Towards More Homogeneous Paragraphs
∼ DocEng 2025 ∼

Didier Verna

EPITA / LRE

didier@lrde.epita.fr

didierverna.net   @didierverna   didier.verna   in/didierverna
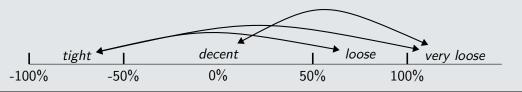
Homogeneity (or lack thereof)

In olden times when wishing still helped one, there lived a king whose daughters were all beautiful; and the youngest was so beautiful that the sun itself, which has seen so much, was astonished whenever it shone in her face. Close by the king's castle lay a great dark forest, and under an old lime-tree in the forest was a well, and when the day was very warm, the king's child went out into the forest and sat down by the side of the cool fountain; and when she was bored she took a golden ball, and threw it up on high and caught it; and this ball was her favorite plaything.

▶ T$_{\text{E}}$X's rendition
  ▶ Line width: 201pt
  ▶ Font: Latin Modern Roman 10pt
▶ Interword blanks scaling
  ▶ Primary tool for justification
  ▶ Typographic color quality depends on it
  ▶ Lines inherently homogeneous
  ▶ Paragraph, not!
🤔 Heterogeneity is surprising
  ▶ The Knuth-Plass is *not* greedy
  ▶ It is equipped to fight against it

Fitness Classes and Adjacent Demerits

- Interword space $s = w_n$ (normal), $+w_{st}$ (stretch), or $-w_{sh}$ (shrink)
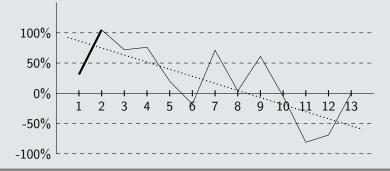- Adjustment ratio $r \in [-1, +\infty[$: used amount of stretching (> 0) or shrinking (< 0)
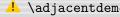
Fitness Classes (*w.r.t.* r)

tight		*decent*	*loose*	*very loose*
-100%	-50%	0%	50%	100%

- \adjacentdemerits $(= 10000)$
- Applied when two consecutive lines are more than one fitness class apart

The Problem

Consecutive lines adjustment ratios

⚠️ `\adjacentdemerits` are applied only once (lines $1 - 2$)

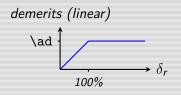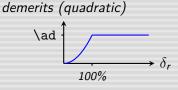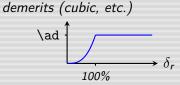▶ Discrepancies ($|r_i - r_{i-1}|$) of less than 100% / 150% are ignored

Gradual Adjacent Demerits

💡 **Idea** (simple !)

▶ Be proportional to the discrepancy (instead of discrete)

▶ Do *not* go above the value of \adjacentdemerits

▶ **Rationale:**
  ▶ Explore alternative layouts instead of ignoring them
  ▶ Use a continuous measure for a continuous typographical trait
  ▶ Do not disturb the algorithm too much, because it "works well in practice"

demerits (linear)

demerits (quadratic)

demerits (cubic, etc.)

Linear Case Definition(s)

In Theory

- $d_{ij} = \min\left(\texttt{\textbackslash ad}, \texttt{\textbackslash ad} \cdot |r_i - r_j|\right)$
- 🚫 Incompatible with the KP's optimization

In Practice

- $c_i = floor(10 r_i + 1/2)$
- $d_{ij} = \min\left(\texttt{\textbackslash ad}, \texttt{\textbackslash ad} \cdot \frac{|c_i - c_j|}{10}\right)$

tight　　　　*decent*　　　　*loose*　　*very loose*

-10　　　-5　　　0　　　5　　　10

September 4, 2025

Demo

In olden times when wishing still helped one, there lived a king whose daughters were all beautiful; and the youngest was so beautiful that the sun itself, which has seen so much, was astonished whenever it shone in her face. Close by the king's castle lay a great dark forest, and under an old lime-tree in the forest was a well, and when the day was very warm, the king's child went out into the forest and sat down by the side of the cool fountain; and when she was bored she took a golden ball, and threw it up on high and caught it; and this ball was her favorite plaything.

