $r$ . Let  $g(q)$  be the point on the large circle  $C$  in  $\mathbb{R}^3$  obtained by

Set

$$g(q) = \langle R + ru, 0, rv \rangle,$$

the point on the circle of radius  $r$  in the  $xz$ -plane having center on the  $x$ -axis at the point  $\langle R, 0, 0 \rangle$ ; set

$$\theta(p) = \text{atan2}(s, t).$$

Finally, define  $h(p, q)$  to be the point in  $\mathbb{R}^3$  obtained by rotating the point  $g(q)$  through angle  $\theta(p)$  around the  $z$ -axis, so that (see [Exercise 1.56](#))

$$\begin{aligned} h(p, q) &= \begin{pmatrix} \cos \theta(p) & -\sin \theta(p) & 0 \\ \sin \theta(p) & \cos \theta(p) & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot g(q) \\ &= \langle x \cos \theta(p) - y \sin \theta(p), y \cos \theta(p) + x \sin \theta(p), z \rangle \\ &= \langle (R + ru) \cos \theta(p), (R + ru) \sin \theta(p), rv \rangle \\ &= \langle (R + ru) \cos \text{atan2}(s, t), (R + ru) \sin \text{atan2}(s, t), rv \rangle. \end{aligned}$$

To see that  $h(p, q) \in T$ , verify that its coordinates satisfy [equation \(\\*\)](#).

The function  $\text{atan2}$  is discontinuous at all points of the horizontal negative semi-axis, where its values jump by  $2\pi$  as you cross that semi-axis; nevertheless, the composites  $\cos \circ \text{atan2}$  and  $\sin \circ \text{atan2}$  are continuous at all points (other than the origin, which is not in their domains; in particular, these composites are continuous on the unit circle). Hence the map  $h: S_1 \times S_1 \rightarrow T$  is continuous.

To see that  $h$  is actually a homeomorphism, reverse the process that gives  $h(p, q)$  from  $\langle p, q \rangle$ . Namely,

The map  $g : S_1 \rightarrow K$  is bijective, with inverse  $g^{-1}$

MORE  $\diamond$

### Topological properties

subsec:top-props

We want to elaborate now on the earlier comment that two topological spaces that are homeomorphic should be regarded from a topological point of view as being essentially the same. Suppose spaces  $X$  and  $Y$  are homeomorphic, so that there is a one-to-one correspondence between the points of  $X$  and the points of  $Y$  inducing a one-to-one correspondence between the open sets in  $X$  and the open sets in  $Y$ . Then regardless of the particular natures of their points,  $X$  and  $Y$  are indistinguishable purely as topological spaces. Hence anything we can correctly say about  $X$  *in so far as it is a topological space*, we can correctly say about  $Y$ , too, and vice versa.

Succinctly put, we are claiming that **homeomorphic topological spaces have the same topological properties**. The preceding was only a heuristic justification for this claim—and motivation for a precise definition of ‘topological property’. Once we give the definition, our claim will become a tautology.

def:top-prop

**3.27 Definition.** Suppose  $P$  is a predicate that is meaningful to assert about topological spaces. Then  $P$  is said to be a **topological property** if  $P$  holds for each topological space that is homeomorphic to a topological space for which  $P$  holds; in other words, for all topological spaces  $X$  and  $Y$ :

$$P(X) \text{ and } X \cong Y \implies P(Y)$$

We may say, then, that a topological property is a property of topological spaces that is *preserved by*, or *invariant under*, *homeomorphism*. Some examples will clarify the definition.

ex:top-props

**3.28 Examples.** (1) We begin with two “nonexamples.” The property of being bounded with diameter at most 1 cannot qualify as a topological property, for it is meaningless to assert it about a topological space (although it is meaningful about a metric space).

The property of being a subspace of the real line is *not* a topological property, even though it does make sense to say whether a given topological space is a subspace of  $\mathbb{R}$ . However, this property is not preserved under homeomorphism. For example,  $\mathbb{R}$  is itself a subspace of  $\mathbb{R}$ , the space  $x$ -axis  $Y = \{ \langle x, 0 \rangle : x \in \mathbb{R} \}$  in the plane is homeomorphic to  $\mathbb{R}$ , but  $Y$  itself is certainly not a subspace of  $\mathbb{R}$ —in fact, not even a subset of  $\mathbb{R}$ .

ex:metrizability-top-prop

(2) **Metrizability is a topological property.** In fact, certainly it is meaningful to ask whether any given topological space is metrizable. Now suppose that  $X$  is a metrizable space and  $Y$  is a topological space with  $X \cong Y$ . There is a metric  $d$  inducing the topology of  $X$  and a homeomorphism  $h : Y \cong X$ . The formula

$$d'(y, z) = d(h(y), h(z)) \quad (y, z \in Y)$$

defines a metric on  $Y$ ; symmetry of  $d'$ , for example, follows from symmetry of  $d$ :

$$d'(y, z) = d(h(y), h(z)) = d(h(z), h(y)) = d'(z, y).$$

Hausdorff space!topological property@and topological property  
 discrete space!topological property@and topological property  
 second-countable space!topological property@and topological property  
 first-countable space!topological property@and topological property  
 separable space!topological property@and topological property

To complete the argument that  $Y$ , too, is metrizable we show that  $d'$  induces the given topology  $\mathcal{T}$  of  $Y$ . First, let  $U$  be an arbitrary  $\mathcal{T}$ -open subset of  $Y$ . If  $y \in U$ , then: there is an  $\varepsilon > 0$  with  $B_\varepsilon(y; d') \subseteq U$ ; the set  $B_\varepsilon(y; d') = B_\varepsilon(h(y); d)$

is open in  $X$  because  $d$  induces the topology of  $X$ ; its inverse image

$$h^{-1}(B_\varepsilon(h(y); d)) = B_\varepsilon(y; d')$$

is  $\mathcal{T}$ -open in  $Y$  because  $h$  is a homeomorphism; and so  $B_\varepsilon(y; d')$  is a  $\mathcal{T}$ -neighborhood of  $y$  contained in  $U$ . Hence  $U$  is  $\mathcal{T}$ -open.

Conversely, let  $U$  be an arbitrary  $\mathcal{T}$ -open subset of  $Y$ . For each  $y \in U$ : the image  $h(U)$  is an open neighborhood of  $h(y)$  in  $X$ ; there is an  $\varepsilon > 0$  with  $B_\varepsilon(h(y); d) \subset h(U)$ ; and so

$$y \in B_\varepsilon(y; d') = h^{-1}(B_\varepsilon(h(y); d)) \subset U.$$

Hence  $U$  is  $d'$ -open.

ex:hausdorff-top-prop (3) **Being a Hausdorff space is a topological property.** In fact, let  $X$  be a Hausdorff space and let  $h: X \cong Y$  be a homeomorphism from  $X$  to a topological space  $Y$ . Let  $y, z \in Y$  with  $y \neq z$ . Since  $h^{-1}(y) \neq h^{-1}(z)$ , there are disjoint open neighborhoods  $U$  and  $V$  of  $h^{-1}(y)$  and  $h^{-1}(z)$ , respectively, in  $X$ . Then  $h(U)$  and  $h(V)$  are disjoint open neighborhoods of  $y$  and  $z$ , respectively, in  $Y$ .

Similar arguments show that being a  $T_0$ -space, being a  $T_1$ -space, and regularity (that is, being regular) are topological properties. (See [Exercise 95](#).)

ex:discrete-top-prop (4) **Discreteness is a topological property.** In fact, let  $X$  be a discrete space and let  $h: X \cong Y$  be a homeomorphism from  $X$  to a topological space  $Y$ . If  $V$  is any subset of  $Y$ , its inverse image  $h^{-1}(V)$  is open in the discrete space  $X$ , and then  $V = h(h^{-1}(V))$  is open in  $Y$ .

ex:2nd-countable-top-prop (5) **Second-countability is a topological property.** To see this, let  $X$  be a second-countable space and let  $h: X \cong Y$  be a homeomorphism. Choose some countable base  $\mathcal{B}$  of  $X$ . Then the countable collection  $\{h(B) : B \in \mathcal{B}\}$  is a base of  $Y$ . In fact, each member  $h(B)$  of this collection is open in  $Y$  because  $h$  is an open map. Now let  $y \in V \subset Y$  with  $V$  open in  $Y$ . Then  $h^{-1}(y) \in h^{-1}(V)$  with  $h^{-1}(V)$  being open in  $X$ , and so there is some  $B \in \mathcal{B}$  with  $h^{-1}(y) \in B \subset h^{-1}(V)$ . Hence  $y \in h(B) \subset V$ .

Similarly, **first-countability is a topological property**. **Separability**—the third countability property considered in [Chapter 2](#)—**is a topological property** as well (see [Exercise 9](#)).  $\diamond$

The two most significant topological properties—compactness and connectedness—and related properties will be studied in depth in [Chapters 4](#) and [5](#), respectively.

A topological property may also be ascribed to topologies as well as to topological spaces. For example, we may refer to a topology  $\mathcal{T}$  on a set  $X$  as being metrizable, Hausdorff, or second-countable as another way to assert that the topological space  $\langle X, \mathcal{T} \rangle$  is metrizable, Hausdorff, or second-countable, respectively.

If we are given a topological space, then by listing enough of its topological properties we could hope to characterize it “up to homeomorphism” in the sense that any other space having those same properties is homeomorphic to the given one. Except for a few special spaces (see, for example, [Exercise 5.38](#) and [Example 4.16](#)), unfortunately, this hope cannot



be fulfilled in any reasonable way. Often the best we can do is to characterize topologically arbitrary *subspaces* of certain given spaces; see, for example, [Corollary 6.46](#).

Topological properties can, however, be used usefully to distinguish between spaces that are *not* homeomorphic: *if one of two spaces has a particular topological property but the other does not, then the two spaces cannot be homeomorphic*. For example, the upper half-plane  $\{(x, y) \in \mathbb{R} \times \mathbb{R} : y \geq 0\}$ , with its usual topology, is *not* homeomorphic to the half-disk space [\[Examples 2.20 \(3\)\]](#) because the first is metrizable but the second is not. Likewise, the real line is *not* homeomorphic to the line with two origins [\[Examples 2.20 \(3\)\]](#) because the former is a Hausdorff space whereas the latter is not [see [Examples 2.99 \(3\)](#)].

### Local homeomorphism

subsec:local-homeo

We begin here with an example of two spaces that are *not* homeomorphic to one another.

x:punctured-line-not-homeo-to-line **3.29 Example.** The “punctured” real line

$$\mathbb{R}^* = \mathbb{R} \setminus \{0\}$$

is *not* homeomorphic to the real line  $\mathbb{R}$ . In fact, just suppose there were some homeomorphism  $f: \mathbb{R}^* \cong \mathbb{R}$ . The rays  $L = ]-\infty, 0[$  and  $M = ]0, +\infty[$  form a partition of  $\mathbb{R}^*$  into open subsets. Then their images  $A = f(L)$  and  $B = f(M)$  form a partition of  $\mathbb{R}$  into two open subsets.

Evidently  $\mathbb{R}$  cannot be partitioned into two open intervals. Then  $A$  and  $B$  cannot both be open intervals. But this does not rule out the existence of a more complicated partition of  $\mathbb{R}$  into open subsets. According to the following lemma, however, there cannot be any such partition.  $\diamond$

lem:R-has-no-separation **3.30 Lemma.** *There does not exist a partition of the real line  $\mathbb{R}$  into two open subsets.*

**Proof.** Just suppose, to the contrary, that there does exist a partition  $\{A, B\}$  of  $\mathbb{R}$  into two open subsets  $A$  and  $B$ .

Choose some  $a \in A$  and  $b \in B$ . Since  $A$  and  $B$  are disjoint, then  $a \neq b$ . Without loss of generality we may assume that  $a < b$ , for otherwise we simply reverse the roles of  $A$  and  $B$ .

Form the subset

$$E = A \cap [a, b] = \{x \in \mathbb{R} : a \leq x \leq b \text{ and } x \in A\}$$

of  $\mathbb{R}$ . Notice that  $E$  is closed in  $\mathbb{R}$  because  $A$ , being the complement of the open subset  $B$  of  $\mathbb{R}$ , is closed.

The set  $E$  is nonempty because  $a \in E$ , and  $E$  is bounded above in  $\mathbb{R}$ , by  $b$ . From the order-completeness of  $\mathbb{R}$  ([Axiom 0.74](#)), the set  $E$  has a supremum in  $\mathbb{R}$ ; let

$$c = \sup E.$$

We are going to show that  $c \in A \cap B$ , which is impossible because  $A$  and  $B$  are disjoint.

We show that  $c \in E$ . Since  $E$  is closed in  $\mathbb{R}$ , it suffices to show that  $c \in \text{cls } E$ . Let  $]u, v[$  be an arbitrary open interval containing  $c$ . Then  $]u, v[$  contains some element of  $E$ . In fact, since  $c$  is the *least* upper bound of  $E$ , then  $u$  is *not* an upper bound of  $E$ . This means there is some  $t \in E$  with  $u < t$ . But  $t \leq c$  because  $c$  is an upper bound of  $E$ . Hence  $t \in ]u, c] \subset ]u, v[$ .

Since  $c \in E$ , already  $c \in A$ .

connected space  
real line! as connected space

To complete the proof, we show now that  $c \in B$  too. If  $c = b$ , we are done. So suppose now that  $c \neq b$ . Then  $c \in ]a, b[$ . Since  $c \in A$  and  $A$  is an open set, there exist numbers  $x$  and  $y$  with

$$a < x < c < y < b, \quad ]x, y[ \subset A.$$

Choose any  $z \in ]c, y[$ . Then  $z \in [a, b] \cap A = E$ . However,  $z \notin E$  because  $c < z$  and  $c$  is an upper bound of  $E$ .  $\square$

In Chapter 5 we shall define a space to be *connected* when it does *not* have a partition into two open subsets. And we shall see that being connected is a topological property. Thus Example 3.29 and the preceding lemma together say that  $\mathbb{R}^*$  is not homeomorphic to  $\mathbb{R}$  because the real line  $\mathbb{R}$  is connected whereas the punctured line  $\mathbb{R}^*$  is not.

Although the punctured real line  $\mathbb{R}^*$  and the real line  $\mathbb{R}$  are not homeomorphic, they are “locally homeomorphic” in the sense of the following definition.

def:local-homeo

**3.31 Definition.** Let  $X$  and  $Y$  be topological spaces. A map  $f: X \rightarrow Y$  is said to be a **local homeomorphism from  $X$  to  $Y$**  when, for each point  $x \in X$ , there is an open neighborhood  $U$  of  $x$  in  $X$  such that  $f(U)$  is open in  $Y$  and the domain-codomain restriction  $f|_{U, f(U)}: U \rightarrow f(U)$  is a homeomorphism. When some local homeomorphism from  $X$  to  $Y$  exists, we say that  $X$  is **locally homeomorphic to  $Y$** .

Every homeomorphism is of course a local homeomorphism. And a local homeomorphism is necessarily a continuous open map. Hence a *bijective* local homeomorphism is just a homeomorphism.

According to Examples 3.25 (14), for each positive integer  $n$  the  $n$ -sphere  $S_n$  is locally homeomorphic to Euclidean  $n$ -space  $\mathbb{R}^n$ . However,  $S_n$  is *not* homeomorphic to  $\mathbb{R}^n$ . We are not in a position to establish this yet; the reason will be that  $S_n$  [Examples 4.35 (5)] has the topological property of being “compact” in the sense of Definition 4.5, whereas  $\mathbb{R}^n$  does not (see Theorem 4.34). The case  $n = 1$  can be handled now, however: **the 1-sphere  $S_1$  is not homeomorphic to the real line  $\mathbb{R}$** —see Exercise 80.

### Locally Euclidean spaces

subsec:locally-euclidean

Examples 3.25 (14) of the preceding subsection leads to the following definition. There and in the sequel, unless otherwise indicated,  $n$  will denote a nonnegative integer.

**3.32 Definition.** A topological space is said to be **locally Euclidean of dimension  $n$** , and is called to be an  **$n$ -dimensional locally Euclidean space**, if it is locally homeomorphic to  $\mathbb{R}^n$ . The nonnegative integer  $n$  is called the **dimension** of the space.

Euclidean  $n$ -space  $\mathbb{R}^n$  is, of course, locally Euclidean of dimension  $n$ , as is every open subset of  $\mathbb{R}^n$ . And according to Examples 3.25 (14), so is the  $n$ -sphere  $S_n$ . Since  $\mathbb{R}^0 = \{0\}$ , a 0-dimensional locally Euclidean space is just a discrete space. Further examples, including the open Möbius strip, the Klein bottle, and projective  $n$ -space  $\mathbb{R}P_n$ , are introduced later.

Notice that the definition of “locally Euclidean” deliberately excludes spaces such as the subspace  $B_2 \cup ([1, 2[ \times \{0\})$  of the plane, in which some points have neighborhoods homeomorphic to  $\mathbb{R}^2$  whereas others have neighborhoods homeomorphic to  $\mathbb{R}^1$ .

In a locally Euclidean space, each point has, in fact, an open neighborhood that is homeomorphic to all of  $\mathbb{R}^n$ .

prop:locally-eucl-has-nbds-homeo-Rn

**3.33 Proposition.** *Let  $X$  be an  $n$ -dimensional locally Euclidean space. Then each point of  $X$  has an open neighborhood that is homeomorphic to  $\mathbb{R}^n$  itself.*

**Proof.** Let  $x \in X$ . There is homeomorphism  $h: U \cong f(U)$  with domain some open neighborhood  $U$  of  $x$  in  $X$  such that  $f(U)$  is an open neighborhood of  $h(x)$  in  $\mathbb{R}^n$ . Choose  $\varepsilon > 0$  such that the ball  $B = B_\varepsilon(h(x); d) \subset f(U)$ , where  $d$  is the Euclidean metric. The set

$$V = h^{-1}(B)$$

is an open subset of  $U$  containing  $x$ , and the domain-codomain restriction  $h|_{V, B}: V \rightarrow B$  is a homeomorphism. But by [Examples 3.25 \(10\)](#),  $B \cong \mathbb{R}^n$ . Hence  $V \cong \mathbb{R}^n$ .  $\square$

According to the preceding proposition, each point of an  $n$ -dimensional locally Euclidean space has some neighborhood that topologically looks exactly like Euclidean  $n$ -space. Something more is true.

euclidean-nbds-in-loc-euclidean-space

**3.34 Proposition.** *Let  $X$  be an  $n$ -dimensional locally Euclidean space. Then for each point  $x$  of  $X$ , each neighborhood  $U$  of  $x$  contains an open neighborhood  $V$  of  $x$  with  $V \cong \mathbb{R}^n$ .*

**Proof.** The proof will be a refinement of that used for [Proposition 3.33](#).

Let  $x \in X$ , and let  $U$  be a neighborhood of  $x$  in  $X$ . To establish existence of such a  $V$ , we break the proof into three steps, corresponding to the three parts of [Figure 3.7](#).

pfstep-loc-euclidean-1

Step 1: There is an open neighborhood  $U_0$  of  $x$  in  $X$  with  $U_0 \subset U$ , and there is an open neighborhood  $U_1$  of  $x$  in  $X$  and a homeomorphism

$$h: U_1 \cong \mathbb{R}^n.$$

pfstep-loc-euclidean-2

Step 2: The set

$$W = U_0 \cap U_1$$

is an open neighborhood of  $x$  in the subspace  $U_1$  of  $X$ , so that its image  $h(W)$  is an open neighborhood of  $h(x)$  in  $\mathbb{R}^n$  (see [Figure 3.7](#)). If  $d$  denotes the Euclidean metric on  $\mathbb{R}^n$ , there is an  $\varepsilon > 0$  with

$$B_\varepsilon(h(x); d) \subset h(W).$$

pfstep-loc-euclidean-3

Step 3: Let

$$V = h^{-1}(B_\varepsilon(h(x); d)).$$

From [Examples 3.25 \(10\)](#) we have  $V \cong \mathbb{R}^n$ , and  $V$  is an open neighborhood of  $x = h^{-1}(h(x))$  in  $W$  and hence in  $X$ .  $\square$

Thus not only does each point in a space that is locally Euclidean of dimension  $n$  have some *sufficiently small* “Euclidean neighborhood”—a neighborhood that topologically looks exactly like Euclidean  $n$ -space—but each point has *arbitrary small* Euclidean neighborhoods.

The important idea about a space being locally Euclidean of dimension  $n$  is that in the neighborhood of each given point  $p$  we can assign “local coordinates.”



Figure 3.7: Steps in proving that each point in an  $n$ -dimensional locally Euclidean space has arbitrarily small neighborhoods that are Euclidean.

fig:manifold-is-locally-euclidean

def:loc-coord

**3.35 Definition.** Let  $X$  be an  $n$ -dimensional locally Euclidean space. An ordered pair  $\langle U, \varphi \rangle$  consisting of an open subset  $U$  of  $X$  and a homeomorphism  $\varphi: U \cong U'$  from  $U$  to an open subset  $U'$  of  $\mathbb{R}^n$  is called a **chart** in  $X$ , the open set  $U$  is called a **coordinate domain**, and the homeomorphism  $\varphi$  is called a **coordinate map**. When  $p \in U$ , we say that  $\langle U, \varphi \rangle$  is a **chart at**  $p$ , etc.; in this case, for each point  $x \in U$ , the successive entries in the  $n$ -tuple  $\varphi(x) = \langle x_1, x_2, \dots, x_n \rangle \in U' \subset \mathbb{R}^n$  are called the **local coordinates of**  $x$  (**with respect to**  $\varphi$ ).

An **atlas** on  $X$  is a collection of charts whose domains cover  $X$ .

If  $\langle U, \varphi \rangle$  is a chart in  $X$  at  $p$ , then there is always another chart  $\langle V, \psi \rangle$  at  $p$  with  $V \subset U$  and  $\varphi(V) = \mathbb{R}^n$  and even with  $\psi(p) = \mathbf{0}$ —see [Proposition 3.34](#) and [Exercise 98](#).

ex:stereographic map

**3.36 Examples.** (1) The stereographic map

$$\pi^+ : S_n \setminus \{\mathbf{p}\} \cong \mathbb{R}^n,$$

from Examples 3.25 (13), and its composition

$$\pi^- : S_n \setminus \{-\mathbf{p}\} \cong \mathbb{R}^n$$

with the antipodal map, from Examples 3.25 (14), give rise to the charts  $\langle S_n \setminus \{\mathbf{p}\}, \pi^+ \rangle$  and  $\langle S_n \setminus \{-\mathbf{p}\}, \pi^- \rangle$  in  $X$ , respectively. Here  $\mathbf{p}$  and  $-\mathbf{p}$  are the north pole and the south pole, respectively. While in general it may take many charts in a locally Euclidean space to have their coordinate domains cover  $X$ , in this example it takes just the two coordinate domains  $S_n \setminus \{\mathbf{p}\}$  and  $S_n \setminus \{-\mathbf{p}\}$ . In other words,  $\{\langle S_n \setminus \{\mathbf{p}\}, \pi^+ \rangle, \langle S_n \setminus \{-\mathbf{p}\}, \pi^- \rangle\}$  is an atlas on  $S_n$ .

- (2) On the plane  $\mathbb{R}^2$ , the identity map gives rise to the atlas  $\{\langle \mathbb{R}^2, \iota_{\mathbb{R}^2} \rangle\}$  consisting of a single chart, and the local coordinates of a point in  $\mathbb{R}^2$  for this chart are just its usual  $x$ - and  $y$ -coordinates.

More generally, for an arbitrary open subset  $U$  of  $\mathbb{R}^n$   $\{\langle U, \iota_U \rangle\}$  is an atlas on  $U$ .

ex:polar-coords

- (3) Recall that each point  $\langle x, y \rangle$  in the plane  $\mathbb{R}^2$  has many pairs of “polar coordinates”  $\langle r, \theta \rangle$ . In fact, each point other than the origin has a whole sequence of polar coordinates; but the origin has as polar coordinates  $\langle 0, \theta \rangle$  for every real  $\theta$ . Then if we are to use polar coordinates as local coordinates for charts in an atlas, we must omit the origin, in other words, use the punctured plane  $\mathbb{R}^2 \setminus \{\langle 0, 0 \rangle\}$  rather than the entire plane  $\mathbb{R}^2$ .

Start with the two maps

$$\begin{aligned} f : ]0, \infty[ \times ]0, 2\pi[ &\rightarrow \mathbb{R}^2, \\ \langle r, \theta \rangle &\mapsto \langle r \cos \theta, r \sin \theta \rangle \\ g : ]0, \infty[ \times ]-\pi, \pi[ &\rightarrow \mathbb{R}^2, \\ \langle r, \theta \rangle &\mapsto \langle r \cos \theta, r \sin \theta \rangle \end{aligned}$$

which are continuous injections with respective ranges

$$\begin{aligned} U &= \mathbb{R}^2 \setminus ]0, \infty[ \times \{0\}, \\ V &= \mathbb{R}^2 \setminus ]-\infty, 0[ \times \{0\} \end{aligned}$$

—the plane with the nonnegative real axis and the nonpositive real axis removed, respectively. Thus  $U$  and  $V$  are open sets that cover the punctured plane  $\mathbb{R}^2 \setminus \{\langle 0, 0 \rangle\}$ . Let  $f'$  and  $g'$  be the codomain restrictions of  $f$  and  $g$ , respectively, to their ranges, so that

$$\begin{aligned} f' : ]0, \infty[ \times ]0, 2\pi[ &\rightarrow U, \\ g' : ]0, \infty[ \times ]-\pi, \pi[ &\rightarrow V \end{aligned}$$

are continuous bijections. Their inverses

$$\begin{aligned} \varphi : U &\rightarrow ]0, \infty[ \times ]0, 2\pi[, \\ \psi : V &\rightarrow ]0, \infty[ \times ]-\pi, \pi[ \end{aligned}$$

are readily seen to be continuous as well, so that they are homeomorphisms. Thus  $\{\langle U, \varphi \rangle, \langle V, \psi \rangle\}$  is an atlas on the punctured plane  $\mathbb{R}^2 \setminus \{\langle 0, 0 \rangle\}$ .  $\diamond$

punctured plane

**local coordinate system** Suppose  $\langle U, \varphi \rangle$  is chart at a point of a locally Euclidean space  $X$  of dimension  $n$ . For simplicity, suppose also that  $\varphi(U) = \mathbb{R}^n$ . Then we may use  $\varphi$  to assign to each point  $x \in U$  its local coordinates  $\varphi(x) = \langle x_1, x_2, \dots, x_n \rangle$ ; and every  $n$ -tuple in  $\mathbb{R}^n$  constitutes the local coordinates of a unique point in  $U$ . Moreover, the local coordinates  $\langle x_1, x_2, \dots, x_n \rangle$  vary continuously as  $x$  varies within  $U$ ; and the point  $x$  in  $U$  having local coordinates  $\langle x_1, x_2, \dots, x_n \rangle$  varies continuously as its local coordinates  $\langle x_1, x_2, \dots, x_n \rangle$  vary in  $\mathbb{R}^n$ .

The idea of local coordinates is most easily understood in the case  $n = 2$ . Fix a point  $p$  in a 2-dimensional locally Euclidean space  $X$  and let  $\langle U, \varphi \rangle$  be a chart at  $p$ . Let  $c = \langle c_1, c_2 \rangle$  be the local coordinates of  $p$  with respect to  $h$ , so that  $c = \varphi(p)$ . (Perhaps  $c = \langle 0, 0 \rangle$ , the origin of  $\mathbb{R}^2$ .) The vertical line  $\{c_1\} \times \mathbb{R}$  and the horizontal line  $\mathbb{R} \times \{c_2\}$  intersect at  $c$ ; the family

$$\langle \{x_1\} \times \mathbb{R} : x_1 \in \mathbb{R} \rangle$$

of parallel vertical lines and the family

$$\langle \mathbb{R} \times \{x_2\} : x_2 \in \mathbb{R} \rangle$$

of parallel horizontal lines fill out  $\mathbb{R}^n$  that coordinatise  $\mathbb{R}^2$  in the usual way. Then the vertical line  $\{c_1\} \times \mathbb{R}$  and the horizontal line  $\mathbb{R} \times \{c_2\}$  give rise to the pair of curves  $\varphi^{-1}(\{c_1\} \times \mathbb{R})$  and  $\varphi^{-1}(\mathbb{R} \times \{c_2\})$ , respectively, in  $U$  intersecting at  $p$ . See Figure 3.8.

And the neighborhood  $U$  of  $p$  is filled out by the family of curves



Figure 3.8: Local coordinates near a point in a 2-dimensional locally Euclidean space.

fig:local-coord-sys-on-surface

$$\langle \varphi^{-1}(\{x_1\} \times \mathbb{R}) : x_1 \in \mathbb{R} \rangle$$

and the family of curves

$$\langle \varphi^{-1}(\mathbb{R} \times \{x_2\}) : x_2 \in \mathbb{R} \rangle$$

where the first coordinate  $x_1$  and the second coordinate  $x_2$  are fixed in the first family and the second family, respectively. These curves are no longer necessarily lines, as their sources in  $\mathbb{R}^2$  were, but may well be “curvy” curves. The two families of curves in  $U$  are sometimes referred to as a **local coordinate system at  $p$** .

Similarly, when  $n = 3$ , we have three families, each family consisting of lines parallel to one of the Euclidean coordinate axes. Then the inverse images of these families under a chart  $\langle U, h \rangle$  will be the corresponding three families of curves in  $U$ .

A given point  $p$  of a locally Euclidean space  $X$  will typically belong to the coordinate domains of several, even of many, different charts for  $X$ . When such a point belongs to the overlap  $U \cap V$  of the coordinate domains of two charts  $\langle U, \varphi \rangle$  and  $\langle V, \psi \rangle$ , then each point  $x$  in that overlap may have quite different local coordinates  $\varphi(x) = \langle x_1, x_2, \dots, x_n \rangle$  and  $\psi(x) = \langle y_1, y_2, \dots, y_n \rangle$  for the two charts. See Figure 3.9.

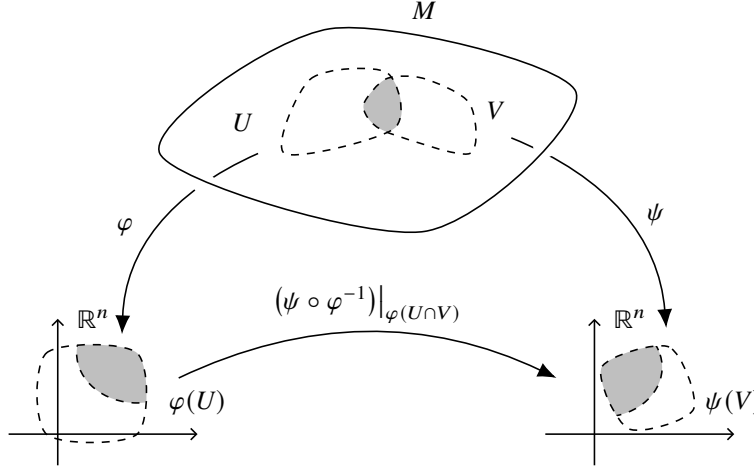


Figure 3.9: Transition map for two charts  $\langle U, \varphi \rangle$  and  $\langle V, \psi \rangle$  in a 2-dimensional locally Euclidean space  $M$ .

fig:transition-map

For two such charts  $\langle U, \varphi \rangle$  and  $\langle V, \psi \rangle$  at the same point  $p$ , the **transition map**

$$(\psi \circ \varphi^{-1})|_{\varphi(U \cap V)} : \varphi(U \cap V) \rightarrow \psi(U \cap V)$$

**from  $\varphi$  to  $\psi$**  is a homeomorphism between open subsets of  $\mathbb{R}^n$ . This map describes a **change of local coordinates**: If

$$\varphi(x) = \langle x_1, x_2, \dots, x_n \rangle$$

are the local coordinates of a point  $x \in U \cap V$  with respect to  $\varphi$  and if

$$\psi(x) = \langle y_1, y_2, \dots, y_n \rangle$$

are the local coordinates of  $x$  with respect to  $\psi$ , then these two  $n$ -tuples are related by

$$\langle y_1, y_2, \dots, y_n \rangle = (\psi \circ \varphi^{-1}) \langle x_1, x_2, \dots, x_n \rangle$$

so that each local coordinate  $y_j$  with respect to  $\psi$  is related to the  $n$ -tuple of local coordinates with respect to  $\varphi$  by

$$y_j = (\psi \circ \varphi^{-1})_j \langle x_1, x_2, \dots, x_n \rangle.$$

ex:transition-map-sterographic

stereographic projection

**3.37 Example.** Let

$$U^+ = S_n \setminus \{\mathbf{p}\}, \quad U^- = S_n \setminus \{-\mathbf{p}\}$$

and consider again the homeomorphisms

$$\begin{aligned} \pi^+ : U^+ &\cong \mathbb{R}^n, \\ \pi^- : U^- &\cong \mathbb{R}^n \end{aligned}$$

constructed in [Examples 3.25 \(13\)](#) and [\(14\)](#), respectively. Then the collection  $\{\langle U^+, \pi^+ \rangle, \langle U^-, \pi^- \rangle\}$  is an atlas on  $S_n$ . Explicitly,

$$\begin{aligned} \pi^+(x_1, x_2, \dots, x_n, x_{n+1}) &= \left\langle \frac{x_1}{1 - x_{n+1}}, \frac{x_2}{1 - x_{n+1}}, \dots, \frac{x_n}{1 - x_{n+1}} \right\rangle, \\ \pi^-(x_1, x_2, \dots, x_n, x_{n+1}) &= \left\langle \frac{x_1}{1 + x_{n+1}}, \frac{x_2}{1 + x_{n+1}}, \dots, \frac{x_n}{1 + x_{n+1}} \right\rangle. \end{aligned}$$

By solving the equation  $\pi^+(x_1, x_2, \dots, x_n, x_{n+1}) = \langle y_1, y_2, \dots, y_n \rangle$  for  $x_1, x_2, \dots, x_n, x_{n+1}$  in terms of  $y_1, y_2, \dots, y_n$ , you will find that

$$(\pi^+)^{-1}(x_1, x_2, \dots, x_n, x_{n+1}) = \left\langle \frac{2y_1}{\|y\|^2 + 1}, \frac{2y_2}{\|y\|^2 + 1}, \dots, \frac{2y_n}{\|y\|^2 + 1}, \frac{\|y\|^2 - 1}{\|y\|^2 + 1} \right\rangle.$$

Now

$$\pi^+(U^+ \cap U^-) = \pi^-(U^+ \cap U^-) = \mathbb{R}^n \setminus \{\mathbf{0}\},$$

and the transition map

$$\pi^- \circ (\pi^+)^{-1} : \mathbb{R}^n \setminus \{\mathbf{0}\} \cong \mathbb{R}^n \setminus \{\mathbf{0}\}$$

is given, for  $y = \langle y_1, y_2, \dots, y_n \rangle \neq \mathbf{0}$ , by:

$$\pi^- \circ (\pi^+)^{-1}(y_1, y_2, \dots, y_n) = \left\langle \frac{y_1}{\|y\|^2}, \frac{y_2}{\|y\|^2}, \dots, \frac{y_n}{\|y\|^2} \right\rangle = \frac{1}{\|y\|^2} y. \quad \diamond$$

When defining “locally Euclidean”, we tacitly assumed that the dimension of a given locally Euclidean space is unique. This is a consequence of the following theorem.

thm:Rn-not-homeo-Rm

**3.38 Theorem.** *Euclidean  $n$ -space  $\mathbb{R}^n$  is not homeomorphic to Euclidean  $m$ -space  $\mathbb{R}^m$  unless  $m = n$ .*

The preceding result is trivial in the case  $m = 0$ ; later [\[Examples 5.23 \(3\)\]](#) we shall use connectedness to give a simple proof in the case  $m = 1 < n$ . The general case is a consequence of the following theorem.

thm:invariance-of-domain

**3.39 Invariance of Domain Theorem.** *Let  $f : U \rightarrow \mathbb{R}^n$  be a continuous injection from an open subset  $U$  of  $\mathbb{R}^n$  to  $\mathbb{R}^n$ . Then the image  $f(U)$  of  $U$  is open in  $\mathbb{R}^n$  and the codomain restriction  $f|_{U, f(U)} : U \rightarrow f(U)$  is a homeomorphism.*

Standard proofs of Invariance of Domain Theorem require tools from algebraic topology that are beyond the scope of this book; see, for example, Hatcher [\[34, Theorem 2B.3\]](#) or Munkres [\[50, Theorem 36.5\]](#).



**Proof of Theorem 3.38.** Let  $m$  and  $n$  be nonnegative integers with  $m < n$  and just suppose there is some homeomorphism  $h: \mathbb{R}^n \cong \mathbb{R}^m$ . If  $m = 0$ , then we are done because  $\mathbb{R}^m = \{0\}$  is a one-point space whereas  $\mathbb{R}^n$  is not. So suppose  $0 < m < n$ . The map

$$g: \mathbb{R}^m \rightarrow \mathbb{R}^n \\ \langle x_1, x_2, \dots, x_m \rangle \mapsto \langle x_1, x_2, \dots, x_m, \underbrace{0, 0, \dots, 0}_{n-m \text{ 0s}} \rangle$$

is a continuous injection with range

$$g(\mathbb{R}^m) = \{x \in \mathbb{R}^n : x_{m+1} = x_{m+2} = \dots = x_n = 0\}.$$

Then the composite  $h = g \circ f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a continuous injection. By Invariance of Domain Theorem (3.39), the image  $h(\mathbb{R}^n)$  must be open in  $\mathbb{R}^n$ . However,  $h(\mathbb{R}^n) = g(\mathbb{R}^m)$ , which is manifestly not open.  $\square$

### Topological manifolds

subsec:manifolds

The locally Euclidean spaces of greatest interest to us are those known as “manifolds.”

def:n-manifold

**3.40 Definition.** For a given nonnegative integer  $n$ , an  **$n$ -dimensional manifold**, or more tersely an  **$n$ -manifold**, is a topological space that:

- is locally Euclidean of dimension  $n$ ;
- is a Hausdorff space; and
- is second-countable.

The nonnegative integer  $n$  is called the **dimension** of the  $n$ -manifold. In particular, a 2-dimensional manifold is also called a **surface**.

The space  $\mathbb{R}^n$  is itself an  $n$ -dimensional manifold, for  $\mathbb{R}^n$  itself is an open neighborhood of each point of  $\mathbb{R}^n$  that is homeomorphic to  $\mathbb{R}^n$ . More generally, an arbitrary open subset  $M$  of  $\mathbb{R}^n$  is an  $n$ -dimensional manifold; in fact, if  $x \in M$ , there is a Euclidean ball  $B_\varepsilon(x; d) \subset M$ , and we know from Examples 3.25 (10) that  $B_\varepsilon(x; d) \cong \mathbb{R}^n$ . According to Examples 3.25 (14), which motivated the preceding definition, the  $n$ -sphere  $S_n$  is an  $n$ -manifold; in particular, the 2-sphere  $S_2$  is a surface. Additional examples appear in the next two sections.

Manifolds in the sense of Definition 3.40 are often called **topological manifolds**, so as to distinguish them from *differentiable manifolds*, which have extra structure allowing us to talk about differentiable functions on them), and from other special kinds of manifolds.

**Usage note.** The requirement in Definition 3.40 that a manifold be a Hausdorff space is a standard one intended to exclude certain pathological examples (such as the line with two origins—see Exercise 81). The requirement that a manifold be second-countable is also a standard one, intended to exclude certain other pathological examples, such as the one that follows. (*Caution:* Some authors do *not* insist that a topological manifold be second-countable.)

The following example is of a locally Euclidean Hausdorff space that is *not* a manifold. Understanding the example requires familiarity with the set of all countable ordinals; the material in subsection “The first uncountable ordinal” of Chapter 0 will suffice as preparation.

topological manifold  
manifold!differentiable  
differentiable manifold  
line with two origins!manifold@and  
ordinal

longline  
longline  
long closed ray  
ordinal

**3.41 Example.** Give the set  $[0, \Omega[$  of all countable ordinals the order described in subsection “The first uncountable ordinal” (page 110), and give the half-open, half-closed unit interval  $[0, 1[$  its usual order. Give the product set  $[0, \Omega[ \times [0, 1[$  its lexicographic ordering (Example 0.70) and then endow it with the corresponding order topology [Examples 2.72 (1)]. The resulting topological space is the **long closed ray**, which we denote by  $L$ . Its subspace

$$L^* = ([0, \Omega[ \times [0, 1[ \setminus \{(0, 0)\})$$

is the **long line**.

While  $\langle 0, 0 \rangle \notin L^*$ , nonetheless  $\langle \alpha, 0 \rangle \in L^*$  for each  $\alpha > 0$ .  $\diamond$

**Usage note.** What we are calling the “long line” some authors call the “long open ray,” and then for them the “long line” is the topological space obtained by giving a copy  $L'$  of the long open ray the order topology induced by the *reverse* of its lexicographic ordering and then forming the Cartesian sum [Examples 3.47 (4)] of  $L'$  with the long closed ray  $L$ .

**3.42 Proposition.** *The long line  $L^*$  is a 1-dimensional locally Euclidean Hausdorff space that is not second countable and hence is not a manifold.*

**Proof.**  $L^*$  is a Hausdorff space: As with any totally ordered set under its order topology, so  $L^*$  is a Hausdorff space [see Examples 2.72 (1) and Exercise 2.89].

$L^*$  is not second-countable: Just suppose  $L^*$  is second-countable. The collection

$$\mathcal{A} = \{ ]\langle 0, 0 \rangle, \langle \alpha, 0 \rangle[ : \alpha \in ]0, \Omega[ \}$$

is a base of  $L^*$ . From Theorem 2.79, there is a countable subset  $B$  of  $]0, \Omega[$  for which the subcollection

$$\mathcal{B} = \{ ]\langle 0, 0 \rangle, \langle \beta, 0 \rangle[ : \beta \in B \}$$

of  $\mathcal{A}$  is also a base of  $L^*$ . From (Q4) of the set  $\Omega^+$ , there is some  $\gamma \in ]0, \Omega[$  with  $\gamma > \beta$  for all  $\beta \in B$ . Then  $\langle \gamma, 0 \rangle \notin \bigcup \mathcal{B}$ . This is impossible.

$L^*$  is locally Euclidean of dimension 1: Let  $p = \langle \gamma, t \rangle \in L^*$  be arbitrary. Let  $\alpha$  be the immediate successor of  $\gamma$  in the well-ordered set  $]0, \Omega[$ , so that  $p \in ]\langle 0, 0 \rangle, \langle \alpha, 0 \rangle[$ . In Lemma 3.43, below, we shall show that there is an order-isomorphism

$$f: ]\langle 0, 0 \rangle, \langle \alpha, 0 \rangle[ \rightarrow [0, 1[.$$

Necessarily  $f(\langle 0, 0 \rangle) = 0$ . Then the domain-codomain restriction

$$f|_{] \langle 0, 0 \rangle, \langle \alpha, 0 \rangle[ \rightarrow [0, 1[}$$

is also an order-isomorphism. Since the topologies of both domain and the codomain of this restriction are their order topologies, this restriction is a homeomorphism [see Exercise 75 (a)].  $\square$

Here is the statement of the result used to show that the long line is locally Euclidean.

**3.43 Lemma.** *For each  $\alpha \in ]0, \Omega[$ , the subinterval  $] \langle 0, 0 \rangle, \langle \alpha, 0 \rangle[$  of the lexicographically ordered product  $[0, \Omega[ \times [0, 1[$  is order-isomorphic to the interval  $[0, 1[$ .*

**Proof.** Just suppose the assertion of the lemma is false. Then the subset

$$E = \{\alpha \in ]0, \Omega[ : [\langle 0, 0 \rangle, \langle \alpha, 0 \rangle[ \text{ is not order-isomorphic to } [0, 1[ \}$$

of  $]0, \Omega[$  is not empty. Since  $]0, \Omega[$  is well-ordered,  $E$  has a least element  $\beta$ . Now consider the two cases.

case:successor-ordinal Case (i):  $\beta$  is not a limit ordinal. This means that  $\beta$  has an immediate predecessor  $\alpha$ . Necessarily,  $\alpha \notin E$ , so that  $[\langle 0, 0 \rangle, \langle \alpha, 0 \rangle[$  is order-isomorphic to  $[0, 1[$ . Also,

$$[\langle \alpha, 0 \rangle, \langle \beta, 0 \rangle[ = \{\alpha, t[ : t \in [0, 1[ \} = \{\alpha\} \times [0, 1[$$

is order-isomorphic to  $[0, 1[$ . From Lemma 3.44, below, it follows that  $[\langle 0, 0 \rangle, \langle \beta, 0 \rangle[$  is order-isomorphic to  $[0, 1[$ . This is impossible since  $\beta \in E$ .

case:limit-ordinal Case (ii):  $\beta$  is a limit ordinal. Since  $\beta < \Omega$ , the first uncountable ordinal, the interval  $]0, \beta[$  is countable; since  $\beta$  has no immediate predecessor, this interval is infinite. Thus  $]0, \beta[$  is denumerable. Then there is a sequence  $\langle \alpha_n \rangle_{n \in \mathbb{N}}$  of distinct elements of  $]0, \beta[$  with  $]0, \beta[ = \{\alpha_n : n \in \mathbb{N}\}$ . We may assume this sequence is strictly increasing. The set  $\{\alpha_n : n \in \mathbb{N}\}$  is bounded above in  $]0, \Omega[$  by  $\beta$  and, in fact,  $\beta = \sup\{\alpha_n : n \in \mathbb{N}\}$ .

For each  $n \in \mathbb{N}$ , the ordinal  $\alpha_n$  is less than the least element  $\beta$  of  $E$ , so that the interval  $[\langle 0, 0 \rangle, \langle \alpha_n, 0 \rangle[$  is order-isomorphic to  $[0, 1[$ . From Lemma 3.45, below, it follows that  $[\langle 0, 0 \rangle, \langle \beta, 0 \rangle[$  is order-isomorphic to  $[0, 1[$ . Again this is impossible since  $\beta \in E$ .  $\square$

The following two lemmas are the facts about order-isomorphic totally ordered sets needed to complete the proof of Lemma 3.43 and hence the proof that the long line is locally Euclidean.

One-iff-two-abutting-subintervals-are **3.44 Lemma.** Let  $X$  be a totally ordered set, and let  $a, b, c \in X$  with  $a < c < b$ . Then the interval  $[a, b[$  is order-isomorphic to  $[0, 1[$  if and only if each of its subintervals  $[a, c[$  and  $[c, b[$  are order-isomorphic to  $[0, 1[$ .

**Proof.** Note that if  $f: [a, b[ \rightarrow [0, 1[$  is an order-isomorphism, then necessarily  $f(a) = 0$ .

Assume first that  $[a, b[$  is order-isomorphic to  $[0, 1[$  and let  $f: [a, b[ \rightarrow [0, 1[$  be an order-isomorphism. Then the maps

$$\begin{aligned} [a, c[ &\rightarrow [0, 1[ & [c, b[ &\rightarrow [0, 1[ \\ x &\mapsto \frac{f(x)}{f(c)}, & x &\mapsto \frac{f(x) - f(c)}{1 - f(c)} \end{aligned}$$

—obtained for the first by just scaling the values of  $f$ , and for the second by first shifting and then scaling the values of  $f$ —are also order-isomorphisms.

Assume, conversely, that both  $[a, c[$  and  $[c, b[$  are order-isomorphic to  $[0, 1[$  and let

$$g: [a, c[ \rightarrow [0, 1[, \quad h: [c, b[ \rightarrow [0, 1[$$

be order-isomorphisms. Then the map  $f: [a, b[ \rightarrow [0, 1[$  given piecewise by

$$f(x) = \begin{cases} \frac{1}{2} g(x) & \text{if } x \in [a, c[, \\ \frac{1}{2} + \frac{1}{2} h(x) & \text{if } x \in [c, b[ \end{cases}$$

—obtained by just scaling the values of  $g$  for  $[a, c[$  but first scaling and then shifting the values of  $h$  for  $[c, b[$ —is an order-isomorphism.  $\square$

longline  
ordinal  
n-manifold  
manifold  
locally Euclidean sp

**3.45 Lemma.** Let  $X$  be a totally ordered set, let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a strictly increasing sequence in a totally ordered set that is bounded above in  $X$ , and let  $b = \sup\{x_n : n \in \mathbb{N}\}$ . Then  $[x_0, b[$  is order-isomorphic to  $[0, 1[$  if and only if each of the intervals  $[x_n, x_{n+1}[$  is order-isomorphic to  $[0, 1[$ .

**Proof.** Assume first that  $[x_0, b[$  is order-isomorphic to  $[0, 1[$  and let  $f: [x_0, b[ \rightarrow [0, 1[$  be an order-isomorphism. Then for each  $n \in \mathbb{N}$  the map

$$f_n: [x_n, x_{n+1}[ \rightarrow [0, 1[$$

given by

$$f_n(x) = \frac{f(x) - f(x_n)}{f(x_{n+1}) - f(x_n)}$$

—obtained by shifting and scaling—is an order-isomorphism.

Assume, conversely, that  $[x_n, x_{n+1}[$  is order-isomorphic to  $[0, 1[$  for each  $n \in \mathbb{N}$  and for each  $n$  let

$$f_n: [x_n, x_{n+1}[ \rightarrow [0, 1[$$

be an order-isomorphism. Observe that the sequence  $\langle [x_n, x_{n+1}[ \rangle_{n \in \mathbb{N}}$  of successive subintervals of  $[x_0, b[$  forms a partition of  $[x_0, b[$ . Correspondingly partition the interval  $[0, 1[$  into its successive subintervals  $[1 - 1/2^n, 1 - 1/2^{n+1}[$  for  $n = 0, 1, 2, \dots$ . For each  $n \in \mathbb{N}$ , scale and then shift the values of  $f_n$  so as to map  $[x_n, x_{n+1}[$  to  $[1 - 1/2^n, 1 - 1/2^{n+1}[$ . In other words, form the map

$$f: [x_0, b[ \rightarrow [0, 1[$$

given piecewise by

$$f(x) = \left(1 - \frac{1}{2^n}\right) + \frac{1}{2^{n+1}} f_n(x) \quad \text{if } x \in [x_n, x_{n+1}[.$$

Then  $f$  is an order isomorphism.  $\square$

## Embedding

subsec:embedding

When a homeomorphic image of a space is a subspace of a larger space of interest, some special terminology is employed.

**3.46 Definition.** A map  $f: X \rightarrow Y$  from a topological space  $X$  to a topological space  $Y$  is called an **embedding of  $X$  into  $Y$**  when the map

$$\begin{aligned} X &\rightarrow f(X) \\ x &\mapsto f(x) \end{aligned}$$

obtained by restricting its codomain to its range is a homeomorphism from  $X$  to the subspace  $f(X)$  of  $Y$ .

When there exists an embedding of a space  $X$  into a space  $Y$ , we say that  $X$  is **embeddable in  $Y$**  and that  $X$  can be embedded in  $Y$ .

Thus an embedding of  $X$  into  $Y$  is in essence a homeomorphism from  $X$  to a subspace of  $Y$ , but regarded as a map into all of  $Y$ . Evidently **a map  $f: X \rightarrow Y$  is an embedding precisely when  $f$  is a continuous injection such that the image  $f(U)$  of each open subset  $U$  of  $X$  is open in  $f(X)$** —but not necessarily open in  $Y$ .

metric-embedding-is-top-embedding

**3.47 Examples.** (1) An isometric embedding of one metric space into another, in the sense of [Exercise 1.59](#), is an embedding of the associated topological spaces. To avoid such ambiguity, we therefore sometimes call an embedding in the sense of [Definition 3.46](#) a **topological embedding** or a **homeomorphic embedding**.

isometric embedding  
embedding!topological  
embedding!homeomorphic  
Cartesian sum!two spaces@of two spaces

ex:embed-Rn-in-Rnplus1

(2) According to [Examples 1.46 \(3\)](#), the map

$$\begin{aligned} \mathbb{R}^n &\rightarrow \mathbb{R}^{n+1} \\ \langle x_1, x_2, \dots, x_n \rangle &\mapsto \langle x_1, x_2, \dots, x_n, 0 \rangle \end{aligned}$$

is a topological embedding of  $\mathbb{R}^n$  into  $\mathbb{R}^{n+1}$  (in fact, an isometric embedding for the Euclidean metrics).

In dimension  $n = 1$  this is the embedding

$$\begin{aligned} h: \mathbb{R} &\rightarrow \mathbb{R} \times \mathbb{R} \\ x &\mapsto \langle x, 0 \rangle \end{aligned}$$

of the real line into the plane with image the  $x$ -axis  $\mathbb{R} \times \{0\}$ .

The subspace  $h(\mathbb{R}) = \mathbb{R} \times \{0\}$  of  $\mathbb{R} \times \mathbb{R}$  is a “homeomorphic copy” of  $\mathbb{R}$  that certainly is not identical to  $\mathbb{R}$ . Nonetheless, by using  $h$  to “identify” each point  $x \in \mathbb{R}$  with the corresponding point  $h(x) = \langle x, 0 \rangle \in \mathbb{R} \times \mathbb{R}$  we may speak about  $\mathbb{R}$  as if it actually were the subspace  $\mathbb{R} \times \{0\}$  of  $\mathbb{R} \times \mathbb{R}$ . We may say, for example, that the subset  $]0, 1[$  of the real line has empty interior in the plane, meaning really that  $]0, 1[ \times \{0\}$  has empty interior there.

This “abuse of language” that identifies the nonidentical objects  $\mathbb{R}$  and  $\mathbb{R} \times \{0\}$  is the topological counterpart to the algebraic identification of  $\mathbb{R}$  with  $\mathbb{R} \times \{0\}$  that treats each real number  $x$  as the complex number  $x + 0i = \langle x, 0 \rangle + \langle 0, 0 \rangle \langle 0, 1 \rangle$ , as explained in the [subsection “Complex numbers”](#) ([page 89](#)).

ex:embed-cube-into-cube

(3) For positive integers  $m$  and  $n$  with  $m < n$ , the map

$$\begin{aligned} l^m &\rightarrow l^n \\ \langle x_1, x_2, \dots, x_m \rangle &\mapsto \langle x_1, x_2, \dots, x_m, 0, 0, \dots, 0 \rangle \end{aligned}$$

is an embedding of  $l^m$  into  $l^n$ .

Similarly, for each  $n$  the rule  $\langle x_1, x_2, \dots, x_n \rangle \mapsto \langle x_1, x_2, \dots, x_n, 0, 0, \dots, 0, \dots \rangle$  defines an embedding of  $l^n$  into the Hilbert cube  $l^\infty$ . (Later, in [Theorem 6.45](#), we shall see that a much wider collection of spaces can be embedded in the Hilbert cube.)

More generally, if  $A$  and  $B$  are sets with  $A \subset B$ , then similarly there is an embedding of  $l^A$  into  $l^B$ .

ex:top-sum-two

(4) Any two topological spaces  $X_1$  and  $X_2$  can simultaneously be embedded into the same topological space  $Z$ .

To see how, consider first the special case when  $X_1$  and  $X_2$  are disjoint. Let

$$Z = X_1 \cup X_2,$$

be their union. Then the collection

$$\{U_1 \cup U_2 : U_1 \text{ is open in } X_1, U_2 \text{ is open in } X_2\}$$

of subsets of  $Z$ , which may also be described as

$$\{U : U \subset Z, U \cap X_1 \text{ open in } X_1, U \cap X_2 \text{ open in } X_2\},$$

is a topology on  $Z$ . With this topology on  $Z$ , both  $X_1$  and  $X_2$  with their given topologies are then subspaces of  $Z$ . [In the language of [Exercise 2.18](#),  $Z$  is the (Cartesian) sum of the family  $\langle X_i \rangle_{i \in \{1, 2\}}$ .]

Cartesian sum of two spaces! Consider now the general case when  $X_1$  need not be disjoint from  $X_2$ . The preceding construction no longer works because  $X_1$  and  $X_2$  might induce different topologies on their overlap  $X_1 \cap X_2$ . To get around this obstacle, make  $X_1$  disjoint from  $X_2$  by labeling each element of  $X_1$  with 1 and each element of  $X_2$  with 2, obtaining the sets

$$X_1 \times \{1\} = \{\langle x, 1 \rangle : x \in X_1\}, \quad X_2 \times \{2\} = \{\langle x, 2 \rangle : x \in X_2\},$$

which are disjoint. Now endow the sets  $X_1 \times \{1\}$  and  $X_2 \times \{2\}$  with the topologies

$$\{U_1 \times \{1\} : U_1 \text{ open in } X_1\}, \quad \{U_2 \times \{2\} : U_2 \text{ open in } X_2\},$$

which make the maps

$$\begin{array}{ll} X_1 \rightarrow X_1 \times \{1\}, & X_2 \rightarrow X_2 \times \{1\} \\ x \mapsto \langle x, 1 \rangle & x \mapsto \langle x, 2 \rangle \end{array}$$

homeomorphisms.

Finally, endow the union

$$Z = (X_1 \times \{1\}) \cup (X_2 \times \{2\})$$

with the topology

$$\{(U_1 \times \{1\}) \cup (U_2 \times \{2\}) : U_1 \text{ open in } X_1, U_2 \text{ open in } X_2\}$$

provided by the case of disjoint spaces. This topology makes  $X_1 \times \{1\}$  and  $X_2 \times \{2\}$  subspaces of  $Z$  and makes the maps

$$\begin{array}{ll} h_1 : X_1 \rightarrow Z & , \quad h_2 : X_2 \rightarrow Z \\ x \mapsto (x, 1) & \quad x \mapsto (x, 2) \end{array}$$

embeddings of  $X_1$  and  $X_2$ , respectively, into  $Z$ . The topological space  $Z$  is called the **(Cartesian) sum** of  $X_1$  and  $X_2$  and is denoted by  $X_1 + X_2$ .

*Abuse of language:* When  $X_1$  and  $X_2$  are disjoint, as we saw, we may use  $X_1$  and  $X_2$  as is, without “tagging” their elements with 1 or 2, respectively, and then the maps  $h_1$  and  $h_2$  are just the inclusion maps  $j_1 : X_1 \rightarrow X_1 \cup X_2$  and  $j_2 : X_2 \rightarrow X_1 \cup X_2$ . However, even when  $X_1$  and  $X_2$  are not disjoint, sometimes we act *as if* they were disjoint and then loosely regard  $X_1$  and  $X_2$  as actual subspaces of  $X_1 + X_2$ , suppressing mention of their “tags” and their embeddings into that Cartesian sum.

In [Exercise 106](#) the preceding construction is generalized to the situation of an arbitrary family of topological spaces.  $\diamond$

This is the place to correct a common misconception of topology that, unfortunately, has been disseminated in certain popularizations of mathematics: that two geometric objects are topologically the same only when one of them can be transformed into the other by “bending and stretching” but without “cutting and pasting.” To see why this is wrong, consider, for example, two embeddings of the circle  $S_1$  into Euclidean 3-space  $\mathbb{R}^3$  whose images are the *trefoil knot*  $X_1$  and the unknotted circle  $X_2$  as depicted in [Figure 3.10](#). The spaces  $X_1$  and  $X_2$ , being both homeomorphic to  $S_1$ , are topologically the same. However,  $X_1$  cannot be transformed into  $X_2$  within  $\mathbb{R}^3$  without cutting it apart, then unknitting it, and finally pasting the cut ends back together. (Mathematically, this means there is no homeomorphism of the ambient space  $\mathbb{R}^3$  onto itself that carries  $X_1$  onto  $X_2$ . Although we do not prove this fact here, a simple experiment with a loop of string should be convincing!)

For more advanced results about embedding, see [Chapter 6](#).

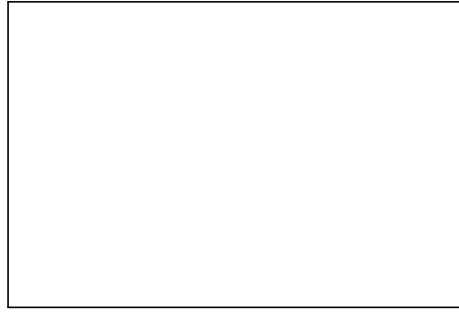


Figure 3.10: Trefoil knot (left) and unknotted circle (right) embedded in 3-space.

topological property!preserved  
Cartesian sum!two spaces@of two spaces

fig:knotted-unknotted-circles-in-R3

### Preservation of topological properties

subsec:preserve-top-props

So far we have seen just three general constructions of new spaces from old ones: subspace of a space; product of two (or finitely many) topological spaces; and Cartesian sum of two (or finitely many) topological spaces. Later, in [Sections 3.3](#) and [3.4](#) we shall encounter two additional constructions: product of an arbitrary family of topological spaces; and quotient of a topological space.

One of the things we shall want to know is whether a particular topological property is *preserved* under the construction of new topological spaces from old ones, that is, whether a space has that property when it is constructed from a given space, or given spaces, already having that property.

For example, here is the situation with preserving countability properties when forming the Cartesian sum [[Examples 3.47 \(4\)](#)] of two spaces.

**3.48 Proposition.** (1) *The Cartesian sum of two first-countable spaces is itself first-countable.*

(2) *The Cartesian sum of two second-countable spaces is itself second-countable.*

(3) *The Cartesian sum of two separable spaces is itself separable.*

**Proof.** Proofs of parts (2) and (3) are requested in [Exercise 93](#).

- (1) Let  $z$  be a point in the Cartesian sum  $X_1 + X_2$  of first-countable spaces  $X_1$  and  $X_2$ . Without loss of generality we may suppose that  $z \in X_1$ . There is a countable local base  $\mathcal{M}$  at  $x$  in  $X_1$ ; then the same collection  $\mathcal{M}$  is also local base at  $z$  in  $X_1 + X_2$ .

More precisely, as in [Examples 3.47 \(4\)](#) let  $Z = (X_1 \times \{1\}) \cup (X_2 \times \{2\})$  and let  $h_1: X_1 \rightarrow X_1 \times \{1\}$  be the embedding  $x \mapsto (x, 1)$ ; then the collection  $\{h_1(M) : M \in \mathcal{M}\}$  is a countable local base at  $h_1(z)$  in  $X_1 + X_2$ .  $\square$

Here is the situation with preserving separation properties when forming the Cartesian sum of two spaces. (Proofs are requested in [Exercise 94](#).)

prop-part:sum-2-preserve-countability

prop-part:sum-2-preserve-2nd-count

prop-part:sum-2-preserve-separable

Cartesian sum of two  
of-2-preserve-some-separation-prop  
topological properties  
line with two origins

- 3.49 Proposition.** (1) *The Cartesian sum of two  $T_0$ -spaces is itself a  $T_0$ -space.*
- (2) *The Cartesian sum of two  $T_1$ -spaces is itself a  $T_1$ -space.*
- (3) *The Cartesian sum of two  $T_2$ -spaces is itself a  $T_2$ -space.*
- (4) *The Cartesian sum of two regular spaces is itself a regular space.*

To say that a topological property is *hereditary*, or that subspaces *inherit* the property, means that the property is preserved under formation of subspaces. To say that a topological property is *preserved by finite products* means that if each of a finite family of topological spaces has the property, then so does their (Cartesian) product; to say that a property is *preserved by finite sums* means that if each of a finite family of topological spaces has the property, then so does their Cartesian sum.

At this point, we know:

- First-countability, second-countability, and metrizability are: hereditary [Proposition 2.78 and Examples 2.10 (1)]; preserved by finite products (Proposition 2.89); and preserved by finite sums (Proposition 3.48).
- Separability is: *not* hereditary (Example 2.88); but is preserved by finite products (Proposition 2.89); and is preserved by finite sums (Proposition 3.48).
- The properties  $T_0$ ,  $T_1$  space,  $T_2$ , and regularity are: hereditary (Proposition 2.102); preserved by finite products (Proposition 2.104); and preserved by finite sums (Proposition 3.49).

The situation for completely regular and normal spaces will be considered later, in Section 6.2.

For a summary of the preceding relationships, and more, see Table 3.1. (References concerning complete regularity and normality are to propositions and examples not appearing until later, in Section 6.2.)

### EXERCISES FOR SECTION 3.2

- 44.** Construct directly a homeomorphism  $f: ]0, 1[ \cong ]1, \infty[$ .
- 45.** Give several examples of continuous bijections that are not homeomorphisms.
- 46.** Give some examples of open and closed surjections that are not a homeomorphism.
- 47.** Prove Proposition 3.24.
- 48.** Let  $f: X \rightarrow Y$  be a bijection from a topological space  $X$  to a set  $Y$ . Then the collection  $\{V \subset Y : f^{-1}(V) \text{ is open in } X\}$  is a topology on  $Y$ , and for this topology the map  $f$  is a homeomorphism.
- Note:* This construction of a topology on  $Y$  will be generalized in Exercise 215.
- 49.** According to Examples 3.25 (8), the unit square is a simple closed curve, and this square is the subset  $\{(x, y) \in \mathbb{R}^2 : d_\infty(x, y) = 1\}$  of the  $xy$ -plane, where  $d_\infty$  is the max metric. Show that the square  $\{(x, y) \in \mathbb{R}^2 : d_1(x, y) = 1\}$ , where  $d_1$  is the taxicab metric, is also a simple closed curve.
- 50.** Show that the two representations of the line with two origins in Examples 2.20 (3)—the first using  $\mathbb{R} \cup \{0'\}$  and the second  $\mathbb{R}^* \cup \{0', 0''\}$ —give essentially the same topological space.



Property	Construction		
	subspace	finite product	finite sum
$T_0$	✓ (Proposition 2.102)	✓ (Exercise 2.150)	✓ (Proposition 3.49)
$T_1$	✓ (Proposition 2.102)	✓ (Exercise 2.150)	✓ (Proposition 3.49)
$T_2$	✓ (Proposition 2.102)	✓ (Exercise 2.150)	✓ (Proposition 3.49)
regular	✓ (Proposition 2.102)	✓ (Exercise 2.150)	✓ (Proposition 3.49)
completely regular	✓ (Proposition 6.18)	✓ (Proposition 6.18)	✓ (Proposition 6.18)
normal	✗ (Example 6.23)	✗ (Example 6.32)	✓ (Proposition 6.24)
first-countable	✓ (Proposition 2.78)	✓ (Proposition 2.89)	✓ (Proposition 3.48)
second-countable	✓ (Proposition 2.78)	✓ (Proposition 2.89)	✓ (Proposition 3.48)
separable	✗ (Example 2.88)	✓ (Proposition 2.89)	✓ (Proposition 3.48)
metrizable	✓ [Examples 2.10 (1)]	✓ (Exercise 2.108)	✓ (Exercise 92)

Table 3.1: Preservation of topological properties:  
separation and countability properties—version 1

figure-eight  
figure-eight

tab:preserve-ver-1

prob:figure-8 **51.** The **figure-eight** is the curve  $E$  in the plane that is the graph of the equation

$$x^2 = 4y^2(1 - y^2)$$

(see Figure 3.11). This curve may be parametrized by

$$\begin{cases} x = \sin 2t, \\ y = \sin t. \end{cases}$$

Show that the figure-eight is homeomorphic to the union of two circles tangent at a single point, for example, the circles with equations

$$x^2 + (y - 1/2)^2 = 1/4, \quad x^2 + (y + 1/2)^2 = 1/4.$$

**52.** (Continuation of Exercise 51.)

(a) Show that, in fact, the maps

$$f: ]0, 2\pi[ \rightarrow E, \quad g: ]-\pi, \pi[ \rightarrow E$$

$$t \mapsto \langle \sin 2t, \sin t \rangle \quad t \mapsto \langle \sin 2t, \sin t \rangle$$

are bijections from the intervals  $]0, 2\pi[$  and  $]-\pi, \pi[$ , respectively, to  $E$ .

figure-eight

Sorgenfreyline!reverse

upper-limit topology

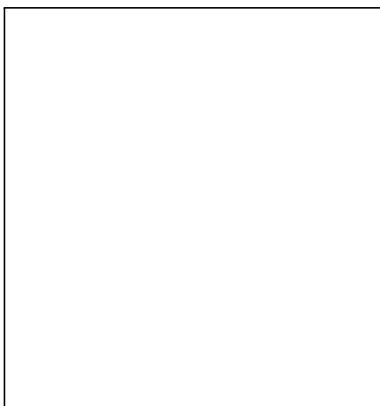


Figure 3.11: The figure-eight.

fig:figure-8

(b) Let  $\mathcal{T}$  and  $\mathcal{S}$  be the topologies on  $E$  that make  $f$  and  $g$ , respectively, homeomorphisms (see [Exercise 48](#)). Is either  $\mathcal{T}$  or  $\mathcal{S}$  the same as usual (Euclidean) topology on  $E$ ? If not, is  $\langle E, \mathcal{T} \rangle$  or  $\langle E, \mathcal{S} \rangle$  homeomorphic to  $E$  with its usual topology?

(c) Is  $\langle E, \mathcal{T} \rangle$  homeomorphic to  $\langle E, \mathcal{S} \rangle$ ?

53. Prove [Proposition 3.24](#).

54. When will two indiscrete spaces be homeomorphic?

55. Provide  $X = \{0, 1, 2\}$  with the topology  $\mathcal{T}$  whose open sets are  $\emptyset$ ,  $\{0, 1\}$ ,  $\{2\}$ , and  $\{0, 1, 2\}$ . What other topologies on the same set make it homeomorphic to  $\langle X, \mathcal{T} \rangle$  and which do not?

prob:Ru-reverse-Sorgenfrey-line

56. The **upper-limit topology** on the set  $\mathbb{R}$  of real numbers is the topology associated with the base consisting of all left-open, right-closed intervals  $]x, y]$ , and the resulting topological space, denoted here by  $\mathbb{R}_u$ , may be called the **reverse Sorgenfrey line**.

(a) Explain the name ‘upper-limit topology’ for this topology in terms of convergence of nets.

(b) Construct a homeomorphism between  $\mathbb{R}_u$  and the Sorgenfrey line  $\mathbb{R}_l$  [[Examples 2.20 \(1\)](#) and [Examples 2.68 \(6\)](#)].

57. Can there be more than one homeomorphism between two homeomorphic spaces? Can there be infinitely many?

58. How does a homeomorphism transform closures, interiors, and boundaries of subsets of its domain?

59. In terms of neighborhoods of points in  $X$  and  $Y$ , when is a bijection  $h: X \rightarrow Y$  between two topological spaces a homeomorphism?

60. Give a homeomorphism  $h: X \cong Y$  and a set  $A \subset X$ , show that  $h$  restricts to homeomorphisms  $A \cong h(A)$  and  $(X \setminus A) \cong (Y \setminus h(A))$ .

61. Prove that each interval in  $\mathbb{R}$  that contains at least two points—including each ray—is homeomorphic to exactly one of the intervals  $[0, 1]$ ,  $[0, 1[$ , and  $]0, 1[$ .

t:homeo-if-homeo-closed-partitions

- 62. (a)** Suppose spaces  $X_1$  and  $X_2$  have partitions  $\{A_1, B_1\}$  and  $\{A_2, B_2\}$ , respectively, into closed subspaces such that  $A_1 \cong A_2$  and  $B_1 \cong B_2$ . Show that then  $X_1 \cong X_2$ . Is the supposition that the sets be closed really needed?

**(b)** Redo (a) but for partitions into open subspaces.

- 63.** Construct a homeomorphism between the closed unit interval  $I$  and the **U**-shaped subspace  $(\{0\} \times I) \cup (I \times \{0\}) \cup (\{1\} \times I)$  of the plane, thereby establishing that the latter space is a 1-cell.

- 64. (a)** Construct a homeomorphism between each pair of the three planar sets that are the graphs of the equations

$$x^2 + y^2 = 1, \quad x^2/a^2 + y^2/b^2 = 1, \quad |x| + |y| = 1,$$

where  $a$  and  $b$  are nonzero constants.

**(b)** Do the same for the graphs of the inequalities

$$x^2 + y^2 \leq 1, \quad x^2/a^2 + y^2/b^2 \leq 1, \quad |x| + |y| \leq 1.$$

**(c)** Do the same for the graphs of the strict inequalities

$$x^2 + y^2 < 1, \quad x^2/a^2 + y^2/b^2 < 1, \quad |x| + |y| < 1.$$

g:graph-continuous-R-to-R-homeo-R

- 65.** Prove that the graph

$$G = \{(x, y) \in \mathbb{R}^2 : y = f(x)\}$$

of a continuous function  $f: \mathbb{R} \rightarrow \mathbb{R}$  is always homeomorphic to the real line  $\mathbb{R}$ . Does this result generalize to continuous functions  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ?

prob:cylinder-annulus-etc

- 66.** Exhibit a homeomorphism from  $X$  to  $Y$  when:

**(a)**  $X$  is the **closed annulus**  $\{x \in \mathbb{R}^2 : 1/2 \leq \|x\| \leq 1\}$  and  $Y$  is the **cylinder**

$$S_1 \times I = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 = 1, 0 \leq x_3 \leq 1\}.$$

(Note that the cylinder  $S_1 \times I$  might more properly be described as a *bounded cylindrical surface*, in contrast to the unbounded cylinder  $S_1 \times \mathbb{R}$ , on the one hand, and the bounded **solid cylinder**  $D_2 \times I$ , on the other hand.)

**(b)**  $X$  is the “punctured plane”  $\mathbb{R}^2 \setminus \{0\}$  and  $Y$  is the unbounded cylinder

$$S_1 \times \mathbb{R} = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 = 1\}.$$

red-plane-and-unbounded-cylinder

prob:bdyJn

- 67.** Verify the claim made in [Examples 3.25 12](#) that the boundary of the centered  $n$ -cube  $J^n$  is the set  $\{x \in \mathbb{R}^n : d_\infty(x, \mathbf{0}) = 1\}$ .

prob:hemispheres-homeo

- 68.** Construct homeomorphism between the two hemispheres

$$S_n^+ = \{x \in S_n : x_{n+1} \geq 0\}, \quad S_n^- = \{x \in S_n : x_{n+1} \leq 0\}$$

(of the  $n$ -sphere  $S_n \subset \mathbb{R}^{n+1}$ ) and the  $n$ -disk  $D_n$ .

prob:open-box-homeo-Bn

- 69.** Show that an “open box”  $]a_1, b_1[ \times ]a_1, b_1[ \times \cdots \times ]a_n, b_n[$  in  $\mathbb{R}^n$  is homeomorphic to the  $n$ -ball  $B_n$ .

prob:count-props-sum-2

- 70.** Show that the Cartesian sum [[Examples 3.47 \(4\)](#)] of two spaces is first-countable, second-countable, or separable, respectively, if each of the spaces has that property.

annulus

cylinder

cylinder!unbounded

cylinder!solid

hemisphere

hemisphere

open box

Cartesian sum!countability properties

71. Must the Cartesian sum of two Lindelöf spaces itself be a Lindelöf space?

72. Show that the standard  $n$ -simplex

$$\Delta_n = \left\{ \langle x_0, x_1, \dots, x_n \rangle \in \mathbb{R}^{n+1} : x_i \geq 0 \text{ for } i = 0, 1, \dots, n, \sum_{i=0}^n x_i = 1 \right\}$$

in  $\mathbb{R}^{n+1}$  is an  $n$ -cell, that is, is homeomorphic to the  $n$ -cube  $I^n$ .

(Hint: You may wish to show first that  $\Delta_n$  is homeomorphic to the subset

$$\Delta_n' = \left\{ \langle y_1, y_2, \dots, y_n \rangle \in \mathbb{R}^n : y_i \geq 0 \text{ for } i = 1, 2, \dots, n, \sum_{i=1}^n y_i = 1 \right\}$$

of  $\mathbb{R}^n$ . Both  $\Delta_n$  and  $\Delta_n'$  are depicted in Figure 3.12 for  $n = 2$ .)



Figure 3.12: The standard 2-simplex in  $\mathbb{R}^3$  (left) and its homeomorph in  $\mathbb{R}^2$  (right).

fig:std-2-simplex

73. Let  $X$  be the subspace of  $\mathbb{R}^3$  obtained by attaching to the sphere  $S_2$  at its north pole a “whisker” pointing outward, and let  $Y$  be obtained by attaching to the same sphere at the same point a whisker pointing inward, as shown in Figure 3.13. Is  $X$



Figure 3.13: Whiskers attached to the 2-sphere outward (left) and inward (right).

fig:2-sphere-with-whiskers

homeomorphic to  $Y$ ?

es-interval-order-top-homeo-to-ray

74. Which, if either, of these topological spaces are homeomorphic to the ray  $[0, +\infty[$ ?(a) The product set  $\mathbb{N} \times [0, 1[$  when provided with its order topology induced by the usual orderings of  $\mathbb{N}$  and  $[0, 1[$ .(b) The product set  $\mathbb{N} \times [0, 1[$  when provided with its product topology.

ordered-sets-and-homeo

75. (a) Prove that an order-isomorphism between two totally ordered sets is a homeomorphism when the sets are provided with their order topologies.

(b) Must a homeomorphism between two totally ordered sets, provided with their order topologies, be an order-isomorphism?

prob:homog

76. A topological space  $X$  is said to be **homogeneous** when, for arbitrary points  $x$  and  $y$  in  $X$ , there exists a homeomorphism  $h: X \rightarrow X$  with  $h(x) = y$ . For example any discrete space is homogeneous.

Show that each of the following spaces is homogeneous:

(a) For arbitrary  $n \geq 1$ , Euclidean  $n$ -space  $\mathbb{R}^n$ (b) For  $n = 1, 2$ , and  $3$ , the  $(n - 1)$ -sphere  $S_{n-1}$ . (Actually,  $S_{n-1}$  is homogeneous for all  $n \geq 1$ .)

(c) Any space that is homeomorphic to a homogeneous space. (Thus homogeneity, that is, being homogeneous, is a topological property.)

*Note:* One example of a space that is *not* homogeneous is the unit interval  $[0, 1]$ . The proof for  $[0, 1]$ , requested in [Exercise 5.14](#), involves “connectedness.” For more such examples, see the note for [Exercise 78](#).

prob:Rn-strongly-locally-homog

77. (Continuation of [Exercise 76](#).) Let  $B = ]a_1, b_1[ \times ]a_1, b_1[ \times \cdots \times ]a_n, b_n[$  be an “open box” in  $\mathbb{R}^n$  and let  $u, v \in B$ . Construct a homeomorphism  $h: \mathbb{R}^n \cong \mathbb{R}^n$  with  $h(u) = v$  and  $h(x) = x$  for all  $x \notin B$ .*Note:* In particular, then each such open box is a homogeneous space.*[Hint:* First show why it suffices to construct, for each index  $j$  and for each two points  $u, v \in B$  with  $u_i = v_i$  for all  $i \neq j$ , a homeomorphism  $h_j$  with  $h_j(u) = v$ . Then for such  $j$  and  $p$  and  $q$ , make the  $j$ th coordinate function  $(h_j)_j$  map  $]a_j, b_j[$  piecewise linearly onto itself.]

prob:n-ball-homog

78. (Continuation of [Exercise 76](#).) Show that the  $n$ -ball  $B_n$  is homogeneous by each of the following methods:(a) By using [Exercises 69](#) and [77](#).(b) Directly: given two points  $u, v \in B_n$ , construct explicitly a homeomorphism of  $B_n$  sending  $u$  to  $v$ . In fact, construct a homeomorphism of  $\mathbb{R}^n$  that maps  $D_n$  homeomorphically onto itself, sends  $u$  to  $v$ , and sends each  $x \in \mathbb{R}^n \setminus B_n$  to itself. (*Hint:* It is enough to construct, for each  $z \in B_n$ , such a homeomorphism sending the origin to  $z$ .)*Note:* By contrast, the  $n$ -disk  $D_n$  is *not* homogeneous. The proof for  $n = 1$  involves the same consideration of “connectedness” mentioned in the note for [Exercise 76](#). For  $n = 2$ , the proof involves the notion of “simple connectedness,” which is discussed in [Section 5.5](#); see [Exercise 5.114](#). For  $n \geq 3$ , the proof involves more advanced machinery.

79. Is the relation “is locally homeomorphic to” an equivalence relation among topological spaces?

order isomorphism!and homeomorp

order topology

homogeneous space

homogeneous space

topological space!homogeneous

box!open

open box

homogeneous space!n-ball@and \$n

n-ball@\$n\$-ball!homogeneous spac

n-disk@\$n\$-disk!homogeneous spac

homogeneous space

- prob:S<sub>1</sub>-not-homeo-to-R 80. You know that the 1-sphere (the unit circle)  $S_1$  is locally homeomorphic to the real line  $\mathbb{R}$ . Show, however, that  $S_1$  is *not* homeomorphic to  $\mathbb{R}$ .  
 spherical coordinates
- origins-loc-euclidean-not-manifold 81. Show that the line with two origins [Examples 2.20 (3)] is locally Euclidean.  
 order-dense set  
 longline  
 longline  
 longline  
 n-manifold@ $S^n$ -manifold!and open subset  
 prob:paths-Haus-S<sub>1</sub> 82. (a) Verify that if  $U^+ = \{\langle x, y \rangle \in S_1 : x > 0\}$ , then the map  
 open annulus  
 annulus
- $$f^+ : U^+ \rightarrow ]-1, 1[$$
- $$\langle x, y \rangle \mapsto y$$
- defines a chart in the circle  $S_1$ . Construct additional charts on  $S_1$  that, together with  $\langle U^+, f^+ \rangle$ , form an atlas on  $S_1$ . Then calculate the transition map between each pair of charts in this atlas.
- (b) Generalize the construction in (a) so as to obtain an atlas on the  $n$ -sphere  $S_n$ . Then calculate the transition map between each pair of charts in this atlas.
- prob:spherical-coords 83. (a) Use spherical coordinates  $\langle \rho, \varphi, \theta \rangle$  so as to construct an atlas on the sphere  $S_2$ .  
 (b) Use spherical coordinates  $\langle \rho, \varphi, \theta \rangle$  so as to construct an atlas on the largest open subspace of  $\mathbb{R}^3$  as possible. [Compare Examples 3.36 (3).]
- prob:long-line-order-dense 84. Show that the long line  $L^*$ , with its lexicographic ordering, is order-dense (Definition 0.73).  
 85. (a) According to Proposition 3.42, the long line  $L^*$  is not second-countable. Establish the stronger result that  $L^*$  is not separable.  
 (b) Is  $L^*$  first-countable?  
 86. (a) Show that each open interval in  $L^*$  is homeomorphic to the open unit interval  $]0, 1[$ .  
 (b) Show that each closed interval in  $L^*$  is homeomorphic to the closed unit interval  $[0, 1]$ .
- prob:0-manifold 87. The definition of ' $n$ -manifold' included the case  $n = 0$ . Which topological spaces are 0-dimensional manifolds? (Hint: Recall that  $\mathbb{R}^0 = \{0\}$ .)
- open-subset-of-manifold-is-manifold 88. Prove that an open subset of an  $n$ -dimensional manifold is itself an  $n$ -dimensional manifold.  
 Note: This fact immediately provides many interesting examples of  $n$ -manifolds, one of which is the **open annulus**  $\{x \in \mathbb{R}^2 : 1/2 < \|x\| < 1\}$ .  
 89. Show that a Hausdorff space is an  $n$ -manifold if each of its points has some open neighborhood that is homeomorphic to some open subset of  $\mathbb{R}^n$ .  
 90. Is the given topological space a topological manifold and, if so, what is its dimension?  
 (a) A line in  $\mathbb{R}^n$ .  
 (b) The solution set (in  $\mathbb{R}^2$ ) of the equation  $y = |x|$ .  
 (c) The solution set (in  $\mathbb{R}^2$ ) of the equation  $|y| = |x|$ .  
 (d) The cylinder  $S_1 \times I$ .  
 (e) The unbounded cylinder  $S_1 \times \mathbb{R}$ .

- (f) The solution set (in  $\mathbb{R}^3$ ) of the equation  $z = x^2 + y^2$ .
- (g) The solution set (in  $\mathbb{R}^3$ ) of the equation  $z^2 = x^2 + y^2$ .
- (h) The torus  $T_2 = S_1 \times S_1$  [Examples 3.25 (15)].
- (i) The **3-torus**  $T_3 = S_1 \times S_1 \times S_1$ .

prob:manifold-with-bdy

- 91.** A second-countable Hausdorff space is said to be an  **$n$ -dimensional manifold-with-boundary**, or simply an  **$n$ -manifold-with-boundary**, when each of its points has either an open neighborhood that is homeomorphic to an open set in  $\mathbb{R}^n$  or an open neighborhood that is homeomorphic to an open set in the subspace

$$\mathbb{R}_+^n = \{x \in \mathbb{R}^n : x_n \geq 0\}$$

of  $\mathbb{R}^n$ .

(Note: That a point cannot have open neighborhoods of both kinds is guaranteed by the **Invariance of Domain Theorem**, which asserts that a subset of  $\mathbb{R}^n$  is open in  $\mathbb{R}^n$  if it is homeomorphic to an open subset of  $\mathbb{R}^n$ . Of course a neighborhood homeomorphic to an open set in  $\mathbb{R}^n$  is already homeomorphic to an open set in  $\mathbb{R}_+^n$ .)

Verify that each of these spaces is an  $n$ -manifold-with-boundary:

- (a) Any  $n$ -manifold.
- (b) The  $n$ -cube  $I^n$ .
- (c) The  $n$ -disk  $D_n$ .
- (d) The cylinder  $S_1 \times I$  of Exercise 66 (a); here  $n = 2$ .
- (e) The standard  $n$ -simplex  $\Delta_n$  (Exercise 72).
- (f) An  $n$ -cell [Examples 3.25 (12)].
- (g) The standard  $n$ -simplex  $\Delta_n$  (Exercise 72).
- (h) A closed-halfspace in  $\mathbb{R}^n$  (Exercise 1.80).
- (i) The **solid torus**  $D_2 \times S_1$ .

rt:n-cell-is-manifold-with-boundary  
ob-part:n-cell-as-manifold-with-bdy

prob:sum-2-preserve-metrizability  
:prove-sum-2-preserve-countability

- 92.** Prove that the Cartesian sum of two metrizable spaces is itself metrizable.
- 93.** (a) Prove that the Cartesian sum of two second-countable spaces is second-countable. [This is Proposition 3.48 (2)].
- (b) Prove that the Cartesian sum of two separable spaces is separable. [This is Proposition 3.48 (3)].
- (c) Must the Cartesian sum of two Lindelöf spaces (Exercise 2.115) be a Lindelöf space?

rob:separation-properties-sum-of-2

- 94.** (a) Prove that the Cartesian sum of two  $T_0$ -,  $T_1$ -,  $T_2$ -, or regular spaces has the same property. (This is Proposition 3.49.).
- (b) Is the Cartesian sum of two normal spaces necessarily normal?

prob:T0-T1-reg-top-props

- 95.** Show that each of the properties  $T_0$ ,  $T_1$ , and regularity is a topological property.
- 96.** Which of the following are topological properties?
- (a) Having a denumerable set of points.
  - (b) Having some complete metric inducing the topology.
  - (c) Zero-dimensionality.
  - (d) Being a Lindelöf space (Exercise 2.115).

torus  
n-torus@\$n\$-torus  
torus  
n-torus@\$n\$-torus  
manifold-with-boundary  
n-manifold-with-boundary@\$n\$-ma  
invariance of domain  
n-cube@\$n\$-cube!and \$n\$-manifold  
n-disk@\$n\$-disk!and \$n\$-manifold-  
n-simplex@\$n\$-simplex  
n-simplex@\$n\$-simplex!standard  
n-cell@\$n\$-cell!as manifold-with-bo  
n-simplex@\$n\$-simplex  
n-simplex@\$n\$-simplex!standard  
closed-halfspace!and \$n\$-manifold-  
solid torus  
countability properties!and Cartesia  
Lindelof space@Lindel\"of space

- rob:bdy-of-manifold-with-boundary 97. (Continuation of Exercise 91.) The **boundary**  $\partial M$  of an  $n$ -manifold-with-boundary  $M$  is the set of those points  $x \in M$  that do *not* have any open neighborhood homeomorphic to an open subset of  $\mathbb{R}^n$ .
- chart (a) What is  $\partial M$  when  $M = D_n$ , the  $n$ -disk?
- n-manifold@n-manifold (b) According to Exercise 91 (e), an  $n$ -cell  $E$  is a manifold-with-boundary. What, then, is  $\partial E$ ?
- half-disk space (c) Give an example for  $n = 2$  of such an  $M$  that is a subspace of  $\mathbb{R}^n$  and for which  $\partial M$  is distinct from the boundary  $\text{bdy } M$  of  $M$  in the topological space  $\mathbb{R}^n$ .
- Cartesian sum!of a family of spaces (d) Prove that, in general,  $\partial M$  is closed in an  $n$ -manifold-with-boundary.
- Cartesian sum!of a family of spaces (e) Prove that the boundary  $\partial M$  of an  $n$ -manifold-with-boundary is an  $(n - 1)$ -manifold-with-boundary.
- ob-part:interior-mfld-w-bdy-is-mfld (f) Explain why the subspace  $M \setminus \partial M$  of an  $n$ -manifold-with-boundary  $M$  is an  $n$ -manifold.
- prob:chart-codomain- $\mathbb{R}^n$  98. Let  $\langle U, \varphi \rangle$  be a chart at  $p$  in an  $n$ -dimensional locally Euclidean space  $X$ . Show that there is another chart  $\langle V, \psi \rangle$  at  $p$  such that  $V \subset U$ ,  $\varphi(V) = \mathbb{R}^n$ , and  $\psi(p) = \mathbf{0}$ .
99. Show that being an  $n$ -dimensional manifold is a topological property. Do the same for being an  $n$ -dimensional manifold-with-boundary (Exercise 91).
100. Use a suitable topological property to demonstrate that  $X$  is not homeomorphic to  $Y$ :
- (a)  $X = \mathbb{R}$  and  $Y = \mathbb{Q}$  each with its usual topology.
  - (b)  $X = \mathbb{R}$  with its usual topology and  $Y = \mathbb{R}$  with its right-interval topology, that is,  $Y$  is the Sorgenfrey line [Examples 2.20 (1)].
  - (c)  $X = \mathbb{R}$  with its usual topology and  $Y = \mathbb{R}$  with its finite-complement topology [Examples 2.3 (7)].
  - (d)  $X = \ell^2$ , the Hilbert sequence space (Example 1.10) and  $Y =$  the set of all functions  $f: \mathbb{R} \rightarrow \mathbb{R}$  with its topology of pointwise convergence [Examples 2.72 (8)].
101. If  $f: X \rightarrow Y$  is a continuous bijection between topological spaces  $X$  and  $Y$  that are homeomorphic, must  $f$  itself be a homeomorphism?
102. Are there topological spaces  $X$  and  $Y$  and continuous bijections  $f: X \rightarrow Y$  and  $g: Y \rightarrow X$  such that neither  $f$  nor  $g$  is a homeomorphism?
- prob:top-analog-cantor-bernstein 103. If each of two topological spaces can be homeomorphically embedded in the other, must they be homeomorphic to one another?
- [Note: This is asking whether there is a topological analog of the purely set-theoretic Cantor-Bernstein Theorem (0.126), according to which there is a bijection between two sets when there are injections of each into the other.]
104. (a) Embed the Euclidean plane  $\mathbb{R}^2$  in the half-disk space  $H \cup L$  of Examples 2.20 (3).  
(b) Can  $H \cup L$  be embedded in  $\mathbb{R}^2$ ?
105. Show that a space  $X$  that, from a topological point of view, “looks like” a subspace of a space  $Y$  actually *is* a subspace of a space that “looks like”  $Y$ . In other words, if  $h: X \rightarrow Y$  is an embedding, then construct a topological space  $Z$  that contains  $X$  as a subspace and is homeomorphic to  $Y$ . [Hint: Replace  $h(X)$  by  $X$ .]
- prob:cartesian-sum 106. Let  $\langle X_i \rangle_{i \in I}$  be a family of topological spaces.



- (a) Construct a topological space  $X$  and embeddings  $h_i: X_i \rightarrow X$  for all  $i \in I$  such that  $X = \bigcup_{i \in I} h(X_i)$  and  $h_i(X_i)$  is open in  $X$  for each  $i$ .
- (b) Show that such an  $X$  is unique up to homeomorphism in the following sense. If  $Y$  is a topological space and  $k_i: X_i \rightarrow Y$  are embeddings with  $k_i(X_i)$  open in  $Y$  for each  $i$ , then there is a *unique* homeomorphism  $f: X \cong Y$  such that  $f \circ h_i = k_i$  for each  $i$ .

countability properties!and Cartesian  
 Cartesian sum!of a family of spaces  
 separation properties!and Cartesian  
 coffee cup  
 doughnut  
 product space

Such a space is called the **(Cartesian) sum of  $\langle X_i \rangle_{i \in I}$**  and is denoted by

$$\bigoplus_{i \in I} X_i.$$

(Some mathematicians prefer the notation  $\bigsqcup_{i \in I} X_i$  or  $\coprod_{i \in I} X_i$  for Cartesian sum.)

[*Note:* This generalizes both the sum of a family of disjoint spaces as constructed in [Exercise 2.18](#) and the sum of two disjoint spaces as constructed in [Examples 3.47 \(4\)](#).]

prob:cont-map-on-sum **107.** (Continuation of [Exercise 106](#).) Prove that a map

$$f: \bigoplus_{i \in I} X_i \rightarrow Y$$

from a Cartesian sum to a topological space  $Y$  is continuous precisely when the maps  $f \circ h_i: X_i \rightarrow Y$  are continuous for all  $i \in I$ .

prob:sum-metrizable **108.** (Continuation of [Exercise 106](#).) Show that the Cartesian sum of an arbitrary family of topological spaces is metrizable if and only if each of the spaces is metrizable. (*Hint:* You may want to use [Proposition 1.40](#) and [Exercise 1.49](#).)

countability-properties-sum-family **109.** (Continuation of [Exercise 106](#).) Which of the countability properties are preserved under formation of the Cartesian sum of an arbitrary family of topological spaces?

b:separation-properties-sum-family **110.** (Continuation of [Exercise 106](#).) Which of the separation properties are preserved under formation of the Cartesian sum of an arbitrary family of topological spaces?

prob:coffee-cup-and-doughnut **111.** Discuss: “A topologist is somebody who cannot tell the difference between a coffee cup and a doughnut.”

### 3.3 Product Spaces

sec:product

The theme of this section and the next is the construction of new spaces from old. We look first at product spaces.

[Examples 2.72 \(5\)](#) showed a way to put a topology on the product  $X \times Y$  of the underlying sets of *two* topological spaces, namely, by taking as base the collection

$$\{U \times V : U \text{ is open in } X \text{ and } V \text{ is open in } Y\}$$

of products of open sets in the respective factors  $X$  and  $Y$ .

Suppose we are given an arbitrary number of topological spaces constituting a family  $\langle X_i \rangle_{i \in I}$ . How can we put a suitable topology on the product  $\times_{i \in I} X_i$  of their underlying sets? Just because this product is a set, several topologies are already at our disposal—the discrete topology, the indiscrete topology, the finite-complement topology, and so forth—but

box topology

metric! product@on product

these not be related in any way to the given topologies on the given sets. We could try to generalize the case of two spaces by forming the collection

$$\left\{ \bigtimes_{i \in I} U_i : U_i \text{ is open in } X_i \text{ for each } i \in I \right\},$$

which is, indeed, the base of a topology on  $\times_{i \in I} X_i$ , called the *box topology*—see [Exercise 140](#). However, it turns out that this is *not* the “natural” topology to use! To see what to do, we turn once again to metric spaces for inspiration.

### Topology of the product of countably many metric spaces

We have met two general instances of metrics on products. The first concerns finitely many metric spaces  $\langle X_1, d_1 \rangle, \langle X_2, d_2 \rangle, \dots, \langle X_n, d_n \rangle$ . Recall from [Example 1.14](#) that on the product set  $X = X_1 \times X_2 \times \cdots \times X_n$  the max metric  $d_\infty$  induced by  $\langle d_1, d_2, \dots, d_n \rangle$  is defined by

$$d_\infty(x, y) = \max_{1 \leq i \leq n} d_i(x_i, y_i).$$

**3.50 Proposition.** *Given finitely many metric spaces  $\langle X_1, d_1 \rangle, \langle X_2, d_2 \rangle, \dots, \langle X_n, d_n \rangle$ , the topology on the product set*

$$X = X_1 \times X_2 \times \cdots \times X_n$$

*induced by the max metric  $d_\infty$  has as a base the collection*

$$\mathcal{B} = \{U_1 \times U_2 \times \cdots \times U_n : U_i \text{ is } d_i\text{-open in } X_i \text{ for each } i = 1, 2, \dots, n\}$$

*of all products of open sets in the factors  $X_1, X_2, \dots, X_n$ .*

**Proof.** We shall use the criterion of [Proposition 2.69](#).

First we show that each  $B \in \mathcal{B}$  is  $d_\infty$ -open. Let  $B \in \mathcal{B}$ . For  $B$  to be  $d_\infty$ -open, it suffices to show that it contains a  $d_\infty$ -ball at each of its points. Write

$$B = U_1 \times U_2 \times \cdots \times U_n$$

with each  $U_i$  being a  $d_i$ -open subset of  $X_i$ . Let  $x = \langle x_1, x_2, \dots, x_n \rangle \in B$ . For each  $i = 1, 2, \dots, n$ , the  $i$ th coordinate  $x_i \in U_i$ , and so

$$B_{\varepsilon_i}(x_i; d_i) \subset U_i$$

for some  $\varepsilon_i > 0$ . Set

$$\varepsilon = \min\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}.$$

Then  $B_\varepsilon(x; d_\infty) \subset B$ . In fact, if  $y = \langle y_1, y_2, \dots, y_n \rangle \in B_\varepsilon(x; d_\infty)$ , then

$$d_i(x_i, y_i) \leq d_\infty(x, y) < \varepsilon \leq \varepsilon_i$$

for each  $i = 1, 2, \dots, n$  and hence

$$\begin{aligned} y = \langle y_1, y_2, \dots, y_n \rangle &\in B_{\varepsilon_1}(x_1; d_1) \times B_{\varepsilon_2}(x_2; d_2) \times \cdots \times B_{\varepsilon_n}(x_n; d_n) \\ &\subset U_1 \times U_2 \times \cdots \times U_n = B. \end{aligned}$$

Now let  $U$  be an arbitrary  $d_\infty$ -open set in  $X$  and let  $x \in U$ . We show that  $x \in B \subset U$  for some  $B \in \mathcal{B}$ . Choose  $\varepsilon > 0$  with

$$B_\varepsilon(x; d_\infty) \subset U.$$

For each  $i = 1, 2, \dots, n$ , the ball  $B_\varepsilon(x_i; d_i)$  is a  $d_i$ -open subset of  $X_i$ , and so the set

$$B = B_\varepsilon(x_1; d_1) \times B_\varepsilon(x_2; d_2) \times \cdots \times B_\varepsilon(x_n; d_n)$$

belongs to  $\mathcal{B}$ . Certainly  $x \in B$ . Finally,  $B \subset U$  because if  $y = \langle y_1, y_2, \dots, y_n \rangle \in X$  with  $d_i(x_i, y_i) < \varepsilon$  for each  $i$ , then  $d_\infty(x, y) < \varepsilon$ , too.  $\square$

In the notation of [Proposition 3.50](#), a basic open set

$$B = U_1 \times U_2 \times \cdots \times U_n$$

—a member of the base  $\mathcal{B}$ —is the intersection of the  $n$  subsets

$$p_k^{-1}(U_k) = X_1 \times X_2 \times \cdots \times X_{k-1} \times U_k \times X_{k+1} \times \cdots \times X_n$$

for  $1 \leq k \leq n$ . Here  $p_k^{-1}(U_k)$  is the inverse image of the  $d_k$ -open subset  $U_k$  of  $X_k$  under the  $k$ th projection map

$$p_k: X_1 \times X_2 \times \cdots \times X_n \rightarrow X_k \\ \langle x_1, x_2, \dots, x_n \rangle \mapsto x_k.$$

Hence the topology induced by the max metric  $d_\infty$  on the product  $X = X_1 \times X_2 \times \cdots \times X_n$  of  $n$  metric spaces  $\langle X_1, d_1 \rangle, \langle X_2, d_2 \rangle, \dots, \langle X_n, d_n \rangle$  has as a *subbase* the collection

$$\begin{aligned} \mathcal{S} &= \{ X_1 \times X_2 \times \cdots \times X_{k-1} \times U_k \times X_{k+1} \times \cdots \times X_n : \\ &\quad 1 \leq k \leq n \text{ and } U_k \text{ is } d_k\text{-open in } X_k \} \\ &= \{ p_k^{-1}(U_k) : 1 \leq k \leq n \text{ and } U_k \text{ is } d_k\text{-open in } X_k \} \end{aligned}$$

The second instance of a metric on a product concerns denumerably many metric spaces  $\langle X_1, d_1 \rangle, \langle X_2, d_2 \rangle, \langle X_3, d_3 \rangle, \dots$ . Without loss of generality we may assume that each of the metric spaces is bounded and of diameter at most 1. In fact, if  $\langle X_i, d_i \rangle$  is not bounded, or bounded but of diameter strictly greater than 1, just replace  $d_i$  by the metric  $d_i^*$  given by  $d_i^*(x, y) = \min\{d_i(x, y), 1\}$ ; according to [Proposition 1.40](#), the metrics  $d_i^*$  and  $d_i$  induce the same topology on  $X_i$ .

Recall from [Exercise 1.51](#) that the formula

$$\text{\{eq:dprime-prod-formula\}} \quad (*) \quad d'(x, y) = \sum_{i=1}^{\infty} \frac{d_i(x_i, y_i)}{2^i}$$

defines a metric  $d'$  on the product set

$$X = \bigtimes_{i=1}^{\infty} X_i.$$

Things are not quite as simple now as they were with finitely many factors, for even a product of balls need not be  $d'$ -open!

p:base-dprime-denumerable-product **3.51 Proposition.** *Given a sequence  $\langle X_1, d_1 \rangle, \langle X_2, d_2 \rangle, \langle X_3, d_3 \rangle, \dots$  of metric spaces each bounded and of diameter at most 1, the topology induced by the metric  $d'$ , given by  $(*)$ , on the product set*

$$X = \bigtimes_{i=1}^{\infty} X_i$$

*has as a base the collection*

$$\mathcal{B} = \{ U_1 \times U_2 \times U_n \times X_{n+1} \times X_{n+2} \times \cdots : \\ n \geq 1, U_i \text{ is } d_i\text{-open in } X_i \text{ for each } i = 1, 2, \dots, n \}$$

*of products.*

**Proof.** Again we use [Proposition 2.69](#).

First we show that each  $B \in \mathcal{B}$  is  $d'$ -open. Let  $B \in \mathcal{B}$  and let  $x \in B$ . We need only show that  $B_\varepsilon(x; d') \subset B$  for some  $\varepsilon > 0$ . Write

$$B = U_1 \times U_2 \times \cdots \times U_n \times X_{n+1} \times X_{n+2} \times \cdots$$

with  $n \geq 1$  and  $U_i$  a  $d_i$ -open set in  $X_i$  for each  $i = 1, 2, \dots, n$ . For each  $i = 1, 2, \dots, n$  there is some  $\varepsilon_i > 0$  with

$$B_{\varepsilon_i}(x_i; d_i) \subset U_i.$$

Define

$$\varepsilon = \min \left\{ \frac{\varepsilon_1}{2}, \frac{\varepsilon_2}{2^2}, \dots, \frac{\varepsilon_n}{2^n} \right\}.$$

Then  $B_\varepsilon(x; d') \subset B$ . In fact, if  $y \in B_\varepsilon(x; d')$ , then

$$\frac{d_i(x_i, y_i)}{2^i} \leq d'(x, y) < \varepsilon \leq \frac{\varepsilon_i}{2^i},$$

$d_i(x_i, y_i) < \varepsilon_i$  for each  $i = 1, 2, \dots, n$ , and so

$$y \in B_{\varepsilon_1}(x_1; d_1) \times B_{\varepsilon_2}(x_2; d_2) \times \cdots \times B_{\varepsilon_n}(x_n; d_n) \times X_{n+1} \times X_{n+2} \times \cdots \subset B.$$

Now let  $U$  be any  $d'$ -open set in  $X$  and let  $x \in U$ . We show that  $x \in B \subset U$  for some  $B \in \mathcal{B}$ . Choose any  $\varepsilon > 0$  with

$$B_\varepsilon(x; d') \subset U$$

and then choose  $n \geq 1$  with

$$\sum_{i=n+1}^{\infty} \frac{1}{2^i} < \frac{\varepsilon}{2}.$$

The set

$$B = B_{\varepsilon/2n}(x_1; d_1) \times B_{\varepsilon/2n}(x_2; d_2) \times \cdots \times B_{\varepsilon/2n}(x_n; d_n) \times X_{n+1} \times X_{n+2} \times \cdots$$

belongs to  $\mathcal{B}$ . Clearly  $x \in B$ . Finally,  $B \subset U$  because if  $y \in B$  with  $d_i(x_i, y_i) < \varepsilon/2n$  for each  $i = 1, 2, \dots, n$ , then

$$\begin{aligned} d'(x, y) &= \sum_{i=1}^n \frac{d_i(x_i, y_i)}{2^i} + \sum_{i=n+1}^{\infty} \frac{d_i(x_i, y_i)}{2^i} \\ &\leq \sum_{i=1}^n d_i(x_i, y_i) + \sum_{i=n+1}^{\infty} \frac{1}{2^i} \\ &< n \cdot \frac{\varepsilon}{2n} + \frac{\varepsilon}{2} = \varepsilon. \quad \square \end{aligned}$$

In the notation of [Proposition 3.51](#), a basic open set

$$B = U_1 \times U_2 \times \cdots \times U_n \times X_{n+1} \times X_{n+2} \times \cdots$$

—a member of the base  $\mathcal{B}$ —is the intersection of the  $n$  subsets

$$p_k^{-1}(U_k) = X_1 \times X_2 \times \cdots \times X_{k-1} \times U_k \times X_{k+1} \times \cdots \times X_n \times \cdots$$

for  $1 \leq k \leq n$ . Here  $p_k^{-1}(U_k)$  is the inverse image of the  $d_k$ -open subset  $U_k$  of  $X_k$  under the  $k$ th projection map

$$\begin{aligned} p_k: X_1 \times X_2 \times X_3 \times \cdots &\rightarrow X_k \\ \langle x_1, x_2, x_3, \dots \rangle &\mapsto x_k. \end{aligned}$$

Hence the topology induced by the metric  $d'$ , given by (\*), on the product  $X = \times_{i=1}^{\infty} X_i$  of a sequence of metric spaces  $\langle X_i, d_i \rangle$  has as a *subbase* the collection

$$\begin{aligned} \mathcal{S} &= \{ X_1 \times X_2 \times \cdots \times X_{k-1} \times U_k \times X_{k+1} \times X_{k+2} \times \cdots : \\ &\quad k \geq 1 \text{ and } U_k \text{ is } d_k\text{-open in } X_k \} \\ &= \{ p_k^{-1}(U_k) : k \geq 1 \text{ and } U_k \text{ is } d_k\text{-open in } X_k \} \end{aligned}$$

The bases  $\mathcal{B}$  in both Proposition 3.50 and Proposition 3.51 can be described in the same way with the aid of some unifying terminology.

**3.52 Definition.** Given a set  $I$ , a property that is meaningful for elements  $i \in I$  will be said to hold for **almost all**  $i \in I$  when it holds for all except a finite number (possibly zero) of values of the index  $i \in I$ , in other words, when the set of  $i \in I$  for which it does *not* hold is finite (and possibly empty).

In both the finite case (Proposition 3.50) and the denumerable case (Proposition 3.51), the product set  $X$  has the form

$$X = \bigtimes_{i \in I} X_i$$

where either  $I = \{1, 2, \dots, n\}$  or else  $I = \{1, 2, 3, \dots\}$ . Then in both cases the base  $\mathcal{B}$  of the metric topology on  $X$  may be described, in the language of Definition 3.52, as

$$\mathcal{B} = \{ \bigtimes_{i \in I} U_i : U_i \text{ is open in } X_i \text{ for all } i \in I, \\ U_i = X_i \text{ for almost all } i \in I \}$$

(In the finite case  $I = \{1, 2, \dots, n\}$ , the ‘almost all’ clause is superfluous.) In both the finite case and the denumerable case, moreover, the metric topology on the product set  $X = \times_{i \in I} X_i$  has as a subbase the collection

$$\mathcal{S} = \{ p_j^{-1}(U) : j \in I \text{ and } U \text{ is open in } X_j \}.$$

**Note.** For the sake of concreteness, in the remainder of this section you may wish to think of the index set  $I$  as a nonempty set of natural numbers—in fact, either as  $\{1, 2, \dots, n\}$  for some positive integer  $n$  or else as  $\mathbb{N}^* = \{1, 2, 3, \dots\}$ . Nonetheless, *everything that follows about topologies of products, unless otherwise indicated explicitly, applies to an arbitrary index set  $I$ .*

### The product topology

Suppose now we are given a family  $\langle X_i \rangle_{i \in I}$  of topological spaces, but with the topologies on the  $X_i$  not necessarily induced by metrics. We wish to endow the product set  $\times_{i \in I} X_i$  with a topology that is “naturally” related to the topologies on the individual spaces  $X_i$ . To do so, we are guided by the unified expression (††) for the subbase  $\mathcal{S}$  in the metric cases described in the preceding subsection.

**3.53 Lemma.** Let  $\langle X_i \rangle_{i \in I}$  be a family of topological spaces. Then the collection

$$\mathcal{S} = \{ p_j^{-1}(U) : j \in I \text{ and } U \text{ is open in } X_j \}$$

covers the product set  $\times_{i \in I} X_i$  and hence is a subbase of a topology on  $\times_{i \in I} X_i$ . Here for each  $j \in I$ , the map  $p_j: \times_{i \in I} X_i \rightarrow X_j$  is the  $j$ th projection.

**Proof.** If  $I$  is empty, there is nothing to show, so suppose  $I$  is nonempty. Let  $X = \langle X_i \rangle_{i \in I} \in \prod_{i \in I} X_i$  be arbitrary. Choose some  $j \in I$ . Then  $p_j(x) = x_j \in X_j$ , and  $x_j \in p_j^{-1}(X_j)$ . But  $X_j$  is open in  $X_j$ . So  $x \in p_j^{-1}(X_j)$ . That the cover  $\mathcal{S}$  is a subbase of a topology on the product set is an immediate consequence of Proposition 2.92.  $\square$

def:product-space

**3.54 Definition.** The **product** of a family  $\langle X_i \rangle_{i \in I}$  of topological spaces is the topological space whose underlying set is  $X = \prod_{i \in I} X_i$  and whose topology has as a subbase the collection

$$\mathcal{S} = \{p_j^{-1}(U) : j \in I \text{ and } U \text{ is open in } X_j\}$$

where, for each  $j \in I$ , the map  $p_j : \prod_{i \in I} X_i \rightarrow X_j$  is the  $j$ th projection. This topology is called the **product topology**, and members of  $\mathcal{S}$  are referred to as **subbasic open sets**. (Such sets are also sometimes called “open cylinder sets.”) For each  $i \in I$ , the individual topological space  $X_i$  is called the  **$i$ th factor space** of the product space.

*Unless otherwise indicated, any future reference to a topology on a product of topological spaces is to the product topology.*

For a family  $\langle X_i \rangle_{i \in I}$  of topological spaces, for an index  $j \in I$ , and an open set  $U$  of  $X_j$ , a typical subbasic open set—a member of the cover  $\mathcal{S}$  that generates the product topology—has the form

$$p_j^{-1}(U) = \bigtimes_{i \in I} U_i$$

where for each  $i \in I$ ,

$$U_i = \begin{cases} U & \text{if } i = j, \\ X_i & \text{if } i \neq j. \end{cases}$$

If now  $J$  is a finite subset of the index set  $I$  and if  $\langle U_j \rangle_{j \in J}$  is a family of sets with  $U_j$  open in  $X_j$  for each  $j \in J$ , then

$$\bigcap_{j \in J} p_j^{-1}(U_j) = \bigtimes_{i \in I} U_i$$

where

$$U_i = X_i \text{ if } i \in I \setminus J.$$

These observations yield the following alternative description of the product topology, in terms of a base rather than a subbase.

prop:prod-base

**3.55 Proposition.** Let  $\langle X_i \rangle_{i \in I}$  be a family of topological spaces. Then the collection

$$\mathcal{B} = \left\{ \bigtimes_{i \in I} U_i : \begin{array}{l} U_i \text{ is open in } X_i \text{ for all } i \in I, \\ U_i = X_i \text{ for almost all } i \in I \end{array} \right\}$$

is a base of a topology of the product topology on  $X = \prod_{i \in I} X_i$ , namely, the base associated with the subbase  $\{p_j^{-1}(U) : j \in I \text{ and } U \text{ is open in } X_j\}$ .

def:prod-basic-open-set

**3.56 Definition.** For a family  $\langle X_i \rangle_{i \in I}$  of topological spaces, a member of the base  $\mathcal{B}$  of the preceding proposition is referred to as a **basic open set** in the product space  $\prod_{i \in I} X_i$ .

The most prominent product space is  $n$ -dimensional Euclidean space  $\mathbb{R}^n$  with its usual topology. As a set,  $\mathbb{R}^n$  is the product of  $n$  copies of the real line  $\mathbb{R}$ . According to [Proposition 3.50](#), the product topology on  $\mathbb{R}^n$  (arising from the usual topology on each of these copies of  $\mathbb{R}$ ) is induced by the max metric, and we know that the max metric induces the usual topology on  $\mathbb{R}^n$ .

When the index set  $I$  is finite, say

$$I = \{1, 2, \dots, n\}$$

for some integer  $n \geq 1$ , the basic open sets in a product space

$$\bigtimes_{i \in I} X_i = \bigtimes_{i=1}^n X_i = X_1 \times X_2 \times \cdots \times X_n$$

are all the products

$$\bigtimes_{i=1}^n U_i = U_1 \times U_2 \times \cdots \times U_n$$

of open sets  $U_1, U_2, \dots, U_n$  in the respective factor spaces  $X_1, X_2, \dots, X_n$ . In particular, when the index set consists of just two elements, say  $I = \{1, 2\}$ , we are dealing with the product  $X_1 \times X_2$  of two topological spaces  $X_1$  and  $X_2$ , and the basic open sets are all the products  $U_1 \times U_2$  with  $U_1$  open in  $X_1$  and  $U_2$  open in  $X_2$ .

**Caution!** Even when there are only two factors, *an open set in a product space need not be a product of open sets* in the factors—that is, need not be a basic open set. For example, the unit ball  $B_2$  in  $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$  is open, but  $B_2 \neq U_1 \times U_2$  for any sets  $U_1 \subset \mathbb{R}, U_2 \subset \mathbb{R}$ . In fact, the points  $\langle 3/4, 0 \rangle$  and  $\langle 0, 3/4 \rangle$  both belong to  $B_2$ ; if  $B_2 = U_1 \times U_2$ , then  $3/4 \in U_1$  and  $3/4 \in U_2$  whence  $\langle 3/4, 3/4 \rangle \in U_1 \times U_2$ , which is impossible since  $\langle 3/4, 3/4 \rangle \notin B_2$ .

In a product space

$$\bigtimes_{i=1}^{\infty} X_i = X_1 \times X_2 \times X_3 \times \cdots$$

of denumerably many factor spaces, the basic open sets are those products

$$U_1 \times U_2 \times \cdots \times U_n \times X_{n+1} \times X_{n+2} \times \cdots$$

of open sets for which the  $i$ th factor is the entire  $i$ th space for all sufficiently large values of  $i$ .

caution:arb-prod-open-not-open

**Caution!** In the product  $\bigtimes_{i=1}^{\infty} X_i$  of a sequence of spaces, **an arbitrary product  $V = \bigtimes_{i=1}^{\infty} V_i$  of open sets need not be open**. For example, for each  $i = 1, 2, 3, \dots$ , take  $X_i$  to be the two-point discrete space  $\{0, 1\}$  and take  $V_i$  to be the open subset  $\{0\}$  of  $X_i$ . The product  $V = \bigtimes_{i=1}^{\infty} V_i$  consists of the single point all of whose coordinates are 0. If  $V$  were open then it would have to contain a basic open set and hence would contain some point almost all of whose coordinates are equal to 1.

In general, the collection  $\mathcal{B}$  of all basic open sets in a product space  $\bigtimes_{i \in I} X_i$  is a base of the product topology. Then according to [Proposition 2.69](#), for each point  $x \in \bigtimes_{i \in I} X_i$ , the collection  $\{B : x \in B \in \mathcal{B}\}$  of all basic open sets containing  $x$  is a local base at  $x$ . Hence neighborhoods in a product space are related to neighborhoods in the factor spaces as follows.

product space!base@and base  
base!product space@of product spa  
basic open set!product topology@ar

neighborhood!prod  
 prod nbd in prod  
 product space!sub  
 product space!sub  
 subspace!product s  
 relative topology!p

**3.57 Proposition.** Let  $x = \langle x_i \rangle_{i \in I}$  be a point in the product of a family  $\langle X_i \rangle_{i \in I}$  of topological spaces. Then a subset  $V$  of  $\prod_{i \in I} X_i$  is a neighborhood of  $x$  if and only if  $V$  contains a product  $\prod_{i \in I} V_i$ , where  $V_i$  is a neighborhood of  $x_i$  for all  $i \in I$  and  $V_i = X_i$  for almost all  $i \in I$ .

**Proof.** If  $V$  is a neighborhood of  $x$ , there is a basic open set  $B$  with  $x \in B \subset V$ .

Conversely, suppose we have such a product  $\prod_{i \in I} V_i$  with  $x \in \prod_{i \in I} V_i \subset V$ . Let  $J = \{i \in I : V_i \neq X_i\}$ , a finite set. For each  $i \in J$ , the set  $V_i$  is a neighborhood of  $x_i$  in  $X_i$  and therefore contains an open neighborhood  $U_i$  of  $x_i$ ; let  $U_i = X_i$  for each  $i \in I \setminus J$ . Then the set  $B$  defined by

$$B = \bigtimes_{i \in I} U_i$$

is a basic open set such that

$$x \in B \subset \prod_{i \in I} V_i \subset V,$$

and so  $V$  is a neighborhood of  $x$ .  $\square$

Nothing is changed in the preceding proposition if we require in addition that each  $V_i$  be open in  $X_i$ . And as usual, the ‘almost all’ clause may be omitted when the index set  $I$  is finite. In particular, then, a neighborhood of a point  $\langle x_1, x_2 \rangle$  in the product  $X_1 \times X_2$  of two spaces is any subset of  $X_1 \times X_2$  that contains a product  $V_1 \times V_2$  of a neighborhood  $V_1$  of  $x_1$  in  $X_1$  with a neighborhood  $V_2$  of  $x_2$  in  $X_2$ , or even a product of such *open* neighborhoods (see Figure 3.14).



Figure 3.14: Neighborhood of a point in the product of two spaces.

fig:nbd-prod-two

Product topologies mesh nicely with relative topologies. Consider, for example, the unit square  $I \times I$ , where as usual  $I = [0, 1]$  with its usual topology. Each of the factors  $I$  has a topology as a subspace of the real line  $\mathbb{R}$ , and so  $I \times I$  may be given the product topology obtained from these factors. At the same time,  $I \times I$  is a subset of the plane  $\mathbb{R} \times \mathbb{R}$  and may be given its relative topology induced by the product topology—the usual topology—on  $\mathbb{R} \times \mathbb{R}$ . These two topologies on  $I \times I$  coincide (since both are induced by the Euclidean metric). The same thing holds, more generally, for any product that is a subset of a larger product.



prop:prod-subspace-of-prod

**3.58 Proposition.** Let  $\langle Y_i \rangle_{i \in I}$  and  $\langle X_i \rangle_{i \in I}$  be families of topological spaces such that  $Y_i$  is a subspace of  $X_i$  for each  $i \in I$ , so that

$$\bigtimes_{i \in I} Y_i \subset \bigtimes_{i \in I} X_i.$$

Then the product topology on  $\bigtimes_{i \in I} Y_i$  is the same as the topology induced by the product topology of  $\bigtimes_{i \in I} X_i$ .

**Proof.** Let

$$Y = \bigtimes_{i \in I} Y_i, \quad X = \bigtimes_{i \in I} X_i.$$

The collection

$$\mathcal{A} = \{B \cap Y : B \text{ is a basic open set in } X\}$$

is a base of the relative topology on  $Y$  induced by the product topology on  $X$ . We need only show that  $\mathcal{A}$  is the collection of basic open sets in the product space  $Y$ .

Let  $B = \bigtimes_{i \in I} U_i$  be a basic open set in  $X$ . Then

$$B \cap Y = \bigtimes_{i \in I} (U_i \cap Y_i)$$

with  $U_i \cap Y_i$  being open in  $Y_i$  for all  $i \in I$  and with  $U_i \cap Y_i = Y_i$  whenever  $U_i = X_i$  and hence for almost all  $i \in I$ . Thus the member  $B \cap Y$  of  $\mathcal{A}$  is a basic open set in  $Y$ .

Conversely, let  $V = \bigtimes_{i \in I} V_i$  be a basic open set in the product space  $Y$ . For each  $i$  belonging to the finite set

$$J = \{i \in I : V_i \neq Y_i\}$$

there is an open subset  $U_i$  of  $X_i$  such that  $U_i \cap Y_i = V_i$ ; let  $U_i = X_i$  for each  $i \in I \setminus J$ . Then  $B = \bigtimes_{i \in I} U_i$  is a basic open set in  $X$  such that  $B \cap Y = V$ . Hence  $V \in \mathcal{A}$ .  $\square$

standing:prod-nonempty

**3.59 Standing Assumption.** From now on we assume that each product  $\bigtimes_{i \in I} X_i$  we consider is nonempty.

This assumption implies that each factor of such a product is nonempty, for if some  $X_j$  were empty there certainly could not exist any family indexed by  $I$  whose  $j$ th coordinate belonged to  $X_i$ .

Conversely, if each factor  $X_i \neq \emptyset$ , then the product  $\bigtimes_{i \in I} X_i \neq \emptyset$ . When  $I$  is infinite, this is a consequence of the Axiom of Choice—see [Corollary 0.27](#).

### Products and maps

subsec:prod-maps

The next several results concern continuity of maps with respect to product topologies.

In [Examples 3.16 \(3\)](#), we saw by using the max metric  $d_\infty$  on  $\mathbb{R}^2$  that the first projection  $p_1: \mathbb{R}^2 \rightarrow \mathbb{R}$  given by  $(x, y) \mapsto x$  is a continuous open map. Since  $d_\infty$  induces the product topology on  $\mathbb{R}^2$ , that example is just a special case of the next theorem.

thm:proj-cont-open-surj

**3.60 Theorem.** Let  $\langle X_i \rangle_{i \in I}$  be a family of topological spaces. Then for each  $j \in I$  the  $j$ th projection

$$p_j: \bigtimes_{i \in I} X_i \rightarrow X_j$$

is a continuous open surjection.

product space!subspace@and subspace!product space@and product space!relative topology!product space@and product space!continuous map@and product space!continuous map!product space@and product space!projection!product space@of product space!projection@and projection

projection!closed map@prod closed map  
closed map!projection@and projection

**Proof.** Let  $X = \prod_{i \in I} X_i$ , and fix  $j \in I$ .

First, the map  $p_j$  is surjective. (This is purely set-theoretic and does not involve any topology.) In fact, let  $y \in X_j$  be arbitrary. By our [Standing Assumption 3.59](#) there exists some  $z = \langle z_i \rangle_{i \in I} \in X$ . Define the element  $x = \langle x_i \rangle_{i \in I}$  of  $X$  by

$$x_i = \begin{cases} y & \text{if } i = j, \\ z_i & \text{if } i \neq j. \end{cases}$$

Then  $p_j(x) = x_j = y$ .

Second, the map  $p_j$  is continuous. In fact, if  $U$  is an open set in  $X_j$ , then its inverse image  $p_j^{-1}(U)$  is a subbasic open set in  $X$ .

Third,  $p_j$  is an open map. In fact, if  $U = \prod_{i \in I} U_i$  is a basic open set in  $X$ , then  $p_j(U) = \emptyset$  if  $U = \emptyset$  and  $p_j(U) = U_j$  if  $U \neq \emptyset$ .  $\square$

However, **projections from a product space need not be closed maps**. We saw this already in [Examples 3.16 \(3\)](#): the first projection  $\langle x, y \rangle \mapsto x$  from  $R \times R$  to  $R$  is not closed map.

We motivated the definition of the product topology by means of metrics. However, there is a purely topological reason for preferring the product topology, as we defined it, over any other topology on a product of topological spaces.

cor:prod-top-least

**3.61 Corollary.** Let  $\langle X_i \rangle_{i \in I}$  be a family of topological spaces. Then the product topology is the coarsest topology on the product set  $X = \prod_{i \in I} X_i$  that makes the projections  $p_j: X \rightarrow X_j$  continuous for all  $j \in I$ .

**Proof.** [Theorem 3.60](#) says that the product topology does make all the projections continuous.

Suppose now that  $\mathcal{T}$  is an arbitrary topology on the set  $X$  that makes all the projections continuous. To show that the product topology is contained in  $\mathcal{T}$  it suffices to show that each subbasic open set belongs to  $\mathcal{T}$ . Let  $S$  be an arbitrary subbasic open set, so that  $S = p_j^{-1}(U)$  for some  $j \in I$  and some open set  $U$  in  $X_j$ . Then this inverse image  $S$  is open for  $\mathcal{T}$  because by assumption  $\mathcal{T}$  makes  $p_j$  continuous.  $\square$

In calculus you learn that a map  $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$  is continuous if and only if its component functions  $f_1: \mathbb{R}^m \rightarrow \mathbb{R}, f_2: \mathbb{R}^m \rightarrow \mathbb{R}, \dots, f_n: \mathbb{R}^m \rightarrow \mathbb{R}$  are continuous, where

$$f(x) = \langle f_1(x), f_2(x), \dots, f_n(x) \rangle \quad (x \in \mathbb{R}^m).$$

This fact can be generalized to any product space once we note that

$$f_1 = p_1 \circ f, f_2 = p_2 \circ f, \dots, f_n = p_n \circ f,$$

where

$$p_1: \mathbb{R}^n \rightarrow \mathbb{R}, p_2: \mathbb{R}^n \rightarrow \mathbb{R}, \dots, p_n: \mathbb{R}^n \rightarrow \mathbb{R}$$

are the projections.

**3.62 Theorem.** Let

$$f: Y \rightarrow \prod_{i \in I} X_i$$

be a map from a topological space  $Y$  into the product of a family  $\langle X_i \rangle_{i \in I}$  of topological spaces. Then  $f$  is continuous if and only if for every  $j \in I$  the composite

$$p_j \circ f: Y \rightarrow X_j$$

of  $f$  with the  $j$ th projection  $p_j: \prod_{i \in I} X_i \rightarrow X_j$  is continuous.

The relationship of the maps in the preceding theorem is shown in the following commutative diagram.

$$\begin{array}{ccc} X & \xrightarrow{f} & \prod_{i \in I} X_i \\ & \searrow p_j \circ f & \downarrow p_j \\ & & X_j \end{array}$$

**Proof.** Since the projections  $p_j$  are continuous, the composites  $p_j \circ f$  are all continuous if  $f$  is.

Conversely, assume that  $p_j \circ f$  is continuous for every  $j \in I$ . Let  $B = \prod_{i \in I} U_i$  be a basic open set in  $\prod_{i \in I} X_i$ . If  $B = \prod_{i \in I} X_i$ , then the set  $f^{-1}(B) = Y$  is open in  $Y$ . Now suppose that  $B \neq \prod_{i \in I} X_i$ . As before we have

$$B = \bigcap_{j \in J} p_j^{-1}(U_j)$$

for some nonempty finite subset  $J$  of  $I$ . Then the set

$$f^{-1}(B) = \bigcap_{j \in J} f^{-1}(p_j^{-1}(U_j)) = \bigcap_{j \in J} (p_j \circ f)^{-1}(U_j)$$

is open in  $Y$  because, by the assumed continuity of  $p_j \circ f$ , the set  $(p_j \circ f)^{-1}(U_j)$  is open in  $Y$  for every  $j \in J$ .  $\square$

In [Exercise 19](#) you were asked to prove the special case of [Theorem 3.62](#) of the product of just two spaces.

The preceding theorem is most often applied in the following situation. Suppose we are given a whole family of continuous maps

$$f_i: Y \rightarrow X_i,$$

one for each  $i \in I$ , all having the common domain  $Y$ . These may be combined into the single map

$$\begin{aligned} f: Y &\rightarrow \prod_{i \in I} X_i \\ y &\mapsto \langle f_i(y) \rangle_{i \in I} \end{aligned}$$

—the unique map  $f$  satisfying

$$p_i \circ f = f_i \quad (i \in I).$$

projection!product space@of produ  
product space!projection@and proje

components of map. Then [Theorem 3.62](#) tells us that  $f$  is continuous. The maps  $f_i$  are called the **components of  $f$** . The relationship of  $f$  to its components is shown in the following commutative diagram.

$$\begin{array}{ccc} X & \xrightarrow{f} & \prod_{i \in I} X_i \\ & \searrow f_j & \downarrow p_j \\ & & X_j \end{array}$$

[Theorem 3.62](#) can also be used if we are given maps having different domains.

**3.63 Corollary.** Let  $\langle Y_i \rangle_{i \in I}$  and  $\langle X_i \rangle_{i \in I}$  be families of topological spaces indexed by the same set  $I$ , and for each  $i \in I$  let

$$f_i: Y_i \rightarrow X_i$$

be a continuous map. Then the map

$$f: \prod_{i \in I} Y_i \rightarrow \prod_{i \in I} X_i$$

defined by

$$y = \langle y_i \rangle_{i \in I} \in \prod_{i \in I} Y_i \implies f(y) = \langle f_i(y_i) \rangle_{i \in I}$$

is continuous. Moreover, if

$$f_i: Y_i \cong X_i$$

for each  $i \in I$ , then

$$f: \prod_{i \in I} Y_i \cong \prod_{i \in I} X_i.$$

**Proof.** Set  $Y = \prod_{i \in I} Y_i$  and  $X = \prod_{i \in I} X_i$ , and for each  $j \in I$  let

$$q_j: Y \rightarrow Y_j, \quad p_j: X \rightarrow X_j$$

be the projections. To show that  $f$  is continuous, it suffices by [Theorem 3.62](#) to show that the composites

$$p_j \circ f: Y \rightarrow X_j$$

are continuous for all  $j \in I$ . Now for  $j \in I$  we have

$$p_j(f(y)) = p_j(\langle f_i(y_i) \rangle_{i \in I}) = f_j(y_j) = f_j(q_j(y))$$

for each  $y = \langle y_i \rangle_{i \in I} \in Y$ , so that

$$p_j \circ f = f_j \circ q_j,$$

and  $f_j \circ q_j$  is continuous because it is a composite of continuous maps.

Suppose now that  $f_i: Y_i \cong X_i$  for each  $i \in I$ . Define

$$g_i = f_i^{-1}: X_i \cong Y_i \quad (i \in I).$$

The map  $g: X \rightarrow Y$  defined by

$$x = \langle x_i \rangle_{i \in I} \implies g(x) = \langle g_i(x_i) \rangle_{i \in I}$$

satisfies

$$f(g(x)) = x \quad (x \in X), \quad g(f(y)) = y \quad (y \in Y).$$

Hence  $f$  is bijective and  $g = f^{-1}$ . Since each  $g_i$  is continuous, the same reasoning as before, but now with the roles of  $X$  and  $Y$  reversed, shows that  $g$  is continuous. Hence  $f$  is a homeomorphism from  $Y$  to  $X$ .  $\square$

The relationship among the maps of [Corollary 3.63](#) and its proof is shown by the following commutative diagram.

$$\begin{array}{ccc} \prod_{i \in I} Y_i & \xrightarrow{f} & \prod_{i \in I} X_i \\ q_j \downarrow & & \downarrow p_j \\ Y_j & \xrightarrow{f_j} & X_j \end{array}$$

The maps  $f_i: Y_i \rightarrow X_i$  are called the **components of the map**

$$f: \prod_{i \in I} Y_i \rightarrow \prod_{i \in I} X_i.$$

When the index set  $I$  consists of only two elements, say  $I = \{1, 2\}$ , the map  $f: Y_1 \times Y_2 \rightarrow X_1 \times X_2$  has two components  $f_1: Y_1 \rightarrow X_1$  and  $f_2: Y_2 \rightarrow X_2$ ; in this case  $f$  is called the **product of  $f_1$  and  $f_2$**  and may be denoted by  $f_1 \times f_2$ , so that

$$(f_1 \times f_2)(y_1, y_2) = \langle f_1(y_1), f_2(y_2) \rangle$$

for each  $\langle y_1, y_2 \rangle \in Y_1 \times Y_2$ .

The examples that follow illustrate the use of [Theorem 3.62](#) and [Corollary 3.63](#). Included are two earlier examples that can now be handled more elegantly, without any special estimates involving metrics.

**Example 3.64 Examples.** (1) Vector addition in  $\mathbb{R}^n$  is the map

$$\alpha: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n \\ \langle x, y \rangle \mapsto x + y$$

having as components the maps

$$\alpha_i: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \\ \langle x, y \rangle \mapsto x_i + y_i \quad (i = 1, 2, \dots, n).$$

For each  $i = 1, 2, \dots, n$ , the component  $\alpha_i$  is the composite of the map

$$p_i \times p_i: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \times \mathbb{R} \\ \langle x, y \rangle \mapsto \langle x_i, y_i \rangle,$$

which is continuous by [Corollary 3.63](#), and the map

$$\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \\ \langle u, v \rangle \mapsto u + v,$$

which is continuous by [Exercise 1.71](#), so that  $\alpha_i$  is continuous. From [Theorem 3.62](#) it follows that  $\alpha$  is continuous.

Similarly, scalar multiplication

$$\mu: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n \\ \langle \lambda, x \rangle \mapsto \lambda x$$

is continuous.

components of map  
product of maps  
vector addition  
scalar multiplication

(2) Let  $f, f_1, f_2, \dots$  be continuous functions with domain the same topological space  $X$ . Their sum

$$\begin{aligned} f_1 + f_2: X &\rightarrow \mathbb{R} \\ x &\mapsto f_1(x) + f_2(x) \end{aligned}$$

is the composite of the two maps

$$\begin{aligned} X &\rightarrow \mathbb{R} \times \mathbb{R} & , & & \mathbb{R} \times \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto \langle f_1(x), f_2(x) \rangle & & & \langle u, v \rangle &\mapsto u + v \end{aligned}$$

The first of these maps is continuous by [Theorem 3.62](#), and the second by [Exercise 1.71](#). Hence the sum  $f_1 + f_2$  is continuous.

Similarly, the product

$$\begin{aligned} f_1 \cdot f_2: X &\rightarrow \mathbb{R} \\ x &\mapsto f_1(x) \cdot f_2(x) \end{aligned}$$

is continuous.

(3) The function  $\max: \mathbb{R} \rightarrow \mathbb{R} \rightarrow \mathbb{R}$  given by

$$\max(x, y) = \begin{cases} x & \text{if } x \geq y, \\ y & \text{if } x < y \end{cases}$$

is continuous. In fact,

$$(*) \quad \max(x, y) = \frac{1}{2} (x + y + |y - x|).$$

The absolute value function is continuous, according to [Examples 3.14 \(1\)](#). The result now follows from repeated application of [Example \(1\)](#).

Similarly, the function  $\min$  given by

$$\min(x, y) = \begin{cases} y & \text{if } x \geq y, \\ x & \text{if } x < y \end{cases}$$

is continuous.

(4) Let  $f$  and  $g$  be continuous real-valued functions on a topological space  $X$ . Then the real-valued function  $\max(f, g)$  given by  $x \mapsto \max\{f(x), g(x)\}$  is continuous. In fact, this function may be written as the composite of the continuous maps

$$\begin{aligned} \max: \mathbb{R} \times \mathbb{R} &\rightarrow \mathbb{R} & , & & f \times g: X \times X &\rightarrow \mathbb{R} \times \mathbb{R} & , & & \delta_X: X &\rightarrow X \times X \\ \langle u, v \rangle &\mapsto \max\{u, v\} & & & \langle x, y \rangle &\mapsto (f(x), g(y)) & & & z &\mapsto \langle z, z \rangle \end{aligned}$$

Similarly, the real-valued function  $\min(f, g)$  is continuous.

(5) The map

$$\begin{aligned} f: \mathbb{R}^n &\rightarrow \mathbb{B}_n \\ x &\mapsto \frac{1}{1 + \|x\|} x \end{aligned}$$

considered in [Examples 3.25 \(9\)](#) is the composite of the maps

$$\begin{aligned} \mathbb{R}^n &\rightarrow \mathbb{R} \times \mathbb{R}^n & , & & \mathbb{R} \times \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ x &\mapsto \left( \frac{1}{1 + \|x\|}, x \right) & & & (\lambda, u) &\mapsto \lambda u \end{aligned}$$

The former of these is continuous by [Theorem 3.62](#), and the latter by [Example \(1\)](#). Hence  $f$  is continuous.

(6) The map

$$h: I^n \rightarrow J^n \\ \langle x_1, x_2, \dots, x_n \rangle \mapsto \langle k(x_1), k(x_2), \dots, k(x_n) \rangle$$

constructed in [Examples 3.25 \(11\)](#) is obtained from  $n$  component functions each of which is the homeomorphism  $k: I = [0, 1] \cong J = [-1, 1]$ . From [Corollary 3.63](#) it follows that  $h$  is a homeomorphism, too.

This is a much easier way to show that  $h$  is a homeomorphism of the  $n$ -cube  $I^n$  with the centered  $n$ -cube  $J^n$  than in that example, where we used metrics.

Hilbert cube!product of intervals@as

ex:Hilbert-cube-prod-intervals

(7) The Hilbert cube [\(b\)](#)

$$I^\infty = \{ \langle x_i \rangle_{i=1,2,3,\dots} : x_i \in \mathbb{R} \text{ and } |x_i| \leq \frac{1}{i} \text{ for each } i = 1, 2, \dots \}$$

can be topologized in two natural ways. First, as a subset of the Hilbert sequence space  $\ell^2$ , the cube  $I^\infty$  has the topology induced by the usual metric  $d_2$  on  $\ell^2$ . Second, as a product

$$I^\infty = [-1, 1] \times [-1/2, 1/2] \times [-1/3, 1/3] \times \cdots = \bigtimes_{i=1}^{\infty} [-1/i, 1/i]$$

of intervals,  $I^\infty$  has its product topology. These two topologies are the same, as we shall demonstrate in a moment. Then since  $[-1/i, 1/i] \cong [0, 1] = I$  for each  $i = 1, 2, 3, \dots$ , we can conclude from [Corollary 3.63](#) that

$$I^\infty \cong I \times I \times \cdots \times I \times \cdots$$

(which justifies the notation  $I^\infty$  for the Hilbert cube).

Let  $X_i = [-1/i, 1/i]$  for each  $i = 1, 2, 3, \dots$ . To see that the metric  $d_2$  on  $\ell^2$  does in fact induce the product topology on the set  $I^\infty = X_1 \times X_2 \times X_3 \times \cdots$ , fix a point  $x = \langle x_1, x_2, x_3, \dots \rangle \in I^\infty$ . The collections

$$\{U : U \text{ is a basic open set in } I^\infty \text{ and } x \in U\}, \\ \{B_\varepsilon(x; d_2) \cap I^\infty : \varepsilon > 0\}$$

are local bases at  $x$  for the product topology and for the topology induced by  $d_2$ , respectively. We shall show now that these collections define the same neighborhood system at  $x$ .

Let

$$U = U_1 \times U_2 \times \cdots \times U_n \times X_{n+1} \times X_{n+2} \times \cdots$$

be a basic open set (for the product topology) in  $I^\infty$  with  $x \in U$ . Then  $U_1 \times U_2 \times \cdots \times U_n$  is a neighborhood of the point  $\langle x_1, x_2, \dots, x_n \rangle$  in the product space  $X_1 \times X_2 \times \cdots \times X_n$ . Because the Euclidean metric induces the product topology on  $X_1 \times X_2 \times \cdots \times X_n$ , there is an  $\varepsilon > 0$  such that for  $\langle y_1, y_2, \dots, y_n \rangle \in X_1 \times X_2 \times \cdots \times X_n$ ,

$$\sum_{i=1}^n (x_i - y_i)^2 \leq \varepsilon^2 \implies \langle y_1, y_2, \dots, y_n \rangle \in U_1 \times U_2 \times \cdots \times U_n.$$

Then  $y = \langle y_1, y_2, \dots, y_n, \dots \rangle \in I^\infty$  with  $d_2(x, y) < \varepsilon$  implies

$$\sum_{i=1}^n (x_i - y_i)^2 \leq \sum_{i=1}^{\infty} (x_i - y_i)^2 = (d_2(x, y))^2 < \varepsilon^2$$

so that  $y \in U$ . Thus  $B_\varepsilon(x; d_2) \cap I^\infty \subset U$ .

Anderson, Richard D.

Now suppose  $\varepsilon > 0$ . We show that  $x \in U \subset B_\varepsilon(x; d_2)$  for some basic open set  $U$  in  $l^\infty$ . The series  $\sum_{i=1}^\infty (2/i)^2$  converges, and so we may choose an  $n$  so large that the “tail” of that series after the  $n$ th term satisfies

$$\sum_{i=n+1}^\infty \left(\frac{2}{i}\right)^2 < \frac{\varepsilon^2}{2}.$$

Since the Euclidean metric induces the product topology on the space  $X_1 \times X_2 \times \cdots \times X_n$ , there are open neighborhoods  $U_i$  of  $x_i$  in  $X_i$  for  $i = 1, 2, \dots, n$  such that

$$\langle y_1, y_2, \dots, y_n \rangle \in U_1 \times U_2 \times \cdots \times U_n \implies \sum_{i=1}^n (x_i - y_i)^2 < \frac{\varepsilon^2}{2}.$$

The basic open set

$$U = U_1 \times U_2 \times \cdots \times U_n \times X_{n+1} \times X_{n+2} \times \cdots$$

in  $l^\infty$  contains  $x$ . Moreover,  $U \subset B_\varepsilon(x; d_2)$ . In fact, let

$$y = \langle y_1, y_2, \dots, y_n, y_{n+1}, \dots \rangle \in U.$$

Then  $\langle y_1, y_2, \dots, y_n \rangle \in U_1 \times U_2 \times \cdots \times U_n$  and  $|x_i - y_i| \leq 2/i$  for every  $i \geq n+1$ , so that

$$[d_2(x, y)]^2 = \sum_{i=1}^n (x_i - y_i)^2 + \sum_{i=n+1}^\infty (x_i - y_i)^2 < \frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{2} = \varepsilon^2,$$

and so  $d_2(x, y) < \varepsilon$ .

space-homeo-prod-denum-copies-R

- (8) A fact related to (7), proved by R. D. Anderson in 1966, is that the Hilbert sequence space  $\ell^2$  is homomorphic to the product of denumerably many lines:

$$\ell^2 \cong \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} \times \cdots$$

[See Anderson [1].]  $\diamond$

Another application of Theorem 3.62 is the commutative law for products, which says that a product space is unchanged topologically if its factors are permuted. We consider here only the special case of two factors; the general case is treated in Exercise 128.

prop:prod-two-commute

**3.65 Proposition.** *Let  $X$  and  $Y$  be topological spaces. Then the bijection*

$$\begin{aligned} h: X \times Y &\rightarrow Y \times X \\ \langle x, y \rangle &\mapsto \langle y, x \rangle \end{aligned}$$

*is a homeomorphism. Thus*

$$X \times Y \cong Y \times X.$$

**Proof.** All four of the projections

$$\begin{aligned} p_1: X \times Y &\rightarrow X, & p_2: X \times Y &\rightarrow Y \\ \langle x, y \rangle &\mapsto x & \langle x, y \rangle &\mapsto y \end{aligned}$$

and

$$\begin{aligned} q_1: Y \times X &\rightarrow Y, & q_2: Y \times X &\rightarrow X \\ \langle y, x \rangle &\mapsto y & \langle y, x \rangle &\mapsto x \end{aligned}$$



are continuous. Then the components  $q_1 \circ h: X \times Y \rightarrow Y$ ,  $q_2 \circ h: X \times Y \rightarrow X$  of  $h$  are continuous because

$$q_1 \circ h = p_2, \quad q_2 \circ h = p_1.$$

Hence  $h$  is continuous. Similarly,  $h^{-1}$  is continuous.  $\square$

There is also an associative law for product spaces, which says that a product space is topologically unchanged if its factors are grouped in any way. Again, we formulate here only a special case; the general case is treated in [Exercise 130](#).

prop:finite-prod-initial-associative

**3.66 Proposition.** *Let  $X_1, X_2, \dots, X_n$  be topological spaces. Then*

$$X_1 \times X_2 \times \cdots \times X_n \cong (X_1 \times X_2 \times \cdots \times X_{n-1}) \times X_n.$$

**Proof.** Let

$$X = X_1 \times X_2 \times \cdots \times X_n, \quad Y = X_1 \times X_2 \times \cdots \times X_{n-1}.$$

An element of  $X$  is an  $n$ -tuple  $\langle x_1, x_2, \dots, x_n \rangle$ , while an element of  $Y \times X_n$  is a pair  $\langle y, x \rangle$  with  $y$  itself an  $(n-1)$ -tuple  $\langle y_1, y_2, \dots, y_{n-1} \rangle$ . We have a natural map

$$h: \begin{array}{ccc} X & \rightarrow & Y \times X_n \\ \langle x_1, x_2, \dots, x_n \rangle & \mapsto & (\langle x_1, x_2, \dots, x_{n-1} \rangle, x_n) \end{array}$$

which is a bijection whose inverse is

$$h^{-1}: \begin{array}{ccc} Y \times X_n & \rightarrow & X \\ (\langle y_1, y_2, \dots, y_{n-1} \rangle, x) & \mapsto & \langle y_1, y_2, \dots, y_{n-1}, x \rangle \end{array}$$

The map  $h$  has the two components

$$\begin{array}{ccc} X & \rightarrow & Y, \\ \langle x_1, x_2, \dots, x_n \rangle & \mapsto & \langle x_1, x_2, \dots, x_{n-1} \rangle \end{array}, \quad \begin{array}{ccc} X & \rightarrow & X_n, \\ \langle x_1, x_2, \dots, x_n \rangle & \mapsto & x_n \end{array}$$

The latter of these two maps is continuous because it is just the  $n$ th projection of the product  $X$ . The former of these is a map into the product  $Y$  that, by [Theorem 3.62](#), is continuous because its own components are the 1st through  $(n-1)$ st projections of the product  $Y$ . Hence by [Theorem 3.62](#) again the map  $h$  is continuous. Similarly, the inverse  $h^{-1}$  map is continuous.  $\square$

The preceding [Proposition 3.66](#) was already used implicitly in [Exercise 66 \(a\)](#) when we wrote

eq:cylinder-prod-circle-and-interval} (\*)

$$S_1 \times I = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 = 1, 0 \leq x_3 \leq 1\}.$$

The space  $S_1 \times I$  on the left-hand side of (\*) is, of course, not really a subspace of  $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ , but rather of  $\mathbb{R}^2 \times \mathbb{R} = (\mathbb{R} \times \mathbb{R}) \times \mathbb{R}$ . It is, however, the subspace of  $\mathbb{R}^2 \times \mathbb{R}$  that corresponds to the set on the right-hand side of (\*) under the homeomorphism  $\mathbb{R}^3 \cong \mathbb{R}^2 \times \mathbb{R}$  provided by the proof of the associative law expressed by [Proposition 3.66](#). Thus the equality sign ‘=’ in (\*) should actually be ‘ $\cong$ ’, but we have implicitly made an identification of the two spaces in (\*). In the sequel, *we reserve the right to make similar implicit identifications that make implicit use of the associative law for product spaces*.

If we choose points  $c_1 \in \mathbb{R}$  and  $c_2 \in \mathbb{R}$ , then the subspace

$$L = \{\langle x_1, x_2, x_3 \rangle \in \mathbb{R}^3 : x_1 = c_1, x_2 = c_2\}$$

of  $\mathbb{R}^3$ , obtained by fixing the first two coordinates of points, is a line that is homeomorphic to the third factor  $\mathbb{R}$  of  $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$  (see [Figure 3.15](#)). In general, a product space

Figure 3.15: The real line embedded as a subspace  $L$  of  $\mathbb{R}^3$ .

fig:line-embedded-in-space

contains copies of each of its factors.

thm:embed-factor-in-prod

**3.67 Theorem.** Let  $\langle X_i \rangle_{i \in I}$  be a family of topological spaces, let  $c = \langle c_i \rangle_{i \in I}$  be a point in the product space  $\prod_{i \in I} X_i$ , and let  $j \in I$  be a particular index. Then the map

$$f: X_j \rightarrow \prod_{i \in I} X_i$$

defined by

$$f(y) = \langle x_i \rangle_{i \in I}$$

where

$$x_i = \begin{cases} y & \text{if } i = j, \\ c_i & \text{if } i \neq j \end{cases}$$

is an embedding.

