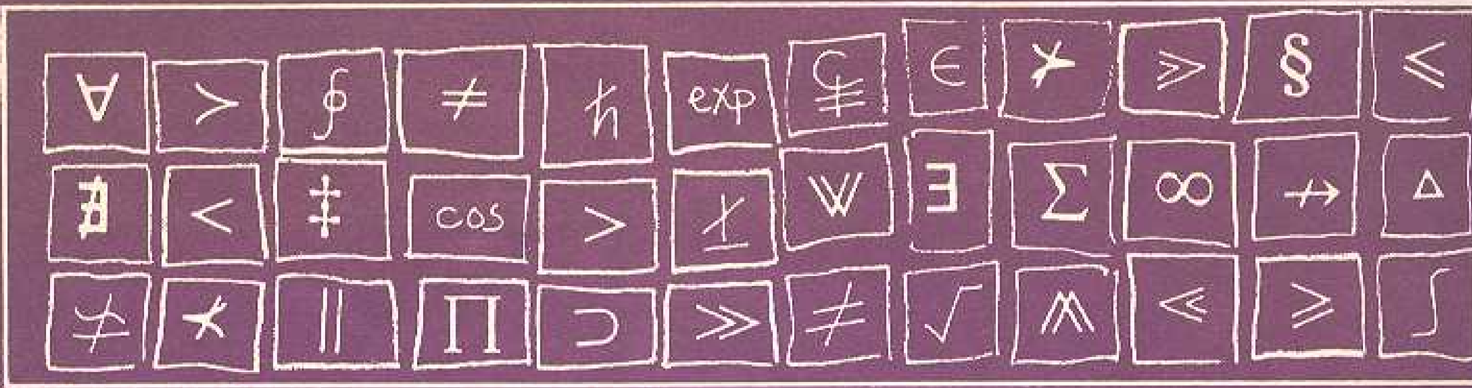
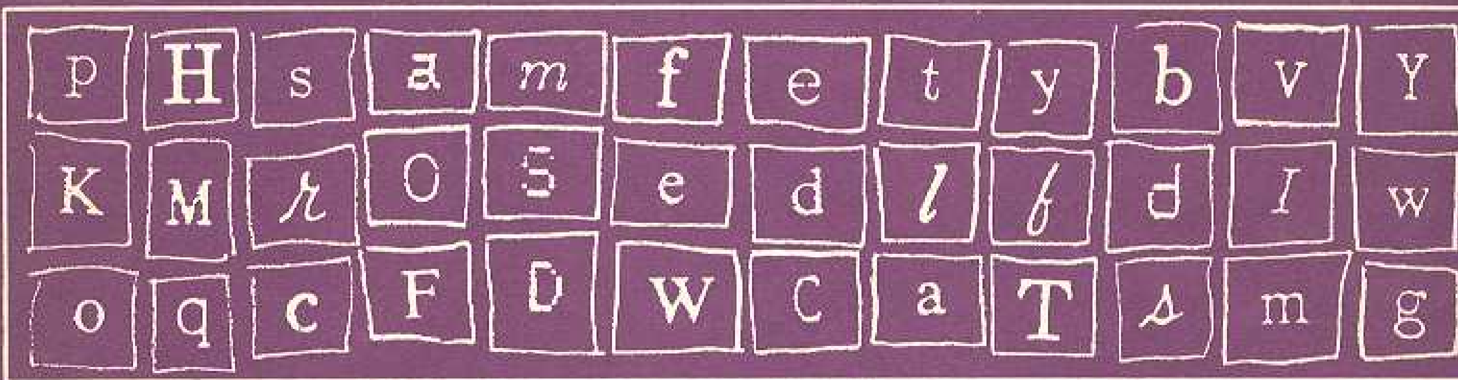
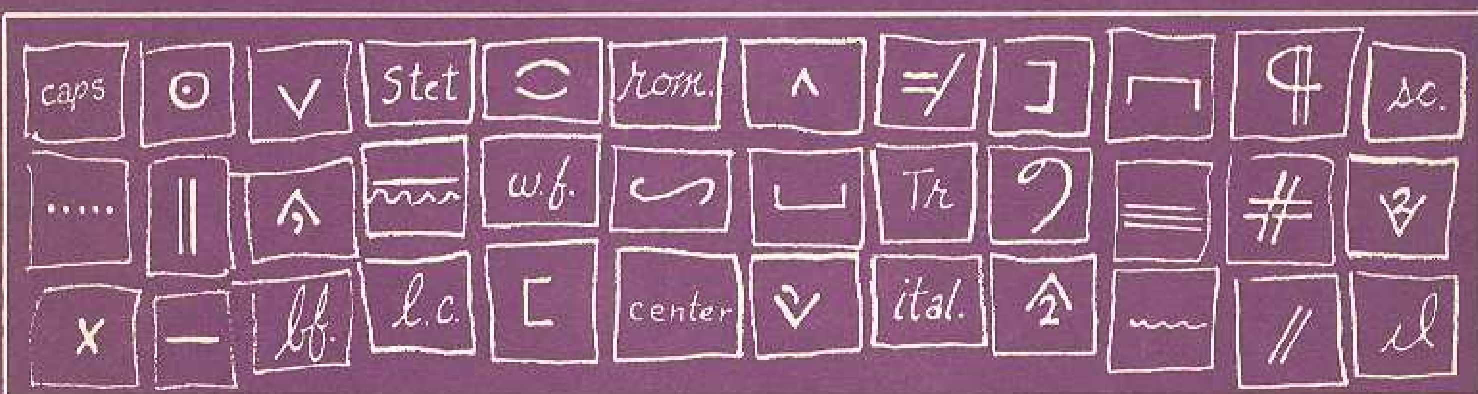


MATHEMATICAL WRITING



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



MAA Notes Number 14

THE MATHEMATICAL ASSOCIATION OF AMERICA

§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:

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

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

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

    and those colons are wrong.

