Can "\parfillskip=0pt" shorten a short paragraph in plain TEX by two lines?

Udo Wermuth

Abstract

The decisions of TEX when it breaks a paragraph into lines are based on numerical calculations of badness values, line demerits, etc. With the help of the formulas that TEX implements, experts can decide questions about possible or impossible tasks.

The question in the title compares the number of lines that TEX produces for a given text, if it is typeset with plain TEX's default values, to the number that one gets with a single change applied to these defaults: `\parfillskip` is set to 0 pt. A problem of this type cannot be solved by the abovementioned formulas alone although they help to find an assumed example in the simplest case of a three-line paragraph. But this example doesn't respect `plain`'s default values, as they aren't captured in the formulas. Additionally, some assumptions about the TEX input are needed to show that the answer to the question in the title is "no" for Computer Modern Roman fonts of sizes 8 pt, 9 pt, 10 pt, and 12 pt.

1 Introduction

In my article about the parameter `\parfillskip` [7] a couple of texts that TEX breaks with the defaults of the plain format in one to three lines demonstrate how different values for `\parfillskip` change the positions of the line breaks. In experiment 4 `\parfillskip` is set to 0 pt and a text previously typeset by TEX in three lines needs only two. Therefore I name this input a *3/2 text*.

I asked myself, does a "normal text" exist that is typeset by TEX two lines shorter, without any warning or error, if the parameter `\parfillskip` is set to 0 pt and all other parameters except `\hsize` (that was set to the column width of *TUGboat*) keep the default values of plain TEX. Thus, my question was: Do 3/1 or 4/2 texts exist? Of course, one can also ask the general question independent of the number of lines: Is there any paragraph that becomes two lines shorter with `\parfillskip=0pt`? This question isn't answered in this article.

Admittedly the reduction of three lines to a single one seems to be very extreme; even 4/2 and 5/3 texts seem to be unlikely. Intuition tells us that it is not possible. Unfortunately, that gives us no arguments to convince a skeptic. To get five lines of a text that fits also in three lines we don't need to extend the width of the three lines by 167% but only by 133% plus a little bit for the fifth line. And we know that certain conditions increase the stretchability of glue. For example, after a period the stretchability is multiplied by 3; see [2, p. 76]. Moreover, in a second pass an additional factor of more than 5/4 for the stretchability is possible because of the higher tolerance; see [2, p. 96]. And, of course, the three lines might be compressed, i.e., they might be *tight* lines, in which the glue shrinks; see [2, p. 97]. So our intuition might not be the best adviser.

Normal texts. I don't accept every valid TEX input; it mustn't be too weird. For example, I require that all glue items stem from spaces or ties, not from `\hskip`, `\hglue`, `\leaders`, etc., outside of hboxes. Section 5 states additional common-sense assumptions. I admit that the word "normal" describes only a vague concept, to which the reader must agree.

The `\hsize`. `\hsize` can be set but not all values are accepted. According to [1, p. 26] the value of `\hsize` should allow the typesetting of justified paragraphs with 45–75 characters per line or 40–50 for two-column layouts. We assume that each line of the shorter paragraphs contains ≈ 50 characters including spaces. The justified lines of the longer paragraphs contain therefore ≈ 25, 33, 37, 40, ... characters if the paragraph has 3, 4, 5, 6, ... lines. One can assume that the last line is very short.

Contents. Section 2 shows that it is possible to get one line with `\parfillskip=0pt` although otherwise a three-line paragraph is typeset, if the constraint that except for `\hsize` all parameters must have the plain TEX default is lifted for one more parameter. So we cannot prove that a 3/1 text is impossible inside plain TEX with the computations of TEX alone; we must look at the used fonts too. Section 3 presents a short introduction to four fonts: 8 pt, 9 pt, 10 pt, and 12 pt Computer Modern Roman. A proof that no 3/1 text exists for these fonts is developed in section 4. Section 5 discusses ways to extend the arguments. They simplify the calculations in section 6 to prove with common-sense assumptions about normal texts that no 4/2, 5/3, 6/4, and 7/5 texts exist for the aforementioned fonts.

Two appendices show how TEX's calculation help to find one of the examples in section 2.

Notation. The `\hsize` value is abbreviated by h.

$W_{\mathrm{nw}}(\mathtt{T})$ is the natural width of an input text `T`. That is the dimension shown by TEX if `T` is placed in box register 0 as an hbox, and then its width is output by `\the\wd0`. The dimensions $S_0^{\circ}(\mathtt{T})$, $S_0^{+}(\mathtt{T})$ and $S_0^{-}(\mathtt{T})$ stand for the sum of the natural width, stretchability, and shrinkability of all glue in `T`. For

doi.org/10.47397/tb/43-3/tb135wermuth-shorten

example, the hbox with the text `T` must be spread by S_0^+(`T`) or $-S_0^-$(`T`) to get a badness of 100. `T` might be one character, for example, W_{nw}(`-`). And W_{nw}(`␣`) is the natural width of the interword glue.

Some `\fontdimen` parameters of the used font specify the three dimensions of the interword glue. The `\fontdimen` number ι is abbreviated f_ι.

The shrinkability as well as the stretchability of interword glue are determined not only by the `\fontdimen` parameters but also by the value of one of TeX's special integers: `\spacefactor`. The quotient `\spacefactor`/1000 is denoted in this article by the symbol σ.

2 Near 3/1 texts

In this section one more parameter of TeX is allowed to change: `\spaceskip`. If nonzero, this glue parameter replaces the usual interword glue built from the font parameters; see page 76 of [2]. The next examples handle paragraphs that have either three lines or just one line. Nevertheless, these paragraphs are not 3/1 texts as the constraints are weakened. One peculiarity of the examples in this section is that the used `\spaceskip` values define interword glue that can only either shrink or stretch.

Example 1: Description

Set the interword glue not via `\fontdimen` parameters but via a nonzero `\spaceskip` that has no stretchability. Show that under this condition an English text that TeX breaks into three lines with the default `\parfillskip` can be typeset in a single line if `\parfillskip=0pt`.

