1 Intro - Coming up with proofs and writing them down.

Writing mathematical proofs is an acquired skill, and a fairly complex one. Complex not in the sense of “difficult”, but in the sense of “made up of multiple parts”.

A proof is a mathematical argument written in a special style. Hence to begin, one must have an argument. Ideally, it should explain why something is true; sometimes knowing how it is true suffices. In either case, in order to write a proof, one must have a proof. There are no rules for coming up proofs (if there were, all mathematicians would retire), but there are some techniques and heuristics. We discuss these in Section 2.1. Combined with a good deal of deliberate practice they can make one a competent proof-creator.

Once you have a proof, you have to write it down. Writing proofs has a lot to do with writing in general, or at least with writing persuasively, but also has a lot of traits peculiar to this particular writing task. They range from basic literacy of mathematical notation and stylistics like “avoid starting a sentence with a symbol”, to meta-level things like “do not clutter your proof with extraneous lines of
reasoning”. Fortunately, this part of the task is more of a collection of best practices, and rarely requires unusual inspiration. We shall try to collect some of these practices in Section 2.2. The list is of course incomplete, but practice, as usual, makes perfect.

Initially you may find that you often “understand what’s going on” but can’t write it down well. With practice and experience writing things down becomes the easy part, and understanding comes to focus more. I hope you will find useful things about both in this text. The text is rather long, but you can read whatever parts you think will be useful.

2 How to solve it? - the four steps.

“How to solve it?” is the title of a classic book on mathematical reasoning by George Polya (of Stanford). Polya was one of the first people to think about problem-solving seriously. He identified four phases of problem solving.

1. First, one has to understand the problem. For a problem asking for proof, ask “What is the hypothesis? What is the conclusion?”. Make sure you understand all the terms; look up definitions if necessary. Understand each part of the hypothesis, try to see how it affects the problem. See how various parts of the hypothesis relate to each other. Think about the conclusion. How may it be related to the hypothesis?

2. Second, devise a plan for solving the problem. This is what Ross does in “discussion” part of the text, before the formal proof, and what we try to do in lectures. Coming up with ideas and plan of the proof is a complicated business. We will have more to say about this in the next section.

3. Third, one has to carry out the plan, that is to write down an argument, either right away as a formal proof, or first as a outline argument to be filled in and turned into a formal proof later. This involves cleaning up the ideas and such things as using correct notation, good style, justifying individual steps and formatting into smaller units - lemmas or propositions - if needed. This part is easier, though by no means obvious. The rules of the game are to be learned, but playing it should become reasonably routine with practice. We will give a few pointers on this as well.

4. Finally, one should look back at the solution. Can you check the argument? What are one or two important ideas/key steps in the proof? Can you make a one sentence summary that will let you reconstruct it? Can you prove the same thing in a different way? Can you use the same method for some other problem? Can you use the result itself to simplify other things? Perhaps do a concrete example?
All four stages are important. I recommend reading first two parts of Polya’s book (it’s 35 pages total) to see more details on each, illustrated with specific examples from “elementary” mathematics. In the next two sections we shall talk about the stages 2 and 3 illustrating them with examples from the class.

2.1 The Art and Craft of Problem Solving - the three levels.

The heading of this section - which is about coming up with proofs, roughly corresponding to the first two steps in Polya’s 4-step process - is taken from the title of a book by Paul Zeitz, a professor at University of San Francisco (there is also a video course of the same title from The Great Courses series).

Zeitz points out that an experienced problem-solver (or proof-creator) operates on three levels:

- **Strategy**: Mathematical and psychological techniques for starting and pursuing proofs.
- **Tactics**: Diverse mathematical methods that work in many settings.
- **Tools**: Narrowly focused techniques and “tricks” for specific situations (analysis has its share of such tools).

