we have $f \circ \sigma \neq f$. If $M(n,k)$ is the number of these and $F(n,k)$ is the number of orbits of mappings $f$ under the action of $C_n$, then evidently

$$F(n,k) = \sum_{d|n} M(d,k) \ (n \geq 1) \tag{1}$$

But since, clearly,

$$\sum_{d|n} dM(d,k) = k^n \ (n \geq 1) \tag{2}$$

we find from (2),

$$M(n,k) = (1/n) \sum_{d|n} \mu(n/d) k^d$$

# Mathematical Writing

*Donald E. Knuth*
*Tracy Larrabee*
*Paul M. Roberts*

MAA Notes Number 14 The Mathematical Association of America

# Mathematical Writing

Donald E. Knuth, Tracy Larrabee,
and Paul M. Roberts

*Published and Distributed by*
The Mathematical Association of America

*Library of Congress Control Number 89-062390*
ISBN 978-0-88385-063-3
*Printed in the United States of America*
Current Printing (last digit):
10 9 8 7 6 5 4 3

47. Exemplary Programs in Introductory College Mathematics: Innovative Programs Using Technology, *Susan Lenker*, Editor.
48. Writing in the Teaching and Learning of Mathematics, *John Meier and Thomas Rishel*.
49. Assessment Practices in Undergraduate Mathematics, *Bonnie Gold*, Editor.
50. Revolutions in Differential Equations: Exploring ODEs with Modern Technology, *Michael J. Kallaher*, Editor.
51. Using History to Teach Mathematics: An International Perspective, *Victor J. Katz*, Editor.
52. Teaching Statistics: Resources for Undergraduate Instructors, *Thomas L. Moore*, Editor.
53. Geometry at Work: Papers in Applied Geometry, *Catherine A. Gorini*, Editor.
54. Teaching First: A Guide for New Mathematicians, *Thomas W. Rishel*.
55. Cooperative Learning in Undergraduate Mathematics: Issues That Matter and Strategies That Work, *Elizabeth C. Rogers, Barbara E. Reynolds, Neil A. Davidson, and Anthony D. Thomas*, Editors.
56. Changing Calculus: A Report on Evaluation Efforts and National Impact from 1988 to 1998, *Susan L. Ganter*.
57. Learning to Teach and Teaching to Learn Mathematics: Resources for Professional Development, *Matthew Delong and Dale Winter*.
58. Fractals, Graphics, and Mathematics Education, Benoit Mandelbrot and Michael Frame, Editors.
59. Linear Algebra Gems: Assets for Undergraduate Mathematics, *David Carlson, Charles R. Johnson, David C. Lay, and A. Duane Porter*, Editors.
60. Innovations in Teaching Abstract Algebra, *Allen C. Hibbard and Ellen J. Maycock*, Editors.
61. Changing Core Mathematics, *Chris Arney and Donald Small*, Editors.
62. Achieving Quantitative Literacy: An Urgent Challenge for Higher Education, *Lynn Arthur Steen*.
63. Women Who Love Mathematics: A Sourcebook of Significant Writings, *Miriam Cooney*, Editor.
64. Leading the Mathematical Sciences Department: A Resource for Chairs, *Tina H. Straley, Marcia P. Sward, and Jon W. Scott*, Editors.
65. Innovations in Teaching Statistics, *Joan B. Garfield*, Editor.
66. Mathematics in Service to the Community: Concepts and models for service-learning in the mathematical sciences, *Charles R. Hadlock*, Editor.
67. Innovative Approaches to Undergraduate Mathematics Courses Beyond Calculus, *Richard J. Maher*, Editor.
68. From Calculus to Computers: Using the last 200 years of mathematics history in the classroom, Amy Shell-Gellasch and Dick Jardine, Editors.
69. A Fresh Start for Collegiate Mathematics: Rethinking the Courses below Calculus, *Nancy Baxter Hastings*, Editor.
70. Current Practices in Quantitative Literacy, *Rick Gillman*, Editor.
71. War Stories from Applied Math: Undergraduate Consultancy Projects, *Robert Fraga*, Editor.