TeX input

```
\spaceskip=0.9\hsize minus 0.87\hsize
\noindent She is the\penalty66\
  granddaughter~of~John's~oldest friend.
```

TeX output

- `\parfillskip=0pt plus 1fil`:

She is the
granddaughter of John's oldest
friend.

- `\parfillskip=0pt`:

She is the granddaughter of John's oldest friend. ▯

(The rectangle in the gutter or margin at the end of the single line signals the end of the example.)

The ties and the penalty in the input are important to get the result; the `\penalty` must be between 58 and 66. But the most important setting is the large shrinkability! Of course, such a `\spaceskip` shouldn't be used for a text. Its *shrink ratio*, i.e., the quotient of the amount of shrinkability and the natural width of spaces, has the exceptionally high value of $0.87/0.9 = 0.96666\ldots$; I name this ratio α_-.

An analysis based on the length of lines gives a necessary condition for the shrink ratio to make the above construction possible. The shrink ratio must be larger than 0.5 to move from a line with width `\hsize` and the maximum of shrinkability to a line that has a natural width larger than 2`\hsize`. A criterion to prevent TeX from typesetting a three-line paragraph vs. a one-line paragraph if spaces can only shrink is therefore to have a shrink ratio of the interword glue ≤ 0.5.

Example 2: Description

Again, set the interword glue not via `\fontdimen` parameters but via a nonzero `\spaceskip`; this time use natural width and stretchability only. Show that a text that TeX breaks in three lines with the plain TeX default value of `\parfillskip` can now be typeset in a single line if `\parfillskip=0pt`.

TeX input

```
\spaceskip=5pt plus 87.3pt
\noindent Now we see\penalty-151\
  {\tt \string\parfillskip}'s top
  contribution\penalty13\ clearly.
```

TeX output

- `\parfillskip=0pt plus 1fil`:

Now we see
`\parfillskip`'s top contribution
clearly.

- `\parfillskip=0pt`:

Now we see `\parfillskip`'s top contribution clearly. ▯

The *stretch ratio*, α_+, in example 2 has the incredible high value of $87.3\,\mathrm{pt}/5\,\mathrm{pt} = 17.46$.

It is much more difficult to find such a text than for glue that can only shrink. Appendix A demonstrates how TeX's formulas for line demerits helped to find the criteria for the text used by example 2. The main points of the construction are: 1) the first line must be *very loose*, 2) it must break with a penalty of -151 or less, 3) the second line must be a "little bit" *loose* so that 4) the join of second and third line is *decent*. See [2, p. 97] for a description of TeX's *fitness classes*: very loose, loose, decent, and tight. (Section 2 of [5] explains them too and its section 3 shows how to look at TeX's line-breaking decisions.) Note, the positive penalty could be zero.

Appendix B proves that, in the case that glue can only stretch, $\alpha_+ \leq \sqrt{2} = 1.41421\ldots$ avoids the scenario that three lines are typeset by TeX with the default `\parfillskip` but only a single line if `\parfillskip=0pt`.

3 Interword glue from fonts

TeX selects the line breaks for the paragraph in a way that minimizes the so-called *total demerits*; see pages 97–98 of [2]. The formula to compute them involves a constant and three variable parameters: the *badness* of each constructed line, the condition under which the line break occurs expressed in a *penalty value*, and *additional demerits* that penalize certain unwanted effects between neighboring lines. All values are integers.

The constructions in section 2 show that the total demerits approach doesn't reflect the available interword glue from the default font in plain TeX, `cmr10`. Its contribution, i.e., the natural width of the spaces as well as the stretch- and shrinkability, must be taken into account for our problem.

Of course, the normal interword glue defined by some `\fontdimen` parameters of the Computer Modern Roman fonts can both shrink and stretch so it is quite a different situation from the examples in section 2. As mentioned above the shrinkability as well as the stretchability of spaces are not only determined by the glue but also by the `\spacefactor` in front of the glue. TeX has a table of codes, the `\sfcode` table, that specifies for every character of a font how it changes the `\spacefactor`. The default value is 1000 but it is lowered to 999 after uppercase letters and it is raised to 3000 after a period in plain TeX, for example. The shrink ratio is divided by $\sigma = \texttt{\char`\\spacefactor}/1000$ and the stretch ratio is multiplied by this value; see [2, p. 76 and the macro `\nonfrenchspacing` on p. 351]. Plain TeX uses only six values different from 0 for σ: 0.999, 1, 1.25, 1.5, 2, and 3.

Four `\fontdimen` parameters of the used font specify the three dimensions of the interword glue. Computer Modern Roman fonts in the sizes 12 pt, 10 pt, 9 pt, and 8 pt obey the following relationships [4, p. 12 and p. 37].

1. `\fontdimen2`, f_2, is the base natural width for the interword glue; it's the width used if $\sigma < 2$.
2. `\fontdimen3`, f_3, equals $f_2/2$. It specifies the stretchability.
3. `\fontdimen4`, f_4, is one third of f_2. Its value specifies the shrinkability.
4. `\fontdimen7`, f_7 is also $f_2/3$ and it is added to form the natural width with f_2 if $\sigma \geq 2$.

For this discussion the contribution to the natural width of interword glue in the case $\sigma \geq 2$, i.e., `\fontdimen7`, is seen as part of $W_{\mathrm{nw}}(\mathtt{T})$, that is, the natural width of the input text `T`. It can vanish at a line break but it doesn't add stretch- or shrinkability to the glue. Thus $W_{\mathrm{nw}}(\textvisiblespace)$ is always f_2.

Now it's possible to transfer the above rules into a formula for the shrink or stretch ratios.

$$\alpha_- = \frac{f_4}{f_2}\Big/\sigma = \frac{1}{3\sigma}. \tag{1}$$

In any case α_- is smaller than 0.5 in plain TeX. Its maximum is reached with $\sigma = 0.999$; then it is $1/2.997 \approx 0.333667$. Similarly,

$$\alpha_+ = \frac{f_3}{f_2} \times \sigma = \frac{\sigma}{2} \tag{2}$$

has its maximum value $1.5 > \sqrt{2}$ with $\sigma = 3$.