Zeitz makes the following analogy

You are standing at the base of a mountain, hoping to climb to the summit. Your first strategy may be to take several small trips to various easier peaks nearby, so as to observe the target mountain from different angles. After this, you may consider a somewhat more focused strategy, perhaps to try climbing the mountain via a particular ridge. Now the tactical considerations begin: how to actually achieve the chosen strategy. For example, suppose that strategy suggests climbing the south ridge of the peak, but there are snowfields and rivers in our path. Different tactics are needed to negotiate each of these obstacles. For the snowfield, our tactic may be to travel early in the morning, while the snow is hard. For the river, our tactic may be scouting the banks for the safest crossing. Finally, we move onto the most tightly focused level, that of tools: specific techniques to accomplish specialized tasks. For example, to cross the snowfield we may set up a particular system of ropes for safety and walk with ice axes. The river crossing may require the party to strip from the waist down and hold hands for balance. These are all tools. They are very specific.

2.1.1 Example.

As an illustration, of both Zeitz’s three levels and Polya’s four steps, let’s take an example that is relatively fresh in our minds, from Lecture 8. In discussing it we
will mention a lot of strategies, tactics, and tools. The list of examples of all three that are commonly used will follow. The discussion is long, but the process itself does not have to be. Sometimes it takes longer to say thing than to do them.

**Proposition 2.1.** If \( s_n \) converges to \( s \), and \( t_n \) converges to \( t \) then \( d_n = s_n t_n \) converges to \( d = st \).

First we must understand the problem. Zeits calls this getting oriented. What is the hypothesis? It has two parts - \( s_n \) converges to \( s \) and \( t_n \) converges to \( t \). What does this mean? If we don’t remember the definition of “converges” we look it up. We translate the hypothesis into something more concrete.

For any \( \varepsilon > 0 \) there is \( N \) such that \( |s_n - s| < \varepsilon \) for all \( n > N \).

For any \( \varepsilon > 0 \) there is \( N \) such that \( |t_n - t| < \varepsilon \) for all \( n > N \).

But the \( N \) in the second part of the hypothesis does not have to be the same as the \( N \) in the first part. Come to think of it, neither does \( \varepsilon \). Let’s give them separate names.

For any \( \varepsilon_1 > 0 \) there is \( N_1 \) such that \( |s_n - s| < \varepsilon_1 \) for all \( n > N_1 \).

For any \( \varepsilon_2 > 0 \) there is \( N_2 \) such that \( |t_n - t| < \varepsilon_2 \) for all \( n > N_2 \).

(This bit of choosing notation is often an important strategy.)

Are the two parts of the hypothesis related? Well, they both refer to what happens for large values of \( n \). So we can get both statements to hold for large \( n \). Ok, that’s something. Otherwise they seem fairly independent. We may fish for some other ideas, or move on to the conclusion.

We translate that as well.

For any \( \varepsilon > 0 \) there is \( N \) such that \( |s_n t_n - st| < \varepsilon \) for all \( n > N \).

We could use \( \varepsilon_3 \) and \( N_3 \) here, but we have freed up \( \varepsilon \) and \( N \), so we can use those, there should be no confusion.

The conclusion is in one statement, no moving parts here. But how does it relate to the hypothesis? Well, it’s again about something happening for large \( n \). Since the hypothesis only makes a statements about large \( n \) and we are free to choose how large. So we can assume we chosen \( n \) large enough that both \( |s_n - s| < \varepsilon_1 \) and \( |t_n - t| < \varepsilon_2 \) hold.

Just by implementing Step 1 carefully we have made some progress.

On to Step 2 - coming up with a plan. To begin, we employ some strategies. One strategy to start is to try things or follow your nose.

