Writing is not a soft skill

Agelos Georgakopoulos

Mathematics Institute, University of Warwick
CV4 7AL, UK

January 25, 2023

Perhaps you are familiar with this situation: you are thinking about a problem with a colleague, and after filling some blackboards with figures you decide that you have solved it. You agree to write it up as a paper, but postpone it in order to think about a different problem. After all, solving problems is fun, but writing is not.

The last sentence could not be any further from the truth, and a very harmful attitude for your career. The purpose of this text is to explain why, and help you enjoy writing and do it efficiently.

Writing a research paper can be fun, but I am not claiming it is easy. Quite the contrary, it is a very challenging endeavour. You have to explain something highly complicated to people of various backgrounds. To make matters worse, you are under time pressure due to other duties, like teaching or other papers waiting to be written.

Let us get the time pressure illusion out of the way. The most prolific mathematicians of all time, say Euler or Erdős, wrote about 30,000 pages in total. Writing with co-authors you can achieve the same by writing about one page per day on average! With practise and the right tools, some of which I hope you may acquire by reading on, you will be able to write a page in an hour or two, leaving enough time to solve problems, read others' papers, and prepare high quality lectures, without ever rushing.

First of all, choose your projects carefully. I mentioned Euler and Erdős above, but most great mathematicians actually wrote much less. Bill Thurston left us less than 80 papers, but they changed mathematics forever. My department often appoints a mathematician to a permanent position with just a handful of very strong papers, while others that have written dozens of less impactful ones are not invited to an interview. Hiring committees and prize juries will never ask how many papers you have, but they will ask you what your best paper is. The volume of your output matters, but its quality matters more. Thus before starting to work on a project, ask yourself if it is worth spending a few months of full-time work writing a paper about it. If the answer is yes, take the time it needs to write it well. Nobody will want to read a paper that you wrote in a rush, and **it is the papers that are read that matter**, not the papers that are written. Moreover, rushing does more to slow you down than speed you up, as you pay for it in the revision stage.

This raises the question of *how* to choose a worthwhile project to work on. There is no algorithm, yet there is a simple answer: **read a lot of papers, and attend many talks**. This will help you become aware of what interests others and what results are potentially useful. On the contrary, if you ignore others, they will ignore you too; not in retaliation, but just because having common interests is a symmetric binary relation.

Having got rid of the time pressure, let me now address the other major cause of difficulty: sitting down to type up the paper too early, i.e. before you really have a proof. Most colleagues agree that '*proving is more fun than writing*', but somewhat paradoxically, cut the former short by starting to write an incomplete proof. This ensures that writing is no fun, as it becomes the stage where one notices gaps or mistakes, and realises that more work is needed than originally believed. One ends up doing most of the real work in front of a computer, where distractions loom. Moreover, much time is spent deleting and rewriting after mistakes are found in other parts of a proof.

There is a solution to this that sounds simple, though it takes experience to learn how to implement: make sure that the details add up before you start typing. We all know that it is essentially impossible to be sure about the correctness of a proof even after writing it up, yet we learn to **recognise the key moment** when *proving* is complete, and it is time to *write*. Check that all details of your proof are sound, by writing an outline on your notepad. This outline should not be too sketchy; it should be a complete formal proof up to straightforward details.

The principle of proving, writing a sketch, then typing, need not be applied to the paper as a whole; it is a helpful principle to apply piecewise. Suppose you have a strategy for attacking a problem that requires an important intermediate lemma. Once you believe you can prove this lemma, fill all its details in your notepad, then type it. This is often more efficient than waiting for the rest of the proof to be completed, for several reasons. Firstly, the latter may happen months later, by which time you may find it hard to remember the details of the proof of your lemma. Secondly, you will often realise in the process that your lemma is not exactly correct. It may be completely wrong, or in need of a refinement. Having got it right will not only help you avoid wasting time trying to prove something you cannot use, it may point you to the right direction.

This last remark brings up another important idea: **the distinction between proving and writing is fuzzy!** Details that materialise in the stage of writing clarify what is actually proved, and lead to strengthenings, generalisations and corollaries that may be more interesting than the original results envisioned. But although some fuzziness is unavoidable and productive, we should be trying to contain it. According to a popular anecdote, Jean Bourgain used to keep his desk empty apart from a pile of empty sheets of paper and a pile of hand-written paper. When the latter became full, he would give it to his secretary to type up. Whether this is true or not, you may want to remember this picture as an ideal situation you could be striving for (accepting that it cannot be reached).