4 Do 3/1 texts exist?

Before the combination of shrink- and stretchability is studied further, we show that we need to consider only the first pass for 3/1 texts.

If the input `T` can be (a) typeset in one line and (b) split into three parts to fill more than two lines then the single line cannot be in the fitness class *very loose*. According to the description on page 97 of [2] `T` in a very loose line obeys:

$$W_{\mathrm{nw}}(\mathtt{T}) + S_0^+(\mathtt{T}) \leq h \tag{3}$$

where equality holds if and only if the badness of the line is 100. Obviously, $S_0^+(\mathtt{T}) < h$.

To get three lines we need a higher badness. But even for a line with badness 200 (the maximal tolerance for a second pass; see page 96 of [2]) in which the stretchability is increased by the factor $\sqrt[3]{2} \approx 1.25992$ it is not possible to stretch `T` wider than $2h$ if h, i.e., the `\hsize`, is not unreasonably short. Let's further assume that the first two lines end with hyphenated words so that the text `T` is *extended* by two inserted hyphens. (Note all four fonts have $W_{\mathrm{nw}}(\textvisiblespace) = W_{\mathrm{nw}}(\texttt{-})$; see [4, p. 37 and p. 143].)

$$\begin{aligned} &W_{\mathrm{nw}}(\mathtt{T}) + \sqrt[3]{2}\, S_0^+(\mathtt{T}) + 2\, W_{\mathrm{nw}}(\texttt{-}) \\ &\quad = W_{\mathrm{nw}}(\mathtt{T}) + S_0^+(\mathtt{T}) + (\sqrt[3]{2} - 1) S_0^+(\mathtt{T}) + 2\, W_{\mathrm{nw}}(\texttt{-})\,. \end{aligned}$$

Using (3), $\sqrt[3]{2} - 1 < 0.26$, $S_0^+(\mathtt{T}) < h$, and our assumption that lines have 50 characters — here amply translated as $2\, W_{\mathrm{nw}}(\texttt{-}) < 0.5h$ — we find

$$\begin{aligned} W_{\mathrm{nw}}(\mathtt{T}) + \sqrt[3]{2}\, S_0^+(\mathtt{T}) + 2\, W_{\mathrm{nw}}(\texttt{-}) &< h + 0.26h + 0.5h \\ &< 2h\,. \end{aligned}$$

We proved:

The fitness class of the single line in a 3/1 text cannot be very loose. (4)

Thus, the single line is typeset in the first pass. With such a solution the three-line paragraph must be typeset in the first pass too as TeX has no reason to execute a second pass. Note this result is also valid for the examples in section 2.

Are there 3/1 texts? Okay, `T` can fit into one line. We assume that all of its shrinkability must be used because then we find the `T` with the largest width:

$$W_{\mathrm{nw}}(\mathtt{T}) - S_0^-(\mathtt{T}) = h\,.$$

Let's assume that there are $\nu_{\mathtt{T}}$ spaces in `T`. These spaces might have different widths, stretchability and shrinkability as these values are influenced by the \spacefactor as explained above. We assume here that all spaces shrink with a single value:

$$W_{\mathrm{nw}}(\mathtt{T}) - \nu_{\mathtt{T}}\alpha_{-}\, W_{\mathrm{nw}}(\sqcup) = h\,. \tag{5}$$

To get a 3/1 text, the text T together with its maximal stretchability must be wider than $2h$. We cannot use more than the maximal stretchability as TEX typesets in the first pass according to (4) and underfull lines are excluded. Let's directly replace $S_0^+(\mathtt{T})$ by the term that involves the number of spaces and one value for the stretchability:

$$W_{\mathrm{nw}}(\mathtt{T}) + \nu_{\mathtt{T}}\alpha_{+}\, W_{\mathrm{nw}}(\sqcup) > 2h = h + h\,. \tag{6}$$

In the inequality (6) one h on the right hand side is replaced by the left hand side of (5)

$$W_{\mathrm{nw}}(\mathtt{T}) + \nu_{\mathtt{T}}\alpha_{+}\, W_{\mathrm{nw}}(\sqcup) > W_{\mathrm{nw}}(\mathtt{T}) - \nu_{\mathtt{T}}\alpha_{-}\, W_{\mathrm{nw}}(\sqcup) + h$$

and simplified to

$$\alpha_{+} + \alpha_{-} > \frac{h}{\nu_{\mathtt{T}}\, W_{\mathrm{nw}}(\sqcup)}\,. \tag{7}$$

Next α_+ and α_- are replaced by their expressions involving σ, i.e., by (1) and (2):

$$\frac{\sigma}{2} + \frac{1}{3\sigma} > \frac{h}{\nu_{\mathtt{T}}\, W_{\mathrm{nw}}(\sqcup)}\,.$$

As there are only six values for σ a table suffices to find the maximum for the left hand side (LHS).

$\sigma =$	0.999	1	1.25	1.5	2	3
LHS $\approx$	0.8332	0.8333	0.8917	0.9722	1.1667	1.6111

Thus, the maximum for the left hand side is reached with $\sigma = 3$. Using this maximum the above relation reads now $1.6111 > h/(\nu_{\mathtt{T}}\, W_{\mathrm{nw}}(\sqcup))$ or after moving $\nu_{\mathtt{T}} \geq 0$ to the left hand side:

$$\nu_{\mathtt{T}} > \frac{h}{1.6111\, W_{\mathrm{nw}}(\sqcup)}\,. \tag{8}$$

This inequality states that the number $\nu_{\mathtt{T}}$ of spaces in a text T must exceed the right hand side to get more than two lines using TEX with plain's default settings. Note, (8) doesn't state that text T is a 3/1 text if the inequality holds; it's only a *necessary condition*, i.e., if T is a 3/1 text then (8) holds.

