Mathematical Writing

UNIVERSITY OF OXFORD
Mathematical Institute

A ramble by Samuel N. Cohen

An underchecked and overambitious summary of thoughts, rants and opinions on the art of writing mathematics, liberally plagiarised from Higham [3] and Knuth [4] (principally), scattered with personal hobbyhorses and grievances.

Mathematical Institute
University of Oxford

2019, Not for general distribution

Oxford
Mathematics

Purposes

Mathematicians who merely think great theorems have no more done their job than painters who merely think great paintings

— Paul Halmos (in [4])

Why do we write mathematics?

Why do we write mathematics?

- To explain some mathematics (to people who don't already know it)
- To get a DPhil
- To publish papers and gain the associated academic credit
- To seem clever at family gatherings with that awkward uncle who always thought you should become a plumber.

Why do we write mathematics?

- To explain some mathematics (to people who don't already know it)
- To get a DPhil
- To publish papers and gain the associated academic credit
- To seem clever at family gatherings with that awkward uncle who always thought you should become a plumber.

Ultimately, the first point has to be the overriding goal, as the others won't be achieved unless the first is accomplished.

To do this, a few broad principles are needed in our writing

- Clarity
- Precision
- Novelty (and clarity and precision about what is novel)
- Engagement with readers' existing knowledge

If in doubt, most issues with writing can be answered by applying the golden rule:

Do to others as you would have them do to you

- If your writing doesn't explain to someone why they should spend their time reading what you have to say, then your writing is not going to do well.
- One of our goals is naturally to demonstrate to other people that we have understood something and are clever. The only way to do this is to explain it well.
- The vast majority of mathematics needs some selling. In most cases, people won't willingly spend time trying to decipher what you've written.
- For many people, muddled writing is taken as a sign of muddled thinking

Some broader ethical reasons why we try and write well:

- Most of you are being funded by society (through governments, grants, etc...) with an obligation to engage with the community. Part of this is explaining your work well.
- You are benefiting from the research community and the mathematics that has been developed and explained in the past. As a member of that community, you should try and make a contribution.
- Other people are supporting you, and part of thanking them for that support is doing the job well.

Know your readership

This section is about addition. The fact that the reader has been told this does not necessarily mean that he knows what the section is about, at all. He still has to know what addition is, and that he may not yet know. It is the author's fond hope that he may not even know it after he has read the whole section. Though addition, in general, is a special case of multiplication, here it is thought enough to consider only the addition of numbers, a very special case.

— Carl E. Linderholm, *Mathematics Made Difficult* (1970)
Chapter 2: Ad astra per aspirin

How should we pitch our work?

- A sensible target for a research article or thesis is someone who is starting/a few months into a doctorate in the area.
- If something is really standard in your area, then you don't need to detail everything, but you need to give enough detail to remind an educated reader who learnt these facts 10 years ago, and to suggest to a less familiar reader where they should go and find references.
- You lose less by explaining too much than by explaining too little. Explaining too much comes across as inelegant, while explaining too little comes across as muddled thinking (and hence bad mathematics)

The other part of knowing your readership is understanding the various conventions that should be used.

- Often problems will have standard notation (even when standardization is not really needed), and it's unkind to the reader to unnecessarily depart from this
- Comparing your work with standard examples often helps a reader to understand what's going on.
- It's easy to aim too high, or to think that what you are doing is so easy it's not worth explaining.

Types of writing

There are various types of writing that mathematicians do:

- Research papers
- Theses
- Survey/Expository/Review articles
- Monographs
- Large scale textbooks
- Opinion pieces
- Blog posts
- Public descriptions/outreach ...

Each of these has their own general style. We're going to focus on research papers (and theses to a lesser extent).

Telling a story

- Mathematical writing is usually a form of narrative.
- Occasionally you can write purely technical specification papers, but most mathematics is not of this type.
- As a consequence, good writing tells a story:
  - What is the problem?
  - Why are we interested?
  - How does it relate to earlier work?
  - How can we address the problem?
  - Does our solution work well?
  - What is the next problem?
  - What does this change about how we view the next/related problems?
- Involving the reader in the story leads to better understanding.
- Narrative $\neq$ History – don't just describe how you did it!

Structure

- Usual structure in my area (stochastic analysis/mathematical finance):
  1. Introduction
     1.1 Literature review
  2. Problem specification and defining terminology/notation
  3. Main result 1
  4. Main result 2
  5. Numerical implementation and discussion
  6. Bibliography
  7. Appendix of irritations which you need to do but don't want in the main text (annoying proofs, more numerics,...)
- Usually you want examples throughout
- Examples should include counterexamples when relevant – these are often a key part of the exposition

How to structure a thesis

- Typical chapter structure:
  1. Introduction and general review
  2. Chapters addressing related problems/variants of a broad problem – stereotypically equivalent to a single paper
  3. Conclusion and future extensions
- As a rule, a thesis should be more self-contained than a paper
  - Proofs should be given in more detail, more presentation of other people's work is appropriate.
  - Summaries of the techniques used can be included
- Usually it's a good idea to aim to write papers during the first couple of years of the doctorate, then rework these mildly into chapters of the thesis
- Try and be consistent with notation from the start.

Titles, Abstracts, Introductions, Conclusions...

Probably not.