**Proof.** The range  $f(X_j)$  of  $f$  is the subspace  $A = \prod_{i \in I} A_i$  of  $\prod_{i \in I} X_i$  given by

$$A_i = \begin{cases} X_j & \text{if } i = j, \\ \{c_i\} & \text{if } i \neq j. \end{cases}$$

We shall show that  $f$ , considered as a map from  $X_j$  onto  $A_j$  is a homeomorphism. Note that by [Proposition 3.58](#) the relative topology on  $A$  is the product topology on  $A$ .

Let  $q_i: A \rightarrow A_i$  be the  $i$ th projection for each  $i \in I$ . Then  $y \in X_j$  implies

$$(q_j \circ f)(y) = y, \quad (q_i \circ f)(y) = c_i \quad (i \neq j).$$

This means that each of the components  $q_i \circ f: X_j \rightarrow A_i$  is a constant map except the  $j$ th, which is the identity map of  $X_j = A_j$ , and so all of these component maps are continuous. Hence  $f: X_j \rightarrow A$  is itself continuous.

The relation  $(q_j \circ f)(y) = y$  above coupled with

$$(f \circ q_j)(a) = a \quad (a \in A)$$

shows that the surjection  $f$  is injective as well, and that

$$f^{-1} = q_j: A \rightarrow A_j = X_j.$$

Hence  $f^{-1}$  is continuous, too.  $\square$

product space!continuous map@and  
continuous map!product space@and

### Further properties of product spaces

subsec:prod-further

[Theorem 3.67](#) of the preceding subsection can sometimes be used to show that if a product space has a certain topological property, then each of its factors has the same property. Of course it is typically more interesting to know that if each of a family of spaces has the property, then their product has the same property—as we say, that the property is “preserved” under the formation of products. The next theorem is such a result.

thm:prod-and-topological-T0

**3.68 Theorem.** (1) A product space  $\times_{i \in I} X_i$  is a  $T_0$ -space if and only if its factors  $X_j$  are  $T_0$ -spaces for all indices  $j \in I$ .

thm-part:prod-and-T1

(2) A product space  $\times_{i \in I} X_i$  is a  $T_1$ -space if and only if its factors  $X_j$  are  $T_1$ -spaces for all indices  $j \in I$ .

thm-part:prod-and-T2

(3) A product space  $\times_{i \in I} X_i$  is a Hausdorff space if and only if its factors  $X_j$  are Hausdorff spaces for all indices  $j \in I$ .

thm-part:prod-and-regular

(4) A product space  $\times_{i \in I} X_i$  is regular if and only if its factors  $X_j$  are regular spaces for all indices  $j \in I$ .

**Proof.** The proofs of (1), (2), and (4) are left as exercises.

(3) Let  $X = \times_{i \in I} X_i$ .

Assume first that the product  $X$  is a Hausdorff space. Then each subspace of  $X$  is a Hausdorff space. If  $j \in I$  is arbitrary, then the factor space  $X_j$  is, by [Theorem 3.67](#), homeomorphic to a subspace of  $X$ , and so  $X_j$  is a Hausdorff space.

Conversely, assume that  $X_j$  is a Hausdorff space for every  $j \in I$ . Let  $x = \langle x_i \rangle_{i \in I}$  and  $y = \langle y_i \rangle_{i \in I}$  be points in  $X$  with  $x \neq y$ . Then  $x_j \neq y_j$  for some index  $j \in I$ . By assumption there are disjoint neighborhoods  $U_j$  and  $V_j$  of  $x_j$  and  $y_j$ , respectively, in  $X_j$ . Then the subbasic open sets  $U = p_j^{-1}(U_j)$  and  $V = p_j^{-1}(V_j)$  are neighborhoods of  $x$  and  $y$ , respectively, in the product space  $X$ . Moreover, the sets  $U$  and  $V$  are disjoint, because no point of  $X$  can have its  $j$ th coordinate both in  $U_j$  and in  $V_j$ .  $\square$

Thus each of the properties of being a  $T_0$ -space, a  $T_1$ -space, a Hausdorff space, or a regular space is preserved under the formation of products. The relationship of being completely regular or normal to the formation of products is considered in [Section 6.2](#).

Each of the topological properties of compactness and connectedness, introduced in the next two chapters, is preserved under the formation of products. See [Theorems 4.31](#) and [4.33](#) (the Tychonoff Product Theorem) and [Theorems 5.24](#) and [5.29](#), respectively.

Metrizability is another topological property that is preserved under the formation of products, provided the index set is suitably restricted.

thm:prod-and-metrizable

**3.69 Theorem.** Let  $I$  be a countable index set. Then a product space  $\times_{i \in I} X_i$  is metrizable if and only if its factors  $X_j$  are metrizable for all indices  $j \in I$ .

**Proof.** Since a subspace of a metrizable space is metrizable, [Theorem 3.67](#) implies that each  $X_j$  is metrizable if  $\times_{i \in I} X_i$  is.

Conversely, assume that  $X_j$  is metrizable for every  $j \in I$ . For each  $j \in I$  choose a metric  $d_j$  inducing the topology of  $X_j$ .

Case (i): the index set  $I$  is finite. Relabeling the factors if necessary, we may assume without loss of generality that  $I = \{1, 2, \dots, n\}$  for some positive integer  $n$ . Then by [Proposition 3.50](#), the max metric induced by  $\langle d_1, d_2, \dots, d_n \rangle$  induces the product topology on  $\times_{i \in I} X_i$ .

Case (ii): the index set  $I$  is infinite. Then  $I$  is denumerable, so we may assume without loss of generality that  $I = \mathbb{N}^* = \{1, 2, 3, \dots\}$ . We may further assume without loss of generality that for each  $j \in I$  the metric  $d_j$  makes  $X_j$  bounded with  $\text{diam } X_i \leq 1$ , because we may replace any  $d_j$  not having that property by an equivalent metric that does have it. Then by [Proposition 3.51](#) the metric  $d'$  defined by [equation \(\\*\)](#) on [page 361](#) induces the product topology on  $\times_{i \in I} X_i$ .  $\square$

The assumption of a countable index set is essential for a product space to be metrizable (except in certain special cases of one-point spaces): see [Exercise 138](#).

Countability properties are also preserved by products of countably many spaces.

prop:product-and-countability

**3.70 Proposition.** Let  $\langle X_i \rangle_{i \in I}$  be a family of topological space having a countable index set. Then:

prop-part:product-and-1st-countable

(1) The product space  $\times_{i \in I} X_i$  is first-countable if and only if its factors  $X_i$  are first-countable for all indices  $i \in I$ .

prop-part:product-and-2nd-countable

(2) The product space  $\times_{i \in I} X_i$  is second-countable if and only if its factors  $X_i$  are second-countable for all indices  $i \in I$ .

prop-part:product-and-separable

(3) The product space  $\times_{i \in I} X_i$  is separable if and only if its factors  $X_i$  are separable for all indices  $i \in I$ .

The proofs are requested in [Exercises 137](#) and [139](#). For first- and second-countability, the assumption of a countable index set is essential (except in certain special cases of one-point spaces): see [Exercise 138](#). For separability, however, the assumption of a countable index set is too restrictive: see the note in [Exercise 139](#).

In a note on [page 363](#) we suggested that you might want to regard the index set  $I$  for products as being countable and, in fact, either the finite set  $\{1, 2, \dots, n\}$  for some positive integer  $n$  or else the denumerable set  $\{1, 2, 3, \dots\}$ . Yet we noted that everything about products following, unless otherwise indicated, applied to an arbitrary index set. For the next example, however, we will definitely want the index set to be arbitrary, not necessarily countable; indeed, for most interesting cases of this example the index set will be uncountable.

top-ptwise-convergence-is-prod-top

**3.71 Example.** Let  $X$  be any set and let  $Y$  be a topological space. Form the set

$$\mathcal{F}(X, Y) = \{f : f \text{ is a map from } X \text{ to } Y\}$$

of all maps  $f : X \rightarrow Y$ . (See [Definition 0.15](#).) Then this set of functions is the  $X$ th power set  $Y^X$  of  $Y$  and so, in a different guise, is just a product

$$\mathcal{F}(X, Y) = \prod_{x \in X} Y_x$$

where  $X$  is the index set and where

$$Y_x = Y \quad (x \in X).$$

A typical element of  $\mathcal{F}(X, Y)$ —a map

$$\begin{aligned} f : X &\rightarrow Y \\ x &\mapsto f_x \end{aligned}$$

—is a family  $f = \langle f_x \rangle_{x \in X}$  indexed by  $X$  such that  $f_x \in Y_x = Y$  for each  $x \in X$ . Nonetheless, we shall revert to the usual functional notation, writing  $f(x)$  for the  $x$ th coordinate  $f_x$  of an  $f \in \mathcal{F}(X, Y)$ .

Give  $\mathcal{F}(X, Y)$  its product topology. Then a subbasic open set  $U$  in this product space is obtained as follows: Choose a point  $x \in X$  and an open subset  $V$  of  $Y$ ; then

$$\{eq:subbasic-set-func-space\} (*) \quad U = p_x^{-1}(V) = \{f \in \mathcal{F}(X, Y) : f(x) \in V\}.$$

The expression  $(*)$  indicates that the subbasic open set  $U$  is precisely what was denoted by  $B(x, V)$  in [Examples 2.93 \(2\)](#). Thus **the product topology on  $\mathcal{F}(X, Y)$  is the topology of pointwise convergence**, as described in that example. [That example generalized the case  $X = Y = \mathbb{R}$  treated in [Examples 2.72 \(8\)](#).]

We shall look again at the topology of pointwise convergence in [Section 3.5](#), where we shall explain its name.  $\diamond$

### EXERCISES FOR SECTION 3.3

**112.** Show by example that for subsets  $A$  and  $B$  of topological spaces  $X$  and  $Y$ , respectively, it need not be the case that  $\text{bdy}(A \times B) = (\text{bdy } A) \times (\text{bdy } B)$ . [Compare formula [\(c\)](#) in [Exercise 2.105](#).]

**113.** Do formulas [\(a\)](#) and [\(b\)](#) in [Exercise 2.105](#) generalize to finitely many factors? to infinitely many factors?

prob:Sorgenfrey-plane **114.** Show that the Sorgenfrey plane [[Examples 2.20 \(1\)](#)] is a separable space but that its subspace  $D = \{\langle x, -x \rangle : x \in \mathbb{R}_I\}$ , its “reverse diagonal,” is not separable.

[*Note:* The previous example of a separable space in which a subspace need not be separable was the half-disk space—see [Examples 2.20 \(3\)](#).]

**115.** Let  $A$  be a subset of a product set  $X = \times_{i \in I} X_i$ , and for each  $i \in I$  let  $p_i : X \rightarrow X_i$  be the  $i$ th projection. Prove or disprove:

**(a)**  $A \subset \times_{i \in I} p_i(A_i)$ .

**(b)**  $A = \times_{i \in I} p_i(A_i)$ .

**(c)** If  $x = \langle x_i \rangle_{i \in I} \notin A$ , then  $x_i \notin p_i(A_i)$  for some  $i \in I$ .

prob:prod-open-closed-dense **116.** For each  $i \in I$  let  $A_i$  be a subset of a topological space  $X_i$

prob-part:prod-open **(a)** Suppose that  $A_i$  is nonempty and open in  $X_i$  for each  $i \in I$ . We know that  $\times_{i \in I} A_i$  is open in  $\times_{i \in I} X_i$  if  $A_i = X_i$  for almost all  $i \in I$ . Show that the converse is also true.

prob-part:prod-closed **(b)** Prove: If  $A_i$  is closed in  $X_i$  for each  $i \in I$ , then  $\times_{i \in I} A_i$  is closed in  $\times_{i \in I} X_i$ . Is the converse also true?

**(c)** When, in general, is  $\times_{i \in I} A_i$  dense in  $\times_{i \in I} X_i$ ?

prob:prod-and-discrete **117.** **(a)** Prove that the product of finitely many discrete spaces is itself discrete.

**(b)** Show that the product of infinitely many discrete spaces cannot be discrete unless almost all of the factor spaces consist of a single point.

prob:max-min-cont **118.** **(a)** Verify the formula  $(*)$  used in [Examples 3.14 \(3\)](#) to show that the function  $\max$  is continuous.

**(b)** Prove the assertion made in [Examples 3.14 \(3\)](#) that the function  $\min$  is continuous.

product topology!topology of pointw  
topology!pointwise convergence@o  
pointwise convergence  
product space

Sorgenfrey plane!separable space@  
separable space!subspace@and sub  
subspace!separable space@of separ  
product space!open set@and open s  
product space!closed set@and close  
discrete space

- 119.** Use the techniques of this section to establish the continuity of the following maps.
- (a) The quotient  $x \mapsto f(x)/g(x)$  of two real-valued functions  $f$  and  $g$  on a topological space  $X$ , where  $g(x) \neq 0$  at all  $x \in X$ .
  - (b) The map  $e: J^n \setminus \{0\} \rightarrow E$  of Examples 3.25 (12).
  - (c) The map  $f: S_n \setminus \{p\} \rightarrow H = \mathbb{R}^n \times \{0\}$ , from the punctured  $n$ -sphere to the “equatorial hyperplane”, of Examples 3.25 (13).
  - (d) The **antipodal map**  $x \mapsto -x$  of  $S_n \rightarrow S_n$  that reflects points of the  $n$ -sphere through the origin.
- 120.** Which of the maps (b)–(d) in Exercise 119 are homeomorphisms?
- 121.** If  $X$  is any topological space, show that the diagonal
- $$\Delta_X = \{\langle x, x \rangle : x \in X\}$$
- of the product  $X \times X$  is homeomorphic to  $X$ .
- 122.** Prove that a topological space  $X$  is a Hausdorff space if and only if the diagonal  $\Delta_X$  of  $X \times X$  is closed in the product  $X \times X$ .
- 123.** Let  $f: X \rightarrow Y$  be a map from a topological space  $X$  to a topological space  $Y$  and let
- $$G = \{\langle x, y \rangle \in X \times Y : y = f(x)\}$$
- be its graph. Prove that if  $f$  is continuous, the its graph  $G$  is homeomorphic to the domain  $X$  of  $f$ .
- 124.** Let  $f: X \rightarrow Y$  be a map from a topological space  $X$  to a topological space  $Y$  and let  $G$  be its graph.
- (a) Suppose  $Y$  is a Hausdorff space. Prove that if  $f$  is continuous, then  $G$  is closed in  $X \times Y$ .  
[Note: The converse of this implication fails—see (c), below. However, the converse is true with the additional condition that  $Y$  be “compact”; this is the *closed-graph theorem*—see Exercise 4.44.]
  - (b) Does the result in (a) still hold when  $Y$  is not necessarily a Hausdorff space?
  - (c) Let  $f: \mathbb{R} \rightarrow \mathbb{R}$  be defined by  $f(x) = 1/x$  if  $x \neq 0$  but  $f(0) = 0$ . Of course  $\mathbb{R}$  is a Hausdorff space and  $f$  is *not* continuous. Show that nonetheless the graph of  $f$  is closed in  $\mathbb{R} \times \mathbb{R}$ . [Thus in general the converse of the implication in (a) does not hold. (But see Exercise 4.44 for a situation where that converse does hold.)]
- 125.** Consider a quadratic function  $q(t) = a(t)x^2 + b(t)x + c(t)$  with variable coefficients  $a(t), b(t), c(t)$  that are continuous real-valued functions with common domain an open interval  $J$  in  $\mathbb{R}$ .
- (a) Assume that  $b(t)^2 - 4a(t)c(t) \geq 0$  for all  $t \in J$ , so that, at each  $t \in \mathbb{R}$ , the function  $q$  has either two distinct real roots or a single real root of multiplicity 2. Show that the largest and smallest of these roots are continuous as functions of  $t$ .
  - (b) What can you say about continuity of roots, as functions of  $t$ , if the condition that  $b(t)^2 - 4a(t)c(t) \geq 0$  for all  $t \in J$  is no longer assumed.

prob:two-topo-spaces

- 126. (a)** Establish continuity of the vector operations

$$\begin{aligned}\ell^2 \times \ell^2 &\rightarrow \ell^2, & \mathbb{R} \times \ell^2 &\rightarrow \ell^2 \\ \langle x, y \rangle &\mapsto x + y & \langle \lambda, x \rangle &\mapsto \lambda x\end{aligned}$$

in the Hilbert sequence space  $\ell^2$ .

[*Note:* That these operations in the vector space  $\ell^2$  are continuous means that  $\ell^2$  is a **topological vector space**. Likewise, [Examples 3.64 \(1\)](#) says that  $\mathbb{R}^n$  is a topological vector space.]

- (b)** Deduce from [\(a\)](#) that the map

$$\begin{aligned}\ell^2 &\rightarrow \ell^2 \\ x &\mapsto -x\end{aligned}$$

is continuous.

prob:n-cell-homeo-n-cube

- 127.** Show that, for a positive integer  $n$ , the product

$$[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$$

of a family  $\langle [a_i, b_i] \rangle_{i=1,2,\dots,n}$  of closed intervals with  $a_i < b_i$  for each  $i$  is homeomorphism to the  $n$ -cube  $I^n$  and hence is an  $n$ -cell.

prob:prod-topo-spaces

- 128. (a)** Prove the general commutative law for product spaces: Let  $\sigma$  be a permutation of the index set  $I$  of a family  $\langle X_i \rangle_{i \in I}$  of topological spaces (that is,  $\sigma$  is a bijection from  $I$  to  $I$ —see [page 29](#)). Then

$$\prod_{i \in I} X_i \cong \prod_{i \in I} X_{\sigma(i)}.$$

- (b)** Obtain [Proposition 3.65](#) as a special case of [\(a\)](#).

part:prove-law-product-of-products

- 129. (a)** Given families  $\langle X_i \rangle_{i \in I}$  and  $\langle Y_i \rangle_{i \in I}$  of topological spaces indexed by the same set  $I$ , construct a homeomorphism

$$\prod_{i \in I} (X_i \times Y_i) \cong \left( \prod_{i \in I} X_i \right) \times \left( \prod_{i \in I} Y_i \right).$$

- (b)** Apply part [\(a\)](#) with  $I = \{1, 2, \dots, n\}$  and  $X_i = Y_i = \mathbb{R}$  to reprove continuity of vector addition  $\langle x, y \rangle \mapsto x + y$  in  $\mathbb{R}^n$ .

prob:prod-topo-spaces

- 130. (a)** Prove the general associative law for product spaces: Let  $\langle I_j \rangle_{j \in J}$  be a partition of the index set  $I$  of a family  $\langle X_i \rangle_{i \in I}$  of topological spaces (so that each  $I_j$  is a nonempty subset of  $I$ , the set  $I_j$  is disjoint from  $I_k$  whenever  $j \neq k$ , and  $\bigcup_{j \in J} I_j = I$ ). Then

$$\prod_{i \in I} X_i \cong \prod_{j \in J} \left( \prod_{i \in I_j} X_i \right).$$

- (b)** Obtain [Proposition 3.66](#) as a special case of [\(a\)](#).

- 131.** Generalize [Theorem 3.67](#) by showing that  $\prod_{j \in J} X_j$  can be embedded in  $\prod_{i \in I} X_i$  for each nonempty subset  $J$  of the index set  $I$ .

prob:prod-manifolds

- 132. (a)** Show that the product of an  $m$ -dimensional manifold and an  $n$ -dimensional manifold is an  $(m + n)$ -dimensional manifold.
- (b)** Deduce from [\(a\)](#) that the product of finitely many manifolds is again a manifold. (*Hint:* Use associativity of products—[Proposition 3.66](#).)

Hilbert sequence space  
topological vector space  
n-cell

product space!commutative law  
product space!associative law  
embedding  
product space!embedding  
manifold!product space  
product space!manifold

**133.** Let  $\langle X_i \rangle_{i \in I}$  be a family of topological spaces and for each  $i \in I$  let  $\mathcal{B}_i$  be a base of  $X_i$ . Show that the collection of all products  $\times_{i \in I} B_i$  such that  $B_i = X_i$  for almost all  $i \in I$  and  $B_i \in \mathcal{B}_i$  whenever  $B_i \neq X_i$  is a base of the product topology. Describe this base in terms of the projections.

**134.** Let  $\langle X_i \rangle_{i \in I}$  be a family of topological spaces and let  $x = \langle x_i \rangle_{i \in I}$  be a point in  $\times_{i \in I} X_i$ . Suppose that for each index  $i \in I$ , the collection  $\mathcal{M}_i$  is a local base at  $x_i$  in  $X_i$ . Use the family  $\langle \mathcal{M}_i \rangle_{i \in I}$  to determine a local base at  $x$  in  $\times_{i \in I} X_i$ .

**135. (a)** Prove: A product space is a  $T_0$ -space if and only if each of its factors is a  $T_0$ -space. [This is Theorem 3.68 (1).]

**(b)** Prove: A product space is a  $T_1$ -space if and only if each of its factors is a  $T_1$ -space. [This is Theorem 3.68 (2).]

**136.** Prove: A product space is regular if and only if each of its factors is regular. [This is Theorem 3.68 (4).]

**137.** Prove Proposition 3.70 parts (1) and (2): the product of *countably* many spaces is

first-countable (respectively, second-countable) if and only if each of its factors is first-countable (respectively, second-countable).

**138.** Let  $\langle X_i \rangle_{i \in I}$  be a family of spaces whose index set  $I$  is uncountable. Suppose  $X_i$  contains

at least two points for uncountably many indices  $i \in I$ . Show that  $\times_{i \in I} X_i$  cannot be first-countable—and hence cannot be either second-countable or metrizable—even if each  $X_i$  is first-countable. [Hint: Compare Examples 2.72 (8).]

**139.** Prove Proposition 3.70 part (3): the product of *countably* many spaces is separable if and only if each of its factors is separable.

[Note: The hypothesis that the index set be countable is more restrictive than is needed: The product  $\times_{i \in I} X_i$  of a family  $\langle X_i \rangle_{i \in I}$  of topological spaces is separable if the index set  $I$  has “cardinality at most  $\mathfrak{c}$ ” and each factor space  $X_i$  is separable; the converse holds if in addition each factor space  $X_i$  is a Hausdorff space with at least two points. (To say that  $I$  has cardinality at most  $\mathfrak{c}$  means that  $I$  can be put into one-to-one correspondence with some subset of  $\mathbb{R}$ .) For a proof, see Willard [74, Theorem 16.4 (c), page 109].)

**140.** Let  $\langle X_i \rangle_{i \in I}$  be a family of topological spaces.

(a) Verify that the collection

$$\mathcal{B} = \left\{ \times_{i \in I} U_i : U_i \text{ is open in } X_i \text{ for each } i \in I \right\},$$

is the base of a topology on  $\times_{i \in I} X_i$ .

The topology having this collection  $\mathcal{B}$  as a base is called the **box topology** on the product set  $\times_{i \in I} X_i$ , and the resulting topological space is called the **box product** of the given family  $\langle X_i \rangle_{i \in I}$  of spaces and may be denoted by  $\square_{i \in I} X_i$ . (When each member of the family is the same set topological space  $S$ , we may denote the box product by  $\square^I S$  and refer to it as a **box power** of  $S$ .) A member of the base  $\mathcal{B}$  is sometimes called a **box**, by analogy with the situation when  $I = \{1, 2, 3\}$ , each  $X_i = \mathbb{R}$ , and each  $U_i$  is an open interval.

*Note:* The box topology coincides with the product topology when the index set  $I$  is finite. In general, the box topology is finer than the product topology; it may be strictly finer, even when the index set is denumerable: see the example on page 365 concerning  $\times_{i=1}^{\infty} \{0, 1\}$ .



- (b) Show that each projection  $p_j: \prod_{i \in I} X_i \rightarrow X_j$  is continuous and open.
- (c) Show that the box topology on a product set need not be the coarsest topology making each projection continuous.
- (d) Show that a map  $f: Y \rightarrow \prod_{i \in I} X_i$  need not be continuous even though each composite  $p_j \circ f: Y \rightarrow X_j$  is continuous. [Hint: Take  $Y = \mathbb{R}$ ,  $I = \mathbb{N}$ , and  $X_i = \mathbb{R}$  for each  $i \in \mathbb{N}$ . Then take  $f: Y \rightarrow \prod_{i \in I} X_i$  to be the diagonal map given by  $x \mapsto \langle x, x, x, \dots \rangle$ .]
- Note:* This deviation from the situation for the product topology (as enunciated in [Theorem 3.62](#)) is the fundamental reason the box topology is *not* the “natural” topology to put on the product of sets underlying a family of topological spaces. (For another reason, see [Exercise 4.43](#)).

box product

topology!pointwise convergence@o

pointwise convergence

topology!pointwise convergence@o

pointwise convergence

metrizable space

initial topology

topology!initial

final topology

topology!final

**141.** (Continuation of [Exercise 140](#).)

- (a) Show that the box product of even denumerably many first-countable spaces need not be first-countable—in fact, that the box product of denumerably many separable metrizable spaces need not be either first-countable (hence not metrizable nor second-countable) or separable.
- (b) Must the box product of denumerably many separable spaces be separable?

**142.** Equip the set  $\mathcal{F}(X, Y)$  of all maps from a set  $X$  to a topological space  $Y$  with its topology of pointwise convergence ([Example 3.71](#)).

- (a) Express a subbasic open set in  $\mathcal{F}(X, Y)$  as a product.
- (b) Describe a basic open set in  $\mathcal{F}(X, Y)$  in a form similar to that of (\*), [page 379](#).
- (c) Express a subbasic open set in  $\mathcal{F}(X, Y)$  as a product.

**143.** Provide the set  $\mathcal{F}(X, Y)$  of all maps from a set  $X$  to a *metrizable* space  $Y$  with its topology of pointwise convergence. When is  $\mathcal{F}(X, Y)$  metrizable?**144.** *Note:* This exercise presents a generalization of both subspaces and product spaces.

Let  $X$  be a set, let  $\langle Y_i \rangle_{i \in I}$  be a family of topological spaces, and for each  $i \in I$  let  $f_i: X \rightarrow Y_i$  be a map. Let  $\mathcal{T}$  to be the topology on  $X$  generated by the collection of all subsets of the form  $f_i^{-1}(V)$  where  $i \in I$  and  $V$  is an open subset of  $Y_i$ .

- (a) Show  $\mathcal{T}$  is the *coarsest* topology  $\mathcal{T}$  on  $X$  that makes every  $f_i$  continuous.

We call  $\top$  the **initial topology on  $X$  induced by  $\langle f_i \rangle_{i \in I}$** .

- (b) Let  $g: Y \rightarrow Z$  be a map to a topological space  $Z$ . Show that  $g$  is continuous for the final topology  $\mathcal{T}$  on  $Y$  if and only if the composite  $g \circ f_i: X_i \rightarrow Z$  is continuous for each  $i \in I$ .

Show, further, that  $\mathcal{T}$  is the unique topology on  $Y$  having the preceding property.

- (c) Show that the relative topology on a subset of a topological space is an initial topology.
- (d) Show that the product topology on the product ([Exercise 106](#)) of a family of topological spaces is an initial topology.

(*Note:* There is a “dual” notion of *final topology* that will be discussed in [Exercise 215](#).)

*Exercises 145–151 require a bit of familiarity with groups.*

prob:topological-group **145.** A **topological group** is a group  $G$  provided with a topology that makes the maps

$$\begin{array}{ll} G \times G \rightarrow G, & G \rightarrow G \\ \langle x, y \rangle \mapsto xy & x \mapsto x^{-1} \end{array}$$

topological group  
circle group  
general linear group  
special linear group  
orthogonal group  
special orthogonal group

continuous when  $G \times G$  is given its product topology induced by the topology on  $G$ . (As is usually the case in algebra, here we are eliding the symbol for the group operation.)

Verify that the following are topological groups:

- (a) The multiplicative group  $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$  with its usual topology.
- (b) Any group with its discrete topology.
- (c) The additive group  $\mathbb{R}$  with its usual topology. (Here, of course, we write  $x + y$  instead of  $xy$  and  $-x$  instead of  $x^{-1}$ .)
- (d) The additive group  $\mathbb{R}^n$  with its usual topology.
- (e) The **circle group**—the multiplicative group  $\{z \in \mathbb{C} : |z| = 1\}$ —with its usual topology as a subset of the plane.  
*Note:* Under the usual identification of  $\mathbb{C}$  with  $\mathbb{R}^2$ , the circle group is just the 1-sphere  $S_1$ . Whenever  $S_1$  is considered as a topological group, it is with respect to multiplication of its elements as complex numbers.
- (f) The **general linear group of degree 2**, that is, the multiplicative group  $\text{GL}(2, \mathbb{R})$  of all nonsingular 2-by-2 real matrices provided with the following topology  $\mathcal{T}$ . Let

$$\varphi: \text{GL}(2, \mathbb{R}) \rightarrow \mathbb{R}^4$$

be the injection defined by

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mapsto \langle a_{11}, a_{12}, a_{21}, a_{22} \rangle.$$

Then  $\mathcal{T}$  is the unique topology on  $\text{GL}(2, \mathbb{R})$  that makes  $\varphi$  an embedding. (*Hint:* Write out explicitly the formulas for the product of two matrices and the inverse of a matrix.)

prob:matrix-top-grps **146.** Verify that the following are topological groups:

- prob-part:GLn-is-top-grp (a) For  $n \geq 1$ , the **general linear group of degree  $n$** , that is, the generalization  $\text{GL}(n, \mathbb{R})$  of [Exercise 145 \(f\)](#) to  $n$ -by- $n$  real matrices.
- prob-part:SLn-is-top-grp (b) For  $n \geq 1$ , the **special linear group of degree  $n$** , that is, the subgroup  $\text{SL}(n, \mathbb{R})$  of  $\text{GL}(n, \mathbb{R})$  consisting of all  $n$ -by- $n$  real matrices having determinant 1.
- prob-part:On-is-top-grp (c) For  $n \geq 1$ , the **orthogonal group of degree  $n$** , that is, the subgroup  $\text{O}(n, \mathbb{R})$  of  $\text{GL}(n, \mathbb{R})$  consisting of all orthogonal  $n$ -by- $n$  matrices.  
*Note:* Each member of  $\text{O}(n, \mathbb{R})$  has determinant 1 or  $-1$ .
- (d) For  $n \geq 1$ , the **special orthogonal group of degree  $n$** , that is, the subgroup  $\text{SO}(n, \mathbb{R})$  of  $\text{O}(n, \mathbb{R})$  consisting of all orthogonal  $n$ -by- $n$  matrices having determinant 1.

**147.** For  $n \geq 1$ , the following subgroup inclusions hold among the topological groups in [Exercise 146](#):

$$\text{SO}(n, \mathbb{R}) \subset \text{O}(n, \mathbb{R}) \subset \text{SL}(n, \mathbb{R}) \subset \text{GL}(n, \mathbb{R})$$

Which of these subgroups are closed in the topological groups containing them? which are open?

:multiplications-homeos-in-top-grp

**148.** Fix an element  $a$  in a topological group  $G$ . Show that each of the following maps from  $G$  to  $G$  is a homeomorphism:

topological group

circle!quotient space@as quotient sp

- (a) Left translation  $x \mapsto ax$  by  $a$ .
- (b) Right translation  $x \mapsto xa$  by  $a$ .
- (c) Conjugation  $x \mapsto axa^{-1}$  by  $a$ .

**149.** (a) Prove that a group  $G$  provided with a topology is a topological group if and only if the single map

$$\begin{aligned} G \times G &\rightarrow G \\ \langle x, y \rangle &\mapsto xy^{-1} \end{aligned}$$

is continuous.

- (b) If  $U$  is a neighborhood of the identity element  $e$  of a topological group  $G$ , show that  $U$  contains a neighborhood  $V$  of  $e$  such that  $V^2 \subset U$  and  $V = V^{-1}$ . Here  $V^2$  denotes the subset

$$V^2 = \{xy : x, y \in V\}$$

of  $G$  and  $V^{-1}$  denotes the subset

$$V^{-1} = \{x^{-1} : x \in V\}$$

of  $G$ .

**150.** Let  $G$  be a topological group and let  $e$  be its identity element. Prove:

- (a) If  $V$  is a neighborhood of  $e$ , then  $xV = \{x \cdot v : v \in V\}$  is a neighborhood of  $x$  for each  $x \in G$ .
- (b) As a topological space,  $G$  is first-countable if there is some local base at  $e$ .
- (c) As a topological space,  $G$  is a Hausdorff space if for each  $x \neq e$  there is some neighborhood  $U$  of  $e$  such that  $x \notin U$ .

prob:subgroup-of-topological-group

**151.** Let  $H$  be a subgroup of a topological group. Prove:

- (a) Under its relative topology  $H$  is a topological group.
- (b) The closure of  $H$  in  $G$  is also a subgroup of  $G$ .
- (c) If  $H$  is open in  $G$ , then  $H$  is also closed in  $G$ .

art:open-subgrp-of-topgrp-is-closed

*Note:* For an application, see [Exercise 5.39](#).

### 3.4 Quotient Spaces

sec:quotient

This section continues the theme of constructing new spaces from old. Suppose we are given a topological space and an equivalence relation on the underlying set of that space. How can we put a suitable topology on the quotient of that set under the equivalence relation?

#### The unit circle as a quotient space

subsec:S1-as-quotient-space

Consider the problem of how to represent mathematically a simple physical experiment: bend a straight piece of very thin, flexible wire around until its ends touch. The resulting shape is a circle or—in case the bending was not done uniformly or introduced some kinks or corners—a shape that is topologically the same as a circle.

Represent a circular shape by the unit circle

$$S_1 = \{ \langle x_1, x_2 \rangle \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1 \}$$

and the original, unbent, wire by the closed interval

$$X = [0, 2\pi]$$

whose length is the same as the circumference of the circle. (We could instead use an interval having a different length, so as to represent stretching the wire while bending it, and this would just introduce a slight modification of the analysis below.) To represent bending the wire until its ends touch, *identify* the endpoints 0 and  $2\pi$  of  $X$  with each other—that is, treat them *as if* they were the same point. In strictly mathematical terms, form the equivalence relation  $\sim$  on  $X$  given by

$$t \sim s \iff (t = 0 \text{ and } s = 2\pi) \text{ or } (t = 2\pi \text{ and } s = 0) \text{ or } (t = s).$$

Thus represent the bent piece of wire by the quotient set (Definition 0.93)

$$Y = X/\sim$$

whose points are the equivalence classes

$$\begin{aligned} [0] &= [2\pi] = \{0, 2\pi\}, \\ [t] &= \{t\} \quad (0 < t < 2\pi). \end{aligned}$$

The mathematical problem, then, is to endow the quotient set  $Y$  with a topology making it homeomorphic to the circle  $S_1$ .

To attack this problem, change the point of view and think of wrapping the wire around an existing circular shape (whose circumference equals the length of the wire). In mathematical terms, consider the map

$$\begin{aligned} f: X &\rightarrow S_1 \\ t &\mapsto \langle \cos t, \sin t \rangle \end{aligned}$$

of the interval  $X$  onto the circle  $S_1$ . Here we rely on the familiar properties of the cosine and sine functions guaranteeing that each point  $\langle x_1, x_2 \rangle \in S_1$  has a parametric representation

$$\begin{cases} x_1 = \cos t, \\ x_2 = \sin t \end{cases}$$

for some  $t$  with  $0 \leq t \leq 2\pi$ . In fact, each point of  $S_1$  has a unique representation of this sort *except* the point  $\langle 1, 0 \rangle$ , which has the two representations

$$\begin{cases} 1 = \cos 0, \\ 0 = \sin 0 \end{cases}, \quad \begin{cases} 1 = \cos 2\pi, \\ 0 = \sin 2\pi. \end{cases}$$

In other words, for  $t, s \in X$  we have

$$\{\text{eq:f-on-interval-iff-sim}\} \quad (*) \quad f(t) = f(s) \iff t \sim s.$$

Property  $(*)$  allows us to relate  $Y = X/\sim$  to  $S_1$  by passing to the quotient (Theorem 0.104 and Definition 0.105), that is, by forming the map

$$h: Y \rightarrow S_1$$

given by

$$h([t]) = f(t) \quad (t \in X),$$

or equivalently,

$$h(y) = f(t) \quad (y \in Y, t \text{ a representative of } y).$$

Since  $f$  is surjective and takes distinct values at representatives of different equivalence classes, by [Theorem 0.104](#) the map  $h$  is bijective.

Now endow  $Y$  with its unique topology  $\mathcal{T}$  making  $h$  a homeomorphism—hence making  $Y$  homeomorphic to  $S_1$ —namely,

$$\mathcal{T} = \{U \subset Y : h(U) \text{ is open in } S_1\}.$$

This solves our problem, but not quite in a form suitable for generalization. We seek a description of  $\mathcal{T}$  that makes no reference to  $S_1$  and is expressed purely in terms of the topology of  $X$  and the equivalence relation  $\sim$  on  $X$ .

To do this, let

$$\begin{aligned} q: X &\rightarrow Y = X/\sim \\ t &\mapsto [t] \end{aligned}$$

be the quotient map, which sends each element of  $X$  to its equivalence class under  $\sim$ . Then  $h$  is the unique map from  $Y$  to  $S_1$  satisfying the equation

$$h(q(t)) = f(t) \quad (t \in X),$$

that is,

$$h \circ q = f.$$

Because  $h$  is bijective,

$$\{\text{eq:f-inv-circ-h-is-p-inverse}\} \quad (**) \quad f^{-1}(h(U)) = q^{-1}(U) \quad (U \subset Y).$$

The crucial observation concerns the special nature of  $f$ :

$$\{\text{eq:V-open-iff-f-inv-V-open}\} \quad (***) \quad V \text{ is open in } S_1 \iff f^{-1}(V) \text{ is open in } X \quad (V \subset S_1)$$

In fact,  $f$  is continuous because its component functions  $t \mapsto \cos t$  and  $t \mapsto \sin t$  are continuous; hence  $f^{-1}(V)$  is open in  $X$  if  $V$  is open in  $S_1$ . Conversely, suppose that  $V$  is a subset of  $S_1$  for which  $f^{-1}(V)$  is open in  $X = [0, 2\pi]$ . To complete the proof of condition [\(\\*\\*\\*\)](#), let  $x \in V$  be arbitrary; we shall show that  $V$  includes an open arc that contains  $x$ .

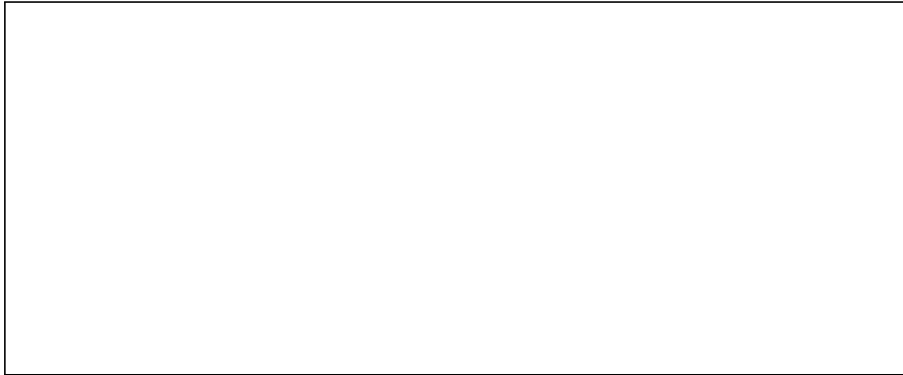


Figure 3.16: Point  $x \in V \subset S_1$  contains an open arc. Case (i):  $x \neq \langle 1, 0 \rangle$ .

fig:arc-case-interior-pt

Case (i):  $x \neq \langle 1, 0 \rangle$ . Then  $f(t) = x$  for a unique  $t \neq 0, 2\pi$ . Since  $f^{-1}(V)$  is an open neighborhood of  $t$  in  $[0, 2\pi]$ , there exist numbers  $s$  and  $u$  with  $0 < s < t < u < 2\pi$  and  $]s, u[ \subset f^{-1}(V)$ . In this case,  $f([s, u])$  is the desired open arc in  $S_1$ . (See [Figure 3.16](#).)

circle!quotient space  
quotient topology  
quotient space

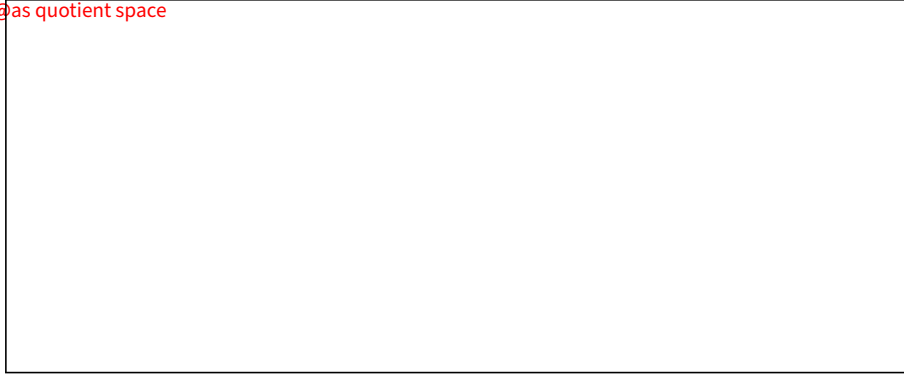


Figure 3.17: Point  $x \in V \subset S_1$  contains an open arc. Case (ii):  $x = \langle 1, 0 \rangle$ .

fig:arc-case-endpt

Case (ii):  $x = \langle 1, 0 \rangle$ . Then  $f(0) = x = f(2\pi)$ ; both numbers 0 and  $2\pi$  belong to  $f^{-1}(V)$ ; and there exist numbers  $s$  and  $u$  with  $0 < s < u < 2\pi$  and  $[0, s[ \cup ]u, 2\pi] \subset f^{-1}(V)$ . In this case,  $f([0, s[ \cup ]u, 2\pi])$  is the desired open arc in  $S_1$ . (See Figure 3.17.)

Combining (\*\*) and (\*\*\*) with the definition of the topology  $\mathcal{T}$  on  $Y$ , we obtain:

$$\begin{aligned} U \in \mathcal{T} &\iff h(U) \text{ is open in } S_1 \\ &\iff q^{-1}(U) = f^{-1}(h(U)) \text{ is open in } X. \end{aligned}$$

Thus

$$\mathcal{T} = \{U \subset Y : q^{-1}(U) \text{ is open in } X\}.$$

This formula does describe—purely in terms of the topological space  $X$  and the equivalence relation  $\sim$  on  $X$ —a topology on the quotient set  $Y = X/\sim$  that makes  $Y$  homeomorphic to  $S_1$ .

### The quotient topology

subsec:quotient-top

We now proceed rapidly to extract from the example in the preceding subsection the general definition of a quotient space, and to formulate some general principles that will allow us to give additional concrete examples.

lem:justify-def-quotient-top

**3.72 Lemma.** Let  $\sim$  be an equivalence relation on a topological space  $X$ , and let  $q: X \rightarrow X/\sim$  be the induced quotient map. Then the collection

$$\mathcal{T} = \{V \subset X/\sim : q^{-1}(V) \text{ is open in } X\}$$

is a topology on the quotient set  $X/\sim$ .

**Proof.** Since  $q^{-1}(\emptyset) = \emptyset$  and  $q^{-1}(X/\sim) = X$ , both  $\emptyset$  and  $X/\sim$  belong to  $\mathcal{T}$ . If  $\langle V_i \rangle_{i \in I}$  is a family of sets belonging to  $\mathcal{T}$ , then  $q^{-1}(V_i)$  is open in  $X$  for each  $i \in I$ , and so

$$q^{-1}\left(\bigcup_{i \in I} V_i\right) = \bigcup_{i \in I} q^{-1}(V_i)$$

is open in  $X$ . Finally, if  $U \in \mathcal{T}$  and  $V \in \mathcal{T}$ , then both  $q^{-1}(U)$  and  $q^{-1}(V)$  are open in  $X$ , and so  $q^{-1}(U \cap V) = q^{-1}(U) \cap q^{-1}(V)$  is open in  $X$ .  $\square$

The preceding lemma justifies the following definition.

def:quot-top

**3.73 Definition.** The **quotient space** of a topological space  $X$  under an equivalence relation  $\sim$  is the topological space whose underlying set is the quotient set  $X/\sim$  and whose topology is the collection

$$\{V \subset X/\sim : q^{-1}(V) \text{ is open in } X\},$$

where  $q: X \rightarrow X/\sim$  is the quotient map. This topology is called the **quotient topology (induced by  $\sim$ )**.

**Convention!** Unless otherwise indicated, *any future reference to a topology on a quotient set of a topological space is to the quotient topology*.

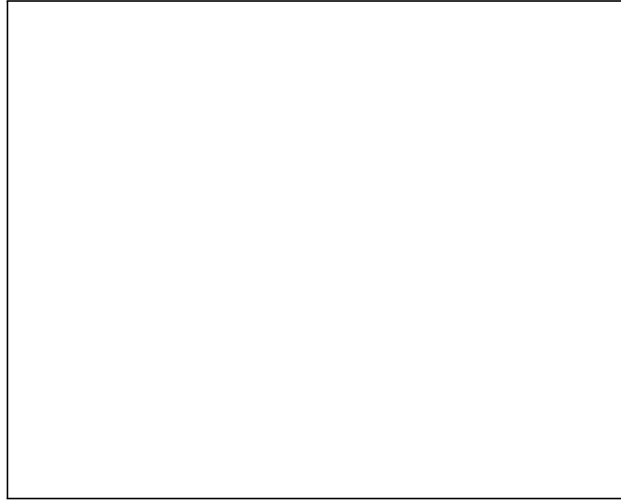


Figure 3.18: An open set in a quotient space and its inverse image under the quotient map.

fig:open-set-quotient-space

A set  $V \subset X/\sim$  has as its points equivalence classes in  $X$  and has as its inverse image

$$q^{-1}(V) = \{x \in X : [x] \in V\} = \bigcup \{[x] : x \in X, [x] \in V\},$$

the union of those equivalence classes belonging to  $V$ . Hence **an open set in the quotient space  $X/\sim$  is just a collection of equivalence classes whose union is an open subset of  $X$** . [The relation between  $V$  and  $q^{-1}(V)$  is indicated schematically in Figure 3.18, where equivalence classes in  $X$  are indicated by vertical line segments.]

Recall from Definition 0.98 that a subset  $U$  of  $X$  is said to be *saturated* by the equivalence relation  $\sim$  when

$$x \in U \text{ and } y \sim x \implies y \in U,$$

that is, when  $U$  is a union of equivalence classes under  $\sim$ . If  $V$  is a subset of  $X/\sim$ , then its inverse image  $q^{-1}(V)$  is saturated by  $\sim$ ; and if  $U$  is a saturated subset of  $X$ , then  $U = q^{-1}(V)$  for a subset of  $X/\sim$ , namely, for  $V = q(U)$ .

The quotient topology may now be described as

$$\{q(U) : U \text{ is a saturated open subset of } X\}.$$

Thus each open set  $V$  in  $X/\sim$  is the image under  $q$  of an open subset of  $X$ , namely of the saturated open subset  $q^{-1}(V)$  of  $X$ .

**Caution!** The image  $q(U)$  of an open subset of  $X$  need not be open in the quotient space  $X/\sim$ , that is, **the quotient map  $q: X \rightarrow X/\sim$  need not be an open map**. For example, the set  $q([0, 2\pi[)$  is not open in the quotient space  $[0, 2\pi]/\sim$  considered in the previous subsection (page 385), because the set  $q^{-1}(q([0, 2\pi[)) = [0, \pi[ \cup \{2\pi\}$  is not open in  $[0, 2\pi]$ .

By definition of the quotient topology, if  $V$  is a subset of  $X/\sim$ , then

$$V \text{ is open in } X/\sim \iff q^{-1}(V) \text{ is open in } X.$$

Dually, if  $E$  is a subset of  $X/\sim$ , then

$$E \text{ is closed in } X/\sim \iff q^{-1}(E) \text{ is closed in } X.$$

To see this, let  $Y = X/\sim$ . Then  $E$  is closed in  $Y$  if and only if  $Y \setminus E$  is open in  $Y$ , that is  $q^{-1}(Y \setminus E)$  is open in  $X$ . But  $q^{-1}(Y \setminus E) = X \setminus q^{-1}(E)$ , and the latter set is open in  $X$  if and only if  $q^{-1}(E)$  is closed in  $X$ .

### Quotient spaces and continuous maps

Our definition of the quotient topology was a generalization of a single example. Nonetheless, among all topologies on a quotient set, the quotient topology is the most natural one, as the following characterization demonstrates.

**3.74 Proposition.** Let  $q: X \rightarrow X/\sim$  be the quotient map induced by an equivalence relation on a topological space  $X$ . Then the quotient topology is the finest topology on  $X/\sim$  that makes  $q$  continuous.

**Proof.** By its very definition, the quotient topology  $\mathcal{T}$  certainly makes  $q$  continuous. Now let  $\mathcal{S}$  be an arbitrary topology on  $X/\sim$  that makes  $q$  continuous. If  $V \in \mathcal{S}$ , then  $q^{-1}(V)$  is open in  $X$  by assumption, and so  $V \in \mathcal{T}$  by definition of  $\mathcal{T}$ . Hence  $\mathcal{S} \subset \mathcal{T}$ .  $\square$

Another reason that the quotient topology is “natural” is that the continuous maps on a quotient of a space are determined by the continuous maps on that space.

**3.75 Theorem.** A map

$$g: X/\sim \rightarrow Y$$

from a quotient space  $X/\sim$  into a topological space  $Y$  is continuous if and only if its composite

$$g \circ q: X \rightarrow Y$$

with the quotient map  $q: X \rightarrow X/\sim$  is continuous.

**Proof.** Let  $g$  be such a map. Since the projection  $q$  is continuous, then  $g \circ q$  is continuous if  $g$  is. Conversely, assume that  $g \circ q$  is continuous. To see that  $g$  is then continuous, let  $W$  be an arbitrary open subset of  $Y$ . By assumption, the set

$$q^{-1}(g^{-1}(W)) = (g \circ q)^{-1}(W)$$



is open in  $X$ . Hence  $g^{-1}(W)$  is open in  $X/\sim$ .  $\square$

passing to the quotient

A composite

$$f = g \circ q: X \rightarrow Y$$

of the kind just considered behaves in a very special way with respect to the equivalence relation  $\sim$ . Namely, if  $x$  and  $t$  are elements of  $X$  that belong to the same equivalence class under  $\sim$ , that is, if  $x \sim t$ , then  $q(x) = q(t)$  and hence  $f(x) = g(q(x)) = g(q(t)) = f(t)$ . In short, ***f is constant on each equivalence class under  $\sim$*** . Conversely, a map on  $X$  that behaves in this way determines a map on  $X/\sim$ . For example, in the subsection “The unit circle as a quotient space” (page 385), the map  $h: [0, 2\pi] \rightarrow S_1$  was constructed from  $f: [0, 2\pi] \rightarrow S_1$  in this very manner.

For such a map  $f$ , we may pass to the quotient as in [Theorem 0.104](#), but now the maps involved will be continuous ones.

thm:pass-to-quotients

**3.76 Theorem (passing to the quotient).** *Let  $q: X \rightarrow X/\sim$  be the quotient map induced by an equivalence relation on a topological space  $X$ . Let  $f: X \rightarrow Y$  be a continuous map from  $X$  to a topological space  $Y$  that is constant on each equivalence class under  $\sim$ . Then there is a unique continuous map*

$$f^*: X/\sim \rightarrow Y$$

*such that*

$$f^* \circ q = f.$$

*Moreover:*

thm-part:fstar-surj-iff-f-is

(1) *The map  $f^*$  is surjective if and only if  $f$  is surjective.*

distinct-vals-on-distinct-equiv-classes

(2) *The map  $f^*$  is injective if  $f$  takes distinct values at representatives of different equivalence classes under  $\sim$ .*

thm-part:when-fstar-is-open

(3) *The map  $f^*$  is open if  $f$  takes distinct values at representatives of different equivalence classes under  $\sim$  and the open subsets of  $Y$  are those subsets  $W$  of  $Y$  for which  $f^{-1}(W)$  is open in  $X$ .*

**Proof.** The existence of a unique map  $f^*$  satisfying  $f^* \circ q = f$  follows from [Theorem 0.104](#). So do assertions (1) and (2). All that is purely set-theoretic and has nothing to do with any topology.

Continuity of such  $f^*$  follows from that of  $f$  by [Theorem 3.75](#).

(3) Assume that  $f$  takes distinct values at representatives of different equivalence classes and that the open sets in  $Y$  are as stated in (3). Let  $V$  be open in  $X/\sim$ . To show that  $f^*(V)$  is open in  $Y$ , we need therefore only check that  $f^{-1}(f^*(V))$  is open in  $X$ . And to do this we need in turn only to show that

$$f^{-1}(f^*(V)) = q^{-1}(V),$$

because  $q^{-1}(V)$  is open in  $X$  by continuity of  $p$ .

passing to the quotient  
quotient map!abstract  
abstract quotient map

If  $x \in q^{-1}(V)$ , then  $q(x) \in V$ ,

$$f(x) = f^*(q(x)) \in f^*(V),$$

and so  $x \in f^{-1}(f^*(V))$ . Conversely, let  $x \in f^{-1}(f^*(V))$ . Then  $f(x) = f^*(v)$  for some  $v \in V$ , so that

$$f^*(q(x)) = f(x) = f^*(v).$$

Now  $q(x) = v$  because  $f^*$  is injective. Hence  $x \in q^{-1}(V)$ .  $\square$

cor:pass-to-quotients-homeo

**3.77 Corollary.** Let  $q: X \rightarrow X/\sim$  be the quotient map induced by an equivalence relation on a topological space  $X$ . Let  $f: X \rightarrow Y$  be a continuous map from  $X$  into a topological space  $Y$ . Then the map obtained by passing to the quotient will be a homeomorphism  $f^*: X/\sim \cong Y$  when  $f$  has the three properties:

cond:f-surj-pass-to-quotient

(i)  $f$  is surjective;

d:f-on-equiv-classes-pass-to-quotient

(ii)  $f$  is constant on each equivalence class under  $\sim$  and takes distinct values at different equivalence classes; and

cond:open-images-pass-to-quotient

(iii) the open subsets of  $Y$  are those whose inverse images under  $f$  are open in  $X$ .

If the conditions (i)–(iii) in the preceding corollary hold and if  $Y$  is, say, a subspace of a Euclidean space, then we shall be able to realize the “abstract” space  $X/\sim$  as the “concrete” space  $Y$ . Several such realizations are exhibited in [Examples 3.81](#), below; additional ones appear in the exercises.

### Quotient maps

subsec:quotient-maps

The observation at the end of the preceding subsection suggests it may be possible to represent a given quotient space by a “concrete” space homeomorphic to it by constructing a map  $f$  having the topological behavior that the open subsets of  $Y$  are those whose inverse images under  $f$  are open in  $X$ . Hence we pause to look at such maps in general, without reference to quotient spaces.

def:quotient-map

**3.78 Definition.** A **quotient map** from a topological space  $X$  to a topological space  $Y$  is a *surjection* such that, for each  $V \subset Y$ ,

$$V \text{ is open in } Y \iff f^{-1}(V) \text{ is open in } X.$$

Equivalently, a surjection  $f: X \rightarrow Y$  will be a quotient map when, for each  $E \subset Y$ ,

$$E \text{ is closed in } Y \iff f^{-1}(E) \text{ is closed in } X.$$

Necessarily, such a quotient map will be continuous.

To distinguish this meaning of ‘quotient map’ from that of [Definition 3.73](#), we may refer to a quotient map in the sense of [Definition 3.78](#) as an *abstract* quotient map.

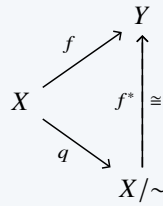
The quotient map  $q: X \rightarrow X/\sim$  induced by an equivalence relation  $\sim$  on a topological space  $X$  is an abstract quotient map. Conversely, an abstract quotient map is *essentially* a quotient map in the original sense of the term. To what this means, recall from

**Examples 0.90 (7)** that the *equivalence kernel* of a map  $f: X \rightarrow Y$  is the equivalence relation  $\sim$  on the domain of the map given by

$$x \sim u \iff f(x) = f(u) \quad (x, u \in X).$$

Then the “is essentially” statement has the following precise meaning; the proof is left to the reader (**Exercise 157**).

**3.79 Proposition.** Let  $f: X \rightarrow Y$  be a quotient map. Let  $\sim$  be the equivalence kernel of  $f$  and let  $q: X \rightarrow X/\sim$  be the quotient map induced by  $\sim$ . Then there is a unique homeomorphism  $f^*: X/\sim \rightarrow Y$  such that the following diagram commutes:



In short, the codomain of an abstract quotient map in the sense of **Definition 3.78** is a quotient space  $X/\sim$  “up to homeomorphism.”

For further examples of abstract quotient maps we may rely on the following proposition.

**3.80 Proposition.** Each continuous open surjection and each continuous closed surjection is a quotient map.

**Proof.** Let  $f: X \rightarrow Y$  be a continuous surjection from a space  $X$  to a space  $Y$ . Already  $f^{-1}(V)$  is open in  $X$  for each open subset  $V$  of  $Y$ .

Assume first that  $f$  is an open map. Let  $V$  be a subset of  $Y$  whose inverse image  $f^{-1}(V)$  is open in  $X$ . Then  $V = f(f^{-1}(V))$  is open in  $Y$ .

Assume next that  $f$  is a closed map. Again let  $V$  be a subset of  $Y$  whose inverse image  $f^{-1}(V)$  is open in  $X$ . Now use an argument like the one for an open map, but take complements.  $\square$

**Caution!** A quotient map need not be either open nor closed: see **Exercise 155**, for example.

As we shall see later, however, a large class of quotient maps  $X \rightarrow Y$  with a  $T_2$ -codomain  $Y$  are, in fact, closed maps; these include such maps whose domain is a closed subspace of some  $\mathbb{R}^n$  that is  $d$ -bounded in  $\mathbb{R}^n$  for the Euclidean metric  $d$ .

### Concrete realizations of quotient spaces

At last we are in a position to discuss additional examples of quotient spaces, many of which we can realize “concretely.”

**3.81 Examples.** (1) Let  $X = I = [0, 1]$ . Identify the endpoints 0 and 1 of  $X$  by means of the equivalence relation  $\sim$  given by

$$t \sim s \iff (t = 0 \text{ and } s = 1) \text{ or } (t = 1 \text{ and } s = 0) \text{ or } (t = s).$$

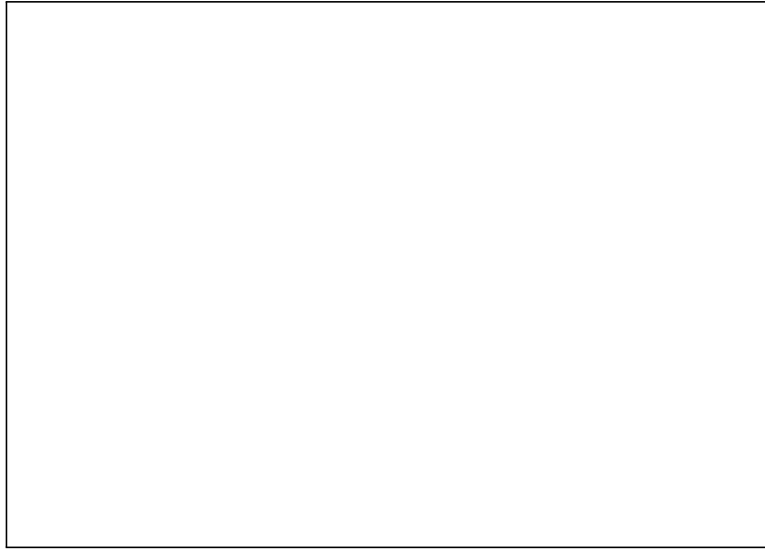


Figure 3.19: Identifying ends of the unit interval  $I$  to obtain the unit circle  $S_1$ .

fig:S1-quotient-of-I

It is geometrically plausible that  $X/\sim \cong S_1$  (see Figure 3.19). To prove, this consider the continuous surjection

$$\begin{aligned} f: X &\rightarrow S_1 \\ t &\mapsto \langle \cos 2\pi t, \sin 2\pi t \rangle \end{aligned}$$

Then

$$f(t) = f(s) \iff t \sim s \quad (t, s \in X),$$

which means that  $f$  is constant on each equivalence class under  $\sim$  and takes distinct values at representatives of different equivalence classes. It remains only to show that  $f$  is a quotient map in the sense of Definition 3.78, for then by Theorem 3.76 the map  $f$  induces a homeomorphism  $f^*: X/\sim \rightarrow S_1$ .

We show in fact that  $f$  is a closed map; from Proposition 3.80 it will follow that  $f$  is a quotient map. Let  $E$  be a closed subset of  $X$  and let  $z \in S_1$  with  $z \notin f(E)$ . We shall find an open arc in  $S_1$  that contains  $z$  and is disjoint from  $f(E)$ .

Case (i):  $z \neq (1, 0)$ . Choose  $t \in X$  with  $f(t) = z$ . Then  $t \neq 0, 1$  and  $t \notin E$ . There is an open interval  $U \subset ]0, 1[$  that is disjoint from  $E$  and contains  $t$ . Then  $f(U)$  is the desired open arc.

Case (ii):  $z = (1, 0)$ . Then  $f(0) = z$  and  $0 \notin E$ . There is an interval  $[0, a[ \subset [0, 1]$  that is disjoint from  $E$ ; similarly, there is an interval  $]b, 1] \subset [0, 1]$  that is disjoint from  $E$ . Then the union  $f([0, a[) \cup f(]b, 1])$  is the desired open arc.

txt:cont-surj-closed-special-arg

*Note:* In this example and in some others below we resort to a special argument to verify that a continuous surjection  $f: X \rightarrow Y$  is closed. Corollary 4.24 will often provide a shortcut for such a verification: A continuous map from a “compact” space to a Hausdorff space is necessarily a closed map. Now it so happens that any subspace of  $\mathbb{R}^n$  that is both closed in  $\mathbb{R}^n$  and  $d$ -bounded for the Euclidean metric  $d$  is compact. Hence: **A continuous surjection from a closed,  $d$ -bounded subspace of  $\mathbb{R}^n$  to a Hausdorff space is a closed map.** Feel free to apply this result to examples even now, before it has been established!

ex:Sone-quotient-of-R

(2) The equivalence relation  $\sim$  on  $I = [0, 1]$  above may also be described by

$$t \sim s \iff t - s = -1 \text{ or } t - s = 1 \text{ or } t - s = 0$$

and hence by

$$t \sim s \iff t - s \in \mathbb{Z}.$$

The latter recipe also defines an equivalence relation on the real line  $\mathbb{R}$  that identifies any two points an integral distance apart, and for which

$$[t]_{\sim} = \{\dots, t-2, t-1, t, t+1, t+2, \dots\}$$

for each  $t \in \mathbb{R}$ .

The quotient space

$$\mathbb{R}/\sim \cong S_1.$$

To see this, consider the continuous surjection

$$\begin{aligned} f: \mathbb{R} &\rightarrow S_1 \\ t &\mapsto \langle \cos 2\pi t, \sin 2\pi t \rangle. \end{aligned}$$

The graphs of the functions  $t \mapsto \cos 2\pi t$  and  $t \mapsto \sin 2\pi t$  are shown in [Figure 3.20](#),

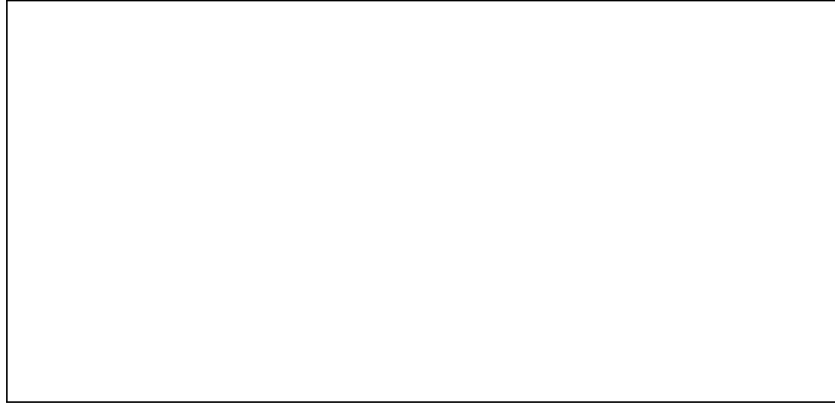


Figure 3.20: Graphs of period 1 cos and sin functions.

fig:cos-sin-to-wrap

are obtained from the graphs of  $t \mapsto \cos t$  and  $t \mapsto \sin t$  by rescaling the  $t$ -axis so that they have period 1 instead of  $2\pi$ . Inspection of the graphs reveals that

$$\cos 2\pi t = \cos 2\pi s \text{ and } \sin 2\pi t = \sin 2\pi s \iff t = s + n \text{ or some integer } n,$$

in other words,

$$f(t) = f(s) \iff t \sim s.$$

Thus  $f$  is constant on each equivalence class under  $\sim$  and takes distinct values at representatives of different equivalence classes. Hence  $f$  induces a continuous bijection  $f^*: \mathbb{R}/\sim \rightarrow S_1$ . The map  $f$  is an open map because it carries each open interval in  $\mathbb{R}$  onto an open arc on  $S_1$ . Hence  $f$  is a quotient map, and so  $f^*$  is a homeomorphism.

ex:bded-cylinder-quotient-of-square (3) Let

$$X = I \times I = [0, 1] \times [0, 1],$$

the (filled-in) unit square in the plane. Define  $\alpha$  to be the relation in  $X$  to itself that identifies each point on the left-hand vertical edge of  $X$  with the point on the right-hand vertical edge at the same height:

$$\langle 0, v \rangle \alpha \langle 1, v \rangle. \quad (v \in I)$$

Now let  $\sim$  be the least equivalence relation on  $X$  containing  $\alpha$ , that is,  $\sim$  is the reflexive, symmetric, transitive closure (Definition 0.91) of  $\alpha$ .

To obtain  $\sim$  requires merely enlarging  $\alpha$  so as to include ordered pairs of identical points of  $X$  along with the reverse  $\langle \langle 1, v \rangle, \langle 0, v \rangle \rangle$  of each ordered pair  $\langle \langle 0, v \rangle, \langle 1, v \rangle \rangle$  already in  $\alpha$ . Thus, for  $\langle t, s \rangle, \langle u, v \rangle \in X$ ,

$$\begin{aligned} \langle t, s \rangle \sim \langle u, v \rangle &\iff \langle t, s \rangle = \langle u, v \rangle \\ &\text{or } (t = 0, u = 1, s = v) \\ &\text{or } (t = 1, u = 0, s = v). \end{aligned}$$

Doing or imagining the simple experiment of rolling a square of paper into a tube should lead you predict that

$$X/\sim \cong S_1 \times I,$$

a (bounded) cylinder, as shown in Figure 3.21.

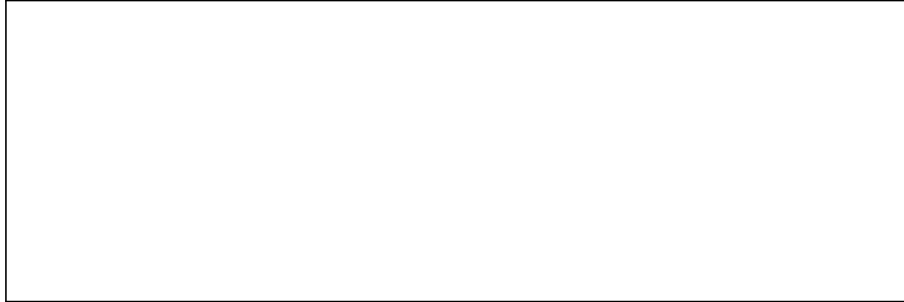


Figure 3.21: Identify opposite vertical edges of a square to obtain a cylinder.

fig:bded-cylinder-quotient-of-square

To confirm this prediction, consider the map

$$\begin{aligned} g: X &\rightarrow S_1 \times I \\ \langle t, s \rangle &\mapsto \langle f(t), s \rangle \end{aligned}$$

where  $f: I \rightarrow S_1$  is the map used in Example (1). Then  $g$  is a continuous surjection such that

$$g(t, s) = g(u, v) \iff \langle t, s \rangle \sim \langle u, v \rangle.$$

Hence by passing to the quotient we obtain from  $g$  a continuous bijection  $g^*: X/\sim \rightarrow S_1 \times I$ . To show that  $g^*$  is a homeomorphism, it suffices to show that  $g$  is a closed map. Although this would follow immediately from the fact that  $S_1 \times I$  is a Hausdorff space and  $X$  is a closed and  $d$ -bounded subset of  $\mathbb{R}^2$  for the euclidean metric  $d$ —see the note on page 393, just after the end of Example (1)—let us verify it directly here.

Let  $E$  be an arbitrary closed subset of  $X$  and let  $\langle z, s \rangle \in S_1 \times I$  with  $\langle z, s \rangle \notin g(E)$ . We shall find an open neighborhood of  $\langle z, s \rangle$  in  $S_1 \times I$  that is disjoint from  $g(E)$ .

Case (i):  $\langle z, s \rangle \notin \{1, 0\} \times I$ . In this case, choose  $t$  with  $f(t) = z$ , so that  $g(t, s) = \langle z, s \rangle$ . Then  $t \neq 0, 1$  and  $\langle t, s \rangle \notin E$ . There is an open neighborhood of  $\langle t, s \rangle$  in  $X$  of the form  $U \times V$ , where  $U$  is an open interval contained in the open interval  $]0, 1[$ . Then  $g(U \times V) = f(U) \times V$ , the product of an open arc in  $S_1$  and an open set in  $I$ , is the desired neighborhood of  $\langle z, s \rangle$ .

Case (ii):  $\langle z, s \rangle \in \{1, 0\} \times I$ . In this case,  $f(0) = z = f(1)$ , so that  $g(0, s) = \langle z, s \rangle = g(1, s)$ . Since  $\langle 0, s \rangle \notin E$  and  $\langle 1, s \rangle \notin E$ , there are half-open, half-closed intervals  $[0, a[$  and  $]b, 1]$  contained in  $[0, 1]$  and an open neighborhood  $V$  of  $s$  in  $I$  such that the product  $([0, a[ \cup ]b, 1]) \times V$  is disjoint from  $E$ . Then the image under  $g$  of this product is the desired neighborhood of  $\langle z, s \rangle$ .

**Convention!** From now on, *whenever we refer to a relation used to form a quotient space, it will be implicit that the reflexive, symmetric, transitive closure of the given relation is the equivalence relation being used.*

ex:torus-quotient-of-square

- (4) Again let  $X = I \times I$ . This time, the equivalence relation  $\sim$  on  $X$  will identify not only points on the vertical edges that were identified in [Example \(3\)](#), but also points on the horizontal edges having the same distance from the left-hand edge, as indicated in [Figure 3.22](#). In view of the convention stated above, this means that  $\sim$  is given by

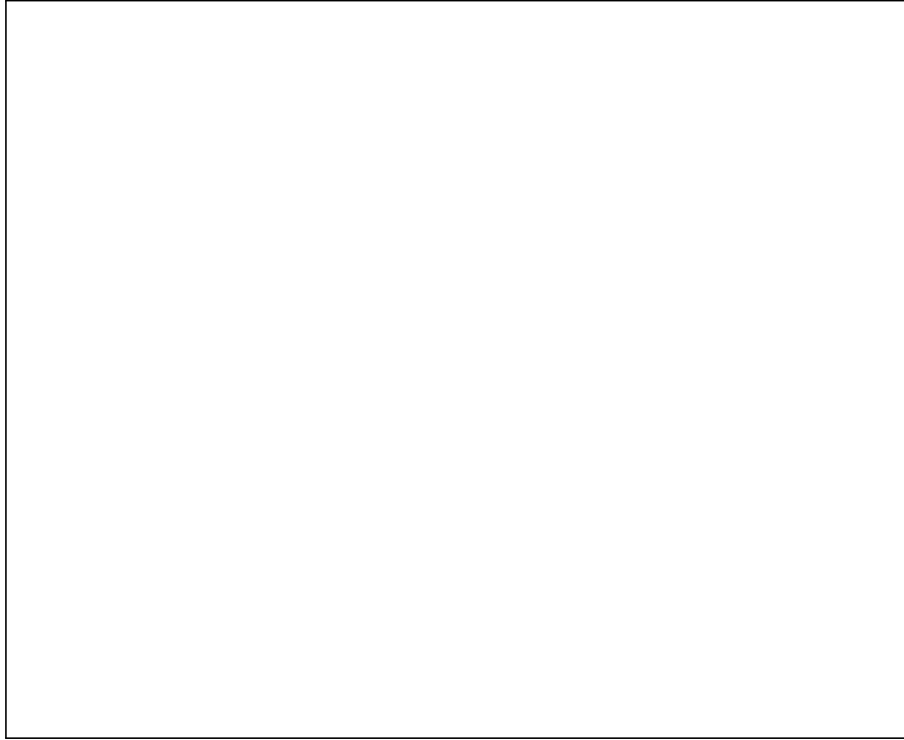


Figure 3.22: Identify opposite edges of a square to obtain a torus.

fig:torus-quotient-of-square

torus  
n-torus@n-torus

$$\begin{aligned} \langle t, s \rangle \sim \langle u, v \rangle &\iff \langle t, s \rangle = \langle u, v \rangle \\ &\text{or } (t = 0, u = 1, s = v) \\ &\text{or } (t = 1, u = 0, s = v) \\ &\text{or } (t = u, s = 0, v = 1) \\ &\text{or } (t = u, s = 1, v = 0), \end{aligned}$$

We have

$$X/\sim \cong S_1 \times S_1.$$

In fact, the map

$$\begin{aligned} g: X &\rightarrow S_1 \times S_1 \\ \langle t, s \rangle &\mapsto \langle f(t), f(s) \rangle \end{aligned}$$

where  $f: I \rightarrow S_1$  is again the map considered in [Example \(1\)](#), is a continuous surjection such that

$$g(t, s) = g(u, v) \iff \langle t, s \rangle \sim \langle u, v \rangle.$$

Hence by passing to the quotient we obtain from  $g$  a continuous bijection  $g^*: X/\sim \rightarrow S_1 \times S_1$ . You may directly show that  $g$  is a closed map: just modify the argument used in [Example \(3\)](#), except that now for  $\langle z, w \rangle \notin g(E)$ , you must consider separately whether  $z$  is or is not  $\langle 1, 0 \rangle$  and whether  $w$  is or is not  $\langle 1, 0 \rangle$ . Thus  $g^*$  is a homeomorphism.

The product  $T_2 = S_1 \times S_1$ , to which our quotient is homeomorphic, is called the **torus**, or the **2-torus**, (the product of  $n$  circles being the  **$n$ -torus**  $T_n$ ).

This product  $S_1 \times S_1$  is a subspace of  $\mathbb{R}^2 \times \mathbb{R}^2 \cong \mathbb{R}^4$ , Euclidean 4-space. Actually, the torus is embeddable in Euclidean 3-space  $\mathbb{R}^3$ : it is homeomorphic to the doughnut-shaped surface  $T$  in  $\mathbb{R}^3$  obtained by rotating the circle

$$C = \{ \langle 0, y, z \rangle \in \mathbb{R}^3 : (y - 2)^2 + z^2 = 1 \}$$

in the  $yz$ -plane about the  $z$ -axis; the surface  $T$  and its generating circle  $C$  are depicted in [Figure 3.23](#).



Figure 3.23: Rotate a circle in the  $yz$ -plane about the  $z$ -axis to obtain a torus.

fig:rotate-circle-obtain-torus

A homeomorphism  $h: S_1 \times S_1 \cong T$  may be constructed explicitly as follows (see [Figure 3.24](#)). Let  $\langle u, v \rangle = \langle \langle u_1, u_2 \rangle, \langle v_1, v_2 \rangle \rangle \in S_1 \times S_1$ . Identify the first factor  $S_1$





Figure 3.24: Constructing a homeomorphism of the torus  $S_1 \times S_1$  with the surface  $T$  in  $\mathbb{R}^3$ .

Möbius strip@Möbius strip  
Möbius, August Ferdinand  
Listing, Johann Benedict  
Klein bottle

fig:construct-homeo-torus-with-surf

of  $S_1 \times S_1$  with the circle  $C$  by identifying each  $u = \langle u_1, u_2 \rangle$  with the corresponding point  $\langle 0, u_1 + 2u_2 \rangle \in C$ . Draw the circle  $C_u$  that passes through this point, is parallel to the  $xy$ -plane, and has its center on the  $z$ -axis. Identify the second factor  $S_1$  of  $S_1 \times S_1$  with the circle  $x^2 + y^2 = 1$  in the  $xy$ -plane by identifying each  $v = \langle v_1, v_2 \rangle$  with the corresponding point  $\langle v_1, v_2, 0 \rangle$  on that circle. Next, construct the half-plane  $P_v$  that passes through this point and has the  $z$ -axis as an edge. Finally, take  $h(u, v)$  to be the unique point at which the circle  $C_u$  intersects the half-plane  $P_v$ .

ex:Möbius-strip-quotient-of-square

- (5) Once again let  $X = I \times I$ . As in [Example \(3\)](#), only points on the vertical edges will be identified, but this time the left and right edges will be identified in *opposite* directions. In other words, we define an equivalence relation  $\sim$  on  $X$  by

$$\begin{aligned} \langle t, s \rangle \sim \langle u, v \rangle &\iff \langle t, s \rangle = \langle u, v \rangle \\ &\text{or } (t = 0, u = 1, v = 1 - s) \\ &\text{or } (t = 1, u = 0, v = 1 - s). \end{aligned}$$

The quotient space  $X/\sim$  is embeddable in Euclidean 3-space  $\mathbb{R}^3$ : it is homeomorphic to the surface  $M$  shown in [Figure 3.25](#). Both the quotient space  $X/\sim$  and the topologically equivalent surface  $M$  are known as the **Möbius strip**, so named after one of its discoverers, August Ferdinand Möbius. (It was discovered independently, in the same year 1858, by Johann Benedict Listing.) To make a physical model of the surface  $M$ , give a half-twist to a rectangular strip of paper and then glue its ends together.

Some of the fascinating, and even surprising, properties of the Möbius strip are explored in the exercises.

Klein bottle  
Klein, Felix  
Klein bottle

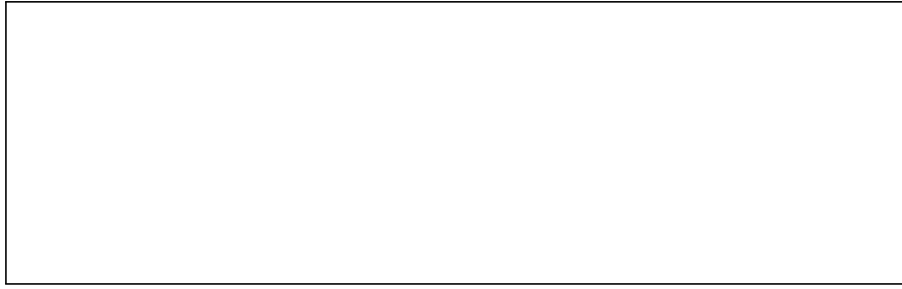


Figure 3.25: Identify edges of a square to obtain the Möbius strip.

fig:Möbius-strip-quotient-of-square

ex:Klein-bottle-quotient-of-square

- (6) For the final time let  $X = I \times I$ . This time the equivalence relation  $\sim$  on  $X$  identifies the vertical edges in opposite directions (as with the Möbius strip) but *also* identifies the horizontal edges in the same direction. In other words,

$$\begin{aligned} \langle 0, s \rangle &\sim \langle 1, 1 - s \rangle & (0 \leq s \leq 1), \\ \langle t, 0 \rangle &\sim \langle t, 1 \rangle & (0 \leq t \leq 1), \end{aligned}$$

as indicated in Figure 3.26. The quotient space  $K = X/\sim$  is the **Klein bottle**, named

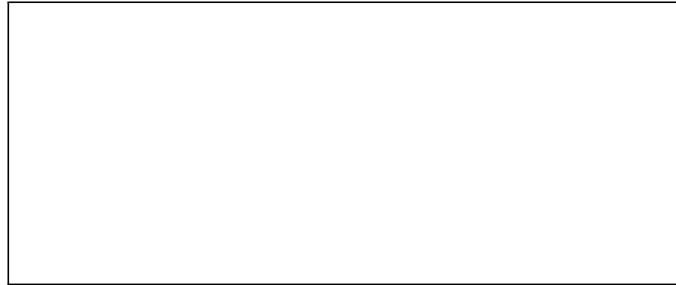


Figure 3.26: How edges of the square are identified in order to the Klein bottle.

fig:identifying-edges-of-square-to-g

after its discoverer Felix Klein.

It is a fact that *the Klein bottle  $K$  is not embeddable in Euclidean 3-space  $\mathbb{R}^3$* , even though  $K$  is a “surface”—a two-dimensional manifold. To see intuitively why not, inspect the cylinder in Figure 3.26, where so far only the top and bottom edges of the square have been identified; the circular ends of this cylinder must still be identified in the directions shown in order to obtain the Klein bottle  $K$ . You might imagine making the latter identification in the way suggested by Figure 3.27. by pushing in the cylinder’s right end, grabbing hold of the cylinder’s left end and pushing it through the cylinder’s side, and then joining the two ends. Of course each point on the small circle at which the pushed cylinder intersects itself on its side represents an identification of two points in the interior of  $I \times I$  that are not supposed to be identified under the equivalence relation  $\sim$ . Hence the subspace of  $\mathbb{R}^3$  shown at the right in Figure 3.27 is *not* the Klein bottle! Rather, it is the image of the Klein bottle  $K$  under a “local homeomorphism”—a map that is a homeomorphism when restricted to a suitably small neighborhood of each point of  $K$ .

Although it is not embeddable in  $\mathbb{R}^3$ , the Klein bottle *is* embeddable in  $\mathbb{R}^4$ —see Exercise 213.

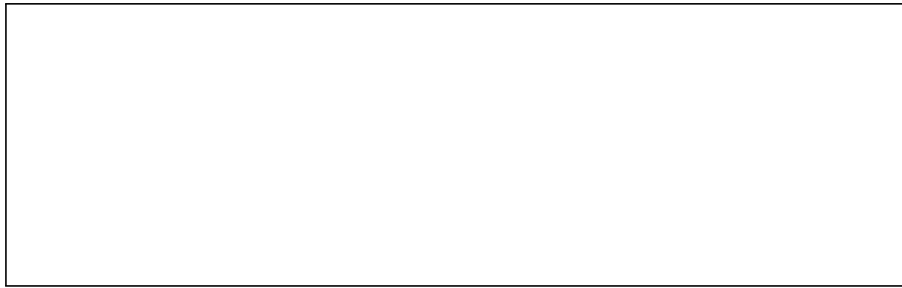


Figure 3.27: Trying to identify the end circles of a cylinder inside  $\mathbb{R}^3$  to obtain the Klein bottle.

projective  $n$ -space  
projective plane  
projective point  
projective line  
affine plane

fig:try-push-cylinder-end-through-la

ex:projective-space

- (7) Fix  $n \geq 1$  and let  $\sim$  be the equivalence relation on the  $n$ -sphere  $S_n \subset \mathbb{R}^{n+1}$  that identifies an arbitrary pair of antipodal (that is, diametrically opposite) points with each other:

$$x \sim y \iff x = y \text{ or } x = -y.$$

The quotient space  $S_n/\sim$  is **(real) projective  $n$ -space**, denoted by  $\mathbb{RP}_n$ . In particular,  $\mathbb{RP}_1$  is the **(real) projective line** and  $\mathbb{RP}_2$  is the **(real) projective plane**.

The projective spaces  $\mathbb{RP}_n$  arise in projective geometry, as we indicate for dimension  $n = 2$ . Let  $q: S_2 \rightarrow \mathbb{RP}_2$  be the quotient map induced by  $\sim$ . Call the image  $q(x) = \{x, -x\}$  of a point  $x \in S_n$  a “*projective point*” and call the image  $q(C)$  of a great circle  $C \subset S_2$  a “*projective line*.” Since any two points of  $S_2$  that are not antipodal lie on a unique great circle, and since two distinct great circles intersect in exactly one pair of antipodal points (see Figure 3.28), it follows that:

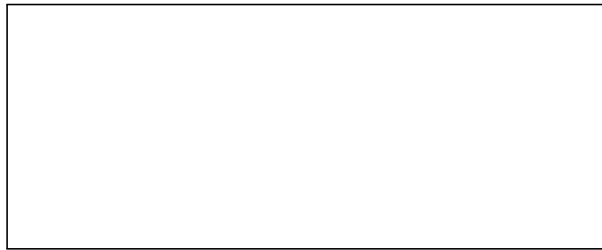


Figure 3.28: Antipodal points and great circles on the 2-sphere  $S_2$ .

fig:antipodal-pts-and-great-circles

projective-pts-determine-projective-line  
projective-lines-intersect-at-projective-pt

(P1) Any two distinct projective points lie on a unique projective line.

(P2) Any two distinct projective lines intersect at a unique projective point.

Property (P1) is just like the corresponding property of “ordinary” points and lines in the plane  $\mathbb{R}^2$  (which, when only the incidence properties of its points and lines are under discussion, should be called the “affine plane”). However, property (P2) exhibits a radical difference between projective lines and ordinary lines in that two lines in the affine plane intersect only if they are not parallel. To explain the difference, we first derive another description of  $\mathbb{RP}_2$ .

hemisphere  
equator

The equivalence relation  $\sim$  on the 2-sphere  $S_2$  induces an equivalence relation  $\simeq$  on its lower (southern) hemisphere

$$S_2^- = \{ \langle x_1, x_2, x_3 \rangle \in S_2 : x_3 \leq 0 \}$$

given by

$$x \simeq y \iff x = y \quad \text{or} \quad x = -y.$$

Let

$$P = S_2^- / \simeq$$

and give  $P$  its quotient topology. Now two distinct points of  $S_2^-$  are antipodal if and only if they are diametrically opposite points on the equator

$$E = \{ \langle x_1, x_2, x_3 \rangle \in S_2 : x_3 = 0 \}$$

(see Figure 3.29).

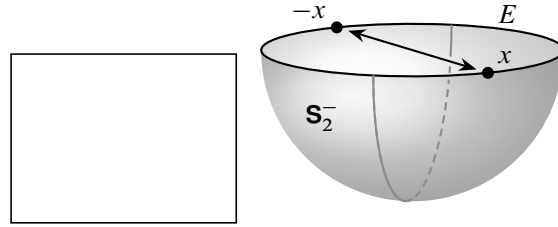


Figure 3.29: Antipodal points on the lower hemisphere are diametrically opposite points on the equatorial circle.

fig:diametrically-opposite-pts-on-eq

Hence a point of  $P$  is either a point  $x$  of  $S_2^-$  lying below the equator—more precisely, a singleton  $\{x\}$  for such an  $x$ —or else a pair  $\{x, -x\}$  of diametrically opposite points on the equator.

A homeomorphism

$$g^* : \mathbb{RP}_2 \cong P$$

is obtained as follows. The map  $k : S_2 \rightarrow S_2^-$  given by

$$k(x) = \begin{cases} x & \text{if } x \in S_2^-, \\ -x & \text{if } x \notin S_2^- \end{cases}$$

is *not* continuous; however, its composite

$$g = q \circ k : S_2 \rightarrow P$$

with the quotient map  $q : S_2^- \rightarrow P$  is continuous, and this composite is constant on each equivalence class under  $\sim$ . Hence  $g$  induces a continuous map

$$g^* : \mathbb{RP}_2 = S_2 / \sim \rightarrow S_2^- / \simeq = P.$$

Similarly, the composite  $f = p \circ j : S_2^- \rightarrow \mathbb{RP}_2$  of the inclusion map  $j : S_2^- \rightarrow S_2$  with the quotient map  $p : S_2 \rightarrow \mathbb{RP}_2$  induces a continuous map  $f^* : P \rightarrow \mathbb{RP}_2$ , and, as is readily checked, both  $g^* \circ f^*$  and  $f^* \circ g^*$  are identity maps. Hence  $g^*$  is indeed a homeomorphism (and  $f^*$  is its inverse).

Let us now refer to  $P$  itself as the projective plane, to the points of  $P$  as projective points, and to the images of projective lines in  $\mathbb{RP}_2$  under  $g^*$  as projective lines. Then

each projective line in  $P$  has the form  $q(C \cap S_2^-)$  for some great circle  $C$  in  $S_2$ . This follows from the commutativity of the diagram

$$\begin{array}{ccc} S_2 & \xrightarrow{k} & S_2^- \\ p \downarrow & & \downarrow q \\ \mathbb{RP}_2 & \xrightarrow{g^*} & P \end{array}$$

and the fact that  $k$  maps each great circle  $C$  of the sphere  $S_2$  onto its intersection  $C \cap S_2^-$  with the lower hemisphere.

Let us also work not with  $\mathbb{R}^2$  itself, but instead with the plane

$$H = \{ \langle x_1, x_2, x_3 \rangle \in \mathbb{R}^3 : x_3 = -1 \}$$

in  $\mathbb{R}^3$  to which it is homeomorphic under the map

$$\langle x_1, x_2 \rangle \mapsto \langle x_1, x_2, -1 \rangle.$$

This map carries lines in  $\mathbb{R}^2$  onto lines in  $H$ , so that  $\mathbb{R}^2$  and  $H$  are not just topologically the same but also geometrically, at least with respect to incidence of points and lines.

To each  $x \in H$  we associate the point  $r(x) \in S_2^- \setminus E$  at which the line joining  $x$  to the origin  $\mathbf{0} = \langle 0, 0, 0 \rangle$  meets  $S_2^-$  (see Figure 3.30). It is easy to see that  $q$  maps the



Figure 3.30: Constructing the embedding  $r$  of the plane  $H$  into the lower hemisphere  $S_2^-$ .

fig:map-r-from-hyperplane-to-lower

subspace  $S_2^- \setminus E = r(H)$  of  $S_2^-$  homeomorphically onto the subspace  $P \setminus q(E)$  of  $P$ . Hence the composite

$$h = q \circ r: H \rightarrow P$$

is a topological embedding of the plane  $H$  into the projective plane  $P$ .

What is really of interest here is the geometric, not the topological, behavior of the map  $h$ . A simple geometric argument shows that the image  $r(L)$  of an arbitrary line  $L \subset H$  has the form

$$r(L) = (C \cap S_2^-) \setminus \{y, -y\}$$

where  $C$  is a great circle and  $y, -y$  are diametrically opposite points of  $C$  on  $E$  (see Figure 3.30 again); hence  $h(L)$  is a projective line in  $P$  with a single projective point  $i(L) = \{y, -y\}$  removed. Thus: the projective plane can be obtained from the ordinary (affine) plane by adjoining to each ordinary line  $L$  a single “ideal” projective point  $i(L)$ , and then  $L \cup \{i(L)\}$  is a projective line. Besides these projective lines, there is a single “ideal” projective line  $q(E)$ , which is the set of all ideal projective points.

For additional information about projective geometry, consult Birkhoff and Mac Lane [5, sec. 9.14] or Rosenbaum [55].

ex:line-with-2-origins-as-quotient-space-as-the-subspace  
collapsing to a point

(8) Let  $X$  be the subspace of the  $\langle x, y \rangle$ -plane consisting of the horizontal lines with equations  $y = 0$  and  $y = 1$ , so that  $X \cong \mathbb{R} \times \{0, 1\}$ , where  $\{0, 1\}$  has its discrete topology. Then the line with two origins  $Y$  [Examples 2.20 (3)] is homeomorphic to a quotient space of  $X$ . In fact, the map  $f: X \rightarrow Y$  given by

$$f(x, 0) = f(x, 1) = x \text{ if } x \neq 0, \quad f(0, 0) = 0, \quad f(0, 1) = 0'$$

is a continuous open surjection. (See Exercise 190.)

ex:collapse-subspace-to-pt

(9) In Example (1) we identified distinct points of the unit interval  $[0, 1]$  only when they belonged to the two-element subset  $\{0, 1\}$  consisting of its endpoints. We generalize that example.

Let  $A$  be a nonempty subset of a topological space  $X$ . Define  $\sim$  to be the equivalence relation on  $X$  that identifies all points of  $A$  with each other, that is:

$$x \sim y \iff (x = y) \text{ or } (x \in A \text{ and } y \in A).$$

Then the points of the quotient set  $X/\sim$  are the singletons  $\{x\}$  for  $x \notin A$  together with the distinguished point  $A$ . (Note that  $A = \llbracket a \rrbracket_{\sim} \in X/\sim$  for each  $a \in A$ .) The quotient space  $X/\sim$ , denoted by  $X//A$ , is said to be obtained by **collapsing  $A$  to a point**. (Some mathematicians use the notation  $X/A$  instead of  $X//A$ , but unfortunately the former notation clashes with the standard notation for the construct of a quotient group when  $X$  is a group and  $A$  is a subgroup that is “normal” in  $X$  in the algebraic sense; compare Exercise 209.)

Let

$$q: X \rightarrow X//A$$

be the quotient map. Then  $q$  maps the subspace  $X \setminus A$  onto the complement  $(X//A) \setminus \{A\}$  of the point  $A$  in the quotient space. Moreover, the restriction of  $q$  to  $X \setminus A$  is injective.

Suppose now that  $A$  is a *closed* subset of  $X$ . Then:

- $X \setminus A$  is open in  $X$ ;
- $(X//A) \setminus \{A\}$  is open in  $X//A$ ; and
- $q$  induces a *homeomorphism*

$$X \setminus A \cong (X//A) \setminus \{A\}.$$

In fact, if  $U$  is open in  $X \setminus A$ , then its saturation  $q^{-1}(q(U)) = U$  is open in  $X$  whence  $q(U)$  is open in  $X//A$ .

Thus in the case of a closed  $A$ , the quotient space  $X//A$  contains a point whose complement is topologically the same as the portion  $X \setminus A$  of  $X$  in which no identifications are made.

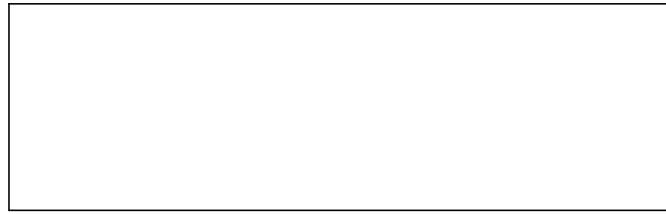
ex:disk-collapse-circle

(10) As a specific example of collapsing a subset to a point, take  $X = D_2$ , the 2-disk in the plane, and  $A = S_1$ , its bounding circle. To see what the quotient  $D_2//S_1$  looks like, imagine a circular piece of cloth with a drawstring around its edge; when the string is drawn tight, a spherical bag results (see Figure 3.31).

Thus we should expect

$$D_2//S_1 \cong S_2.$$

To find a homeomorphism  $h: D_2//S_1 \cong S_2$ , it suffices to construct a continuous closed surjection  $f: D_2 \rightarrow S_2$  that sends each point of  $S_1$  to the north pole  $\mathbf{p}$  of  $S_2$  and maps



collapsing to a point

Figure 3.31: Collapsing the bounding circle  $S_1$  of the disk  $D_2$  to a point to obtain the sphere  $S_2$ .

fig:collapse-bdy-of-disk-obtain-sphere

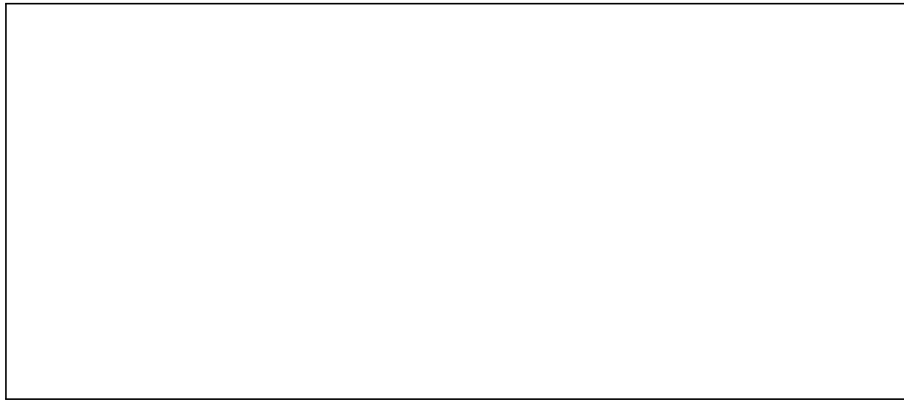


Figure 3.32: Constructing a continuous closed surjection  $f$  from the disk  $D_2$  to the sphere  $S_2$  that maps the inside of the disk injectively onto the complement of the north pole.

fig:map-D2-to-S2

the inside  $D_2 \setminus S_1$  of the disk injectively onto the complement  $S_2 \setminus \{\mathbf{p}\}$  of the north pole, for then we may take  $h$  to be the map  $f^*$  obtained by passing to the quotient. One such  $f$ , indicated in [Figure 3.32](#), is the composite of two maps:

- the homeomorphism  $f_1$ , given by  $\langle x_1, x_2 \rangle \mapsto \langle \pi x_1, \pi x_2, -1 \rangle$ , from  $D_2$  onto the disk

$$D = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq \pi^2, x_3 = -1\}$$

tangent to  $S_2$  at the south pole; and

- the map  $f_2$  that wraps each radial segment of  $D$  upward onto a meridian of  $S_2$  so that an arbitrary  $x \in D$  with cylindrical coordinates  $\langle r, \theta, -1 \rangle$  is sent to the point on  $S_2$  with spherical coordinates  $\langle \rho, \varphi, \theta \rangle = \langle 1, \pi - r, \theta \rangle$ .

That the composite  $f = f_2 \circ f_1$  is actually a closed map is most easily established by the shortcut mentioned at the end of [Example \(1\)](#).

(Another way of constructing the desired  $f$ , more amenable to generalization to higher dimensions, is indicated in [Exercise 169](#).)

ex:collapse-two-subsets-to-pts

- (11) If  $A$  and  $B$  are disjoint nonempty subsets of a space  $X$ , then we may form the quotient space obtained by collapsing  $A$  and  $B$  to distinct points: see [Exercise 172](#).

As a specific example, take  $A = [0, 1]$ ,  $B = [2, 3]$ , and  $X = A \cup B$ ; then the space obtained by collapsing  $A$  and  $B$  to distinct points is homeomorphic to a two-point discrete space.

ex:attach-spaces-by-map (12) Let  $X$  and  $Y$  be topological spaces and let  $f: A \rightarrow Y$  be a continuous map on a nonempty *closed* subset  $A$  of  $X$ . Imagine joining  $X$  and  $Y$  together by gluing each point  $a \in A$  to the point  $f(a)$  to which  $f$  sends it; continuity of  $f$  means that nearby points in  $A$  are glued to nearby points in  $Y$ . The result should be a topological space  $Z$  which contains (homeomorphic copies of)  $X \setminus A$  and  $Y$  and in which each  $y \in f(A)$  represents an identification of all points  $a \in f^{-1}(y)$  with  $y$ . We shall show how to construct such a space  $Z$ .

Temporary simplifying assumption:  $X$  is disjoint from  $Y$ . Endow the set  $X \cup Y$  with the topology constructed in Examples 3.47 (4): a subset of  $X \cup Y$  is open in this topology exactly when it is the union of an open (possibly empty) subset of  $X$  and an open (possibly empty) subset of  $Y$ . Then the topological space  $X \cup Y$  contains  $X$  and  $Y$  as open, closed subspaces. Define  $\sim$  to be the equivalence relation on  $X \cup Y$  given by

$$\begin{aligned} u \sim v \iff & (u = v) \\ & \text{or } (u \in A \text{ and } v = f(u)) \\ & \text{or } (v \in A \text{ and } u = f(v)) \\ & \text{or } (u \in A \text{ and } v \in A \text{ and } f(u) = f(v)). \end{aligned}$$

The desired  $Z$  is the quotient space  $(X \cup Y)/\sim$ . Define

$$X \cup_f Y = (X \cup Y)/\sim.$$

We call  $X \cup_f Y$  the **adjunction space obtained by attaching  $X$  to  $Y$  by  $f$**  and refer to  $f$  as the **attaching map**.

We show that  $X \cup_f Y$  has the desired properties. Let

$$q: X \cup Y \rightarrow X \cup_f Y$$

be the quotient map. Clearly

$$\{eq:attach-properties\} \quad (*) \quad \left. \begin{aligned} X \cup_f Y &= q(X) \cup q(Y) = q(X \setminus A) \cup q(A) \cup q(Y), \\ q(X) \cap q(Y) &= q(A) = q(f(A)), \\ q(X \setminus A) \cap q(Y) &= \emptyset. \end{aligned} \right\}$$

Moreover:

- (i) the set  $q(X \setminus A)$  is open in  $X \cup_f Y$ , and  $q$  maps  $X \setminus A$  homeomorphically onto  $q(X \setminus A)$ ; and
- (ii) the set  $q(Y)$  is closed in  $X \cup_f Y$ , and  $q$  maps  $Y$  homeomorphically onto  $q(Y)$ .

In fact, the set  $q(Y)$  is closed because  $q^{-1}(q(Y)) = A \cup Y$  is closed in the space  $X \cup Y$ . (Recall the stipulation at the start that  $A$  is closed in  $X$ !) The restriction  $q|_Y$  is continuous and injective; if  $E$  is an arbitrary closed subset of  $Y$ , then its image  $q(E)$  is closed in  $q(Y)$  because  $q^{-1}(q(E)) = A \cup E$  is closed in  $X \cup Y$ . This proves (ii). The proof of (i) is similar but uses open sets.

A specific example. Take  $X = [0, 1]$ ,  $A = \{0, 1\}$ , and  $Y = D_2$ ; and take  $f: A \rightarrow Y$  to be the map given by  $f(0) = f(1) = \mathbf{0} \in D_2$ . Then  $X \cup_f Y$  is homeomorphic to the subspace

$$\begin{aligned} W = & \{x \in \mathbb{R}^3 : x_3 = 0, x_1^2 + x_2^2 \leq 1\} \\ & \cup \{x \in \mathbb{R}^3 : x_1 = 0, x_2^2 + (x_3 - 1)^2 = 1\} \end{aligned}$$

of  $\mathbb{R}^3$  shown in Figure 3.33. To prove this, form the map  $g: X \cup Y \rightarrow W$  given by



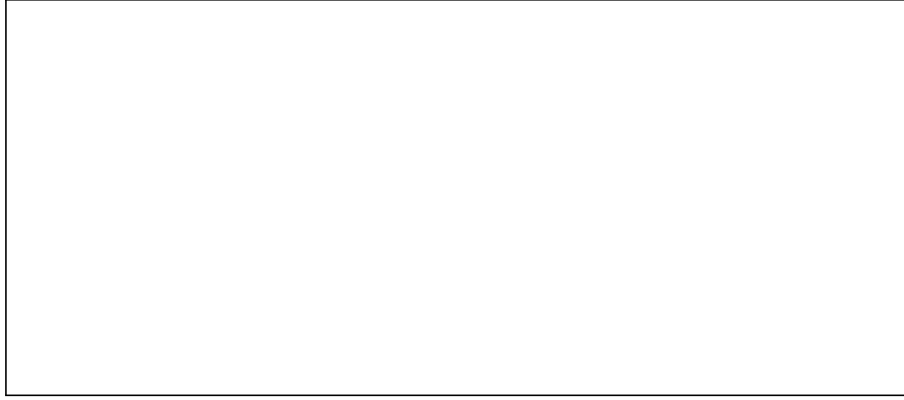


Figure 3.33: Attaching the unit interval  $[0, 1]$  to the disk  $D_2$  by the map  $f$  that sends the endpoints of  $[0, 1]$  to the origin.

fig:attach-l-to-D2-endpots-to-origin

$$\begin{aligned} g(x) &= \langle 0, \sin 2\pi x, 1 - \cos 2\pi x \rangle & (x \in X), \\ g(y) &= \langle y_1, y_2, 0 \rangle & (y \in Y) \end{aligned}$$

and then pass to the quotient.

The general case. To construct the desired  $Z = X \cup_f Y$  in the case that  $X$  is not necessarily disjoint from  $Y$ , the preceding procedure must be modified. Form the Cartesian sum  $X + Y$  of  $X$  and  $Y$  and the embeddings

$$h: X \rightarrow X + Y, \quad k: Y \rightarrow X + Y$$

as in [Examples 3.47 \(4\)](#). Define  $\sim$  to be the equivalence relation on  $X + Y$  given by

$$\begin{aligned} u \sim v \iff & (u = v) \\ & \text{or } [u = h(a) \text{ and } v = k(f(a)) \text{ for some } a \in A] \\ & \text{or } [v = h(a), u = k(f(a)) \text{ for some } a \in A] \\ & \text{or } [u = h(a), v = h(b) \text{ for } a, b \in A \text{ with } f(a) = f(b)]. \end{aligned}$$

Finally, define

$$X \cup_f Y = (X + Y) / \sim$$

and let  $q: X + Y \rightarrow (X + Y) / \sim$  be the quotient map. Then

$$\begin{aligned} X \cup_f Y &= (q \circ h)(X) \cup (q \circ k)(Y) \\ &= (q \circ h)(X \setminus A) \cup (q \circ h)(A) \cup (q \circ k)(Y), \\ (q \circ h)(X) \cap (q \circ k)(Y) &= (q \circ h)(A) = (q \circ k)(f(A)), \\ (q \circ h)(X \setminus A) \cap (q \circ k)(Y) &= \emptyset. \end{aligned}$$

Moreover,  $q \circ h$  embeds  $X \setminus A$  as an open subspace of  $X \cup_f Y$  and  $q \circ k$  embeds  $Y$  as a closed subspace of  $X \cup_f Y$ .

*Abuse of language:* When  $X$  and  $Y$  are disjoint, the maps  $h$  and  $k$  are superfluous. However, even when they are not disjoint, sometimes we act *as if* they were disjoint and then we somewhat loosely write the same relations (\*) as in the disjoint case, suppressing mention of  $h$  and  $k$ , and still say that  $q$  embeds  $X \setminus A$  as an open subspace of  $X \cup_f Y$  and embeds  $Y$  as a closed subspace  $q(Y)$  of  $X \cup_f Y$ . Even more loosely, sometimes we regard  $X$  and  $Y$  as *as if* they were subspaces of  $X + Y$  and then regard  $X \setminus A$  and  $Y$  as open and closed subspaces, respectively, of  $X \cup_f Y$ .

ex:attach-cell (13) Let  $X$  be a topological space. Let  $h: D_n \cong D$  be a homeomorphism from the  $n$ -disk to a topological space  $D$ , let  $f: S_{n-1} \rightarrow X$  be a continuous map. Then the adjunction space  $D \cup_{f \circ h} X$  is said to be **obtained by attaching an  $n$ -cell to  $X$** .

adjunction space  
countability properties!quotient space@of quotient space  
quotient space!countability properties@and countability properties  
real line with integers collapsed to a point

### Countability and separation properties of quotient spaces

subsec:count-sep-quotients

For any particular construction of new topological spaces from old ones, we want to know which topological properties are preserved. We already know that every subspace of a first-countable or second-countable space has the same property (Proposition 2.102). However, a subspace of a separable space need not be separable: see Example 2.88. Likewise, we already know that every subspace of a  $T_0$ ,  $T_1$ , Hausdorff, or regular space has the same property (Proposition 2.102). However, a subspace of a normal space need not be normal, as we shall establish later: see Example 6.23.

Unfortunately, the situation with quotient spaces is not so nice as that with subspaces. Consider first the countability properties. We already know that a continuous image of a separable space is separable (Proposition 3.19), so in particular the following is true.

prop:quot-separable **3.82 Proposition.** Any quotient space of a separable space is itself separable.

The remaining countability properties fail to be preserved under formation of quotients.

ex:quotient-and-countability **3.83 Example.** Although the real line  $\mathbb{R}$  is both second-countable and, *a fortiori*, first-countable, its quotient space  $\mathbb{R}/\mathbb{Z}$ , obtained by collapsing the set of integers to a point [Examples 3.81 (9)], is neither.

pg:visualize-R-collapse-Z To see this, it may be helpful to visualize  $\mathbb{R}/\mathbb{Z}$  as being a doubly infinite sequence of loops all issuing from, and returning to, a common point (that point being the element  $\mathbb{Z}$  of the quotient). Each loop with the common point removed is homeomorphic to an open interval in  $\mathbb{R}$ . At each point on one of the loops there is a local base consisting of open arcs short enough so as not to include the common point; at the common point, there is a local base consisting of infinitely many open arcs each having that point in their interior.

Just suppose that  $\mathbb{R}/\mathbb{Z}$  has a countable local base  $\{V_n : n \in \mathbb{N}\}$  at the point  $q(0) = \mathbb{Z}$  in  $\mathbb{R}/\mathbb{Z}$ , where  $q: \mathbb{R} \rightarrow \mathbb{R}/\mathbb{Z}$  is the quotient map; without loss of generality we may assume that each  $V_n$  is open. To reach a contradiction we shall use a “diagonal argument” on  $\mathbb{N} \times \mathbb{Z}$ .

Temporarily fix  $n \in \mathbb{N}$ . The inverse image  $q^{-1}(V_n)$  is an open neighborhood of the subset  $\mathbb{Z}$  in  $\mathbb{R}$ . Then for each integer  $k$  there is some  $\varepsilon_{n,k}$  with  $0 < \varepsilon_{n,k} < 1/2$  such that

$$]k - \varepsilon_{n,k}, k + \varepsilon_{n,k}[ \subset q^{-1}(V_n).$$

Observe that each of the intervals  $]k - \varepsilon_{n,k}, k + \varepsilon_{n,k}[$  contains exactly one integer and that the collection of all of these intervals is pairwise disjoint. Now let  $\delta_{n,k} = \varepsilon_{n,k}/2$  for each  $k$ .

Define

$$U = \bigcup_{n \in \mathbb{N}} ]n - \delta_{n,n}, n + \delta_{n,n}[.$$

Then  $U$  is a saturated open neighborhood of the subset  $\mathbb{Z}$  in  $\mathbb{R}$ , so that  $q(U)$  is an open neighborhood of the point  $\mathbb{Z}$  in  $\mathbb{R}/\mathbb{Z}$ . Since  $\{V_n : n \in \mathbb{N}\}$  is a local base at the point  $\mathbb{Z}$  in  $\mathbb{R}/\mathbb{Z}$ , there must be some  $n \in \mathbb{N}$  for which  $V_n \subset q(U)$ . For such  $n$ , then

$$]n - \varepsilon_{n,n}, n + \varepsilon_{n,n}[ \subset q^{-1}(V_n) \subset q^{-1}(q(U)) = U.$$

But this is impossible because the only open interval centered at  $n$  that is contained in  $U$  is  $]n - \delta_{n,n}, n + \delta_{n,n}[$ , which is shorter than  $]n - \varepsilon_{n,n}, n + \varepsilon_{n,n}[$ .  $\diamond$

The situation for separation properties of quotient spaces is not nice, either.

exs:quotient-not-Hausdorff

**3.84 Examples.** (1) The quotient of a  $T_0$ -,  $T_1$ -,  $T_2$ -, or regular space need not have the same property. For example, give  $\mathbb{R}$  its usual topology and give  $Y = \{0, 1\}$  its *indiscrete* topology. Then the characteristic function of  $\mathbb{Q}$  in  $\mathbb{R}$  is an open (but not closed) surjection from  $\mathbb{R}$  to  $Y$ , but  $Y$  is not regular and not  $T_0$ , hence not  $T_1$  or  $T_2$ , either.

(2) The line with two origins [Examples 3.81 (8)] is a  $T_1$ -space, but not a  $T_2$ -space [Examples 2.99 (3)], that is a continuous open image of the  $T_2$ -space  $\mathbb{R} \times \{0, 1\}$ , where  $\{0, 1\}$  has its discrete topology.  $\diamond$

Here is one special yet important case where a quotient *is* necessarily a  $T_2$ -space; the proof is left to the reader [see Exercise 181 (b)].

prop:collapse-closed-set-in-reg-get-T2

**3.85 Proposition.** *The quotient space  $X//A$  obtained from a regular  $T_0$ -space  $X$  by collapsing a nonempty closed subset  $A$  to a point is a  $T_2$ -space.*

ex:R-collapse-Z-T2

**3.86 Example.** Apply the preceding Proposition 3.85 to  $X = \mathbb{R}$  and  $A = \mathbb{Z}$  to see that the quotient space  $\mathbb{R}//\mathbb{Z}$  is a Hausdorff space.

$\diamond$

The following positive result on preservation of  $T_2$  under formation of quotients is akin to Proposition 3.85; its proof is left to the reader (Exercise 182).

prop:open-closed-image-reg-then-T2

**3.87 Proposition.** *A continuous, open, and closed image of a regular  $T_0$ -space is itself regular  $T_0$ , hence  $T_2$ .*

The following example is due to Chaber [13, Example 3.2], as modified by Scott [59]. In light of the preceding proposition, it is a “best possible” counterexample concerning preservation of  $T_2$  under quotient maps.

er-open-closed-map-T2-onto-nonT2

**3.88 Example (a  $T_1$  non-Hausdorff open and closed quotient of a  $T_2$ -space).** Let  $X = X_{-1} \cup X_0 \cup X_1$  where  $X_i = \mathbb{R} \times \{i\}$  for  $i = -1, 0, 1$ . Provide  $X$  with the topology such that, for each  $x \in \mathbb{R}$ :

- the singleton  $\{\langle x, 0 \rangle\}$  is open (so that the topology induced on  $X_0$  is discrete);
- a local base at  $\langle x, 1 \rangle$  consists of all sets of the form

$$B(\langle x, 1 \rangle, a, A) = \{\langle x, 1 \rangle\} \cup ([x, a[ \setminus A) \times \{0\})$$

for  $x < a$  and  $A$  a countable subset of  $\mathbb{R}$ ; and

- a local base at  $\langle x, -1 \rangle$  consists of all sets of the form

$$B(\langle x, -1 \rangle, b, A) = \{\langle x, -1 \rangle\} \cup ([b, x[ \setminus b) \times \{0\})$$

for  $b < x$  and  $B$  a countable subset of  $\mathbb{R}$ .

countability properties!quotient space!  
quotient space!countability property!  
separation properties!quotient space!  
quotient space!separation properties!  
quotient space! $T_2$ -space@as \Ttwo-s  
Hausdorff space!quotient space@an  
line with two origins!quotient space@  
line with two origins!separation prop  
real line with integers collapsed to a  
collapsing to a point!separation prop  
Hausdorff space!collapsing to a poin  
Chaber, J.  
Scott, Brian M.

Let  $\sim$  be the equivalence relation on  $X$  for which:  $\langle x, -1 \rangle \sim \langle y, -1 \rangle$  for all  $x, y \in \mathbb{R}$ ;  $\langle x, 1 \rangle \sim \langle y, 1 \rangle$  for all  $x, y \in \mathbb{R}$ ; and  $\langle x, 0 \rangle \sim \langle y, 0 \rangle$  if and only if  $x - y \in \mathbb{Q}$ . In other words, the equivalence classes under  $\sim$  are  $X_{-1}$  and  $X_1$  together with all subsets of  $X_0$  of the form

$$\langle x, 0 \rangle + (\mathbb{Q} \times \{0\}) = \{\langle x + r, 0 \rangle : r \in \mathbb{Q}\}.$$

Let  $Y = X/\sim$  and let  $q: X \rightarrow Y$  be the quotient map.

Then  $X$  is a  $T_2$ -space, the quotient map  $q: X \rightarrow Y$  is both open and closed, but the quotient  $Y$  is a  $T_1$ -space that is *not* a  $T_2$ -space.  $\diamond$

Again let  $A$  be a nonempty closed subset of a  $T_2$ -space  $X$ . Then the graph of the equivalence relation  $\sim$  induced by the partition  $X//A$  may be expressed in the form

$$\text{graph}(\sim) = (A \times A) \cup \Delta_X.$$

Now  $A \times A$  is closed in  $X \times X$  since  $A$  is closed in  $X$ ; and from the [closed diagonal criterion](#) ([Proposition 2.74](#)), the diagonal  $\Delta_X$  is closed in  $X \times X$ . Hence the graph of  $\sim$  is a closed subset of  $X \times X$ . This situation generalizes.

**3.89 Proposition.** A necessary condition for the quotient  $X/\sim$  of a topological space  $X$  to be a  $T_2$ -space is that the graph of the equivalence relation  $\sim$  be closed in  $X \times X$ .

**Proof.** Assume that  $X/\sim$  is a  $T_2$ -space. We shall deduce that the complement of  $\text{graph}(\sim)$  in  $X \times X$  is open. Let  $\langle x, y \rangle \in X \times X$  with  $x \not\sim y$ . Then  $\llbracket x \rrbracket$  and  $\llbracket y \rrbracket$  are distinct points of  $X/\sim$ , so they have disjoint neighborhoods  $V'$  and  $W'$  in  $X/\sim$ . Let  $V$  and  $W$  be the inverse images of  $V'$  and  $W'$ , respectively, under the quotient map  $q: X \rightarrow X/\sim$ . Then  $V \times W$  is an open neighborhood of  $\langle x, y \rangle$  in  $X \times X$  that is disjoint from  $\text{graph}(\sim)$ .  $\square$

In general, that condition is *not* sufficient for the quotient to be a  $T_2$ -space: for a counterexample, see [Exercise 197](#). In one important situation, though, the condition is sufficient, namely when  $X$  is also “compact” in the sense of [Definition 4.5](#); this is the case, in particular, when  $X$  is a subspace of a Euclidean space that is closed and bounded for the Euclidean metric: see [Theorem 4.39](#).

For a summary of the preceding results on preserving topological properties for forming quotient spaces, and the corresponding results for other topological constructions, see [Table 3.2](#). (References concerning complete regularity and normality are to propositions and examples not appearing until later, in [Section 6.2](#).)

### EXERCISES FOR SECTION 3.4

**152.** Let  $q: X \rightarrow X/\sim$  be the quotient map induced by an equivalence relation  $\sim$  on a space  $X$ . Suppose  $\mathcal{T}$  is a topology on  $X/\sim$  such that:

- $q$  is continuous with respect to  $\mathcal{T}$ ; and
- an arbitrary map  $g: X/\sim \rightarrow Y$  is continuous with respect to  $\mathcal{T}$  precisely when its composite  $g \circ q: X \rightarrow Y$  is continuous.

Must  $\mathcal{T}$  be the quotient topology?

**153. (a)** Let  $f: X \rightarrow Y$  be a continuous map. Let  $\sim$  and  $\simeq$  be equivalence relations on  $X$  and  $Y$ , respectively, such that  $x \sim u$  always implies  $f(x) \simeq f(u)$ . Use  $f$  to construct a continuous map  $f': X/\sim \rightarrow Y/\simeq$ .

**(b)** If  $f$  is a quotient map, show that the constructed  $f'$  must be a quotient map.

Property	Construction			
	subspace	finite product	finite sum	quotient
$T_0$	✓ (Proposition 2.102)	✓ (Exercise 2.150)	✓ (Proposition 3.49)	✗ (Examples 3.84)
$T_1$	✓ (Proposition 2.102)	✓ (Exercise 2.150)	✓ (Proposition 3.49)	✗ (Examples 3.84)
$T_2$	✓ (Proposition 2.102)	✓ (Exercise 2.150)	✓ (Proposition 3.49)	✗ (Examples 3.84)
regular	✓ (Proposition 2.102)	✓ (Exercise 2.150)	✓ (Proposition 3.49)	✗ (Examples 3.84)
completely regular	✓ (Proposition 6.18)	✓ (Proposition 6.18)	✓ (Proposition 6.18)	✗ (Example 6.19)
normal	✗ (Example 6.23)	✗ (Example 6.32)	✓ (Proposition 6.24)	✗ (Exercise 6.37)
first-countable	✓ (Proposition 2.78)	✓ (Proposition 2.89)	✓ (Proposition 3.48)	✗ (Example 3.83)
second-countable	✓ (Proposition 2.78)	✓ (Proposition 2.89)	✓ (Proposition 3.48)	✗ (Example 3.83)
separable	✗ (Example 2.88)	✓ (Proposition 2.89)	✓ (Proposition 3.48)	✓ (Proposition 3.82)
metrizable	✓ [Examples 2.10 (1)]	✓ (Exercise 2.108)	✓ (Exercise 92)	✗ (Example 3.83)

Table 3.2: Preservation of topological properties: separation and countability properties—version 2

equivalence kernel!quotient map@  
quotient map!equivalence kernel@

tab:preserve-ver-2

154. What can be said about a quotient space  $X/\sim$  if  $\sim$  is the equality relation  $=$  on  $X$ ?
- quot-Sierpinski-not-open-not-closed155. Let  $X$  be the closed unit interval  $[0, 1]$ , let  $Y$  be the Sierpinski space [Examples 2.3 (6)], and let  $f: X \rightarrow Y$  be the characteristic function of  $[1/2, 1]$ . Verify that  $f$  is quotient map that is neither open nor closed.
- prob:quot-map-equiv-via-closed156. Verify the equivalence with Definition 3.78 of the criterion on page 392 for a surjection  $f: X \rightarrow Y$  to be a quotient map: for each  $E \subset Y$ , the set  $E$  is closed in  $Y$  if and only if its inverse image  $f^{-1}(E)$  is closed in  $X$ .
- quot-map-is-concrete-quotient-map157. Let  $f: X \rightarrow Y$  be an arbitrary quotient map. Recall from Exercise 0.150 that the equivalence kernel of  $f$  is the equivalence relation  $\sim$  on  $X$  given by

$$x \sim u \iff f(x) = f(u)$$

Obviously  $f$  is constant on each equivalence class under  $\sim$ , and so by passing to the quotient we obtain from  $f$  the map  $f^*: X/\sim \rightarrow Y$  such that  $f^* \circ q = f$ , where  $q: X \rightarrow X/\sim$  is the quotient map induced by  $\sim$ . Prove that  $f^*: X/\sim \cong Y$ .  
(Thus any “abstract” quotient map, in the sense of Definition 3.78, with domain a space  $X$  is essentially the same as the “concrete” quotient map induced by an equivalence relation on  $X$ .)

158. Find a quotient space of the plane  $\mathbb{R}^2$  that is homeomorphic to the torus.

torus!cylinder@and 159. Show that the torus is a 2-dimensional manifold.

Klein bottle!cylinder@and cylinder

prob-part:torus-from-cylinder 160. (a) Let  $\sim$  and  $\sim^*$  be two equivalence relations on the same space  $X$  such that  $x \sim^* u$  always implies  $x \sim u$ . Prove that  $\sim$  induces an equivalence relation  $\simeq$  on the quotient  $Y = X/\sim^*$  such that  $Y/\simeq$  is homeomorphic to  $X/\sim$ .

solid torus

manifold-with-boundary

prob-part:torus-from-cylinder

Möbius strip@Möbius strip

edge of Möbius strip@edge of M

path

Möbius strip@Möbius strip

prob-part:Klein-bottle-from-cylinder

projective n-space@projective n-space

projective n-space@projective n-space

prob:solid-torus

projectiveline

(b) Figure 3.22 suggests that the torus can be obtained from a cylinder by identifying corresponding points on the circular ends of the cylinder. Show that this is indeed the case by applying (a) with  $X = I \times I$  and  $\sim$  the equivalence relation of Examples 3.81 (4).

(c) Show similarly that the Klein bottle is a quotient space of a cylinder.

161. The product  $S_1 \times D_2$  is the **solid torus**.

(a) Show that the solid torus is a 3-manifold-with-boundary and determine its boundary  $\partial(S_1 \times D_2)$ . (Refer to Exercises 91 and 97 for the definitions.)

(b) Embed the solid torus as a solid in  $\mathbb{R}^3$  that is bounded by the doughnut-shaped surface  $T$  described in Examples 3.81 (4).

(c) Represent the 3-sphere  $S_3$  as the union of two subspaces  $X_1$  and  $X_2$  each of which is homeomorphic to the solid torus and whose intersection  $X_1 \cap X_2$  is homeomorphic to the (ordinary, non-solid) torus.

prob:Möbius-strip-as-mfld-with-bdy 162. Let  $M$  be the Möbius strip. The **edge** of  $M$  is, in the notation of Examples 3.81 (5), the image  $E$  of the set  $I \times \{0, 1\}$  under the quotient map  $q: I \times I \rightarrow M = (I \times I)/\sim$ .

(a) Show that  $E \cong S_1$ .

(b) Prove that  $M$  is a 2-dimensional manifold-with-boundary having  $E = \partial M$  as its boundary. (Refer to Exercises 91 and 97.)

163. (Continuation of Exercise 162.) Given a space  $X$  and points  $x, y \in X$ , a *path* in  $X$  from  $x$  to  $y$  is a continuous map  $\sigma: I \rightarrow X$  such that  $\sigma(0) = x$  and  $\sigma(1) = y$ . We may interpret such a path as a way to move continuously in  $X$  from point  $x$  to point  $y$ . (See Section 5.3 for more about paths.)

(a) Show that we can move continuously on the Möbius strip from any point not on its edge to any other such point without ever crossing the edge. More precisely: If  $x$  and  $y$  are arbitrary points of  $M$  that are not on its edge  $E$ , construct a path in  $M \setminus E$  from  $x$  to  $y$ . [Hint: Choose  $a \in q^{-1}(x)$  and  $b \in q^{-1}(y)$ , construct a suitable map  $\rho: I \rightarrow I \times I$ , and pass to the quotient.]

(b) Let  $K = q(I \times \{1/2\})$ , the image under  $q$  of the middle horizontal line segment  $I \times \{1/2\}$  of the unit square. If  $x$  and  $y$  are arbitrary points in  $M \setminus K$ , must there exist a path in  $M \setminus K$  from  $x$  to  $y$ ?

(Project: Represent the Möbius strip  $M$  as a twisted strip of paper with its ends glued together. Draw on it the curve representing  $K$ . What happens if you cut the paper along this curve?)

quotient-of-Rnplus1-without-origin 164. Noting that each pair of antipodal points of  $S_n$  determines a unique line in  $\mathbb{R}^{n+1}$  through the origin  $\mathbf{0}$ , show that projective  $n$ -space  $\mathbb{RP}_n$  is homeomorphic to the quotient  $(\mathbb{R}^{n+1} \setminus \{\mathbf{0}\})/\simeq$ , where  $x \simeq y$  exactly when  $x$  and  $y$  line on the same line in  $\mathbb{R}^{n+1}$  through the origin.

prob-part:RPn-as-quotient 165. (a) Show that  $\mathbb{RP}_n \cong D_n/\sim$ , where  $\sim$  identifies each pair of antipodal points of the subset  $S_{n-1}$  of  $D_n$ .

prob-part:projective-line-homeo-circle (b) Deduce that the projective line  $\mathbb{RP}_1$  is homeomorphic to the unit circle  $S_1$ .

prob:RP2-less-pt-is-open-M **166.** In view of [Exercise 97 \(f\)](#), the subspace  $M \setminus E$  of the Möbius strip  $M$  is a 2-manifold, which is called the **open Möbius strip**.

(a) Show that the open Möbius strip is homeomorphic to the space obtained by deleting one point from the projective plane  $\mathbb{RP}_2$ .

(b) Does it make any difference which point of  $\mathbb{RP}_2$  is deleted?

ob:embed-projective-plane-into-R4 **167.** Use the map

$$f: S_2 \rightarrow \mathbb{R}^4 \\ \langle x_1, x_2, x_3 \rangle \mapsto \langle x_1^2 - x_2^2, x_1x_2, x_1x_3, x_2x_3 \rangle$$

to induce an embedding of the projective plane into  $\mathbb{R}^4$ .

**168.** Starting with the embedding

$$f: S_n \rightarrow S_{n+1} \\ x \mapsto \langle x_1, x_2, \dots, x_{n+1}, 0 \rangle,$$

use [Exercise 153](#) to construct an embedding  $f': \mathbb{RP}_n \rightarrow \mathbb{RP}_{n+1}$  such that the complement  $\mathbb{RP}_{n+1} \setminus f'(\mathbb{RP}_n)$  is homeomorphic to the  $n$ -ball  $B_n$ .

ob:another-pf-D2-modS1-homeo-S2 **169.** Let  $g: B_2 \cong \mathbb{R}^2$  be the homeomorphism of [Examples 3.25 \(9\)](#), and let  $k: \mathbb{R}^2 \cong S_2 \setminus \{\mathbf{p}\}$  be the inverse of the stereographic projection [[Examples 3.25 \(13\)](#), [page 329](#)], where  $\mathbf{p}$  is the north pole. Then  $k \circ g$  defines an embedding of  $B_2$  into  $S_2$ , which may be extended to a continuous map  $f: D_2 \rightarrow S_2$  by taking  $f(x) = \mathbf{p}$  for all  $x \in S_1$ . Using this data, demonstrate anew that the quotient  $D_2 // S_1$  obtained by collapsing the bounding circle of the disk to a point is homeomorphic to the sphere  $S_2$ .

(Note: See [Exercise 4.117](#) for a generalization.)

**170. (a)** Show that the space obtained from the cylinder  $S_1 \times I$  by collapsing its top  $T = S_1 \times \{1\}$  to a point is homeomorphic to the conical surface  $K$  in  $\mathbb{R}^3$  having the point  $(0, 0, 1)$  as its vertex and the set  $\{x \in \mathbb{R}^3 : x_3 = 0, x_1^2 + x_2^2 = 1\}$  as its base.

(b) If, instead, the bottom  $B = S_1 \times \{0\}$  is collapsed to a point, is the space obtained still homeomorphic to the conical surface  $K$ ?

**171.** What happens if, in the definition of  $X // A$  [[Examples 3.81 \(9\)](#)], we allow the subset  $A$  of  $X$  to be empty?

prob:collapse-two-subsets **172. (a)** Give a precise definition for the construction alluded to in [Examples 3.81 \(11\)](#), namely, the quotient space obtained by collapsing disjoint nonempty subsets  $A$  and  $B$  of a space  $X$  to distinct points.

(b) Verify the claim, made in [Examples 3.81 \(11\)](#), that the quotient space obtained by collapsing the subsets  $[0, 1]$  and  $[2, 3]$  of  $[0, 1] \cup [2, 3]$  to distinct points is homeomorphic to a two-point discrete space.

(c) Let  $X$  be the space obtained from the cylinder  $S_1 \times I$  by collapsing its top  $S_1 \times \{1\}$  and its bottom  $S_1 \times \{0\}$  to distinct points. To what subspace of  $\mathbb{R}^3$ , if any, is  $X$  homeomorphic?

**173.** Use the map  $\langle x, t \rangle \mapsto tx$  from  $S_{n-1} \times I$  into  $D_n$  to induce a homeomorphism  $(S_{n-1} \times I) // (S_{n-1} \times \{0\}) \cong D_n$ .

**174.** Show that the quotient map  $q: \mathbb{R} \rightarrow \mathbb{R} // \mathbb{Z}$  is closed but not open.



**175.** Show that the domain-codomain restriction  $\mathbb{R} \setminus \mathbb{Q} \rightarrow (\mathbb{R}/\mathbb{Q}) \setminus \{\mathbb{Q}\}$  of the quotient map  $\mathbb{R} \rightarrow \mathbb{R}/\mathbb{Q}$  is not an open map and hence not a homeomorphism.

**176.** (a) The subspace  $(0, \infty)$  of  $\mathbb{R}$  is first-countable, second-countable, and metrizable. Show, however, that its quotient space  $(0, \infty)/\mathbb{N}^*$  has none of these properties.

(b) Is  $\mathbb{R}/\mathbb{Q}$  first-countable? second-countable?

(c) Is  $\mathbb{Q}/\mathbb{Z}$  first-countable? second-countable?

**177.** (a) Explain why the space  $[0, 1] \times \mathbb{N}$  is second-countable.

(b) Visualize the quotient  $([0, 1] \times \mathbb{N})/(\{0\} \times \mathbb{N})$  in a manner analogous to how  $\mathbb{R}/\mathbb{Z}$  was visualized on page 408, in Example 3.83.

(c) Is the quotient  $([0, 1] \times \mathbb{N})/(\{0\} \times \mathbb{N})$  first-countable? second-countable?

**178.** Prove: The image of a  $T_1$ -space under a closed (but not necessarily continuous) map is itself a  $T_1$ -space.

**179.** Which separation properties do the following quotient spaces possess?

figure-eight

(a) The quotient space  $\mathbb{R}/\sim$  where the equivalence relation  $\sim$  on  $\mathbb{R}$  is given by  $x \sim y$  if and only if  $x = y$  or  $xy > 0$ .

(b) The quotient space  $D_2/\sim$  where the equivalence relation  $\sim$  on  $D_2$  is given by  $\langle x, u \rangle \sim \langle y, v \rangle$  if and only if  $\langle y, v \rangle = \langle tx, tu \rangle$  for some  $t \neq 0$ .

(c) The quotient space  $[0, 1]/\sim$  where, under the equivalence relation  $\sim$  on  $[0, 1]$ , the equivalence classes are the doubletons  $\{x, 1 - x\}$  for  $x \neq 0, 1/2, 1$  together with the singletons  $\{0\}$ ,  $\{1/2\}$ , and  $\{1\}$ .

**180.** Let  $A$  be a nonempty closed subset of a topological space  $X$ . Prove:

(a) Prove: If  $X$  is a  $T_0$ -space, then so is  $X/A$ .

(b) Prove: If  $X$  is a  $T_1$ -space, then so is  $X/A$ .

**181.** Let  $A$  be a nonempty closed subset of a topological space  $X$ .

(a) Exhibit such an  $X$  and  $A$  for which  $X$  is a  $T_2$ -space but  $X/A$  is not.

(b) Prove: If  $X$  is a regular  $T_0$ -space, then  $X/A$  is a  $T_2$ -space. (This is Proposition 3.85.)

**182.** Prove Proposition 3.87: A continuous, open, and closed image of a regular  $T_0$ -space is itself regular  $T_0$ .

(Note: For examples of such maps, see Exercise 46.)

**183.** Give a new proof of Proposition 3.89 by establishing and applying that  $\sim = (q \times q)^{-1}(\Delta_{X/\sim})$ , where  $q: X \rightarrow X/\sim$  is the quotient map.

**184.** Prove: If  $\sim$  is an equivalence relation on a topological space  $X$  whose graph is closed in  $X \times X$  and for which the quotient map  $X \rightarrow X/\sim$  is open, then the quotient space  $X/\sim$  is a  $T_2$ -space.

**185.** Show that the sphere with a whisker attached, shown in Figure 3.13, is homeomorphic to the space  $I \cup_f S_2$ , where  $f: \{0\} \rightarrow S_2$  sends the endpoint 0 of  $I$  to the north pole.

**186.** Given points  $a$  and  $b$  in spaces  $X$  and  $Y$ , respectively, the one-point union of  $X$  with  $Y$  is the space  $X \cup_f Y$ , where  $f: \{a\} \rightarrow Y$  takes the value  $b$ .

Show that for any choices of points  $a$  and  $b$  on the circle  $S_1$ , the one-point union of  $S_1$  with itself is homeomorphic to the figure-eight (Exercise 51).



- prob:collapse-to-pt-is-attachment **187.** Let  $A$  be a nonempty closed subset of a space  $X$ . Show that the space  $X//A$  obtained by collapsing  $A$  to a point can be obtained by attaching  $X$  to a one-point space.
- 188.** If  $f: S_{n-1} \rightarrow D_n$  is the inclusion map, to what familiar space is  $D_n \cup_f D_n$  homeomorphic?
- 189.** Let  $M$  be the Möbius strip and let  $\partial M$  be its edge (Exercise 162). If  $j: \partial \rightarrow M$  is the inclusion map, show that  $M \cup_j M$  is homeomorphic to the Klein bottle.
- prob:cont-open-image-T2-that-is-not-T2 **190.** Verify the assertion made in Examples 3.81 (8) that the map  $f$  from  $\mathbb{R} \times \{0, 1\}$  to the line with two origins is a continuous open surjection.

- prob:cone **191.** The **cone**  $K(X)$  over a topological space  $X$  is the quotient space

$$K(X) = (X \times I) / (X \times \{1\})$$

obtained by collapsing the “top”  $X \times \{1\}$  of the “cylinder”  $X \times I$  to a point.

*Note:* We use the notation  $K(X)$  rather than  $C(X)$  for the cone over  $X$  so as to avoid confusion with the notation  $C(X)$  for the space of continuous real-valued functions on  $X$ .

- (a) Prove that, in general,  $K(X)$  can be obtained by attaching  $X \times I$  to a one-point space by means of a suitable map.
- prob-part:bottom-of-cone (b) The “bottom” of the cone over a topological space  $X$  is its subspace  $q(X \times \{0\})$ , where  $q: X \times I \rightarrow K(X)$  is the quotient map. Show that the bottom of the cone is homeomorphic to  $X$ .
- (c) Find subspaces of the lowest possible dimension Euclidean space homeomorphic to the following: the cone over a one-point space; the cone over a two-point space; the cone over a closed interval on the real line.
- (d) Embed  $K(S_1)$  into  $\mathbb{R}^3$  as a surface. Sketch this surface and identify its subspace that is its bottom in the sense of (b).
- (e) Construct a homeomorphism  $K(S_1) \cong D_2$ .
- (f) Show that  $K(S_n) \cong D_{n+1}$ .
- (g) Show that  $K(D_n) \cong D_{n+1}$ .

- prob:suspension **192.** The **suspension** of a topological space  $X$  is the quotient space

$$S(X) = (X \times I) / \sim,$$

where  $\sim$  identifies all points of  $X \times \{0\}$  with one another and  $X \times \{1\}$  with one another—in other words, the space obtained by collapsing each of  $X \times \{-1\}$  and  $X \times \{1\}$  to a point in the sense of Exercise 172.

- (a) Prove that, in general,  $S(X)$  can be obtained by attaching  $X \times I$  to a two-point space by means of a suitable map.
- (b) Verify that  $S(S_n) \cong S_{n+1}$ .
- (c) Show that  $S(D_n) \cong D_{n+1}$ .

- prob:zones-on-S2-and-RP2 **193.** Represent the 2-sphere as  $S_2 = D^+ \cup C \cup D^-$ , where:  $D^+$  is the north polar cap consisting of all points above and on the meridian that is the intersection of the sphere with the plane  $x_3 = 1/2$ ;  $D^-$  is the south polar cap consisting of all points below and on the meridian that is the intersection of the sphere with the plane  $x_3 = -1/2$ ; and  $C$  is the equatorial zone consisting of points between and on those two meridians. Let  $q: S_2 \rightarrow \mathbb{RP}_2$  be the quotient map. Verify:

collapsing to a point!attaching@and  
adjunction space  
collapsing to a point  
Möbius strip@M"obius strip!Klein bo  
Klein bottle  
line with two origins  
Hausdorff space!continuous image@  
continuous image!Hausdorff space@  
cone over a space  
adjunction space  
n-sphere@\$n\$-sphere!cone@and co  
n-disk@\$n\$-disk!cone@and cone  
suspension  
suspension!adjunction space@and a  
adjunction space  
n-sphere@\$n\$-sphere!suspension@  
n-disk@\$n\$-disk!suspension@and s  
sphere  
Möbius strip@M"obius strip

projective plane

(a)  $D^+ \cong D_2$  and  $C \cong S_1 \times I$ .

adjunction space

(b)  $q|_{D^+}$  is an embedding of the north polar cap into the projective plane.

quotient space!separation properties@and separation properties

separation properties!quotient space@of quotient space

(c)  $q(C) \cong M$ , the Möbius strip, and  $q(C^+) \cong \partial M$ , the edge of  $M$  (Exercise 162), where  $C^+$  is the meridian that is the intersection of the sphere with the plane

$$x_3 = 1/2.$$

open equivalence relation

closed equivalence relation

(d)  $RP_2 = q(C) \cup q(D^+)$  with  $q(C) \cap q(D^+) \cong S_1$ .

equivalence relation!open

prob:RP2-by-attaching-disk-to-M

equivalence relation!closed

open equivalence relation

closed equivalence relation

194. (Continuation of Exercise 193.) Deduce that the projective plane  $RP_2$  can be obtained by attaching the disk  $D_2$  to the Möbius strip  $M$  by an embedding  $f: S_1 \rightarrow M$  with  $f(S_1) = \partial M$ .

prob:separation-props-of-quotients

equivalence relation!open

equivalence relation!closed

195. Let  $\sim$  be an equivalence relation on a topological space  $X$ . Prove:

(a) The quotient space  $X/\sim$  is a  $T_1$ -space if and only if the equivalence class of each  $x \in X$  is closed in  $X$ .

(b) The quotient space  $X/\sim$  is a  $T_2$ -space if and only if each two distinct equivalence classes for  $\sim$  have disjoint open *saturated* neighborhoods.

196. Let  $\sim$  be an equivalence relation on a space  $X$ .

part:graph-open-then-quot-discrete

(a) Prove: If the graph of  $\sim$  is open in  $X \times X$ , then the quotient space  $X/\sim$  is discrete.

(b) Does the converse of (a) hold?

prob:quot-not-T2-yet-graph-closed

197. Let  $A = \mathbb{Q}^*$  and  $B = \mathbb{R}^* \setminus \mathbb{Q}$ , whence  $A$  and  $B$  are disjoint dense subsets of  $\mathbb{R}^*$ , and let

$$X = (A \times [0, +\infty[) \cup (B \times ]-\infty, 0]) \cup (\{0\} \times \mathbb{R}^*),$$

so that  $X$  is metrizable. Define  $\sim$  to be the equivalence relation on  $X$  that identifies  $\langle x, y \rangle$  with  $\langle u, v \rangle$  when  $x = u$  and, if in addition  $x = u = 0$ , when  $y$  and  $v$  have the same sign. Show that  $X/\sim$  is not a  $T_2$ -space even though the graph of  $\sim$  is closed in  $X \times X$ .

*Note:* Thus this example shows that the condition in Proposition 3.89 is not sufficient for a quotient space to be  $T_2$ .

prob:saturate-subsets-of-square

198. Determine the saturation (Definition 0.100) of the specified subsets  $E$  of  $X = I \times I$  by the given equivalence relation  $\sim$  on  $X$ :

(a) Let  $\sim$  collapse  $\{0\} \times I$  to a point. Take  $E$  to be an arbitrary subset of  $X$ , but consider separately the cases that  $E$  does or does not intersect  $\{0\} \times I$ .

(b) Let  $\sim$  collapse  $]0, 1] \times I$  to a point. Take  $E$  to be an arbitrary subset of  $X$ , but consider separately the cases that  $E$  is or is not contained in  $\{0\} \times I$ .

(c) Let  $x \sim u$  if and only if  $x = u$  or  $x_1 = u_1 \in I \setminus \mathbb{Q}$ . First take  $E = I \times [0, 1/2[$ ; then take  $E = I \times [0, 1/2]$ .

ob:examples-open-closed-equiv-rel

199. (Continuation of Exercise 198.) An equivalence relation  $\sim$  on a topological space  $X$  is said to be **open** (respectively, **closed**) when the saturation of each open (respectively, closed) subset of  $X$  by  $\sim$  is itself open (respectively, closed) in  $X$ .

Determine which of the equivalence relations on  $I \times I$  from Exercise 198 are open and which are closed.

prob:open-closed-equiv-rel-criteria

200. (Continuation of Exercise 199.) Let  $\sim$  be an equivalence relation on a topological space  $X$ .

closed-equiv-rel-and-quotient-map

- (a) Explain why  $\sim$  is open (respectively, closed) precisely when the quotient map  $q: X \rightarrow X/\sim$  is an open (respectively, closed) map.

quotient space!separation properties  
separation properties!quotient space

rel-closed-via-nbds-of-equiv-classes

- (b) Show that  $\sim$  is closed if and only if, for each equivalence class  $A$  under  $\sim$  and each open set  $U$  in  $X$  with  $A \subset U$ , there is some saturated open set  $V$  in  $X$  with  $A \subset V \subset U$ .

open equivalence relation  
closed equivalence relation  
equivalence relation!open

- 201.** (Continuation of [Exercise 199](#).) Let  $\sim$  be an equivalence relation on a space  $X$ . Prove or disprove:

equivalence relation!closed  
Kolmogoroff quotient  
wedge sum  
pointed space

ph-eq-rel-closed-then-eq-rel-closed

- (a) If the graph of  $\sim$  is a closed subset of  $X \times X$ , then  $\sim$  is closed in the sense of [Exercise 199](#).

graph-eq-rel-open-then-eq-rel-open

- (b) If the graph of  $\sim$  is an open subset of  $X \times X$ , then  $\sim$  is open in the sense of [Exercise 199](#).

- (c) The converse of (a).

- (d) The converse of (b).

aph-criterion-when-quot-map-open

- 202.** (Continuation of [Exercise 199](#).)

ce-T2-iff-rel-open-and-graph-closed

- (a) Prove: If  $X$  is a  $T_2$ -space and if the equivalence relation  $\sim$  is open, then the quotient space  $X/\sim$  is a  $T_2$ -space if and only if the graph of  $\sim$  is a closed subset of  $X \times X$ .

- (b) Show that the hypothesis that  $X$  be a  $T_2$ -space is essential in (a).

prob:kolmogoroff-quotient

- 203.** Let  $X$  be an arbitrary topological space. Define an equivalence relation  $\sim$  on  $X/\sim$  by  $x \sim y$  if and only if  $\{x\}^- = \{y\}^-$ , and let  $X_0 = X/\sim$  be the corresponding quotient space.

- (a) Show that the quotient space  $X_0$  is a  $T_0$ -space. (This space may be called the **Kolmogoroff quotient of  $X$** , or the  **$T_0$ -ization of  $X$** .)

- (b) If  $X$  is already a  $T_0$ space, show that  $X_0$  is homeomorphic to  $X$ .

- 204.** Which of the quotient spaces in [Exercise 176](#) are Hausdorff spaces?

prob:wedge-sum

- 205.** Let  $\langle X, x_0 \rangle$  and  $\langle Y, y_0 \rangle$  be pointed spaces ([Definition 2.4](#)). Their **wedge sum** is the pointed space

$$\langle X, x_0 \rangle \vee \langle Y, y_0 \rangle = \langle (X + Y)/\sim, z_0 \rangle,$$

where  $(X + Y)/\sim$  is the quotient space of the Cartesian sum  $X + Y$  [[Examples 3.47 \(4\)](#)] obtained by identifying  $x_0$  with  $y_0$  and where the base point  $z_0$  is the point of  $(X + Y)/\sim$  obtained from  $x_0$  and  $y_0$ .

*Abuse of language:* It is common to suppress mention of the particular base points and to write simply  $X \vee Y$  instead of  $\langle X, x_0 \rangle \vee \langle Y, y_0 \rangle$ . [See (e).]

[*Technical note:* What happens if  $x_0 = y_0$ ? In this case, the spaces  $X$  and  $Y$  are not disjoint, so that by definition the Cartesian sum  $X + Y = (X \times \{1\}) \cup (Y \times \{2\})$ . Then it is actually the *distinct* points  $\langle x_0, 1 \rangle$  and  $\langle y_0, 2 \rangle$  that are identified when forming  $(X + Y)/\sim$ .]

- (a) Write explicitly the equivalence relation  $\sim$  on  $X + Y$  for which the space of the wedge sum is  $(X + Y)/\sim$ .

- (b) Verify that  $X \vee Y$  is homeomorphic to the adjunction space obtained by attaching the base point of  $X$  to  $Y$ .

- (c) To what subspace of the plane is the wedge sum  $S_1 \vee S_1$  of the circle with itself homeomorphic?

smash product  
pointed space  
double

(d) Show that  $S_n \vee S_n \cong S_n // E$ , where  $E$  is the “equator” given by

$$E = \{ \langle x_1, x_2, \dots, x_n, 0 \rangle : \| \langle x_1, x_2, \dots, x_n \rangle \| = 1 \}.$$

part:wedge-sum-depends-base-pt  
manifold-with-boundary

(e) Show that, when both  $X$  and  $Y$  are homogeneous in the sense of Exercise 76, then  $\langle X, x_0 \rangle \vee \langle Y, y_0 \rangle$  is independent of the choice of base points  $x_0$  and  $y_0$ , that is,  $\langle X, x_0 \rangle \vee \langle Y, y_0 \rangle \cong \langle X, x_1 \rangle \vee \langle Y, y_1 \rangle$  for all  $x_1 \in X$  and all  $y_1 \in Y$ . Show further, that when  $X$  or  $Y$  is not homogeneous, then  $\langle X, x_0 \rangle \vee \langle Y, y_0 \rangle$  need not be independent of the choice of base points.

prob:smash-prod 206. Let  $\langle X, x_0 \rangle$  and  $\langle Y, y_0 \rangle$  be pointed spaces (Definition 2.4). Their **smash product** is the pointed space

$$\langle X, x_0 \rangle \wedge \langle Y, y_0 \rangle = ((X \times Y) / \sim, z_0),$$

where  $\sim$  identifies  $\langle x_0, y \rangle$  with  $\langle x, y_0 \rangle$  for every  $x \in X$  and every  $y \in Y$  and where  $z_0$  is the point of  $(X \times Y) / \sim$  obtained from all those identified points.