The denominator of the right hand side (RHS) of (8) can easily be computed for the above listed Computer Modern Roman fonts.

font:	cmr8	cmr9	cmr10	cmr12
$W_{\mathrm{nw}}(\sqcup)\,/\mathrm{pt} =$	2.83337	3.08331	3.33333	3.91663
$\times\, 1.6111 =$	4.56484	4.96752	5.37033	6.31008

The only unknown value is the \hsize h.

Our general assumption about the \hsize is that a line holds ≈ 50 characters. Thus the number of spaces in such a line must be less than $50/2 = 25$.

Digression. For my article [7] $h = 225\,\mathrm{pt}$ and the font was cmr9. As $225/4.96752 \approx 45.3$ one needs 46 spaces and therefore 46 punctuation marks in front of them and one at the end of the line, plus 46 times additional white space of width f_7 as $\sigma \geq 2$, and then there should be some text too. This is impossible. (*End of digression.*)

The table on page 28 of [1] helps to find the line length that's related to a given number of characters for a given font. The table requires as input the lowercase alphabet length of this font.

font name:	cmr8	cmr9	cmr10	cmr12
lc alphabet length in pt:	108.4	118.0	127.6	149.9
requires according to [1]				
this line width: $h/\mathrm{pt} =$	192	216	240	264
to get this no. of chars:	49	50	52	51
making the RHS of (8) $\approx$	42.1	43.5	44.7	41.8

The numbers in the last line are much larger than the half of the corresponding value in the previous line! Thus, no 3/1 texts exist with the Computer Modern Roman fonts of sizes 12 pt, 10 pt, 9 pt, and 8 pt if the line width allows ≈ 50 characters.

5 Improvements

The computations in the section 4 applied only arguments about the length of a line when the interword spaces have to shrink or stretch. The computations are valid for all input if glue items are spaces. But one might ask if the computations should be limited to these arguments as "normal texts" are considered.

Of course, section 4 establishes a clear result so that the argument doesn't require an additional complication. On the other hand, what we learn in the smallest case might be useful in other cases.

Line breaks I. The single line shrinks so much that its badness is 100 and thus the total demerits that TEX computes for this paragraph becomes $(10+100)^2 = 110^2 = 12100$; see [2, p. 98].

In a three-line paragraph that contains two lines that stretch so that they each have badness 100, the total demerits receive from these badness values a contribution of at least $2\times110^2+100$; the 100 comes from the last line. So TEX never considers typesetting the three-line paragraph if there isn't another contribution that reduces the total demerits. Such a contribution can only come from a negative penalty at a line break; this effect was used in example 2.

Yes, plain.tex contains macros that use a negative penalty but except for one these macros operate in vertical mode, not in a paragraph. The exception is \break that uses \penalty-10000 to force a line break. (This macro cannot be used in a 3/1 text as otherwise the single line couldn't be typeset. In general this macro splits a paragraph into two parts that can be considered independently.)

I think the majority of texts don't use the primitive `\penalty` in horizontal mode directly. Usually it occurs only through the macro executed by the tilde to prevent a line break. Thus one can question that an input with an explicitly stated negative penalty creates a "normal text". So, let's assume that negative penalties aren't part of the input.

Thus, the badness of the two complete lines in a three-line paragraph must be smaller than 100 to make the paragraph's total demerits smaller than the total demerits of the one-line paragraph. With other words (6) shouldn't use α_+ but a smaller value. Let's compute the maximum factor $\phi < 1$ that has to be applied to α_+ in (6).

We assume that both lines have the same badness $\beta < 100$; then the following inequality must hold to make TeX typeset the three-line paragraph with the defaults.

$$(10+100)^2 > 2(10+\beta)^2 + 100$$
$$\iff \quad 0 > (10+\beta-\sqrt{6000})(10+\beta+\sqrt{6000})\,.$$

Simple arithmetic transformations and an application of the well-known equation $a^2-b^2=(a+b)(a-b)$ lead to the last inequality. It can only hold if the first factor is less than 0 as $\beta \geq 0$. Thus

$$\beta < \sqrt{6000} - 10 \approx 77.46 - 10 = 67.46$$

or $\beta \leq 67$, as badness is an integer. (It's possible to take one line with badness 67 and one with 68. But if one line has a much higher badness then the badness of the other must be reduced more drastically. For example, if one line has badness 87 then the other can only have badness 31 and their combined width is shorter than with nearly equal badness values.)

Page 97 of [2] explains the relationship between ϕ and the badness β as an equation but it's only an approximation: $\phi \approx \sqrt[3]{\beta/100}$. Thus, with $\beta = 67$ we get $\phi \approx 0.875$.

Using $0.875\alpha_+$ instead of α_+ in (7) the table for the LHS changes:

$\sigma =$	0.999	1	1.25	1.5	2	3
LHS $\approx$	0.7707	0.7708	0.8135	0.8785	1.0417	1.4236

so that 1.6111 gets replaced by 1.4236; that is, (8) becomes

$$\nu_{\mathtt{T}} > \frac{h}{1.4236\,W_{\mathrm{nw}}(\sqcup)}\,. \qquad (*)$$

The new value implies a change to the table for the fonts as the RHS of (∗) is larger than the RHS of (8).

Line breaks II. And one can go even further as the lines have breakpoints. A line break must occur either at glue (including kerns and the end of inline math, if followed by glue) or at a penalty item, for example, at an explicit hyphen or inside inline math. A break at a penalty increases the total demerits, a break at glue reduces the number of spaces. Let's assume the breaks are at glue. Then inequality (6) is changed to

$$W_{\mathrm{nw}}(\mathtt{T}) - 2f_7[\sigma \geq 2] + (\nu_{\mathtt{T}} - 2)\phi\alpha_+\,W_{\mathrm{nw}}(\sqcup) > 2h$$

where $[\sigma \geq 2] = 1$ if and only if the relation $\sigma \geq 2$ is true; otherwise it's 0. This leads to the following replacement of (8):

$$\nu_{\mathtt{T}} > \frac{h}{1.4236\,W_{\mathrm{nw}}(\sqcup)} + \frac{2f_7}{1.4236\,W_{\mathrm{nw}}(\sqcup)} + \frac{2\phi\alpha_+}{1.4236}$$

and as $f_7/\,W_{\mathrm{nw}}(\sqcup) = 1/3$, $\phi = 0.875$, and $\alpha_+ = 3/2$

$$\nu_{\mathtt{T}} > \frac{h}{1.4236\,W_{\mathrm{nw}}(\sqcup)} + 0.47 + 1.84$$

which increases the right hand side compared to (∗) by 2.31 spaces.

