HOW TO WRITE A CLEAR MATH PAPER: SOME 21ST CENTURY TIPS

IGOR PAK*

Abstract. In this note we explain the importance of clarity and give other tips for mathematical writing. Some of it is mildly opinionated, but most is just common sense and experience.

1. Be clear!

This is the *golden rule*, really. It's absolutely paramount. Let me explain.

1.1. **What does it mean to be clear?** This might seem like an obvious question, but it's not. Most people think it's about clarity in phrasing, that's all. For example, one should of course write

Abelian groups have trivial center.[1]

rather than

It was discovered by Galois, and later proved formally by Jordan in 1870 (see [Struik]*),*
that having the identity being the only fixed element commuting with any other element
is implied by the abeliannness of a given group.

In fact, this type of clarity is hard to achieve and even harder to teach. While, of course, one should make an effort and try to avoid some easy pitfalls, that's not exactly what I am talking about. The rest of the paper is really a long answer to this question. But let us first take a step back and answer more basic questions.

1.2. **Being clear – how hard can that be?** Well, it can be easy. But it can also be pretty hard, especially if you are an inexperienced writer. The trouble with being clear as a concept, is that most people think it doesn't take time. They think one naturally becomes a better writer. Quite the opposite is true. Making your paper clearer takes time and a lot of effort. You learn to do this faster of course, but it's still a slow process. I once asked Noga Alon how did he get to be so good (and so fast!) at writing. He said "it gets easier after the first 300 papers".

Now, as it always happens, the real test of your commitment to clarity is not when it's easy, but when it's hard. Imagine the following scenario. While finishing your paper you realized that in some sections you use h as a variable, and in other sections h is a function. And on the very last page you had to write $h(h)$ which is just awful. What should you do? Should you spend maybe 30 min going over every instance of h in the

*Department of Mathematics, UCLA, Los Angeles, CA, 90095. Email: pak(@math).ucla.edu.

[1]Mathematically, this statement is completely false. But that's part of my point – how would anyone even know that in the second version? When you are unclear, all claims look reasonably true.

paper and renaming it accordingly? Can't you just make a disclaimer at the beginning of every section "In this section, h is function" and be done? After all, there might be only 2-3 people getting far enough in the paper to be confused, and it would take them only 1 min each to be unconfused, so the arithmetic seems to favor the lazy approach. The answer is **NO**, you should definitely spend these 30 min and fix the notation issue. Yes, really. Let me explain.

1.3. **Why be clear?** Now that we framed it as a tradeoff between your time and effort, and that of the readers, this is no longer an obvious question and it deserves a full explanation. And the key observation is – being clear is not about you! You must think of the reader and how they will read your paper.

Imagine a graduate student at a small university with poor English skills. He is reading your paper. If confused on page 3, he is likely to give up and never finish the reading. He might use an older paper with a weaker result for his research, just because it's better written. *Conclusion*: you didn't make him spend 1 extra min – you just lost a significant fraction of your readership.

Or imagine a postdoc at a major research university. She has a clear project to finish and her supervisor gave her 20 possible papers to "check if they might be helpful". She is quickly looking through your paper. Not noticing your "notation explanation" she is becoming completely confused about the notation and consequently the main results. Rather than making an effort, she assures herself that your paper is irrelevant to the project and moves on to read the other 19 potentially helpful papers. As a result, some theorems do not get proved and the project never gets finished. *Conclusion*: you didn't make her spend 1 extra min – you lost both the citation and a chance to advance the area.

Let me mention two more reasons which are variations on the same theme. For junior mathematicians: *clear writing will make people take you seriously.* It is pretty easy for lazy senior scientists to brush off a paper on the subject with ambiguous results and uncertain proofs. But when you are clear they have no excuse. Don't give them one! Forget that they themselves have been publishing sloppy writing for decades. You are not competing on the same level (yet). In fact, there is an actual checklist on what it takes for senior people to read your paper [1]. Study the checklist and make sure you get an easy pass.

Finally, for all mathematicians: *clear writing will give you a competitive advantage.* It is often the case that the same or nearly the same result is obtained in several papers. If your paper is clear and your competitors' are not, you will get the credit. I know, this is unfair. Think about it differently – you outworked your competition and created a better product. Sometimes it's not about the substance but the presentation.[2] As everyone knows, recording of the same symphony by different orchestra can have very different values. In the era of winner-take-all society it shouldn't be surprising that the same happens to math papers.

1.4. **Can't journals help?** In a word, **NO**. In my experience the copy editors can point out some sentences which are unclear. But these are linguistics rather than math issues. It's like when you are editing a literary book in an unfamiliar foreign language.

[2]I wrote in [10] how Sylvester's "fish-hook" bijection was rediscovered in over a dozen papers. Most authors were aware of other versions, yet all claimed their presentation to be superior over others.

Sometimes you can still find some hanging sentences, sentences without a verb, etc., even if you have no clue what is being said.

But more importantly, who cares? You are likely going to be posting your paper on the `arXiv` anyway, where most people will find it (or on your web page, either way). So the journals are cut out of the process, and you yourself should strive to make your paper as clear and you possibly can.

1.5. **For the sake of clarity, ignore all rules!** This is motivated by the *"Ignore All Rules"* guideline page for *Wikipedia* editors.[3] Roughly, I am saying that when the rules of style and grammar make math unclear, you should simply ignore these rules. Try rewording the sentence first, of course, but if nothing works, go for it, no matter how fundamental the rule is. I will expound on this a little more later, in §5.2. For now, let me mention an example where even the most basic rule — "end all sentences with a period" — leads to a mathematical confusion (intentionally amusing, of course); see Exc. 7.1 in [13]. My point: don't do this unless you are aiming for a comedic effect in a textbook.

2. Where to start

2.1. **Not with this article, but with other literature.** Mathematical writing tends to be so poor, no wonder there are so many very good guides. These include famous essays by Halmos et al. [4], and nice books by Higham [5], Knuth [8] and Krantz [9]. More recent guides we want to mention are by Berndt [2], Goldreich [3] and S. P. Jones [7]. Further essays and resources are included on Terry Tao's blog [14].