*Abuse of language:* As with the wedge sum, so with the smash product it is common to suppress mention of the particular base points and to write simply  $X \wedge Y$  instead of  $\text{opair} X, x_0 \wedge \langle Y, y_0 \rangle$ .

- (a) Write explicitly the equivalence relation  $\sim$  on  $X \times Y$  for which the space of the smash product is  $(X \times Y) / \sim$ .
- (b) To what familiar subspace of a Euclidean space is  $\langle [0, 1], 0 \rangle \wedge \langle [0, 1], 1 \rangle$  homeomorphic?
- (c) Show that  $S_1 \wedge S_1 \cong S_2$ .
- (d) Pick a point  $p \in S_n$ . Show that  $\langle S_n, p \rangle \wedge \langle S_n, p \rangle \cong S_n // E$ , where  $E$  is the equator of the  $n$ -sphere, given by

$$E = \{ \langle x_1, x_2, \dots, x_n, x_{n+1} \rangle \in \mathbb{R}^{n+1} : x_{n+1} = 0 \}.$$

- (e) Is  $\langle X, x_0 \rangle \wedge \langle Y, y_0 \rangle$  independent of the choice of base points when  $X$  and  $Y$  are both homogeneous? when they are not homogeneous? [Compare Exercise 205 (e).]
- (f) Show that  $X \wedge Y \cong (X \times Y) / ((X \times \{y_0\}) \cup (\{x_0\} \times Y))$ .
- (g) Generalize the definition of smash product to the case of an arbitrary family of pointed spaces.

prob:dbl-mfld-with-bdy 207. Let  $M$  be an  $n$ -manifold-with-boundary (see Exercise 91). The **double** of  $M$  is the topological space  $M \cup_j M$ , where  $j: \partial M \rightarrow M$  is the inclusion map of the boundary  $\partial M$  of  $M$  (see Exercise 97).

- (a) Show that the double of the unit interval  $I$  is homeomorphic to  $S_1$  and that the double of the disk  $D_2$  is homeomorphic to the sphere  $S_2$ . Is the double of  $D_n$  homeomorphic to  $S_n$ ?
- (b) Show that the double of the closed-halfspace  $\mathbb{R}_+^n = \{x \in \mathbb{R}^n : x_n \geq 0\}$  is homeomorphic to  $\mathbb{R}^n$ .
- (c) Deduce that the double of an arbitrary  $n$ -manifold-with-boundary is an  $n$ -manifold in which the map  $h: M \rightarrow M + M$  induces an embedding of  $M$  as a closed subspace of that double.

[Note: This result shows that every topological manifold-with-boundary is, in essence, a closed subspace of a manifold (without boundary).]

prob:attach-mflds-with-bdy-along-bdy

**208.** Let  $M$  and  $N$  be  $n$ -manifolds-with-boundary (Exercise 91) having homeomorphic boundaries  $\partial M$  and  $\partial N$ , respectively (Exercise 97). Let  $f: \partial M \cong \partial N$ . We say that the adjunction space  $M \cup_f N$  is obtained by attaching  $M$  and  $N$  along their boundaries.

Show that  $M \cup_f N$  is an  $n$ -manifold in which the maps  $h: M \rightarrow M \cup_f N$  and  $k: N \rightarrow M \cup_f N$  are embeddings with closed images.

(Notes: Such an adjunction space generalizes the double of a manifold-with-boundary as defined in Exercise 207. You may wonder whether this situation is really more general, that is: whether there are two  $n$ -manifolds-with-boundary that have homeomorphic boundaries yet are not themselves homeomorphic? The answer is yes: for example, remove a subspace of the sphere  $S_2$  homeomorphic to the ball  $B_2$  to obtain a 2-manifold-with-boundary  $M$  and remove a subspace of the torus  $T_2$  likewise homeomorphic to the same ball to obtain another 2-manifold-with-boundary  $N$ ; then  $M$  is not homeomorphic to  $N$ , although we shall not develop the tools to prove that until Chapter 5 (Connectedness).)

prob:quotient-space-G-mod-H

**209.** Let  $G$  be a topological group (Exercise 145). and let  $H$  be a subgroup of  $G$ . Denote by  $G/H$  not the space  $G//H$  obtained by collapsing  $H$  to a point, but rather the collection  $\{gH : g \in G\}$  of all left cosets of  $G$  in  $H$ , where  $gH = \{gh : h \in H\}$  for each  $g \in G$ . Thus  $G/H$  is a partition of  $G$  whose associated equivalence relation  $\sim$  is given by

$$x \sim y \iff xH = yH.$$

Let  $q: G \rightarrow G/H = G/\sim$  be the quotient map. Prove:

- (a) The equivalence relation  $\sim$  is open (Exercise 199).
- (b) If  $x \in G$ , then the neighborhood system at  $q(x) = xH$  in the quotient space  $G/H$  is the collection

$$\{q(U) : U \text{ is open in } G \text{ and } x \in UH\}.$$

Here  $UH$  denotes the set  $\{uh : u \in U, h \in H\}$  of all products of elements of  $U$  with elements of  $H$ .

- (c) The space  $G/H$  is *homogeneous* in the sense of Exercise 76.

**210.** (Continuation of Exercise 209.) What can be said about the space  $G/H$  when  $H$  is open in  $G$ ? when  $H$  is closed in  $G$ ? when  $H$  is dense in  $G$ ?

prob:quotient-top-group

**211.** (Continuation of Exercise 209.)

- (a) Suppose now that  $H$  is a **normal** subgroup of  $G$  in the algebraic sense, that is:  $gH = Hg$  for all  $g \in G$ . (This is always the case when the group operation of  $G$  is commutative, that is, when  $G$  is an Abelian group.) Thus the quotient set  $G/H$  is a group under the operation  $(xH) \cdot (yH) = (xy)H$ .

Show that the quotient topology makes  $G/H$  into a topological group.

- (b) What is the topology on the quotient topological group  $\mathbb{R}/\mathbb{Q}$ ?
- (c) Show that  $\mathbb{R}/\mathbb{Z}$  is, as a topological group, essentially the same as the circle group  $S_1$  [Exercise 145 (e)] both algebraically and topologically by constructing a group isomorphism from  $\mathbb{R}/\mathbb{Z}$  onto the circle group that is also a homeomorphism.
- (d) Show that  $\mathbb{R}^2/\mathbb{Z}^2$  is, as a topological group, essentially the same as the topological group  $S_1 \times S_1$  (the torus).

prob:induced-map-quotient-G-mod-ker

**212.** (Continuation of Exercise 211.)

manifold-with-boundary

topological group

topological group!quotient space@

homogeneous space

topological group!quotient group@a

normal subgroup

circle group

circle group

torus!topological group@as topolog

torus

continuous map!group homomorph

**(a)** Let  $\varphi: G \rightarrow H$  be a continuous group homomorphism from a topological group  $G$  to a topological group  $H$ . Form the kernel  $\ker(\varphi)$  of  $\varphi$ , defined as  $\ker(\varphi) = \{x \in G : \varphi(x) = e\}$ , where  $e$  is the identity element of the group  $H$ . Since  $\ker(\varphi)$  is a normal subgroup of  $G$ , the quotient topology makes the quotient group  $G/\ker(\varphi)$  into a topological group. Show that  $\varphi$  is constant on each coset  $x\ker(\varphi)$ , and hence that the map  $\varphi^*: G/\ker(\varphi) \rightarrow H$  obtained by passing to the quotient is also a continuous group homomorphism. Show, further, that  $\varphi^*$  is injective—hence is a continuous monomorphism.

**(b)** Show by example that  $\varphi^*$  need not be an embedding. Deduce that even when  $\varphi$  is a continuous epimorphism—so that  $\varphi^*$  is a continuous group isomorphism—the map  $\varphi^*$  need not be a homeomorphic isomorphism.

**213.** Verify that for  $R = 1$  and  $r = 1/2$  the map  $F: [0, 2\pi] \times [0, 2\pi] \rightarrow \mathbb{R}^4$  given by

$$F(\theta, \varphi) = \langle (r \cos \theta + R) \cos \varphi, (r \cos \theta + R) \sin \varphi, r \sin \theta \cos(\varphi/2), r \sin \theta \sin(\varphi/2) \rangle$$

is an embedding whose image is homeomorphic to the Klein bottle.

**214.** Must a quotient space of a Lindelöf space (Exercise 2.115) itself be a Lindelöf space?

**215.** *Note:* This exercise presents a generalization of both quotient spaces and Cartesian sums (Exercise 106). This generalization is “dual” to the notion of *initial topology* that was discussed in Exercise 144.

Let  $Y$  be a set, let  $\langle X_i \rangle_{i \in I}$  be a family of topological spaces, and for each  $i \in I$  let  $f_i: X_i \rightarrow Y$  be a map. Let  $\mathcal{T}$  be the collection of those subsets  $V$  of  $Y$  for which  $f_i^{-1}(V)$  is open in  $X_i$  for each  $i \in I$ .

**(a)** Verify that  $\mathcal{T}$  is a topology on  $Y$  and is, in fact, the *finest* topology on  $Y$  that makes every  $f_i$  continuous.

We call  $\mathcal{T}$  the **final topology on  $Y$  induced by  $\langle f_i \rangle_{i \in I}$** .

**(b)** Let  $g: Y \rightarrow Z$  be a map to a topological space  $Z$ . Show that  $g$  is continuous for the final topology  $\mathcal{T}$  on  $Y$  if and only if the composite  $g \circ f_i: X_i \rightarrow Z$  is continuous for each  $i \in I$ .

Show, further, that  $\mathcal{T}$  is the unique topology on  $Y$  having the preceding property.

**(c)** Explain why the topology on a quotient space is a final topology.

**(d)** Show that the topology on the Cartesian sum (Exercise 106) of a family of topological spaces is a final topology.

### 3.5 Convergence

sec:converge

In Section 3.1 we generalized the notion of continuity from the setting of metric spaces to that of topological spaces; now we do the same with the notion of sequential convergence.

#### Sequential convergence in topological spaces

subsec:seq-conv-topological

Our definition of sequential convergence in an arbitrary topological space is based upon Corollary 1.63, which said that convergence of a sequence in a metric space  $\langle X, d \rangle$  depends only on  $d$ -neighborhoods and hence only on neighborhoods as defined by the topology induced by  $d$ .

def:seq-converge

**3.90 Definition.** Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence of points of a topological space  $X$ . If  $x \in X$ , then we say that  $\langle x_n \rangle_{n \in \mathbb{N}}$  **converges to  $x$  in  $X$**  and write “ $\langle x_n \rangle_{n \in \mathbb{N}} \rightarrow x$  in  $X$ ” to mean that for each neighborhood  $V$  of  $x$  in  $X$  there exists some  $m \in \mathbb{N}$  such that

$$n \geq m \implies x_n \in V.$$

When the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to some point in  $X$ , we say that  $\langle x_n \rangle_{n \in \mathbb{N}}$  **converges in  $X$** .

Often we omit the “in  $X$ ” qualification when there is no ambiguity as to which space is at issue.

Since a sequence converges to a point in a metric space precisely when it converges to the same point in the associated topological space, we already have an adequate supply of examples from [Section 1.4](#). The first thing to do, then, is to generalize some of the properties of sequential convergence to arbitrary topological spaces. To do so in the absence of any metric, however, requires some topological assumptions.

thm:in-cls-via-sequences

**3.91 Theorem.** Let  $x$  be a point and let  $A$  be a subset of a first-countable space  $X$ . Then  $x \in \text{cls } A$  if and only if there is some sequence of points of  $A$  that converges to  $x$  in  $X$ .

**Proof.** Assume first that there is some sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  of points of  $A$  that converges to  $x$  in  $X$ . Let  $V$  be an arbitrary neighborhood of  $x$  in  $X$ . Then there is some  $m \in \mathbb{N}$ —in fact, infinitely many such  $m$ —such that  $x_m \in V$ . Then  $x_m \in V \cap A$ , and so  $V \cap A \neq \emptyset$ . Hence  $x \in \text{cls } A$ . (First-countability was not used so far.)

Conversely, assume that  $x \in \text{cls } A$ . Choose a *decreasing* sequence  $\langle V_n \rangle_{n \in \mathbb{N}}$  of subsets of  $X$  such that  $\{V_n : n \in \mathbb{N}\}$  is a local base at  $x$  (see [Exercise 2.83](#)). For each  $n \in \mathbb{N}$  we may choose some point

$$x_n \in A \cap V_n$$

whence  $x_n \in V_j$  for each  $j = 0, 1, \dots, n$ . Then the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  of points of  $A$  converges to  $x$  in  $X$ . In fact, if  $V$  is an arbitrary neighborhood of  $x$ , then

$$V_m \subset V$$

for some  $m \in \mathbb{N}$ , and hence

$$n \geq m \implies x_n \in V_m \subset V. \quad \square$$

The first part of the corollary below generalizes [Theorem 1.67](#); the second part generalizes [Corollary 1.68](#).

or:open-closed-via-seq-1st-countable

**3.92 Corollary.** Let  $A$  be a subset of a first-countable space  $X$ . Then:

cor-part:closed-via-sequences

(1) The set  $A$  is closed in  $X$  if and only if each point of  $X$  to which some sequence of points of  $A$  converges in  $X$  itself belongs to  $A$ .

cor-part:open-via-sequences

(2) The set  $A$  is open in  $X$  if and only if each sequence of points of  $X$  that converges in  $X$  to a point of  $A$  is eventually in  $A$ .

**Proof.** (1) The set  $A$  is closed in  $X$  if and only if  $\text{cls } A \subset A$ . Now apply [Theorem 3.91](#).

(2) Apply (1) to  $X \setminus A$ .  $\square$



countable-complement topology and convergence  
eventually constant sequence!  
eventually constant sequence!

The knowledge of which sequences converge to which points thus completely determines the topology of first-countable spaces. Hence sequential convergence should determine which functions between first-countable spaces are continuous (compare Theorem 1.69).

thm:cont-via-seq-conv-1st-countable

**3.93 Theorem.** Let  $X$  and  $Y$  be first-countable spaces. Then a necessary and sufficient condition for a map  $f: X \rightarrow Y$  to be continuous at a point  $x \in X$  is that for each sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converging to  $x$  in  $X$ , the sequence  $\langle f(x_n) \rangle_{n \in \mathbb{N}}$  of images must converge to  $f(x)$  in  $Y$ .

**Proof.** Necessity. Assume that  $f$  is continuous at  $x$ . Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be an arbitrary sequence that converges to  $x$  in  $X$ . Let  $V$  be an arbitrary neighborhood of  $f(x)$  in  $Y$ . Then  $f^{-1}(V)$  is a neighborhood of  $x$  in  $X$ , and so there is an  $m \in \mathbb{N}$  with  $x_n \in f^{-1}(V)$  for all  $n \geq m$ . Hence  $f(x_n) \in V$  for all  $n \geq m$ .

Sufficiency. Assume that the condition holds. Just suppose there is some neighborhood  $V$  of  $f(x)$  in  $Y$  such that  $f^{-1}(V)$  is not a neighborhood of  $x$  in  $X$ . Choose a decreasing sequence  $\langle U_n \rangle_{n \in \mathbb{N}}$  of sets in  $X$  such that the collection  $\{U_n : n \in \mathbb{N}\}$  is a local base at  $x$ . Then for each  $n \in \mathbb{N}$  we have  $U_n \not\subset f^{-1}(V)$ , and so there is some  $x_n \in U_n$  for which  $x_n \notin f^{-1}(V)$ . Clearly the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  so obtained converges to  $x$  in  $X$ , and so by assumption  $\langle f(x_n) \rangle_{n \in \mathbb{N}}$  converges to  $f(x)$  in  $Y$ . Then  $f(x_m) \in V$  for some  $m$ . Hence  $x_m \in f^{-1}(V)$ , which is impossible.  $\square$

Now that we have shown how easily the definition of sequential convergence generalizes to topological spaces and how nicely sequential convergence works in first-countable spaces, we have the sad duty to report that *sequential convergence is inadequate for describing the topology and continuous maps of arbitrary topological spaces*.

ex:conv-seq-in-countable-compl-top

**3.94 Example.** Let  $X$  be an infinite set provided with its countable-complement topology (Exercise 2.7), so that a subset  $U$  of  $X$  is open if and only if either  $U = \emptyset$  or  $X \setminus U$  is countable.

It is easy to determine all convergent sequences in  $X$ . If  $\langle x_n \rangle_{n \in \mathbb{N}}$  is a sequence that is *eventually constant*—that is, there is some  $x \in X$  and some  $m \in \mathbb{N}$  with

$$\{eq:seq-eventually-cst\} \quad (*) \quad x_n = x \quad \text{for all } n \geq m$$

—then obviously  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x$  in  $X$ . Conversely, suppose a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to a point  $x$  in  $X$ . Since the set  $\{x_n : n \in \mathbb{N}, x_n \neq x\}$  is countable, its complement

$$V = X \setminus \{x_n : n \in \mathbb{N}, x_n \neq x\}$$

is an open neighborhood of  $x$  in  $X$ . By convergence, there is some  $m \in \mathbb{N}$  such that  $x_n \in V$  for all  $n \geq m$ . Thus  $(*)$  holds.

Now suppose that  $X$  is in fact uncountable. Choose some  $y \in X$  and let

$$A = X \setminus \{y\}.$$

- No sequence in  $A$  can converge to  $y$  according to what we just established. Nonetheless,  $y \in \text{cls } A$ .

In fact, let  $V$  be an arbitrary neighborhood of  $y$ . Then  $X \setminus V$  is countable, and so the set  $(X \setminus V) \cup \{y\}$  is also countable. Now the set

$$V \cap A = V \cap (X \setminus \{y\}) = X \setminus ((X \setminus V) \cup \{y\})$$

is uncountable because  $X$  is uncountable. Hence  $V \cap A$  is nonempty.



- The set  $A$  is not closed in  $X$  because it is not countable (also, because  $y \in \text{cls } A$  whereas  $y \notin A$ ). Nonetheless, if a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  of points of  $A$  converges in  $X$  to a point  $x$ , then  $x \in A$  because  $x = x_n$  for all sufficiently large  $n$ .
- The set  $\{y\}$  is not open because its complement  $A$  is not countable (also, because  $A$  is not closed). Nonetheless, if a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to the unique point  $y$  of this set, then  $x_n = y$  and so  $x_n \in \{y\}$  for all sufficiently large integers  $n$ .

Moore, E. Hastings  
Smith, Hermank.

Now let  $Y$  be the set  $X$  but provided with its discrete topology and let  $f: X \rightarrow Y$  be the identity map. Then:

- The map  $f$  is not continuous because the inverse image  $\{y\} = f^{-1}(\{y\})$  of the open subset  $\{y\}$  of  $Y$  is not open in  $X$ . Nonetheless, if a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to a point  $x$  in  $X$ , then there is an  $m$  with  $x_n = x$  for all  $n \geq m$ , and so  $\langle f(x_n) \rangle_{n \in \mathbb{N}} = \langle x_n \rangle_{n \in \mathbb{N}}$  does converge to the point  $f(x) = x$  in  $Y$  [see Examples 1.61 (9).]  $\diamond$

Thus in the absence of first-countability, the properties in Theorems 3.91–3.93 about sequential convergence can cease to hold. The intuitive notion of sequential convergence—points getting closer and closer to a given point—is so appealing, however, that it is desirable to find a more general notion than sequential convergence for which analogs of Theorems 3.91–3.93 do hold. Just such a notion was found by E. H. Moore and H. L. Smith (but with different motivation), who in 1922 initiated the theory of convergence of *nets*.

Sequential convergence is inadequate simply because a sequence has only countably many values, whereas in an arbitrary topological space it may take uncountably many neighborhoods of a point to determine all the neighborhoods of the point. Hence we need to replace the countable set  $\mathbb{N}$  that indexes sequences with more general index sets. Like  $\mathbb{N}$ , these sets will need some kind of ordering but, it turns out, not necessarily a total ordering, nor even a partial ordering. All that we need for ordering such an index set is a *direction* in the sense of Definition 3.95.

### Directed sets and nets

subsec:directed-sets

“Directed” preorderings are another type of preordering playing a role in topology, along with partial orderings, total orderings, and well-orderings.

def:directed-preorder

**3.95 Definition.** A relation  $\leq$  in a set  $X$  is said to **direct**  $X$  and is called a **directed preordering of  $X$** , or more briefly simply a **direction on  $X$** , if it has all three of the following properties:

def:directed-reflexive

(D1) Reflexivity: For each  $x \in X$ ,

$$x \leq x.$$

def:directed-transitive

(D2) Transitivity: For all  $x, y \in X$ ,

$$x \leq y \quad \text{and} \quad y \leq z \implies x \leq z.$$

ef:upper-bound-property-of-direction

(D3) For each  $x \in X$  and  $y \in X$ , there exists some  $z \in X$  such that  $x \leq z$  and  $y \leq z$ .

In other words  $\leq$  directs a set if it preorders it and if, in addition, each set of two elements from the set has an upper bound for the relation  $\leq$  in the set. A set together with a specific directed ordering of it is called a **directed set**.

naturals direction  
direction!natural  
natural direction  
direction!natural

**3.96 Examples.** (1) Any total ordering of a set directs that set. In fact, the first two properties of a directed preordering, reflexivity and transitivity, are shared by any total ordering; the third is a substitute for comparability.

In particular, the usual ordering of  $\mathbb{N}$  directs the set of natural numbers. (The usual ordering of  $\mathbb{N}$  is precisely the preordering we shall use to generalize from sequences to arbitrary “nets.”)

(2) Define the relation  $\leq$  in  $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$  by

$$x \leq y \iff |x| \leq |y|,$$

where the ‘ $\leq$ ’ symbol in ‘ $|x| \leq |y|$ ’ denotes the usual ordering relation of  $\mathbb{R}$ . Then  $\leq$  directs  $\mathbb{R}^*$ .

The direction  $\leq$  does *not* partially order  $\mathbb{R}^*$  because it is not antisymmetric: for example,  $-1 \leq 1$  and  $1 \leq -1$

reverse-interval-inclusion-as-direction

(3) Let  $\mathcal{J}$  be the set of all open intervals in  $\mathbb{R}$  that include 0, that is, all sets of the form  $]a, b[ = \{x \in \mathbb{R} : a < x < b\}$  such that  $a < 0 < b$ . Let  $\leq$  be the *reverse* inclusion relation in  $\mathcal{J}$ , that is,

$$]a, b[ \leq ]c, d[ \iff ]a, b[ \supset ]c, d[$$

for all  $]a, b[, ]c, d[ \in \mathcal{J}$ . That  $\leq$  preorders  $\mathcal{J}$  is easy to check. Property (D3) also holds—so that  $\leq$  directs  $\mathcal{J}$ —because

$$]a, b[ \supset ]\max\{a, c\}, \min\{b, d\}[ , \quad ]c, d[ \supset ]\max\{a, c\}, \min\{b, d\}[$$

for all  $]a, b[, ]c, d[ \in \mathcal{J}$ .

Although the direction  $\leq$  partially orders  $\mathcal{J}$ , it does *not* totally order  $\mathcal{J}$  because, for example, the intervals  $] -1, 1/2[$  and  $] -1/2, 1[$  are not comparable.

The following examples are especially relevant to topology.

ex:natural-dir-on-nbds

(4) Let  $x$  be a point in a topological space  $X$ . The neighborhood system  $\mathcal{N}_x$  at  $x$  is directed by the relation  $\leq$  of “reverse inclusion”:

$$U \leq V \iff U \supset V \quad (U, V \in \mathcal{N}_x).$$

In fact, this relation is clearly reflexive and transitive; and if  $U, V \in \mathcal{N}_x$ , then  $W = U \cap V \in \mathcal{N}_x$  with  $U \leq W$  and  $V \leq W$ . This direction of  $\mathcal{N}_x$  is called the **natural direction**.

Notice that comparability does *not* hold for this direction.

Whenever the neighborhood system at a point in a space is considered as a directed set, it will be with respect to its natural direction.

ex:nbd-pt-directed-set

(5) Again let  $x$  be a point in a space  $X$ . Let

$$I_x = \{\langle U, t \rangle : U \in \mathcal{N}_x, t \in U\}.$$

Then the relation  $\leq$  defined by

$$\langle U, t \rangle \leq \langle V, z \rangle \iff U \supset V$$

is a direction of  $I_x$ , called the **natural direction**. Whenever such a set  $I_x$  is considered as a directed set, it will be with respect to this natural direction.  $\diamond$

As you might expect, the reverse of a direction  $\leq$  on a set is denoted by  $\geq$ .

preordering!directed  
direction

def:net **3.97 Definition.** A **net** is a family indexed by some nonempty directed set. In accordance with the terminology for arbitrary families, a net  $(x_i)_{i \in I}$  is **in** a set  $A$  when

net!eventually constant  
eventually constant net

$$i \in I \implies x_i \in A.$$

In particular, a net is in a set  $X$  when that set is the codomain of the net.

When there is no need to refer to particular or generic entries  $x_i$  of a net  $(x_i)_{i \in I}$ , we may denote the net by a single letter, and then we shall use a lower-case Greek letter such as  $\xi$  or  $\eta$ . [Some mathematicians like to denote a net  $(x_i)_{i \in I}$  by  $x_\bullet$ , a net  $(y_j)_{j \in J}$  by  $y_\bullet$ , etc.]

### Convergence of nets

subsec:nets-converge

For defining convergence of nets, the following language is suggestive.

def:eventually **3.98 Definition.** Let  $(x_i)_{i \in I}$  be a net. Given a subset  $A$  of  $X$ , we say that  $(x_i)_{i \in I}$  is **eventually in**  $A$  to mean:

There exists some  $i \in I$  such that  $x_j \in A$  for all  $j \in I$  with  $j \geq i$ .

ver-n-eventually-in-epsilon-interval **3.99 Example.** Take  $I = \mathbb{N}^*$  with its usual ordering, which directs the set because it totally orders it. Take  $X = \mathbb{R}$ . For each real number  $\varepsilon > 0$ , the sequence  $\langle 1/n \rangle_{n \in \mathbb{N}}$  is eventually in the open interval  $] -\varepsilon, \varepsilon[$  and, a fortiori, eventually in the open interval  $] 0, \varepsilon[$ . This is just a restatement of the fact that  $\lim_{n \rightarrow \infty} 1/n = 0$ , as established in [Examples 1.58 \(1\)](#).  $\diamond$

A sequence is just a special kind of net. Then the definition that a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in a topological space  $X$  converge to a point  $x$  in  $X$  may be restated more concisely as:  $\langle x_n \rangle_{n \in \mathbb{N}}$  is eventually in each neighborhood of  $x$  in  $X$ . To define convergence for arbitrary nets, just mimic that form of the definition for sequences.

def:net-converge **3.100 Definition.** Let  $(x_i)_{i \in I}$  be a net in a topological space  $X$ . If  $x \in X$ , we say that  $(x_i)_{i \in I}$  **converges to**  $x$  (**in**  $X$ ) and write

$$(x_i)_{i \in I} \rightarrow x$$

or, more tersely,

$$x_i \rightarrow x$$

to mean that  $(x_i)_{i \in I}$  is eventually in each neighborhood of  $x$ . We say that  $(x_i)_{i \in I}$  **converges (in**  $X$ ) when it converges to some point  $x$  in  $X$ .

Thus a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges in the sense of [Definition 3.90](#) to a point  $x$  in a topological space  $X$  precisely when it converges in the sense of this definition to  $x$  in  $X$ .

ex:eventually-const:netsconverge **3.101 Examples.** (1) Let  $(x_i)_{i \in I}$  be a net in a space that is **eventually constant**, that is, for which there is an  $x \in X$  and an  $i \in I$  such that

$$j \in I \text{ and } j \geq i \implies x_j = x.$$

Then  $(x_i)_{i \in I} \rightarrow x$ .

ex:net-pts-nbds-converge-to-pt and convergent net (2) Let  $\mathcal{N}_x$  be the neighborhood system at a point  $x$  of a topological space  $X$ . Suppose  $\langle x_U \rangle_{U \in \mathcal{N}_x}$  is any net indexed by  $\mathcal{N}_x$  (with its natural direction) such that

$$x_U \in U \quad (U \in \mathcal{N}_x).$$

Then

$$x_U \rightarrow x.$$

In fact, if  $U$  is an arbitrary neighborhood of  $x$ , then

$$V \in \mathcal{N}_x \text{ and } V \geq U \implies x_V \in V \subset U.$$

(3) Let  $x$  be a point in a topological space  $X$ . Let  $I_x$  be the directed set of [Examples 3.96 \(5\)](#). Then the family  $\langle t \rangle_{\langle U, t \rangle \in I_x}$  is a net that converges to  $x$  in  $X$ .

ex:convergence-in-RI (4) Recall that the Sorgenfrey line  $\mathbb{R}_l$  of [Examples 2.20 \(1\)](#) has as a base the collection of all left-closed, right-open intervals  $[a, b[$ . Then a net in  $\mathbb{R}_l$  converges to a point  $x \in \mathbb{R}_l$  if and only if and only if the net “approaches  $x$  from the right” in the sense that, for each  $\varepsilon > 0$ , the net is eventually in the interval  $[x, x + \varepsilon[$ .

This explains the alternate name ‘lower-limit topology’ for the right-interval topology of the Sorgenfrey line .

ex:Riemann-sums-convergence (5) Let  $a$  and  $b$  be real numbers with  $a < b$  and let

$$f: [a, b] \rightarrow \mathbb{R}$$

be a real-valued function on the closed interval  $[a, b]$  that is bounded, that is, whose range is bounded both below and above in  $\mathbb{R}$ . We shall describe the Riemann integral of  $f$  in terms of convergent nets.

For a positive integer  $n$ , each  $(n + 1)$ -tuple  $\langle x_0, x_1, \dots, x_n \rangle$  of real numbers such that

$$a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$$

will be called a *partition* of  $[a, b]$ . (A partition in this sense is *not* a partition in the set-theoretic sense of [Definition 0.96](#), although the collection  $\{[x_{i-1}, x_i[ : i = 0, 1, \dots, n\}$  of successive subintervals is a partition of  $[a, b[$  in that set-theoretic sense.) Form the collection  $\mathcal{P}$  of all such partitions of  $[a, b]$ . For partitions

$$P = \langle x_0, x_1, \dots, x_n \rangle, \quad Q = \langle y_0, y_1, \dots, y_m \rangle$$

of  $[a, b]$ , define

$$P \leq Q$$

to mean that  $Q$  *refines*  $P$  in the sense that

$$\{x_0, x_1, \dots, x_n\} \subset \{y_0, y_1, \dots, y_m\},$$

in other words,  $Q$  is obtained from  $P$  by inserting zero or more additional points between points of  $P$ .

The relation  $\leq$  certainly is reflexive and transitive. If  $P, Q \in \mathcal{P}$ , then by arranging in increasing order all the points of  $P$  and  $Q$  together, we obtain a partition  $R \in \mathcal{P}$  such that  $P \leq R$  and  $Q \leq R$ . Thus  $\leq$  directs  $\mathcal{P}$ .

Let

$$P = \langle x_0, x_1, \dots, x_n \rangle \in \mathcal{P}.$$

Riemann sum

For each  $i = 1, 2, \dots, n$ , the image

$$f([x_{i-1}, x_i]) = \{f(x) : x_{i-1} \leq x \leq x_i\}$$

is a nonempty set of real numbers which, by the assumption that  $f$  is bounded, has both a lower bound and an upper bound, and so it is meaningful to form the numbers

$$m_i = \inf f([x_{i-1}, x_i]), \quad M_i = \sup f([x_{i-1}, x_i]).$$

Then form the lower Riemann sum

$$L_P = \sum_{i=1}^n m_i (x_i - x_{i-1})$$

and the upper Riemann sum

$$U_P = \sum_{i=1}^n M_i (x_i - x_{i-1})$$

of  $f$  with respect to  $P$ . In geometric terms, when  $f$  is nonnegative, the lower sum  $L_P$  is the area under the step function whose value, for each  $i$ , is  $m_i$  on the  $i$ th interval  $[x_{i-1}, x_i]$  and the upper sum  $U_P$  is the area under the step function whose value is  $M_i$  on the  $i$ th interval (see Figure 3.34).



Figure 3.34: Lower and upper Riemann sums  $L_P$  and  $U_P$  for a partition  $P$  of the closed interval  $[a, b]$ .

fig:Riemann-sums

We now have two nets,  $\langle L_P \rangle_{P \in \mathcal{P}}$  and  $\langle U_P \rangle_{P \in \mathcal{P}}$ , in  $\mathbb{R}$ . If there is a number  $c$  such that

$$L_P \rightarrow c \quad \text{and} \quad U_P \rightarrow c,$$

then  $f$  is said to be *Riemann integrable*, and the *Riemann integral of  $f$*  is defined to be that number  $c$ , which then is denoted by

$$\int_a^b f(x) \, dx.$$

When  $f$  is nonnegative, this number is interpreted as the area under the graph of  $f$ .

It is a fact that the two nets do converge, and to the same number, in case  $f$  is continuous on  $[a, b]$ . In other words, a continuous function on a closed interval is Riemann integrable.

Additional descriptions of the Riemann integral appear below, in [Exercises 222](#) and [223](#).  $\diamond$

By proving generalizations of [Theorem 3.91](#), [Corollary 3.92](#), and [Theorem 3.93](#) and we now demonstrate that nets are indeed adequate for determining the topologies and continuous maps of arbitrary topological spaces.

thm:in-cls-via-nets **3.102 Theorem.** *Let  $x$  be a point and let  $A$  be a subset of a topological space  $X$ . Then  $x \in \text{cls } A$  if and only if there is some net in  $A$  that converges to  $x$  in  $X$ .*

**Proof.** Assume first that there is some net  $(x_i)_{i \in I}$  in  $A$  that converges to  $x$  in  $X$ . Let  $V$  be an arbitrary neighborhood of  $x$  in  $X$ . There is some index  $i \in I$  such that  $x_j \in V$  for all  $j \in I$  for which  $j \geq i$ ; in particular,  $x_i \in V$ . Then  $x_i \in V \cap A$ , and so  $V \cap A \neq \emptyset$ . Hence  $x \in \text{cls } A$ .

Conversely, assume that  $x \in \text{cls } A$ . Then each neighborhood  $U$  of  $x$  in  $X$  intersects  $A$ , and we may choose some point  $x_U$  with

$$x_U \in U \cap A.$$

The points so chosen constitute a net  $\langle x_U \rangle_{U \in \mathcal{N}_x}$  in  $A$  indexed by the neighborhood system  $\mathcal{N}_x$  at  $x$ . According to [Examples 3.101 \(2\)](#), this net converges to  $x$  in  $X$ .  $\square$

In the preceding proof, simultaneous choice of all the points  $x_U$  is an implicit application of the Axiom of Choice ([0.26](#)). The Axiom of Choice may be avoided by a slightly more complicated construction: see [Exercise 228](#).

In the same way that [Corollary 3.92](#) was deduced from [Theorem 3.91](#) the following criteria may deduced from [Theorem 3.102](#).

cor:open-closed-via-nets **3.103 Corollary.** *Let  $A$  be a subset of a topological space  $X$ . Then:*

- cor-part:closed-via-nets (1) *The set  $A$  is closed in  $X$  if and only if each point of  $X$  to which some net in  $A$  converges in  $X$  itself belongs to  $A$ .*
- cor-part:open-via-nets (2) *The set  $A$  is open in  $X$  if and only if each net in  $X$  that converges in  $X$  to a point of  $A$  is eventually in  $A$ .*

thm:cont-via-nets **3.104 Theorem.** *Let  $X$  and  $Y$  be topological spaces. Then a necessary and sufficient condition for a map  $f: X \rightarrow Y$  to be continuous at a point  $x \in X$  is that for each net  $(x_i)_{i \in I}$  in  $X$  converging to  $x$  in  $X$ , the net  $(f(x_i))_{i \in I}$  of images must converge to  $f(x)$  in  $Y$ , in other words:*

$$x_i \rightarrow x \implies f(x_i) \rightarrow f(x).$$

For a net  $\xi = (x_i)_{i \in I}$  in  $X$ , the net  $(f(x_i))_{i \in I}$  of images is, as a map, just the composite  $f \circ \xi$ . Then the condition in [Theorem 3.104](#) for  $f: X \rightarrow Y$  to be continuous at  $x$  is that, for every net  $\xi$  in  $X$ ,

$$\xi \rightarrow x \implies f \circ \xi \rightarrow f(x).$$

**Proof of Theorem 3.104.** Necessity is proved as in the same way as for [Theorem 3.93](#).

For sufficiency, assume that the condition holds. Just suppose there is some neighborhood  $V$  of  $f(x)$  in  $Y$  such that  $f^{-1}(V)$  is *not* a neighborhood of  $x$  in  $X$ . Then for each neighborhood

$U$  of  $x$  in  $X$  we may choose some point  $x_U \in U$  with  $x_U \notin f^{-1}(V)$ . As in [Examples 3.101 \(2\)](#), finite-complement topology we obtain a net  $\langle x_U \rangle_{U \in \mathcal{N}_x}$  that converges to  $x$  in  $X$ . By assumption,  $f(x_U) \rightarrow f(x)$  in  $Y$ , and so  $f(x_U) \in V$  for some  $U \in \mathcal{N}_x$ . This means that  $x_U \in f^{-1}(U)$ , which contradicts the way that  $x_U$  was chosen.  $\square$

As an application of [Theorem 3.104](#), let us prove anew, but without any explicit use of neighborhoods or open sets, that the composite of continuous maps is continuous. Let  $f: X \rightarrow Y$  be continuous at a point  $x \in X$  and let  $g: Y \rightarrow Z$  be continuous at the point  $f(x)$ . We show that the composite  $g \circ f: X \rightarrow Z$  is continuous at  $x$ . Let  $(x_i)_{i \in I}$  be an arbitrary net in  $X$  such that  $(x_i)_{i \in I} \rightarrow x$  in  $X$ . By continuity of  $f$  at  $x$ , the net  $\langle f(x_i) \rangle_{i \in I} \rightarrow f(x)$  in  $Y$ . Then by continuity of  $g$  at  $f(x)$ , the net  $\langle g(f(x_i)) \rangle_{i \in I} \rightarrow g(f(x))$  in  $Z$ . Thus  $\langle (g \circ f)(x_i) \rangle_{i \in I} \rightarrow (g \circ f)(x)$ .

In an arbitrary space there is nothing to prohibit a net—even a sequence—from converging to two different points.

q-conv-2-pts-finite-complement-top

**3.105 Example.** Let  $X$  be an infinite set provided with its finite-complement topology [[Examples 2.3 \(7\)](#)]. Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be any sequence in  $X$  with  $x_i \neq x_j$  whenever  $i \neq j$ . Then

$$x \in X \implies \langle x_n \rangle_{n \in \mathbb{N}} \rightarrow x.$$

In fact, let  $x \in X$  and let  $V$  be an arbitrary neighborhood of  $x$ . Then  $X \setminus V$  is finite whereas  $\{x_n : n \in \mathbb{N}\}$  is infinite, and so there exists some  $i \in \mathbb{N}$  such that  $x_j \notin X \setminus V$  for every  $j \geq i$ . Hence  $x_j \in V$  for every  $j \geq i$ .  $\diamond$

The net in the preceding example behaved in the worst possible way: it converged to every point of the space. In a wide class of spaces, however, a net cannot converge to more than one point.

thm:limits-unique-in-T2-space

**3.106 Theorem (uniqueness of net limits in a Hausdorff space).** Let  $(x_i)_{i \in I}$  be a net in a Hausdorff space  $X$  and let  $x, y \in X$  with

$$x_i \rightarrow x \quad \text{and} \quad x_i \rightarrow y.$$

Then

$$x = y.$$

**Proof.** Just suppose  $x \neq y$ . There are disjoint neighborhoods  $U$  and  $V$  of  $x$  and  $y$ , respectively. By hypothesis there exists an  $i \in I$  such that

$$k \in I \text{ and } k \geq i \implies x_k \in U$$

and a  $j \in I$  such that

$$k \in I \text{ and } k \geq j \implies x_k \in V.$$

Since the index set  $I$  is directed, there is some  $k \in I$  with  $k \geq i$  and  $k \geq j$ . then  $x_k \in U \cap V$ , which is impossible because  $U$  and  $V$  are disjoint.  $\square$

The preceding theorem guarantees, for example, that in [Examples 3.101 \(5\)](#) there is only one number  $c$  to which the net of lower Riemann sums and the net of upper Riemann sums each converges, and hence that the Riemann integral of a Riemann integrable function is uniquely defined. More generally, the theorem justifies the following definition.

net in Hausdorff space  
def: net limit in T2  
Hausdorff space

**3.107 Definition.** Let  $(x_i)_{i \in I}$  be a net that converges in a Hausdorff space  $X$ . Then the unique  $x \in X$  for which

$$x_i \rightarrow x$$

is called the **limit of**  $(x_i)_{i \in I}$  and is denoted by

$$\lim_{i \in I} (x_i)_{i \in I}$$

or by

$$\lim_{i \in I} x_i$$

or, even more tersely, by

$$\lim x_i.$$

Recall from [Example 1.64](#) that a sequence in the product space  $\mathbb{R}^n$  converges exactly when it converges “coordinatewise.” The next theorem will assert that the same thing is true for a net in any product space.

In [Section 3.31](#), and later, we denote the index set of a product space by ‘ $I$ ’ and the indices of its individual factor spaces by lower-case Roman letters ‘ $i$ ’, ‘ $j$ ’, etc. Yet in the present section we have been using ‘ $I$ ’ for the index set of a net. In the theorem below we therefore depart temporarily from our customary notation for nets and instead denote the index set of a net by ‘ $A$ ’ and indices by lower-case Greek letters ‘ $\alpha$ ’, ‘ $\beta$ ’, etc. Then for a net  $\langle x_\alpha \rangle_{\alpha \in A}$  in the product of a family  $\langle X_i \rangle_{i \in I}$  of spaces and an index  $\alpha \in A$ , for each  $i \in I$  we use the shorthand  $x_\alpha^i$  for the  $i$ th coordinate  $(x_\alpha)_i$  of the point  $x_\alpha$  in the product space.

converge-in-product-iff-coordinatewise

**3.108 Theorem (convergence of a net in a product).** Let  $\langle x_\alpha \rangle_{\alpha \in A}$  be a net in the product  $X$  of a family  $\langle X_i \rangle_{i \in I}$  of topological spaces, so that

$$x_\alpha = \langle x_\alpha^i \rangle_{i \in I} \quad (\alpha \in A).$$

Let

$$y = \langle y^i \rangle_{i \in I}$$

be a point of  $X$ . Then

$$\langle x_\alpha \rangle_{\alpha \in A} \rightarrow y \text{ in } X$$

if and only if

$$\langle x_\alpha^i \rangle_{\alpha \in A} \rightarrow y^i \text{ in } X_i \text{ for each } i \text{ in } I.$$

**Proof.** Assume first that  $\langle x_\alpha \rangle_{\alpha \in A} \rightarrow y$  in  $X$ . For each  $i \in I$ , the projection  $p_i: X \rightarrow X_i$  is continuous, so by [Theorem 3.104](#)

$$\langle x_\alpha^i \rangle_{\alpha \in A} = \langle p_i(x_\alpha) \rangle_{\alpha \in A} \rightarrow p_i(y) = y^i$$

in  $X_i$ , as desired.

Conversely, assume that  $\langle x_\alpha^i \rangle_{\alpha \in A} \rightarrow y^i$  in  $X_i$  for each  $i$  in  $I$ . Let  $V$  be an arbitrary neighborhood of  $y$  in  $X$ . Then  $V$  contains some basic open neighborhood

$$U = \bigcap_{i \in I} U_i$$

of  $y$ . Let

$$J = \{j \in I : U_j \neq X_j\}.$$

If  $J = \emptyset$ , then  $V = U = X$  and surely  $x_\alpha \in V$  for all  $\alpha \in A$ . Suppose now that  $J \neq \emptyset$ . For each  $j \in J$  the set  $U_j$  is a neighborhood of  $y^j$  in  $X_j$ , and so by assumption there is an index



$\alpha_j \in A$  such that

$$\beta \in A \quad \text{and} \quad \beta \geq \alpha_j \implies x_{\alpha^j} \in U_j.$$

Since  $J$  is finite and  $A$  is directed, there is some  $\gamma \in A$  with

$$\gamma \geq \beta_j \quad (j \in J).$$

Then

$$\begin{aligned} \beta \in A \text{ and } \beta \geq \gamma &\implies x_{\beta^j} \in U_j \text{ for each } j \in J \\ &\implies x_{\beta} \in U. \end{aligned}$$

Hence  $x_{\beta} \in V$  for all  $\beta \geq \gamma$ . This prove that  $\langle x_{\alpha} \rangle_{\alpha \in A} \rightarrow y$ .  $\square$

The topology of pointwise convergence [Examples 2.93 (2)] on the set of functions from one space to another is, according to (Example 3.71), a product topology. Then the preceding Theorem 3.108 tells us which nets converge to which functions for this topology. It is nevertheless instructive to examine convergence for the topology of pointwise convergence directly, without explicit reference to product spaces.

**3.109 Examples.** (1) Let  $\mathcal{F}(X, Y)$  be the set of all functions from a topological space  $X$  to a topological space  $Y$ . Recall from Examples 2.93 (2)—and from Examples 2.72 (8) in the case  $X = Y = \mathbb{R}$ —that the topology of pointwise convergence has as a base the collection of all sets of the form

$$\begin{aligned} B(x_1, x_2, \dots, x_n; V_1, V_2, \dots, V_n) \\ = \{f \in \mathcal{F}(X, Y) : f(x_k) \in V_k \text{ for each } k = 1, 2, \dots, n\} \end{aligned}$$

where  $x_1, x_2, \dots, x_n$  are points of the domain  $X$  and  $V_1, V_2, \dots, V_n$  are open sets in the codomain  $Y$ .

Let  $f \in \mathcal{F}(X, Y)$  and let  $(f_i)_{i \in I}$  be a net in  $\mathcal{F}(X, Y)$ . We claim that

$$(f_i)_{i \in I} \rightarrow f \text{ in } \mathcal{F}(X, Y)$$

for the topology of pointwise convergence if and only if  $(f_i)_{i \in I}$  **converges pointwise** in the sense that

$$\text{for each } x \in X, \quad (f_i(x))_{i \in I} \rightarrow f(x) \text{ in } Y.$$

To verify that claim, assume first that  $f_i \rightarrow f$  in  $\mathcal{F}(X, Y)$ . Let  $x \in X$  be arbitrary. If  $V$  is an arbitrary open neighborhood of  $f(x)$  in  $Y$ , then  $B(x; V)$  is an open neighborhood of  $f$  in  $\mathcal{F}(X, Y)$ , so by assumption there is an  $i \in I$  with  $f_j \in B(x; V)$  for all  $j \geq i$ , that is,  $f_j(x) \in V$  for all  $j \geq i$ . Thus  $f_i(x) \rightarrow f(x)$  in  $Y$ .

Conversely, assume that  $f_i(x) \rightarrow f(x)$  in  $Y$  for each  $x \in X$ . Consider a basic open neighborhood  $W$  of  $f$  of the form

$$W = B(x_1, x_2, \dots, x_n; V_1, V_2, \dots, V_n)$$

as above. Then for each  $k = 1, 2, \dots, n$ ,

$$f(x_k) \in V_k$$

and since  $f_i(x_k) \rightarrow f(x_k)$  in  $Y$ , there is an index  $i_k \in I$  such that

$$j \geq i_k \implies f_j(x_k) \in V_k.$$

Since  $I$  is directed, there is an  $i \in I$  with

$$i \geq i_k \quad (k = 1, 2, \dots, n).$$

pointwise convergence  
topology!of pointwise convergence  
converge!pointwise  
pointwise convergence!and nets

pointwise convergence Then

topology!of pointwise convergence

topology!of uniform convergence

uniform convergence

Hence  $f_i \rightarrow f$  in  $\mathcal{F}(X, Y)$ .

topology!of uniform convergence

prob:top-unif-conv (2) Let  $X, Y$ , and  $\mathcal{F}(X, Y)$  be as in (1), and suppose now that the topology of  $Y$  is induced by a *bounded* metric  $d$ . Then for functions  $f$  and  $g$  belonging to  $\mathcal{F}(X, Y)$ , the set  $\{d(f(x), g(x)) : x \in X\}$  of nonnegative real numbers has an upper bound, and so it is meaningful to form

$$d_\infty(f, g) = \sup_{x \in X} d(f(x), g(x)).$$

This formula defines a metric  $d_\infty$  on  $\mathcal{F}(X, Y)$  (compare [Example 1.16](#) and the discussion preceding [Theorem 1.83](#)), and the topology on  $\mathcal{F}(X, Y)$ —and on any subset of  $\mathcal{F}(X, Y)$  as well—is called the **topology of uniform convergence**. Note that

$$d_\infty(f, g) \leq \varepsilon \iff d(f(x), g(x)) \leq \varepsilon \text{ for all } x \in X.$$

A net  $(f_i)_{i \in I}$  will converge to a function  $f$  in  $\mathcal{F}(X, Y)$  for this topology precisely when  $(f_i)_{i \in I}$  is eventually in each  $d_\infty$ -ball of  $f$ . Hence

$$(f_i)_{i \in I} \rightarrow f \text{ in } \mathcal{F}(X, Y) \text{ for the topology of uniform convergence}$$

if and only if  $(f_i)_{i \in I}$  **converges uniformly to**  $f$  in the sense that

for each  $\varepsilon > 0$ , there exists an  $i \in I$  such that

$$d(f_i(x), f(x)) < \varepsilon \text{ for all indices } j \geq i \text{ and all points } x \in X.$$

Uniform convergence in the above sense is exactly the same uniform convergence used for sequences of functions in advanced calculus. Compare what it means for  $(f_i)_{i \in I}$  to converge *pointwise* to  $f$ :

For each  $\varepsilon > 0$  and for each  $x \in X$ , there exists an  $i \in I$  such that

$$d(f_j(x), f(x)) < \varepsilon \text{ for all } j \geq i.$$

Thus for pointwise convergence, the value of  $i$  depends on both the given  $\varepsilon$  and the particular point  $x$ . By contrast, for uniform convergence, the value of  $i$  depends solely on the given  $\varepsilon$ , so that the same value of  $i$  “works” for all points  $x \in X$ .  $\diamond$

It is worth recording the following observation about nets of functions which holds, in particular, for sequences of functions.

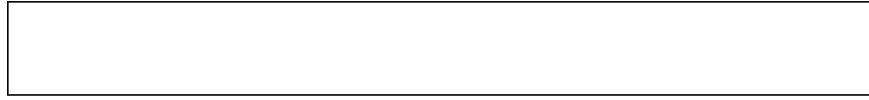
**3.110 Proposition.** Let  $X$  be a topological space, let  $(Y, d)$  be a bounded metric space, let  $\mathcal{F}(X, Y)$  be the set of all functions from  $X$  to  $Y$ . If a net  $(f_i)_{i \in I}$  in  $\mathcal{F}(X, Y)$  converges uniformly in  $\mathcal{F}(X, Y)$ , then it converges pointwise there and the pointwise limit of the net is the uniform limit of that net.

The converse of [Proposition 3.110](#) fails—even when the net is a sequence: see [Exercise 236](#).

The sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in  $\mathbb{R}$  given by

$$x_n = \begin{cases} n & \text{if } n \text{ is even,} \\ 1 - \frac{1}{n+1} & \text{if } n \text{ is odd} \end{cases}$$

does not converge in  $\mathbb{R}$ , yet its values do “pile up” at the point 1 (see [Figure 3.35](#)). In order



clustering net  
net! and clustering

Figure 3.35: A sequence in  $\mathbb{R}$  that clusters but does not converge.

fig:non-conv-clustering-seq

to study nets behaving in such a way, we shall introduce next the notion of *clustering*, which is weaker than convergence. In the sequel, clustering will be applied only to sequences in metric spaces [in [Section 4.2](#) of [Chapter 4](#) (Compactness)].

Just as the notion of being eventually in a set ([Definition 3.98](#)) helped in defining net convergence, so the notion of being “frequently” in a set will help in defining net clustering. And to express the notion of “frequently,” it is convenient to introduce the auxiliary notion of cofinality.

**3.111 Definition.** Let  $D$  be a set directed by  $\leq$ . A subset  $C$  of  $D$  is said to be **cofinal** in the directed set  $\langle D, \leq \rangle$  when for each  $d \in D$  there exists some  $c \in C$  for which  $d \leq c$ .

For example, when  $\mathbb{N}$  is directed by its usual ordering, the set of all even nonnegative integers is cofinal in  $\mathbb{N}$ , as is the set of all odd positive integers. Likewise, when  $\mathbb{R}$  is directed by its usual ordering, the set  $\mathbb{N}$  is cofinal in  $\mathbb{R}$ , as is the set  $\mathbb{N}$ ; however, the interval  $]0, 10^{100}[$  is *not* cofinal in  $\mathbb{R}$ .

**3.112 Definition.** Let  $(x_i)_{i \in I}$  be a net in a set  $X$  and let  $A \subset X$ . We say that  $(x_i)_{i \in I}$  is **frequently in**  $A$  when the set  $\{j \in I : x_j \in A\}$  is cofinal in the index set  $I$ . In other words,  $(x_i)_{i \in I}$  is frequently in  $A$  exactly when:

For every  $i \in I$ , there exists some  $j \in I$  such that  $j \geq i$  and  $x_j \in A$ .

Evidently, if a net is eventually in  $A$  ([Definition 3.98](#)), then it is frequently in  $A$ .

**3.113 Example.** Recall [Example 3.99](#), where we saw that, for each  $\varepsilon > 0$ , the sequence  $\langle 1/n \rangle_{n \in \mathbb{N}^*}$  is eventually in the open interval  $]0, \varepsilon[$ .

Consider now, instead, the sequence  $\langle (-1)^{n+1}/n \rangle_{n \in \mathbb{N}^*}$  in  $X = \mathbb{R}$ , which is also a net indexed by  $\mathbb{N}^*$  with its usual ordering. For each  $\varepsilon > 0$ , this sequence is frequently—but *not* eventually—in the open interval  $]0, \varepsilon[$ . This is an easy consequence of [Example 3.99](#).  $\diamond$

**3.114 Definition.** A net  $\xi$  in a topological space  $X$  is said to **cluster at** a point  $x$  (**in**  $X$ ) when it is frequently in each neighborhood of  $x$ . We say that  $\xi$  **clusters (in**  $X$ ) when it clusters at some point of  $X$ .

Roughly speaking, a net  $\xi = (x_i)_{i \in I}$  clusters at a point  $x$  when for each neighborhood  $V$  of  $x$  there are arbitrarily large indices  $j \in I$  for which  $x_j \in V$ . In the case of a sequence in a metric space, clustering in this sense is the same as clustering in the sense of [Exercise 1.100](#).

Although in a Hausdorff space a net cannot converge to more than one point ([Theorem 3.106](#)), even in a Hausdorff space a net can cluster at more than one point. For example, the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  given by

$$x_n = \begin{cases} 0 & \text{if } n \text{ is even,} \\ 1 - \frac{1}{n} & \text{if } n \text{ is odd} \end{cases}$$

eq:non-conv-seq-clustering-at-2-pts} (\*)

clustering net!and convergent net  
convergent net!and clustering net

clusters at both 0 and 1 in  $\mathbb{R}$ .

**If a net converges to a point, then it clusters at that point.** The converse, however, is false: the preceding sequence (\*) is a counterexample. Nevertheless, if a net  $(x_i)_{i \in I}$  clusters at a point  $x$ , there will be *some* net converging to  $x$  whose values are among the values of  $(x_i)_{i \in I}$ , as we shall demonstrate in the subsection “Subnets” that follows. Here we handle just the special case of sequences.

def:subseq **3.115 Definition.** A **subsequence** of a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  is a sequence of the form  $\langle x_{n_j} \rangle_{j \in \mathbb{N}}$  where  $\langle n_j \rangle_{j \in \mathbb{N}}$  is a *strictly* increasing sequence of natural numbers, that is,

$$n_0 < n_1 < n_2 < \cdots < n_j < n_{j+1} < \cdots .$$

For example, the sequence  $\langle \frac{1}{2j+1} \rangle_{j \in \mathbb{N}}$  is a subsequence of the sequence  $\langle \frac{1}{n+1} \rangle_{n \in \mathbb{N}}$ , but the sequence

$$\left\langle \frac{1}{j + (-1)^j + 1} \right\rangle_{j \in \mathbb{N}} = \left\langle \frac{1}{2}, 1, \frac{1}{4}, \frac{1}{3}, \frac{1}{6}, \frac{1}{5}, \dots \right\rangle$$

is *not* a subsequence of  $\langle \frac{1}{2j+1} \rangle_{j \in \mathbb{N}}$ .

For later use we record here an observation about subsequences.

lem:subseq-indices-cofinal **3.116 Lemma.** If  $\langle x_{n_j} \rangle_{j \in \mathbb{N}}$  is a subsequence of a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$ , then the set  $\{n_j : j \in \mathbb{N}\}$  is *cofinal* in  $\mathbb{N}$ .

prop:seq-cluster-if-subseq-converge **3.117 Proposition.** Let  $x$  be a point and  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence in a topological space  $X$ . If  $\langle x_n \rangle_{n \in \mathbb{N}}$  has a subsequence converging to  $x$ , then  $\langle x_n \rangle_{n \in \mathbb{N}}$  clusters at  $x$ .

**Proof.** Assume that  $\langle x_{n_j} \rangle_{j \in \mathbb{N}}$  is a subsequence of  $\langle x_n \rangle_{n \in \mathbb{N}}$  that converges to  $x$ . We show that  $\langle x_n \rangle_{n \in \mathbb{N}}$  clusters at  $x$ . Let  $V$  be an arbitrary neighborhood of  $x$  and let  $i$  be an arbitrary nonnegative integer. By assumption there is some  $k \in \mathbb{N}$  such that

$$j \geq k \iff x_{n_j} \in V.$$

Since  $\{j \in \mathbb{N} : j \geq k\}$  is infinite whereas  $\{j \in \mathbb{N} : n_j < i\}$  is only finite, there is some  $j \in \mathbb{N}$  with  $j \geq k$  and  $n_j \geq i$ . Then  $x_{n_j} \in V$ .  $\square$

ter-iff-subseq-converge-1st-countable **3.118 Theorem.** Let  $x$  be a point and  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence in a first-countable space  $X$ . Then  $\langle x_n \rangle_{n \in \mathbb{N}}$  clusters at  $x$  in  $X$  if and only if it has some subsequence converging to  $x$  in  $X$ .

**Proof.** Assume that  $\langle x_n \rangle_{n \in \mathbb{N}}$  clusters at  $x$ . We shall construct a subsequence that converges to  $x$ . There is a sequence  $\langle V_n \rangle_{n \in \mathbb{N}}$  of open sets in  $X$  for which  $\{V_n : n \in \mathbb{N}\}$  is a local base at  $x$ . By replacing each  $V_n$  with  $V_0 \cap V_1 \cap \cdots \cap V_n$  if necessary, we may assume without loss of generality that

$$V_0 \supset V_1 \supset \cdots \supset V_n \supset V_{n+1} \supset \cdots .$$

Construct a subsequence  $\langle x_{n_j} \rangle_{j \in \mathbb{N}}$  of  $\langle x_n \rangle_{n \in \mathbb{N}}$  recursively as follows. Choose  $n_0 \in \mathbb{N}$  with

$$x_{n_0} \in V_0.$$

Next, choose  $n_1 \in \mathbb{N}$  with

$$n_1 > n_0, \quad x_{n_1} \in V_1.$$

In general, once  $n_0, n_1, \dots, n_{j-1}$  have been obtained, choose  $n_j \in \mathbb{N}$  with

$$n_j > n_{j-1}, \quad x_{n_j} \in V_j.$$

All these choices are possible because  $\langle x_n \rangle_{n \in \mathbb{N}}$  clusters at  $x$ . Then the sequence  $\langle x_{n_j} \rangle_{j \in \mathbb{N}}$  converges to  $x$ . In fact, if  $V$  is an arbitrary neighborhood of  $x$ , there is some  $k \in \mathbb{N}$  with  $V_k \subset V$ , and then

$$j \geq k \implies x_{n_j} \in V_j \subset V_k \subset V.$$

The converse, which does not require first-countability, is [Proposition 3.117](#).  $\square$

The preceding theorem will be used in [Chapter 4](#) (Compactness). On the other hand, *nothing in the remainder of this section will be needed for the sequel!*

### Subnets

subsec:subnets

Our objective now is to generalize subsequences to the realm of nets—to define a notion of “subnet” so as to be able to generalize [Theorem 3.118](#), in other words, to be able to prove that a net clusters at a point if and only if it has some subnet converging to the point.

If  $\langle x_n \rangle_{n \in \mathbb{N}}$  is a sequence in a set  $X$ , then a subsequence of  $\langle x_n \rangle_{n \in \mathbb{N}}$  can be written in the form  $\langle x_{\sigma(j)} \rangle_{j \in \mathbb{N}}$  by letting  $\sigma(j) = n_j$  for each  $j \in \mathbb{N}$ . Thus a subsequence of  $\langle x_n \rangle_{n \in \mathbb{N}}$  is just a net  $\langle y_j \rangle_{j \in \mathbb{N}}$  that has the *same* index set  $\mathbb{N}$  as the original sequence and that takes the form

$$y_j = x_{\sigma(j)} \quad (j \in \mathbb{N})$$

where

$$\sigma: \mathbb{N} \rightarrow \mathbb{N}$$

is a strictly increasing map.

By dropping the requirement that the index sets be the same and by modifying the requirement that  $\sigma$  be strictly increasing, we arrive at a suitable generalization of subsequences. (Alternative definitions are possible for which the corresponding generalizations of [Theorem 3.118](#) are still true. See [Exercise 257](#).)

**3.119 Definition.** Let  $(x_i)_{i \in I}$  be a net in a set  $X$ . A **subnet** of  $(x_i)_{i \in I}$  is a net  $\langle y_j \rangle_{j \in J}$  such that

$$y_j = x_{\sigma(j)} \quad (j \in J)$$

for some map

$$\sigma: J \rightarrow I$$

having the two properties:

(SN1) *The map  $\sigma$  is order-preserving, that is, for each  $j, j' \in J$ , if  $j \leq j'$ , then  $\sigma(j) \leq \sigma(j')$ .*

(SN2) *The range  $\sigma(J)$  of  $\sigma$  is cofinal in  $I$ , that is, for each  $i \in I$ , there exists some  $j \in J$  such that  $\sigma(j) \geq i$ .*

subnet  
subsequence!and subnet  
subnet!and subsequence  
subsequence

subsequence!and subnet  
subnet!and subsequence



Figure 3.36: Relationship between a net with index set  $I$  in a set  $X$  and a subnet with index set  $J$ .

fig:subnet-of-net

The relationship between a net  $(x_i)_{i \in I}$  in a set  $X$  and a subnet  $\langle y_j \rangle_{j \in J}$  of that net is indicated in Figure 3.36. Although the directed set that indexes the subnet need not be the same as the one that indexes the net, the values of the subnet are among the values of the net. Moreover, property (SN2) says that there are arbitrarily large indices  $i \in I$  for which  $x_i$  is one of the values of the subnet.

A strictly increasing map  $\sigma: \mathbb{N} \rightarrow \mathbb{N}$  certainly has property (SN1). According to Lemma 3.116, it has property (SN2) as well. Hence **a subsequence**  $\langle x_{\sigma(j)} \rangle_{j \in \mathbb{N}}$  **of a sequence**  $\langle x_i \rangle_{i \in \mathbb{N}}$  **is a subnet of**  $\langle x_i \rangle_{i \in \mathbb{N}}$ .

We can now establish the desired generalization of Theorem 3.118.

**3.120 Theorem.** *Let  $x$  be a point and  $(x_i)_{i \in I}$  be a net in a topological space  $X$ . Then  $(x_i)_{i \in I}$  clusters at  $x$  in  $X$  if and only if it has some subnet of  $(x_i)_{i \in I}$  converging to  $x$  in  $X$ .*

**Proof.** Assume first that  $(x_i)_{i \in I}$  has some subnet  $\langle y_j \rangle_{j \in J}$  converging to  $x$ . Just suppose that  $(x_i)_{i \in I}$  does *not* cluster at  $x$ . Then  $x$  has some neighborhood  $V$  such that the net  $(x_i)_{i \in I}$  is *not* frequently in  $V$ , so that  $(x_i)_{i \in I}$  is eventually in  $X \setminus V$ . It follows that the subnet  $\langle y_j \rangle_{j \in J}$  is also eventually in  $X \setminus V$ , and so  $\langle y_j \rangle_{j \in J}$  is *not* eventually in  $V$ . This contradicts the assumption that  $\langle y_j \rangle_{j \in J}$  converges to  $x$ .

Conversely, assume that  $(x_i)_{i \in I}$  clusters at  $x$ . Define

$$J = \{ \langle i, V \rangle : i \in I, V \text{ is a neighborhood of } x, \text{ and } x_i \in V \}.$$

Direct the set by the relation  $\leq$  given by

$$\langle i, V \rangle \leq \langle i', V' \rangle \iff i \leq i' \text{ and } V \supset V'.$$

Form the map

$$\begin{aligned} \sigma: J &\rightarrow I, \\ (i, V) &\mapsto i \end{aligned}$$

This map  $\sigma$  clearly satisfies (SN1); it also satisfies (SN2) because if  $i \in I$  is arbitrary, then  $(i, X) \in J$  with  $\sigma(i, X) = i \geq i$ . Thus  $\langle x_i \rangle_{(i, V) \in J}$  is a subnet of  $(x_i)_{i \in I}$ . Finally, as is readily verified, this subnet converges to  $x$ .  $\square$

Another topological connection between nets and subnets is as follows.

thm:net-cluster-iff-subnet-converge

prop:subnet-converge-if-net-converge

**3.121 Proposition.** *If a net  $\xi$  converges to a point  $x$  in a topological space  $X$ , then every subnet of  $\xi$  converges to  $x$ .*

subnet  
limit of a map

**Proof.** Assume that the net  $\xi = (x_i)_{i \in I}$  converges to  $x$  in  $X$ . Let  $\eta = (y_j)_{j \in J}$  be a subnet of  $\xi$ . There is an increasing map  $\sigma: J \rightarrow I$  for which  $J$  is cofinal in  $I$ . Let  $V$  be an arbitrary neighborhood of  $x$  in  $X$ . We shall show that  $\eta$  is eventually in  $V$ . Since  $\xi$  is eventually in  $V$ , there is some  $i \in I$  with  $x_k \in V$  for all  $k \geq i$ . By cofinality of  $J$  in  $I$ , there is some  $j \in J$  with  $\sigma(j) \geq i$ . If  $p \in J$  with  $p \geq j$ , then  $\sigma(p) \geq \sigma(j) \geq i$ , and so  $y_p = x_{\sigma(p)} \in V$ .  $\square$

### Limit of a map

The final topic involving convergence we look at is the notion of limit of a function. Our aim here is not to prove any profound theorems, but merely to show that the diverse kinds of limits met in calculus—including one-sided limits, limits at infinity, and infinite limits—are subsumed by one general notion of limit.

In calculus we often compute the limit of a function  $f$  at a point not belonging to the domain of the function. For example,

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$$

even though 0 does not belong to the domain  $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$  of  $(\sin x)/x$ . Hence we will want to define the limit of a function  $f: D \rightarrow Y$  at a point  $c$  of a space  $X$  containing the domain  $D$  of  $f$  even when  $c \notin D$ .

The idea of the definition below is to call a point  $p \in Y$  a limit of  $f$  at  $c$  when  $f(t)$  can be made arbitrarily close to  $p$  by taking  $x \in D$  sufficiently close to, but different from,  $c$ . Then it is reasonable to suppose there *are* points of  $D$  other than  $c$  itself that are arbitrarily close to  $c$ —in other words, to suppose that  $c \in \text{cls}(D \setminus \{c\})$ .

def:lim-of-map

**3.122 Definition.** Let

$$f: D \rightarrow Y$$

be a map from a subset  $D$  of a topological space  $X$  into a topological space  $Y$ , and let  $c$  be a point of  $X$  with

$$c \in \text{cls}(D \setminus \{c\}).$$

A point  $p$  of  $Y$  is said to be **a limit of  $f$  at  $c$**  when, for each net  $(x_i)_{i \in I}$  in  $D \setminus \{c\}$ ,

$$x_i \rightarrow c \implies f(x_i) \rightarrow p.$$

If there is exactly one such point  $p$ , then it is called **the limit of  $f$  at  $c$**  and is denoted by

$$\lim_c f$$

or, more suggestively, by

$$\lim_{x \rightarrow c} f(x)$$

(or similarly, but with some other variable used in place of  $x$ ).

A map may have more than one limit at a given point. This will be the case, for example, if the codomain  $Y$  is indiscrete: in this case, *every* point of  $Y$  will be a limit of  $f$  at  $c$ . Imposing a separation condition on  $Y$  will guarantee uniqueness of a limit.

lem:unique-fn-lim-in-T2-codom

**3.123 Lemma.** Suppose that the codomain  $Y$  of a map  $f: D \rightarrow Y$  is a Hausdorff space. Let  $c \in X$  with  $c \in \text{cls}(D \setminus \{c\})$ . Then at most one point of  $Y$  can be a limit of  $f$  at  $c$ .

**Proof.** Since  $c \in \text{cls}(D \setminus \{c\})$ , there is at least one net  $(x_i)_{i \in I}$  in  $D \setminus \{c\}$  that converges to  $c$  in  $X$ . If points  $p$  and  $q$  of  $Y$  are both limits of  $f$  at  $c$ , then  $f(x_i) \rightarrow p$  and  $f(x_i) \rightarrow q$ , whence  $p = q$ .  $\square$

Again let  $c \in X$  with  $c \in \text{cls}(D \setminus \{c\})$ . Observe:

- When  $c \in D$ , the value  $f(c)$  of  $f$  at  $c$  has nothing to do with the limit (if any) of  $f$  at  $c$ . Indeed, if  $g: D \rightarrow Y$  is another map such that

$$x \in D \setminus \{c\} \implies f(x) = g(x),$$

then  $f$  has a limit at  $c$  precisely when  $g$  does, and in this case (if a limit exists and is unique)

$$\lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c} g(x).$$

- When  $c \notin D$ , then

$$c \in \text{cls}(D \setminus \{c\}) \iff c \in \text{cls } D.$$

The following criterion, whose proof involves only minor modifications to the proof of [Theorem 3.104](#), shows that [Definition 3.122](#) generalizes the notion of limit familiar from calculus.

prop:lim-of-fn-via-nbds

**3.124 Proposition.** Let  $f: D \rightarrow Y$  be a map from a subset  $D$  of a topological space  $X$  to a Hausdorff space  $Y$  and let  $c \in X$  with  $c \in \text{cls}(D \setminus \{c\})$ . A necessary and sufficient condition for

$$p = \lim_{x \rightarrow c} f(x)$$

to hold is that:

for each neighborhood  $V$  of  $p$  in  $Y$ , there exists some neighborhood  $U$  of  $c$  in  $X$  such that

$$c \neq x \in U \cap D \implies f(x) \in V.$$

When the domain  $D$  of  $f: D \rightarrow Y$  is the entire space  $X$ ,

$$c \in \text{cls}(D \setminus \{c\}) \iff c \text{ is not isolated in } X$$

Then just as in calculus the following corollary holds.

**3.125 Corollary.** Let  $f: X \rightarrow Y$  be a map from a topological space  $X$  to a Hausdorff space  $Y$ . Let  $c \in X$  be a non-isolated point of  $X$ . Then  $f$  is continuous at  $c$  if and only if

$$\lim_{x \rightarrow c} f(x) = f(c).$$

Next we exhibit some limits from calculus as just special instances of our general definition.



epsilon-delta-def-of-lim-in-calculus

**3.126 Examples.** (1) Let  $f$  be a real-valued function on a subset  $D$  of  $\mathbb{R}$ , let  $c \in \mathbb{R}$ , and suppose

$$]a, c[ \cup ]c, b[ \subset D$$

for some  $a$  and  $b$  with  $a < c < b$ . Then  $c \in \text{cls}(D \setminus \{c\})$ . According to [Proposition 3.124](#),

$$p = \lim_{x \rightarrow c} f(x)$$

if and only if

for each  $\varepsilon > 0$ , there exists some  $\delta > 0$  such that  
 $x \in D$  and  $0 < |x - c| < \delta \implies |f(x) - p| < \varepsilon$ .

This  $\varepsilon$ - $\delta$  condition is the usual calculus definition for  $p = \lim_{x \rightarrow c} f(x)$ .

delta-def-of-one-sided-lim-in-calculus

(2) Let  $g$  be a real-valued function on a subset  $E$  of  $\mathbb{R}$ , let  $c \in \mathbb{R}$ , and suppose

$$]a, c[ \subset E$$

for some  $a < c$ . Take

$$D = E \cap ]-\infty, c[, \quad f = g|_D.$$

Then  $c \in \text{cls}(D \setminus \{c\})$  because  $(a, c) \subset D$ . According to [Proposition 3.124](#),

$$p = \lim_{x \rightarrow c} f(x)$$

if and only if

for each  $\varepsilon > 0$ , there exists some  $\delta > 0$  such that  
 $x \in E$  and  $c - \delta < x < c \implies |f(x) - p| < \varepsilon$ .

Thus

$$\lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c^-} g(x),$$

the left-hand limit (that is, limit from below) of  $g$  at  $c$  in the usual calculus sense.

A similar discussion holds for the right-hand limit (that is, limit from above)  $\lim_{x \rightarrow c^+} g(x)$ .

ex:limits-at-infty

(3) Let  $f$  be a real-valued function on a subset  $D$  of  $\mathbb{R}$  that is not bounded above in  $\mathbb{R}$ . Take

$$X = \widehat{\mathbb{R}},$$

the extended real line ([Example 1.41](#)). According to [Example 2.40](#),

$$+\infty \in \text{cls } D = \text{cls}(D \setminus \{+\infty\}).$$

By [Lemma 1.43](#), a subset of  $X$  containing  $+\infty$  is a neighborhood of  $+\infty$  precisely when it contains some ray  $]u, +\infty[$ . Then by [Proposition 3.124](#), a real number  $y$  satisfies

$$p = \lim_{x \rightarrow +\infty} f(x)$$

if and only if

for each  $\varepsilon > 0$ , there exists some  $u \in \mathbb{R}$  such that  
 $x \in D$  and  $x > u \implies |f(x) - p| < \varepsilon$ .

limit-one-sided  
limit at infinity

filter  
Cartan, Henri

Hence  $\lim_{x \rightarrow +\infty} f(x)$  in the sense given by [Definition 3.122](#) is the limit of  $f$  at  $+\infty$  in the usual calculus sense.

A similar discussion holds for  $\lim_{x \rightarrow -\infty} f(x)$ .

ex:infinite-limits

- (4) Let  $g$  be a real-valued function on a subset  $D$  of  $\mathbb{R}$ , let  $c \in \mathbb{R}$ , and suppose as in (1) that

$$]a, c[ \cup ]c, b[ \subset D$$

for some  $a$  and  $b$  with  $a < c < b$ . Form the map

$$f: D \rightarrow \widehat{\mathbb{R}}$$

that has the same graph as the given  $g$  but whose codomain is obtained by enlarging the codomain  $\mathbb{R}$  of  $g$  to the extended real line  $\widehat{\mathbb{R}}$ . By [Proposition 3.124](#),

$$+\infty = \lim_{x \rightarrow c} f(x)$$

if and only if

for each  $u \in \mathbb{R}$ , there exists some  $\delta > 0$  such that  
 $x \in D$  and  $|x - c| < \delta \implies g(x) > u$ .

Thus  $+\infty$  is the limit of  $f$  at  $c$  precisely when  $+\infty$  is the limit of  $g$  at  $c$  in the usual calculus sense.

A similar discussion holds for  $\lim_{x \rightarrow c} g(x) = -\infty$ .  $\diamond$

Additional examples appear in the following exercises.

## Filters

subsec:filters

For describing convergence, nets are intuitively appealing generalizations of sequences. However, subnets may at first seem awkward as generalizations of subsequences. For that reason, among others, there is another theory of convergence, using *filters*, which was created in 1937 by H. Cartan. While perhaps less intuitively natural, this other theory avoids directed sets and families indexed by them. Moreover, the theory of filters is not only parallel to the theory of nets but is, in a sense that shall be made precise, equivalent to it.

Convergence to a point  $x$  in a space involves arbitrarily small neighborhoods of  $x$  and hence concerns the neighborhood system  $\mathcal{N}_x$  of  $x$ . By generalizing (N2) and dropping (N5) from the properties (N1)–(N5) [pages 237–237] for a neighborhood system we obtain the more general notion of a filter.

**3.127 Definition.** A **filter** on a set  $X$  is a collection  $\mathcal{F}$  of subsets of  $X$  having the properties:

property:F1 (F1) *The collection  $\mathcal{F}$  is nonempty.*

property:F2 (F2) *Each  $F \in \mathcal{F}$  is nonempty.*

property:F3 (F3) *If  $F_1, F_2 \in \mathcal{F}$ , then also  $F_1 \cap F_2 \in \mathcal{F}$ .*

property:F4 (F4) *If  $E$  is a subset of  $X$  for which  $F \subset E$  for some  $F \in \mathcal{F}$ , then also  $E \in \mathcal{F}$ .*

This definition implies, first, that there is no filter whatsoever on the empty set, and second, that the entire set  $X$  must belong to every filter on a set  $X$ .

Property (F3) implies that the intersection of finitely many members of  $\mathcal{F}$  is also a member of  $\mathcal{F}$ . In short, then, a filter on a set  $X$  is a nonempty collection of nonempty subsets of  $X$  that is closed under the operations of forming finite intersections and supersets in  $X$ .

In the presence of property (F4), property (F3) is equivalent to:

property:F3prime (F3') If  $F_1 \in \mathcal{F}$  and  $F_2 \in \mathcal{F}$ , there is some  $F \in \mathcal{F}$  such that  $F \subset F_1 \cap F_2$ .

ex:nbd-systems:filter 3.128 Examples. (1) For a point  $x$  of  $X$ , the neighborhood system  $\mathcal{N}_x$  at  $x$  is a filter on  $X$ , called the **neighborhood filter at  $x$  (in  $X$ )**. More generally, for a nonempty subset  $A$  of a topological space  $X$ , the collection of all neighborhoods of  $A$  in  $X$  is a filter on  $X$ .

ex:net-eventuality-filter (2) Every net gives rise to an associated filter. To see how, consider first a net  $\xi$  in a topological space  $X$ . That the net converges to a point  $x$  in  $X$  means, by definition, that  $\xi$  is eventually in every neighborhood of  $x$ .

Consider more generally now a net  $\xi = (x_i)_{i \in I}$  in a set  $X$ . Then the collection consisting of all those subsets of  $X$  that  $\xi$  is eventually in is a filter on  $X$ , which is called the **eventuality filter** of the net  $\xi$  and will be denoted in this book by  $\Phi(\xi)$ . In symbols,

$$\Phi(\xi) = \{F : F \subset X \text{ and } \xi \text{ is eventually in } F\}.$$

In fact, this collection is nonempty because  $X \in \Phi(\xi)$ , and each member of  $\Phi(\xi)$  is nonempty because it includes some values  $x_i$  of the net. The collection  $\Phi(\xi)$  is closed under formation of supersets in  $X$  because if  $\xi$  is eventually in a member  $F$  of  $\Phi(\xi)$  and if  $F \subset E \subset X$ , then also  $\xi$  is eventually in  $E$ .

We verify property (F3) for  $\Phi(\xi)$ . Let  $F_1, F_2 \in \Phi(\xi)$ . Since  $\xi = (x_i)_{i \in I}$  is eventually in  $F_1$  and eventually in  $F_2$ , there are indices  $j_1, j_2 \in I$  for which

$$\begin{aligned} \{x_i : i \in I, i \geq j_1\} &\subset F_1, \\ \{x_i : i \in I, i \geq j_2\} &\subset F_2. \end{aligned}$$

From property (D3) of the direction on the index set  $I$ , there is some  $k \in I$  with  $k \geq j_1$  and  $k \geq j_2$ . Then

$$\{x_i : i \in I, i \geq k\} \subset F_1 \cap F_2,$$

so that  $\xi$  is eventually in  $F_1 \cap F_2$ . This means that  $F_1 \cap F_2 \in \Phi(\xi)$ , too.

Another description of  $\Phi(\xi)$  is as follows. Call a subset of  $X$  of the form  $\{x_j : j \in I, j \geq k\}$  for some  $k \in I$  a **tail of  $\xi$** . (The term “**section**” is sometimes used as a synonym for “tail.”) To say that  $\xi$  is eventually in a subset  $F$  of  $X$  is to say that  $F$  contains some tail of  $\xi$ . Thus

$$\Phi(\xi) = \{F : F \subset X \text{ and } F \text{ contains some tail of } \xi\}.$$

ex:principal-filter (3) If  $x$  is a point in a set  $X$ , then the collection  $\{B : x \in B \subset X\}$  is a filter on  $X$ , called the **principal filter (on  $X$ ) generated by  $x$**  and denoted by  $\mathcal{U}_x$ .

ex:induced-filter (4) Let  $A$  be a nonempty subset of a set  $X$  and let  $\mathcal{F}$  be a filter on the whole set  $X$ . Then, unlike the situation with a topology, the collection  $\{A \cap F : F \in \mathcal{F}\}$  of subsets of  $A$  need not be a filter on  $A$ .

For example, let  $A$  be a proper subset of a set  $X$ , take an  $x \in X \setminus A$ , and form the principal filter  $\mathcal{U}_x$  generated by  $x$ . The collection  $\mathcal{F}' = \{A \cap F : F \in \mathcal{U}_x\}$  cannot be a filter on  $A$  because  $\{x\} \in \mathcal{U}_x$  and yet the member  $A \cap \{x\}$  of  $\mathcal{F}'$  is empty.

neighborhood system!as filter  
filter!and neighborhood system  
neighborhood filter  
eventuality filter  
filter!eventuality  
tail of a net  
principal filter  
filter!principal  
filter!and subset

**filter!induced on subset** If, however, each member of the filter  $\mathcal{F}$  intersects  $A$ . Then  $\{A \cap F : F \in \mathcal{F}\}$  is a filter on  $A$ , which is said to be **induced on  $A$  by  $\mathcal{F}$** .

**filter!extended**  
**filter!on a subset**  
**ex:extended-filter** (5) Let  $\mathcal{F}'$  be a filter on a subset  $A$  of a set  $X$ . Then the collection

$$\mathcal{F} = \{E : E \subset X \text{ and } F \subset E \text{ for some } F \in \mathcal{F}'\}$$

**finite-complement filter**  
**filter base**  
**Frechet filter@Fr\'echet filter** is a filter on the larger set  $X$ , which is called the **extended filter on  $X$  induced by  $\mathcal{F}'$** . Moreover, the original filter  $\mathcal{F}'$  is just the filter induced on  $A$  by  $\mathcal{F}$  in the sense of Example (4).

**filter base!and net**  
**tail of a net**  
**ex:cofinite filter** (6) Let  $X$  be an infinite set. Then the collection  $\mathcal{F}$  of all subsets of  $X$  whose complements are finite is a filter on  $X$ , called the **finite-complement filter on  $X$** . Note that this collection  $\mathcal{F}$  is *not* quite the finite-complement topology [Examples 2.3 (7)], because  $\emptyset \notin \mathcal{F}$ .  $\diamond$

**filter!eventuality**

The collection of all tails of a given net is not quite a filter, since it need not be closed under formation of supersets; it is what we shall call a *filter base*.

**def:filter-base** **3.129 Definition.** A **filter base** on a set  $X$  is a collection  $\mathcal{B}$  of subsets of  $X$  having the properties:

**property:FB1** (FB1) *The collection  $\mathcal{B}$  is nonempty.*

**property:FB2** (FB2) *Each  $B \in \mathcal{B}$  is nonempty.*

**property:FB3** (FB3) *If  $B_1, B_2 \in \mathcal{B}$ , then there is some  $B \in \mathcal{B}$  such that  $B \subset B_1 \cap B_2$ .*

Just as the collection of tails of a net give rise to a filter, so does any filter base.

**lem:filter-generated-by-fb** **3.130 Lemma.** Let  $\mathcal{B}$  be a filter base on a set  $X$ . Then the collection

$$\{F : F \subset X \text{ and there is some } B \in \mathcal{B} \text{ with } B \subset F\}$$

of supersets in  $X$  of members of  $\mathcal{B}$  is a filter on  $X$ .

**def:filter-generated-by-fb** **3.131 Definition.** If  $\mathcal{B}$  is a filter base on a set  $X$ , then the filter on  $X$  consisting of all supersets in  $X$  of members of  $\mathcal{B}$  is said to be **generated by  $\mathcal{B}$** , and  $\mathcal{B}$  is said to **generate  $\mathcal{F}$**  and to be a **base of  $\mathcal{F}$** .

Notice that a filter base  $\mathcal{B}$  on a subset  $A$  of a set  $X$  is a filter base on  $X$ ; and a filter base on a set  $X$  all of whose members are subsets of  $A$  is a filter base on  $A$ . However, if  $A$  is a proper subset of  $X$  and if  $\mathcal{B}$  is a filter base on  $A$ , then the filter on  $A$  generated by  $\mathcal{B}$  will be different from the filter on  $X$  generated by  $\mathcal{B}$ .

**exs:filter-base** **3.132 Examples.** (1) A filter on a set is itself a filter base on that set.

**ex:Frechet-filter** (2) The collection  $\{n, n+1, n+2, \dots\} : n \in \mathbb{N}\}$  is a filter base on  $\mathbb{N}$ . The filter it generates is called the **Fréchet filter on  $\mathbb{N}$** .

Similarly, the collection  $\{]a, +\infty[ : a \in \mathbb{R}\}$  of open rays is a filter base on  $\mathbb{R}$ . The filter it generates is called the **Fréchet filter on  $\mathbb{R}$** .

- ex:fb-generated-by-net (3) The discussion in [Examples 3.128 \(2\)](#) tells us that the collection of all tails of a net  $\xi$  in a set  $X$  is a filter base on  $X$ , and the filter it generates is the eventuality filter  $\Phi(\xi)$  of the net  $\xi$ . principal filter!and filter base  
filter!on a subset  
filter!on a subset
- ex:subset-as-fb (4) For a given nonempty subset  $A$  of a set  $X$ , the one-member collection  $\{A\}$  is a filter base on  $X$ .  
In particular, for an element  $x$  of  $X$ , the collection  $\{\{x\}\}$  consisting solely of the singleton  $\{x\}$  is a filter base on  $X$ , and the filter it generates is precisely the principal filter  $\mathcal{U}_x$  [[Examples 3.128 \(3\)](#)] generated by the point  $x$ .
- ex:local-base-as-fb (5) Given a point  $x$  in a topological space  $X$ , any local base at  $x$  ([Definition 2.55](#)) is a filter base on  $X$ . Then also the entire neighborhood system  $\mathcal{N}_x$  at  $x$  is a filter base on  $X$ .  
In particular, for  $x \in \mathbb{R}$ , the collection  $\{]x - \varepsilon, x + \varepsilon[ : \varepsilon > 0\}$  of all open intervals centered at  $x$  is a filter base on  $\mathbb{R}$ , as is its denumerable subcollection  $\{]x - 1/n, x + 1/n[ : n = 1, 2, 3, \dots\}$ .
- ex:filter-setwise-intersecting-subset (6) Let  $A$  be a nonempty subset of a set  $X$  and let  $\mathcal{F}$  be a filter on  $X$  each member of which intersects  $A$ . Then  $\mathcal{F}$  is a filter base on  $A$ , and the filter on  $A$  that it generates is the one induced on  $A$  by  $\mathcal{F}$  in the sense of [Examples 3.128 \(4\)](#).
- ex:fb-of-extended-filter (7) Let  $\mathcal{F}'$  be a filter on a subset  $A$  of a set  $X$ . Then  $\mathcal{F}'$  is a filter base on the larger set  $X$ , and it is a base of the extended filter

$$\mathcal{F} = \{E : E \subset X \text{ and } F \subset E \text{ for some } F \in \mathcal{F}'\}$$

induced on  $X$  by  $\mathcal{F}'$  [[Examples 3.128 \(5\)](#)]. In other words,  $\mathcal{F}$  is the filter on  $X$  generated by  $\mathcal{F}'$ .  $\diamond$

### Convergence of filters

subsec:filter-convergence

The intuitive idea that a filter on a topological space converge to a point in the space is that there are small members of the filter arbitrarily close to the point.

def:filter-conv **3.133 Definition.** Let  $x$  be a point in a topological space  $X$ . A filter base  $\mathcal{B}$  on  $X$  is said to **converge to  $x$  (in  $X$ )** when, for each neighborhood  $V$  of  $x$  in  $X$ , there exists some member  $B \in \mathcal{B}$  such that  $B \subset V$ , and then we write  $\mathcal{B} \rightarrow x$ .  
In particular, a filter  $\mathcal{F}$  on  $X$  is said to **converge to  $x$  (in  $X$ )** when, for each neighborhood  $V$  of  $x$  in  $X$ , there exists some member  $F \in \mathcal{F}$  such that  $F \subset V$ , and then we write  $\mathcal{F} \rightarrow x$ . To indicate that a filter converges to a point  $x$  in the space  $X$ , we write  $\mathcal{F} \xrightarrow{X} x$ . A filter or filter base on  $X$  is said to **converge**, and to be **convergent, (in  $X$ )** when it converges to some point of  $X$ .

There is an equivalent way to describe convergence of filters.

prop:filter-fb-conv-equiv **3.134 Proposition.** Let  $x$  be a point in a topological space  $X$ . Then a filter  $\mathcal{F}$  on  $X$  converges to  $x$  in  $X$  if and only if each neighborhood of  $x$  belongs to  $\mathcal{F}$ . In other words,  $\mathcal{F} \rightarrow x$  if and only if  $\mathcal{N}_x \subset \mathcal{F}$ .

The relation between convergence of a filter and convergence of a base of that filter is as follows.

neighborhood filter  
prop:conv-filter-vs-fb  
principal filter  
Fréchet filter@Fr\`e

**3.135 Proposition.** Let  $\mathcal{B}$  be a base of a filter  $\mathcal{F}$  on a topological space  $X$  and let  $x \in X$ . Then

$$\mathcal{B} \rightarrow x \iff \mathcal{F} \rightarrow x.$$

convergent net!and convergent filter

The proposition implies that if one base of a given filter converges to a particular point, then every base of that filter converges to the same point.

exs:filter-conv

**3.136 Examples.** (1) For a point  $x$  in a topological space, its neighborhood filter converges to  $x$ ; in symbols,  $\mathcal{N}_x \rightarrow x$ .

(2) For a point  $x$  in a topological space  $X$ , the principal filter  $\mathcal{U}_x$  on  $X$  generated by  $x$  [Examples 3.128 (3)] converges to  $x$ .

(3) The Fréchet filter  $\mathcal{F}$  does *not* converge in  $\mathbb{N}$  (with its discrete topology), because if it converged to a point  $n$  in  $\mathbb{N}$ , then  $\{n\}$  would be a member of  $\mathcal{F}$ .

(4) Let  $\xi = (x_i)_{i \in I}$  be a net in a topological space  $X$  and let  $x$  be a point in  $X$ . Further, let  $\Phi(\xi)$  be the eventuality filter of this net [Examples 3.128 (2)], so that  $\Phi(\xi)$  is generated by the collection of all tails of  $\xi$ . Recall that, by definition, the net  $\xi$  converges to  $x$  in  $X$  if and only if each neighborhood of  $x$  contains some tail of  $\xi$ . Since the tails of the net form a base of the filter  $\Phi(\xi)$ , this means that

$$\xi \rightarrow x \iff \Phi(\xi) \rightarrow x.$$

Thus *a net converges to a point if and only if its eventuality filter converges to that point.*  $\diamond$

A filter  $\mathcal{F}$  on a subset  $A$  of a topological space  $X$  is a filter base on that subset  $A$ . Further, a filter base on the subset  $A$  is also a filter base on the entire set  $X$ . Convergence of these two filter bases is connected as follows.

lem:conv-fb-subspace-vs-space

**3.137 Lemma.** Let  $\mathcal{B}$  be a filter base on a subspace  $A$  of a topological space  $X$  and let  $x \in X$ . Then  $\mathcal{B}$  converges to  $x$  in  $A$  if and only if  $\mathcal{B}$  converges to  $x$  in  $X$ .

In the notation introduced at the end of Definition 3.133, the lemma says that

$$\mathcal{B} \xrightarrow{A} x \iff \mathcal{B} \xrightarrow{X} x.$$

In particular, if  $\mathcal{F}$  is a *filter* on subspace  $A$  of  $X$ , then it is a filter base on the entire space  $X$ , and then

$$\mathcal{F} \xrightarrow{A} x \iff \mathcal{F} \xrightarrow{X} x.$$

cor:conv-filter-subspace-vs-space

**3.138 Corollary.** Let  $A$  be a subspace of a topological space  $X$  and let  $x$  be a point in the subspace  $A$ .

(1) If  $\mathcal{F}$  is a filter on  $X$  each member of which intersects  $A$  and if  $\mathcal{B}$  is the filter base it generates on  $A$  [Examples 3.132 (6)], then  $\mathcal{F}$  converges to  $x$  in the entire space  $X$  if and only if  $\mathcal{B}$  converges to  $x$  in  $A$ .

(2) If  $\mathcal{F}'$  is a filter on the subspace  $A$ , then  $\mathcal{F}'$  converges to  $x$  in  $A$  if and only if the extended filter on  $X$  of which is a base [Examples 3.132 (7)] converges to  $x$  in  $X$ .

We are now going to establish the analogs for filters of [Theorem 3.102](#), [Corollary 3.103](#), [Theorem 3.104](#), thereby demonstrating that filters, like nets, are adequate for determining the topologies and continuous maps of topological spaces. image of a filter base

thm:in-cls-via-filters **3.139 Theorem.** *Let  $x$  be a point and let  $A$  be a subset of a topological space  $X$ . Then the following conditions are equivalent:*

cond:in-cls (i)  $x \in \text{cls } A$ .

cond:fb-conv-to-pt (ii) *There is some filter base on  $A$  that converges to  $x$  in  $X$ .*

cond:extended-filter-conv-to-pt (iii) *There is some filter on  $A$  for which the extended filter induced by it on  $X$  converges to  $x$  in  $X$ .*

**Proof.** In view of [Lemma 3.137](#) and [Corollary 3.138](#), conditions (ii) and (iii) are equivalent.

We show that (i)  $\implies$  (ii). Assume condition (i). Since each neighborhood of  $x$  in  $X$  intersects  $A$ , then the collection  $\mathcal{B} = \{V \cap A : V \in \mathcal{N}_x\}$  is a filter base on  $A$ . Then  $\mathcal{B} \rightarrow x$  in  $X$  (and the filter on  $A$  generated by  $\mathcal{B}$  converges to  $x$  as well).

We show, conversely, that (iii)  $\implies$  (i). Assume condition (iii). There is some filter  $\mathcal{F}'$  on  $A$  for which the extended filter  $\mathcal{F}$  induced by  $\mathcal{F}'$  on  $X$  converges to  $x$  in  $X$ . Let  $V$  be an arbitrary neighborhood of  $x$  in  $X$ . There is a member  $F$  of  $\mathcal{F}$  such that  $F \subset V$ , and then there is a member  $F'$  of  $\mathcal{F}'$  with  $F' \subset F$ , and so also  $F' \subset V$ . Since  $F'$  is a nonempty subset of  $A$ , the neighborhood  $V$  intersects  $A$ .  $\square$

cor:open-closed-via-filters **3.140 Corollary.** *Let  $A$  be a subset of a topological space  $X$ . Then:*

- (1) *The set  $A$  is closed in  $X$  if and only if each point of  $X$  to which some filter base on  $A$  converges in  $X$  itself belongs to  $A$ .*
- (2) *The set  $A$  is open in  $X$  if and only if there is no point of  $A$  to which a filter base on  $X \setminus A$  converges in  $X$ .*

In order to obtain an analog of [Theorem 3.104](#), we need the notion of the image of a filter base.

lem:image-filter **3.141 Lemma.** *Let  $f: X \rightarrow Y$  be a map.*

lem-part:images-fb-is-fb (1) *If  $\mathcal{B}$  is a filter base on the domain  $X$  of  $f$ , then the collection  $\{f(B) : B \in \mathcal{B}\}$  of set images is a filter base on  $Y$ .*

lem-part:images-under-surj-filter-filter (2) *If  $\mathcal{F}$  is a filter on the domain  $X$  of  $f$ , then the filter base  $\{f(F) : F \in \mathcal{F}\}$  will be a filter on the codomain  $Y$  if and only if  $f$  is surjective.*

The lemma justifies the following definition.

def:image-of-fb **3.142 Definition.** *Let  $f: X \rightarrow Y$  be a map. If  $\mathcal{F}$  is a filter on  $X$ , then the filter on  $Y$  generated by the filter base  $\{f(F) : F \in \mathcal{F}\}$  on  $Y$  is called the **image of  $\mathcal{F}$  (under  $f$ )** and is denoted by  $f[\mathcal{F}]$ .*

thm:cont-via-filters

**3.143 Theorem.** Let  $X$  and  $Y$  be topological spaces. Then a necessary and sufficient condition for a map  $f: X \rightarrow Y$  to be continuous at a point  $x \in X$  is that for each filter  $\mathcal{F}$  on  $X$  converging to  $x$  in  $X$ , the image  $f[\mathcal{F}]$  of  $\mathcal{F}$  must converge to  $f(x)$  in  $Y$ , that is, in symbols:

$$\mathcal{F} \rightarrow x \implies f[\mathcal{F}] \rightarrow f(x).$$

**Proof.** Fix a point  $x \in X$ . Assume first that  $f$  is continuous at  $x$ . Let  $\mathcal{F}$  be an arbitrary filter on  $X$  converging to  $x$ . To show that  $f[\mathcal{F}]$  converges to  $f(x)$ , let  $V$  be an arbitrary neighborhood of  $f(x)$  in  $Y$ . By continuity,  $f^{-1}(V)$  is a neighborhood of  $x$  in  $X$ . Since  $\mathcal{F}$  converges to  $x$ , there is some member  $F \in \mathcal{F}$  with  $F \subset f^{-1}(V)$ . Then the member  $f(F)$  of  $f[\mathcal{F}]$  satisfies  $f(F) \subset f(U) \subset V$ .

Conversely, assume the condition holds. To show that  $f$  is continuous at  $x$ , let  $V$  be an arbitrary neighborhood of  $f(x)$  in  $Y$ . The neighborhood system  $\mathcal{N}_x$  at  $x$  in  $X$  is a filter on  $X$  that converges to  $x$ . From the condition, the image  $f[\mathcal{N}_x]$  converges to  $f(x)$  in  $Y$ . Then  $f(U) \subset V$  for some  $U \in \mathcal{N}_x$ .  $\square$

As with nets, so with filters: a filter may converge to more than one point.

**3.144 Example.** Let  $X$  be an infinite set provided with its finite complement topology [Examples 2.3 (7)] and let  $\mathcal{F}$  be the filter on  $X$  consisting of all subsets of  $X$  whose complements are finite [Examples 3.128 (2)]. Then  $\mathcal{F} \rightarrow x$  for every  $x \in X$ .

In fact, let  $x \in X$  and let  $V$  be an arbitrary neighborhood of  $x$ . There is some open neighborhood  $U$  of  $x$  with  $V \subset U$ . This means that  $X \setminus U$  is finite, and so  $U \in \mathcal{F}$ . Hence the superset  $V$  of  $U$  also belongs to  $\mathcal{F}$ .  $\diamond$

The filter in the preceding example behaved in the worst possible way: it converged to every point of the space. In Hausdorff spaces, however, a filter cannot converge to more than one point.

thm:filter-limits-unique-in-T2-space

**3.145 Theorem (uniqueness of filter limits in a Hausdorff space).** Let  $\mathcal{F}$  be a filter or a filter base on a Hausdorff space  $X$  and let  $x, y \in X$  such that

$$\mathcal{F} \rightarrow x \text{ and } \mathcal{F} \rightarrow y.$$

Then

$$x = y.$$

**Proof.** Just suppose  $x \neq y$ . There are disjoint neighborhoods  $U$  and  $V$  of  $x$  and  $y$ , respectively. By hypothesis  $U \in \mathcal{F}$  and  $V \in \mathcal{F}$ . Then also  $U \cap V \in \mathcal{F}$ . But this is impossible because members of a filter must be nonempty.  $\square$

Notice how “clean” the preceding proof is compared with the proof of the corresponding result (Theorem 3.106) for nets.

The theorem justifies the following definition.



def:filter-limit-in-T2

**3.146 Definition.** Let  $\mathcal{F}$  be a filter or filter base that converges in a *Hausdorff* space  $X$ . Then the unique  $x \in X$  for which

$$\mathcal{F} \rightarrow x$$

is called the **limit of  $\mathcal{F}$**  and is denoted by

$$\lim \mathcal{F}.$$

image of a filter base  
clustering net

We shall apply the notion of image of a filter to the case of projections from a product.

converge-in-product-iff-coordinatewise

**3.147 Theorem (convergence of a filter in a product).** Let  $\mathcal{F}$  be a filter in the product  $X$  of a family  $\langle X_i \rangle_{i \in I}$  of topological spaces and let

$$y = \langle y_i \rangle_{i \in I}$$

be a point of  $X$ . Then

$$\mathcal{F} \rightarrow y \text{ in } X$$

if and only if

$$p_i[\mathcal{F}] \rightarrow y_i \text{ in } X_i \text{ for each } i \in I,$$

where for each  $i$ , the map  $p_i: X \rightarrow X_i$  is the  $i$ th projection.

**Proof.** For the “only if” implication, apply [Theorem 3.143](#).

Conversely, assume that  $p_i[\mathcal{F}] \rightarrow y_i$  in  $X_i$  for each  $i \in I$ . Let  $V$  be an arbitrary neighborhood of  $y$  in  $X$ . Then  $V$  contains some basic open neighborhood

$$U = \bigcap_{i \in I} U_i$$

of  $y$ . Let

$$J = \{j \in I : U_j \neq X_j\}.$$

If  $J = \emptyset$ , then  $V = U = X$  and surely there is some member  $F \in \mathcal{F}$  with  $F \subset V$ ; in fact, every member of the nonempty collection  $\mathcal{F}$  has this property. Suppose now that  $J \neq \emptyset$ . For each  $j \in J$  the set  $U_j$  is a neighborhood of  $y_j$  in  $X_j$ , and so by assumption there is some member  $F_j \in \mathcal{F}$  for which  $p_j(F_j) \subset U_j$ . Let  $F = \bigcap_{j \in J} F_j$ , so that  $F \in \mathcal{F}$  also. Then  $F \subset U$  whence  $F \subset V$ , as desired.  $\square$

Let  $\xi = (x_i)_{i \in I}$  be a net in a topological space  $X$  and let  $x$  be a point in  $X$ . Form the eventuality filter  $\mathcal{F}$  of the net  $\xi$  [[Examples 3.128 \(2\)](#)]. According to ([Definition 3.114](#)), that  $\xi$  clusters at  $x$  amounts to saying that each tail of  $\xi$  intersects each neighborhood of  $x$ , in other words, that each member of  $\mathcal{F}$  intersects each neighborhood of  $x$ . This suggests the following general definition of clustering for a filter.

def:filter-cluster

**3.148 Definition.** A filter base  $\mathcal{B}$  on a topological space  $X$  is said to **cluster at a point  $x$  (in  $X$ )** when each member of  $\mathcal{B}$  intersects each neighborhood of  $x$ . In particular, a filter  $\mathcal{F}$  on  $X$  is said to **cluster at a point  $x$  (in  $X$ )** when each member of  $\mathcal{F}$  intersects each neighborhood of  $x$ . We say that a filter base  $\mathcal{B}$  or filter  $\mathcal{F}$  **clusters (in  $X$ )** when it clusters at some point of  $X$ .

clustering filter  
clustering filter!and convergent filter  
finer filter!strictly  
coarser filter!strictly

That a filter base  $\mathcal{B}$  clusters at  $x$  is equivalent to saying:

$$x \in \bigcap_{B \in \mathcal{B}} \text{cls } B.$$

principal ultrafilter

filter!generated by a filter base  
follows.

The relation between clustering of a filter and clustering of a base of that filter is as follows.

prop:cluster-filter-vs-fb

**3.149 Proposition.** Let  $\mathcal{B}$  be a filter base of a filter  $\mathcal{F}$  on a topological space  $X$  and let  $x \in X$ . Then  $\mathcal{B}$  clusters at  $x$  if and only if  $\mathcal{F}$  clusters at  $x$ .

The proposition implies that if one base of a given filter clusters at a particular point, then every base of that filter clusters at the same point.

Although in a Hausdorff space a filter cannot converge to more than one point ([Theorem 3.145](#)), even in a Hausdorff space a filter can cluster at more than one point. For example, the filter

{eq:filter-cluster-2-pts-in-R} (\*)

$$\mathcal{F} = \{F : F \subset \mathbb{R} \text{ and } 0 \in F \text{ and } 1 \in F\}$$

on  $\mathbb{R}$  clusters at both 0 and 1 in  $\mathbb{R}$ .

*If a filter converges to a point, then it clusters at that point.* The converse, however, is false: the preceding filter (\*) is a counterexample. Nonetheless, as we shall see in the following [subsection “Comparison of filters”](#), if a filter  $\mathcal{F}$  clusters at a point  $x$ , there will be *some* filter that converges to  $x$  and is related to  $\mathcal{F}$ .

### Comparison of filters

subsec:cf-filters

As collections of subsets, filters on a given set may be compared with respect to inclusion.

def:compare-filters

**3.150 Definition.** Let  $\mathcal{F}_1$  and  $\mathcal{F}_2$  be filters on a set  $X$ . We say that  $\mathcal{F}_2$  is **finer than**  $\mathcal{F}_1$ , and that  $\mathcal{F}_1$  is **coarser than**  $\mathcal{F}_2$  when  $\mathcal{F}_1 \subset \mathcal{F}_2$ , that is, when each member of  $\mathcal{F}_1$  is also a member of  $\mathcal{F}_2$ .

A filter that is finer than a filter  $\mathcal{F}$  is called a **refinement of**  $\mathcal{F}$ .

The relation “is finer than” is a *partial ordering* of the class of all filters on a given set. Accordingly, we have a corresponding “strict” ordering: for filters  $\mathcal{F}_1$  and  $\mathcal{F}_2$  on a set  $X$ , we say that  $\mathcal{F}_2$  is **strictly finer than**  $\mathcal{F}_1$ , and that  $\mathcal{F}_1$  is **strictly coarser than**  $\mathcal{F}_2$ , when  $\mathcal{F}_1 \subset \mathcal{F}_2$  but  $\mathcal{F}_1 \neq \mathcal{F}_2$ .

The term “finer” is being used because a filter  $\mathcal{F}_2$  that is *strictly* finer than a filter  $\mathcal{F}_1$  will include more, presumably “smaller,” sets.

ex:conv-filter-finer-than-nbd-filter

**3.151 Examples.** (1) If a filter  $\mathcal{F}$  on a topological space  $X$  converges to a point  $x$  in  $X$ , then by [Proposition 3.134](#),  $\mathcal{F}$  is finer than the neighborhood filter  $\mathcal{N}_x$  at  $x$  in  $X$ .

ex:nbd-filter-finer-than-principal-uf

(2) For a point  $x$  in a topological space  $X$ , the principal ultrafilter  $\mathcal{U}_x$  generated by  $x$  is finer than the neighborhood filter  $\mathcal{N}_x$  at  $x$ .

(3) Given a filter base  $\mathcal{B}$ , on a set  $X$ , the filter generated by  $\mathcal{B}$  ([Definition 3.131](#)) is the coarsest filter that contains  $\mathcal{B}$ .

ex:subnet-produces-finer-filter

(4) Let  $\eta = \langle y_j \rangle_{j \in J}$  be a subnet of a net  $\xi = (x_i)_{i \in I}$  in a set  $X$ . Let  $\mathcal{F}_\eta$  and  $\Phi(\xi)$  be the eventuality filters of  $\eta$  and  $\xi$ , respectively [[Examples 3.128 \(2\)](#)]. Then  $\mathcal{F}_\eta$  is finer than  $\Phi(\xi)$  because each tail of  $\xi$  contains some tail of  $\eta$ .  $\diamond$

Next, we show as promised that if a filter clusters at a point, then there is a filter that converges to the same point.

**3.152 Proposition.** Let  $\mathcal{F}$  be a filter on a topological space  $X$  that clusters at a point  $x$  in  $X$ . Then:

(1) The collection

$$\mathcal{B} = \{F \cap V : F \in \mathcal{F} \text{ and } V \in \mathcal{N}_x\}$$

is a filter base on  $X$ .

(2) The filter  $\mathcal{F}'$  generated by  $\mathcal{B}$  converges to  $x$  in  $X$ .

(3) The filter  $\mathcal{F}'$  is the coarsest filter finer than  $\mathcal{F}$  and converging to  $x$  in  $X$ .

**Proof.** (1) Since  $\mathcal{F}$  clusters at  $x$ , the members of  $\mathcal{B}$  are all nonempty. The other conditions for a filter base are readily checked.

(2) Let  $V$  be an arbitrary neighborhood of  $x$ . Pick some  $F \in \mathcal{F}$ . Then  $V \cap F$  is a member of  $\mathcal{B}$ , hence of  $\mathcal{F}'$ , with  $V \cap F \subset V$ .

(3) Let  $\mathcal{G}$  be another filter that is finer than  $\mathcal{F}$  and converges to  $x$ . We must show that  $\mathcal{F}' \subset \mathcal{G}$ . It suffices to show that  $\mathcal{B} \subset \mathcal{G}$ . Let  $B \in \mathcal{B}$ , so that  $B = F \cap V$  for some  $F \in \mathcal{F}$  and some neighborhood  $V$  of  $x$ . Since  $\mathcal{G} \rightarrow x$ , there is some  $G \in \mathcal{G}$  with  $G \subset V$ . Now both  $F$  and  $G$  belong to  $\mathcal{G}$ , so that  $F \cap G \in \mathcal{G}$ , too. But  $F \cap V$  is a superset of  $F \cap G$ , so that  $B = F \cap V \in \mathcal{G}$ , too.  $\square$

**3.153 Theorem.** Let  $x$  be a point in a topological space  $X$  and let  $\mathcal{F}$  be a filter on  $X$ . Then  $\mathcal{F}$  clusters at  $x$  in  $X$  if and only if some filter finer than  $\mathcal{F}$  converges to  $x$  in  $X$ .

**Proof.** The “only if” part follows from [Proposition 3.152](#).

Conversely, assume there is some filter  $\mathcal{F}'$  finer than  $\mathcal{F}$  that converges to  $x$ . To see that  $\mathcal{F}$  clusters at  $x$ , let  $V$  be an arbitrary neighborhood of  $x$  and let  $F$  be an arbitrary member of  $\mathcal{F}$ . There is some member  $F' \in \mathcal{F}'$  with  $F' \subset V$ . Since both  $F$  and  $F'$  are members of the filter  $\mathcal{F}'$ , they intersect. Since  $F \cap F' \subset F \cap V$ , it follows that  $F \cap V \neq \emptyset$ .  $\square$

Of special interest among filters on a given set are *ultrafilters*—filters than which there are no finer filters (for example, the principal filter generated by a point). Ultrafilters are studied in [Section 6.1](#), where they are used to provide a short, elegant proof of the Tychonoff Theorem about the product of spaces that are “compact.”

### Filters vs. nets

From [Examples 3.128 \(2\)](#) we already know that each net  $\xi = (x_i)_{i \in I}$  in a set  $X$  gives rise to a filter on  $X$ , namely, its *eventuality filter*

$$\begin{aligned} \Phi(\xi) &= \{A : A \subset X \text{ and } \xi \text{ is eventually in } A\} \\ &= \{A : A \subset X \text{ and there is some } i \in I \text{ with } x_j \in A \text{ for all } j \geq i\}. \end{aligned}$$

In the reverse direction, each filter on a set gives rise to a net in that set. To see how, first we construct the index set for such a net, in a manner akin to [Examples 3.96 \(5\)](#).

convergent filter!and clustering filter  
clustering filter!and convergent filter  
convergent filter!and clustering filter  
clustering filter!and convergent filter  
ultrafilter  
eventuality filter  
filter!eventuality

convergence

**3.154 Lemma.** Let  $\mathcal{F}$  be a filter on a set  $X$ . Let

$$I = \{\langle F, x \rangle : F \in \mathcal{F} \text{ and } x \in F\}$$

and let  $\leq$  be the relation on  $I$  given by

$$\langle F, x \rangle \leq \langle E, y \rangle \iff F \supset E \quad (\langle F, x \rangle, \langle E, y \rangle \in I).$$

Then  $\leq$  directs  $I$ .

**Proof.** The relation  $\leq$  is certainly reflexive and transitive. To see that it has the third property, (D3), required of a direction, let  $\langle F, x \rangle, \langle E, y \rangle \in I$ . Then  $\emptyset \neq F \cap E \in \mathcal{F}$ , and so there is some  $z \in F \cap E$ . Consequently,  $\langle F \cap E, z \rangle \in I$  with  $\langle F, x \rangle \leq \langle F \cap E, z \rangle$  and  $\langle E, y \rangle \leq \langle F \cap E, z \rangle$ .  $\square$

def:net-from-filter

**3.155 Definition.** Let  $\mathcal{F}$  be a filter on a set  $X$ . Define the set  $I$  and the relation  $\leq$  in  $I$  as in the preceding lemma. Then the family indexed by  $I$  that assigns to each element  $\langle F, x \rangle \in I$  its second coordinate  $x$  is called the **net associated with** the filter  $\mathcal{F}$  and will be denoted in this book by  $\Psi(\mathcal{F})$ . In symbols,

$$\Psi(\mathcal{F}) = \langle x \rangle_{\langle F, x \rangle \in I}.$$

Convergence of filters is related to convergence of nets as follows.

prop:net-vs-filter-conv

**3.156 Proposition.** Let  $x$  be a point in a topological space  $X$ . Then:

prop-part:conv-net-iff-eventuality-filter

(1) A net  $\xi$  in  $X$  converges to  $x$  in  $X$  if and only if its eventuality filter  $\Phi(\xi)$  converges to  $x$  in  $X$ .

prop-part:conv-filter-iff-associated-net

(2) A filter  $\mathcal{F}$  on  $X$  converges to  $x$  in  $X$  if and only if its associated net  $\Psi(\mathcal{F})$  converges to  $x$  in  $X$ .

Fineness of filters is related to subnets of nets as follows.

prop:subnet-then-eventuality-filter-finer

**3.157 Proposition.** If  $\eta$  is a subnet of a net  $\xi$  in a set  $X$ , then its eventuality filter  $\Phi(\eta)$  is finer than the eventuality filter  $\Phi(\xi)$ .

**Proof.** Let  $\eta = \langle y_j \rangle_{j \in J}$  be a subnet of a net  $\xi = \langle x_i \rangle_{i \in I}$  in  $X$ . There is an increasing map  $\sigma : J \rightarrow I$  such that

$$y_j = x_{\sigma(j)} \quad (j \in J)$$

and

for each  $i \in I$ , there is some  $j \in J$  with  $i \leq \sigma(j)$ .

Let  $F \in \Phi(\xi)$ . To show that  $F \in \Phi(\eta)$  it suffices to show that some tail of  $\eta$  is contained in  $F$ . There is an  $i \in I$  such that the tail  $T = \{x_k : k \in I \text{ and } i \leq k\}$  of  $\xi$  is a subset of  $F$ . Hence it suffices to show that some tail of  $\eta$  is contained in  $T$ . There is a  $j \in J$  such that  $i \leq \sigma(j)$ . Then  $p \in J$  with  $j \leq p$  implies  $i \leq \sigma(j) \leq \sigma(p)$  and so  $y_p = x_{\sigma(p)} \in T$ , as desired.  $\square$

It is tempting to believe that the converse of Proposition 3.157 holds as well; and that, more generally, if a filter  $\mathcal{F}$  is finer than a filter  $\mathcal{E}$  on  $X$ , then the net  $\Psi(\mathcal{F})$  associated with  $\mathcal{F}$  is a subnet of the net  $\Psi(\mathcal{E})$  associated with  $\mathcal{E}$ . Unfortunately, this is *not* so: see Exercises 267 and 268.

Of special interest among all nets in a given set are the *ultranets*, which are considered in Exercises 6.17–6.22.

ultranet  
filter

### EXERCISES FOR SECTION 3.5

**216.** Which sequences converge to which points in the following spaces?

- (a) The Sorgenfrey line [Examples 2.20 (1)].
- (b) The line with two origins [Examples 2.20 (3)].
- (c) The half-disk space [Examples 2.20 (3)].

**217.** Let  $f: X \rightarrow Y$  be a map from a first-countable space  $X$  to a topological space  $Y$ . Use Corollary 3.92 (2)—but *not* the definition of convergence or Theorem 3.93—to prove anew:  $f$  is continuous if and only if for each  $x \in X$  and each sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  converging to  $x$  in  $X$ , the sequence  $\langle f(x_n) \rangle_{n \in \mathbb{N}}$  converges to  $f(x)$  in  $Y$ .

**218.** Describe the interior of a set in a first-countable space by means of sequences and sequential convergence.

**219.** Let  $X$  and  $Y$  be directed sets with directions  $\leq_X$  and  $\leq_Y$ . Preorder the product set  $X \times Y$  by the relation  $\leq$  defined “coordinatewise”, as in Examples 0.60 (6) and Exercise 0.109. Must  $\leq$  direct  $X \times Y$ ?

- 220.** (a) Must a *finite* directed set have a greatest element? a least element?
- (b) Must an *infinite* directed set have a greatest element? a least element?

**221.** Let  $\langle x_i \rangle_{i \in I}$  be a family in a directed set  $\langle X, \leq \rangle$ .

- (a) Prove: If  $\langle x_i \rangle_{i \in I}$  is eventually in each of two subsets  $A$  and  $B$  of  $X$ , then it is eventually in their intersection  $A \cap B$ .
- (b) Prove the generalization of (a) to the case of finitely many subsets of  $X$ .
- (c) Does the analog of (a) hold for “frequently in” instead of “eventually in”?

**222.** Let  $f: [a, b] \rightarrow \mathbb{R}$  be a bounded real-valued function on a closed interval  $[a, b]$  in  $\mathbb{R}$ . Let  $\mathcal{P}$  be the directed set of all partitions of  $[a, b]$  in the sense of Examples 3.101 (5).

- (a) If  $P, Q \in \mathcal{P}$  with  $P \leq Q$ , show that  $L_P \leq L_Q$  and  $U_Q \leq U_P$ .
- (b) If  $P, Q \in \mathcal{P}$ , show that  $L_P \leq U_Q$ .
- (c) Prove that  $f$  is Riemann integrable if and only if for each  $\varepsilon > 0$  there exist  $P, Q \in \mathcal{P}$  for which  $|U_Q - L_P| < \varepsilon$ .
- (d) Show that  $L = \sup\{L_P : P \in \mathcal{P}\}$  and  $U = \inf\{U_P : P \in \mathcal{P}\}$  both exist. Deduce that  $f$  is Riemann integrable if and only if  $L = U$ , and in that case  $L = U = \int_a^b f(x) \, dx$ .

**223.** (Continuation of Exercise 222.) Suppose  $f: [a, b] \rightarrow \mathbb{R}$  is a bounded, Riemann integrable function. For each partition  $P = \langle x_0, x_1, \dots, x_n \rangle$  of  $[a, b]$  and each  $n$ -tuple  $z = \langle z_1, z_2, \dots, z_n \rangle$  of “sample points” with  $z_i \in [x_{i-1}, x_i]$  for  $i = 1, 2, \dots, n$ , there is an associated *Riemann sum*

$$S_{P,z} = \sum_{i=1}^n f(z_i) (x_i - x_{i-1}).$$

By suitably directing the set  $I$  of such pairs  $(P, z)$ , obtain a net  $\langle S_{P,z} \rangle_{(P,z) \in I}$  that converges to  $\int_a^b f(x) \, dx$ .

convergence! box topology @ and box topology

where 'almost all' here means 'all but finitely many'. Construct a function  $g \in \mathcal{F}$  such that  $g \in \text{cls } \mathcal{A}$  yet no sequence in  $\mathcal{A}$  converges to  $g$  in  $\mathcal{F}$ .

- 235.** Let  $\mathcal{F} = \mathcal{F}(\mathbb{R}, \mathbb{R})$  of all functions from a topological space  $X$  into itself. Let  $g: X \rightarrow X$  be one of these functions. Use nets to determine which of the maps

$$\begin{array}{ccc} \mathcal{F} \rightarrow \mathcal{F} & , & \mathcal{F} \rightarrow \mathcal{F} \\ f \mapsto f \circ g & & f \mapsto g \circ f \end{array}$$

are continuous for the topology of pointwise convergence.

prob:ptwise-conv-non-unif-conv

- 236. (a)** For each nonnegative integer  $n$ , let  $f_n: \mathbb{R} \rightarrow \mathbb{R}$  be the function defined by

$$f(x) = \begin{cases} n & \text{if } 1/(2n+2) \leq x \leq 1/(n+1), \\ 0 & \text{otherwise.} \end{cases}$$

Show that the sequence  $\langle f_n \rangle_{n \in \mathbb{N}}$  converges pointwise in  $\mathcal{F}(\mathbb{R}, \mathbb{R})$  but does not converge uniformly (when the codomain  $\mathbb{R}$  is provided with any bounded metric inducing the usual topology).

- (b)** Construct a sequence of *continuous* functions that converges pointwise in  $\mathcal{F}(\mathbb{R}, \mathbb{R})$  but does not converge uniformly (when the codomain  $\mathbb{R}$  is provided with any bounded metric inducing the usual topology).
- 237.** Let  $\langle f_n \rangle_{n \in \mathbb{N}}$  be a sequence of *continuous* functions from  $\mathbb{R}$  to  $\mathbb{R}$  that converges uniformly to a function  $f$ . Must  $f$  be continuous?
- 238.** Let  $\mathcal{F}$  be the set of all functions from a set  $X$  to a metric space  $\langle Y, d \rangle$ , where  $d$  is not necessarily bounded. Let  $\mathcal{B}$  consist of those  $f \in \mathcal{F}$  that are bounded with respect to  $d$ , that is, for which the range  $f(X)$  is  $d$ -bounded. Show that the topologies  $\mathcal{T}_1$  and  $\mathcal{T}_2$  described below are the same.

- (i) Let  $d^*$  be the bounded metric on  $Y$  obtained from  $d$  as in [Proposition 1.40](#). Form the topology of uniform convergence for  $\mathcal{F}$  using  $d^*$ . Then  $\mathcal{T}_1$  is the topology induced on the subset  $\mathcal{B}$  of  $\mathcal{F}$ .
- (ii) Let  $d_\infty$  be the sup metric on  $\mathcal{B}$  defined by

$$d_\infty(f, g) = \sup_{x \in X} d(f(x), g(x))$$

(compare the discussion preceding [Theorem 1.83](#)). Then  $\mathcal{T}_2$  is the topology induced by the metric  $d_\infty$ .

- 239.** Do two equivalent bounded metrics on a set  $Y$  necessarily determine the same topology of uniform convergence on the set  $\mathcal{F}(X, Y)$  of all functions from a set  $X$  to  $Y$ ?

prob:cofinal-subset

- 240. (a)** Show that the power set of a set  $S$  directed by *reverse* inclusion has a cofinal subset but that the set of all *nonempty* subsets of  $S$ , again directed by *reverse* inclusion, is not cofinal there.
- (b)** Prove that a subset of  $\mathbb{N}$  is cofinal in  $\mathbb{N}$  if and only if it is infinite. [Note: This means, in particular, that if  $\langle n_k \rangle_{k=0}^\infty$  is a strictly increasing sequence in  $\mathbb{N}$ , then  $\{n_k : k = 0, 1, 2, \dots\}$  is cofinal in  $\mathbb{N}$ .]

- 241.** Let  $(x_i)_{i \in I}$  be a net that converges to a point  $x$  in a Hausdorff space  $X$ , so that  $(x_i)_{i \in I}$  necessarily clusters at  $x$  in  $X$ . Can  $(x_i)_{i \in I}$  also cluster at a point of  $X$  different from  $x$ ?

- 242.** Construct a net that clusters at every point of  $\mathbb{R}$ .

- 243.** Let  $f: X \rightarrow Y$  be a map that is continuous at a point  $x \in X$ . If a net  $(x_i)_{i \in I}$  clusters at  $x$  in  $X$ , must the net  $(f(x_i))_{i \in I}$  of values cluster at  $f(x)$  in  $Y$ ?

**244.** Construct nets  $(x_i)_{i \in I}$  and  $(y_i)_{i \in I}$  clustering at points  $x$  and  $y$  in  $\mathbb{R}$ , respectively, such that  $((x_i, y_i))_{i \in I}$  does not cluster at  $\langle x, y \rangle$  in  $X \times Y$ .

**245.** Let  $X = \mathbb{N} \times \mathbb{N}$ . By applying [Theorem 2.19](#), show that there is a topology on  $X$  whose neighborhood systems are the following. If  $\langle m, n \rangle \neq \langle 0, 0 \rangle$ , then

$$V \in \mathcal{N}_{\langle m, n \rangle} \iff \langle m, n \rangle \in V \subset X;$$

and

$$V \in \mathcal{N}_{\langle 0, 0 \rangle} \iff \langle 0, 0 \rangle \in V \text{ and, for almost all } m \in \mathbb{N}, \\ \langle m, n \rangle \in V \text{ for almost all } n \in \mathbb{N}.$$

Construct a sequence in  $X \setminus \{\langle 0, 0 \rangle\}$  that clusters at  $\langle 0, 0 \rangle$  in  $X$  but has no subsequence converging to  $\langle 0, 0 \rangle$  in  $X$ .

**246.** Exhibit a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in  $\mathbb{R}$  and a subnet of this sequence that is not a subsequence of  $\langle x_n \rangle_{n \in \mathbb{N}}$ . If possible, make  $\langle x_n \rangle_{n \in \mathbb{N}}$  convergent.

prob:strictly-incr-seq-ints-is-cofinal **247.** Let  $\langle n_j \rangle_{j \in \mathbb{N}}$  be a strictly increasing sequence of nonnegative integers. Prove that the set  $\{n_j : j \in \mathbb{N}\}$  is cofinal in  $\mathbb{N}$ .

**248.** Let  $I$  and  $J$  be directed sets. Prove that a map  $\sigma : J \rightarrow I$  has properties (SN1) and (SN2) of [Definition 3.119](#) if and only if it has the single property:

property:SN (SN) For each  $i \in I$  there exists  $j \in J$  such that for all  $j' \in J$ ,  
 $j \leq j' \implies i \leq \sigma(j')$ .

**249.** Let  $\langle y_j \rangle_{j \in J}$  be a subnet of a net  $(x_i)_{i \in I}$  in a topological space  $X$ .

(a) If  $(x_i)_{i \in I}$  converges to  $x$ , must  $\langle y_j \rangle_{j \in J}$  necessarily also converge to  $x$ ? If not, must  $\langle y_j \rangle_{j \in J}$  converge to some point of  $X$ ?

(b) If  $(x_i)_{i \in I}$  clusters at  $x$ , must  $\langle y_j \rangle_{j \in J}$  necessarily also cluster at  $x$ ? If not, must  $\langle y_j \rangle_{j \in J}$  cluster at some point of  $X$ ?

**250.** Let  $f : D \rightarrow Y$  and  $x$  be as in [Definition 3.122](#), and let  $\langle t \rangle_{\langle U, t \rangle \in I}$  be the net defined in [Exercise 226](#). Prove that a point  $y \in Y$  is a limit of  $f$  at  $x$  precisely when the net  $\langle f(t) \rangle_{\langle U, t \rangle \in I}$  converges to  $y$  in  $Y$ .

prob:calc-limits-involve-inf **251.** Explain each of the following kinds of limits from calculus as instances of [Definition 3.122](#):

(a)  $\lim_{x \rightarrow +\infty} f(x) = +\infty$ .

(e)  $\lim_{x \rightarrow c^-} f(x) = +\infty$ .

(b)  $\lim_{x \rightarrow +\infty} f(x) = -\infty$ .

(f)  $\lim_{x \rightarrow c^-} f(x) = -\infty$ .

(c)  $\lim_{x \rightarrow -\infty} f(x) = +\infty$ .

(g)  $\lim_{x \rightarrow c^+} f(x) = +\infty$ .

(d)  $\lim_{x \rightarrow -\infty} f(x) = -\infty$ .

(h)  $\lim_{x \rightarrow c^+} f(x) = -\infty$ .

**252.** Construct a map  $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  that is discontinuous at  $\langle 0, 0 \rangle$  and yet for which

$$\lim_{x \rightarrow 0} f(x, y) = f(0, y) \quad (y \in \mathbb{R}),$$

$$\lim_{y \rightarrow 0} f(x, y) = f(x, 0) \quad (x \in \mathbb{R}).$$

**253.** Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a convergent sequence in a Hausdorff space. Explain why  $\lim_{n \rightarrow \infty} x_n$  is the limit of a function in the sense of [Definition 3.122](#). (*Hint:  $N \subset \mathbb{R}^*$ .*)



- 254.** Let  $f: X \rightarrow Y$  and  $g: Y \rightarrow Z$  be maps between Hausdorff spaces. Let  $c \in X$  with  $c \in \text{cls}(X \setminus \{c\})$ . Assume that  $f$  has a limit at  $c$  and that

$$p = \lim_{x \rightarrow c} f(x) \in \text{cls}(Y \setminus \{p\}).$$

- (a) If  $g$  is continuous at  $p$ , show that

$$\lim_{x \rightarrow c} g(f(x)) = g\left(\lim_{x \rightarrow c} f(x)\right).$$

- (b) Give an example where  $g \circ f$  has a limit at  $c$  and yet

$$\lim_{x \rightarrow c} g(f(x)) \neq g\left(\lim_{x \rightarrow c} f(x)\right).$$

Kelley subnet  
subnet!Kelley  
subnet

- prob:Cauchy-eq-and-cont-at-one-pt **255.** Let  $f: \mathbb{R} \rightarrow \mathbb{R}$  be a function that is continuous at some number and satisfies *Cauchy's equation*  $f(a + b) = f(a) + f(b)$  for all real numbers  $a$  and  $b$ . Prove that  $f$  is continuous everywhere.

(Note: In view of [Exercise 26](#), it follows that  $f$  is just multiplication by a constant.)

- prob:adjoin-infty-to-net **256.** Let  $\langle y \rangle_{i \in I}$  be a net in a topological space  $Y$ . Let  $\infty$  be some object that does not belong to  $Y$  (for example,  $\infty = \{Y\}$ ). Let  $X = I \cup \{\infty\}$ .

- (a) Verify that  $X$  has a unique topology such that:

- for each  $i \in I$ , the collection  $\{\{i\}\}$  is a neighborhood base at  $i$ , that is, every subset of  $X$  containing  $i$  is an open neighborhood of  $i$  in  $X$ ; and
- a neighborhood base at  $\infty$  in  $X$  consists of all sets of the form  $\{\infty\} \cup \{j \in I : j \geq j_0\}$  for some  $j_0 \in I$ .

- (b) Fix  $y \in Y$ . Prove that  $\langle y \rangle_{i \in I} \rightarrow y$  in  $Y$  if and only if  $\langle (i, y_i) \rangle_{i \in I} \rightarrow (\infty, x)$  in the product space  $X \times Y$ .]

- prob:Kelley-subnets **257.** This problem concerns an alternative definition of subnets, as presented in Kelley [40]. In this alternative to [Definition 3.119](#), property (SN1) is no longer required, but property (SN2) is strengthened.

Given nets  $\eta = \langle y_j \rangle_{j \in J}$  and  $\xi = \langle x_i \rangle_{i \in I}$  in a set  $X$ , say that  $\eta$  is a **Kelley subnet** of  $\xi$  when there is a map  $\sigma: J \rightarrow I$  such that:

for each  $i$  in  $I$  there is some  $j \in J$  such that, for all  $k \in J$ ,  
 $j \leq k$  implies  $i \leq \sigma(k)$ .

Which, if any, of the following results remain true when Kelley subnets are used instead of subnets in the original, more restrictive, sense?

- (a) [Theorem 3.120](#).  
 (b) [Proposition 3.121](#).

- 258.** (a) Verify that the Fréchet filter on  $\mathbb{N}$  [[Examples 3.132 \(2\)](#)] is the same as the finite-complement filter on  $\mathbb{N}$  [[Examples 3.132 \(2\)](#)].

- (b) Which of the following, if any, is a base of the Fréchet filter on  $\mathbb{R}$  [[Examples 3.132 \(2\)](#)]: the collection  $\{[a, +\infty[ : a \in \mathbb{R}\}$  of closed rays; the collection  $\{[a, +\infty[ : a \in \mathbb{R}, a > 0\}$ ; the collection  $\{]a, +\infty[ : a \in \mathbb{R}, a > 0\}$ ?

- 259.** Let  $\mathcal{E}$  be a nonempty collection of subsets of a set  $X$  with the property that the intersection of any nonempty finite subcollection of  $\mathcal{E}$  is nonempty. Show that there exists a least filter base  $\mathcal{B}$  on  $X$  containing  $\mathcal{E}$  and that the filter generated by  $\mathcal{B}$  is the coarsest filter containing  $\mathcal{E}$ .

Kelley subnet  
subnet!Kelley  
subnet

**260.** Let  $\mathcal{B}_1$  and  $\mathcal{B}_2$  be bases of filters  $\mathcal{F}_1$  and  $\mathcal{F}_2$  on a set  $X$ , respectively. Show that  $\mathcal{F}_2$  is finer than  $\mathcal{F}_1$  if and only if, for each  $B_1 \in \mathcal{B}_1$ , there exists some  $B_2 \in \mathcal{B}_2$  with  $B_2 \subset B_1$ .

**261.** Which filters converge to which points in the following spaces?

- (a) A discrete space?
- (b) An indiscrete space?
- (c) The set  $\mathbb{N}$  provided with its finite-complement topology [Examples 2.3 (7)]?

**262. (a)** At which points of the discrete space  $\mathbb{N}$ , if any, does the finite-complement filter [Examples 3.128 (6)] cluster? To which points, if any, does it converge?

**(b)** At which points in  $\mathbb{R}$ , if any, does the filter on  $\mathbb{R}$  generated by  $\{[0, 1]\}$  cluster? To which points, if any, does it converge?

**(c)** At which points in  $\mathbb{R}^2$ , if any, does the filter on  $\mathbb{R}^2$  generated by the two-point set  $\{(0, 0), (1, 0)\}$  cluster? To which points, if any, does it converge?

prob:cluster-filter-vs-net **263.** Do the analogs of parts (1) and (2) of Proposition 3.156 hold for clustering?

prob:non-unique-filter-clustering **264.** Can a filter on a Hausdorff space cluster at every point of the space?

**265.** The proof of Theorem 3.147 implicitly used the Axiom of Choice. Make its use there explicit.

**266.** Prove that the intersection of all filters converging to a point  $x$  in a topological space is the neighborhood system at  $x$ .

$\gamma$ -filter-finer-not-imply-net-is-subnet **267.** Disprove the converse of Proposition 3.157 by finding nets  $\eta$  and  $\xi$  in a set  $X$  such that  $\Phi(\eta)$  is finer than  $\Phi(\xi)$  yet  $\eta$  is not a subnet of  $\xi$ .

prob:finer-filter-not-imply-subnet **268.** Let  $\mathcal{F}$  and  $\mathcal{E}$  be filters on the set  $X$  with  $\mathcal{F}$  finer than  $\mathcal{E}$ . Show that the net  $\Psi(\mathcal{F})$  associated with  $\mathcal{F}$  need not be a subnet of the net  $\Psi(\mathcal{E})$ .

**269. (a)** Prove: if  $\mathcal{F}$  is a filter on a set  $X$ , then the eventuality filter of the net associated with  $\mathcal{F}$  is the same as the original  $\mathcal{F}$ , that is,  $\Phi(\Psi(\mathcal{F})) = \mathcal{F}$ . Deduce that distinct filters have distinct eventuality filters.

**(b)** If  $\xi$  is a net in a set  $X$ , must the net associated with the eventuality filter of  $\xi$  be the same as  $\xi$ , that is, must  $\Psi(\Phi(\xi)) = \xi$ ?

prob:Kelley-subnets-and-filters **270.** (Continuation of Exercise 257.)

**(a)** Show that the converse of Proposition 3.157 no longer holds when applied to a Kelley subnet.

**(b)** Is an example such as in Exercise 268 still possible when Kelley subnets are used instead of subnets in the original, more restrictive sense?

# Compactness

chap:compact

## Chapter Contents

4.1 Compact Spaces . . . . .	458	Countability properties in compact metrizable spaces . . . . .	493
Compactness of the closed unit interval . . . . .	458	Totally bounded metric spaces . . . . .	494
Open covers and compact spaces . . . . .	459	Convergence in compact metrizable spaces . . . . .	495
Closed sets and compact sets . . . . .	462	More about totally bounded metric spaces . . . . .	498
The Cantor set . . . . .	466	More metric consequences of compactness . . . . .	501
Separation properties in compact spaces . . . . .	473	Exercises for Section 4.2 . . . . .	504
Continuous maps on compact spaces . . . . .	475	4.3 Locally Compact Spaces . . . . .	508
Products of compact spaces . . . . .	477	Locally compact spaces . . . . .	508
Separation in quotients of compact spaces . . . . .	480	Local compactness and continuity . . . . .	511
Compactness and convergence . . . . .	482	One-point compactification . . . . .	513
Exercises for Section 4.1 . . . . .	484	Local compactness and countability . . . . .	517
4.2 Compact Metric Spaces . . . . .	493	Exercises for Section 4.3 . . . . .	518

## Introduction

We have already met several topological properties—metrizable, separated, second-countable, first-countable, and separable. Among the various properties a topological space may possess, one of the most significant in its consequences, particularly those concerning the continuous functions on the space, is that of being *compact*. Just as a real-valued function on a finite set attains a maximum value, so a real-valued continuous function on a compact space attains a maximum value. Indeed, compactness is a topological analog of finiteness and can often be exploited to reduce consideration of all points in a space to the consideration of only finitely many points. That is why so many theorems in calculus involve closed and bounded intervals in  $\mathbb{R}$ , which all turn out to be compact spaces.

The first section of this chapter concerns compact spaces in general; their separation properties; the continuous functions on them; and their subspaces, quotient spaces, and product spaces. The special features of compact spaces that are metrizable are examined in the second section, where compactness is related to completeness. Many important spaces,

although not necessarily compact, have enough compact subspaces to allow compactness arguments to be used for them; these spaces are studied in the third section.

## 4.1 Compact Spaces

sec:compact

To motivate the definition of a compact space, we begin with a very concrete situation, namely, the closed unit interval.

### Compactness of the closed unit interval

subsec:l-is-cpt

thm:unit-interval-compact

**4.1 Proposition.** *Let  $\mathcal{U}$  be an arbitrary collection of open subsets of the unit interval  $[0, 1]$  such that each point of  $I$  belongs to at least one member of  $\mathcal{U}$ . Then there exist finitely many sets from the collection  $\mathcal{U}$  such that each point of  $I$  belongs to at least one of those finitely many sets.*

**Proof.** Define  $A$  to be the set of those points  $x \in [0, 1]$  having the property that  $[0, x]$  is a subset of the union of finitely many sets belonging to  $\mathcal{U}$ . We want to show that  $1 \in A$ . We proceed in several steps.

Step 1: the set  $A$  has a supremum  $b$ . In fact,  $0$  belongs to some member of  $\mathcal{U}$ , so that  $0 \in A$ ; thus  $A$  is nonempty. And the set  $A$  is bounded above, by  $1$ , because  $A \subset [0, 1]$ . By order-completeness of  $\mathbb{R}$  (Axiom 0.74), it is meaningful to define

$$b = \sup A.$$

Step 2:  $b \in A$ . In fact, just suppose  $b \notin A$ . Then since  $b \in [0, 1]$ , there is some  $U_0 \in \mathcal{U}$  with  $b \in U_0$ . Now  $b > 0$  because  $0 \in A$  and  $b \notin A$ . Since  $U_0$  is open, there is an  $\varepsilon > 0$  such that

$$]b - \varepsilon, b] \subset U_0.$$

Since  $b - \varepsilon$  is *not* an upper bound of  $A$ , there is some  $a \in A$  with

$$b - \varepsilon < a < b.$$

By definition of  $A$ , there are finitely many sets  $U_1, U_2, \dots, U_n \in \mathcal{U}$  with

$$[0, a] \subset U_1 \cup U_2 \cup \dots \cup U_n.$$

Then

$$[0, b] = [0, a] \cup [a, b] \subset (U_1 \cup U_2 \cup \dots \cup U_n) \cup U_0.$$

This means that  $b \in A$  after all, which is a contradiction,

Step 3:  $b = 1$ . In fact, just suppose  $b < 1$ . By definition of  $A$ , there are finitely many sets  $U_1, U_2, \dots, U_n \in \mathcal{U}$  with

$$[0, b] \subset U_1 \cup U_2 \cup \dots \cup U_n.$$

Then there is some  $j$  with  $1 \leq j \leq n$  and  $b \in U_j$ . Since  $b < 1$  and  $U_j$  is open, there is an  $\varepsilon > 0$  with

$$[b, b + \varepsilon] \subset U_j.$$

Then

$$[0, b + \varepsilon] = [0, b] \cup [b, b + \varepsilon] \subset (U_1 \cup U_2 \cup \dots \cup U_n) \cup U_j.$$

This means that  $b + \varepsilon \in A$ , which contradicts the definition of  $b$  as the *least* upper bound of the set  $A$ .  $\square$

The preceding proposition, first proved at the end of the nineteenth century by Eduard Heine and Émile Borel, has significant consequences. The following corollary—which was already used in [Chapter 1](#) (see [Example 1.8](#))—is just one of them

Heine, Eduard  
Borel, Émile

prop:cont-interval-to-R-bded

**4.2 Corollary.** *A continuous function  $f: [0, 1] \rightarrow \mathbb{R}$  is bounded.*

**Proof.** Let  $f: [0, 1] \rightarrow \mathbb{R}$  be continuous. It will suffice to find finitely many points  $x_1, x_2, \dots, x_n \in [0, 1]$  having the property that, for each  $x \in [0, 1]$ , there is an index  $i$  with  $1 \leq i \leq n$  and

$$|f(x) - f(x_i)| < 1.$$

Indeed, if  $x \in [0, 1]$  and if  $i$  is such an index, then

$$\begin{aligned} |f(x)| &\leq |f(x) - f(x_i)| + |f(x_i)| \\ &< 1 + \max_{1 \leq i \leq n} |f(x_i)|. \end{aligned}$$

Thus the boundedness of  $f$  reduces to the fact that the finitely many numbers  $|f(x_1)|, |f(x_2)|, \dots, |f(x_n)|$  have a maximum.

To show that points  $x_1, x_2, \dots, x_n$  with the above-stated property actually exist, we use the continuity of  $f$ . Denote the Euclidean metric on  $[0, 1]$  as usual by  $d$ . For each  $x \in [0, 1]$  there is some number  $\delta(x) > 0$  such that

$$\{eq:balls-cover-I\} \quad (*) \quad u \in B_{\delta(x)}(x; d) \implies |f(x) - f(u)| < 1.$$

Now apply [Proposition 4.1](#) to the collection

$$\mathcal{U} = \{B_{\delta(x)}(x; d) : x \in [0, 1]\}.$$

There exist points  $x_1, x_2, \dots, x_n \in [0, 1]$  such that

$$[0, 1] = \bigcup_{i=1}^n B_{\delta(x_i)}(x_i; d).$$

If  $x \in [0, 1]$ , then  $x \in B_{\delta(x_i)}(x_i; d)$  for some  $i$ , and so from [\(\\*\)](#) it follows that

$$|f(x) - f(x_i)| < 1,$$

as desired.  $\square$

The preceding proof used [Proposition 4.1](#), and that proposition may be stated more succinctly as:

Let  $\mathcal{U}$  be an arbitrary collection of open subsets of the unit interval  $[0, 1]$  whose union is  $[0, 1]$ . Then there exists a finite subcollection of  $\mathcal{U}$  whose union is still  $[0, 1]$ .

The argument in the proof of [Corollary 4.2](#) will still be valid for any metrizable space—and, in fact, for any topological space—that shares with  $[0, 1]$  this property about open sets. Accordingly, we shall next embark upon a systematic study of such spaces.

## Open covers and compact spaces

To facilitate studying the kind of spaces just discussed, we introduce some simplifying terminology, which extends that of [Definition 2.81](#). Recall from [Definition 0.23](#) that a collection  $\mathcal{A}$  of sets is a **cover** of a set  $S$  when  $S \subset \bigcup \mathcal{A}$ .

fix: To rewrite  
open cover of  $S$  in  
 $X$ , given earlier  
def of open cover  
of  $X$

open cover  
def:open-cover-in  
Alexandroff, Paul  
Aleksandrov, Pavel  
Urysohn, Pavel  
Fréchet, Maurice

**4.3 Definition.** Let  $S$  be a subset of a topological space  $X$ . Then a collection  $\mathcal{U}$  of subsets of  $X$  is called an **open cover of  $S$  in  $X$**  if  $\mathcal{U}$  covers  $S$  and each  $U \in \mathcal{U}$  is an open subset of  $X$ . An open cover of  $X$  in itself is called simply an **open cover of  $X$** .

closed interval  
compact space  
open interval

The following continuation of [Example 2.82](#) on page 279 illustrates this extended terminology.

ex:covers-of-unit-interval-in- $\mathbb{R}$

**4.4 Example.** As in [Example 2.82](#), let  $\varepsilon_0$  and  $\varepsilon_1$  be real numbers with

$$0 < \varepsilon_0 < 1/2, \quad 0 < \varepsilon_1 < 1/2.$$

And for each  $x \in ]0, 1[$  let  $\varepsilon_x$  be any real number with

$$0 < \varepsilon_x < \min\{x, 1 - x\}.$$

Then:

- the collection

$$\{]x - \varepsilon_x, x + \varepsilon_x[ : 0 \leq x \leq 1\}$$

is an open cover of  $[0, 1]$  in  $\mathbb{R}$ , but it is *not* an open cover of  $[0, 1]$  since its member  $]0 - \varepsilon_0, 0 + \varepsilon_0[$  is not a subset of  $[0, 1]$ ; and

- the collection

$$\{[0, \varepsilon_0[, ]1 - \varepsilon_1, 1]\} \cup \{]x - \varepsilon_x, x + \varepsilon_x[ : 0 < x < 1\}$$

is an open cover of  $[0, 1]$ .  $\diamond$

It is time to name the property that [Proposition 4.1](#) says the closed interval  $[0, 1]$  possesses.

def:cpct

**4.5 Definition.** A topological space  $X$  is said to be **compact** when each open cover of  $X$  contains some *finite* cover of  $X$ . And a topological space is said to be **noncompact** when it is not compact.

This definition was first applied to arbitrary topological spaces by Alexandroff and Urysohn in 1924. [However, they used the term ‘bcompact’ to distinguish it from the term ‘compact’ that had been introduced by Fréchet 18 years earlier for metric spaces having a special property equivalent to that of [Definition 4.5](#): see [Section 4.2](#) and specifically condition (iv) of [Theorem 4.51](#).]

**Usage note.** Contemporary authors often stipulate as part of the definition of a compact space that it be a Hausdorff space and reserve the term “**quasicompact**” for what we have called “compact.” In order to keep our terminology unburdened by extra assumptions, however, we shall *not* follow that practice.

ex:closed-interval-compact

**4.6 Examples.** (1) With only minor modifications, the proof of [Proposition 4.1](#) also shows that *every closed interval  $[a, b]$  in  $\mathbb{R}$  is a compact space*.

ex:open-intervals-rays-noncpt

- (2) **No nonempty open interval in  $\mathbb{R}$  is compact.** To see that this is so, let  $a, b \in \mathbb{R}$  with  $a < b$ . Define  $L = b - a$  and consider the collection

$$\mathcal{U} = \{U_n : n = 1, 2, 3, \dots\}$$

of open subsets of  $]a, b[$  given by

$$U_n = \left] a + \frac{1}{n}, b \right[ \quad (n = 1, 2, 3, \dots).$$

Then  $\mathcal{U}$  is a cover of  $]a, b[$ ; in fact,  $a < x < b$  implies  $x - a > L/n$  for some positive integer  $n$ , and then  $x \in U_n$ . However, if  $\mathcal{F}$  is any finite subcollection of  $\mathcal{U}$ , then

$$a + \frac{L}{m+1} \in ]a, b[, \quad a + \frac{L}{m+1} \notin U_n \quad (U_n \in \mathcal{F})$$

for

$$m = \max\{n : U_n \in \mathcal{F}\},$$

and so  $\mathcal{F}$  is not a cover of  $]a, b[$ .