— M. V. Berry, N. Brunner, S. Popescu, P. Shukla
Abstract to their paper

Can apparent superluminal neutrino speeds be explained as a quantum weak measurement?
arXiv:1110.2832

- Usually written fairly close to the end of the process
- Four principles from Herbert Wilf (in [4])
  - Get the attention of your readers immediately – snappy titles, arresting first sentences, lucid initial paragraphs...
  - Get everything up front. Tell your reader what you are going to do, so they can decide whether they are interested.
  - Remember people scan papers when they read them. Summarize your statements well. Make sure your notation is intuitively clear when the paper is read out of order.
  - A little motivation is good, but readers don't like too much.
- According to Kerkut (cited in [3]), for every person who reads the whole text of a scientific paper, five hundred read only the title.
- The title should be informative, but also snappy.

Abstracts

Mathematical Institute

- The abstract is there to summarize the contents of the paper. It should allow the reader to decide whether to read the rest.
    - Abstracts shouldn't make claims that the paper doesn't live up to!
    - Your abstract should make it clear what your new results are
    - Try not to start the abstract with 'In this paper' or 'This paper' or similar.

Introductions

- Introductions need to give a more detailed summary of the context and content of the paper than the abstract.
  - Generally an introduction is fairly short – a few hundred words. If I see a longer introduction, I often end up skipping most of it at a first read through.
  - Longer introductions may be needed if there isn't a good review available of the area.
  - It should define the problem, explain what the work attempts to do, and should outline the plan of attack. Unless there is good reason not to do so, it is advisable to summarize the results achieved.
  - You are not Agatha Christie. This is not a murder mystery. Tell the reader what you are going to do. Then tell them when you do it. Then remind them afterwards of what you've done.
  - Beginning with a question can be a good strategy.

Introductions

- Beginning with a list of your notation is not a good start to the introduction – leave this until later, when the reader is hooked!
- Some notation can be introduced, but be very sparing (this may vary between subfields)
- Often a literature review (of a length appropriate to the paper) is included immediately after/at the end of the introduction, or is incorporated to the introduction.

Exercise!

Go to arXiv.org and look at the recent papers in your area of mathematics. Read the abstracts of the 10 most recent papers and try and answer these questions:

- Do you know what the paper is about?
- Do you know what results the paper contains?
- Do you want to read the paper?
- How would you improve the abstract (not having read the paper)?
- Try and identify what in these makes for a good/bad abstract

Now try the same by taking your most and least favourite abstract, looking up the paper, and reading the introductions.

Exercise!

Oxford University of Oxford Mathematical Institute

Here's the basis for the proof of uniqueness of ODEs with a one-sided Lipschitz condition:

- Suppose $y' \leq ky$, then apply the chain rule to $e^{kx}y(x)$ to show $y(x) \leq y(0)e^{kx}$.
- Suppose $y \cdot y' \leq ky^2$. Then $(y^2)' \leq 2ky^2$, so we can apply the first result. This gives $|y(x)| \leq |y(0)|e^{kx}$ (Grönwall's inequality)
- Suppose $(g(y) - g(z))/(y - z) \leq k$, and consider $y' = g(y)$ and $z' = g(z)$. Using the above, get a bound on $|y(x) - z(x)|$, and apply this with $y(0) = z(0)$.

Write up a first draft of a proper proof (take no more than 20 minutes)

Exercise!

Give a short presentation of the Quicksort algorithm (read up on it on Wikipedia if you don't know it already). Present it two ways: in a format where you can derive the worst-case complexity and as pseudocode.

Give a short presentation of the Lotka–Volterra predator-prey model, say as a model for hares and wolves on an island. In particular, aim to describe the connection between the populations and the equations, and motivate the dynamics of the mathematical model.

Drafts and versioning

When you are a Bear of Very Little Brain, and you Think of Things, you find sometimes that a Thing which seemed very Thingish inside you is quite different when it gets out into the open and has other people looking at it.

— A.A. Milne, Winnie the Pooh

- Getting into good habits with drafts is critical.
- You need to try and write up results on the go, even if you rewrite them completely for the final version.
- Trying to get things down in full allows you to see where your thinking is faulty.
- Write the easy bits first! If a result is really standard, then feel free to leave it until later (but beware!). Getting at least a clear statement of the result is often needed.
- Get into a good system with backing up old versions of work (good tools help) and off-site storage!

Kill your darlings

– Attributed to Faulkner
(but also to Ginsberg, Wilde, Welty, Chesterton, Chekov...)

Kill your darlings, kill your darlings, even when it breaks your egocentric little scribbler's heart, kill your darlings.

– Stephen King (*On Writing*, 1999)

If you here require a practical rule of me, I will present you with this: 'Whenever you feel an impulse to perpetrate a piece of exceptionally fine writing, obey it—wholeheartedly—and delete it before sending your manuscript to press. Murder your darlings.'

– Arthur Quillen-Couch
(*On style*, 1914; apparently the first!)

Revising drafts

Mathematical Institute

Good drafting is the key to good writing. I usually find that the first draft is hard to write, then I need to leave it for a few weeks, then I am able to edit it again.

- Rewriting vs editing depends on your preference.
- Editing must be ruthless. Very few people write well as a stream-of-consciousness; the rest of us have to edit.
- In many cases, less is more. Write plenty, then aim to cut drastically. Remove unnecessary words, irrelevant ideas, ...
- Dead ends that you thought would be interesting but didn't work out can be cut or moved to an appendix
- In the end, a short good paper is much better than a long mediocre paper, even if the long paper has more ideas in it.
- I find it easiest to edit on paper, with a big red pen.