No gradation

In olden times when wishing still helped one, there lived a king whose daughters were all beautiful; and the youngest was so beautiful that the sun itself, which has seen so much, was astonished whenever it shone in her face. Close by the king's castle lay a great dark forest, and under an old lime-tree in the forest was a well, and when the day was very warm, the king's child went out into the forest and sat down by the side of the cool fountain; and when she was bored she took a golden ball, and threw it up on high and caught it; and this ball was her favorite plaything.

Linear Gradation

In olden times when wishing still helped one, there lived a king whose daughters were all beautiful; and the youngest was so beautiful that the sun itself, which has seen so much, was astonished whenever it shone in her face. Close by the king's castle lay a great dark forest, and under an old lime-tree in the forest was a well, and when the day was very warm, the king's child went out into the forest and sat down by the side of the cool fountain; and when she was bored she took a golden ball, and threw it up on high and caught it; and this ball was her favorite plaything.

Quadratic Gradation

September 4, 2025

Validation

▶ Grimm Brothers "Frog King" §1 @ 379 $\neq$ widths

▶ 1279 § from Moby Dick @ 284pt ($\approx$ 10cm)

▶ 1658 experiments in total

KP Disruption

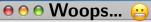|	
All algorithms agree	45%
Quadratic > linear	60%
Gradual $\in$ KP's top 10%	70%

Adjacency

KP average A_{rms}	0.5
Gradual average A_{rms}	0.46
Gradual A_{rms}< KP A_{rms}	87%

$$A_{rms} = \sqrt{\frac{\sum_{i=1}^{n-1}\left(r_i - r_{i+1}\right)^2}{n-1}}$$

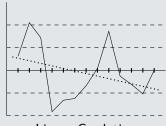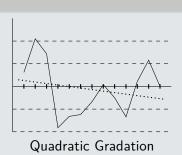💯 Success! But wait…

September 4, 2025

Woops… 😬

Adjustment Ratios

No Gradation   Linear Gradation   Quadratic Gradation
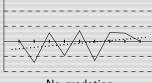
- 💥 Our A_{rms} are worse!
- ▶ Further hypotheses ($\implies$ further analyses, *cf.* paper)
  - ▶ Global trend (linear regression)
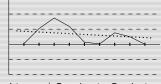 - ▶ Oscillation (amplitude, frequency, *etc.*)

Counter Counter-Example (307pt)

In olden times when wishing still helped one, there lived a king whose daughters were all beautiful; and the youngest was so beautiful that the sun itself, which has seen so much, was astonished whenever it shone in her face. Close by the king's castle lay a great dark forest, and under an old lime-tree in the forest was a well, and when the day was very warm, the king's child went out into the forest and sat down by the side of the cool fountain; and when she was bored she took a golden ball, and threw it up on high and caught it; and this ball was her favorite plaything.

In olden times when wishing still helped one, there lived a king whose daughters were all beautiful; and the youngest was so beautiful that the sun itself, which has seen so much, was astonished whenever it shone in her face. Close by the king's castle lay a great dark forest, and under an old lime-tree in the forest was a well, and when the day was very warm, the king's child went out into the forest and sat down by the side of the cool fountain; and when she was bored she took a golden ball, and threw it up on high and caught it; and this ball was her favorite plaything.

No gradation

Linear / Quadratic Gradation

▶ Better A_{rms}

▶ Linear regression: flatter (slope), better quality (R^2)

▶ Oscillation: fewer peaks, lesser amplitude

September 4, 2025

Conclusion

- Gradual / clamped adjacent demerits
  - Easily done in production engines
  - No change to the user-level API
  - Unimpactful performance-wise
- Remaining questions
  - Pertinence of the "global trend" concept?
  - Correlation(s) between oscillation frequency, amplitude, and linearity?
  - What is homogeneity (*cf.* cognitive science)?

Trait	Impact
Disruption	✔
A_{rms} (adjacency)	✔
Peaks (oscillation)	✔
Linearity (R^2)	(✔)
Global Trends (slope)	✗

Last Minute

- LuaMetaTEX now has "granular fitness classes"
- *Cf.* TUGboat, Volume 46 (2025), No. 1, p. 114