MAA Service Center
P.O. Box 91112
Washington, DC 20090-1112
1-800-331-1MAA FAX: 1-301-206-9789

## Table of Contents


§1. Minicourse on technical writing . . . . . . . . . . . . . . . . . . . 1
§2. An exercise on technical writing . . . . . . . . . . . . . . . . . . 7
§3. An answer to the exercise . . . . . . . . . . . . . . . . . . . . . 8
§4. Comments on student answers (1) . . . . . . . . . . . . . . . . . 9
§5. Comments on student answers (2) . . . . . . . . . . . . . . . . . 11
§6. Preparing books for publication (1) . . . . . . . . . . . . . . . . 14
§7. Preparing books for publication (2) . . . . . . . . . . . . . . . . 15
§8. Preparing books for publication (3) . . . . . . . . . . . . . . . . 18
§9. Handy reference books . . . . . . . . . . . . . . . . . . . . . . 19
§10. Presenting algorithms . . . . . . . . . . . . . . . . . . . . . . 20
§11. Literate Programming (1) . . . . . . . . . . . . . . . . . . . . 22
§12. Literate Programming (2) . . . . . . . . . . . . . . . . . . . . 26
§13. User manuals . . . . . . . . . . . . . . . . . . . . . . . . . . 28
§14. Galley proofs . . . . . . . . . . . . . . . . . . . . . . . . . . 30
§15. Refereeing (1) . . . . . . . . . . . . . . . . . . . . . . . . . 31
§16. Refereeing (2) . . . . . . . . . . . . . . . . . . . . . . . . . 34
§17. Hints for Referees . . . . . . . . . . . . . . . . . . . . . . . 36
§18. Illustrations (1) . . . . . . . . . . . . . . . . . . . . . . . . 37
§19. Illustrations (2) . . . . . . . . . . . . . . . . . . . . . . . . 40
§20. Homework: Subscripts and superscripts . . . . . . . . . . . . . . 40
§21. Homework: Solutions . . . . . . . . . . . . . . . . . . . . . . 43
§22. Quotations . . . . . . . . . . . . . . . . . . . . . . . . . . . 47
§23. Scientific American Saga (1) . . . . . . . . . . . . . . . . . . 49
§24. Scientific American Saga (2) . . . . . . . . . . . . . . . . . . 51
§25. Examples of good style . . . . . . . . . . . . . . . . . . . . . 54
§26. Mary-Claire van Leunen on 'hopefully' . . . . . . . . . . . . . . 57
§27. Herb Wilf on Mathematical Writing . . . . . . . . . . . . . . . . 59
§28. Wilf's first extreme . . . . . . . . . . . . . . . . . . . . . . 61
§29. Wilf's other extreme . . . . . . . . . . . . . . . . . . . . . . 62
§30. Jeff Ullman on Getting Rich . . . . . . . . . . . . . . . . . . . 66
§31. Leslie Lamport on Writing Papers . . . . . . . . . . . . . . . . 69
§32. Lamport's handout on unnecessary prose . . . . . . . . . . . . . 71
§33. Lamport's handout on styles of proof . . . . . . . . . . . . . . 72
§34. Nils Nilsson on Art and Writing . . . . . . . . . . . . . . . . . 73
§35. Mary-Claire van Leunen on Calisthenics (1) . . . . . . . . . . . 77
§36. Mary-Claire's handout on Composition Exercises . . . . . . . . . 81
§37. Comments on student work . . . . . . . . . . . . . . . . . . . . 89
§38. Mary-Claire van Leunen on Which vs. That . . . . . . . . . . . . 93
§39. Mary-Claire van Leunen on Calisthenics (2) . . . . . . . . . . . 98
§40. Computer aids to writing . . . . . . . . . . . . . . . . . . . . 100
§41. Rosalie Stemer on Copy Editing . . . . . . . . . . . . . . . . . 102
§42. Paul Halmos on Mathematical Writing . . . . . . . . . . . . . . 106
§43. Final truths . . . . . . . . . . . . . . . . . . . . . . . . . . 112

## §1. Notes on Technical Writing

Stanford's library card catalog refers to more than 100 books about technical writing, including such titles as *The Art of Technical Writing*, *The Craft of Technical Writing*, *The Teaching of Technical Writing*. There is even a journal devoted to the subject, the *IEEE Transactions on Professional Communication*, published since 1958. The American Chemical Society, the American Institute of Physics, the American Mathematical Society, and the Mathematical Association of America have each published "manuals of style." The last of these, *Writing Mathematics Well* by Leonard Gillman, is one of the required texts for CS 209.