Collaboration Tools

When a paper has more than one author, it's important to find ways of interacting which everyone is comfortable with.

- Paper copies
- Email
- Dropbox
- Github
- Overleaf

Work with whatever makes life easiest and leads to the best results.

Mathematical style

Vidit Nanda
@viditnanda

Follow

Kashiwara-Schapira: making grown men scream into their pillows in despair since 1990.

$Y \times \{z_\circ\} \subset \{(x_\circ, y_\circ, z_\circ)\}$ on a neighborhood of $(x_\circ, y_\circ, z_\circ)$, (K_1, K_2) satisfies the conditions (7.3.3) and (7.3.4), and "$\varinjlim_{K'_1, K'_2}$" $K'_1 \circ K'_2 \xrightarrow{\sim}$ "$\varinjlim_{V}$" $(K_1)_{X \times V} \circ K_2$ belongs to $\mathbf{D}^b(X \times Z; (p_X, p_Z^a))$. $\square$

Proposition 7.3.4. *Let q'_1 and q'_2 denote the projection from $(X \times Z) \times (Y \times Z)$ to $X \times Z$ and $Y \times Z$ respectively and let i denote the diagonal embedding $X \times Y \times Z \hookrightarrow (X \times Z) \times (Y \times Z)$. Let $(K_1, K_2) \in \mathrm{Ob}(\mathbf{D}^b(X \times Y; (p_X, p_Y^a))) \times \mathrm{Ob}(\mathbf{D}^b(Y \times Z; (p_Y, p_Z^a)))$ be a microlocally composable pair at (p_X, p_Y, p_Z). Then $(i_*(K_1 \boxtimes A_Z), K_2) \in \mathrm{Ob}(\mathbf{D}^b(X \times Z \times Y \times Z; (p_X, p_Z^a, p_Y^a, p_Z))) \times \mathrm{Ob}(\mathbf{D}^b(Y \times Z \times \{\mathrm{pt}\}; (p_Y, p_Z^a, \mathrm{pt})))$ is microlocally composable at $((p_X, p_Z^a), (p_Y, p_Z^a), \mathrm{pt})$ and:*

$$K_1 \underset{\mu}{\circ} K_2 \simeq i_*(K_1 \boxtimes A_Z) \underset{\mu}{\circ} K_2 .$$

The proof is straightforward.

Proposition 7.3.5. *Let X, Y, Z and W be four manifolds, $p_X \in T^*X$, $p_Y \in T^*Y$,*

10:13 AM - 18 Nov 2019

- Mathematics is (usually) written in full prose.
  - There are exceptions: Maxwell wrote a poem called ‘A problem in dynamics’
  - A more common exception is (pseudo)code.
- This means that is subject to all the rules of grammar from the language in which it is written.
- Mathematical writing needs to be precise, so it is written in a more careful manner than most writing.
- Getting a reputation for being sloppy is really bad in mathematics. Being incomprehensible is only marginally better.

- Mathematical expressions are part of the sentence and so should be punctuated. This means full stops or commas should often appear at the end of a displayed equation.
- Symbols in different formulae should be separated by words, e.g. "Consider S_p, $p < q$" is worse than "Consider S_p, where $p < q$".
- Don't start a sentence with a symbol.
- Short sentences, with one clear idea, are easier to process.
- Run a spell checker. Seriously.
- Usually, the logic of a paper is to be expressed almost entirely in words. Avoid using $\Rightarrow, \forall, \exists$ and above all, unless you're doing pure logic, don't use $\therefore$ to indicate your conclusion.

One significant difference between mathematical writing and general prose is the large number of explicit internal references.

- When referencing an equation/result/lemma/etc... it is much more convenient to the reader if you describe the nature of what you are describing. So, it is better to say 'From the fundamental inequality (6.3)' rather than 'From (6.3)'.
- Remember you want to make it easy for your reader. Note that just describing it as an 'equation' doesn't really help.
- If you wish to point out that a theorem doesn't apply, be careful not to say that the theorem itself is invalid – you are usually claiming that it is inapplicable. Rarely, you will want to say that it is invalid, but only when you are constesting it.

From Higham:

Mathematicians are supposed to like numbers and symbols, but I think many of us prefer words. If we had to choose between reading a paper dominated by symbols and one dominated by words then, all other things being equal, most of us would choose the wordy paper, because we would expect to be easier to understand. One of the decisions constantly facing the mathematical writer is how to express ideas: in symbols, in words or both. I suggest some guidelines.

- *Use symbols if the idea would be too cumbersome to express in words, or if it is important to make a precise mathematical statement.*
- *Use words as long as they do not take up much more space than the corresponding symbols.*
- *Explain in words what the symbols mean if you think the reader might have difficulty grasping the meaning or essential feature.*

Broad guideline: In some sense, think of your writing as if you were giving a very long and detailed lecture.

- There are no absolute rules for deciding on the tense of mathematical writing, but consistency is king.
- The present tense can make your writing easier to follow, with the past tense used to indicate what the reader has already read.
- The past tense should be used for ‘facts’ that we now know are false.

Many people have strong opinions about we/I as the pronoun in mathematical writing.