A similar argument shows that no half-open interval  $]a, b]$  or  $[a, b[$  is compact.

ex:R-noncpt

- (3) **The real line  $\mathbb{R}$  is noncompact,** because the collection

$$\{]-n, n[ : n = 1, 2, 3, \dots\}$$

is an open cover of  $\mathbb{R}$  that contains no finite cover of  $\mathbb{R}$ .

Similar arguments show that no open ray  $]-\infty, b[$  or  $]a, +\infty[$  is compact and no closed ray  $]-\infty, b]$  or  $[a, +\infty[$  is compact.

ex:cpt-intervals-in-R

- (4) As a consequence of (1)–(3), **the only compact intervals in  $\mathbb{R}$  are the empty set and closed intervals.**

ex:finite-space-cpt

- (5) Any finite topological space is compact. In fact, clearly the empty space  $\emptyset$  is compact. Now let  $\mathcal{U}$  be an open cover of a nonempty compact space  $X = \{x_1, x_2, \dots, x_n\}$ . For each  $i = 1, 2, \dots, n$  choose a set  $U_i \in \mathcal{U}$  with  $x_i \in U_i$ . Then the finite collection  $\{U_1, U_2, \dots, U_n\} \subset \mathcal{U}$  is also a cover of  $X$ .

ex:discrete-space-cpt-iff-finite

- (6) For a discrete space  $X$  to be compact it is both necessary and sufficient that  $X$  be a finite set. Sufficiency follows from (5). To prove necessity, suppose that  $X$  is discrete and infinite. Then the infinite collection

$$\mathcal{U} = \{\{x\} : x \in X\}$$

is an open cover of  $X$ , but no subcollection of  $\mathcal{U}$  is a cover of  $X$  except the entire collection  $\mathcal{U}$  itself.

This example and the preceding one show that **compactness may be regarded as a generalization of finiteness.**

ex:nonT2-cpt

- (7) Most of the interesting compact spaces we shall encounter are Hausdorff spaces. However, it is easy to find compact spaces that are *not* Hausdorff spaces. For example, any space  $X$  whose topology is indiscrete [Examples 2.3 (5)] is compact, because it has the single open cover  $\{X\}$ , but  $X$  will *not* be a Hausdorff space if it has at two or more points.

real line!noncompact space@as non  
discrete space  
indiscrete space

(8) A more substantial example of a compact space that is not a Hausdorff space is an infinite set  $X$  provided with its finite-complement topology [Examples 2.3 (7)].

To see that such an  $X$  is not a Hausdorff space, let  $U$  and  $V$  be any two nonempty open sets in  $X$ , so that their complements  $X \setminus U$  and  $X \setminus V$  are both finite; then  $U$  cannot be disjoint from  $V$ , for otherwise

$$X = X \setminus \emptyset = X \setminus (U \cap V) = (X \setminus U) \cup (X \setminus V)$$

would be finite.

To see that such an  $X$  is compact, let  $\mathcal{U}$  be an open cover of  $X$ . Arbitrarily choose some nonempty  $U_0 \in \mathcal{U}$ . If  $U_0 = X$ , then already  $\{U_0\}$  is a finite cover of  $X$ . If, on the other hand,  $U_0 \neq X$ , then  $X \setminus U_0$  is a finite set  $\{x_1, x_2, \dots, x_n\}$ . For each  $i = 1, 2, \dots, n$  there is some  $U_i \in \mathcal{U}$  with  $x_i \in U_i$ . Then the collection  $\{U_0, U_1, \dots, U_n\}$  is a finite cover of  $X$ .

(9) Let  $X$  be an order-complete totally ordered set that has both a least element and a greatest element. Then  $X$  is compact in its order topology. In particular, **both of the ordinal spaces**  $\Omega^+ = [0, \Omega]$  **and**  $\omega^+ = [0, \omega]$  **are compact Hausdorff spaces.**

To see that such  $X$  is compact, let  $\mathcal{U}$  be an open cover of  $X$ . Denote by  $x_0$  and  $x_1$  the least and greatest elements of  $X$ , respectively. Define  $A$  to be the set of those points  $x \in X$  having the property that  $[x_0, x]$  is covered by some finite subcollection of  $\mathcal{U}$ . The aim is to show that  $x_1 \in A$ . Proceed now as when proving compactness of  $[0, 1]$  (Proposition 4.1); existence of  $\sup A$  is guaranteed because  $A$  is bounded above in  $X$ , by  $x_1$  and the totally ordered set  $X$  is order-complete.  $\diamond$

The few examples above are but a meager sample of the totality of compact spaces. We will be able to present many more examples by using the general theory of compact spaces that we shall now begin to develop.

**4.7 Definition.** A subset  $K$  of a topological space  $X$  is said to be **compact** if it is a compact topological space when it is provided with its relative topology induced by the topology of  $X$ .

A subset  $K$  of a topological space  $X$  is a compact subset if and only if each open cover of  $K$  in the entire space  $X$  contains a finite subcollection that also covers  $K$ .

### Closed sets and compact sets

The definition of compactness of a topological space  $X$  may be reformulated as: Given any collection  $\mathcal{U}$  of open subsets of  $X$ , if each finite subcollection  $\mathcal{F}$  of  $\mathcal{U}$  fails to be a cover of  $X$ , then  $\mathcal{U}$  is not a cover of  $X$ . Now a collection  $\mathcal{A}$  of subsets of  $X$  is *not* a cover of  $X$  when  $X \neq \bigcup \mathcal{A}$ , that is, when

$$\bigcap_{A \in \mathcal{A}} (X \setminus A) = X \setminus \bigcup_{A \in \mathcal{A}} A \neq \emptyset.$$

Hence we are led to the following definition and the next theorem.

**4.8 Definition.** A collection  $\mathcal{E}$  of sets is said to have the **finite intersection property** if  $\mathcal{E}$  is nonempty and if each nonempty finite subcollection of  $\mathcal{E}$  has nonempty intersection.



If  $\mathcal{E}$  does have the finite intersection property, then in particular for each  $E \in \mathcal{E}$  the collection  $\{E\}$  has nonempty intersection, that is,  $E \neq \emptyset$ .

Cantor Product Theorem  
Cantor, Georg  
compact space!subspace@and subspace!  
subspace!compact space@of compact space

thm:cpt-via-fip

**4.9 Theorem (finite intersection property criterion).** *A topological space  $X$  is compact if and only if each collection of closed subsets of  $X$  having the finite intersection property itself has nonempty intersection.*

**Proof.** Assume that  $X$  is compact. Let  $\mathcal{E}$  be a collection of closed subsets of  $X$  that has the finite intersection property. Define

$$\mathcal{U} = \{X \setminus E : E \in \mathcal{E}\},$$

so that  $\mathcal{U}$  is a collection of open subsets of  $X$ . If  $\mathcal{F}$  is any *nonempty* finite subcollection of  $\mathcal{U}$ , then  $\{X \setminus U : U \in \mathcal{F}\}$  is a nonempty finite subcollection of  $\mathcal{E}$ ,

$$X \setminus \bigcup_{U \in \mathcal{F}} U = \bigcap_{U \in \mathcal{F}} (X \setminus U) \neq \emptyset,$$

and so  $\mathcal{F}$  is not a cover of  $X$ . The empty collection is also not a cover of  $X$ , because the fact that  $\mathcal{E}$  has a nonempty member tells us that  $X$  itself is nonempty. Thus no finite subcollection of  $\mathcal{U}$  is a cover of  $X$ . Since  $X$  is compact, the collection  $\mathcal{U}$  cannot be a cover of  $X$ . Hence

$$\bigcap_{E \in \mathcal{E}} E = \bigcap_{E \in \mathcal{E}} (X \setminus (X \setminus E)) = X \setminus \bigcup_{E \in \mathcal{E}} (X \setminus E) = X \setminus \bigcup_{U \in \mathcal{U}} U \neq \emptyset.$$

The proof of the converse is left as an exercise.  $\square$

For a simple application of [Theorem 4.9](#), observe that if  $\langle E_n \rangle_{n \in \mathbb{N}}$  is a decreasing sequence of nonempty sets, then the collection  $\{E_n : n \in \mathbb{N}\}$  does have the finite intersection property, because a nonempty finite collection  $\mathcal{F} \subset \{E_n : n \in \mathbb{N}\}$  has as its intersection the nonempty set  $E_m$ , where  $m = \max \{n \in \mathbb{N} : E_n \in \mathcal{F}\}$ . Hence **a decreasing sequence of nonempty closed subsets of a compact space has nonempty intersection**. When applied to  $[0, 1]$  and other compact subspaces of  $\mathbb{R}$ , this result is the classic **Cantor product theorem** ('product' is an obsolete word for 'intersection')—a generalization of the Nested Interval Property ([Theorem 0.83](#)).

To tell whether a subspace of a given topological space is compact we have the following criterion.

subspace-cpt-via-cover-of-whole-space

**4.10 Lemma.** *A subspace  $K$  of a topological space  $X$  is compact if and only if each open cover of  $K$  in  $X$  contains some finite cover of  $K$ .*

**Proof.** Assume first that  $K$  is compact. Let  $\mathcal{U}$  be an open cover of  $K$  in  $X$ . Then  $\{U \cap K : U \in \mathcal{U}\}$  is an open cover of  $K$  (in  $K$ ), and so there is a finite  $\mathcal{F} \subset \mathcal{U}$  for which  $\{U \cap K : U \in \mathcal{F}\}$  is a cover of  $K$ . Hence the finite subcollection  $\mathcal{F}$  of  $\mathcal{U}$  is a cover of  $K$ .

Conversely, assume that each open cover of  $K$  in  $X$  contains some finite cover of  $K$ . Let  $\mathcal{V}$  be an open cover of  $K$  (in  $K$ ). Then  $\mathcal{V} = \{U \cap K : U \in \mathcal{U}\}$  for a collection  $\mathcal{U}$  of open subsets of the entire space  $X$ . By assumption some finite  $\mathcal{F} \subset \mathcal{U}$  is a cover of  $K$ . Hence the finite subcollection  $\{U \cap K : U \in \mathcal{F}\}$  of  $\mathcal{V}$  is a cover of  $K$ .  $\square$

Not every subset of a compact space is compact. For example, the closed unit interval  $[0, 1]$  is compact according to [Proposition 4.1](#), but by [Examples 4.6 \(2\)](#) its subset  $]0, 1[$  is not. There is one case, however, in which a subset of a compact space is necessarily compact: when the subset is closed.

compact space!sub  
thm:closed-in-compact-is-compact

**4.11 Theorem.** *A closed subset of a compact space is itself compact.*

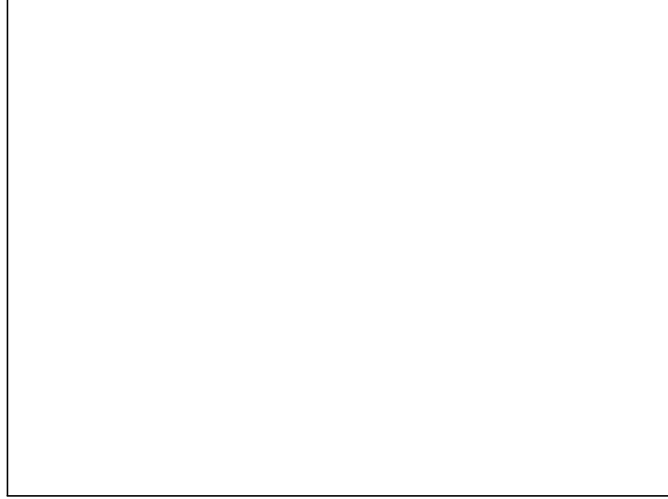


Figure 4.1: A cover of a closed subset  $K$  of a topological space  $X$  by open subsets of  $X$ .

fig:open-cover-closed-subspace-of-c

**Proof.** Let  $K$  be a closed subset of a compact space  $X$ . We use [Lemma 4.10](#). Let  $\mathcal{U}$  be an open cover of  $K$  in  $X$ . The only points of  $X$  that could not belong to any member of  $\mathcal{U}$  must belong to the open subset  $X \setminus K$  of  $K$  (see [Figure 4.1](#)). Hence  $\mathcal{U} \cup \{X \setminus K\}$  is an open cover of  $X$ . Since  $X$  is compact, this larger collection  $\mathcal{U} \cup \{X \setminus K\}$  contains a finite cover  $\mathcal{E}$  of  $X$ . Then the collection  $\mathcal{F} = \mathcal{E} \setminus \{X \setminus K\}$  is a finite subcollection of the original  $\mathcal{U}$ . Moreover,  $\mathcal{F}$  is a cover of  $K$  because no point of  $K$  belongs to the member  $X \setminus K$  that was deleted in forming  $\mathcal{F}$  from  $\mathcal{E}$ .  $\square$

*A compact subset  $K$  of a topological space  $X$  need not be closed in  $X$ .* For example, take  $X$  to be an infinite set with its finite-complement topology [Examples 2.3 \(7\)](#) and take  $K$  to be a subset of  $X$  such that both  $K$  and  $X \setminus K$  are infinite. Then  $K$  is not closed in  $X$ , but  $K$  is compact according to [Examples 4.6 \(8\)](#). In this counterexample the space  $X$  is not a Hausdorff space.

thm:cpt-subset-of-T2-is-closed

**4.12 Theorem.** *Each compact subset of a Hausdorff space is closed in the space.*

**Proof.** Let  $K$  be a compact subset of a Hausdorff space  $X$ . If  $K$  is empty, there is nothing to prove, so we assume that  $K$  is nonempty. Let  $x \in X \setminus K$ . We shall find an open neighborhood  $V$  of  $x$  in  $X$  that is disjoint from  $K$ . Since  $X$  is a Hausdorff space, for each  $y \in K$  there are open subsets  $U_y$  and  $V_y$  of  $X$  with

$$y \in U_y, \quad x \in V_y, \quad U_y \cap V_y = \emptyset$$

(see [Figure 4.2](#)). Then  $\{U_y : y \in K\}$  is an open cover of the compact set  $K$  in  $X$ , and so there exists a finite subset  $F$  of  $K$  such that

$$K \subset \bigcup_{y \in F} U_y,$$

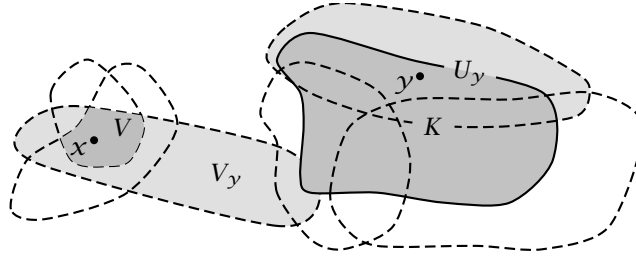


Figure 4.2: Proving that a compact subset of a Hausdorff space is closed.

compact space!Hausdorff space@an  
Hausdorff space!compact set@and c

fig:pf-cpt-in-T2-closed

Let

$$V = \bigcap_{y \in F} V_y.$$

Then  $V$ , being the intersection of finitely many open sets each containing  $x$ , is itself an open subset of  $X$  containing  $x$ .

We show that  $V$  is disjoint from  $K$ . Just suppose there is some  $z \in V \cap K$ . Since  $z \in K$ , there is some  $y$  with

$$y \in F, \quad z \in U_y.$$

Since  $z \in V$ , then  $z \in V_y$ . This is impossible because  $U_y$  is disjoint from  $V_y$ .  $\square$

The preceding proof uses a *standard compactness argument*. We shall see the same sort of argument again, for example, in the proofs of [Lemma 4.18](#) and [Theorem 4.19](#).

page:prototypical-cpt-arg

cor:subset-of-cptT2-cpt-iff-closed

**4.13 Corollary.** Let  $X$  be a compact Hausdorff space. Then a subset of  $X$  is compact if and only if it is closed in  $X$ .

With the aid of the following special result we shall be able to characterize all compact subsets of the real line  $\mathbb{R}$ .

lem:cpt-subset-metric-is-bded

**4.14 Lemma.** Let  $K$  be a compact set in a metric space  $\langle X, d \rangle$ . Then  $K$  is  $d$ -bounded.

**Proof.** We may suppose that  $K$  is nonempty. Since the collection  $\{B_1(x; d) : x \in K\}$  of all balls of radius 1 is an open cover of  $K$  in  $X$ , there are points  $x_1, x_2, \dots, x_n$  in  $K$  with

$$K \subset B_1(x_1; d) \cup B_1(x_2; d) \cup \dots \cup B_1(x_n; d).$$

Set

$$m = \max\{d(x_i, x_j) : 1 \leq i \leq n, 1 \leq j \leq n\}.$$

Then for any two points  $x$  and  $y$  in  $K$  there are indices  $i$  and  $j$  with

$$x \in B_1(x_i; d), \quad y \in B_1(x_j; d),$$

and hence

$$d(x, y) \leq d(x, x_i) + d(x_i, x_j) + d(x_j, y) < 1 + m + 1. \quad \square$$

A much stronger boundedness property of compact metric spaces will be obtained later (see [Proposition 4.48](#)).

Cantor set  
thm:Heine-Borel-thm  
Cantor set

**4.15 Heine–Borel–Lebesgue Theorem.** Let  $d$  be the Euclidean metric on the real line  $\mathbb{R}$ . Then a subset  $K$  of  $\mathbb{R}$  is compact if and only if it is both closed in  $\mathbb{R}$  and  $d$ -bounded.

**Proof.** If  $K$  is compact, then it is closed in  $\mathbb{R}$  by [Theorem 4.12](#) and  $d$ -bounded by the preceding [Lemma 4.14](#).

Conversely, assume that  $K$  is closed in  $\mathbb{R}$  and  $d$ -bounded. Since  $K$  is  $d$ -bounded, it is contained in some closed interval  $[a, b]$ . Now  $[a, b]$  is compact by [Examples 4.6 \(1\)](#). Since  $K$  is closed in  $[a, b]$ , it follows from [Theorem 4.11](#) that  $K$  is compact.  $\square$

The Heine–Borel–Lebesgue Theorem will be extended to Euclidean  $n$ -space  $\mathbb{R}^n$  below (see [Theorem 4.34](#)). However, **a  $d$ -bounded closed subset of a metric space  $\langle X, d \rangle$  need not be compact, even when  $X = \mathbb{R}$  and  $d$  induces the Euclidean topology.** In fact, take  $d$  to be a bounded metric on  $\mathbb{R}$  that is equivalent to the Euclidean metric (such exists by [Proposition 1.40](#)); then  $\mathbb{R}$  itself is  $d$ -bounded and closed in  $\mathbb{R}$ , but by [Examples 4.6 \(3\)](#) the space  $\mathbb{R}$  is not compact. An analog of the Heine–Borel–Lebesgue Theorem involving a strong form of boundedness is true and will be proved in [Section 4.2](#) (see [Corollary 4.60](#)).

### The Cantor set

subsec:Cantor-K

To illustrate how strange a compact subset of the real line can be, we examine in some detail a famous example described in 1884 by Cantor (but previously discovered by several others).

**4.16 Example.** The **Cantor set** (or the *Cantor ternary set*, or the *Cantor middle-third set*) is the subspace  $K$  of  $\mathbb{R}$  constructed as follows:

- from the closed interval  $[0, 1]$  remove its open middle-third subinterval, leaving two disjoint closed intervals of length  $1/3$ ;
- next, from each of those remaining closed intervals remove its open middle-third subinterval, leaving four pairwise disjoint closed intervals each of length  $1/9$ ;
- next, from each of these remaining closed intervals remove its open middle third subinterval, leaving eight pairwise disjoint closed intervals each of length  $1/27$ ; and
- continue this process “forever.”

Then  $K$  is the set of points that remain.

In more formal terms, the Cantor set  $K$  is constructed as follows. For any closed interval  $[a, b]$  in  $\mathbb{R}$ , let

$$[a, b]^* = \left[ a, a + \frac{1}{3}(b - a) \right] \cup \left[ a + \frac{2}{3}(b - a), b \right]$$

be the set obtained from  $[a, b]$  by deleting its subinterval  $\left] a + \frac{1}{3}(b - a), a + \frac{2}{3}(b - a) \right[$ . For a union

$$E = \bigcup_{i=1}^k [a_i, b_i]$$

of finitely many disjoint closed intervals, let

$$E^* = \bigcup_{i=1}^k [a_i, b_i]^*,$$

so that  $E^*$  is a closed set consisting of  $2k$  pairwise disjoint closed intervals.

Starting with the closed unit interval

$$K_0 = [0, 1]$$

successively construct sets

$$K_1 = K_0^* = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right],$$

$$K_2 = K_1^* = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right],$$

and so on, and in general

$$K_{n+1} = K_n^* \quad (n \in \mathbb{N})$$

(see Figure 4.3). This construction produces a decreasing sequence  $\langle K_n \rangle_{n \in \mathbb{N}}$  of closed subsets of  $[0, 1]$ , with  $K_n$  being the union of  $2^n$  pairwise disjoint closed intervals each of length  $1/3^n$ . Finally, define

$$K = \bigcap_{n=0}^{\infty} K_n$$

and provide  $K$  with its relative topology in  $\mathbb{R}$ .

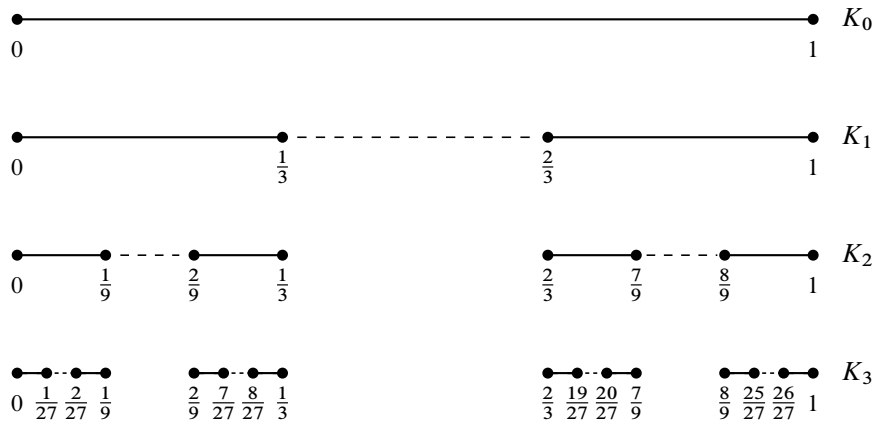


Figure 4.3: The first three steps in forming the Cantor set  $K$ .

fig:Cantor-K

Since  $K$  is closed in  $\mathbb{R}$  and is contained in  $[0, 1]$ , it follows from [Proposition 4.1](#) and [Theorem 4.11](#), or from the Heine–Borel–Lebesgue Theorem ([4.15](#)), that **the Cantor set  $K$  is compact**. As a subspace of the real line,  **$K$  is metrizable**.

The set  $K$  definitely does contain some points—lots of them! Call an  $x \in [0, 1]$  an *endpoint* of  $K$  if for some  $n \geq 0$  it is an endpoint of one of the open intervals that was deleted from  $K_n$  to obtain  $K_{n+1}$ . Then every endpoint of  $K$  actually belongs to  $K$ . Among these endpoints are

$$0, 1, \frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{2}{9}, \frac{7}{9}, \frac{8}{9}, \frac{1}{27}, \frac{2}{27}, \frac{7}{27}, \frac{8}{27}, \frac{19}{27}, \frac{20}{27}, \frac{25}{27}, \frac{26}{27}.$$

Moreover, these are the *only* irreducible fractions  $m/3^n$  that are endpoints of  $K_n$  for  $n = 0, 1, 2, 3$ . Thus neither of the numbers  $4/9$  and  $5/9$  is an endpoint of  $K_2$  and hence neither is an endpoint of  $K$ ; likewise, none of the numbers  $4/27, 5/27, 10/27, 11/27, 13/27, 14/27, 16/27, 17/27, 21/27, 22/27$ , and  $23/27$  is an endpoint of  $K_3$  and hence none is an endpoint of  $K$ . We shall see shortly that there are also many *nonendpoints* of  $K$ , that is, points of  $K$  that are not endpoints of  $K$ .

self-dense space  
Hausdorff, Felix  
Alexandroff, Paul  
Hausdorff, Felix  
Cantor set!

subset-symmetric interval about  $x$  and continuous image

**The Cantor set  $K$  is “self-dense”** in the sense that none of its points are isolated, that is, no singleton is one of its open sets. In fact, just suppose to the contrary that  $\{x\}$  is open in  $K$  for some  $x$ . Then

$$]x - \varepsilon, x + \varepsilon[ \cap K = \{x\}$$

for some  $\varepsilon > 0$ . Choose a positive integer  $n$  with  $1/3^n < \varepsilon$ . Since  $x \in K_n$ , we have  $x \in [a, b]$ , where  $[a, b]$  is one of the  $2^n$  intervals whose union is  $K_n$ , and

$$[a, b] \subset ]x - \varepsilon, x + \varepsilon[.$$

At least one endpoint  $y$  of  $[a, b]$  is different from  $x$ . Since  $y$  is an endpoint of  $K$ , necessarily  $y \in K$ , which contradicts (\*).

In the next chapter (in the subsection “Totally disconnected spaces”) we shall see that **the Cantor set  $K$  is “totally disconnected.”** This means, roughly speaking, that the only “connected pieces” of  $K$  are singletons.

The properties of  $K$  already mentioned characterize this space topologically! It is a theorem of Hausdorff that every totally disconnected, self-dense, compact, metrizable space containing at least two points is homeomorphic to the Cantor set. Moreover, a theorem of Alexandroff and Hausdorff asserts that every nonempty compact metrizable space is a continuous image of the Cantor set. Hence from Corollary 4.24, below, and Proposition 3.80, every nonempty compact metrizable space is homeomorphic to a quotient space of the Cantor set. (Some concrete instances of the Alexandroff–Hausdorff theorem appear in Example 4.17 and Application 6.41. Proofs of both theorems are given in Willard [74, sec. 30, pp. 216–218].)

We already know that every endpoint of  $K$  belongs to  $K$ . For the endpoints of  $K_0, K_1, K_2, K_3$  listed earlier, observe:

$$\begin{aligned} 0 &= \frac{0}{3^0}, & 1 &= \frac{1}{3^0}, \\ \frac{1}{3} &= \frac{1}{3^1}, & \frac{2}{3} &= \frac{2}{3^1}, \\ \frac{1}{9} &= \frac{0}{3^1} + \frac{1}{3^2}, & \frac{2}{9} &= \frac{0}{3^1} + \frac{2}{3^2}, & \frac{7}{9} &= \frac{2}{3^1} + \frac{1}{3^2}, & \frac{8}{9} &= \frac{2}{3^1} + \frac{2}{3^2}, \\ \frac{1}{27} &= \frac{0}{3^1} + \frac{0}{3^2} + \frac{1}{3^3}, & \frac{2}{27} &= \frac{0}{3^1} + \frac{0}{3^2} + \frac{2}{3^3}, & \frac{7}{27} &= \frac{0}{3^1} + \frac{2}{3^2} + \frac{1}{3^3}, \\ \frac{8}{27} &= \frac{0}{3^1} + \frac{2}{3^2} + \frac{2}{3^3}, & \frac{19}{27} &= \frac{2}{3^1} + \frac{0}{3^2} + \frac{1}{3^3}, & \frac{20}{27} &= \frac{2}{3^1} + \frac{0}{3^2} + \frac{2}{3^3}, \\ \frac{25}{27} &= \frac{2}{3^1} + \frac{2}{3^2} + \frac{1}{3^3}, & \frac{26}{27} &= \frac{2}{3^1} + \frac{2}{3^2} + \frac{2}{3^3}. \end{aligned}$$

These representations suggest that, in order to determine exactly which of the points in  $[0, 1]$  belong to  $K$ , we use the ternary (base 3) expansions of real numbers in  $[0, 1]$ , as follows.

Given a sequence  $\langle x_i \rangle_{i=1,2,3,\dots}$  with each  $x_i = 0, 1$ , or  $2$ , by comparing the series  $\sum_{i=1}^{\infty} x_i/3^i$  with the convergent geometric series  $\sum_{i=1}^{\infty} 2/3^i$  we see that  $\sum_{i=1}^{\infty} x_i/3^i$  converges to a real number  $x \in [0, 1]$ , and then we write

$$x = (.x_1x_2 \dots x_i \dots)_3.$$

Some numbers in  $[0, 1]$  have two such distinct ternary expansions; for example,

$$\begin{aligned} \frac{1}{3} &= (.010 \dots 0 \dots)_3 = (.002 \dots 2 \dots)_3, \\ \frac{2}{3} &= (.200 \dots 0 \dots)_3 = (.122 \dots 2 \dots)_3. \end{aligned}$$

We are about to choose systematically one of the two ternary expansions for all those numbers having two.

Let  $x \in [0, 1]$ . Construct recursively a certain sequence

$$s(x) = \langle x_i \rangle_{i=1,2,3,\dots}$$

in  $\{0, 1, 2\}$  with

$$x = (.x_1x_2 \dots x_i \dots)_3 = \sum_{i=1}^{\infty} \frac{x_i}{3^i}$$

as follows. If  $x = 0$ , take  $x_i = 0$  for all  $i$ . Now let  $x > 0$ . Take  $x_1$  to be the unique  $k \in \{0, 1, 2\}$  such that

$$\frac{k}{3^1} < x \leq \frac{k+1}{3^1}.$$

Once  $x_1, x_2, \dots, x_{i-1}$  have been constructed, take  $x_i$  to be the unique  $k \in \{0, 1, 2\}$  such that

$$\frac{k}{3^i} < x - \sum_{j=1}^{i-1} \frac{x_j}{3^j} \leq \frac{k+1}{3^i}.$$

Since  $0 < \sum_{j=1}^{i-1} x_j/3^j \leq 3/3^{i+1}$  for each  $i \geq 1$ , it follows that  $x = \sum_{i=1}^{\infty} x_i/3^i$ .

The preceding construction yields, for example, the expansions

$$\begin{aligned} \frac{1}{3} &= (.022 \dots 2 \dots)_3, & \frac{2}{3} &= (.122 \dots 2 \dots)_3, & 1 &= (.222 \dots 2 \dots)_3, \\ \frac{1}{9} &= (.002 \dots 2 \dots)_3, & \frac{2}{9} &= (.012 \dots 2 \dots)_3, & \frac{4}{9} &= (.102 \dots 2 \dots)_3. \end{aligned}$$

The sequence  $s(x) = \langle x_i \rangle_{i=1,2,3,\dots}$  associated with an  $x \in [0, 1]$  was so constructed that it is *not* the case that, for some  $n$ ,

$$x_n \neq 0 \text{ and } x_i = 0 \text{ for all } i > n.$$

In fact, if  $x_n = 1$  and  $x_i = 0$  for all  $i > n$ , then the definition of  $x_n$  gives

$$\frac{1}{3^n} < x - \sum_{j=1}^{n-1} \frac{x_j}{3^j} = \frac{1}{3^n} \leq \frac{2}{3^n},$$

which is impossible; the case that  $x_n = 2$  and  $x_i = 0$  for all  $i > n$  is likewise impossible.

Make one adjustment to the definition of  $s(x) = \langle x_i \rangle_{i=1,2,3,\dots}$ : If for some  $n$  we have  $x_n = 1$  and  $x_i = 2$  for all  $i > n$ , then redefine  $x_n, x_{n+1}, \dots$  by

$$x_n = 2 \text{ and } x_i = 0 \text{ for all } i > n.$$

It is still true that

$$x = (.x_1x_2 \dots x_i \dots)_3,$$

but now *both the cases*

$$\begin{aligned} x_n = 1 \quad \text{and} \quad x_i = 0 \text{ for all } i > n, \\ x_n = 1 \quad \text{and} \quad x_i = 2 \text{ for all } i > n \end{aligned}$$

are excluded. Thus the adjusted construction yields, for example, the expansions

$$\begin{aligned} \frac{1}{3} &= (.022 \dots 2 \dots)_3, & \frac{2}{3} &= (.200 \dots 0 \dots)_3, & 1 &= (.222 \dots 2 \dots)_3, \\ \frac{1}{9} &= (.002 \dots 2 \dots)_3, & \frac{2}{9} &= (.020 \dots 0 \dots)_3, & \frac{4}{9} &= (.102 \dots 2 \dots)_3. \end{aligned}$$

Which of the sequences  $s(x)$  arise from points in  $K$ ? Let  $x \in [0, 1]$  be arbitrary and let  $s(x) = \langle x_i \rangle_{i=1,2,3,\dots}$ . Induction on  $n$  shows that to each  $n \geq 1$ ,

$$\text{eq:which-ternary-expansionsfor-Kn} \quad (**) \quad x \in K_n \iff x_1 \neq 1, x_2 \neq 1, \dots, x_n \neq 1.$$

Cantor, George  
Cantor function  
Cantor set!continuous image continuous image

To start the induction, observe that each  $x \in K_1 = [0, 1/3] \cup [2/3, 1]$  has the ternary expansion  $x = (.0x_2x_3\ldots)_3$  or  $x = (.2x_2x_3)_3$ , whereas each  $x \in ]1/3, 2/3[$  has the ternary expansion  $x = (.1x_2x_3\ldots)_3$ . From (\*\*) it follows that

$$x \in K \iff x_n \in \{0, 2\} \text{ for all } n = 1, 2, 3, \dots$$

Thus  $s(x) \in \{0, 2\}^{\mathbb{N}^*}$  for all  $x \in K$ .

Restrict the domain and codomain of  $s: [0, 1] \rightarrow \{0, 1, 2\}^{\mathbb{N}^*}$  to obtain the *bijection*

$$\begin{aligned} \sigma: K &\rightarrow \{0, 2\}^{\mathbb{N}^*} \\ x &\mapsto s(x) \end{aligned}$$

Exactly as in [Example 0.53](#), the set  $\{0, 2\}^{\mathbb{N}^*}$  of all sequences of 0s and 2s is uncountable. Hence **the Cantor set is uncountable**, too.

Provide  $\{0, 2\}$  with its discrete topology and  $\{0, 2\}^{\mathbb{N}^*}$  with its product topology. We claim that

$$\sigma: K \cong \{0, 2\}^{\mathbb{N}^*}.$$

To prove this, it suffices to show that the inverse

$$\begin{aligned} \tau: \{0, 2\}^{\mathbb{N}^*} &\rightarrow K \\ \langle x_i \rangle_{i=1,2,3,\dots} &\mapsto \sum_{i=1}^{\infty} \frac{x_i}{3^i} \end{aligned}$$

of  $\sigma$  is a homeomorphism. Let  $x = \langle x_i \rangle_{i=1,2,3,\dots} \in \{0, 2\}^{\mathbb{N}^*}$  be arbitrary. The space  $\{0, 2\}^{\mathbb{N}^*}$  has as a local base at  $x$  the collection of all sets of the form

$$V_n = \{ \langle y_i \rangle_{i=1,2,3,\dots} : y_1 = x_1, y_2 = x_2, \dots, y_n = x_n \}$$

for some  $n \geq 1$  (see the paragraph preceding [Proposition 3.57](#)). Let  $\varepsilon > 0$  be arbitrary. Choose  $n \geq 1$  with

$$\frac{1}{3^n} < \varepsilon.$$

Then  $y = \langle y_i \rangle_{i=1,2,3,\dots} \in V_n$  implies

$$|f(y) - f(x)| = \left| \sum_{i=n+1}^{\infty} \frac{y_i - x_i}{3^i} \right| \leq \sum_{i=n+1}^{\infty} \frac{2}{3^i} = \frac{1}{3^n} < \varepsilon.$$

Hence  $\tau$  is continuous at  $x$ . A similar argument shows that  $\tau$  is open at  $x$ .  $\diamond$

In 1883, Cantor used his eponymous set to construct a surprising example of a continuous function.

**4.17 Example.** The **Cantor function**  $\kappa: [0, 1] \rightarrow [0, 1]$ , also known more picturesquely as the **Cantor staircase function**, to be defined below, has the remarkable properties:

- $\kappa$  is continuous and increasing;
- $\kappa$  maps  $[0, 1]$  *onto* itself; and, in fact
- $\kappa$  maps the Cantor set  $K$  onto  $[0, 1]$ .

While the Cantor set is zero-dimensional in the sense of [Definition 2.61](#) nonetheless the 1-dimensional space  $[0, 1]$  is its continuous image. The Cantor function destroys any naive notion of the dimension of a space as being “the number of continuous parameters needed to describe it.” So do *plane-filling curves*: see [Example 5.65](#) and the discussion preceding it.



In fact, as [Application 6.41](#) will show, the Cantor function furnishes a way of obtaining a plane-filling curve different from the construction in [Example 5.65](#).

The Cantor function is of particular interest in analysis because of this additional property:

- $\kappa$  has derivative 0 at all points of  $[0, 1] \setminus K$ .

Thus, despite its being continuous and taking on all values from 0 to 1, nevertheless the function has slope 0 “almost everywhere,” namely, at all points of the interval except for those in the zero-dimensional, nowhere dense subset  $K$ !

This function is defined first on the open intervals that are removed from  $[0, 1]$  to obtain the Cantor set  $K$ . For each  $n = 1, 2, 3, \dots$ , denote by  $J_{n,1}, J_{n,2}, \dots, J_{n,2^{n-1}}$ , in order from left to right, the open intervals removed from  $K_{n-1}$  to obtain  $K_n$  and let

$$J_n = \bigcup_{i=1}^{2^{n-1}} J_{n,i},$$

so that

$$[0, 1] \setminus K_n = \bigcup_{n=1}^{\infty} J_n.$$

Thus

$$\begin{aligned} J_{1,1} &= ]\frac{1}{3}, \frac{2}{3}[ , \\ J_{2,1} &= ]\frac{1}{9}, \frac{2}{9}[ , \quad J_{2,2} = ]\frac{7}{9}, \frac{8}{9}[ , \\ J_{3,1} &= ]\frac{1}{27}, \frac{2}{27}[ , \quad J_{3,2} = ]\frac{7}{27}, \frac{8}{27}[ , \quad J_{3,3} = ]\frac{19}{27}, \frac{20}{27}[ , \quad J_{3,4} = ]\frac{25}{27}, \frac{26}{27}[ . \end{aligned}$$

Now define

$$\begin{aligned} \kappa_0(x) &= \frac{1}{2} \text{ if } x \in J_{1,1}, \\ \kappa_0(x) &= \frac{1}{4} \text{ if } x \in J_{2,1}, \quad \kappa_0(x) = \frac{3}{4} \text{ if } x \in J_{2,2}, \\ \kappa_0(x) &= \frac{1}{8} \text{ if } x \in J_{3,1}, \quad \kappa_0(x) = \frac{3}{8} \text{ if } x \in J_{3,2}, \quad \kappa_0(x) = \frac{5}{8} \text{ if } x \in J_{3,3}, \quad \kappa_0(x) = \frac{7}{8} \text{ if } x \in J_{3,4}, \end{aligned}$$

and, in general,  $\kappa_0(x)$  takes successive constant values  $1/2^n, 3/2^n, \dots, (2^n - 1)/2^n$  on the intervals  $J_{n,1}, J_{n,2}, \dots, J_{n,2^{n-1}}$  constituting  $J_n$ .

Since  $\kappa_0$  is constant on each  $J_{n,i}$ , it now extends continuously to the union of the closures of all the intervals  $J_{n,i}$  by giving it at the endpoints of each  $J_{n,i}$  the same constant value it has on  $J_{n,i}$ ; and then it extends also to the endpoints 0 and 1 of  $[0, 1]$  by  $\kappa_0(0) = 0$  and  $\kappa_0(1) = 1$ . In this way,  $\kappa_0$  becomes defined at all the endpoints of  $K$ .

Observe that

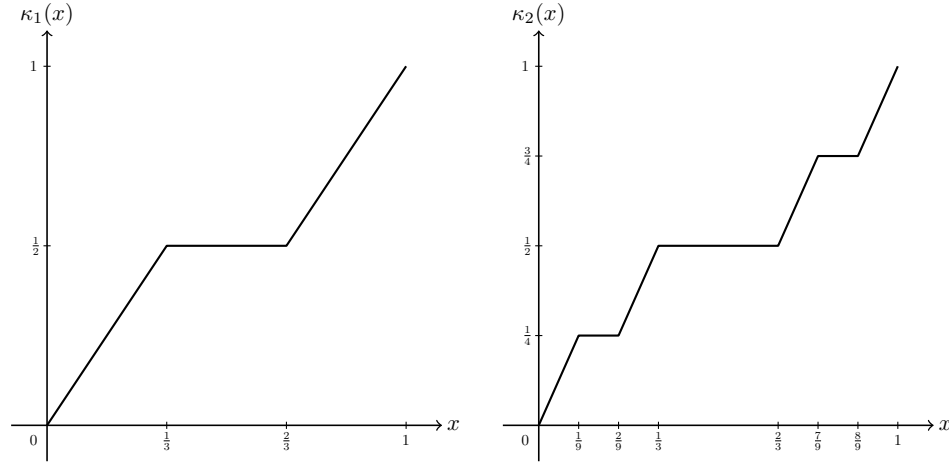
$$|x - u| < \frac{1}{3^n} \implies |\kappa_0(x) - \kappa_0(u)| < \frac{1}{2^n} \quad (x, u \in [0, 1] \setminus K),$$

which means that  $\kappa_0$  is uniformly continuous on  $[0, 1] \setminus K$ . Since  $[0, 1] \setminus K$  is dense in  $[0, 1]$ , the function  $\kappa_0$  has a unique continuous extension  $\kappa$  to  $[0, 1]$ . The **Cantor function** is this extension  $\kappa$ .

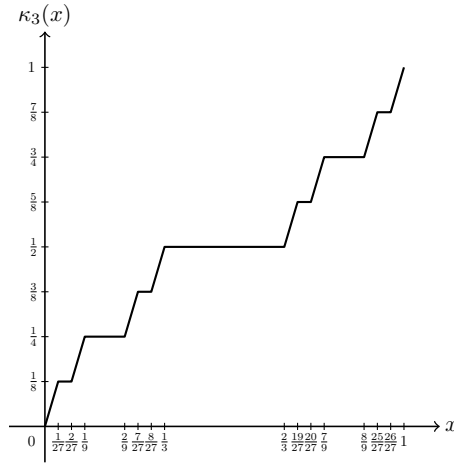
In order to establish properties of the Cantor function, we shall obtain  $\kappa$  explicitly. If the process of obtaining the Cantor set as an intersection of the complements  $K_n$  of the sets  $J_n$  would terminate at some finite number  $m$  of steps, then  $\kappa_0$  could be extended to a function  $\kappa$  on all of  $[0, 1]$  by simply “connecting the dots,” that is, by interpolating linearly between each consecutive pair of endpoints of  $J_m$ . However, the process does not terminate.

Instead, for each  $n = 1, 2, 3, \dots$ , extend  $\kappa_0$  to  $[0, 1]$  by making it *linear* on each interval of  $K_n$ , thereby obtaining a function  $\kappa_n: [0, 1] \rightarrow [0, 1]$  that is piecewise linear, increasing,

continuous, and surjective. (See Figure 4.4 for the graphs of  $\kappa_1$ ,  $\kappa_2$ , and  $\kappa_3$ . Note that the horizontal segments of the graphs are from the graph of  $\kappa_0$ .)



(a) First approximation  $\kappa_1$  to Cantor's function. Second approximation  $\kappa_2$  to Cantor's function.



(c) Third approximation  $\kappa_3$  to Cantor's function.

Figure 4.4: The first three approximations to Cantor's function  $\kappa$ .

fig:kappa-ns

For each  $n$ :

$$\{eq:kappas-restrict\} \quad (*) \quad \kappa_{n+1}|_{J_n} = \kappa_n|_{J_n},$$

and

$$\{eq:kappas-diff\} \quad (**) \quad |\kappa_{n+1}(x) - \kappa_n(x)| < \frac{1}{2^n} \quad (x \in [0, 1]).$$

From  $(**)$  it follows that, for every  $n = 1, 2, 3, \dots$ ,

$$|\kappa_{n+k}(x) - \kappa_n(x)| < \sum_{j=n}^{n+k-1} \frac{1}{2^j} < \sum_{j=n}^{\infty} \frac{1}{2^j} = \frac{1}{2^{n-1}} \quad (x \in [0, 1], k = 1, 2, 3, \dots).$$

Thus

$$d_\infty(\kappa_{n+k}, \kappa_n) < \frac{1}{2^{n-1}} \quad (n = 1, 2, 3, \dots, k = 1, 2, 3, \dots),$$

where  $d_\infty$  is the sup metric on the space  $C([0, 1])$  of all continuous real-valued functions on  $[0, 1]$ . This means that  $\langle \kappa_n \rangle_{n \in \mathbb{N}}$  is a Cauchy sequence in  $\langle C([0, 1]), d_\infty \rangle$ . According to [Corollary 1.87](#), that metric space is complete, and so the sequence converges there—that is, converges *uniformly*—to a continuous function  $\kappa: [0, 1] \rightarrow [0, 1]$ . (The graph of  $\kappa$  is represented in [Figure 4.5](#).) Since each  $\kappa_n$  extends  $\kappa_0$  and is increasing and surjective, the

Cantor function  
Cantor set

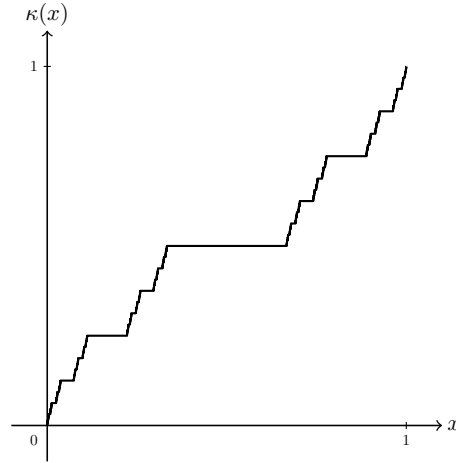


Figure 4.5: Cantor's function  $\kappa$ .

fig:Cantor-fn

same is true of  $\kappa$ .

An explicit formula for  $\kappa(x)$  for all  $x \in [0, 1]$  is possible. In terms of ternary expansion,

$$x = \sum_{i=1}^{\infty} \frac{x_i}{3^i} \quad \text{if } x \in K \text{ with } x_i \in \{0, 2\} \text{ for each } i,$$

then

$$\kappa(x) = \sum_{i=1}^{\infty} \frac{x_i/2}{2^i} = \frac{1}{2} \sum_{i=1}^{\infty} \frac{x_i}{2^i},$$

the latter being a *binary* expansion of  $\kappa(x)$ . Moreover,

$$\kappa(x) = \sup\{\kappa(t) : t \in [0, 1] \setminus K, t \leq x\} \quad \text{if } x \in [0, 1] \setminus K. \quad \diamond$$

### Separation properties in compact spaces

subsec:cpt-and-separation

Let us return to general properties of compact spaces. The “standard compactness argument” used in the proof of [Theorem 4.12](#) actually yields a stronger conclusion than the asserted theorem. Recall that, in that proof, for the compact subset  $K$  of the Hausdorff space  $X$  and a point  $x \in X \setminus K$  we obtained families  $\langle U_y : y \in F \rangle$  and  $\langle V_y : y \in F \rangle$  of open sets such that

$$\begin{aligned} x \in V &= \bigcap_{y \in F} V_y, & K &\subset \bigcup_{y \in F} U_y, \\ U_y \cap V_y &= \emptyset & (y \in F). \end{aligned}$$

Then the set

$$U = \bigcup_{y \in F} U_y$$

is disjoint from  $V$ . This proves the following lemma.

lem:separate-pt-from-cpt-in-T2

**4.18 Lemma.** *Let  $K$  be a compact subset of a Hausdorff space  $X$  and let  $x \in X \setminus K$ . Then there exist disjoint open subsets  $U$  and  $V$  of  $X$  with*

$$x \in V, \quad K \subset U.$$

Thus in a Hausdorff space, any compact subset can be separated (by open sets) from any point not in the subset. Something even stronger is true—and again is proved by using a “standard compactness argument.”

thm:separate-disj-cpts-in-T2

**4.19 Theorem (separation of compact sets in a Hausdorff space).** *Let  $K$  and  $L$  be disjoint compact subsets of a Hausdorff space  $X$ . Then there exist disjoint open subsets  $U$  and  $V$  of  $X$  such that*

$$K \subset U, \quad L \subset V.$$

**Proof.** We may suppose that neither  $K$  nor  $L$  is empty. By [Lemma 4.18](#), for each  $x \in L$  there are open subsets  $U_x$  and  $V_x$  of  $X$  such that

$$x \in V_x, \quad K \subset U_x, \quad U_x \cap V_x = \emptyset.$$

Since  $\{V_x : x \in L\}$  is an open cover of  $L$  in  $X$ , there exists a nonempty finite set  $E \subset L$  with

$$L \subset \bigcup_{x \in E} V_x.$$

The desired sets  $U$  and  $V$  separating  $K$  and  $L$  are now obtained by taking

$$U = \bigcap_{x \in E} U_x, \quad V = \bigcup_{x \in E} V_x. \quad \square$$

Notice that the preceding proof used a “standard compactness argument” (see the comment on [page 465](#) following the proof of [Theorem 4.12](#)).

By combining [Theorem 4.11](#) and [Theorem 4.19](#) we obtain the following theorem.

thm:cpt-T2-is-normal

**4.20 Theorem (normality of compact  $T_2$ -space).** *Each compact Hausdorff space is normal.*

The preceding theorem says that disjoint closed subsets of a compact Hausdorff space can be separated by open sets; that is, for any two disjoint closed subsets  $A$  and  $B$  of a compact Hausdorff space  $X$ , then there exist disjoint open subsets  $U$  and  $V$  such that  $A \subset U$  and  $B \subset V$ . Something much more is true—and not just for compact Hausdorff spaces but for arbitrary normal spaces—namely, disjoint closed subsets of a normal space can be separated by a continuous real-valued function on the space; that is, for any two disjoint closed subsets  $A$  and  $B$  of a normal space  $X$ , then there exists a continuous function  $f: X \rightarrow [0, 1]$  such that  $f(a) = 0$  for all  $a \in A$  and  $f(b) = 1$  for all  $b \in B$ . This is the content of *Urysohn’s Lemma* ([6.26](#)). (For a special case, see [Exercise 25](#).)

### Continuous maps on compact spaces

compact space!continuous map@an  
continuous map!compact space@on

subsec:cpt-and-cont

As with any topological property, we want to know whether compactness is preserved under the formation of subspaces, quotient spaces, and product spaces. We have already discussed subspaces of compact spaces—see [Theorem 4.11](#). Concerning quotient spaces we have the key, more general, result that any continuous image of a compact space is itself compact.

thm:cont-image-cpt-is-cpt **4.21 Theorem.** *The image of a compact space under a continuous map is itself compact.*

**Proof.** Let  $f: X \rightarrow Y$  be a continuous map from a compact space  $X$  to an arbitrary topological space  $Y$ . Let  $\mathcal{V}$  be an open cover of  $f(X)$  in  $Y$ . By continuity of  $f$ , the inverse image  $f^{-1}(V)$  is open in  $X$  for each  $V \in \mathcal{V}$ . Moreover,  $\{f^{-1}(V) : V \in \mathcal{V}\}$  is a cover of  $X$  because

$$X = f^{-1}(f(X)) \subset f^{-1}\left(\bigcup_{V \in \mathcal{V}} V\right) = \bigcup_{V \in \mathcal{V}} f^{-1}(V).$$

Then there exists a finite subcollection  $\mathcal{F}$  of  $\mathcal{V}$  such that  $\{f^{-1}(V) : V \in \mathcal{F}\}$  is a cover of  $X$ . Hence

$$f(X) = f\left(\bigcup_{V \in \mathcal{F}} f^{-1}(V)\right) = \bigcup_{V \in \mathcal{F}} f(f^{-1}(V)) = \bigcup_{V \in \mathcal{F}} V,$$

which means that  $\mathcal{F}$  is a cover of  $f(X)$ .  $\square$

cor:quotient-of-cpt-is-cpt **4.22 Corollary.** *Any quotient space of a compact space is itself compact.*

cor:homeomorph-of-cpt-is-cpt **4.23 Corollary.** *Compactness is a topological property: a topological space that is homeomorphic to a compact space is also compact.*

For an application, recall that

$$\widehat{\mathbb{R}} \cong [0, 1], \quad \mathbb{R} \not\cong ]0, 1[.$$

From the compactness of  $[0, 1]$  and the noncompactness of  $]0, 1[$ , it follows that ***the extended real line  $\widehat{\mathbb{R}}$  is compact, but the real line  $\mathbb{R}$  is noncompact.***

It was emphasized earlier that continuity is not in general a sufficient condition for a bijection to be a homeomorphism. However, it will follow from the next corollary that continuity *is* sufficient under special circumstances.

cor:cont-from-cpt-to-T2-is-closed **4.24 Corollary (closed map principle).** *Let  $f: X \rightarrow Y$  be a continuous map from a compact space  $X$  into a Hausdorff space  $Y$ . Then  $f$  is a closed map.*

In particular, under the conditions of the preceding corollary,  $f$  will be a quotient map if it is surjective.

continuous bijection  
 cor:cont.bij.principle  
 compact space!  
 continuous map!  
 compact space!  
 Hausdorff space!  
 compact space!  
 Hausdorff space!

**4.25 Corollary (continuous bijection principle).** Let  $f: X \rightarrow Y$  be a continuous bijection from a compact space  $X$  into a Hausdorff space  $Y$ . Then  $f$  is a homeomorphism.

Applying the continuous bijection principle will often shorten the work of showing that a given bijection is a homeomorphism.

It also follows from the closed map principle that a continuous injection from a compact space into a Hausdorff space is an embedding.

The first two parts of the following proposition are straightforward to establish; the second two follow from previous results. (See [Exercise 26](#).)

prop:cptT2-max-cpt-min-T2

**4.26 Proposition.** Let  $\mathcal{T}$  be a topology on a set  $X$ .

- (1) If  $\mathcal{T}$  is compact, then every coarser topology on  $X$  is also compact.
- (2) If  $\mathcal{T}$  is Hausdorff, then every finer topology on  $X$  is also Hausdorff.
- (3) If  $\mathcal{T}$  is compact Hausdorff, then there is no compact topology on  $X$  that is strictly finer than  $\mathcal{T}$ .
- (4) If  $\mathcal{T}$  is compact Hausdorff, then there is no Hausdorff topology on  $X$  that is strictly coarser than  $\mathcal{T}$ .

prop-part:cptT2-max-cpt

prop-part:cptT2-min-T2

Thus among all topologies on a given set, **a compact Hausdorff topology is maximally compact and minimally Hausdorff**.

At the beginning of this section we proved that any continuous real-valued function on  $[0, 1]$  is bounded. We can now both generalize and strengthen that result.

cor:extreme-value-thm

**4.27 Corollary (Extreme Value Theorem).** Let  $f: X \rightarrow \mathbb{R}$  be a continuous real-valued function on a nonempty compact space  $X$ . Then  $f$  is bounded. Moreover,  $f$  attains both a minimum value and a maximum value on  $X$ , that is, there exist  $x_1, x_2 \in X$  such that

$$f(x_1) \leq f(x) \leq f(x_2) \quad (x \in X).$$

**Proof.** Boundedness of  $f$  is a consequence of [Theorem 4.21](#) and [Lemma 4.14](#). Let

$$m_1 = \inf f(X), \quad m_2 = \sup f(X).$$

Now the image  $f(X)$  is closed in  $\mathbb{R}$  by [Corollary 4.24](#), and so  $m_1 \in f(X)$  and  $m_2 \in f(X)$ . Finally, take  $x_1$  and  $x_2$  to be any points of  $X$  for which  $f(x_1) = m_1$  and  $f(x_2) = m_2$ .  $\square$

In the special case that  $X$  is a closed interval  $[a, b]$  in  $\mathbb{R}$ , this is a familiar fact having many uses in calculus. For an application to arbitrary metric spaces, consider a nonempty subset  $A$  of a metric space  $\langle X, d \rangle$  and a point  $x \in X$ . We know that in general the distance  $d(x, A)$  from  $x$  to  $A$  ([Definition 1.28](#)) need not equal  $d(x, a)$  for any  $a \in A$  whatsoever. However, if  $A$  is actually compact, then it is true

$$d(x, A) = d(x, a)$$

for some  $a \in A$ ; this follows from [Corollary 4.27](#) and the continuity of the map

$$\begin{aligned} A &\rightarrow \mathbb{R} \\ a &\mapsto d(x, a) \end{aligned}$$

[see [Exercise 1.81 \(a\)](#)].

## Products of compact spaces

Tychonoff Theorem

The proof that compactness is preserved under the formation of products is considerably more difficult than the proof that it is preserved under the formation of quotients.

As an aid in the proof, the following terminology will be suggestive.

def:slice-tube

**4.28 Definition.** Let  $X$  and  $Y$  be topological spaces and let  $x \in X$ . Then the **slice** of  $X \times Y$  **through**  $x$  is the subset  $\{x\} \times Y$ .

A **tube around** this slice is a subset of  $X \times Y$  having the form  $U \times Y$  for some open neighborhood  $U$  of  $x$  in  $X$ .

Similarly one may define the slice of  $X \times Y$  through a point of  $Y$  and then a tube around such a slice.

Although by definition such a tube is open in the product space, nonetheless for emphasis we shall sometimes redundantly refer to it as an “open tube.”

The slice  $\{x\} \times Y$  through a point  $x$  in  $X$  is just the *fiber*  $p^{-1}(x)$  over  $x$  for the first projection  $p: X \times Y \rightarrow X$ . (Such a slice is exactly the same as what earlier, in the context of totally ordered sets, we described as a “stalk”—see [Example 0.70](#).) Moreover, a tube  $U \times Y$  around this slice is just the inverse image  $p^{-1}(U)$  of  $U$  under the projection  $p$ .

For example, take  $X = \mathbb{R}$  and  $Y = [0, 1]$ . Then the slice of  $X$  through a point  $x \in [0, 1]$  is just the closed vertical line segment joining  $\langle x, 0 \rangle$  to  $\langle x, 1 \rangle$ ; for any  $a$  and  $b$  with  $a < x < b$ , the rectangular region  $]a, b[ \times [0, 1]$  is a tube around that slice. (See [Figure 4.6](#).)

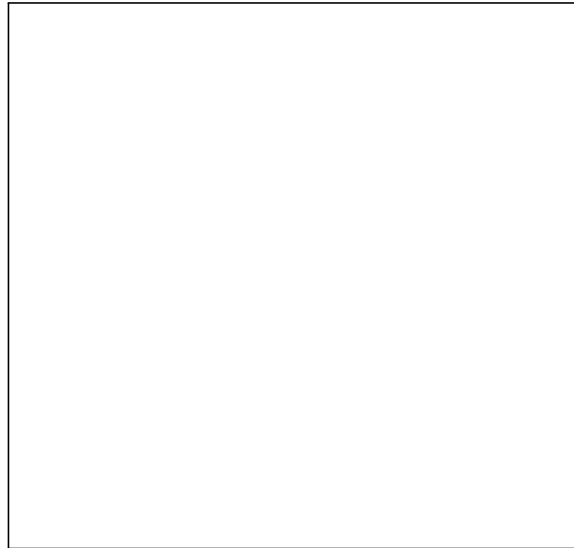
Figure 4.6: A tube around a slice in  $\mathbb{R} \times [0, 1]$ .

fig:tube-slice-R-times-I

The following lemma will be used below in the proof that the product of two compact spaces is compact ([Theorem 4.31](#)).

lem:slice-homeomorph

**4.29 Lemma.** *Let  $X$  and  $Y$  be topological spaces. Then for each  $x \in X$ , the slice  $\{x\} \times Y$  of  $X \times Y$  through  $x$  is homeomorphic to  $Y$ .*

**Proof.** Let  $q: X \times Y \rightarrow Y: \langle x, y \rangle \mapsto y$  be the second projection. Now  $q$  is a continuous open surjection. Then for each  $x \in X$ , the restriction  $q|_{\{x\} \times Y}: \{x\} \times Y \rightarrow Y$  to the slice through  $x$ , which is injective, is a homeomorphism.  $\square$

thm:tube-lemma

**4.30 The Tube Lemma.** *Let  $X$  be a topological space, let  $Y$  be a compact space, and let  $x \in X$ . Let  $W$  be a neighborhood of the slice  $\{x\} \times Y$  through  $x$ . Then there is a tube  $U \times Y$  around the slice  $\{x\} \times Y$  such that  $U \times Y \subset W$ .*

**Proof.** Since  $W$  is a neighborhood of the slice  $\{x\} \times Y$ , for each  $y \in Y$  it is a neighborhood of the point  $\langle x, y \rangle$  in  $X \times Y$ , and so there are open sets  $U_y$  and  $V_y$  in  $X$  and  $Y$ , respectively, with

$$U_y \times V_y \subset W.$$

The collection  $\{V_y : y \in Y\}$  is an open cover of  $Y$ . Since  $Y$  is compact, there is a finite subset  $F$  of  $Y$  such that the subcollection  $\{V_y : y \in F\}$  is still a cover of  $Y$ .

Define

$$U = \bigcap_{y \in F} U_y.$$

Since  $F$  is finite, the set  $U$  is open in  $X$ . Moreover, since  $x \in U$ ,

$$\{x\} \times Y \subset U \times Y = U \times \bigcup_{y \in F} V_y = \bigcup_{y \in F} (U \times V_y) \subset W. \quad \square$$

thm:Tychonoff-for-2

**4.31 Theorem (“Tychonoff-for-two”).** *The product of two compact spaces is itself compact.*

**Proof.** Let  $X$  and  $Y$  be compact spaces. Let  $\mathcal{W}$  be an arbitrary open cover of  $X \times Y$ .

Let  $x \in X$ . Then  $\mathcal{W}$  covers the slice  $\{x\} \times Y$  through  $x$ . Now this slice is compact, being homeomorphic to the compact space  $Y$ , so there is some finite subcollection  $\mathcal{W}_x$  of  $\mathcal{W}$  that is still a cover of  $\{x\} \times Y$ . By The Tube Lemma (4.30), there is an open neighborhood  $U_x$  of  $x$  in  $X$  such that  $\{x\} \times Y \subset U_x \times Y \subset \bigcup \mathcal{W}_x$ .

Since  $X$  is compact, there is a finite subset  $F$  of  $X$  such that  $X = \bigcup_{x \in F} U_x$ . Then  $\bigcup_{x \in F} \mathcal{W}_x$  is a subcollection of  $\mathcal{W}$  that covers the entire product space  $X \times Y$ .  $\square$

The preceding theorem generalizes from just two factors to any finite number of factors.

thm:tychonoff-finite-family

**4.32 Theorem (Tychonoff Product Theorem—finite case).** *The product of a finite family of nonempty topological spaces is compact if and only if each of the individual spaces is compact.*

**Proof.** Let  $\langle X_i : i = 1, 2, \dots, n \rangle$  be a finite family of nonempty topological spaces.

Assume first that the product  $\times_{i=1}^n X_i$  is compact. For each  $j = 1, 2, \dots, n$ , the  $j$ th projection

$$p_j: \bigtimes_{i=1}^n X_i \rightarrow X_j$$

is a continuous surjection, and so by Theorem 4.21 the space  $X_j$  is compact.



The converse is proved by induction on  $n$  using “Tychonoff-for-two” (Theorem 4.31) and the relation

$$\bigtimes_{i=1}^n X_i \cong \left( \bigtimes_{i=1}^{n-1} X_i \right) \times X_n$$

(see Proposition 3.66).  $\square$

In the next section we shall prove the analog of Theorem 4.32 for the product of a sequences of spaces under the additional assumption that each of the spaces is metrizable (Theorem 4.61). Both of these results—for a finite family of topological spaces and for a denumerable family of metrizable spaces—are particular cases of the following general result.

**4.33 Tychonoff Product Theorem.** *The product of an arbitrary family of nonempty compact topological spaces is compact if and only if each of the factor spaces is compact.*

This general Tychonoff Product Theorem is often called simply “The Tychonoff Theorem.” The “only if” part follows from Theorem 4.21, just as in the case of a finite family of spaces. Proving the “if” part is not at all simple, in that it involves Zorn’s Lemma (0.115) or another maximal principle. With one exception, the general theorem will not be used until Chapter 6 (Embedding). (The exception is in the proof of Theorem 4.77, concerning the product of an arbitrary family of *locally* compact spaces.) Accordingly, *proof of the “if” part is deferred until Section 6.1.*

The analog of the Tychonoff Product Theorem for the box topology (Exercise 3.140), instead of the product topology, fails: see Exercise 43.

Using Theorem 4.32 allows us to generalize the Heine–Borel–Lebesgue Theorem (4.15) to  $n$ -dimensional Euclidean space.

**4.34 Theorem (compact subspaces of Euclidean space).** *Let  $d$  be the Euclidean metric on  $\mathbb{R}^n$ . Then a subset of  $\mathbb{R}^n$  is compact if and only if it is both closed in  $\mathbb{R}^n$  and  $d$ -bounded.*

**Proof.** Let  $K \subset \mathbb{R}^n$ . If  $K$  is compact, then  $K$  is closed in  $\mathbb{R}^n$  by Theorem 4.12 and is  $d$ -bounded by Corollary 4.13.

Conversely, assume that  $K$  is both closed in  $\mathbb{R}^n$  and  $d$ -bounded. Since  $K$  is  $d$ -bounded, it is contained in some  $d$ -disk and hence in some cube

$$E = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n].$$

By Examples 4.6 (1) each of the closed intervals  $[a_1, b_1], [a_2, b_2], \dots, [a_n, b_n]$  in  $\mathbb{R}$  is compact, and so by Theorem 4.32 their product  $E$  is a compact subset of  $\mathbb{R}^n$ . However,  $K$  is a closed subset of  $E$ , and so  $K$  is compact by Theorem 4.11.  $\square$

**4.35 Examples.** (1) Because the Euclidean metric is not bounded, *Euclidean  $n$ -space  $\mathbb{R}^n$  is not compact.*

(2) The  $n$ -cube

$$I^n = [0, 1]^n$$

is compact.

Tychonoff Theorem  
box topology  
Tychonoff Theorem  
Heine-Borel-Lebesgue Theorem@He  
n-cube@\$n\$-cube!compact space@

~~ex:n-disk-is-cpt~~ (3) The  $n$ -disk

~~n-cell@\$n\$-cell!compact space@as compact space~~

$$D_n = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$$

~~n-cell@\$n\$-cell!compact space@as compact space~~

~~n-sphere@\$n\$-sphere!compact space@as compact space~~

~~ex:n-cell-is-cpt~~

~~n-torus@\$n\$-torus!compact space@as compact space~~

is compact. By contrast, the  $n$ -ball  $B_n = \{x \in \mathbb{R}^n : \|x\| < 1\}$  is noncompact.

(4) More generally than (2) and (3), any  $n$ -cell is compact.

~~ex:n-sphere-is-cpt~~ (5) The  $n$ -sphere

~~project n-space@projective \$n\$-space!compact space@as compact space~~

$$S_n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$$

~~Klein bottle!compact space@as compact space~~

~~Möbius strip@\$M\$-obius strip!compact space@as compact space~~

~~\$M\$-obius strip!compact space@as compact space~~

is compact. In particular, the unit circle

$$S_1 = \{x \in \mathbb{R}^2 : \|x\| = 1\}$$

is compact.

~~ex:n-torus-is-cpt~~ (6) The  $n$ -torus

$$T_n = (S_1)^n$$

is compact. In particular, the 2-torus

$$T_2 = S_1 \times S_1$$

is compact.

~~ex:projective-space-is-cpt~~ (7) The  $n$ -dimensional real projective space  $\mathbb{RP}_n$  is compact, for according to [Examples 3.81 \(7\)](#) it is a quotient space of the  $n$ -sphere  $S_n$ .

~~ex:klein-bottle-and-Möbius-strip-are-cpt~~ (8) The Klein bottle [[Examples 3.81 \(6\)](#)] and the Möbius strip [[Examples 3.81 \(5\)](#)] are compact, each being a quotient space of the square  $I^2$ .  $\diamond$

### Separation in quotients of compact spaces

~~subsec:sep-quot-cpt~~

Recall that, in general, a quotient of a  $T_2$ -space need not be a  $T_2$ -space: see [Examples 3.84 \(1\)](#). In fact, a quotient of a compact  $T_2$ -space need not be a  $T_2$ -space.

~~ex:non-T1-quot-of-cpt-T2~~ **4.36 Example.** The closed interval  $[-1, 1]$  is a compact  $T_2$ -space. Form the three-point quotient space  $[-1, 1]/\sim$  obtained by collapsing the interior of this closed interval to a point. Of course this quotient space is compact, but it is *not* a  $T_2$ -space. In fact, it is not even a  $T_1$ -space because the complement of the singleton  $\{[-1, 1]\}$  in  $[-1, 1]/\sim$  is the two-point subset  $\{-1, 1\}$ , which is not open.  $\diamond$

Recall also, from [Proposition 3.89](#), that a necessary condition for the quotient  $X/\sim$  of a topological space  $X$  to be a  $T_2$ -space is that the graph of the equivalence relation  $\sim$  be closed in the product  $X \times X$ . We are going to prove a theorem showing that this condition is also sufficient—provided that  $X$  is a compact  $T_2$ -space.

~~ex:fibers-cpt-for-closed-surj-on-cpt-T2~~ **4.37 Lemma.** If  $f: X \rightarrow Y$  is a closed continuous surjection whose domain is a compact  $T_2$ -space, then the fiber  $f^{-1}(y)$  over each point  $y$  of  $Y$  is compact.

**Proof.** Let  $y \in Y$  be arbitrary. Since  $f$  is surjective, there is some  $x \in X$  with  $f(x) = y$ . Since  $X$  is a  $T_2$ -space, the singleton  $\{x\}$  in  $X$  is closed in  $X$ , and so its image  $f(\{x\})$  is closed in  $Y$ . By continuity of  $f$  the inverse image  $f^{-1}(y) = f^{-1}(f(\{x\}))$  is closed in  $X$ . Since  $X$  is compact, the fiber  $f^{-1}(y)$  is compact.  $\square$

**4.38 Proposition.** Let  $f: X \rightarrow Y$  be a closed continuous surjection whose domain  $X$  is a compact  $T_2$ -space. Then the codomain  $Y$  is also a  $T_2$ -space.

**Proof.** Let  $v$  and  $w$  be arbitrary distinct points in  $Y$ . According to Lemma 4.37, the fibers  $f^{-1}(v)$  and  $f^{-1}(w)$  are compact. From Theorem 4.19, these fibers have disjoint open neighborhoods  $V$  and  $W$  in  $X$ , respectively. Then the sets

$$Y \setminus f(X \setminus V) = \{y \in Y : f^{-1}(y) \subset V\}$$

and

$$Y \setminus f(X \setminus W) = \{y \in Y : f^{-1}(y) \subset W\}$$

are disjoint open neighborhoods of  $v$  and  $w$  in  $Y$ .  $\square$

Here is the promised theorem.

**4.39 Theorem (closed-graph criterion for a  $T_2$  quotient).** Let  $\sim$  be an equivalence relation on a compact  $T_2$ -space. Then the following statements are equivalent:

- (i) The quotient space  $X/\sim$  is a  $T_2$ -space.
- (ii) The graph of  $\sim$  is closed in  $X \times X$ .
- (iii) The quotient map  $X \rightarrow X/\sim$  is a closed map.

**Proof.** That (i) implies (ii) is Proposition 3.89. And that (iii) implies (i) is just Proposition 4.38.

(ii)  $\implies$  (iii): Assume (ii). Let  $E$  be an arbitrary closed subset of  $X$ . The image  $q(E)$  under the quotient map  $q: X \rightarrow X/\sim$  will be closed if  $q^{-1}(q(E))$  is closed in  $X$ . Now

$$q^{-1}(q(E)) = p(\text{graph}(\sim) \cap (X \times E)),$$

where  $p: X \times X \rightarrow X$  is the first projection. Moreover,  $p$  is a closed map because  $X$  is compact and  $T_2$ . Since both  $\text{graph}(\sim)$  and  $X \times E$  are closed in  $X \times X$ , so is  $p(\text{graph}(\sim) \cap (X \times E))$ . Thus  $q^{-1}(q(E))$  is closed in  $X$ , and consequently  $q(E)$  is closed in  $X/\sim$ .  $\square$

**4.40 Examples.** (1) We use the closed-graph criterion for a  $T_2$  quotient of the preceding theorem to show that the Möbius strip is a  $T_2$ -space. Recall from Examples 3.81 (5), that the Möbius strip is the quotient  $I \times I/\sim$  where  $\sim$  is given by

$$\begin{aligned} \langle t, s \rangle \sim \langle u, v \rangle &\iff \langle t, s \rangle = \langle u, v \rangle \\ &\text{or } (t = 0, u = 1, v = 1 - s) \\ &\text{or } (t = 1, u = 0, v = 1 - s). \end{aligned}$$

Then

$$\text{graph}(\sim) = \Delta_{I \times I} \cup E \cup E'$$

where

$$\begin{aligned} E &= \{ \langle \langle 0, s \rangle, \langle 1, 1 - s \rangle \rangle : s \in I \}, \\ E' &= \{ \langle \langle 1, s \rangle, \langle 0, 1 - s \rangle \rangle : s \in I \}. \end{aligned}$$

Since  $I \times I$  is a  $T_2$ -space, the diagonal  $\Delta_{I \times I}$  is closed there. Next,  $E$  is the image in the  $T_2$ -space  $I \times I$  of the compact space  $I$  under the continuous map  $s \mapsto \langle \langle 0, s \rangle, \langle 1, 1 - s \rangle \rangle$ , whence  $E$  is closed in  $I \times I$ . Similarly  $E'$  is closed in  $I \times I$ . Thus  $\text{graph}(\sim)$  is closed in  $I \times I$ .

- ex:RPn-T2 (2) We use the [closed-graph criterion for a  \$T\_2\$  quotient](#) of the preceding theorem to show that real projective  $n$ -space  $\mathbb{RP}_n$  is a  $T_2$ -space. Recall from [Examples 3.81 \(7\)](#) that  $\mathbb{RP}_n$  is the quotient  $S_n/\sim$  where

$$x \sim y \iff x = y \text{ or } x = -y.$$

Then

$$\text{graph}(\sim) = \Delta_{S_n} \cup \Delta_{S_n}^*,$$

where  $\Delta_{S_n}^* = \{\langle x, -x \rangle : x \in S_n\}$ . Since  $S_n$  is a Hausdorff space, already  $\Delta_{S_n}$  is closed in  $S_n \times S_n$ . Further,  $\Delta_{S_n}^*$  is the image of  $\Delta_{S_n}$  under the continuous map  $\langle x, x \rangle \mapsto \langle x, -x \rangle$ ; since  $S_n \times S_n$  is compact  $T_2$ , that image is also closed in  $S_n \times S_n$ . Hence  $\text{graph}(\sim)$  is closed in  $S_n \times S_n$ .  $\diamond$

For an “abstract” quotient map  $f: X \rightarrow Y$  (in the sense of [Definition 3.78](#)), [Theorem 4.39](#) may be restated in terms of the equivalence kernel of  $f$  [[Examples 0.90 \(7\)](#)] as follows.

thm:quot-T2-equiv-ker-closed-map

**4.41 Theorem.** *Let  $f: X \rightarrow Y$  be a quotient map whose domain  $X$  is a compact  $T_2$  space. The the following statements are equivalent:*

thm-condition:quot-via-equiv-ker-T2

(i) *The quotient space  $Y$  is a  $T_2$ -space.*

n-condition:quot-via-equiv-ker-closed

(ii) *The equivalence kernel  $\{\langle x, u \rangle \in X \times X : f(x) = f(u)\}$  of  $f$  is closed in  $X \times X$ .*

dition:quot-via-equiv-ker-closed-map

(iii) *The quotient map  $f$  is a closed map.*

### Compactness and convergence

To conclude our discussion of compact spaces in general, we characterize compactness in terms of convergence.

thm:cpt-via-nets

**4.42 Theorem (compactness and net convergence).** *A topological space is compact if and only if each net in the space clusters in it.*

**Proof.** We shall use the [finite intersection property criterion](#) ([Theorem 4.9](#)). Assume first that the topological space  $X$  is compact. Let  $(x_i)_{i \in I}$  be an arbitrary net in  $X$ . Form the collection  $\{E_i : i \in I\}$  given by

$$E_i = \text{cls } A_i, \\ A_i = \{x_j : j \in I, i \leq j\}$$

for each  $i \in I$ . If  $J$  is any nonempty finite subset of  $I$ , then since  $I$  is directed, there is some  $k \in I$  with  $j \leq k$  for all  $j \in J$ , and so

$$\emptyset \neq A_k \subset \bigcap_{j \in J} A_j \subset \bigcap_{j \in J} E_j.$$

Thus  $\{E_i : i \in I\}$  has the finite intersection property. Since  $X$  is compact, by the [finite intersection property criterion](#) there exists some

$$x \in \bigcap_{i \in I} E_i.$$

We show that  $(x_i)_{i \in I}$  clusters at the point  $x$ . Let  $V$  be an arbitrary neighborhood of  $x$  and let  $i \in I$ . Since  $x \in E_i$ , then  $V \cap A_i \neq \emptyset$ . Hence  $x_j \in V$  for some  $j \geq i$ .

Conversely, assume that each net in the topological space  $X$  clusters in  $X$ . Let  $\mathcal{E}$  be an arbitrary collection of closed subsets of  $X$  having the finite intersection property. Construct a net  $(x_i)_{i \in I}$  as follows. The index set  $I$  is defined by

$$I = \{ \langle \mathcal{F}, x \rangle : \mathcal{F} \subset \mathcal{E}, \mathcal{F} \text{ is finite, and } x \in \bigcap \mathcal{F} \}.$$

The direction  $\leq$  of  $I$  is defined by

$$\langle \mathcal{F}_1, x_1 \rangle \leq \langle \mathcal{F}_2, x_2 \rangle \iff \mathcal{F}_1 \subset \mathcal{F}_2.$$

Clearly  $\leq$  is reflexive on  $I$  and transitive. To see that  $\leq$  actually directs  $I$ , note that if  $\langle \mathcal{F}_1, x_1 \rangle, \langle \mathcal{F}_2, x_2 \rangle \in I$ , then  $\mathcal{F}_1 \cup \mathcal{F}_2$  is a nonempty finite subset of  $\mathcal{E}$ , and so there is some  $x \in \bigcap (\mathcal{F}_1 \cup \mathcal{F}_2)$ ; then  $\langle \mathcal{F}_1 \cup \mathcal{F}_2, x \rangle$  is an element of the index set  $I$  such that

$$\langle \mathcal{F}_1, x_1 \rangle \leq \langle \mathcal{F}_1 \cup \mathcal{F}_2, x \rangle, \quad \langle \mathcal{F}_2, x_2 \rangle \leq \langle \mathcal{F}_1 \cup \mathcal{F}_2, x \rangle,$$

as desired.

Now for each index

$$i = \langle \mathcal{F}, x \rangle \in I,$$

define the point  $x_i$  of  $X$  by

$$x_i = x.$$

By assumption, the net  $(x_i)_{i \in I}$  clusters at some point  $x$  in  $X$ . We show that  $x \in \bigcap \mathcal{E}$ . Let  $E$  be an arbitrary member of  $\mathcal{E}$ . To show that  $x$  belongs to the closed set  $E$ , it is enough to show that each neighborhood of  $x$  intersects  $E$ . Let  $V$  be an arbitrary neighborhood of  $x$ . Since  $E \in \mathcal{E}$ , there is some  $y \in E$ . Then  $\langle \{E\}, y \rangle \in I$ , and so there exists some index  $\langle \mathcal{F}, z \rangle \in I$  with

$$\langle \{E\}, y \rangle \leq \langle \mathcal{F}, z \rangle, \quad z \in V.$$

But  $E \in \mathcal{F}$ , and so

$$z \in \bigcap \mathcal{F} \subset E.$$

Thus  $z \in V \cap E$ .  $\square$

In view of [Theorem 3.120](#), the preceding [Theorem 4.42](#) has the equivalent formulation: *a topological space  $X$  is compact if and only each net in  $X$  has some convergent subnet.*

There is a counterpart to [Theorem 4.42](#) for filters.

thm:cpt-via-filters

**4.43 Theorem (compactness and filter convergence).** *A topological space is compact if and only if each filter on the space clusters in it.*

The proof is left as an exercise: see [Exercise 39](#).

In view of [Theorem 3.153](#), the preceding [Theorem 4.43](#) has the equivalent formulation: *a topological space is compact if and only each filter on the space has a convergent refinement.*

The compactness criterion of [Theorem 4.43](#) may be sharpened, as will be seen in the subsection “Proof of the Tychonoff Theorem using ultrafilters” of [Section 6.1](#): instead of clustering of all filters, it suffices to have just convergence of all *ultrafilters*—filters than which there are no finer filters.

### EXERCISES FOR SECTION 4.1

extended real line  
countable-complement topology  
Sorgenfrey line  
line with two origins  
half-disk space  
tangent disk space  
order topology!  
compact space!  
lexicographically ordered square!  
order topology!  
compact space!  
lexicographic ordering  
lexicographically ordered square!  
Lindelöf space!  
compact space!  
base!  
compact space!

1. Show that the set  $A$  constructed in the proof of [Proposition 4.1](#) is both open and closed in  $[0, 1]$ .

[*Note:* In [Chapter 5](#) (Connectedness) it will be proved that the space  $[0, 1]$  is “connected” in the sense that it does not have a partition into two proper subsets each of which is open in the space, and so the only nonempty subset of  $[0, 1]$  that is both open and closed in this space is the entire space  $[0, 1]$ . Hence compactness of  $[0, 1]$  is a consequence of its connectedness. Conversely, connectedness of  $[0, 1]$  can be deduced from its compactness—see [Exercise 5.34](#).]

2. Working directly with the definition of compactness, show that the extended real line  $\widehat{\mathbb{R}}$  is compact.

3. Give examples of:

- (a) A compact space that is not pseudometrizable ([Exercise 2.11](#)).
- (b) A compact Hausdorff space that is not metrizable.
- (c) A compact  $T_1$ -space [Definition 2.94 \(1\)](#) that is not a  $T_2$ -space.

4. Determine which of the following spaces are compact:

- (a) An uncountable set provided with its countable-complement topology [[Exercise 2.7](#)].
- (b) The Sorgenfrey line [[Examples 2.20 \(1\)](#)].
- (c) The interval  $[0, 1]$  with its relative topology in the Sorgenfrey line.
- (d) The line with two origins [[Examples 2.20 \(3\)](#)].
- (e) The half-disk space [[Examples 2.20 \(3\)](#)].
- (f) The tangent disk space [[Exercise 2.37](#)].

5. Fill in the missing details of the proof for [Examples 4.6 \(9\)](#): an order-complete totally ordered set having both a least element and a greatest element is compact in its order topology.

6. Let  $X$  be a totally ordered set provided with its order topology [[Examples 2.72 \(1\)](#)]. [Examples 4.6 \(9\)](#) gave a sufficient condition for  $X$  to be compact. Is this condition also necessary for compactness? if not, can it be modified to become necessary?

(*Hint:* : Which properties of the usual ordering on  $[0, 1]$  were used in the proof of [Proposition 4.1](#)?)

7. Show that the lexicographically ordered square  $I_{\text{lex}}^2$  ([Exercise 2.90](#)) is compact.

8. (*Continuation of [Exercise 2.115](#).*)

- (a) Show that the lexicographically ordered square  $I_{\text{lex}}^2$  ([Exercise 2.90](#)) is a Lindelöf space.
- (b) Show that  $[0, 1] \times ]0, 1[$  is an *open* subspace of  $I_{\text{lex}}^2$  that is not a Lindelöf space.

9. Prove: If a topological space  $X$  has some base  $\mathcal{B}$  with the property that each cover of  $X$  by members of  $\mathcal{B}$  contains a finite cover of  $X$ , then  $X$  must be compact.

*Note:* Through use of Zorn’s Lemma, this result may be strengthened to require only a *subbase* with the stated property: see [Theorem 6.1](#).

10. (a) Prove that the union of finitely many compact subsets of a topological space is itself compact. Cartesian sum!family of spaces@of a  
 (b) What can be said about compactness of the intersection of two compact subsets of a topological space? locally finite collection  
relatively compact set  
precompact set
11. When is the Cartesian sum (Exercise 3.106) of a family of topological spaces compact? Cantor set  
Cantor set!continuous image@and c
12. A collection  $\mathcal{U}$  of subsets of a topological space  $X$  is said to be **locally finite** when each point of  $X$  has some neighborhood that intersects only finitely many members of  $\mathcal{U}$ . [In other words, the collection  $\mathcal{U}$  is locally finite exactly when the family  $\langle U \rangle_{U \in \mathcal{U}}$  is locally finite in the sense of Exercise 2.63.]
- (a) Construct a locally finite cover of  $\mathbb{R}$  having infinitely many members.  
 (b) Show that each locally finite cover of a compact space is necessarily finite.
13. A subset  $A$  of a topological space  $X$  is said to be **relatively compact (in  $X$ )** when it is contained in some compact subset of  $X$ , and  $A$  is said to be **precompact** when  $\text{cls } A$  is compact. Clearly  $A$  will be relatively compact if it is relatively compact. Under what circumstances does the converse hold?
14. Characterize the endpoints of the Cantor set  $K$  in terms of their ternary expansions. How many endpoints does  $K$  have? How many nonendpoints?
- (a) Show that  $1/4 \in K$  but is not an endpoint of  $K$ .  
 (b) Find three more rational numbers belonging to  $K$  that are not endpoints of it.
16. Show that  $K$  is nowhere dense in  $[0, 1]$  (see Definition 2.50).
17. Show that the Cantor set  $K = \bigcap_{n \in \mathbb{N}} (f^n)^{-1}([0, 1/3] \cup [2/3, 1])$  where, for each  $n$ , the function  $f^n$  is the  $n$ th iterate of the function  $f: [0, 1] \rightarrow [0, 1]$  given by
- $$f(x) = \begin{cases} 3x & \text{if } 0 \leq x \leq 1/3, \\ 0 & \text{if } 1/3 < x < 2/3, \\ 3x - 2 & \text{if } 2/3 \leq x \leq 1. \end{cases}$$
18. Given a positive integer  $k$ , construct a continuous map from  $K$  onto the product space  $[0, 1]^k$ .  
 [Hint: Show first that  $(\{0, 2\}^{\mathbb{N}^*})^k \cong \{0, 2\}^{\mathbb{N}^*}$ .]
- (a) Show that  $3/2$  is the sum of two elements of the Cantor set  $K$ .  
 (b) Show more generally that each real number in the closed interval  $[0, 2]$  is the sum of two (not necessarily distinct) elements of  $K$ .
20. Suppose in the construction of the Cantor set we still start with the unit interval but now we remove at each step the open middle-*fifth* of each subinterval instead of the middle-third. Let  $F_n$  denote the set obtained at the  $n$ th step and let  $F = \bigcap_{n \in \mathbb{N}} F_n$ .
- (a) Carry out the first three steps, describing and sketching the resulting sets  $F_1, F_2, F_3$  at those steps. In particular, what are the endpoints of each?  
 (b) How many subintervals constitute  $F_n$  and what is the length of each? What is the total length of the intervals constituting  $F_n$ ?  
 (c) Characterize those numbers that belong to  $F$ , and those number in  $F$  that are endpoints of any  $F_n$ , in terms of an expansion similar to ternary expansion.

Cantor function

Cantor set

regular space

Urysohn Metrization Theorem

prob:Cantor-fn

metrizable space

convex set

diameter

distance!between subsets

(d) Which topological properties of the Cantor set discussed in [Example 4.16](#) are shared by  $F$ , the intersection of the sets obtained at all the steps?

(e) Repeat [Exercise 20](#) but with removing the middle *half* of each subinterval.

21. Refer to the construction of the Cantor function in [Example 4.17](#).

(a) Verify the relations (\*) and (\*\*) for the sequence  $\langle \kappa_n \rangle_{n=1,2,3,\dots}$  of approximations to  $\kappa$ .

(b) Justify the assertions that  $\kappa$  is increasing and maps the Cantor set  $K$  onto  $[0, 1]$ .

22. Let  $\langle K_n \rangle_{n \in \mathbb{N}}$  be a decreasing sequence of nonempty compact subsets of a Hausdorff space. If  $U$  is an open set in  $X$  with  $\bigcap_{n=0}^{\infty} K_n \subset U$ , show that  $K_n \subset U$  for some  $n$ .

23. Let  $X$  be a regular space [[Definition 2.94 \(3\)](#)]. If  $A$  is a compact subset of  $X$  and  $B$  is a closed subset of  $X$  disjoint from  $A$ , show that  $A$  and  $B$  have disjoint neighborhoods in  $X$ .

24. Let  $\sim$  be the equivalence relation on the compact Hausdorff space  $[0, 1]$  from [Exercise 3.179 \(c\)](#). Show that each of the points  $\{0\}$  and  $\{1\}$  of the quotient space  $[0, 1]/\sim$  has a compact  $T_2$  neighborhood that is not closed in  $Y$ .

prob:urysohn-metrizable

25. Prove the following special case of Urysohn's Lemma ([6.26](#)): If  $A$  and  $B$  are disjoint closed sets in a metrizable space  $X$ , then there is a continuous function  $f: X \rightarrow \mathbb{R}$  with  $f(a) = 0$  for all  $a \in A$ ,  $f(b) = 1$  for all  $b \in B$ , and  $0 \leq f(x) \leq 1$  for all  $x \in X$ .

[Hint: If  $d$  induces the topology of  $X$ , consider the function defined by  $f(x) = d(x, A)/(d(x, A) + d(x, B))$ .]

prob:cptT2-max-cpt-min-T2

26. (a) Given topologies  $\mathcal{T}$  and  $\mathcal{S}$  on the same set  $X$  with  $\mathcal{T}$  finer than  $\mathcal{S}$ , prove: If  $\langle X, \mathcal{T} \rangle$  is compact, then  $\langle X, \mathcal{S} \rangle$  is compact; and if  $\langle X, \mathcal{S} \rangle$  is Hausdorff, then  $\langle X, \mathcal{T} \rangle$  is Hausdorff. Moreover, if  $\langle X, \mathcal{T} \rangle$  is compact and  $\langle X, \mathcal{S} \rangle$  is a Hausdorff space, then  $\mathcal{T} = \mathcal{S}$ . (This is a restatement of [Proposition 4.26](#).)

(b) Can a set have two different topologies each of which make it a compact Hausdorff space?

ob-part:cpt-via-cont-f-max-and-min

27. (a) Must a subset  $K$  of  $\mathbb{R}$  be compact if every continuous function  $f: K \rightarrow \mathbb{R}$  is bounded for the Euclidean metric and attains a maximum value on  $K$ .

(b) Same question as in (a) but for a subset  $K$  of  $\mathbb{R}^n$  where  $n > 1$ .

28. Let  $K$  be a nonempty subset of the plane  $\mathbb{R}^2$  and provide  $\mathbb{R}^2$  with its Euclidean metric  $d$ . If  $x \in \mathbb{R}^2$ , then a point  $y \in K$  is called a “nearest point in  $K$  to  $x$ ” when  $d(x, K) = d(x, y)$ . Prove that a necessary and sufficient condition for  $K$  to be both closed and convex is that for each  $x \in \mathbb{R}^2$ , there exists a *unique* closest point in  $K$  to  $x$ .

[Hint: To establish necessity, you may want to use continuity of the function  $y \mapsto d(x, y)$  on  $K$ .]

prob:diam-and-dist-cpt

29. Let  $K$  be a nonempty compact subset of a metric space  $\langle X, d \rangle$ .

(a) Show that  $\text{diam } K = d(x, y)$  for some  $x, y \in K$ .

prob-part:diam-cpt-set-attained

(b) If  $L$  is another nonempty compact subset of  $X$ , show that there are points  $x \in K$  and  $y \in L$  such that  $d(x, y) = d(K, L)$ , the distance from  $K$  to  $L$  in the sense of [Exercise 1.43](#).

part:dist-between-cpt-sets-attained

(c) Does the conclusion of (b) still hold if  $L$  is only assumed to be a nonempty closed subset of  $X$ ?



- per-semicont-fn-on-cpt-attains-max 30. Let  $f: X \rightarrow Y$  be a continuous injection from a compact space  $X$  to a Hausdorff space  $Y$  whose image is dense in  $Y$ . Prove that  $f$  is in fact a homeomorphism of  $X$  with  $Y$ . semicontinuous map  
upper semicontinuous map  
lower semicontinuous map  
maximum!and semicontinuous func  
minimum!and semicontinuous func
- prob:adjoin-pt-at-infty-to-lattice-pt 31. (a) If  $X$  is a nonempty compact space, show that every function  $f: X \rightarrow \mathbb{R}$  that is upper semicontinuous ([Exercise 3.43](#)) attains a maximum value on  $X$ .  
(b) Prove and analog of (a) for lower semicontinuous functions. slice  
tube
32. Let  $X = (\mathbb{N} \times \mathbb{N}) \cup \{\infty\}$  where  $\infty$  is some object that is not a member of  $\mathbb{N} \times \mathbb{N}$  (for example,  $\infty = \mathbb{N}$ ). compact set!nets@and nets  
net!compact space@in compact spa  
compact space!filters@and filters  
filter!compact space@in compact sp
- (a) Let  $\mathcal{T}$  be the collection of all subsets  $U$  of  $X$  such that either  $\infty \notin U$  or else for almost all  $m \in \mathbb{N}$ , the set  $U$  contains almost all points of  $\{m\} \times \mathbb{N}$ . Verify that  $\mathcal{T}$  is a topology on  $X$ . (Obviously, then,  $\mathbb{N} \times \mathbb{N}$  is a discrete subspace of  $X$ .) clustering filter
- (b) Show that when  $X$  is endowed with the topology  $\mathcal{T}$ , the resulting topological space is a Hausdorff space in which the only compact subsets are finite subsets. filter!clustering@and clustering  
compact space!filter base@and filter  
filter base!compact space@in compa
33. (a) Sketch a slice of  $\mathbb{R} \times D_2$  through a point  $x \in \mathbb{R}$  together with a tube around this slice.
- (b) Sketch a slice of  $[0, 1] \times D_2$  through a point  $x \in ]0, 1[$  together with a tube around this slice. Do the same thing when  $x = 0$  and then again when  $x = 1$ .
- (c) Sketch a slice of the solid torus  $D_2 \times S_1$  through a point  $x \in B_2$  together with a tube around this slice. Do the same thing when  $x \in S_1$ .
- prob:cover-diag-cpt-times-itself 34. Give a compact space  $X$ , let  $U$  be an open subset of  $X \times X$  containing the diagonal  $\{\langle x, x \rangle : x \in X\}$  of  $X \times X$ . Show that there is a finite open cover  $\{V_1, V_2, \dots, V_n\}$  of  $X$  with  $\bigcup_{i=1}^n (V_i \times V_i) \subset U$ .  
(Note: For an application to the case when  $X$  is metrizable, see [Exercise 88](#).)
35. Let  $X$  be a topological space, let  $Y$  be a compact space, let  $x \in X$ . and let  $\mathcal{U}$  be an open cover of the slice  $\{x\} \times Y$  in  $X \times Y$ . Prove that there there is an open neighborhood  $U$  of  $x$  in  $X$  and a finite subcollection of  $\mathcal{U}$  that is a cover of the tube  $U \times Y$ .
- prob:open-prod-around-prod-2-cpt 36. Let  $K$  and  $L$  be compact subsets of topological spaces  $X$  and  $Y$ , respectively. If  $W$  is a neighborhood of  $K \times L$  in  $X \times Y$ , show that there are open sets  $U$  in  $X$  and  $V$  in  $Y$  such that  $K \times L \subset U \times V \subset W$ .
- more-cpt-T2-quot-via-closed-graph 37. Use the closed-graph criterion for a  $T_2$ quotient ([Theorem 4.39](#)) to show that each of the following quotient spaces is a  $T_2$ -space:
- (a) the Klein bottle [[Examples 3.81 \(6\)](#)].
- (b) the homeomorph  $(I \times I)/\sim$  of the torus as constructed in [Examples 3.81 \(4\)](#).
- prob:net-union-lim-cptQ 38. (a) Let  $(x_i)_{i \in I}$  be a net in a topological space  $X$  that converges to a point  $x \in X$ . Must  $\{x\} \cup \{x_i : i \in I\}$  be compact?
- (b) Must  $\{x\} \cup \{x_i : i \in I\}$  be compact if  $(x_i)_{i \in I}$  just clusters at  $x$ ?
- prob:filter-and-filter-base-cpt 39. (a) Prove [Theorem 4.43](#): A topological space is compact if and only if each filter on the space clusters in it.
- (b) From [Theorem 4.43](#), if a topological space is compact, then each filter base on it clusters in it. Is the converse true, too?
- prob:cpt-via-fb 40. Let  $\mathcal{B}$  be the collection of all subsets of a topological space  $X$  whose complements are compact.

**(a)** Show that  $\mathcal{B}$  is a filter base on  $X$  if and only if  $X$  is not compact.

**(b)** If  $X$  is not compact, must  $\mathcal{B}$  be a filter on  $X$ ?

**41.** In the proof of the converse part of [Theorem 4.42](#), suppose we had constructed a net  $(x_i)_{i \in I}$  from  $\mathcal{E}$  by taking  $I$  to be the set of all nonempty finite subsets of  $\mathcal{E}$ , with  $I$  being directed by inclusion, and by taking  $x_i \in \bigcap_{E \in i} E$  for each  $i \in I$ .

**Closed-graph Theorem**

**(a)** If  $x$  is a point in  $X$  at which  $(x_i)_{i \in I}$  clusters, show that  $x \in \bigcap \mathcal{E}$ .

**continuous map!closed graph!and closed graph**

**(b)** Discuss why such a net need exist.

**closed map!projection!and projection**

**42.** Prove or disprove: If a filter  $\mathcal{F}$  on a compact space clusters at exactly one point  $x$ , then  $\mathcal{F}$  must converge to that point  $x$ .

**closed map!projection!and projection**

**43.** Show that the box product ([Exercise 3.140](#)) of a sequence of copies of the two-point discrete space  $\{0, 1\}$  is not compact by doing one or both of the following:

**closed map**

**(a)** Construct a sequence in that box product that does not cluster there.

**(b)** Construct an open cover of that box product that contains no finite cover.

*Note:* Thus formation of box products does *not* preserve compactness. This is yet another reason that the box topology is not the “natural” topology to put on a product. [For the more fundamental reason, see the note in [Exercise 3.140 \(d\)](#).]

**44.** Prove the **closed-graph theorem**: If  $f: X \rightarrow Y$  is a map from a topological space  $X$  to a compact space  $Y$  whose graph  $\{\langle x, f(x) \rangle : x \in X\}$  is closed in  $X \times Y$ , then  $f$  is continuous.

[*Hint:* Show that if  $(x_i)_{i \in I}$  is a net converging to  $x$  in  $X$ , then  $\langle f(x_i) \rangle_{i \in I}$  converges to  $f(x)$  in  $Y$  by applying [Theorem 4.42](#) to  $\langle f(x_i) \rangle_{i \in I}$ .]

**45.** Prove: If  $Y$  is a compact space, then the first projection  $p: X \times Y \rightarrow X$  is a closed map for every topological space  $X$ .

[*Note:* [Examples 3.16 \(3\)](#) showed that a projection from a product space need not be a closed map.]

**46.** Prove the converse of [Exercise 45](#): A topological space  $Y$  is compact if, for every topological space  $X$ , the first projection  $X \times Y \rightarrow X$  is a closed map.

[*Hint:* Show that an arbitrary net  $\langle y \rangle_{i \in I}$  in  $Y$  clusters. Take  $X = I \cup \{\infty\}$  to be the space whose topology is defined as in [Exercise 3.256](#).]

(*Note:* This result may seem surprising: its conclusion concerns compactness, yet its hypothesis seems to have nothing to do with compactness!)

**47. (a)** Let  $f: X \rightarrow Y$  be a *closed* surjection whose fibers  $f^{-1}(y)$  are compact for all points  $y$  of  $Y$ . (*Note:* For example, according to [Exercise 45](#), the first projection  $p: X \times Y \rightarrow X$  is such a map when the second factor  $Y$  is compact.) Prove that  $X$  is compact if  $Y$  is compact. *Note:* It is *not* assumed here that  $f$  is continuous.

**(b)** Give an example of a continuous surjection  $f: X \rightarrow Y$  whose codomain  $Y$  is compact and whose fibers  $f^{-1}(y)$  are compact for all points  $y$  of  $Y$ , and yet not all inverse images of compact subsets of  $Y$  are compact.

**48.** Let  $f: X \rightarrow Y$  be a *closed* map whose fibers  $f^{-1}(y)$  are compact for all points  $y$  of  $Y$ . Prove that the inverse image  $f^{-1}(K)$  of each compact subset  $K$  of  $Y$  is compact. *Note:* It is *not* assumed here that  $f$  is continuous or that  $f$  is surjective.

(*Hint:* Reduce the situation to the case where  $f$  is surjective and the codomain  $Y$  itself is compact, then apply [Exercise 47](#).)

49. Form the quotient  $D_2//B_2$  of the compact  $T_2$ -space  $D_2$  obtained by collapsing the interior of  $D_2$  to a point. Is this quotient a  $T_2$ -space? perfect map  
perfect map  
proper map
- prob:perfect-map-and-T2 50. A continuous map  $f: X \rightarrow Y$  is said to be **perfect** when  $f$  is a closed surjection and the inverse image  $f^{-1}(y)$  is compact for each  $y \in Y$ . (For example, according to [Exercise 45](#), the first projection  $p: X \times Y \rightarrow X$  is perfect if the second factor  $Y$  of the domain is compact.)
- (a) Show that a continuous map from a compact space onto a Hausdorff space is perfect.
- (b) Must a continuous closed surjection from a Hausdorff space to a compact Hausdorff space be perfect?
- prob-part:perfect-preserves-T2 (c) If  $f: X \rightarrow Y$  is perfect and  $X$  is a Hausdorff space, prove that  $Y$  is also a Hausdorff space.
- (d) Show that a perfect map need not be open.
- prob:perfect-bis 51. (*Continuation of Exercise 50.*) Show that perfect maps preserve second-countability; that is, if  $f: X \rightarrow Y$  is a perfect map whose domain  $X$  is second-countable, then its range  $Y$  is also second-countable.
- [*Note:* Thus perfect maps preserve second-countability. According to [Exercise 50 \(c\)](#), perfect maps preserve being a  $T_2$ -space. In some later exercises you will be asked to show that perfect maps preserve connectedness as well as local compactness. Thus perfect maps preserve a number of important topological properties even though they need not be homeomorphisms.]
- prob:perfect-iff-universally-closed 52. (*Continuation of Exercise 50.*) Let  $f: X \rightarrow Y$  be a continuous surjection. Prove that  $f$  is perfect if and only if, for every topological space  $Z$ , the product map  $f \times \text{id}_Z: X \times Z \rightarrow Y \times Z$  of  $f$  with the identity map of  $Z$  is a closed map.
- prob:proper-map-and-cpt 53. A continuous map  $f: X \rightarrow Y$  is said to be **proper** when the inverse image  $f^{-1}(K)$  of each compact subset  $K$  of  $Y$  is itself compact.
- (a) Show that a continuous map from a compact space to a Hausdorff space is necessarily proper.
- (b) Give an example of a continuous injection whose domain is compact Hausdorff yet which is not proper.
- (c) According to [Exercise 48](#), every continuous map that is perfect ([Exercise 50](#)) is proper. By considering a Fortissimo space ([Exercise 2.111](#)) or otherwise, show that a continuous proper surjection need not be closed—and hence need not be perfect.
- (*Note:* For a fairly general situation where a continuous proper map *is* closed, see [Exercise 56](#).)
- prob-part:cont-map-cpt-to-T2-is-proper 54. (*Continuation of Exercises 50 and 53.*) Which of the continuous functions of [Exercise 3.34](#) are perfect? which are proper?
- prob:polys-and-proper-maps 55. (a) Prove: If  $f: \mathbb{R}^n \setminus A \rightarrow \mathbb{R}^m$  be a continuous map from the complement of a finite (possibly empty) subset  $A$  of  $\mathbb{R}^n$  such that  $\lim_{x \rightarrow \infty} |f(x)| = \infty$  and (if  $A$  is nonempty)  $\lim_{x \rightarrow a} |f(x)| = \infty$  for all  $a \in A$ , then  $f$  is a proper map ([Exercise 53](#)).
- (b) Show that every *nonconstant* polynomial  $\sum_{j=1}^k a_j x^j$  in a single real variable  $x$  and with real coefficients  $a_j$  defines a proper function  $x \mapsto \sum_{j=1}^k a_j x^j$  from  $\mathbb{R}$  to  $\mathbb{R}$ .
- ob-part:noncst-poly-1-var-is-proper

compactly generated space (c) Is the analog of (b) true for nonconstant polynomials in more than one variable?  
 compactly generated space (d) Repeat (b) but for polynomials in a single complex variable with complex coefficients.  
 Sierpinski carpet

Sierpiński, Wacław (e) Let  $p(z)$  and  $q(z)$  be polynomials with complex coefficients in a single complex variable  $z$ , and suppose  $\deg p(z) > \deg q(z)$ . Show that  $z \mapsto p(z)/q(z)$  defines a proper function on  $\mathbb{C} \setminus A$ , where  $A$  is the set of zeros of  $q$  that are not also zeros of  $p$ .

prob:compactly-generated 56. Say that a topological space  $X$  **compactly generated** when an arbitrary subset  $U$  is open in  $X$  exactly when the intersection  $U \cap K$  is open in  $K$  for each compact subset  $K$  of  $X$ .

(a) Show that  $X$  is compactly generated if and only if an arbitrary subset  $E$  of  $X$  is closed whenever  $E \cap K$  is closed in  $K$  for each compact subset  $K$  of  $X$ .

(b) Show that every first-countable space is compactly generated.

(Note: As a consequence, every metrizable space is compactly generated; in particular,  $\mathbb{R}^n$  is compactly generated for every  $n$ . Later we shall study another class of compactly generated topological spaces which also includes all the Euclidean spaces  $\mathbb{R}^n$ —namely, the *locally compact* spaces. For examples of spaces that are *not* compactly generated, see Exercise 57.)

(c) Let  $X$  be a compactly generated space. Show that a map  $f: X \rightarrow Y$  is continuous if and only if the restriction  $f|_K: K \rightarrow Y$  is continuous for every compact subset  $K$  of  $X$ .

per-to-cptly-generated-T2-is-closed

(d) Prove that if  $f: X \rightarrow Y$  is a continuous proper map to a compactly generated Hausdorff space  $Y$ , then  $f$  is a closed map (and hence is a perfect map if it is also surjective).

prob:non-cptly-generated-spaces

57. (Continuation of Exercise 56.) Show that the following spaces are *not* compactly generated:

(a) The space of Exercise 32.

ble-power-of-R-not-cptly-generated

(b) The product space  $\mathbb{R}^I$  when  $I$  is an uncountable set (for example, when  $I = \mathbb{R}$ ).

prob:Urysohn-like-cpt-subsets

58. Let  $X$  be a topological space having the property that, for any two distinct points  $a, b \in X$ , there is a continuous function  $h: X \rightarrow [0, 1]$  with  $h(a) = 0$  and  $h(b) = 1$ .

(a) Let  $B$  be a compact subset of  $X$ . If  $a \in X \setminus B$ , show that there is a continuous function  $g: X \rightarrow [0, 1]$  with  $g(a) = 0$  and  $g(b) = 1$  for all  $b \in B$ .

[Hint: First obtain a continuous function  $g': X \rightarrow [0, 1]$  with  $g'(a) = 0$  and  $g'(b) > 1/2$  for all  $b \in B$ . To do this, note that for  $b \in B$  and for a continuous function  $h: X \rightarrow [0, 1]$  with  $h(b) = 1$ , the set  $h^{-1}([1/2, 1])$  is a neighborhood of  $b$ . You will probably want to use Examples 3.25 (4).]

(b) Prove: If  $A$  and  $B$  are disjoint compact subsets of a topological space  $X$ , then there is a continuous function  $f: X \rightarrow [0, 1]$  such that  $f(a) = 0$  for all  $a \in A$  and  $f(b) = 1$  for all  $b \in B$ .

Exercises 59 and 60 involve analogs of the Cantor set in two- and three-dimensional space.

prob:Sierpinski-carpet

59. The **Sierpinski carpet** is a two-dimensional analog of the Cantor set, formed as follows. Start with the unit square  $S_0 = [0, 1] \times [0, 1]$ . Form a 3-by-3 grid of horizontal and vertical lines in  $S_0$ , with the edges of the grid on the edges of the

square, thereby obtaining 9 congruent smaller “subsquares.” Remove the open central subsquare and denote by  $S_1$  the subspace of  $S_0$  so obtained.

Next, remove the middle-ninth subsquare of each of the remaining 8 subsquares constituting  $S_1$  and denote the result by  $S_2$ . (The sets  $S_1$  and  $S_2$  are shown in Figure 4.7.) Repeat this process, obtaining successively  $S_3, S_4, S_5, \dots$ . The Sierpinski carpet is the intersection  $S = \bigcap_{n \in \mathbb{N}} S_n$ .



(a) The first stage  $S_1$ . (b) The second stage  $S_2$ .

Figure 4.7: The first and second stages in constructing the Sierpinski carpet.

fig:SierpinskiCarpet-levels1and2

- (a) Sketch  $S_3$ . (The next set,  $S_4$ , is shown in Figure 4.8.)

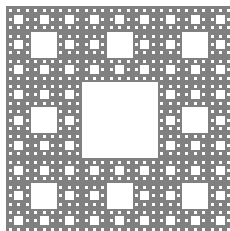


Figure 4.8: The fourth stage  $S_4$  in constructing the Sierpinski carpet.

fig:Sierpinski-carpet-S4

- (b) Which points are corners of the squares of  $S_2$  or  $S_3$ ?
- (c) How many squares constitute  $S_n$  and what are the dimensions of each? What is the total area of these squares?
- (d) Verify that the boundary of  $S_n$ , as a subset of the plane  $\mathbb{R}^2$ , is the union of the sides of the squares constituting  $S_n$ . What is the length of this boundary of  $S_n$ , that is, the total length of all these squares?
- (e) Verify that for each  $n \in \mathbb{N}$ ,

$$S_{n+1} = \left\{ \langle x, y \rangle \in \mathbb{R}^2 : \langle 3x - i, 3y - j \rangle \in S_n \text{ for some } \langle i, j \rangle \in \{0, 1, 2\}^2 \text{ with at most one of } i, j \text{ equal to } 1 \right\}$$

In terms of this description of  $S_n$ , or otherwise, which points are corners of the squares constituting  $S_n$ ? which points are on the sides of these squares?

- (f) Using (e), or otherwise, describe in terms of ternary expansions which points of  $[0, 1]^2$  belong to  $S_n$ , which belong to the boundary of  $S_n$ , which points belong to  $S$ , and which belong to the boundary of  $S$ .

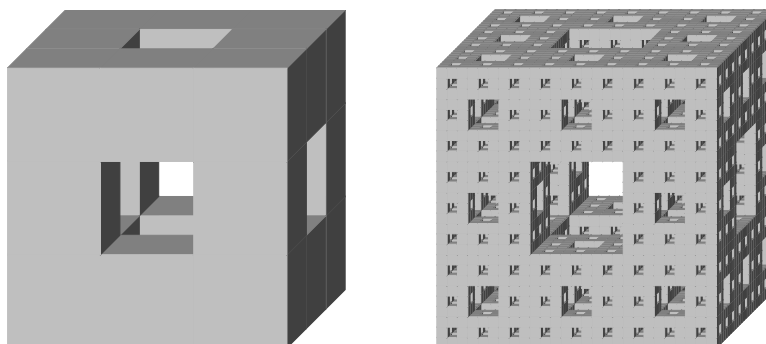
Menger sponge

topological group!compact set@and compact set

compact set!topological group@in topological group  
prob:Menger-sponge

(g) Which topological properties of the Cantor set discussed in Example 4.16 are shared by the Sierpinski carpet  $F$ ?

60. The Menger sponge, first described by Karl Menger, is a three-dimensional analog of the Cantor set and the Sierpinski carpet (Exercise 59). It is the intersection  $M = \bigcap_{n \in \mathbb{N}} M_n$  of the nested sequence of subsets  $M_n$  of the unit cube  $M_0 = [0, 1]^3$  obtained by successively removing open central 27th subcubes. (The sets  $M_1$  and  $M_3$  are shown in Figure 4.9.)



(a) The first stage  $M_1$ .subfig:MengerSponge-level1 (b) The third stage  $M_3$ .subfig:MengerSponge-level3

Figure 4.9: The first and third stages in constructing the Menger sponge.

fig:Menger-sponge

- (a) Sketch  $M_2$ .
- (b) Which points are corners of the cubes of  $M_1$ ? of  $M_2$ ?
- (c) How many cubes constitute  $M_n$  and what are the dimensions of each? What is the total volume of these cubes?
- (d) Verify that the boundary of  $M_n$ , as a subset of  $\mathbb{R}^3$ , is the union of the faces of the cubes constituting  $M_n$ . What is the total area of this boundary?
- (e) Give a description of  $M_{n+1}$  in terms of  $M_n$  analogous to the description in Exercise 59 (e). In terms of this description, or otherwise, which points are corners of the subcubes constituting  $M_n$ ? which points are on the faces of these subcubes?
- (f) Using (e), or otherwise, describe in terms of ternary expansions which points of  $[0, 1]^3$  belong to  $M_n$ , which belong to the boundary of  $M_n$ , which points belong to  $M$ , and which belong to the boundary of  $M$ .
- (g) Which topological properties of the Cantor set discussed in Example 4.16 are shared by the Menger sponge  $M$ ?

Exercises 61–63 concern topological groups. (See Exercises 3.145–3.151.)

prob-part:Menger-sponge-pts

prob:which-top-grps-cpt

61. For  $n \geq 1$ , which of the following topological groups are compact?

- (a) The general linear group  $GL(n, \mathbb{R})$
- (b) The special linear group  $SL(n, \mathbb{R})$ ?
- (c) The orthogonal group  $O(n, \mathbb{R})$ ?
- (d) The special orthogonal group  $SO(n, \mathbb{R})$ ?

prob:cpt-sets-in-top-grps

62. Let  $G$  be a topological group (Exercise 3.145).

- (a) If  $K$  and  $L$  are compact subsets of  $G$ , show that the set  $KL$  of all products of elements of  $K$  with elements of  $L$  is also compact. (Hint: Use Theorem 4.31.)
- (b) If  $K$  and  $L$  are compact subsets of  $G$  and if  $W$  is an open subset of  $G$  containing the set  $KL$ , show that there exist open subsets  $U$  and  $V$  of  $G$  for which  $K \subset U$ ,  $L \subset V$ , and  $UV \subset W$ . (Hint: Use Exercise 36.)
- (c) If  $A$  is a closed subset of  $G$  and  $B$  is a compact subset of  $G$ , prove that the product  $AB$  is closed in  $G$ . [Hint: If  $x \in G \setminus AB$ , then  $xB^{-1} \subset G \setminus A$ ; use (b).]
- (d) By considering the subsets  $A = \mathbb{Z}$  and  $B = \{\sqrt{2}n : n \in \mathbb{Z}\}$ , show that  $AB$  need not be closed in  $G$  when  $A$  and  $B$  are both closed subsets of  $G$ . (Note that when  $G$  is a group, such as  $\mathbb{R}$ , whose operation is addition, then  $AB$  actually means the set  $A + B = \{a + b : a \in A, b \in B\}$ .)

topological group!compact set@and  
compact set!topological group@in t

art:prod-nbd-of-prod-cpt-in-top-grp

63. Prove: If  $H$  is a compact normal subgroup of a topological group  $G$  for which the quotient group  $G/H$  is compact, then  $G$  itself is compact. Here “normal” means in the algebraic sense; see Exercise 3.211.

-grp-cpt-if-subgrp-and-quotient-cpt

## 4.2 Compact Metric Spaces

sec:compactmetric

Among the compact spaces we have already encountered are: the closed unit interval  $I = [0, 1]$  in  $\mathbb{R}$ ; the  $n$ -cube  $I^n = [0, 1]^n$ , the  $n$ -disk  $D_n$ , and the  $(n - 1)$ -sphere  $S_{n-1}$  in  $\mathbb{R}^n$ ; and, more generally, every closed bounded subset of  $\mathbb{R}^n$ . Indeed, many of the most important compact spaces are subspaces of Euclidean spaces and hence are metrizable. For this reason we now explore the special properties that result from the presence of a metric that induces the topology of a compact topological space.

### Countability properties in compact metrizable spaces

subsec:count-in-cpt-metric

A compact subset of  $\mathbb{R}^n$ , like any subspace of  $\mathbb{R}^n$ , is second-countable. The same is true of any compact metrizable space.

cpt-metrizable-implies-2nd-countable

**4.44 Theorem.** A compact metrizable space is second-countable.

**Proof.** Let  $X$  be a compact metrizable space. Choose any metric  $d$  that induces the topology of  $X$ . For each positive integer  $n$ , the collection  $\{B_{1/n}(x; d) : x \in X\}$  of all balls of radius  $1/n$  is an open cover of  $X$ , and so there is a finite set  $F_n \subset X$  with

$$X = \bigcup_{x \in F_n} B_{1/n}(x; d).$$

Let

$$\mathcal{B} = \{B_{1/n}(x; d) : n = 1, 2, \dots, x \in F_n\},$$

whence  $\mathcal{B}$  is a countable collection of open subsets of  $X$ .

To see that  $\mathcal{B}$  is actually a base of  $X$ , let  $U$  be an arbitrary open set in  $X$  and let  $x \in U$ . Choose  $n$  large enough so that

$$B_{2/n}(x; d) \subset U.$$

For this  $n$  there is some  $y \in F_n$  with

$$x \in B_{1/n}(y; d).$$

Then  $B_{1/n}(y; d) \in \mathcal{B}$  with  $x \in B_{1/n}(y; d) \subset U$ .  $\square$



Urysohn Metrization Theorem  
 Urysohn, Pavel  
 Hilbert cube  
 Hilbert cube

**Theorem 4.44** has a converse: Any second-countable compact Hausdorff space is metrizable. This is a special case of the Urysohn Metrization Theorem (6.47), which is proved later, in Section 6.4.

Recall that a metrizable space is second-countable precisely when it is separable (Theorem 2.87). Recall also that any separable metrizable space can be homeomorphically embedded in the Hilbert cube [Exercise 1.36 (b)]. Since the Hilbert cube is a Hausdorff space, Theorem 4.44 provides the following “concrete” representation of compact metrizable spaces.

tot-metrizable-embeds-in-Hilbert-cube

**4.45 Corollary.** Any compact metrizable space is homeomorphic to a closed subspace of the Hilbert cube.

Conversely, each closed subspace of the Hilbert cube is a compact metrizable space, for the Hilbert cube itself is metrizable and, as will follow from Theorem 4.61, below, compact as well.

### Totally bounded metric spaces

subsec:tot-bded

Suppose  $X$  is a compact metrizable space, and let  $d$  be a metric inducing the topology of  $X$ . Then  $X$  is  $d$ -bounded (Lemma 4.14). We are going to prove that  $X$  must actually have a very strong form of  $d$ -boundedness—namely, that the space can be written as the union of finitely many sets whose diameters are as small as we wish.

def:total-bded

**4.46 Definition.** A metric space  $\langle X, d \rangle$  is said to be **totally bounded** when, for each real number  $\varepsilon > 0$ , there is some finite cover of  $X$  consisting of  $d$ -bounded sets each of diameter at most  $\varepsilon$ .

The following criterion provides an equivalent formulation of total boundedness that is often simpler to work with than the preceding definition.

lem:total-bded-criterion

**4.47 Lemma.** A metric space  $\langle X, d \rangle$  is totally bounded if and only if for each  $\varepsilon > 0$ , there is a finite set  $F \subset X$  that is “ $\varepsilon$ -dense in  $X$ ” in the sense that each point of  $X$  is at a distance less than  $\varepsilon$  from some point of  $F$ .

**Proof.** If  $\varepsilon > 0$  and  $\mathcal{A}$  is a finite cover of  $X$  consisting of  $d$ -bounded sets each of  $d$ -diameter at most  $\varepsilon/2$ , then a set consisting of exactly one point from each member of  $\mathcal{A}$  will be finite and  $\varepsilon$ -dense in  $X$ .

Conversely, if  $F \subset X$  is a finite set that is  $\varepsilon/2$ -dense in  $X$ , then the collection  $\{B_{\varepsilon/2}(y; d) : y \in F\}$  will be a finite cover of  $X$  such that  $\text{diam } B_{\varepsilon/2}(y; d) \leq \varepsilon$  for each  $y \in F$ .  $\square$

prop:cpt-metrizable-is-totally-bded

**4.48 Proposition.** A compact metric space is totally bounded.

**Proof.** Let  $\langle X, d \rangle$  be a compact metric space. Let  $\varepsilon > 0$  be arbitrary. The collection  $\{B_\varepsilon(x; d) : x \in X\}$  of all  $d$ -balls of radius  $\varepsilon$  is an open cover of  $X$ , so that  $X$  has a finite subset  $F$  with

$$X = \bigcup_{y \in F} B_\varepsilon(y; d).$$

This means that  $F$  is  $\varepsilon$ -dense in  $X$  in the sense of Lemma 4.47.  $\square$



Notice that from the topological property of compactness we have deduced what appears to be a nontopological conclusion concerning the metric  $d$ . This point is clarified by the following examples.

**4.49 Examples.** (1) If a metric space  $\langle X, d \rangle$  is isometric to a metric space  $\langle Y, d' \rangle$ , then  $\langle X, d \rangle$  is totally bounded if and only if  $\langle Y, d' \rangle$  is totally bounded. Thus total boundedness is a metric property.

(2) If  $Y$  is a subset of a totally bounded metric space  $\langle X, d \rangle$  and if  $d'$  is the metric on  $Y$  induced by  $d$ , then  $\langle Y, d' \rangle$  is also totally bounded.

(3) Let  $d'$  be the Euclidean metric on the open unit interval  $]0, 1[$ . Since the closed interval  $[0, 1]$  is compact, it follows from the preceding [Proposition 4.48](#) and ((2)) that the metric space  $\langle ]0, 1[, d' \rangle$  is totally bounded. However, the open interval  $]0, 1[$  is *not* compact.

(4) Let  $d$  be the Euclidean metric on  $\mathbb{R}$ . Then  $\langle \mathbb{R}, d \rangle$  is *not* totally bounded. In fact, if  $\langle \mathbb{R}, d \rangle$  were totally bounded, we could write  $\mathbb{R}$  as the union of finitely many sets each of diameter at most 1, which would mean that  $\mathbb{R}$  were  $d$ -bounded.

(5) Let  $d'$  and  $d$  be as in (3) and (4), respectively. Then:

- the metric space  $\langle ]0, 1[, d' \rangle$  is topologically equivalent to  $\langle \mathbb{R}, d \rangle$  [because  $]0, 1[$  is homeomorphic to  $\mathbb{R}$ ];
- the metric space  $\langle ]0, 1[, d' \rangle$  is totally bounded; but
- the metric space  $\langle \mathbb{R}, d \rangle$  is *not* totally bounded.  $\diamond$

Although the class of all totally bounded metric spaces strictly contains the class of all compact metric spaces, the two classes share the property of being second-countable.

**4.50 Proposition.** A totally bounded metric space is second-countable.

**Proof.** Let  $\langle X, d \rangle$  be a metric space that is totally bounded.. By [Theorem 2.87](#) it suffices to show that  $X$  is separable. For each positive integer  $n$  there is a finite subset  $F_n$  of  $X$  that is  $(1/n)$ -dense in  $X$  in the sense of [Lemma 4.47](#). Then  $\bigcup_{n=1}^{\infty} F_n$  is a countable set that is easily seen to be dense in  $X$ .  $\square$

A totally bounded metric space need not be compact. However, from [Proposition 4.50](#) and the [Lindelöf Theorem \(2.84\)](#), such a space is close to being compact: each open cover of the space contains a *countable* cover of the space.

### Convergence in compact metrizable spaces

According to [Theorem 4.42](#), an arbitrary topological space is compact precisely when each net in the space clusters. We know that sequences suffice to describe the topology of a metrizable space. Then it is not unreasonable to expect that, for a metrizable space, compactness can be characterized in terms of sequences. This is happily the case and provides the principal tool for dealing with compactness in metrizable spaces.

**4.51 Theorem.** For a metrizable space  $X$  the following conditions are equivalent:

- (i)  $X$  is compact.
- (ii) Each countable open cover of  $X$  contains some finite cover of  $X$ .
- (iii) Each sequence in  $X$  clusters in  $X$ .
- (iv) Each sequence in  $X$  has a convergent subsequence.

**Proof.** Clearly (i) implies (ii). Moreover, conditions (iii) and (iv) are equivalent because in a first-countable space a sequence clusters at a point  $x$  if and only if some subsequence converges to  $x$  (Theorem 3.118).

(ii)  $\implies$  (iii). Assume (ii). We deduce (iii). Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence in  $X$ , and just suppose this sequence does not cluster in  $X$ . For each  $n \in \mathbb{N}$  define an open subset  $U_n$  of  $X$  by

$$U_n = \bigcup \{V : V \text{ is open in } X, x_i \in V \text{ for no } i \geq n\}.$$

Then the countable collection  $\{U_n : n \in \mathbb{N}\}$  of open sets is a cover of  $X$ . In fact, if  $x \in X$ , then  $\langle x_n \rangle_{n \in \mathbb{N}}$  does not cluster at  $x$ , so there is a neighborhood  $V$  of  $x$  and an  $n \in \mathbb{N}$  with  $x_i \in V$  for no  $i \geq n$ , and hence  $x \in U_n$ .

By assumption, the countable open cover  $\{U_n : n \in \mathbb{N}\}$  contains a finite cover of  $X$ . Because  $\langle U_n \rangle_{n \in \mathbb{N}}$  is an increasing sequence of sets, it follows that  $X = U_n$  for some  $n \in \mathbb{N}$ . Since  $x_n \in X$ , we have  $x_n \in U_n$ . But this contradicts the definition of  $U_n$ .

(iii)  $\implies$  (ii). Assume (iii). We deduce (ii). Let  $\{U_n : n \in \mathbb{N}\}$  be a countable open cover of  $X$  and just suppose that  $\{U_n : n \in \mathbb{N}\}$  contains no finite cover of  $X$ . For each  $n \in \mathbb{N}$  we may choose some point

$$x_n \in X \setminus \bigcup_{i=0}^n U_i.$$

By assumption, the sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  clusters at some  $x \in X$ . Let  $i \in \mathbb{N}$  with  $x \in U_i$ . Since  $\langle x_n \rangle_{n \in \mathbb{N}}$  clusters at  $x$ , there is an  $n \geq i$  with  $x_n \in U_n$ . But this contradicts the choice of  $x_n$ .

(ii)  $\implies$  (i). Finally, assume (ii). We deduce (i). Let  $d$  be a metric that induces the topology of  $X$ . [This is the first place we use the full strength of the hypothesis that  $X$  be metrizable.] In view of Proposition 4.50 and the Lindelöf Theorem (2.84), it suffices to show that  $\langle X, d \rangle$  is totally bounded.

Just suppose that  $\langle X, d \rangle$  is not totally bounded. Then there is an  $\varepsilon > 0$  such that no finite subset of  $X$  is  $\varepsilon$ -dense in  $X$  (see Lemma 4.47). Starting with any point  $x_0 \in X$ , we may successively choose points  $x_1, x_2, \dots, x_n, \dots$  in  $X$  with

$$\begin{aligned} d(x_1, x_0) &\geq \varepsilon, \\ d(x_2, x_0) &\geq \varepsilon, \quad d(x_2, x_1) \geq \varepsilon, \\ &\dots\dots\dots \\ d(x_n, x_0) &\geq \varepsilon, \quad d(x_n, x_1) \geq \varepsilon, \quad \dots, \quad d(x_n, x_{n-1}) \geq \varepsilon \end{aligned}$$

The sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  certainly cannot cluster in  $X$ . By the consequence (iii) of our assumption (ii), however, this sequence must cluster in  $X$ .  $\square$

Conditions (ii) and (iv) in the preceding theorem have names.

**4.52 Definition.** A topological space  $X$  is said to be **countably compact** when each countable open cover of  $X$  contains a finite cover.

Evidently, a second countable, countably compact space is compact.

def:seq-cpt

**4.53 Definition.** A topological space  $X$  is said to be **sequentially compact** when each sequence in  $X$  has a convergent subsequence.

Sequential compactness, the form of compactness originally defined by Fréchet for arbitrary metric spaces, is not in general the same as compactness. There are compact Hausdorff spaces that are not sequentially compact: see [Exercise 75](#). And there are first-countable sequentially compact Hausdorff spaces that are not compact: see [Example 4.55](#) and [Exercise 2.102](#).

Although countable compactness and sequential compactness are topological properties that can be studied in their own right, in contrast to compactness they suffer the limitation of not being preserved under the formation of products: see [Exercises 75](#) and [99](#).

For the next example we need a purely set-theoretic result. The term “monotonic” in its name is used to describe a sequence that is either increasing or decreasing.

lem:monotonic-subseq

**4.54 Lemma (Monotonic subsequence theorem).** A sequence in a totally ordered set has a subsequence that is increasing or a subsequence that is decreasing.

**Proof of Lemma 4.54.** Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a sequence in a totally ordered set  $X$ . Call an index  $n$  a “peak index” if  $x_m < x_n$  for all  $m > n$ . Consider two cases.

case:inf-many-peaks

Case (i): The sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  has infinitely many peak indices. In this case there is a strictly increasing sequence  $\langle n_k \rangle_{k \in \mathbb{N}}$  of peak indices. Then  $\langle x_{n_k} \rangle_{k \in \mathbb{N}}$  is a (strictly) decreasing subsequence of the original sequence.

case:finitely-many-peaks

Case (ii): The sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  has only finitely many peak indices. In this case we shall use recursion to obtain an increasing subsequence of  $\langle x_n \rangle_{n \in \mathbb{N}}$  as follows. First, there is an index  $n_0 \in \mathbb{N}$  that is strictly greater than every peak index. In particular,  $n_0$  is not a peak index, and so there is some index  $n_1 > n_0$  with  $x_{n_1} \geq x_{n_0}$ . Next, since  $n_1 > n_0$ , then  $n_1$  is not a peak index either, and so there is some index  $n_2 > n_1$  with  $x_{n_2} \geq x_{n_1}$ . Continuing in this way, we obtain the desired increasing subsequence  $\langle x_{n_k} \rangle_{k \in \mathbb{N}}$  of the original sequence.  $\square$

:Omega-noncpt-T2-sequentially-cpt

**4.55 Example.** Consider the ordinal space  $\Omega = [0, \Omega[$  consisting of all countable ordinals, that is, all ordinals strictly less than the first uncountable ordinal; see [Examples 2.72 \(4\)](#) and [Theorem 0.109](#). **The ordinal space  $[0, \Omega[$  is a sequentially compact Hausdorff space that is not compact.**

$[0, \Omega[$  is not compact: In fact, if  $[0, \Omega[$  were compact, then it would be a closed subspace of the larger (compact) Hausdorff space  $[0, \Omega]$ . But  $[0, \Omega[$  is not a closed subset of  $[0, \Omega]$  because  $\Omega \in \text{cls } [0, \Omega[$ .

$[0, \Omega[$  is a Hausdorff space: In fact, the topology on  $[0, \Omega[$  is its order topology, and any totally ordered set becomes a Hausdorff space when provided with its order topology: see [Examples 2.72 \(1\)](#).

$[0, \Omega[$  is sequentially compact: In fact, let  $\langle \alpha_n \rangle_{n \in \mathbb{N}}$  be an arbitrary sequence in  $[0, \Omega[$ . By the Monotonic subsequence theorem, this sequence either has a subsequence  $\langle \alpha_{n_k} \rangle_{k \in \mathbb{N}}$  that is increasing or else has a subsequence that is decreasing; let  $\langle \alpha_{n_k} \rangle_{k \in \mathbb{N}}$  be such a subsequence.

Suppose first that  $\langle \alpha_{n_k} \rangle_{k \in \mathbb{N}}$  is increasing. According to [Proposition 0.113](#), the range  $\{\alpha_{n_k} : k \in \mathbb{N}\}$  of this subsequence has a supremum  $\beta$  in  $[0, \Omega[$ , and then the subsequence converges to  $\beta$ .

Fréchet, Maurice  
sequential compactness  
compact space!sequences@and seq  
ordinal space!compact space@and c

Bolzano, Bernard  
Weierstrass, Karl

Suppose, on the other hand, that  $\langle \alpha_{n_k} \rangle_{k \in \mathbb{N}}$  is decreasing. Then its range has a least element  $\beta$ , and so the subsequence converges to  $\beta$ .  $\diamond$

To relate [Theorem 4.51](#) to a result in classical analysis, we recall from [Definition 2.23](#) the notion of a *limit point*: a point  $x$  in a space  $X$  is a limit point in  $X$  of a subset  $A$  of  $X$  when each neighborhood of  $x$  contains at least one element of  $A$  different from  $x$ .

cor:limit-pt-in-cpt-metrizable

**4.56 Corollary.** *A metrizable space is compact if and only if each of its infinite subsets has a limit point in the space.*

**Proof.** Let  $X$  be a metrizable space.

First, assume that  $X$  is compact. Let  $A$  be an infinite subset of  $X$ . There exists a sequence of *distinct* points of  $A$ ; choose such a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$ . By [Theorem 4.51](#), this sequence clusters at some  $x \in X$ . Then  $x$  is a limit point of  $A$  in  $X$ . In fact, let  $V$  be any neighborhood of  $x$ . There is an  $n \in \mathbb{N}$  with  $x \neq x_i$  for all  $i \geq n$  (possibly  $n = 0$ ). Since  $\langle x_n \rangle_{n \in \mathbb{N}}$  clusters at  $x$ , there is an  $i \geq n$  with  $x_i \in V$ . Then  $x \neq x_i \in V \cap A$ , as desired.

Conversely, assume that each infinite subset of  $X$  has a limit point in  $X$ . To show that  $X$  is compact, by [Theorem 4.51](#) it suffices to show that each sequence in  $X$  clusters in  $X$ . Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be any sequence in  $X$ . If the range  $\{x_n : n \in \mathbb{N}\}$  of this sequence is finite, the sequence certainly clusters in  $X$ . Now suppose that the range  $\{x_n : n \in \mathbb{N}\}$  is an infinite set. By assumption, this set has a limit point  $x$  in  $X$ .

We complete the proof by showing that  $\langle x_n \rangle_{n \in \mathbb{N}}$  clusters at  $x$ . Let  $V$  be any neighborhood of  $x$  and let  $n \in \mathbb{N}$ . The neighborhood

$$V \setminus \{x_i : 1 \leq i < n, x_i \neq x\}$$

of  $x$  contains some point  $x_i$  of  $\{x_n : n \in \mathbb{N}\}$  that is different from  $x$ . Then  $i \geq n$  and  $x_i \in V$ , as desired.  $\square$

From that corollary and the Heine–Borel–Lebesgue Theorem ([4.15](#)) we obtain the following classical result, named after Bernard Bolzano, who first proved it (as a step toward proving the Intermediate-value Theorem ([5.13](#))), and Karl Weierstrass, who subsequently emphasized its importance.

thm:Bolzano-Weierstrass

**4.57 Bolzano–Weierstrass Theorem.** *Each bounded infinite set of real numbers has a limit point in  $\mathbb{R}$ .*

Proof of the following consequence of Bolzano–Weierstrass Theorem is left to the reader ([Exercise 68](#)).

or:uncountable-subset-R-then-limit-pt

**4.58 Corollary.** *An uncountable subset of  $\mathbb{R}$  has a limit point in  $\mathbb{R}$ .*

subsec:tot-bded-bis

### More about totally bounded metric spaces

For our first substantial application of [Theorem 4.51](#), we at last establish the exact relationship between compactness and boundedness.

thm:cpt-iff-complete-and-totally-bded

**4.59 Theorem.** *Let  $d$  be any metric inducing the topology of a metrizable space  $X$ . Then a necessary and sufficient condition for the topological space  $X$  to be compact is that the metric space  $\langle X, d \rangle$  be both complete and totally bounded.*

**Proof.** Necessity. Assume that  $X$  is compact. That  $\langle X, d \rangle$  is totally bounded is just Proposition 4.48. To see that  $\langle X, d \rangle$  is also complete, let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be a Cauchy sequence in  $\langle X, d \rangle$ . By Theorem 4.51 (or by Theorem 4.42), this sequence clusters at some  $x$  in  $X$ .

totally bounded subset

We show that  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges to  $x$ . Let  $\varepsilon > 0$  be arbitrary. Since  $\langle x_n \rangle_{n \in \mathbb{N}}$  is a Cauchy sequence, there is an  $n \in \mathbb{N}$  with

$$d(x_i, x_j) < \frac{\varepsilon}{2} \quad (i \geq n, j \geq n),$$

and since the sequence clusters at  $x$ , there is an  $m \geq n$  with

$$d(x_m, x) < \frac{\varepsilon}{2}.$$

Then  $d(x_i, x) < \varepsilon$  for all  $i \geq n$ .

(Observe that this part of the proof did not actually require 4.51, only 4.42.)

Sufficiency. Assume that  $\langle X, d \rangle$  is complete and totally bounded. We are going to show that each sequence in  $X$  clusters there, and from Theorem 4.51 it will follow that  $X$  is compact. Let  $\langle x_n \rangle_{n \in \mathbb{N}}$  be an arbitrary sequence in  $X$ . If the range

$$R = \{x_n : n \in \mathbb{N}\}$$

of the sequence is finite, then already the sequence will cluster in  $X$ , and so we may now suppose that  $R$  is in fact infinite. To obtain a point at which the sequence will cluster, we shall use the total boundedness of  $\langle X, d \rangle$  to construct a certain sequence of sets and then apply the Nested Set Theorem (1.88) to this sequence of sets.

Since  $\langle X, d \rangle$  is totally bounded, for each  $n \in \mathbb{N}$  there is a finite cover  $\mathcal{E}_n$  of  $X$  consisting of  $d$ -bounded sets with

$$\text{diam } E \leq \frac{1}{n+1} \quad (E \in \mathcal{E}_n).$$

Since the diameter of the closure of a  $d$ -bounded set is the same as the diameter of the set itself (Exercise 2.49), we may assume that each member of each  $\mathcal{E}_n$  is closed in  $X$ .

Now we construct a decreasing sequence  $\langle E_n \rangle_{n \in \mathbb{N}}$  of closed subsets of  $X$  such that

$$\text{diam } E_n \leq \frac{1}{n+1}, \quad E_n \cap R \text{ is finite} \quad (n \in \mathbb{N}).$$

To begin, since the cover  $\mathcal{E}_0$  of  $X$  is finite while  $R$  is infinite, we may choose some  $E_0 \in \mathcal{E}_0$  such that  $E_0 \cap R$  is infinite. Now suppose  $n > 0$  and we have already constructed  $E_0, E_1, \dots, E_{n-1}$ . Since the cover  $\mathcal{E}_n$  of  $X$  is finite while  $E_{n-1} \cap R$  is infinite, we may choose some  $E'_n \in \mathcal{E}_n$  such that  $E'_n \cap (E_{n-1} \cap R)$  is infinite, and then we take  $E_n = E'_n \cap E_{n-1}$ .

By the Nested Set Theorem there is some point  $x \in \bigcap_{n=0}^{\infty} E_n$ . To complete the proof we show that  $\langle x_n \rangle_{n \in \mathbb{N}}$  clusters at  $x$ . Let  $\varepsilon > 0$  be arbitrary and let  $m \in \mathbb{N}$ . Choose  $n \in \mathbb{N}$  with  $1/(n+1) < \varepsilon$ . Since  $E_n \cap R$  is infinite, there is an  $i \geq m$  with  $x_i \in E_n$ . But  $x \in E_n$  with  $\text{diam } E_n \leq 1/(n+1) < \varepsilon$ , and so  $d(x, x_i) < \varepsilon$ .  $\square$

The remarkable thing about this theorem is its assertion that, in a metrizable space, the topological property of compactness is equivalent to the metric property of completeness and total boundedness. Note that neither completeness nor total boundedness of  $\langle X, d \rangle$  alone suffices to ensure the compactness of  $X$ : consider  $\mathbb{R}$  and  $]0, 1[$  with their Euclidean metrics.

We shall say that a subset  $K$  of a metric space  $\langle X, d \rangle$  is  **$d$ -totally bounded** to mean, of course, that the metric space  $\langle K, d' \rangle$  is totally bounded, where  $d'$  is the metric on  $K$  induced by  $d$ . By combining Theorem 4.59 with Theorem 1.81, we obtain the following corollary.

ete-metric-cpt-iff-closed-and-tot-bded

**4.60 Corollary.** *Let  $\langle X, d \rangle$  be a complete metric space. Then a subset  $K$  of  $X$  is compact if and only if  $K$  is both closed in  $X$  and  $d$ -totally bounded.*

This corollary is an analog of the Heine–Borel–Lebesgue Theorem (4.15) and at the same time is an explanation of why that theorem does not generalize intact to arbitrary complete metric spaces.

For a second application of Theorem 4.51, we now prove, as promised in the preceding section, that the product of a sequence of compact metrizable spaces is itself compact. (Actually, our proof uses Theorem 4.59, but the proof of that result rested in turn on Theorem 4.51.)

honoff-thm-denumerable-metric-case

**4.61 Theorem (Tychonoff Product Theorem—denumerable metrizable case).** *The product of a sequence of nonempty metrizable spaces is compact if and only if each of the spaces is compact.*

**Proof.** Let  $\langle X_i \rangle_{i=1,2,3,\dots}$  be a sequence of nonempty metrizable spaces and let  $X = \prod_{i=1}^{\infty} X_i$ .

$X$  compact  $\implies X_i$  compact for each  $i$ . The proof of this implication is the same as in the case of a finite family (Theorem 4.32) and requires no metrizability assumption whatsoever.

$X_i$  compact for each  $i \implies X$  compact. For this converse implication we do use metrics. Assume that  $X_i$  is compact for each  $i$ . By Proposition 1.40, for each  $i$  there is a bounded metric  $d_i$  that induces the topology of  $X_i$  and for which  $\text{diam } X_i \leq 1$ . Then the formula

$$d(x, y) = \sum_{i=1}^{\infty} \frac{d_i(x_i, y_i)}{2^i}$$

defines a metric on  $X$  that induces the product topology there (see Proposition 3.51).

Assume that  $X_i$  is compact for each  $i = 1, 2, 3, \dots$ . By Theorem 4.59, the metric space  $\langle X_i, d_i \rangle$  is complete and totally bounded for each  $i$ . By Exercise 1.109 the product metric space  $\langle X, d \rangle$  is complete. It now remains only to establish that  $\langle X, d \rangle$  is also totally bounded, for then from Theorem 4.59 it will follow that  $X$  is compact.

Let  $\varepsilon > 0$  be arbitrary. We shall construct a finite subset of  $X$  that is  $\varepsilon$ -dense in  $X$  in the sense of Lemma 4.47. Choose  $n$  so large that

$$\sum_{n+1}^{\infty} \frac{1}{2^i} < \frac{\varepsilon}{2}.$$

Fix any point

$$\langle z_{n+1}, z_{n+2}, z_{n+3}, \dots \rangle \in X_{n+1} \times X_{n+2} \times X_{n+3} \times \dots.$$

For each  $i = 1, 2, \dots, n$ , the totally bounded metric space  $\langle X_i, d_i \rangle$  has a finite subset  $F_i$  that is  $(\varepsilon/2)$ -dense in  $X_i$ . Define

$$F = F_1 \times F_2 \times \dots \times F_n \times \{z_{n+1}\} \times \{z_{n+2}\} \times \{z_{n+3}\} \times \dots,$$

so that  $F$  is a finite subset of  $X$ .

To complete the proof, we show that this set  $F$  is  $\varepsilon$ -dense in  $X$ . Consider an arbitrary point  $x = \langle x_i \rangle_{i=1,2,3,\dots} \in X$ . For each  $i = 1, 2, \dots, n$  there is a point  $y_i \in F_i$  with

Hilbert cube  
Hilbert cube

$$d_i(x_i, y_i) < \frac{\varepsilon}{2}.$$

Define  $y_i = z_i$  for each  $i = n+1, n+2, n+3, \dots$ . Then the point  $y = \langle y_i \rangle_{i=1,2,3,\dots}$  of  $F$  satisfies

$$\begin{aligned} d(x, y) &= \sum_{i=1}^n \frac{d_i(x_i, y_i)}{2^i} + \sum_{i=n+1}^{\infty} \frac{d_i(x_i, y_i)}{2^i} \\ &< \sum_{i=1}^n \frac{\varepsilon}{2^{i+1}} + \sum_{i=n+1}^{\infty} \frac{1}{2^i} \\ &< \frac{\varepsilon}{2} \sum_{i=1}^n \frac{1}{2^i} + \frac{\varepsilon}{2} \\ &= \varepsilon. \quad \square \end{aligned}$$

It is possible to prove that the product of an arbitrary sequence of compact spaces is compact without any assumption of metrizable—yet without applying the general Tychonoff Product Theorem: see [Exercise 6.5](#).

Recall that the Hilbert cube  $l^\infty$ , defined in [Exercise 1.36](#) as a certain subspace of Hilbert sequence space  $\ell^2$ , is homeomorphic to the product of a sequence of copies of the unit interval  $I$ . Hence the following is an immediate consequence of [Theorem 4.61](#).

cor:Hilbert-cube-cpt **4.62 Corollary.** *The Hilbert cube is compact.*

### More metric consequences of compactness

The main thrust of our discussion of compact metrizable spaces so far has been the use of metrics to provide criteria for compactness. In the remainder of the discussion we proceed in the opposite direction, deducing metric properties from the topological assumption of compactness.

Suppose  $\mathcal{U}$  is an open cover of a compact space  $X$ . Although some (possibly very large) finite number of members of  $\mathcal{U}$  will still form a cover of  $X$ , a given subset of  $X$  may require many of these members to cover all its points. When the topology of  $X$  is induced by a metric  $d$ , however, any subset of  $X$  having sufficiently small  $d$ -diameter will lie completely within a single one of these members. That is the content of the following result.

thm:Lebesgue-covering-lemma **4.63 Lebesgue Covering Lemma.** *Let  $\mathcal{U}$  be an open cover of a compact metric space  $\langle X, d \rangle$ . Then there exists a real number  $\delta > 0$  such that any two points  $x$  and  $t$  of  $X$  having distance  $d(x, t) < \delta$  belong to the same member of  $\mathcal{U}$ .*

**Proof.** For each  $x \in X$  we may choose some set  $U_x \in \mathcal{U}$  with  $x \in U_x$  and then choose a number  $\delta(x) > 0$  with

$$B_{\delta(x)/2}(x; d) \subset U_x.$$

The collection  $\{B_{\delta(x)/2}(x; d) : x \in X\}$  is an open cover of  $X$ , and so there is a finite subset  $F$  of  $X$  such that

$$X = \bigcup_{x \in F} B_{\delta(x)/2}(x; d)$$

uniformly continuous map  
 continuous map!uniformly continuous map@and uniformly continuous map  
 map!uniformly continuous  
 Lebesgue Covering Lemma  
 Heine, Eduard

(we may assume that  $F \neq \emptyset$ , for otherwise  $X = \emptyset$  and there is nothing to prove). Define

$$\delta = \min_{x \in F} \frac{\delta(x)}{2}.$$

Let  $x, t \in X$  with  $d(x, t) \leq \delta$ . There is a  $z \in F$  with  $x \in B_{\delta(x)/2}(x; d)$ . Then

$$d(t, z) \leq d(t, x) + d(x, z) < \delta + \frac{\delta(z)}{2} \leq \frac{\delta(z)}{2} + \frac{\delta(z)}{2} = \delta(z).$$

Hence  $x$  and  $t$  both belong to  $B_{\delta(z)}(z; d)$  and so to the member  $U_z$  of  $\mathcal{U}$ .  $\square$

An evidently equivalent way of stating the property of  $\delta$  is that if  $A \subset X$  with  $\text{diam } A \leq \delta$ , then  $A \subset U$  for some  $U \in \mathcal{U}$ .

def:Lebesgue-number

**4.64 Definition.** Let  $\mathcal{U}$  be an open cover of a metric space  $X$ . By a **Lebesgue number of  $\mathcal{U}$**  is meant any positive number  $\delta$  having the property that whenever  $A$  is a subset of  $X$  for which  $\text{diam } A \leq \delta$ , then  $A \subset U$  for some  $U \in \mathcal{U}$ .

We are going to apply the Lebesgue Covering Lemma to prove an important result concerning uniform continuity, a notion introduced in [Exercise 1.101](#) and which we now recall. Given metric spaces  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$ , a map  $f: X \rightarrow Y$  is said to be  $\langle d, d' \rangle$ -**uniformly continuous**—or simply **uniformly continuous**—when the metrics are understood—provided that for each  $\varepsilon > 0$ , there is some  $\delta > 0$  such that

$$\text{for all } x, t \in X, \quad d(x, t) < \delta \implies d'(f(x), f(t)) < \varepsilon.$$

Contrast this definition with what it means to say that  $f$  is continuous in the ordinary sense: the map  $f$  is continuous when it is  $\langle d, d' \rangle$ -continuous at each  $x \in X$ —in other words, when for each  $\varepsilon > 0$  and each  $x \in X$ , there is some  $\delta(x) > 0$  such that

$$d(x, t) < \delta(x) \implies d'(f(x), f(t)) < \varepsilon.$$

Corresponding to a given  $\varepsilon$ , then, ordinary continuity of  $f$  says that there is a separate  $\delta(x)$  for each  $x \in X$  which “works at that individual  $x$ ”, whereas uniform continuity says that there is a single  $\delta$  which “works at every  $x \in X$ .”