2.2. **Read a good guide on writing nonfiction.** I strongly recommend Zinsser's book [16] in part because I don't know any other, but in part because it's so well written I can't imagine a better guide. To get the taste, here is a short section on how to organize your paragraphs, see [16, p. 80]. Most of this applies to math papers with minor adjustments:

> "Keep your paragraphs short. Writing is visual—it catches the eye before it has a chance to catch the brain. Short paragraphs put air around what you write and make it look inviting, whereas a long chunk of type can discourage a reader from even starting to read.
>
> "Newspaper paragraphs should be only two or three sentences long; newspaper type is set in a narrow width, and the inches quickly add up. You may think such frequent paragraphing will damage the development of your point. Obviously *The New Yorker* is obsessed by this fear—a reader can go for miles without relief. Don't worry; the gains far outweigh the hazards.
>
> "But don't go berserk. A succession of tiny paragraphs is as annoying as a paragraph that's too long. I'm thinking of all those midget paragraphs—verbless wonders—written by modern journalists trying to make their articles quick 'n' easy. Actually they make the reader's job harder by chopping up a natural train of thought."

[3] See `https://en.wikipedia.org/wiki/Wikipedia:IAR`

Let me tailor my advice. If you are a native English speaker, read Zinsser before anything else and take his advice to heart. Think of it this way: Zinsser's book is to mathematical writing as good *foundation* is to a perfect makeup. Now, if you are are not a native English speaker, read Halmos and other short pieces first. Come back to Zinsser when you gain more experience. After a few more years, read it again – you will most likely find something useful you missed the first time around.

2.3. **So why do we need this new guide then?** I don't have a concise answer for that. I think the world is changing too fast. With the ever increasing competition for jobs, publishing in top journals, etc., some of the old advice needs to be calibrated and adjusted for modern times. This is particularly true about typesetting in LaTeX which is universal and represents its own advantages and challenges. While most advice in [8] still applies, it feels overwhelming and somewhat stale, while the TeX-nology part is surprisingly incomplete.

To make further contrast with older works, one no longer expects their papers to be all that interesting to survive decades. It's the short term goals that became all too important. Thus the emphasis should be on a modest goal of *clear* rather than *perfect* writing.

The same applies to reading. With the rapidly increasing growth in the number of publications, nobody has time nor patience to read all relevant papers. Some people read many titles on the `arXiv`, only occasionally reading the abstracts. Some quickly skim most papers in their areas, but read none carefully. Some just read the introductions. Some read whatever is suggested by `Google Scholar` with its obvious bias towards citations of their own papers. Some skip everything in the paper and go straight to main results; if sufficiently interested, they then go back to read what's it all about. Some read nothing at all and learn about new work at seminars, conferences, etc. So if you want to increase your readership and enhance their reading experience, the papers need to be written in a new manner compared to old style guides, to appeal to all these diverse readership styles.

3. Macro tips

3.1. **Structure of the paper.** Every newspaper writing guide, including the above mentioned [16], will advise to write an article in a Matryoshka doll manner – start with a super brief summary, then make a longer summary, and only then, once the reader is hooked and interested in details, proceed to give a complete set of facts. Over the years, math articles developed a similar structure with a progression of the title, abstract, introduction, the main part, final remarks and references. I feel that modern practices make some corrections here when compared to old guidelines. Let me discuss each part separately.

3.2. **Title.** This is super important. Read about how to write a good title *everywhere*. Think about it a long time. Try different versions on your colleagues. Then think again. Your title shouldn't be too long, too short, too vague or generic (as in "On some problems in group theory"), but should be the first approximation to contents of your paper. These are often contradictory constraints and there are no general rules which apply in all cases.

Some trickery is useful sometimes. Say, you introduce a some cumbersome class of permutations and give their asymptotic analysis. Give them a name! Say, these

permutations are inspired by Alice Munro's book. Call them *Munro permutations* right in the beginning of the paper and make the title "Asymptotic analysis of Munro's permutations". The reader may or may not find this title appealing enough to click on the article, but at least it conveys *some* sense of what's in the paper. In fact, if you don't actually like the name, you can denote this set A_n, and use the notation for the rest of the paper.

There are drawbacks in this approach. If others find the name useful they will always attribute the objects to Munro. For example, some years ago I introduced the *iterated Dyson's map*, and people are using it now without ever mentioning me. I lost that battle. Also, this approach might raise some eyebrows of the referees. At one point, my coauthor and I invented *Gayley polytopes* named after a street I lived on, and to rhyme with *Cayley polytopes* of which they were Generalizations to all Graphs G (get it?) The referee was annoyed, but we kept the name just because it's amusing and memorable.

Finally, let me self-quote the title naming advice I gave on `MathOverflow`, with some possibly useful examples of titles:[4]

> "You should emphasize not the length but the content. If you prove that all tennis balls are white make the title "All tennis balls are white". If you prove that some tennis balls are white, title your note "On white tennis balls", or "New examples of white tennis balls", or whatever. If your note is a new simple proof, and this is what you want to emphasize, make the title "Short proof that all tennis balls are white". If there was a conjecture that all tennis balls were white and you found a counterexample, use "Not all tennis balls are white". If you study further properties of white tennis balls, use "A remark on white tennis balls". You see the idea.
>
> "On the other hand, if you wrote a survey, it's important to emphasize that, regardless whether it's long or short. That's because this is a property of the content and style of presentation. For example, "A survey on white tennis balls" or "White tennis balls, a survey in colored pictures", etc. In fact, if your title is "A short survey on tennis ball colors", that would mean that your survey is short in content, as in "brief, incomplete", rather than in length – an important info for the reader to know."