- Most of us write as we, with the justification that this carries the reader along with you – as if you were giving a lecture.
- 'I' is reserved for more directly personal points.
- For example, I would write a transfer thesis in 'we' for the mathematics, but 'I' for when I get to the stage of writing about what I plan to do next.
- This rule changes quite a lot when writing less formal work (for example, a blog post)

- Do not distract the reader. Kill your darlings.
- The language and notation you use will depend on the audience. For example,
  - linear algebraists would write $\min_x \|Ax - b\|_2$ for the least squares problem,
  - statisticians would more naturally write $\min_b \|Xb - y\|_2$ for the linear regression problem,
  - probabilists would write $E[Y|X]$ for the conditional expectation of X given Y
  - geometers would write $P_{\mathrm{span}(X)}(Y)$ for the projection of Y on the subspace generated by X in a Hilbert space.
- Using natural notation will make it easier for your intended reader to grasp your meaning.

- Poor grammar is a distraction to many people, as well as leading to confusion.
- In order to get the reader to interact, most of us need to work hard not to distract them.
- Classic works to look at include Fowler's *Modern English Usage* (1926) [2] and *Hart's Rules* [9] (see also the *Oxford Guide to Style* [8]).
- The survey-based approach in Peters [6] is also interesting.

While this is not meant to be a course on English, a few common or mathematically sensitive errors are as follows:

- A vs An
- The active voice adds life and movement to writing. Look out for phrases like 'was performed/experienced/carried out/conducted/accomplished'. The sentence "Each of the Δz's are then multiplied by 2" is often easier to follow if you write "Multiply each Δz by 2"
- The adverbs *very, rather, quite, nice* and *interesting* should be used with a lot of caution. Are they needed? do they add anything? could a more specific alternative be used? Do they lead to confusion in different cultures (e.g. US vs UK)?
- Contractions such as *it's (not its), let's, can't, don't* are not usually used in formal works.

- Ambiguous *This* and *it*. Phrases like 'This is a consequence of Theorem 2' should be used with caution. The reader needs to backtrack to find out what 'this' is, and if you add an intermediate sentence later, it can modify the meaning. You want to be clear, so ambiguity is to be avoided.
- *Like* is used to denote similarity, while *such as* is used to denote a special case.
- Dangling participles are dangerous: "A bug was found in the program using random test data." While these often seem innocuous, and can be interpreted reasonably in nice cases, it becomes difficult when an author writes "The linear operator T acts on A-sets B from C using D." The same problem comes up when the dependence of variables is unclear.

- Elegant variation (not using the same word repeatedly to represent different things) is often used in English, and helps readers separate different ideas.
- Repetition can be used for emphasis, and helps to link words. So use parallelism when parallel concepts are being discussed. For example (varied from Strunk and White, in [4]),
  - Don't write: " Formerly, science was taught by the textbook method, while now the laboratory approach is employed."
  - Instead write: "Formerly, science was taught by the textbook method; now it is taught by the laboratory method."
  - On the other hand, don't write: "This method for studying the textbook method depends on observing student's methods."

- Enumeration/bullet points should still make sense grammatically. Avoid cases like:

  > *The use of PDEs in probability theory is common because they*
  >
  > - *allow for deterministic computational methods;*
  > - *naturally exploit the Markov assumption.*
  > - *The Feynman–Kac and Kolmogorov equations provide a clean link between Markov processes and PDEs.*

- If part of a single sentence, bullet points often end with semicolons. If separate sentences, they usually are preceded by a colon.

Punctuation is there to make meanings clear. This is the ultimate principle of English punctuation.

- Commas are often useful, but, it is possible, in cases, to over-use them. However avoiding them completely makes your writing completely incomprehensible to most readers and difficult to read even when it makes otherwise perfectly good grammatical sense and sounds OK in your head.
- The Oxford, or serial, comma is generally a good idea, as it improves clarity.
- For more discussion of the correct use of commas and other punctuation marks, *Hart's rules* [9] is very useful.

- Words that are derived from a person's name are capitalized: Gaussian, Hamiltonian, Hermitian, Jacobian, Lagrangian, Eulerian, Lax pairs, Desarguesian, Cartesian
- A hyphen - is used for a single idea/person's name, while an endash – (in latex made by - -) is used for two names or two link two distinct ideas.
  - So it is the Birch–Swinnerton-Dyer conjecture (2 people, Bryan Birch and Peter Swinnerton-Dyer)
- A hyphen may appear when a word is broken at the end of a line, while an endash is used to separate a clause in a sentence – for example, like this – as an explanation.

- Semicolons are often not used extensively in mathematical writing, but this is personal preference.
- I like footnotes but many mathematicians don't. If you use them, use them sparingly.
- The word 'only' is often dangerous for a mathematician, as we need it for precision. To see why this is an issue, consider inserting 'only' in an arbitrary position in the sentence 'I hit him in the eye yesterday'.

- While not as strong in English as in other languages (e.g. Latin or Greek), the first words in the sentence usually guide the force of the meaning. For example,
  - 'PDEs are used to model the flow of heat in various materials'
  - 'The flow of heat in various materials is modelled using PDEs'
  - 'In various materials, the flow of heat is modelled using PDEs'
- Reordering a sentence can strengthen it and remove ambiguity, e.g.

  The limit point is only a stationary point when the regularity conditions are satisfied.

  vs

  The limit point is a stationary point only when the regularity conditions are satisfied.