24. The opening paragraph should be your best paragraph, and its first sentence should be your best sentence. If a paper starts badly, the reader will wince and be resigned to a difficult job of fighting with your prose. Conversely, if the beginning flows smoothly, the reader will be hooked and won't notice occasional lapses in the later parts.

    Probably the worst way to start is with a sentence of the form "An x is y." For example,

    Bad: An important method for internal sorting is quicksort.
    Good: Quicksort is an important method for internal sorting, because ...
    Bad: A commonly used data structure is the priority queue.
    Good: Priority queues are significant components of the data structures needed for many different applications.

25. The normal style rules for English say that commas and periods should be placed inside quotation marks, but other punctuation (like colons, semicolons, question marks, exclamation marks) stay outside the quotation marks unless they are part of the quotation. It is generally best to go along with this illogical convention about commas and periods, because it is so well established, except when you are using quotation marks to describe some text as a specific string of symbols. For example,

    Good: Always end your program with the word "end".

    On the other hand, punctuation should always be strictly logical with respect to parentheses and brackets. Put a period inside parentheses if and only if the sentence ending with that period is entirely within the parentheses. The punctuation within parentheses should be correct, independently of the outside context, and the punctuation outside the parentheses should be correct if the parenthesized statement would be removed.

    Bad: This is bad, (although intentionally so.)

26. Resist the temptation to use long strings of nouns as adjectives: consider the packet switched data communication network protocol problem.

    In general, don't use jargon unnecessarily. Even specialists in a field get more pleasure from papers that use a nonspecialist's vocabulary.

    Bad: "If $\mathbf{L}^+(P, N_0)$ is the set of functions $f\colon P \to N_0$ with the property that

    $$\mathop{\exists}_{n_0 \in N_0} \ \mathop{\forall}_{p \in P} \ p \geq n_0 \Rightarrow f(p) = 0$$

    then there exists a bijection $N_1 \to \mathbf{L}^+(P, N_0)$ such that if $n \to f$ then

    $$n = \prod_{p \in P} p^{f(p)}$$

    Here P is the prime numbers and $N_1 = N_0 \sim \{0\}$."

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!]

§2. An Exercise on Technical Writing

In the following excerpt from a term paper, N denotes the nonnegative integers, N^n denotes the set of n-tuples of nonnegative integers, and $A_n = \{(a_1, \ldots, a_n) \in N^n \mid a_1 \geq \cdots \geq a_n\}$. If $C, P \subset N^n$, then $L(C, P)$ is defined to be $\{c + p_1 + \cdots + p_m \mid c \in C,\ m \geq 0,$ and $p_j \in P$ for $1 \leq j \leq m\}$. We want to prove that $L(C, P) \subseteq A_n$ implies $C, P \subseteq A_n$.

The following proof, directly quoted from a sophomore term paper, is mathematically correct (except for a minor slip) but stylistically atrocious:

$L(C, P) \subset A_n$
$C \subset L \Rightarrow C \subset A_n$
Spse $p \in P$, $p \notin A_n \Rightarrow p_i < p_j$ for $i < j$
$c + p \in L \subset A_n$
$\therefore\ c_i + p_i \geq c_j + p_j$ but $c_i \geq c_j \geq 0, p_j \geq p_i \ \because\ (c_i - c_j) \geq (p_j - p_i)$
but $\exists$ a constant $k \ni c + kp \notin A_n$
let $k = (c_i - c_j) + 1 \qquad c + kp \in L \subset A_n$
$\therefore\ c_i + kp_i \geq c_j + kp_j \Rightarrow (c_i - c_j) \geq k(p_j - p_i)$
$\Rightarrow k - 1 \geq k \cdot m \qquad k, m \geq 1 \qquad$ Contradiction
$\therefore\ p \in A_n$
$\therefore\ L(C, P) \subset A_n \Rightarrow C, P \subset A_n$ and the lemma is true.

A possible way to improve the quality of the writing:

Let N denote the set of nonnegative integers, and let

$$N^n = \{\, (b_1, \ldots, b_n) \mid b_i \in N \text{ for } 1 \leq i \leq n \,\}$$

be the set of n-dimensional vectors with nonnegative integer components. We shall be especially interested in the subset of "nonincreasing" vectors,

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

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

$$L(C, P) = \{c + p_1 + \cdots + p_m \mid c \in C, m \geq 0, \text{ and } p_j \in P \text{ for } 1 \leq j \leq m\} \tag{2}$$

be the smallest subset of N^n that contains C and is closed under the addition of elements of P. Since A_n is closed under addition, $L(C, P)$ will be a subset of A_n whenever C and P are both contained in A_n. We can also prove the converse of this statement.

Lemma 1. *If $L(C, P) \subseteq A_n$ and $C \neq \emptyset$, then $C \subseteq A_n$ and $P \subseteq A_n$.*

Proof. (Now it's your turn to write it up beautifully.)

§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 $c_1 \geq \cdots \geq c_n$.