Writing Mathematics Papers

This essay is intended to help your senior conference paper. It is a somewhat hastily produced amalgam of advice I have given to students in my PDCs (Math 4 and Math 9), so it's not fined tuned for you, but it still should be useful.

Mathematics writing is different from ordinary writing and harder — in addition to all the requirement of ordinary good writing, there are additional constraints and conventions in mathematics. An additional constraint is that mathematics follows much more demanding rules of logic than ordinary discourse, and your writing must follow and display this logic. (However, this does *not* mean that there is only one right way to present a mathematical argument.) Some of the additional conventions are those for defining new concepts and those for organizing the material through theorems and examples.

Titles

Every paper should have one. It should be informative without being too long. Choosing a good title is not so easy. Often a math paper hinges on a concept that is only defined within the paper itself, so to use the name of that concept in the title will convey no meaning at all. Incidentally, whereas in some literary fields an obscure word-play is considered a good title, it is generally not well regarded in mathematics.

Theorems

A math paper is not complete without some theorems and proofs. Each theorem in your paper should be highlighted — I suggest indenting it, putting extra vertical space around it, and using a different type style if that's available to you. (Other formats are allowed; browse and see.) Why highlight theorems (and other key items such as examples)? Because mathematical arguments can be so complex. Formal statements of theorems (and other key items) serve as touchstones, like having an outline within the text body. They also makes it easy for the reader to refer back to key specifics. For the same reasons, for each proof in your paper you should indicate clearly where it begins and where it ends.

Definitions

Definitions are a major way that math writing differs from general writing. Most disciplines don't deal with definitions nearly as much as math — they don't need to be so precise nor do they deal so regularly with situations outside common experience. Along with the need to define words (derivative) comes the need to introduce notation ($\frac{d}{dx}$): if you use a new concept frequently, you need a shorthand way to refer to it.

How and where definitions should be presented. For the most part, you will be tempted to define a word immediately after you first used it. This shows lack of planning and lack of thinking about what will read best. When this policy leads to definitions being parceled out in twos and threes in the early part of the paper, it works fine. But if it means you interrupt a difficult, key proof late in the paper to give two more definitions, it is disruptive and confusing.

Before you start writing a math paper, you should make a list of all the definitions you think you need, and then try to see to it that they all get introduced relatively early. (Of course, you will not yet realize some of the definition you will need, so after the first draft you will have to go back and add more definitions.) The extreme version of this approach is to put all the definitions in a definitions section near the beginning. This isn't always best either: a long section of definitions can be deadly boring.

When you give a definition, you can do it "in-line" (within a paragraph), but the word or phrase being defined should be highlighted, by underlining, by *italic print*, or by **boldface**. Boldface is now the most common format, since underlining is not common in typeset material and italic already has a use in both math and ordinary writing to indicate *emphasis*. Here's an example of a definition: "A **prime number** is a positive integer with no positive integer divisors other than 1 and itself." Another format is to display definitions just like theorems are displayed. The display format should be reserved for the most important definitions. For instance, in Math 4 the definition of derivative might be displayed, but the definition of polynomial (if you need it at all) can be done in-line.

Once you define a word, don't ever substitute another word which is a synonym in ordinary English. For instance, if you have defined "slope", don't afterwards sometimes say "steepness". The whole point of mathematics definitions is temporarily to give ordinary words some extra precision needed for a mathematical discussion. But there may be several different precise meanings that are mathematically useful and that one can attach to the same looser ordinary concept. A trained mathematical reader, upon seeing "steepness" in your paper after you have defined "slope", will assume you are using both words because you need two different refinements of the ordinary slope concept. Such a reader will then attempt to find where you defined steepness and will get frustrated.

Among the methods of indicating that a word is being defined, I did not include quotes. Quotes may be used to indicate that a word is being used in a somewhat special way, but no declaration is thereby made that this special way applies for the rest of the paper. Here is an incorrect use of quotes:

> The initial case $P(n_0)$ of an induction is called the "basis".

This is incorrect because basis is being defined. Here is a correct use:

> In order to verify the inductive step, we first "simplify" $\sum_{k=1}^{n} k^2$ to $\left(\sum_{k=1}^{n-1} k^2\right) + n^2$.

The point is, it is a bit of a surprise to call this a simplification (although it does simplify in the sense of leading to a solution) and we don't claim we will always use "simplify" in this sense.