The nicest little reference for a quick tutorial is *The Elements of Style*, by Strunk and White (Macmillan, 1979). Everybody should read this 85-page book, which tells about English prose writing in general. But it isn't a required text—it's merely recommended.

The other required text for CS 209 is *A Handbook for Scholars* by Mary-Claire van Leunen (Knopf, 1978). This well-written book is a real pleasure to read, in spite of its unexciting title. It tells about footnotes, references, quotations, and such things, done correctly instead of the old-fashioned "op. cit." way.

Mathematical writing has certain peculiar problems that have rarely been discussed in the literature. Gillman's book refers to the three previous classics in the field: An article by Harley Flanders, *Amer. Math. Monthly*, 1971, pp. 1–10; another by R. P. Boas in the same journal, 1981, pp. 727–731. There's also a nice booklet called *How to Write Mathematics*, published by the American Mathematical Society in 1973, especially the delightful essay by Paul R. Halmos on pp. 19–48.

The following points are especially important, in your instructor's view:

1. Symbols in different formulas must be separated by words.

   Bad: Consider $S_q$, $q < p$.

   Good: Consider $S_q$, where $q < p$.

2. Don't start a sentence with a symbol.

   Bad: $x^n - a$ has $n$ distinct zeroes.

   Good: The polynomial $x^n - a$ has $n$ distinct zeroes.

3. Don't use the symbols $\therefore$, $\Rightarrow$, $\forall$, $\exists$, $\ni$; replace them by the corresponding words. (Except in works on logic, of course.)

4. The statement just preceding a theorem, algorithm, etc., should be a complete sentence or should end with a colon.

   Bad: We now have the following
   **Theorem**. $H(x)$ is continuous.

   This is bad on three counts, including rule 2. It should be rewritten, for example, like this:

   Good: We can now prove the following result.
   **Theorem.** The function $H(x)$ defined in (5) is continuous.

   Even better would be to replace the first sentence by a more suggestive motivation, tying the theorem up with the previous discussion.

5. The statement of a theorem should usually be self-contained, not depending on the assumptions in the preceding text. (See the restatement of the theorem in point 4.)
6. The word "we" is often useful to avoid passive voice; the "good" first sentence of example 4 is much better than "The following result can now be proved." But this use of "we" should be used in contexts where it means "you and me together", *not* a formal equivalent of "I". Think of a dialog between author and reader.

   In most technical writing, "I" should be avoided, unless the author's persona is relevant.
7. There is a definite rhythm in sentences. Read what you have written, and change the wording if it does not flow smoothly. For example, in the text *Sorting and Searching* it was sometimes better to say "merge patterns" and sometimes better to say "merging patterns". There are many ways to say "therefore", but often only one has the correct rhythm.
8. Don't omit "that" when it helps the reader to parse the sentence.

   Bad: Assume $A$ is a group.

   Good: Assume that $A$ is a group.

   The words "assume" and "suppose" should usually be followed by "that" unless another "that" appears nearby. But *never* say "We have that $x = y$," say "We have $x = y$." And avoid unnecessary padding "because of the fact that" unless you feel that the reader needs a moment to recuperate from a concentrated sequence of ideas.
9. Vary the sentence structure and the choice of words, to avoid monotony. But use parallelism when parallel concepts are being discussed. For example (Strunk and White #15), don't say this:

   > Formerly, science was taught by the textbook method, while now the laboratory method is employed.

   Rather:

   > Formerly, science was taught by the textbook method; now it is taught by the laboratory method.

   Avoid words like "this" or "also" in consecutive sentences; such words, as well as unusual or polysyllabic utterances, tend to stick in a reader's mind longer than other words, and good style will keep "sticky" words spaced well apart. (For example, I'd better not say "utterances" any more in the rest of these notes.)
10. Don't use the style of homework papers, in which a sequence of formulas is merely listed. Tie the concepts together with a running commentary.
11. Try to state things twice, in complementary ways, especially when giving a definition. This reinforces the reader's understanding. (Examples, see §2 below: $N^n$ is defined twice, $A_n$ is described as "nonincreasing", $L(C, P)$ is characterized as the smallest subset of a certain type.) All variables must be defined, at least informally, when they are first introduced.

12. Motivate the reader for what follows. In the example of §2, Lemma 1 is motivated by the fact that its converse is true. Definition 1 is motivated only by decree; this is somewhat riskier.