Some examples extended from Higham [3]

- (Combinations) also, and, as well as, besides, both, furthermore, in addition to, likewise, moreover, similarly
- (Implications/explanations) as, because, conversely, due to, for example, given, in other words, in particular, in view of, it follows that, otherwise, owing to, since, specifically, that is, thus, unlike
- (Consequences) accordingly, consequently, hence, therefore, thus, results in, leads to

Linking words

University of Oxford
Mathematical Institute

- (Modifications/restrictions) although, alternatively, but, despite, except, however, in contrast, in spite of, nevertheless, of course, on the contrary, on the other hand(*), though, unfortunately, whereas, yet
- (Emphasis) actually, certainly, clearly, in fact, indeed, obviously, surely

Bullshit

- Don't. (Porter [7])
- The people who you need to impress will not be impressed.
- There are few things more disappointing than reading a paper only to find that it doesn't address the problem that it claims to address.
- Be careful not to oversell your work. The readers whose opinion you need will often be expert enough to spot this, and it doesn't look good.
  - For academic writing, being careful to be precise is critical to being taken seriously.

Theorems and Proofs

If you write clearly, then your readers may understand your mathematics and conclude that it isn't profound. Worse, a referee may find your errors.

— J.S. Milne

See `www.jmilne.org/math/tips.html` for "some tips for avoiding these awful possibilities"

- Proofs are there to support the theorems, and are usually not of direct interest in their own right (but there are notable exceptions)
- Most of the people who will read a proof in detail will want all the details, as they are trying to use/extend the result.
- There is a balance to be found between patronizing the reader and not skipping things. Getting a good feel for this is a matter of experience.

What are the differences between

- Theorems, Lemmas, Propositions, Corollaries, and Conjectures?
- Assumptions, Hypotheses, Settings, and Definitions?
- Algorithms and Methods?
- Examples and Implementation results?

- A theorem is a major result of independent interest.
  - The proof is usually nontrivial; the statement should be fairly self-contained (it shouldn't depend too much on the discussion beforehand, the notation is often (re-)defined as part of the statement, etc..)
  - Often theorems are results that you hope someone else will cite.
- A lemma (plural lemmata or lemmas) is an auxiliary result, a stepping stone to a theorem.
  - More tolerance is given to statements not being completely self-contained.
  - Sometimes historical results are known as lemmas even though they are effectively theorems (and vice versa).
  - Proofs may be easy or difficult, but are usually given. If difficult or distracting, proofs may be relegated to an appendix.

- A proposition often means something between a theorem and a lemma.
  - Sometimes it is used to indicate a true statement for which the proof is omitted (but it's better to be explicit and give a proof).
  - A proposition is still certainly true.
- A corollary is a (usually easy) consequence of a theorem (or more rarely a lemma or proposition).
  - Some corollaries are of significant independent interest.
  - Proofs are often omitted when the result is a clear application of the parent result.
- A conjecture is a statement which there is usually some reason to believe is true, but no proof. Some famous conjectures are well known.

- A setting (I can't think of a better word for this) is not usually specified explicitly, but forms the basic framework for the paper.
  - For example you could be considering a certain class of PDEs, and so you write the equation you wish to consider, the types of solutions you need, etc...
  - Sometimes these will be expressed in a definition, but more often they just appear in the general text.

- An assumption is a statement, usually not part of the basic framework, which you will use.
  - Assumptions need to be things which could be either true or false.
  - You should specify whether an assumption will be taken to hold for the remainder of the work, or whether you will be invoking it whenever needed (e.g. 'Under assumption 1, we know that the ODE (...) admits unique solutions.')
  - Even when you have specified that an assumption holds throughout your paper, it may be worth reminding the reader of this when you get to a significant result which you want to be self-contained (I have seen the negative consequences of authors failing to do this...)

- A hypothesis is a statement that is taken as a basis for further reasoning, usually in a proof (e.g. for induction).
  - Rarely hypotheses stand on their own (e.g. Riemann or Continuum, which are somewhere between assumptions or conjectures).
- A definition is usually there to give a key concept a name or notation.
  - Definitions should be used sparingly, as they require the reader's memory.
  - At the same time, a good definition can save a thousand repeated qualifications.
  - In definitions, by convention, if is taken to mean 'if and only if'.

- An algorithm is a formal statement, often in pseudocode, of a computational method to determine something of interest
- A method is a technique, often more general than an algorithm, to perform some task
- An example is usually an application of a result or method, which may be of independent interest, or may be a toy setting where the key phenomena are displayed.
  - Theorems often have examples; lemmas usually don't
- An (implementation) result is of independent, often primary, interest. Usually this will not be a toy problem.

> *It was said of Jordan's writings that if he had 4 things on the same footing as* (a, b, c, d) *they would appear as* $a, M_3', \epsilon_2, \Pi_{1,2}''$.
>
> — J.E. Littlewood. Littlewood's Miscellany

- Readers are often not very interested in the details of a proof, but want to know the outline and key ideas. This guides the presentation of good proofs.
- At the same time, the readers who will go through the details are often trying to extend/apply/modify the argument, so missing steps are poor form (and very annoying).
- Many readers will only read the theorem statements on the first iteration, so try and make these understandable without reading the proofs.

Say what you mean

- Be careful of saying things like "According to Theorem 1.1, a single trajectory $X(t, x)$ passes through almost every point in phase space", when you mean that for each point in the space, there is a unique trajectory passing through that point.
- Be particularly careful to make the dependence of terms clear – 'for every $\epsilon > 0$ there exists a $\delta > 0$' is not the same as 'there exists a $\delta > 0$ such that for every $\epsilon > 0$'
- Similar concerns arise in expressions where you need to indicate what constants depend on what previously defined constants.