Local Definitions. So far I have been talking about "global" definitions, those that apply throughout your paper. Most global definitions introduce words or notation. In contrast, local definitions apply only briefly, say, to the current example or current paragraph. Usually local definitions are definitions of symbols, for instance, the f in "let $f(x) = x^2$." Local definitions don't require special highlighting, although they should be displayed if they involve complicated formulas (see the section on displays later). In any event, the item being defined should always come on the left-hand side of the equation; e.g., "let $x^2 = f(x)$" is *wrong*. Why?

Because when an equal sign is used with "let" to make a definition, the equal sign has a special meaning: the quantity on the right is being assigned to the quantity on the left. This is quite different from regular equality, where it makes no difference if you write $a = b$ or $b = a$.

Examples

Examples are like definitions, in that they can appear in-line or be highlighted by indentation and extra spacing. For a very brief example, the in-line method is fine. However, a lengthier example should be displayed, especially if it is a key item of your work. Not only does this format draw attention to your example, but is also makes it easy to refer back to later (e.g., see Example 3). If an example is a sample problem, make clear where the solution begins and ends.

Figures

Figures can be extremely helpful. Each figure should be numbered, referred to by number, and (preferably) be positioned in the paper shortly after the first reference to it. (Another convention is to put all figures at the end. If you use this alternative, please say so the first time you reference a figure.) I'd prefer each figure to have a caption as well; e.g., "The steeper the line, the greater the slope". If you know how to use computer software to produce a good figure and place it into your paper, great, but I'm perfectly happy if you insert your figures by hand. You'll probably need more space than you first think.

Footnotes

Generally, footnotes are *not used* in mathematics papers. First, the superscripts used to number them might be confused with exponents. Second, it used to be that mathematical symbols weren't always available in the smaller type used for footnotes. (Some math texts, including yours, do use footnotes occasionally, but published paper still almost never do.)

So what do you do? If you need to refer to an outside source in the text, write $[x]$ (where x is the number of the reference in the reference list at the end). For instance, if Goldstein et al. (the text for Math 4) is reference 1 in the list at the end of your paper, you might write: "The following application is treated in [1]", or "The following is treated in Goldstein et al [1]". If you want to refer to a specific place in your reference, a theorem or example number is at least as valuable as a page number: "The following application is treated in [1, Thm 3.5].

When to Give Credit

The conventions for giving credit to others are somewhat different in mathematics than in general academic writing. Briefly put, direct quotes must be credited, but paraphrase generally need not be. Furthermore, math papers rarely include direct quotes. Thus, while a humanities paper is full of footnotes giving credit, in a mathematics paper the credits are less frequent (and they appear in text, not in footnotes, as explained earlier). This is all elaborated below. Pay close attention, since failure to give credit when it should be given amounts to plagiarism and is a major academic sin.

Rightly or wrongly, mathematics is regarded as having an existence independent of the words used to describe it. Thus Goldstein may describe the Chain Rule theorem in slightly different words than anyone else, but that doesn't give Goldstein any special credit. If you use the Chain Rule, and you learned about it from the Goldstein text, you should *not* reference Goldstein at the point where you introduce the Chain Rule. This would be an embarrassment to Goldstein and his coauthors, because they would never claim credit for the Chain Rule themselves. This same principle applies to definitions; don't reference Goldstein for the definition of derivative.

Note. Theorems have two sorts of names: descriptive names, such as "Chain Rule"; and sequential names, such as "Theorem 6" and "Limit Rule II". A text gives a descriptive name only if it has become widely used, so you can use the name "Chain Rule" too. But sequential names are specific to an individual text. Thus, if you must refer to Limit Rule II, you have to reference Goldstein and give a page number — otherwise readers won't have a clue what rule you are referring to. But why make your readers look this up when you could simply state the rule in your paper?

The "independent existence" perspective also explains why mathematicians rarely quote each other directly — your original words are generally not perceived to have an advantage over my paraphrase. This attitude is quite different from that in humanities, where the nuances in somebody's verbalization of an idea may make all the difference. (If you feel you should nonetheless use a direct quote, then write: "As Smith says [3, p. 43]" and then display the quote as an indented paragraph.)