Distributions I. Another approach for improvement addresses the use of normal texts. The assumption that in such a text all spaces stretch and shrink by the same amount is of course not very likely. So instead of having 100% spaces with $\sigma = 3$ one can define a less extreme distribution.

I don't apply the distribution of punctuation marks in a large corpus as that might not reflect the special aspects of the short text that we are looking for. So let's use a reasonable but also somewhat extreme distribution. For example, 70% of the spaces have $\sigma = 1$ and are therefore not preceded by any punctuation mark; 10% occur with $\sigma = 1.25$ after a comma; 5% with $\sigma = 1.5$ after a semicolon; 5% with $\sigma = 2$ after a colon; and 10% with $\sigma = 3$ after a period, an exclamation or question mark.

If the left hand side of (7) is split in this way one gets a new value instead of 1.6111. Using the values of the table for the LHS of (7) we compute

$$0.7\cdot 0.8333 + 0.1\cdot 0.8917 + 0.05\cdot 0.9722$$
$$+\,0.05\cdot 1.1667 + 0.1\cdot 1.6111 = 0.940535 \qquad (**)$$

which increases the right hand side of (8) when used instead of 1.6111.

Distributions II. Another questionable assumption for a normal text is that between two printed characters a space is output. A text with spaces that are always preceded by a punctuation mark can therefore not contain any letter between the spaces.

Again without citing data from a corpus we assume — somewhat extremely — that the output has no more than 25% spaces. Thus, a line with 50 characters contains at most 12 spaces or 13 words or word parts if hyphenation is used in the line.

6 The general case

Let's extend the analysis to 4/2, 5/3, etc., texts; in general, to a $\mu/\mu-2$ text for $\mu > 3$.

Both paragraphs might be typeset in the second pass. If one hyphen is inserted the shorter paragraph can be typeset with maximally tight lines all ending in a hyphen adding $(\mu-4)$`\doublehyphendemerits` and `\finalhyphendemerits` to the total demerits. With breaks at glue and lines that stretch $> 50\%$ in the longer paragraph, its badness values must obey

$$\beta < \sqrt{\frac{\mu-3}{\mu-1}14600 + \frac{\mu-4}{\mu-1}10000 + \frac{17000}{\mu-1}} - 10\,;$$

see section 5, "Line breaks I". Thus, $0.97 < \phi < 1.1 < \sqrt[3]{2}$ if $4 \le \mu \le 11$; at best a tiny gain for α_+.

Therefore we can assume to get the best candidates for small μ if the longer paragraph needs the first pass, the shorter the second pass with an inserted hyphen, and all other line breaks in both paragraphs are at explicit hyphens to keep the glue.

The formulas (5) and (6) are easily changed

$$W_{\rm nw}(\mathtt{T}) + W_{\rm nw}(\mathtt{-}) - \nu_{\mathtt{T}}\alpha_-\, W_{\rm nw}(\textvisiblespace) = (\mu-2)h$$
$$W_{\rm nw}(\mathtt{T}) + \nu_{\mathtt{T}}\alpha_+\, W_{\rm nw}(\textvisiblespace) > (\mu-1)h$$

and lead to an inequality like (8); note $W_{\rm nw}(\mathtt{-}) = W_{\rm nw}(\textvisiblespace)$. Now apply the distributions of the previous section; that is, with $(**)$ (8) is replaced by

$$\nu_{\mathtt{T}} > \frac{h}{0.940535\, W_{\rm nw}(\textvisiblespace)} + \frac{1}{0.940535}\,. \qquad (8')$$

The denominator of the RHS' first term computes to

font:	cmr8	cmr9	cmr10	cmr12
$W_{\rm nw}(\textvisiblespace)\,/\mathrm{pt} =$	2.83337	3.08331	3.33333	3.91663
$\times\, 0.940535 =$	2.66488	2.89996	3.13511	3.68373

which gives the following values for the RHS if used with h suggested in [1]; see the end of section 4.

font:	cmr8	cmr9	cmr10	cmr12
RHS of $(8') \approx$	73.1	75.6	77.6	72.7

Results: a) As all RHS are > 50 no 4/2 texts exist. b) The RHS for `cmr9` and `cmr10` are so large that no 5/3 text exists for them. c) There are no 5/3, 6/4, and 7/5 texts for the four fonts if 25% of all characters are spaces. The shorter paragraph needs at least 6 lines as the text has 292–312 characters.

Appendix A: Conditions for example 2

We need to work with the calculation that TeX performs to find line breaks. Thus, we need to state several formulas and do some math. These tasks become much simpler if we have a consistent notation. We develop it bit by bit in this appendix.

The input line number ι is named $\mathtt{L}_\iota$ and its line demerits are represented by Λ_ι. This value is computed from other numbers; see [2, p. 98]:

$$\Lambda_\iota = (\lambda + \beta_\iota)^2 + \mathrm{sgn}(\pi_\iota)\pi_\iota^2 + \delta_\iota \qquad \text{(A1)}$$

where λ is the `\linepenalty` (a constant in the output paragraph), β_ι is the badness assigned by TeX to this line, $\mathrm{sgn}(\pi_\iota)$ is the signum function that returns the sign of its argument, π_ι is the penalty that occurs at the line break, and δ_ι is the sum of the parameters `\adjdemerits`, `\doublehyphendemerits`, and `\finalhyphendemerits` that TeX assigns based on a comparison of lines ι and $\iota-1$; line ι gets these demerits. Note two special cases: A last line cannot get `\doublehyphendemerits` and only a very loose first line has $\delta_1 > 0$ because of `\adjdemerits`; see [6].