Signposting

- It is kind to help the reader keep track of where they are in a proof.
- Without this, we easily get confused, distracted, and find it difficult to follow.
- Bigger proofs should be broken up into smaller results – rarely do you need more than two pages!
- Even if you break things up, telling the reader up-front how the proof works will often be helpful.
- Don't just signpost once – tell the reader how the proof is progressing on-the-fly!

Some phrases which might help are (from Higham and references therein):

- The aim/idea is to
- Our first goal is to show that
- Now for the harder part
- The trick of the proof is to find
- ... is the key relation
- The only, but crucial use of ... is that ...
- To obtain ... a little manipulation is needed
- The essential observation is that

When you omit part of a proof it is best to indicate what is omitted, via phrases such as:

- It is easy/simple/straightforward to show that
- Some tedious manipulation yields
- An straightforward induction gives
- After two applications of ... we find
- An argument similar to that used in ... shows that

Be very careful with this! Two key risks:

- If you haven't done the work, you could be wrong, and this is very embarrassing.
- I usually recommend against using easy/trivial/obvious/etc...

54

Nous pouvons maintenant définir l'intégrale stochastique. Soit $X \in \underline{\underline{S}}$ et H un processus prévisible. On veut que l'intégrale stochastique $H \bullet X$ soit un élément de $\underline{\underline{S}}$ vérifiant $(H \bullet X)^c = H \bullet X^c$ et $\Delta(H \bullet X) = H \Delta X$, et qui en plus coïncide avec l'intégrale de Stieltjes lorsque $X \in \underline{\underline{V}}$ (voir (2.51)) Compte tenu de ces contraintes, et étant donnés (2.64) et (2.66), il est clair que la plus grande classe possible d'intégrands prévisibles est:

$$(2.67) \quad L(X) = \{H \text{ processus prévisible: } H \in L^1_{loc}(X^c) \text{, il existe } D \in \underline{\underline{D}}(X) \cap \underline{\underline{D}}'(H\Delta X) \text{ avec } H \bullet \check{X}^D \in \underline{\underline{V}}\}.$$

Remarquer que si on localise $L(X)$ à la manière de (2.41), on a $L_{loc}(X) = L(X)$.

— J. Jacod, *Calcul Stochastique et Problèmes de Martingales* (1979)

It took me three days of work to get my head around why 'il est clair que' (it is clear that, before (2.67)).

Keeping track to where you are helps a reader. Useful phrases include:

- First we establish that
- Our task is now to...
- Our problem reduces to
- It remains to show that
- We are almost ready to invoke
- We are now in a position to
- Finally, we have to show that

Examples

- Examples can be very effective in explaining the idea/problem/setting/motivation.
- The practice of writing papers which do all the theory before any examples is to be avoided – you want to explain what you are interested in.
- Examples need to be well chosen though – they should demonstrate your main point, without too many distractions.
- Counterexamples can also be very useful in helping a reader to understand the conditions of a theorem.
- The style of examples varies depending on the paper: If you're mainly proving theorems, then examples will demonstrate the theorems. If you're building a model, then examples may be special/interesting cases...

Exercise!

Look at the two example papers.

- The first is deliberately written badly. Pick out key flaws in the presentation and make some recommendations.
- The second is not deliberately badly written (I wrote as a student). There are still plenty of points where the presentation could be improved! Go through with a red pen and suggest changes.

Exercise!

Mathematical Institute

Take the ODE uniqueness proof that you wrote and improve it.

Mathematical Modelling

Presenting a model of something in the real world is difficult.

- You need to translate people's prior knowledge and intuition into explicit mathematics.
- The mathematics needs to be motivated.
  - How does the basic physics/economics/biology/chemistry of the problem give rise to the question you are answering?
  - Tell me why equation (2) is there – is it just Newton's laws of motion? Bellman's Principle? The Hamiltonian of the system?
  - Try and express this with reference to the real system studied
- Define your variables clearly, and at some point give a precise statement of the mathematical model(s) you are using to consider the problem.
- Give the model before doing any analysis.

Data and numerical results

- Well chosen graphs and other output are key to conveying understanding.
  - A graph should never be presented without discussion.
  - It's easier to compare columns than rows in a table.
- If you need to include a large amount of data, consider putting it in an appendix (or supplementary information), and summarizing it in the main text.
- Be careful not to mislead the reader with the output that you present. It will come back to bite you.
- Have a look at Tufte's *The visual display of quantitative information* [11] for ideas on how to present data clearly
- Presentation of algorithms etc. is subdiscipline specific, but usually some sort of pseudocode is used.

Citations

> *To write a reference, you must have the work you're referring to in front of you. Do not rely on your memory. Do not rely on your memory. Just in case the idea ever occurred to you, do not rely on your memory.*
>
> — Mary-Claire van Leunen,
> *A Handbook for Scholars* (1992, cited in [3])

- Ultimately, the key rule for citations is the golden rule: 'Do to others as you would have them do to you'.
- You lose no credibility by citing generously, provided it's not ridiculous. You do lose credibility by omitting people.
- Keeping good records is critical to doing this in practice. Use a bibliographic manager (e.g. Jabref, Endnote,...)
- To ensure your own citations are not lost, use the same name (and form of name) on your mathematical work.
  - I am Samuel N. Cohen in all my published academic work.
  - Not Sam Cohen, S Cohen, SN Cohen, Samuel Cohen, Samuel Nicholas Cohen or other variants.
  - The initial helps distinguish me from Serge Cohen (Toulouse, also a probabilist), but I still sometimes get confused with the Samuel Cohen who works on Alzheimers.