The original *discoverer* of a mathematical result or method *is* given credit for it, and if the result is fairly recent, the paper in which the result first appeared must be referenced. However, if the result is classical, e.g., Euler's graph theorem, it suffices to call it Euler's Theorem; no reference to a source is necessary. (In mathematics, a result is "classical" if it is a well known and already appears in many written sources; occasionally results proved as recently as 20 years ago have already become classical.)

So much for giving credit for definitions and theorems. Should examples be credited? Again, the answer is "usually no". If the example is very unusual and involved, give credit. Also, if you use the exact numbers and words of an example from some book, then give credit; but in general you shouldn't be quoting problems verbatim anyway. In order to write about an example with understanding you should internalize it and make it your own before using it. And if you have internalized it, you can recreate it (perhaps with different numbers) without looking it up.

Bibliography format. Several disciplines seem to have special conventions for references. The style used for most mathematics is different from the style I was taught in high school and from the style in the Chicago Style Manual (followed by some other disciplines). Here are standard math examples for journals and books.

1. J. E. Skeath, Holomorphic functions in 2 variables, J. Complex Analysis 31 (1966) 137-143.
2. D. Rosen and E. Mullins, "Calculus with Probability," Boden & Quigley, Philadelphia, 1968.

(In the first, “Holomorphic Functions in 2 Variables” was an article appearing on pages 137-143 of the Journal of Complex Analysis, Volume 31, which volume appeared in 1966.)

Fonts for Math

A font is a style of type, e.g., roman or italic. In books, letters representing mathematical quantities, such as a, x and $f(z)$, are *not* set in the same roman font as ordinary text. The default for math letters is italic, as just done. However, certain quantities are traditionally printed in other fonts, for instance, the vector $\mathbf{x}$ is usually boldface, the angle θ is Greek, large sets $\mathcal{S}$ are often script capitals, and certain authors like to use old-German letters, like $\Re$. More generally, if you have several quantities, and each has an English name beginning with “M”, you might want to represent each of them mathematically by the single letter m. Then you need to use a different font for each meaning.

The previous paragraph describes what books do. What should you do? If you have access to a good word-processing system, you will have italic fonts available and you should use them. If you don’t have italics (for instance, you are using a traditional typewriter), then put extra blank space around mathematical letters and expressions. For instance,

```
Consider the variables  x  and  y  in the equation  y = 2x+3.
```

Such extra space is especially important around the letter a, since it could be either an English word or a math symbol.

While letters in mathematical expressions are in italic, punctuation (like parentheses, brackets and exclamation points) are not. This makes writing a little tedious when (as with a basic Macintosh system) all you have are ordinary roman and ordinary italic fonts. In ordinary italic, punctuation is slanted too, so you have to go back and forth from roman to italic just to write a single equation. What you really need are “math italic” fonts, but none of the standard Macintosh fonts are math italic. A basic Mac system does have subscripts and superscripts; learn how to employ them.

Warning: There is a special problem when Macintosh MacWrite documents containing italic are printed with a laser printer. When you sit at a Mac and type italic f followed by roman (, the slanted f will overlap the (on the screen. You will be inclined to fix this by putting a space between them. If you do this *the laser output won’t look right* — there will be too much space between the f and the (.

The problem is that the laser output and the screen image are not exact duplicates. Fortunately, there is a fairly simple fix. First, always write you math papers using a “laser font”. The fonts named after cities (Geneva, Chicago, etc.) are *not* laser fonts; most of the others available with MacWrite (Times Roman, Courier, Helvetica) *are* laser fonts. Since Geneva is the default font, you have to “pull down” the font window to change the font before you start writing. I especially recommend switching to Times Roman. Second, *don’t correct* the overlaps you see on the screen. When you print things out, it will look fine.

I suspect the same problem, and the same solution, obtain if you use the word-processing software “Word” on a Macintosh, but I haven’t tried it.

Handwriting Mathematics. If your method of producing your paper gives you no special capabilities for printing mathematics, you will need to insert the more complicated mathematical displays of your paper by hand. Even if you have a computer with math italic fonts, complicated mathematical displays are still hard to "set". If you haven't set any complicated mathematics by computer before, it's not worth it for you to spend hours learning how to do it now. (On the other hand, any practicing mathematician should probably learn at least one such system eventually, preferably the most powerful, which currently is TEX.) If a display is going to take a lot of time to set, then leave a blank space in your paper and write it in by hand. At first, you will probably grossly underestimate the amount of space you need for hand insertion. But if the rest of your paper is written on a computer, it will be easy to output another copy with more space.