The badness β is an approximation based on how much the glue of the line must use either its stretchability or shrinkability to make a line of a given length. It is a function of two dimensions but they are usually dropped in the notation.

$$\beta(u,a) \approx 100\left(\frac{u}{a}\right)^3, \quad a > 0\,\mathrm{pt} \qquad \text{(A2)}$$

where u is the amount of either stretchability or shrinkability that was used in the construction of the line and a is the available amount of stretchability or shrinkability, resp., in this line; see [2, p. 97].

Badness is a function that is monotone increasing in u and monotone decreasing in a; see §108 of [3]. This means for $x, y \ge 0\,\mathrm{pt}$:

$$\beta(u,a) \le \beta(u+x,a) \wedge \beta(u,a) \ge \beta(u,a+y)\,.$$

Thus, if material is added to a line, in which the glue can only stretch, its badness is reduced.

The function $(\lambda + \beta(u,a))^2$ has the same property if $\lambda > 0$; in plain TeX $\lambda = 10$. So a difference built for two lines only from their first term in (A1) is positive if the line whose badness is subtracted contains more material. And it also means that such a difference is larger than the difference of the same lines both extended by more identical material.

One line. First we show that a negative penalty must be involved if a text that TeX can typeset in one line is split and *the glue can only stretch*. By (A1) and as $\lambda = 10$ in plain TeX

$$\Lambda_1 = \lambda^2 = 100$$

as the badness must be 0, the final break has no penalty contribution, and no additional demerits are applied to a single line if `\parfillskip` has TeX's default value. For a pair of lines we have by (A1)

$$\begin{aligned}\Lambda_1 + \Lambda_2 &= (\lambda+\beta_1)^2 + \mathrm{sgn}(\pi_1)\pi_1^2 + \delta_1 \\ &\quad + (\lambda+\beta_2)^2 + \mathrm{sgn}(\pi_2)\pi_2^2 + \delta_2 \\ &> \lambda^2 + \mathrm{sgn}(\pi_1)\pi_1^2 + \lambda^2\end{aligned}$$

as both badness values are ≥ 0, both additional demerits are ≥ 0, and $\pi_2 = 0$. Thus π_1 must be less than 0 to make the left hand side $\le \lambda^2$. A similar argument shows that with two line breaks at least one must have a negative penalty. Thus, both lines in such a paragraph don't break with a hyphen, i.e., no `\doublehyphendemerits` are involved.

To distinguish between a last line that is produced with the default \parfillskip and one that is output with \parfillskip=0pt, a prime is attached to the variables in the second case. Primed non-last line variables have identical values to their unprimed version.

With \parfillskip=0pt we have the single line as the best solution so that any two-line paragraph must have higher demerits:

$$\Lambda'_{1/1} < \Lambda'_{1/2} + \Lambda'_{2/2}\,. \tag{A3}$$

The subscript to identify line demerits and the associated parameters is written here as a pair of the line number and the total lines in a set of line breaks.

Three lines. TeX typesets three lines only if the sum of the demerits for these three lines is smaller than the demerits for the one-line solution as well as for any two-line solution.

$$\Lambda_{[1]/3} + \Lambda_{[2]/3} + \Lambda_{[3]/3} < \Lambda_{1/1} \tag{A4}$$

$$\Lambda_{[1]/3} + \Lambda_{[2]/3} + \Lambda_{[3]/3} < \Lambda_{1/2} + \Lambda_{2/2} \tag{A5}$$

Here the line numbers on the left hand sides are set in brackets to mark them as fixed. If two or three lines from these fixed lines are joined together then we write, for example, [2&3]/2 or [1&2&3]/1.

The single line is unique, of course, and therefore (A4) is better written as

$$\Lambda_{[1]/3} + \Lambda_{[2]/3} + \Lambda_{[3]/3} < \Lambda_{[1\&2\&3]/1}\,.$$

Construction: Step 1. Because of (A5) we know

$$\Lambda_{[1]/3} + \Lambda_{[2]/3} + \Lambda_{[3]/3} < \Lambda_{[1\&2]/2} + \Lambda_{[3]/2}$$

where $\Lambda_{[1\&2]/2}$ stands for the line demerits of the line $\mathtt{L}_{[1\&2]/2}$ that consists of the input for the first two lines, $\mathtt{L}_{[1]/3}$ and $\mathtt{L}_{[2]/3}$, of the three-line solution as explained above. Note: $\mathtt{L}_{[1]/3}$ must end in a negative penalty as shown in subsection "One line".

$\Lambda_{[3]/3}$ and $\Lambda_{[3]/2}$ are identical except if the line $\mathtt{L}_{[2]/3}$ is very loose; then the first term has the additional demerits \adjdemerits. In order to remember this situation better we write $\delta_a[\mathtt{L}_{[2]/3}\,\mathrm{v}]$ instead of $\delta_{[3]/3}$. Here δ_a is the value of \adjdemerits and the bracket has the value 1 if the stated condition is true, otherwise it's 0. Thus instead of

$$\Lambda_{[1]/3} + \Lambda_{[2]/3} + \delta_{[3]/3} < \Lambda_{[1\&2]/2}$$

we write

$$\Lambda_{[1]/3} + \Lambda_{[2]/3} + \delta_a[\mathtt{L}_{[2]/3}\,\mathrm{v}] < \Lambda_{[1\&2]/2}\,.$$

Next we apply (A1) and it follows that

$$\begin{aligned} & \Lambda_{[1]/3} < \Lambda_{[1\&2]/2} - \Lambda_{[2]/3} - \delta_a[\mathtt{L}_{[2]/3}\,\mathrm{v}] \\ \iff & \Lambda_{[1]/3} < (\lambda + \beta_{[1\&2]/2})^2 + \mathrm{sgn}(\pi_{[1\&2]/2})\pi^2_{[1\&2]/2} \\ & \qquad + \delta_{[1\&2]/2} \\ & \qquad - (\lambda + \beta_{[2]/3})^2 - \mathrm{sgn}(\pi_{[2]/3})\pi^2_{[2]/3} \\ & \qquad - \delta_{[2]/3} - \delta_a[\mathtt{L}_{[2]/3}\,\mathrm{v}]\,. \end{aligned}$$