Like every complex construction, **writing a paper requires careful planning**. Once the key moment has arrived to start writing, do not rush to move from the notepad to the computer. Instead, *plan* the structure of your write-up. The longer your proof is, the more important it is to split it into pieces that can be read independently. A good example is the proof of the Robertson–Seymour Graph Minor Theorem. This monumental theorem was proved in a series of twenty papers spanning over 500 pages [3]–[4]. But many of these twenty papers provide intermediate steps that are stand-alone results with independent impact. Further examples come from the many famous theorems that are called *lemmas* because their authors used them to prove something else, but they had taken the time to structure their proofs well. If you have a long proof, try to identify any elements that can stand alone as a lemma. This may increase the impact of your paper, and it will definitely ease the work of the referees and readers. Bear in mind that only a small fraction of your readership will want to read everything, while others will want to see a rough overview of your proofs. Give the latter opportunities to skip the technical parts by keeping the structure transparent and offering high-level overviews.

At this stage of structuring your write up, it is also important to ask yourself if you have found the best way to prove your result. Could there be a more elegant approach, potentially leading to a stronger result? Can any of the intermediate steps be improved to yield a stand-alone lemma? Apart from leading to a better paper, asking these questions will help you realise if the key moment has really come. If you start writing too early, you will not be able to make up your mind on whether your proof is best possible.

Once you have enough details of your proof to convince yourself, it still takes effort to convince the readers. The major difficulty is putting yourself in their shoes: your proof looks easier to you than it is because you know it too well. Imagine having to guide a stranger (the reader) to your bedroom (the proof) in complete darkness. You have a clear mental image of the room, but the reader is in the dark. But you are holding a torch, with which you can illuminate parts of the room one by one. Start by illuminating the corners to give them an overview of the shape, then the major furniture, then the details. **Your task is not just to *write* your proof, but to *explain* it.**

A common mistake is to describe the difficult part of an argument well enough while neglecting to say how it combines to yield the desired conclusion. No matter how well you illuminate the bed, the stranger will not be able to see its position if they are not shown where the walls are.

The aforementioned mistake is an example of a more general one, namely not paying enough attention to ‘trivial’ elements. To explain this, let me use an example from another highly demanding mental activity. If you look up images of former chess world champion Garry Kasparov, you will notice a strange coincidence: his knights are always facing left. This must be intentional, as

Figure 1: Kasparov's knights always face left.

tournament organisers usually place the knights facing each other when setting up the boards, and it calls for an explanation. Some attribute it to superstitions, but I think there is a better explanation: the pattern-matching task of recognising a figure of a knight becomes slightly easier if that figure is always the same, and this may be why some chess players arrange their knights with purpose. Players at this level are able to win against a dozen of strong opponents playing simultaneously and blindfolded. It may surprise you that they pay attention to such a trivial detail as the direction of their knights, yet when one such player faces another, they would do anything to relieve their brain from unnecessary hurdles. For a similar reason, some players will respond to a draw offer with a shake of the head rather than by saying 'no', to avoid activating their brain's language centre.

Whether you are convinced by the above explanation or not, these examples convey an important lesson: **get the 'trivial' obstructions out of the way!** Keep the logical structure of your proof perfectly clear, e.g. by using expressions of the form *'this completes Case I, and we now proceed to Case II'*. For each conclusion made, say clearly why it is true, e.g. *'by Lemma 3.4 and the pigeonhole principle'*. Avoid combining two logical steps or calculations in one. Use precise sentences. I remember when helping me with my first paper, my advisor (Reinhard Diestel) rewrote the sentence *'The set X contains...'* as *'The set X consists of...'*. After the initial shock, I recognised that the latter is slightly more precise, and learned a lot from it. Precision is not achieved by using more words, but by using them carefully. You may ask: how much time is it worth investing to make all those tiny improvements to a paper? The answer is: a lot of time for your first paper; with practise it will become your second nature and you will be writing the precise version at the first go.

Getting the trivial obstructions out of the way is not only helpful for your readers, but also for you: if you are writing down a lemma to be used in the most complex part of your proof, state your lemma as cleanly as possible before starting with the latter. If knights facing right bothered Kasparov, imagine what notification pop-ups from your devices do to you. This is another reason why you want to do most of your proving, and some of the writing, on your notepad rather than your computer. By the way, I insist that my notepads have blank white pages.

The issue of precision brings up another important question: one often wonders whether to present an argument by formal or colloquial language. Writing with precision does not mean imitating the style of a Volkswagen instruction manual. The aim is to be both precise and elegant. As an example, here is an excerpt from a proof of Zeeman [6, Chapter III, p. 261]:
"choose a spine in the interior; expand each edge like a banana and collapse from one side; then expand each vertex like a pineapple and collapse from one face".
This evocative description does not lack precision, and its elegance helps the reader immediately *see* what the author means more efficiently than any formal definition would have done. As Halmos [1] put it, **"Clarity is what's wanted, not pedantry; understanding, not fuss"**. Elegance is not just an aesthetic matter, you want to offer your readers a mental picture that they will remember, not a cumbersome definition that they will have to re-read several times.

There will be cases where the notion or picture you want to explain is complicated, and you do not see a way to present it in an elegant and precise way. In this case you could present it in two different ways, or you could offer a colloquial overview followed by a precise definition and an example.