Clearly  $f$  will be continuous if it is  $\langle d, d' \rangle$ -uniformly continuous, but the converse fails (see [Exercise 1.101](#)). A partial converse, however, does hold.

thm:cont-on-cpt-is-unif-cont

**4.65 Theorem.** A continuous map from a compact metric space  $\langle X, d \rangle$  to a metric space  $\langle Y, d' \rangle$  is  $\langle d, d' \rangle$ -uniformly continuous.

**Proof.** Let  $f: X \rightarrow Y$  be a continuous map from a compact metric space  $\langle X, d \rangle$  to a metric space  $\langle Y, d' \rangle$ . Let  $\varepsilon > 0$  be arbitrary. The collection

$$\{f^{-1}(B_{\varepsilon/2}(y; d')) : y \in Y\}$$

is an open cover of  $X$ . By Lebesgue Covering Lemma ([4.63](#)), this open cover has a Lebesgue number  $\delta > 0$ . Let  $x$  and  $t$  be arbitrary points in  $X$  with  $d(x, t) < \delta$ . There is some  $y \in Y$  such that  $x$  and  $t$  both belong to  $f^{-1}(B_{\varepsilon/2}(y; d'))$ . Thus both  $f(x)$  and  $f(t)$  belong to  $B_{\varepsilon/2}(y; d')$ , and consequently  $d'(f(x), f(t)) < \varepsilon$ .  $\square$

This theorem was first proved by Eduard Heine for the case of a continuous real-valued function on a closed interval  $[a, b]$  in  $\mathbb{R}$ . In that case, especially, it has many significant applications to the theory of integration and other areas of analysis. We give here just one simple application, which was already exploited in [Application 1.92](#).



app:close-piecewise-linear

**4.66 Application.** Consider a closed interval  $[a, b]$  in  $\mathbb{R}$ . A function  $g: [a, b] \rightarrow \mathbb{R}$  which, for suitable points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b,$$

is linear on each of the *open* intervals  $]x_0, x_1[, ]x_1, x_2[, \dots, ]x_{n-1}, x_n[$  is said to be **piecewise linear**. Such a function will be continuous precisely when it is linear on each of the *closed* intervals  $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$ .

Let  $f: [a, b] \rightarrow \mathbb{R}$  be a given continuous function, and let  $\varepsilon > 0$  be a given number. We claim:

There is a continuous *piecewise linear* function  $g: [a, b] \rightarrow \mathbb{R}$  with

$$|f(x) - g(x)| < \varepsilon \quad (a \leq x \leq b),$$

in other words, with  $g$  approximating  $f$  to within  $\varepsilon$  *uniformly* over  $[a, b]$ .

To justify our claim we first invoke the uniform continuity of  $f$  to obtain a  $\delta > 0$  such that

$$x, t \in [a, b] \quad \text{and} \quad |x - t| < \delta \implies |f(x) - f(t)| < \frac{\varepsilon}{2}.$$

Next we choose a positive integer  $n$  with  $(b - a)/n < \delta$  and define points  $x_0, x_1, x_2, \dots, x_n$  in  $[a, b]$  by

$$x_i = a + i \frac{b - a}{n} \quad (i = 0, 1, 2, \dots, n).$$

Then for each  $i = 1, 2, \dots, n$  we have  $|x_i - x_{i-1}| = (b - a)/n < \delta$ , so that

$$x, t \in [x_{i-1}, x_i] \implies |f(x) - f(t)| < \frac{\varepsilon}{2}.$$

Define  $g: [a, b] \rightarrow \mathbb{R}$  to be the function that is linear on each of the closed intervals  $[x_{i-1}, x_i]$  and takes values

$$g(x_i) = f(x_i) \quad (i = 0, 1, 2, \dots, n)$$

at the endpoints of these intervals (see Figure 4.10).

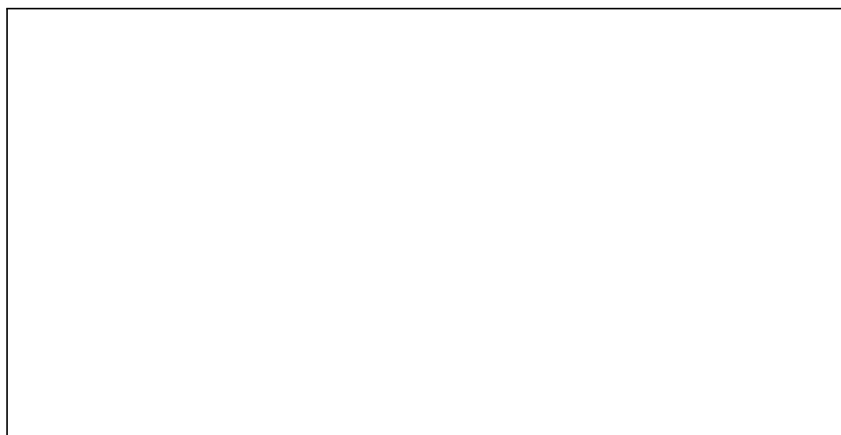


Figure 4.10: Constructing a piecewise linear function  $g$  to uniformly approximate a continuous function  $f: [a, b] \rightarrow \mathbb{R}$ .

fig:piecewise-approx-cont-closed-int

Let  $x \in [a, b]$  be arbitrary. There is an  $i$  with  $1 \leq i \leq n$  for which  $x \in [x_{i-1}, x_i]$ . Then  $g(x)$  lies between the numbers  $g(x_{i-1}) = f(x_{i-1})$  and  $g(x_i) = f(x_i)$ , so that

Lindelöf Theorem

$$|f(x_i) - g(x)| \leq |f(x_i) - f(x_{i-1})| < \frac{\varepsilon}{2}.$$

Lindelöf Theorem

Since also

$$|f(x) - f(x_i)| < \frac{\varepsilon}{2},$$

we conclude that  $|f(x) - g(x)| < \varepsilon$ , as desired.  $\diamond$

## EXERCISES FOR SECTION 4.2

64. Show that a metric space  $\langle X, d \rangle$  is totally bounded precisely when each infinite subset of  $X$  contains distinct points that are an arbitrarily small distance apart.
65. If  $d$  is the Euclidean metric on  $\mathbb{R}^n$ , show that a subset of  $\mathbb{R}^n$  is  $d$ -totally bounded if and only if it is  $d$ -bounded. Is the same thing true if  $d$  is instead the max metric (Definition 1.7) on  $\mathbb{R}^n$ ? the taxicab metric (Definition 1.6)?
66. Given two equivalent metrics  $d$  and  $d'$  on the same set  $X$  with  $\langle X, d \rangle$  totally bounded, must  $\langle X, d' \rangle$  also be totally bounded?
67. Suppose a metric space  $\langle X, d \rangle$  has a dense  $d$ -totally bounded subset. Will  $\langle X, d \rangle$  necessarily be totally bounded?

68. Prove Corollary 4.58: an uncountable subset of  $\mathbb{R}$  has a limit point in  $\mathbb{R}$ . (Do this *without* applying the stronger result Exercise 2.120.)

69. Prove the following converse of the Lindelöf Theorem (2.84): If each open cover of a metrizable space  $X$  contains a countable cover of  $X$ , then  $X$  is second-countable.

[Note: From the definitions, a compact space is a Lindelöf space; and according to the Lindelöf Theorem (2.84), a second-countable space is a Lindelöf space. Thus the property of being a Lindelöf space generalizes both compactness and second-countability.]

70. (Continuation of Exercise 2.115.)

Must a quotient of a Lindelöf space be a Lindelöf space?

71. A topological space is said to be  $\sigma$ -compact when it is the union of a sequence of compact subsets. For example, any compact space is  $\sigma$ -compact; the noncompact space  $\mathbb{R}$  is  $\sigma$ -compact since  $\mathbb{R} = \bigcup_{n=0}^{\infty} [-n, n]$ .

(a) Show that the Lindelöf Theorem (2.84) remains valid if the hypothesis that  $X$  be second-countable is changed to  $X$  being  $\sigma$ -compact. In other words, show that a  $\sigma$ -compact space is necessarily a Lindelöf space (Exercise 2.115).

(b) Let  $\mathcal{T}$  be the usual topology on  $\mathbb{R}$  and let  $S = \{1/n : n = 1, 2, 3, \dots\}$ . Verify that the collection

$$\mathcal{S} = \{U \setminus A : U \in \mathcal{T}, A \subset S\}$$

is a topology on  $\mathbb{R}$  and that the topological space  $\langle \mathbb{R}, \mathcal{S} \rangle$  is a noncompact,  $\sigma$ -compact Hausdorff space that is second-countable but not metrizable.

(c) Prove that an arbitrary  $\sigma$ -compact metrizable space is necessarily second-countable.

fix: Move  
problem on quot  
of Lindelöf  
somewhere

prob:sigma-cpt



- prob:part-rel-compact-iff-seq-cluster
80. (a) Show that a subset  $A$  of a metrizable space  $X$  is relatively compact (Exercise 13) if and only if each sequence in  $A$  clusters in  $X$ .  
 (b) Establish an analog of (a) concerning limit points.
- prob:lim-pt-cpt
81. Call a topological space  $X$  **limit-point compact** when each infinite subset of  $X$  has a limit point in  $X$ .  
 (a) Prove that a countably compact space is limit-point compact.  
*Note:* In particular, then, every compact space is limit-point compact.  
 (b) Prove that a  $T_1$ -space that is limit-point compact is countably compact.  
 (c) Verify that the countable collection  $\{2n, 2n + 1\} : n \in \mathbb{Z}\}$  is a base for a topology  $\mathcal{T}$  on  $\mathbb{Z}$  such that  $\langle \mathbb{Z}, \mathcal{T} \rangle$  is limit point compact but not compact.  
 (d) Must a limit-point compact Hausdorff space be compact?  
 (e) What is the relationship of limit-point compactness to sequential compactness?
82. Must an infinite totally bounded subset of a metric space have a limit point in the space? If not, what about the case when the space is complete?
- prob:cptication-intro
83. A **compactification** of a topological space  $X$  is a compact space  $Y$  containing  $X$  as a dense subspace. For example, the extended real line (Example 1.41) is a compactification of the real line.  
 Prove that a separable metrizable space always has a compactification.  
*(Hint: Look at Corollary 6.46.)*
84. Let  $X$  be a second-countable metrizable space. Show that there is a totally bounded metric on  $X$  that induces the given topology of  $X$ .
85. Prove that a metric space is totally bounded if and only if each sequence in it has a subsequence that is a Cauchy sequence.
86. Is the Hilbert sequence space (Example 1.10) compact?
87. Let  $K$  be a compact subset of the product of a sequence of noncompact metrizable spaces. Prove that  $K$  has empty interior in  $X$ .
- prob:cover-diag-cpt-metric-times-itself
88. Give a compact metric space  $\langle X, d \rangle$ , let  $U$  be an open subset of  $X \times X$  containing the diagonal  $\{\langle x, x \rangle : x \in X\}$  of  $X \times X$ . Show that there is an  $\varepsilon > 0$  such that  $\{\langle x, y \rangle \in X \times X : d(x, y) < \varepsilon\} \subset U$ .  
*(Hint: See Exercise 34.)*
89. Let  $d'$  be the Euclidean metric on  $\mathbb{R}$ . Prove that the following is a necessary and sufficient condition for a real-valued function  $f$  on a metric space  $\langle X, d \rangle$  to be  $\langle d, d' \rangle$ -uniformly continuous: If  $\langle t_n \rangle_{n \in \mathbb{N}}$  and  $\langle x_n \rangle_{n \in \mathbb{N}}$  are arbitrary sequences in  $X$  with  $\lim_{n \rightarrow \infty} d(t_n, x_n) = 0$ , then  $\lim_{n \rightarrow \infty} |f(t_n) - f(x_n)| = 0$ .
90. (a) Let  $f : [0, 1] \rightarrow \mathbb{R}$  be continuous and let  $\varepsilon > 0$  be given. Show that there is a positive integer  $n$  and a continuous function  $g : [0, 1] \rightarrow \mathbb{R}$  that is linear on each of the subintervals  $[0, 1/n], [1/n, 2/n], \dots, [(n-1)/n, 1]$ , takes rational values at the  $n+1$  endpoints of these subintervals, and approximates  $f$  uniformly within  $\varepsilon$  over  $[0, 1]$ .  
 (b) Deduce that the space  $C([0, 1])$  of all continuous real-valued functions on  $[0, 1]$  is separable when given the topology induced by its max matrix  $d_\infty$  (Example 1.8).

- 91.** Again let  $f: [0, 1] \rightarrow \mathbb{R}$  be continuous and let  $\varepsilon > 0$  be given. Establish the existence of a positive integer  $n$  and real-valued functions  $g$  and  $h$  on  $[0, 1]$  that are constant on each of the open subintervals  $]0, 1/n[, ]1/n, 2/n[, \dots, ](n-1)/n, 1[$  such that  $g(x) \leq f(x) \leq h(x)$  and  $|h(x) - g(x)| < \varepsilon$  for all  $x \in [0, 1]$ . (The “step functions”  $g$  and  $h$  need not be continuous, of course.)

(Note: This result is crucial for proving that  $f$  is integrable.)

- 92.** Let  $d'$  be the Euclidean metric on  $\mathbb{R}$  and let  $\langle X, d \rangle$  be a metric space having the property that every continuous function  $f: X \rightarrow \mathbb{R}$  is  $\langle d, d' \rangle$ -uniformly continuous.

- (a) If  $A$  and  $B$  are any two disjoint nonempty closed subsets of  $X$ , show that  $0 < d(A, B)$ , the distance between  $A$  and  $B$  (Exercise 1.43).

(Hint: Use Exercise 25.)

- (b) Prove that  $\langle X, d \rangle$  must be complete.

[Hint: If  $\langle x_n \rangle_{n \in \mathbb{N}}$  is a Cauchy sequence of distinct points in  $X$ , apply (a) taking  $A = \{x_{2n} : n \in \mathbb{N}\}$  and  $B = \{x_{2n+1} : n \in \mathbb{N}\}$ .]

- 93.** Let  $\langle X, d \rangle$  be a metric space, let  $d_1$  be the metric induced by  $d$  on a dense subset  $D$  of  $X$ , and let  $\langle Y, d' \rangle$  be a complete metric space. Let  $f: D \rightarrow Y$  be a map.

- (a) If  $f$  is  $\langle d_1, d' \rangle$ -uniformly continuous, prove that  $f$  has a unique continuous extension  $F: X \rightarrow Y$  and that  $F$  is in fact  $\langle d, d' \rangle$ -uniformly continuous.

[Hint: If  $x \in X$ , there is a sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in  $D$  converging to  $x$  in  $X$ ; if such a continuous extension does exist, then necessarily  $\langle f(x_n) \rangle_{n \in \mathbb{N}}$  must converge to  $F(x)$ .]

- (b) Must  $f$  have a continuous extension to  $X$  if it is  $\langle d_1, d' \rangle$ -uniformly continuous but  $\langle Y, d' \rangle$  is not complete? If  $f$  is only continuous but  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$  are both complete?

- 94.** Let  $\langle X, d \rangle$  and  $\langle Y, d' \rangle$  be metric spaces. A surjection  $f: X \rightarrow Y$  is said to be “ $\langle d, d' \rangle$ -uniformly open” when for each  $\delta > 0$ , there is some  $\varepsilon > 0$  such that  $y, z \in Y$  with  $d'(y, z) < \varepsilon$  implies  $y = f(x)$  and  $z = f(t)$  for some  $x, t \in X$  with  $d(x, t) < \delta$ .

- (a) Show that a  $\langle d, d' \rangle$ -continuous open surjection  $f: X \rightarrow Y$  must be an open map, but that the converse need not hold.

- (b) If  $X$  is compact, prove that each continuous open surjection  $f: X \rightarrow Y$  is  $\langle d, d' \rangle$ -uniformly open.

- 95.** Let  $\langle E_n \rangle_{n \in \mathbb{N}}$  be a decreasing sequence of nonempty closed subsets of a compact metric space. Compare  $\text{diam}(\bigcap_{n=0}^{\infty} E_n)$  with  $\inf_{n=0,1,2,\dots} \text{diam } E_n$ .

- 96.** Let  $\mathcal{F}$  be the collection of all nonempty closed subsets of a compact metric space  $\langle X, d \rangle$  and provide  $\mathcal{F}$  with its topology induced by the Hausdorff metric (Exercise 1.54). Show that the map  $E \mapsto \text{diam } E$  from  $\mathcal{F}$  to  $\mathbb{R}$  is continuous.

- 97.** (a) Prove: A Hausdorff space that is the continuous image of a compact metrizable space is second-countable.

[Hint: You may want to use Exercise 53 (a).]

- (b) Deduce from (a) and the Urysohn Metrization Theorem (6.47) (see also the paragraph following Theorem 4.44) that the quotient space of a compact metrizable space under a closed equivalence relation is itself metrizable.

- prob:isom:metrically embed compact space  
sequentially compact space  
Lindelof space  
locally compact space  
locally compact space! Hausdorff space and Hausdorff space
98. (a) Prove that an isometric embedding (Exercise 1.59) of a compact metric space  $\langle X, d \rangle$  into itself actually maps  $X$  onto  $X$  and hence is an isometry from  $\langle X, d \rangle$  to  $\langle X, d \rangle$ .  
(b) If each of two compact metric spaces can be isometrically embedded in the other, prove that the two metric spaces are isometric to one another. (Compare Exercise 1.61.)
- prob-part:ccpt-times prob:the-2-ccpt
99. (a) Prove that the product of a countably compact space with a compact space is countably compact.  
[Suggestion: Use Exercise 45 to show that the first projection is a perfect map (Exercise 50).]  
(b) Prove that the product of a countably compact space with a first-countable, countably compact space is also countably compact.  
(c) Prove that the product of a countably compact space with a sequentially compact, countably compact space is also countably compact.  
*Note:* It is known that the product of a countably compact space with itself need not be countably compact: see Novák [52].
100. Must the product of a Lindelöf space (Exercise 2.115) with a compact space be a Lindelöf space?

### 4.3 Locally Compact Spaces

sec:localcompact

It is a fact of life that numerous topological spaces, including Euclidean spaces, fail to be compact. The pleasant properties that compact spaces enjoy can still be exploited in many of these noncompact spaces—provided they are “locally” compact in the sense of having arbitrarily small compact neighborhoods at each point.

#### Locally compact spaces

subsec:loc-cpt

def:loc-cpt **4.67 Definition.** A topological space  $X$  is said to be **locally compact** if at each point  $x \in X$  there is a local base consisting of compact sets—in other words, if for each neighborhood  $U$  of  $x$  there is a compact neighborhood  $V$  of  $x$  with  $V \subset U$ .

**Usage notes.** (1) Some mathematicians define a locally compact space to be, instead, a topological space in which each point has *some* compact neighborhood. When the space is a Hausdorff space, as we are about to see, there is no distinction between that weaker definition and that of Definition 4.67.

- (2) Many mathematicians stipulate as part of the definition of a locally compact space that it be a Hausdorff space. So as to be consistent with our usage of “compact” and to keep our terminology unburdened by extra assumptions, however, we shall *not* so stipulate.

Nonetheless, most of the locally compact spaces one encounters *are* Hausdorff spaces, and for such spaces the following proposition provides an especially simple criterion for local compactness.

prop:criterion-loc-cpt-if-T2

**4.68 Proposition.** For a Hausdorff space  $X$ , the following conditions are equivalent:

cond:loc-cpt-if-T2

(i) The space  $X$  is locally compact.

pt-if-T2-small-open-nbds-with-cpt-clc

(ii) For each  $x \in X$  and each neighborhood  $U$  of  $x$  there is an open neighborhood  $V$  of  $x$  such that  $\text{cls } V$  is compact and  $\text{cls } V \subset U$ .

cond:loc-cpt-if-T2-and-exists-cpt-nbd

(iii) For each  $x$  in  $X$ , there is some compact neighborhood of  $x$  in  $X$ .

locally compact space! Hausdorff spa

**Proof.** Assume (i). We deduce (ii). Let  $x \in X$  and let  $U$  be an arbitrary neighborhood of  $x$ . By (i), there is a compact neighborhood  $W$  of  $x$  with  $W \subset U$ . Take

$$V = \text{int } W,$$

so that  $V$  is an open neighborhood of  $x$ . Now  $W$ , being a compact subset of the Hausdorff space  $X$ , is closed in  $X$ . Hence  $\text{cls } V \subset \text{cls } W = W \subset U$ , and  $\text{cls } V$  is compact because it is a closed subset of the compact set  $W$ .

Clearly (ii) implies (iii).

Assume (iii). We deduce (i). Let  $x \in X$  and let  $U$  be an arbitrary neighborhood of  $x$ . Without loss of generality we may assume that  $U$  is open. By assumption, there exists some compact neighborhood  $K$  of  $x$  in  $X$ . Since  $X$  is a Hausdorff space,  $K$  is closed in  $X$ .

The complement  $K \setminus U$  is a closed subset of the compact set  $K$ , and so  $K \setminus U$  is itself compact. By Lemma 4.18, there are disjoint open sets  $M$  and  $N$  in  $X$  with

$$x \in M, \quad K \setminus U \subset N.$$

Let

$$V = M \cap \text{int } K.$$

We are going to show that  $\text{cls } V$  is a compact neighborhood of  $x$  with  $\text{cls } V \subset U$ . Certainly  $\text{cls } V$  is a neighborhood of  $x$  because  $V$  is an open neighborhood of  $x$ .

Since  $V$  is a subset of  $K$  that is disjoint from  $K \setminus U$ ,

$$V \subset U.$$

Since  $V \subset K$  and  $K$  is closed in  $X$ ,

$$\text{cls } V \subset K.$$

Since  $V \subset M$  and  $M$  is disjoint from  $N$ ,

$$(\text{cls } V) \cap (K \setminus U) \subset (\text{cls } V) \cap N.$$

But  $\text{cls } V$  is disjoint from  $N$ ; in fact, if there is some  $y \in (\text{cls } V) \cap N$ , then the neighborhood  $N$  of  $y$  must meet  $V$ , and this is impossible because  $V \subset M$  whereas  $M$  is disjoint from  $N$ . Hence  $(\text{cls } V) \cap (K \setminus U) = \emptyset$ , and so

$$\text{cls } V \subset U. \quad \square$$

The main import of the preceding proposition is that a Hausdorff space will be locally compact as soon as each of its points has *at least one* compact neighborhood. Since a topological space is itself a neighborhood of each of its points, we obtain in particular the following corollary.

**4.69 Corollary.** A compact Hausdorff space is locally compact.

*Hilbert sequence space*  
*cor: cpt-ℓ<sup>2</sup> is loc-cpt*  
*finite-complement topology: locally compact space @ and locally compact space*  
*rational numbers! locally compact space @ and locally compact space*

**4.70 Examples.** (1) Any discrete space  $X$  is locally compact, since for  $x \in X$  the singleton  $\{x\}$  is a neighborhood of  $x$ . Hence any infinite discrete space ( $\mathbb{Z}$ , for example) is a locally compact Hausdorff space that is not compact.

*ex: R<sup>n</sup>-loc-cpt* (2) Euclidean space  $\mathbb{R}^n$ , which is a noncompact Hausdorff space, is locally compact. In fact, if  $d$  is the Euclidean metric on  $\mathbb{R}^n$ , then for each  $x \in \mathbb{R}^n$  and each  $\varepsilon > 0$ , the  $d$ -disk  $D_\varepsilon(x; d)$  is a neighborhood of  $x$  that, by [Theorem 4.34](#), is compact.

*Hilbert sequence space is not loc-cpt* (3) The Hilbert sequence space  $\ell^2$  ([Example 1.10](#)) is *not* locally compact, for no neighborhood of a point in  $\ell^2$  is compact. To see why, just suppose some point

$$x = \langle x_1, x_2, x_3, \dots \rangle \in \ell^2$$

had a compact neighborhood. Then for sufficiently small  $\varepsilon > 0$ , the closed neighborhood  $D_\varepsilon(x; d_2)$  of  $x$  would also be compact.

We show, to the contrary, that the disk  $D_\varepsilon(x; d_2)$  is not compact. In view of [Theorem 4.51](#), it suffices to construct a sequence  $\langle x^n \rangle_{n \in \mathbb{N}}$  in  $D_\varepsilon(x; d_2)$  that does not cluster in  $\ell^2$ . (Here we are using superscripts to denote entries in the sequence of points because subscripts are used to denote coordinates of points in  $\ell^2$ .) Define the sequence  $\langle x^n \rangle_{n \in \mathbb{N}}$  by

$$\begin{aligned} x^0 &= x^1 = \langle x_1 + \sqrt{\varepsilon/2}, x_2, x_3, \dots \rangle, \\ x^2 &= \langle x_1, x_2 + \sqrt{\varepsilon/2}, x_3, \dots \rangle, \\ &\vdots \\ x^n &= \langle x_1, x_2, \dots, x_{n-1}, x_n + \sqrt{\varepsilon/2}, x_{n+1}, \dots \rangle \\ &\vdots \end{aligned}$$

Then  $d_2(x^n, x^m) = \sqrt{\varepsilon}$  for all  $n, m \geq 1$  with  $n \neq m$ . Hence the sequence  $\langle x^n \rangle_{n \in \mathbb{N}}$  cannot cluster in  $\ell^2$ .

*ex: finite-compl-top-loc-cpt* (4) Let  $X$  be an infinite set provided with its finite-complement topology [[Examples 2.3 \(7\)](#)]. By [Examples 4.6 \(8\)](#),  $X$  is a compact space that is not a Hausdorff space. Nevertheless,  $X$  is locally compact. In fact, every subset  $A$  of  $X$  is compact because the relative topology on  $A$  is the finite-complement topology on  $A$ .

*ex: Q-not-loc-cpt* (5) The space  $\mathbb{Q}$  of rational numbers is not locally compact. In fact, by the Heine–Borel–Lebesgue Theorem ([4.15](#)), no closed neighborhood in  $\mathbb{Q}$  of the form  $\mathbb{Q} \cap [x - \varepsilon, x + \varepsilon]$  with  $\varepsilon > 0$  can be compact.

Likewise, the subspace  $\mathbb{Q} \cap [0, 1]$  is not locally compact.

*ex: Q-with-cofinite-at-added-pt* (6) Let  $z$  be an object with  $z \notin \mathbb{Q}$  and let  $X = \mathbb{Q} \cup \{z\}$ . Define

$$\mathcal{N}_z = \{X \setminus F : F \subset \mathbb{Q}, F \text{ is finite}\},$$

and for each  $x \in \mathbb{Q}$  define

$$\begin{aligned} \mathcal{N}_x &= \{V : V \text{ is a neighborhood of } x \text{ in } \mathbb{Q}\} \\ &\cup \{V \cup \{z\} : V \text{ is a neighborhood of } x \text{ in } \mathbb{Q}\}. \end{aligned}$$

It is easy to check that the conditions (N1)–(N5) of [Theorem 2.18](#) are satisfied. Then [Theorem 2.19](#) furnishes a topology on  $X$  such that for each  $y \in X$  the collection



$\mathcal{N}_y$  is the neighborhood system at  $y$ . Provided with this topology, the space  $X$  is compact—the argument is like that in [Examples 4.6 \(8\)](#)—and so each point of  $X$  has a compact neighborhood. However,  $X$  is not locally compact because no point  $x \in \mathbb{Q}$  has arbitrarily small compact neighborhoods in  $X$ .  $\diamond$

Additional examples are furnished by the following proposition.

**4.71 Proposition.** *A subspace of a locally compact space  $X$  that is either open or closed in  $X$  is itself locally compact.*

**Proof.** Let  $Y$  be a subset of a locally compact space  $X$ , let  $y \in Y$ , and let  $U$  be an arbitrary neighborhood of  $y$  in  $Y$ .

Case (1):  $Y$  is open in  $X$ .

In this case  $U$  is also a neighborhood of  $y$  in  $X$ , and so there is some compact neighborhood  $V$  of  $y$  in  $X$  with  $V \subset U$ . Since  $V \subset Y$ , then  $V$  is also a neighborhood of  $y$  in  $Y$ .

Case (2):  $Y$  is closed in  $X$ .

Choose a neighborhood  $U_0$  of  $y$  in  $X$  such that  $U = U_0 \cap Y$ . There is some compact neighborhood  $V_0$  of  $y$  in  $X$  with  $V_0 \subset U_0$ . Set

$$V = V_0 \cap Y.$$

Then  $V$  is a neighborhood of  $y$  in  $Y$  with  $V \subset U$ . Now  $V_0$  is compact, and  $V$  is closed in  $V_0$  because  $Y$  is closed in  $X$ . Hence  $V$  is compact.  $\square$

The preceding result is the analog of [Theorem 4.11](#) for locally compact spaces. However, [Theorems 4.12](#) and [4.21](#) and [Corollaries 4.24](#) and [4.25](#) do *not* admit direct analogs for locally compact spaces: *a locally compact subspace of a Hausdorff space need not be closed*, and *the continuous image of a locally compact space need not be locally compact*.

**4.72 Examples.** (1) By [Examples 4.70 \(2\)](#) and [Proposition 4.71](#), the open interval  $]0, 1[$  in  $\mathbb{R}$  is a locally compact subspace of the Hausdorff space  $\mathbb{R}$ , but  $]0, 1[$  is *not* closed in  $\mathbb{R}$ .

(2) Let  $X$  be the set of rational numbers provided with its discrete topology and let  $Y$  be the same set but provided with its Euclidean topology. Then  $X$  is a locally compact Hausdorff space,  $Y$  is a Hausdorff space, and the identity map  $f: X \rightarrow Y$  is continuous. By [Examples 4.70 \(5\)](#), the image  $Y$  of  $X$  under  $f$  is *not* locally compact, and the map  $f$  is *not* closed.  $\diamond$

### Local compactness and continuity

Although the image of a locally compact space under an arbitrary continuous map thus need not be locally compact, we do have a limited result in this direction.

**4.73 Proposition.** *Let  $f: X \rightarrow Y$  be a continuous open surjection from a locally compact space  $X$  to a topological space  $Y$ . Then  $Y$  is locally compact.*

**Proof.** Let  $V$  be an arbitrary neighborhood of a point  $y \in Y$ . Choose some  $x \in X$  with  $f(x) = y$ . Then  $f^{-1}(V)$  is a neighborhood of  $x$  in  $X$ , and so there is some compact neighborhood  $W$  of  $x$  in  $X$  with  $W \subset f^{-1}(V)$ . Since  $f$  is an open map, the image  $f(W)$  is

real line with integers collapsed to a point, a neighborhood of  $y$  in  $Y$ , and clearly  $f(W) \subset V$ . Finally,  $f(W)$  is compact according to Theorem 4.21.  $\square$

It follows from the preceding proposition that **local compactness is a topological property**.

Despite Proposition 4.73, local compactness is *not* in general preserved under the formation of quotients. In fact, a closed continuous image of a locally compact space need not itself be locally compact.

ex:RcollapseZ-not-loc-cpt **4.74 Example.** Consider again the quotient space  $\mathbb{R}/\mathbb{Z}$  obtained from the real line  $\mathbb{R}$  by collapsing the set  $\mathbb{Z}$  of integers to a point (Example 3.83). Although  $\mathbb{R}$  is a locally compact Hausdorff space and, from Example 3.86, the quotient  $\mathbb{R}/\mathbb{Z}$  is also a Hausdorff space, the quotient map  $\mathbb{R} \rightarrow \mathbb{R}/\mathbb{Z}$  is a closed map, but  $\mathbb{R}/\mathbb{Z}$  is *not locally compact*. (The proof is requested in Exercise 110.)

In fact, just suppose there was a compact neighborhood  $K$  of the point  $\mathbb{Z}$  of  $\mathbb{R}/\mathbb{Z}$ . Let  $W$  be an open neighborhood of  $\mathbb{Z}$  in  $\mathbb{R}/\mathbb{Z}$  with  $W \subset K$ . Let  $q: \mathbb{R} \rightarrow \mathbb{R}/\mathbb{Z}$  be the quotient map. Then  $W = (V \setminus \mathbb{Z}) \cup \{\mathbb{Z}\}$  where  $V = q(U)$  for some open  $U \subset \mathbb{R} \setminus \mathbb{Z}$ . Want actually  $\mathbb{Z} \subset U$  in  $\mathbb{R}$ !! For each  $n \in \mathbb{Z}$  there is some  $\varepsilon_n$  with  $0 < \varepsilon_n < 1$  and  $]n - \varepsilon_n, n + \varepsilon_n[ \subset U$ . Define  $U' = \bigcup_{n \in \mathbb{Z}} ]n - \varepsilon_n/2, n + \varepsilon_n/2[$ . Further, define  $U'' = q(U')$ , so that  $U''$  is open in  $\mathbb{R}/\mathbb{Z}$ . then  $\{U''\} \cup \{]n, n + 1[ : n \in \mathbb{Z}\}$  is an open cover of  $K$ . By compactness of  $K$ , there is a positive integer  $m$  such that the collection  $\{U''\} \cup \{]n, n + 1[ : n \in \mathbb{Z}, |n| \leq m\}$  still covers  $K$ . But this is impossible because since this collection does not even cover  $W$ .  $\diamond$

Local compactness is preserved under the formation of finite products.

thm:finite-prod-loc-cpt **4.75 Theorem.** *The product of a finite family of nonempty topological spaces is locally compact if and only if each member of the family is locally compact.*

**Proof.** Let  $\langle X_i \rangle_{i=1,2,\dots,n}$  be a finite family of topological spaces and let  $X = \prod_{i=1}^n X_i$ .

Assume first that  $X$  is locally compact. For each index  $i$ , the  $i$ th projection  $p_i: X \rightarrow X_i$  is a continuous open surjection, and so from Proposition 4.73 the space  $X_i$  is locally compact.

Conversely, assume that  $X_i$  is locally compact for each  $i$ . Let  $U$  be an arbitrary neighborhood of a point  $x = \langle x_1, x_2, \dots, x_n \rangle \in X$ . There are neighborhoods  $V_1$  of  $x_1$  in  $X_1$ ,  $V_2$  of  $x_2$  in  $X_2$ ,  $\dots$ ,  $V_n$  of  $x_n$  in  $X_n$  such that  $V_1 \times V_2 \times \dots \times V_n \subset U$ . By assumption, for each  $i = 1, 2, \dots, n$  there is a compact neighborhood  $W_i$  of  $x_i$  in  $X_i$  such that  $W_i \subset V_i$ . Then  $W = W_1 \times W_2 \times \dots \times W_n$  is a neighborhood of  $x$  in  $X$  with  $W \subset U$ , and  $W$  is compact by the Tychonoff Product Theorem—finite case (Theorem 4.32).  $\square$

Unlike compactness, local compactness is *not* preserved under the formation of products of arbitrarily many spaces—not even of denumerably many spaces.

thm:denumerable-prod-loc-cpt **4.76 Theorem.** *Let  $\langle X_i \rangle_{i=1,2,3,\dots}$  be a sequence of nonempty metrizable spaces. Then the product space  $\prod_{i=1}^{\infty} X_i$  is locally compact if and only if  $X_i$  is locally compact for every  $i$  and  $X_i$  is compact for almost all  $i$ .*

Recall that “almost all” means “all except finitely many.”

**Proof.** Let  $X = \prod_{i=1}^{\infty} X_i$ . Assume first that  $X$  is locally compact. That  $X_i$  is locally compact for all  $i \in I$  follows exactly as in the proof of Theorem 4.75. Next, choose some point  $x \in X$

and some compact neighborhood  $V$  of  $x$  in  $X$ . Replacing  $V$  if necessary by a smaller closed neighborhood of  $x$ , we may assume without loss of generality that  $V$  has the form

$$V = V_1 \times V_2 \times \cdots \times V_n \times X_{n+1} \times X_{n+2} \times \cdots$$

for some  $n \geq 1$  and some neighborhoods  $V_1$  of  $x_1$  in  $X_1$ ,  $V_2$  of  $x_2$  in  $X_2$ ,  $\dots$ ,  $V_n$  of  $x_n$  in  $X_n$ . Then for all  $i > n$ , the  $i$ th projection  $p: X \rightarrow X_i$  maps the compact set  $V$  onto  $X_i$ , whence  $X_i$  is compact.

Conversely, assume that  $X_i$  is locally compact for all  $i$  and is compact for almost all  $i$ . Choose  $n \geq 1$  such that  $X_i$  is compact for all  $i > n$ . Let  $x \in X$  and let  $U$  be an arbitrary neighborhood of  $x$  in  $X$ . There is an  $m \geq n$  such that  $U$  contains a neighborhood  $V$  of  $x$  of the form

$$V = V_1 \times V_2 \times \cdots \times V_m \times X_{m+1} \times X_{m+2} \times \cdots,$$

where  $V_1$  is a neighborhood of  $x_1$  in  $X_1$ ,  $V_2$  is a neighborhood of  $x_2$  in  $X_2$ ,  $\dots$ ,  $V_m$  is a neighborhood of  $x_m$  in  $X_m$ . By assumption, for each  $i = 1, 2, \dots, m$  there is a compact neighborhood  $W_i$  of  $x_i$  in  $X_i$  such that  $W_i \subset V_i$ . Set

$$W = W_1 \times W_2 \times \cdots \times W_m \times X_{m+1} \times X_{m+2} \times \cdots.$$

Then  $W$  is a neighborhood of  $x$  in  $X$  such that  $W \subset V \subset U$ . From the Tychonoff Product Theorem—denumerable metrizable case (Theorem 4.61), it follows that  $W$  is compact.  $\square$

The only place metrizability of the spaces  $X_i$  was exploited in the preceding proof was in applying the denumerable metric case (Theorem 4.61) of the Tychonoff Product Theorem. Since the product of a family of compact spaces is compact in general, without any assumption of metrizability or any assumption about the size of the index set, so is the analog of Theorem 4.76.

thm:prod-loc-cpt

**4.77 Theorem.** Let  $\langle X_i \rangle_{i \in I}$  be a family of nonempty topological spaces. Then the product space  $\prod_{i \in I} X_i$  is locally compact if and only if  $X_i$  is locally compact for every  $i$  in  $I$  and  $X_i$  is compact for almost all  $i$  in  $I$ .

**Proof.** The proof is similar to that of Theorem 4.76 except that it uses the general Tychonoff Product Theorem (4.33) instead of its denumerable metrizable case (Theorem 4.61). See Exercise 127.  $\square$

According to Theorems 4.76 and 4.77, the product of denumerably many copies of  $\mathbb{R}$  is *not* locally compact. Actually, this product is homeomorphic to the Hilbert sequence space  $\ell^2$ —see Examples 3.64 (8)—and we have already proved directly that  $\ell^2$  is not locally compact [Examples 4.70 (3)].

### One-point compactification

The very definition of locally compact spaces refers to compact spaces, and the preceding theorems about locally compact spaces are proved by using corresponding theorems about compact ones. We now look at another way that locally compact spaces are related to compact spaces.

For motivation, consider the locally compact space  $\mathbb{R}$ . Although  $\mathbb{R}$  itself is noncompact, it is a dense subspace of the compact space  $\widehat{\mathbb{R}}$ , the extended real line—see Example 1.41, Proposition 1.42, and Examples 2.43 (4)—and the complement  $\widehat{\mathbb{R}} \setminus \mathbb{R} = \{-\infty, +\infty\}$ . Thus by adjoining *two* points to the locally compact space  $\mathbb{R}$  we obtain a compact space. But we can do even better than this: we can obtain a compact space in which  $\mathbb{R}$  is dense by adjoining just *one* point to  $\mathbb{R}$ .

locally compact space! product space!  
product space! locally compact space!  
one-point compactification

To see how, form the quotient space

$$Z = \widehat{\mathbb{R}} / \{-\infty, +\infty\}$$

obtained from  $\widehat{\mathbb{R}}$  by collapsing the closed two-element subset  $\{-\infty, +\infty\}$  of  $\widehat{\mathbb{R}}$  to a single point. (Imagine grabbing hold of the extended real line at its ends  $-\infty$  and  $+\infty$  and pulling them together, as suggested in [Figure 4.11](#).) Denote the point  $\{-\infty, +\infty\}$  of  $Z$  by  $\infty$  (with



Figure 4.11: Collapsing the “ends” of the extended real line to a single point.

fig:collapse-extended-R-to-1-pt-cpt

no plus or minus sign). The space  $Z$  is compact because  $\widehat{\mathbb{R}}$  is. By [Examples 3.81 \(9\)](#), the quotient map

$$q: \widehat{\mathbb{R}} \rightarrow Z$$

maps  $\mathbb{R}$  homeomorphically onto the open subspace

$$q(\mathbb{R}) = Z \setminus \{\infty\}$$

of  $Z$ . Moreover, the homeomorphic image  $q(\mathbb{R})$  of  $\mathbb{R}$  is dense in  $Z$  since  $\mathbb{R}$  is dense in  $\widehat{\mathbb{R}}$ .

As in [Examples 3.47 \(2\)](#), we may now regard  $\mathbb{R}$  as actually being a subspace of  $Z$ , namely, by replacing  $q(\mathbb{R})$  with its homeomorphic copy  $\mathbb{R}$ . With this replacement,  $\mathbb{R}$  becomes a dense open subspace of  $Z$  whose complement in  $Z$  consists of the single point  $\{\infty\}$ . Since  $\mathbb{R}$  is an open subspace of  $Z$ , the open subsets of  $Z$  that *do not* contain  $\infty$  are precisely the open subsets of the real line  $\mathbb{R}$ .

Let us next determine the open subsets of  $Z$  that *do* contain  $\infty$ . Let  $U$  be an open set in  $Z$  with  $\infty \in U$ . Then  $V = q^{-1}(U)$  is an open set in  $\widehat{\mathbb{R}}$  with  $-\infty, +\infty \in V$ . By [Lemma 1.43](#), the set  $[-\infty, a) \cup (b, +\infty] \subset V$  for suitable real  $a$  and  $b$  with  $a < b$ . Then the closed subset  $\mathbb{R} \setminus V$  of  $\mathbb{R}$  satisfies  $\mathbb{R} \setminus V \subset [a, b]$ . Hence  $\mathbb{R} \setminus V$  is compact. Setting

$$K = q(Z \setminus \mathbb{R}),$$

we see that  $K$  is a compact subset of  $Z \setminus \{\infty\} = \mathbb{R}$  with

$$(*) \quad U = Z \setminus K.$$

Now  $Z$  is easily seen to be a Hausdorff space, and so conversely each set  $U$  of this form  $(*)$  is an open subset of  $Z$  containing  $\infty$ .

To summarize,

$$Z = \mathbb{R} \cup \{\infty\}, \quad \infty \notin \mathbb{R},$$

and the topology of  $Z$  is just

$$\{U : U \text{ is open in } \mathbb{R}\} \cup \{Z \setminus K : K \text{ is a compact subset of } \mathbb{R}\}.$$

With this topology,  $Z$  is a compact space that contains  $\mathbb{R}$  as a dense subspace.

This description of the compact space obtained by adjoining a single point to  $\mathbb{R}$  can immediately be generalized if we start with an arbitrary locally compact Hausdorff space.

**4.78 Theorem (Alexandroff one-point compactification).** *Let  $X$  be a noncompact, locally compact Hausdorff space whose topology is  $\mathcal{T}$ . Set*

$$X_\infty = X \cup \{\infty\}$$

*where  $\infty \notin X$ . Then the collection*

$$\mathcal{T}_\infty = \mathcal{T} \cup \{X_\infty \setminus K : K \text{ is a compact subset of } X\}$$

*is a topology on  $X_\infty$  making  $X_\infty$  a compact Hausdorff space in which  $X$  (with its original topology) is a dense subspace.*

thm:1-pt-cptn

**Proof.** Note first that if  $V \in \mathcal{T}$  and if  $K$  is a compact subset of  $X$ , then  $V \setminus K \in \mathcal{T}$  because  $X$  is a Hausdorff space.

$\mathcal{T}_\infty$  is a topology on  $X_\infty$ : Clearly  $\emptyset \in \mathcal{T}_\infty$  and  $X_\infty \in \mathcal{T}_\infty$ . We leave as an exercise verifying that the union of any collection of members of  $\mathcal{T}_\infty$  again belongs to  $\mathcal{T}_\infty$ . Now let  $V_1, V_2 \in \mathcal{T}_\infty$ . To show that  $V_1 \cap V_2 \in \mathcal{T}_\infty$ , we consider three possibilities.

If  $V_1, V_2 \in \mathcal{T}$ , then  $V_1 \cap V_2 \in \mathcal{T} \subset \mathcal{T}_\infty$ .

If  $V_1 \in \mathcal{T}$  but  $V_2 \notin \mathcal{T}$ , then  $V_2 = X_\infty \setminus K$  for some compact  $K \subset X$ , and so

$$V_1 \cap V_2 = V_1 \cap (X_\infty \setminus K) = V_1 \setminus K \in \mathcal{T} \subset \mathcal{T}_\infty.$$

(If  $V_2 \in \mathcal{T}$  but  $V_1 \notin \mathcal{T}$ , then just reverse the roles of  $V_1$  and  $V_2$  in the preceding sentence.)

Finally, if  $V_1 \notin \mathcal{T}$  and  $V_2 \notin \mathcal{T}$ , then  $V_1 = X_\infty \setminus K_1$ ,  $V_2 = X_\infty \setminus K_2$  for some compact sets  $K_1, K_2 \subset X$ , and so

$$V_1 \cap V_2 = (X_\infty \setminus K_1) \cap (X_\infty \setminus K_2) = X_\infty \setminus (K_1 \cup K_2) \in \mathcal{T}_\infty$$

since  $K_1 \cup K_2$  is compact.

$\langle X, \mathcal{T} \rangle$  is a subspace of  $\langle X_\infty, \mathcal{T}_\infty \rangle$ : Since  $\mathcal{T} \subset \mathcal{T}_\infty$ , we need only show that the set  $V \cap X \in \mathcal{T}$  for each  $V \in \mathcal{T}_\infty$ . Let  $V \in \mathcal{T}_\infty$ . If  $V \in \mathcal{T}$ , we are done. Otherwise, if  $V \notin \mathcal{T}$ , write  $V = X_\infty \setminus K$  for some compact  $K \subset X$ ; then  $V \cap X = X \setminus K \in \mathcal{T}$ .

$X_\infty$  is compact: Let  $\mathcal{U}$  be an open cover of  $X_\infty$ . Choose some  $U_\infty \in \mathcal{U}$  for which  $\infty \in U_\infty$ . We have

$$U_\infty = X_\infty \setminus K$$

for some compact  $K \subset X$ . Since  $\mathcal{U}$  covers  $K$ , some finite  $\mathcal{E} \subset \mathcal{U}$  also covers  $K$ . Then  $\mathcal{E} \cup \{U_\infty\}$  is a finite cover of  $X_\infty$  contained in  $\mathcal{U}$ .

$X_\infty$  is a Hausdorff space: Since  $\mathcal{T} \subset \mathcal{T}_\infty$ , any two distinct points of  $X$  have disjoint neighborhoods in  $X_\infty$ . Now let  $x \in X_\infty$  with  $x \neq \infty$ , that is,  $x \in X$ . By hypothesis there is at least one compact neighborhood  $W$  of  $x$  in  $X$ . Then  $X_\infty \setminus W$  and  $W$  are disjoint neighborhoods of  $\infty$  and  $x$ , respectively, in  $X_\infty$ .

$X$  is dense in  $X_\infty$ : Let  $V$  be an arbitrary nonempty open subset of  $X_\infty$ . If, on the one hand,  $V \in \mathcal{T}$ , then  $V \cap X = V \neq \emptyset$ . If, on the other hand,  $V \notin \mathcal{T}$ , then

$$V = X_\infty \setminus K$$

for some compact subset  $K$  of  $X$ , the set  $K \neq X$  because  $X$  is not compact, and hence

$$V \cap X = X \setminus K \neq \emptyset. \quad \square$$

def:1-pt-cptn

**4.79 Definition.** The space  $X_\infty$  constructed in [Theorem 4.78](#) is called the **one-point compactification of  $X$** , and  $\infty$  is called the **point at infinity in  $X_\infty$** . The same space is also called the **Alexandroff compactification of  $X$**  and then is typically denoted by  $\alpha(X)$  or  $\alpha X$ .

Is there another way to construct a compact Hausdorff space in which  $X$  is a dense subspace whose complement is a single point? No, not really: the one-point compactification of  $X_\infty$  as just constructed is, up to homeomorphism, the only such space.

prop:1-pt-cptn-only-such

**4.80 Proposition.** Let  $X$  be a noncompact, locally compact Hausdorff space. Suppose  $Y$  is a compact Hausdorff space and  $p \in Y$  is a point with

$$X \cong Y \setminus \{p\}.$$

Then there is a homeomorphism

$$g: X_\infty \cong Y$$

with

$$g(\infty) = p.$$

**Proof.** Let  $f: X \cong Y \setminus \{p\}$ . Let  $g: X_\infty \rightarrow Y$  be the domain-codomain extension of  $f$  for which  $g(\infty) = p$ , so that  $g$  is bijective. Since  $X_\infty$  is compact and  $Y$  is a Hausdorff space, in view of continuous bijection principle ([Corollary 4.25](#)), it remains only to show that  $g$  is continuous.

Let  $x \in X$ . We show that  $g$  is continuous at  $x$ . Let  $V$  be an arbitrary open neighborhood of  $g(x)$  in  $Y$ . Since  $g(x) = f(x) \neq p$ , there is an open neighborhood  $W$  of  $g(x)$  in  $Y$  with  $p \notin W \subset V$ . Then the set  $U = f^{-1}(W)$  is an open neighborhood of  $x$  in  $X$ , and hence in  $X_\infty$ , for which  $g(U) = f(U) \subset V$ .

We show that  $g$  is continuous at  $\infty$ . Let  $V$  be an arbitrary open neighborhood of  $p = g(\infty)$  in  $Y$ . Now  $Y \setminus V$  is compact, and so the set

$$K = f^{-1}(Y \setminus V)$$

is a compact subset of  $X$ . Then the set

$$U = X_\infty \setminus K$$

is an open neighborhood of  $\infty$  in  $X_\infty$  for which  $g(U) \subset V$ .  $\square$

Observe that, in the notation of [Proposition 4.80](#), there is a *unique* homeomorphism  $g: X_\infty \cong Y$  extending a given homeomorphism  $X \cong Y \setminus \{p\}$  for which  $g(\infty) = p$ .

It is at least intuitively evident that the space  $Z$  constructed from  $\mathbb{R}$  at the beginning of this subsection is, in fact, a circle. so that the one-point compactification of the real line “is” (that is, is homeomorphic to) the circle  $S_1$ . Moreover, this example generalizes to higher dimensions.

ex:1-pt-compactification-of-Rn

**4.81 Example.** For each integer  $n \geq 1$ , the one-point compactification  $(\mathbb{R}^n)_\infty$  of Euclidean  $n$ -space  $\mathbb{R}^n$  is homeomorphic to the  $n$ -sphere  $S_n$ .

In fact let  $\mathbf{p}$  be the north pole  $(0, 0, \dots, 0, 1) \in S_n$ . The earlier discussion of the stereographic projection [Examples 3.25 (13)] established that

$$\mathbb{R}^n \cong S_n \setminus \{\mathbf{p}\}.$$

It follows from Examples 4.35 (5) and Proposition 4.80 that

$$(\mathbb{R}^n)_\infty \cong S_n. \quad \diamond$$

The case  $n = 2$  is of particular importance in complex analysis, for as a topological space, the sphere  $\mathbb{R}^2$  is the complex plane  $\mathbb{C}$ . In this context the one-point compactification of  $\mathbb{R}^2$  is called the *Riemann sphere*. (The Alexandroff one-point compactification of an arbitrary locally compact space is the natural generalization of the Riemann sphere.) We can often describe the behavior of a complex function  $f: \mathbb{C} \rightarrow \mathbb{C}$  at complex numbers  $z$  of large modulus  $|z|$  in terms of the behavior at the point  $\infty \in \mathbb{C}_\infty$  of a suitable extension of  $f$  to the Riemann sphere (see (124)–(126)).

n-sphere@n-sphere!one-point compactification of  $\mathbb{R}^n$   
Riemann sphere  
Riemann, Bernhard

### Local compactness and countability

subsec:loc-cpt-count

The development in Section 4.2 indicates that, among all compact spaces, the metrizable ones are especially nice. We noted there that the second-countable compact Hausdorff spaces are precisely the compact metrizable spaces. It is therefore of interest to know that the one-point compactification of a second-countable locally compact Hausdorff space is itself second-countable. To prove this, we need a bit of preparation.

loc-cpt-T2-has-open-nbd-with-cpt-cls

**4.82 Lemma.** Let  $K$  be a compact subset of a locally compact Hausdorff space  $X$ . Then  $K \subset U$  for some open subset  $U$  of  $X$  with  $\text{cls } U$  compact.

**Proof.** For each  $x \in X$  there is, by Proposition 4.68, an open set  $U_x$  in  $X$  with  $x \in U_x$  and  $\text{cls } U_x$  compact. Since  $K$  is compact, it has a finite subset  $F$  such that  $K \subset \bigcup_{x \in F} U_x$ . Then  $U = \bigcup_{x \in F} U_x$  is the desired open set.

□

ng-seq-opens-in-2nd-count-loc-cpt-T2

**4.83 Proposition.** Let  $X$  be a second-countable locally compact Hausdorff space. Then there is a sequence  $\langle U_n \rangle_{n \in \mathbb{N}}$  of open sets in  $X$  such that  $\langle U_n \rangle_{n \in \mathbb{N}}$  covers  $X$  and

$$\text{cls } U_n \subset U_{n+1} \text{ and } \text{cls } U_n \text{ is compact} \quad (n \in \mathbb{N}).$$

**Proof.** For each  $x \in X$  there is, by Proposition 4.68, an open neighborhood  $V_x$  of  $x$  with  $\text{cls } V_x$  compact. By the Lindelöf Theorem (2.84), the open cover  $\{V_x : x \in X\}$  contains a countable cover of  $X$ . Replacing each member of this countable cover by its closure, we obtain a sequence  $\langle K_n \rangle_{n \in \mathbb{N}}$  of compact subsets of  $X$  with

$$X = \bigcup_{n=0}^{\infty} K_n.$$

By Lemma 4.82 there is an open set  $U_0$  in  $X$  such that

$$K_0 \subset U_0, \quad \text{cls } U_0 \text{ is compact.}$$

locally compact space  
topologist's sine curve

Now  $K_1 \cup \text{cls } U_0$  is compact, so by [Lemma 4.82](#) again, there is an open set  $U_1$  in  $X$  such that

$$K_1 \cup \text{cls } U_0 \subset U_1, \quad \text{cls } U_1 \text{ is compact.}$$

Continuing in this way, we construct open sets  $U_1, U_2, U_3, \dots$  in  $X$  such that

$$K_n \cup \text{cls } U_{n-1} \subset U_n, \quad \text{cls } U_n \text{ is compact}$$

for each  $n = 1, 2, 3, \dots$ . As  $\{K_n : n \in \mathbb{N}\}$  is a cover of  $X$ , so is  $\{U_n : n \in \mathbb{N}\}$ .  $\square$

When  $X = \mathbb{R}^k$ , we can construct such a sequence  $\langle U_n \rangle_{n \in \mathbb{N}}$  directly by taking  $U_n = B_{n+1}(0; d)$  with  $d$  being the Euclidean metric.

**4.84 Theorem.** A noncompact, locally compact Hausdorff space is second-countable if and only if its one-point compactification is second-countable.

**Proof.** If  $X_\infty$  is second-countable, then its subspace  $X$  is second-countable, too. Conversely, assume that  $X$  is second-countable. Let  $\langle U_n \rangle_{n \in \mathbb{N}}$  be as in [Proposition 4.83](#). Let  $V$  be an arbitrary open neighborhood of  $\infty$  in  $X_\infty$ . We show that

$$X_\infty \setminus \text{cls } U_m \subset V$$

for some  $m \in \mathbb{N}$ . There is a compact subset  $K$  of  $X$  with

$$V = X_\infty \setminus K.$$

Since  $\langle U_n \rangle_{n \in \mathbb{N}}$  is increasing and  $\langle U_n \rangle_{n \in \mathbb{N}}$  is an open cover of  $K$  in  $X$ , there is an  $m \in \mathbb{N}$  such that  $K \subset U_m$ . Then  $K \subset \text{cls } U_m$ , and so  $X_\infty \setminus \text{cls } U_m \subset X_\infty \setminus K = V$ .

Now let  $\mathcal{U}$  be a countable base of  $X$ . It follows from what has been proved so far that

$$\mathcal{U} \cup \{X_\infty \setminus \text{cls } U_n : n \in \mathbb{N}\}$$

is a countable base of  $X_\infty$ .  $\square$

### EXERCISES FOR SECTION 4.3

**101.** Which of the following spaces are locally compact?

- (a) An uncountable set provided with its countable-complement topology ([Exercise 2.7](#)).
- (b) The Sorgenfrey line [[Examples 2.20 \(1\)](#)].
- (c) The line with two origins [[Examples 2.20 \(3\)](#)].
- (d) The half-disk space [[Examples 2.20 \(3\)](#)].
- (e) The tangent disk space ([Exercise 2.37](#)).
- (f) The cone  $K(\mathbb{Z})$  over the discrete space  $\mathbb{Z}$  (see [Exercise 3.191](#)).

**102.** The **topologist's sine curve** is the subspace  $G \cup L$  of  $\mathbb{R}^2$ , where

$$G = \{(x, \sin(1/x)) : 0 < x \leq 1\},$$

$$L = \{0\} \times [-1, 1] = \{(0, y) : -1 \leq y \leq 1\}$$

(see [Figure 5.11](#), in which different scales are used on the vertical and horizontal axes).

Which of these subspaces of the plane are compact? which are locally compact?



- (a)  $G$ . (b)  $G \cup \{(0, 0)\}$ . (c)  $G \cup L$ .

[Note: Connectedness properties of these spaces are considered in [Examples 5.57 \(2\)](#).]

prob:mfld-loc-cpt **103.** Prove that an arbitrary topological manifold is locally compact.

**104.** Verify that the collection

$$\{\emptyset, ]0, 1[ \} \cup \{ \langle 1/n, 1 \rangle : n = 2, 3, 4, \dots \}$$

is a topology on the open interval  $]0, 1[$ . Then show that  $]0, 1[$ , provided with this topology, is a locally compact space but such that  $\text{cls } U$  is *not* compact for any neighborhood  $U$  of any point of this space.

(Hint: Show that every proper open set in the space is compact but that no nonempty closed set is compact.)

**105.** Verify that the collection

$$\{\emptyset\} \cup \{U : U \subset \mathbb{R}, \mathbb{R} \setminus U \text{ is compact}\}$$

is a topology on  $\mathbb{R}$ . Is  $\mathbb{R}$  provided with this topology compact? Is it locally compact? Is it a Hausdorff space?

**106.** Let  $Y$  be a topological space that is *not* locally compact but is a dense subspace of some compact Hausdorff space  $X$ . Can  $X \setminus Y$  be a singleton? a finite set? a denumerable set?

**107.** Let  $X$  be a Hausdorff space. According to [Proposition 4.71](#), if  $X$  is locally compact, then each subspace  $Y$  of  $X$  of the form  $Y = U \cap E$  with  $U$  open in  $X$  and  $E$  closed in  $X$  must be locally compact. Prove that, conversely, each locally compact subspace  $Y$  of  $X$  has this form.

(Hint: Find an open subset of  $X$  containing  $Y$  in which  $Y$  is closed.)

**108. (a)** Let  $f: X \rightarrow Y$  a continuous *open* surjection whose domain  $X$  is a locally compact Hausdorff space. Prove that  $f$  must be proper ([Exercise 53](#)).

**(b)** Must the map  $f$  still be proper when its locally compact domain is not a Hausdorff space?

**109.** Prove that if a continuous map from a topological space to a locally compact Hausdorff space is proper ([Exercise 53](#)), then it is a closed map.

prob:pf-RcollapseZ-not-loc-cpt **110.** Prove the claim made in [Example 4.74](#), namely, that the quotient space of the real line obtained by collapsing the set of integers to a point is a Hausdorff space, the quotient map is closed, but the quotient is not locally compact. (Hint: See the argument in [Example 3.83](#).)

(Note: Compare [Exercise 111](#).)

ob:certain-quot-by-closed-equiv-rel **111.** Let  $\sim$  be a *closed* equivalence relation ([Exercise 3.199](#)) on a locally compact Hausdorff space  $X$  for which each equivalence class is compact. Prove that the quotient space  $X/\sim$  is then a locally compact Hausdorff space.

[Hint: Use [Exercise 50 \(c\)](#).]

**112.** Let  $X = ]0, 1[$  with its *discrete* topology.

**(a)** Give a topology on  $X \cup \{0\}$  for which the resulting space is homeomorphic to the one-point compactification  $\alpha X$  of  $X$ .

manifold!locally compact space@as  
locally compact space!manifold@an  
locally compact space!open subspac  
locally compact space!closed subspa  
subspace!locally compact space@of  
proper map!locally compact space@  
locally compact space!proper map@  
proper map!locally compact space@  
locally compact space!proper map@  
locally compact space!quotient spac  
collapsing to a point!locally compact  
collapsing to a point!separation prop  
locally compact space!quotient spac  
quotient space!locally compact spac

one-point compactification! (b) If  $f$  is a real-valued continuous function on  ${}^\alpha X$ , show that the set  $\{x \in {}^\alpha X : f(x) \in \mathbb{Z}\}$  is countable.

one-point compactification! integer  $f(x) \in \mathbb{Z}$

n-sphere@ $S_n$ -sphere!quotient of  $n$ -disk@as quotient of  $S_n$ -disk.  
collapsing to a point  
sigma-compact space@ $\sigma$ -compact space  
113. Suppose in Theorem 4.78 that  $X$  is actually a compact Hausdorff space. Show that  $\mathcal{T}_\infty$  is still a topology making  $X_\infty$  a compact Hausdorff space containing  $X$  as a subspace, but that  $X$  is no longer dense in  $X_\infty$ .

114. Let  $X$  be any topological space with topology  $\mathcal{T}$ , let  $X_\infty = X \cup \{\infty\}$  with  $\infty \notin X$ , and let

$$\mathcal{T}_\infty = \mathcal{T} \cup \{X_\infty \setminus K : K \text{ is a compact subset of } X\}.$$

Verify that  $\mathcal{T}_\infty$  is a topology on  $X_\infty$ . Provide  $X_\infty$  with this topology. Prove:

- (a) The original space  $X$  is a subspace of  $X_\infty$ .
  - (b) The space  $X_\infty$  is compact.
  - (c) The space  $X_\infty$  is a Hausdorff space if and only if  $X$  is a locally compact Hausdorff space.
115. (a) Embed the one-point compactification of each of the discrete spaces  $\mathbb{N}$  and  $\mathbb{Z}$  in the real line.
- (b) Embed the one-point compactification of  $]0, 1[ \cup ]2, 3[$  in the plane.
- (c) Can the one-point compactification of  $]0, 1[ \cup ]2, 3[$  be embedded in the real line?

116. Does the real line have a “3-point compactification”?

one-collapsed-to-pt-homeo-sphere 117. Show that the  $n$ -sphere  $S_n$  is homeomorphic to the quotient space of the  $n$ -disk  $D_n$  obtained by collapsing the  $(n-1)$ -sphere  $S_{n-1}$  to a point.

[Hint: First show that  $(D_n // S_{n-1}) \setminus \{S_{n-1}\} \cong B_n$ .]

1-pt-cptn-homeo-vs-spaces-homeo 118. Let  $X$  and  $Y$  be two noncompact, locally compact Hausdorff spaces. Prove:  $X_\infty \cong Y_\infty$  if  $X \cong Y$ . Is the converse true?

prob:1-pt-cptn-punctured-plane 119. Find a suitable quotient space of the 2-sphere  $S_2$  that is homeomorphic to the one-point compactification of the “punctured plane”  $\mathbb{R}^2 \setminus \{0\}$ .

120. Prove: If  $U$  is an open neighborhood of a compact subset  $K$  in a locally compact Hausdorff space  $X$ , then there is an open set  $V$  for which  $K \subset V \subset \text{cls } V \subset U$  and  $\text{cls } V$  is compact.

(Note: This strengthens Lemma 4.82.)

121. In the notation of Proposition 4.83, show that if  $K$  is any compact subset of  $X$ , then  $K \subset U_n$  for some  $n$ .

122. (a) Prove that a noncompact, locally compact Hausdorff space  $X$  is  $\sigma$ -compact (Exercise 71 if and only if there is a countable local base at the point  $\infty$  in  $X_\infty$ . (Hint: The conclusion of Proposition 4.83 still holds if  $X$  is  $\sigma$ -compact but not necessarily second-countable.)

(b) Deduce that a second-countable locally compact Hausdorff space is  $\sigma$ -compact.

123. Let  $f: X \rightarrow Y$  be a continuous map between noncompact, locally compact Hausdorff spaces. Define  $f_\infty: X_\infty \rightarrow Y_\infty$  be the map such that

$$f_\infty(x) = \begin{cases} f(x) & \text{if } x \in X, \\ \infty_Y & \text{if } x = x_\infty, \end{cases}$$

where  $x_\infty$  and  $y_\infty$  are the points-at-infinity in  $X_\infty$  and  $Y_\infty$ , respectively. Prove that  $f$  is proper (Exercise 53) if and only if  $f_\infty$  is continuous.

prob:vanish-at-inf **124.** Let  $X$  be a noncompact, locally compact Hausdorff space. A real-valued function  $f$  on  $X$  is said to **vanish at infinity** when for each  $\varepsilon > 0$  there is a compact subset  $K$  of  $X$  for which  $|f(x)| < \varepsilon$  for all  $x \in X \setminus K$ . limit at infinity  
vanish at infinity  
vanish at infinity

- (a) Exhibit a continuous function  $f: \mathbb{R} \rightarrow \mathbb{R}$  that vanishes at infinity but with  $f(x) \neq 0$  for all  $x \in \mathbb{R}$ . limit at infinity
- (b) Prove that  $f: X \rightarrow \mathbb{R}$  vanishes at infinity precisely when its extension  $f_\infty: X_\infty \rightarrow \mathbb{R}$  with  $f_\infty(\infty) = 0$  is continuous at the point-at-infinity  $\infty$ .
- (c) Show that a continuous function  $f: X \rightarrow \mathbb{R}$  that vanishes at infinity is necessarily bounded. Must such a function attain a maximum or a minimum value on  $X$ ?

**125.** Let  $f: X \rightarrow Y$  be a map between Hausdorff spaces and let  $y \in Y$ . We write “ $\lim_{x \rightarrow \infty} f(x) = y$ ” to mean that for each neighborhood  $V$  of  $y$  in  $Y$ , there is some compact subset  $K$  of  $X$  such that  $f(x) \in V$  for all  $x \in X \setminus K$ .

- (a) When  $Y = \mathbb{R}$ , show that  $f$  vanishes at infinity ([Exercise 124](#)) precisely when  $\lim_{x \rightarrow \infty} f(x) = 0$ .
- (b) Denoting the collection of all compact subsets of  $X$  by  $\mathcal{K}$ , verify that the set

$$I = \{ \langle x, K \rangle : K \in \mathcal{K}, x \in X \setminus K \}$$

is directed by the relation  $\leq$  defined by

$$\langle x, K \rangle \leq \langle y, L \rangle \iff L \subset K.$$

Then show that  $\lim_{x \rightarrow \infty} f(x) = y$  if and only if the net  $\langle f(x) \rangle_{\langle x, K \rangle \in I}$  converge to  $y$  in  $Y$ .

- (c) If  $X$  is noncompact and locally compact, prove that  $\lim_{x \rightarrow \infty} f(x) = y$  if and only if the extension  $f_\infty: X_\infty \rightarrow Y$  of  $f$  with  $f_\infty(\infty) = y$  is continuous at the point  $\infty$  of  $X_\infty$ .
- (d) Explain convergence of sequences in  $Y$  in the context of this exercise. (Hint: Compare [Remark 1.70](#).)

prob:inf-lim-at-inf-loc-cpt **126.** Let  $X$  and  $Y$  be noncompact locally compact Hausdorff spaces and let  $f: X \rightarrow Y$  be a given map. We write “ $\lim_{x \rightarrow \infty} f(x) = \infty$ ” to mean that for each compact  $K \subset Y$ , there is some compact  $F \subset X$  for which  $f(x) \notin K$  for all  $x \notin F$ .

- (a) When  $Y = \mathbb{R}$ , show that  $\lim_{x \rightarrow \infty} f(x) = \infty$  if and only if for each  $c > 0$  there is some compact set  $F \subset X$  for which  $|f(x)| \geq c$  for all  $x \notin F$ . Generalize that result to  $Y = \mathbb{R}^n$ .
- (b) In the case that  $X = Y = \mathbb{R}$ , show that  $\lim_{x \rightarrow \infty} f(x) = \infty$  if and only if  $\lim_{x \rightarrow -\infty} f(x) = \infty = \lim_{x \rightarrow +\infty} f(x)$  in the usual calculus meaning of the latter two limits. (Note: Compare [Examples 3.126](#) and [Exercise 3.251](#).)
- (c) Prove that, in general,  $\lim_{x \rightarrow \infty} f(x) = \infty$  if and only if  $f^{-1}(K)$  is compact for each compact subset  $K$  of  $Y$ .
- (d) Prove also that  $\lim_{x \rightarrow \infty} f(x) = \infty$  if and only if the domain-codomain extension  $f_\infty: X_\infty \rightarrow Y_\infty$  of  $f$  with  $f_\infty(\infty) = \infty$  is continuous at the point  $\infty$  of  $X_\infty$ .
- (e) Take again  $Y = \mathbb{R}$ . Suppose that  $f: X \rightarrow \mathbb{R}$  is continuous and  $\lim_{x \rightarrow \infty} f(x) = \infty$ . Must  $f$  attain a minimum value on  $X$ ? a maximum value?

prob:prod-loc-cpt **127.** Prove in detail the generalization [Theorem 4.77](#) of [Theorems 4.75](#) and [4.76](#).

- prob: Baire category thm 128. Establish the following analog of the Baire Category Theorem (1.91): The intersection of any sequence of dense open subsets of a locally compact Hausdorff space is itself dense in the space.
- pointwise bounded family Urysohn's lemma compactly generated space [Hint: Mimic the construction used in the proof of the Baire Category Theorem, replacing  $d$ -balls by open neighborhoods and  $d$ -disks by compact closures of open neighborhoods. What is a substitute for the Nested Set Theorem (1.88)?]
129. Show that any denumerable locally compact Hausdorff space has a subspace homeomorphic to the discrete space  $\mathbb{N}$ .
- (Hint: Prove that the set  $E$  of all isolated points of  $X$  is denumerable by applying Exercise 128 to the subspace  $X \setminus E$  of  $X$ .)
130. Let  $\langle f_i \rangle_{i \in I}$  be a nonempty family of continuous real-valued functions on a locally compact Hausdorff space  $X$ . Suppose this family is “pointwise bounded” in the sense that the set  $\{f_i(x) : i \in I\}$  of real numbers is bounded for each individual  $x \in X$ . Prove that there exists a nonempty open set  $U$  in  $X$  and a constant  $c > 0$  such that  $|f_i(x)| \leq c$  for all  $x \in U$  and all  $i \in I$ .
- (Hint: Consider the sets  $A_n = \{x \in X : |f_i(x)| \leq n \text{ for all } i \in I\}$  for all  $n \in \mathbb{N}$ .)
131. It is known that for any two disjoint closed subsets  $A$  and  $B$  of a compact Hausdorff space  $Y$  there is a continuous function  $g : Y \rightarrow [0, 1]$  with  $g(a) = 0$  for all  $a \in A$  and  $g(b) = 1$  for all  $b \in B$  [compare Exercise 58 and see Urysohn's Lemma (6.26)]. Assuming this, prove that for any two disjoint subsets  $K$  and  $E$  of  $X$  with  $K$  compact and  $E$  closed in  $X$ , there is a continuous function  $f : X \rightarrow [0, 1]$  with  $f(x) = 0$  for all  $x \in K$  and  $f(x) = 1$  for all  $x \in E$ .
- locally-cpt-imply-cptly-generated 132. Prove that every locally compact space is compactly generated in the sense of Exercise 56.

## CHAPTER

# 5

## Connectedness

chap:connected

### Chapter Contents

5.1 Connected Spaces . . . . .	524	Paths . . . . .	558
Connected spaces . . . . .	524	Path-connectedness . . . . .	561
Connected subsets of topological spaces . . . . .	526	Curves . . . . .	564
Continuous maps on connected spaces . . . . .	527	Exercises for Section 5.3 . . . . .	570
Covering by connected sets . . . . .	532	5.4 Homotopy . . . . .	571
Products of connected spaces . . . . .	536	Contractible spaces . . . . .	572
Exercises for Section 5.1 . . . . .	541	Retracts and deformation retracts . . . . .	574
5.2 Components and Locally Connected Spaces . . . . .	545	Homotopies . . . . .	576
Components . . . . .	545	Path-homotopies . . . . .	577
Totally disconnected spaces . . . . .	549	Exercises for Section 5.4 . . . . .	585
Locally connected spaces . . . . .	552	5.5 Simply connected spaces . . . . .	587
Exercises for Section 5.2 . . . . .	555	Simple connectedness . . . . .	587
5.3 Path-Connected Spaces . . . . .	558	The circle and simple connectedness . . . . .	593
		Exercises for Section 5.5 . . . . .	602

### Introduction

Of the many topological properties a space may have, the most geometrical are those invoking the notion of connectedness. A *connected* space resembles an interval on the real line in consisting of a single piece, as opposed, say to the union of  $\{\langle x, 0 \rangle : 0 \leq x \leq 1\} \cup \{\langle x, 1 \rangle : 0 \leq x \leq 1\}$  of two parallel line segments in the plane. Actually, that a space  $X$  be connected may be defined in two conceptually different ways: the presence of just one “component” of  $X$  (*connectedness*), or the existence, for any two points of  $X$ , of a path joining the two points (*path-connectedness*). Fortunately, these two definitions turn out to be equivalent for a large class of spaces, including all manifolds.

Connectedness of a real-valued function’s domain provides a proof of the Intermediate-value Theorem used so often in calculus. It also allows us to distinguish topologically between a circle and a sphere, and between the real line and higher-dimensional Euclidean spaces.

We shall be examining spaces such as Euclidean spaces that have arbitrarily small connected neighborhoods of each point (*locally connected* spaces) and spaces such as the Cantor set whose only connected subspaces are points (*totally disconnected* spaces).

fix: Will material  
on homotopy  
move to separate  
chapter?

The final two sections in this chapter deal with continuous deformations of one map into another. Here the focus is on *simply connected* spaces, such as convex subspaces of Euclidean spaces, in which each closed path can be continuously shrunk to a point. The machinery developed to prove the intuitively evident fact that a circle is not simply connected will provide as a bonus a proof of the Fundamental Theorem of Algebra.

## 5.1 Connected Spaces

sec:connected

The notion of connectedness we wish to make precise here is that a connected space consists of a single “piece.” It is geometrically reasonable, for example, to deem each of the spaces  $\{0\}$ ,  $[0, 1]$ , and  $\mathbb{R}$  to be connected. However, it would be unreasonable to deem any of the spaces  $\{0, 1\}$ ,  $[0, 1] \cup [2, 3]$ , or  $\mathbb{R} \setminus \{0\}$  to be connected in this sense. Each of the latter three spaces can be split apart into more than one nonempty open subset, and it is plausible that this cannot be done with any of the former three spaces.

### Connected spaces

**5.1 Definition.** A **separation** of a topological space  $X$  is a collection  $\{A, B\}$  of  $X$  of two nonempty disjoint open subsets  $A$  and  $B$  such that  $A \cup B = X$ . A topological space  $X$  is said to be **connected** if there does *not* exist any separation of  $X$ . And  $X$  is said to be **disconnected** if it is not connected—that is, if there *does* exist some separation of  $X$ . A subset  $A$  of a topological space  $X$  is said to be **connected** if it is a connected space when given its relative topology.

exs:connected-disconnected-spaces

**5.2 Examples.** (1) Neither the empty space nor any space consisting of just a single point has a separation, so these spaces are connected.

(2) A discrete space  $X$  consisting of more than one point is disconnected, because the collection  $\{\{x\}, X \setminus \{x\}\}$  is a separation of it.

(3) The space  $[0, 1] \cup [2, 3]$  is disconnected, because  $\{[0, 1], [2, 3]\}$  is a separation of it.

ex:R-connected (4) According to [Lemma 3.30](#), *the real line  $\mathbb{R}$  is connected*.

ex:I-connected (5) By an argument similar to the proof of [Lemma 3.30](#), *the unit interval  $[0, 1]$  is connected*. This is a special case of [Theorem 5.4](#), below.  $\diamond$

A separation  $\{A, B\}$  of a topological space  $X$  is just a partition of  $X$  consisting of two open subsets of  $X$ . If  $\{A, B\}$  is any partition of  $X$  into two subsets, then  $A = X \setminus B$  and  $B = X \setminus A$ . Moreover, if  $A$  is any subset of  $X$  with  $\emptyset \neq A \neq X$ , then  $\{A, X \setminus A\}$  is a partition of  $X$ . Now a subset of  $X$  is open in  $X$  precisely when its complement is closed in  $X$ . From these observations we see that *both members of a separation of  $X$  are closed as well as open in  $X$* , and so we obtain the following criterion.

lem:disconnected-criteria

**5.3 Lemma.** *The following assertions about a topological space  $X$  are equivalent:*

- (i) *The space  $X$  is disconnected; that is, there exist disjoint nonempty open subsets  $U$  and  $V$  of  $X$  with  $X = U \cup V$ .*
- (ii) *There exist disjoint nonempty closed subsets  $E$  and  $F$  of  $X$  with  $X = E \cup F$ .*
- (iii) *There exist disjoint nonempty open, closed subsets  $A$  and  $B$  of  $X$  with  $X = A \cup B$ .*
- (iv) *There exists a nonempty proper open and closed subset of  $X$ .*

By definition, a topological space  $X$  is connected when it cannot be partitioned into *two* open subsets. Then a connected space  $X$  cannot be partitioned into more than two open subsets, either. In fact, if  $\mathcal{A}$  is a partition of  $X$  consisting of more than two open subsets of  $X$ , then we obtain a separation  $\{A, B\}$  by choosing as  $A$  any member of  $\mathcal{A}$  and taking  $B$  to be the union of the remaining members.

With the aid of [Lemma 5.3](#) we can discuss the connectedness of  $[0, 1]$  and other subspaces of the real line. Recall that an interval in  $\mathbb{R}$  is by definition a subset  $X$  of  $\mathbb{R}$  having the property that  $a, b \in X$  with  $a < b$  implies  $[a, b] \subset X$ . According to [Theorem 0.87](#), a subset  $X$  of  $\mathbb{R}$  is an interval in  $\mathbb{R}$  precisely when it has one of the following forms:

$$\emptyset, \quad ]a, b[, \quad [a, b[, \quad ]a, b], \quad [a, b], \quad ]a, +\infty[, \quad [a, +\infty[, \quad ]-\infty, b[, \quad ]-\infty, b], \quad \mathbb{R}.$$

thm:R-subspace-conn-iff-interval

**5.4 Theorem.** *A subspace  $X$  of  $\mathbb{R}$  is connected if and only if  $X$  is an interval.*

**Proof.** Assume that  $X$  is a connected subspace of  $\mathbb{R}$ . We show that  $X$  is an interval. Let  $a, b \in X$  be arbitrary with  $a < b$ . If  $c \in [a, b]$  but  $c \notin X$ , then

$$\{]-\infty, c[ \cap X, \quad ]c, +\infty[ \cap X\}$$

would be a separation of  $X$ . Hence  $[a, b] \subset X$ .

Conversely, assume that  $X$  is an interval in  $\mathbb{R}$ . We show that  $X$  is connected. Just suppose that  $X$  is disconnected, so that there exists a separation  $\{A, B\}$  of  $X$ . Now proceed much as in the proof of [Lemma 3.30](#). Choose some  $a \in A$  and  $b \in B$ , so that  $a \neq b$ ; without loss of generality we may assume that  $a < b$ . The set  $A \cap [a, b]$  is nonempty, and  $b$  is an upper bound of this set in  $\mathbb{R}$ . Invoking the order-completeness of  $\mathbb{R}$ , we may let

$$c = \sup(A \cap [a, b]).$$

We show  $c \in A$ . Clearly  $c \in \text{cls } A$  (this closure being taken in  $\mathbb{R}$ ), and  $c \in X$  because  $a \leq c \leq b$  and  $X$  is an interval in  $\mathbb{R}$ . Thus  $c \in X \cap \text{cls } A$ . But this set is the closure of  $A$  in the space  $X$ , and  $A$  is closed in  $X$ , and so  $X \cap \text{cls } A = A$ . Hence  $c \in A$ .

Since  $A$  is also open in the subspace  $X$  of  $\mathbb{R}$ , there exist  $x, y \in \mathbb{R}$  with

$$x < y < b, \quad c \in ]x, y[ \cap X \subset A.$$

Choose any  $z \in ]c, y[$ . Then  $z \in A$  because both  $c$  and  $b$  belong to the interval  $X$ . However,  $z \notin A$  because  $c < z$  and  $c$  is an upper bound of  $A \cap [a, b]$  in  $\mathbb{R}$ . We have reached a contradiction.  $\square$

From the preceding theorem we recover, in particular, the following result already established in [Lemma 3.30](#), cited in [Examples 5.2 \(4\)](#).

cor:R-connected

**5.5 Corollary.** *The real line  $\mathbb{R}$  is connected.*

extended real line! connected space @ as connected space

### Connected subsets of topological spaces

Recall that a subset of a topological space  $X$  is said to be connected when it is connected when provided with its relative topology making it a subspace of  $X$ . Thus we may rephrase [Theorem 5.4](#) as: the connected subsets of  $\mathbb{R}$  are precisely the intervals in  $\mathbb{R}$ .

A subset of a connected space need not itself be connected—even if it is open or is closed in the larger space. For example, the space  $[0, 3]$  is connected, but its subsets  $]0, 1[ \cup ]2, 3[$  and  $\{0, 2\}$  are not. Nonetheless, many connected spaces do occur naturally as subspaces of (not necessarily connected) topological spaces. For testing whether a subset of a space is connected, the following relativized version of [Lemma 5.3](#) is frequently useful.

lem:subset-conn-criteria

**5.6 Lemma.** *Each of the two conditions below is both necessary and sufficient for a subset  $A$  of a topological space  $X$  to be connected:*

(i) *There do not exist two open subsets  $U$  and  $V$  of  $X$  with*

$$U \cap A \neq \emptyset \neq V \cap A, \quad U \cap V \cap A = \emptyset, \quad \text{and} \quad A \subset U \cup V.$$

(ii) *There do not exist two closed subsets  $E$  and  $F$  of  $X$  with*

$$E \cap A \neq \emptyset \neq F \cap A, \quad E \cap F \cap A = \emptyset, \quad \text{and} \quad A \subset E \cup F.$$

One application of this lemma follows.

thm:cls-conn-is-conn

**5.7 Theorem.** *Let  $A$  be a connected subset of a topological space  $X$ . Then  $\text{cls } A$  is connected. More generally, each set  $B$  with*

$$A \subset B \subset \text{cls } A$$

*is connected.*

**Proof.** Let  $A \subset B \subset \text{cls } A$  and just suppose that  $B$  is not connected. By [Lemma 5.6](#), there exist open subsets  $U$  and  $V$  of  $X$  with

{eq:separating-subset-of-cls} (\*) 
$$U \cap B \neq \emptyset \neq V \cap B, \quad U \cap V \cap B = \emptyset, \quad \text{and} \quad B \subset U \cup V.$$

Then  $U \cap \text{cls } A$  and  $V \cap \text{cls } A$  are nonempty open subsets of the space  $\text{cls } A$ , so that

$$U \cap A \neq \emptyset, \quad V \cap A \neq \emptyset.$$

Since  $A \subset B$ , from (\*) we obtain

$$U \cap V \cap A = \emptyset, \quad A \subset U \cup V.$$

According to [Theorem 5.7](#), the set  $A$  is therefore not connected, which contradicts the hypothesis about  $A$ .  $\square$

cor:conn-if-conn-dense-set

**5.8 Corollary.** *A topological space that has a connected dense subset is itself connected.*

The real line  $\mathbb{R}$  is dense in the extended real line  $\widehat{\mathbb{R}}$ , and so from [Corollaries 5.5](#) and [5.8](#) it follows that  $\widehat{\mathbb{R}}$  is also connected.



Notice that a dense subset of a connected space need not be connected. For example, the real line  $\mathbb{R}$  is connected; however, its dense subspace  $\mathbb{Q}$  is disconnected, because from the irrationality of  $\sqrt{2}$  we see that

$$\{\mathbb{Q} \cap ]-\infty, \sqrt{2}[, \mathbb{Q} \cap ]\sqrt{2}, +\infty[ \}$$

is a separation of  $\mathbb{Q}$ .

### Continuous maps on connected spaces

Although connectedness is not preserved under the formation of arbitrary subspaces, it is preserved under the formation of quotient spaces. This is a consequence of the following more general fact.

**5.9 Theorem.** *The image of a connected space under a continuous map is itself connected.*

**Proof.** Let  $f: X \rightarrow Y$  be a continuous map from a connected space  $X$  into a topological space  $Y$ . By replacing  $Y$  with  $f(X)$  if necessary, we may assume without loss of generality that  $f(X) = Y$ . Just suppose that  $Y$  is disconnected, and let  $\{A, B\}$  be a separation of  $Y$ . Define

$$U = f^{-1}(A), \quad V = f^{-1}(B).$$

We claim that  $\{U, V\}$  is a separation of  $X$ , in contradiction to the hypothesis that  $X$  is connected. In fact, both  $U$  and  $V$  are open in  $X$  because  $f$  is continuous and the sets  $A$  and  $B$  are open in  $Y$ . Both  $U$  and  $V$  are nonempty because  $f$  maps  $X$  onto  $Y$  and the sets  $A$  and  $B$  are nonempty. The sets  $U$  and  $V$  are disjoint because

$$U \cap V = f^{-1}(A) \cap f^{-1}(B) = f^{-1}(A \cap B) = f^{-1}(\emptyset) = \emptyset.$$

Finally,

$$U \cup V = f^{-1}(A) \cup f^{-1}(B) = f^{-1}(A \cup B) = f^{-1}(Y) = X. \quad \square$$

One application of [Theorem 5.9](#) is that **the 1-sphere**

$$S_1 = \{ \langle x_1, x_2 \rangle \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1 \}$$

**is connected.** In fact, by [Exercise 3.20](#), the 1-sphere is the image of the unit interval  $I = [0, 1]$  under the continuous map  $t \mapsto \langle \cos 2\pi t, \sin 2\pi t \rangle$ , and by [Theorem 5.4](#) the unit interval  $I$  is connected. Later [[Examples 5.21 \(3\)](#)] we shall see that the  $n$ -sphere  $S_n$  is connected for every  $n \geq 1$ .

By contrast, the 0-sphere  $S_0 = \{-1, 1\}$ , being a two-point discrete space, is disconnected.

**5.10 Corollary.** *A quotient space of a connected space is itself connected.*

**5.11 Corollary.** *Connectedness is a topological property: If  $Y$  is a topological space that is homeomorphic to a connected space, then  $Y$  is connected.*

[Corollary 5.11](#) provides some examples that will shortly aid us in obtaining still further, more interesting, examples of connected spaces.

Intermediate-value Theorem  
Bolzano, Bernard  
polynomial  
root!polynomial@of a polynomial

**5.12 Examples.** (1) Given any two distinct points  $x, y \in \mathbb{R}^n$ , the line

$$L = \{(1-t)x + ty : t \in \mathbb{R}\}$$

passing through  $x$  and  $y$  is connected, for by Examples 3.25 (7) the map

$$\begin{aligned}\mathbb{R} &\rightarrow L \\ t &\mapsto (1-t)x + ty\end{aligned}$$

is a homeomorphism, and by Corollary 5.11 the real line  $\mathbb{R}$  is connected.

ex:line-segments-are-conn

(2) The restriction of the above map to  $[0, 1]$  gives a homeomorphism from the connected space  $[0, 1]$  onto the line segment

$$\{(1-t)x + ty : t \in \mathbb{R}, 0 \leq t \leq 1\}$$

joining  $x$  to  $y$ , and so again by Corollary 5.11, this line segment is connected.  $\diamond$

A particularly important case of Theorem 5.9 arises when the continuous function  $f$  on the connected space  $X$  is real-valued. In that case the theorem implies that if  $f$  assumes two distinct values, then it must assume as values all numbers lying between those two.

thm:IVT

**5.13 Intermediate-value Theorem.** Let  $f: X \rightarrow \mathbb{R}$  be a continuous real-valued function on a connected space. Let  $a, b \in X$  with  $f(a) < f(b)$ . Then for each real number  $c$  with

$$f(a) < c < f(b)$$

there exists some  $x \in X$  with

$$f(x) = c.$$

**Proof.** Let  $c \in \mathbb{R}$  with  $f(a) < c < f(b)$ . By Theorem 5.9, the image  $f(X)$  of  $X$  under  $f$  is a connected subset of  $\mathbb{R}$ , and by Theorem 5.4 this image is an interval in  $\mathbb{R}$ . Now  $f(a), f(b) \in f(X)$ , so that the closed interval  $[f(a), f(b)] \subset f(X)$ . In particular,  $c \in f(X)$ .  $\square$

A closed interval in  $\mathbb{R}$  is connected. Hence when we take in Theorem 5.13  $X = [a, b]$ , an interval in  $\mathbb{R}$ , we obtain the classical intermediate-value theorem of Bernard Bolzano familiar from calculus:

A continuous function  $f: [a, b] \rightarrow \mathbb{R}$  takes all numbers between  $f(a)$  and  $f(b)$  as values.

Suppose, in the notation of Theorem 5.13 that

$$f(a) < 0, \quad f(b) > 0.$$

Then the theorem asserts the existence of at least one solution  $x \in X$  to the equation

$$f(x) = 0.$$

Which polynomials are guaranteed to have roots? Of course, the quadratic formula provides an explicit formula for the roots of a degree 2 polynomial; and there are analogous “algebraic” formulas for finding roots of cubic (degree 3) and quartic (degree 4) polynomials—see, for example, Wikipedia [70, 72]. The following application of the Intermediate-value Theorem guarantees that some other polynomials have roots (although it does not provide a formula for finding them).

ex:odd-deg-poly-has-real-root **5.14 Example.** Any polynomial function

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

with real coefficients  $a_0, a_1, \dots, a_n$  and of *odd* degree has at least one real root; that is, the equation

$$p(x) = 0$$

has some solution  $x \in \mathbb{R}$ . To prove this, it suffices to show that

$$p(-b) < 0, \quad p(b) > 0$$

for some real number  $b > 0$ , because the result will then follow by applying the [Intermediate-value Theorem](#) to  $p$ , which is continuous.

To say that  $p$  is of degree  $n$  means that the leading coefficient  $a_n \neq 0$ . Now for a real number  $x$ , we have  $p(x) = 0$  if and only if  $a_n^{-1} p(x) = 0$ . By replacing  $p(x)$  with  $a_n^{-1} p(x)$ , if necessary, we may therefore assume without loss of generality that

$$a_n = 1.$$

Then for arbitrary  $x \neq 0$  we may write

$$p(x) = x^n q(x)$$

where

$$q(x) = 1 + \frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + \cdots + \frac{a_0}{x^n}.$$

We shall find a  $b > 0$  with

$$q(b) > 0, \quad q(-b) > 0.$$

It will then follow that  $p(b) > 0$  and, since the degree  $n$  is odd,  $p(-b) < 0$ .

For arbitrary  $x$  with  $|x| > 1$  we have

$$\begin{aligned} |q(x) - 1| &\leq \left| \frac{a_{n-1}}{x} \right| + \left| \frac{a_{n-2}}{x^2} \right| + \cdots + \left| \frac{a_0}{x^n} \right| \\ &< \frac{|a_{n-1}|}{|x|} + \frac{|a_{n-2}|}{|x|} + \cdots + \frac{|a_0|}{|x|} \\ &= \frac{A}{|x|} \end{aligned}$$

where

$$A = |a_0| + |a_1| + \cdots + |a_{n-1}|.$$

Then

$$|x| > \max\{1, 2A\}$$

implies

$$|q(x) - 1| < \frac{1}{2}$$

and therefore

$$q(x) > 0.$$

Now choose any  $b > \max\{1, 2A\}$ . Since  $|-b| = b$ , the numbers  $q(b)$  and  $q(-b)$  are both positive, as desired.  $\diamond$

The existence of (possibly complex) roots of polynomials of even as well as of odd degree is guaranteed by the Fundamental Theorem of Algebra (5.106), which is proved later. One sometimes wants to solve an equation not of the form  $f(x) = 0$ , but instead of the form  $f(x) = x$ ; in other words, to find a *fixed-point*  $x \in X$  of a map  $f: X \rightarrow X$  (see page 21). This is the case, for example, in establishing the existence of solutions of a first-order ordinary differential equation, where the Banach Contraction Mapping Principle (Exercise 1.124) may be used. Of course, not every continuous map  $f: X \rightarrow X$  has a fixed-point, even when the space  $X$  is connected: consider, for example, a rotation of  $S_1$  through an angle that is not an integral multiple of  $2\pi$ . The existence of a fixed-point of a continuous map  $f: X \rightarrow X$  can, however, be inferred from the connectedness of special spaces  $X$ .

cor:Brouwer-fixed-pt-dim-1

### 5.15 Corollary (Brouwer Fixed-point Theorem in dimension 1). Let

$$f: [0, 1] \rightarrow [0, 1]$$

be a continuous map from the closed unit interval to itself. Then  $f$  has a fixed-point; that is, there exists an  $x \in [0, 1]$  such that

$$f(x) = x.$$

**Proof.** To say that  $f(x) = x$  is equivalent to saying that  $x - f(x) = 0$ . Accordingly, introduce the function

$$g: [0, 1] \rightarrow \mathbb{R}$$

defined by

$$g(x) = x - f(x).$$

If  $f(0) = 0$  or  $f(1) = 1$ , we are done, so we now suppose that  $f(0) \neq 0$  and  $f(1) \neq 1$ . Since  $f$  maps into  $[0, 1]$ , we have

$$f(0) > 0, \quad f(1) < 1,$$

so that

$$g(0) = 0 - f(0) < 0, \quad g(1) = 1 - f(1) > 0.$$

Since  $g$  is continuous and the interval  $[0, 1]$  is connected, the Intermediate-value Theorem implies that  $g(x) = 0$  for some  $x \in [0, 1]$ . Then  $f(x) = x$ .  $\square$

Geometrically, the Brouwer Fixed-point Theorem in dimension 1 (Corollary 5.15) says that the graph of a continuous function  $f: [0, 1] \rightarrow [0, 1]$  must cross the line  $y = x$  (see Figure 5.1).

It is a fact more general than Corollary 5.15 that for arbitrary dimension  $n \geq 1$ , any continuous map  $f: I^n \rightarrow I^n$  from the  $n$ -cube to itself must have a fixed-point. This is the *Brouwer Fixed-point Theorem*. We shall give a proof for the case of dimension  $n = 2$  in Section 5.5: see Theorem 5.105.

The Intermediate-value Theorem provides a surprising result concerning real-valued functions on the circle  $S_1$ . Observe that if  $x = \langle x_1, x_2 \rangle \in S_1$ , then  $-x = \langle -x_1, -x_2 \rangle \in S_1$ , too; in fact,  $-x$  diametrically opposite  $x$ , that is, it is the point on  $S_1$  at which the line through  $x$  and the origin intersects  $S_1$  (see Figure 5.2).

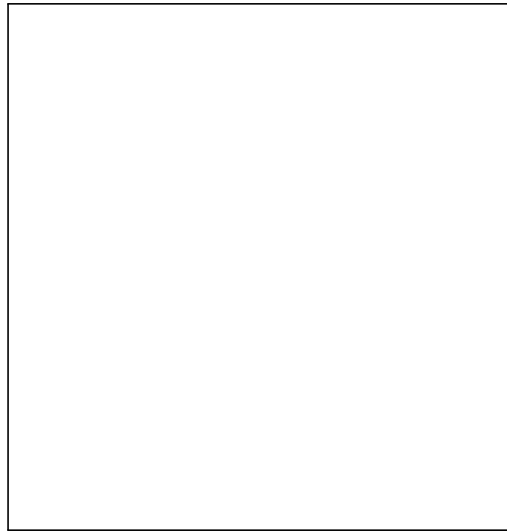


Figure 5.1: Geometric meaning of the Brouwer Fixed-point Theorem in dimension 1.

fig:Brouwer-fixed-pt-dim-1

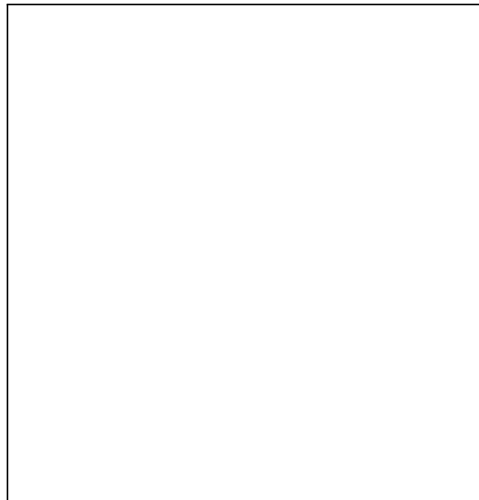


Figure 5.2: Antipodal points on the circle  $S_1$ .

fig:antipodal-pts-S1

cor:Borsuk-Ulam-dim-1

**5.16 Corollary (Borsuk–Ulam theorem in dimension 1).** *Let*

$$f: S_1 \rightarrow \mathbb{R}$$

*be a continuous function from the unit circle to the real line. Then there exists some  $x \in S_1$  such that*

$$f(x) = f(-x).$$

**Proof.** Define the continuous map

$$g: S_1 \rightarrow \mathbb{R}$$

Borsuk-Ulam Theorem  
 Borsuk, Karol  
 Ulam, Stanisław  
 Borsuk-Ulam Theorem

by

$$g(x) = f(x) - f(-x)$$

We seek an  $x \in S_1$  for which  $g(x) = 0$ . Clearly

$$g(-x) = -g(x) \quad (x \in S_1),$$

so to find a solution of  $g(x) = 0$  we need only look at points in the upper semicircle  $S_1^+ = \{x \in S_1 : x_2 \geq 0\}$  of  $S_1$ . Accordingly, we introduce the map

$$\begin{aligned} p: [0, 1] &\rightarrow S_1 \\ t &\mapsto \langle \cos \pi t, \sin \pi t \rangle \end{aligned}$$

which is continuous and maps  $[0, 1]$  onto this upper semicircle  $S_1^+$ . To show that  $g(x) = 0$  for some  $x$ , and hence to complete the proof, we need only show that  $g(p(t)) = 0$  for some  $t \in [0, 1]$ .

Form the continuous map

$$F = g \circ p: [0, 1] \rightarrow \mathbb{R}.$$

Direct calculation gives

$$F(1) = -F(0).$$

If  $F(0) \neq 0$ , then one of the numbers  $F(0)$  and  $F(1)$  is positive while the other is negative, so by the Intermediate-value Theorem (5.13) there is some  $t \in ]0, 1[$  for which  $F(t) = 0$ . Thus for such  $t$ ,

$$g(p(t)) = F(t) = 0,$$

as needed.  $\square$

The [Borsuk–Ulam theorem in dimension 1](#) has an amusing physical interpretation. At a given moment of time, the temperature at all points on a given meridian of the earth may be modeled by a function  $f: S_1 \rightarrow \mathbb{R}$ . It is a reasonable physical assumption that temperature varies continuously around this meridian, in other words, that  $f$  is continuous. The conclusion of [Corollary 5.16](#) then says that at a given moment there are two diametrically opposite points on the meridian at which the temperatures are the same. (Question: Could one experimentally verify—or refute—this conclusion or the physical assumption leading to it?)

Like the Brouwer Fixed-point Theorem in dimension 1 ([Corollary 5.15](#)), [Corollary 5.16](#) can be extended to arbitrary dimension  $n \geq 1$ : the **Borsuk–Ulam theorem** asserts that if  $f: S_n \rightarrow \mathbb{R}^n$  is a continuous map, then  $f(x) = f(-x)$  for some  $x \in S_n$ .

### Covering by connected sets

The next theorem and its corollaries will allow us to enlarge our repertoire of connected spaces. First, a bit of preparation.

lem:conn-in-open-closed-it-meets

**5.17 Lemma.** *Let  $C$  be a connected subset of a topological space  $X$ , and let  $U$  be an open, closed subset of  $X$  that intersects  $C$ . Then  $C \subset U$ .*

**Proof.** The set  $C \cap U$  is a nonempty subset of  $C$  that is both open and closed in  $C$ . By [Lemma 5.3](#),  $U \cap C = C$ . Hence  $C \subset U$ .  $\square$

**5.18 Theorem.** Let  $X$  be a topological space. Suppose there is a cover  $C$  of  $X$  consisting of connected sets such that, for each two members  $A$  and  $B$  of  $C$ , there are finitely many sets  $A_1, A_2, \dots, A_n$  such that

$$A_1 = A, \quad A_n = B,$$

and

$$A_i \cap A_{i+1} \neq \emptyset \quad (i = 1, 2, \dots, n-1).$$

Then  $X$  is connected.

**Proof.** Just suppose there exists some separation  $\{U, V\}$  of  $X$ . Since  $U$  and  $V$  are nonempty and  $C$  covers  $X$ , there are some  $A, B \in C$  with

$$A \cap U \neq \emptyset, \quad B \cap V \neq \emptyset.$$

By hypothesis there are sets  $A_1, A_2, \dots, A_n \in C$  with  $A_1 = A, A_n = B$ , and

$$A_i \cap A_{i+1} \neq \emptyset \quad (i = 1, 2, \dots, n-1).$$

We are going to show that  $B = A_n \subset U$ ; this will produce the desired contradiction, for then  $B$  will intersect both of the disjoint sets  $U$  and  $V$ .

To show that  $A_n \subset U$ , we show something more, namely, that  $A_i \subset U$  for each  $i = 1, 2, \dots, n$ . Since  $A_1$  is connected and intersects the open, closed subset  $U$  of  $X$ , it follows from [Lemma 5.17](#) that  $A_1 \subset U$ . Now let  $1 \leq i < n$  and assume that  $A_i \subset U$ . Then  $A_{i+1}$  is a connected set which, since

$$\emptyset \neq A_i \cap A_{i+1} \subset U \cap A_{i+1},$$

intersects  $U$ . By [Lemma 5.17](#) again,  $A_{i+1} \subset U$ , too.  $\square$

Suppose, in the notation of [Theorem 5.18](#), that the cover  $C$  of  $X$  actually has nonempty intersection. Then *a fortiori* any two members  $A$  and  $B$  of  $C$  will intersect, so that the hypothesis of [5.18](#), will hold with  $n = 2$ . This proves the following corollary.

**5.19 Corollary.** A topological space is connected if it has a cover consisting of connected sets with nonempty intersection.

**5.20 Corollary.** Suppose that for each two points of a topological space there is a connected subset containing them both. Then the space is connected. Then  $X$  is connected.

**Proof.** Let  $X$  be a topological space. If  $X$  is empty, there is nothing to prove, and so we assume that  $X \neq \emptyset$ . Assume that for each two points  $x$  and  $y$  of  $X$ , there is a connected subset  $C_{x,y}$  of  $X$  for which  $x \in C_{x,y}$  and  $y \in C_{x,y}$ .

Arbitrarily choose some  $y \in X$ . Then

$$x \in \bigcup_{x \in X} C_{x,y}, \quad y \in \bigcap_{x \in X} C_{x,y}.$$

Hence the preceding corollary applies with  $C = \{C_{x,y} : x \in X\}$ .  $\square$

**Example 5.21** (1) Any convex subset of  $\mathbb{R}^n$  is connected. In fact, if  $X$  is such a set, then for each  $x, y \in X$  the line segment

$$L_{x,y} = \{(1-t)x + ty : 0 \leq t \leq 1\}$$

is a subset of  $X$  containing both  $x$  and  $y$  and which, by [Examples 5.12 \(2\)](#), is connected. Hence by [Corollary 5.20](#),  $X$  is connected.

In particular, the  $n$ -cube

$$I^n = [0, 1]^n,$$

the  $n$ -disk

$$D_n = \{x \in \mathbb{R}^n : \|x\| \leq 1\},$$

the  $n$ -ball

$$B_n = \{x \in \mathbb{R}^n : \|x\| < 1\},$$

and  $\mathbb{R}^n$  itself are all connected spaces.

(2) The Klein bottle [\[Examples 3.81 \(6\)\]](#) and the Möbius strip [\[Examples 3.81 \(5\)\]](#) are connected, being quotient spaces of  $I^2$ .

(3) For every  $n \geq 1$  the  $n$ -sphere

$$S_n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$$

is connected. To see this, let

$$\mathbf{p} = \langle 0, 0, \dots, 1 \rangle$$

be the north pole of  $S_n$ , so that  $-\mathbf{p}$  is the south pole  $\langle 0, 0, -1 \rangle$ . Define

$$A = S_n \setminus \{\mathbf{p}\}, \quad B = S_n \setminus \{-\mathbf{p}\}.$$

Now

$$A \cong \mathbb{R}^n$$

by [Examples 3.25 \(13\)](#), and likewise

$$B \cong \mathbb{R}^n.$$

In (1) we proved that  $\mathbb{R}^n$  is connected, so by [Corollary 5.11](#) both  $A$  and  $B$  are connected. Since

$$A \cup B = S_n, \quad A \cap B \neq \emptyset,$$

it follows from [Corollary 5.19](#) that  $S_n$  is connected.

(4) The  $n$ -dimensional real projective space  $\mathbb{RP}_n$  is connected because it is a quotient space of  $S_n$  [see [Examples 3.81 \(7\)](#)].

(5) For  $n > 1$ , the complement  $\mathbb{R}^n \setminus E$  of a countable subset  $E$  of  $\mathbb{R}^n$  is connected. To prove this, we shall show that an arbitrary pair of points  $x, y \in \mathbb{R}^n \setminus E$  belong to some connected subset of  $\mathbb{R}^n \setminus E$ . Fix  $x, y \in \mathbb{R}^n \setminus E$  with  $x \neq y$ . For arbitrary  $u, v \in \mathbb{R}^n$  let

$$L_{u,v} = \{(1-t)u + tv : 0 \leq t \leq 1\},$$

the line segment joining  $u$  to  $v$ .



Choose any point  $z$  with

$$z \in L_{x,y}, \quad x \neq z \neq y$$

(for example, take the midpoint  $z = 1/2 u + 1/2 v$ ). Since  $n > 1$ , there exists some  $w \in \mathbb{R}^n \setminus L_{x,y}$  for which

$$L_{z,w} \cap L_{x,y} = \{z\}$$

(see Figure 5.3); this is where we are using the hypothesis that  $n > 1$ . For each  $u \in L_{z,w}$ , let

$$C_u = L_{x,u} \cup L_{u,y},$$

so that  $x \in C_u$  and  $y \in C_u$ . For each  $u \in L_{z,w}$ , the set  $C_u$  is connected because both the line segments  $L_{x,u}$  and  $L_{u,y}$  are connected and contain the point  $u$ . Hence it suffices to show that  $C_u \subset \mathbb{R}^n \setminus E$  for some  $u \in L_{z,w}$ .

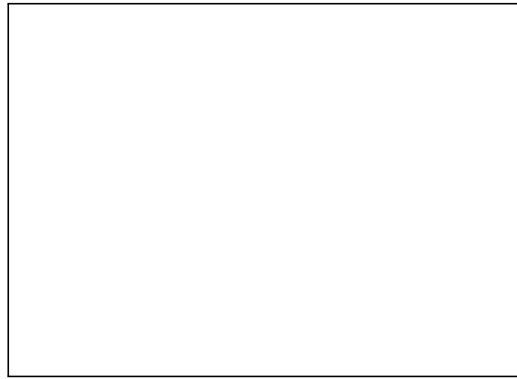


Figure 5.3: Configuration of line segments for proving that the complement of a countable set in  $\mathbb{R}^n$  is connected.

fig:pf-Rn-less-countable-is-conn

Just suppose that

$$C_u \cap E \neq \emptyset \quad (u \in L_{z,w}).$$

We have

$$C_u \cap C_v = \{x, y\} \quad (u, v \in L_{z,w}; u \neq v)$$

and  $\{x, y\} \cap E = \emptyset$ . Therefore for distinct  $u, v \in L_{z,w}$  the points in  $C_u \cap E$  are all different from the points in  $C_v \cap E$ . Since  $L_{z,w}$  is an uncountable set, we conclude that the set

$$E \cap \bigcup_{u \in L_{z,w}} C_u$$

is uncountable. But this is impossible because  $E$  was supposed to be countable.  $\diamond$

cted space!covering@and covering])

The fact that connectedness is a topological property tells us that a connected space cannot be homeomorphic to a disconnected space. For example, the closed unit interval  $[0, 1]$  is not homeomorphic to the union  $[0, 1] \cup [2, 3]$  of two disjoint closed intervals. Sometimes this fact can even be used to show that two connected spaces  $X$  and  $Y$  are not homeomorphic to one another by applying it to suitable subspaces of  $X$  and  $Y$ .

**5.22 Definition.** A subset  $A$  of a topological space  $X$  is said to **disconnect**  $X$ , and  $X$  is said to be **disconnected by**  $A$ , when the complement  $X \setminus A$  of  $A$  is disconnected.

exs:disconnecting sets **5.23 Examples.** (1) An interval  $J$  in  $\mathbb{R}$  containing at least two points is disconnected by the singleton  $\{c\}$  for any nonendpoint  $c \in J$ , because then  $J \setminus \{c\}$  is not an interval.

- (2) Let  $z$  be an arbitrary point in  $S_1$ . Then  $\{z\}$  does *not* disconnect  $S_1$ , because by [Examples 3.25 \(13\)](#), [Exercise 1.55](#), and [Examples 3.25 \(1\)](#),

$$S_1 \setminus \{z\} \cong \mathbb{R},$$

and  $\mathbb{R}$  is connected.

We are now in a position to use connectedness for proving that *the circle  $S_1$  is not homeomorphic to the real line  $\mathbb{R}$  or, more generally, to any interval in  $\mathbb{R}$* . (That  $S_1$  is not homeomorphic to  $\mathbb{R}$  or to any non-closed interval in  $\mathbb{R}$  follows already from compactness considerations: the circle  $S_1$  is compact [[Examples 4.35 \(5\)](#)] whereas neither the real line nor any nonempty, non-closed interval in  $\mathbb{R}$  is compact [[Examples 4.6 \(4\)](#)]. However, the following proof uses only connectedness.)

Just suppose that there exists some interval  $J$  in  $\mathbb{R}$  (possibly  $\mathbb{R}$  itself) and some homeomorphism

$$f: S_1 \cong J.$$

Since  $S_1$  contains at least three points, the same thing is true of  $J$ , which means that  $J$  includes a non-endpoint. Choose any  $z \in S_1$  with the number  $x = f(z)$  *not* an endpoint of  $J$ . By restricting  $f$ , we obtain a homeomorphism

$$S_1 \setminus \{z\} \cong J \setminus \{x\}.$$

This is impossible because  $S_1 \setminus \{z\}$  is connected whereas  $J \setminus \{x\}$  is not.

- ex:Rn-not-homeo-R (3) For  $n > 1$ , Euclidean  $n$ -space  $\mathbb{R}^n$  is not homeomorphic to the real line  $\mathbb{R}$ . In fact, if there existed some  $f: \mathbb{R}^n \cong \mathbb{R}$ , then we would have

$$\mathbb{R}^n \setminus \{x\} \cong \mathbb{R} \setminus \{y\}$$

for  $x \in \mathbb{R}^n$  and  $y = f(x) \in \mathbb{R}$ . But this is impossible because by [Examples 5.21 \(5\)](#) the singleton  $\{x\}$  does not disconnect  $\mathbb{R}^n$ , whereas  $\{y\}$  does disconnect  $\mathbb{R}$ .

This example is a special case of [Theorem 3.38](#), whose proof rests upon Invariance of Domain Theorem ([3.39](#)).  $\diamond$

### Products of connected spaces

Having looked at subspaces and quotient spaces of connected spaces, we now turn to products of connected spaces.



Figure 5.4: Intersecting slices in a product of connected spaces.

fig:slices-to-prove-product-connecte

thm:prod-connected-two

**5.24 Theorem.** *The product of two connected spaces is itself connected.*

**Proof.** Let  $X$  and  $Y$  be connected spaces. Take arbitrary points  $a = \langle a_1, a_2 \rangle$  and  $b = \langle b_1, b_2 \rangle$  in  $X \times Y$ . According to [Corollary 5.20](#), it suffices to find a connected subset of  $X \times Y$  containing both  $a$  and  $b$ . Form the slices

$$A = \{a_1\} \times Y, \quad B = X \times \{b_2\}$$

of  $X \times Y$  (see [Figure 5.4](#)). Now

$$A \cong Y, \quad B \cong X,$$

and so by assumption the sets  $A$  and  $B$  are both connected. Moreover,

$$\langle a_1, b_2 \rangle \in A \cap B$$

and so by [Corollary 5.19](#) the union  $A \cup B$  is connected. Now  $a \in A$  and  $b \in B$ . Thus  $A \cup B$  is the connected subset of  $X \times Y$  we seek.  $\square$

cor:prod-connected-finite

**5.25 Corollary.** *The product of a finite family of nonempty topological spaces is connected if and only if each factor space is connected.*

**Proof.** Let  $\langle X_i \rangle_{i=1,2,\dots,n}$  be a finite family of nonempty topological spaces. Assume first that the product  $\prod_{i=1}^n X_i$  is connected. Then for each index  $j = 1, 2, \dots, n$ , the  $j$ th projection  $p_j: \prod_{i=1}^n X_i \rightarrow X_j$  is a continuous surjection, and so by [Theorem 5.9](#) the space  $X_j$  is connected.

The converse is proved by induction on  $n$ , using [Theorem 5.24](#) and the relation

$$\prod_{i=1}^n X_i \cong \left( \prod_{i=1}^{n-1} X_i \right) \times X_n$$

(see [Proposition 3.66](#)).  $\square$

From [Corollary 5.25](#) and the connectedness of the circle  $S_1$ , it follows that the  $n$ -torus

$$(S_1)^n = S_1 \times S_1 \times \cdots \times S_1$$

is connected.

More generally than [Corollary 5.25](#), the product of an arbitrary family of connected spaces is connected. Because the proof is harder than that for a finite family, it is worth looking separately at the proof for a sequence of connected spaces. The proof for the general case follows that.

lem:dense-special-set-in-seq-prod

**5.26 Lemma.** Let  $a = \langle a_i \rangle_{i=1,2,3,\dots}$  be a point in the product  $X$  of a sequence  $\langle X_i \rangle_{i=1,2,3,\dots}$  of topological spaces. Then the set

$$D = \{x \in X : \text{for some } n \in \mathbb{N}^*, x_i = a_i \text{ for all } i \geq n\}$$

is dense in  $X$ .

**Proof.** Let  $U$  be an arbitrary open neighborhood of an arbitrary point  $y$  in  $X$ . We show that  $U$  intersects  $D$ . There exists an  $n \geq 1$  and open sets  $V_1$  in  $X_1$ ,  $V_2$  in  $X_2$ ,  $\dots$ ,  $V_n$  in  $X_n$  such that

$$y \in V_1 \times V_2 \times \cdots \times V_n \times X_{n+1} \times X_{n+2} \times \cdots \subset U.$$

Define the point  $x \in X$  by

$$x_i = \begin{cases} y_i & \text{if } i = 1, 2, \dots, n, \\ a_i & \text{if } i = n+1, n+2, \dots \end{cases}$$

Then  $y \in D \cap U$ .  $\square$

thm:prod-seq-conn-is-conn

**5.27 Theorem.** The product of a sequence of nonempty connected spaces is itself connected.

**Proof.** Let  $\langle X_i \rangle_{i=1,2,3,\dots}$  be a sequence of nonempty connected spaces and let  $X = \prod_{i=1}^{\infty} X_i$ . Fix any  $a \in X$ . Define

$$C = \{x \in X : \text{there is a connected } E \subset X \text{ with } a \in E \text{ and } x \in E\}.$$

Note that  $a \in C$  because  $\{a\}$  is connected. The set  $C$  is connected according to [Corollary 5.20](#), so by [Corollary 5.8](#) the subset  $\text{cls } C$  of  $X$  is also connected. To establish that  $X$  is connected we are going to show that  $C$  is dense in  $X$ .

Define  $D$  as in [Lemma 5.26](#). Then it suffices to prove that

$$D \subset C,$$

for then

$$X = \text{cls } D \subset \text{cls } C \subset X.$$

For each integer  $n \geq 1$ , define

$$D_n = \{x \in X : x_i = a_i \text{ for all } i \geq n\}$$

Then

$$D = \bigcup_{n=1}^{\infty} D_n.$$

To prove that  $D \subset C$ , it remains only to show that  $D_n \subset C$  for all  $n \geq 1$ . We use induction on  $n$ . First,

$$D_1 = \{a\} \subset C.$$

Now let  $n \geq 1$  and assume that  $D_n \subset C$ . We deduce that  $D_{n+1} \subset C$ , too.

Let  $x \in D_{n+1}$ . We must show that  $x \in C$ . We have

$$x_i = a_i \quad (i = n+2, n+3, \dots).$$

Define  $y \in X$  by

$$y_i = \begin{cases} x_i & \text{if } i \neq n+1, \\ a_{n+1} & \text{if } i = n+1. \end{cases}$$

(see Figure 5.5). Then  $y \in D_n$ , and so by the inductive assumption  $y \in C$ , that is, there is a connected subset  $A$  of  $X$  with

$$a \in A, \quad y \in A.$$

Now the set

$$\begin{aligned} B &= \{z \in X : z_i = x_i \text{ for all } i \neq n+1\} \\ &= \{x_1\} \times \{x_2\} \times \cdots \times \{x_n\} \times X_{n+1} \times \{x_{n+2}\} \times \{x_{n+3}\} \times \cdots, \end{aligned}$$

being homeomorphic to the connected space  $X_{n+1}$  is connected. Moreover,

$$y \in B, \quad x \in B.$$

Then  $a, x \in A \cup B$ , and by Corollary 5.19 the set  $A \cup B$  is connected. Hence  $x \in C$ . This completes the proof of the inductive step that  $D_{n+1} \subset C$ .  $\square$

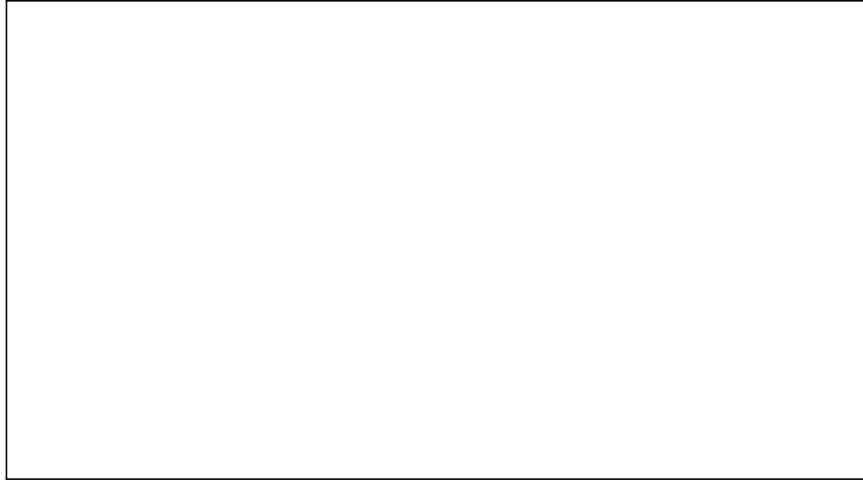


Figure 5.5: Constructing a point  $y$  at the inductive step in proving that the product of a sequence of connected spaces is connected.

fig:pf-induction-prod-seq-conn

An immediate consequence of Theorem 5.27 is that the Hilbert cube

$$I^\infty = \{x \in \ell^2 : |x_i| \leq 1/i \text{ for each } i = 1, 2, 3, \dots\} = \bigcap_{i=1}^{\infty} [-1/i, 1/i]$$

is connected.

For a family of topological spaces indexed by an arbitrary set  $I$ , the counterpart to the condition “for all  $i \geq n$ ” used above is “for almost all  $i \in I$ .” Accordingly, to prove that the

product of an arbitrary family of connected spaces is connected we shall use the following generalization of [Lemma 5.26](#).

lem:dense-special-set-in-arb-prod

**5.28 Lemma.** Let  $a = \langle a_i \rangle_{i \in I}$  be a point in the product  $X$  of a family  $\langle X_i \rangle_{i \in I}$  of topological spaces. Then the set

$$D = \{x \in X : x_i = a_i \text{ for almost all } i \in I\}$$

is dense in  $X$ .

**Proof.** Let  $U$  be an arbitrary open neighborhood of an arbitrary point  $y$  in  $X$ . We show that  $U$  intersects  $D$ . There is a basic open set

$$V = \bigcap_{j \in J} p_j^{-1}(V_j),$$

where  $J$  is a finite subset of  $I$  and  $\langle V_j \rangle_{j \in J}$  is a family with  $V_j$  an open subset of  $X_j$  for each  $j \in J$ , such that

$$y \in V \subset U.$$

Define the point  $x \in X$  by

$$x_i = \begin{cases} y_i & \text{if } i \in J, \\ a_i & \text{if } i \in I \setminus J. \end{cases}$$

Then  $x \in D \cap U$ .  $\square$

thm:prod-connected

**5.29 Theorem.** The product of a family of nonempty topological spaces is connected if and only if each factor space is connected.

**Proof.** Let  $\langle X_i \rangle_{i \in I}$  be a family of nonempty topological spaces and let  $X = \times_{i \in I} X_i$  be their product. Assume first that the product  $X$  is connected. Then for each index  $j \in I$  the  $j$ th projection  $p_j : X \rightarrow X_j$  is a continuous surjection, and so by [Theorem 5.9](#) the space  $X_j$  is connected.

Conversely, assume that  $X_i$  is connected for each  $i \in I$ . If  $I$  is finite, then [Corollary 5.25](#) applies and we are done. So assume that  $I$  is infinite. Fix any  $a = \langle a_i \rangle_{i \in I} \in X$ . Define

$$C = \{x \in X : \text{there is a connected } E \subset X \text{ with } a \in E \text{ and } x \in E\}.$$

Then  $C$  is connected according to [Corollary 5.20](#), so by [Corollary 5.8](#) the subset  $\text{cls } C$  of  $X$  is also connected. To establish that  $X$  is connected we are going to show that  $C$  is dense in  $X$ .

Define  $D$  as in [Lemma 5.28](#). Then it suffices to prove that

$$D \subset C,$$

for then

$$X = \text{cls } D \subset \text{cls } C \subset X.$$

For each integer  $n \geq 1$ , define

$$D_n = \{x \in X : x_i = a_i \text{ for all except at most } n - 1 \text{ values of } i\}.$$

Then

$$D = \bigcup_{n=1}^{\infty} D_n.$$

To prove that  $D \subset C$ , it remains only to show that  $D_n \subset C$  for all  $n \geq 1$ . We use induction on  $n$ . First, separated sets

$$D_1 = \{a\} \subset C.$$

Now let  $n \geq 1$  and assume that  $D_n \subset C$ . We deduce that  $D_{n+1} \subset C$ , too.

Let  $x \in D_{n+1}$ . We must show that  $x \in C$ . Let

$$J = \{j \in I : x_j \neq a_j\},$$

so that  $J$  is a finite subset of  $I$  having at most  $n$  members with

$$x_i = a_i \quad (i \in I \setminus J).$$

Since  $I$  is infinite and  $J$  is finite, we may choose some  $k \in I \setminus J$ . (Using such a  $k$  is, for a general index set  $I$ , the analog of using  $n + 1$  for the case of index set  $\mathbb{N}^*$ .) Define  $y \in X$  by

$$y_i = \begin{cases} x_i & \text{if } i \neq k, \\ a_k & \text{if } i = k. \end{cases}$$

Then  $y \in D_n$ , and so by the inductive assumption  $y \in C$ , that is, there is a connected subset  $A$  of  $X$  with

$$a \in A, \quad y \in A.$$

Now the set

$$B = \{z \in X : z_i = x_i \text{ for all } i \neq k\},$$

being homeomorphic to the connected space  $X_k$  is connected. Moreover,

$$y \in B, \quad x \in B.$$

Then  $a, x \in A \cup B$ , and by [Corollary 5.19](#) the set  $A \cup B$  is connected. Hence  $x \in C$ . This completes the proof of the inductive step that  $D_{n+1} \subset C$ .  $\square$

### EXERCISES FOR SECTION 5.1

1. Which topological spaces having at most three points are connected?
2. Which of the following spaces are connected?
  - (a) An infinite set provided with its finite-complement topology [[Examples 2.3 \(7\)](#)].
  - (b) An uncountable set provided with its countable-complement topology ([Exercise 2.7](#)).
  - (c) The real line provided with its right-interval topology [[Examples 2.72 \(2\)](#)].
  - (d) The line with two origins [[Examples 2.20 \(3\)](#)].
  - (e) The half-disk space [[Examples 2.20 \(3\)](#)].
  - (f) The tangent disk space ([Exercise 2.37](#)).
3. (a) Show that a topological space is connected if and only if each of its nonempty proper subsets has nonempty boundary.  
 (b) If a connected subset  $C$  of a topological space  $X$  intersects both a subset  $A$  of  $X$  and  $X \setminus A$ , must  $C$  intersect  $\text{bdy } A$ ?
4. Given subsets  $A$  and  $B$  of a topological space  $X$ , one says that  $A$  is **separated from  $B$  in  $X$**  when

$$(\text{cls } A) \cap B = \emptyset = A \cap (\text{cls } B)$$

prob:spaces-to-test-for-conn

prob:separated sets

separated sets  
cut point

- (a) Can  $\text{cls } A$  intersect  $\text{cls } B$  when  $A$  is separated from  $B$  in  $X$ ?
- (b) Suppose the topology of  $X$  is induced by a metric  $d$ . Show that  $A$  is separated from  $B$  in  $X$  in case  $d(A, B) > 0$ , where  $d(A, B)$  is the distance between the sets in the sense of [Exercise 1.43](#). Does the converse hold?

5. (Continuation of [Exercise 4](#).)

- (a) If  $A$  and  $B$  are nonempty subsets of a topological space  $X$ , prove that  $A$  is separated from  $B$  in  $X$  precisely when  $\{A, B\}$  is a separation of  $A \cup B$ .
- (b) Deduce from (a) that a subset  $C$  of  $X$  is connected if there do *not* exist nonempty subsets  $A$  and  $B$  of  $X$  such that  $C = A \cup B$  and  $A$  is separated from  $B$  in  $X$ .

6. Let  $X$  be a topological space with  $X = A \cup B$  for nonempty subsets  $A$  and  $B$  of  $X$ .

- (a) Can  $X$  be connected even though  $(\text{cls } A) \cap B = \emptyset$ ?
- (b) Show that  $X$  is connected in case both  $A$  and  $B$  are connected and  $(\text{cls } A) \cap B = \emptyset$ .

7. (a) Let  $X$  be a totally ordered with provided with its order topology [[Examples 2.72 \(1\)](#)]. Find a necessary and sufficient condition on the total ordering for  $X$  to be connected.

(Hint: Which properties of the usual ordering on intervals in  $\mathbb{R}$  were used in the proof of [Theorem 5.4](#)?)

- (b) Apply this condition to the total ordering on  $\mathbb{R} \times \mathbb{R}$  that was defined in [Exercise 2.95](#).

prob:non-homeo-via-conn

8. (a) Show that the real line  $\mathbb{R}$  is not homeomorphic to the union of the  $x$ - and  $y$ -axes in the plane.
- (b) Show that the unit interval  $I$  is not homeomorphic to the unit square  $I^2$ .

prob:cut-point

9. If  $X$  is a connected space, a **cut-point** of  $X$  is a point  $c \in X$  such that  $X \setminus \{c\}$  is disconnected. Any other point is a **non-cut point**. For example,  $1/2$  is a cut-point of  $[0, 1]$ , but  $0$  is not.

ration-of-complement-of-noncut-pt

- (a) Let  $c$  be a cut-point of a connected Hausdorff space  $X$ . If  $\{U, V\}$  is a separation of  $X \setminus \{c\}$ , show that  $U \cup \{c\}$  and  $V \cup \{c\}$  are connected.

prob-part:conn-T2-splits-into-conn

- (b) Let  $X$  be a connected Hausdorff space containing at least two points. Prove that  $X = A \cup B$  for some nonempty connected subsets  $A$  and  $B$  of  $X$  with  $A \neq B$ .

(Hint: Distinguish the cases that  $X$  has or does not have a cut-point.)

- (c) In (b), can  $A$  and  $B$  always be found so that  $A$  is disjoint from  $B$ ?

10. Must the inverse image of a connected space under a continuous surjection be connected?

a-no-surj-to-nondegenerate-discrete

11. (a) Prove that a topological space  $X$  is connected if and only if there does *not* exist any continuous surjection  $f: X \rightarrow \{0, 1\}$ , where  $\{0, 1\}$  has its discrete topology.

- (b) Use (a) to give a new proof that the continuous image of a connected space is connected ([Theorem 5.9](#)).

- (c) Suppose that every continuous real-valued function on a topological space  $X$  has the “intermediate-value property” of [Theorem 5.13](#). Show that  $X$  must be connected.



12. Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  be a continuous function, where  $n \geq 2$ . Suppose  $f(a) < 0$  and  $f(b) > 0$  for some  $a, b \in \mathbb{R}$ . Show that  $f(x) = 0$  for uncountably many points  $x \in \mathbb{R}^n$ . locally constant function  
starlike set  
unicoherent space  
figure-eight
13. Prove that a continuous injection  $f: [a, b] \rightarrow [a, b]$  of a closed interval in  $\mathbb{R}$  must be either increasing or decreasing.
- prob:closed-interval-not-homog 14. Prove that a homeomorphism  $f: [a, b] \cong [a, b]$  of a closed interval in  $\mathbb{R}$  must either map the endpoints  $a$  and  $b$  to themselves or else map them to  $b$  and  $a$ , respectively.  
*Note:* This shows that a closed interval is not homogeneous (Exercise 3.76).
15. Consider a surjection  $f: [a, b] \rightarrow [c, d]$ , where  $a, b, c, d \in \mathbb{R}$  with  $a < b$  and  $c < d$ .  
(a) Prove that  $f$  is a homeomorphism if it is either strictly increasing or strictly decreasing.  
(b) Is the converse of (a) true, too?
16. A map  $f: X \rightarrow Y$  of topological spaces is said to be **locally constant** if each  $x \in X$  has a neighborhood on which  $f$  is constant.  
(a) Give an example of a nonconstant but locally constant continuous function  $f: X \rightarrow \mathbb{R}$  if  $X$  is an appropriate subspace of  $\mathbb{R}$ .  
(b) Show that a locally constant continuous function  $f: X \rightarrow \mathbb{R}$  on a *connected* domain  $X$  must be constant on  $X$ .  
(c) Can (b) be generalized to codomains other than  $\mathbb{R}$ ?
17. Construct a decreasing sequence of connected subsets of the plane  $\mathbb{R}^2$  whose intersection is disconnected, or else prove that no such sequence can exist.
18. Let  $X$  be a topological space having a cover consisting of connected sets each pair of which are *not* separated from each other (Exercise 4). Prove that  $X$  must be connected.
- prob:starlike 19. A set  $E \subset \mathbb{R}^n$  is said to be **starlike** if there is a point  $c \in E$  such that for each  $x \in E$ , the line segment joining  $c$  and  $x$  is contained in  $E$ .  
(a) Show that each convex subset of  $\mathbb{R}^n$  is starlike. Determine all starlike subsets of  $\mathbb{R}^1$ . Describe several nonconvex but starlike subsets of each of  $\mathbb{R}^2$  and  $\mathbb{R}^3$ .  
(b) Prove that every starlike subset of  $\mathbb{R}^n$  is connected
20. A connected space  $X$  is said to be **unicoherent** when the intersection of any two closed connected subsets of  $X$  that cover  $X$  is itself connected. Verify that the unit interval  $I$  is unicoherent but that  $S_1$  is not. Intuitively, does it seem that the 2-sphere  $S_2$  is unicoherent?
- closed-sets-conn-implies-both-conn 21. (a) Let  $A$  and  $B$  be subsets of a topological space  $X$  such that  $A \cup B$  and  $A \cap B$  are connected. Prove: If  $A$  and  $B$  are both closed in  $X$ , then  $A$  and  $B$  are both connected.  
(b) Is the hypothesis that  $A$  and  $B$  be closed really needed in (a)?
22. Is the set of points in the plane at least one of whose coordinates is irrational connected?
23. Show that the Hilbert sequence space  $\ell^2$  (Example 1.10) is connected.
24. If  $n \geq 2$ , show that the sphere  $S_{n-1}$  disconnects  $\mathbb{R}^n$  but that the disk  $D_n$  does not.
25. Show that a circle is not homeomorphic to the figure-eight (Exercise 3.51).

**26.** Prove that  $S_1$  is not homeomorphic to  $S_n$  for any  $n > 1$ .

**27.** Classify the ten digits

0 1 2 3 4 5 6 7 8 9

into distinct homeomorphism classes (that is, into equivalence classes under the relation ‘is homeomorphic to’). You should regard these as “idealized” shapes, in other words, as curves not having any physical width.

**28.** Given proper subsets  $A$  and  $B$  of connected spaces  $X$  and  $Y$ , respectively, prove that  $(X \times Y) \setminus (A \times B)$  is connected.

**29.** Is the real line homeomorphic to the product  $X \times X$  of a topological space with itself?

**30.** Show that the Bing triangle space ([Exercise 2.128](#)) is connected.

*Note:* This shows that there exist denumerable connected Hausdorff spaces!

**31.** Let  $X$  be a countable connected space.

(a) Show that each continuous real-valued function on  $X$  must be constant.

(b) Prove that  $X$  cannot be metrizable if it contains more than one point.

[*Hint:* If the metric  $d$  induces the topology of  $X$ , consider for a fixed point  $a \in X$  the function  $x \mapsto d(x, a)$ .]

prob:conn-mfld-homog **32.** Prove that a connected manifold is homogeneous in the sense of [Exercise 3.76](#).

(*Hint:* Use [Exercise 3.78](#).)

prob:epsilon-chainable **33.** Let  $\langle X, d \rangle$  be a metric space. Given points  $x$  and  $y$  in  $X$  and a number  $\varepsilon > 0$ , an  **$\varepsilon$ -chain from  $x$  to  $y$**  is a finite family  $\langle x_1, x_2, \dots, x_n \rangle$  of points of  $X$  such that

$$x_1 = x, \quad x_n = y, \quad \text{and} \quad d(x_i, x_{i+1}) < \varepsilon \text{ for } i = 1, 2, \dots, n-1.$$

For  $\varepsilon > 0$ , the space  $\langle X, d \rangle$  is said to be  **$\varepsilon$ -chainable** when for all  $x, y \in X$ , there exists some  $\varepsilon$ -chain from  $x$  to  $y$ .

Prove that if  $X$  is connected, then  $\langle X, d \rangle$  is  $\varepsilon$ -chainable for every  $\varepsilon > 0$ .

(*Hint:* Given  $\varepsilon > 0$  and  $x \in X$ , consider the set of all those  $y \in X$  for which there exists an  $\varepsilon$ -chain from  $x$  to  $y$ .)

*Note:* For a converse, see [Exercise 35](#).

(34)–(38) involve compactness along with connectedness.

prob:l-cpt-implies-connected **34.** Deduce the connectedness of the unit interval  $[0, 1]$  from its compactness as follows (or otherwise). Suppose that  $\{A, B\}$  is a separation of  $[0, 1]$ . Define  $f: A \times B \rightarrow \mathbb{R}$  by  $f(x, y) = |x - y|$ . Then  $f$  must attain a minimum value at some point  $\langle a, b \rangle \in A \times B$ . Consider the midpoint  $c$  of the interval with endpoints  $a$  and  $b$ , which is also in  $[0, 1]$ .

conn-if-epsilon-chainable-all-epsilon **35.** Prove that a compact metric space is connected if, for every  $\varepsilon > 0$ , it is  $\varepsilon$ -chainable in the sense of [Exercise 33](#).

**36.** Let  $X$  be a compact Hausdorff space. Show that the intersection of all open, closed subsets of  $X$  containing a given point  $x$  of  $X$  is connected.

(*Hint:* Let  $C$  be that intersection and suppose that  $\{A, B\}$  is a separation of  $C$ . Then  $A$  and  $B$  are disjoint closed subsets of  $X$ , and so there are disjoint open subsets  $U$  and  $V$  of  $X$  such that  $A \subset U$  and  $B \subset V$ , respectively. Show that  $E \subset U \cup V$  for some open, closed subset  $E$  of  $X$  containing  $x$ .)

*Note:* For an application, see [Exercise 61](#).

- 37.** Prove that a countable locally compact Hausdorff space cannot be connected unless it is empty or consists of a single point.

(Hint: If  $X$  is connected and contains more than a single point, then  $\{x\}$  is not open in  $X$  for any  $x \in X$ .)

continuum

non-cut point

Moore, Robert L.

topological group!connected space

prob:continua-and-noncut-pts

- 38.** A **continuum** is a compact, connected Hausdorff space.

Show that a continuum containing at least two points must contain at least two non-cut points (Exercise 9).

[Hint: Assume that  $X$  has a cut-point  $c$  and let  $\{U, V\}$  be a separation of  $X \setminus \{c\}$ . Show that each of  $U$  and  $V$  contains a non-cut point of  $X$ . To see that  $U$  contains a non-cut point of  $X$ , assume the contrary and apply Exercise 9 (a).]

Note: A theorem of R. L. Moore says that a metrizable continuum having exactly two non-cut points must be homeomorphic to the closed unit interval  $[0, 1]$ .

of-e-in-conn-top-grp-generates-grp

- 39.** Let  $G$  be a topological group that is connected. If  $V$  is an arbitrary neighborhood of the identity element of  $G$ , prove that  $G$  is generated by  $V$  in the sense that each  $x \in G$  can be written as a finite product

$$x = v_1 \cdot v_2 \cdot \cdots \cdot v_n$$

of elements of  $V$ .

[Hint: Let  $U = V \cap V^{-1}$ . Show that  $\bigcup_{n=1}^{\infty} U^n$  is an open subgroup of  $G$ , where  $U^n$  is the group power of  $U$  given by

$$U^n = \{u_1 \cdot u_2 \cdot \cdots \cdot u_n : u_1, u_2, \dots, u_n \in U\}.$$

Then use Exercise 3.151 (c).]

## 5.2 Components and Locally Connected Spaces

sec:components

In this section we examine two special kinds of connected subsets of a space: the largest connected pieces into which the space can be divided, and connected neighborhoods of points.

### Components

subsec:components

Even when a topological space  $X$  is disconnected, it is always possible to decompose it into pairwise disjoint connected subsets. For example, since each singleton is connected, we can always decompose  $X$  as

$$X = \bigcup_{x \in X} \{x\}.$$

When  $X$  is discrete, this will be the only way to decompose  $X$  into connected pieces. If  $X$  is not discrete, however, there may be other—and even many other—ways of so decomposing  $X$ . For example, the disconnected space

$$X = \{x \in \mathbb{R} : 0 \leq x \leq 1 \text{ or } 2 < x < 3\}$$

has the following three distinct decompositions into connected subsets:

$$X = [0, 1/2[ \cup [1/2, 1] \cup ]2, 7/3] \cup ]7/3, 3[,$$

$$X = \{0\} \cup ]2, 3[ \cup \bigcup_{n=1}^{\infty} ]1/(n+1), 1/n],$$

$$X = [0, 1] \cup ]2, 3[.$$

Geometric intuition tells us that only in the third of these decompositions have we really broken  $X$  into “the” connected pieces of  $X$ . What distinguishes this third decomposition from the other two is that neither of the connected subsets  $[0, 1]$  and  $]2, 3[$  can be enlarged to a new connected subset of  $X$ .

def:component **5.30 Definition.** A connected subset of a topological space  $X$  that is not *properly* contained in any connected subset of  $X$  is called a **component** of  $X$ .

Among all the connected subsets of a topological space  $X$ , a component is thus one that is *maximal* with respect to inclusion. (We say ‘maximal’ rather than ‘maximum’ because a component need not contain every connected subset of  $X$ , but it just cannot be contained in any connected subset of  $X$  different from itself.)

thm:components-closed **5.31 Theorem.** Each component of a topological space is closed in the space.

**Proof.** Let  $C$  be a component of a topological space  $X$ . Since  $C$  is connected, by [Theorem 5.7](#) the subset  $\text{cls } C$  of  $X$  is also connected. Since  $C \subset \text{cls } C$  and  $C$  is maximal with respect to inclusion, it follows that  $C = \text{cls } C$ . Thus  $C$  is closed in  $X$ .  $\square$

Trivially, the empty space has  $\emptyset$  as its unique component. If a topological space  $X$  is nonempty, then  $\emptyset$  is not a component of  $X$ , because it is properly contained in each singleton and a singleton is connected. The following theorem says that  $X$  can really be “decomposed” into its components.

thm:components-partition-space **5.32 Theorem.** Let  $X$  be a nonempty topological space. Then:

- thm-part:pt-in-unique-component
- (1) Each point  $x$  of  $X$  belongs to a unique component of  $X$ , namely, the union of all connected subsets of  $X$  that contain  $x$ .
  - (2) The collection of all components of  $X$  is a partition of  $X$ .

**Proof.** (1) Let  $x \in X$ . To see that  $x$  belongs to at most one component of  $X$ , just suppose that  $A$  and  $B$  are components of  $X$  with  $x \in A$  and  $x \in B$ . By [Corollary 5.19](#), the set  $A \cup B$  is connected. Since  $A \subset A \cup B$  and  $A$  is a component of  $X$ , then  $A = A \cup B$ ; similarly,  $B = A \cup B$ . Hence  $A = B$ .

Now define

$$C_x = \bigcup \{A : x \in A \text{ and } A \text{ is connected}\}.$$

We show that  $C_x$  is a component of  $X$  that contains  $x$ . Since  $\{x\}$  is connected,  $\{x\} \subset C_x$ , so that  $x \in C_x$ . By [Corollary 5.19](#), the set  $C_x$  is connected. Finally, if  $A$  is a connected subset of  $X$  with  $C_x \subset A$ , then  $x \in A$ ,  $A \subset C_x$  by definition of  $C_x$ , and hence  $C_x = A$ .

- (2) We have already observed that no component of  $X$  is empty. By (1),  $X$  is the union of all its components, and distinct components of  $X$  are disjoint.  $\square$

According to (1), the component of  $X$  containing a given point  $x$  of  $X$  is actually the largest connected subset of  $X$  that contains  $x$ .

cor:conn-set-in-unique-component

**5.33 Corollary.** A nonempty connected subset  $A$  of a topological space  $X$  is contained in a unique component of  $X$ , namely, the component of  $X$  to which each point of  $A$  belongs.

component!homeomorphism@and  
homeomorphism!components@and

exs:components **5.34 Examples.** (1) A topological space  $X$  is connected precisely when  $X$  is the one and only component of  $X$ .

- (2) If  $A$  is a nonempty connected subset of a topological space  $X$  that is both open and closed in  $X$ , then necessarily  $A$  is a component of  $X$ . In fact, by [Corollary 5.33](#) the connected set  $A$  is contained in some component  $C$  of  $X$ , and so from [Lemma 5.17](#) it follows that  $C = A$ .

For example, the components of  $X = [0, 1] \cup ]2, 3[$  are  $[0, 1]$  and  $]2, 3[$ ; the components of the punctured line  $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$  are the open rays  $] -\infty, 0[$  and  $]0, +\infty[$ .

- (3) The components of a nonempty discrete space  $X$  are just the singletons  $\{x\}$  for all  $x \in X$ . (This class of examples will be generalized by “totally disconnected” spaces—see the [subsection “Totally disconnected spaces”](#).)

ex:Q-components

- (4) Although the space  $\mathbb{Q}$  is not discrete, nonetheless the components of  $\mathbb{Q}$  are the singletons  $\{x\}$  for all  $x \in \mathbb{Q}$ . In fact, a component of  $\mathbb{Q}$ , being a nonempty connected subset of  $\mathbb{R}$ , is by [Theorem 5.4](#) an interval in  $\mathbb{R}$ , and a nonempty interval cannot be contained in  $\mathbb{Q}$  unless it is a singleton.

Similarly, the components of the space  $\mathbb{R} \setminus \mathbb{Q}$  of all irrational numbers are the singletons  $\{x\}$  for all  $x \in \mathbb{R} \setminus \mathbb{Q}$ .  $\diamond$

Let  $f: X \rightarrow Y$  be a continuous map. If  $C$  is a component of  $X$ , then its image  $f(C)$  is connected, and so  $f(C) \subset D$  for some component  $D$  of  $Y$ . Suppose now that  $f$  maps  $X$  homeomorphically onto  $Y$ . In the notation just used, the inverse image  $f^{-1}(D)$  is a connected subset of  $X$  such that  $C \subset f^{-1}(D)$ ; then  $C = f^{-1}(D)$  by the maximality of  $C$  among connected sets; and hence  $f(C) = D$ . Thus  $f$  maps each component of  $X$  homeomorphically onto a component of  $Y$ . The same argument applies to  $f^{-1}: Y \rightarrow X$ . These observations establish the following result.

m:homeo-induces-bij-of-components

**5.35 Theorem.** Let  $\mathcal{C}$  and  $\mathcal{D}$  denote the collections of all components of topological spaces  $X$  and  $Y$ , respectively, and let

$$f: X \cong Y$$

be a homeomorphism. Then the map

$$\begin{aligned} C &\rightarrow \mathcal{D} \\ C &\mapsto f(C) \end{aligned}$$

is a bijection between  $\mathcal{C}$  and  $\mathcal{D}$  with

$$C \cong f(C) \quad (C \in \mathcal{C}).$$

This theorem can sometimes be used to establish that two spaces are not homeomorphic to one another, as in the following examples.

exs:components-and-homeomorphisms  
 ki-non-homeo-unions-seq-intervals  
 5.36 Examples. (1) The spaces  
 $X = [0, 1] \cup [2, 3], \quad Y = [0, 1] \cup [2, 3] \cup [4, 5]$   
 are not homeomorphic, because  $X$  has two components whereas  $Y$  has three.

(2) This example is due to Kuratowski. Let

$$X = \bigcup_{n=0}^{\infty} (]3n, 3n+1[ \cup \{3n+2\}),$$

$$Y = ]0, 1] \cup \bigcup_{n=1}^{\infty} (]3n, 3n+1[ \cup \{3n+2\})$$

(see Figure 5.6). The  $X$  is not homeomorphic to  $Y$ , because the components of

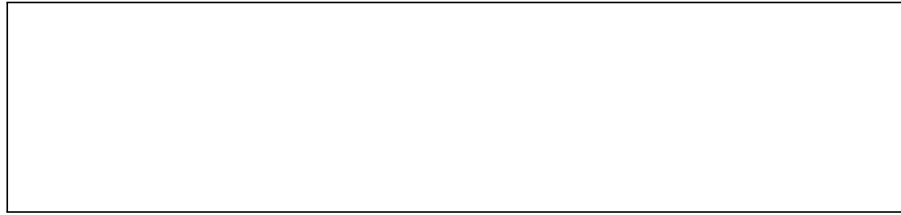


Figure 5.6: Kuratowski's example of a continuous bijection  $f$  between non-homeomorphic unions of sequences of intervals.

fig:Kuratowski-cont-bij-f-of-non-hom

$X$  are all open intervals and singletons, whereas the component  $]0, 1]$  of  $Y$  is not homeomorphic to either an open interval or a singleton. Nevertheless, there are continuous bijections

$$f: X \rightarrow Y, \quad g: Y \rightarrow X$$

given by

$$f(x) = \begin{cases} x & \text{if } x \neq 2, \\ 1 & \text{if } x = 2 \end{cases}$$

(see Figure 5.6 again), and

$$g(y) = \begin{cases} \frac{1}{2}y & \text{if } y \in ]0, 1], \\ \frac{1}{2}(y-2) & \text{if } y \in ]3, 4[, \\ y-3 & \text{otherwise} \end{cases}$$

(see Figure 5.7)  $\diamond$



Figure 5.7: Kuratowski's example of a continuous bijection  $g$  between non-homeomorphic unions of sequences of intervals.

fig:Kuratowski-cont-bij-g-of-non-hom

### Totally disconnected spaces

subsec:totally-disconnected

The largest component a space  $X$  can possibly have is the entire set  $X$ , namely, when  $X$  is itself connected. At the opposite extreme, the smallest components  $X$  can have, if  $X$  is nonempty, are the singletons  $\{x\}$  for  $x \in X$ .

def:totally-disconnected

**5.37 Definition.** A topological space  $X$  is said to be **totally disconnected** if no component of  $X$  consists of more than a single point.