Citing technical results

Often, you will need to cite other people's arguments as part of making your own.

- The simplest case of this is when using someone else's result.
- Here the important thing is to give a precise reference, and to make it clear that you have verified all the assumptions.
- For example, saying "Bardi and Da Lio [*] shows that this equation has a unique viscosity solution" is not very useful to most readers.
- A better option is to say "As we have verified that the function f is uniformly continuous, and our space is locally compact, we can apply Theorem 6.28 of Bardi and Da Lio [*] to verify that our PDE (2) has a unique viscosity solution".

- More complicated is when you are referencing a larger body of work. Here you get statements like "The interested reader is referred to Amari [*], and references therein, for a general presentation of the development of information geometry."
- Very rarely, you will want to reference someone else's proof technique. You can say things like "The proof now follows as in Dellacherie and Meyer [*], proof of Theorem 2.37, with the assumption of convexity guaranteeing the continuity of the function *h*."
- This is a bit riskier, but sometimes is needed to avoid reproducing a whole proof. Be very careful.

- Journals have different styles for citations, but the most common practice in mathematics seems to be cite-by-number (rather than by name-date). It's good style to include the author's name(s) when citing by number, as
  1. people like seeing their name in print, and you may as well make people happy
  2. it enables readers to more easily remember what you are referring – many readers will know a good number of your references, and others they can remember more easily by name than by number.
- Write authors' names out in full (at least for ≤ 3 authors) the first time, then use 'et al.' after.
- Cite authors in the order they appear, and with the spellings, which are on the paper (not the journal contents page!)

- You will need to produce a reference list.
- Use a good reference manager.
  - I like jabref (which is just a front-end for bibtex), and try and put everything I read in it asap.
  - Others use endnote etc, but do use something.
  - Don't just write it directly in the tex file, as that is more work.
  - You will save a lot of time by not leaving it to the end.

- Plagiarism is a major and serious academic issue. (Look up Karl-Theodor zu Guttenberg if you don't believe me)
- As stated by Higham, if you copy a sentence or more, you should either place it in quotes and acknowledge the source as a citation, or give an explicit reference such as "As stated by Higham ..."
- The shortest viva I have heard of began with asking for a definition of plagiarism.
- If you want to paraphrase, then it's best to do so with a little distance from the source material.

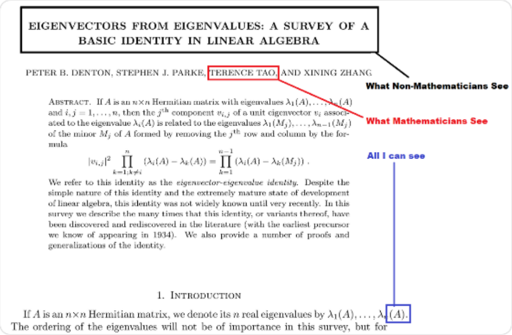
- Reworking your own publications without acknowledgement, and pretending they are new, is self-plagiarism, and is also impermissible.
- There is a special place in Hell for those who submit the same work to multiple journals.

Typesetting

Vidit Nanda
@viditnanda

Follow

Overfull \hbox is my tragic flaw.



EIGENVECTORS FROM EIGENVALUES: A SURVEY OF A BASIC IDENTITY IN LINEAR ALGEBRA

PETER B. DENTON, STEPHEN J. PARKE, TERENCE TAO, AND XINING ZHANG

Abstract. If A is an $n \times n$ Hermitian matrix with eigenvalues $\lambda_1(A), \dots, \lambda_n(A)$ and $i, j = 1, \dots, n$, then the j^{th} component $v_{i,j}$ of a unit eigenvector v_i associated to the eigenvalue $\lambda_i(A)$ is related to the eigenvalues $\lambda_1(M_j), \dots, \lambda_{n-1}(M_j)$ of the minor M_j of A formed by removing the j^{th} row and column by the formula

$$|v_{i,j}|^2 \prod_{k=1; k \neq i}^{n} (\lambda_i(A) - \lambda_k(A)) = \prod_{k=1}^{n-1} (\lambda_i(A) - \lambda_k(M_j)) .$$

We refer to this identity as the *eigenvector-eigenvalue identity*. Despite the simple nature of this identity and the extremely mature state of development of linear algebra, this identity was not widely known until very recently. In this survey we describe the many times that this identity, or variants thereof, have been discovered and rediscovered in the literature (with the earliest precursor we know of appearing in 1934). We also provide a number of proofs and generalizations of the identity.