Displays

Any long expression with mathematical symbols is displayed — centered on a line by itself, with extra vertical space around it. There are three reasons for displays. First, if an expression is particularly important, you draw attention to it by displaying it. Second, longer mathematical expressions tend to be tall. To fit something like $\displaystyle\int_2^5 \frac{dx}{x^2+1}$ in-line requires adding disconcerting interline space (see it?), or else making the symbols disconcertingly small, like this: $\int_2^5 \frac{dx}{x^2+1}$. Third, it is distracting to break an expression across a line. You can hyphenate words, but math expressions are usually longer than words and don't bear hyphenation well. So if an expression would have to be broken if it appeared in the middle of a paragraph, it should instead be put in a display. For instance, if you wish to say

$$y' = \frac{d}{dx}\,(x^2 + 3x)$$

within a line, it would be bad form to break this after the + and *terrible* form to break it after $\frac{d}{dx}$. According to some writers, it is even bad form to break it after "=".

A display should be numbered if you find that you refer to it anywhere other than within a few lines of it. The numbering can appear either on the left, as in

(1) $$y' = \frac{d}{dx}\,(x^2 + 3x),$$

or on the right, as in

$$y' = \frac{d}{dx}\,(x^2 + 3x), \qquad (2)$$

but be consistent. Just below a display you can refer to it as the "previous display", but two pages later it's too long-winded to refer to "the third display on page 7". Besides, if you modify your paper, you will have to rewrite such a reference completely. That's why we use numbering.

Multiline Displays. In most cases, these should be lined up so that the main connectives (usually an =, but maybe $\leq$ or $\Longrightarrow$) fall right on top of each other:

$$\begin{aligned} 2x+1 &= 2+3, \\ 2x &= 4, \\ x &= 2. \end{aligned} \tag{3}$$

Similarly, one writes

$$\begin{aligned} x^2+2x &< x^2+2x+1 \\ &= (x{+}1)^2. \end{aligned} \tag{4}$$

Notice that display (3) has expressions on both sides of the equal sign on each line, but display (4) has only a right-hand side after the first line. When a display could have been written as one long line (e.g., $a = b < c = d = e$), but is written in several lines for emphasis or because it won't all fit on one line, this "right side only" format is appropriate.

The fact that calculations *can* be written as one long display without words doesn't mean that they *should* be written this way. Mathematics does not consist of calculations alone! Very elementary algebra, as in (3), need not be explained, but more complicated calculations should be, especially if the calculations are justified by mathematics that the reader is just learning. There are two ways to provide the explanation. One (your text's way) is to put it before, between and after the lines of calculation. For instance, suppose you have defined $f(x) = x^2$, and want to show that $[f(x{+}h) - f(x)]/h = 2x{+}h$. You could write:

> Substituting the definition $f(x) = x^2$,
>
> $$\frac{f(x{+}h)-f(x)}{h} = \frac{(x{+}h)^2-x^2}{h}\,.$$
>
> Then expanding the first square, combining like terms, and finally canceling the common factor h, we obtain
>
> $$\begin{aligned} \frac{f(x{+}h)-f(x)}{h} &= \frac{(x^2+2xh+h^2)-x^2}{h} \\ &= \frac{2xh+h^2}{h} \\ &= 2x+h. \end{aligned}$$

An alternative approach is to put all the reasons in comments on the right:

$$\begin{aligned} \frac{f(x{+}h)-f(x)}{h} &= \frac{(x{+}h)^2-x^2}{h} && \text{[def. of } f(x)\text{]} \\ &= \frac{(x^2+2xh+h^2)-x^2}{h} && \text{[expand]} \\ &= \frac{2xh+h^2}{h} && \text{[combine terms]} \\ &= 2x+h. && \text{[divide]} \end{aligned}$$

Actually, for you seniors all of these steps are standard enough that you could skip the commentary. But if you were a calculus student you shouldn't skip it. In your own papers, any calculations that weren't immediately obvious when you first saw them in your reading should be explained.