Since any nonempty connected subset of a topological space  $X$  is contained in some component of  $X$ , the space  $X$  is *totally disconnected* if and only if *the only connected subsets of  $X$  are the empty set  $\emptyset$  and the singletons  $\{x\}$  for  $x \in X$ .*

exs:totally-disconnected

**5.38 Examples.** (1) Any discrete space is totally disconnected.

ex:Q-and-RlessQ-tot-disconn

(2) By [Examples 5.34 \(4\)](#), the nondiscrete spaces  $\mathbb{Q}$  and  $\mathbb{R} \setminus \mathbb{Q}$  are totally disconnected.

(3) The space  $[0, 1] \cup [2, 3]$  is disconnected but not totally disconnected.

ex:Cantor-K-totally-disconnected

(4) The Cantor set  $K$  (see [Example 4.16](#)) is totally disconnected.

To see that  $K$  is totally disconnected, just suppose it has a component containing points  $a$  and  $b$  with  $a < b$ . Then  $[a, b] \subset K$  by [Theorem 5.4](#). We shall show this to be impossible.

Choose a positive integer  $n$  so large that  $3^{-n} < b - a$ . In the notation of [Example 4.16](#), the points  $a$  and  $b$  both belong to the set  $K_n$ , which is the union of finitely many pairwise disjoint closed intervals each of length  $3^{-n}$ . Then  $a$  and  $b$  must belong to two different ones of these intervals, which means that  $[a, b] \not\subset K_n$ . Since  $K \subset K_n$ , we conclude that  $[a, b] \not\subset K$ .

ex:explosion-pt

(5) In 1921, Bronisław Knaster and Kazimierz Kuratowski constructed an uncountable connected subspace of the plane having the remarkable property that removing just a single point from it leaves a totally disconnected space. Such a point in a connected space is called an **explosion point** (or *dispersion point*) of the space. We proceed to describe the Knaster–Kuratowski example.

Identify the real line  $\mathbb{R}$  with the  $x$ -axis  $\mathbb{R} \times \{0\}$  in the Euclidean plane under the homeomorphism  $x \mapsto \langle x, 0 \rangle$ . Then the Cantor set  $K$  is identified with the subset  $K \times \{0\}$  of the  $x$ -axis. Denote by  $E$  the set of all endpoints of  $K$ —that is, the set of all endpoints of all the intervals comprising the sets  $K_n$  (compare [Exercise 4.14](#))—and let

$$F = K \setminus E,$$

the set of all nonendpoints of  $K$ . Let  $p$  be the point

$$p = \langle 1/2, 1/2 \rangle$$

lying above the  $x$ -axis; this is going to become the explosion point. For each  $t \in K$  let  $L_t$  be the line segment in  $\mathbb{R}^2$  joining  $t$  and  $p$ . Form the set

$$X_E = \{ \langle x, y \rangle \in L_t : t \in E, y \in \mathbb{Q} \}$$

consisting of all points having *rational* ordinate and lying on the line segment joining  $p$  to some endpoint of  $K$ ; and form the set

$$X_F = \{ \langle x, y \rangle \in L_t : t \in F, y \notin \mathbb{Q} \}$$

Cantor set!totally disconnected@is t  
explosion point  
Knaster-Kuratowski space@Knaster-  
Knaster, Bronis\law  
Kuratowski, Kazimierz  
explosion point

consisting of all points having *irrational* ordinate and lying on the line segment joining  $p$  to some *nonendpoint* of  $K$ . Finally, let

$$X = X_E \cup X_F$$

(see Figure 5.8). This is the **Knaster–Kuratowski space**.



Figure 5.8: The Knaster–Kuratowski example of an uncountable connected space having an explosion point.

fig:Knaster-Kuratowski-explosion-pt

Note that each  $t \in F$  has the rational ordinate 0, and so

$$F \cap X = \emptyset.$$

The easy thing to show is that the complement  $X \setminus \{p\}$  of the point  $p$  is totally disconnected. In fact, let  $C$  be a component of  $X \setminus \{p\}$ . Suppose, first, that  $C$  intersects two line segments  $L_t$  and  $L_s$  for  $t, s \in K$  with  $t < s$ . Choose  $u \in \mathbb{R}$  with

$$t < u < s, \quad u \notin K;$$

such a  $u$  exists because  $[t, s] \not\subset K$  [see (4)]. Then the line  $L$  in  $\mathbb{R}^2$  through  $p$  and  $u$  contains no points of  $X \setminus \{p\}$ , and so the intersections with  $C$  of the two open half-planes into which  $L$  divides  $\mathbb{R}^2$  form a separation of  $C$ . Thus  $C \subset L_t$  for some one  $t \in K$ . Since  $L_t$  is totally disconnected [compare (2)], it follows that  $C$  consists of just a single point.

The hard thing to show is that  $X$  is connected; our proof will exploit the Baire Category Theorem (1.91). Just suppose there is some separation  $\{A, B\}$  of  $X$ , say with

$$p \in A.$$

Since  $A$  and  $B$  are disjoint and closed in  $X$ ,

$$\{eqn:A-B-cl-rel-KK\} \quad (*) \quad (\text{cls } A) \cap B = \emptyset = A \cap (\text{cls } B).$$

[Relation (\*) holds no matter whether the closures are taken in  $X$  or in  $\mathbb{R}^2$ ; but here and below all closures are taken in the plane  $\mathbb{R}^2$ .] We shall eventually contradict (\*) by showing that  $X \subset \text{cls } A$ .



For each  $t \in K$ , define a point

$$f(t) \in L_t$$

as follows. Let

$$f(t) = t \text{ in case } L_t \cap B = \emptyset.$$

However, in case  $B$  intersects  $L_t$ , let  $f(t)$  be the “least upper bound of points belonging to  $B \cap L_t$ ,” or more precisely,

$$f(t) = \text{the point } (x, y) \in L_t \text{ for which } y = y_t$$

where

$$y_t = \sup\{y : \langle x, y \rangle \in L_t \cap B \text{ for some } x\}.$$

Then in either case,

$$f(t) \neq p \quad (t \in K)$$

because  $p \in A$  and  $A$  is open in  $X$ .

Let  $t \in K$ . If  $f(t) \neq t$ , then  $f(t) \in (\text{cls } A) \cap (\text{cls } B)$ , and so  $f(t) \notin X$  by (\*). Thus

$$f(t) \notin X \quad \text{or} \quad f(t) = t.$$

Moreover, since  $F \cap X = \emptyset$ ,

$$f(t) \in X \implies f(t) = t \in E.$$

Define

$$S = \{f(t) : t \in F\}.$$

By the preceding paragraph,  $t \in F$  implies  $f(t) \in L_t \setminus X$ , and so  $f(t)$  has rational ordinate. Denoting the set of rational numbers in  $\langle 0, 1/2 \rangle$  by  $Q$ , we can therefore write

$$S = S_0 \cup \bigcup_{q \in Q} S_q$$

where for each  $q \in Q \cup \{0\}$ ,

$$S_q = \{f(t) : t \in F \text{ and the ordinate of } f(t) \text{ is } q\}.$$

Then

$$\text{cls } S_q \subset \mathbb{R} \times \{q\} \quad (q \in Q \cup \{0\}).$$

Clearly

$$S_0 \subset F.$$

Let  $q \in Q$ . Then  $S_q \subset (\text{cls } A) \cap (\text{cls } B)$ , and so  $(\text{cls } S_q) \cap X = \emptyset$  by (\*). Now a point  $\langle x, y \rangle \in L_t$  for  $t \in E$  belongs to  $X$  when  $y$  is rational, and each point of  $\text{cls } S_q$  has a rational ordinate. Hence

$$\{eqn:cls-S_q-meetL_t-KK\} \quad (**) \quad (\text{cls } S_q) \cap L_t = \emptyset \quad (q \in Q, t \in E).$$

For each  $q \in Q$ , define

$$T_q = \{t \in K : (\text{cls } S_q) \cap L_t \neq \emptyset\}.$$

From (\*\*),

$$T_q \subset F \quad (q \in Q),$$

explosion point

Since  $t \in F$  implies  $f(t) \in S_q$  for some  $q \in Q \cup \{0\}$ , we have

Knaster-Kuratowski space@Knaster-Kuratowski space  
component

$$F \subset S_0 \cup \bigcup_{q \in Q} T_q.$$

It follows that

{eqn:K-is-E-union-S0-union-Tqs-KK} (\*\*\*)

$$K = E \cup S_0 \cup \bigcup_{q \in Q} T_q.$$

We are going to show that  $S_0$  is dense in  $K$ . Then the set  $D$  defined by

$$D = X \cap \bigcup_{t \in S_0} L_t$$

will be dense in  $X$ . The purpose for showing that is as follows. When  $t \in S_0$ , then  $f(t) = t$  and hence  $X \cap L_t \subset A$  according to the definition of  $f(t)$ . Thus we will be able to conclude that  $D \subset A$  and hence that  $X \subset \text{cls } A$ , as was desired.

To prove that  $S_0$  is in fact dense in  $K$  we shall apply the Baire Category Theorem (1.91) to  $K$  provided with its Euclidean metric. (By Theorem 1.76 and Theorem 1.81 this metric is complete because  $K$ , being the intersection of closed subsets of  $\mathbb{R}$ , is closed in  $\mathbb{R}$ .) From (\*\*\*) we obtain the representation

$$S_0 = \left( \bigcap_{t \in E} (K \setminus \{t\}) \right) \cap \left( \bigcap_{q \in Q} (K \setminus T_q) \right)$$

of  $S_0$  as the intersection of countably many subsets of  $K$ . It remains now only to show that each of these subsets is open and dense in  $K$ , or equivalently, that the complement of each is closed in  $K$  and contains no nonempty open subset of  $K$ .

If  $t \in E$ , then  $\{t\}$  is closed in  $K$  and by Example 4.16, not open in  $K$ . Now let  $q \in Q$ . One can show directly that  $T_q$  is closed in  $K$  (or can argue that  $T_q$  is the image of the compact set  $\text{cls } S_q$  under the continuous map that, for each  $t \in K$ , projects all points on the line segment  $L_t$  to the point  $t$ ). Since  $E$  is dense in  $K$  and disjoint from  $F$ , then  $F$  cannot contain a nonempty open subset of  $K$ , and so the same is true of its subset  $T_q$ .  $\diamond$

The argument just given should dispel any lingering belief that specific, concrete examples are necessarily easier to understand than general, abstract theorems!

Even though it is uncountable, the Cantor set can, in two senses, be thought of as “dust,” not having much in it: it is zero-dimensional and it is totally connected. These two topological properties are related.

prop:0-dim-T2-then-tot-disc

**5.39 Proposition.** *A zero-dimensional Hausdorff space is totally disconnected.*

### Locally connected spaces

In Euclidean space  $\mathbb{R}^n$ , any neighborhood  $U$  of any point  $x$  contains some connected neighborhood of  $x$ , for  $U$  contains some  $d$ -ball  $B_x(\varepsilon; d)$ , where  $d$  is the Euclidean metric, and  $B_x(\varepsilon; d)$  is connected by Examples 5.21 (1) because it is convex.

def:loc-conn

**5.40 Definition.** A topological space  $X$  is said to be **locally connected** if at each point  $x \in X$  there is a local base consisting of connected sets—in other words, if for each neighborhood  $U$  of  $x$  there is a connected neighborhood  $V$  of  $x$  with  $V \subset U$ .

Thus a locally connected space is one having “arbitrarily small” neighborhoods at each point.

Euclidean  $n$ -space@Euclidean  $\mathbb{R}^n$ -space  
discrete space!locally connected space  
topologist's comb

exs:loc-conn-spaces **5.41 Examples.** (1) For each  $n \geq 1$ , Euclidean space  $\mathbb{R}^n$  is both connected and locally connected.

(2) A discrete space  $X$  is locally connected, since for each  $x \in X$  the singleton  $\{x\}$  is a connected neighborhood of  $x$  that is contained in every neighborhood of  $x$ . As soon as  $X$  contains at least two points, however, it is not connected.

ex:comb (3) A connected space need not be locally connected, even if it is a Hausdorff space. Hence **a Hausdorff space need not be locally connected even though each of its points has a connected neighborhood.** (Contrast this with the fact that a Hausdorff space is locally compact if each of its points has a compact neighborhood: see [Proposition 4.68](#).)

For an example, start with the set

$$K = \{0\} \cup \{1/n : n = 1, 2, 3, \dots\}$$

on the  $x$ -axis and let  $C$  to be the subspace of the plane  $\mathbb{R}^2$  given by

$$C = ([0, 1] \times \{0\}) \cup (K \times [0, 1])$$

(see [Figure 5.9](#)). The topological space  $C$  is called the **topologist's comb**. It consists of its horizontal “shaft”

$$S = [0, 1] \times \{0\}$$

along with its vertical “teeth”

$$T_0 = \{0\} \times [0, 1]$$

and

$$T_n = \{1/n\} \times [0, 1] \quad (n = 1, 2, 3, \dots).$$

The shaft and teeth of the topologist's comb are all connected, each being homeomorphic to the unit interval  $[0, 1]$ . Since the shaft  $S$  intersects the  $n$ th tooth  $T_n$  for each  $n$ , it follows from [Theorem 5.18](#) that  $C$  is connected.

To see that  $C$  is *not* locally connected, consider the point

$$p = \left\langle 0, \frac{1}{2} \right\rangle$$

half-way up on the left-most tooth  $T_0$ . Consider a neighborhood  $U$  of  $p$  in  $C$  having the form

$$U = W \cap C$$

where

$$W = ]-\varepsilon, \varepsilon[ \times \left] \frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon[ , \quad 0 < \varepsilon < \frac{1}{2}$$

(such a neighborhood is highlighted in [Figure 5.9](#)). Just suppose that  $U$  contains a connected neighborhood  $V$  of  $p$  in  $C$ . For some sufficiently large positive  $n$ , the set  $V$  intersects  $T_n$ . Now for such  $n$ , the set

$$U \cap T_n = W \cap T_n = \left\{ \frac{1}{2} \right\} \times \left] \frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon[$$

is both open and closed in  $U$ . By [Lemma 5.17](#), the set  $V \subset U \cap T_n$ , but this is impossible because  $p \in V$ .  $\diamond$

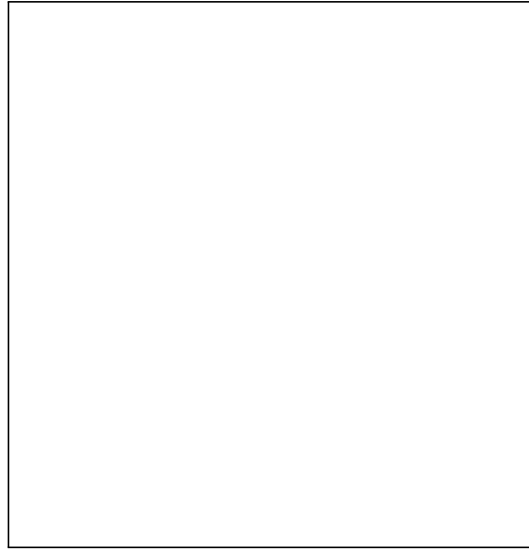


Figure 5.9: The topologist's comb: a connected space that is not locally connected.

fig:comb

The preceding example of the topologist's comb also shows that **a subspace  $Y$  of a locally connected space  $X$  need not be locally connected**, even if  $Y$  is closed in  $X$ . In one case, however, we can deduce that  $Y$  is locally connected.

**5.42 Proposition.** *An open subspace of a locally connected space is itself locally connected.*

**Proof.** Let  $Y$  be an open subspace of a locally connected space  $X$ . Let  $U$  be an arbitrary neighborhood in  $Y$  of an arbitrary point  $x$  of  $Y$ . Then  $U$  is also a neighborhood of  $x$  in  $X$ , and so there exists a connected neighborhood  $V$  of  $x$  in  $X$  with  $V \subset U$ . But  $V$  is a neighborhood of  $x$  in  $Y$ , too.  $\square$

As the next example demonstrates, **the continuous image of a locally connected space need not itself be locally connected**.

**5.43 Example.** Let

$$Y = \{0\} \cup \{1/n : n = 1, 2, 3, \dots\}.$$

The discrete space  $\mathbb{N}$  is locally connected, and the map  $f : \mathbb{N} \rightarrow Y$  given by

$$f(n) = \begin{cases} 0 & \text{if } n = 0, \\ 1/n & \text{if } n > 0 \end{cases}$$

is a continuous bijection. However, the space  $Y$  is not locally connected, because the point 0 of  $Y$  has no connected neighborhood in  $Y$ —the argument is similar to the one used for the topologist's comb [Examples 5.41 (3)].  $\diamond$

Local connectedness is, however, preserved under a large class of continuous maps. To prove this we shall use the following result that relates the two main notions, local connectedness and components, being considered in this section.

thm:loc-conn-via-components

**5.44 Theorem.** *A necessary and sufficient condition for a topological space to be locally connected is that each component of each open subset of the space be open in the entire space.*

**Proof.** Let  $X$  be a topological space.

Necessity. Assume that  $X$  is locally connected. Let  $U$  be an open subset of  $X$  and let  $C$  be a component of  $U$ . To show that  $C$  is open in  $X$ , we show that it is a neighborhood of each of its points. Let  $x \in C$  be arbitrary. Since  $U$  is a neighborhood of  $x$  in  $X$ , there is a connected neighborhood  $V$  of  $x$  in  $X$  with  $V \subset U$ . Since  $V$  is connected and  $C$  is the component of  $U$  containing  $x$ , then  $V \subset C$ . Hence  $C$  is a neighborhood of  $x$  in  $X$ .

Sufficiency. Conversely, assume that the condition on components of open subsets holds. Let  $x \in X$  and let  $W$  be an arbitrary neighborhood of  $x$  in  $X$ . Choose an open set  $U$  with  $x \in U \subset W$ . Then the component of  $U$  to which  $x$  belongs is a connected open neighborhood of  $x$  in  $X$  that is contained in  $W$ .  $\square$

thm:quot-map-preserves-loc-conn

**5.45 Theorem.** *If the domain of a quotient map is locally connected, then so is its range.*

**Proof.** Let  $f: X \rightarrow Y$  be a quotient map. Let  $V$  be an arbitrary open subset of  $Y$  and let  $E$  be a component of  $V$ . In view of the preceding [Theorem 5.44](#), it suffices to show that  $E$  is open in  $Y$ . Since  $f$  is a quotient map, it therefore suffices to show that the inverse image  $f^{-1}(E)$  is open in  $X$ . Now by [Theorem 5.44](#), each component of  $f^{-1}(V)$  is open in  $X$ , and so it suffices to show that  $f^{-1}(E)$  is a union of components of  $f^{-1}(V)$ .

Let  $x \in f^{-1}(E)$  and let  $C$  be an arbitrary component of  $f^{-1}(V)$  such that  $x \in C$ . It remains to show that  $C \subset f^{-1}(E)$ , that is,  $f(C) \subset E$ . Now  $f(C)$  is a connected subset of  $V$  containing  $f(x)$ . But  $E$  is a component of  $V$  with  $f(x) \in E$ . Hence indeed  $f(C) \subset E$ .  $\square$

or:quot-space-of-loc-conn-is-loc-conn

**5.46 Corollary.** *A quotient space of a locally connected space is itself locally connected.*

For example, the torus [[Examples 3.81 \(4\)](#)], the Klein bottle [[Examples 3.81 \(6\)](#)], and the Möbius strip [[Examples 3.81 \(5\)](#)] are all locally connected, being quotient spaces of the square  $I^2$ .

:cont-open-or-closed-image-loc-conn

**5.47 Corollary.** *The image of a locally connected space under a continuous open map or a continuous closed map is locally connected.*

cor:loc-conn-top-prop

**5.48 Corollary.** *Local connectedness is a topological property.*

The preservation of local connectedness under the formation of product spaces is left to the exercises (see Exercises [58–59](#)).

## EXERCISES FOR SECTION 5.2

- 40.** Show that the complement  $\mathbb{R}^2 \setminus S_1$  of the unit circle has two components of which  $S_1$  is the common boundary in  $\mathbb{R}^2$ .

component!product space@Note: The *Jordan curve theorem* generalizes this result from the unit circle to any simple closed curve, that is, any homeomorphic image of the unit circle. (For more about this, see [Theorem 5.98](#) and the discussion accompanying it.)

zero-dimensional space!zero-dimensional space@and totally disconnected space  
totally disconnected space!zero-dimensional space@and zero-dimensional space

41. Find the components of each of the spaces in [Exercise 2](#).

zero-dimensional space!zero-dimensional space@and totally disconnected space  
zero-dimensional space!zero-dimensional space@and totally disconnected space

42. According to [Examples 1.22 \(3\)](#), a proper open subset  $U$  of  $\mathbb{R}$  is the union of countably many pairwise disjoint bounded open intervals and open rays. Show that these intervals and rays are the components of  $U$ .

43. Suppose that a topological space  $X$  can be partitioned into a finite number  $n$  of connected sets  $C_1, C_2, \dots, C_n$  but cannot be partitioned into any smaller number of connected sets. Must  $C_1, C_2, \dots, C_n$  be the components of  $X$ ?

prob:equiv-rel-from-components 44. Define a binary relation  $\sim$  on a nonempty topological space  $X$  by the rule:

$$x \sim y \iff \text{there exists a connected set } C \subset X \text{ with } x \in C \text{ and } y \in C.$$

Verify that  $\sim$  is an equivalence relation on  $X$  and that the equivalence classes of  $X$  under  $\sim$  are precisely the components of  $X$ .

45. Construct homeomorphic embeddings of each of the spaces  $X$  and  $Y$  of [Examples 5.36 \(2\)](#) into the other.

46. Show that the image of a component of a topological space  $X$  under a continuous surjection  $f: X \rightarrow Y$  need not be a component of  $Y$ .

prob:components-of-prod-2 47. Prove that the components of the product of two topological space are the products of components in the individual spaces.

[Hint: Let  $p_1: X_1 \times X_2 \rightarrow X_1$  and  $p_2: X_1 \times X_2 \rightarrow X_2$  be the projections of the product space. If  $C$  is a component of  $X_1 \times X_2$ , then  $C \subset p_1(C) \times p_2(C)$ . If  $C_1$  and  $C_2$  are components of  $X_1$  and  $X_2$ , respectively, then  $C_1 \times C_2 \subset C$  for some component  $C$  of  $X_1 \times X_2$ , and so  $C_1 = p_1(C_1 \times C_2) \subset p_1(C)$ .]

48. (a) Generalize [Exercise 47](#) to the product of any finite family of spaces.

(b) Generalize it to the product of a sequence of metrizable spaces.

(c) Generalize it to the product of an arbitrary family of topological spaces.

49. In the notation of [Exercise 44](#), prove that the quotient space  $X/\sim$  is totally disconnected.

50. (a) Prove that total disconnectedness is preserved under the formation of subspaces

(b) Prove that total disconnectedness is preserved under the formation of product spaces.

Note: By contrast, **a quotient space of a totally disconnected space need not be totally disconnected**: this follows from the total disconnectedness of the Cantor set [[Examples 5.38 \(4\)](#)] together with the theorem of Alexandroff and Urysohn cited within [Example 4.16](#).

prob-part-of-tot-disc-not-tot-disc 51. (a) Prove [Proposition 5.39](#): a zero-dimensional Hausdorff space must be totally disconnected.

(Note: For a converse, see the next problem.)

(b) By considering the subspace  $X \setminus \{p\}$  of the Knaster–Kuratowski space [[Examples 5.38 \(5\)](#)], show that the converse of (a) fails. (Note: Compare [Exercise 61](#).)

prob:tot-disc-loc-cpt-then-0-dim 52. Prove that a totally disconnected *locally compact*  $T_2$ -space must be zero-dimensional.

53. Prove or disprove: Each convex subspace of  $\mathbb{R}^n$  is locally connected. quasicomponent

54. Prove that a topological space is locally connected if and only if the collection of all its connected open subsets is a base of its topology.

55. (a) Show that an  $n$ -manifold (Definition 3.40) is locally connected.

(b) Is the same true of any  $n$ -manifold-with-boundary (Exercise 3.91)?

56. Let  $p = \langle 0, 0 \rangle \in \mathbb{R}^2$  and let  $L = \{0\} \times [-1, 1] \subset \mathbb{R}^2$ .

(a) Which, if any, of the spaces  $G$ ,  $G \cup \{p\}$ , and  $G \cup L$  are locally connected if

$$G = \{ \langle x, \sin(1/x) \rangle : 0 < x \in \mathbb{R} \}$$

(see Figure 5.11 in Section 5.3).

(b) Repeat (a) if now

$$G = \{ \langle x, x \sin(1/x) \rangle : 0 < x \in \mathbb{R} \}.$$

57. By constructing a continuous map from the ray  $[0, +\infty[$  in  $\mathbb{R}$  onto the topologist's comb [(3)], show that a connected space that is the continuous image of a locally connected space need not be locally connected.

(Note: Contrast Exercise 62.)

58. (a) Prove that the product of two nonempty topological spaces is locally connected if and only if both of the spaces are locally connected.

(b) Generalize (a) to the product of any finite family of nonempty spaces.

59. (a) Show that the product space  $\times_{n=1}^{\infty} X_n$  is *not* locally connected if  $X_n = \{0, 1\}$ , the two-point discrete space, for every  $n \geq 1$ . Show, by contrast, that the product *is* locally connected if  $X_1 = X_2 = \{0, 1\}$  again but now  $X_n = ]-1/n, 1/n[$  for all  $n \geq 3$ .

(b) What generalization of Exercise 58 does (a) suggest? Prove that generalization.

60. The **quasicomponent** of a point  $x$  in a space  $X$  is the intersection of all open-closed subsets of  $X$  that contain  $x$ . (Note that such a quasicomponent is necessarily closed in  $X$ . A **quasicomponent** of  $X$  is the quasicomponent of some point of  $X$ .)

(a) A quasicomponent of  $X$  is an equivalence class for the equivalence relation  $\approx$  defined by  $x \approx y$  if and only if there is *no* open partition  $\{U, V\}$  of  $X$  with  $x \in U$  and  $y \in V$ .

(b) Show that the quasicomponents of all the points in  $X$  form a partition of  $X$ .

(c) Prove that the component of  $X$  containing a given point is a subset of the quasicomponent of  $x$ .

Each component of  $X$  is contained in some quasicomponent, and each quasicomponent of  $X$  is a union of components of  $X$ .

However, a quasicomponent need not be a component.

Example:  $X = (K \times [0, 1]) \cup \{ \langle 0, 0 \rangle \} \cup \{ \langle 0, 1 \rangle \}$  where the set  $K$  is given by  $K = \{1/n : n = 1, 2, 3, \dots\}$ . Then the points  $\langle 0, 0 \rangle$  and  $\langle 0, 1 \rangle$  belong to different components of  $X$  but belong to the same quasicomponent.

An open component is a quasicomponent

(d) In a locally connected space  $X$ , each component is open, and the components of  $X$  are the same as the quasicomponents of  $X$ .

The following exercises invoke compactness.

topological group!compactness of topological group

61. (a) Prove that the component of a compact  $T_2$ -space containing a given point is the intersection of all open, closed subsets of the space containing that point.
- (b) Deduce from (a) that a compact  $T_2$ -space is zero-dimensional (Definition 2.61) if it is totally disconnected.

62. Show that a compact locally connected space has only finitely many components.

63. Prove that a Hausdorff space that is the continuous image of a compact locally connected space is necessarily locally connected.

*Note:* Hence any Hausdorff space that is the continuous image of the unit interval  $I$  is locally connected.

64. (a) Let  $\langle X, d \rangle$  be a compact, locally connected metric space. Given any  $\varepsilon > 0$ , establish the existence of some  $\delta > 0$  such that any two points  $x$  and  $y$  of  $X$  satisfying  $d(x, y) < \delta$  both belong to a connected subset  $C$  of  $X$  for which  $\text{diam } C < \varepsilon$ .

- (b) Is the hypothesis that  $X$  be compact actually needed in (a)?

item

Let  $G$  be a topological group (Exercise 3.145) and let  $C$  be the component of the identity element in  $G$ .

- (a) Show that  $C$  is a normal subgroup of  $G$ . (Here “normal” is meant in the group-theoretic sense, not in the separation property sense.)
- (b) Show that the quotient group  $G/C$  is the set of all components of  $G$ .

### 5.3 Path-Connected Spaces

sec:pathconnected

To say that a space is “connected” in the sense defined earlier is by no means the only way of capturing the intuitive idea that a space consists of a single piece. Another, equally natural, way is to say that one can travel continuously in the space from any given point to any other. In order to give a precise definition of this latter kind of connectedness, it is necessary first to formalize the notion of traveling continuously from one point to another.

#### Paths

subsec:paths

Let  $x$  and  $y$  be points of a space  $X$ . Starting at some initial time at the point  $x$ , we want to travel continuously in the space until, at some later time, we arrive at the point  $y$ . Let us represent times by real numbers, denoted generically by  $s$ . (We use ‘ $s$ ’ rather than ‘ $t$ ’ for time here because in the next section we shall be using ‘ $t$ ’ to represent time for a different purpose.) Further, represent the initial time by 0, and the arrival time by 1. (Any two real numbers  $s_0$  and  $s_1$  with  $s_0 < s_1$  would do to represent the initial and arrival times, but the choice  $s_0 = 0$  and  $s_1 = 1$  is a simplifying standardization.) At each time  $s$  between the initial and arrival times we are to be at some point  $x_s$  in  $X$ , where  $x_0 = x$  and  $x_1 = y$ . The requirement that we move continuously means that the point  $x_s$  depends continuously on the time  $s$ . Thus our continuous trip from  $x$  to  $y$  may be described by the continuous function  $s \mapsto x_s$  from  $I = [0, 1]$  into  $X$ .



def:path **5.49 Definition.** A **path** in a topological space  $X$  is a continuous map

$$\sigma: I \rightarrow X$$

where, as usual,  $I$  denotes the closed unit interval  $[0, 1]$  in  $\mathbb{R}$ . The points

$$x = \sigma(0), \quad y = \sigma(1)$$

are called the **initial point** and **terminal point** of  $\sigma$ , respectively, and then  $\sigma$  is said to be a path in  $X$  **from  $x$  to  $y$** . The path  $\sigma$  is also said to **start at  $x$**  and **end at  $y$** . The range  $\sigma(I)$  is also referred to as the **trace of  $\sigma$** .

**Usage note.** The term “path” is perhaps unfortunate in that in everyday speech it connotes a physical road or way over which one travels. A more suggestive term would be “trip” or “journey,” which connotes the action of traveling in time over such a physical road. Unfortunately, the term “path” is standard, and so we are stuck with it!

ex:lineas:path **5.50 Examples.** (1) For two points  $x$  and  $y$  in a convex subset  $X$  of  $\mathbb{R}^n$ , the map  $s \mapsto (1-s)x + sy$  is a path in  $X$  from  $x$  to  $y$ .

ex:circular-path (2) For a point  $z = \langle \cos \theta, \sin \theta \rangle$  in  $S_1$ , the map  $s \mapsto \langle \cos s\theta, \sin s\theta \rangle$  is a path in  $S_1$  starting at  $\langle 1 \rangle$  and ending at  $z$ .

ex:re-parametrization (3) If  $\sigma: [0, 1] \rightarrow X$  is a path in a space  $X$  from a point  $x$  to a point  $y$  and if  $\phi: I \rightarrow I$  is any continuous surjection with  $\phi(0) = 0$  and  $\phi(1) = 1$ —for example, if  $\phi$  is a strictly increasing continuous surjection—then  $\sigma \circ \phi$  is another path in  $X$  from  $x$  to  $y$ . We call such a composite  $\sigma \circ \phi$  a **reparametrization of the path  $\sigma$**  (and of the curve that is the trace of  $\sigma$ ).  $\diamond$

New paths can be constructed from given ones in the ways described by the following two lemmas. The first lemma shows how we may travel continuously in a space  $X$  from a point  $x$  to a point  $z$  in unit time by traversing in succession, but with twice the original speed, first a given route from  $x$  to a point  $y$  and then a given route from that  $y$  to  $z$ , passing through the intermediate point  $y$  at time  $1/2$ .

**5.51 Lemma.** Let  $\sigma$  be a path in a topological space  $X$  from a point  $x$  to a point  $y$ , and let  $\tau$  be a path in  $X$  from  $y$  to a point  $z$ . Then the map  $\gamma: I \rightarrow X$  given by

$$\gamma(t) = \begin{cases} \sigma(2t) & \text{if } 0 \leq t \leq \frac{1}{2}, \\ \tau(2t - 1) & \text{if } \frac{1}{2} \leq t \leq 1. \end{cases}$$

lem:path-prod is a path in  $X$  from  $x$  to  $z$ .

def:path-prod **5.52 Definition.** Using the notation of the lemma, we call  $\gamma$  the **(path-) product of  $\sigma$  and  $\tau$**  and denote it by  $\sigma * \tau$ .

Construction of the product of paths is indicated schematically in [Figure 5.10](#).

In the notation  $\sigma * \tau$ , it is the path  $\sigma$  on the left that is traversed first and the path  $\tau$  on the right that is traversed second. This is the opposite of the meaning of function composition

reverse of path

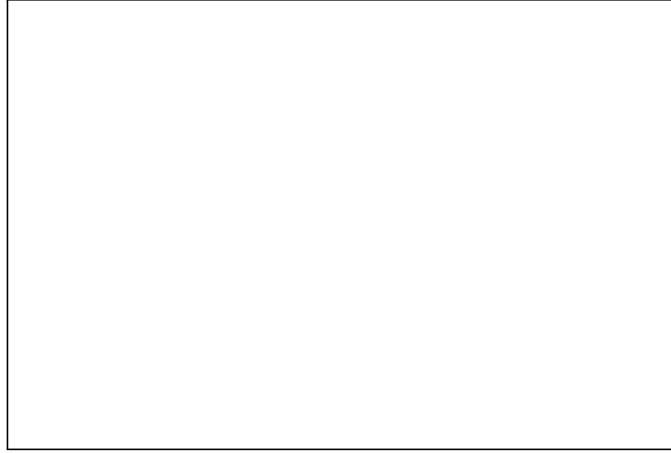


Figure 5.10: The product of two paths.

fig:path-prod

notation  $g \circ f$ , where it is the function  $f$  on the right that is applied first and then the function  $g$  on the left that is applied second.

**Proof of Lemma 5.51.** First,  $\gamma$  is a well-defined map on  $I$  because  $\sigma(2 \cdot 0) = y = \tau(2 \cdot 0 - 1)$  (see Proposition 0.9). Next, the composite of the homeomorphism  $t \mapsto 2t$  of  $[0, 1/2] \rightarrow [0, 1]$  with  $\sigma$  is continuous, as is the composite of the homeomorphism  $t \mapsto 2t - 1$  of  $[1/2, 1] \rightarrow [0, 1]$  with  $\tau$ . By the Gluing Lemma (Theorem 3.13), the map  $\gamma$  is continuous. Finally,  $\gamma(0) = \sigma(0) = x$  and  $\gamma(1) = \tau(1) = z$ , so that  $\gamma$  is indeed a path from  $x$  to  $z$ .  $\square$

The second lemma says that, by reversing direction on a path from a point  $x$  to a point  $y$ , we obtain a path going backwards from  $y$  to  $x$ .

lem:reverse-path

**5.53 Lemma.** Let  $\sigma$  be a path in a topological space  $X$  from a point  $x$  to a point  $y$ . Then the map  $\tau: I \rightarrow X$  given by

$$\tau(t) = \sigma(1 - t)$$

is a path in  $X$  from  $y$  to  $x$ .

def:reverse-path

**5.54 Definition.** Using the notation of the preceding lemma, we call  $\tau$  the **reverse of**  $\sigma$  and denote it by  $\bar{\sigma}$ .

A path  $\sigma$  in a space  $X$  may describe a route that:

- it:path-intersects-itself (i) intersects itself in the sense that  $\sigma(s_1) = \sigma(s_2)$  for times  $s_1 \neq s_2$ ;
- it:path-repeats (ii) traverses a portion of its trace more than once, even doubling back on itself; or
- it:path-returns-to-start (iii) ends at the point where it started in the sense that  $\sigma(1) = \sigma(0)$ .

ex:peculiar-path-behavior **5.55 Example.** Let  $\sigma$  be the path in the plane  $\mathbb{R}^2$  defined by

$$\sigma(s) = e^{\pi i s} = \langle \cos \pi s, \sin \pi s \rangle,$$

which traces out the upper unit semicircle in a counterclockwise direction, starting at  $\langle 1, 0 \rangle$  and ending at  $\langle -1, 0 \rangle$ . Then the path

$$\tau = \sigma * \bar{\sigma}$$

exhibits all three phenomena (i)–(iii) mentioned above. In fact, it traces out that same semicircle, again in the counterclockwise direction, as time increases from 0 to  $1/2$  but then reverses course and retraces the same circle, now in the clockwise direction, as time increases from  $1/2$  to 1.  $\diamond$

reverse of path  
topologist's sine curve

### Path-connectedness

Now we are in a position to define the second kind of “connectedness,” expressing the notion of being able to travel continuously in the space from any point to any other.

def:path-connected **5.56 Definition.** A topological space  $X$  is said to be **path-connected** when, for every  $x \in X$  and  $y \in X$ , there exists a path in  $X$  from  $x$  to  $y$ .

The phrase *pathwise connected* is often seen as a synonym for ‘path-connected.’

ex:convex-sets-path-connected **5.57 Examples.** (1) From [Examples 5.50 \(1\)](#), a convex subspace  $X$  of  $\mathbb{R}^n$  is path-connected. In particular, Euclidean  $n$ -space  $\mathbb{R}^n$ , the  $n$ -ball  $B_n$ , and the  $n$ -disk  $D_n$  are path-connected for all  $n \geq 1$ .

ex:topologist-sine-curve (2) The **topologist's sine curve**, introduced in [Exercise 4.102](#), is the subspace  $G \cup L$  of  $\mathbb{R}^2$ , where

$$G = \left\{ \left\langle x, \sin \frac{1}{x} \right\rangle : 0 < x \leq 1 \right\},$$

$$L = \{0\} \times [-1, 1] = \{ \langle 0, y \rangle : -1 \leq y \leq 1 \}.$$

(See [Figure 5.11](#), in which different scales are used on the vertical and horizontal axes). The graph  $G$  oscillates between the lines  $y = -1$  and  $y = 1$  infinitely often, the oscillation becoming more rapid the closer we get to the  $y$ -axis.

The space  $G$  is connected, being the image of the interval  $]0, 1]$  under the continuous map  $x \mapsto \langle x, \sin 1/x \rangle$ ; here, as always, we take for granted the continuity of the sine function.

We shall show that *the topologist's sine curve  $X$  is connected but not path-connected*. (Local compactness of  $X$  was at issue in [Exercise 4.102](#).)

That  $X$  is connected will follow from [Theorem 5.7](#) if we can show that  $X \subset \text{cls } G$ , the closure being taken in  $\mathbb{R}^2$  (it is easy to show that  $\text{cls } G \subset X$ , too). To show that  $X \subset \text{cls } G$ , we need only show that  $L \subset \text{cls } G$ . Let  $\langle 0, y \rangle \in L$ . If  $0 < \varepsilon < 1$ , there exists some  $t \in [2/\varepsilon, 2\pi + 2/\varepsilon]$  with  $\sin t = y$ , and then the point

$$\left\langle x, \sin \frac{1}{x} \right\rangle = \left\langle \frac{1}{t}, y \right\rangle \in G$$

belongs to the Euclidean ball of radius  $\varepsilon$  at  $\langle 0, y \rangle$ , as indicated in [Figure 5.12](#). Hence  $\langle 0, y \rangle \in \text{cls } G$ .



Figure 5.11: The topologist's sine curve is connected but not locally connected.

fig:sin-curve-with-lim-line

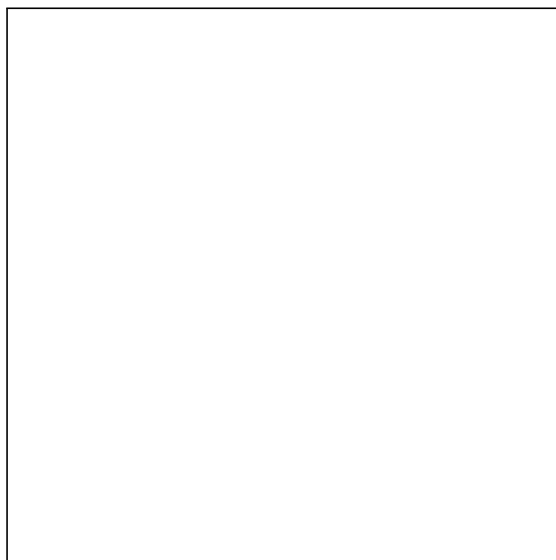


Figure 5.12: The  $\varepsilon$ -ball at a point  $\langle 0, y \rangle$  on the limiting line  $L$  of the topologist's sine curve.

fig:nbd-on-L-for-closed-sine-curve

We are going to show now that  $X$  is *not* path-connected. Just suppose there exists a path  $\sigma$  in  $X$  from the origin  $\langle 0, 0 \rangle$  to the point  $\langle 1/\pi, 0 \rangle$  on the  $x$ -axis. Denote by  $p$  and  $q$  the restrictions to  $X$  of the projections from  $\mathbb{R}^2$  onto  $\mathbb{R}$ ; that is,  $p$  and  $q$  are essentially the projections onto the  $x$ - and  $y$ -axes with each axis identified with the real line  $\mathbb{R}$ . Then the maps

$$p \circ \sigma: I \rightarrow \mathbb{R}, \quad q \circ \sigma: I \rightarrow \mathbb{R}$$

are continuous. Let us assume that  $\sigma(s)$  does not remain in the line segment  $L$  for an entire interval  $[0, t]$  with  $t > 0$ , that is, assume that

$$0 < t < 1 \implies p(\sigma(s)) > 0 \text{ for some } s \in ]0, t].$$

Since  $q \circ \sigma$  is continuous at 0 and  $q(\sigma(0)) = 0$ , there exists some  $b$  with  $0 < b < 1$  such that

$$\{eqn:qsigma-in-interval\} (*) \quad 0 < t < b \implies q(\sigma(t)) \in ]-\frac{1}{2}, \frac{1}{2}[.$$

By our assumption, such a  $b$  exists with

$$p(\sigma(b)) > 0.$$

Since  $p \circ \sigma$  is continuous on  $[0, b]$ , there exists some  $t$  with  $0 < t < b$  and

$$p(\sigma(t)) = \frac{2}{n\pi}$$

for some positive integer  $n$ . Then

$$|q(\sigma(t))| = \left| \sin\left(\frac{1}{2/(n\pi)}\right) \right| = 1,$$

which contradicts (\*).

If our assumption concerning  $\sigma$  does not hold, however, now define

$$a = \sup\{t \in [0, 1] : p(\sigma(s)) = 0 \text{ for all } s \in ]0, t]\}.$$

By continuity of  $p \circ \sigma$  at  $a$ , we have  $p(\sigma(a)) = 0$ ; since  $p(\sigma(1)) = 1/\pi > 0$ , then  $a < 1$ . Now argue as in the preceding paragraph, but using the continuity of  $p \circ \sigma$  at  $a$  instead of at 0.  $\diamond$

Path-connectedness is a topological property. This is a consequence of the following, more general, fact.

**5.58 Theorem.** *The continuous image of a path-connected space is path-connected.*

**Proof.** Let  $f: X \rightarrow Y$  be a continuous surjection with path-connected domain  $X$ . Let  $y_0, y_1 \in Y$ . Choose any  $x_0$  and  $x_1$  in  $X$  with  $y_0 = f(x_0)$  and  $y_1 = f(x_1)$ , respectively. There is some path  $\sigma$  in  $X$  from  $x_0$  to  $x_1$ . Then  $f \circ \sigma$  is a path in  $Y$  from  $y_0$  to  $y_1$ .  $\square$

From Examples 5.57 (1) and the preceding Theorem 5.58 it follows, for example, that the circle  $S_1$ , the Möbius strip [Examples 3.81 (5)], and the Klein bottle [Examples 3.81 (6)] are path-connected.

An immediate corollary of Theorem 5.58 is that **being path-connected is a topological property**. Accordingly, we speak of the property of **path-connectedness**.

**5.59 Theorem.** *Every path-connected space is connected.*

**Proof.** Let  $X$  be a path-connected space. If  $x, y \in X$ , then there is a path  $\sigma$  in  $X$  from  $x$  to  $y$ , and then the trace  $\sigma(I)$  of this path is a connected subset of  $X$  containing both  $x$  and  $y$ . By Corollary 5.20, the space  $X$  is connected.  $\square$

topologist's sine curve  
Möbius strip  
Klein bottle

thm:cont-map-preserves-path-conn

thm:conn-implies-path-conn

The preceding proof does *not* say that the path-connected space  $X$  itself is a continuous image of the unit interval  $I$ ; in fact, a noncompact path-connected space, such as  $\mathbb{R}$ , could not be a continuous image of the compact space  $I$ .

As shown by Examples 5.57 (2), a connected space need not be path-connected. Nonetheless, there is a large class of topological spaces, including manifolds, for which connectedness is equivalent to path-connectedness.

**5.60 Theorem.** *A space is path-connected if it is connected and each of its points has at least one path-connected neighborhood.*

**Proof.** Let  $X$  be a connected space in which each point has at least one path-connected neighborhood. Fix a point  $x \in X$ . Define

$$K = \{y \in X : \text{there exists a path in } X \text{ from } x \text{ to } y\}.$$

It suffices to show that  $K = X$ , for then if  $y, z \in X$ , there will exist paths in  $X$  from  $x$  to  $y$  and from  $x$  to  $z$ , and hence by Lemma 5.51 a path in  $X$  from  $y$  to  $z$ . Since the constant map from  $I$  to  $X$  with value  $x$  is a path in  $X$  from  $x$  to  $x$ , then  $x \in K$ . Hence  $K \neq \emptyset$ . Since  $X$  is connected, to show that  $K = X$  it will therefore suffice to show that  $K$  is both open and closed in  $X$ .

We show that  $K$  is open in  $X$ . Let  $y \in K$ . By hypothesis there exists some path-connected neighborhood  $U$  of  $y$  in  $X$ . We need only show that  $U \subset K$ . Let  $z \in U$  be arbitrary. There is a path in  $U$ —and therefore in  $X$ —from  $y$  to  $z$ ; since  $y \in K$ , there is a path in  $X$  from  $x$  to  $y$ . By Lemma 5.51, there is a path in  $X$  from  $x$  to  $z$ , and so  $z \in K$ . Hence  $U \subset K$ .

We show that  $K$  is closed in  $X$ . Let  $y \in X \setminus K$ . There exists a path-connected neighborhood  $V$  of  $y$  in  $X$ . We need only show that  $V$  is disjoint from  $K$ . Just suppose that there is some  $z \in V \cap K$ . Since  $z \in K$ , there is a path from  $x$  to  $z$ ; since  $z \in V$ , there is a path from  $z$  to  $y$ . By Lemma 5.51 again, there is a path from  $x$  to  $y$ , and so  $y \in K$ . This is a contradiction.  $\square$

**5.61 Corollary.** *A necessary and sufficient condition for a topological manifold to be path-connected is that it be connected.*

**Proof.** Let  $X$  be an  $n$ -manifold. If  $x \in X$ , then there is a neighborhood  $U$  of  $x$  in  $X$  such that  $U \cong \mathbb{R}^n$ , and  $U$  is path-connected since  $\mathbb{R}^n$  is. Sufficiency now follows from Theorem 5.60. Necessity follows from Theorem 5.59.  $\square$

**5.62 Example.** For  $n \geq 1$ , the  $n$ -sphere  $S_n$  is path-connected. This is an immediate application of Corollary 5.61, since by Examples 5.21 (3) and Examples 3.25 (14), such an  $n$ -sphere is a connected manifold.

Of course, we could prove that  $S_n$  is path-connected the hard way, by explicitly constructing a path between an arbitrary pair of points on it (compare Exercise 65).  $\diamond$

## Curves

Earlier we distinguished between a path, as a map, and its trace, as a subset of a topological space. The trace of a path is a special kind of curve. A **curve**, in the most general sense of the term, is a continuous image  $C$  of a nondegenerate interval  $J$  in  $\mathbb{R}$ , and

then any continuous *surjection*  $J \rightarrow C$  is called a **parametrization of  $C$**  and is said to **parametrize  $C$** .

For example, the subset  $C$  of the plane with equation  $y = x^2$  is a curve because it is the image of the map  $x \mapsto x^2$  of the interval  $\mathbb{R} = ]-\infty, \infty[$  to  $\mathbb{R}^2$ .

When the domain of a parametrization  $\varphi$  of a curve  $C$  is a closed, bounded interval  $[a, b]$ , then the curve will be the trace of a path in  $C$  (and hence in any topological space  $X$  having  $C$  as a subspace), namely, the trace of the path  $\varphi \circ \alpha$  where  $\varphi: I \rightarrow [a, b]$  is the linear map  $s \mapsto a + (b - a)s$ .

parametrization  
figure-eight  
arc-connected space  
dimension!paramters@and paramet  
Peano, Giuseppe  
plane-filling curve

ex:figure-eight-path-curve **5.63 Example.** The figure-eight, shown in [Figure 3.11](#)—the graph of the equation

$$x^2 = 4y^2(1 - y^2)$$

—is parametrized by  $\varphi: [0, 2\pi] \rightarrow \mathbb{R}^2$  given by  $\varphi(t) = (x(t), y(t))$  where

$$\begin{cases} x(t) = \sin 2t, \\ y(t) = \sin t. \end{cases}$$

The corresponding path  $\sigma$  whose trace is the same as this curve is the map  $I \rightarrow \mathbb{R}^2$  given by  $s \mapsto \varphi(2\pi s)$ .

Notice that this curve, and the associated path whose trace it is, not only ends where it starts but also “crosses itself.”  $\diamond$

The phenomena of a path traversing some or all of its trace again, or returning at the end to its start, are avoided when  $\sigma$  is an embedding, for then  $\sigma$  is injective (since  $I$  is compact, a continuous injection from  $I$  to  $X$  is automatically an embedding whenever  $X$  is a Hausdorff space).

def:arc-conn **5.64 Definition.** Let  $X$  be a topological space. For points  $x$  and  $y$  of a topological space  $X$ , a path  $\sigma$  from  $x$  to  $y$  that is actually an embedding of  $I$  into  $X$  is called an **arc in  $X$  from  $x$  to  $y$** .

When for any two distinct points  $x$  and  $y$  of  $X$  there is an arc in  $X$  from  $x$  to  $y$ , the space  $X$  is said to be **arc-connected**.

The phrase *arcwise connected* is often seen as a synonym for “arc-connected.”

It can be proved that a path-connected Hausdorff space is arc-connected; for the proof, which uses R. L. Moore’s characterization of  $I$  mentioned in [Exercise 38](#), see Hocking and Young [[36](#), Theorem 2-27] or Willard [[74](#), Corollary 31.6].

According to [Theorem 5.58](#), a continuous image of the closed unit interval  $I$  is necessarily path-connected. But which spaces are continuous images of  $I$ ? With the exception of a one-point space, such a space would seem to be “one-dimensional” in the sense that it takes just a single continuously varying parameter  $t$ , for  $0 \leq t \leq 1$ , to describe its points. And it seems intuitively obvious—does it not?—that the square  $I \times I$  is “two-dimensional” in the sense that it requires two continuously varying parameters  $0 \leq t \leq 1$  and  $0 \leq s \leq 1$  to describe its points. This intuition is wrong! In 1890, G. Peano published the following surprising example of a **plane-filling curve**—a continuous map from the interval  $I$  onto the square  $I \times I$ . ([Application 6.41](#) will present a different, but non-constructive, way of obtaining a plane-filling curve.)

ex:Peano-plane-filling-curve **5.65 Example (Peano’s plane-filling curve).** For each sequence  $\langle i_1, i_2, i_3, \dots \rangle$  in the set  $\{0, 1, 2, \dots, 8\}$  we are going to obtain a certain point  $z(i_1, i_2, i_3, \dots)$  in the unit square  $I \times I$  and a certain number  $t(i_1, i_2, i_3, \dots)$  in the unit interval  $I$ , and then map this number to that point.

To do that, we shall: describe a basic method for subdividing a square; apply the basic method to successively subdivide the unit square in order to obtain the points  $z(i_1, i_2, i_3, \dots)$ ; use a similar method to successively subdivide the unit interval in order to obtain the numbers  $t(i_1, i_2, i_3, \dots)$ ; define the map; and check the map's properties.

Basic method for subdividing a square. Let  $S$  be any square in the plan having its sides parallel to the axes and in which two diagonally opposite vertices  $\alpha S$  and  $\omega S$  have been designated. Divide  $S$  into 9 congruent squares by subdividing each edge of  $S$  into 3 subintervals and then drawing horizontal and vertical lines connecting pairs of opposite ends of those subintervals. There is a unique way of numbering these squares  $S_0, S_1, \dots, S_8$  and of designated two opposite vertices  $\alpha S_i$  and  $\omega S_i$  of each  $S_i$  in such a way that  $S_0$  and  $S_1$  share a common vertical side and that

$$\{\text{eqn:8-vertex-pairs}\} \quad (*) \quad \alpha S_0 = \alpha S, \quad \omega S_i = \alpha S_{i+1} \quad (i = 0, 1, \dots, 7), \quad \omega S_8 = \omega S.$$

Since there are four possible choices of the pair  $\langle \alpha S, \omega S \rangle$ , there are four possible configurations for  $S_0, S_1, \dots, S_8$ . These configurations are depicted in [Figure 5.13](#), where we have indicated  $\alpha S_i$  and  $\omega S_i$  by drawing an arrow from  $\alpha S_i$  to  $\omega S_i$ . Observe that

$$\{\text{eqn:diam-square-Si}\} \quad (**) \quad \text{diam}(S_i) = \frac{1}{3} \text{diam}(S) \quad (i = 0, 1, \dots, 8).$$

Successively dividing the unit square. Starting with the square  $B = I \times I$  with designated vertices  $\alpha B = \langle 0, 0 \rangle$  and  $\omega B = \langle 1, 1 \rangle$ , apply the preceding basic method so as to obtain 9 squares

$$B(0) = B_0, B(1) = B_1, \dots, B(8) = B_8$$

with each  $B(i)$  having two designated opposite vertices  $\alpha B(i)$  and  $\omega B(i)$ —see [Figure 5.14](#).

Next, apply the basic method to each of those squares  $B(i)$  so as to obtain squares

$$B(i, 0) = B(i)_0, B(i, 1) = B(i)_1, \dots, B(i, 8) = B(i)_8 \quad (i = 0, 1, \dots, 8),$$

yielding  $9^2 = 81$  smaller squares  $B(i, j)$ —see [Figure 5.15](#). Continue this process repeatedly, so that once  $9^n$  squares  $B(i_1, i_2, \dots, i_n)$  have already been constructed for all choices of  $i_1, i_2, \dots, i_n \in \{0, 1, 2, \dots, 8\}$ , apply the basic method to each of these squares so as to obtain squares,

$$\begin{aligned} B(i_1, i_2, \dots, i_n, 0) &= B(i_1, i_2, \dots, i_n)_0, \\ B(i_1, i_2, \dots, i_n, 1) &= B(i_1, i_2, \dots, i_n)_1, \\ &\vdots \\ B(i_1, i_2, \dots, i_n, 8) &= B(i_1, i_2, \dots, i_n)_8. \end{aligned}$$

Each of the  $9^n \cdot 9 = 9^{n+1}$  squares so obtained has, of course, two designated opposite vertices.

For each  $n \geq 1$  and each  $n$ -tuple  $\langle i_1, i_2, \dots, i_n \rangle$  in  $\{0, 1, 2, \dots, 8\}$  we have a square  $B(i_1, i_2, \dots, i_n)$ . By (\*),

$$\begin{aligned} \alpha B(i_1, i_2, \dots, i_n, 0) &= \alpha B(i_1, i_2, \dots, i_n), \\ \alpha B(i_1, i_2, \dots, i_n, i) &= \alpha B(i_1, i_2, \dots, i_n, i+1) \quad (i = 0, 1, 2, \dots, 7), \\ \omega B(i_1, i_2, \dots, i_n, 8) &= \omega B(i_1, i_2, \dots, i_n). \end{aligned}$$

From (\*\*),

$$\text{diam } B(i_1, i_2, \dots, i_n) = \frac{\sqrt{2}}{3^n}.$$

Now

$$B(i_1, i_2, \dots, i_n, i) \subset B(i_1, i_2, \dots, i_n) \quad (i = 0, 1, 2, \dots, 7).$$



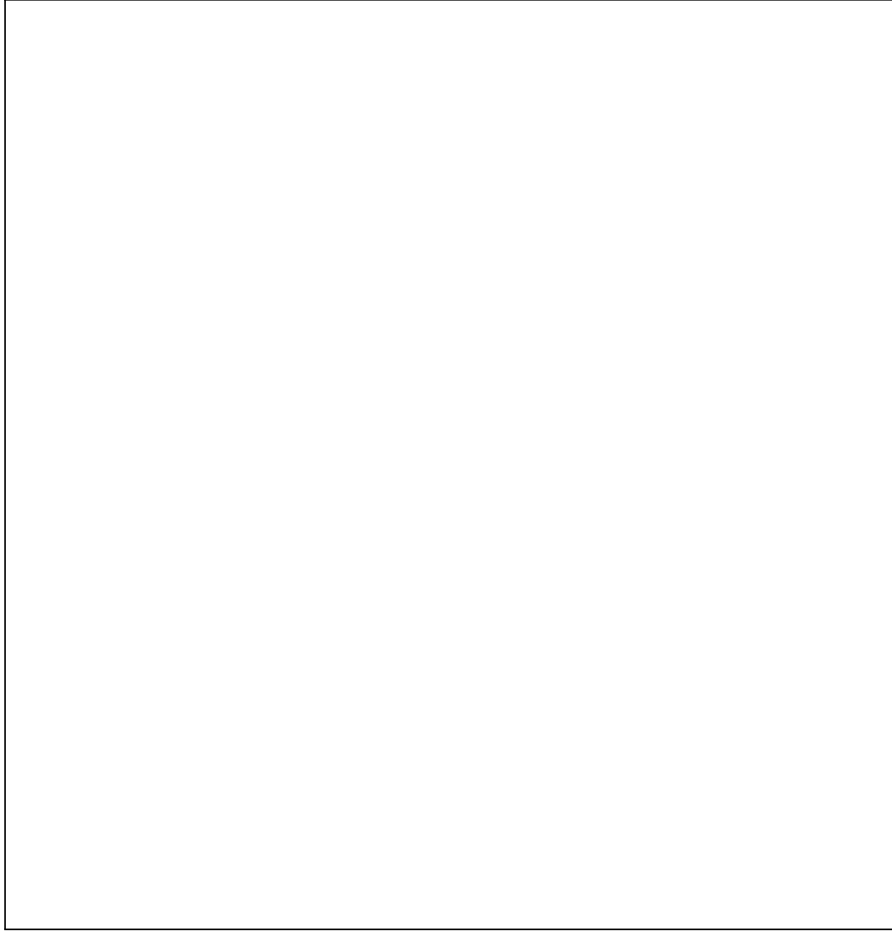


Figure 5.13: Configurations of opposite corners of the 8 subsquares of a given square.

fig:configs-8-subsquares

From the Nested Set Theorem (1.88) it follows that for each infinite sequence  $\langle i_1, i_2, i_3, \dots \rangle$  in  $\{0, 1, 2, \dots, 8\}$ , there is a *unique* point

$$z(i_1, i_2, i_3, \dots) \in \bigcap_{n=1}^{\infty} B(i_1, i_2, \dots, i_n).$$

**Successively subdividing the unit interval.** Divide the interval  $I$  into 9 closed subintervals  $A(0), A(1), \dots, A(8)$  each of length  $9^{-1}$ , numbered from left to right. Next, for each  $i \in \{0, 1, 2, \dots, 8\}$ , divide  $A(i)$  into 9 intervals  $A(i, 0), A(i, 1), \dots, A(i, 8)$  each of length  $9^{-2}$ , numbered from left to right. Continuing this process, for each  $n \geq 1$  we obtain  $9^n$  intervals  $A(i_1, i_2, \dots, i_n)$  each of length  $9^{-n}$ . For each infinite sequence  $\langle i_1, i_2, i_3, \dots \rangle$  in the set  $\{0, 1, 2, \dots, 8\}$  there is a *unique* point

$$t(i_1, i_2, i_3, \dots) \in \bigcap_{n=1}^{\infty} A(i_1, i_2, i_3, \dots).$$

plane-filling curve  
space-filling curve

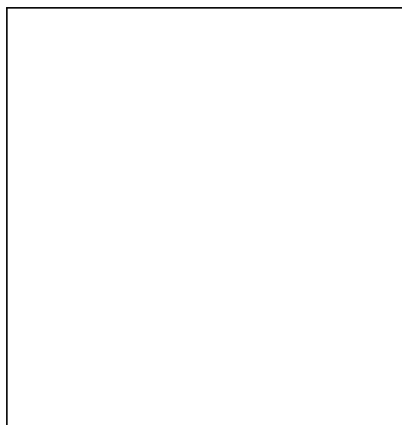


Figure 5.14: The first 9 squares and their designated opposite corners.

fig:construct-first-9-squares

Constructing the map from the interval to the square. Define a map  $\sigma: I \rightarrow I \times I$  as follows. Let  $t \in I$ . There is a sequence  $\langle i_1, i_2, i_3, \dots \rangle$  in  $\{0, 1, 2, \dots, 8\}$  with

$$t = t(i_1, i_2, i_3, \dots),$$

and we set

$$\sigma(t) = z(i_1, i_2, i_3, \dots).$$

The map  $\sigma$  is well-defined, because if also  $t = t(j_1, j_2, j_3, \dots)$  for another sequence  $\langle j_1, j_2, j_3, \dots \rangle$ , then  $z(i_1, i_2, i_3, \dots) = z(j_1, j_2, j_3, \dots)$ —see [Exercise 79](#).

**Properties of the map.** We leave to the reader to verify that each point  $z \in I \times I$  is  $z(i_1, i_2, i_3, \dots)$  for some  $\langle i_1, i_2, i_3, \dots \rangle$ , in other words, that  $\sigma$  is surjective.

We show that  $\sigma$  is continuous. Let  $\varepsilon > 0$  be arbitrary. Choose  $n$  so large that

$$\frac{2\sqrt{2}}{3^n} < \varepsilon.$$

Suppose  $t, s \in I$  with  $|t - s| < 9^{-n}$ . If  $t, s \in A(i_1, i_2, \dots, i_n)$  for some  $\langle i_1, i_2, \dots, i_n \rangle$ , then  $\sigma(t), \sigma(s) \in B(i_1, i_2, \dots, i_n)$  and so

$$\|\sigma(t) - \sigma(s)\| \leq \frac{\sqrt{2}}{3^n} < \varepsilon.$$

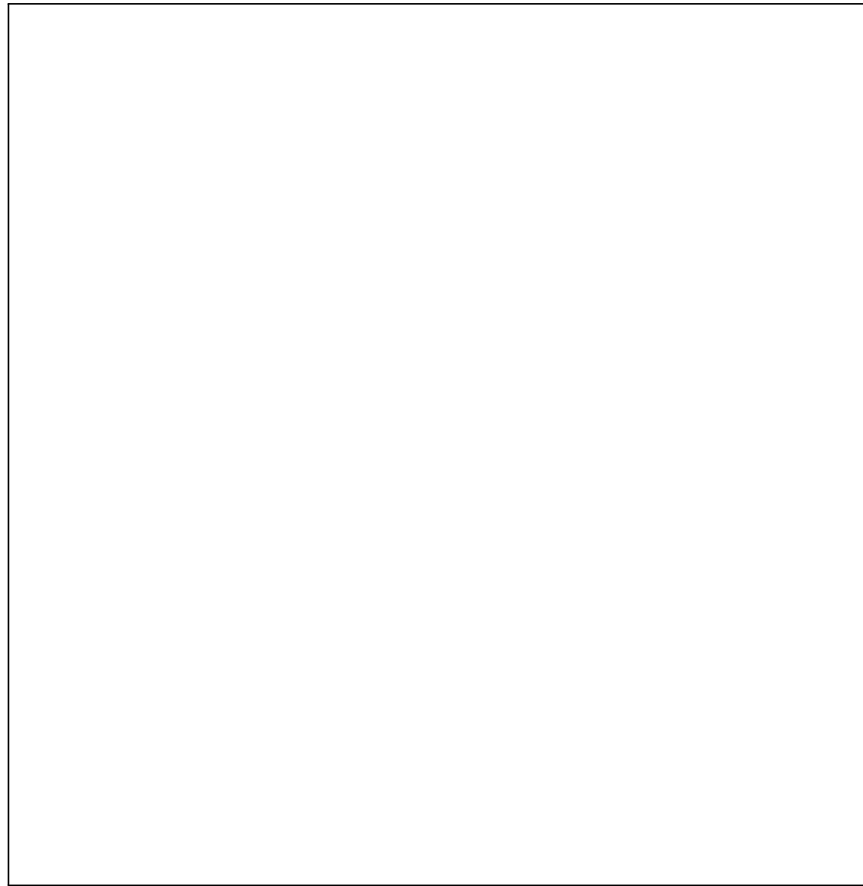
If not, then  $t \in A(i_1, i_2, \dots, i_n)$  and  $s \in A(i_1, i_2, \dots, i_n + 1)$  for some  $\langle i_1, i_2, \dots, i_n \rangle$ , and so  $\sigma(t) \in B(i_1, i_2, \dots, i_n)$  and  $\sigma(s) \in B(i_1, i_2, \dots, i_n + 1)$ ; then

$$\|\sigma(t) - \sigma(s)\| \leq \frac{2\sqrt{2}}{3^n} < \varepsilon. \quad \diamond$$

A different plane-filling curve is described in [Application 6.41](#).

Similarly, there are **space-filling curves**—continuous maps from the unit interval  $I$  onto the unit cube  $I^3$ : see [Exercise 81](#).

Peano's work on plane-filling curves and space-filling curves did not settle the question of exactly which spaces are continuous images of  $I$ . A necessary condition for a nonempty metrizable space to be a continuous image of  $I$  is that it be compact, connected, and locally connected (see [Theorem 4.21](#), [Theorem 5.9](#), and [Corollary 5.46](#)). In 1914–1920, H. Hahn



Hahn, Hans  
Mazurkiewicz, Stefan

Figure 5.15: The first 81 squares and their designated corners.

fig:first-81-squares

and S. Mazurkiewicz proved that this condition is also sufficient. Now a Hausdorff space that is the continuous image of a compact metrizable space is itself metrizable (see [Exercise 4.97](#) and the paragraph following [Theorem 4.44](#)). Hence we may state the characterization of continuous images of  $I$  as follows.

thm:Hahn-Mazurkiewicz

**5.66 Hahn–Mazurkiewicz Theorem.** *A nonempty Hausdorff space is a continuous image of the closed unit interval if and only if it is:*

- *compact,*
- *connected,*
- *locally connected, and*
- *metrizable.*

For a proof, see Willard [74, 34, section 31]; the idea of the proof is first to obtain a continuous map from the Cantor set  $K$  onto the given compact, connected, locally connected, metrizable space and then extend that map to a continuous map on  $I$ .

starlike set

topologist's sine curve  
path and  $S^1$   
polygonal path

polygonally connected set  
 exhibit-path-on-S2  
 starlike set

## path-component

prohibited paths between S1 and S2

prob-part:exhibit-path-on-S2

prob:topologist-sine-etc-path-connQ

rob-part-finite-conn-then-path-conn

prob:polygonal-conn

part:polygonal-path-range-of-path

n-open-in-Rn-implies-polygon-conn

prob:path-component

65. (a) Given an arbitrary pair of points on the circle  $S_1$ , exhibit a path from the first to the second.  
 (b) Given an arbitrary pair of points on the 2-sphere  $S_2$ , exhibit a path from the first to the second.
66. Write an explicit formula for the function  $\sigma \circ \overleftarrow{\sigma} : I \rightarrow \mathbb{R}^2$  of Example 5.55. (The formula will have to be given piecewise.)
67. Must a subset of  $\mathbb{R}^n$  be path-connected if it is starlike (Exercise 19)?
68. Which, if any, of the following subspaces of the topologist's sine curve [Examples 5.57 (2)] are path-connected?  
 (a) The subspace  $G = \{ \langle x, \sin(1/x) \rangle : 0 < x \leq 1 \}$ ?  
 (b) The subspace  $G \cup \{ \langle 0, 0 \rangle \}$ ?
69. Construct a path-connected subset of the plane containing two points  $a$  and  $b$  having the property that no path in the subset from  $a$  to  $b$  has finite length.
70. Analyze the path-connectedness of product spaces:
71. (a) Prove that a connected space is path-connected if it is also finite.  
 (b) Does (a) generalize if “finite” is replaced by “compact”?
72. Let  $X \subset \mathbb{R}^n$ . For points  $x, y \in X$ , a **polygonal path in  $X$  from  $x$  to  $y$**  is the union of finitely many line segments  $L(x_0, x_1), L(x_1, x_2), \dots, L(x_{n-1}, x_n)$  contained in  $X$ , where  $L(x_i, x_{i+1})$  denotes a line segment joining  $x_i$  to  $x_{i+1}$ , such that  $x_0 = x$  and  $x_n = y$ .  
 Call  $X$  **polygonally connected** if, for each two points  $x, y$  of  $X$ , there exists a polygonal path in  $X$  from  $x$  to  $y$ . For example, any convex subset of  $\mathbb{R}^n$  is polygonally connected.  
 (a) Show that a polygonal path in  $X$  from  $x$  to  $y$  is the trace of some path in  $X$  from  $x$  to  $y$  (in the sense of Definition 5.49).  
 (b) Must a subset of  $\mathbb{R}^n$  be polygonally connected if it is starlike (Exercise 19)?  
 (c) From (a) and Theorem 5.59 it follows that a subspace  $X$  of  $\mathbb{R}^n$  is connected if it is polygonally connected. Prove that the converse holds in case  $X$  is open in  $\mathbb{R}^n$  but otherwise need not hold.
73. Let  $X$  be a nonempty open connected subset of the plane  $\mathbb{R}^2$ . Prove that if  $x, y$ , and  $z$  are any three distinct points of  $X$ , then there is a path  $\sigma$  in  $X$  from  $x$  to  $y$  with  $z \notin \sigma(I)$ . Use this to show that  $X$  is *not* homeomorphic to any subset of the real line  $\mathbb{R}$ .
74. If  $B$  is a subset of a compact, path-connected, metrizable space  $X$  and if  $A \subset B$ , must  $d(x, \text{bdy } A) \leq d(x, \text{bdy } B)$  for all  $x \in A$ ?
75. Let  $X$  be a nonempty space. A **path-component of  $X$**  is a nonempty path-connected subset of  $X$  that is not properly contained in any other path-connected subset of  $X$ .  
 What are the path-components of each of the spaces in Exercise 68?
76. (Continuation of Exercise 75.)  
 (a) Prove that two points of  $X$  belong to the same path-component of  $X$  if and only if there exists a path in  $X$  from one of the points to the other.

- (b) Formulate and prove the analog of [Theorem 5.32](#) for path-components.
- (c) Is the analog of [Theorem 5.31](#) for path-components true?
- prob:locally-path-conn **77.** A space is said to be **locally path-connected** if each neighborhood of each point  $x$  of the space contains some path-connected neighborhood of  $x$ .
- (a) Prove that a space  $X$  is locally path-connected if for each neighborhood  $U$  of each point  $x \in X$ , there is a neighborhood  $V$  of  $x$  such that  $V \subset U$  and such that, for each  $y \in V$ , there is a path in  $V$  from  $x$  to  $y$ .
- (b) What can be said about the path-components ([Exercise 75](#)) of a locally path-connected space?
- (c) Find a path-connected space that is not locally path-connected.
- 78.** Give an example of a path-connected space that is not arc-connected ([Definition 5.64](#)). (By the discussion following [Definition 5.64](#), such a space cannot be a Hausdorff space.)
- prob:detail-re-Peano-curve **79.** Let  $t \in I$ . In the notation of [Example 5.65](#), how many sequences  $\langle i_1, i_2, i_3, \dots \rangle$  can there be such that  $t = t(i_1, i_2, i_3, \dots)$ ? For which  $t$  is there a unique such sequence? (*Hint:* Use base 9 expansions.)
- prob:Peano-map-and-inj **80.** (a) Show that Peano's map  $\sigma: I \rightarrow I \times I$  ([Example 5.65](#)) is not injective by exhibiting  $t, s \in I$  with  $t \neq s$  but  $\sigma(t) = \sigma(s)$ .
- (b) Can any continuous map from  $I$  onto  $I \times I$  be injective?
- (c) Can any map from  $I$  onto  $I \times I$  be injective?
- prob:space-filling-curve **81.** Construct a *space-filling curve*—a continuous map from  $I$  onto  $I^3$ . You may want to use the method of [Example 5.65](#) or else look ahead and use the method of [Application 6.41](#).

locally path-connected space  
space-filling curve  
curve

## 5.4 Homotopy

sec:homotopy

The disk

$$D_2 = \{x \in \mathbb{R}^2 : \|x\| \leq 1\}$$

and the annulus

$$A = \{x \in \mathbb{R}^2 : 1/2 \leq \|x\| \leq 1\}$$

do not appear to be homeomorphic (see [Figure 5.16](#)). Yet none of the topological properties we have already studied distinguish between these two spaces: both are path-connected, locally connected, compact, separable, and metrizable—and hence also connected, locally compact, second-countable, first-countable, and  $T_2$ .

It is intuitively plausible that the disk  $D_2$  can be continuously shrunk to a point whereas the annulus cannot, that  $D_2$  cannot be continuously deformed into a circle whereas  $A$  can be, and that  $D_2$  does not have a “hole” whereas  $A$  does. Elaborating these ideas forms part of *homotopy theory*, which deal with the question of whether one space can be continuously deformed into another space and whether one map can be continuously transformed into another map.

convex set!  
punctured sphere

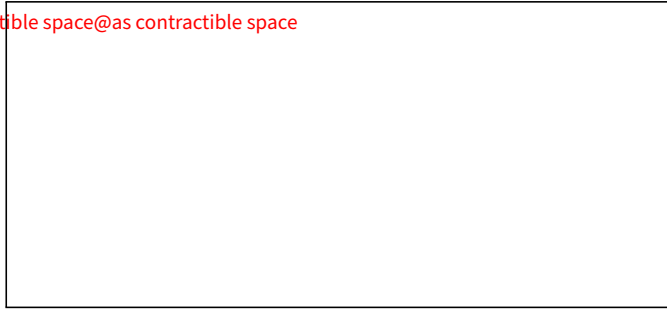


Figure 5.16: The 2-disk  $D_2$  and the annulus  $A$  do not seem to be homeomorphic.

fig:disk-and-annulus

### Contractible spaces

subsec:contractible

What should it mean to say that a space  $X$  can be continuously shrunk to a point  $y \in X$ ? We cannot simply mean that  $X$  can be mapped continuously onto its one-point subspace  $\{y\}$ , for that is always possible. It should mean, instead, that a continuous process can be performed over a time interval during which each point  $x \in X$  moves toward  $y$  and ends up there. To make this more precise, as usual we parametrize time by real numbers  $0 \leq t \leq 1$ , with 0 being the start time and 1 the end time. If we denote by  $h(x, t)$  the position of point  $x \in X$  at time  $t \in [0, 1]$ , then the continuous process is to satisfy the conditions  $h(x, 0) = x$  and  $h(x, 1) = y$ . Thus shrinking  $X$  continuous to a point may be defined as follows.

def:contraction **5.67 Definition.** A **contraction** of a topological space  $X$  to a point  $y \in X$  is a continuous map

$$h: X \times I \rightarrow X$$

such that

$$h(x, 0) = x, \quad h(x, 1) = y \quad (x \in X).$$

If a contraction of  $X$  to a point  $y \in X$  exists, then we say that  $X$  is **contractible to**  $y$ . If some contraction of  $X$  to some point of  $X$  exists, then  $X$  is said to be **contractible**.

Note that the present meaning of ‘contraction’ is quite different from its meaning as a kind of map of a metric space in the sense of [Exercise 1.123](#).

ex:convex-set:contractible **5.68 Examples.** (1) A convex subset  $X$  of  $\mathbb{R}^n$  for  $n \geq 1$  is contractible, for if we choose some  $y \in X$ , then the map

$$h: X \times I \rightarrow X \\ \langle x, t \rangle \mapsto (1 - t)x + ty$$

is a contraction of  $X$  to  $y$ . Observe that  $h$  moves each  $x \in X$  to  $y$  along the line segment joining  $x$  to  $y$ .

In particular, all the spaces  $\mathbb{R}^n$ ,  $D_n$ , and  $B_n$  are, for  $n \geq 1$ , are contractible.

(2) The punctured  $n$ -sphere  $X$ —the subspace  $X = S_n \setminus \{\mathbf{p}\}$  of the  $n$ -sphere  $S_n$ , where  $\mathbf{p}$  is the north pole—is not convex (unless  $n = 0$ ). However,  $X$  is contractible. To see this, use a homeomorphism

$$f: X \cong \mathbb{R}^n$$

[see [Examples 3.25 \(13\)](#)] together with a contraction

$$H: \mathbb{R}^n \times I \rightarrow \mathbb{R}^n$$

of  $\mathbb{R}^n$  to the origin  $\mathbf{0}$  to form the map

$$h: X \times I \rightarrow X$$

given by

$$h(x, t) = f^{-1}(H(f(x), t)).$$

Then  $h$  is a contraction of  $X$  to the point  $f^{-1}(\mathbf{0})$  of  $X$ .

It is a fact that the entire  $n$ -sphere  $S_n$  is *not* contractible. The case  $n = 1$  will be proved later. The case  $n = 0$  follows from the next proposition.  $\diamond$

contractible space  
contraction  
gluing maps  
Gluing Lemma

prop:contractible-then-path-conn

**5.69 Proposition.** *A contractible space is path-connected.*

**Proof.** Let  $h: X \times I \rightarrow X$  be a contraction of the space  $X$  to its point  $y$ . In view of , to show that  $X$  is path-connected it suffices to show that for each  $x \in X$  there is a path in  $X$  from  $x$  to  $y$ . And, in fact, for arbitrary  $x \in X$ , the continuous map

$$\begin{aligned} \sigma: I &\rightarrow X \\ t &\mapsto h(x, t) \end{aligned}$$

is a path in  $X$  from  $x = h(x, 0)$  to  $y = h(x, 1)$ .  $\square$

The choice of the point to which a contractible space is shrunk is immaterial.

prop:contractible-pt-immaterial

**5.70 Proposition.** *A contractible space is contractible to each of its points.*

**Proof.** Let  $h: X \times I \rightarrow X$  be a contraction of  $X$  to some point  $y \in X$  and let  $z$  be an arbitrary point of  $X$ . To construct a contraction of  $X$  to  $z$ , first move each  $x \in X$  to  $y$  along the path  $t \mapsto h(x, t)$  at twice the original speed, so that  $x$  arrives at  $y$  at time  $t = 1/2$ ; then move  $y$  to  $z$  backward along the path  $t \mapsto h(z, t)$ , again at twice the original speed. In other words, define  $H: X \times I \rightarrow X$  by

$$H(x, t) = \begin{cases} h(x, 2t) & \text{if } 0 \leq t \leq 1/2, \\ h(z, 2 - 2t) & \text{if } 1/2 \leq t \leq 1. \end{cases}$$

The map  $H$  is well-defined, for on

$$X \times \{1/2\} = (X \times [0, 1/2]) \cap (X \times [1/2, 1])$$

we have

$$h(x, 2 \cdot 1/2) = y = h(z, 2 - 2 \cdot 1/2).$$

By the Gluing Lemma ([Theorem 3.13](#)),  $H$  is continuous, as its restrictions to the closed subspaces  $X \times [0, 1/2]$  and  $X \times [1/2, 1]$  of  $X \times I$  are continuous, each being a composition of continuous maps. Finally,

$$H(x, 0) = h(x, 0) = x, \quad H(x, 1) = h(z, 0) = z$$

for each  $x \in X$ .  $\square$

Explicit use of the Gluing Lemma ([Theorem 3.13](#)), such as was made in the preceding proof, will usually be omitted below.

annulus

punctured disk

disk!punctured

**Retracts and deformation retracts**

A space that is not contractible can sometimes still be continuously deformed onto a proper subspace without the points of that subspace being disturbed.

def: def-retraction

**5.71 Definition.** Let  $E$  be a subspace of a topological space  $X$ . A **deformation retraction of  $X$  onto  $E$**  is a continuous map

$$h: X \times I \rightarrow X$$

such that

$$\begin{aligned} h(x, 0) &= x, & h(x, 1) &\in E & (x \in X), \\ h(y, 1) &= y & & & (y \in E). \end{aligned}$$

When some deformation retraction of  $X$  onto  $E$  exists, then  $E$  is called a **deformation retract of  $X$** .

exs: def-retracts

**5.72 Examples.** (1) The circle

$$S_1 = \{x \in \mathbb{R}^2 : \|x\| = 1\}$$

is a deformation retract of the annulus

$$A = \{x \in \mathbb{R}^2 : 1/2 \leq \|x\| \leq 1\},$$

for we may slide each  $x \in A$  radially outward toward the point  $\|x\|^{-1}x \in S_1$  (see [Figure 5.17](#)). Formally, define the deformation retraction  $h: A \times I \rightarrow A$  of  $A$  onto  $S_1$

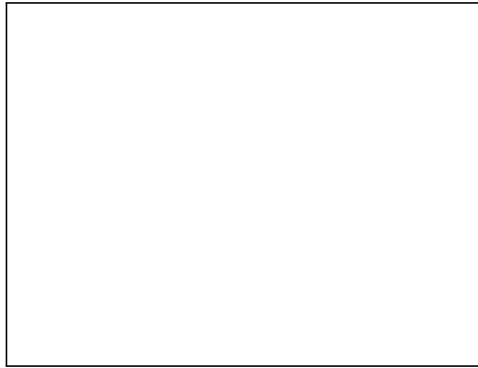


Figure 5.17: A deformation retraction of the annulus onto the circle.

fig: deformation-retraction-annulus-c

by the formula

$$h(x, t) = (1 - t)x + t \cdot \frac{1}{\|x\|} x.$$

The same formula defines a deformation retraction of the punctured disk  $D_2 \setminus \{0\}$  onto the circle  $S_1$ .

(2) The map

$$h: \mathbb{R}^n \times I \rightarrow \mathbb{R}^n$$



defined by

$$h(x, t) = \begin{cases} (1-t)x + t \cdot \frac{1}{\|x\|} x & \text{if } \|x\| \geq 1, \\ x & \text{if } \|x\| \leq 1 \end{cases}$$

is a deformation retraction of Euclidean  $n$ -space  $\mathbb{R}^n$  onto the  $n$ -disk  $D_n$ .  $\diamond$

ex:contraction-as-def-retraction

- (3) A contraction of a space  $X$  to a point  $y \in X$  is just a deformation retraction of  $X$  onto its subspace  $\{y\}$ . Thus deformation retractions are generalizations of contractions.

A deformation retraction  $h: X \times I \rightarrow X$  of a space  $X$  onto a subspace  $E$  gives rise to the map  $x \mapsto h(x, 1)$  of  $X$  onto  $E$  which leaves each  $x \in E$  fixed.

def:retraction

**5.73 Definition.** Let  $E$  be a subspace of a topological space  $X$ . A **retraction of  $X$  onto  $E$**  is a continuous map

$$r: X \rightarrow E$$

such that

$$r(x) = x \quad (x \in E)$$

(whence  $r$  maps  $X$  onto  $E$ ); in other words, a retraction of  $X$  onto  $E$  is a continuous extension to  $X$  of the identity map of  $E$ . When some retraction of  $X$  exists,  $E$  is called a **retract of  $X$** .

A deformation retract  $E$  of  $X$  is thus a retract of  $X$  (but a deformation retraction of  $X$  onto  $E$  is not a retraction of  $X$  onto  $E$ ). However, **a retract of a space need not be a deformation retract of that space**. In fact, a one-point subspace of a space  $X$  is always a retract of  $X$ , but a one-point subspace cannot be a deformation retract of  $X$  unless  $X$  is contractible.

The next proposition explains why [Examples 5.72](#) did not include, for example, a deformation retraction of the plane  $\mathbb{R}^2$  onto the ball  $B_2$ .

prop:retract-of-T2-is-closed

**5.74 Proposition.** A retract of a Hausdorff space is closed in that space.

**Proof.** Let  $r: X \rightarrow E$  be a retraction of a space  $X$  onto a subspace  $E$ . Let  $x \in X \setminus E$  be arbitrary. Then  $r(x) \neq x$ , so that  $r(x)$  and  $x$  have disjoint neighborhoods  $V$  and  $W$  in  $X$ . Since  $r$  is continuous at  $x$ , there is a neighborhood  $U_1$  of  $x$  with  $r(U_1) \subset V$ . Then the neighborhood  $U = U_1 \cap W$  of  $x$  is disjoint from  $E$ .  $\square$

Earlier we saw that the disk  $D_n$  is contractible, and later we shall prove that the circle  $S_1$  is not. From the next result it will then follow that  $S_1$  is not a retract of  $D_2$ .

retract-of-contractible-is-contractible

**5.75 Proposition.** A retract of a contractible space is itself contractible.

**Proof.** Let  $c: X \times I \rightarrow X$  be a contraction of the space  $X$  to a point  $z \in X$ , and let  $r: X \rightarrow E$  be a retraction of  $X$  onto a subspace  $E$ . Define  $h: E \times I \rightarrow E$  by

$$h(x, t) = r(c(x, t)).$$

Then  $x \in E$  implies

$$h(x, 0) = r(x) = x \quad h(x, 1) = r(z),$$

and so  $h$  is a contraction of  $E$  onto the point  $r(z)$  of  $E$ .  $\square$

contraction!deformation retraction@  
deformation retraction!contraction@  
deformation retract  
deformation retraction  
retract  
retraction

deformation retract  
subsec:homotopy  
contractible space

## Homotopies

A deformation retraction  $h: X \times I \rightarrow X$  of  $X$  onto  $E \subset X$  may be viewed as a continuous family of paths  $t \mapsto h(x, t)$  in  $X$  ending in  $E$ , one path for each point  $x \in X$  (compare the proof of Proposition 5.69). It may also be viewed as a continuous family of continuous maps

$$h_t: X \rightarrow X, \\ x \mapsto h(x, t)$$

one map for each  $t \in I$ , where  $h_0$  is the identity map of  $X$  and  $h_1$  sends  $X$  onto  $E$  while leaving each  $x \in E$  fixed. From this latter viewpoint,  $h$  may then be regarded as a continuous process that deforms the map  $h_0$  into the map  $h_1$ . This idea may be generalized.

def:homotopy **5.76 Definition.** Let  $f, g: X \rightarrow Y$  be continuous maps. We call a continuous map

$$h: X \times I \rightarrow Y$$

a **homotopy from  $f$  to  $g$**  when

$$h(x, 0) = f(x), \quad h(x, 1) = g(x) \quad (x \in X)$$

and then we write

$$h: f \simeq g.$$

When such an  $h$  exists, we say that  $f$  is **homotopic to  $g$**  and write  $f \simeq g$ .

In terms of the maps

$$h_t: X \rightarrow Y \quad (t \in I), \\ x \mapsto h(x, t)$$

the condition that a continuous map  $h: X \times I \rightarrow Y$  must satisfy to be a homotopy from  $f$  to  $g$  is simply that

$$h_0 = f, \quad h_1 = g.$$

ex: def-retract

**5.77 Examples.** (1) Let  $j: E \rightarrow X$  be the inclusion map of a subspace  $E$  of a space  $X$  into  $X$  and let  $i: X \rightarrow X$  be the identity map. Then a deformation retraction of  $X$  onto  $E$  is just a homotopy  $h: i \simeq j \circ r$  for some retraction  $r: X \rightarrow E$  of  $X$  onto  $E$ .

(2) A space  $X$  is contractible precisely when the identity map of  $X$  is homotopic to a constant map of  $X$  into  $X$ .

(3) Any two continuous maps  $f, g: X \rightarrow Y$  from the same space into a contractible space  $Y$  are homotopic. In fact, let  $c: Y \times I$  be a contraction of  $Y$  to a point  $z \in Y$ . Define  $h: X \times I \rightarrow Y$  by

$$h(x, t) = \begin{cases} c(f(x), 2t) & \text{if } 0 \leq t \leq 1/2, \\ c(g(x), 2 - 2t) & \text{if } 1/2 \leq t \leq 1. \end{cases}$$

Then  $h: f \simeq g$ .

In particular, any two continuous maps from the same space into  $\mathbb{R}^n$ , or more generally into a convex subspace of  $\mathbb{R}^n$ , are homotopic.

(4) The identity map of  $S_1$  is homotopic to the reflection  $x \mapsto -x$  of  $S_1$  through the origin. To construct a homotopy  $h$ , notice that the two maps are rotations of  $S_1$  through angles

of 0 and  $\pi$ , respectively. Then for each  $t \in I$  take  $h_t$  to be the rotation through an angle of  $t\pi$ . More precisely, represent each  $x \in S_1$  in the form

$$x = \langle \cos \pi s, \sin \pi s \rangle$$

with  $s \in \mathbb{R}$  and set

$$h(x, t) = \langle \cos \pi(s + t), \sin \pi(s + t) \rangle.$$

[Of course, it must be checked that this definition of  $h(x, t)$  is independent of the particular choice of  $s$  used to represent  $x$ .]  $\diamond$

thm:homotopy-equiv-rel

**5.78 Theorem.** The homotopy relation  $\simeq$  is an equivalence relation on the set of all continuous maps of a topological space  $X$  to a topological space  $Y$ .

**Proof.** Reflexivity. Let  $f: X \rightarrow Y$  be continuous. Then the map  $h: X \times I \rightarrow Y$  defined by  $h(x, t) = f(x)$  is a homotopy from  $f$  to itself.

Symmetry. Let  $h$  be a homotopy from  $f$  to  $g$ . Then the map  $H: X \times I \rightarrow Y$  defined by

$$H(x, t) = h(x, 1 - t)$$

is a homotopy from  $g$  to  $f$ .

Transitivity. Let  $h$  be a homotopy from  $f$  to  $g$  and  $k$  be a homotopy from  $g$  to  $j$ .

$$F: f_0 \simeq f_1, \quad G: f_1 \simeq f_2$$

be homotopies of continuous maps  $f_0, f_1, f_2: X \rightarrow Y$ . Define  $H: X \times I \rightarrow Y$  by

$$H(x, t) = \begin{cases} F(x, 2t) & \text{if } 0 \leq t \leq 1/2, \\ G(x, 2t - 1) & \text{if } 1/2 \leq t \leq 1. \end{cases}$$

Then  $H: f_0 \simeq f_2$  since  $H_0 = f_0$  and  $H_1 = G_1$ .  $\square$

### Path-homotopies

Aside from contractions and deformation retractions, the homotopies of most interest are deformations of paths that leave endpoints fixed. Suppose  $\sigma$  and  $\tau$  are paths in a space  $X$ , both from a point  $x_0$  to a point  $x_1$ . To deform  $\sigma$  into  $\tau$  while leaving their ends  $x_0$  and  $x_1$  fixed is to prescribe a continuous family  $\langle h_t \rangle_{t \in I}$  of paths in  $X$  from  $x_0$  to  $x_1$  with  $h_0 = \sigma$  and  $h_1 = \tau$ .

def:homotop-paths

**5.79 Definition.** Let  $\sigma, \tau: I \rightarrow X$  be paths in a topological space  $X$  with

$$\sigma(0) = \tau(0) = x_0, \quad \sigma(1) = \tau(1) = x_1.$$

We call a homotopy  $h: \sigma \simeq \tau$  for which

$$h_t(0) = x_0, \quad h_t(1) = x_1 \quad (t \in I)$$

a **path-homotopy from  $\sigma$  to  $\tau$  (in  $X$ )** and write

$$h: \sigma \sim \tau.$$

When such an  $h$  exists we say that  $\sigma$  is **path-homotopic to  $\tau$  (in  $X$ )** and write  $\sigma \sim \tau$ .

**change in parameter** A path-homotopy from  $\sigma$  to  $\tau$  is thus just a continuous map

$$h: I \times I \rightarrow X$$

on the unit square for which

$$\begin{aligned} h(s, 0) &= \sigma(s), & h(s, 1) &= \tau(s) & (s \in I), \\ h(0, t) &= x_0, & h(1, t) &= x_1 & (t \in I), \end{aligned}$$

as represented in Figure 5.18. The first variable, being denoted by  $s$ , parametrizes each

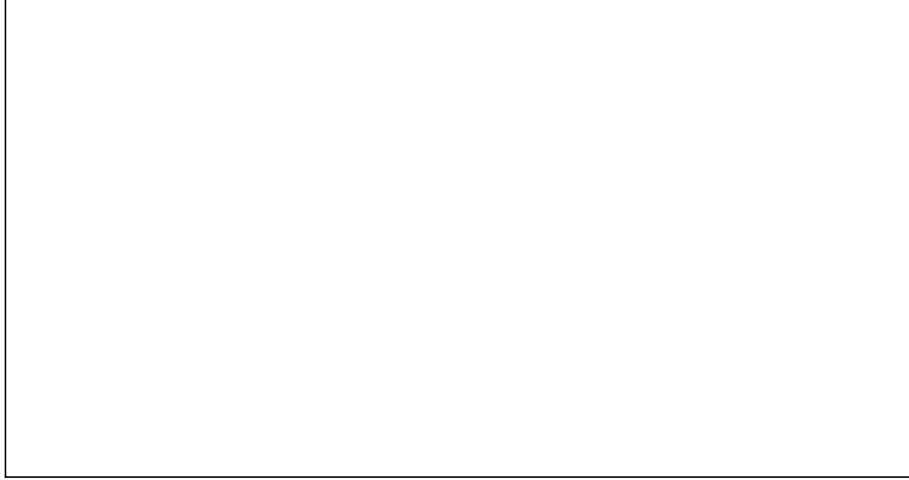


Figure 5.18: A path-homotopy as a map on the unit square.

fig:path-homotopy-as-map-on-square

individual path  $h_t$  from  $x_0$  to  $x_1$ ; and the second variable, being denoted by  $t$ , parametrizes the family  $\langle h_t \rangle_{t \in I}$  of paths  $h_t$ .

If in the proof of Theorem 5.78 we start with path homotopies between paths having the same endpoints, then the homotopies constructed there will also be path-homotopies. Hence **the path-homotopy relation  $\sim$  is an equivalence relation on the set of all paths in a space from a point  $x_0$  to a point  $x_1$ .**

**5.80 Examples.** (1) Let  $\sigma$  and  $\tau$  be paths from a point  $x_0$  to a point  $x_1$  in a convex subspace  $X$  of  $\mathbb{R}^n$ . Then  $\sigma \sim \tau$ , with a path-homotopy from  $\sigma$  to  $\tau$  being the map

$$h(s, t) = (1 - t)\sigma(s) + t\tau(s),$$

which from time  $t = 0$  to time  $t = 1$  moves  $\sigma(s)$  along a straight line to  $\tau(s)$ .

- (2) Let  $\sigma$  be any path in a space  $X$ . A strictly increasing function  $\alpha: I \rightarrow I$  of  $I$  onto  $I$ , such as  $s \mapsto s^2$ , yields a new path  $\tau = \sigma \circ \alpha$  in  $X$  from  $\sigma(0)$  to  $\sigma(1)$  by a “change in parameter.” The formula

$$h(s, t) = \sigma((1 - t)s + t\alpha(s))$$

defines a path-homotopy  $h: \sigma \sim \tau$  for which  $h_t(I) = \sigma(I)$  for each  $t \in I$ .  $\diamond$

The next two lemmas concern what happens to products of paths and reverses of paths when those paths are replaced by others path-homotopic to them.

lem:hom-preserve-path-prod

**5.81 Lemma (path-homotopy preserves path-products).** *Let  $\sigma, \sigma'$  be paths in a topological space from a point  $x$  to a point  $y$  and let  $\tau, \tau'$  be paths in from  $y$  to a point  $z$  in the same space. If  $\sigma \sim \sigma'$  and  $\tau \sim \tau'$ , then  $\sigma * \tau \sim \sigma' * \tau'$ .*

**Proof.** Denote by  $X$  the topological space involved. Let  $f: \sigma \sim \sigma'$  and  $g: \tau \sim \tau'$  be path-homotopies, so that for all  $s, t \in I$ ,

$$\begin{aligned} f(s, 0) &= \sigma(s), & f(s, 1) &= \sigma'(s), & f(0, t) &= x, & f(1, t) &= y, \\ g(s, 0) &= \tau(s), & g(s, 1) &= \tau'(s), & g(0, t) &= y, & g(1, t) &= z. \end{aligned}$$

Divide the unit square  $I \times I$  into two rectangles,  $L = [0, 1/2] \times I$  on the left and  $R = [1/2, 1] \times I$  on the right, by the vertical line segment  $S$  joining the midpoints  $\langle 1/2, 0 \rangle$  and  $\langle 1/2, 1 \rangle$  of its top and bottom, respectively, as indicated in Figure 5.19.

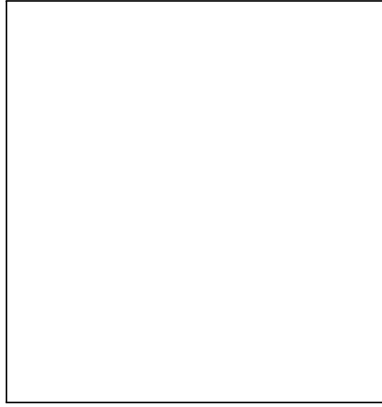


Figure 5.19: Dividing the unit square into two rectangles to make  $\sigma * \tau \sim \sigma' * \tau'$ .

fig:divide-square-2-rectangles

To construct the desired path-homotopy  $h$  from  $\sigma * \tau$  to  $\sigma' * \tau'$ , expand the left-hand rectangle  $L$  rightward to the unit square by means of the linear map  $\phi: L \rightarrow I \times I$  given by

$$\phi(s, t) = \langle 2s, t \rangle \quad (0 \leq s \leq 1/2, 0 \leq t \leq 1),$$

and expand the right-hand rectangle  $R$  leftward to the unit square by means of the linear map  $\psi: R \rightarrow I \times I$  given by

$$\psi(s, t) = \langle 2s - 1, t \rangle \quad (1/2 \leq s \leq 1, 0 \leq t \leq 1),$$

The rectangles  $L$  and  $R$  have intersection  $S = L \cap R = \{1/2\} \times I$ , the vertical line segment joining the midpoints  $\langle 1/2, 0 \rangle$  and  $\langle 1/2, 1 \rangle$  of the square's top and bottom, respectively. On this intersection, the maps  $\phi$  and  $\psi$  agree, that is,

$$(f \circ \phi)|_S = (g \circ \psi)|_S.$$

In fact,

$$(f \circ \phi)(1/2, t) = f(\phi(1/2, t)) = f(1, t) = y = g(1, t) = g(\psi(1/2, t)) = (g \circ \psi)(1/2, t)$$

for all  $t \in I$ . By the Gluing Lemma (Theorem 3.13), there is a unique continuous map  $h: I \times I \rightarrow X$  for which

$$h|_L = f \circ \phi, \quad h|_R = g \circ \psi.$$

We claim that  $h$  is the desired path-homotopy. In terms of coordinates,

$$h(s, t) = \begin{cases} f(2s, t) & \text{if } 0 \leq s \leq 1/2, \\ g(2s - 1, t) & \text{if } 1/2 \leq s \leq 1. \end{cases}$$

path-product!associativity@and associativity

Then for all  $s \in I$

$$h(s, 0) = \begin{cases} f(2s, 0) = \sigma(2s) = (\sigma * \tau)(s) & \text{if } 0 \leq s \leq 1/2, \\ g(2s - 1, 0) = \tau(2s - 1) = (\sigma * \tau)(s) & \text{if } 1/2 \leq s \leq 1, \end{cases}$$

that is,

$$h(s, 0) = (\sigma * \tau)(s).$$

Similarly, for all  $s \in I$ ,

$$h(s, 1) = (\sigma' * \tau')(s).$$

Further, for all  $t \in I$ ,

$$h(0, t) = f(0, t) = x,$$

$$h(1, t) = g(1, t) = y.$$

Thus  $h: \sigma * \tau \sim \sigma' * \tau'$ , as desired.  $\square$

The analog of [Lemma 5.81](#) holds for reversing paths.

lem:hom-preserve-path-reverse

**5.82 Lemma (path-homotopy preserves path reverse).** *Let  $\sigma, \sigma'$  be paths in a topological space having the same initial point and the same terminal point. If  $\sigma \sim \sigma'$ , then  $\bar{\sigma} \sim (\sigma')^{\leftarrow}$ .*

We now examine the algebraic properties “up to homotopy” of products of paths.

Path-products themselves do not satisfy such customary algebraic laws as associativity. If  $\sigma, \tau$ , and  $\eta$  are paths in a space  $X$  with  $\tau$  starting where  $\sigma$  ends and  $\eta$  starting where  $\tau$  ends, then it is not true in general that  $\sigma * (\tau * \eta) = (\sigma * \tau) * \eta$ . For example, define paths  $\sigma, \tau$ , and  $\eta$  in  $X = \mathbb{R}$  by

$$\sigma(s) = s, \quad \tau(s) = s + 1, \quad \eta(s) = s + 2.$$

The graphs of the two triple-products  $\sigma * (\tau * \eta)$  and  $(\sigma * \tau) * \eta$ , shown in [Figure 5.20](#), are

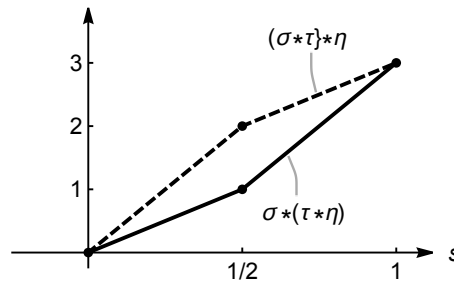


Figure 5.20: Example of non-associativity of path-products.

fig:triple-path-product

clearly not the same. In particular,

$$(\sigma * (\tau * \eta))(1/2) = 2 \neq 1 = ((\sigma * \tau) * \eta)(1/2).$$

The *traces* of two triple-products of three paths are the same—here, the interval  $[0, 3]$ —and this is true in general; it is just that the two triple-product paths pass through particular points on their common trace at different times.

In general, the two triple-products of three paths are essentially the same from the viewpoint of path-homotopy: they are path-homotopic. This will be the first of several results below saying that the operations of path-product and path-reverse behave nicely *up to homotopy*.

path-product!associativity@and asso  
path-product!commutativity@and co  
path-product!commutativity@and co  
loop

prop:path-homotopy-assoc

**5.83 Proposition (associativity of path-product up to homotopy).** Let  $\sigma$ ,  $\tau$ , and  $\eta$  be paths in a space with  $\tau$  starting where  $\sigma$  ends and  $\eta$  starting where  $\tau$  ends. Then

$$\sigma * (\tau * \eta) \sim (\sigma * \tau) * \eta.$$

Proposition 5.83 justifies the following notational convention.

convention:triple-path-prod

**Convention!** Whenever it is solely a matter of path-homotopy, *we may unambiguously denote either of such triple products  $\sigma * (\tau * \eta)$  and  $(\sigma * \tau) * \eta$  simply by  $\sigma * \tau * \eta$ .*

**Proof of Proposition 5.83.** Let  $\sigma$ ,  $\tau$ , and  $\eta$  go from point  $x$  to point  $y$ , from  $y$  to  $z$ , and from  $z$  to  $w$  in the space  $X$ . Figure 5.21 suggest how to construct the desired path-homotopy  $h: \sigma * (\tau * \eta) \sim (\sigma * \tau) * \eta$ . On the bottom of the square drawn there,  $\sigma$  is to operate in the left-hand half-interval  $[0, 1/2]$  while  $\tau * \eta$  is to operate in the right-hand half-interval  $[1/2, 1]$ ; the latter means that  $\tau$  is to operate in the quarter-length subinterval  $[1/2, 3/4]$  while  $\eta$  is to operate in the quarter-length subinterval  $[3/4, 1]$ . Similarly, on the top of the square,  $\sigma$  is to operate in  $[0, 1/4]$  while  $\tau$  is to operate in  $[1/4, 1/2]$ ; and  $\eta$  is to operate in  $[1/2, 1]$ . Accordingly, join by slanted lines the indicated interior points  $1/2$  and  $3/4$  on the bottom to the points  $1/4$  and  $1/2$ , respectively.

The path-homotopy  $h$  will need to take constant value  $x$  on the square's left side,  $z$  on the right side,  $y$  along the first of those two slanted lines and  $z$  on the second of those slanted lines, as indicated on the figure.

To complete constructing the desired  $h$ , linearly squeeze or compress what  $h$  must do on each of the three bottom subintervals to the corresponding top subinterval. Details are left to an exercise.  $\square$

If  $\sigma$  and  $\tau$  are paths for which both path products  $\sigma * \tau$  and  $\tau * \sigma$  can be formed, that is, with  $\sigma$  ending where  $\tau$  starts and  $\tau$  ending where  $\sigma$  starts, then it can happen that  $\sigma * \tau \neq \tau * \sigma$ . But the situation is even worse than that.

**Caution!** Suppose that  $\sigma$  and  $\tau$  are paths for which both path products  $\sigma * \tau$  and  $\tau * \sigma$  can be formed. Then *it is not the case in general that  $\sigma * \tau \sim \tau * \sigma$* . In short, commutativity of path-product up to homotopy is *not* the case in general.

Of particular interest, especially later, will be paths that start and end at the same point.

def:loop

**5.84 Definition.** Let  $x$  be a point in a topological space  $X$ . Then a **loop at  $x$  (in  $X$ )** is a path from  $x$  to  $x$  in  $X$ . In particular, the **null loop at  $x$  (in  $X$ )**, denoted by  $\nu_x$ , is the constant path in  $X$  with  $\nu_x(s) = x$  for all  $s \in I$ .

loop!map on  $S^1$  as map on  $S^1$  Sone



Figure 5.21: Constructing a path-homotopy  $h: \sigma * (\tau * \eta) \sim (\sigma * \tau) * \eta$ .

fig:path-homotopy-assoc

undamental-loop-around-unit-circle

**5.85 Examples.** (1) The path  $\omega$  in  $\mathbb{R}^2$  given by  $\omega(s) = e^{2\pi i s} = \langle \cos 2\pi s, \sin 2\pi s \rangle$  is a loop at the point  $\langle 1, 0 \rangle$  in  $\mathbb{R}^2$ . This loop winds around the unit circle once, in the counterclockwise direction.

(2) The product  $\sigma * \tau$  of two loops at a point  $x$  in a space  $X$  is again a loop at  $x$ .

In particular, for  $\sigma = \tau = \omega$ , the loop of (1), the the path product  $\omega * \omega$  of the  $\omega$  with itself is another loop at  $\langle 1, 0 \rangle$  in  $\mathbb{R}^2$ . This loop winds around the unit circle twice, again in the counterclockwise direction.

(3) The reverse of a loop at a point  $x$  in a space  $X$  is again a loop at  $x$ .

In particular, the reverse  $\bar{\omega}$  of the loop of (1) is another loop at  $\langle 1, 0 \rangle$  in  $\mathbb{R}^2$ , this loop winding once around the unit circle but in the clockwise direction.  $\diamond$

rem:2-paths-to-loop-and-vice-versa

**5.86 Remark.** Two paths  $\sigma$  and  $\tau$  from a point  $x$  in a space  $X$  to a point  $y$  in  $X$  may be combined into a *loop* at  $x$  by traversing  $\sigma$  forward and then  $\tau$  in reverse (see [Definitions 5.52](#) and [5.54](#)); this loop

$$\omega = \sigma * \bar{\tau}$$

is given by

$$\omega(s) = \begin{cases} \sigma(2s) & \text{if } 0 \leq s \leq 1/2, \\ \tau(2 - 2s) & \text{if } 1/2 \leq s \leq 1. \end{cases}$$

Conversely, an arbitrary loop  $\omega$  at  $x$  may be obtained in this manner starting with paths  $\sigma, \tau$  from  $x$  to the point  $y = \omega(1/2)$ , namely,

$$\begin{aligned} \sigma(s) &= \omega(s/2) & (0 \leq s \leq 1), \\ \tau(s) &= \omega(1 - s/2) & (0 \leq s \leq 1). \end{aligned}$$

A loop at a point  $x$  in a space  $X$  has the same range as a continuous map  $S^1 \rightarrow X$  sending  $1 = \langle 1, 0 \rangle$  to  $x$ , and vice versa.



prop:loops-vs-maps-on-circle

**5.87 Proposition.** Let  $x$  be a point in a topological space  $X$ . Let  $q: I \rightarrow S_1$  be the restriction of the standard map, that is, the quotient map  $s \mapsto e^{2\pi i s} = \langle \cos 2\pi s, \sin 2\pi s \rangle$ . Then:

- (1) For each loop  $\omega$  at  $x$  in  $X$ , there is a unique continuous map  $f_\omega$  satisfying  $f_\omega \circ q = \omega$ , that is, making the diagram

$$\begin{array}{ccc} I & & \\ q \downarrow & \searrow \omega & \\ S_1 & \xrightarrow{f_\omega} & X \end{array}$$

commutative.

- (2) if  $f: S_1 \rightarrow X$  is a continuous map sending  $\langle 1, 0 \rangle$  to  $x$ , then there is a unique loop  $\omega_f$  in  $X$  at  $x$  satisfying  $f \circ q = \omega_f$ , that is, making the diagram

$$\begin{array}{ccc} I & & \\ q \downarrow & \searrow \omega_f & \\ S_1 & \xrightarrow{f} & X \end{array}$$

commutative, namely,  $\omega_f = f \circ q$ .

**Proof.** (1) Pass to quotients: see [Theorem 3.76](#) and [Examples 3.81 \(1\)](#).  $\square$

def:set-of-loops

**5.88 Definition.** Given a point  $x$  in a topological space  $X$ , the set of all loops in  $X$  at  $x$  is denoted by  $\Omega(X, x)$ .

Thus we have a one-to-one correspondence  $\omega \mapsto f_\omega$  between the set  $\Omega(X, x)$  of all loops in  $X$  at  $x$  and the set  $C[(S_1, 1), (X, x)]$  of all continuous maps from the pointed space  $\langle S_1, 1 \rangle$  to the pointed space  $\langle X, x \rangle$ .

Let us return to the matter of path-homotopies of loops and, more generally, paths.

prop:path-homotopy-identities

**5.89 Proposition (null loops as one-sided identities for paths up to homotopy).** Let  $\sigma$  be a path from a point  $x$  to a point  $y$  in a space. Then:

$$\begin{aligned} \nu_x * \sigma &\sim \sigma, \\ \sigma * \nu_y &\sim \sigma. \end{aligned}$$

Of particular interest later will be the special case of [Proposition 5.89](#) when the path is a loop.

cor:path-homotopy-identity-loop

**5.90 Corollary (null loop as identity for loops up to homotopy).** Let  $\omega$  be a loop at a point  $x$  in a space. Then

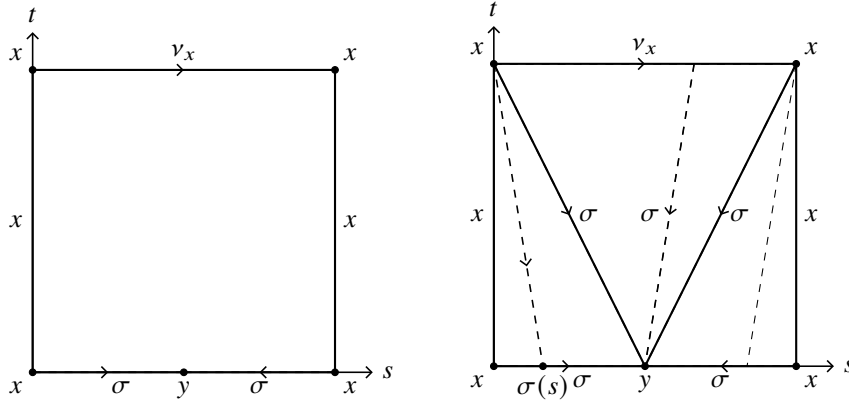
$$\nu_x * \omega \sim \omega \sim \omega * \nu_x.$$

**5.91 Proposition (reverse of loop as inverse up to homotopy).** Let  $\sigma$  be a path from a point  $x$  to a point  $y$  in a space. Then

$$\sigma * \overleftarrow{\sigma} \sim \nu_x,$$

$$\overleftarrow{\sigma} * \sigma \sim \nu_y.$$

prop:path-homotopy-inverses-path



(a) Action of  $h$  on square's sides. (b) Action of  $h$  on slanted lines.

Figure 5.22: Constructing a path-homotopy  $h: \sigma * \overleftarrow{\sigma} \sim \nu_x$ .

fig:path-homotopy-sigma-ast-revers

**Proof.** To construct the desired path-homotopy  $h: \sigma * \overleftarrow{\sigma} \sim \nu_x$ , first consider how it must map the sides of the unit square, as depicted in Figure 5.22(a). Next, consider how it might then map the two solid slanted lines shown there so as to keep  $h$  continuous; what will work is to map those two slanted lines as indicated in Figure 5.22(b), that is, to run through all the values of  $\sigma$  as you move down each line from top to bottom.

To see how to define  $h$  between those two solid slanted lines, that is, on the triangular region in the center, fill the region with line segments drawn from points  $\langle s, 1 \rangle$  on its top to the point  $\langle 1/2, 0 \rangle$  at its bottom vertex. As a point starting at  $\langle s, 1 \rangle$  moves downward along such a segment, let  $h$  run through all the values of  $\sigma$ , as indicated in Figure 5.22(b).

To see how to define  $h$  on the triangular region to the lower left of the left-hand slanted line, fill the region with line segments drawn from the upper left corner  $\langle 1, 1 \rangle$  to points  $\langle s, 0 \rangle$  on the  $s$ -axis for  $0 \leq s \leq 1/2$ . As a point starting at the corner moves downward along such a segment, make  $h$  run through values  $\sigma(0)$  through  $\sigma(s)$  only, as indicated on the dashed line shown in Figure 5.22(b). Similarly define  $h$  on the triangular region to the lower right of the right-hand slanted line.  $\square$

Again, the special case when the path is a loop will be of particular interest later.

**5.92 Corollary (reverse of loop is as two-sided inverse up to homotopy).** Let  $\omega$  be a loop at a point  $x$  in a space. Then

$$\omega * \overleftarrow{\omega} \sim \nu_x \sim \overleftarrow{\omega} * \omega.$$

cor:path-homotopy-inverse-loop

From the preceding results, it appears that on the set  $\Omega(X, x)$  of all loops at a given point  $x$  in a space  $X$ , the operation  $*$  of path-product is *almost* a group operation, with the null loop  $\nu_x$  being the identity and the reverse  $\bar{\sigma}$  of a loop  $\sigma$  at  $x$  being its inverse; indeed, if path-homotopy  $\sim$  could be replaced by equality  $=$  in those results, then we would have a group operation. In [Section 7.1](#) we shall obtain a genuine group operation on the *quotient set*  $\Omega(X, x)/\sim$  of  $\Omega(X, x)$  under the path-homotopy equivalence relation  $\sim$  on  $\Omega(X, x)$ .

loop  
path-homotopy  
punctured line  
punctured plane

fix: material on  
simply connected  
moved to next  
section

### EXERCISES FOR SECTION 5.4

- 82.** Construct a contraction of the subspace

$$(\{1\} \times \{0\}) \cup (\{0\} \times \{1\}) \cup \bigcup_{n=1}^{\infty} (\{1/n\} \times \{1\})$$

of the plane to its point  $\langle 0, 1 \rangle$ .

- 83. (a)** Prove that each subset of  $\mathbb{R}^n$  that is starlike ([Exercise 19](#)) is contractible, thereby extending [Examples 5.68 \(1\)](#).  
**(b)** Describe or draw several contractible subsets of the plane  $\mathbb{R}^2$  that are not starlike (and hence not convex, either).  
**84.** Show that contractibility is a topological property.  
**85.** Must the product of two contractible spaces be contractible? the product of infinitely many spaces?  
**86.** Must the Cartesian sum of two contractible spaces be contractible?  
**87.** Show that a topological space  $X$  can always be embedded in the cone  $K(X)$  over  $X$  ([Exercise 3.191](#)) and that  $K(X)$  is contractible.

prob-part: def-retr-square

- 88. (a)** Construct a deformation retraction of the square  $[-1, 1] \times [1, 2]$  onto the union of its left side, right side, and bottom.  
*(Hint: Move each point of the square along the line passing through the point and on the portion of the line segment within the square of the line drawn from the point  $\langle 0, 2 \rangle$ ; you will want to distinguish the case that this line segment intersects the left and right sides from the case that it intersects the bottom of the square.)*

prob-part: def-retr-cylinder

- (b)** Construct a deformation retraction of the solid cylinder  $D_2 \times [1, 2]$  onto its subspace consisting of its lateral surface together with its bottom.  
**(c)** Generalize [\(a\)](#) and [\(b\)](#) to arbitrary dimension.

prob-part: def-retract-punc-line-to-S0

- 89. (a)** Construct a deformation retraction of the “punctured line”  $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$  onto its subspace  $\{-1, 1\}$ .

prob-part: def-retract-punc-plane-to-S1

- (b)** Construct a deformation retraction of the “punctured plane”  $(\mathbb{R}^2)^* = \mathbb{R}^2 \setminus \{0\}$  onto the unit circle.

prob-part: def-retract-punc-n-space-to-Snminus1

- (c)** Generalize [\(a\)](#) and [\(b\)](#) to arbitrary dimension  $n$ .

- 90.** Find a subspace of the Möbius strip  $M$  [[Examples 3.81 \(5\)](#)] homeomorphic to  $S_1$  that is a deformation retract of  $M$ .

prob-part: retract-of-retract

- 91. (a)** If  $A \subset X$  is a retract of  $X$  and  $B \subset A$  is a retract of  $A$ , show that then  $B$  is a retract of  $X$ .

path-connected space! retract (b) Does the analog of (a) for deformation retracts hold?

retract! path-connected space @ of path-connected space  
 prob: retract-path-conn 92. Must a retract of a path-connected space be path-connected?

93. Prove that a subspace  $E$  of a space  $X$  is a retract of  $X$  if and only if each continuous map  $f: E \rightarrow Y$  into each topological space  $Y$  has a continuous extension to  $X$ .

94. Call a space  $Y$  “nice” if each continuous map  $f: A \rightarrow Y$  on each closed subspace  $A$  of each topological space  $X$  can be extended continuously to  $X$ . Prove that a retract of a nice space is also nice.

95. Show that a space  $Y$  is contractible if each two continuous maps  $f, g: X \rightarrow Y$  from each space  $X$  into  $Y$  are homotopic to one another.

96. Let  $f, f': X \rightarrow Y$  and  $g, g': Y \rightarrow Z$  be continuous maps with  $f \simeq f'$  and  $g \simeq g'$ . Deduce that  $g \circ f \simeq g' \circ f'$ .

97. (a) Let  $f, g: X \rightarrow S_n$  be continuous maps such that for all  $x \in X$ , the image  $f(x)$  is not antipodal to  $g(x)$ , that is,  $f(x) \neq -g(x)$ . Show that  $f \simeq g$ .

[Hint: Project the line segment joining  $f(x)$  and  $g(x)$  onto  $S_n$ .]

(b) If  $f: X \rightarrow S_n$  is a continuous map that is *not* surjective, show that  $f$  must be homotopic to a constant map.

98. Prove that a continuous map  $f: S_n \rightarrow Y$  is homotopic to a constant map if and only if  $f$  has a continuous extension to  $D_{n+1}$ .

(Hint: For inspiration, draw a picture for  $n = 1$ .)

99. Suppose a continuous map  $f: S_n \rightarrow Y$  is homotopic to a constant map  $c: S_n \rightarrow Y$ . Given a point  $z \in S_n$ , construct a homotopy  $h$  from  $f$  to a constant map with value some  $y \in Y$  such that  $h(x, t) = y$  for all  $t \in I$ .

100. Draw diagrams illustrating the proof that  $\sim$  is an equivalence relation on the set of all paths in a space from a point  $x_0$  to a point  $x_1$ .

101. (a) The discussion in Remark 5.86 showed how a pair  $\langle \sigma, \tau \rangle$  of paths, both starting at a point  $x$  and ending at a point  $y$  in a space  $X$ , give rise to a loop  $\omega$  at  $x$ . Show that if  $\langle \sigma', \tau' \rangle$  is another such pair of paths, giving rise to a loop  $\omega'$  at  $x$ , then  $\omega \sim \omega'$ .

(b) The discussion in Remark 5.86 also showed how a loop  $\omega$  at a point  $x$  in a space  $X$  gives rise to a pair  $\langle \sigma, \tau \rangle$  of paths for which  $\sigma * \tau = \omega$ . If  $\omega'$  is another loop at  $x$  and if  $\langle \sigma', \tau' \rangle$  is the pair of paths similarly obtained, must  $\sigma \sim \sigma'$  and  $\tau \sim \tau'$ ?

b: homotopy-is-family-of-cont-maps 102. (Note: This exercise shows that under some restrictions on the spaces involved, a homotopy really is a continuous family of continuous maps.)

Let  $\langle X, d \rangle$  be a *compact* metric space, let  $\langle Y, d' \rangle$  be any bounded metric space, and let  $h: X \times I \rightarrow Y$  be a continuous map. Then for each  $t \in I$ , the map  $h_t: X \rightarrow Y$  given by  $x \mapsto h(t, x)$  is continuous. Provide the set  $C(X, Y)$  of all continuous maps from  $X$  to  $Y$  with its topology of uniform convergence [Examples 3.109 (2)]. Prove that the map

$$\begin{aligned} I &\rightarrow C(X, Y) \\ t &\mapsto h_t \end{aligned}$$

is continuous.

**103.** (Continuation of [Exercise 102](#).) In the notation of [Exercise 102](#), let  $f, g: X \rightarrow Y$  be continuous maps. Show that  $f \simeq g$  if and only if there is a path in the space  $C(X, Y)$  from  $f$  to  $g$ .

simply connected space  
convex set! simply connected space@  
simply connected space! line integral  
simply connected space! conservative  
loop

**104.** Let  $X$  be a topological space and let  $\langle Y, d' \rangle$  be a bounded metric space. Give the set  $C(X, Y)$  of all continuous maps from  $X$  to  $Y$  its topology of uniform convergence [[Examples 3.109 \(2\)](#)]. Let  $\langle h_t \rangle_{t \in I}$  be a family of continuous maps from  $X$  to  $Y$  for which the map

$$\begin{aligned} I &\rightarrow C(X, Y) \\ t &\mapsto h_t \end{aligned}$$

is continuous. Show that the map

$$\begin{aligned} h: X \times I &\rightarrow Y \\ (x, t) &\mapsto h_t(x) \end{aligned}$$

need not be continuous—so that  $h$  need not be a homotopy from  $h_0$  to  $h_1$ .

fix: Which  
topology on  
 $C(X, Y)$ ?

fix: Need more  
basic homotopy  
problems!

## 5.5 Simply connected spaces

sec:simplyconnected

Examination of [Figure 5.18](#) suggests that, for a path-homotopy  $h: \sigma \sim \tau$  between two paths joint the same points, as  $t$  varies from 0 to 1, the paths  $h_t$  must move through the points “surrounded by” the union  $\sigma(I) \cup \tau(I)$  of the traces of the given paths. Hence the requirement that any two points in a space with the same endpoints be path-homotopic says, in a way, that *the space  $X$  has no “holes.”*

### Simple connectedness

subsec:simple-connected

def:simply-conn

**5.93 Definition.** A topological space  $X$  is said to be **simply connected** if  $X$  is path-connected and if  $\sigma \sim \tau$  whenever  $\sigma$  and  $\tau$  are paths in  $X$  for which  $\sigma(0) = \tau(0)$  and  $\sigma(1) = \tau(1)$ .

Together, [Examples 5.80 \(1\)](#) and [Examples 5.57 \(1\)](#) say that **a convex subset of  $\mathbb{R}^n$  is simply connected**. In particular, the plane  $\mathbb{R}^2$  and the ball  $B_2$  are simply connected. It is an extremely attractive conjecture—and is true—that neither a circle nor an annulus is simply connected, but the proofs will take some effort.

Simply connected open sets in the plane play an important part in the theory of line integrals; see, for example, Apostol [[2](#), Chapter 10 and pages 383–385]. Suppose

$$F(x, y) = \langle P(x, y), Q(x, y) \rangle$$

is a continuously differentiable vector field on a connected open set  $S$  in the plane satisfying the condition  $\partial P / \partial y = \partial Q / \partial x$ . If  $S$  is simply connected, then the value of the line integral  $\int F \cdot d\gamma$  along a piecewise smooth path  $\gamma$  in  $S$  from a point  $\langle x_0, y_0 \rangle$  to a point  $\langle x_1, y_1 \rangle$  will be independent of the particular path used, and then  $F$  will be the gradient of some scalar field on  $S$ . (In physical terms, if  $F$  represents a force field on the simply connected region  $S$ , then  $F$  will be conservative.)

**Remark 5.86** discussed how a pair  $\langle \sigma, \tau \rangle$  of paths, both starting at a point  $x$  and ending at a point  $y$  in a space  $X$ , give rise to a loop  $\omega$  at  $x$ ; and how a loop  $\omega$  at a point  $x$  in a space  $X$  gives rise to a pair  $\langle \sigma, \tau \rangle$  of paths for which  $\sigma * \tau = \omega$ . Accordingly, simple connectedness should be expressible purely in terms of loops.

loop!nullhomotopic  
def:loop-nullhomotopic

**5.94 Definition.** A loop  $\omega$  at a point  $x$  in a space  $X$  is said to be **nullhomotopic (in  $X$ )** when  $\omega \sim \nu_x$  in  $X$ .

Loosely speaking, a loop at  $x$  in  $X$  is nullhomotopic when it can be shrunk within the space  $X$  to the single point  $x$  through a continuous family of loops at  $x$  in  $X$  (see Figure 5.23).



Figure 5.23: A nullhomotopic loop.

fig:nullhomotopic-loop

thm:simply-conn-via-loops

**5.95 Theorem.** A necessary and sufficient condition for a path-connected space to be simply connected is that each loop at each point of the space be nullhomotopic.

**Proof.** Let  $X$  be a path-connected space.

**Necessity.** Assume that  $X$  is simply connected. Let  $\omega$  be an arbitrary loop at an arbitrary point  $x$  in  $X$ . Then both  $\omega$  and the null loop  $\nu_x$  are paths in  $X$  from  $x$  to  $y = x$ , and so  $\omega \sim \nu_x$ .

**Sufficiency.** Assume that the condition holds. Let  $\sigma$  and  $\tau$  be paths in  $X$  from  $x$  to  $y$ . Then  $\sigma * \tilde{\tau}$  is a loop at  $x$ , and so by assumption

$$\sigma * \tilde{\tau} \sim \nu_x.$$

By the algebraic properties in Lemma 5.81 and Propositions 5.83, 5.89, and 5.91, it follows successively that

$$\begin{aligned} (\sigma * \tilde{\tau}) * \tau &\sim \nu_x * \tau, \\ \sigma * (\tilde{\tau} * \tau) &\sim \nu_x * \tau, \\ \sigma * \nu_y &\sim \nu_x * \tau, \end{aligned}$$

and finally,  $\sigma \sim \tau$ .  $\square$

Relying on Theorem 5.95, when treating simple connectedness we shall actually work with loops and not arbitrary paths. For example, we next explain why our definition of a simply connected space stipulated that the space be path-connected.

**5.96 Proposition.** A path-connected space is simply connected if it contains some point  $x$  for which each loop at  $x$  is nullhomotopic.

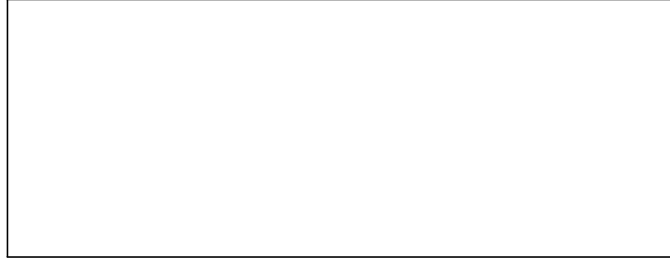


Figure 5.24:

fig:nullhomotopic-loops-at-some-pt

**Proof.** Assume that  $X$  is path-connected and has such a point  $x$ . Let  $\omega$  be a loop at an arbitrary point  $y$  in  $X$ . We wish to show that  $\omega \sim \nu_y$ . Since  $X$  is path-connected, there is some path  $\sigma$  in  $X$  from  $x$  to  $y$ . By traversing first  $\sigma$  forward, then  $\omega$  forward, and finally  $\sigma$  backward (see Figure 5.24) we obtain the loop  $\sigma * \omega * \bar{\sigma}$ . (Here, and below, we are following the Convention from page 581.)

By hypothesis,

$$\{\text{eq:loops-at-x-nullhomotopic}\} \quad (*) \quad \sigma * \omega * \bar{\sigma} \sim \nu_x.$$

In the same way we just used  $\sigma$  to transform the loop  $\omega$  at  $y$  into the loop  $\sigma * \omega * \bar{\sigma}$  at  $x$ , so we now use  $\bar{\sigma}$  to transform the loops  $\sigma * \omega * \bar{\sigma}$  and  $\nu_x$  into the loops  $\bar{\sigma} * (\sigma * \omega * \bar{\sigma}) * \sigma$  and  $\bar{\sigma} * \nu_x * \sigma$ , respectively, at  $y$ . From \*,

$$\bar{\sigma} * (\sigma * \omega * \bar{\sigma}) * \sigma \sim \bar{\sigma} * \nu_x * \sigma.$$

Now

$$\{\text{eq:eq:loops-at-x-nullhomotopic-2}\} \quad (**) \quad \bar{\sigma} * (\sigma * \omega * \bar{\sigma}) * \sigma \sim \omega,$$

$$\{\text{eq:eq:loops-at-x-nullhomotopic-3}\} \quad (***) \quad \bar{\sigma} * \nu_x * \sigma \sim \nu_y.$$

Hence  $\omega \sim \nu_y$ , as desired.  $\square$

The preceding proposition enables us to expand our repertory of simply connected spaces beyond convex subsets of  $\mathbb{R}^n$  to arbitrary contractible spaces.

**5.97 Theorem.** Every contractible space is simply connected.

**Proof.** Let  $X$  be a contractible space and let  $\omega$  be a loop at a point  $x$  of  $X$ . We shall show that  $\omega \sim \nu_x$ . According to Proposition 5.69, already  $X$  is path-connected. Choose a contraction  $c: X \times I \rightarrow X$  of  $X$  to one of its points  $x$ . (Note that, according to Proposition 5.70, the space will be contractible to each of its points.)

It is tempting to take for the desired homotopy  $g: \omega \sim \nu_x$  the map  $g: I \times I \rightarrow X$  given by

$$g(s, t) = c(\omega(s), t).$$

*This would be incorrect!* To be sure,  $g(s, 0) = \omega(s)$  and  $g(s, 1) = x$  for all  $s \in I$ , so that  $g$  is a homotopy from  $\omega$  to  $\nu_x$ . Unfortunately,  $g$  need not take the constant value  $x$  on the vertical sides  $\{0\} \times I$  and  $\{1\} \times I$  of the square  $I \times I$ ; in other words,  $g$  need not be a *path*-homotopy.

Our strategy for a correct proof is to construct a certain loop  $\sigma$  at  $x$  for which

$$\sigma * \omega * \overleftarrow{\sigma} \sim \nu_x.$$

From that it will follow that

$$\overleftarrow{\sigma} * (\sigma * \omega * \overleftarrow{\sigma}) * \sigma \sim \overleftarrow{\sigma} * \nu_x * \sigma,$$

and so

$$\omega \sim \nu_x.$$

To construct  $\sigma$ , we shall first construct a homotopy—but *not* a path-homotopy—

$$h: \omega \simeq \nu_x$$

for which

$$(*) \quad h(0, t) = h(1, t) \quad (t \in I).$$

Then the restriction of  $h$  to the left edge  $\{0\} \times I$  of the square may be regarded as a loop at  $x$ , and  $\sigma$  is defined to be the inverse of this loop; in other words,

$$\sigma(t) = h(0, 1 - t) \quad (t \in I)$$

(see [Figure 5.25](#)).



Figure 5.25: Constructing the loop  $\sigma$  in the proof of [Theorem 5.97](#).

fig:new-loop-sigma-from-h

Once  $h$  and then  $\sigma$  have been constructed, it is easy to construct a *path*-homotopy  $F: \sigma * \omega * \overleftarrow{\sigma} \sim \nu_x$  in the way suggested by [Figure 5.26](#), namely: take

$$F(s, t) = \begin{cases} x & \text{if } t \geq 3s, \\ \sigma(3s - t) & \text{if } t \leq 3s \end{cases}$$

on the left-third rectangle  $[0, 1/3] \times I$ ; take

$$F(s, t) = h(3s - 1, t)$$

on the middle-third rectangle  $[1/3, 2/3] \times I$ ; and take

$$F(s, t) = \begin{cases} \sigma(3 - t - 3s) & \text{if } t \leq 3 - 3s, \\ x & \text{if } t \geq 3 - 3s \end{cases}$$

on the right-third rectangle  $[2/3, 1] \times I$ .





simple closed curve  
Jordan Curve Theorem

Figure 5.26: Constructing the path-homotopy  $F: \sigma * \omega * \bar{\sigma} \sim \nu_x$  in the proof of Theorem 5.97.

fig:path-homotop-F-contractible-is-s

To construct  $h: \sigma \simeq \nu_x$  satisfying equation (\*), we regard  $\omega$  as a map  $f: S_1 \rightarrow X$  sending  $\langle 1, 0 \rangle \in S_1$  to  $x$ . More precisely, in the notation of Proposition 5.87,  $f = f_\omega$  is the unique continuous map such that  $f \circ q = \omega$ , where  $q: I \rightarrow S_1$  is the restriction of the standard map  $s \mapsto e^{2\pi si}$  to  $I$ . (Of course,  $f$  is not actually a loop at  $x$  since its domain is not  $I$ .) Since  $X$  is contractible, by cref, the map  $f$  is homotopic to the constant map taking value  $x$  on  $S_1$ . Hence there is a homotopy

$$h': S_1 \times I \rightarrow X$$

for which

$$h'(z, 0) = f(z), \quad h'(z, 1) = x \quad (z \in S_1).$$

By cutting the cylinder  $S_1 \times I$  along the line segment  $\langle 1, 0 \rangle \times I$  and unwrapping it, we obtain the desired homotopy  $h: \omega \simeq \nu_x$ . More precisely, define  $h: I \times I \rightarrow X$  by

$$h(s, t) = h'(q(s), t).$$

Finally, condition equation (\*) is satisfied because

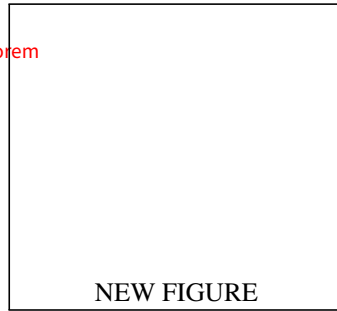
$$h(0, t) = h'(q(0), t) = h'(q(1), t) = h(1, t). \quad \square$$

Let  $\omega$  be a loop in a Hausdorff space  $X$  that does not pass through any point more than once except at times 0 and 1, that is,  $\omega(s) \neq \omega(s')$  whenever  $s, s' \in ]0, 1[$  with  $s \neq s'$ . Let  $f_\omega: S_1 \rightarrow X$  be the unique continuous map for which  $f_\omega \circ q = \omega$ , where  $q: I \rightarrow S_1$  is the restriction of the standard map  $s \mapsto e^{2\pi si}$  to  $I$  (see Proposition 5.87). Then  $f_\omega$  is injective and maps  $S_1$  homeomorphically onto  $f_\omega(S_1)$ . In other words, the range  $\omega(I)$  is a **simple closed curve**—a homeomorphic image of the circle [see Figure 5.27(a)].

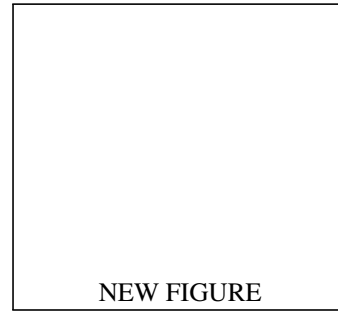
Let us now specialize to the plane. We want to assert that a simple closed curve  $C$  in the plane divides the plane into two regions, one inside and the other outside  $C$ . This assertion is geometrically evident when  $C$  is a circle, a square, or a triangle. But is it evident for more general  $C$ ? For example, which points are “inside” the simple closed curve shown in Figure 5.28 and which are “outside” it?

The precise assertion we want is the following.

Jordan, Camille  
 Veblen, Oswald  
 Jordan Curve Theorem  
 inside  
 outside



(a)  $\omega(I)$  is a simple closed curve. subfig:simple-closed-curve



(b)  $\omega(I)$  is *not* a simple closed curve. subfig:non-simple-closed-curve

Figure 5.27: Simple and non-simple closed curves.

fig:simple-and-non-simple-closed-curve

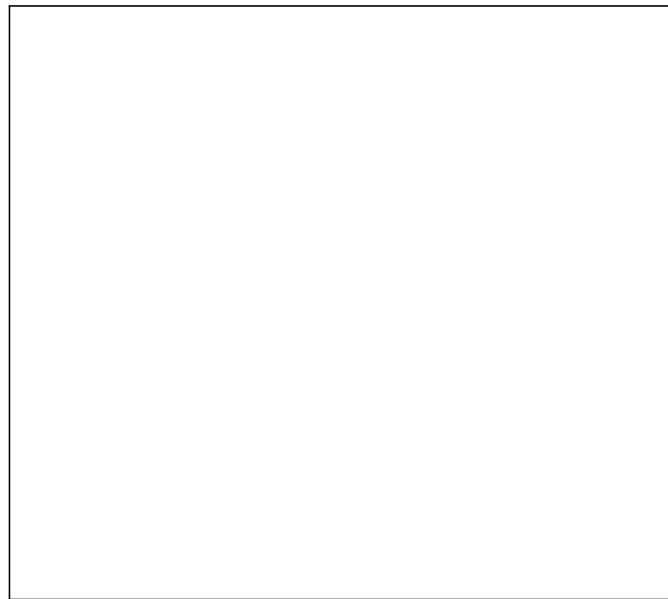


Figure 5.28: Which points are inside, and which are outside, this simple closed curve?

fig:weird-simple-closed-curve

thm:Jordan-curve-thm

**5.98 Jordan Curve Theorem.** *A simple closed curve  $C$  in the plane has a complement consisting of exactly two components, one of them bounded and the other unbounded for the Euclidean metric, with  $C$  being the boundary of both.*

fix: Will I include  
 proof of Jordan  
 Curve Theorem  
 using fundamental  
 group?

The explicit statement of this assertion as a theorem requiring proof is due to C. Jordan, who in 1893 gave an incomplete proof. The first correct proof was given only later, in 1905, by O. Veblen. Known proofs require methods outside the scope of this book; an accessible one having a very geometric flavor may be found in Hall and Spencer II [32, Chapter 5, Sections 3–4].

In terms of the Jordan Curve Theorem, the bounded complement of a simple closed curve in the plane is known as its **inside** and the unbounded complement as its **outside**.

Let us now combine the preceding idea to give a nice geometric criterion for the simple

connectedness of certain spaces. Consider a subspace  $X$  of the plane that is open and connected—and hence by [Corollary 5.61](#) path-connected. For  $X$  to be simply connected, according to [Theorem 5.95](#) it is necessary that each loop  $\omega$  in  $X$  whose range is a simple closed curve be nullhomotopic; then the inside of  $\omega(l)$  ought to be contained in  $X$ . And this is correct (although we offer no proof here): **an open and connected subspace  $X$  of the plane is simply connected if and only if the inside of each simple closed curve in  $X$  is contained in  $X$ .** Informally stated: *a simply connected region in the plane is just a region having no holes.*

### The circle and simple connectedness

subsec:circ-and-simply-conn

In this subsection we shall present the promised proof that the unit circle  $S_1$  is *not* simply connected and then shall deduce from that several important consequences, including Fundamental Theorem of Algebra.

State by considering the loop  $s \mapsto e^{2\pi i s} = \langle \cos 2\pi s, \sin 2\pi s \rangle$  that wraps around  $S_1$  once in the counterclockwise direction. To show that  $S_1$  is not simply connected it will suffice to establish that this one particular loop is not nullhomotopic. However geometrically evident this fact may seem, its proof is surprisingly difficult. To approach the proof, we study homotopies of loops in  $S_1$  by studying homotopies of paths in the real line  $\mathbb{R}$ .

For the remainder of this section, we denote by  $q$  the standard map

$$\begin{aligned} q: \mathbb{R} &\rightarrow S_1 \\ s &\mapsto e^{2\pi i s} = \langle \cos 2\pi s, \sin 2\pi s \rangle. \end{aligned}$$

We have  $q(n) = 1 = \langle 1, 0 \rangle$  for every integer  $n$ . This implies that an arbitrary path  $\sigma$  in  $\mathbb{R}$  from 0 to an integer  $n$  induces a loop  $q \circ \sigma: I \rightarrow S_1$  at  $\langle 1, 0 \rangle$  in  $S_1$ . For example, the path  $\sigma(s) = s$  from 0 to 1 in  $\mathbb{R}$  induces the loop considered in the preceding paragraph; and the path  $\sigma(s) = 0$  from 0 to 0 induces the null loop at  $\langle 1, 0 \rangle$ . More generally, for each positive integer  $n$ , the path  $\sigma(s) = ns$  from 0 to  $n$  in  $\mathbb{R}$  induces a loop at  $\langle 1, 0 \rangle$  in  $S_1$  that wraps around the circle  $n$  times in the counterclockwise direction.

Conversely, let a loop  $\tau$  at  $\langle 1, 0 \rangle$  in  $S_1$  be given. We claim that  $\tau$  is induced in the preceding manner by some unique path  $\sigma$  in  $\mathbb{R}$  with  $\sigma(0) = 0$ . (*Caution:* The path  $\sigma$  need not be a loop at 0 in  $\mathbb{R}$ , as the preceding examples demonstrate.) In other words, we want to **lift** the loop  $\tau$  in a unique way to a continuous map  $\sigma: I \rightarrow \mathbb{R}$  with  $\sigma(0) = 0$  and making the following diagram commutative.

$$\begin{array}{ccc} & & \mathbb{R} \\ & \nearrow \sigma & \downarrow q \\ I & \xrightarrow{\tau} & S_1 \end{array}$$

To proceed, we need the elementary property of  $q$  is a **covering map**: each  $z \in S_1$  has an open neighborhood  $V$  in  $S_1$  that is **evenly covered** in the sense that  $q^{-1}(V)$  is the union of a collection of disjoint open subsets of  $\mathbb{R}$  each of which is mapped homeomorphically onto  $V$  by  $q$ .

lem:q-R-to-S1-covering-map

**5.99 Lemma.** The map  $q: \mathbb{R} \rightarrow S_1$  given by  $s \mapsto e^{2\pi i s} = \langle \cos 2\pi s, \sin 2\pi s \rangle$  is a covering map.

**Proof.** Let  $z \in S_1$  be arbitrary. Take  $V$  to be any proper open arc  $q(]a, b[)$  in  $S_1$  containing  $z$ . Since

$$q(s) = q(s') \iff s - s' \in \mathbb{Z},$$

we have

$$q^{-1}(V) = \bigcup_{n \in \mathbb{Z}} \{n + s : a < s < b\}$$

with each of the translates  $\{n + s : a < s < b\}$  of  $]a, b[$  being open in  $\mathbb{R}$  and mapped homeomorphically onto  $V$  by  $q$ . (See Figure 5.29).  $\square$

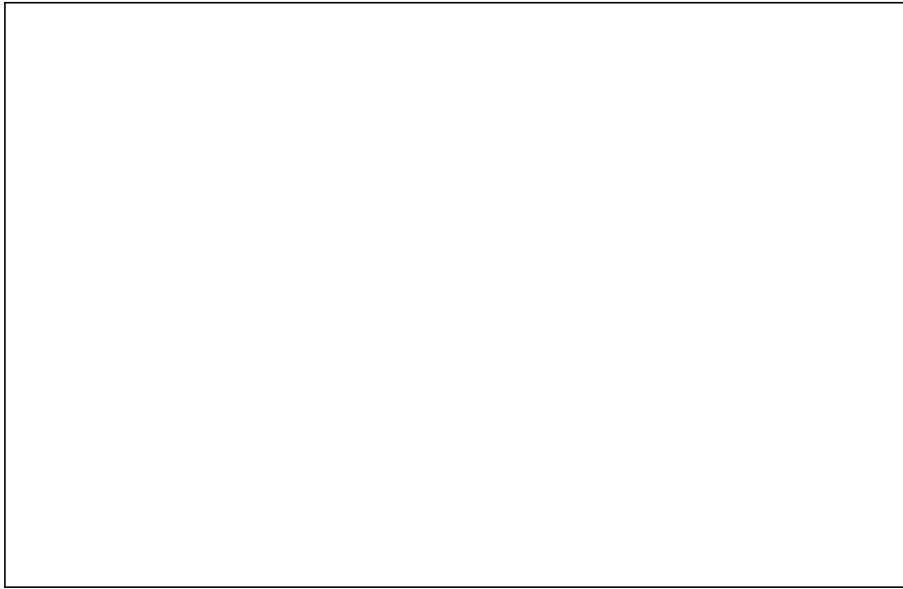


Figure 5.29: Evenly covering an open arc on the unit circle.

fig:even-cover-open-arc

The following result generalizes the discussion of this subsection's beginning that each loop in the circle arises from a path in  $\mathbb{R}$ .

lem:path-lift-S1-to-R

**5.100 Lemma (path-lifting property for the circle).** Let  $\tau: I \rightarrow S_1$  be any path in  $S_1$  and let  $u \in \mathbb{R}$  be any point for which

$$q(u) = \tau(0).$$

Then there is a unique path  $\sigma: I \rightarrow \mathbb{R}$  in  $\mathbb{R}$  such that

$$q \circ \sigma = \tau, \quad \sigma(0) = u.$$

**Proof.** Existence. We shall suitably choose points

{eq:suitable-s-pf-path-lifting-circle} (\*)

$$0 = s_0 < s_1 < \cdots < s_i < s_{i+1} < \cdots < s_n = 1$$

in  $I$  and then successively define  $\sigma$  on the subintervals  $[0, 0] = [0, s_0]$ ,  $[0, s_1]$ ,  $\dots$ ,  $[0, s_n] = I$  of  $I$ , extending the definition on each subinterval to the next one.

The points  $s_0, s_1, \dots, s_n$  satisfying (\*) are so chosen that for each  $i = 0, 1, \dots, n-1$ ,

$$\tau([s_i, s_{i+1}]) \subset V_i$$

for some evenly covered open subset  $V_i$  of  $S_1$ . This choice can be made using the compactness of  $I$  as follows. Each  $z \in S_1$  has an evenly covered open neighborhood  $V_z$ . By Lemma 4.63, the open cover  $\{\tau^{-1}(V_z) : z \in S_1\}$  of  $I$  has a Lebesgue number  $\varepsilon > 0$ . Arbitrarily choose  $s_0, s_1, \dots, s_n \in I$  satisfying (\*) with  $|s_{i+1} - s_i| < \varepsilon$  for each  $i$ ; then for each  $i$  take  $V_i = V_z$  for some  $z \in S_1$  with  $[s_i, s_{i+1}] \subset \tau^{-1}(V_z)$ .

Using induction we now show that for each  $i = 0, 1, \dots, n$  there is a continuous map

$$\sigma_i : [0, s_i] \rightarrow \mathbb{R}$$

such that

$$\{eq:sigma-i-pf-path-lifting-circle\} \quad (**) \quad q \circ \sigma_i = \tau|_{[0, s_i]}, \quad \sigma_i(0) = u.$$

Then  $\sigma = \sigma_n$  will be the desired lift of  $\tau$ . To begin the induction, simply set  $\sigma(0) = u$ .