    Perhaps the most important principle of good writing is to keep the reader uppermost in mind: What does the reader know so far? What does the reader expect next and why?

    When describing the work of other people it is sometimes safe to provide motivation by simply stating that it is "interesting" or "remarkable"; but it is best to let the results speak for themselves or to give *reasons* why the things seem interesting or remarkable.

    When describing your own work, be humble and don't use superlatives of praise, either explicitly or implicitly, even if you are enthusiastic.

13. Many readers will skim over formulas on their first reading of your exposition. Therefore, your sentences should flow smoothly when all but the simplest formulas are replaced by "blah" or some other grunting noise.

14. Don't use the same notation for two different things. Conversely, use consistent notation for the same thing when it appears in several places. For example, don't say "$A_j$ for $1 \le j \le n$" in one place and "$A_k$ for $1 \le k \le n$" in another place unless there is a good reason. It is often useful to choose names for indices so that $i$ varies from 1 to $m$ and $j$ from 1 to $n$, say, and to stick to consistent usage. Typographic conventions (like lowercase letters for elements of sets and uppercase for sets) are also useful.

15. Don't get carried away by subscripts, especially when dealing with a set that doesn't need to be indexed; set element notation can be used to avoid subscripted subscripts. For example, it is often troublesome to start out with a definition like "Let $X = \{x_1, \ldots, x_n\}$" if you're going to need subsets of $X$, since the subset will have to defined as $\{x_{i_1}, \ldots, x_{i_m}\}$, say. Also you'll need to be speaking of elements $x_i$ and $x_j$ all the time. Don't name the elements of $X$ unless necessary. Then you can refer to elements $x$ and $y$ of $X$ in your subsequent discussion, without needing subscripts; or you can refer to $x_1$ and $x_2$ as specified elements of $X$.

16. Display important formulas on a line by themselves. If you need to refer to some of these formulas from remote parts of the text, give reference numbers to all of the most important ones, even if they aren't referenced.

17. Sentences should be readable from left to right without ambiguity. Bad examples: "Smith remarked in a paper about the scarcity of data." "In the theory of rings, groups and other algebraic structures are treated."

18. Small numbers should be spelled out when used as adjectives, but not when used as names (i.e., when talking about numbers as numbers).

    Bad: The method requires 2 passes.

    Good: Method 2 is illustrated in Fig. 1; it requires 17 passes. The count was increased by 2. The leftmost 2 in the sequence was changed to a 1.

19. Capitalize names like Theorem 1, Lemma 2, Algorithm 3, Method 4.

20. Some handy maxims:

Watch out for prepositions that sentences end with.
When dangling, consider your participles.
About them sentence fragments.
Make each pronoun agree with their antecedent.
Don't use commas, which aren't necessary.
Try to never split infinitives.

21. Some words frequently misspelled by computer scientists:

| | | |
|---|---|---|
| implement | not | impliment |
| complement | not | compliment |
| occurrence | not | occurence |
| dependent | not | dependant |
| auxiliary | not | auxillary |
| feasible | not | feasable |
| preceding | not | preceeding |
| referring | not | refering |
| category | not | catagory |
| consistent | not | consistant |
| PL/I | not | PL/1 |
| descendant (noun) | not | descendent |
| its (belonging to it) | not | it's (it is) |

The following words are no longer being hyphenated in current literature:

nonnegative
nonzero

22. Don't say "which" when "that" sounds better. The general rule nowadays is to use "which" only when it is preceded by a comma or by a preposition, or when it is used interrogatively. Experiment to find out which is better, "which" or "that", and you'll understand this rule.

Bad: Don't use commas which aren't necessary.
Good: Don't use commas that aren't necessary.

Another common error is to say "less" when it should be "fewer".

23. In the example at the bottom of §2 below, note that the text preceding displayed equations (1) and (2) does not use any special punctuation. Many people would have written

... of "nonincreasing" vectors:

$$A_n = \{(a_1, \ldots, a_n) \in N^n \mid a_1 \geq \cdots \geq a_n\}\,. \qquad (1)$$

If $C$ and $P$ are subsets of $N^n$, let:

$$L(C, P) = \ldots$$

and those colons are wrong.