3.3. **Abstract.** This is the easiest section to write. Just think of a short `MathSciNet` summary (not a longer more careful review they have sometimes). The abstract should have nothing personal, just dry fact about the results. State key results first and briefly mention the existence of others, including some generalizations, but no need for precise statements. Provide no details and no connections to other works unless absolutely necessary. Some journal guidelines advise not to include any citations, though I personally see no harm is writing "We disprove a conjecture stated by the author in [Pak12]," since this is more precise than "stated by the author in 2012" (is this the date of the idea? of the talk where the conjecture was first stated? of the `arXiv` preprint, or what?)

Either way, no need to worry about the abstract too much, but do put some minor effort into it. Remember – large fraction of `MathSciNet` reviews are just the abstracts, so make it clear, precise, plain and uninventive. As a rule of thumb, the number of lines in the abstract should be at 0.3–0.5 times the number of pages. An abstract with 10 lines for a paper of 10 pages looks way too excessive.

[4]See `https://mathoverflow.net/questions/81128`

3.4. **Table of contents.** Don't include it unless your paper is over 60 pages. But then you probably need a different style guide. Either way, `Adobe Reader` already has this feature and most people read papers in .pdf anyway. So skip on that.

3.5. **Introduction.** This is the hardest section to write. It's probably the only part of your paper that will be read by all but a few most devoted readers. If you have a senior coauthor, ask her or him to write this. If you don't, ask a senior colleague to read it and comment on your draft. Start writing your paper by writing the first draft of the Introduction, so you have an idea what's in the paper, and completely rewrite it after the rest of the paper is written. More often than not, the paper turns up differently than you initially imagined it. This could be for technical reasons, or since you proved more results, or now understand your own results much better than when you started writing. Then let the paper stew for a week or two while you show it to your closest and trusted colleagues, and after their comments on the contents of the paper rewrite it again, perhaps with a new emphasis.

To underscore the importance of the Introduction, here is a helpful quote by Rota [12], who gets things half-right in my opinion:

> "Nowadays reading a mathematics paper from top to bottom is a rare event. If we wish our paper to be read, we had better provide our prospective readers with strong motivation to do so. A lengthy introduction, summarizing the history of the subject, giving everybody his due, and perhaps enticingly outlining the content of the paper in a discursive manner, will go some of the way towards getting us a couple of readers."

As I explained earlier, the problem with this approach is that "nowadays" some people don't even have patience to read a long introduction. So what should you do? Well, make a lengthy introduction, Rota–style, extract a few pages to keep in the introduction and the rest put into *Final Remarks*. Alternatively, use a *Foreword*. More on these later.

What to include into the introduction: Start by setting up the problem and statements of the main results. If there are only a few technical definitions needed for these results – include them. If you are resolving a conjecture or a question – state it (with attribution). But do aim to have your first theorem on the first page, or at worst on the second page.

Sometimes this doesn't work. For example, there are too many details in the definitions, the theorems are long and cumbersome to state, the main result could be a bijection which takes very long to state, the context or the history is too long, etc. Sometimes it's a tradition in a particularly technical area. Well, I have seen the introductions with no stated results. They work only if they are short (under 1.5 pages), and the paper is itself a note (at most 10 pp).

The best way to get around stating technicalities in the main result is to skip the main theorem altogether and include interesting, nontrivial, but easy to state corollaries of main results. Cook them up if necessary and think of them as an advertisement of your paper, even if nobody ever cared to ask about this special case. This corollary is all that the paper passers-by will remember and when prompted can tell other people. If there is no such result in the Introduction, they remember nothing other than "perhaps this recent preprint is relevant to Q's work", which is much too weak as a clue to tell to Q.

What not to include: Technical definitions, examples, big figures illustrating some special cases, etc. Instead, whenever relevant and you feel like including them, use "(see Figure 5.1)" or "(see the exact definition in §3.4)" to get your point across. The interested reader will click on the link to take a look and use the `Back button` in the `Reader` to get back to the Introduction.

Also, ignore Rota's "giving everybody his due" advice – it's no longer applicable as stated. Most likely, there are too many relevant papers, so it's impossible to do this in the introduction and control its length. Instead, explain the history that's directly relevant to your main result. For example, "Paper [A] asked about XYZ and proved *weak*-XYZ. Last year, paper [B] showed that *strong*-XYZ is false. In this paper we refine the tools in [B] to show that XYZ is also false. We conjecture that the *weak*-XYZ is the strongest possible result in this direction. We also analyze the examples in [C] and show that..." You get the idea.

In the last paragraph or subsection of the Introduction, outline the structure of the paper in your own words. In the absence of the table of content, this helps the reader to navigate the paper and use section links visible in the `Adobe Reader`.

3.6. **Foreword.** If the Introduction is relatively short (say, under 3 pages), you are probably ok. But if you followed the rules above and it's still over 4 pages, that probably means your paper is quite long, you have too many results, and/or the paper spans several sub-areas of mathematics which all have plenty of relevant background. In this case you should divide your Introduction into subsections, and I suggest using Foreword as the first subsection. Think of it as a nontechnical introduction to your Introduction. Ordinarily, this function would be played by the Abstract, but we already mentioned that it's governed by its own very constraining rules.

Consider putting in the Foreword some highly literary description of what you are doing. If it's beautiful or sufficiently memorable, it might be quoted in other papers, sometimes on barely related subject, and bring some extra clicks to your work. Feel free to discuss the big picture, NSF project outline style, mention some motivational examples in other fields of study, general physical or philosophical principles underlying your work, etc. There is no other place in the paper to do this, and I doubt referees would object if you keep your Foreword under one page. For now such discussions are relegated to surveys and monographs, which is a shame since as a result some interesting perspectives of many people are missing.

Note: even if your paper is short, you can still get away with writing the first paragraph of the Introduction in this style. I suggest you use this opportunity. Nothing is less inspiring than a paper which starts "Let $G = (V, E)$ be a loopless graph on n vertices." Read [16] and other writing guides about how to write the first sentence. Note that since math writing tends to be so rigid, this is your only place to shine. Use it! It sets you aside as a better (math) writer.

In some sense, this is an exact reverse of Zinsser's quip [16, p. 21] :

> "It's amazing how often an editor can throw away the first three or four paragraphs of an article, or even the first few pages, and start with the paragraph where the writer begins to sound like himself or herself."