Now let  $0 \leq i < n$  and assume that a map  $\sigma_i : [0, s_i] \rightarrow \mathbb{R}$  satisfying (equation (\*\*)) exists. The next map  $\sigma_{i+1} : [0, s_{i+1}] \rightarrow \mathbb{R}$  is obtained by glueing together  $\sigma_i$  and a continuous map  $\eta : [s_i, s_{i+1}] \rightarrow \mathbb{R}$  for which

$$\eta(s_i) = \sigma_i(s_i), \quad q \circ \eta = \tau|_{[s_i, s_{i+1}]}.$$

Such an  $\eta$  exists by the choice of  $s_i$  and  $s_{i+1}$ . In fact, since  $(q \circ \sigma_i)(s_i) = \tau(s_i)$ , we have  $\sigma_i(s_i) \in q^{-1}(V_i)$ . Since  $V_i$  is evenly covered, there is a unique open set  $U$  in  $\mathbb{R}$  that contains  $\sigma_i(s_i)$  and is mapped homeomorphically by  $q$  onto  $V_i$ . Denote by  $q'$  the inverse of the homeomorphism  $U \cong V_i$  defined by  $q$ . Finally, take

$$\eta(s) = q'(\tau(s)) \quad s \in [s_i, s_{i+1}].$$

This completes the extension of  $\sigma_i$  to  $\sigma_{i+1}$  and concludes the induction.

Uniqueness. Suppose that  $\sigma, \sigma'$  are two paths in  $\mathbb{R}$  for which

$$q \circ \sigma = \tau = q \circ \sigma', \quad \sigma(0) = u = \sigma'(0).$$

We shall show that the subset

$$E = \{s \in I : \sigma(s) = \sigma'(s)\}$$

of  $I$  is the whole of  $I$ . Now  $E$  is nonempty because  $0 \in E$ . By continuity of  $\sigma$  and  $\sigma'$ , the set  $E$  is closed in  $I$ . In view of the connectivity of  $I$ , it remains only to show that  $E$  is also open in  $I$ .

Let  $s \in E$  be arbitrary. Choose an evenly covered open neighborhood  $V$  of  $(q \circ \sigma)(s)$  in  $S_1$ . There is an open subset  $U$  of  $\mathbb{R}$  containing  $\sigma(s)$  which  $q$  maps homeomorphically onto  $V$ . Let

$$W = \sigma^{-1}(U) \cap (\sigma'^{-1})(U).$$

Then  $W$  is an open neighborhood of  $s$  in  $I$ . We claim that  $W \subset E$ . In fact, if  $t \in W$  is arbitrary, then:  $\sigma(t), \sigma'(t) \in U$ ; by hypothesis  $q(\sigma(t)) = q(\sigma'(t))$ ; and hence  $\sigma(t) = \sigma'(t)$  since  $q$  is injective on  $U$ .  $\square$

For proving that the circle is not simply connected, the crucial result is the fact that a path-homotopy in the circle can be lifted to a path-homotopy in the real line.

lem:homotopy-lift-prop-S1-to-R

**5.101 Lemma (homotopy-lifting property for the circle).** *Let*

$$h: \omega_0 \sim \omega_1$$

*be a path-homotopy of loops  $\omega_0, \omega_1: \mathbb{I} \rightarrow S_1$  at  $\langle 1, 0 \rangle$  in  $S_1$ . Let  $\sigma_0, \sigma_1: \mathbb{I} \rightarrow \mathbb{R}$  be the unique paths for which*

$$q \circ \sigma_0 = \omega_0, \quad \sigma_0(0) = 0 = \sigma_1(0), \quad q \circ \sigma_1 = \omega_1.$$

*Then*

$$\sigma_0(1) = \sigma_1(1)$$

*and there is a path-homotopy  $H: \sigma_0 \sim \sigma_1$  such that  $q \circ H = h$ .*

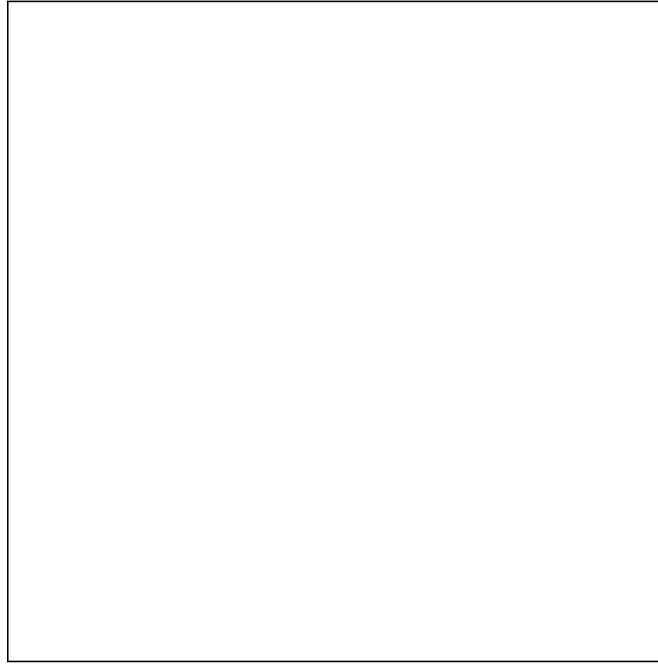


Figure 5.30: Constructing the map  $G_j$  to prove the homotopy-lifting property for the circle. fig:construct-Gj-pf-homotopy-lift-cir

**Proof.** We shall use the method of the preceding proof to obtain a continuous map  $H: \mathbb{I} \times \mathbb{I} \rightarrow \mathbb{R}$  such that

$$H_0 = \sigma_0, \quad q \circ H = h.$$

We can then conclude the proof by repeatedly invoking the uniqueness assertion of [Lemma 5.100](#) as follows. The path  $t \mapsto H(0, t)$  and the null loop  $\nu_0$  at 0 in  $\mathbb{R}$  satisfy

$$\begin{aligned} (q \circ H)(0, t) &= h(0, t) = \langle 1, 0 \rangle = (q \circ \nu_0)(t) & (t \in \mathbb{I}), \\ H(0, 0) &= \sigma_0(0) = 0 = \nu_0(0), \end{aligned}$$

and so by [Lemma 5.100](#),

$$H(0, t) = 0 \quad (t \in \mathbb{I}).$$

In particular,  $H(0, 1) = 0$ . Next, the paths  $H_1$  and  $\sigma_1$  in  $\mathbb{R}$  satisfy

$$\begin{aligned}(q \circ H_1)(s) &= q(H(s, 1)) = h(s, 1) = \omega_1(s) = (q \circ \sigma_1)(s) & (s \in I), \\ H_1(0) &= H(0, 1) = 0 = \sigma_1(0),\end{aligned}$$

and so  $H_1 = \sigma_1$ , that is,

$$H(s, 1) = \sigma_1(s) \quad (s \in I).$$

Next, the path  $t \mapsto H(1, t)$  and the null loop  $t \mapsto \sigma_0(1)$  at  $\sigma_0(1)$  in  $\mathbb{R}$  satisfy

$$\begin{aligned}(q \circ H)(1, t) &= h(1, t) = (1, 0) = \omega_0(1) = (q \circ \sigma_0)(1) & (t \in I), \\ H(1, 0) &= H_0(1) = \sigma_0(1),\end{aligned}$$

and so

$$H(1, t) = \sigma_1(1) \quad (t \in I).$$

Hence

$$\sigma_0(1) = H(1, 1) = \sigma_1(1)$$

and  $H: \sigma_0 \sim \sigma_1$ .

To construct such  $H$ , we shall apply the same kind of reasoning as in the proof of [Lemma 5.100](#). First, divide the square  $I \times I$  into rectangles determined by points

$$\begin{aligned}0 &= s_0 < s_1 < \cdots < s_i < s_{i+1} < \cdots < s_n = 1, \\ 0 &= t_0 < t_1 < \cdots < t_j < t_{j+1} < \cdots < t_m = 1\end{aligned}$$

in such a way that for each  $i = 0, 1, \dots, n-1$  and  $j = 0, 1, \dots, m-1$ ,

$$h([s_i, s_{i+1}] \times [t_j, t_{j+1}]) \subset V_{i,j}$$

for some evenly covered open subset  $V_{i,j}$  of  $S_1$ .

What we are about to do now is use an induction on  $j$  to establish the existence of the restrictions of the desired  $H$  on the successively taller horizontal rectangles  $I \times [0, t_j]$  until we reach  $j = m$ , which will give  $H$  on the entire square  $I \times I$ . Further, to do each  $j$ th step of the induction, we shall use an induction on  $i$  to establish existence of extension of  $H$  beyond the horizontal rectangle  $I \times [0, t_j]$  so as to include also successively longer horizontal strips  $[0, s_i] \times [0, t_{j+1}]$ ; see [Figure 5.31](#). Thus we shall use a *double induction*—an induction within an induction!

Starting with the map  $G_0: I \times \{0\} \rightarrow \mathbb{R}$  defined by  $G_0(s, 0) = \sigma_0(s)$  we use induction to show that for each  $j = 0, 1, \dots, m$ , there is a continuous map

$$G_j: I \times [0, t_j] \rightarrow \mathbb{R}$$

on the horizontal strip  $I \times [0, t_j]$  such that

$$(G_j)_0 = \sigma_0, \quad q \circ G_j = h|_{(I \times [0, t_j])}.$$

Suppose  $0 \leq j < m$  and that  $G_j$  has already been shown to exist. To show that  $G_{j+1}$  exists, use a new induction (within the induction on  $j$ ) to show that for each  $i = 0, 1, \dots, n$ , there is a continuous map

$$G_{j+1}^i: (I \times [0, t_j]) \cup ([0, s_i] \times [t_j, t_{j+1}]) \rightarrow \mathbb{R}$$

that extends  $G_j$  and satisfies  $(q \circ G_{j+1}^i)(s, t) = h(s, t)$  on its domain; the domains of  $G_j$  and  $G_{j+1}^i$  are indicated in [Figure 5.31](#). This induction is similar to the one in the proof of [Lemma 5.100](#).  $\square$

retraction!unit disk@and unit disk

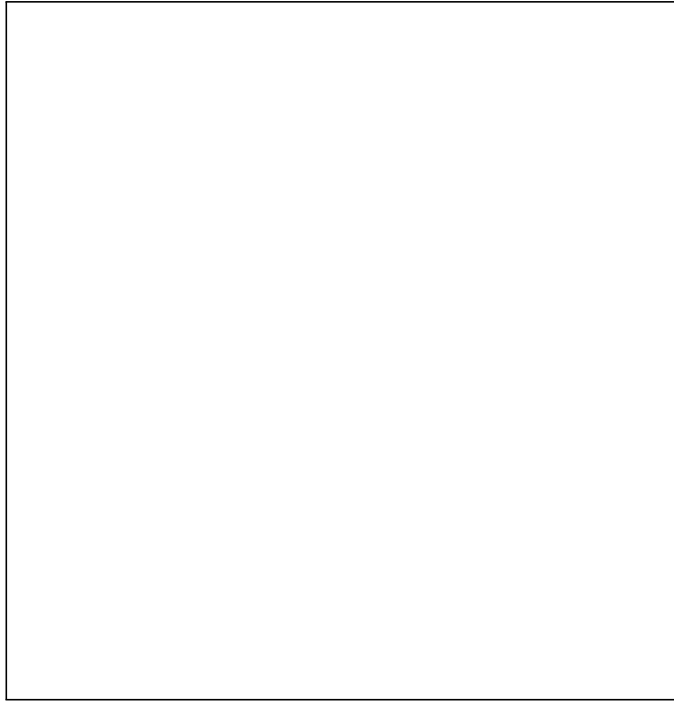


Figure 5.31: The domains of  $G_j$  and  $G_{j+1}^i$  in the proof of the homotopy-lifting property. fig:doms-Gj-Gjplus1toi

Our main theorem about loops in the circle is an immediate consequence of [Lemma 5.101](#).

thm:loops-in-S1 **5.102 Theorem.** *Let  $m$  and  $n$  be distinct integers and let  $\sigma_n, \sigma_m: I \rightarrow \mathbb{R}$  be the paths given by*

$$\sigma_n(s) = ns, \quad \sigma_m(s) = ms \quad (s \in I).$$

*Then the loops  $q \circ \sigma_n$  and  $q \circ \sigma_m$  at the point  $\langle 1, 0 \rangle$  in  $S_1$  are not path-homotopic.*

In particular, the null loop  $q \circ \sigma_0$  at  $\langle 1, 0 \rangle$  in  $S_1$  is not path-homotopic to the loop  $q \circ \sigma_n$  whenever  $n \neq 0$ . At last we have proved the following theorem.

thm:S1-not-simply-conn **5.103 Theorem.** *The circle  $S_1$  is not simply connected.*

It is geometrically plausible that it is not possible to have a continuous map of the 2-disk onto its boundary, the unit circle  $S_1$ , leaving all points of  $S_1$  in place. And it is true!

thm:no-retract **5.104 No-Retraction Theorem.** *The circle  $S_1$  is not a retract of the disk  $D_2$ .*

**Proof.** According to [Examples 5.68 \(1\)](#), the disk  $D_2$  is contractible. If  $S_1$  were a retract of  $D_2$ , then by [Proposition 5.75](#) the circle  $S_1$  would be contractible and hence, by [Theorem 5.97](#), would be simply connected.  $\square$

Recall that a point  $x$  of a space  $X$  is called a **fixed-point** of a map  $f: X \rightarrow X$  when  $f(x) = x$ . Earlier (see [Corollary 5.15](#)) we deduced from the connectedness of the unit interval  $I$  that each continuous map  $f: I \rightarrow I$  has a fixed-point. From the No-Retraction



Theorem we now deduce an analogous result for dimension 2, namely: each continuous map  $f: I^2 \rightarrow I^2$  has a fixed-point. Given such an  $f$ , the map  $g: D_2 \rightarrow D_2$  defined by  $g = h^{-1} \circ f \circ h$ , where  $h: D_2 \cong I^2$ , is also continuous, and  $h(y)$  will be a fixed-point of  $f$  whenever  $y \in D_2$  is a fixed point of  $g$ . Hence it suffices to consider maps from  $D_2$  to itself.

thm:Brouwer-fixed-pt-dim-2

**5.105 Brouwer Fixed-point Theorem in Dimension 2.** *Each continuous map from the 2-disk  $D_2$  to itself has a fixed-point.*

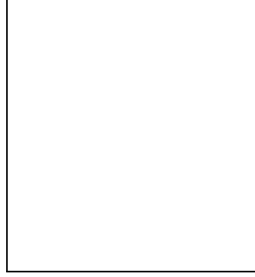


Figure 5.32: Construction of  $r$  in the proof of the Brouwer Fixed-point Theorem in Dimension 2.

fig:r-for-pf-Brouwer-dim-2

**Proof.** Let  $g: D_2 \rightarrow D_2$  be a continuous map and just suppose that  $g(x) \neq x$  for each  $x \in D_2$ . Then for each  $x \in D_2$ , there is a unique line  $L_x$  passing through the distinct points  $x$  and  $g(x)$ , and  $L_x$  intersects the bounding circle  $S_1$  of  $D_2$  in exactly two points; denote by  $r(x)$  that one of these two points lying on the same side of  $g(x)$  as does  $x$  (see Figure 5.32). We shall show that the map  $r: D_2 \rightarrow S_1$  so defined is a retraction, thereby contradicting the No-Retraction Theorem.

By construction,  $r(x) = x$  when  $x \in S_1$ . It remains only to establish the continuity of  $r$ . This is an easy consequence of the continuity of  $g$ . In fact, for each  $x \in D_2$ , the point  $r(x)$  is the unique  $y = tx + (1 - t)g(x)$  for which  $t \geq 1$  and  $\|y\| = 1$ . By explicitly solving the quadratic equation  $\|y\|^2 = 1$  for  $t$  in terms of  $x$  we see that its root  $t$  satisfying  $t \geq 1$  depends continuously on  $x$  and  $g(x)$ , hence on  $x$  alone. Therefore the value  $r(x)$  depends continuously on  $x$ .  $\square$

Brouwer's theorem has a surprising physical interpretation. Imagine a circular dish filled with a thin layer of some liquid. Suppose the liquid is stirred for a while and suppose that when all motion stops the molecules on the surface, and only those molecules, remain on the surface. Then some molecule on the surface will be where it was before the stirring.

One mathematical application of Brouwer's theorem concerns a system

$$\{\text{eq:2-2dim-eqns}\} \quad (*) \quad \begin{cases} h_1(x_1, x_2) = 0, \\ h_2(x_1, x_2) = 0 \end{cases}$$

of two simultaneous equations, where  $h_1$  and  $h_2$  are continuous real-valued functions of two real variables. Introduce the auxiliary map  $f$  defined by

$$f(x_1, x_2) = \langle h_1(x_1, x_2) + x_1, h_2(x_1, x_2) + x_2 \rangle.$$

A point  $x = \langle x_1, x_2 \rangle \in \mathbb{R}^2$  will be a solution of  $*$  if and only if it is a fixed-point of the map  $f$ . If now  $f(E) \subset E$  for some subset  $E$  of  $\mathbb{R}^2$  homeomorphic to  $D_2$ , then  $f$  will have a fixed-point  $x$  in  $E$ , and such  $x$  will be a solution of the original system  $*$ .

**Gauss, Karl Friedrich** [Example 5.14](#), a consequence of the [Intermediate-value Theorem](#), showed that a polynomial with real coefficients and of *odd* degree has at least one real root. The next theorem generalizes that result to all degrees and to complex as well as real coefficients.

thm:FTAlg **5.106 Fundamental Theorem of Algebra.** *A nonconstant polynomial with complex coefficients has at least one complex root.*

The Fundamental Theorem of Algebra was first proved by Karl Friedrich Gauss in his 1799 doctoral dissertation. A number of different proofs are possible—Gauss himself offered four—including one using Liouville’s theorem about “entire” complex functions (see Mathews and Howell [46, Theorem 6.19, pp. 252–253] or another book on complex analysis) and another using only minimal amounts of analysis and topology (see Lang [43, Chapter VIII, Section 3]). For an extended presentation of seven of the proofs, including a version of Gauss’s, see Fine and Rosenberger [24].

It might seem strange that topological considerations concerning the line, plane, and circle should serve to provide a proof of an algebraic result. The reasons they do are that the set  $\mathbb{C}$  of complex numbers is precisely the set of points of the plane  $\mathbb{R}^2$ , and the modulus  $|z|$  of a complex number  $z = x + yi$  is just the norm  $\|\langle x, y \rangle\|$  of the point  $\langle x, y \rangle$  in  $\mathbb{R}^2$ .

Along with the elementary facts about the algebra of complex numbers—see the subsection “Complex numbers” (page 89)—we shall need De Moivre’s Formula ([Proposition 0.88](#)), which may be reformulated in terms of the standard map  $q: \mathbb{R} \rightarrow S_1$  as

$$(q(s))^n = q(ns) \quad (n = 1, 2, 3, \dots)$$

for each real number  $s$ .

The Fundamental Theorem of Algebra applies, in particular, to polynomials all of whose coefficients are real. Even in this case, the theorem only guarantees only that there is some *complex* root, which need not be real. For example,  $x^2 + 1$  has real coefficients, yet its only roots are the nonreal complex numbers  $i$  and  $-i$ .

Notice that the Fundamental Theorem of Algebra is a pure *existence* result: it provides no method for actually finding any root of a polynomial, only the assurance that there is one! As noted earlier (see the discussion preceding [Example 5.14](#)), there are explicit “algebraic” formulas for finding roots of quadratic, cubic, and quartic equations. However, it can be proved that no algebraic formula exists for finding roots of polynomials that are quintic (degree 5)—see, for example, Birkhoff and Mac Lane [5, sec. 15.9]—or, more generally, of degree  $n$  when  $n \geq 5$ !

**Proof of the Fundamental Theorem of Algebra.** Let

$$f(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$$

be a nonconstant polynomial of degree  $n$  having complex coefficients  $a_0, a_1, \dots, a_n$ . The leading coefficient  $a_n \neq 0$  because  $f$  has degree  $n$ . Since  $f(z) = 0$  if and only if  $a_n^{-1} f(z) = 0$ , we may assume without loss of generality that  $a_n = 1$ .

We begin by comparing the behavior of the functions  $z \mapsto f(z)$  and  $z \mapsto z^n$  for large values of  $z$ , that is for complex numbers  $z$  for which  $|z|$  is large. When

$$|z| > c = \max\{1, |a_{n-1}| + |a_{n-2}| + \cdots + |a_1| + |a_0|\},$$

we have

$$\begin{aligned}
 |f(z) - z^n| &= |a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \cdots + a_1z + a_0| \\
 &\leq |a_{n-1}||z|^{n-1} + |a_{n-2}||z|^{n-2} + \cdots + |a_1||z| + |a_0| \\
 &= |z|^{n-1} \left( |a_{n-1}| + \frac{|a_{n-2}|}{|z|} + \cdots + \frac{|a_1|}{|z|^{n-2}} + \frac{|a_0|}{|z|^{n-1}} \right) \\
 &\leq |z|^{n-1} (|a_{n-1}| + |a_{n-2}| + \cdots + |a_1| + |a_0|) \\
 &< |z|^n = |z^n|,
 \end{aligned}$$

so that the line joining  $f(z)$  and  $z^n$  cannot pass through the origin. A fortiori,

$$\{eq:abs-z-big-then-f-nonzero\} \quad (*) \quad |z| > c \implies f(z) \neq 0.$$

Next we compare the behavior of  $z \mapsto f(z)$  and  $z \mapsto z^n$  on the circle  $\{rq(s) : s \in I\}$  in the complex plane  $\mathbb{C}$  having center at the origin and radius  $r > 0$ ; as usual,  $q: I \rightarrow S_1$  is the restriction of the standard map. The two maps

$$s \mapsto f(rq(s)), \quad s \mapsto (rq(s))^n = r^n q(ns)$$

from  $I$  to  $\mathbb{C} = \mathbb{R}^2$  are loops in  $\mathbb{C}$  at the points  $f(r)$  and  $r^n$ , respectively. The map

$$h: I \times I \rightarrow \mathbb{C}$$

given by

$$h(s, t) = tr^n q(ns) + (1-t)f(rq(s))$$

is a homotopy between these two maps satisfying

$$h(0, t) = h(1, t) = tr^n + (1-t)f(r) \quad (t \in I).$$

According to the preceding paragraph, moreover,

$$\{eq:h-nonzero-when-r-big\} \quad (**) \quad r > c \implies h(s, t) \neq 0 \text{ for all } s \in I, t \in I.$$

Now take any  $r > c$ . Then the loops in  $\mathbb{C}$  just considered may be transformed into loops  $\omega_r, \tau$  at  $1 = (1, 0)$  in  $S_1$  through the “normalizations”

$$\omega_r(s) = \frac{|f(r)|}{f(r)} \cdot \frac{f(rq(s))}{|f(rq(s))|}, \quad \tau(s) = q(ns).$$

Then

$$\omega_r \sim \tau.$$

In fact, in view of **\*\*** we obtain a path-homotopy  $H: \omega_r \sim \tau$  by setting

$$H(s, t) = \frac{|h(0, t)|}{h(0, t)} \cdot \frac{h(s, t)}{|h(s, t)|}.$$

For the first time now, just suppose that  $f(z) \neq 0$  for all  $z \in \mathbb{C}$ . Then the loop  $\omega_r$  is nullhomotopic for each  $r > 0$ , since a path-homotopy  $G$  of it to the null loop at  $(1, 0)$  in  $S_1$  is given by

$$G(s, t) = \frac{|f(tr)|}{f(tr)} \cdot \frac{f(trq(s))}{|f(trq(s))|}.$$

Since  $\omega_r \sim \tau$  for  $r > c$ , we conclude that  $\tau$  is also nullhomotopic. But this contradicts [Theorem 5.102](#).  $\square$

## EXERCISES FOR SECTION 5.5

retract!simply connected space@and simply connected space  
 prob:retract-simply-conn 105. Prove that a retract of a simply connected space must itself be simply connected.  
 homogeneous space!n-disk@and \$n\$-disk  
 n-disk@\$n\$-disk!homogeneous space@and homogeneous space  
 Note: A retract of a simply connected space is already path-connected: compare Exercise 92.

106. Give detailed justifications for the path-homotopy relations (\*\*) and (\*\*\*) in the proof of Proposition 5.96.

107. Describe the path  $q \circ \sigma$  in  $S_1$  where  $\sigma: \mathbb{I} \rightarrow \mathbb{R}$  is given by  $\sigma(s) = ns$  for a negative integer  $n$  and  $q: \mathbb{R} \rightarrow S_1$  is standard map.

108. Write an explicit algebraic formula for the map  $r: D_2 \rightarrow S_1$  in the proof of Theorem 5.105 and establish that  $r$  is, in fact, continuous.

109. The quadratic formula together with the fact that each complex number has two complex square-roots shows that every quadratic polynomial with real coefficients has a complex (possibly real) root. Without using Fundamental Theorem of Algebra prove now that every polynomial of odd degree with real coefficients has a real root. UH-OH: This was Example 5.14

every quadratic polynomial with real coefficients has a complex (possibly real) root.

prob-part:non-cst-poly-surj 110. (a) Prove that every non-constant polynomial function  $\mathbb{C} \rightarrow \mathbb{C}$  is surjective.  
 [Hint: Let  $K$  be the set of all critical points of such a polynomial function  $p$ , so that  $K$  is finite. Set  $B = \mathbb{C} \setminus p(K)$  and  $A = \mathbb{C} \setminus p^{-1}(p(K))$ . Show that the domain-codomain restriction  $r: A \rightarrow B$  of  $p$  is an open map. Deduce that  $r$  is also a closed map (since every polynomial is a proper map—the inverse image of each compact set is compact). (See Exercise 4.53.) Conclude that  $B$  is connected and hence that  $p$  is surjective. ]

(b) Give a new proof of FTA by deducing it from the result in (a).

111. (a) May the complement of a simple closed curve in the Möbius strip have just one component? two components?

(b) Assuming the Jordan Curve Theorem (5.98), what can you say about the complement of a simple closed curve in the 2-sphere?

112. Does Proposition 5.96 remain true if the hypothesis that  $X$  be path-connected is dropped (and the requirement of path-connectedness is omitted from the definition of simple connectedness)?

113. Prove that a path-connected space is simply connected precisely when each continuous map  $S_1 \rightarrow X$  has a continuous extension to  $D_2$ .

prob:D2-not-homogeneous 114. Prove that the 2-disk  $D_2$  is not homogeneous (Exercise 3.76).

Note: By contrast, the 2-ball  $B_2$ —and, in fact, the  $n$ -ball for every  $n \geq 1$ —is homogeneous: see Exercise 3.78.

(Hint: In fact, according to Exercise 3.78, for any two points  $u, v \in B_n$ , there is a homeomorphism of  $D_n$  sending  $u$  to  $v$ . Thus the proof requires showing that  $D_n$  has no homeomorphism of itself sending a point on  $S_{n-1}$  to a point in  $B_n$ .)

115. (a) something

(b) something

116. something

- 117. (a)** Give an example of a continuous map from the unit square  $J \times J$  to itself with the property that, for a given  $y \in J$  the map  $x \mapsto f(x, y)$  of  $J \rightarrow J$  map have more than one fixed-point.
- (b)** Let  $f$  be a continuous map from the unit square  $J \times J$  to itself. Suppose that, for each  $y \in J$  there is a *unique* fixed-point of the map  $x \mapsto f(x, y)$  of  $J \rightarrow J$ . Deduce that  $f$  must have a fixed-point.



## CHAPTER

# 6

## Embedding

chap:embedding

### Chapter Contents

6.1 The Tychonoff Product Theorem . . . . .	605	Exercises for Section 6.2 . . . . .	627
Proof of the Tychonoff Theorem		6.3 Embedding theorems . . . . .	631
using subbases . . . . .	606	Embedding in Tychonoff cubes . . . . .	631
Proof of the Tychonoff Theorem		Exercises for Section 6.3 . . . . .	637
using ultrafilters . . . . .	607	6.4 Urysohn Metrization Theorem . . . . .	638
Exercises for Section 6.1 . . . . .	610	Exercises for Section 6.4 . . . . .	638
6.2 Complete regularity and normality . . . . .	612	6.5 Compactification . . . . .	638
Completely regular spaces . . . . .	612	Exercises for Section 6.5 . . . . .	638
Normal spaces . . . . .	616	6.6 Dimension . . . . .	638
Urysohn's Lemma . . . . .	618	Exercises for Section 6.6 . . . . .	642
The Tietze Extension Theorem . . . . .	623		

### Introduction

The central theme in this chapter of advanced topics is to characterize when certain classes of topological spaces can be embedded in special kinds of spaces. Some of the “ambient” spaces into which the embeddings occur will be products of uncountable families of compact spaces, and so the chapter begins with proving the Tychonoff Theorem in its full generality, for arbitrary families of spaces. The most prominent class of spaces for which embedding will be established are the completely regular spaces, and so the second section studies completely regular spaces in depth; the same section resumes the study of normal spaces, and in Urysohn’s Lemma demonstrates that normal  $T_1$ -spaces are necessarily completely regular. Some of the ambient spaces will be metrizable, and so the fourth section establishes criteria for metrizability; notable among these criteria will be that in the Urysohn Metrization Theorem.

### 6.1 The Tychonoff Product Theorem

sec:tychonoff-proof

The following theorem, arguably the single most important one in topology, was stated in [Section 4.1](#) but not proved there.

**4.33 Tychonoff Product Theorem.** *The product of an arbitrary family of nonempty compact topological spaces is compact if and only if each of the factor spaces is compact.*

As is commonly done, we shall refer to this result more briefly as “*the Tychonoff Theorem*.” Our aim in this section is to prove it. We are going to offer two proofs, the first using subbases and the second using special kinds of filters called “ultrafilters.” Both proofs use a consequence of Axiom of Choice. In fact, every proof of the Tychonoff Theorem uses some consequence, or equivalent, of the Axiom of Choice, and most typically Zorn’s Lemma (0.115). And that is no accident: the Tychonoff Theorem is *equivalent* to the Axiom of Choice! [See also the comment “Tychonoff Product Theorem” (page 655) in the [Additional Readings](#).]

### Proof of the Tychonoff Theorem using subbases

subsec:Tychonoff-pf-subbases

According to the definition of compactness, a space is compact when each cover of it by members of its topology contains a finite cover. According to [Exercise 4.9](#), it suffices to consider covers by sets only from some base of the topology. And, in fact, it suffices to consider covers by sets only from some subbase.

thm:AlexanderSubbase

**6.1 Alexander Subbase Theorem.** *A topological space  $X$  is compact if it has some subbase  $\mathcal{S}$  with the property that each cover of  $X$  by members of  $\mathcal{S}$  contains a finite cover of  $X$ .*

**Proof.** Assume that  $X$  has such a subbase  $\mathcal{S}$ . Call an open cover of  $X$  *inadequate* if it contains *no* finite subcover of  $X$ . Just suppose that  $X$  is not compact. Then there is at least one inadequate open cover of  $X$ . If  $\beta$  is a chain of inadequate open covers of  $X$  (with open covers being ordered by subcollection inclusion), then  $\bigcup \beta$  is also an inadequate open cover of  $X$  that contains each member of the chain. By Zorn’s Lemma (0.115), there is a maximal inadequate open cover  $\mathcal{M}$  of  $X$ .

If  $\mathcal{S} \cap \mathcal{M}$  covered  $X$ , then it would have a finite subcover of  $X$ , contrary to the fact that  $\mathcal{M}$  is inadequate. Hence there is some point  $x \in X$  belonging to no member of  $\mathcal{S} \cap \mathcal{M}$ . The point  $x$  belongs to some member  $M$  of  $\mathcal{M}$ . Since  $\mathcal{S}$  is a subbase of  $X$ , it contains a finite subcollection  $\mathcal{V}$  for which  $x \in \bigcap \mathcal{V} \subset M$ .

Temporarily fix  $V \in \mathcal{V}$ . According to the choice of  $x$ , the set  $V$  is not a member of  $\mathcal{M}$ , so that  $\mathcal{M} \cup \{V\}$  strictly contains  $\mathcal{M}$ . By maximality of  $\mathcal{M}$ , the collection  $\mathcal{M} \cup \{V\}$ , which is still an open cover of  $X$ , *does* have a finite subcollection covering  $X$ . Thus there is a finite subcollection  $\mathcal{N}_V$  of  $\mathcal{M}$  for which  $\mathcal{N}_V \cup \{V\}$  is a finite cover of  $X$ .

It follows that  $\{\mathcal{M}\} \cup \bigcup_{V \in \mathcal{V}} \mathcal{N}_V$  is a finite cover of  $X$  each member of which belongs to  $\mathcal{M}$ . But this is impossible because  $\mathcal{M}$  is inadequate.  $\square$

Recall that the product topology on the product  $X = \prod_{i \in I} X_i$  of a family of spaces has as a subbase the collection

$$\{\text{eq: def-prod-subbase}\} \quad (*) \quad \mathcal{S} = \{p_j^{-1}(U) : j \in I \text{ and } U \text{ is open in } X_j\}$$

where, for each  $j \in I$ , the map  $p_j: \prod_{i \in I} X_i \rightarrow X_j$  is the  $j$ th projection. Accordingly, It is this subbase to which we shall apply the Alexander Subbase Theorem (6.1) in order to prove the Tychonoff Product Theorem.

**Proof of the Tychonoff Theorem—using Alexander Subbase Theorem.** Let  $X$  be the product  $\prod_{i \in I} X_i$  of a nonempty family of topological spaces. We already know that if  $X$  is compact, then each  $X_i$ , being the image of  $X$  under the continuous map  $p_i$ , is itself compact.



Conversely, suppose that  $X_i$  is compact for each  $i \in I$ . Let  $\mathcal{S}$  be the subbase of  $X$  given by equation (\*), above.. Let  $\mathcal{U}$  be a cover of  $X$  consisting solely of subbasic open sets. In view of the Alexander Subbase Theorem (6.1), it suffices to show that  $\mathcal{U}$  contains some finite cover of  $X$ .

For each  $j \in I$ , define

$$\mathcal{U}_j = \{U : U \text{ is open in } X_j \text{ and } p_j^{-1}(U) \in \mathcal{U}\}.$$

There is some  $j \in I$  for which  $\mathcal{U}_j$  covers  $X$ . In fact, if no such  $j$  exists, necessarily for each  $i \in I$  there is some point  $x_i \in X_i$  with  $x_i \notin \bigcup \mathcal{U}_i$ . Then the point  $\langle x_i \rangle_{i \in I}$  of  $X$  belongs to no member of  $\mathcal{U}$ , which is impossible since  $\mathcal{U}$  covers  $X$ .

For such  $j$ , there is a finite cover  $\mathcal{V}_j$  of  $X_j$  by members of  $\mathcal{U}_j$ . Then  $\{p_j^{-1}(V) : V \in \mathcal{V}_j\}$  is a finite cover of  $X$  by members of  $\mathcal{U}$ , as desired.  $\square$

Recall that a power  $Y^S$  of a set  $Y$  is the set  $\mathcal{F}(S, Y)$  of all maps from  $S$  to  $Y$ . When  $Y$  is a topological space, the product topology on the power  $Y^S$  is just the topology of pointwise convergence on  $\mathcal{F}(S, Y)$ . [See Examples 2.93 (2).]

In topology itself, the most important application of Tychonoff Product Theorem is to the powers of the unit interval  $I = [0, 1]$ .

**6.2 Definition.** A **Tychonoff cube** is any power  $I^S$  of the unit interval  $I = [0, 1]$  when provided with its product topology.

**6.3 Corollary.** Every Tychonoff cube is a compact Hausdorff space.

In fact, it was for this case alone that A. N. Tychonoff in 1930 originally stated and proved the theorem bearing his name; in 1937 Eduard Čech gave the first published proof for the general case.

### Proof of the Tychonoff Theorem using ultrafilters

The machinery of ultrafilters, developed in 1937 by H. Cartan, provides a different, deceptively simple, proof of the Tychonoff Product Theorem; this proof appears in Bourbaki [8, §9.5, p. 88] and is commonly known as “the Bourbaki proof.”

For the prerequisite material about filters, see subsections “Filters” (page 440), “Convergence of filters” (page 443), and “Comparison of filters” (page 448).

Among all filters on a given topological space, we distinguish a special class, namely, the “ultrafilters.”

**6.4 Definition.** An **ultrafilter** (or **universal filter**) on a set  $X$  is a filter that is maximal with respect to the partial ordering “finer than.” In other words, a filter  $\mathcal{U}$  on  $X$  is an ultrafilter when there is no filter on  $X$  that is strictly finer than  $\mathcal{U}$ .

**6.5 Example.** Given a point  $x$  in a set  $X$ , the principal filter  $\mathcal{U}_x$  generated by  $x$  [Examples 3.128 (3)] is an ultrafilter.

In fact, just suppose there is a filter  $\mathcal{F}$  that is strictly finer than  $\mathcal{U}$ . This means that  $\mathcal{U} \subset \mathcal{F}$  and there is some  $F \in \mathcal{F}$  with  $F \notin \mathcal{U}$ . The singleton  $\{x\}$  is a member of the principal filter  $\mathcal{U}$ , and so also  $\{x\} \in \mathcal{F}$ . Since then both  $F$  and  $\{x\}$  are members of  $\mathcal{F}$ , then

topology!pointwise convergence@o  
pointwise convergence  
Tychonoff, Andrey Nikolayevich  
Čech, Eduard  
Cartan, Henri  
Tychonoff Theorem!ultrafilters@and  
ultrafilter!Tychonoff Theorem@and  
ultrafilter!principal filter@and princi  
principal filter!ultrafilter@as ultrafilt

by property (F2) for filters, also  $\{x\} \cap F \in \mathcal{F}$ . A fortiori,  $\{x\} \cap F \neq \emptyset$ , that is,  $x \in F$ . But this means  $F \in \mathcal{U}$ , which is impossible.  $\diamond$

Why did we just give only one kind of example of an ultrafilter? Because principal filters are the *only* kind whose existence can be established without applying the Axiom of Choice!

The following theorem gives a criterion for a filter to be an ultrafilter that does not overtly involve the partial ordering of filters.

thm:uf-iff-set-or-compl-belongs

**6.6 Theorem.** A filter  $\mathcal{F}$  on a set  $X$  is an ultrafilter if and only if, for each subset  $A$  of  $X$ , either  $A \in \mathcal{F}$  or else  $X \setminus A \in \mathcal{F}$ .

**Proof.** Assume first that  $\mathcal{F}$  is an ultrafilter on  $X$ . Let  $A$  be an arbitrary subset of  $X$  and just suppose that  $X \setminus A \notin \mathcal{F}$ . If some  $F \in \mathcal{F}$  had empty intersection with  $A$  it would be contained in  $X \setminus A$ , and then its superset  $X \setminus A$  would be a member of  $\mathcal{F}$  after all. Thus each  $F \in \mathcal{F}$  intersects  $A$ . Consequently, the collection  $\{A \cap F : F \in \mathcal{F}\}$  is a filter base whose generated filter  $\mathcal{G}$  is finer than  $\mathcal{F}$ . Since  $\mathcal{F}$  is maximal among filters,  $\mathcal{G} = \mathcal{F}$ . Now  $A \in \mathcal{G}$ , and so  $A \in \mathcal{F}$  as well.

Conversely, assume that each subset of  $X$  or its complement belongs to a given filter  $\mathcal{F}$  on  $X$ . Just suppose there is some filter  $\mathcal{G}$  that is strictly finer than  $\mathcal{F}$ . There is some  $G \in \mathcal{G}$  with  $G \notin \mathcal{F}$ , and so  $X \setminus G \in \mathcal{F}$ . Then  $X \setminus G \in \mathcal{G}$ , so both  $G$  and  $X \setminus G$  belong to the filter  $\mathcal{G}$  and hence so does their intersection  $G \cap (X \setminus G)$ . But this is impossible because  $G \cap (X \setminus G) = \emptyset$ .  $\square$

The next two propositions will show some of the way in which ultrafilters are special among filters with respect to topologies. The first shows that, *for ultrafilters, clustering is equivalent to convergence*.

prop:uf-converge-if-cluster

**6.7 Proposition.** If an ultrafilter clusters at a point in a topological space, then it converges to that point.

**Proof.** Let  $\mathcal{U}$  be an ultrafilter on a topological space  $X$ . Assume that  $\mathcal{U}$  clusters at a point  $x \in X$ . Let  $V$  be an arbitrary neighborhood of  $x$  and just suppose there is no  $F \in \mathcal{U}$  with  $F \subset V$ . In particular,  $V \notin \mathcal{U}$ . From Theorem 6.6,  $X \setminus V \in \mathcal{U}$ . Since  $\mathcal{U}$  clusters at  $x$ , the neighborhood  $V$  of  $x$  must intersect the member  $X \setminus V$  of  $\mathcal{U}$ , and this is manifestly impossible.  $\square$

Next we prove the existence of non-principal ultrafilters. The proof uses Zorn's Lemma (0.115), an equivalent of the Axiom of Choice.

thm:uf-exists

**6.8 Theorem (existence of non-principal ultrafilters).** For each filter on a set there is a finer ultrafilter.

**Proof.** Let  $\mathcal{F}$  be a filter on a set  $X$ . Define  $\Phi$  to be the class of all filters on  $X$  that are finer than  $\mathcal{F}$ , partially ordered as usual by subcollection inclusion.

Let  $\Gamma$  be an arbitrary chain in  $\Phi$ . We show that  $\Gamma$  is bounded above in  $\Phi$ . Form the collection  $\mathcal{G} = \bigcup \Gamma$  of subsets of  $X$ , so that  $\mathcal{G}$  is a nonempty collection of nonempty subsets of  $X$ . If  $G_1, G_2 \in \mathcal{G}$ , there are filters  $\mathcal{G}_1, \mathcal{G}_2$  in  $\Gamma$  with  $G_1 \in \mathcal{G}_1$  and  $G_2 \in \mathcal{G}_2$ , respectively. Since  $\Gamma$  is totally ordered, one of  $\mathcal{G}_1$  and  $\mathcal{G}_2$  is contained in the other, say  $\mathcal{G}_1 \subset \mathcal{G}_2$ ; then  $G_1$  and  $G_2$  both belong to the filter  $\mathcal{G}_2$ , and so  $G_1 \cap G_2$  belongs to  $\mathcal{G}_2$ , too, and hence to  $\mathcal{G}$ . Further, if  $G \in \mathcal{G}$  and if  $A \subset X$  with  $G \subset A$ , then there is some member  $\mathcal{F}$  of  $\Gamma$  with

$G \in \mathcal{F}$ ; the set  $A$  is a member of the filter  $\mathcal{F}$  and hence of  $\mathcal{G}$ . This proves that  $\mathcal{G}$  is a filter. Clearly  $\mathcal{G}$  is finer than  $\mathcal{F}$ . Thus  $\mathcal{G}$  is the desired upper bound of  $\Gamma$  in  $\Phi$ .

From Zorn's Lemma it follows that  $\Phi$  has a maximal member  $\mathcal{U}$ , in other words,  $\mathcal{U}$  is an ultrafilter that is finer than the original  $\mathcal{F}$ .  $\square$

**6.9 Proposition.** A filter on a set  $X$  is the intersection of all ultrafilters on  $X$  that are finer than it.

**Proof.** Let  $\mathcal{F}$  be a filter on  $X$  and let  $\Phi$  be the class of all ultrafilters on  $X$  that are finer than  $\mathcal{F}$ . Certainly  $\mathcal{F} \subset \bigcap \Phi$ .

To establish the reverse inclusion, let  $A$  be a subset of  $X$  such that  $A \notin \mathcal{F}$ . We shall show there is some ultrafilter finer than  $\mathcal{F}$  to which  $A$  does not belong. Necessarily each  $F \in \mathcal{F}$  intersects  $X \setminus A$ . Then the collection  $\{F \cap (X \setminus A) : F \in \mathcal{F}\}$  is a filter base, and the filter  $\mathcal{F}'$  it generates is finer than  $\mathcal{F}$ . Moreover,  $X \setminus A \in \mathcal{F}'$ . By Theorem 6.8, there is an ultrafilter  $\mathcal{U}$  on  $X$  that is finer than  $\mathcal{F}'$ . Since  $X \setminus A \in \mathcal{F}'$ , then also  $X \setminus A \in \mathcal{U}$ . From Theorem 6.6,  $A \notin \mathcal{U}$ .  $\square$

**6.10 Lemma.** Let  $f: X \rightarrow Y$  be a map and let  $\mathcal{U}$  be an ultrafilter on the domain  $X$  of  $f$ .

- (1) The image  $f[\mathcal{U}]$  is an ultrafilter on the codomain  $Y$ .
- (2) If  $f$  is surjective, then  $f[\mathcal{U}] = \{f(U) : U \in \mathcal{U}\}$ , and  $f[\mathcal{U}]$  is an ultrafilter on  $Y$ .

**Proof.** (1) Recall that the image  $f[\mathcal{U}]$  of  $\mathcal{U}$  is the filter generated by the filter base  $\mathcal{B} = \{f(U) : U \in \mathcal{U}\}$ . Just suppose there is a filter  $\mathcal{F}$  on  $Y$  that is strictly finer than  $f[\mathcal{U}]$ . There is some member  $F \in \mathcal{F}$  such that  $f(U) \not\subset F$  for all  $U \in \mathcal{U}$ . Then  $U \not\subset f^{-1}(F)$  and so  $U \cap f^{-1}(F) \neq \emptyset$  for all  $U$  in  $\mathcal{U}$ . This means that the collection  $\{U \cap f^{-1}(F) : U \in \mathcal{U}\}$  is a filter base on  $X$ , and the filter  $\mathcal{F}'$  it generates is finer than  $\mathcal{U}$ . Since its member  $f^{-1}(F) = X \cap f^{-1}(F)$  does not belong to  $\mathcal{U}$ , this filter  $\mathcal{F}'$  is strictly finer than  $\mathcal{U}$ . But this is impossible.

- (2) The proof is left as an exercise (Exercise 16).  $\square$

**6.11 Proposition.** Let  $X$  and  $Y$  be topological spaces. Then a necessary and sufficient condition for a map  $f: X \rightarrow Y$  be continuous at a point  $x \in X$  is that for each ultrafilter  $\mathcal{U}$  on  $X$  converging to  $x$  in  $X$ , the image  $f[\mathcal{U}]$  of  $\mathcal{U}$  must converge to  $f(x)$  in  $Y$ .

**Proof.** Necessity of the condition follows from Theorem 3.143.

To prove sufficiency, assume the condition holds. Just suppose there is some neighborhood  $W$  of  $f(x)$  in  $Y$  for which  $f^{-1}(W)$  is not a neighborhood of  $x$  in  $X$ . Let  $V = f^{-1}(W)$ . By supposition,  $N \not\subset V$  for each neighborhood  $N$  of  $x$  in  $X$ , and so  $N \cap (X \setminus V) \neq \emptyset$  for each such  $N$ . Then the collection  $\{N \cap (X \setminus V) : N \in \mathcal{N}_x\}$  is a filter base on  $X$ ; let  $\mathcal{F}$  be the filter it generates. There is an ultrafilter  $\mathcal{U}$  on  $X$  that is finer than  $\mathcal{F}$ . If  $N$  is an arbitrary neighborhood of  $x$ , it contains the member  $N \cap (X \setminus V)$  of  $\mathcal{F}$ , which is also a member of  $\mathcal{U}$ . Thus  $\mathcal{U} \rightarrow x$  in  $X$ . By assumption  $f[\mathcal{U}] \rightarrow f(x)$  in  $Y$ . Consequently, the neighborhood  $W$  of  $f(x)$  contains  $f(U)$  for some  $U \in \mathcal{U}$ . This means that  $U \subset f^{-1}(W) = V$ , and so  $V \in \mathcal{U}$ . This is impossible because  $X \setminus V = X \cap (X \setminus V) \in \mathcal{U}$  and  $V \cap (X \setminus V) = \emptyset$ .  $\square$

compact space!filters@and filters  
filter!compact space@in compact space  
ultrafilter  
Tychonoff Theorem  
Tychonoff Theorem!Axiom of Choice@and Axiom of Choice

Recall from Theorem 4.43 that a topological space is compact if and only if each filter on it clusters in it. Since an ultrafilter converges to a point exactly when it clusters at that point (Proposition 6.7), we obtain the following criterion for compactness.

cor:Tychonoff Converge  
finite intersection p

**6.12 Corollary (ultrafilter characterization of compactness).** A topological space is compact if and only if each ultrafilter on it converges.

Tychonoff Theorem!two spaces@for two spaces  
Tychonoff Theorem!sequences of spaces@for sequence of spaces

Using ultrafilters now makes proving the Tychonoff Theorem almost trivial!

**Proof of the Tychonoff Theorem—using ultrafilters.** Apply the following three facts:

- a topological space is compact if each ultrafilter on it converges. (Corollary 6.12);
- the image of an ultrafilter under a surjection is itself an ultrafilter. (Lemma 6.10); and
- a filter on a product space converges if its image under each of the projections converges (Theorem 3.147).  $\square$

## EXERCISES FOR SECTION 6.1

1. The proof of the Tychonoff Product Theorem above that applied the Alexander Subbase Theorem implicitly used the Axiom of Choice. Make the use there explicit.

prob:WallaceThm

2. Let  $\langle X_i \rangle_{i \in I}$  be a family of topological spaces and let  $\langle K_i \rangle_{i \in I}$  be a family with  $K_i$  a compact subset of  $X_i$  for each  $i$ . Let  $V$  be a neighborhood of  $\times_{i \in I} K_i$  in  $\times_{i \in I} X_i$ . Show that there exists a family  $\langle W_i \rangle_{i \in I}$  of sets such that:  $W_i$  is open in  $X_i$  for each  $i \in I$ ;  $W_i = X_i$  for all except finitely many  $i \in I$  (so that  $\times_{i \in I} W_i$  is a basic set in  $\times_{i \in I} X_i$ ); and  $\times_{i \in I} K_i \subset \times_{i \in I} W_i \subset V$ .

3. Give a proof of the Tychonoff Product Theorem that uses the finite intersection property criterion for compactness (Theorem 4.9).

[Hint: Given a collection  $\mathcal{F}$  of closed subsets of the product space having the finite-intersection property, begin by obtaining a maximal collection of (not necessarily closed) subsets containing  $\mathcal{F}$ .]

4. *Note:* This problem furnishes an alternative proof of “Tychonoff-for-two” (Theorem 4.31) that suggests a proof for more general situations.

Let  $\mathcal{W}$  be a collection of open subsets of the product  $X \times Y$  of two compact spaces. Suppose that *no* finite subcollection of  $\mathcal{W}$  covers  $X \times Y$ . Deduce that  $\mathcal{W}$  cannot be a cover of  $X \times Y$  by establishing the following:

- (a) There is some  $x_0 \in X$  such that, for each neighborhood  $U$  of  $x_0$  in  $X$ , no finite subcollection of  $\mathcal{W}$  covers the tube  $U \times Y$ .  
(Hint: Assume the contrary and reach a contradiction.)
- (b) For such  $x_0$ , there is some  $y_0$  in  $Y$  such that, for each basic open neighborhood  $U \times V$  of  $\langle x_0, y_0 \rangle$  in  $X \times Y$ , no finite subcollection of  $\mathcal{W}$  covers  $U \times V$ .  
(Hint: Again, assume the contrary and reach a contradiction.)
- (c) The point  $\langle x_0, y_0 \rangle$  belongs to no member of  $\mathcal{W}$ .

5. Extend the reasoning used in Exercise 4 to show that the product of a sequence of compact spaces is compact.

(Note: This problem removes the metrizability assumption made in Theorem 4.61.)

prob:tychonoff-seq

- product-of-uncountably-many-seq-cpt
- of-uncountably-many-countably-cpt
- prob-part:uf-3-set-partition
- prob:free-fixed-filters
- prob:surj-image-uf-is-uf
- prob:ultranets
- prob-part:image-of-ultranet
- prob:ultranets-and-subnets
6. After well-ordering the index set, apply the method of [Exercise 5](#) to prove anew the general Tychonoff Product Theorem. Tychonoff Theorem!Well-ordering TH  
Well-ordering Theorem!Tychonoff TH
  7. Show that the product of uncountably many sequentially compact spaces need not be sequentially compact. Tychonoff Theorem  
ultrafilter
  8. Show that the product of uncountably many countably compact spaces need not be countably compact. ultrafilter  
filter!free  
free filter
  9. Prove that a filter  $\mathcal{U}$  on a set  $X$  is an ultrafilter if and only if each subset of  $X$  that intersects each member of  $\mathcal{U}$  belongs to  $\mathcal{U}$ . filter!fixed  
fixed filter  
ultrafilter
  10. Generalize [Theorem 6.6](#) by showing that a filter  $\mathcal{F}$  on a set  $X$  is an ultrafilter if and only if: for all subsets  $A$  and  $B$  of  $X$ , if  $A \cup B \in \mathcal{F}$ , then  $A \in \mathcal{F}$  or  $B \in \mathcal{F}$ . ultrafilter  
ultranet
  11. (a) Prove: A filter  $\mathcal{F}$  on a set  $X$  is an ultrafilter if and only if: for every finite partition  $\mathcal{A}$  of  $X$ , exactly one of the members of  $\mathcal{A}$  belongs to  $\mathcal{F}$ . ultrafilter!ultranet@and ultranet  
ultranet!ultrafilter@and ultrafilter
  - (b) Prove or disprove: A filter  $\mathcal{F}$  on a set  $X$  is an ultrafilter if and only if: for every partition  $\mathcal{A}$  of  $X$  into three sets, exactly one of the members of  $\mathcal{A}$  belongs to  $\mathcal{F}$ .
  - (c) Prove or disprove: A filter  $\mathcal{F}$  on a set  $X$  is an ultrafilter if and only if: for every partition  $\mathcal{A}$  of  $X$  into two sets, exactly one of the members of  $\mathcal{A}$  belongs to  $\mathcal{F}$ .
  12. (a) Prove that an ultrafilter on a finite set is necessarily principal.
  - (b) Prove, more generally, that an ultrafilter is principal if and only if some member of it is a finite set.
  13. A filter  $\mathcal{F}$  on a set  $X$  is said to be **free** when  $\bigcap \mathcal{F} = \emptyset$  and is said to be **fixed** otherwise. For example, the principal filter  $\mathcal{U}_x$  generated by an element  $x$  of  $X$  is fixed.
    - (a) Show that the Fréchet filter on  $\mathbb{N}$  is free.
    - (b) Show, more generally, that a filter  $\mathcal{F}$  on  $\mathbb{N}$  is free if and only if it is finer than the Fréchet filter.
  14. (Continuation of [Exercise 13](#).)
    - (a) Is the finite-complement filter [[Examples 3.128 \(6\)](#)] on an infinite set necessarily free ([Exercise 13](#))?
    - (b) Is the finite-complement filter on an infinite set necessarily an ultrafilter?
  15. Show that an ultrafilter having a minimal element with respect to set inclusion must be a principal ultrafilter for some element of the set.
  16. Prove part (2) of [Lemma 6.10](#).
  17. A net  $\xi$  in a set  $X$  is said to be an **ultranet** (or **universal net**) when, for each subset  $A$  of  $X$ , either  $\xi$  is eventually in  $A$  or  $\xi$  is eventually in  $X \setminus A$ . For example, a net  $(x_i)_{i \in I}$  that is eventually constant [[Examples 3.101 \(1\)](#)] is an ultranet; in particular, any constant net is an ultranet.
    - (a) Show that the net associated with an ultrafilter on a set  $X$  is an ultranet.
    - (b) Must the eventuality filter of an ultranet be an ultrafilter?
    - (c) Let  $f: X \rightarrow Y$  be a map and let  $\xi$  be an ultranet in  $X$ . Show that the net  $f \circ \xi$  in  $Y$  is also an ultranet.
  18. (Continuation of [Exercise 17](#).)

ultranet!compact space@and compact space (a) Show that a subnet of an ultranet is itself an ultranet.

ultranet!Tychonoff Theorem and Tychonoff Theorem (b) Must a net  $\xi$  in a set  $X$  be an ultranet if and only if it has no subnet other than itself?

ultranet!compact space@and compact space (c) Prove that every net has a subnet that is an ultranet.

ultranet

prob:ultranet-and-convergence 19. (Continuation of [Exercise 17](#).) Prove or disprove:

(a) An ultranet in a topological space converges to each point at which it clusters.

(b) If  $x$  is a point in the domain of a map  $f: X \rightarrow Y$  between topological spaces, then  $f$  is continuous at  $x$  if and only if, for each ultranet  $\xi$  converging to  $x$  in  $X$ , the net  $f \circ \xi$  converges to  $f(x)$  in  $Y$ . [Note: See [Exercise 17 \(c\)](#).]

prob:ultranets-and-cpt 20. (Continuation of [Exercise 17](#).)

Formulate and prove a criterion for compactness in terms of ultranets.

21. (Continuation of [Exercise 17](#).)

Give a proof of the Tychonoff Product Theorem that uses ultranets ([Exercise 17](#)) instead of ultrafilters.

prob:ultranet-re-Kelley-subnets 22. (Continuation of [Exercise 18](#).) Redo [Exercise 18](#) but for Kelley subnets ([Exercise 3.257](#)) instead of subnets in the original sense.

## 6.2 Complete regularity and normality

sec:separation-bis

Now we resume the study of the “higher” separation properties, complete regularity and normality.

### Completely regular spaces

subsec:compl-reg-bis

Recall from [Definition 2.94](#) that a space is said to be *completely regular* when each closed subset and each point not in that subset can be separated by a continuous real-valued function with values in  $[0, 1]$ . Equivalently, a space is completely regular when, for each open subset and each point in that subset, there is a continuous real-valued function on the space taking value 0 at the point and value 1 on the complement of the open subset.

Complete regularity, like the other separation properties, is a topological property.

The following proposition restates the result of [Exercise 1.82](#).

prop:metrizable-is-compl-reg **6.13 Proposition.** A metrizable space is completely regular  $T_0$ -space.

Already that proposition shows that examples of completely regular spaces abound. Later we shall see also that every locally compact  $T_2$ -space, and in particular every compact  $T_2$ -space, is completely regular.

In the definition of complete regularity, the unit interval  $[0, 1]$  may be replaced by any closed and bounded interval  $[a, b]$  or, for that matter, by  $\mathbb{R}$ .

prop:nasc-cr

**6.14 Proposition.** Each of the following is a necessary and sufficient condition for a space  $X$  to be completely regular.

prop-part:nasc-cr-i

(i) If  $[a, b]$  is a nondegenerate closed interval in  $\mathbb{R}$ , then for each point  $x \in X$  and each closed subset  $E$  of  $X$  with  $x \notin E$ , there is a continuous function  $f: X \rightarrow [a, b]$  such that  $f(x) = a$  and  $f(y) = b$  for every  $y \in E$ .

prop-part:nasc-cr-ii

(ii) For each point  $x \in X$  and each closed subset  $E$  of  $X$  with  $x \notin E$ , there is a continuous function  $f: X \rightarrow \mathbb{R}$  such that  $f(x) = 0$  and  $f(y) = 1$  for every  $y \in E$ .

**Proof.** (i) This condition is equivalent to complete regularity because: the linear function  $t \mapsto (b - a)t + a$  of  $[0, 1] \rightarrow [a, b]$  is a homeomorphism carrying 0 and 1 to  $a$  and  $b$ , respectively; and the inverse of that function is a homeomorphism carrying  $a$  and  $b$  to 0 and 1, respectively.

pf-part:max-min-convert-to-01

(ii) Clearly complete regularity implies (ii). Conversely, assume (ii) and let  $x$  be a point of  $X$  and  $E$  be a closed subset of  $X$  not containing  $x$ . Let  $f: X \rightarrow \mathbb{R}$  be a continuous function for which  $f(x) = 0$  and  $f(y) = 1$  for every  $y \in E$ . The function  $g: X \rightarrow [0, 1]$  given by

$$g(y) = \max\{0, \min\{f(y), 1\}\}.$$

is continuous according to Examples 3.64 (4). Finally,  $g(x) = 0$  and  $g(y) = 1$  for all  $y \in E$ .  $\square$

The following criterion for complete regularity is analogous to the criterion in Proposition 2.96 (1) for regularity.

prop:equivalent-regular

**6.15 Proposition.** A topological space  $X$  is completely regular if and only if for each point  $x$  of  $X$  and each open neighborhood  $U$  of  $x$ , there exists a continuous function  $f: X \rightarrow [0, 1]$  such that  $f(x) = 0$  and  $f(y) = 1$  for each  $y \in X \setminus U$ .

Again in the preceding criterion, any nondegenerate closed interval  $[a, b]$  or  $\mathbb{R}$  may be used in place of  $[0, 1]$ .

Complete regularity is relatively strong among the separation properties, according to the following result, which supplements Proposition 2.98.

prop:completely-regular-implies-regular

**6.16 Proposition.** (1) Every completely regular space is regular.

completely-regular-T0-implies-regular-T1

(2) Every completely regular  $T_0$ -space is a regular  $T_1$ -space and hence a  $T_2$ -space.

**Proof.** Let  $X$  be a completely regular space.

completely-regular-implies-regular

(1) We show that  $X$  must be regular. Let  $x$  be a point in  $X$  and let  $E$  be a closed subset of  $X$  such that  $x \notin E$ . By assumption, there is some continuous function  $f: X \rightarrow [0, 1]$  such that

$$f(x) = 0, \quad f(y) = 1 \text{ for all } y \in E.$$

Then the sets

$$U = f^{-1}([0, 1/2]), \quad V = f^{-1}([1/2, 1]).$$

are disjoint open subsets of  $X$  that are neighborhoods of  $x$  and  $E$ , respectively.



Mysior, Adam  
Mysior's space

- (2) Assume now that  $X$  is also  $T_0$ . To show that  $X$  must be  $T_1$ , we prove the stronger result that it is  $T_2$ . Let  $x, y \in X$  with  $x \neq y$ . Because  $X$  is  $T_0$ , there is an open neighborhood of  $x$  that does not contain  $y$  or an open neighborhood of  $y$  that does not contain  $x$ . Say there is an open neighborhood  $W$  of  $x$  that does not contain  $y$ . Then the set  $E = X \setminus W$  is a closed subset of  $X$  with  $x \notin E$ . Now argue as in the proof of part (1).  $\square$

Most regular spaces that one encounters are, in fact, completely regular. Nonetheless, the converse of the implication [Proposition 6.16 1](#) fails: **a regular space need not be completely regular**. Unfortunately, it is not so easy to construct counterexamples. The one that follows is due to Adam Mysior. (For further counterexamples, see the comment “Regular spaces that are not completely regular” ([page 656](#)) in the [Additional Readings](#).)

ex:Mysior-reg-not-cr

**6.17 Example (Mysior).** Let  $H = \mathbb{R} \times ]0, \infty[$ , the open upper half-plane in  $\mathbb{R}^2$ , let  $L = \mathbb{R} \times \{0\}$ , the  $x$ -axis, and let  $\bar{H} = H \cup L = \mathbb{R} \times [0, \infty[$ , the closed upper half-plane. Define  $X = \bar{H} \cup \{a\}$  to be the set obtained by adjoining a new point  $a$  to  $\bar{H}$ . Provide  $X$  with the topology for which:

- at each  $\langle x, y \rangle \in H$  the collection  $\{\{\langle x, y \rangle\}\}$  is a local base, that is,  $\langle x, y \rangle$  is isolated;
- at each  $\langle x, 0 \rangle \in L$ , a local base consists of all subsets of  $\{\langle x, 0 \rangle\} \cup I_x \cup J_x$  that include almost all points from  $I_x \cup J_x$ , where  $I_x$  and  $J_x$  are the vertical and  $45^\circ$ -slanted “whiskers” given by

$$I_x = \{\langle x, y \rangle : 0 \leq y < 2\},$$

$$J_x = \{\langle x + y, y \rangle : 0 \leq y < 2\},$$

respectively; and

- at  $a$  a local base consists of the sets  $V_n$  given by

$$V_n = \{a\} \cup \{\langle x, y \rangle : n < x, 0 \leq y < 2\} \quad (n = 1, 2, 3, \dots).$$

We shall call the resulting space  $X$  **Mysior's space**. This space is regular because each basic neighborhood of each point of  $\bar{H}$  is closed in  $X$  and  $\text{cls } V_{n+2} \subset V_n$  for each  $n$ .

Just suppose that  $X$  were completely regular. Then there must exist a continuous  $f: X \rightarrow [0, 1]$  with  $f(a) = 0$  and  $f(z) = 1$  for all  $z \in E$ , where  $E$  is the closed set given by

$$E = [0, 1] \times \{0\} = \{\langle x, 0 \rangle : 0 \leq x \leq 1\}.$$

Since  $f(a) < 1$ , there is some  $m$  for which  $f(V_m) \subset [0, 1[$ . Define

$$L_n = \{x : n - 1 \leq x \leq n \text{ and } f(x, 0) = 1\} \quad (n = 1, 2, 3, \dots).$$

We are going to show that  $L_n$  is infinite for every  $n$ . In particular,  $L_{m+1}$  will be infinite, and, *a fortiori*, there will be some  $x \in L_{m+1}$ . This will yield a contradiction because, on the one hand,  $\langle x, 0 \rangle \in V_m$  so that  $f(x, 0) = 1$  by definition of  $L_{m+1}$ , and, on the other hand,  $f(x, y) < 1$  since  $\langle x, 0 \rangle \in V_m$ .

We use induction to prove that  $L_n$  is infinite for each  $n$ . First,  $L_1 = [0, 1]$  because  $[0, 1] \subset E \subset f^{-1}(1)$ , and so  $L_1$  is infinite.

Now let  $n \geq 1$  and assume that  $L_n$  is infinite. Then  $L_n$  has a denumerable subset  $D$ .



Temporarily fix  $d \in D$ , so that  $\langle d, 0 \rangle \in L_n$ . By continuity of  $f$  at  $\langle d, 0 \rangle$ , for each  $k = 1, 2, 3, \dots$ , the set

$$G_k = \{z \in X : |f(z) - 1| < 1/k\}$$

is an open neighborhood of  $\langle d, 0 \rangle$  in  $X$ , and so  $J_d \setminus G_k$  is a *finite* set not containing  $\langle d, 0 \rangle$ . Since  $f^{-1}(1) = \bigcap_{k=1}^{\infty} G_k$ , its complement

$$J_d \setminus f^{-1}(1) = \bigcup_{k=1}^{\infty} (J_d \setminus G_k)$$

in  $J_d$  is *countable*.

Since  $D$  is denumerable, the set

$$C = \bigcup_{d \in D} (J_d \setminus f^{-1}(1))$$

is countable, and so its projection

$$P = \{\langle x, 0 \rangle : \langle x, y \rangle \in C \text{ for some } y\}$$

onto the  $x$ -axis is also countable. It follows that  $L_{n+1} \setminus P$  is infinite. This concludes the induction.  $\diamond$

What do the standard constructions of new spaces from old do to complete regularity? The following proposition—which supplements [Proposition 2.102](#), [Proposition 3.49](#), and [Theorem 3.68](#)—together with the example following it provides an answer.

**6.18 Proposition.** (1) A subspace of a completely regular space is itself completely regular.

(2) The product of an arbitrary family of topological spaces is completely regular if and only if each member of the family is completely regular.

(3) The sum of two completely regular spaces is itself completely regular.

**Proof.** (1) Let  $F$  be a closed subset and  $x$  a point in a subspace  $Y$  of a completely regular space  $X$  with  $x \notin F$ . There is a closed set  $E$  in  $X$  with  $F = E \cap Y$ . Now apply complete regularity of  $X$  to  $x$  and  $F$  and restrict the function obtained.

(2) Let  $X = \langle X_i \rangle_{i \in I}$  be the product of a family of completely regular spaces. Let  $F$  be a closed subset of  $X$ , and let  $x \in X$  with  $x \notin F$ . There is a basic open neighborhood  $U = \times_{i \in I} U_i$  of  $x$  in  $X$  that is disjoint from  $F$ . There is a finite subset  $J$  of  $I$  such that  $U_i = X_i$  for all  $i \notin J$ . For each  $j \in J$  there is a continuous function  $f_j: X_j \rightarrow [0, 1]$  such that

$$g_j(x_j) = 0, \quad g_j(y) = 1 \quad (y \in X_j \setminus U_j).$$

For each  $j \in J$ , define  $f_j = g_j \circ p_j$ , where  $p_j: X \rightarrow X_j$  is the  $j$ th projection. Then the formula

$$f(y) = \prod_{j \in J} g_j(y) \quad (y \in X)$$

defines a function  $f: X \rightarrow [0, 1]$  that separates  $x$  from  $E$ .

The converse is left as an exercise.

(3) Let  $E$  be a closed subset and  $x$  be a point of the Cartesian sum  $X + Y$  of two completely regular spaces with  $x \notin E$ . Without loss of generality we may assume that  $X$  and  $Y$

line with two origins!separation properties@and separation properties  
 line with two origins!quotient space@as quotient space  
 normal space  
 locally compact space!normal space@and normal space  
 normal space!locally compact space@and normal space  
 completely regular space!normal space@and normal space  
 normal space!completely regular space@and normal space  
 regular space!normal space@and normal space  
 normal space!regular space@and regular space

are disjoint and that  $x \in X$ . Now  $E = F + K$  for closed subsets  $F$  and  $K$  of  $X$  and  $Y$ , respectively. There is a continuous function  $g: X \rightarrow [0, 1]$  with  $g(x) = 0$  and  $g(u) = 1$  for all  $u \in F$ . Then the extension  $f: X + Y \rightarrow [0, 1]$  with  $f(y) = 0$  for all  $y \in Y$  is continuous and separates  $x$  from  $E$  in  $X + Y$ .  $\square$

As the following example shows, a quotient of a completely regular space—even a  $T_1$ -quotient—need not be completely regular!

ex:non-cr-quot-cr **6.19 Example.** The line with two origins  $Y$  [Examples 2.20 (3)] is a  $T_1$ -space that is a quotient of a completely regular  $T_0$ -space but is *not* completely regular. In fact, according to Examples 3.81 (8),  $Y$  is a continuous open image of the product  $\mathbb{R} \times \{0, 1\}$ , where  $\{0, 1\}$  has its discrete topology; as a subspace of  $\mathbb{R} \times \mathbb{R}$ , the space  $\mathbb{R} \times \{0, 1\}$  is metrizable and hence is a completely regular  $T_0$ -space. Now  $Y$  is a  $T_1$ -space that is not a  $T_2$ -space [Examples 2.99 (3)]. In view of Proposition 6.15,  $Y$  cannot be completely regular.  $\diamond$

The quotient space in the preceding example is a  $T_1$ -space that is not a  $T_2$ -space. It is even possible for the continuous *closed* image of a completely regular  $T_0$ -space to be a  $T_2$ -space yet fail to be completely regular: see Exercise 35.

## Normal spaces

subsec:normal

Recall from Definition 2.94 that a topological space is said to be **normal** when each two disjoint closed subsets have disjoint neighborhoods. When a space is normal, disjoint closed subsets will in fact have disjoint *open* neighborhoods.

The class of normal spaces is wide in that it includes:

- all metrizable spaces (Proposition 2.100);
- all compact Hausdorff spaces (Theorem 4.20); and, as we are about to prove,
- all second-countable regular spaces (Theorem 6.20, below).

But, as we shall see in Example 6.23, a locally compact Hausdorff space need not be normal, and a regular, or even completely regular, space need not be normal.

thm:regular-2nd-count-then-normal

**6.20 Theorem.** A second-countable regular space is normal.

**Proof.** Let  $E$  and  $F$  be disjoint closed subsets of a second-countable regular space  $X$ . By regularity, for each  $x \in E$  there are disjoint open neighborhoods  $M_x$  and  $N_x$  of the point  $x$  and the set  $F$ , respectively. For each  $x \in E$ ,

$$M_x \subset \text{cls } M_x \subset X \setminus N_x \subset X \setminus F.$$

The collection  $\{M_x : x \in E\}$  is an open cover of  $E$  in  $X$ . Now  $E$  is second-countable because  $X$  is, and so by the Lindelöf Theorem (2.84) there is a sequence  $\langle U_n \rangle_{n \in \mathbb{N}}$  of open sets such that

$$E \subset \bigcup_{n \in \mathbb{N}} U_n, \quad F \subset X \setminus \text{cls } U_n \text{ for each } n \in \mathbb{N}.$$

Similarly, there is a sequence  $\langle V_n \rangle_{n \in \mathbb{N}}$  of open sets such that

$$F \subset \bigcup_{n \in \mathbb{N}} V_n, \quad E \subset X \setminus \text{cls } V_n \text{ for each } n \in \mathbb{N}.$$

Although the sets  $\bigcup_{n \in \mathbb{N}} U_n$  and  $\bigcup_{n \in \mathbb{N}} V_n$  are disjoint open neighborhoods of  $E$  and  $F$ , respectively, they need not be disjoint. We shall remove closed portions of each  $U_n$  and

each  $V_n$  so as to leave sequences of open sets whose unions are still neighborhoods of  $E$  and  $F$  but that are now disjoint. For each  $n = 0, 1, 2, \dots$ , define

$$G_n = U_n \setminus \bigcup_{j=0}^n \text{cls } V_j = U_n \cap \bigcap_{j=0}^n (X \setminus \text{cls } V_j),$$

$$H_n = V_n \setminus \bigcup_{j=0}^n \text{cls } U_j = V_n \cap \bigcap_{j=0}^n (X \setminus \text{cls } U_j),$$

so that  $G_n$  and  $H_n$  are open sets. Then the sets  $\bigcup_{m \in \mathbb{N}} G_m$  and  $\bigcup_{n \in \mathbb{N}} H_n$  are also open. These two sets contain  $E$  and  $F$  respectively, because

$$E \subset \bigcup_{n=0}^{\infty} \left( U_n \cap \bigcap_{j=0}^n (X \setminus \text{cls } V_j) \right) = \bigcup_{n \in \mathbb{N}} G_n,$$

and similarly for  $F$ . Finally, for  $m \geq n$  the sets  $G_m$  and  $H_n$  are disjoint, because  $G_m \cap H_n \subset (X \setminus \text{cls } V_n) \cap V_n = \emptyset$ , and similarly for  $n \geq m$ ; thus  $\bigcup_{m \in \mathbb{N}} G_m$  and  $\bigcup_{n \in \mathbb{N}} H_n$  are disjoint.  $\square$

For reference below, we recall the normality criterion from [Proposition 2.96 \(2\)](#).

**6.21 Proposition.** *A topological space  $X$  is normal if and only if for each closed subset  $E$  of  $X$  and each neighborhood  $U$  of  $E$  in  $X$ , there is an open neighborhood  $V$  of  $E$  with  $\text{cls}(V) \subset U$ .*

Next we consider which subspaces of normal spaces are themselves normal. The proof of the following proposition is straightforward and is left to the reader ([Exercise 26](#)).

**6.22 Proposition.** *A closed subspace of a normal space is itself normal.*

Most constructions of new spaces from old ones fail to preserve normality. The next example shows that ***an open subspace of a normal  $T_1$ -space need not itself be normal.***

**6.23 Example.** The **Tychonoff plank** is the product

$$P = \Omega^+ \times \omega^+,$$

or the ordinal spaces  $\Omega^+ = [0, \Omega]$  and  $\omega^+ = [0, \omega]$  [see [Examples 2.72 \(4\)](#)]. From [Examples 4.6 \(9\)](#) and “[Tychonoff-for-two](#)” ([Theorem 4.31](#)), the Tychonoff plank is compact. In addition, this space is a Hausdorff space, because by [Examples 2.72 \(1\)](#) each of its factors is a Hausdorff space. Hence ***the Tychonoff plank  $P$  is normal.***

The “**chipped**” Tychonoff plank

$$P^* = (\Omega^+ \times \omega^+) \setminus \{\langle \Omega, \omega \rangle\}$$

is formed by removing the “upper-right corner”  $\langle \Omega, \omega \rangle$  from  $P$ . Since it is an open subspace of  $P$ , the chipped Tychonoff plank  $P^*$  is a locally compact Hausdorff space. However, as we are about to see, ***the chipped Tychonoff plank is not normal.*** Thus ***a locally compact Hausdorff space need not be normal.***

Being a compact  $T_2$ -space, the Tychonoff plank  $P$  is a completely regular  $T_0$ -space; then its subspace  $P^*$  is also a completely regular  $T_0$ -space. Thus ***a completely regular  $T_0$ -space need not be normal.***

normal space!subspace@and subspa  
subspace!normal space@of normal s  
Tychonoff plank  
Tychonoff, Andrey Nikolayevich  
chipped Tychonoff plank  
Tychonoff plank!chipped  
locally compact space!normal space  
normal space!locally compact space  
completely regular space!normal spa  
normal space!completely regular spa

prop:normal-characterization-via-cls

prop:closed-subspace-of-normal

ex:Tychonoff-plank

normal space!subspace@and subspace  
subspace!normal space@of normal space  
edges

To see that the chipped Tychonoff plank is not normal, consider the “top” and “right-hand” edges

$$E = [0, \Omega] \times \{\omega\},$$

$$F = \{\Omega\} \times [0, \omega]$$

of the Tychonoff plank  $P$ . These are disjoint subsets of the chipped Tychonoff plank  $P^*$ , and they are closed in  $P^*$  because their complements in  $P^*$  are open in  $P$  and hence in  $P^*$ . We claim that  $E$  and  $F$  do *not* have disjoint neighborhoods in  $P^*$ .

To see why, let  $V$  be any neighborhood of  $F$  in  $P^*$ . For each  $n \in [0, \Omega[$ , the set  $V$  is a neighborhood of the point  $\langle \Omega, n \rangle$  of  $F$ , and so there is some  $\alpha_n \in [0, \Omega[$  with

$$]\alpha_n, \Omega] \times \{n\} \subset V.$$

From [Proposition 0.113](#), the countable subset  $\{\alpha_n : n \in [0, \Omega[$  of  $[0, \Omega[$  has an upper bound  $\beta$  in  $[0, \Omega[$ . Then

$$]\beta, \Omega] \times [0, \omega[ \subset V.$$

Then each neighborhood  $U$  in  $P^*$  of the point  $\langle \beta, \omega \rangle$  of  $E$  must intersect  $V$ , and so there is no neighborhood of  $E$  that is disjoint from  $V$ .  $\diamond$

**The product of even two normal spaces need not itself be normal!** In fact, as [Examples 6.29](#) and [6.32](#) together will show, the product of two normal  $T_1$ -spaces can be completely regular  $T_0$  and yet be nonnormal.

By contrast, normality behaves nicely with respect to Cartesian sums.

prop:normal-sum-of-2-normal **6.24 Proposition.** *The Cartesian sum of two normal spaces is itself normal.*

For a generalization, see [Exercise 32](#).

**A quotient space of a normal space need not be normal!** In fact, a *nonnormal*  $T_1$ -space can be the quotient of a normal  $T_1$ -space, as the next example shows.

ex:nonnormal-quot-normal-T1 **6.25 Example.** The line with two origins [[Examples 2.20 \(3\)](#)] is a  $T_1$ -space that is not a  $T_2$ -space and, *a fortiori*, is *not* normal. According to [Examples 3.81 \(8\)](#), this space is a quotient of the product  $\mathbb{R} \times \{0, 1\}$  of the real line with its discrete subspace  $\{0, 1\}$ . The latter space  $\mathbb{R} \times \{0, 1\}$  is metrizable and therefore, by [Proposition 2.100](#), a normal  $T_1$ -space.  $\diamond$

### Urysohn's Lemma

subsec:urysohn

Recall that:

- a space is regular when neighborhoods separate each closed subset from each point not in the subset;
- a space is completely regular when continuous functions separate each closed subset from each point not in the subset; and
- whereas a completely regular space is necessarily regular, a regular space need not be completely regular.

Now:

- a space is normal when neighborhoods separate disjoint closed subsets.

Analogously to the situation with regularity and complete regularity, we may want to introduce the notion that. . .

- a space is “utterly normal”<sup>1</sup> when continuous functions separate disjoint closed subsets. . .

. . . and then expect that:

- whereas an utterly normal space is necessarily normal, a normal space need not be utterly normal.

It turns out, however, that no such stronger property of “utterly normality” is needed: a normal space necessarily has this property! That is the content of the next result, [Urysohn’s Lemma](#). Although it has the designation “lemma” because it is a preliminary to proving the Tietze Extension Theorem (6.34), the result is a major theorem in topology. And it is remarkable because it says that, from a purely “internal” property of a topological space—one concerning solely the points and open and closed subsets in the space—there arises an “external” property involving the real numbers.

lem:urysohn **6.26 Urysohn’s Lemma.** *Let  $A$  and  $B$  be disjoint closed subsets of a normal topological space  $X$ . Then there exists a continuous function  $f: X \rightarrow [0, 1]$  such that*

$$f(x) = 0 \text{ for all } x \in A, \quad f(y) = 1 \text{ for all } y \in B.$$

Such a function  $f$  is said to be a **Urysohn function for  $A$  and  $B$  (in  $X$ )** and is said to **separate  $A$  and  $B$** . Thus Urysohn’s Lemma asserts that any two disjoint closed subsets of a normal space can be separated by a continuous real-valued function.

In the statement of Urysohn’s Lemma, the interval  $[0, 1]$  as codomain may be replaced by any nondegenerate closed interval  $[a, b]$ , with  $f$  taking value  $a$  on  $A$  and  $b$  on  $B$ .

Actually, the property that continuous functions separate disjoint closed subsets characterizes normal spaces.

prop:normal-iff-fns-separate-closed **6.27 Proposition.** *A necessary and sufficient condition that a topological space be normal is that for each pair of its disjoint closed subsets there exists a Urysohn function.*

**Proof.** Necessity is just Urysohn’s Lemma, which is about to be proved. The proof of sufficiency is left to the reader ([Exercise 46](#)). □

**Proof of Urysohn’s Lemma.** Our strategy is to begin by defining  $f$  to take values 0 on  $A$  and 1 on  $B$  and then step-by-step extend  $f$  to larger and larger sets containing  $A$ , with  $f$  taking successively larger and larger values, while also extending  $f$  to larger and larger sets containing  $B$ , with  $f$  taking successively smaller and smaller values. In order to obtain those sets, we shall repeatedly invoke the normality criterion from [Proposition 2.96 \(2\)](#).

<sup>1</sup>We use the temporary term “utterly normal” in order to distinguish it from the term “completely normal.” The latter term appears in the literature with two quite different meanings, namely: (i) each subspace of the space is normal; and (ii) the space is a normal  $T_1$ -space.

To begin, there are open subsets  $U_0$  and  $U_1$  of  $X$  with

$$A \subset U_0 \subset \text{cls } U_0 \subset X \setminus B, \quad \text{cls } U_0 \subset U_1 \subset \text{cls } U_1 \subset X \setminus B,$$

so that

$$A \subset U_0 \subset \text{cls } U_0 \subset U_1 \subset \text{cls } U_1 \subset X \setminus B.$$

Since  $\text{cls } U_0 \subset U_1$ , there is an open subset  $U_{1/2}$  of  $X$  with

$$\text{cls } U_0 \subset U_{1/2} \subset \text{cls } U_{1/2} \subset U_1,$$

so that

$$A \subset U_0 \subset \text{cls } U_0 \subset U_{1/2} \subset \text{cls } U_{1/2} \subset U_1 \subset \text{cls } U_1 \subset X \setminus B.$$

(We chose the subscript  $1/2$  because that is a number between 0 and 1, just as the closure of this new set  $U_{1/2}$  falls “between”  $U_0$  and  $U_1$ .) Next, since  $\text{cls } U_0 \subset U_{1/2}$  and  $\text{cls } U_{1/2} \subset U_1$ , there are open subsets  $U_{1/4}$  and  $U_{3/4}$  with

$$\text{cls } U_0 \subset U_{1/4} \subset \text{cls } U_{1/4} \subset U_{1/2}, \quad \text{cls } U_{1/2} \subset U_{3/4} \subset U_1,$$

so that

$$U_0 \subset \text{cls } U_0 \subset U_{1/4} \subset \text{cls } U_{1/4} \subset U_{1/2} \subset \text{cls } U_{1/2} \subset U_{3/4} \subset \text{cls } U_{3/4} \subset U_1.$$

(As before, so now we chose the intermediate subscripts  $1/4$  and  $3/4$  for the way the new sets fall “between” the previous ones.)

We want to continue in this way so as to obtain a family  $\langle U_q \rangle_{q \in \mathcal{A}}$  of open subsets of  $X$  indexed by the set

$$\mathcal{A} = \left\{ \frac{m}{2^n} : n \in \mathbb{N}, m = 0, 1, 2, \dots, 2^n \right\}$$

of dyadic rationals in  $[0, 1]$ —see [Example 0.82](#)—such that:

$$\begin{aligned} A &\subset U_0, & B &\subset X \setminus U_1, \\ \text{cls } U_q &\subset U_r \text{ for all } q, r \in \mathcal{A} \text{ with } 0 \leq q < r \leq 1. \end{aligned}$$

In order to carry out the recursion, we introduce the sets

$$\mathcal{A}_n = \left\{ \frac{m}{2^n} : m = 0, 1, 2, \dots, 2^n \right\}$$

for all  $n \in \mathbb{N}$ . Thus the sequence  $\langle \mathcal{A}_n \rangle_{n \in \mathbb{N}}$  satisfies

$$\begin{aligned} \{0, 1\} &= \mathcal{A}_0 \subset \mathcal{A}_1 \subset \dots \subset \mathcal{A}_n \subset \mathcal{A}_{n+1} \subset \dots, \\ \mathcal{A} &= \bigcup_{n=0}^{\infty} \mathcal{A}_n. \end{aligned}$$

As before, we begin the recursion by choosing open sets  $U_0$  and  $U_1$  with

$$\{eq:U0-U1-for-urysohn\} \quad (0) \quad A \subset U_0 \subset \text{cls } U_0 \subset X \setminus B, \quad \text{cls } U_0 \subset U_1 \subset \text{cls } U_1 \subset X \setminus B.$$

Thus we have a family  $\langle U_q \rangle_{q \in \mathcal{A}_0}$  such that

$$\text{cls } U_q \subset U_r \text{ whenever } q, r \in \mathcal{A}_0 \text{ with } q < r.$$

Now let  $n > 0$  and suppose we have obtained a family  $\langle U_q \rangle_{q \in \mathcal{A}_n}$  of open sets satisfying

$$\text{cls } U_q \subset U_r \text{ whenever } q, r \in \mathcal{A}_n \text{ with } q < r.$$

In order to obtain the next family  $\langle U_q \rangle_{q \in \mathcal{A}_{n+1}}$ , we need only obtain the new sets indexed by elements of  $\mathcal{A}_{n+1} \setminus \mathcal{A}_n$ , namely, for each  $k = 0, 1, 2, \dots, 2^k - 1$  the index  $(2k + 1)/2^{n+1}$

from  $\Delta_{n+1}$  midway between the consecutive indices  $k/2^n$  and  $(k+1)/2^n$  from  $\Delta_n$ . For each  $k = 0, 1, 2, \dots, 2^k - 1$ , there is an open subset  $U_{(2k+1)/2^{n+1}}$  of  $X$  with

$$\text{cls } U_{k/2^n} \subset U_{(2k+1)/2^{n+1}} \subset \text{cls } U_{(2k+1)/2^{n+1}} \subset U_{(k+1)/2^n}.$$

It now follows that

$$\{\text{eq:U-rel-Delta-for-urysohn}\} \quad (1) \quad \text{cls } U_q \subset U_r \text{ whenever } q, r \in \Delta_{n+1} \text{ with } q < r.$$

With the family  $\langle U_q \rangle_{q \in \Delta}$  in hand, we are going to define a function  $f: X \rightarrow [0, 1]$  in such a way that

$$f(x) \leq q \text{ whenever } x \in U_q.$$

To accomplish that, we set

$$f(x) = \begin{cases} 1 & \text{if } x \notin U_1, \\ \inf\{q \in \Delta : x \in U_q\} & \text{if } x \in U_1. \end{cases}$$

Already  $f(x) = 0$  if  $x \in A$  and  $f(y) = 1$  if  $y \in B$ .

Finally, we show that  $f$  is continuous. From the definition of  $f$ :

$$\{\text{ineq-urysohn-a}\} \quad (2) \quad \text{If } x \notin \text{cls } U_q, \text{ then } f(x) \geq q.$$

$$\{\text{ineq-urysohn-b}\} \quad (3) \quad \text{If } x \notin \text{cls } U_q, \text{ then } f(x) \geq q.$$

Let  $x \in X$ . We show that  $f$  is continuous at  $x$ . Let  $\varepsilon > 0$  be arbitrary. Now consider several cases.

$\text{urysohn-pf-cont-case-i}$  Case (i):  $f(x) = 0$ : In this case, by density of  $\Delta$  in  $[0, 1]$ , there is some  $q \in \Delta$  with  $0 < q < \varepsilon$ . Then  $U_q$  is a neighborhood of  $x$  for which  $f(U_q) \subset [0, \varepsilon]$ .

Case (ii):  $f(x) = 1$ : This case is similar to case [Case \(i\)](#).

Case (iii):  $0 < f(x) < 1$ : In this case, there are some  $q, r \in \Delta$  with

$$f(x) - \varepsilon < q < f(x) < r < f(x) + \varepsilon.$$

Then  $U_r \setminus \text{cls } U_q$  is a neighborhood of  $x$  for which

$$f(U_r \setminus \text{cls } U_q) \subset (f(x) - \varepsilon, f(x) + \varepsilon). \quad \square$$

The following immediate consequence of Urysohn's Lemma completes the list of implications among the separation properties as enumerated in [Proposition 2.98](#) and displayed in [Tables 2.1](#) and [2.2](#).

$\text{cor:normal-T1-cr}$  **6.28 Corollary.** A normal  $T_1$ -space is completely regular.

$\text{ex:Sorgenfrey-plane-cr}$  **6.29 Example.** The Sorgenfrey line  $\mathbb{R}_l$  [[Examples 2.20 \(1\)](#)] is a normal  $T_1$ -space: see [Examples 2.97 \(3\)](#). From the preceding corollary, the Sorgenfrey line is, *a fortiori*, a completely regular  $T_0$ -space. By [Proposition 6.18](#), the Sorgenfrey plane—the product  $\mathbb{R}_l \times \mathbb{R}_l$  of the Sorgenfrey line  $\mathbb{R}_l$  with itself [[Examples 2.72 \(7\)](#)—is therefore a completely regular  $T_0$ -space. However, the Sorgenfrey plane is *not* normal: see either [Exercise 2.154](#) or [Example 6.32](#).  $\diamond$

normal space!discr  
thm:loc-cnt-T2-cr  
normal space!sepa

**6.30 Theorem.** A locally compact  $T_2$ -space is completely regular.

separable space!normal space@and normal space  
normal space!cardinality@and cardinality  
Jones, F. Burton

Sorgenfrey plane!normal space@and normal space  
product space!normal space@and product space

normal space!product space@and product space  
normal space!Sorgenfrey plane@and Sorgenfrey plane

normal space!discrete subspace@and discrete subspace

normal space!separable space@and separable space

separable space!normal space@and normal space

normal space!cardinality@and cardinality  
Urysohn's lemma

**Proof.** Let  $X$  be a locally compact  $T_2$ -space. If  $X$  is already compact, then it is normal and hence, by Corollary 6.28, completely regular. Now assume that  $X$  is noncompact. The one-point compactification  $\alpha(X)$  (Definition 4.79) is a compact  $T_2$ -space, hence completely regular. Being homeomorphic to a subspace of  $\alpha(X)$ , the space  $X$  itself is completely regular.  $\square$

One application of Urysohn's Lemma, due to F. Burton Jones, is to establish a constraint upon the cardinality of certain subspaces of normal spaces.

thm:jones

**6.31 Theorem (Jones's Lemma).** Let  $D$  be a discrete closed subspace of a separable normal space  $X$ . Then

$$\text{card } D < \text{card } \mathbb{R}.$$

**Proof.** For each subset  $E$  of  $D$ , the sets  $E$  and  $D \setminus E$  are disjoint closed subsets of  $X$ , and so by Urysohn's Lemma we may choose some continuous function  $f_E: X \rightarrow \mathbb{R}$  for which  $f_E(x) = 1$  if  $x \in E$  and  $f_E(x) = 0$  if  $x \in D \setminus E$ . Then the map

$$\begin{aligned} \mathcal{P}(D) &\rightarrow C(X) \\ E &\mapsto f_E \end{aligned}$$

is injective. Hence

$$\{\text{eq:jones-lem-ineq-1}\} \quad (*) \quad \text{card } 2^D \leq \text{card } C(X).$$

Now choose some countable dense subset  $G$  of  $X$ . By extension of identities (Theorem 3.12), the map

$$\begin{aligned} C(X) &\rightarrow \mathbb{R}^G \\ f &\mapsto f|_G \end{aligned}$$

is injective. Hence

$$\{\text{eq:jones-lem-ineq-2}\} \quad (**) \quad \text{card } C(X) \leq \text{card } \mathbb{R}^G.$$

From Corollary 0.129,

$$\{\text{eq:jones-lem-ineq-3}\} \quad (***) \quad \text{card } \mathbb{R}^G \leq \text{card } \mathbb{R}^{\mathbb{N}} = \text{card } (2^{\mathbb{N}}) = 2^{\aleph_0}.$$

From equations (\*), (\*\*), and (\*\*\*) it follows that:

$$\text{card } 2^D \leq 2^{\aleph_0}.$$

But  $\text{card } D \leq \text{card } 2^D$  by Cantor's Theorem (0.125), and  $\text{card } 2^{\aleph_0} = \text{card } \mathbb{R}$  by Example 0.121.  $\square$

Jones's Lemma may be used to produce counterexamples: if a separable space has a discrete closed subspace with the same cardinality as  $\mathbb{R}$ , then the space cannot be normal.

Sorgenfrey-plane-nonnormal-Jones

**6.32 Example.** The Sorgenfrey plane  $\mathbb{R}_l \times \mathbb{R}_l$  (Exercise 3.114) is separable, having  $\mathbb{Q} \times \mathbb{Q}$  as a dense subset. Its main diagonal  $\Delta = \{\langle x, x \rangle : x \in \mathbb{R}_l\}$  is a closed discrete subspace with  $\text{card } \Delta = \text{card } \mathbb{R}$ . According to the Jones's Lemma (Theorem 6.31),  $\mathbb{R}_l \times \mathbb{R}_l$  cannot be normal.  $\diamond$

Some quotients of normal spaces are normal.



prop:cont-closed-image-normal

**6.33 Proposition.** *A continuous closed image of a normal space is itself normal.*

**Proof.** The proof is left to the reader ([Exercise 27](#)).  $\square$

In general a quotient of a normal space need not be normal. Indeed, as in [Example 6.25](#), a  $T_1$ -quotient of a normal  $T_1$ -space need not be normal. But the situation is worse than that: **a completely regular  $T_0$ -space that is the quotient of a normal  $T_1$ -space need not itself be normal.** However, finding examples to show that is not so easy. One such anomalous example, involving the tangent disk space  $\Gamma$  ([Exercise 2.37](#)), is constructed in [Exercise 37](#). Another, involving the “beta compactification”  $\beta\mathbb{N}$  of the discrete space of natural numbers, is presented in [Exercise 67](#).

### The Tietze Extension Theorem

subsec:tietze

In the [subsection “Extending and gluing maps”](#) we raised the question of when a continuous map on a subspace of given space can be extended continuously to the entire space. We saw that it is not in general possible, even for a real-valued map [[Examples 3.11 \(2\)](#)], but that it *is* always possible for a continuous function  $A \rightarrow [0, 1]$  on a *closed* subspace of a *metrizable* space [[Examples 3.14 \(2\)](#)]. Now metrizable spaces are among the class of normal spaces, and existence of continuous extensions holds for this more general class.

thm:tietze

**6.34 Tietze Extension Theorem.** *A continuous function  $f: A \rightarrow [0, 1]$  on a closed subspace  $A$  of a normal space  $X$  has a continuous extension  $F: X \rightarrow [0, 1]$  to the entire space.*

In the statement of the Tietze Extension Theorem, any nondegenerate closed interval  $[a, b]$  may be used in place of  $[0, 1]$ .

Actually, the property of such extensions existing is equivalent to normality.

prop:normal-iff-cont-ext

**6.35 Proposition.** *A necessary and sufficient condition for a topological space  $X$  to be normal is that each continuous function  $A \rightarrow [0, 1]$  on a closed subset of  $X$  have a continuous extension to the entire space  $X$ .*

**Proof.** Necessity is just the Tietze Extension Theorem, to be proved below. The proof of sufficiency is left as an exercise ([Exercise 48](#)).  $\square$

The method for continuously extending a given continuous function  $f$  on  $A$  to the entire space  $X$  will be to build a sequence of continuous functions on  $X$  that successively better approximate  $f$  on  $A$  and then to take the limit of the sequence. Instead of working with the interval  $[0, 1]$  as codomain, it will be more convenient to work with the interval  $[-1, 1]$ , which is symmetric around the origin. In fact, in the proof we will work more generally with various closed intervals  $[-r, r]$  symmetric around the origin.

normal space!quotient space@and c  
quotient space!normal space@of no

lem:approx-f-on-A-by-fn-on-X

**6.36 Lemma.** Let  $r > 0$  and let  $f$  be a continuous real-valued function on a closed subset  $A$  of a normal space  $X$  such that

$$|f(x)| \leq r \text{ for all } x \in A.$$

Then there exists a continuous real-valued function  $g$  with domain the entire space  $X$  for which

$$\begin{aligned} |g(x)| &\leq \frac{1}{3} r \text{ for all } x \in X, \\ |f(x) - g(x)| &\leq \frac{2}{3} r \text{ for all } x \in A. \end{aligned} \quad (*)$$

**Proof.** Partition  $[-r, r]$  into three subintervals of equal length. The inverse images

$$C = f^{-1}\left(\left[-r, -\frac{1}{3}r\right]\right), \quad D = f^{-1}\left(\left[\frac{1}{3}r, r\right]\right)$$

of the extreme left and right thirds are disjoint closed subsets of  $A$  and hence of  $X$ . By Urysohn's Lemma there is a Urysohn function  $g: X \rightarrow \left[-\frac{1}{3}r, \frac{1}{3}r\right]$  for  $C$  and  $D$ .

Condition [equation \(\\*\)](#) holds because, for any given  $x \in A$ , both  $f(x)$  and  $g(x)$  belong to the same one of the three subintervals of  $[-r, r]$ .  $\square$

**Proof of the Tietze Extension Theorem.** Let  $f: A \rightarrow [0, 1]$  be a continuous function on a closed subspace  $A$  of a normal space  $X$ . Form the linear homeomorphism  $h: [0, 1] \cong [-1, 1]$  mapping 0 to  $-1$  and 1 to 1, respectively, and let  $f^* = h \circ f: A \rightarrow [-1, 1]$ . We shall obtain a continuous extension  $F^*: X \rightarrow [-1, 1]$  of  $f^*$ , and then  $h^{-1} \circ F^*$  will be the desired continuous extension of the original  $f$ . Accordingly, to simplify the notation, we assume that  $f$  itself has codomain  $[-1, 1]$  rather than the original  $[0, 1]$ .

By [Lemma 6.36](#) applied to  $f$  with  $r = 1$ , we obtain an approximating function  $g_1: X \rightarrow \mathbb{R}$  to  $f$  with

$$\begin{aligned} |g_1(x)| &\leq \frac{1}{3} \text{ for all } x \in X, \\ |f(x) - g_1(x)| &\leq \frac{2}{3} \text{ for all } x \in A. \end{aligned}$$

Next, by the same lemma applied to  $f - g_1$  with  $r = 2/3$ , we obtain a function  $g_2: X \rightarrow \mathbb{R}$  with

$$\begin{aligned} |g_2(x)| &\leq \frac{1}{3} \frac{2}{3} \text{ for all } x \in X, \\ |f(x) - g_1(x) - g_2(x)| &\leq \left(\frac{2}{3}\right)^2 \text{ for all } x \in A, \end{aligned}$$

so that  $g_1 + g_2$  is a better approximating function on  $A$  to  $f$  than was  $g_1$ . Continuing in this way, we obtain a sequence  $\langle g_n \rangle_{n=1,2,\dots}$  of approximating continuous real-valued functions on  $X$  with

$$\begin{aligned} |g_n(x)| &\leq \frac{1}{3} \left(\frac{2}{3}\right)^{n-1} \text{ for all } x \in X, \\ |f(x) - g_1(x) - g_2(x) - \cdots - g_n(x)| &\leq \left(\frac{2}{3}\right)^n \text{ for all } x \in A. \end{aligned}$$

For each  $n$ , define  $s_n: X \rightarrow \mathbb{R}$  by

$$s_n(x) = \sum_{j=1}^n g_j(x),$$

so that  $s_n$  is continuous and satisfies

$$(*) \quad |f(x) - s_n(x)| \leq \left(\frac{2}{3}\right)^n \text{ for all } x \in A.$$

{eq:sn-approx-f-tietze}

For each  $n$ ,

$$\{eq:tietze-pf-bd-sn\} \quad (**) \quad |s_n(x)| \leq \sum_{j=1}^n \frac{1}{3} \left(\frac{2}{3}\right)^{j-1} = 1.$$

Thus each  $s_n$  in fact maps  $X$  into  $[-1, 1]$  and is therefore bounded, that is, belongs to the space  $\mathcal{B}(X)$  of bounded real-valued functions on  $X$ .

We are going to show that the sequence  $\langle s_n \rangle_{n=1,2,\dots}$  converges uniformly on  $X$  to a continuous extension  $F$  of  $f$  to  $X$ . (In other terms, we are going to show that the series  $\sum_{n=1}^{\infty} g_n$  of functions converges uniformly to the desired  $F$ .) More precisely, we are going to show that the sequence  $\langle s_n \rangle_{n=1,2,\dots}$  converges in  $(\mathcal{B}(X), d_{\infty})$ , where  $d_{\infty}$  is the sup metric, and that its limit is continuous, extends  $f$  to  $X$ , and maps  $X$  into  $[-1, 1]$ .

We show that  $\langle s_n \rangle_{n=1,2,\dots}$  is a Cauchy sequence with respect to  $d_{\infty}$ . Let  $\varepsilon > 0$  be arbitrary. For all positive integers  $m$  and  $n$  with  $m > n$  we have

$$|s_m(x) - s_n(x)| \leq \sum_{j=n+1}^m |g_j(x)| \leq \sum_{j=n+1}^m \frac{1}{3} \left(\frac{2}{3}\right)^{j-1} \leq \left(\frac{2}{3}\right)^n$$

for all  $x \in X$ , which means that

$$d_{\infty}(s_m, s_n) \leq \left(\frac{2}{3}\right)^n.$$

Thus for  $k$  so large that  $\left(\frac{2}{3}\right)^k < \varepsilon$ , we have  $d_{\infty}(s_m, s_n) < \varepsilon$  for all  $m, n$  with  $m \geq k, n \geq k$ .

Since  $\mathcal{B}(X)$  is complete with respect to  $d_{\infty}$  (Theorem 1.83), the sequence  $\langle s_n \rangle_{n=1,2,\dots}$  does converge in  $(\mathcal{B}(X), d_{\infty})$ . Let

$$F = \lim_{n \rightarrow \infty} s_n.$$

According to Lemma 1.85  $F$  is continuous on  $X$ . From equation (\*) it follows that  $F(x) = f(x)$  at all  $x \in A$ .

To complete the proof, we show that  $F$  maps  $X$  into  $[-1, 1]$ . Let  $x \in X$ . For each  $n = 1, 2, 3, \dots$ , we have

$$|F(x)| \leq \|F\|_{\infty} \leq \|s_n\|_{\infty} + \|F - s_n\|_{\infty} \leq 1 + \|F - s_n\|_{\infty}$$

Since  $\langle s_n \rangle_{n=1,2,\dots} \rightarrow F$  with respect to  $d_{\infty}$ , for sufficiently large  $n$  we have  $\|F - s_n\|_{\infty} < \eta$  and so  $|F(x)| \leq \|F\|_{\infty} < 1 + \eta$ . Since  $\eta$  was arbitrary,  $|F(x)| \leq 1$ , as needed.  $\square$

It is possible to prove Tietze Extension Theorem without constructing such a sequence of approximating functions: see the comment “Proof of the Tietze Extension Theorem” (page 656) in the Additional Readings.

cor:tietze-R-valued **6.37 Corollary.** *A continuous real-valued function on a closed subset of a normal space has a continuous extension to the entire space.*

**Proof.** Use the same method as in the proof of (ii) of Proposition 6.14 to convert a given continuous function  $A \rightarrow \mathbb{R}$  to a continuous function  $A \rightarrow [0, 1]$ .  $\square$

Since a compact  $T_2$ -space is normal, the next corollary follows at once from the Tietze Extension Theorem.

cor:tietze-cpt **6.38 Corollary.** *A continuous function  $A \rightarrow [0, 1]$  on a closed subspace  $A$  of a compact  $T_2$ -space  $X$  has a continuous extension  $X \rightarrow [0, 1]$  to the entire space.*

The situation is more complicated when the space is only locally compact.

uniformly convergent sequence of fu

ex:non-extendable-f-on-loc-cpt  
plane-filling curve  
Cantor set!continuous image@and continuous image  
ternary expansion

**6.39 Example.** Let  $X$  be a nonnormal locally compact Hausdorff space. For example, take  $X$  to be the chipped Tychonof plank (Example 6.23). In view of Proposition 6.35, there exists a continuous function  $X \rightarrow [0, 1]$  having no continuous extension to  $X$ . (Exercise 49 requests an explicit example of such a function.)  $\diamond$

However, if we strengthen the condition on the subspace of the locally compact space, then again a continuous extension exists.

cor:tietze-cpt-in-loc-cpt-T2

**6.40 Corollary.** Let  $f: K \rightarrow [0, 1]$  be a continuous function on a compact subset of a locally compact  $T_2$ -space  $X$ . Then  $f$  has a continuous extension  $F: X \rightarrow [0, 1]$  that vanishes outside some compact neighborhood of  $K$ .

**Proof.** The compact set  $K$  has an open neighborhood  $V$  for which  $\text{cls } V$  is compact and hence normal. By the Tietze Extension Theorem,  $f$  has a continuous extension  $f: \text{cls } V \rightarrow [0, 1]$ . By Urysohn's Lemma, there is a continuous function  $g: \text{cls } V \rightarrow [0, 1]$  for which

$$g(x) = 0 \text{ for all } x \in (\text{cls } V) \setminus V, \quad g(x) = 1 \text{ for all } x \in K.$$

Then the desired  $F: X \rightarrow [0, 1]$  is given by

$$F(x) = \begin{cases} g(x)f(x) & \text{if } x \in \text{cls } V, \\ 0 & \text{if } x \in X \setminus \text{cls } V. \end{cases} \quad \square$$

A plane-filling curve—a continuous surjection from the unit interval  $I = [0, 1]$  onto the unit square  $I^2 = I \times I$ —was explicitly constructed in Example 5.65. An alternative way to establish existence of such a curve is to invoke Tietze Extension Theorem.

app:plane-filling-curve-tietze

**6.41 Application.** There is a continuous surjection of the unit interval  $I = [0, 1]$  onto the unit square  $I^2 = I \times I$ . To establish this, begin with the continuous surjection

$$g: K \rightarrow I,$$

constructed in Example 4.17, given by

$$g\left(\sum_{n=1}^{\infty} \frac{x_n}{3^n}\right) = \sum_{n=1}^{\infty} \frac{x_n/2}{2^n},$$

where  $x_n \in \{0, 2\}$  for every  $n$ .

Next, define

$$h: K \rightarrow K \times K$$

by

$$h(x) = \left( \sum_{i=1}^{\infty} \frac{x_{2i-1}}{3^{2i-1}}, \sum_{i=1}^{\infty} \frac{x_{2i}}{3^{2i}} \right),$$

where again  $x_n \in \{0, 2\}$  for every  $n$ . The function  $h$  is surjective. In fact, each sequence in  $\{0, 2\}$  occurs as the sequence of odd-index digits in the ternary expansion of some  $x \in K$ , just as each such sequence occurs as the sequence of even-index digits in the ternary expansion of some  $x \in K$ . The function  $h$  is also continuous. In fact, it suffices to establish that its composites with the first and second projections  $p: K \times K \rightarrow K$  and  $q: K \times K \rightarrow K$ ,

respectively, are continuous. Consider  $q \circ h$ ; the proof for  $p \circ h$  is similar. If  $n$  is a positive integer, then for all  $u, v \in K$  with  $|u - v| < 1/3^n$ , we have  $u_i = v_i$  for all  $i \leq n$ , so that

$$|(q \circ h)(u) - (q \circ h)(v)| = \left| \sum_{i=n+1}^{\infty} \frac{u_{2i} - v_{2i}}{3^{2i}} \right| \leq \sum_{i=n+1}^{\infty} \frac{2}{3^{2i}} = \frac{2}{3^{2n}}.$$

(We have actually shown that  $h$  is uniformly continuous with respect to the metrics on  $K$  and  $K \times K$  induced by the Euclidean metrics on  $\mathbb{R}$  and  $\mathbb{R} \times \mathbb{R}$ .)

It follows that the composite  $f = (g \times g) \circ h: K \rightarrow I \times I$  is a continuous surjection. The unit interval  $I$ , being a compact Hausdorff space, is normal. By the Tietze Extension Theorem,  $f$  has a continuous extension  $F: I \rightarrow I^2$ , and this extension is surjective since already  $f$  is surjective.  $\diamond$

normal space

Sorgenfrey line!completely regular space

quotient space!separation properties

separation properties!quotient space

collapsing to a point!separation properties

Mysior's space

normal space!quotient space@and compact

one-point compactification

completely regular space!normal space

normal space!completely regular space

## EXERCISES FOR SECTION 6.2

prob:RI-compl-reg **23.** Example 6.29 used Corollary 6.28 together with Examples 2.97 (3) to reach the result that the Sorgenfrey line  $\mathbb{R}_l$  is completely regular. Establish the same result directly from the definition of complete regularity.

prob:Q-collapse-R-not-cr **24.** Show that the quotient space  $\mathbb{R}/\mathbb{Q}$  obtained by collapsing the rationals to a point is a  $T_0$ -space (but not a  $T_1$ -space) that is *not* completely regular, even though the real line  $X = \mathbb{R}$  is completely regular.

**25.** This exercise refers to Mysior's space of Example 6.17.

- (a) Verify that the given local bases do, in fact, have the requisite properties to define a topology on  $X$ .
- (b) Verify that each basic neighborhood of a point of  $\bar{H}$  is closed and that  $\text{cls } V_{n+2} \subset V_n$  for each  $n$ .
- (c) Show that the sets  $J_x$  are closed in  $X$  for all  $x \in \mathbb{R}$ . For  $x \in \mathbb{R}$ , is the set  $I_x$  also closed in  $X$ ?
- (d) Is the subspace  $\bar{H}$  of  $X$  completely regular?
- (e) What, if anything, changes in the analysis of Mysior's space if, instead of  $H$  and  $\bar{H}$ , one uses  $\mathbb{R} \times [0, 2[$  and  $\mathbb{R} \times [0, 2]$ ?
- (f) Suppose we adjoin to  $X$  a new point  $b \neq a$  and take as a local base at  $b$  the sets  $W_n$  given by

$$W_n = \{b\} \cup \{(x, y) : x < -n, x + n + 2 < y\} \quad (n = 1, 2, 3, \dots),$$

thereby obtaining a new space  $Y$ . Show that, like  $X$ , the space  $Y$  is regular. Show further that  $f(a) = f(b)$  for each continuous  $f: Y \rightarrow [0, 1]$ , so that  $Y$  is not completely regular.

prob:closed-subspace-of-normal **26.** Prove Proposition 6.22: a closed subspace of a normal space is itself normal.

prob:pf-cont-closed-image-normal **27.** Prove Proposition 6.33: a closed continuous image of a normal space is normal. (Hint: You may want to apply Exercise 2.143.)

prob:chipped-prod-alphas **28.** Let  $D$  be the set of real numbers with its discrete topology. Let  $\infty_0$  and  $\infty_1$  be the points-at-infinity of the one-point compactifications  $N_\infty$  and  $D_\infty$  of  $\mathbb{N}$  and  $D$ , respectively. Show that the subspace

$$(\mathbb{N}_\infty \times D_\infty) \setminus \{(\infty_0, \infty_1)\}$$

of  $\mathbb{N}_\infty \times D_\infty$  is a completely regular  $T_0$ -space that is not normal.

**29.** Prove that a  $T_1$ -space is completely regular if and only if it has a base  $\mathcal{B}$  such that:

- for each  $x \in X$  and each  $B \in \mathcal{B}$ , there is some  $B' \in \mathcal{B}$  with  $x \notin B'$  and  $B \cup B' = X$ ; and
- for all  $A, B \in \mathcal{B}$  with  $A \cup B = X$ , there are disjoint  $A', B' \in \mathcal{B}$  with  $X \setminus A \subset B'$  and  $X \setminus B \subset A'$ .

**30.** The Sorgenfrey line  $\Gamma$  is a regular  $T_0$ -space (and *a fortiori*, a  $T_2$ -space): see Exercise 2.153.

(a) Verify that  $\Gamma$  is completely regular.

[Hint: Fix a point  $z$  and a closed set  $E$  in  $\Gamma = H \cup L$  with  $z \notin E$ . In the case  $z = \langle x, 0 \rangle \in L$ , there is an  $r > 0$  with the neighborhood  $\{z\} \cup B$  of  $z$  disjoint from  $E$ , where  $B = B_r(\langle x, r \rangle)$ . Define  $g: \{z\} \cup (\Gamma \setminus B) \rightarrow [0, 1]$  by  $g(z) = 0$  and  $g(w) = 1$  for all  $w \in \Gamma \setminus B$ . Now extend  $g$  to  $\Gamma$ .]

(b) Use Jones's Lemma (Theorem 6.31) to prove anew that  $\Gamma$  is not normal.

[Note: A different proof, using the Baire Category Theorem (1.91), was outlined in Exercise 2.153. A more direct proof, avoiding use of both cardinality and the Baire Category Theorem but applying instead the Nested Set Theorem, is possible: see Vetterlein [66].]

**31.** Strengthen Theorem 6.20 by proving: A regular Lindelöf space is normal.

[Note: According to Examples 2.97 (2) and Exercise 2.115 (b), the Sorgenfrey line  $\mathbb{R}_l$  is a regular Lindelöf space. Thus the result here provides an alternative way to establish normality of  $\mathbb{R}_l$ .]

**32.** Show that the Cartesian sum  $+_{i \in I} X_i$  (Exercise 3.106) of an arbitrary family  $\langle X_i \rangle_{i \in I}$  of normal spaces is itself normal.

**33.** Must the Cartesian sum  $+_{i \in I} X_i$  (Exercise 3.106) of an arbitrary family  $\langle X_i \rangle_{i \in I}$  of completely regular spaces be completely regular?

**34.** Prove that, when provided with its box topology,  $\mathbb{R}^I$  is a completely regular space.

**35.** Recall that the tangent disk space  $\Gamma$  (Exercise 2.37) is a completely regular  $T_0$ -space: see Exercise 30.

(a) Form the quotient space  $\Gamma \rightarrow \Gamma//Q$  obtained by collapsing to a point the subset  $Q = \mathbb{Q} \times \{0\}$ , the set of rational points on the  $x$ -axis. Show that the quotient map  $\Gamma \rightarrow \Gamma//Q$  is a closed map. Then determine which separation properties does this quotient space have?

(b) Repeat (a) but for the quotient space of  $\Gamma$  obtained by collapsing  $Q$  and  $L \setminus Q$  to distinct points (Exercise 3.172).

**36.** Must the continuous image of a completely regular  $T_0$ -space under a map that is both open and closed be completely regular?

**37.** (a) From Exercise 2.100 (or from Exercise 58, below), the tangent disk space  $\Gamma$  has an open cover  $\langle U_i \rangle_{i \in I}$  consisting of metrizable subspaces. Deduce that the Cartesian sum  $+_{i \in I} U_i$  is a normal  $T_1$ -space.

(b) Show that the natural map  $q: +_{i \in I} U_i \rightarrow \Gamma$  given by  $q(x) = x$  is a continuous open surjection.

(Note: This provides, as promised, an example of a nonnormal but completely regular  $T_0$ -quotient of a normal  $T_1$ -space.)

prob:Thomas-plank-normality

- 38.** Show that the Thomas plank [Examples 2.60 (4)], which is a regular  $T_0$ -space, is not normal.

Thomas plank!separation properties

Thomas plank

prob:non-cr-quot-cr-T2

- 39.** (Continuation of Exercise 38.)

completely regular space!quotient sp

corkscrew

Let  $P$  be the Thomas plank [Examples 2.60 (4)]. Make a doubly infinite “stack” of copies of  $P$  by forming the product space  $P \times \mathbb{Z}$ . Next, adjoin to  $P \times \mathbb{Z}$  new points  $p^+$ , a “point at  $+\infty$ ,” and  $p^-$ , a “point at  $-\infty$ ,” as follows. Take two objects  $p^+$  and  $p^-$  not already in  $P \times \mathbb{Z}$  and form the set

$$X = (P \times \mathbb{Z}) \cup \{p^+, p^-\}.$$

Endow  $X$  with the topology such that: at a point  $(\langle x, y \rangle, k) \in P \times \mathbb{Z}$ , the usual collection of open neighborhoods in the product is a local base; at  $p^+$ , the collection of the sets  $(P \times [n, +\infty[) \cup \{p^+\}$  for  $n \in \mathbb{Z}$  is a local base; and at  $p^-$ , the collection of the sets  $(P \times ]-\infty, n]) \cup \{p^-\}$  for  $n \in \mathbb{Z}$  is a local base.

- (a) Verify that the collections defined above do specify a topology on  $X$ .  
 (b) Show that  $X$  is a completely regular  $T_0$ -space. [Note: Recall from Examples 2.60 (4) and Examples 2.97 (1) that the Thomas plank  $P$  is a regular  $T_0$ -space and, *a fortiori*, a  $T_2$ -space.]  
 (c) Let  $A$  be the “left-hand edge” and  $B$  be the “bottom edge” of  $P$ , that is,

$$A = \{0\} \times \{1/n : n \in \mathbb{N}^*\} = \{\langle 0, 1/n \rangle : n = 1, 2, 3, \dots\},$$

$$B = [0, 1[ \times \{0\} = \{\langle x, 0 \rangle : 0 < x \leq 1\}.$$

The **corkscrew over the Thomas plank**  $P$  is the quotient space  $Y$  obtained by gluing the left-hand edge of each even-numbered copy of  $P$  with the left-hand edge of the odd-numbered copy stacked above it and gluing the bottom edge of each odd-numbered copy of  $P$  with the bottom edge of the even-numbered copy stacked above it; in other words,  $Y = X/\sim$  is obtained by making the identifications

$$\begin{aligned} \langle a, 2k \rangle &\sim \langle a, 2k + 1 \rangle && \text{for all } a \in A \text{ and all } k \in \mathbb{Z}, \\ \langle b, 2k + 1 \rangle &\sim \langle b, 2k + 2 \rangle && \text{for all } b \in B \text{ and all } k \in \mathbb{Z}. \end{aligned}$$

Show that  $Y$  is a regular  $T_0$ -space that is not completely regular.

(Hint: To show that  $Y$  is not completely regular, show that any continuous real-valued function on  $Y$  must take the same value at  $p^+$  as at  $p^-$ .)

prob:urysohn-lem-scott-way

- 40.** Prove: If  $A$  and  $B$  be  $F_\sigma$ -sets (Exercise 2.19) in a normal space with  $A$  disjoint from  $B$  and  $B$  disjoint from  $A$ , then  $A$  and  $B$  have disjoint neighborhoods in  $X$ .

prob:rational-seq-top

- 41.** For each irrational number  $x$ , choose some sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  of distinct rational numbers converging in  $\mathbb{R}$  to  $x$  for the usual topology on  $\mathbb{R}$ . The **rational sequence topology** is the unique topology on  $\mathbb{R}$  such that at each rational  $x$ , the collection  $\{\{x\}\}$  is a local base and at each irrational number  $x$ , the collection  $\{\{x\} \cup \{x_k : k > n\} : n = 0, 1, 2, \dots\}$  is a local base. Show that the resulting space is a locally compact and completely regular  $T_0$ -space that is zero-dimensional but not normal.
- 42.** Establish the following version of Tietze Extension Theorem: Let  $f : A \rightarrow ]a, b[$  be a continuous function on a closed subset  $A$  of a normal space  $X$  to an open interval in  $\mathbb{R}$ . Then  $f$  has a continuous extension  $F : X \rightarrow ]a, b[$ .

- 43. (a)** Deduce the Tietze Extension Theorem from the following special case: A continuous function  $f: A \rightarrow [0, 1]$  on a closed subset  $A$  of a *connected* normal space  $X$  can be has a continuous extension  $F: X \rightarrow [0, 1]$ .
- (b)** Deduce Urysohn's Lemma from the following special case: If  $A$  and  $B$  are disjoint closed subsets of a *connected* normal space  $X$ , then there exists a continuous function  $f: X \rightarrow [0, 1]$  such that  $f(x) = 0$  for all  $x \in A$  and  $f(y) = 1$  for all  $y \in B$ .

**44.** Verify that the extension  $F$  obtained in the proof of [Corollary 6.40](#) is in fact continuous.

**45.** Prove that an arbitrary topological space  $Y$  is the continuous open image of a Hausdorff space. (In particular, any non-Hausdorff space is the continuous open image of a Hausdorff space.)

[Hint: For each  $y \in Y$ , let  $Z_y$  be the subspace of the product space  $\mathbb{R}^Y$  given by

$$Z_y = \{ \langle z_t \rangle_{t \in Y} : z_y \in \mathbb{Q} \text{ and } z_t \in \mathbb{R} \setminus \mathbb{Q} \text{ for each } t \neq y. \}$$

Let  $Z$  be the subspace of  $\mathbb{R}^Y$  given by

$$Z = \bigcup_{y \in Y} Z_y$$

and let  $X$  be the subspace of  $Y \times Z$  given by

$$X = \bigcup_{y \in Y} (\{y\} \times Z_y) = \{ \langle y, z \rangle : y \in Y \text{ and } z \in Z_y \}.$$

Finally, let  $p: X \rightarrow Y$  be the restriction to  $X$  of the first projection  $Y \times Z \rightarrow Y$ .

Show:

- (i) The family  $\langle Z_y \rangle_{y \in Y}$  partitions  $Z$  into dense subsets;
- (ii)  $X$  is a Hausdorff space; and
- (iii) the map  $p$  is a continuous open surjection.]

prove-normal-if-fns-separate-closed

**46.** Prove the sufficiency part of [Proposition 6.27](#).

**47.** Let  $f: [0, 1] \rightarrow \mathbb{R}$  be the map  $x \mapsto 1/(x+1)$  with domain the subspace  $[0, 1]$  of  $\mathbb{R}$ . Give an explicit continuous extension of  $f$  to  $\mathbb{R}$ .

prob:prove-normal-if-cont-ext

**48.** Prove the sufficiency part of [Proposition 6.35](#).

non-extendable-f-in-chipped-plank

**49.** Exhibit a continuous function  $f: E \rightarrow [0, 1]$  on a closed subset  $E$  of the chipped Tychonoff plank  $P^*$  having no continuous extension to  $P^*$ .

(Note: This strengthens [Example 6.39](#), which only established existence of such an  $E$  and  $f$ .)

prob-part:extend-lim-bdded-seq-R

**50. (a)** Endow the set  $\ell^\infty$  of all bounded sequences of reals provided with the topology induced by its sup metric. (See [Exercise 1.17](#).) Show that there is a continuous map  $\Phi: \ell^\infty \rightarrow \mathbb{R}$  such that  $\Phi(\langle x_n \rangle_{n \in \mathbb{N}}) = \lim_{n \rightarrow \infty} x_n$  whenever  $\langle x_n \rangle_{n \in \mathbb{N}}$  converges.

**(b)** Repeat **(a)** but for sequences of complex numbers.

**51.** Let  $f: E \rightarrow \mathbb{R}$  be a continuous map with domain a closed and bounded subspace of  $\mathbb{R}^n$ . Show that  $f$  has an extension  $F: \mathbb{R}^n \rightarrow \mathbb{R}$  for which  $\{x \in \mathbb{R}^n : f(x) \neq 0\}$  is bounded.



52. (a) Show that the Tietze Extension Theorem fails if the codomain  $[0, 1]$  is replaced by the unit circle  $S_1$ .  
 (b) Prove, however, that the Tietze Extension Theorem does remain true if the codomain  $[0, 1]$  is replaced by the unit circle  $S_1$  and the map  $f$  with domain  $A$  is *not* surjective.  
 (c) What is the situation if the codomain is replaced, instead, by the unit disk  $D_1$ ?
53. If  $F: X \rightarrow \mathbb{R}$  is a continuous extension to a normal space  $X$  of a continuous map  $f: A \rightarrow \mathbb{R}$  on a closed subset  $A$  of  $X$ , must  $F(X) = f(A)$ ?
54. We deduced the Tietze Extension Theorem from Urysohn's Lemma. Instead, deduce Urysohn's Lemma from the Tietze Extension Theorem.
55. We know that, for a metrizable space  $X$ , if  $X$  is compact, then every continuous real-valued function on  $X$  is constant. (In fact, metrizability is not needed.) Prove the converse.  
 [Hint: Suppose there is some sequence  $\langle x_n \rangle_{n \in \mathbb{N}}$  in  $X$  having no convergent subsequence. Consider the function  $g: \{x_n : n \in \mathbb{N}\} \rightarrow \mathbb{R}$  given by  $g(n) = n$ .]
56. Establish the following modification of the Tietze Extension Theorem. If  $f: A \rightarrow Y$  is a continuous map from a closed subset  $A$  of a normal space  $X$  to a completely regular space  $Y$ , then  $f$  has a continuous domain-codomain extension  $F: X \rightarrow Z$  with  $Z$  being a compact Hausdorff space containing  $Y$  as a subspace.  
 (Hint: Take  $Z = I^{C(Y, I)}$ .)

embedding  
Hilbert cube

### 6.3 Embedding theorems

sec:embedding

Certain kinds of topological spaces can be embedded into specific “concrete” spaces such as Tychonoff cubes (that is, powers of  $I = [0, 1]$ ) and Euclidean spaces  $\mathbb{R}^n$ . Existence of such embeddings will be shown by applying existence of sufficiently many continuous real-valued functions, as established in the preceding section

#### Embedding in Tychonoff cubes

subsec:embed-cubes

Sometimes a single space can be found into which all members of a wide class of spaces can simultaneously be embedded. One such space is the *Hilbert cube*

$$I^\infty = \{ \langle x_i \rangle_{i=1,2,3,\dots} : x_i \in \mathbb{R} \text{ and } |x_i| \leq 1/i \text{ for each } i = 1, 2, 3, \dots \},$$

which (according to [Exercise 1.36](#)) is a subset of the Hilbert sequence space  $\ell^2$  and therefore has a topology induced by the Hilbert metric  $d_2$  as defined in [Example 1.10](#). Furthermore, the Hilbert cube is the product

$$I^\infty = [-1, 1] \times [-1/2, 1/2] \times [-1/3, 1/3] \times \cdots = \bigtimes_{i=1}^{\infty} [-1/i, 1/i]$$

of closed intervals, and, according to [Examples 3.64 \(7\)](#), its topology is the product topology, so that

$$I^\infty \cong I \times I \times I \times \cdots = I^{\mathbb{N}}.$$

ed-separable-metrizable-Hilbert-cube **6.42 Theorem.** Every separable metrizable space is embeddable in the Hilbert cube  $l^\infty$ .

**Proof.** Assume that the topological space  $X$  is both separable and metrizable. Choose a sequence  $\langle x_i \rangle_{i=1,2,3,\dots}$  such that the set  $\{x_i : i = 1, 2, 3, \dots\}$  is dense in  $X$ , and choose a metric  $d$  on  $X$  inducing the topology of  $X$  such that  $d(x, y) \leq 1$  for all  $x, y \in X$ .

Step 0: Define a map  $h: X \rightarrow l^\infty$ . The formula

$$h(x) = \left\langle \frac{1}{i} d(x, x_i) : i = 1, 2, 3, \dots \right\rangle$$

defines a map

$$h: X \rightarrow l^\infty.$$

We shall use repeatedly below the following property of  $h$ : for  $x, y \in X$  and for  $j = 1, 2, 3, \dots$  we have

$$\begin{aligned} \frac{1}{j} |d(x, x_j) - d(u, x_j)| &\leq \left( \sum_{i=1}^{\infty} \frac{1}{i^2} |d(x, x_i) - d(u, x_i)|^2 \right)^{1/2} \\ &= d_2(h(x), h(u)) \\ &\leq \left( \sum_{i=1}^{\infty} \frac{1}{i^2} (d(x, u))^2 \right)^{1/2}. \end{aligned}$$

{pf-eq:inequality-embed-in-elltwo} (\*)

Step 1: The map  $h$  is injective. To prove this, let  $x, u \in X$  with  $x \neq u$ . There is a  $j$  with  $d(x, x_j) < \frac{1}{2} d(x, u)$ . Now  $d(x, x_j) \neq d(u, x_j)$ , for otherwise we would have

$$d(x, u) \leq d(x, x_j) + d(x_j, u) = 2d(x, x_j) < d(x, u),$$

which is absurd. Then from (\*),

$$d_2(h(x), h(u)) \geq \frac{1}{j} |d(x, x_j) - d(u, x_j)| > 0,$$

and so  $h(x) \neq h(u)$ .

Step 2: The injection  $h$  is continuous. To prove this, fix  $x \in X$  and let  $\varepsilon > 0$  be arbitrary. Since the series  $\sum_{i=1}^{\infty} 1/i^2$  converges, we may choose  $n$  so large that

$$\sum_{i=n+1}^{\infty} \frac{1}{i^2} < \frac{\varepsilon^2}{2}.$$

If  $u \in X$  with

$$d(x, u) < \left( \frac{\varepsilon^2}{2n} \right)^{1/2},$$

then from (\*) we obtain

$$\begin{aligned} (d_2(h(x), h(u)))^2 &\leq \sum_{i=1}^{\infty} \frac{1}{i^2} (d(x, u))^2 + \sum_{i=n+1}^{\infty} \frac{1}{i^2} (d(x, u))^2 \\ &< n \cdot \frac{\varepsilon^2}{2n} + \sum_{i=n+1}^{\infty} \frac{1}{i^2} < \frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{2} = \varepsilon^2, \end{aligned}$$

and so  $d_2(h(x), h(u))$ , as desired.

Step 3: The continuous injection  $h$  maps each open subset of  $X$  onto an open subset of  $h(X)$  in  $l^\infty$ . To prove this, fix  $x \in X$  and let  $\varepsilon > 0$  be arbitrary. It suffices to find a  $\delta > 0$  such that

$$B_\delta(h(x); d_2) \cap h(X) \subset h(B_\varepsilon(x; d)).$$

There exist a  $j$  with  $d(x, x_j) < \varepsilon/3$ . We claim that

$$\delta = \frac{\varepsilon}{3j}$$

will do. In fact, suppose that  $y \in h(X)$  with

$$d_2(h(x), y) < \delta.$$

Since  $h$  is injective, there is a unique  $u \in X$  with  $y = h(u)$ . From (\*),

$$|d(x, x_j) - d(u, x_j)| \leq j \cdot d_2(h(x), h(u)) < \frac{\varepsilon}{3}.$$

Then

$$\begin{aligned} d(u, x_j) &< d(x, x_j) + \frac{\varepsilon}{3} < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2\varepsilon}{3}, \\ d(u, x) &\leq d(u, x_j) + d(x_j, x) < \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \end{aligned}$$

and so  $u \in B_\varepsilon(x; d)$ . Hence  $y = h(u) \in h(B_\varepsilon(x; d))$ , as desired.  $\square$

**6.43 Corollary.** A necessary and sufficient condition for a topological space to be separable and metrizable is that it be embeddable into the Hilbert cube  $l^\infty$ .

**Proof.** Sufficiency of the condition is just the preceding theorem. To establish necessity, recall that the Hilbert sequence space  $\ell^2$ , which is metrizable, is separable [Examples 2.86 (4)]. Now a metrizable space is separable if and only if it is second-countable (Theorem 2.87), and a subspace of a second-countable space is itself second-countable (Proposition 2.78). Hence each subspace of  $l^\infty$  is separable and metrizable, and so the same is true of each space that is homeomorphic to a subspace of  $l^\infty$ .  $\square$

Recall from Definition 6.2 that a Tychonoff cube is any power  $I^S$  of the unit interval  $I = [0, 1]$ , provided with its product topology. In particular, the Hilbert cube  $l^\infty$ , being homeomorphic to the product space  $I^{\mathbb{N}^*}$ , is essentially a Tychonoff cube. It turns out that subspaces of Tychonoff cubes provide models for a wide class of topological spaces.

**6.44 Tychonoff Embedding Theorem.** A necessary and sufficient condition for a topological space to be a completely regular  $T_0$ -space is that it be embeddable into a Tychonoff cube.

**Proof.** Necessity. As a product of copies of the compact Hausdorff space  $I = [0, 1]$ , a Tychonoff cube is also a compact Hausdorff space and, a fortiori, a completely regular  $T_0$ -space. A subspace of a completely regular  $T_0$ -space is also a completely regular  $T_0$ -space. Hence a space homeomorphic to such a subspace is also a completely regular  $T_0$ -space.

Hilbert cube!embedding@and embe  
separable space!Hilbert cube@and H  
metrizable space!Hilbert cube@and  
Hilbert cube  
Tychonoff cube  
Hilbert cube!Tychonoff cube@and T

ze-subspaces-hilbert-cube-metrizable

thm:tychonoff-embedding

evaluation map  
normal space  
Urysohn's lemma  
evaluation map

**Sufficiency.** Let  $X$  be a completely regular  $T_0$ -space. Form the Tychonoff cube  $I^{C(X,I)}$ , where as usual  $C(X, I)$  denotes the set of all continuous functions from  $X$  to the unit interval  $I$ . We shall show that the *evaluation map*

$$e: X \rightarrow I^{C(X,I)} \\ x \mapsto \langle f(x) \rangle_{f \in C(X,I)}$$

is an embedding.

The evaluation map is injective: Let  $x$  and  $y$  be distinct points of  $X$ . Since continuous maps separate points in a completely regular  $T_0$ -space, there is some  $g \in C(X, I)$  with  $g(x) = 0$  and  $g(y) = 1$ . Then  $e(x) \neq e(y)$  because  $e(x)_g = g(x) \neq g(y) = e(y)_g$ .

The evaluation map is continuous: For each index  $f \in C(X, I)$  of the product space  $I^{C(X,I)}$ , the composite  $p_f \circ e: X \rightarrow I$  of the evaluation map with the  $f$ th projection  $p_f: I^{C(X,I)} \rightarrow I$  is just the map  $f$ , which is continuous. By [Theorem 3.62](#), the evaluation map  $e$  is continuous.

The evaluation map sends open sets to open sets in  $e(X)$ : We show, equivalently, that the image of each closed subset of  $X$  is closed in the range of  $e$ . Let  $F$  be a closed subset of  $X$ . Let  $z$  be an arbitrary point of  $e(X)$  not belonging to  $e(F)$ . Since  $e$  is injective, there is a unique  $x \in X$  with  $e(x) = z$ , so that  $z = \langle f(x) \rangle_{f \in C(X,I)}$ . Since  $X$  is completely regular, there is some  $g \in C(X, I)$  with  $g(x) = 0$  but  $g(y) = 1$  for all  $y \in F$ . Form the basic open set  $V = \times_{f \in C(X,I)} V_f$  with  $V_g = [0, 1[$  and  $V_f = I$  for all  $f \neq g$ . Then  $z = e(x) \in V$  and  $V$  is disjoint from  $e(F)$ .  $\square$

In the preceding proof, the power  $C(X, I)$  of the unit interval used to form the Tychonoff cube may be unnecessarily large. For example, if  $X = I$  itself, the proof embeds  $X$  into  $I^{C(I,I)}$ , whereas  $I$  can be embedded in  $I^1$ . In some situations, the power can be cut down drastically. This is the case, for example, when the space  $X$  is also second-countable.

thm:embed-2nd-count-reg

**6.45 Theorem.** A second-countable regular space can be embedded in the Tychonoff cube  $I^{\mathbb{N}^*}$ , or equivalently, in the Hilbert cube  $I^\infty$ .

**Proof.** Let  $\mathcal{B}$  be a countable base of the second-countable regular space  $X$ . Form the countable collection

$$\mathcal{S} = \{ \langle U, V \rangle : U, V \in \mathcal{B} \text{ and } \text{cls } U \subset V \}.$$

From [Theorem 6.20](#), the space  $X$  is normal. Then by Urysohn's Lemma ([6.26](#)), for each  $\langle U, V \rangle \in \mathcal{S}$ , there is a continuous function  $f_{U,V}: X \rightarrow I$  with  $f_{U,V}(x) = 0$  for all  $x \in \text{cls } U$  and  $f_{U,V}(y) = 1$  for all  $y \in X \setminus V$ . Form the Tychonoff cube  $I^{\mathcal{S}}$ . We shall show that the evaluation map

$$e: X \rightarrow I^{\mathcal{S}} \\ x \mapsto \langle f_{U,V}(x) \rangle_{\langle U,V \rangle \in \mathcal{S}}$$

is an embedding.

The evaluation map is injective: Let  $x$  and  $y$  be distinct points of  $X$ . Since  $X$  is a  $T_1$ -space, there is some  $V \in \mathcal{B}$  with  $x \in V$  but  $y \notin V$ . Since  $X$  is regular, there is some  $U \in \mathcal{B}$  with  $x \in U \subset \text{cls } U \subset V$ , so that  $\langle U, V \rangle \in \mathcal{S}$ . Then  $e(x) \neq e(y)$  because  $e(x)_{\langle U,V \rangle} = f_{U,V}(x) = 0$  whereas  $e(y)_{\langle U,V \rangle} = f_{U,V}(y) = 1$ .

The evaluation map is continuous: The proof is similar to that of the corresponding assertion in the proof of the Tychonoff Embedding Theorem ([6.44](#)). (See [Exercise 59](#).)

The evaluation map sends open sets to open sets in  $e(X)$ : We show, equivalently, that the image of each closed subset of  $X$  is closed in the range of  $e$ . Let  $F$  be a closed subset of  $X$ . Let  $z$  be an arbitrary point of  $e(X)$  not belonging to  $e(F)$ . Since  $e$  is injective, there is a unique  $x \in X$  with  $e(x) = z$ , so that  $z = \langle f_{U,V}(x) \rangle_{\langle U,V \rangle \in \mathcal{S}}$ . There is some  $\langle U', V' \rangle \in \mathcal{S}$  with  $x \in U'$  and  $\text{cls } V' \subset X \setminus F$ . Form the basic open set  $W = \times_{\langle U,V \rangle \in \mathcal{S}} W_{U,V}$  in  $l^{\mathcal{S}}$  given by  $W_{U',V'} = [0, 1[$  and  $W_{U,V} = ]$  for all  $\langle U, V \rangle \in \mathcal{S}$  with  $\langle U, V \rangle \neq \langle U', V' \rangle$ . Then  $e(x) \in W$  and  $W$  is disjoint from  $e(F)$ .  $\square$

embedding  
Alexandroff double circle

The Hilbert cube  $l^{\infty}$  is second-countable and regular, and so the same is true of all of its subspaces. This fact, together with the preceding theorem, establishes the following supplement to [Corollary 6.43](#).

r:characterize-subspaces-hilbert-cube

**6.46 Corollary.** *A necessary and sufficient condition for a topological space to be a second-countable regular space is that it be embeddable in the Hilbert cube  $l^{\infty}$ .*

If a space is separable and metrizable space, then it is a second-countable regular  $T_0$ -space. The converse follows immediately from the preceding [Theorem 6.45](#) and the fact that the Hilbert cube is metrizable.

thm:Urysohn-metrization

**6.47 Urysohn Metrization Theorem.** *A second-countable regular  $T_0$ -space is metrizable.*

Thus for a topological space  $X$ , the following are equivalent:

- (i)  $X$  is a second-countable regular  $T_0$ -space;
- (ii)  $X$  is a separable metrizable space; and
- (iii)  $X$  is embeddable in the Hilbert cube.

Regularity is certainly needed in the hypothesis of the Urysohn Metrization Theorem ([6.47](#)), since a metrizable space is necessarily regular. Second-countability is also needed there, as the following example demonstrates.

ex:Alexandroff-double-circle

**6.48 Example.** Let

$$C_1 = S_1(\mathbf{0}; d) = S_1, \quad C_2 = S_1(\mathbf{0}; d) = 2C_1$$

denote the concentric circles of radius 1 and 2, respectively, centered at the origin (where  $d$  is the Euclidean metric on the plane). Let  $h: C_1 \cong C_2$  be the homeomorphism  $w \mapsto 2w$  that maps each point on  $C_1$  radially outward to the corresponding point on  $C_2$ . The **Alexandroff double circle** is the set

$$X = C_1 \cup C_2$$

provided with its unique topology for which:

- each  $z \in C_2$  is isolated; and
- each  $z \in C_1$  has as a local base the collection  $\mathcal{B}_z = \{U_n(z) : n = 1, 2, 3, \dots\}$  where, for each positive integer  $n$ ,

$$U_n(z) = A_n(z) \cup h(A_n(z)) \setminus \{h(z)\}$$

with  $A_n(z)$  being the open arc on  $C_1$  of length  $2/n$  centered at  $z$ . (A typical  $U_n(z)$  is depicted in [CREFig:basic-nbd-C1-double-circle](#).)

metrizable space!second-countable space@and second-countable space  
 second-countable space!metrizable space@and metrizable space  
 plane with two origins  
 half-disk space!metrizable space@and metrizable space



Figure 6.1: A basic neighborhood  $U_n(z)$  at a point  $z \in C_1$  in the Alexandroff double circle. fig:basic-nbd-C1-double-circle

Evidently  $X$  is a Hausdorff space. Since the relative topology on  $C_1$  in  $X$  is its usual topology, which is compact, the entire space  $X$  is compact (see [Exercise 61](#)). Then  $X$  is a regular  $T_0$ -space. Since the subspace  $C_2$  of  $X$  is an uncountable discrete space,  $X$  is *not* separable and, a fortiori, not second-countable. Now a compact metrizable space is necessarily second-countable ([Theorem 4.44](#)). Hence  $X$  cannot be metrizable.  $\diamond$

Both hypotheses of regularity and second-countability are needed in the Urysohn Metrization Theorem ([6.47](#)), as the following examples show.

exs:2nd-count-T2-not-metrizable

**6.49 Examples.** (1) The **plane with two origins** is a second-countable Hausdorff space that is not regular and is not metrizable. Thus regularity

Let  $X = \mathbb{R}^2 \cup \{0'\}$  where  $0' \notin \mathbb{R}^2$ . Provide  $X$  with its topology in which:

- for  $x \neq 0, 0'$ , has as a local base the usual neighborhood system  $\mathcal{N}(x)$  at  $x$  and so also the collection  $\{B_\varepsilon(x; d) : \varepsilon > 0\}$  of  $d$ -balls, where  $d$  is the Euclidean metric;
- the point  $0$  has as a local base the collection  $\mathcal{B}_x = \{M_n : n = 1, 2, 3, \dots\}$  where

$$M_n = \{0\} \cup \{(x, y) \in B_0(1/n; d) : y > 0\}$$

for each positive integer  $n$ ; and

- the point  $0'$  has as a local base the collection  $\mathcal{B}'_x = \{M'_n : n = 1, 2, 3, \dots\}$  where

$$M'_n = \{0'\} \cup \{(x, y) \in B_0(1/n; d) : y < 0\}$$

for each positive integer  $n$ .

- (2) The half-disk space  $X = H \cup L$  [[Examples 2.25 \(3\)](#)] is a Hausdorff space that is both separable ([Example 2.88](#)) and first countable [[Examples 2.66 \(2\)](#)]. However, the space is not metrizable because it is not regular [[Examples 2.99 \(4\)](#)].  $\diamond$

The continuous image of a metrizable space, even of a separable metrizable space, need not be metrizable. For example, let  $X = \mathbb{N}$  with its usual (discrete) topology and let  $Y = \mathbb{N}$

with its finite-complement topology. Then  $X$  is a separable metrizable space and the identity map  $X \rightarrow Y$  is a continuous surjection, but  $Y$  is not metrizable since it is not a Hausdorff space.

EXAMPLE: continuous surjection  $f: X \rightarrow Y$  of separable metrizable space to non-metrizable Hausdorff space??

What about  $i: \mathbb{R} \rightarrow \mathbb{R}_l$ ?

**6.50 Theorem.** *A continuous compact Hausdorff image of a separable metrizable space is itself metrizable.*

**6.51 Theorem.** *A continuous Hausdorff image of a compact metrizable space is itself metrizable.*

**Proof.** Let  $f: X \rightarrow Y$  be a continuous surjection from a compact metrizable space  $X$  to a Hausdorff space  $Y$ . Then  $Y$  is a compact and so is a regular  $T_0$ -space. In view of the Urysohn Metrization Theorem (6.47), it suffices to prove that  $Y$  is second-countable.

Choose a countable base  $\mathcal{B}$  of  $X$  and define  $\mathcal{C} = \{f(B) : B \in \mathcal{B}\}$ .

Although members of  $\mathcal{C}$  need not be open in  $Y$ , the collection  $\mathcal{C}$  has the following property:

{eq:image-network} (\*) for each  $y \in Y$  and each neighborhood  $V$  of  $y$ , there is a  $C \in \mathcal{C}$  with  $y \in C \subset V$ .

In fact, equation (\*), let  $V$  be a neighborhood of a point  $y \in Y$ . Choose some  $x \in X$  with  $f(x) = y$ . Then  $f^{-1}(V)$  is a neighborhood of  $x$  in  $X$ , and so there is a member  $B$  of the base  $\mathcal{B}$  with  $x \in B \subset f^{-1}(V)$ . It follows that  $y \in f(B) \subset V$ .

Now define

$\mathcal{P} = \{\langle C, D \rangle \in \mathcal{C} \times \mathcal{C} : \text{there are disjoint open neighborhoods } V, W \text{ of } C, D, \text{ respectively}\}.$

For each pair  $\langle C, D \rangle$  in  $\mathcal{P}$ , choose disjoint open neighborhoods  $V_{C,D}$ ,  $W_{C,D}$  of  $C, D$ , respectively. Define  $\mathcal{S}$  to be the collection of all these open sets  $V_{C,D}$  and  $W_{C,D}$ . Since  $\mathcal{B}$  covers  $X$ , the collection  $\mathcal{C}$  covers  $Y$ . Then  $\mathcal{S}$  is a subbase of a topology  $\mathcal{T}'$  on  $Y$ , that is,  $\mathcal{T}'$  be the topology generated by  $\mathcal{S}$ . Since  $\mathcal{S}$  is countable, the base generated by  $\mathcal{S}$  is also countable. Thus  $(Y, \mathcal{T}')$  is second-countable.

It remains to show only that  $\mathcal{T}'$  is the same as the given topology  $\mathcal{T}$  of  $Y$ . Now  $\langle Y, \mathcal{T}' \rangle$  is a Hausdorff space, whereas  $\langle Y, \mathcal{T} \rangle$  is a compact Hausdorff space. From equation (\*), the topology  $\mathcal{T}'$  contains the given topology  $\mathcal{T}$  of  $Y$ . Since a compact Hausdorff topology on a given set is minimally Hausdorff (Proposition 4.26), it follows that  $\mathcal{T} = \mathcal{T}'$ .  $\square$

### EXERCISES FOR SECTION 6.3

**57.** Prove that a nonnormal topological space can be embedded into a normal space by adding a single point to the space. If the nonnormal space is a  $T_1$ -space, will that normal space be a  $T_1$ -space?

-space-loc-metrizable-Urysohn-thm

**58.** (Continuation of Exercise 30.)

Apply the Urysohn Metrization Theorem (6.47) to prove anew that the tangent disk space  $\Gamma$  (Exercise 2.37) is locally metrizable. (Note: A direct construction of compatible metrics on neighborhoods in  $\Gamma$  was outlined in the hint to Exercise 2.100.)

(Hint: Use Exercise 2.152.)

(Note: A direct construction of compatible metrics on neighborhoods of points of  $\Gamma$  was outlined in the hint to [Exercise 2.100](#).)

59. Verify that the evaluation map used in the proof of [Theorem 6.45](#) is continuous.

60. The deleted sequence space  $\mathbb{R}_K$  [[Exercise 2.92](#)] is a second-countable Hausdorff space that is not regular and hence is not metrizable.

61. Verify that the Alexandroff double circle space is compact, as asserted in [Example 6.48](#).

## 6.4 Urysohn Metrization Theorem

MORE

### EXERCISES FOR SECTION 6.4

62.

63. (a)  
(b)

64.

## 6.5 Compactification

MORE

### EXERCISES FOR SECTION 6.5

65.

66. (a)  
(b)

67.

## 6.6 Dimension

Peano's unexpected example ([5.65](#)) of a plane-filling curve shattered the intuitive notion of the dimension as being the least number of continuous parameters needed to describe a space, and it precipitated a search for a rigorous definition of dimension.