This may involve the strategy of restating the problem - perhaps to make hypothesis and conclusion look more like each other. Perhaps we realize that we need to control \( |s_n t_n - st| \) so we should try to make it look more like \( |t_n - t| \) and \( |s_n - s| \). Here the tactics of factoring and decomposing may come up - we see that \( s_n t_n - st \) is quadratic and we need to control it in terms of linear terms, so it would be good to factor; however it does not factor “as is” so perhaps we may gain something by factoring part of it and controlling that, and then figuring out what to do with the rest. We may also employ the strategy of looking at penultimate
that is trying to see what kind of statement would imply the conclusion in one step. Maybe one gets an idea that \( |s_n t_n - st| < A|s_n - s| + B|t_n - t| \) for some constants \( A \) and \( B \) would do it (one is more likely to get such an idea after some practice with analysis proofs. But practice is more of a meta-method rather than a tactic or strategy). We can continue reasoning backwards, noticing that by the triangle inequality (one of the analysis-specific tools), the last part would be implied by \( s_n t_n - st = A(s_n - s) + B(t_n - t) \). We still may need a bit of inspiration to decide that \( B = s \) is a good constant to take - we may notice that is gives the \( st \) term we want. It may seem imperfect because there are probably other things it would create, but here is a good place to employ the all-important strategy of mental toughness, and the principle that anything that furthers the investigation is good and try it. If \( s_n t_n - st = A(s_n - s) + B(t_n - t) \) and \( B = s \) we do some algebra and get \( A = t_n \), which is good because it is simple, but bad because it is not a constant. We may decide one of two things - either this line of reasoning should be abandoned OR we should be tougher and persevere. The fact that we are making progress and that we seem close both suggest we try more. This is a good time to reflect on what we did (reflection is of course step 4 in the whole proof-writing/problem-solving process, where we look back at the proof as a whole, but it is also indispensable during the proof, especially when some progress has been made but an impasse is reached). We are trying to say something along the lines of

\[
s_n t_n - st = A(s_n - s) + B(t_n - t), \text{ so } |s_n t_n - st| \leq |A||s_n - s| + |B||(t_n - t)|,
\]

and if \( A \) and \( B \) are constants we can make \(|(s_n - s)| \) and \(|(t_n - t)| \) small enough to make the terms \(|A||s_n - s|\) and \(|B||(t_n - t)|\) small, so that \( |s_n t_n - st| \) is small as well.

Our problem is that \(|A|\) is not a constant. But if we engage in the strategy of wishful thinking and consider how \(|A|\) would be used in our proof (if it were, indeed, constant), we may notice that the only thing we care about is that \(|A|\) is bounded! So if we knew \(|t_n|\) is bounded we would be done. Now we employ the final strategy of don’t start from scratch and use the lemma that said that convergent sequences like \( t_n \) are bounded (and hence so are their absolute values \(|t_n|\)). Note that if we did not know this from before, we still could separate a lemma and so reduce our problem to this lemma, to which all the steps and levels of problem solving would apply as well.

By now we have the plan, so we have completed step 2.

In step 3 we must write the proof. We shall have more to say on this process in Section 2.2 but here is one possible outcome. (See also Appendix A).

**Proof.** Let \( \varepsilon \) be any positive number.

Since \( t_n \) converges, there exists an upper bound on \(|t_n|\). Let \( B_1 \) be a (non-zero) such bound. Let also \( B_2 = |s| \).

Let \( \varepsilon_1 \) and \( \varepsilon_2 \) be two positive numbers such that \( \varepsilon_1 B_1 + \varepsilon_2 B_2 = \varepsilon \).

Since \( s_n \) converges to \( s \), there exists \( N_1 \) such that for all \( n > N_1 \) we have \(|s_n - s| < \varepsilon_1 \) (error for \( s_n \) is small).

*Penultimate means the one before last.*
Since \( t_n \) converges to \( t \), there exists \( N_2 \) such that for all \( n > N_2 \) we have \( |t_n - t| < \varepsilon_2 \) (error for \( t_n \) is small).

Then for all \( n > \max\{N_1, N_2\} \) we have

\[
|d_n - d| = |t_n s_n - st| = |t_n s_n - t_n s + t_n s - ts| = |t_n(s - s_n) + s(t_n - t)|
\]
\[
\leq |t_n| |s_n - s| + |s| |t_n - t| \leq |t_n| \varepsilon_1 + |s| \varepsilon_2
\]
\[
\leq B_1 \varepsilon_1 + B_2 \varepsilon_2 = \varepsilon
\]

(Note that we have used the triangle inequality in going from the first line above to the second). This means that \( d_n \) converges to \( d \), as wanted.

\[\square\]

We are now ready for step 4. What made the proof work? After translating everything we saw that we needed a decomposition of the term to be estimated into summands, and of each of the summands into a product, with each resulting piece investigated and bound separately. This kind of divide and conquer approach is a very useful tool tool in analysis. Another thing we may notice is that in estimating things bounded terms can be treated as constants in appropriate sense. This useful here, and is in fact a repeated theme in analysis.

