Notes on Writing Mathematics

A. Parts of the paper

Mathematics papers generally follow a standard pattern.

Abstract. Many journals demand an Abstract, and many mathematicians are in the habit of providing one even when it is not required. An Abstract is set off in small type just below the title and author lines. It is a very terse summary of what is accomplished in the paper, normally lacking definitions and background.

Introduction. All mathematics papers begin with an Introduction, which serves three distinct purposes: (1) it introduces the area or problem to be studied in the paper, succinctly and with a minimum of definition; (2) it describes the main accomplishments of the paper, in some detail, and gives some indication of the methods of proof; (3) and it summarizes the structure of the paper, typically listing the sections in order and describing the content of each in a phrase or sentence. The Introduction of a mathematics paper fulfills some of the functions of a Conclusion in papers in many other disciplines. Mathematicians do not like suspense: we want to hear the general lessons right off the top.

Body. This is followed by the body of the paper. Generally, good mathematical writing adheres to the inverted pyramidal form of a newspaper article: push up front as much of the most important and generally interesting material as possible. Sometimes it is necessary to begin with a section on notation, or on background, or on previous work. But never begin with an unstructured series of preparatory lemmas! Lemmas should be integrated into the narrative, in a context which makes it clear how they are used. A weak but commonly used strategy is to defer proofs of lemmas to the end, or at least to a later section of the paper.

Mathematics papers normally lack a Conclusion. The last word may be "QED." Sometimes the author may want to end with some speculation about extensions of the work. But a recap is never part of a mathematics paper.

B. Mathematical style

There is a highly developed standard international style of writing mathematics, which this brief will only hint at.

Attribution. Anything not referenced or explicitly labeled as known is assumed to be original. It is plagiarism to copy phrases or sentences from a source without attribution. But it is also inadmissible to reproduce an argument, even in your own words, without some form of attribution. It is common to refer to secondary sources for arguments, so one rarely cites Gauss or Abel explicitly today, but one often cites secondary literature or textbooks as sources for arguments.

Citation. LaTeX supports several standard citation systems. The easiest uses the command "`\begin{thebibliography}{99}`" where you want the bibiliography. Each bibliograhic item is created by a command like

```
\bibitem{fenn-rourke}
R.~Fenn and C.~Rourke,
Racks and links in codimension two,
J.~of Knot Theory and its Ramifications 1 (1992) 343--406.
```

The bibliography is ended with "`\end{thebibliography}`." The item is referenced in the text by "`\cite{fenn-rourke}`." You need to execute LaTeX twice on the document in order to link the citation to the bibliographic item.

Reference. Rather than "Equation 5," refer to "(5)." This can be automated in LaTeX by inserting "`\label{name-of-this-equation}`" in the range of a displayed equation, and then "`(\ref{name-of-this-equation})`" when reference is needed. You need to execute LaTeX twice on the document in order to link the reference to the label.

Proof. Be clear about the distinction between what you observe empirically, or believe to be true, and what you claim to be able to prove, or to know to have been proved.

C. A miscellany of common errors

What follows is an unordered list of common errors made in writing math papers. Examples are taken from actual papers. Any list like this will reflect the personal preferences of the author to some extent, but I think there is a consensus on most points mentioned.

1. Use of articles. Do not use an article if a noun is modified by a possessive: "an Euler's constant" and "the Euler's constant" are both incorrect; rather one says "Euler's constant" without an article.

In many cases one uses instead a name as an adjective, as in "the Euler characteristic." In this case, one uses an article, and "Euler's characteristic" is incorrect. This distinction is conventional and can only be deduced by studying established usage.

2. Use of parallel constructions. The punctuation of parallel constructions in mathematical English carries information about the role of the construction.

> "Consider C_n^2, the set of 2-incompatible permutations."

This use of language indicates that the notation C_n^2 has already been introduced, and we are being reminded of what it signifies. If instead this is a definition of the notation, one says "Consider the set C_n^2 of 2-incompatible permutations." The absence of commas is important. Surrounding "C_n^2" with commas would give us back the first meaning. This usage can also be used in the first sense, referring to a prior definition.

3. Category errors. A common error in mathematical writing is what I call a "category error." Here is an example:

> "Since $x > 0$, let $y = \sqrt{x}$."