We can simplify this inequality as $\pi_{[1\&2]/2} = \pi_{[2]/3}$ (it is the same line break) and $\delta_{[1\&2]/2} = 0$ (because in the first pass a very loose line that gets more text isn't very loose any more so \adjdemerits can't be charged). As \doublehyphendemerits can't occur, $\delta_{[2]/3}$ can only be \adjdemerits if either the first line is decent and the second very loose or the first line is very loose and the second decent. That is, $\delta_{[2]/3} = \delta_a[\mathtt{L}_{[1]/3}\,\mathrm{v}][\mathtt{L}_{[2]/3}\,\mathrm{d}] + \delta_a[\mathtt{L}_{[2]/3}\,\mathrm{v}][\mathtt{L}_{[1]/3}\,\mathrm{d}]$ or short: $\delta_{[2]/3} = \delta_a[\mathtt{L}_{[1]/3}/\mathtt{L}_{[2]/3}\,\mathrm{d/v}]$. Therefore:

$$\begin{aligned} \Lambda_{[1]/3} < {} & (\lambda + \beta_{[1\&2]/2})^2 - (\lambda + \beta_{[2]/3})^2 \\ & - \delta_a[\mathtt{L}_{[1]/3}/\mathtt{L}_{[2]/3}\,\mathrm{d/v}] - \delta_a[\mathtt{L}_{[2]/3}\,\mathrm{v}]\,. \end{aligned} \tag{A6}$$

Construction: Step 2. All we know for sure is that the text can be broken at two places. The third line of the three-line solution is short and so lines two and three must be joined for \parfillskip=0pt to create a two-line solution. Thus, by (A3)

$$\Lambda'_{[1\&2\&3]/1} < \Lambda'_{[1]/2} + \Lambda'_{[2\&3]/2} \tag{A7}$$

must be true. For the left hand side we have

$$\Lambda'_{[1\&2\&3]/1} = (\lambda + \beta'_{[1\&2\&3]/1})^2 \tag{A8}$$

by (A1) because a last line has no countable penalty at its end and only \adjdemerits could be charged for a first line. But as a join of more than two lines, the line isn't very loose and $\delta'_{[1\&2\&3]/1} = 0$.

Of course,

$$\begin{aligned} \Lambda'_{[1]/2} &= \Lambda_{[1]/3} \quad \text{as well as} \\ \beta_{[1\&2]} &= \beta'_{[1\&2]} \quad \text{and} \quad \beta_{[2]/3} = \beta'_{[2]/3} \end{aligned} \tag{A9}$$

as in both cases the same break is used and all lines stretch to h with the same badness. By (A1)

$$\Lambda'_{[2\&3]/2} = (\lambda + \beta'_{[2\&3]/2})^2 + \delta'_{[2\&3]/2}$$

as there is no penalty for the break. Next, $\delta'_{[2\&3]/2}$ can only be \adjdemerits as shown above. So if $\delta'_{[2\&3]/2} \neq 0$ then the first line must be very loose and this line decent. Thus with

$$\Lambda'_{[2\&3]/2} = (\lambda + \beta'_{[2\&3]/2})^2 + \delta_a[\mathtt{L}_{[1]/3}\,\mathrm{v}][\mathtt{L}_{[2\&3]/2}\,\mathrm{d}]$$

and together with (A9), (A6), and (A8) we compute

$$\begin{aligned} & \Lambda'_{[1]/2} + \Lambda'_{[2\&3]/2} - \Lambda'_{[1\&2\&3]/1} \\ & \quad < (\lambda + \beta'_{[1\&2]/2})^2 - (\lambda + \beta'_{[2]/3})^2 \\ & \qquad - \delta_a[\mathtt{L}_{[1]/3}/\mathtt{L}_{[2]/3}\,\mathrm{d/v}] - \delta_a[\mathtt{L}_{[2]/3}\,\mathrm{v}] \\ & \qquad + (\lambda + \beta'_{[2\&3]/2})^2 + \delta_a[\mathtt{L}_{[1]/3}\,\mathrm{v}][\mathtt{L}_{[2\&3]/2}\,\mathrm{d}] \\ & \qquad - (\lambda + \beta'_{[1\&2\&3]/1})^2 \\ & \quad < \delta_a[\mathtt{L}_{[1]/3}\,\mathrm{v}][\mathtt{L}_{[2\&3]/2}\,\mathrm{d}] - \delta_a[\mathtt{L}_{[1]/3}/\mathtt{L}_{[2]/3}\,\mathrm{d/v}] \\ & \qquad - \delta_a[\mathtt{L}_{[2]/3}\,\mathrm{v}] =: \Delta\,. \end{aligned}$$

(Note, Δ is a short-cut for the sum of the three δ_a.) We discussed that badness is a monotone function after (A2) and that this property can be extended to the first term of (A1). Here

$$(\lambda + \beta'_{[2]/3})^2 - (\lambda + \beta'_{[2\&3]/2})^2 \ge (\lambda + \beta'_{[1\&2]/2})^2 - (\lambda + \beta'_{[1\&2\&3]/1})^2$$

holds and thus a negative value was dropped above.

Construction: Final step. If $\Delta \le 0$ then

$$\Lambda'_{[1]/2} + \Lambda'_{[2\&3]/2} - \Lambda'_{[1\&2\&3]/1} < 0$$
$$\iff \Lambda'_{[1]/2} + \Lambda'_{[2\&3]/2} < \Lambda'_{[1\&2\&3]/1}$$

and this is a contradiction to (A7): There is a two-line solution in the case `\parfillskip=0pt` that has less total demerits than the single line which is output by TeX by assumption.