A one sentence summary may be “express the difference of the products in terms of difference of the terms, and then use divide and conquer”.

Now we can go and see if another approach is possible, whether the argument or write-up can be improved. Or we could decide that we did enough with this problem and move on to another one. Rarely, if ever is a problem exhausted on the first or the second pass.

2.1.2 Strategies.

I list some common ones, but an exhaustive inventory is hardly possible.

The primary, and probably most important strategy has little to do with mathematics per se, but has a lot to do with problem solving in general. It is the strategy of mental toughness.

A quote form one of the greatest mathematicians of the 20th century, Alexandre Grothendieck, illustrates this notion:

If you think of a theorem to be proved as a nut to be opened, so as to reach “the nourishing flesh protected by the shell”, then the hammer and chisel principle is: “put the cutting edge of the chisel against the shell and strike hard. If needed, begin again at many different points until the shell cracks - and you are satisfied”.

Similar thought is echoed in a Polya quote cited by Zeitz:
That is the way to solve problems. We must try and try again until eventually we recognize the slight difference between the various openings on which everything depends. We must vary our trials so that we may explore all sides of the problem. Indeed, we cannot know in advance on which side is the only practicable opening where we can squeeze through. The fundamental method [...] is the same: to try, try again, and to vary the trials so that we do not miss the few favorable possibilities. [...] A man need not throw himself bodily against the obstacle, he can do so mentally; a man can vary his trials more and learn more from the failure of his trials.

Of course many times one or two strikes may be enough, but sometimes (more and more often as you go on in math), mental toughness becomes a core strategy. Now as for the other strategies:

(i) “Get your hands dirty” - just try stuff. Write some immediate consequences of the hypothesis. Follow your nose.

A substrategy is “try an example” - if the problem is about “a function” try \( f(x) = x \), or \( \sin(x) \) or a step function, or \( f(x) = 0 \) for rational \( x \) and \( f(x) = 1 \) for irrational \( x \) etc etc. Your choice of examples would be driven by the proble.

As a substrategy of that, if the problem involves an integer \( n \), try small cases \( n = 1, 2, \ldots \). If it involves a sequence, write out a few terms.

Try the contrapositive† Try proof by contradiction. Do things look easier that way?

(ii) “Don’t start from scratch” - do you know any theorems or lemmas that may be relevant? Do they help? Reading the book and lecture notes and related problems in the problem section often helps.

(iii) “Penultimate step” - what would imply the conclusion in one step. More generally, “think backwards” - what would imply the thing that implies the conclusion? And so on. One has to be careful that the implications go the right direction, but this is an extremely powerful strategy.

(iv) Restate the problem. Part of is writing out definitions, but there are other ways. Another strategy is to use words instead of symbols. I find “eventually” easier to understand than “there exists an \( N \) such that for all \( n > N \) ... and “arbitrarily small” easier than “\( \varepsilon \) for any \( \varepsilon > 0 \) “. Other examples abound.

(v) Draw a picture! - what is says. Another form of restatement. This can do wonders. Be careful that you use the picture as inspiration/to generate insights, and not as substitute for proofs.

†This is the method of proving \( A \) implies \( B \) by showing (not \( B \)) implies (not \( A \)). If you have not heard about it before you may benefit from looking it up.
(vi) Make a diagram - a personal favorite. I put the hypothesis to the left, and conclusion on the right. Any relevant theorems/lemmas/definitions get a node that floats close to whatever it is they are related to. Anything that is implied by some statement or group of statements already on the diagram gets an arrow into it from the relevant nodes. The game is then to construct a directed path from the hypothesis and theorems/properties/lemmas to the conclusion. This strategy allows you to visualize which links are missing and what the overall shape of the argument is. It also comes in handy when it’s time to write the proof (see Section 2.2).