Having discussed the basic principles, let me comment on the structure of a paper. I usually first write the results and proofs, then the introduction, then the preliminaries. Before you start writing your first paper, ask your advisor for a copy of their templates, and for pointers to well-written specimens in your area (I would recommend [2]). Imagine that your uncle offered you to work in his tavern. You would not tell him that you are going to bring your own chairs because you prefer the colour. Likewise, try to **profit from the accumulated experience of centuries of mathematical writing**.

The *Introduction* is used to announce the main results of your paper, but also to motivate them. As Persi Diaconis put it, 'the big questions in mathematics are **so what?** and **who cares?**'. Do not underestimate the latter, which is the reason why anyone would want to read your paper, or why a journal would want to publish it. The standard way to give a convincing answer is by appropriate citations. Ideally, you want to cite several recent articles published in quality journals that are related to your results. (And this is something that can help you decide which topic to write about before you start.) If you only cite the textbooks and the couple of papers that you absolutely need, your editors may think that your paper will not interest anybody, and they would struggle to find willing referees.

The introduction should get straight to the point, making it clear that something substantial is achieved in the paper. Sometimes this is challenging, as the results may involve a lot of terminology. A potential solution is to state the results in rough terms, and postpone the details to later sections. An example

Three-manifolds are greatly more complicated than surfaces, and I think it is fair to say that until recently there was little reason to expect any analogous theory for manifolds of dimension 3 (or more)—except perhaps for the fact that so many 3-manifolds are beautiful. The situation has changed, so that I feel fairly confident in proposing the

1.1. CONJECTURE. *The interior of every compact* 3-*manifold has a canonical decomposition into pieces which have geometric structures.*

In §2, I will describe some theorems which support the conjecture, but first some explanation of its meaning is in order.

Figure 2: Thurston's geometrization conjecture as stated in his introduction [5].

is given in Figure 2, an excerpt from Thurston's work [5] stating one of most influential conjectures in the history of mathematics[1]. To state the conjecture formally, Thurston would have had to first define his 8 geometries. Instead, he wisely spends the introduction to explain where the conjecture comes from, without forgetting to mention that it would imply the Poincaré conjecture. He introduces the now famous 8 geometries in the subsequent pages, along with illuminating examples and figures. All in all, a deep technical conjecture that would take pages to state formally becomes an enjoyable read even for the non-expert. This highlights another important principle: rather than offering the reader the end result of your efforts, guiding them through the **milestones of your thought process** can make your work much more understandable and enjoyable.

It is common practise to follow the introduction with a section of *Preliminaries and Definitions*. This is helpful, but should be kept short. This section is there to be skipped rather than read. Sequences of definitions are boring, and boredom distracts. Moreover, it is difficult to remember something if you do not know what it is used for. The purpose of this section is to fix standard terminology in case any doubts arise later. Any new notions that you introduce for your proofs, and some of the standard ones, are better placed right before the first time they are used. For longer papers, consider using a preliminaries subsection at the beginning of each chapter.

Some authors place the less elegant parts of a proof in an *Appendix*. I disagree with this practise, as it leads to an awkward structure, and does more to emphasize the messy part than to hide it.

Once you think that your paper is ready, try to read it from end to end with the eyes of somebody seeing your results for the first time. Polish every detail. The referee's work is not to check the correctness of your proof, it is to confirm it. The right time to revise your paper is now, not a year later when the referee reports arrive. It is possible to spend half a career revising older papers, but also to get all your papers accepted with a handful of minor edits. **To become a great writer, you first need to become a careful**

[1]This is the *geometrization conjecture*, now a theorem of Perelman.

reader! (When refereeing, you can be generous with your optional comments, but you should be more lenient with others than with yourself; we should not be imposing our style on each other.)

Once done, wait 24 hours before showing your paper to anyone, including your co-authors. You will be surprised how often you will discover omissions and little mistakes within that period.

You now know enough of the basic principles to get going. Writing well will be a big challenge, but once you realise its importance and engage you will be able to overcome all its difficulties and gradually become a master, enjoying it as much as proving.

References

[1] P. R. Halmos. *How to Write Mathematics.* American Math. Society, 1973.

[2] T. R. Riley and W. P. Thurston. The Absence of Efficient Dual Pairs of Spanning Trees in Planar Graphs. *Electronic J. Comb.*, pages N13–N13, 2006.

[3] N. Robertson and P.D. Seymour. Graph Minors. I. Excluding a Forest. *J. Combin. Theory (Series B)*, 35:39–61, 1983.

[4] N. Robertson and P.D. Seymour. Graph Minors. XX. Wagner's Conjecture. *J. Combin. Theory (Series B)*, 92(2):325–357, 2004.

[5] W. P. Thurston. Three dimensional manifolds, Kleinian groups and hyperbolic geometry. *Bull. Amer. Math. Soc. (N.S.)*, 6(3):357–381, 1982.

[6] E. C. Zeeman. *Seminar on combinatorial topology.* IHES notes, 1963.