Better: "According to the 'fundamental theorem of arithmetic' (proved in ex. 1.2.4–21), each positive integer $u$ can be expressed in the form

$$u = 2^{u_2}3^{u_3}5^{u_5}7^{u_7}11^{u_{11}}\ldots = \prod_{p\ \text{prime}} p^{u_p},$$

where the exponents $u_2, u_3, \ldots$ are uniquely determined nonnegative integers, and where all but a finite number of the exponents are zero."

[The first quotation is from Carl Linderholm's neat satirical book *Mathematics Made Difficult*; the second is from D. Knuth's *Seminumerical Algorithms*, Section 4.5.2.]

27. When in doubt, read *The Art of Computer Programming* for outstanding examples of good style.

[That was a joke. Humor is best used in technical writing when readers can understand the joke only when they also understand a technical point that is being made. Here is another example from Linderholm:

> "... $\emptyset D = \emptyset$ and $N\emptyset = N$, which we may express by saying that $\emptyset$ is absorbing on the left and neutral on the right, like British toilet paper."

Try to restrict yourself to jokes that will not seem silly on second or third reading. And don't overuse exclamation points!]

## §3. An Answer

Here is one way to complete the exercise in the previous section. (But please try to WORK IT YOURSELF BEFORE READING THIS.) Note that a few clauses have been inserted to help keep the reader synchronized with the current goals and subgoals and strategies of the proof. Furthermore the notation $(b_1, \ldots, b_n)$ is used instead of $(p_1, \ldots, p_n)$, in the second paragraph below, to avoid confusion with formula (2).

*Proof.* Assume that $L(C, P) \subseteq A_n$. Since $C$ is always contained in $L(C, P)$, we must have $C \subseteq A_n$; therefore only the condition $P \subseteq A_n$ needs to be verified.

If $P$ is not contained in $A_n$, there must be a vector $(b_1, \ldots, b_n) \in P$ such that $b_i < b_j$ for some $i < j$. We want to show that this leads to a contradiction.

Since the set $C$ is nonempty, it contains some element $(c_1, \ldots, c_n)$. We know that the components of this vector satisfy $c_1 \geq \cdots \geq c_n$, because $C \subseteq A_n$.

Now $(c_1, \ldots, c_n) + k(b_1, \ldots, b_n)$ is an element of $L(C, P)$ for all $k \geq 0$, and by hypothesis it must therefore be an element of $A_n$. But if we take $k = c_i - c_j + 1$, we have $k \geq 1$ and

$$c_i + kb_i \geq c_j + kb_j,$$

hence

$$c_i - c_j \geq k(b_i - b_j). \tag{3}$$

This is impossible, since $c_i - c_j = k - 1$ is less than $k$, yet $b_j - b_i \geq 1$. It follows that $(b_1, \ldots, b_n)$ must be an element of $A_n$. ∎

Note that the hypothesis $C \neq \emptyset$ is necessary in Lemma 1, for if $C$ is empty the set $L(C, P)$ is also empty regardless of $P$.

[This was the "minor slip."]

BUT ... don't always use the first idea you think of. The proof above actually commits another sin against mathematical exposition, namely the unnecessary use of proof by contradiction. It would have been better to use a direct proof:

Let $(b_1, \ldots, b_n)$ be an arbitrary element of $P$, and let $i$ and $j$ be fixed subscripts with $i < j$; we wish to prove that $b_i \geq b_j$. Since $C$ is nonempty, it contains some element $(c_1, \ldots, c_n)$. Now the vector $(c_1, \ldots, c_n) + k(b_1, \ldots, b_n)$ is an element of $L(C, P)$ for all $k \geq 0$, and by hypothesis it must therefore be an element of $A_n$. But this means that $c_i + kb_i \geq c_j + kb_j$, i.e.,

$$c_i - c_j \geq k(b_j - b_i), \tag{3}$$

for arbitrarily large $k$. Consequently $b_j - b_i$ must be zero or negative.

We have proved that $b_j - b_i \leq 0$ for all $i < j$, so the vector $(b_1, \ldots, b_n)$ must be an element of $A_n$. ∎

This form of the proof has other virtues too: It doesn't assume that the $b_i$'s are integer-valued, and it doesn't require stating that $c_1 \geq \cdots \geq c_n$.

## §4. Excerpts from class, October 7 [notes by TLL]

Our first serious business involved examining "the worst abusers of the 'Don't use symbols in titles' rule." Professor Knuth (hereafter known as Knuth) displayed a paper by Gauss that had a long displayed formula in the title. He showed us a bibliography he's preparing that references not only that paper but another with even more symbols in the title. (Such titles make more than bibliographies difficult; they make bibliographic data retrieval systems and keyword-in-context produce all sorts of hiccups.)