We are accustomed to referring to a line as being one-dimensional and a plane as being two-dimensional, but what does that mean? A Euclidean space  $\mathbb{R}^n$  is said to be  $n$ -dimensional because, among other reasons, as a vector space it has a basis consisting of  $n$  vectors. However, here we are interested in a different sense of "dimension," one that is meaningful even for topological spaces that do not happen to be vector spaces.

It is not unreasonable to expect a topological dimension function,  $\text{Dim}$ , with domain a suitable class of topological spaces, to have at least the following properties:

(Dim-1)  $\text{Dim}(X)$  is an integer or the symbol  $\infty$ ;<sup>2</sup>

(Dim-2) Invariance property: if  $X \cong Y$ , then  $\text{Dim}(X) = \text{Dim}(Y)$ ;

<sup>2</sup>For some purposes it is desirable to allow even fractional dimensions: see CREF



property:dim-subspace	(Dim-3) Subspace property: <i>If <math>Y</math> is a subspace of <math>X</math>, then <math>\text{Dim}(Y) \leq \text{Dim}(X)</math>;</i>
property:dim-Rn	(Dim-4) $\text{Dim}(\mathbb{R}^n) = n$ for every $n \geq 0$ .
property:dim-sum-2	(Dim-5) Sum property: <i>If <math>X = E \cup F</math> for closed subspaces <math>E</math> and <math>F</math> with <math>\text{Dim}(E) \leq n</math> and <math>\text{Dim}(F) \leq n</math>, then <math>\text{Dim}(X) \leq n</math>.</i>

Brouwer, Luitzen Egbertus Jan  
Lebesgue, Henri  
Urysohn, Pavel  
Menger, Karl

Note: Hilbert cube is not the union of a sequence of zero-dimensional subspaces, so the decomposition property does not hold when  $n = \infty$ .

The product of a sequence of regular  $T_0$ -spaces is zero-dimensional if and only if each of the factors is zero-dimensional.

We have  $\mathbb{R}^n = \bigcup_{j=0}^n Q_j$  where  $Q_j$  is the set of those points in  $\mathbb{R}^n$  having exactly  $j$  rational coordinates. Now  $Q_j$  is zero-dimensional for each  $j$ . Hence  $\text{ind}(\mathbb{R}^n) \leq n$ .

The first satisfactory definition of such a dimension function was given in 1913 by L. E. J. Brouwer, who developed an idea of H. Lebesgue. A different, but equivalent, definition, which coincides with Brouwer's for a large class of spaces (namely, locally connected, compact metrizable spaces), was discovered independently by P. Urysohn and K. Menger in 1922 and 1923, respectively.

Here we just introduce that Urysohn-Menger definition of dimension. (A detailed development of appears in Hurewicz and Wallman [37]; a more up-to-date treatment, including other versions of topological dimension, appears in Engelking [23]. The application of dimension theory to curves can be found in Blumenthal and Menger [6, Part 4].)

intuit:ind **Intuitive idea—Urysohn-Menger dimension.** Let us deem that a finite discrete space has dimension 0, as each point there is isolated from all others. Then we deem the real line  $\mathbb{R}$  to have dimension 1, because any given point  $x \in \mathbb{R}$  can be “trapped” inside an arbitrarily small neighborhood that is an interval  $]a, b[$ , and the boundary of  $]a, b[$  is the two-point discrete space  $\{a, b\}$ , which has dimension 0. Similarly, we may deem a circle or a square to have dimension 1.

Next, we deem the plane  $\mathbb{R}^2$  to have dimension 2, because any given point  $z \in \mathbb{R}^2$  can be “trapped” inside an arbitrarily small neighborhood that is a disk (or a square region), whose boundary is a circle (or the union of the four sides of a square, respectively), which has dimension 1. Similarly, we deem a sphere or the boundary of a cube in  $\mathbb{R}^3$  to have dimension 2.

Next, we deem space  $\mathbb{R}^3$  to have dimension 3, because any given point  $z \in \mathbb{R}^3$  can be “trapped” inside an arbitrarily small neighborhood that is a 3-disk (or a cube), whose boundary is a sphere (or the union of the eight planar faces of a cube, respectively), which has dimension 2.

Dimension 0 fits nicely into this scheme: each point in a finite discrete space has arbitrarily small neighborhoods whose boundaries are empty. Accordingly, we should assign dimension  $-1$  to the empty space!

The definition of dimension will, accordingly, be recursive. (Perversely, the function to be defined is known as an *inductive* dimension.)

large inductive dim  
dimension!inductiv  
dimension!large inc  
zero-dimensional s  
small inductive dim  
Euclidean n-space@  
real line!dimension!  
Euclidean n-space@  
invariant!topologic  
topological invarian

**6.52 Definition.** The **small inductive** (or **Urysohn-Menger**) **dimension** of a topological space  $X$ , denoted by  $\text{ind}(X)$ , is defined recursively as follows.

The empty topological space is said to have **small inductive dimension**  $-1$ , that is,  $\text{ind}(\emptyset) = -1$ .

For  $n \geq 0$ , a topological space  $X$  is said to have **small inductive dimension**  $n$ , written  $\text{ind}(X) = n$ , when  $n$  is the *least* integer  $k \geq -1$ —if there is such a  $k$ —such that at each point of  $X$ , there is a local base consisting of sets whose boundaries have small inductive dimension at most  $k - 1$ . In this case,  $X$  is said to have **finite** small inductive dimension. In case there is no such  $k$ , the space  $X$  is said to have **infinite** small inductive dimension, written  $\text{ind}(X) = \infty$ .

The qualifier “small” in “small inductive dimension,” and the lower-case “i” in “ind,” are used because there is also a *large inductive dimension*, denoted by  $\text{Ind}$ , whose definition we do not give here. (But see [Exercise 69](#).)

By its very definition, the function  $\text{ind}$  satisfies property [\(Dim-1\)](#).

**6.53 Examples.** (1) A space  $X$  is **zero-dimensional** in the sense of [Definition 2.61](#) *precisely when*  $\text{ind}(X) = 0$ .

(2) The real line has small inductive dimension 1:  $\text{ind}(\mathbb{R}) = 1$ .

To see this, first note that  $\text{ind}(\mathbb{R}) \neq -1$  since  $\mathbb{R}$  is nonempty. Next, from [Examples 2.63 \(4\)](#),  $\text{ind}(\mathbb{R}) \neq 0$ . Finally,  $\text{ind}(\mathbb{R}) \leq 1$  because each point of  $\mathbb{R}$  has arbitrarily small neighborhoods that are open intervals  $]a, b[$ , and  $\text{ind}(\text{bdy } ]a, b[) = \text{ind}(\{a, b\}) = 0$ .

(3) A circle  $S_r(z; d)$  in  $\mathbb{R}^2$ , where  $d$  is the Euclidean metric, has small inductive dimension 1. The reason is similar to that in (2), except that circular arcs take the place of intervals.

(4) The plane  $\mathbb{R}^2$  has  $\text{ind}(\mathbb{R}^2) = 2$ .

In fact, it is easy to see that  $\text{ind}(\mathbb{R}^2) \leq 2$ : each point  $z \in \mathbb{R}^2$  has arbitrarily small circular neighborhoods  $B_\varepsilon(z; d)$ , where  $d$  is the Euclidean metric, and from (3) we have  $\text{ind}(\text{bdy}(B_\varepsilon(z; d))) = \text{ind}(S_\varepsilon(z; d)) = 1$ .

It is *not* easy to show that  $\text{ind}(\mathbb{R}^2) \not\leq 1$ . Certainly  $\text{ind}(\mathbb{R}^2) \neq -1$ ; that  $\text{ind}(\mathbb{R}^2) \neq 0$  will follow from the fact that  $\mathbb{R}^2$  is connected [[Examples 5.21 \(1\)](#)] along with [Proposition 2.64](#). However, showing that  $\text{ind}(\mathbb{R}^2) \neq 1$  requires a lot more work.

(5) In general, Euclidean  $n$ -space  $\mathbb{R}^n$  does have  $\text{ind}(\mathbb{R}^n) = n$ , in other words, the function  $\text{ind}$  satisfies property [\(Dim-4\)](#).

In fact, it is easy to see that  $\text{ind}(\mathbb{R}^n) \leq n$ : use induction. However, as for the case  $n = 2$ , it is not easy to show that  $\text{ind}(\mathbb{R}^n) \not\leq n - 1$ ; doing so requires techniques beyond those at our disposal here!  $\diamond$

**Having small inductive dimension  $n$  is, as desired, a topological property:** if  $X$  is homeomorphic to  $Y$ , then  $\text{ind } X = \text{ind } Y$ ; in other words,  $\text{ind}$  satisfies [\(Dim-2\)](#). Now the small inductive dimension of a space is a numeric object, namely, an integer at least  $-1$  or else  $\infty$ ; since having small inductive dimension  $n$  is a topological property, we say that  $\text{ind}$  is a **topological invariant**.

For the purpose of comparing small inductive dimensions of spaces, we stipulate that  $n < \infty$  for each integer  $n$  with  $n \geq -1$ . In other words, if  $Y$  is a space having finite small

inductive dimension and if  $X$  is a space having infinite small inductive dimension, then  $\text{ind}(Y) < \text{ind}(X)$ .

According to the next proposition, the function  $\text{ind}$  satisfies property (Dim-3).

prop:cf-ind-subspace

**6.54 Proposition.** *If  $Y$  is a subspace of a topological space  $X$ , then  $\text{ind}(Y) \leq \text{ind}(X)$ .*

**Proof.** There is nothing to prove in the case  $\text{ind}(X) = \infty$ . We use complete induction on integers  $n \geq -1$  to prove that, for every regular space  $X$  with  $\text{ind}(X) = n$ , we have  $\text{ind}(Y) \leq n$  for every subspace  $Y$  of  $X$ .

Base step:  $n = -1$ . That  $\text{dim}(X) = -1$  means  $X$  is empty; in this case the same is true of each subspace of  $X$ .

Inductive step. Now let  $n \geq 0$  and suppose that  $\text{ind}(S) \leq \text{ind}(Z)$  for every subspace  $S$  of every space  $Z$  with  $\text{ind}(Z) < n$ . Let  $X$  be a topological space with  $\text{ind}(X) = n$  and let  $Y$  be a subspace of  $X$ . We want to deduce that  $\text{ind}(Y) \leq n$ .

Let  $y$  be an arbitrary point in  $Y$  and let  $V$  be an open neighborhood of  $y$  in  $Y$ . There is an open set  $U$  in  $X$  with  $V = Y \cap U$ . Since  $\text{ind}(X) = n$ , there is an open set  $W'$  in  $X$  with

$$y \in W' \subset U, \quad \text{ind}(\text{bdy}_X W') \leq n - 1.$$

Let  $W = Y \cap W'$ , so that  $W$  is a neighborhood of  $y$  in  $Y$  with  $W \subset V$ . Now

$$\begin{aligned} \text{bdy}_Y W &\subset \text{bdy}_X W = \text{bdy}_X (Y \cap W') \\ &= \text{cls}_X (Y \cap W') \cap \text{cls}_X (Y \setminus W') \\ &\subset \text{cls}_X W' \cap \text{cls}_X (X \setminus W') = \text{bdy}_X W'. \end{aligned}$$

Since  $\text{ind}(X) = n$ , we have  $\text{ind}(\text{bdy}_X W') \leq n - 1$ . By the inductive assumption,  $\text{ind}(\text{bdy}_Y W) \leq n - 1$ , too. This completes the proof that  $\text{ind}(Y) \leq n$ .  $\square$

prop:ind-sum-2-thm

**6.55 Proposition.** *If  $E$  and  $F$  are closed subspaces of a space  $X$  with  $X = E \cup F$  and  $\text{ind}(E) \leq n$  and  $\text{ind} F \leq n$ , then  $\text{ind}(X) \leq n$ .*

Thus the function  $\text{ind}$  satisfies property (Dim-5). In fact, small inductive dimension satisfies a stronger sum property:

property:dim-sum

(Dim-5\*) *If  $X = \bigcup_{j=0}^{\infty} F_j$  for a sequence  $\langle F_j \rangle_{j \in \mathbb{N}}$  of closed subspaces of  $X$  with  $\text{Dim}(F_j) \leq n$  for each  $j$ , then  $\text{dim}(X) \leq n$ .*

The classical theory of inductive dimension concerned separable metrizable spaces. When mathematicians tried to extend results about dimension to a wider class of topological spaces, it turned out that small inductive dimension did not allow this. Instead, a different definition of dimension, *covering dimension*, which agrees with small inductive dimension for separable metric spaces, was appropriate.

It so happens that, for topological purposes, neither small inductive dimension ( $\text{ind}$ ) nor large inductive dimension ( $\text{Ind}$ ) suffices. Rather, the preferable notion of dimension is the following one, introduced by Hausdorff and developed by Besicovitch.

dimension

one-dimensional space  
dimension

**6.56 Definition.** The order (or ply) of a cover  $\mathcal{U}$  of a set  $X$  is the smallest integer  $k \geq -1$ , if such  $k$  exists, such that each point of the space belongs to at most  $k$  members of the cover. In other terms it is the smallest such  $k$  such that the intersection of each  $k + 1$  members of  $\mathcal{U}$  have empty intersection. If such  $k$  does not exist, then it is said that the cover has infinite order (or ply).

The **covering dimension** of a topological space  $X$ , denoted by  $\dim(X)$  (or sometimes by  $\text{cov}(X)$ ) is the least integer  $n$ , if such  $n$  exists, such that every open cover of  $X$  has a refinement of order  $n + 1$ .

...

### EXERCISES FOR SECTION 6.6

prob:1-dim **68.** Verify that the following are one-dimensional in the sense of having small inductive dimension 1:

- (a) A polygonal path in  $\mathbb{R}^2$  (Exercise 5.72).
- (b) The topologist's sine curve [Examples 5.57 (2)].
- (c) A manifold that is 1-dimensional in the sense of Definition 3.40.

*Note:* Small inductive dimension can serve to distinguish  $\mathbb{I}^2$  from  $\mathbb{I}$  topologically: compare Examples 6.53. It is possible to prove, for example, that a compact connected subset of the plane  $\mathbb{R}^2$  is one-dimensional if and only if it is nowhere dense in  $\mathbb{R}^2$ : see Blumenthal and Menger [6, Section 14]. Of course, there is a simpler way to distinguish  $\mathbb{I}^2$  from  $\mathbb{I}$ : see Exercise 5.8.

prob:Ind **69.** The **large inductive dimension**  $\text{Ind}(X)$  of a topological space is defined recursively as follows.

Fundamental Group and Covering Spaces

chap:fundamentalcovering

Chapter Contents

7.1 The fundamental group . . . . .	643	7.2 Covering spaces . . . . .	643
Definition of the fundamental group . .	643	Definition of a covering space . . . . .	643
Exercises for Section 7.1 . . . . .	643	Exercises for Section 7.2 . . . . .	643

Introduction

This chapter . . .

fix: Write fund'l group & covering spaces intro

7.1 The fundamental group

sec:fund-grp

Definition of the fundamental group

fix: Write section "The fundamental group"

EXERCISES FOR SECTION 7.1

- 1. problem
- 2.

7.2 Covering spaces

sec:covering-spaces

Definition of a covering space

fix: Write section "Covering spaces"

EXERCISES FOR SECTION 7.2

- 3. problem
- 4.

statement@statement  
 $\mathbb{R} @ \mathbb{R} @ \mathbb{R} @ \mathbb{R}$   
 real numbers!zzzz@  
 collection!zzzz@  
 class!zzzz@  
 set!zzzz@  
 inclusion!zzzz@  
 binary relation@binary relation  
 relation!zzzz@  
 relation!inverse@relation!inverse  
 induction!zzzz@  
 axioms!zzzz@  
 proof by induction!zzzz@  
 Well-ordering Principle!zzzz@  
 composite of maps!zzzz@  
 function!zzzz@  
 injective map!zzzz@  
 one-to-one@one-to-one  
 surjective map!zzzz@  
 map onto@map onto  
 bijective map!zzzz@  
 one-to-one correspondence!zzzz@  
 direct image@direct image  
 image of a set!inverse@image of a set!inverse  
 lifting@lifting  
 Euclidean space@Euclidean space  
 Euclidean  $n$ -space@Euclidean  $n$ -space!zzzz@  
 $n$ -dimensional Euclidean space@ $n$ -dimensional Euclidean space@ $n$ -dimensional Euclidean space@ $n$ -dimensional Euclidean space  
 Euclidean space  
 Euclidean plane!zzzz@  
 covering@covering  
 Cartesian product@Cartesian product  
 Cartesian product@Cartesian product  
 disjoint sets!zzzz@  
 Principle of Ordinary Recursion!zzzz@  
 recursion!strong@recursion!strong  
 Principle of Complete Recursion!zzzz@  
 recursive definition@recursive definition  
 sequence!recursive@sequence!recursive  
 infinite set!zzzz@  
 Cantor-Schroeder-Bernstein Theorem@Cantor-Schröder-Bernstein Theorem@Cantor-Schroeder-Bernstein Theorem@Cantor-Schröder-Bernstein Theorem  
 base expansion!zzzz@  
 order relation@order relation  
 ordering relation@ordering relation  
 preordered set!zzzz@  
 preorder@preorder  
 partially ordered set!zzzz@  
 reflexivity@reflexivity  
 transitivity@transitivity  
 irreflexivity@irreflexivity  
 irreflexivity@irreflexivity  
 strict ordering@strict ordering  
 weak order@weak order  
 partial order@partial order  
 greatest element!zzzz@  
 least element!zzzz@  
 largest element@largest element  
 smallest element@smallest element  
 order-preserving map!zzzz@  
 order-reversing map!zzzz@  
 interval!zzzz@  
 upper bound!least!zzzz@  
 lower bound!greatest!zzzz@  
 expansion@expansion  
 totally ordered set!zzzz@  
 lexicographically ordered product!zzzz@  
 chain!zzzz@  
 bound@bound  
 dictionary ordering@dictionary ordering  
 complete@complete  
 quotient set!zzzz@  
 partition!zzzz@  
 ordinal number@ordinal number

## Guide to the Exercises

chap:exercise-guide

This guide lists exercises containing definitions, results, and examples that are not included in the body of the text but are needed to solve subsequent exercises or to read subsequent sections of the text. Not listed here are exercises needed for immediately following exercises in the same end-of-section set.

Exercise	Needed for Exercises	Needed for Text
1.99	2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99	pages 235–999, 1058–2456
1.99	2.99	page 23
1.99	2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99	pages 235–999, 1058–2456
1.99	2.99	page 23
1.99	2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99	pages 235–999, 1058–2456
1.99	2.99	page 23
1.99	2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99	pages 235–999, 1058–2456
1.99	2.99	page 23
1.99	2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99	pages 235–999, 1058–2456
1.99	2.99	page 23
1.99	2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99	pages 235–999, 1058–2456
1.99	2.99	page 23
1.99	2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99	pages 235–999, 1058–2456
1.99	2.99	page 23
1.99	2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99	pages 235–999, 1058–2456
1.99	2.99	page 23
1.99	2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99, 2.99	pages 235–999, 1058–2456
1.99	2.99	page 23

[illegible]



# Bibliography

bib:main

- RDAndersonHandLines1966 [1] R. D. Anderson, *Hilbert space is homeomorphic to the countable infinite product of lines*, Bulletin American Mathematical Society **72** (1966), 515–519. [Cited on page 374.]
- Apostol1969vol2 [2] Tom M. Apostol, *Calculus: Multi-Variable Calculus and Linear Algebra with Applications to Differential Equations and Probability*, 2 ed., vol. 2, Wiley, 1969. [Cited on page 587.]
- Bernays1991 [3] Paul Bernays, *Axiomatic Set Theory*, 2 ed., Dover, New York, 1991; republication of 2 ed., North-Holland, 1958. [Cited on pages 2, 651.]
- Besalgić1985 [4] Amer Bešliagić, *A Dowker product*, Transactions of the American Mathematical Society **292** (1985), 519–530.
- BirkhoffMacLane1997 [5] Garrett Birkhoff and Saunders Mac Lane, *A Survey of Modern Algebra*, 5 ed., A K Peters/CRC Press, 1997. [Cited on pages 88, 403, 600.]
- BlumenthalMenger1970 [6] Leonard M. Blumenthal and Karl Menger, *Studies in Geometry*, W. H. Freeman, San Francisco, 1970. [Cited on pages 639, 642.]
- BoasPrimerReal1996 [7] Ralph P. Boas, Jr. and Harold P. Boas, *A Primer of Real Functions*, 4 ed., Carus Mathematical Monographs 13, Mathematical Association of America, Washington, D. C., 1996. [Cited on page 215.]
- BourbakiTopology1989 [8] Nicholas Bourbaki, *General Topology*, 2 vols., Elements of Mathematics, Springer-Verlag, New York, 1989. [Cited on pages 607, 651, 656.]
- BourbakiSets2004 [9] ———, *Theory of Sets*, Elements of Mathematics, Springer-Verlag, New York, 2004. [Cited on pages 2, 15, 651.]
- topologieAlgebriqueChaps1to4\_2016 [10] ———, *Topologie Algébrique, Chapitres 1 à 4*, Éléments de Mathématique, Springer-Verlag, New York, 2016.
- BoyceDiPrimaElemDiffEqs11 [11] William E. Boyce, Richard C. DiPrima, and Douglas B. Meade, *Elementary Differential Equations*, Wiley, New York, 2017. [Cited on page 221.]
- britannica:Topology [12] Sephan C. Carlson, *Topology*, 2016, <https://www.britannica.com/topic/topology> [accessed 2016-09-21]. [Cited on page 651.]
- Chaber1972 [13] J. Chaber, *Remarks on open-closed mappings*, Fundamenta Mathematicae **74** (1972), no. 3, 197–208. [Cited on page 409.]
- KConradIsometries [14] Keith Conrad, *Isometries of  $\mathbf{R}^n$* , <http://www.math.uconn.edu/~kconrad/blurbs/grouptheory/isometryRn.pdf> [accessed 2017-05-13]. [Cited on pages 173, 174.]
- MacTutorMathHistory [15] ———, *Robertson, Edmund and O'Connor, John*, <https://mathshistory.st-andrews.ac.uk> [accessed 2020-07-30].

- ChrisCooperIsometries [16] Christopher Cooper, *Isometries*, 2016, <http://maths.mq.edu.au/~chris/geometry/CHAP06%5C%20Isometries.pdf> [accessed 2017-05-13]. [Cited on page 173.]
- ChrisCooperIsometries3D [17] ———, *Isometries in 3 dimensions*, 2016, <http://maths.mq.edu.au/~chris/geometry/CHAP08%5C%20Isometries%5C%20in%5C%203%5C%20Dimensions.pdf> [accessed 2017-05-13]. [Cited on page 173.]
- Dowker1951 [18] C. H. Dowker, *On countably paracompact spaces*, Canadian Journal of Mathematics **3** (1951), 219–224.
- Dugundji1966 [19] James Dugundji, *Topology*, Allyn and Bacon, Boston, 1966.
- EdgarMeasureTopFractal2008 [20] Gerald Edgar, *Measure, Topology, and Fractal Geometry*, 2 ed., Springer Science-Business Media, New York, 2008. [Cited on page 651.]
- EisenbergSets1971 [21] Murray Eisenberg, *Axiomatic Theory of Sets and Classes*, Holt, Rinehart and Winston, New York, 1971. [Cited on page 651.]
- EisenbergMathMethod1996 [22] ———, *The Mathematical Method: A Transition to Advanced Mathematics*, Prentice Hall, Upper Saddle River, NJ, 1996. [Cited on pages 2, 88, 651.]
- EngelkingDimension1978 [23] Ryszard Engelking, *Dimension Theory*, North-Holland, Amsterdam, 1978; revised and enlarged translation from Polish original published by WPN-Polish Scientific Publishers, 1977. [Cited on pages 639, 651.]
- FineRosenberger1997 [24] Benjamin Fine and Gerhard Rosenberger, *The Fundamental Theorem of algebra*, Springer-Verlag New York, Springer-Verlag New York, 1997. [Cited on page 600.]
- FlemingSeveralVars2Ed [25] Wendell Fleming, *Functions of Several Variables*, 2 ed., Springer-Verlag New York, 1977. [Cited on pages 173, 174.]
- GaillardZorn2017 [26] Pierre-Yves Gaillard, *Selected Texts*, 2017, [http://www.iecl.univ-lorraine.fr/~Pierre-Yves.Gaillard/DIVERS/Selected\\_Texts/selected-texts-m.pdf](http://www.iecl.univ-lorraine.fr/~Pierre-Yves.Gaillard/DIVERS/Selected_Texts/selected-texts-m.pdf) [accessed 2017-06-21]. [Cited on page 113.]
- GamellnAndGreene1999 [27] Theodore W. Gamelin and Robert Everist Greene, *Introduction to Topology*, 2 ed., Dover, Mineola, NY, 1999.
- GarlingAnalysisCourseVol1 [28] D. J. H. Garling, *A Course in Mathematical Analysis, Vol. I: Foundations and Elementary Real Analysis*, Cambridge University Press, Cambridge, 2013. [Cited on page 88.]
- GoldbergReal1976 [29] Richard R. Goldberg, *Methods of Real Analysis*, 2 ed., Wiley, New York, 1976. [Cited on pages 215, 217.]
- GraysonZorn2007 [30] Daniel R. Grayson, *Zorn's Lemma*, 2007, <http://www.math.uiuc.edu/~dan/ShortProofs/Zorn.pdf> [accessed 2017-06-21]. [Cited on page 113.]
- GreeverTheoryExamplesTopology1969 [31] John Greever, *Theory and Examples of Point-Set Topology*, Brooks/Cole, Belmont, CA, 1968. [Cited on page 651.]
- HallSpencer2018 [32] Dick Wick Hall and Guilford L. Spencer II, *Elementary Topology: A First Course in Point Set Topology*, Blue Collar Scholar/Kindle Publishing, 2018; new introduction and references by K. Maestro; originally published by Wiley, 1955. [Cited on page 592.]
- HatcherRecommend [33] Allen Hatcher, *A List of Recommended Books in Topology*, <https://www.math.cornell.edu/~hatcher/Other/topologybooks.pdf> [accessed 2016-09-21]. [Cited on page 651.]
- HatcherAT2002 [34] ———, *Algebraic Topology*, Cambridge University Press, 2002. [Cited on page 342.]
- HildebrandtContNoDeriv1933 [35] T. H. Hildebrandt, *A simple continuous function with a finite derivative at no point*, American Mathematical Monthly **40** (1933), no. 9, 547–548. [Cited on page 217.]

- HockingYoung1988 [36] John G. Hocking and Gail S. Young, *Topology*, Dover, New York, 1988; originally published by Addison-Wesley, 1961. [Cited on pages 565, 651, 655.]
- HurewiczWallman1941 [37] Witold Hurewicz and Henry Wallman, *Dimension Theory*, Princeton Univ. Press, Princeton, NJ, 1941. [Cited on pages 639, 651.]
- JacobBaileyLinearAlgebra [38] Henry G. Jacob and Duane W. Bailey, *Linear Algebra*, Houghton Mifflin, 1971. [Cited on page 173.]
- JamesHistory1999 [39] I. M. James, ed., *History of Topology*, Elsevier, Amsterdam, 1999. [Cited on page 651.]
- KelleyTopology2017 [40] John L. Kelley, *General Topology*, Dover, New York, 2017; originally published by Van Nostrand Reinhold, 1955. [Cited on pages 455, 651, 656.]
- KneserZorn1950 [41] Helmuth Kneser, *Eine direkte Ableitung des Zornschen Lemmas aus dem Auswahlaxiom*, *Mathematische Zeitschrift* **53** (1950), 110–113, <http://eudml.org/doc/169173>. [Cited on page 113.]
- LakinsMathReasoning2016 [42] Tamara J. Lakins, *The Tools of Mathematical Reasoning*, Pure and Applied Undergraduate Texts 26, American Mathematical Society, Providence, RI, 2016. [Cited on pages 2, 651.]
- LangAnalysisI1968 [43] Serge Lang, *Analysis I*, Addison-Wesley, Reading, MA, 1968. [Cited on page 600.]
- LeeTopManifolds2011 [44] John M. Lee, *Introduction to Topological Manifolds*, 2 ed., Graduate Texts in Mathematics 202, Springer, New York, 2011. [Cited on page 651.]
- ManheimGenesis1964 [45] Jerome H. Manheim, *The Genesis of Point Set Topology*, Pergamon/Macmillan, New York, 1964. [Cited on page 651.]
- MathewsHowell2012 [46] John H. Mathews and Russell W. Howell, *Complex Analysis for Mathematics and Engineering*, 6 ed., Jones & Bartlett Learning, Sudbury, MA, 2012. [Cited on page 600.]
- McCarthyContNowhereDiffFn1953 [47] John McCarthy, *An everywhere continuous nowhere differentiable function*, *American Mathematical Monthly* **60** (1953), no. 10, 709. [Cited on page 217.]
- McCartyTopology1988 [48] George McCarty, *Topology: An Introduction with Application to Topological Groups*, Dover, New York, 1988; originally published by McGraw-Hill, 1967.
- MesserTopologyNow2006 [49] Robert Messer and Philip Philip Straffin, *Topology Now!*, Mathematical Association of America, Washington, DC, 2006.
- MunkresAlgebraicTopology2018 [50] James R. Munkres, *Elements of Algebraic Topology*, CRC Press, Boca Raton, FL, 2018; originally published by Addison Wesley, 1984. [Cited on page 342.]
- MunkresTopology2000 [51] ———, *Topology*, 2 ed., Pearson, New York, 2000; originally published by Prentice Hall. [Cited on pages 651, 655, 656.]
- Novak1953 [52] J. Novák, *On the Cartesian product of two compact spaces*, *Fundamenta Mathematicae* **40** (1953), 106–112. [Cited on page 508.]
- PinterSets2014 [53] Charles C. Pinter, *A Book of Set Theory*, Dover, Mineola, NY, 2014; revision of original published by Addison-Wesley, 1971. [Cited on pages 2, 651.]
- Roe2015 [54] John Roe, *Winding Around: The Winding Number in Topology, Geometry, and Analysis*, Student Mathematical Library 76, American Mathematical Society, Providence, RI, 2015.
- RosenbaumProjectiveGeometry1963 [55] Robert A. Rosenbaum, *Introduction to Projective Geometry and Modern Algebra*, Addison-Wesley, San Francisco, 1963. [Cited on page 403.]
- RubinSets1967 [56] Jean E. Rubin, *Set theory for the mathematician*, Holden-Day, Reading, MA, 1967. [Cited on pages 2, 123.]

- RudinMathAnalysis1976 [57] Walter Rudin, *Principles of Mathematical Analysis*, 3 ed., McGraw-Hill, New York, 1976. [Cited on pages 88, 220.]
- ScottBM1978 [58] Brian M. Scott, A “more topological” proof of the Tietze-Urysohn theorem, *American Mathematical Monthly* **85** (1978), 192–193. [Cited on page 656.]
- ScottOnChaber2015 [59] ———, *An example of open-closed continuous image of  $T_2$ -space that is not  $T_2$* , version 2023-06-21, 2015, <https://math.stackexchange.com/a/1413128/32337> [accessed 2023-07-14]. [Cited on page 409.]
- SinghHausdorffDim2022 [60] Satvik Singh, *A brief discourse on Hausdorff dimension and self-similarity*, *monthly* **129** (2022), 831–845.
- SteenSeebach1995 [61] Lynn Arthur Steen and J. Arthur Seebach, Jr., *Counterexamples in Topology*, 2 ed., Dover, New York, 1995; originally published by Springer-Verlag, 1978. [Cited on pages 266, 651.]
- SundstromReasoning [62] Ted Sundstrom, *Mathematical Reasoning: Writing and Proof*, version 2.1, Ted Sundstrom, Allendale, MI, 2020, <https://scholarworks.gvsu.edu/books/9/>; PDF version available for free download. [Cited on page 2.]
- Suppes1972 [63] Patrick Suppes, *Axiomatic Set Theory*, Dover, New York, 1972; originally published by D. Van Nostrand, 1960. [Cited on pages 2, 651.]
- TaylorAnalysisOneVar2017 [64] Michael E. Taylor, *Introduction to Analysis in One Variable*, AMS Open Math Notes, American Mathematical Society, 2017, <https://www.ams.org/open-math-notes/omn-view-listing?listingId=110664> [accessed 2017-10-22]. [Cited on page 88.]
- VellemanProvelt [65] Daniel J. Velleman, *How to Prove It: A Structured Approach*, 3 ed., Cambridge Univ. Press, New York, 2019. [Cited on page 2.]
- Vetterlein1995 [66] Donald H. Vetterlein, *Another proof that the Niemytzki plane is not normal*, *Pi Mu Epsilon Journal* **10** (1995), 119–121. [Cited on pages 300, 628.]
- WallGeomIntro1993 [67] C. T. C. Wall, *A Geometric Introduction to Topology*, Dover, Mineola, NY, 1993; originally published by Addison-Wesley, 1972.
- mathworld:Topology [68] Eric W. Weisstein, *Topology*, from *MathWorld—A Wolfram Web Resource*, 2016, <http://mathworld.wolfram.com/Topology.html> [accessed 2016-09-21]. [Cited on page 651.]
- wiki:ConstructingReals [69] Wikipedia, *Construction of the real numbers—Wikipedia, The Free Encyclopedia*, 2017, [https://en.wikipedia.org/wiki/Construction\\_of\\_the\\_real\\_numbers](https://en.wikipedia.org/wiki/Construction_of_the_real_numbers) [accessed 2017-08-10]. [Cited on page 88.]
- wiki:CubicEquation [70] ———, *Cubic equation—Wikipedia, The Free Encyclopedia*, 2020, [https://en.wikipedia.org/wiki/Cubic\\_equation](https://en.wikipedia.org/wiki/Cubic_equation) [accessed 2020-11-28]. [Cited on page 528.]
- wiki:GeneralTopology [71] ———, *General Topology—Wikipedia, The Free Encyclopedia*, 2016, [https://en.wikipedia.org/wiki/General\\_topology](https://en.wikipedia.org/wiki/General_topology) [accessed 2016-09-21]. [Cited on page 651.]
- wiki:QuarticFunction [72] ———, *Quartic function—Wikipedia, The Free Encyclopedia*, 2020, [https://en.wikipedia.org/wiki/Quartic\\_function](https://en.wikipedia.org/wiki/Quartic_function) [accessed 2020-11-28]. [Cited on page 528.]
- wiki:Topology [73] ———, *Topology—Wikipedia, The Free Encyclopedia*, 2016, <https://en.wikipedia.org/wiki/Topology> [accessed 2016-09-21]. [Cited on page 651.]
- WillardTopology2004 [74] Stephen Willard, *General Topology*, Dover, Mineola, NY, 2004; originally published by Addison-Wesley, 1970. [Cited on pages 382, 468, 565, 569, 651, 656.]

- YandlBowers2016 [75] André L. Yandl and Adam Bowers, *Elementary Point-Set Topology: A Transition to Advanced Mathematics*, Dover, Mineola, NY, 2016. [Cited on page 651.]

Guidance on proving mathematical statements is given in Lakins [42]. The set-theoretic preliminaries sketched in Chapter 0 are treated in greater detail, at an elementary level, in Eisenberg [22], Lakins [42], and Pinter [53]. A full axiomatic development of set theory, making a distinction between “classes” and “sets”, is provided in Bourbaki [9], Eisenberg [21], and Bernays [3]; a condensed axiomatic development appears in the Appendix to Kelley [40]. A different axiomatic foundation appears in Suppes [63].

For an overview of topology, see Carlson [12], Weisstein [68], Wikipedia [73], or Wikipedia [71].

Accounts of the historical development of topology appear in the several Historical Notes of Bourbaki [8] and in Manheim [45]. Extensive historical notes, with references to the original sources, are included at the end of Willard [74]. Numerous citations of the literature appear in the footnotes of Greever [31]. Histories of various advanced areas of topology are provided by experts in James [39].

Both Greever [31] and Steen and Seebach [61] are replete with examples, especially pathological ones.

Yandl and Bowers [75] provides an introduction to topology aimed especially at students who are just learning how to write proofs.

Today perhaps “the” standard text in topology—general topology along with some algebraic topology—is Munkres [51]. Other standard advanced texts include Bourbaki [8], Hocking and Young [36], Kelley [40], and Willard [74]. Kelley [40] contains a rich supply of exercises, some of which develop entire theories, along with an extensive bibliography.

Hatcher [33] gives an extensive list of recommended books ranging from intuitive and introductory treatments to specialized works on advanced topics.

Lee [44] covers topological manifolds more broadly and deeply than the treatment here, including through the use of both homotopy and homology tools from algebraic topology.

Hurewicz and Wallman [37] presents a detailed development of the Uryshon–Menger (small inductive) definition of dimension and begins with a readable historical account. Engelking [23] gives a more up-to-date treatment and includes extensive historical and bibliographical notes about dimension. Edgar [20] emphasizes covering dimension along with the Hausdorff-Besicovitch measure for fractals.

[compare algebraic books with what I cover

MORE!



## Additional Readings

bib:read The following additional readings can be used for individual or group projects; others should be suggested by references throughout the text.

- |                               |      |  |
|-------------------------------|------|--|
| AmsterHalfPower2021           | [76] | Pablo Amster, <i>The power of the half power</i> , American Mathematical Monthly <b>128</b> (2021), no. 7, 655–657. [Cited on page <a href="#">655</a> .]  |
| ArthreyaReznickTyson2019      | [77] | Jayadev S. Arthreya, Bruce Reznick, and Jeremy T. Tyson, <i>Cantor set arithmetic</i> , American Mathematical Monthly <b>126</b> (2019), no. 1, 4–17. [Cited on page <a href="#">654</a> .]  |
| BajpaiVelleman2016            | [78] | Dvij Bajpai and Daniel J. Velleman, <i>Anonymity in predicting the future</i> , American Mathematical Monthly <b>123</b> (2016), no. 8, 777–788.   |
| BingElusiveFixedPt1969        | [79] | R. H. Bing, <i>The elusive fixed point property</i> , American Mathematical Monthly <b>76</b> (1969), 119–132. [Cited on page <a href="#">655</a> .]   |
| BlaszykSierpinskiThm2019      | [80] | Aleksander Blaszyk, <i>A simple proof of Sierpiński’s theorem</i> , American Mathematical Monthly <b>126</b> (2019), no. 5, 464–466. [Cited on page <a href="#">655</a> .]   |
| BowersMengerEmbed2017         | [81] | John C. Bowers and Philip I. Bowers, <i>A Menger redux: embedding metric spaces isometrically in Euclidean space</i> , American Mathematical Monthly <b>124</b> (2017), no. 7, 621–636. [Cited on page <a href="#">655</a> .]            |
| Cirstea2021                   | [82] | Florica C. Cîrstea, <i>Proofs of Urysohn’s lemma and the Tietze extension theorem via the Cantor function</i> , Bulletin of the Australian Mathematical Society <b>103</b> (2021), no. 2, 326–332. [Cited on page <a href="#">656</a> .] |
| Conover2014                   | [83] | Robert A. Conover, <i>A First Course in Topology: An Introduction to Mathematical Thinking</i> , Dover, Mineola, NY, 2014; originally published by Williams & Wilkins, 1975. [Cited on page <a href="#">656</a> .]                       |
| DoyleHockingInvertible1961    | [84] | P. H. Doyle and J. G. Hocking, <i>Invertible spaces</i> , American Mathematical Monthly <b>68</b> (1961), 959–965. [Cited on page <a href="#">655</a> .]   |
| FlapanKnotsMolecules2016      | [85] | Erica Flapan, <i>Knots, Molecules, and the Universe: An Introduction to Topology</i> , American Mathematical Society, Providence, RI, 2016. [Cited on page <a href="#">656</a> .]  |
| Flapan2000                    | [86] | ———, <i>When Topology Meets Chemistry: A Topological Look at Molecular Chirality</i> , Mathematical Association of America, Washington, DC, 2000. [Cited on page <a href="#">656</a> .]  |
| GallierXuSurfaces2013         | [87] | Jean Gallier and Dianna Xu, <i>A Guide to the Classification Theorem for Compact Surfaces</i> , Springer, New York, 2013. [Cited on page <a href="#">655</a> .]  |
| Gantner1971                   | [88] | T. E. Gantner, <i>A regular space on which every continuous real-valued function is constant</i> , American Mathematical Monthly <b>78</b> (1971), no. 1, 52–53. [Cited on page <a href="#">656</a> .]                                   |
| HardinAndTaylorPredicting2008 | [89] | Christopher S. Hardin and Alan D. Taylor, <i>A peculiar connection between the Axiom of Choice and predicting the future</i> , American Mathematical Monthly <b>115</b> (2008), no. 2, 91–96. [Cited on page <a href="#">656</a> .]      |



- HewittCompactness1960 [90] E. Hewitt, *The rôle of compactness in analysis*, American Mathematical Monthly **67** (1960), 499–516. [Cited on page 655.]
- ConnectedPlaneSetsCauchyEqn1942 [91] F. B. Jones, *Connected and disconnected plane sets and the functional equation  $f(x) + f(y) = f(x + y)$* , Bulletin American Mathematical Society **48** (1942), 115–120. [Cited on page 655.]
- KelleyTychonoffImpliesAC1950 [92] J. L. Kelley, *The Tychonoff product theorem implies the axiom of choice*, Fundamenta Mathematicae **37** (1950), 75–76. [Cited on page 656.]
- KleeAffineTopSpaces1955 [93] V. L. Klee, *Some finite-dimensional affine topological spaces*, Portugaliae Mathematica **14** (1955), 27–30. [Cited on page 654.]
- KottmanEuclideanAndAffine1971 [94] C. A. Kottman, *A characterization of the Euclidean topology among the affine topologies*, Israel Journal of Mathematics **10** (1971), 212–217. [Cited on page 654.]
- KuhnTopLemmas1960 [95] H. W. Kuhn, *Some combinatorial lemmas in topology*, IBM Journal of Research and Development **4** (1960), 518–524. [Cited on page 654.]
- vanMunsterHausdorff1967 [96] Bart van Munster, *Hausdorffization and homotopy*, American Mathematical Monthly **124** (2017), 81–82. [Cited on page 655.]
- OsborneHausdorff2014 [97] M. Scott Osborne, *Hausdorffization and such*, American Mathematical Monthly **121** (2014), 727–733. [Cited on page 655.]
- ShermanKuratowskiVars2010 [98] David Sherman, *Variations on Kuratowski's 14-set theorem*, American Mathematical Monthly **117** (2010), no. 2, 113–123. [Cited on page 655.]
- MHStoneWeierstrassApprox1962 [99] M. H. Stone, *A generalized Weierstrass approximation theorem*, in: *Studies in Modern Analysis*, **1**, Studies in Mathematics, Math, 1962, 30–87. [Cited on page 655.]
- JohnThomas1969 [100] John Thomas, *A regular space, not completely regular*, American Mathematical Monthly **76** (1969), no. 2, 181–182. [Cited on page 656.]
- TuckerDiskSphere1945 [101] A. W. Tucker, *Some topological properties of disk and sphere*, in: *Proceedings of the First Canadian Mathematical Congress (Montreal, 1945)*, 1945, 285–309. [Cited on page 654.]
- JHCWhiteheadBorsukNote1948 [102] J. H. C. Whitehead, *Note on a theorem of Borsuk*, Bulletin American Mathematical Society **54** (1948), 1125–1132. [Cited on page 655.]
- CWolfram2016 [103] Catherine Wolfram, *Recursive derivation of  $2n$  gon topologies*, 2016, <http://math.uchicago.edu/~may/REU2016/REUPapers/Wolfram.pdf> [accessed 2018-12-30]. [Cited on page 654.]

**Affine topologies.** The topology on the plane consisting of the “linearly open” sets, as defined in [Exercise 2.10](#), is an example of an affine topology. The peculiar properties such topologies can have are discussed in Klee [93]. A related paper is Kottman [94].

**Brouwer's fixed-point theorem.** Brouwer's Theorem and its relatives in low dimensions are treated in Tucker [101]. Brouwer's Theorem in arbitrary dimension is derived from a combinatorial lemma in Kuhn [95].

**Cantor set arithmetic.** In a topological sense, the Cantor set  $K$  is “thin” within the unit interval  $[0, 1]$ . Nonetheless, each real number in the interval  $[0, 2]$  is the sum of two elements of  $K$ : see [Exercise 4.19](#). Other surprising arithmetic properties of  $K$  are established in Arhrey, Reznick, and Tyson [77].

**Compact surfaces.** The torus and Klein bottle are compact surfaces that may be obtained as quotient spaces, by identifying opposite edges of the unit square. More generally, every compact surface may be obtained as a quotient space by identifying edges of a  $2n$ -gon, for some  $n$ . Wolfram [103] gives a recursive method for classifying such identifications and



derives a computational formula for counting how many such identifications yield a given compact surface. (For a nearly self-contained treatment of the complete classification of compact surfaces, see Gallier and Xu [87].)

Menger, Karl

**Compactness.** The thesis that compactness is a generalization of finiteness is illustrated in Hewitt [90]. At least the first two examples of Hewitt should be accessible to readers of this text. A much more complete treatment of Hewitt's second example, the Stone-Weierstrass Approximation Theorem, is given in Stone [99].

**Connected sets in the plane.** Weird connected subsets of the plane that arise as graphs of solutions to Cauchy's functional equation are discussed in Jones [91].

**Isometric embedding in Euclidean space.** In 1928, Karl Menger characterized those metric spaces that can be isometrically embedded in Euclidean  $n$ -space. Bowers and Bowers [81] presents a nearly self-contained proof of Menger's result using linear algebra.

**Fixed-point property.** A space  $X$  is said to have the "fixed-point property" when each continuous map  $f: X \rightarrow X$  has at least one fixed-point. An expository account, with many examples, is Bing [79]. Bing discusses the famous unsolved problem of whether a continuum in the plane that does not disconnect the plane must necessarily have the fixed-point property.

**Hausdorff quotients.** Exercise 3.203 gave a construction that, for an arbitrary topological space  $X$ , provides a quotient space of  $X$  that is  $T_0$ . Osborne [97] presents a different construction that provides, in a certain "natural" way, a Hausdorff quotient of an arbitrary topological space  $X$ , and a generalization that provides a  $T_1$  quotient. And Munster [96] proves that this "Hausdorffization" preserves homotopy of maps; this proof uses a result of Whitehead [102] that the product of quotient maps is itself a quotient map if XXXX.

**Invertible spaces.** The  $n$ -sphere has the property that the complement of each of its nonempty open subsets  $U$  can be mapped into  $U$  itself by a suitable homeomorphism of the space. Topological features and examples of other spaces with this property are discussed in Doyle and Hocking [84].

**Variations on the Kuratowski closure-complement theorem.** You are asked in Exercise 2.71 to prove what is known as Kuratowski's closure-complement theorem, namely, that successively applying the operations of taking closures and complements, starting with a given subset of a topological space  $X$ , produces at most 14 distinct subsets of  $X$ . The paper by Sherman [98] deals with variations of the problem Kuratowski solved, for example, the maximum number of distinct subsets obtained by applying the operations of taking closures, interiors, and intersections.

**Sierpiński's Theorem.** Sierpiński's Theorem is that every countable self-dense metric space is homeomorphic to the space of rational numbers. The article Blaszyk [80] presents a proof that relies upon Cantor's theorem (Exercise 0.117) that any two countable, order-dense, totally ordered sets, each of which has neither an upper bound nor a lower bound, are order-isomorphic.

bibcomment:no-retract-sqrt **No-Retraction Theorem.** Our proof of the No-Retraction Theorem (5.104), as a corollary to the fact that the circle is not simply connected, employs the machinery of homotopy and covering spaces. Amster [76] presents a quite elementary proof that avoids such machinery and instead uses complex square-root functions.

bibcomment:TychonoffTheorem **Tychonoff Product Theorem.** As noted in the text, all proofs of this theorem depend on some variant of the Axiom of Choice (0.26). The first one given in Section 6.1 applied Zorn's Lemma (0.115) to prove the Alexander Subbase Theorem (6.1). Both Hocking and Young [36, pages 25–28] and Munkres [51, pages 228–233], which carry out the proof using

the Alexander Subbase Theorem, analyze the difficulties in trying naively to use the [finite intersection property criterion](#) ([Theorem 4.9](#)) without invoking Zorn’s Lemma.

Kelley [40, pages 143–144] also uses the Alexander Subbase Theorem and gives a different proof using the finite intersection property criterion. In an exercise, Munkres [51, page 234] outlines a proof that applies the Well-Ordering Theorem to the index set.

Once the necessary apparatus about filters and ultrafilters has been developed, the simplest proof uses ultrafilters; see Conover [83, pages 200–203] or Bourbaki [8, chap. I, sec. 9.5]. An alternative to using filters and ultrafilters is to use the dual notions of nets and “universal nets”; see Willard [74, Definition 11.10 and Theorems 17.4–17.8].

It is interesting to note that the Tychonoff Product Theorem is logically equivalent to the Axiom of Choice: see Kelley [92] for the proof that the Tychonoff Product Theorem implies the Axiom of Choice.

bibcomment:reg-not-cr **Regular spaces that are not completely regular.** Although a completely regular space must be regular, the converse fails. [Example 6.17](#) presented Mysior’s construction of a counterexample. For additional counterexamples, see Thomas [100] and Gantner [88]. The latter gives a method for manufacturing, from any given regular space, a regular but not completely regular space.

bibcomment:thm-tietze-scott-pf **Proof of the Tietze Extension Theorem.** The proof appearing in [Section 6.2](#) uses the traditional method of building the desired extension as the limit of a sequence of successively better continuous approximations. Scott [58] presents a different proof that avoids using such approximations and instead resembles the method of proof of Urysohn’s Lemma itself. Scott’s proof applies the more sophisticated version of Urysohn’s Lemma given in [Exercise 6.2.6.40](#).

Completely different proofs of both the Tietze Extension Theorem and Urysohn’s Lemma, using the Cantor staircase function, appear in Ćirstea [82].

**Predicting the future.** Regard a function  $f: \mathbb{R} \rightarrow S$  as representing, at each time  $t \in \mathbb{R}$ , the “state”  $f(t) \in S$  of some system. No matter what  $f$  is, it is possible to establish existence of some “strategy” for “predicting” the state at time  $t$  from knowing the states at times earlier than  $t$ —except for a nowhere dense set of times  $t$ . An article by Hardin and Taylor [89] proposes precise meanings for ‘strategy’ and ‘predict’ and offers a proof of the result that relies upon the Axiom of Choice.

**Topology in molecular chemistry.** Knots and other topological objects are used to model molecules and other entities in chemistry. See Flapan [86] and Flapan [85].

## List of Named Results

0.1	Axiom of Induction . . . . .	11
0.2	Principle of Mathematical Induction . . . . .	11
0.3	Well-ordering Principle . . . . .	12
0.9	Construction of a map from two pieces . . . . .	25
0.10	Compositions of injective and surjective maps . . . . .	29
0.11	Injective and surjective compositions . . . . .	29
0.13	Inverse of map with two-sided identities . . . . .	31
0.16	Correspondence between subsets and characteristic functions . . . . .	38
0.19	De Morgan's Laws . . . . .	40
0.21	Piecewise construction of a map . . . . .	41
0.26	Axiom of Choice . . . . .	44
0.27	Axiom of Choice—product version . . . . .	45
0.28	Axiom of Choice—choice function version . . . . .	45
0.29	Axiom of Choice—choice set version . . . . .	45
0.32	Principle of Ordinary Recursion . . . . .	48
0.34	Principle of Primitive Recursion . . . . .	50
0.35	Principle of Complete Recursion . . . . .	52
0.39	Finiteness of union of two disjoint finite sets . . . . .	55
0.40	Finiteness of subset of finite set . . . . .	55
0.41	Finiteness of union of finitely many finite sets . . . . .	56
0.42	Finiteness of product of finitely many finite sets . . . . .	56
0.43	Finiteness of range of surjection from finite set . . . . .	57
0.44	Finiteness of power set of finite set . . . . .	57
0.46	Denumerability of infinite subset of countable set . . . . .	58
0.47	Infinite set has a denumerable subset . . . . .	58
0.48	Countability of range of a map on a countable set . . . . .	58
0.50	Countability of union of countably many countable sets . . . . .	59
0.52	Countability of product of finitely many countable sets . . . . .	60
0.74	Order-completeness . . . . .	79
0.75	Existence of infimum of set of reals bounded below . . . . .	79
0.77	Base expansion of a real number . . . . .	81
0.78	Archimedean Ordering Property . . . . .	82
0.81	Order-density of the rationals in the reals . . . . .	83
0.83	Nested Interval Property . . . . .	84
0.84	Uncountability of closed interval in $\mathbb{R}$ . . . . .	85
0.85	Uncountability of the set of real numbers . . . . .	86
0.86	Order-density of irrational numbers in the reals . . . . .	86
0.87	Characterization of intervals in $\mathbb{R}$ . . . . .	86
0.88	De Moivre's Formula . . . . .	91
0.95	Characterization of equivalent elements . . . . .	97

0.97	Characterization of quotient sets as partitions . . . . .	98
0.102	Saturation of subset by equivalence relation . . . . .	100
0.104	Passing to the quotient . . . . .	101
0.108	Order-completeness of well-ordered set . . . . .	109
0.109	The first uncountable ordinal . . . . .	110
0.115	Zorn's Lemma . . . . .	113
0.117	Well-ordering Theorem . . . . .	116
0.123	Cardinality of range of a map . . . . .	120
0.124	Cardinality transitivity . . . . .	120
0.125	Cantor's Theorem . . . . .	121
0.126	Cantor-Bernstein Theorem . . . . .	121
0.127	Cardinal comparability . . . . .	122
0.128	Power law of cardinality . . . . .	123
0.129	Cardinality of $\mathbb{R}^{\mathbb{N}}$ . . . . .	123
1.3	Properties of Euclidean norm . . . . .	130
1.4	Cauchy-Schwarz Inequality . . . . .	131
1.5	Properties of Euclidean metric . . . . .	131
1.21	Properties of $d$ -open sets . . . . .	153
1.24	Properties of $d$ -closed sets . . . . .	157
1.26	Separating a point from a $d$ -closed set . . . . .	159
1.29	Distance of points from $d$ -closed set . . . . .	160
1.30	Separating $d$ -closed sets . . . . .	160
1.32	Characterization of $d$ -open sets by neighborhoods . . . . .	161
1.35	Characterization of equivalent metrics by balls . . . . .	164
1.36	Equivalence of Euclidean and max metrics . . . . .	165
1.40	Equivalence of metric with bounded metric . . . . .	168
1.43	Balls at infinity as rays in extended real line . . . . .	172
1.47	Properties of isometries . . . . .	174
1.50	Properties of topological equivalence . . . . .	176
1.53	Characterization of continuity in metric spaces in terms of neighborhoods . . . . .	182
1.55	Continuity of equivalent metrics . . . . .	184
1.56	Continuity in terms of open sets . . . . .	184
1.62	Sequential convergence in terms of neighborhoods . . . . .	190
1.63	Preservation of convergence for equivalent metrics . . . . .	190
1.65	Uniqueness of sequential limits . . . . .	192
1.67	Closedness in terms of sequential convergence . . . . .	193
1.68	Openness in terms of sequential convergence . . . . .	194
1.69	Characterization of continuity in terms of sequential convergence . . . . .	194
1.72	Sequential convergence and Cauchy sequence . . . . .	201
1.76	Completeness of the real line . . . . .	202
1.77	Completeness of max metric . . . . .	204
1.78	Completeness of Euclidean metric . . . . .	205
1.81	Completeness and closed sets . . . . .	206
1.82	Completeness of $\ell^2$ . . . . .	206
1.83	Completeness of sup metric on bounded functions . . . . .	208
1.85	Continuity of uniform limit of uniformly convergent sequence . . . . .	210
1.86	Completeness of sup metric on continuous bounded functions . . . . .	212
1.88	Nested Set Theorem . . . . .	212
1.91	Baire Category Theorem . . . . .	213
2.18	Properties of neighborhood systems . . . . .	237

2.19	Constructing a topology from neighborhood systems . . . . .	238
2.44	Closure-interior duality . . . . .	254
2.58	Properties of local bases . . . . .	265
2.59	Constructing a topology from local bases . . . . .	265
2.71	Constructing a topology from a base . . . . .	271
2.74	Closed diagonal criterion . . . . .	276
2.84	Lindelöf Theorem . . . . .	280
3.8	Composite of continuous maps . . . . .	304
3.12	Extension of identities . . . . .	308
3.13	Gluing Lemma . . . . .	309
3.39	Invariance of Domain Theorem . . . . .	342
3.76	Passing to the quotient . . . . .	391
3.106	Uniqueness of net limits in a Hausdorff space . . . . .	429
3.108	Convergence of a net in a product . . . . .	430
3.145	Uniqueness of filter limits in a Hausdorff space . . . . .	446
3.147	Convergence of a filter in a product . . . . .	447
4.9	Finite intersection property criterion . . . . .	463
4.15	Heine–Borel–Lebesgue Theorem . . . . .	466
4.19	Separation of compact sets in a Hausdorff space . . . . .	474
4.20	Normality of compact $T_2$ -space . . . . .	474
4.24	Closed map principle . . . . .	475
4.25	Continuous bijection principle . . . . .	476
4.27	Extreme Value Theorem . . . . .	476
4.30	The Tube Lemma . . . . .	478
4.31	“Tychonoff-for-two” . . . . .	478
4.32	Tychonoff Product Theorem—finite case . . . . .	478
4.33	Tychonoff Product Theorem . . . . .	479
4.34	Compact subspaces of Euclidean space . . . . .	479
4.39	Closed-graph criterion for a $T_2$ quotient . . . . .	481
4.42	Compactness and net convergence . . . . .	482
4.43	Compactness and filter convergence . . . . .	483
4.54	Monotonic subsequence theorem . . . . .	497
4.57	Bolzano–Weierstrass Theorem . . . . .	498
4.61	Tychonoff Product Theorem—denumerable metrizable case . . . . .	500
4.63	Lebesgue Covering Lemma . . . . .	501
4.78	Alexandroff one-point compactification . . . . .	515
5.13	Intermediate-value Theorem . . . . .	528
5.15	Brouwer Fixed-point Theorem in dimension 1 . . . . .	530
5.16	Borsuk–Ulam Theorem in dimension 1 . . . . .	531
5.66	Hahn–Mazurkiewicz Theorem . . . . .	569
5.81	Path-homotopy preserves path-products . . . . .	579
5.82	Path-homotopy preserves path reverse . . . . .	580
5.83	Associativity of path-product up to homotopy . . . . .	581
5.89	Null loops as one-sided identities for paths up to homotopy . . . . .	583
5.90	Null loop as identity for loops up to homotopy . . . . .	583
5.91	Reverse of loop as inverse up to homotopy . . . . .	584
5.92	Reverse of loop is as two-sided inverse up to homotopy . . . . .	584
5.98	Jordan Curve Theorem . . . . .	592
5.100	Path-lifting property for the circle . . . . .	594

5.101	Homotopy-lifting property for the circle . . . . .	596
5.104	No-Retraction Theorem . . . . .	598
5.105	Brouwer Fixed-point Theorem in Dimension 2 . . . . .	599
5.106	Fundamental Theorem of Algebra . . . . .	600
4.33	Tychonoff Product Theorem . . . . .	605
6.1	Alexander Subbase Theorem . . . . .	606
6.8	Existence of non-principal ultrafilters . . . . .	608
6.12	Ultrafilter characterization of compactness . . . . .	610
6.26	Urysohn's Lemma . . . . .	619
6.31	Jones's Lemma . . . . .	622
6.34	Tietze Extension Theorem . . . . .	623
6.44	Tychonoff Embedding Theorem . . . . .	633
6.47	Urysohn Metrization Theorem . . . . .	635

# Index of Notation

Within each topical group below, items are listed in the order they appear in the body of the text.

## Set theory

$x \in X$	$x$ is an element of set $X$ , 7
$x \notin X$	$x$ is not an element of $X$ , 7
$X = Y$	sets $X$ and $Y$ have the same elements, 7
$\{a, b, c, \dots, s\}$	explicitly given set, 8
$X \neq Y$	sets $X$ and $Y$ are not equal, 7
$\{x : P(x)\}$	set of objects having a given property, 8
$\{x \in A : P(x)\}$	set of objects in a set having a property, 8
$\{x \mid P(x)\}$	variant set-builder notation, 9
$\{x\}$	single-element set, 9
$\emptyset$	empty set, 9
$Y \subset X$	$Y$ is a subset of $X$ , 10
$X \supset Y$	$Y$ is a subset of $X$ , 10
$Y \subsetneq X$	$Y$ is a proper subset of $X$ , 11
$A \cup B$	union of two sets, 13
$A \cap B$	intersection of two sets, 13
$\mathcal{P}(X)$	collection of all subsets, 13
$\langle x, y \rangle$	ordered pair, 14
$X \times Y$	product of two sets, 15
$(x, y, z)$	ordered triple, 15
$x \alpha y$	$x$ is related by relation $\alpha$ to $y$ , 15
$x \not\alpha y$	$x$ is not related by $\alpha$ to $y$ , 15
$\Delta_X$	diagonal of $X \times X$ , 15
$\alpha^{-1}$	reverse of relation, 17
$\beta \circ \alpha$	composite of relations, 18
$f: X \rightarrow Y$	map from $X$ to $Y$ , 19
$X \xrightarrow{f} Y$	map from $X$ to $Y$ , 19
$f(x)$	value of function $f$ at $x$ , 19
$\iota_X$	identity map of set, 20
$x \mapsto y$	$x$ maps to $y$ , 20
$f: X \rightarrow Y$ $x \mapsto f(x)$	map from $X$ to $Y$ sending $x$ to $f(x)$ , 20
$f: X \rightarrow Y: x \mapsto f(x)$	map from $X$ to $Y$ sending $x$ to $f(x)$ , 20
$\chi_A$	characteristic function of subset, 20
$ x $	absolute value of $x$ , 20
$ \cdot $	absolute-value function, 20
$\text{dom}(f)$	domain $X$ of map $f: X \rightarrow Y$ , 21

- $\text{codom}(f)$   
 $\text{graph}(f)$   
 $\text{range}(f)$   
 $f(x, y)$   
 $f|_E$   
 $f|E$   
 $f|_{E, f(E)}$   
 $g \circ f$   
 $f(A)$   
 $f\langle A \rangle$   
 $f[A]$   
 $f^{-1}(D)$   
 $f^{-1}\langle D \rangle$   
 $f^{-1}[D]$   
 $f^{-1}(y)$   
 $f^{-1}\langle y \rangle$   
 $f^{-1}[y]$   
 $f^{-1}$   
 $x_i$   
 $\langle x_i \rangle_{i \in I}$   
 $\langle x_i : i \in I \rangle$   
 $\langle a_{i,j} \rangle_{(i,j) \in I \times J}$   
 $\langle a_{i,j} \rangle_{i \in I, j \in J}$   
 $\langle x_i \rangle_{i=1,2,\dots,n}$   
 $\langle x_i \rangle_{i=1}^n$   
 $\langle x_1, x_2, \dots, x_n \rangle$   
 $\langle x_i \rangle_{i=0,1,2,\dots}$   
 $\langle x_i \rangle_{i=0}^\infty$   
 $\langle x_0, x_1, x_2, \dots \rangle$   
 $X^n$   
**0**  
 $x + y$   
 $\alpha x$   
 $-x$   
 $X^I$   
 $\mathcal{F}(X, Y)$   
 $2^X$   
 $\bigcup_{i \in I} X_i$   
 $\bigcap_{i \in I} X_i$   
 $\bigcup \mathcal{A}$   
 $\bigcap \mathcal{A}$   
 $\times_{i=1}^n X_i$   
 $\times_{i=1}^\infty X_i$   
 $\times_{i \in I} X_i$   
 $\prod_{i \in I} X_i$   
 $p_j$   
 $f^n$   
 $\#(X)$   
 $\text{card } X = \text{card } Y$   
 $\text{card } X \leq \text{card } Y$   
 $\text{card } X < \text{card } Y$
- codomain  $Y$  of map  $f: X \rightarrow Y$ , 21  
 graph of map, 21  
 range of map, 21  
 value  $f(\langle x, y \rangle)$  of function of two variables, 21  
 restriction of map, 23  
 restriction of map, 23  
 domain-codomain restriction of map, 24  
 composite of maps, 25  
 image of set under map, 27  
 image of set under map, 27  
 image of set under map, 27  
 inverse image of set under map, 27  
 inverse image of set under map, 27  
 inverse image of set under map, 27  
 inverse image of point under map, 27  
 inverse image of point under map, 27  
 inverse image of point under map, 27  
 inverse of bijective map, 30  
 $i$ th coordinate of family, 34  
 family, 34  
 family, 34  
 doubly-indexed family, 34  
 doubly-indexed family, 34  
 $n$ -tuple, 35  
 $n$ -tuple, 35  
 $n$ -tuple, 35  
 sequence, 35  
 sequence, 35  
 sequence, 35  
 $n$ th power of set, 35  
 origin of  $\mathbb{R}^n$ , 36  
 vector sum, 36  
 product of scalar and vector, 36  
 negative of vector, 36  
 set of all families in  $X$  indexed by  $I$ , 37  
 set of all maps from  $X$  to  $Y$ , 37  
 set of characteristic functions of subsets of  $X$ , 38  
 union of family of sets, 38  
 intersection of family of sets, 38  
 union of collection of sets, 41  
 intersection of collection of sets, 41  
 product of family of  $n$  sets, 42  
 product of sequence of sets, 43  
 product of family of sets, 43  
 product of family of sets, 43  
 $j$ th projection from product set, 43  
 $n$ th iterate of map, 49  
 number of elements in finite set, 54  
 same cardinality, 118  
 same or smaller cardinality, 120  
 strictly smaller cardinality, 120



$\aleph_0$	cardinality of denumerable set, 118
$\mathfrak{c}$	cardinality of $\mathbb{R}$ , 382
$m \mid n$	$m$ divides $n$ , 63
$x \leq y$	precedence, 63
$x < y$	strict precedence, 64
$x \geq y$	reverse precedence, 65
$x > y$	reverse strict precedence, 65
$x \leq y$	less than or equal, 66
$x < y$	(strictly) less than, 67
$y \geq x$	greater than or equal, 68
$y > x$	(strictly) greater than, 68
$\max A$	maximum of set, 68
$\min A$	minimum of set, 68
$]a, \rightarrow[$	open ray, 70
$]\leftarrow, a[$	open ray, 70
$[a, \rightarrow[$	closed ray, 70
$]\leftarrow, a]$	closed ray, 70
$[a, b]$	closed interval, 70
$]a, b[$	open interval, 70
$]a, b]$	half-open interval, 70
$[a, b[$	half-open interval, 70
$]a, +\infty[$	open ray in $\mathbb{R}$ , 70
$]-\infty, b[$	open ray in $\mathbb{R}$ , 70
$[a, +\infty[$	closed ray in $\mathbb{R}$ , 70
$]-\infty, b]$	closed ray in $\mathbb{R}$ , 70
$\sup A$	supremum of set, 71
$\inf A$	infimum of set, 71
$(0.d_1 d_2 d_3 \dots d_n \dots)_b$	base expansion, 80
$x \sim y$	equivalent elements, 93
$n \equiv k \pmod{m}$	congruence of integers modulo $m$ , 94
$\llbracket x \rrbracket \sim$	equivalence class under $\sim$ , 96
$\llbracket x \rrbracket$	equivalence class, 96
$X/\sim$	quotient set, 96
$\text{card } X$	cardinality of set, 122
$\alpha(x)$	value of functional relation $\alpha$ at $x$ , 16

### Special sets

$\mathbb{N}$	set of all natural numbers, 9
$\mathbb{N}^*$	set of all <i>positive</i> natural numbers, 10
$\mathbb{Z}$	set of all integers, 9
$\mathbb{Q}$	set of all rational numbers, 9
$\mathbb{R}$	set of all real numbers, 9
$\mathbb{R}^*$	set of all nonzero real numbers, 10
$\mathbb{R}^+$	set of nonnegative real numbers, 10
$\mathbb{C}$	set of all complex numbers, 9
$\mathbb{C}$	set of all complex numbers, 89
$i$	complex number with square $-1$ , 9
$I$	closed unit interval $[0, 1]$ , 9
$J$	centered interval $[-1, 1]$ , 9
$\mathbb{R}^0$	Euclidean 0-space, 36
$\mathbb{R}^1$	Euclidean line $\mathbb{R}$ , 36

$\mathbb{R}^2$	Euclidean plane, 36
$\mathbb{R}^3$	Euclidean 3-space, 36
$\mathbb{R}^n$	Euclidean $n$ -space, 36
$\mathbb{C}$	set of complex numbers, 89
$i$	complex number with square $-1$ , 89
$\bar{z}$	conjugate of complex number, 90
$ z $	modulus of complex number, 90
$e^{\theta i}$	complex exponential, 90
$\mathbb{R}/\mathbb{Z}$	quotient set of $\mathbb{R}$ mod $\mathbb{Z}$ , 104
$\mathbb{R}/\mathbb{Q}$	quotient set of $\mathbb{R}$ mod $\mathbb{Q}$ , 104
$\Omega$	first uncountable ordinal, 110
$\Omega^+$	set of countable ordinals together with $\Omega$ , 110
$\omega$	first infinite ordinal, 110
$\omega^+$	set of finite ordinals together with $\omega$ , 110
$C(X)$	set of continuous real-valued functions on $X$ , 135
$\mathcal{B}(X)$	set of bounded real-valued functions on $X$ , 141
$\ell^2$	Hilbert sequence space, 138
$B_n$	$n$ -ball in $\mathbb{R}^n$ , 149
$D_n$	$n$ -disk in $\mathbb{R}^n$ , 149
$S_{n-1}$	$(n-1)$ -sphere in $\mathbb{R}^n$ , 149
$I^n$	$n$ -cube in $\mathbb{R}^n$ , 325
$J^n$	centered $n$ -cube in $\mathbb{R}^n$ , 325
$\widehat{\mathbb{R}}$	extended real line, 168
$L$	long closed ray, 344
$L^*$	long line, 344
$T_n$	$n$ -torus, 398
$\mathbb{RP}_n$	real projective $n$ -space, 401

### Metrics and metric spaces

$d$	Euclidean metric, 129
$\ x\ $	Euclidean norm of $x$ , 130
$d_1$	taxicab metric on $\mathbb{R}^n$ , 133
$d_\infty$	max metric on $\mathbb{R}^n$ , 134
$\ f\ _\infty$	sup norm, 141
$d_\infty$	sup metric on $C([0, 1])$ , 136
$d_1$	$L_1$ -metric on $C([0, 1])$ , 136
$\ x\ _2$	Hilbert norm (of square-summable sequence), 138
$d_2$	Hilbert metric on $\ell^2$ , 138
$\ell^\infty$	space of bounded sequences, 144
$I^\infty$	Hilbert cube, 139
$d$	any metric, 139
$\delta$	discrete metric, 140
$d_\infty$	max metric on product space, 140
$d_\infty$	sup metric on set of bounded functions, 141
$ x _p$	$p$ -adic norm of $x$ in $\mathbb{Q}$ , 145
$d_p$	$p$ -adic metric on $\mathbb{Q}$ , 146
$B_\varepsilon(x; d)$	$d$ -ball of radius $\varepsilon$ at $x$ , 147
$D_\varepsilon(x; d)$	$d$ -disk of radius $\varepsilon$ at $x$ , 147
$S_\varepsilon(x; d)$	$d$ -sphere of radius $\varepsilon$ at $x$ , 147
$d(x, A)$	distance from point to set, 160
$d(A, B)$	distance between sets, 163

$\text{diam}(A)$	diameter of set, 167
$\langle x y \rangle$	Euclidean inner product, 178
$\langle x_n \rangle_{n \in \mathbb{N}} \xrightarrow{d} x$	sequence converges to $x$ for metric, 185
$\lim_{n \rightarrow \infty} x_n = +\infty$	infinite limit of sequence, 200
$\lim_{n \rightarrow \infty} x_n = -\infty$	infinite limit of sequence, 200
$C^*(X)$	set of continuous bounded real-valued functions on $X$ , 211

### Topological spaces

$\mathcal{N}_x$	neighborhood system at a point $x$ , 236
$\mathcal{N}_x(X)$	neighborhood system at a point $x$ in space $X$ , 236
$I_{\text{lex}}^2$	lexicographically ordered square, 286
$\mathbb{R}_K$	deleted sequence space, 287
$\mathcal{T}(d)$	topology induced by metric, 227
$(X, x_0)$	pointed topological space, 228
$\text{bdy } A$	boundary of set, 247
$\text{int } A$	interior of set, 250
$A^\circ$	interior of set, 250
$A$	interior of set, 250
$\text{cls } A$	closure of set, 252
$\overline{A}$	closure of set, 252
$\overline{A}$	closure of set, 252
$\mathbb{R}_I$	Sorgenfrey line, 239
$\Gamma$	tangent disk space, 246
$\delta_X$	diagonal map of $X$ into $X \times X$ , 306
$C(X, Y)$	set of all continuous maps between spaces, 307
$C[(X, x), (Y, y)]$	set of all continuous maps between pointed spaces, 307
$f: X \cong Y$	homeomorphism, 320
$X \cong Y$	homeomorphic spaces, 320
$X + Y$	Cartesian sum, 348
$S_n^+$	upper $n$ -hemisphere, 353
$S_n^-$	lower $n$ -hemisphere, 353
$\Delta_n$	standard $n$ -simplex, 354
$\partial M$	boundary of manifold-with-boundary, 358
$\bigcup_{i \in I} X_i$	Cartesian sum, 359
$\bigsqcup_{i \in I} X_i$	Cartesian sum, 359
$\coprod_{i \in I} X_i$	Cartesian sum, 359
$\times_{i \in I} X_i$	product of family of topological spaces, 364
$f_1 \times f_n$	product of maps, 371
$\square_{i \in I} X_i$	box product, 382
$\square^I S$	box power, 382
$\text{GL}(n, \mathbb{R})$	general linear group of degree $n$ , 384
$\text{SL}(n, \mathbb{R})$	special linear group of degree $n$ , 384
$\text{O}(n, \mathbb{R})$	orthogonal group of degree $n$ , 384
$\text{SO}(n, \mathbb{R})$	special orthogonal group of degree $n$ , 384
$X//A$	quotient space obtained by collapsing $A$ to a point, 404
$K(X)$	cone over space, 415
$S(X)$	suspension, 415
$\langle X, x_0 \rangle \vee \langle Y, y_0 \rangle$	wedge sum, 417
$X \vee Y$	wedge sum, 417
$\langle X, x_0 \rangle \wedge \langle Y, y_0 \rangle$	smash product, 418
$X \wedge Y$	smash product, 418

$\langle x_n \rangle_{n \in \mathbb{N}} \rightarrow x$	sequence converges to $x$ in topological space, <a href="#">421</a>
$(x_i)_{i \in I} \rightarrow x$	net converges to $x$ , <a href="#">425</a>
$\lim (x_i)_{i \in I}$	limit of net, <a href="#">430</a>
$\lim_c f$	limit of map at point, <a href="#">437</a>
$\lim_{x \rightarrow c} f(t)$	limit of map at point, <a href="#">437</a>
$\Phi(\xi)$	eventuality filter of $a$ with net $\xi$ , <a href="#">441</a>
$\mathcal{U}_x$	principal filter generated by $x$ , <a href="#">441</a>
$\mathcal{B} \rightarrow x$	filter base converges to $x$ , <a href="#">443</a>
$\mathcal{F} \rightarrow x$	filter converges to $x$ , <a href="#">443</a>
$\mathcal{F} \xrightarrow{X} x$	filter converges to $x$ in $X$ , <a href="#">443</a>
$\lim \mathcal{F}$	limit of a filter, <a href="#">447</a>
$f[\mathcal{F}]$	image of filter under a map, <a href="#">445</a>
$\Psi(\mathcal{F})$	net associated with filter, <a href="#">450</a>
$\mathbb{K}$	Cantor set, <a href="#">466</a>
$X_\infty$	one-point compactification, <a href="#">516</a>
$\alpha X$	one-point compactification, <a href="#">516</a>
$\sigma * \tau$	path product, <a href="#">559</a>
$\bar{\sigma}$	reverse of path, <a href="#">560</a>
$\Omega(X, x)$	set of all loops in a space at a point, <a href="#">583</a>
$h: f \simeq g$	homotopy from $f$ to $g$ , <a href="#">576</a>
$f \simeq g$	$f$ is homotopic to $g$ , <a href="#">576</a>
$h: \sigma \sim \tau$	path-homotopy, <a href="#">577</a>
$\sigma \sim \tau$	path-homotopic paths, <a href="#">577</a>
$\nu_x$	null loop at $x$ , <a href="#">581</a>
$\text{ind}(X)$	small inductive dimension, <a href="#">640</a>
$\text{cov}(X)$	covering dimension, <a href="#">642</a>
$\text{dim}(X)$	covering dimension, <a href="#">642</a>
$\text{Ind}(X)$	large inductive dimension, <a href="#">642</a>

# Index of Names

## Symbols

Čech, Eduard, [607](#)

## A

Aleksandrov, Pavel (*see* Alexandroff, Paul)  
Alexandroff, Paul, [460](#), [468](#), [516](#)  
Anderson, Richard D., [374](#)

## B

Baire, René-Louis, [214](#), [257](#)  
Banach, Stefan, [215](#), [220](#)  
Bernstein, Felix, [121](#)  
Bing, R. H., [292](#)  
Bolzano, Bernard, [498](#), [528](#)  
Borel, Émile, [459](#), [466](#)  
Borsuk, Karol, [531](#), [532](#)  
Brouwer, Luitzen Egbertus Jan, [530](#), [639](#)

## C

Cantor, Georg, [60](#), [77](#), [116](#), [121](#), [122](#), [222](#), [466](#)  
Cartan, Henri, [440](#), [607](#)  
Cauchy, Augustin-Louis, [131](#), [200](#)  
Chaber, J., [409](#)

## D

De Morgan, Augustus, [14](#)  
Dedekind, Richard, [72](#), [92](#)

## F

Fort, Marion K., Jr., [268](#)  
Fréchet, Maurice, [139](#), [241](#), [294](#), [460](#), [497](#)

## G

Gauss, Karl Friedrich, [600](#)

## H

Hahn, Hans, [569](#)  
Hausdorff, Felix, [113](#), [125](#), [212](#), [222](#), [239](#), [241](#), [294](#), [468](#)  
Heine, Eduard, [459](#), [466](#), [502](#)  
Hilbert, David, [138](#)

## J

Jones, F. Burton, [622](#)  
Jordan, Camille, [592](#)

## K

Klein, Felix, [400](#)  
Knaster, Bronisław, [34](#), [549](#)  
Kneser, Hellmuth, [113](#)  
Kolmogoroff, Andrey, [294](#)  
Kuratowski, Kazimierz, [113](#), [262](#), [263](#), [548](#), [549](#)

## L

Lebesgue, Henri, [466](#), [501](#), [639](#)  
Lindelöf, Ernst, [280](#)  
Lipschitz, Rudolf, [221](#)  
Listing, Johann Benedict, [399](#)

## M

Möbius, August Ferdinand, [399](#)  
Mazurkiewicz, Stefan, [569](#)  
Menger, Karl, [639](#), [655](#)  
Moore, E. Hastings, [423](#)  
Moore, Robert L., [246](#), [545](#)  
Mysior, Adam, [614](#)

## N

Niemytzki, Viktor Vladimirovich, [246](#)

**P**

Peano, Giuseppe, [565](#)

Picard, Émile, [221](#)

**R**

Riemann, Bernhard, [517](#)

Riesz, Frigyes, [241](#), [294](#)

**S**

Schr oder, Ernst, [121](#)

Schwarz, Hermann, [131](#)

Scott, Brian M., [409](#)

Sierpiński, Waław, [228](#), [490](#)

Smith, Herman L., [423](#)

Sorgenfrey, Robert, [239](#)

**T**

Tarski, Alfred, [34](#)

Teichmüller, Oswald, [125](#)

Thomas, John, [266](#)

Tietze, Heinrich, [294](#)

Tukey, John W., [113](#), [125](#)

Tychonoff, Andrey Nikolayevich, [478](#), [607](#),  
[617](#)

**U**

Ulam, Stanisław, [531](#), [532](#)

Urysohn, Pavel, [294](#), [460](#), [494](#), [619](#), [639](#)

**V**

Veblen, Oswald, [592](#)

Vietoris, Leopold, [294](#)

**W**

Weierstrass, Karl, [498](#)

**Z**

Zermelo, Ernst, [116](#), [122](#)

Zorn, Max, [113](#)

# Index

For  $d$ -ball,  $d$ -bounded set,  $d$ -open set,  $(d, d')$ -continuous map, etc., see the corresponding entry ball, bounded set, open set, continuous map, etc.

## A

absolute-value function, 20, 310  
 abstract quotient map, 392  
 accumulation point, 241, 244  
     *see also* adherent point, condensation point, limit point  
 addition, 22  
 additive inverse, 22  
 adjunction space, 406, 406–408, 415, 416  
 affine plane, 401  
 ajar subset, 229  
 Alexander Subbase Theorem, 606  
 Alexandroff compactification, 516  
 Alexandroff compactification (*see* one-point compactification)  
 Alexandroff double circle, 635  
 algebraic number, 62  
 almost all, 363  
 ambient space, 230  
 angle of rotation, 142, 177  
 annulus, 353, 356, 574  
 antipodal map, 329, 380  
 antisymmetric relation, 66, 73  
 arc, 565  
 arc-connected space, 565, 565  
 Archimedean Ordering Property, 82  
 arcwise connected space (*see* arc-connected space)  
 associative operation, 22  
 asymmetric relation, 67, 75  
 atlas, 338  
 attaching, 406–408  
     *see also* adjunction space  
 attaching map, 406  
 Axiom of Choice, 44, 44–45  
 Axiom of Induction, 11  
     and Well-ordering Principle, 12, 19

## axioms

order-completeness of reals, 79  
*see also* Axiom of Choice, Axiom of Induction

## B

Baire Category Theorem, 213, 213–217, 257, 522  
 ball, 147, 147  
     and induced metric, 151  
     and homeomorphism, 324  
     *see also*  $n$ -ball  
 base, 263, 269, 269  
     associated with a subbase, 283  
     binary, 188  
     in compact space, 484  
     and cover, 270  
     of a filter, 442  
     *see also* filter base  
     and product of metric spaces, 360–363  
     of product space, 364–365, 382  
     ternary, 188  
     *see also* subbase  
 base expansion, 79–81, 188  
     binary, 80  
     decimal, 80  
     and limit, 80  
     of positive integers, 19  
     and series, 80  
     terminating, 80  
     ternary, 80  
     nonterminating, 80  
     *see also* base- $b$  digits, base for expansion, binary expansion, decimal expansion, ternary expansion  
 base for expansion, 19, 79

base step, 11  
 base- $b$  digits, 79  
 basic open set, 269  
     of product space, 364  
     and product topology, 364–365  
 belonging to, 7  
 bijection, 28, 28–31  
 bijective map, 28  
     *see also* bijection  
 binary base, 188  
 binary bisequence, 143  
 binary expansion, 61, 80, 92  
 binary operation, 22  
 binary relation (*see* relation)  
 binary sequence, 141  
 Bing triangle space, 291, 292, 299  
 bisequence, 143  
 Bolzano–Weierstrass theorem, 498  
 Borsuk–Ulam Theorem, 531–532  
 bound (*see* bounded; lower bound; upper bound)  
 boundary, 247, 247–250  
     of ball, 248  
     and closed set, 250  
     and closure, 254–255, 259, 260  
     and complement, 254  
     of complement, 250  
     of disk, 249  
     and interior, 260  
     and local base, 264  
     of  $n$ -manifold-with-boundary, 358  
     and open set, 250  
     of product, 288  
     in a subspace, 258  
 bounded above, 71  
 bounded below, 71  
 bounded function, 141, 143, 208  
     *see also* bounded map  
 bounded map, 476  
 bounded metric  
     and equivalent metric, 168–171  
 bounded metric space, 167  
 bounded sequence, 186, 186–201  
 bounded set, 167, 167–168  
     and compact space, 479  
 box, 382  
     open, 355  
 box power, 382  
 box product, 382, 382–383  
     and compact space, 488  
 box topology, 360, 382, 479  
     and convergence, 452

Brouwer Fixed-point Theorem, 530, 599

## C

cancellation, 23  
 Candor set, 473  
 Cantor discontinuum (*see* Cantor set)  
 Cantor function, 470, 473, 486  
 Cantor middle-third set (*see* Cantor set)  
 Cantor Product Theorem, 463  
 Cantor set, 257, 261, 268, 466, 466–486  
     and continuous image, 468, 485, 626  
     continuous image continuous image, 470  
     is totally disconnected, 549  
 Cantor staircase function (*see* Cantor function)  
 Cantor ternary set (*see* Cantor set)  
 Cantor's theorem, 77  
 Cantor–Bernstein Theorem, 121, 199  
     and Knaster–Tarski Fixed-point Theorem, 126  
 cardinal number, 122  
     and ordinal, 122, 123  
 cardinality, 118, 118–123, 382  
     of denumerable set, 118  
     of finite set, 54, 118  
     of interval, 119  
     of  $\mathbb{R}$ , 120  
     of the real numbers, 119  
     of a set, 122  
     and surjection, 120  
 carries, 20  
 Cartesian sum  
     and normal space, 628  
 Cartesian product (*see* product)  
 Cartesian sum  
     and completely regular space, 615  
     and countability properties, 353  
     of a family of spaces, 235, 246, 289, 420, 485  
     of a family of spaces, 358, 358–359  
     of two spaces, 350  
     and separation properties, 615  
     of two spaces, 348, 347–349  
 Cauchy complete metric space (*see* complete metric space)  
 Cauchy sequence, 201, 200–201  
     and bounded sequence, 201  
 Cauchy–Schwarz Inequality, 131  
 centered  $n$ -cube, 325  
     and homeomorphism, 324



- see also*  $n$ -cube
- chain, [74](#)
- change in parameter, [578](#)
- characteristic function, [20](#)
- chart, [338](#), [358](#)
- chipped Tychonoff plank, [617](#)
- choice function
  - and well-ordered set, [123](#)
- circle
  - as quotient space, [385–388](#), [393–395](#)
  - see also*  $n$ -sphere, unit circle
- circle group, [384](#), [419](#)
- class, [13](#), [13](#)
  - proper, [13](#)
  - see also* set
- clopen subset, [229](#)
- closed diagonal criterion, [276](#)
- closed equivalence relation, [416](#), [416–417](#)
- closed interval, [70](#), [153](#), [458](#)
  - as compact space, [460](#)
  - and homeomorphism, [321](#), [322](#)
- closed map, [313](#), [488](#)
  - and projection, [313](#), [368](#), [488](#)
  - and quotient map, [393](#)
- closed ray, [70](#), [154](#)
- closed set, [229](#), [229](#), [229](#)
  - and convergent sequences, [421](#)
  - and distances, [160](#)
  - and filters, [445](#)
  - and neighborhoods, [161](#)
  - and nets, [428](#)
  - and boundary, [250](#)
  - and closure, [256](#)
  - and compact space, [464](#), [479](#)
  - in metric space, [157](#), [157–160](#)
  - in subspace, [232](#)
- closed unit interval, [9](#)
- Closed-graph Theorem, [380](#), [488](#)
- closed-halfspace, [198](#)
  - and  $n$ -manifold-with-boundary, [357](#)
- closure, [252](#), [252–316](#)
  - and convergent sequence, [421](#)
  - and filters, [445](#)
  - and nets, [428](#)
  - and axioms, [262](#)
  - and boundary, [254–255](#), [259](#), [260](#)
  - and closed set, [256](#)
  - and closed set, [255](#)
  - and complement, [254](#)
  - of connected set, [526](#)
  - and derived set, [260](#)
  - and inclusion, [256](#)
  - and interior, [255](#)
  - and intersection, [256](#), [261](#)
  - and limit point, [260](#), [261](#)
  - in metric space, [254](#)
  - and neighborhood, [253](#)
  - of product, [288](#)
  - in a subspace, [258](#)
  - and union, [256](#), [261](#), [262](#)
- cluster point, [217](#)
- clustering filter, [447](#), [447–448](#), [483](#), [487](#)
  - and convergent filter, [448–449](#)
- clustering net, [433](#), [433–447](#)
  - and convergent net, [434](#)
  - and convergent subnet, [436](#)
- clustering sequence, [217](#)
  - and convergent subsequence, [434](#)
- co-finite topology (*see* finite-complement topology)
- coarser filter, [448](#)
  - strictly, [448](#)
- coarser topology, [306](#)
- codomain, [19](#), [21](#)
- codomain restriction, [24](#)
- coffee cup, [359](#)
- cofinal subset, [433](#), [453](#)
  - see also* frequently in a set
- cofinite filter (*see* finite-complement filter)
- collapsing to a point, [404–405](#), [413–415](#), [520](#)
  - and attaching, [415](#)
  - and countability properties, [414](#)
  - and locally compact space, [519](#)
  - and separation properties, [409](#), [414](#), [519](#), [627](#)
- collection, [13](#)
  - see also* set
- comb (*see* topologist's comb)
- commutative diagram, [26–27](#)
- commutative operation, [22](#)
- compact set
  - and nets, [487](#)
  - in topological group, [492–493](#)
- compact set (*see* compact space)
- compact space, [460](#), [460](#)
  - and base, [484](#)
  - and bounded set, [479](#)
  - and box product, [488](#)
  - and closed set, [464](#), [479](#)
  - and continuous map, [475–476](#)
  - and countable open cover, [496](#)
  - and filter bases, [487](#)

- and filters, 483–610
- and finite intersection property, 463
- and Hausdorff space, 464–465, 476
- and limit points, 498
- and locally compact space, 510
- and metrizable space, 496
- and nets, 482–483
- as normal space, 474
- and product space, 478
- and quotient space, 475
- and second-countable space, 493
- and sequences, 496, 497
- and subbase, 606
- and subspace, 463–464
- and ultrafilters, 610
- see also* countably compact space, limit-point compact space, sequentially compact space
- compactification, 506
  - compactification
    - Alexandroff (*see* one-point compactification)
- compactly generated space, 490, 522
- compactness (*see* compact space)
- comparability, 73
- comparison test, 187
- complement, 14
  - and boundary, 250
  - and closure, 254
  - and interior, 254
  - of intersection, 14, 40
  - of union, 14, 40
- complete (*see* complete metric space; order-completeness)
- complete metric space, 202, 202–217
- complete regularity (*see* completely regular space)
- completely regular space, 293, 612
  - and locally compact space, 622
  - and Cartesian sum, 615
  - and embedding, 633
  - and metrizable space, 198
  - and normal space, 616, 617, 627
  - and product space, 615
  - and quotient space, 629
  - as regular space, 613
  - and subspace, 615
  - and  $T_2$ -space, 613
  - and Tychonoff cube, 633
- completion of a metric space, 222, 222–223
- complex exponential function, 90–91, 100
- complex numbers, 9, 37, 89, 89–91, 100
- complex-valued function, 90
- component, 546, 546–548, 552
  - and connected sets, 547
  - and homeomorphism, 547–548
  - in product space, 556
  - of topological group, 558
- components of map, 370, 371
- composite of maps, 25, 25–27
  - and continuity, 304
  - see also* composite of relations
- composite of relations, 18
- condensation point, 290, 290
  - two-sided, 290
- cone over a space, 415
- congruence modulo  $m$ , 94, 99
- conjugate, 90
- connected set, 524
  - and closure, 526
  - and components, 547
  - in real line, 525
  - see also* connected space
- connected space, 336, 524, 524
  - and continuous map, 527
  - and path-connected space, 563–564
  - and quotient space, 527
  - see also* totally disconnected space
- connectedness (*see* connected space)
- constant map, 21
- containment, 10
- continuum, 545
- continuous bijection principle, 476
- continuous function
  - and bounded function, 459, 476
- continuous function (*see* continuous map)
- continuous image, 314
  - of Hausdorff space, 415
  - of second-countable space, 319
  - and separable space, 315
- continuous map, 302, 302
  - and convergent sequences, 422
  - and filters, 446
  - and limit, 438
  - and nets, 428
  - and base, 303
  - and closed graph, 380, 488
  - and closed sets, 303
  - and closure, 304, 316, 319
  - and codomain, 305
  - on compact metric space, 502
  - on compact space, 475–476
  - and composite of maps, 304, 305

- and connected space, 527
- and convergent sequence, 194
- from Euclidean  $n$ -space to real line, 128
- and extension, 307
- and gluing, 309–312
- and group homomorphism, 419
- and inverse images, 303
- and local base, 316
- between metric spaces, 180, 180–185
- and neighborhoods, 302
- and open sets, 184
- at a point, 302, 316
- and product space, 367–377
- and quotient map, 390–393
- and restriction, 305
- and subbase, 303
- uniformly, 218
- and uniformly continuous map, 502
- see also* closed map, homeomorphism
  - open map, quotient map
- continuous map principle, 475–476
- contractible space, 572, 573
  - and homotopy, 576
  - as path-connected space, 573
  - as simply connected space, 589
- contraction, 572, 572–573
  - as deformation retraction, 575
  - see also* contractible space, deformation retraction, homotopy
- Contraction Mapping Principle, 220–221, 530
- contraction of metric space, 220, 221
  - see also* contraction mapping principle, Picard existence theorem
- contraction of metric space(, 220
- converge
  - pointwise, 431
  - uniformly, 189
- convergence, 420–450
  - and box topology, 452
- convergent filter, 443, 443
  - and clustering filter, 449
  - and convergent net, 444
- convergent filter base, 443, 443
- convergent net, 425, 425
  - and clustering net, 434
  - and convergent filter, 444
  - and limit point, 452
- convergent sequence, 185, 185–195, 421
  - and closed set, 421
  - and closure, 421
  - and continuous map, 422
  - and open set, 421
  - and bounded sequence, 186
  - and Cauchy sequence, 201
  - and continuous map, 194–195
  - see also* pointwise convergent sequence of functions, uniformly convergent sequence of functions
- convergent series, 187
- convergent subnet
  - and clustering net, 436
- convergent subsequence
  - and clustering sequence, 434
- convex set, 37, 198, 261, 486
  - as contractible space, 572
  - as simply connected space, 587
- coordinate, 34
- coordinate domain, 338
- coordinate map, 338
- corkscrew, 629
- countability properties, 268, 277
  - and Cartesian sum, 357, 359
  - and product space, 378
  - of quotient space, 408–409, 414
- countable family, 59
- countable set, 57
  - and infinite subset, 58
  - and sequence, 59
  - and surjection, 58
- countable-complement topology, 234, 484
  - and convergence, 422
  - and second-countable space, 287
- countably compact space, 496, 496, 505, 508
  - and sequentially compact space, 505
- cover, 40, 40–41, 42, 42, 309
  - and base, 270
  - of compact metric space, 501
  - and open cover, 279, 460
  - and subbase, 283
- covering (*see* cover, covering space)
- covering dimension, 642
- cube (*see* centered  $n$ -cube, Hilbert cube,  $n$ -cube)
- curve, 564, 564–571
  - see also* loop, path, plane-filling curve, simple closed curve, space-filling curve, topologist's sine curve
- cut point, 542
- cylinder, 353

as quotient space, 395  
 solid, 353  
 unbounded, 353  
 cylinder set, 364

## D

de Moivre's Formula, 91  
 De Morgan's Laws, 14, 40  
 decimal expansion, 80  
 decreasing map, 68, 68–69  
 decreasing sequence, 69, 69, 76  
 decreasing sequence of sets, 84  
 Dedekind cut, 92, 126  
 deformation retract, 574, 574–575  
 deformation retraction, 574, 574–575  
   and contraction, 575  
   as homotopy, 576  
   *see also* deformation retract, retract,  
     retraction  
 deleted sequence space, 287  
   and metrizable space, 638  
   and second-countable space, 287  
   and separation properties, 299  
 dense set, 254, 280  
   and Baire Category Theorem, 213  
   and connected space, 526  
   in discrete space, 254  
   in metric space, 213, 213  
   in metric space, 217  
   and self-dense space, 260  
 dense-in-itself space, 245  
   *see also* self-dense space  
 denumerable set, 57  
   and sequence, 58  
   as subset, 58  
 derived set, 244  
   and closure, 260  
 diagonal, 15, 18, 380  
   and  $T_2$ -space, 276  
 diagonal map, 306  
 diagram, 26–27  
   commutative, 26–27  
 diameter, 486  
 diameter of a set, 167  
 dictionary ordering  
   *see also* reverse dictionary ordering  
 dictionary ordering (*see* lexicographic  
   ordering)  
 differentiable manifold, 343  
   *see also* manifold  
 digits

base expansion  
   of base expansion, 80  
 dimension, 266–642  
   and locally Euclidean space, 336  
   covering, 642  
   inductive, 640  
   large inductive, 640  
   of a manifold, 343  
   and parameters, 565, 638  
   small inductive, 640  
   *see also* covering dimension, small  
     inductive dimension, large in-  
     ductive dimension, zero-dimensional  
 direct image (*see* image of a set)  
 directed set, 423  
   *see also* direction  
 direction, 423, 423–425  
   natural, 424  
 disconnected space, 524, 525  
   *see also* connected space; separation;  
     totally disconnected space  
 disconnects, 535  
 discontinuous map, 180, 302  
   *see also* continuous map  
 discrete metric, 140, 140, 150  
   and open sets, 153  
   *see also* discrete space, discrete topol-  
     ogy  
 discrete space, 227, 236, 263, 379, 461  
   as locally connected space, 553  
   and topological property, 334  
 discrete topology, 227  
   and local base, 265  
 disjoint collection, 40  
 disjoint family, 40  
 disjoint sets, 13, 40  
   *see also* disjoint collection, disjoint  
     family  
 disk, 147, 147  
   punctured, 574  
   *see also*  $n$ -disk  
 dispersion point (*see* explosion point)  
 distance, 129–139  
   between sets, 163  
   between subsets, 486  
   from point to set, 160  
   max, 134  
   from point to set, 198  
   in  $\mathbb{R}^n$ , 129  
   taxicab, 133  
   *see also* metric  
 distance-preserving map, 172, 173

*see also* isometry  
 distinguished point, 228  
 distributive operation, 22  
 divides, 63  
 divisibility relation, 63  
 domain, 21  
     of a relation, 15  
 domain of a map, 19  
 domain-codomain restriction, 24  
 double, 418  
 doubleton, 9  
 doubly-indexed family, 34  
 doughnut, 359  
 duality, 254  
 dyadic rational, 83

**E**

edge of Möbius strip, 412  
 element, 7  
 embedding, 346, 346–635  
     homeomorphic, 347  
     topological, 347  
     and Tychonoff cube, 634  
 empty relation, 16  
 empty set, 9  
     and relation, 16  
 end of a path, 559  
 endow set with a topology, 226  
 endpoint, 72  
 entry, 34  
 $\varepsilon$ -chain, 544  
 equal maps, 21  
 equal sets, 7  
 equality  
     of maps, 21  
 equator, 402  
 equivalence class, 96, 96  
     and representative, 96  
 equivalence kernel, 95, 105  
     and quotient map, 411  
 equivalence relation, 16, 17, 94, 94–103  
     closed, 416, 416–417  
     and compatible operation, 106  
     and compatible relation, 106  
     open, 416, 416–417  
     induced by partition, 98  
     and saturated set, 99  
     and saturation, 99  
 equivalent metrics, 164, 164–166, 286  
 Euclidean  $n$ -space  
     as second-countable space, 277

    and usual topology, 227  
 Euclidean 0-space, 36, 129, 140, 227  
 Euclidean 3-space, 36, 128  
 Euclidean distance, 129  
 Euclidean inner product, 178, 198  
 Euclidean metric, 128–140  
     and balls, 147  
     and disks, 147  
     and spheres, 147  
     and max metric, 165  
     on  $\mathbb{R}^0$ , 129  
     and taxicab metric, 176  
 Euclidean  $n$ -space, 36, 36  
     and base, 286  
     and compact space, 479  
     as complete space, 205  
     as connected space, 534  
     and dimension, 640  
     and homeomorphism, 324, 327  
     and local base, 264  
     as locally connected space, 553  
     *see also*  $\mathbb{R}^n$   
 Euclidean norm, 130  
 Euclidean plane, 36, 128  
     *see also*  $\mathbb{R}^2$   
 Euclidean space (*see* Euclidean  $n$ -space)  
 evaluation map, 634  
 eventuality filter, 441, 442, 449  
 eventually constant net, 425  
 eventually constant sequence, 189, 422  
 eventually in a set, 425  
 ex:attach- $n$ -cell, 408  
 ex:wo-pf-sqrt-2-irrational, 12  
 excluded point topology, 233  
 expansion  
     for a base, 80  
     repeating, 92  
 expansion (*see* base expansion)  
 explosion point, 549–552  
 extended real line, 168, 168–172, 195, 218, 273, 484  
     as connected space, 526  
     and dense set, 254  
     homeomorphic to closed interval, 321  
     and second-countable space, 287  
 extends a map, 24  
 extension  
     of continuous map, 307–312  
     of uniformly continuous map, 218  
 extension of map, 24, 24–25, 41  
 exterior, 261

extreme (*see* maximum, minimum)  
 Extreme Value Theorem, 476

## F

factor space, 364  
 family, 34  
     doubly-indexed, 34  
     finite, 55  
     and functional relation, 34  
     infinite, 55  
 fiber, 27  
 field, 23  
     ordered, 78  
 figure-eight, 351, 351–352, 414, 543, 565  
 filter, 440, 440, 440–451  
     and closed set, 445  
     and closure, 445  
     and clustering, 447  
     and continuity, 446  
     and convergence, 443  
     and neighborhood system, 441  
     and open set, 445  
     and product space, 447  
     and refinement, 448  
     and subset, 441  
     and associated net, 450  
     and clustering, 483, 487  
     coarser, 448  
     in compact space, 483–610  
     eventuality, 441, 442, 449  
     extended, 442  
     finer, 448  
     finite-complement, 442  
     fixed, 611  
     free, 611  
     generated by a filter base, 442, 442–448  
     in Hausdorff space, 446  
     induced on subset, 442  
     on a subset, 442–443  
     principal, 441  
     ultrafilter, 607  
     *see also* filter base, neighborhood filter, principal filter, ultrafilter  
 filter base, 442  
     and convergence, 443  
     and net, 442  
     in compact space, 487  
     *see also* filter, net  
 final topology, 383, 420  
 finer filter, 448

    strictly, 448  
 finer topology, 306  
 finite family, 55  
 finite intersection property, 462  
     and Tychonoff Theorem, 610  
     *see also* compact space  
 finite set, 53, 53  
 finite-complement filter, 442  
 finite-complement topology, 228, 236, 315, 429  
     and compact space, 462  
     and locally compact space, 510  
     and second-countable space, 287  
     and separation properties, 296  
 first-category set (*see* meager set)  
 first-countable space, 268, 268–269  
     and convergent sequences, 421, 422  
     and continuous image, 315  
     and product space, 378, 382  
     and second-countable space, 278  
     and topological property, 334  
 fixed filter, 611  
 fixed-point, 21, 34, 220, 530, 599  
 fixed-point theorem, 530  
 filter base, 442  
 Fort space, 268, 297  
     *see also* Fort topology  
 Fort topology, 268, 286, 299, 315, 317  
 Fortissimo space, 289  
     *see also* Fortissimo topology  
 Fortissimo topology, 289, 317  
 Fréchet filter, 442, 444  
 free filter, 611  
 frequently in a set, 433  
 $F_\sigma$ -set, 163, 235  
 function, 19  
     bounded, 141, 143  
     *see also* bounded map  
     function  
         continuous (*see* continuous map)  
         *see also* map  
 function space, 135, 136, 274–276, 281  
 functional relation, 16, 16  
     as family, 34  
     and map, 19  
 fundamental system of neighborhoods  
     *see also* local base  
 Fundamental Theorem of Algebra, 530

## G

$G_\delta$ -set, 163, 235

general linear group, 384  
 Gluing Lemma, 309, 573  
 gluing maps, 309, 573  
 graph, 19, 19–20  
   of continuous map, 380, 488  
   of a relation, 15  
 graph of map, 21  
 greatest element, 68  
   *see also* maximum

## H

Hahn–Mazurkiewicz Theorem, 569  
 half-disk space, 242, 242–243, 264, 268, 316, 358, 484  
   and first-countable space, 268  
   and local bases, 285  
   and metrizable space, 636  
   and second-countable space, 287  
   and separable space, 282  
 half-open interval, 70  
 halfspace (*see* closed-halfspace, open-halfspace)  
 Hausdorff Maximal Principle, 125  
 Hausdorff metric, 177  
 Hausdorff space, 241, 241–293  
   and filters, 446  
   and limit of a map, 438  
   and nets, 429–430  
   and collapsing to a point, 409  
   and compact set, 464–465  
   and compact space, 476  
   and continuous image, 415  
   and diagonal, 380  
   and metric space, 159–160  
   and metrizable space, 242  
   and product space, 377, 380  
   and quotient space, 409–410  
   and subspace, 243  
   and  $T_1$ -space, 296, 298  
   and topological property, 334  
 Heine–Borel–Lebesgue Theorem, 466, 479  
 hemisphere, 353, 402  
 Hilbert cube, 139, 162, 177, 494–501, 631–633  
   as compact space, 501  
   as connected space, 539  
   and embedding, 632–633  
   as product of intervals, 373  
   and Tychonoff cube, 633  
 Hilbert distance, 138

Hilbert metric, 138  
 Hilbert norm, 138, 143  
 Hilbert sequence space, 138, 138–139, 162, 177, 197, 280, 381, 510  
   as complete space, 206  
 Hilbert space (*see* Hilbert sequence space)  
 Hilbert’s hotel, 57  
 Hilbert cube  
   and embedding, 634  
 homeomorphic embedding (*see* embedding)  
 homeomorphic spaces, 320  
   *see also* topologically equivalent metric spaces  
 homeomorphism, 320, 320  
   and components, 547  
   local, 336  
   and passing to the quotient, 392  
 homogeneous space, 355, 355, 419  
   and  $n$ -ball, 355  
   and  $n$ -disk, 602  
 homotopy, 576  
   *see also* contraction, deformation retraction, path-homotopy  
 homotopy of maps, 576  
 hyperplane, 198

## I

identity element, 22  
 identity map, 20, 23  
   as continuous map, 305  
 image  
   and intersection, 40  
   and union, 40  
 image of a filter base, 445, 445–447  
 image of a set, 27, 27–28  
   and intersection, 28  
   image of a set  
     inverse (*see* inverse image of a set)  
   and union, 28  
 imaginary part, 89  
 included point topology, 233  
 inclusion, 11  
   *see also* element, member, subset  
 inclusion map, 23  
   as continuous map, 305  
 increasing map, 68, 68–69  
 increasing sequence, 69, 69, 76, 186  
 increasing sequence of sets, 84  
 index, 34



- index origin, 35
  - index set, 34
  - indiscrete pseudometric, 144
    - see also* indiscrete topology
  - indiscrete space, 236, 461
  - indiscrete topology, 227
  - induced metric, 140
    - and balls, 151
  - induction
    - transfinite, 124
    - see also* Axiom of Induction, mathematical induction
  - inductive step, 11
  - inequivalent metrics, 164, 166
  - infimum, 71, 71–79
  - infinite family, 55
  - infinite limit
    - of sequence, 200
  - infinite series (*see* series)
  - infinite set, 53, 53
    - and denumerable subset, 58
    - see also* countable set, uncountable set
  - initial condition, 47, 48
  - initial point of a path, 559
  - initial topology, 383, 420
    - see also* Cartesian sum, quotient space
  - injection, 28, 28–30
  - injective map, 28
    - see also* injection
  - inner product (*see* Euclidean inner product)
  - inside, 592
  - integers, 9
  - interior, 250, 250
    - and boundary, 260
    - and closure, 255
    - of complement, 254
    - and inclusion, 252
    - and intersection, 252, 261
    - and open set, 251–252
    - of product, 288
    - in a subspace, 258
    - and union, 252, 261
  - interior point, 250
    - and neighborhood, 251
    - see also* interior
  - Intermediate-value Theorem, 528, 528–530
  - intersecting sets, 13
  - intersection
    - of a collection of sets, 41, 41–42
    - of a family of sets, 38
    - and image, 40
    - of interiors, 252
    - and inverse image, 28, 40
    - of open sets, 156
    - of two sets, 13, 13–14
  - interval, 70, 86
    - and open sets, 153
    - and cardinality, 119
    - closed, 70
    - as connected set, 524, 525
    - and dense set, 254
    - and endpoints, 72
    - half-open, 70
    - nondegenerate, 70, 71
    - open, 70
    - see also* ray
  - invariance of domain, 357
  - Invariance of Domain Theorem, 342
  - invariant
    - topological, 640
  - inverse, 30, 30–31
    - additive, 22
  - inverse image, 27, 27
    - and intersection, 40
    - and union, 40
  - inverse image of a set, 28
    - and intersection, 28
    - and set difference, 28
    - and union, 28
  - irrational numbers
    - as dense set, 254
  - irreflexive relation, 64, 67, 75
  - irreflexivity (*see* irreflexive relation)
  - isolated point, 241, 241–244
  - isometric embedding, 178, 347
    - see also* embedding
  - isometric metric spaces, 173
  - isometry, 173, 173–174
  - iterate of map, 49
- J
- Jordan Curve Theorem, 592, 591–592
- K
- Kelley subnet, 455, 456
  - Klein bottle, 400, 399–400, 415, 420, 563
    - as compact space, 480
    - and cylinder, 412
  - Knaster–Kuratowski space, 549–552



Knaster-Tarski Fixed-point Theorem, 34  
 and Cantor-Bernstein Theorem, 126  
 knot  
   trefoil, 348  
 Kolmogoroff quotient, 417  
 Kuratowski closure axioms, 262  
 Kuratowski closure-complement problem, 263

## L

$L_1$ -metric, 136  
 large inductive dimension, 640  
 larger topology (*see* finer topology)  
 largest element (*see* greatest element)  
 least element, 68  
   *see also* minimum  
 Lebesgue Covering Lemma, 501, 501–502  
 Lebesgue number, 501  
 left-closed, right-open interval, 70  
 left-open, right-closed interval, 70  
 Lebesgue number, 502  
 lexicographic ordering, 73, 286, 287, 484  
   and order-complete set, 76  
 lexicographically ordered product, 73  
   *see also* lexicographic ordering  
 lexicographically ordered square, 286  
   as compact space, 484  
   and Lindelöf space, 484  
 lift of a map, 32, 93  
 liftable map, 32, 93  
 lifting (*see* lift)  
 limit  
   and base expansion, 80  
   one-sided, 439  
 limit at infinity, 439, 521  
 limit of a filter, 446  
 limit of a map, 437, 437  
   and continuity, 438  
   and Hausdorff space, 438  
 limit of a net, 430  
 limit of a sequence  
   infinite, 200  
 limit ordinal, 111  
 limit point, 241, 261  
   and closure, 260  
   in compact space, 498  
   and convergent net, 452  
   in  $\mathbb{R}$ , 290, 498, 504  
   in  $T_1$ -space, 244  
   vs.  $\omega$ -accumulation point, 244

*see also* adherent point, condensation point, derived set, perfect space  
 Lindelöf space, 280, 289, 289–290, 295, 357, 508  
   and product space, 289  
   and quotient space, 420  
   and subspace, 289, 290, 484  
   *see also* Lindelöf Theorem  
 Lindelöf Theorem, 280, 504  
   and second-countable space, 289  
   *see also* Lindelöf space  
 line, 37  
 line segment, 37, 37  
 line with two origins, 240, 242, 265, 270, 316, 350, 415, 484  
   and manifold, 343, 356  
   as quotient space, 404, 409, 616  
   and separation properties, 296, 409, 616  
 linearly open set, 234  
 Lipschitz condition, 221  
 local base, 263, 263–269  
   and boundary, 264  
   and continuity, 316  
   in product space, 382  
   and topology, 265–266  
 local coordinate system, 340  
 local coordinates, 338  
 local homeomorphism, 336  
 locally compact space, 508, 508–518  
   and closed subspace, 511, 519  
   and compact space, 510  
   as completely regular space, 622  
   and continuous image, 511  
   and continuous image, 511  
   and Hausdorff space, 508–509  
   and manifold, 519  
   and normal space, 616, 617  
   and open subspace, 511, 519  
   and product space, 512–513  
   and proper map, 519  
   and quotient space, 512, 519  
   and second-countable space, 518  
 locally connected space, 552  
 locally constant function, 543  
 locally Euclidean space, 329, 336, 336–346  
 locally finite collection, 485  
   *see also* locally finite family  
 locally finite family, 262  
   *see also* locally finite collection

locally homeomorphic spaces, 336  
 locally path-connected space, 571  
 long closed ray, 344  
 long line, 344, 344–346, 356  
   and locally Euclidean space, 344  
 loop, 581, 581–585, 587  
   as map on  $S_1$ , 582–583  
   nullhomotopic, 588  
   *see also* null loop, nullhomotopic loop  
 lower bound, 71, 71  
   greatest, 71  
   *see also* infimum  
   and infimum, 79  
 lower semicontinuous map, 319, 487  
 lower-limit topology  
   *see also* right-interval topology, Sorgenfrey line, upper-limit topology

## M

Möbius strip  
   as compact space, 480  
 manifold, 329, 343, 343–346  
   and connected space, 564  
   differentiable, 343  
   as locally compact space, 519  
   and path-connected space, 564  
   and product space, 381  
 manifold  
   with boundary (*see* manifold-with-boundary)  
 manifold-with-boundary, 357, 412, 418–419  
   and double, 418  
 map, 19, 19  
   constant on equivalence class, 101  
   defined piecewise, 41, 309  
   and functional relation, 19  
   liftable, 32, 93  
   defined piecewise, 24–25  
 map  
   proper (*see* proper map)  
   uniformly continuous, 218, 502  
   well-defined, 101  
 map onto (*see* surjection)  
 maps (verb), 20  
 mathematical induction, 11–12, 18  
 max distance, 134  
 max metric, 134, 140, 150, 360  
   and Euclidean metric, 165

  mainref, 136  
   topology induced by, 227  
   *see also* sup metric  
 max norm, 142, 143  
   *see also* max metric  
 maximal chain, 74, 125  
 maximal element, 113, 113, 113  
 maximal principle, 113  
 maximum, 68  
   of semicontinuous function, 487  
   of continuous function, 476  
 meager set, 257  
 member, 7  
 Menger sponge, 492  
 metric, 139, 129–139  
   and ultrametric, 145  
   bounded, 167–171  
   compatible, 228  
   discrete, 140, 140, 150  
   and equivalence, 164  
   Euclidean, 129, 140  
   examples of, 129, 132–141, 146  
   induced, 140, 231  
   max, 134, 150  
   on product, 140  
   on subset, 140  
   on product, 176, 360–363  
   on  $\mathbb{R}^n$ , 129  
   taxicab, 133  
   *see also* pseudometric  
 metric space, 139  
   and closed sets, 157–159  
   and open sets, 152–153, 158–159  
   bounded, 167  
   complete, 202  
   and dense subset, 213, 213–217  
   incomplete, 202  
   and isometry, 173  
   and local base, 265  
   as normal space, 160–161, 163  
   as regular space, 159  
   as  $T_1$ -space, 158  
   as  $T_2$ -space, 159–160  
   *see also* bounded metric space, complete metric space, etc.  
 metric spaces  
   metrically equivalent, 174  
   topologically equivalent, 174, 174–176  
 metrically equivalent metric spaces, 174  
 metrizable space, 228, 228, 229, 263, 383, 486

and compact space, 496  
 as completely regular space, 198  
 as first-countable space, 268  
 and Hausdorff space, 242  
 as a Hausdorff space, 242  
 and Hilbert cube, 632–633  
 as normal space, 297  
 and product space, 377  
 and regular space, 297, 299  
 and second-countable space, 282, 493, 635–636  
 and separable space, 282  
 and subspace, 231  
 and topological property, 333  
 metrization theorem, 635  
 minimal element, 113  
 minimum, 68  
     of semicontinuous function, 487  
     of continuous function, 476  
 Möbius strip, 399, 412, 415, 563  
     as compact space, 480  
     open, 413  
 modulus, 90, 94  
 Moore plane (*see* tangent disk space)  
 multiplication, 22  
 multiplicative inverse, 23  
 Mysior's space, 614–615, 627

## N

$n$ -ball, 149  
     as connected space, 534  
     as homogeneous space, 355  
 $n$ -cell, 327, 381  
     as manifold-with-boundary, 357, 358  
     as compact space, 480  
 $n$ -cube  
     and  $n$ -manifold-with-boundary, 357  
     as compact space, 479  
     and homeomorphism, 324, 325  
     *see also* centered  $n$ -cube  
 $n$ -dimensional Euclidean space, 336  
 $n$ -dimensional manifold  
     *see also* manifold,  $n$ -manifold  
 $n$ -disk, 149  
     and  $n$ -manifold-with-boundary, 357  
     as compact space, 480  
     and cone, 415  
     as connected space, 534  
     and homeomorphism, 325  
     and homogeneous space, 355, 602  
     and suspension, 415  
 $n$ -manifold, 329, 343, 343–346, 358  
     and open subset, 356  
     and dimension, 343  
 $n$ -manifold-with-boundary, 357, 358  
 $n$ -simplex, 354, 357  
     standard, 354, 357  
 $(n - 1)$ -sphere, 149  
 $n$ -sphere, 149  
     as compact space, 480  
     and cone, 415  
     and homeomorphism, 327  
     and one-point compactification, 517  
     as quotient of  $n$ -disk, 520  
     and suspension, 415  
     *see also*  $(n - 1)$ -sphere  
 $n$ -torus, 357, 398  
     as compact space, 480  
     *see also* torus  
 $n$ -tuple, 35, 35, 37  
 $n$ -vector, 37  
 natural direction, 424  
 natural numbers, 9  
 negative of vector, 36  
 neighborhood, 236, 236–238, 263  
     and closed set, 161  
     and open sets, 161  
     and open set, 237  
     in product space, 366  
     of a set, 237, 244  
     in subspace, 236  
     typical, 263  
 neighborhood base, 263  
     *see also* local base  
 neighborhood filter, 441, 444  
 neighborhood system, 236  
     as filter, 441  
     *see also* neighborhood filter  
 Nested Interval Property, 84, 212  
 Nested Set Theorem, 212  
 net, 425  
     and closed set, 428  
     and closure, 428  
     and clustering, 433  
     and continuity, 428  
     and convergence, 425  
     and open set, 428  
     and product space, 430  
     associated with a filter, 450  
     in compact space, 482–483, 487  
     eventually constant, 425  
     eventually in a set, 425  
     frequently in a set, 433

- in Hausdorff space, 429–430
  - see also* filter, subnet, ultranet
- Niemytzki plane (*see* tangent disk space)
- No-Retraction Theorem, 598
- non-cut point, 545
- noncompact, 460
- noncompact space
  - see also* compact space
- nondegenerate interval, 70, 71
- nondegenerate set, 9
- nonnegativity
  - of a metric, 139
- nonnegative, 10
- nonterminating base expansion, 80
- norm, 130–138
  - Euclidean, 130, 130–131
  - see also* max norm, sup norm, taxi-cab norm
- normal space, 293, 299–300, 310, 616–627, 634
  - and cardinality, 622
  - and closure, 299
  - and compact space, 474
  - and completely regular space, 616, 617, 627
  - and discrete subspace, 622
  - and Lindelöf space, 628
  - and locally compact space, 616, 617
  - and metrizable space, 160–161, 163, 297
  - and product space, 296, 300, 622
  - and quotient space, 623, 627, 628
  - and regular second-countable space, 616
  - and regular space, 299, 616
  - and separable space, 622
  - and Sorgenfrey line, 295
  - and Sorgenfrey plane, 296, 622
  - and subspace, 617–618
  - and tangent disk space, 300
  - and Tietze Extension Theorem, 623
- normal subgroup, 419
- north pole, 327
- nowhere dense set, 257, 262
- null loop, 581
  - and path-homotopy, 583, 584
- nullhomotopic, 588
- nullhomotopic loop
  - and simply connected space, 588
- number of elements, 53, 54
- of finite character, 125
- $\omega$ -accumulation point, 244
- one-dimensional space, 642
- one-point compactification, 516, 513–627
  - of integers, 520
  - of natural numbers, 520
- one-point union, 414
- one-to-one (*see* injective map)
- one-to-one correspondence, 28
  - see also* bijection
- open annulus, 356
- open box, 353, 355
- open cover, 279, 279–280, 460, 460
  - in second-countable space, 280
- open equivalence relation, 416, 416–417
- open interval, 70, 460
  - and open sets, 153
  - and homeomorphism, 322
  - as incomplete space, 202
- open map, 313
  - and base, 314
  - and quotient map, 393
- open Möbius strip, 413
- open ray, 70
  - as open set, 154
- open set, 226
  - and convergent sequences, 421
  - and filters, 445
  - and intersection, 156
  - and neighborhoods, 161
  - and nets, 428
  - basic, 269
  - and boundary, 250
  - in metric space, 152, 152–153, 158–159
  - in real line, 153
  - in subset, 155
  - and neighborhood, 237
- open-halfspace, 198
- operation, 22, 22–23
  - and additive inverse, 22
  - associative, 22
  - binary, 22
  - and cancellation, 23
  - commutative, 22
  - distributive, 22
  - and identity element, 22
  - and multiplicative inverse, 23
- order isomorphism
  - and homeomorphism, 355
- order relation (*see* ordering relation)

order topology, 272, 272–274, 287, 355, 462  
     and compact space, 484  
     and Hausdorff space, 272, 286  
 order-complete set, 72, 79  
     and compact space, 462  
     and lexicographic ordering, 76  
 order-dense set, 75, 83, 356  
 order-isomorphic sets, 69, 69–70  
 order-isomorphism, 69, 69–70  
 order-preserving map, 65, 65, 70  
     *see also* increasing map  
 order-reversing map, 65  
     *see also* decreasing map  
 ordered field, 78  
 ordered pair, 14  
     as 2-tuple, 35  
 ordering relation, 16  
 ordering relation (*see* directed preordering; partial ordering; preordering; strict ordering; total ordering; weak ordering; well-ordering)  
 ordinal, 110–112, 273, 343–346  
     countable, 110  
     finite, 110  
     first infinite, 110, 273  
     first uncountable, 110, 273  
     initial, 122  
     limit, 111  
     *see also* long line  
 ordinal number (*see* ordinal)  
 ordinal product, 124  
 ordinal space, 273, 288  
     and compact space, 462, 497, 505  
     and countability properties, 288  
     and Lindelöf space, 290  
 ordinal sum, 124  
 origin, 36  
 orthogonal group, 384  
 outside, 592

## P

$p$ -adic metric, 146, 199  
 $p$ -adic norm, 145  
 parametrization, 565  
 partial order (*see* partial ordering)  
 partial ordering, 66, 66–70  
     coordinatewise, 67  
     and product, 67  
 partially ordered set, 66

*see also* partial ordering  
 partition, 98, 98–99  
     and saturated set, 99  
     and saturation, 99  
     *see also* equivalence relation, quotient set  
 partition of an interval, 426  
 passing to the quotient, 102, 100–103, 391–392  
 path, 412, 559  
     reverse, 560  
     path  
         reverse (*see* reverse of path)  
         *see also* arc, path-connected space  
 path product, 559  
 path-component, 570  
 path-connected space, 561, 561, 561  
     as connected space, 563–564  
     and continuous map, 563  
     and contractible space, 573  
     and retract, 586  
     *see also* arc-connected space, polygonally connected set, simply connected space  
 path-connectedness (*see* path-connected space)  
 path-homotopic, 577  
 path-homotopy, 577, 577–585  
     and path-product, 579  
     and reverse of path, 580  
 path-product, 559  
     and associativity, 580–581  
     and commutativity, 581  
     and path-homotopy, 579  
 pathwise connected space (*see* path-connected space)  
 perfect map, 489, 489  
     *see also* proper map  
 perfect set, 261  
 perfect space, 261  
 permutation, 29  
 pseudometric  
     and completeness, 221  
 Picard Existence Theorem, 221  
 piecewise, 24–25, 41, 309  
 piecewise linear function, 215, 503  
 plane with two origins, 636  
 plane-filling curve, 565–568, 626, 638  
 point, 7  
 point at infinity, 516  
 pointed space, 228, 417, 418  
 pointwise bounded family, 522

- pointwise convergence, 276, 284, 290, 379, 383, 431–432, 607
  - and nets, 431
  - and uniform convergence, 432
- pointwise convergent sequence of functions, 209, 209–210
- polygonal path, 570
- polygonally connected set, 570
- polyhedron, 198
- polynomial, 528–530
- positive, 10
- positivity
  - of a metric, 139
- power of set, 35, 37, 37–38, 44
  - and cardinality, 119
- power set, 13, 57
  - and cardinality, 119
- precedes, 63
  - immediately, 65
  - strictly, 65
- precompact set, 485
- predecessor, 63
  - immediate, 65
  - strict, 65
- predicate, 4, 8
- preorder (*see* preordering)
- preordered set, 63
  - see also* preordering
- preordering, 63, 63
  - directed, 423–425
  - induced, 64
  - inherited, 64
  - strict, 64
  - weak, 64
- preserve distance, 172
- principal filter, 444
- principal filter, 441
  - and filter base, 443
  - as ultrafilter, 607
- principal ultrafilter, 448
- Principle of Complete Recursion, 52
  - see also* Principle of Ordinary Recursion
- Principle of Ordinary Recursion, 48
  - see also* Principle of Complete Recursion
- Principle of Primitive Recursion, 50, 50–51
- Principle of Strong Induction, 18
- Principle of Transfinite Induction, 124
- product
  - of a family of sets, 43, 42–44
  - and partial ordering, 67
  - of paths, 559
  - of real-valued functions, 372
  - of sequence of sets, 43
  - of topological spaces, 364
  - of two sets, 15, 43
  - of two topological spaces, 274
  - see also* path product, product of maps, product space, product topology
- product of maps, 371
- product of paths, 559
- product space, 274, 288, 364, 359–379
  - and  $T_1$ -space, 377
  - and filters, 447
  - and nets, 430
  - and associative law, 381
  - and base, 364–365, 382
  - and basic open set, 364
  - and boundary, 288
  - and closed set, 379
  - and closure, 288
  - and commutative law, 374, 381
  - and compact space, 478, 488
  - and completely regular space, 615
  - and continuous map, 367–377
  - and countability properties, 382
  - and diagonal, 380
  - and embedding, 376, 381
  - and first-countable space, 378
  - and Hausdorff space, 377, 380
  - and interior, 288
  - and local base, 276, 382
  - and locally compact space, 512–513
  - and manifold, 381
  - and metrizable space, 377
  - and neighborhoods, 366
  - and neighborhoods, 366
  - and normal space, 296, 300, 622
  - and open set, 379
  - and projection, 364, 367–369
  - and regular space, 377
  - and second-countable space, 378
  - and separable space, 378, 382
  - and separation properties, 377, 615
  - and subbase, 363–364
  - and subbasic open set, 364
  - and subspace, 366–367
  - and  $T_0$ -space, 377
- product topology, 274, 288, 364
  - and topology of pointwise convergence, 379

projective  $n$ -space  
     as compact space, 480  
 projection, 27, 43, 197  
     and closed map, 313, 368, 488  
     of product space, 364, 367–369  
 projective line, 401, 412  
 projective  $n$ -space, 401, 412–413  
     *see also* projective plane  
 projective plane, 401–403, 413, 416  
 projective point, 401  
 proof by induction, 11  
     *see also* Principle of Mathematical Induction  
 proper class, 13  
 proper map, 489  
     and locally compact space, 519  
     *see also* perfect map  
 proper subset, 11  
 provide set with a topology, 226  
 pseudometric, 144, 144–145, 164, 221–222  
     and metric, 144  
 pseudometrizable space, 234, 298  
 punctured disk, 574  
 punctured line, 585  
 punctured plane, 105, 339, 585  
 punctured sphere, 327, 572

## Q

quasicompact space (*see* compact space)  
 quasicomponent, 557  
 quotient map, 96, 313  
     abstract, 392  
     and closed map, 393  
     and equivalence kernel, 411  
     mainref, 392  
     and open map, 390, 393  
     and quotient topology, 389  
     and saturated set, 100  
 quotient set, 96, 96–97  
     *see also* quotient space  
 quotient space, 389, 388–410  
     and closed graph, 410  
     of compact space, 475  
     of connected space, 527  
     and continuous map, 390–393  
     and countability properties, 408–409, 414  
     and homeomorphism, 392  
     and locally compact space, 519  
     of locally compact space, 512

    of normal space, 623, 628  
     and separation properties, 409–410, 414, 416, 417, 627  
     as  $T_2$ -space, 410  
     as  $T_2$ -space, 409–410  
     *see also* attaching a cell, attaching two spaces, collapsing a subspace to a point, cone over a space, connected sum, smash product, suspension of a space, wedge sum  
 quotient topology, 389, 388–390  
     and final topology, 420

## R

range, 21, 27  
     of a relation, 15  
 rare set (*see* nowhere dense set)  
 rational numbers, 9  
     as dense set, 254  
     as incomplete space, 202  
     and locally compact space, 510  
     and order-density, 83  
     and repeating expansion, 92  
 ray, 70  
     and open sets, 154  
 real line, 36  
     and closed sets, 157  
     and open sets, 153–154  
     as connected space, 336  
     as complete metric space, 202  
     and connected set, 525  
     as connected space, 524, 525  
     as dense set, 254  
     and dimension, 267, 640  
     and extended real line, 231  
     homeomorphic to open interval, 322  
     and local base, 264  
     as noncompact space, 461  
 real line with integers collapsed to a point, 408, 409, 413, 512  
 real numbers, 9  
     and Archimedean Ordering Property, 82  
     and cardinality, 120, 126  
     and Dedekind cuts, 92, 126  
     and Nested Interval Property, 84  
     and order-completeness, 79, 79–85  
     and rational numbers, 83  
     as uncountable set, 85, 86, 92  
     *see also* real line



real part, 89  
 real-valued map, 21  
 recurrence relation, 47, 48  
 recursion, 47–52  
     complete, 52  
     ordinary, 48, 47–51  
     primitive, 50, 50–51  
 recursive definition (*see* recursion)  
 refinement of a cover, 42, 42  
 refinement of a filter, 448  
 reflection, 173, 329  
 reflexive relation, 63, 66, 73, 94, 423  
 reflexive, symmetric, transitive closure of  
     relation, 96, 104  
 reflexivity (*see* reflexive relation)  
 regular spac, 299  
 regular space, 293, 299, 486  
     and closure, 299  
     and metric space, 159  
     and metrizable space, 299  
     and normal space, 299, 616  
     and product space, 377  
     and Sorgenfrey line, 295  
     and subspace, 297  
 regularity (*see* regular space)  
 relation, 15  
     antisymmetric, 66, 73  
     asymmetric, 67, 75  
     and comparability, 73  
     empty, 16  
     functional, 16, 16  
     and graph, 15  
     irreflexive, 64, 67, 75  
     reflexive, 63, 66, 73, 94, 423  
     reverse, 17  
     symmetric, 94  
     transitive, 63, 64, 66, 67, 73, 75, 94,  
         423  
     trivial, 16  
     *see also* equivalence relation, func-  
         tional relation, order relation  
 relative topology, 230, 230–233  
     and ordered space, 272, 286  
     and product space, 366–367  
     *see also* subspace  
 relatively compact set, 485  
 repeating expansion, 92  
 representative, 96  
 residual set, 262  
 restriction of map, 23, 23–24, 41  
     codomain, 24  
     and continuous map, 305

    domain-codomain, 24  
 retract, 575, 575  
     of path-connected space, 586  
     and simply connected space, 602  
 retraction, 575, 575  
     and unit disk, 598  
 reverse lexicographic ordering, 124  
 reverse of path, 560, 560, 561  
     and null loop, 584  
     and path-homotopy, 580, 584  
 reverse relation, 17  
     and partial ordering, 68  
     and preordering, 65  
 Riemann sphere, 517  
 Riemann sum, 427  
 right-interval topology, 239  
     right-interval topology  
         on real numbers (*see* lower-limit  
         topology, Sorgenfrey line)  
 rigid motion, 174  
     *see also* isometry  
 root  
     of a function, 21, 90  
     of a polynomial, 528–530  
 rotation, 142, 172, 177  
 rotation-invariant metric, 142, 172  
 rule of a map, 19

## S

saturated set, 99  
     and map, 33, 105  
 saturation, 99  
     and quotient map, 100  
 scalar multiplication, 371  
 scattered space, 245, 261, 288  
 second-category set (*see* nonmeager set)  
 second-countable space, 277, 277  
     and continuous image, 315, 319  
     and countable cover, 280  
     and embedding, 634  
     and first-countable space, 278  
     and locally compact space, 518  
     and metrizable space, 282, 493, 635–  
         636  
     and normal space, 616  
     and product space, 378, 382  
     and separable space, 282  
     and subspace, 278  
     and topological property, 334  
     and total boundedness, 495  
 section (*see* tail of a net)



- section of a surjection, 45
- self-dense space, 245, 245–468
  - and dense set, 260
- semicontinuous map, 319, 487
- sends, 20
- separable space, 280, 280
  - and metrizable space, 282
  - and continuous image, 315
  - and Hilbert cube, 632–633
  - and normal space, 622
  - and product space, 378, 382
  - and second-countable space, 282
  - and subspace, 282, 379
  - and topological property, 334
- separated sets, 541–542
- separated space (*see* Hausdorff space)
- separation, 524, 524
- separation axioms, 293, 298
  - see also* separation properties
- separation properties, 241, 293, 292–293
  - and Cartesian sum, 350, 359
  - and product space, 349, 377
  - of quotient space, 409–410, 414, 416, 417, 627
- sequence, 35, 38
  - sequence
    - and convergence (*see* convergent sequence)
  - and countable set, 59
  - decreasing, 69, 69, 76, 84
  - and denumerable set, 58
  - eventually constant, 189, 422
  - eventually in a set, 425
  - frequently in a set, 433
  - increasing, 69, 69, 76, 84, 186
  - index origin 0, 35
  - index origin1, 35
  - sequence
    - recursive (*see* recursion)
    - strictly decreasing, 69
    - strictly increasing, 69
- sequence of sets
  - decreasing, 69
  - increasing, 69
- sequential compactness, 496, 497
- sequential convergence (*see* convergent sequence)
- sequentially compact, 497
- sequentially compact space, 505, 508
  - and countably compact space, 505
- series, 187–188
  - and base expansion, 80
- set, 7, 13, 7–13
  - and list of elements, 8
  - and set-builder notation, 8
  - see also* class, proper class
- set-builder notation, 8, 13
- Sierpinski carpet, 490
- Sierpinski space, 228, 234
- $\sigma$ -compact space, 504, 520
- simple closed curve, 323, 323, 591
- simply connected space, 587
  - and nullhomotopic loop, 588
  - and conservative field, 587
  - and contractible space, 589
  - and line integral, 587
  - and retract, 602
- sine curve (*see* topologist's sine curve)
- singleton, 9
- simple connectedness (*see* simply connected space)
- slice, 477, 487
- small inductive dimension, 640
  - and zero-dimensional space, 640
- smaller topology (*see* coarser topology)
- smallest element (*see* least element)
- smash product, 418
- Smirnov space (*see* deleted sequence space)
- solid torus, 357, 412
- Sorgenfrey line, 239, 279, 287, 484, 628
  - and convergence, 426
  - base, 270
  - and completely regular space, 627
  - as Lindelöf space, 289, 290
  - local base, 264
  - and metrizable space, 297
  - as normal space, 295, 297
  - as regular space, 295
  - reverse, 352
  - see also* lower-limit topology, upper-limit topology
- Sorgenfrey plane, 274
  - and separation properties, 300
  - as completely regular space, 621
  - and Lindelöf space, 289
  - and normal space, 296, 622
  - and regular space, 296
  - and separable space, 379
- space (*see* topological space, metric space)
- space-filling curve, 568, 571
- special linear group, 384
- special orthogonal group, 384

sphere, [147](#), [147–415](#)  
     as boundary, [248](#), [249](#)  
     as connected space, [527](#)  
     as path-connected space, [564](#)  
     with whiskers, [354](#)  
     *see also*  $n$ -sphere  
 spherical coordinates, [356](#)  
 square-summable sequence, [137](#), [138](#)  
 standard map from  $\mathbb{R} \rightarrow S_1$ , [91](#), [100](#)  
 starlike set, [543](#), [570](#)  
 start of a path, [559](#)  
 statement (*see* predicate, sentence)  
 stereographic projection, [329](#), [342](#)  
 strict ordering (*see* strict directed ordering; strict partial ordering; strict preordering; strict partial ordering)  
 strict preordering, [64](#)  
 strictly decreasing map, [68](#)  
 strictly increasing map, [68](#)  
 stronger topology (*see* finer topology)  
 subbase, [283](#), [283–285](#)  
     in compact space, [606](#)  
     and cover, [283](#)  
     of product space, [363–364](#)  
     *see also* base  
 subbasic open set  
     of product space, [364](#)  
 subcover, [42](#), [42](#)  
 subnet, [435](#), [435–437](#), [455](#), [456](#)  
     and clustering net, [436](#)  
     and subsequence, [435–436](#)  
     Kelley, [455](#), [456](#)  
     *see also* subsequence  
 subscript, [34](#)  
 subsequence, [434](#), [435](#)  
     and subnet, [435–436](#)  
     *see also* subnet  
 subset, [10](#)  
     proper, [11](#)  
 subspace, [230](#), [230–233](#)  
     and open sets, [155](#)  
     and closed set, [232](#)  
     of compact space, [463](#)  
     and completely regular space, [615](#)  
     of Euclidean space, [231](#)  
     of Hausdorff space, [243](#)  
     of locally compact space, [511](#), [519](#)  
     and neighborhood, [236](#)  
     of normal space, [617–618](#)  
     of ordered space, [272](#), [286](#)  
     and product space, [366–367](#)

    of regular space, [297](#)  
     and second-countable space, [278](#)  
     of separable space, [282](#), [379](#)  
     and separation properties, [615](#)  
     of  $T_0$ -space, [297](#)  
     of  $T_1$ -space, [297](#)  
     of  $T_2$ -space, [297](#)  
 succeeds, [63](#)  
     immediately, [65](#)  
     strictly, [65](#)  
 successor, [63](#)  
     immediate, [65](#)  
     strict, [65](#)  
 sum  
     of real-valued functions, [372](#)  
 sum (*see* Cartesian sum)  
 sup metric, [136](#), [135–136](#), [141](#), [144](#), [166](#), [208](#)  
 sup norm, [141](#)  
     *see also* sup metric  
 superset, [10](#)  
 supremum, [71](#), [71–79](#)  
 surface, [343](#)  
 surjection, [28](#), [28–30](#)  
     and cardinality, [120](#)  
     and section, [45](#)  
 surjective map, [28](#)  
     *see also* surjection  
 suspension, [415](#)  
     and adjunction space, [415](#)  
 symmetric relation, [94](#)  
 symmetry  
     of a metric, [139](#)

## T

$T_0$ -space, [293](#), [298](#)  
     and product space, [377](#)  
     and subspace, [297](#)  
 $T_1$ -space, [293](#), [298](#)  
     and metric space, [158](#)  
     and product space, [377](#)  
     and singletons, [294](#)  
     and subspace, [297](#)  
 in  $T_1$ -space  
     and limit points, [244](#)  
 $T_2$ -space, [241](#), [241–293](#)  
     and diagonal, [276](#)  
     and metric space, [159–160](#)  
     and subspace, [297](#)  
     and  $T_1$ -space, [296](#), [298](#)  
 $T_2$ -space

- see also* Hausdorff space
- tail of a net, 441, 442
- tangent disk space, 246, 296, 316, 484, 628
  - and completely regular space, 628
  - and local bases, 285
  - and metrizable space, 288, 637
  - and nonnormal space, 628
  - as nonnormal space, 300, 628
  - and quotient space, 628
  - and second-countable space, 288
  - separation properties
    - and separation properties, 628
  - and separation properties, 246, 299–300, 628
- taxicab distance, 133
- taxicab metric, 133
  - and Euclidean metric, 176
- taxicab norm, 142, 143
- Teichmüller–Tukey Lemma, 125
- terminal point of a path, 559
- terminating base expansion, 80
- terminating expansion, 92
- ternary base, 188
- ternary expansion, 468, 626
  - mainref, 80
- Thomas plank, 266, 286, 629
  - and normal space, 299
  - as regular space, 295
  - and separation properties, 629
- Tietze Extension Theorem, 310, 623
- topological embedding (*see* embedding)
- topological equivalence, 174
- topological group, 384, 384–385, 419–420
  - and compact set, 492–493
  - and component, 558
  - and connected space, 545
  - and quotient group, 419–420
  - and quotient space, 419
- topological invariant, 640
- topological manifold, 343
  - see also* differentiable manifold, manifold, manifold-with-boundary  $n$ -manifold
- topological property, 333
  - hereditary, 243, 350
  - inherited by subspace, 243
  - preserved, 349
  - see also* compact space, connected space, Hausdorff space, locally compact space, metrizable space, etc.; topological invariant
- topological space, 226, 226–228
  - $n$ -dimensional, 640
  - associated with metric space, 227
  - and attaching, 406–408
  - and dense set, 254
  - Hausdorff, 241
  - homogeneous, 355
  - locally Euclidean, 336
  - metrizable, 228
  - zero-dimensional, 266
  - see also* compact space, connected space, pointed space, product space, quotient space, subspace, etc.
- topological transformation
  - see* homeomorphism, 320
- topological vector space, 381
- topologically equivalent metric spaces, 174, 174–176
- topologies
  - comparison of, 306
- topologist's comb, 553
- topologist's sine curve, 518, 561–563, 570
- topology, 226, 226–228
  - associated with a base, 271
  - coarser, 306
  - countable-complement, 234
  - discrete, 227
  - excluded point, 233
  - final, 383, 420
  - finer, 306
  - finite-complement, 228
  - generated by a cover, 284
  - included point, 233
  - indiscrete, 227
  - induced by a partial ordering, 272
  - induced by metric, 227
  - induced on subset, 230, 231
  - inherited, 230
  - initial, 383, 420
  - topology
    - larger (*see* coarser topology)
  - and local base, 265–266
  - and neighborhood, 238–240
  - of pointwise convergence, 431–432
  - of uniform convergence, 432, 432
  - of pointwise convergence, 276, 284, 284–607
  - of pointwise convergence, 290
  - right-interval, 239

topology  
     smaller (*see* coarser topology)  
 topology  
     stronger (*see* finer topology)  
 topology  
     weaker (*see* coarser topology)  
*see also* product topology, quotient topology, relative topology  
 torus, [329](#), [357](#), [398](#), [419](#)  
     as compact space, [480](#)  
     and cylinder, [412](#)  
     solid, [412](#)  
     as topological group, [419](#)  
     *see also*  $n$ -torus  
 total ordering, [73](#), [73](#)  
 total surface, [330](#)  
 totally bounded metric space, [494](#)  
     and second-countable space, [495](#)  
     *see also* totally bounded set  
 totally bounded set  
     *see also* totally bounded metric space  
 totally bounded subset, [499](#)  
 totally disconnected space, [549](#), [549](#)  
     and zero-dimensional space, [552](#), [556](#)  
 totally ordered set, [73](#), [272](#)  
     *see also* chain, total ordering  
 trace of a path, [559](#)  
 transcendental number, [62](#)  
 transfinite induction, [124](#)  
 transition map, [341](#)  
 transitive relation, [63](#), [64](#), [66](#), [67](#), [73](#), [75](#), [94](#), [423](#)  
 transitivity (*see* transitive relation)  
 translation, [142](#), [172](#), [173](#)  
 translation vector, [142](#)  
 translation-invariant metric, [142](#), [172](#)  
 trefoil knot, [348](#)  
 triangle inequality, [132](#), [139](#)  
 triangularity, [139](#)  
     *see also* triangle inequality  
 trichotomy  
     for cardinalities, [122](#)  
 trichotomy, [75](#)  
 trivial relation, [16](#)  
 tube, [477](#), [487](#)  
     *see also* Tube Lemma  
 Tube Lemma, [478](#)  
 Tukey's Lemma (*see* Teichmüller–Tukey Lemma)  
 Tychonoff cube, [607](#), [633](#)  
     as compact space, [607](#)

    and embedding, [633](#), [634](#)  
 Tychonoff plank, [297](#), [617](#)  
     chipped, [617](#)  
 Tychonoff Product Theorem (*see* Tychonoff Theorem)  
 Tychonoff Theorem, [477](#), [479](#), [479](#), [605](#), [605–611](#)  
     and Axiom of Choice, [610](#)  
     and finite intersection property, [610](#)  
     for finitely many spaces, [478](#)  
     for sequence of spaces, [500](#), [610](#)  
     for two spaces, [478](#), [610](#)  
     and ultrafilters, [607](#)  
     and ultranets, [612](#)  
     and Well-ordering Theorem, [611](#)  
 typical neighborhood, [263](#)

## U

ultrafilter, [449](#), [607](#), [607–611](#)  
     in compact space, [610](#)  
     and principal filter, [607](#)  
     and Tychonoff Theorem, [607](#)  
     and ultranet, [611](#)  
     *see also* ultranet  
 ultrametric, [145](#), [146](#), [164](#)  
 ultrametric inequality, [145](#)  
 ultrametric space, [145](#)  
 ultranet, [451](#), [611–612](#)  
     and compact space, [612](#)  
     and Tychonoff Theorem, [612](#)  
     and ultrafilter, [611](#)  
     *see also* ultrafilter  
 uncountable set, [57](#), [382](#)  
     and real numbers, [85](#), [86](#), [92](#)  
 unicoherent space, [543](#)  
 uniform convergence, [189](#), [432](#)  
     and pointwise convergence, [432](#)  
     of net, [432](#)  
 uniformly continuous map, [218](#), [218](#), [502–504](#)  
     on compact metric space, [502](#)  
 uniformly convergent sequence of functions, [210](#), [210–212](#), [625](#)  
     and continuous limit function, [210](#)  
 union  
     and closure, [256](#)  
     of a collection of sets, [41](#), [41–42](#)  
     of a family of sets, [38](#)  
     and image, [28](#), [40](#)  
     and interior, [252](#)  
     and inverse image, [40](#)

of two sets, [13](#), [13–14](#)  
 unit circle, [149](#)  
 unit interval, [9](#)  
 unit square, [323](#)  
 universal filter (*see* ultrafilter)  
 universal net (*see* ultranet)  
 upper bound, [71](#), [71](#)  
     least, [71](#)  
         *see also* supremum  
     and supremum, [72](#), [79](#)  
 upper semicontinuous map, [319](#), [487](#)  
 upper-limit topology, [352](#)  
 Urysohn Metrization Theorem, [486](#), [494](#),  
     [635](#)  
 Urysohn's Lemma, [522](#), [619](#), [619](#), [622](#),  
     [634](#)  
     for metrizable space, [198](#)  
 Urysohn-Menger dimension, [640](#)  
 usual ordering, [16–17](#)  
 usual topology  
     subset of Euclidean space, [227](#)

## V

value  
     of functional relation, [16](#)  
 vanish at infinity, [521](#)  
 vector, [37](#)  
 vector addition, [371](#)  
 vector product by scalar, [36](#), [36](#)  
 vector sum, [36](#), [36–37](#)

## W

weak order (*see* weak preordering)  
 weak preordering, [64](#)  
 weaker topology (*see* coarser topology)  
 wedge sum, [417](#)  
 well-defined map, [101](#)  
 well-ordered set, [109](#)  
     and order-complete set, [109](#)  
     *see also* well-ordering, Well-ordering  
         Theorem  
 well-ordering, [12](#), [109](#), [109](#)  
 Well-ordering Principle, [12](#), [12](#)  
     and Axiom of Induction, [12](#), [19](#)  
     *see also* Well-ordering Theorem  
 Well-ordering Theorem, [116](#), [116–117](#)  
     and Tychonoff Theorem, [611](#)

## Z

zero of a function, [21](#), [90](#)  
 zero-dimensional space, [266](#), [266–285](#)  
     and small inductive dimension, [640](#)  
     and totally disconnected space, [552](#),  
         [556](#)  
 Zorn's Lemma, [113](#), [113–116](#), [125](#)



## List of to-dos

converting qqal to aligned and inserting shortintertext here...	50
Add fig for strips & rectangles as base of plane?	283
Work on attaching an n-cell!	408
This table will be extended to other constructions and additional properties later on, in WHERE?	410
To rewrite open cover of $S$ in $X$ , given earlier def of open cover of $X$	459
Move problem on quot of Lindelof somewhere	504
Will material on homotopy move to separate chapter?	523
material on simply connected moved to next section	585
Which topology on $C(X,Y)$ ?	587
Need more basic homotopy problems!	587
Will I include proof of Jordan Curve Theorem using fundamental group?	592
Write fund'l group & covering spaces intro	643
Write section "The fundamental group"	643
Write section "Covering spaces"	643

**Number of problems:** 1064