(vii) “Make it easier” - assume something you are not allowed and see if it helps. For example, in proving that $s_n t_n$ converges to $st$ in lecture we first assumed that $s = 0$ - for no other reason than it made the problem much simpler. We used this simpler problem to gain insight - we learned that we can treat the bounded term $t_n$ as a constant - which made our life easier when we tried the general case.

(viii) Wishful thinking (similar to “make it easier”) - suppose you had some property or fact and see what would happen. The fact or property may be something you believe true (like “convergent sequences are bounded”) or something you know to false (like “$A = t_n$ actually is constant). Then see what about this property you are really using, or see if that is likely true and whether you can prove it.

(ix) Run through all the strategies tactics and tools again. Any time your understanding of the problem improves, any time progress has been made a strategy or a tool that seemed irrelevant before may turn out to be useful or even key.

2.1.3 Tactics.

(i) Extreme principle. Given a set, look for its member with some extreme property - in analysis, often the smallest or the largest. This animates the definitions of $\sup$ and $\inf$, but is also useful in other situations. For example, when proving that for each $x$ there is an $n$ with $n \leq x < n + 1$ it is good to consider the smallest element of the set $\{k \in \mathbb{N} | x < k\}$.

(ii) Symmetry - the property of remaining invariant under certain changes, like the fact that $|a|$ does not change if we replace $a$ by $-a$. If we are proving, say $|a| + |b| \geq |a + b|$ we may notice that replacing $a$ by $-a$ and $b$ by $-b$ leaves the inequality to be proved the same, so we may assume that $a \geq 0$ in our proof without loss of generality.

(iii) Invariants - an aspect of the problem that does not change, even as other properties change (note that symmetry is a type of invariance). An exam-
ple is that given by the fact that if \( s_n \) converges, then \( t_n = s_n + 1000 \) also converges and to the same limit.

(iv) Equivariance - a lesser known but often used sister of symmetry. Literally means “changing equally”. For example, if \( s_n \) converges to \( s \), the sequence \( -s_n \) converges not to \( s \) (which would be invariance), but to \( -s \), which is equivariance. Similarly, if for a subset \( S \) of \( \mathbb{R} \) we define its translate \( S+x = \{ s+x \in \mathbb{R} | s \in \mathbb{R} \} \) then the supremum of \( S+x \) is not \( \sup S \), but \( (\sup S)+x \). Equivariance allows you to transport the problem to a new, simpler problem, and then transport it back. We used this to show that any set of integers bounded below has a minimum (by transpoing them to natural numbers, where we have the well-ordering principle). In some sense all limit theorems can be thought of as expressing equivariance.

2.1.4 Tools.

These constitute much of the knowledge in the specific area.

Some of the tools appear on the “Tricki” website [http://www.tricki.org/](http://www.tricki.org/), which is a good, though unfortunately rather spotty resource. Some relevant pages are real analysis front page and in particular the page on solving “routine” problems in analysis and pages about convergence of sequences. Here are some of the tools from that site and some others.

(i) “Give yourself an epsilon of room.” This has many manifestations. One is the following:

Let \( a \) and \( b \) be two real numbers. Then \( a \leq b \) if and only if \( a < b + \epsilon \) for every \( \epsilon > 0 \).

(ii) “Divide and conquer” - to control some term express it as a sum or product (or sum of products) of other terms and control each separately. This came up over and over in the proofs of limit theorems.

(iii) Bounded multiplicative terms can be treated like constants, and so don’t matter. This is expressing things like the fact that if \( s_n \) gets small then \( 1000s_n \) also gets small and even \( t_n s_n \) gets small as long as \( |t_n| < 1000 \). Also came up repeatedly in limit theorems, and should come up again and again.

(iv) Look at a tail and initial part separately, and the initial part often does not matter. Many analysis statements say things like “for all \( n > N \)”. You can divide your sequence into \( n < N \) part and \( n > N \) part. The \( n < N \) part is finite, and so often does not matter (sometimes because it is bounded and so the “bounded things don’t matter” principle applies). For example, we used this split to show that convergent sequences are bounded.