Think about how good we have it compared to non-fiction writers. We only have to be good writers, or at least "sound like ourselves", in the first few lines or few paragraphs. Skipping out on this is a missed opportunity.

3.7. **Final Remarks.** This is the least understood section, in my opinion. I feel that most people use it as a place to include a mix of open problems, examples, applications, references, whatever is left not included in the main part of the paper. The result is always like a paella – sometimes good, but you never know what are you going to find there. While the intention is right, for longer papers this lacks coherence and structure.

Let's start by explaining what this section is for.[5] It is really an expanded footnote section now usually called *endnotes*. Indeed, take any serious monograph in Humanities or Social Sciences. Or, for example, a brick sized presidential biography you can find in an airport bookstore. Or even the infamous *"Infinite Jest"* by David Foster Wallace, if you are into that kind of postmodern literature. In all of them, you will find at the end several hundred pages printed in smaller font, annotating the material in the main part of the paper, describing and quoting the original sources, providing additional context to material in the main part, etc. *My point is simple*: Final Remarks section must play exactly the same role.

The Final Remarks section should be neatly divided into unnamed subsections, each between one-two paragraphs and a page at the most (writing `\subsection{}` will produce a number without the title). The subsections need to have no relation to each other, but should be ordered in reverse order of importance. Typically, the first subsection would deal with expanded history of the subject, citing lots of other papers and "giving everybody his due". In the next subsection, write about where do you go from here, what potential applications of the results you are studying. Then mention your own more speculative conjectures, then other people's speculative conjectures, etc. Take your time and skip on nothing unless you want to be secretive about this kind of matters.

When writing the Introduction or the main part of the paper, whenever you feel there is a need for more explanation, context, related refs, etc., make a placeholder subsection in the Final Remarks, which you can fill in later. When done, go over the whole paper and insert "(for more on this, see §6.1)", "(cf. §6.4)", "We postpone the discussion on this until §6.8", etc. The interested e-reader will click on the internal link and read. The paper reader will flip pages. Then they will go back and continue reading. This is exactly how the endnotes work.

Over the years, I have seen many objections to this approach. "Devoting subsections to each paragraph makes this last section feel very disjointed," wrote one referee. This is correct but also misses the point – you are not supposed to read the Final Remarks in order. Sometimes referees and editors object to the size of Final Remarks section, which tends to grow rather large. There are several ways of dealing with this which are best used in combination.

First, you can extract some lengthy subsections and form new sections titled "Open problems" or "Historical overview" which would be placed right before the Final Remarks. Second, you can simply remove some of them to appease the journal, but keep the full length version on the `arXiv` or your webpage. Remember – nobody cares what version you publish in the journal as long as the theorems/proofs are the same. Third, you can preclude the objections proactively by changing the font size of the Final Remarks section to `\small` or even `\footnotesize`. I can see you squinting, but it's fine, really. Human eye is a fantastic instrument. If people can read *YouTube* comments on

[5] *Warning*: in CS Theory papers, the Conclusions section plays a different role (cf. [3]). Here we stick with the math paper traditions.

their phones in a crowded subway, they can understand your conjectures even when they are typed in 8 pt rather than 11 pt.

A side benefit of this is a dynamic semi-survey feature your paper achieves. Since the `arXiv` is easily updateable, you can continue adding new subsections to Final Remarks without changing anything else in the paper. This allows you to stake or communicate your new ideas even before you get an opportunity to write a new paper. For example, adding an outline of a solution of a (non-major) conjecture can help you fend off competition, a flexible version of "added in print" feature the traditional journals have.

3.8. **Acknowledgements.** Rota advises "Give lavish acknowledgments" and this is mostly correct. Let me expound on that a bit. You still need to make choices when you thank everyone. Make a short unnumbered section at the end of the paper, right before the References. It's ok to have it in a smaller font.

Proceed in the order of increasing importance. First, thank everyone you remember discussing the results in the paper whether or not you remember them actually giving helpful suggestions. Do this by name, in alphabetical order, as in "We are grateful to Freddie Mercury, Ozzy Osbourne, Axl Rose and Stephen Tyler for helpful conversations". Note that it's ok to use nicknames and commonly used names rather than full legal names (in this example, Freddie is actually Farrokh, Axl's real name is William, and "Ozzy" is a nickname). When in doubt, consult emails from a person and use the name as it appears in the sign off line.

Second, single out people who gave you some really helpful suggestions. Say so, as in "We are especially thankful to Adam Smith for informing us about his *invisible hand* Lemma; this result played a crucial role in our laissez-faire equilibrium analysis in Section 3, leading to a much shorter proof of Theorem 3.4." Do this one by one, for all such people, in any order (the explanations stand for themselves). Some guides advise you to email each person to ask permission before thanking. Ignore that. Ask by email for permission if you are using any private information, like e.g. the conjecture this person once told you. Do not get upset if you receive a reply "please remove that conjecture, I am still working on it". Do as requested then.

Conclude with thanking the institutions that hosted you which facilitated collaboration (if any), and then all granting agencies supporting you.

4. References

4.1. **Why so important?** Really, are the references important enough to warrant a separate section in this guide? **Absolutely!** In fact, I always felt this is self-evident, but apparently it's not so to everyone. Even though this guide is more "How-to" than "Why?", this deserves an exception.