In his bibliography Knuth has tried to keep his citations true to the original sources. The bibliography contains mathematical formulas, full name spellings (even alternate spellings when common), and completely spelled-out source journal names. (This last may be unusual enough that some members of a field may be surprised to see the full journal name written out, but it's a big help to novices who want to find it in the library.)

We spent the rest of class going over some of solutions that students had turned in for the exercise of §2 (each sample anonymous). He cautioned us that while he was generally pleased by the assignments, he was going to be pointing out things that could be improved. The following points were all made in the process of going through these samples.

> In certain instances, people did not understand what constitutes a proof. Fluency in mathematics is important for Computer Science students but will not be taught in this class.
>
> Not all formulas are equations. Depending on the formula, the terms 'relation', 'definition', 'statement', or 'theorem' might be used.
>
> Computer Scientists must be careful to distinguish between mathematical notation and programming language notation. While it may be appropriate to use $p[r]$ in a program, in a formal paper it is probably better to use $p$ with a subscript of $r$. As another example, it is not appropriate to use a star ($*$) to denote multiplication in a paper about mathematics. Just say $xy$.
>
> Some people called $p$ an element of $P$ and $p_r$ an element of $p$. Everything was an "element." It's better to call $p_r$ a "component" of $p$, thus distinguishing two kinds of subsidiary relationships.
>
> It is natural in mathematics to hold off some aspects of your definition — to "place action before definition" (as in '$p(x) < p(y)$ for some $x < y$'). But it is possible to carry this too far, if too much is being held back. The best location for certain definitions is a subjective matter.
>
> Remember to put words between adjacent formulas.
>
> When you use ellipses, such as $(P_1, \ldots, P_n)$, remember to put commas before and after the three dots. When placing ellipses between commas the three dots belong on the same level as the commas, but when the ellipsis is bracketed by symbols such as '+' or '<' the dots should be at mid-level.

> Be careful with the spacing around ellipses. The character string '. . .;' looks strange (it should have more space after the last dot). All kinds of accidents happen concerning spaces in formulas. Typesetting systems like TEX have built-in rules to cover 99% of the cases, but if you write a lot of mathematics you will get bitten.

> Linebreaks in the middle of formulas are undesirable. There are ways to enforce this with TEX (as well as other text formatting systems). People who use TEX and wish to use the vertical bar and the empty set symbol in notation like '$\{c \mid c \in \emptyset\}$' should be aware of the TEX commands `\mid` and `\emptyset`.

> Comments such as, "We demonstrate the second conclusion by contradiction," and "There must be a witness to the unsortedness of P," are useful because they tell the reader what is going on or bring in new and helpful vocabulary.

> Numbering all displayed formulas is usually a bad idea; number the important ones only. Extraneous parentheses can also be distracting. For example, in the phrase "let $k$ be $(c_i - c_j) + 1$," the parentheses should omitted.

> You can overdo the use of any good tool. For instance, you could overuse typographic tools by having 20 different fonts in one paper.

Two more topics were touched on (and are sure to be discussed further): the use of 'I' in technical writing, and the use of past or present tense in technical writing.

Knuth says that Mary-Claire van Leunen defends the use of 'I' in scholarly articles, but that he disagrees (unless the identity of the author is important to the reader). Knuth likes the "teamwork" aspect of using 'we' to represent the author and reader together. If there are multiple authors, they can either "revel in the ambiguity" of continuing to use 'we', or they can use added disambiguating text. If one author needs to be mentioned separately, the text can say 'one of the authors (DEK)', or 'the first author', but not 'the senior author'.

Knuth (hereafter known as Don) recommends that one of two approaches be used with respect to tenses of verbs: Either use present tense throughout the entire paper, or write sequentially. Sequential writing means that you say things like, "We saw this before. We will see this later." The sequential approach is more appropriate for lengthy papers. You can use it even more effectively by using words of duration: "We observed this long ago. We saw the other thing recently. We will prove something else soon."