---

\( ^{‡} \) “Control” in analysis means a number of things like “show that it is bounded” or “show it goes to zero” or “show it does not grow too quickly” or “show grows at least as fast as some other fixed expression”, depending on the circumstances.
(v) Convert “every s” into a single arbitrary $s$. Often useful in boundedness/suprema problem. For example, to show that if $S \subset T$ then $\sup S \leq \sup T$ we want to show that all $s \in S$ are $\leq \sup T$. Of course we pick one arbitrary $s \in S$ and see that $s \leq \sup T$.

(vi) If, among the statements you are allowed to assume, you have a statement that begins “for any $\varepsilon$ then make a sensible choice of $\varepsilon$. Usually, there will be a natural choice. Sometimes this choice will only be apparent afterwards.

(More generally, if you want to use an assumption that starts with “for any $x$” then you need to make a sensible choice of $x$).

In all limit theorems, we are given $s_n$ converges, that is for all $\varepsilon > 0$ we have $|s_n - s| < \varepsilon$, and in the proof we choose $\varepsilon$ to be some function of other $\varepsilon$’s.

(vii) Special case of above: divide your epsilon into parts (works with conjunction with divide and conquer). Refers to things like $\varepsilon = \varepsilon_1 + \varepsilon_2$ where $\varepsilon_1$ and $\varepsilon_2$ bound terms that add to the term we want to control.

The advanced version is to note that each epsilon is actually infinitely many(!) epsilons: $\varepsilon = \varepsilon/2 + \varepsilon/4 + \varepsilon/8 + \ldots$

2.2 Writing proofs.

After some practice with proof-writing, a well-thought out argument should not be hard to write down.

It is important to have basic mathematical literacy. Part of it is simply knowing algebra an arithmetic and being careful in their use. Another part is more abstract, but not too deep. This includes basic notions about sets: membership, $\in$, $\subset$, union of sets, intersection of sets, complement of a set. Basic notions of logic and quantifiers: implies, negation, converse, contrapositive, equivalence, for any, there exists, unique. If any of these terms are unfamiliar - look them up. Internet is a good start; I also recommend the book “Reading, Writing and Proving” which covers a lot of this material (and some real analysis as well).

As a general approach, I recommend the following:

Write down a plan - a less detailed overview of the proof indicating main steps. Better yet, start with making a diagram indicating how these steps relate to each other. See the last page, Appendix A for an example. That diagram was made using yEd, but you can just draw such diagrams by hand. Then you can “project this diagram into a linear text” - decide what to write in what order, sort of like an outline of a paper (in fact, when I write a math paper, this is what I do: make a general plan of which sections there are, then as I write each section make a plan for what lemmas, propositions and theorems will go in in what order). For the example of the product theorem the result may be (it may be more or less detailed, whatever you need to write a proof afterwards):

(i) Let $\varepsilon > 0$
(ii) Pick $\varepsilon_1 B_1 + \varepsilon_2 B_2 = \varepsilon$.

(iii) $t_n$ converges, $|t_n| < B_1$.

(iv) $B_2 = |s|$.

(v) $s_n$ converges to $s$, there exists $N_1$ such that for all $n > N_1$ we have $|s_n - s| < \varepsilon_1$.

(vi) $t_n$ converges to $t$, there exists $N_2$ such that for all $n > N_2$ we have $|t_n - t| < \varepsilon_2$.

(vii) $|ts_n - st| = |t(s - s_n) + s(t_n - t)|$

(viii) $|ts_n - st| \leq B_1 \varepsilon_1 + B_2 \varepsilon_2 = \varepsilon$

(ix) QED

After that you simply write the proof in that order filling in details of each step (and correcting things as needed).

There is not a hard rule about how much detail to supply. The standard is that justification for each step should be obvious for the reader. As you know, what is obvious varies from person to person; err on the side of more details. Imagine presenting a statement to another student in the class (perhaps one who is not too friendly) and see if they would think it is obvious. If in doubt, go one step further in justifying. No one ever got points deducted for being too clear.

Here are some more tips, based on “Reading, Writing, Proving”.

(i) Introduce your notation explicitly. Say things like “let $\varepsilon > 0$ be a real number” or, assuming a convergent sequence $s_n$ has been introduced previously, “Denote the limit of $s_n$ by $s$”. Do this before you use the corresponding variable. This is surprisingly effective in spotlighting issues with your proof.