Once I already wrote an answer to this question on my blog, but that may have been buried in the nature of that blog post. So please forgive me for quoting myself again:[6]

[6]See `https://igorpak.wordpress.com/2014/09/12/how-not-to-reference-papers/`

> "First, let's talk about something obvious. Why do we do what we do? I mean, why do we study for many years how to do research in mathematics, read dozens or hundreds of papers, think long thoughts until we eventually figure out a good question. We then work hard, trial-and-error, to eventually figure out a solution. Sometimes we do this in a matter of hours and sometimes it takes years, but we persevere. Then write up a solution, submit to a journal, sometimes get rejected (who knew this was solved 20 years ago?), and sometimes sent for revision with various lemmas to fix. We then revise the paper, and if all goes well it gets accepted. And published. Eventually.
>
> "So, why do we do all of that? For the opportunity to teach at a good university and derive a reasonable salary? Yes, sure, to some degree. But mostly because we like doing this. And we like having our work appreciated. We like going to conferences to present it. We like it when people read our paper and enjoy it, or simply find it useful. We like it when our little papers form building stones towards bigger work, perhaps eventually helping to resolve an old open problem. All this gives us purpose, a sense of accomplishment, a "social capital" if you like fancy terms.
>
> "But all this hinges on a tiny little thing we call citations. They tend to come at the end, sometimes footnote size and is the primary vehicle for our goal. If we are uncited, ignored, all hope is lost. [..] So for anyone out there who thinks the references are in the back because they are not so important – think again. They are of utmost importance – they are what makes the whole system work."

In summary, everyone pays attention as to where and in what way they are cited. Treat this as carefully as you can. This costs almost nothing, takes relatively little time and effort, can have a minor upside of making some people happier, but can also have a great downside if done poorly (people do hold grudges, sometimes for decades).

4.2. **How to cite a single paper.** The citation rules are almost as complicated as Chinese honorifics, with an added disadvantage of never being discussed anywhere. Below we go through the (incomplete) list of possible ways in the decreasing level of citation importance and/or proof reliability.[7]

(1) "*Roth proved Murakami's conjecture in* [Roth]." Clear.

(2) "*Roth proved Murakami's conjecture* [Roth]." Roth proved the conjecture, possibly in a different paper, but this is likely a definitive version of the proof.

(3) "*Roth proved Murakami's conjecture, see* [Roth]." Roth proved the conjecture, but [Roth] can be anything from the original paper to the followup, to some kind of survey Roth wrote. Very occasionally you have "*see* [Melville]", but that usually means that Roth's proof is unpublished or otherwise unavailable (say, it was given at a lecture, and Roth can't be bothered to write it up), and Melville was the first to publish Roth's proof, possibly without permission, but with attribution and perhaps filling some minor gaps.

(4) "*Roth proved Murakami's conjecture* [Roth]*, see also* [Woolf]." Apparently Woolf also made an important contribution, perhaps extending it to greater generality, or fixing some major gaps or errors in [Roth].

[7] *Note*: this list should be taken with a grain of salt, since others might have a somewhat different POV in each particular case.

(5) "*Roth proved Murakami's conjecture in* [Roth] (*see also* [Woolf])." Looks like [Woolf] has a complete proof of Roth, possibly fixing some minor errors in [Roth].

(6) "*Roth proved Murakami's conjecture* (*see* [Woolf])." Here [Woolf] is a definitive version of the proof, e.g. the standard monograph on the subject.

(7) "*Roth proved Murakami's conjecture, see e.g.* [Faulkner, Fitzgerald, Frost]." The result is important enough to be cited and its validity confirmed in several books/surveys. If there ever was a controversy whether Roth's argument is an actual proof, it was resolved in Roth's favor. Still, the original proof may have been too long, incomplete or simply presented in an old fashioned way, or published in an inaccessible conference proceedings, so here are sources with a better or more recent exposition. Or, more likely, the author was too lazy to look for the right reference so overcompensated with three random textbooks on the subject.

(8) "*Roth proved Murakami's conjecture* (*see e.g.* [Faulkner, Fitzgerald, Frost])." The result is probably classical or at least very well known. Here are books/surveys which all probably have statements and/or proofs. Neither the author nor the reader will ever bother to check.

(9) "*Roth proved Murakami's conjecture.*[7] Footnote 7: *See* [Mailer]." Most likely, the author never actually read [Mailer], nor has access to that paper. Or, perhaps, [Mailer] states that Roth proved the conjecture, but includes neither a proof nor a reference. The author cannot verify the claim independently and is visibly annoyed by the ambiguity, but felt obliged to credit Roth for the benefit of the reader, or to avoid the wrath of Roth.

(10) "*Roth proved Murakami's conjecture.*[7] Footnote 7: *Love letter from H. Fielding to J. Austen, dated December 16, 1975.*" This means that the letter likely exists and contains the whole proof or at least an outline of the proof. The author may or may not have seen it. Googling will probably either turn up the letter or a public discussion about what's in it, and why it is not available.

(11) "*Roth proved Murakami's conjecture.*[7] Footnote 7: *Personal communication.*" This means Roth has sent the author an email (or said over beer), claiming to have a proof. Or perhaps Roth's student accidentally mentioned this while answering a question after the talk. The proof may or may not be correct and the paper may or may not be forthcoming.

(12) "*Roth claims to have proved Murakami's conjecture in* [Roth]." Paper [Roth] has a well known gap which was never fixed even though Roth insists on it to be fixable; the author would rather avoid going on record about this, but anything is possible after some wine at a banquet. Another possibility is that [Roth] is completely erroneous as explained elsewhere, but Roth's work is too famous not to be mentioned; in that case there is a followup sentence clarifying the matter, sometimes in parentheses as in "(see, however, [Atwood])". Or, perhaps, [Roth] is a 3 page note published in *Doklady Acad. Sci. USSR* back in the 1970s, containing a very brief outline of the proof, and despite considerable effort nobody has yet to give a complete proof of its Lemma 2; there wouldn't be any followup to this sentence then, but the author would be happy to clarify things by email.

4.3. **How to cite a list of papers.** It is a disservice to the community to write "See [2–19] for some relevant work", as I see in some sloppy papers. This helps neither the readers to find anything at all, nor the authors of 2–19 to get credit; in fact, it pits them

against each other by unfairly equalizing their research contributions. What you should do is go over the papers individually, starting with the most important reference, and describe their contribution. Stop when you are tired. Here is an example of how this works:

> Our Theorem 3 studies partitions into *even primes*. A strongly related result in [A] studies partitions into into *primes which are powers of two*. This paper builds on the tools in [B], which studies partitions into *primes which are squares*. In a different direction, partitions into *even Fermat primes* have been studied in [C,D], and more recently in [E], which under the GRH established Wigner's semicircle law in this case. In an unusual development, last year [F] proved uncomputability of the partition function into *odd perfect numbers*, but the paper remains under scrutiny (my colleague Kiran X. has been unsuccessful in independently verifying the result). We should also mention a series of papers [G1,G2,G3] on partitions into *prime Catalan numbers*, which successively improves on the classical bound in [H] by using...