1. Introduction

If A is an $n \times n$ Hermitian matrix, we denote its n real eigenvalues by $\lambda_1(A), \dots, \lambda_n(A)$. The ordering of the eigenvalues will not be of importance in this survey, but for

1:56 PM - 4 Dec 2019

- Use LaTeX
- It's probably worth learning to write in basic LaTeX, but some people prefer to use a WYSIWYG editor, e.g. Lyx, bakoma, texmacs, texshop or scientific word.
- The AMS-LaTeXtools are useful, particularly `align` and `split` are your friends.
    - `enumerate` is good for bullet points
    - `operator` (from AMS-LaTeX) gives nicely defined functions
    - *The Short Math guide to LaTeX* (by the AMS, available on CTAN) is worth looking through
    - *A Not So Short Introduction to LaTeX* is good for more general points

- As a rule, don't mess with the standard LaTeX setup. The page widths, spacings, margins, ... have been chosen because they generally work well.
- In my experience, most journals accept papers in standard formatting. You can reformat later after acceptance. Similarly for bibliography style. (This varies by subfield)
- For more discussion of general typography and page design, see Bringhurst [1].

- An equation is displayed when
  - it needs to be numbered,
  - it would be hard to read if placed in-line, or
  - it merits special attention, perhaps because it contains the first occurrence of an important variable.
- I think that not every equation needs to be numbered. Equations that you need to reference, or you think someone else might like to reference, should be numbered.
  - Chris Breward thinks you can't tell what equations someone else might like, so number them all.

- It is a bad idea to display false equations, i.e. $1+2=4$ is better than
$$1+2=4.$$
The reader's eye will see them, without necessarily noticing the surrounding text (where you hopefully make it clear that it is false).

- When long formulas don't fit, try to break the lines logically, or rework the notation. Don't make the text size smaller.
- Moving a formula to be displayed makes it easier to split it up logically.
- If you break a line within a parenthesis, alignment should be modified to make it as clear as possible to the reader.
- A new line of a multi-line equation will usually begin with a binary operator (e.g. $+$)

Displayed equations should be aligned at the $=$ sign (or similar).

- Other alignments are rare, and are usually much less clear to the reader
- Try and avoid writing cases like this:

$$\begin{aligned} A = B \leq C &= D \qquad (1) \\ &= E \qquad (2) \end{aligned}$$

  as the reader naturally sees $A = E$ as a conclusion.
- It may be worth considering defining (temporary) notation for the sake of reducing a long equation to a more manageable format, particularly if this helps to highlight a relevant structure.

- When writing inline text, it's usually better to use a slash $(1+2)/4$ than a fraction $\frac{1+2}{4}$.
- This gets harder and harder to read when you get $\frac{C_{t_{k_i}+17}^{(4^{\frac{1}{2}})}}{f(18^2; s_{k_i+j}+t_{k_i})}$, it's also very hard to pick up typos, and messes up the line spacing (which makes reading harder).
- The key is always to make the notation as simple as possible to read the paper.
  - Don't assume your reader will remember your notational conventions
  - If your notation is getting too convoluted, then you may need to simplify your methods

- Personal preference (and the AMS's recommendation) is to manually decide the sizes of parentheses – the common practice of using `\left` and `\right` leads to unnatural sizes, and behaves badly when you have an equation over multiple lines.
- The cardinal rule is that every parenthesis should be paired and sized sensibly to make it easy to see what is inside.
- Pages of equations make me sad and unwilling to read. So do pages of block text. Mix it up to keep me interested.

Summary

There is no way of being almost funny or mildly funny or fairly funny or tolerably funny. You are either funny or not funny and there is nothing in between. And usually it is the writer who thinks he is funny and the reader who thinks he isn't.

— I. Asmiov, in *Buy Jupiter and Other Stories* (1975)

- Be clear.
- Be kind to your reader and don't assume too much of them.
- Good writing takes effort, but is worthwhile.
- Writing your mathematics so that someone else can understand it is an essential part of being a researcher.
- Read a lot, and try and keep track of what is clear to you and why.
- Practice writing well, and work with each other to improve your writing. Don't be afraid to get criticism and help.

For further reading:

- Higham [3] – very good overall presentation
- Knuth [4] – lecture notes from a course
- Krantz [5] – new book on mathematical writing (on arXiv)
- Porter [7] – lecture notes
- Steenrod, Halmos, Schiffer and Dieudonné [10] – a classic
- Terry Tao's blog also has links to good resources.

[1] Robert Bringhurst.
The Elements of Typographic Style.
Hartley & Marks, 1992.

[2] H.W. Fowler.
A dictionary of Modern English Usage.
1926.

[3] Nicholas J. Higham.
Handbook of Writing for the Mathematical Sciences.
SIAM, second edition edition, 1998.

[4] Donald E. Knuth, Tracy Larrabee, and Paul M. Roberts.
Mathematical writing.
Available at http://www.jmlr.org/reviewing-papers/knuth_mathematical_writing.pdf, 1987.

[5] Steven G. Krantz.
A primer of mathematical writing, second edition.
https://arxiv.org/abs/1612.04888, 2016.

[6] Pam Peters.
The Cambridge Guide to English Usage.
Cambridge University Press, 2004.

[7] Mason A. Porter.
Writing papers in applied mathematics: An opinionated tutorial and discussion.
`www.youtube.com/watch?v=oNgqQyF0GfY`, 2018.

[8] Robert Ritter.
The Oxford Guide to Style.
OUP, 2002.

[9] Rosemary Roberts, editor.
New Hart's Rules.
Oxford University Press, 2005.

[10] N. E. Steenrod, P. R. Halmos, M. M. Schiffer, and J. A. Dieudonné.
How to Write Mathematics.
AMS, 1973.

[11] Edward R. Tufte.
The Visual Display of Quantitative Information.
Graphics Press, 1983.