This is poor usage because it reflects faulty logic. We don't let $y = \sqrt{x}$ *because* $x > 0$; we let $y = \sqrt{x}$ because this notation will be useful to us later in the argument. This notation *makes sense* because $x > 0$. One can avoid this error by saying, for example, "Since $x > 0$, we can let $y = \sqrt{x}$."

Here is another example:

> "Since the matrix A is diagonalizable, let $A = PDP^{-1}$."

This sentence has other flaws as well: there is an unstated assumption that D is a diagonal matrix. Fix: "The matrix A is diagonalizable: there is an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$." "Since" is to be avoided here since it is inappropriate to use it in the first clause of a tautology. Note again the complete absence of commas.

And another example:

> "If we draw a line between P and Q, it is easy to show that ...."

One is compelled to ask: what if we don't draw the line? Is it still easy?

5. Comma splices. A comma splice connects two sentences with a comma:

> There is nothing special about 10, this analysis can be generalized by replacing a for 10 as the base."

You can avoid this in many ways—by replacing the comma by a colon or semicolon, or by an m-dash (built as "`---`" in TEX), or by separating the two sentences with a period. Which is best depends on the relationship between the clauses. What follows a colon should be an elaboration of what precedes it, while semicolons separate parallel clauses. Don't overuse semicolons.

6. Conjunctions and commas. "Thus," "Hence," "So," and "Therefore," can all begin a sentence, but don't set them off with a comma. "Hence $x = 4$," not "Hence, $x = 4$."

7. "Such that." In mathematical writing, "such that" has a technical meaning; it is putting a restriction on the noun, as in: "Let $f(x)$ be a function such that $f(0) = 0$." In common language, "such that" is used in several other ways, all of which should be avoided in mathematical writing.

For example, in

> "The trick is to create matrices X and Y such that they form...,"

what is meant is "...that form...."

Another example:

> "We can slowly decrease the spring constant such that the system's natural frequency slowly increases."

Here a simple implication is probably intended: "We can slowly decrease the spring contant, resulting in a slow increase in the system's natural frequency."

Or the author may have intended to put a restriction on the manner in which the spring constant decrease is accomplished. In this case one could say, "We can slowly decrease the spring constant in such a way that the system's ...."

8. Equality, and other mathematical terms. Reserve the notation "=" for true equality! So "$\sqrt{2} = 1.414$" is wrong. There is no universally agreed symbol for approximate equality, partly because the notion itself is ill-defined, but generally $\simeq$ is understood if used in this way.

There are many other common language words used with special meaning in mathematical writing. In the sentence "... for some finitely small ϵ," for example, "finite" is used in the way it is often used in engineering but never in mathematics: it indicates a nonzero quantity. In mathematics, zero is a finite quantity. The word "finite" stands in contrast to "infinite," not "zero." Here the wording could be "... some positive ϵ" or "... some $\epsilon > 0$."

9. Definitions, etc. One often sets off some text as a definition. But it has to be truly a precise definition. Thus

> "**Definition.** The condition number of a matrix is a measure of how sensitive it is to numerical operations."

may give the sense of what the (or a) condition number is, but it is not yet a definition.

There are also Theorems, Propositions, Lemmas, and Corollaries. Choose your names carefully, to make relative importance clear. A theorem should consist of a statement which is of interest and may be referred to from outside the current section. The statement of a theorem should therefore be essentially self contained. Lemmas are small supporting facts, and their statement often uses notations established in the preceeding text. Propositions are technical facts which may be of use elsewhere in the paper, and should generally also be self-contained.

10. Notational consistency and logic. Choose notation carefully and consistently. Try to avoid nested sub- or superscripts: so "$\exp(k^2)$" may be better than "e^{k^2}." But if you use "$\exp(k^2)$" you had better use "$\exp(2k-1)$" rather than "e^{2k-1}" if it comes up nearby.

Avoid burdening symbols for elements by baggage indicating the set it's taken from. Rather than $s_{n+1} \in C^1_{n+1}$, for example, just use $s \in C^1_{n+1}$ (unless of course it happens to be the $(n+1)$st term in a sequence!).

11. Sets and elements. There is a difference between a set and its elements. Thus

> "We let a_n be the size of the complete permutations."

indicates that each "complete permutation" has a "size," which is the same for all of them. This is different from "We let a_n be the cardinality of the set of complete permutations," which is what was intended. Note by the way that "size" is vague—it could be measured in many ways—and we are better off replacing it with a more precise term here.