As we mentioned earlier, this type of description should go into Final Remarks. In the example, the Introduction should have only [A] and [B] papers mentioned.[8]

4.4. **Where to cite a paper?** That's actually easy. Only the most relevant papers should be cited in the Introduction. The rest are cited in the Final Remarks. Nothing is cited in the main part of the paper. The reason is simple – there is no way to be both fair, clear and complete when citing in the middle of the paper, and without breaking the flow of the arguments. *Major Exception*: you are using somebody's result as a lemma in the proof. Then a precise reference is required, but don't elaborate on it. If you want to tell a story of this result, who did what towards it, etc. – make a separate Remark *after* the proof, or postpone it until Final Remarks.

There is also the dynamic updating advantage in this approach. Once you make your paper public, inevitably your colleague XYZ will send you a barely related recent paper you may want to cite since you actually like XYZ and respect his opinion. Or perhaps PQR tells you about some simple special case she worked out and published as a lemma back in 1995. While PQR's tools do not extend to your most general version, you should definitely cite it and explain why they don't. In all these cases, you know exactly where these updates go in the paper.

4.5. **Forming your reference.** Whenever a specific result is quoted in a paper over 5 pages or a monograph, one should include a specific location of that result. Traditionally this is done with either page numbers or names of the results in the published version, as in "Theorem 3.2" or "Eq. (3.12)". Unfortunately, there could be many preliminary electronic versions of the paper, or sometimes post-publication versions, say after the authors fix some minor errors. How do you cite then? Of course, the `arXiv` paging is completely off, but numbering can also shift sometimes, say because a definition is moved from one section to another. So I recommend using section and subsection numbering in the `arXiv` version, since they tend to be the most stable, as in "see [A, §3.1]".[9]

[8] *Warning*: Again, CS Theory style is different. In the example above, all of it goes into the Introduction there, possibly condensed for space.

[9] This is explicitly contradicting many standard guidelines by the publishers. See e.g. *Elsevier*: `http://tinyurl.com/wlr73as`.

4.6. **Style of references.** First and foremost, do NOT use BibTeX. Even if you have hundreds of references, or especially if you do. `MathSciNet` citations tend to be bloated with unnecessary details and have inconsistent style and author's spelling. For example, some papers have "Paul Erdős", while others "Erdös, P.", making it unclear if that's the same person. You should pick a style and stick to it. Make each reference as concise as possible, yet complete enough to be found. For example, omit names of book series, issue numbers, etc. These are all redundant. Physics paper are champions at this; they even omit the titles. I feel it's going too far.

Second, for longer papers use alphanumeric style, as in [SY09,Woo96]. It helps the reader to know the name and date of the publication. For unpublished papers, use [Con17+]. For papers with 5 or more authors, use [A+13] in place of [ABCDEF13]. For papers with the same names and dates use [Tra15a], [Tra15b], etc.

Thirdly, this is perhaps obvious, but *do no emulate Knuth*. While he strives for perfection, you need to worry about clarity. No need to have fully expounded second names which make it more rather than less confusing (as in "Richard Peter Stanley"). Similarly, skip Chinese names written in characters, Russian titles in Cyrillic, etc. Keep it simple.

Finally, for unpublished papers include the `arXiv` number, or a link to the free version if that's not available. Use `hyperref` and `url` packages to make weblinks clickable.

5. Micro tips

5.1. **Basic grammar, syntaxis, punctuation, etc.** The existing guides do a good job. Start with [3, 4, 11, 15] and the first section of [8], which are all relatively short. Then continue with [5] and [9] which are quite different, but both very thorough. We will not repeat these rules here.

5.2. **Downshift your style.** Berndt [2] advises: "You should aspire to the same literary levels as William Shakespeare, J. K. Rowling, and Leo Tolstoy". I strongly disagree.[10] Your audience will likely include readers with limited English, short attention span and no patience. Thus it is perfectly ok to be repetitive and have ten therefore's and be clear, rather than vary the style and have difficult words like "henceforth". In fact, it is a good style to remind the reader of notation you introduced earlier, and skip some easy logical steps confident that the reader with catch up.[11]

In general, outside of the Introduction and Final Remarks, use only present indefinite and occasionally past indefinite. Write short sentences. If you must have longer sentences, separate each clause with commas. Ignore whatever punctuation rules you learned in *Strunk & White*. Instead, use the following rule of my own: *commas exist to make the sentence structure clearer*, rather than to indicate where the English speakers

[10]In response to this writeup, Bruce Berndt writes: "I think that the meaning that I was trying to convey is different from the interpretation that you gave. These writers wrote in a captivating style which engaged readers. They thought deeply about how and what they were going to say. Writers of mathematics should adopt these same principles. Clarity and a captivating style are very important" (personal communication). I completely agree with that, of course.

[11]To make further contrast with Tolstoy's writings, we advise to express yourself concisely rather than on hundreds of pages, which can include side philosophical discussions and page long foreign language quotations. To put the shoe on the other foot, Tolstoy himself couldn't even get the basic arithmetic right; see e.g. his Gematria style calculations in "*War and Peace*": `http://www.jukuu.com/portal/novels/warpeace/war09/ewar19.htm`.

make a pause when reading. An example: "Consider an integer x, whose square is 4, and which is positive." In other words, every time you make a clause, subclause, etc., put with commas on both sides. Ignore the Oxford comma debate, and place commas whichever way you think make the sentence structure clear. Do not use a semicolon – it implies a logical connection between the clauses, which is best made explicit between the sentences. Do not use triple dashes as in "obsessed by this fear" Zinsser's quote above (they are distracting). Instead, use double sashes separated by single spaces from the text as in two sentences above.[12]

Always aim for the most common spelling of the words, which tends to be in Standard American English. For example, do not use "coöperation" or "colouring" no matter how much you like *The New Yorker* or the *BBC*. Sometimes it's a close call, like "dominoes" vs. "dominos". Let `Google Search` break the tie.