A variable that is used without introduction often signals a missing step.

(ii) Choose notation carefully.

(iii) A variable should only be assigned one meaning. If $S$ is a subset of $\mathbb{R}$ do not use $S$ to denote a number or (even worse) an element of $S$.

(iv) Do not string too many symbols together in a row. This includes things like “$n > 0, r < 1$” (use something like “$n > 0$ and $r < 1$”); it also includes starting any sentence with a symbol.

(v) All rules of grammar apply. Sentences should start with capital letters and end with periods. Avoid fragments. This is not some pedantic grammar nazi injunction - a fragment is an incomplete mathematical thought; the reader will most likely be confused.
(vi) Explicitly reference all properties and results you are using. It can be by their “nicknames” - “limit theorem for addition”, “compatibility of $<$ with multiplication” etc, or by their number in the book or in lecture notes. This has two purposes - convinces reader that your reasoning is sound, and trains you to see what it is you are using in the proof - both to avoid using what you don’t have, and to notice when a simpler proof may be possible.

(vii) Explicitly indicate the structure of your proof. It is not a bad idea to write a short outline (in one or two sentences) of the proof before the proof itself. Guide the reader with phrases like “Assume A. We shall now show B” and “Suppose A. We shall see that it leads to contradiction”.

(viii) You can give your proof to a classmate to read, to see if they understand you. It is nice to offer to read one of their proofs in return.

3 Concluding remarks - thinking mathematically.

“Think deeply about simple things.”

Ross Mathematics Summer Program motto

If you are one of the people who are not philosophically inclined and whose eyes glaze over when they encounter musings on “depth” and “ways of thinking” you may want to stop reading now. For the rest of you are these remarks.

Math in general and analysis in particular may strike you as anything but simple. But from a certain perspective much of math is simple. One aspect of this simplicity is control of your environment. When you are operating within a formal mathematical system, your assumptions are clearly delineated, everything else is proved. You may have felt this as a constraint while working on the first homework for this class, and relying only on ordered field axioms, yet once you get used to it, it can be quite liberating. Questions have right and wrong answers. Compared with questions like “how do I design this house?” or “how much anesthetic should I administer to this patient?” or “what should we do about the economy?” the question of “what is the limit of $(1 + \frac{1}{n})^n$" seems positively straightforward.

Another aspect comes from the fact that mathematics stems from simple questions with deep implications: what is a number? what is infinity? how do we measure area? what is continuity? what structure one needs to do arithmetic? to measure distances? - etc. etc. Thinking deeply about these questions underlies much of modern mathematics. It may seem intimidating - there is always something else you don’t know, some more math to learn - but it does not have to. You

\[ \text{We have seen something of a number line in the class so far. There are still some more surprises even in real numbers - for example that a random real number can not be written down! Indeed the set of things that can be written down is infinite but countable, while the set of real numbers is uncountable. A random real number will land outside of the “writable” set! No wonder they call some numbers transcendental.} \]
can discover interesting and useful things at each stage, and enjoy the process of
doing so.

So we should endeavor to think deeply and clearly. Messy thinking may suf-
fice as a start, but the deeper you go the harder it is to continue, unless you have
clarity. You can probably follow a two-sentence explanation even if it’s written in
a haphazard manner, but for a 20-step proof some structure is necessary. To read
mathematics is to comprehend the internal structure of mathematical statements
and their relationships to each other. To write mathematics - and mathematical
proofs in particular - is to recreate this structure on a page. Proofs are like ma-
chines - various pieces have to fit with each other in the right way, or the whole
contraption would fail. There is often more than one machine that does the same
thing, but just switching one cog in a given apparatus will likely break it. Similarly,
there are many proofs for various mathematical statements, but each one must be
carefully constructed. Learning to do this is the art and craft of proving.

References

[DG] Ulrich Daepp and Pamela Gorkin. Reading, writing, and proving: a closer


[Zeitz] Paul Zeitz. The Art and Craft of Problem Solving. John Wiley and Sons,
A The Diagram