Here's another example:

> "The induced cycles of a graph are a basis of the graph cycle space."

The object of this sentence is the plural noun "cycles," but then it is claimed that they are a basis, which is singular. This is wrong; rather, one could say "The *set of* induced cycles is [or, better, "forms"] a basis for the cycle space."

Notice how much better this sentence reads with "graph" omitted. The second occurance of "graph" is based on a further sloppiness of thought; as it is written, it distinguishes the graph cycle space from some more general notion of cycle space, while what is probably intended is "the cycle space of the graph."

Another example:

> "For the equation $X^3 + Y^3 = A$, with X, Y, and A positive integers, there exist infinite numbers A such that there are at least two solutions $\{X, Y\}$ that satisfy this equation."

There is no such thing as an infinite number; it's a certain *set* of A's which is infinite. There are other problems with this sentence: What is meant by the phrase "for the equation"? The symbols X and A play different roles, but are introduced in parallel. Ending the sentence with "this equation" is repetitive and signals clumsy construction. Better: "There are infinitely many positive integers A for which there are at least two unordered pairs of positive integers X, Y such that $X^3 + Y^3 = A$."

12. Sequencing. Sequence introduction of objects logically.

> "Let $X \subseteq Y$. We say that a function $f : Y \to \mathbb{R}$ is an extension of some function $g : X \to \mathbb{R}$ provided that $f(x) = g(x)$ for all $x \in X$."

What is intended is that g is given to us in advance, and the author is explaining what it means for f to be an extension of it. Better: "Let X be a subset of Y. An extension of $g : X \to \mathbb{R}$ to Y is a function $f : Y \to \mathbb{R}$ such that $f(x) = g(x)$ for all $x \in X$."

13. Symbols as verbs, sentences as nouns. In **12** we have also replaced the symbol "$\subseteq$" with the phrase "be a subset of." Generally it is better not to burden a piece of notation with a grammatical role (a verb in this case), but this rule is often bent.

Here's another example:

> "Hence, we know that the eigenvalues $\lambda_j \in \mathbb{Z}$ for all $j = 1, \ldots, n$."

This sentence contains other problems: there should not be a comma after "Hence," and "we know that" and "all" are both redundant. So a corrected version might read "Hence the eigenvalues λ_j are integers for $j = 1, \ldots, n$." This example comes close to providing an example of the good practice of indicating the set to which a named element belongs, as in "The eigenvalues $\lambda_j \in \mathbb{Z}$ are all even."

Don't use a sentence as a noun:

> "We use $a|b$ to denote "*a divides b*"."

As a matter of standard English punctuation, commas and periods always slide left under quotation marks; and there is a difference between left quotes and right quotes. So: "Write $a|b$ to indicate that a divides b."

13. Standard notations. There are a few pieces of standard mathematical notation. It is perverse to introduce new notations for objects or structures with standard notation, or to use the standard notations for a new purpose. There are also conventions regarding the use of fonts, which are flexible but should be bent sparingly. Think twice before using "x" for a set or for a constant, for example.

14. Avoid the ugly, the cute, and the wasteful. One of the ugliest phrases in mathematical writing is "...we have that...," as in "Because n is even we have that n^2 is even." Unfortunately, it is also quite common. Avoid it. Another ugly phrase: "...and we are done," at the end of a proof.

Don't be cute: Don't use "ergo," "thusly," "herein," "aforementioned," or other such Britishisms. It's "First, ..." and "Second, ...," not "Firstly, ..." and "Secondly, ..."

Use phrases like “It can be shown that” or “It is easy to show that” with care. Often the text can be improved by simply omitting them.

15. This and that. In using words like “this” and “that,” be very careful that the antecedent is clear. The sentence

> “In order to acheive that, multiple stategies exist.”

occured at the start of a new section; it was impossible to determine what “that” referred to. These words are better used as adjectives: note how much clearer the last sentence was than if I had chosen to say “This was at the start of a new section.”

The quote also illustrates the importance of thinking about phrase order. Better: “Many strategies exist to acheive this goal.”

Finally, this example provides another example of undesirable cuteness. One says “multiple births” but “many strategies.”

Haynes Miller
August, 2007