Finally, if you introduce a mathematical term, do not use its variations even if to you they make perfect sense. For example, if you define a property "nice graph", then do not use "niceness of graphs imply..." Stick to basic forms. This all may be hard to accept to native English speakers. I feel it's a small price to pay for not having to master a foreign language.

5.3. **LaTeX tips.** Create a *ton of macros*. For everything. For example, I have `\al` for `\alpha`, `\rT` for `\textrm{T}`, `\lra` for `\leftrightarrow`, etc. Remember them and always use the same macros in all your papers. Note that the `arXiv` keeps LaTeX files for all papers, so you can check out and copy macros of other people.

There are several advantages of this system. Obviously, it's faster to type even though it's a bit less logical. More importantly, if you don't like your notation and want to change, say, every α into γ, you just replace the definition of `\al` and you are done. Before the change, you might also check if you have any γ in the text - just comment out the definition of `\ga` and see if the file compiles. That's all there is.

Warning: not all letters are created equal. For example, letters Ξ, ι, ϖ, $\imath$, $\jmath$ and $\varkappa$ look weird, so don't use them unless you are out of options. Letters κ and υ look too much like k and v; you are better off using English letters and playing with the fonts. Make sure to use $\varnothing$ in place of $\emptyset$, as the latter looks too much like a computer zero. Similarly, always use ℓ in place of l, and ε in place of ϵ. While I personally favor φ over ϕ, both letters look nice and sufficiently different, so can be used in the same paper to mean different things.[13]

It pays to create mnemonics for what letters mean. For example, I like to have letters with "tails" to be functions: $f, g, \varphi, \phi, \zeta, \xi, \eta, \rho$. English letters early in the alphabet are constants: a, b, c, d, etc., while letters at the end are variables: x, y, z. Letters ℓ, m, n are always integers, while whatever is left can be anything – integers, vectors, variables, etc.: p, q, r, s, t, u, v.

When you are unsure about the notation, create a macro placeholder and play with different fonts once the paper is finished. Doing this at the end helps to avoid similar looking letters to have completely different meanings. Do not be afraid to play a little

[12]These are very mild suggestions when compared to the *English-language spelling reform* proposals (see e.g. Anatoly Liberman's blog `https://blog.oup.com/category/series-columns/oxford_etymologist/`.

[13]In general, Greek letters have so many different meanings, you can't keep track. See e.g. this helpful list: `https://en.wikipedia.org/wiki/Greek_letters_used_in_mathematics,_science,_and_engineering`

with sizes of the letters if that helps to emphasize the difference. For example, here are letters P written in many different fonts, paired with the same letter of smaller size:[14]

$$\mathbf{P}\,\mathbf{P}\quad \mathbb{P}\,\mathbb{P}\quad \boldsymbol{P}\,\boldsymbol{P}\quad \mathcal{P}\,\mathcal{P}\quad P\,P\quad P\,P\quad \mathrm{P}\,\mathrm{P}\quad \mathtt{P}\,\mathtt{P}\quad \mathtt{P}\,\mathtt{P}\quad \mathrm{p}\,\mathrm{p}\quad p\,p$$

Similarly, feel free to use letters from other alphabets, e.g. ב and ג look sufficiently pretty and there is no reason they must always denote large cardinals. *Warning*: avoid Gothic fonts. No one wants to distinguish $\mathfrak{G}$ vs. $\mathfrak{S}$, or $\mathfrak{F}$ vs. $\mathfrak{T}$, not to mention that copying them on the board is a nightmare. If you feel you need more choices, there are many other fonts and alphabets available on the web – download and play with them until you are happy.

If your paper is short or you need to tag a local equation which you want to emphasize, feel free to use some unusual LaTeX symbols, e.g. $(*)$, $(**)$, $(\diamond)$, $(\divideontimes)$, $(\circledast)$, $(\lozenge)$, or even $(\heartsuit)$ if you especially like that formula. Why be constrained by old conventions?

5.4. **Do NOT trust LaTeX.** We all know that LaTeX does a great job laying out the formulas, so there is a tendency not to second guess what it outputs. This is a wrong attitude. Sometimes, LaTeX imperfections can lead to an ambiguity or a general lack of clarity. In practice, you need to take care of the spacings sometimes, as in this example:

$$\frac{\Phi(2a+c)\Phi(c-2a)\Phi(a)\Phi(3a)}{\Phi(a+c)\Phi(c-a)\Phi(2a)\Phi(4a)}$$

Here the confusion comes from the unclear role of Φ, which looks like a strange $\oplus$-like symbol separating linear forms. But once the "`\,`" are inserted between the terms, it becomes clear that we are taking a product of multiple values of $\Phi(\cdot)$:

$$\frac{\Phi(2a+c)\ \Phi(c-2a)\ \Phi(a)\ \Phi(3a)}{\Phi(a+c)\ \Phi(c-a)\ \Phi(2a)\ \Phi(4a)}$$

Similarly, an integer partition $12 + 10 + 10 + 2 + 1$ when written in a short form becomes 1210^221, which is rather difficult to parse. In fact, it looks more like a partition $1210 + 1210 + 21$. Again, inserting spaces does the job: $12\,10^2\,2\,1$.

5.5. **Figures.** Make a lot of them, but make them small. People have tendency to make BIG figures, sometimes with too many details. Having lots of such figures can distract the reader as they break the natural flow of the paper. Let me repeat – human eye is a very good instrument. If you can read my references you should be able to understand a figure the size of 3-5 cm. Whoever can't, will figure out how to use `Ctrl+` and view them bigger. If you think you need all that many details, create several copies of the same picture, align them next to each other and add different levels of details on each copy. If you *must* have a huge, over half a page size figure, consider moving it to the appendix of the paper, or even better to your website, and add a link directing the reader straight there.