## §5. Excerpts from class, October 9 [notes by TLL]

"I'm thinking about running a contest for the best Pascal program that is also a sonnet," was the one of the first sentences out of Don's mouth on the topic of the exact definition of "Mathematical Writing." He admitted that such a contest was "probably not the right topic for this course." However, a program (presumably even an iambic pentameter program) is among the documents that he will accept as the course term paper. He will accept articles for professional journals, chapters of books or theses, term papers for other courses, computer programs, user manuals or parts thereof: anything that falls into a definite genre where you have a specific audience in mind and the technical aspect is significant.

We spent the rest of class continuing to examine the homework assignment. In the interest of succinct notes, I have replaced many literal phrases by their generic equivalents. For example, I might have replaced '$A > B$' by '⟨relation⟩'. This time I have divided the comments into two sets: those dealing with what I will call "form" (parentheses, capitalization, fonts, etc.) and those dealing with "content" (wording, sentence construction, tense, etc.).

First, the comments concerning form:

> Don't overdo the use of colons. While the colon in 'Define it as follows:' is fine, the one in 'We have: ⟨formula⟩' should be omitted since the formula just completes the sentence. Some papers had more colons than periods.
>
> Should the first word after a colon be capitalized? Yes, if the phrase following the colon is a full sentence; No, if it is a sentence fragment. (This is not "yet" a standard rule, but Don has been trying it for several years and he likes it.)
>
> While too many commas will interfere with the smooth flow of a sentence, too few can make a sentence difficult to read. As examples, a sentence beginning with 'Therefore, ' does not need the comma following 'therefore'. But 'Observe that if ⟨symbol⟩ is ⟨formula⟩ then so is ⟨symbol⟩ because ⟨reasoning⟩' at least needs a comma before 'because'.
>
> Putting too many things in parentheses is a stylistic thing that can get very tiring. (When Don moves from his original, handwritten draft to a typed, computer-stored version his most frequent change is to remove extra parentheses.)
>
> Among the parentheses most in need of removal are nested parentheses. To this end, it is better to write '(Definition 2)' than '(definition (2))'. Unfortunately, however, you can't use the former if the definition was given in displayed formula (2). Then it's probably best to think of a way to avoid the outer parentheses altogether.
>
> In some cases your audience may expect nested parentheses. In this case (or in any other case when you feel you must have them), should the outer pair be changed to brackets (or curly-braces)? This was once the prevailing convention, but it is now not only obsolete but potentially dangerous; brackets and curly braces have semantic content for many scientific professionals. ("The world is short of delimiters," says

Don't get hung up on one or two styles of sentences. The following sort of thing can become very monotonous:

Thus, – – – – .
Consequently, – – – – .
Therefore, – – – – .
And so, – – – – .

On the other hand, parallelism should be used when it is the point of the sentence.

Now the comments involving content:

Try to make sentences easily comprehensible from left to right. For example, "We prove that ⟨grunt⟩ and ⟨snort⟩ implies ⟨blah⟩." It would be better to write "We prove that the two conditions ⟨grunt⟩ and ⟨snort⟩ imply ⟨blah⟩." Otherwise it seems at first that ⟨grunt⟩ and ⟨snort⟩ are being proved.

While guidelines have been given for the use of the word 'that', the final placement must be dictated by cadence and clarity. Read your words aloud to yourself.

The word 'shall' seems to be a natural word for definitions to many mathematical readers, but it is considered formal by younger members of the audience.

Be precise in your wording. If you mean "not nonincreasing," don't say "increasing"; the former means that $p_j < p_{j+1}$ for *some* $j$, while the latter that $p_j < p_{j+1}$ for *all* $j$.

Mixed tenses on the same subject are awkward. For example, "We assume now ⟨grunt⟩, hoping to show a contradiction," is better than, "We assume now ⟨grunt⟩, and will show that this leads to a contradiction."

Many people used the ungainly phrase "Assume by contradiction that ⟨blah⟩." It is better to say, "The proof that ⟨blah⟩ is by contradiction," and even better to say "To prove ⟨grunt⟩, let us assume the opposite and see what happens."

In general, a conversational tone giving signposts and clearly written transition paragraphs provides for pleasant reading. One especially easy-to-read proof contained the phrases "The operative word is zero," "The lemma is half proved," and "We divide the proof into two parts, first proving ⟨blah⟩ and then proving ⟨grunt⟩."

You can give relations in two ways, either saying '$p_i < p_j$' or '$p_j > p_i$'. The latter is for "people who are into dominance," Don says, but the former is much easier for a reader to visualize after you've just said '$p = (p_1, p_2, \ldots, p_n)$ and $i < j$'. Similarly, don't say '$i < j$ and $p_j < p_i$'; keep $i$ and $j$ in the same relative position.