A solution to the problem might only be found with $\Delta > 0$. This can only happen if the first line is very loose, the second combined with the third decent, and the second line itself loose. We know that $\pi_{[1]/3} < 0$. It's best to have it ≤ -151 as

$$\begin{aligned} 151^2 &> \text{line demerits of a very loose first line} \\ &\quad + \text{minimal line demerits of a loose line} \\ &= \big((10+100)^2 + 10000\big) + (10+13)^2 \\ &= 22629 > 150^2. \end{aligned}$$

So all elements of a solution are found. *Q.E.D.*

Appendix B: A bound for the stretch ratio

The badness values of a successfully broken paragraph must not be greater than the tolerance for the pass. The first pass sets the tolerance to 100, the second and third use 200 [2, p. 96].

In the first pass a very loose line, `L`, needs all of its stretchability to fill the line. With a badness larger than 100 it needs more; see (A2).

$$\begin{aligned} W_{\mathrm{nw}}(\mathtt{L}) + S_0^+(\mathtt{L}) &= h \text{ if badness is 100;} \\ W_{\mathrm{nw}}(\mathtt{L}) + \sqrt[3]{2}S_0^+(\mathtt{L}) &= h \text{ if badness is 200.} \end{aligned} \tag{B1}$$

A decent line can either stretch or shrink but it uses at most half of the related capacity to fill the line; see [2, p. 97]. That is, if the line `L` is decent and its glue stretches, then

$$W_{\mathrm{nw}}(\mathtt{L}) + \frac{1}{2}S_0^+(\mathtt{L}) \ge h \ge W_{\mathrm{nw}}(\mathtt{L})\ . \tag{B2}$$

The bound. Let α_+ be the max. stretch ratio. The first line, $\mathtt{L}_1$, must be very loose, so it needs all of its stretchability to reach h. Note: the text is typeset in the first pass as shown in section 4. Thus by (B1)

$$\begin{aligned} & W_{\mathrm{nw}}(\mathtt{L}_1) + S_0^+(\mathtt{L}_1) = h \\ \Longrightarrow\quad & W_{\mathrm{nw}}(\mathtt{L}_1) + \alpha_+ S_0^\circ(\mathtt{L}_1) \ge h \\ \Longrightarrow\quad & (1+\alpha_+)\, W_{\mathrm{nw}}(\mathtt{L}_1) \ge h \\ \Longrightarrow\quad & W_{\mathrm{nw}}(\mathtt{L}_1) \ge \frac{h}{1+\alpha_+}\ . \end{aligned}$$

Here we use $S_0^\circ(\mathtt{L}_1) \le W_{\mathrm{nw}}(\mathtt{L}_1)$.

The line $\mathtt{L}_{2\&3}$ is decent and its glue stretches. This means by (B2)

$$\begin{aligned} & W_{\mathrm{nw}}(\mathtt{L}_{2\&3}) + \frac{1}{2}S_0^+(\mathtt{L}_{2\&3}) \ge h \\ \Longrightarrow\quad & W_{\mathrm{nw}}(\mathtt{L}_{2\&3}) + \frac{\alpha_+}{2}S_0^\circ(\mathtt{L}_{2\&3}) \ge h \\ \Longrightarrow\quad & (1+\frac{\alpha_+}{2})\, W_{\mathrm{nw}}(\mathtt{L}_{2\&3}) \ge h \\ \Longrightarrow\quad & W_{\mathrm{nw}}(\mathtt{L}_{2\&3}) \ge \frac{h}{1+\alpha_+/2}\ . \end{aligned}$$

If the single line gets longer than h then the above construction fails. Thus we must have

$$\begin{aligned} h \ge W_{\mathrm{nw}}(\mathtt{L}_{1\&2\&3}) &\ge W_{\mathrm{nw}}(\mathtt{L}_1) + W_{\mathrm{nw}}(\mathtt{L}_{2\&3}) \\ &\ge \frac{h}{1+\alpha_+} + \frac{h}{1+\alpha_+/2}\ . \end{aligned}$$

Therefore, $1/(1+\alpha_+) + 1/(1+\alpha_+/2)$ must be at most 1. This means

$$\begin{aligned} & \frac{1}{1+\alpha_+} + \frac{1}{1+\alpha_+/2} = 1 \\ \iff\quad & 1 + \frac{\alpha_+}{2} + 1 + \alpha_+ = 1 + \frac{3}{2}\alpha_+ + \frac{1}{2}\alpha_+^2 \\ \iff\quad & 2 = \alpha_+^2 \end{aligned}$$

so that $\alpha_+ = \sqrt{2} \approx 1.41$ is the maximal allowed factor for the stretchability compared to the natural width. *Q.E.D.*

References

[1] Robert Bringhurst, *The Elements of Typographic Style*, 4th edition, version 4.2, Seattle, Washington: Hartley & Marks, 2016.

[2] Donald E. Knuth, *The TeXbook*, Volume A of *Computers & Typesetting*, Boston, Massachusetts: Addison-Wesley, 1984.

[3] Donald E. Knuth, *TeX: The Program*, Volume B of *Computers & Typesetting*, Boston, Massachusetts: Addison-Wesley, 1986.

[4] Donald E. Knuth, *Computer Modern Typefaces*, Volume E of *Computers & Typesetting*, Boston, Massachusetts: Addison-Wesley, 1986.

[5] Udo Wermuth, "Tracing paragraphs", *TUGboat* **37**:3 (2016), 358–373; errata in *TUGboat* **38**:3 (2017), 414. `tug.org/TUGboat/tb37-3/tb117wermuth.pdf`

[6] Udo Wermuth, "TeX's 'additional demerits' parameters", *TUGboat* **39**:1 (2018), 81–87. `tug.org/TUGboat/tb39-1/tb121wermuth-adem.pdf`

[7] Udo Wermuth, "Experiments with `\parfillskip`", *TUGboat* **39**:3 (2018), 276–303. `tug.org/TUGboat/tb39-3/tb123wermuth-parfillskip.pdf`

◇ Udo Wermuth
Dietzenbach, Germany
`u dot wermuth (at) icloud dot com`