[14]All these fonts are different and standard in LaTeX. It's ok if you can't name them all. But if you can't even tell them apart, train yourself. Perhaps, watch [6] first.

6. What to do when you are done

6.1. **You are never really done.** When you finish writing the paper, let it stew for a week or so. Do something else in the meantime. Come back to it then and reread all of it with a fresh eye, looking for typos, bad wording, imprecise statements, etc. Don't put it on the `arXiv` right away, but instead put it on your website and email about it a few colleagues. Not only they might give you some helpful mathematical advice and point out some missing references, they might also help you with the typos. A week later, once you attend all these, it's ok to put it on the `arXiv`. But even afterwards, every time you find or are told about a typo, keep updating the paper on your own website and occasionally on the `arXiv`. You can (and should) do this even after the paper is published. But be careful: if the published version has a wrong proof of a lemma, or is missing a key reference (possibly, you are rediscovering somebody's result), you can't whitewash it in the updated version. These warrant a separate file titled "Erratum" or "Acknowledgement of attribution".

6.2. **Advertise and popularize.** There are several ways you can make the paper more accessible to larger audiences. This is a little area dependent, but let me suggest just a few. If you ever give a recorded talk on the paper, make a link on your website to the talk right next to the paper. Keep the `.pdf` file with slides alongside.[15] Try to explain the gist on your blog, even if it's mostly dormant and has no subscribers. Again, link the blog entry to your website alongside the paper.

There are a few more things you can do, all of them harder. First, write a survey on the subject. If you want to do a good job, this takes a lot of hard work. But you can learn a great deal and are able to place your results into context where they would be better understood. According to Rota, "You are more likely to be remembered by your expository work" [12]. While I don't fully agree, there is a lot of truth to this.

Second, if at all possible, consider the simplest yet still nontrivial special case of your result. If you can extract a little result which can be stated in the language high school or college students can understand and possibly solve by a direct argument, make it into an olympiad style problem. Send it to somebody on your national math olympiad committee (which comprise IMO), or the Putnam committee. They are always looking for good and unusual problems, and many problems are in fact created this way.

Third, if the problem you extract is too difficult for the Putnam Competition, but your argument in this case substantially simplifies, write it up in a separate note. Try to make it accessible to undergraduates. Submit it to *Amer. Math. Monthly*, *Math. Intelligencer*, *Math. Magazine*, or another journal targeting students and teachers. If you explain at the end how this special case fits into your paper, this will clarify and bring more exposure to your work.

6.3. **Rewrite and republish.** Finally, suppose you already wrote and published the paper. You think it's being ignored because it's unclear. Well, rewrite it. Simplify all the arguments, maybe generalize the main result a bit, find some new examples or applications. And publish this new version, even if just on the `arXiv`.

Rota advises to publish the same paper several times, and gives some famous examples [12]. While I disagree with this practice, I think there is a room for compromise. Say, for example, in your latest paper the argument is cumbersome because you are

[15] See by blog post what to do about giving talks and what to do with the videos: `https://igorpak.wordpress.com/2015/05/02/you-should-watch-combinatorics-videos/`

overgeneralizing. It is then ok to start with a proof of a known theorem "for completeness" and then explain how this proof needs to be modified to work in greater generality. That way you get the best of both worlds: a clear presentation of your own old result and a clear pathway to your new generalization.

Acknowledgements. I am grateful to Bruce Berndt, Steven Krantz, Bjorn Poonen and Adam Sheffer, for useful remarks and help with the references. I am especially grateful to Laci Babai for teaching me to distinguish different types of references (all misunderstandings are mine, of course). Special thanks to all my students, postdocs and collaborators who helped me to formulate these simple rules. The author was partially supported by the NSF.

References

[1] S. Aaronson, *Ten Signs a Claimed Mathematical Breakthrough is Wrong*, blog post from January 5, 2008; available at `http://www.scottaaronson.com/blog/?p=304`

[2] B. C. Berndt, *How to Write Mathematical Papers*, `http://www.math.uiuc.edu/~berndt/writingmath.pdf`

[3] O. Goldreich, *How to write a paper*, preprint; `http://www.wisdom.weizmann.ac.il/~oded/writing.html`

[4] P. R. Halmos, *How to Write Mathematics*, AMS, 1973, 64 pp.; included also essays by N. E. Steenrod, M. M. Schiffer and J. A. Dieudonne.

[5] N. J. Higham *Handbook of writing for the mathematical sciences*, SIAM, Philadelphia, PA, 1998, 302 pp.

[6] G. Hustwit (dir.), *Helvetica*, Plexifilm, 2007, 80 min.

[7] S. P. Jones, *How to write a great research paper*, talk available at `https://tinyurl.com/y7p3aoj7`

[8] D. E. Knuth, T. L. Larrabee and P. M. Roberts, *Mathematical Writing*, MAA, Washington, 1989, 115 pp.; see also course notes at `https://tinyurl.com/y9mm4xcq`

[9] S. G. Krantz, *A Primer of Mathematical Writing*, `arXiv:1612.04888`, 276 pp.

[10] I. Pak, Partition Bijections, a Survey, *Ramanujan Journal* **12** (2006), 5–75.

[11] B. Poonen, *Practical suggestions for mathematical writing*, `https://math.mit.edu/~poonen/papers/writing.pdf`

[12] G.-C. Rota, Ten lessons I wish I had been taught, *Notices AMS* **44** (1997), 22–25.

[13] R. P. Stanley, *Enumerative Combinatorics*, vol. 2, Cambridge Univ. Press, 1999.

[14] T. Tao, *On writing*, blog post; `https://terrytao.wordpress.com/advice-on-writing-papers/`

[15] D. B. West, *The Grammar According to West*, `http://www.math.illinois.edu/~dwest/grammar.html`

[16] W. Zinsser, *On Writing Well: The Classic Guide to Writing Nonfiction* (30th Ed.), Harper, 2016, 336 pp.