Don, obviously a fan of rewriting in general, told us that he knows of many computer programs that were improved by scrapping everything after six months and starting from scratch. He said that this exact approach was used at Burroughs on their Algol compiler in 1960 and the result was what Don considers to be one of the best computer programs he has ever seen. On the limits of the usefulness of rewriting, he did say, "Any writing can be improved. But eventually you have to put something out the door."

The last part of class was spent discussing the font used in the coming book: Euler. The Euler typeface was designed by Hermann Zapf ("probably the greatest living type designer") and is an especially appropriate font to use in a book that is all about Euler's work. The idea of the face is to look a bit handwritten. For example, the zero to be used for mathematics has a point at the apex because "when people write zeros, they never really close them". This zero is different from the zero used in the text (for example, in a date), so book preparation with Euler needs more care than usual. You have to distinguish mathematical numerals from English-language numerals in the manuscript.

Somebody asked about 'all' versus 'all of'. Which should it be? Answer: That's a very good question. Sometimes one way sounds best, sometimes the other. You have to use your ear. Another tricky business is the position of 'only' and 'also'; Don says he keeps shifting those words around when he edits for flow.

## §7. Excerpts from class, October 14 [notes by PMR]

Don discussed the labours of the book designer and showed us specimen "page plans" and example pages. The former are templates for the page and show the exact dimensions of margins, paragraphs, etc. His designer also suggested a novel scheme for equations: They are to be indented much like paragraphs rather than being centered in the traditional way. We also saw conventions for the display of algorithms and tables. Although Don is doing his own typesetting, he is using the services of the designer and copy editor. These professionals are well worth their keep, he said. Economists in the audience were not surprised to hear that the prices of books bear almost no relation to their production costs. Hardbacks are sometimes cheaper to produce than paperbacks. For those interested in such things, Don recommended a paperback entitled *One Book / Five Ways* (available in the Bookstore) that describes the entire production process by means of actual documents.

Returning to the editing of his Concrete Maths text, Don went through more of the Before and After pages he began to show us on Monday, picking out specific examples that illustrate points of general interest.

He exhorted writers to try to put themselves in their readers' shoes: "Ask yourself what the reader knows and expects to see next at some point in the text." Ideally, the finished version reads so simply and smoothly that one would never suspect that had been rewritten at all. For example, part of the Concrete Math draft said

> (Before) The general rule is ( ... ) and it is particularly valuable because ( ... ). The transformation in (5.12) is called ( ... ). It is easily proved since ( ... and ... ).

we have $f \circ \sigma \neq f$. If $M(n,k)$ is the number of these and $F(n,k)$ is the number of orbits of mappings $f$ under the action of $C_n$, then evidently

$$F(n,k) = \sum_{d|n} M(d,k) \ (n \geq 1) \tag{1}$$

But since, clearly,

$$\sum_{d|n} dM(d,k) = k^n \ (n \geq 1) \tag{2}$$

we find from (2)

$$M(n,k) = (1/n) \sum_{d|n} \mu(n/d) k^d$$

# Mathematical Writing


*Donald E. Knuth*
*Tracy Larrabee*
*Paul M. Roberts*

With contributions by Paul Halmos, Leslie Lamport, Mary-Claire van Leunen, Nils Nilsson, Rosalie Stemer, Jeff Ullman, and Herbert S. Wilf


**Mathematical Writing** is an all-out attack on the problem of teaching people the art of mathematical writing. We learn by doing, and in these pages, the authors have captured the spirit and substance of Knuth's Stanford course *CS 209, Mathematical Writing—Issues of technical writing and the effective presentation of mathematics and computer science. Preparation of theses, papers, books, and "literate" computer programs...*

Notable here is the mix of mathematical content, with the techniques of writing and rewriting. The participants ranged from ordinary students to experienced mathematical and computer science authors to journal editors to professionals in writing and editing. Together they illuminate the process of writing and editing (for to revise is to edit one's own work).

**Mathematical Writing** will give aid and encouragement for others wishing to teach courses in technical writing, and will be useful to those writing themselves. Those supervising students or employees who write reports or papers will want to have a few copies of **Mathematical Writing** available.


ISBN: 978-0-88385-063-3

9 780883 850633