Punctuation

Punctuation in displays. Notice the commas and the period in display (3). Mathematics is written in sentences, and this display is a sentence consisting of a sequences of clauses, the equations. Thus the clauses are separated by commas (some writers would use semicolons) until the sentence ends with a period. However, there is an alternative convention (followed by about half the publishers in the US) which leaves off punctuation at the end of displays. Personally, I'm not so concerned about the commas (the separation into different lines tells me to pause) but I do feel strongly about the period. Sometimes a sentence does *not* end at the end of the display; that is, I am supposed to keep reading in order to understand why the display makes sense. On the other hand, if the display ends is a period, that is a signal that I am supposed to figure out for myself why the display is legitimate. In other words, the presence or lack of a period at the end of a display tells me whether this is a point at which I need to stop and digest what has just been said. Since it is very important for the author to give the reader cues like that, the punctuation at the end of a display is crucial.

In display (4) there is no comma after the first line and there never should be. Why?

Punctuation in text. Consider the sentence "If Mary goes, John, Sam and Alice will too." The word "then" is missing, but the meaning is clear and so such an elipsis is quite acceptable English. However, if we leave "then" out from

$$\text{If } x = 3\text{, then 4, 5 and 6 are the next cases.}$$

we get

$$\text{If } x = 3\text{, 4, 5 and 6 are the next cases.}$$

Now we have a major ambiguity: which numbers are possible values for x and which are the next cases? Moral: never separate mathematical expressions that are distinct (not merely part of the same list) by commas alone. Similarly, don't separate distinct mathematical expressions by space alone. For instance, although the sentence

$$\text{We will call the value of } P\ P(n)$$

is perfectly reasonable in meaning, it is hard to read. Are P and $P(n)$ being multiplied? So rewrite the sentence to get some words between P and $P(n)$:

$$\text{This value of } P \text{ we will call } P(n).$$

Paragraph and Sentence Divisions

In reading student papers, I often find cases where a paragraph should have been broken up, or the last sentence of one paragraph belonged at the start of the next. Remember, writing isn't formatted into paragraphs to look nicer on the page or to indicate where you took a deep breath. Paragraphing is used to separate the development of one idea from the next. Similarly, a sentence should make one point. If there are several different points, they should go in several sentences.

In a proof, a good rule of thumb is that a sentence should give just one deduction. A paragraph in a proof should represent a major segment of a proof. For instance, if a proof of $A \Longleftrightarrow B$ is short, then one paragraph should be $A \Longrightarrow B$ and the other $B \Longrightarrow A$. If, however, $A \Longrightarrow B$ is long, then it too should be broken in parts. For instance, if you are presenting a long proof that

$$\text{Euler cycle} \implies [\text{connected \& even degree}],$$

then devote one paragraph to

$$\text{Euler cycle} \implies \text{connected},$$

and another to

$$\text{Euler cycle} \implies \text{even degree}.$$

Two Common Mistakes

The first mistake is not to rely on mathematical symbols enough. If, for instance, you refer to the same function more than once, give it a name (let $f(x) = x^2$). Even if you are just discussing functions in general, you need to give the generic function a name if you are going to say anything about it. (e.g., for any function $f(t)$, $f'(t)$ measures the rate at which...). More generally, if you are working towards the precise formula in some definition or theorem, it may be wise to state the formula early on, before you have even explained all the symbols that occur in it. This way at least you have something concrete to refer to as you work through your explanation. Otherwise you bog down in vague verbiage.

The second mistake is to rely on mathematical symbols too much. This happens when you present several lines of computation without any commentary and was discussed in the section on displays.

Revising Papers

This is tougher than you might think. If you have to add or move some definitions, or add or move a theorem, you will have to look over the whole paper to see what else is affected. Given the tight structure of good mathematics writing, a change in one place can and should require adjustments in many other places. Indeed, there is even a riddle/joke about this.

> Riddle: what is the order in which you write the chapters of a math book?
>
> Answer: 1,2,1,2,3,1,2,3,4,...

Miscellaneous

Please number all your pages. This is especially helpful if you should staple them out of order (which happens) or if I get them out of order while reading (this also happens).

If you have any questions or comments on the above, don't hesitate to check with me. If your question/comment is a good one (and most of them are), I will send it and my reply to all your classmates via computer (unless you ask me to keep your inquiry and my response private).

Finally, my writing isn't perfect either. If you find fault with the writing above, please point this out to me. We can all benefit from each other's criticism.

— end —