Math 79SI Notes

Mason Rogers

7 Common Proof Mistakes and Good Communication Habits

7.1 Background

In the previous sections, we looked at four primary proof techniques: direct proof, proof by contrapositive, proof by contradiction, and mathematical induction. We saw several examples using each technique to construct a logically sound proof. In practice, however, coming up with correct and well-articulated proofs is a more difficult task than understanding examples that are known to be correct. It is easy to make mistakes in reasoning, and even if the idea behind a proof is correct, writing the argument precisely and coherently is a challenge. We therefore dedicate this section to avoiding common mistakes and developing good mathematical writing habits.

Different techniques are prone to different sets of mistakes. Here is a brief list of common errors as well as when they are most likely to arise.

- Circular argument—This is assuming what you are trying to prove. It can vary from compressing the bulk of the statement you are trying to prove into a bit of misapplied notation to visibly having result A depend on result B and result B depend on result A. An example of the latter type would be using Euclid's lemma ($p \mid ab \implies p \mid a$ or $p \mid b$) to prove the fundamental theorem of arithmetic (uniqueness of prime factorizations), then using the fundamental theorem of arithmetic to prove Euclid's lemma (which is what experience and intuition might lead one to think).

- Converse fallacy—The mistake of proving the converse of a statement rather than the statement itself can occur for a number of reasons, including incorrectly taking a contrapositive and trying to translate an intuitive "work backwards" argument into a proof.

- Same proof twice—The opposite of the converse fallacy can occur when proving an 'if and only if' statement, where it is necessary to take the converse but possible to take the contrapositive accidentally.

- Incorrect base case—In induction arguments, it is important to check that the number of verified base cases matches the required inputs to the inductive step.

- Special case mishandling—Frequent instances of this fallacy include dividing by 0, taking "real" square roots of negative numbers, or even failing to account for a number having two distinct square roots. This mistake is the traditional source for gaps in geometry proofs (e.g. forgetting degenerate cases of a geometric construction) and can often be resolved with a bit of careful casework.

- Sloppy symbol usage—Symbols that have special properties in some cases may not have those properties in all cases. For instance, $\sqrt[n]{x^n} = \sqrt[n]{y^n}$ does not imply $x = y$ for all n; consider $x = -1, y = 1, n = 2$.

- Mismatching outputs and inputs—If you want to use a previously-proven result, you need to ensure that the case at hand satisfies all of its hypotheses.

- **Imprecise language**—Intuition is a double-edged sword; it is helpful in brainstorming how to prove a statement, but attempting to write a proof based on how the details "should" work out can sometimes be dangerous. You need to apply relevant definitions and theorems precisely, not try to describe roughly what they mean and work with vague wording. In general, if your proof does not genuinely use the definitions of the relevant mathematical terms or invoke previous results resting on those definitions then it is not precise enough.

- **Logical leaps**—Writing a proof is generally more difficult than deciding whether or not a statement is plausible. Be careful not to assert that statement B follows from statement A unless you have a proof. There are many examples of "obvious" statements that are totally false. For example, for a sequence of differentiable functions $f_n : \mathbb{R} \to \mathbb{R}$ which converge at every x as n tends to ∞ to the function $f(x) = \lim_{n\to\infty} f_n(x)$, it is not generally true that $\lim_{n\to\infty} f'_n(x)$ exists, and even if it does, it is not generally equal to $f'(x)$. In fact, $f(x)$ may not even be differentiable!

In addition to thinking about what technical mistakes to avoid, it is also helpful to think of writing habits that allow you to communicate your work efficiently to others. At the end of the day, in addition to answering questions, writing good proofs allows mathematicians to communicate their ideas to each other. Here is a list of some writing tips.

- Aim for clarity—Extreme usage of symbols (e.g. $\forall x \in \mathbb{R}_{\geq 0}, \exists y \in \{s \in \mathbb{R} : s^2 = x\}$) or of words (e.g. for all nonnegative real numbers, there exists another real number whose square is the first real number) makes math difficult to read. In general, use a mixture of both that makes both the big picture and the technical details easy to follow (e.g. for all $x \in \mathbb{R}_{\geq 0}$, there exists $y \in \mathbb{R}$ such that $y^2 = x$).

- Remember the audience—The reader will not know what x means unless you tell them what x means. Additionally, the amount of assumed background should depend on who will be reading your work.

- Cue the reader—After spending a long time solving a problem, you may take certain insights for granted that a reader may not realize. For complicated proofs, it is helpful to indicate in non-rigorous prose the intuition behind what you will be doing as well as what the main steps are. You can then return to the big picture after each step to give the reader a sense of where they are in the proof. Transition words and motivating exposition go a long way in making your work readable. As an analogy, it is much easier to drive from point A to point B if you can see the route on a map before embarking rather than just trying to listen to a list of directions.

- Solve the problem before writing it up—An eraser-scrubbed sheet with arrows indicating the flow of logic across the page and a few crossed-out paragraphs is difficult to read. Different people have different systems, but I like to solve each problem in a scratch-work notebook before typing it up or writing it out. Then I can reorder the details how I see fit, and I am forced to double-check my reasoning. Similarly, if you type your work, it is often helpful to start with a pencil and paper or a chalkboard, then type your work once you have a solution.

As a final note, mistakes in math are often an essential part of the process of finding a correct answer. Failing to prove a statement with one approach may inform your approach for the next attempt, and if a completed proof is shown to be incorrect, it is possible that the intuition is still valid and the mistake can be remedied.

7.2 Examples

Rather than a list of examples of correct proofs, in this section we will examine a collection of incorrect proofs and their mistakes. Some will read like serious attempts at proofs, and others will be more like mnemonics for avoiding the mistakes in the future. In each example, we will identify the logical issues as well as the extent to which they can be resolved.

We begin with a famous example of proving a statement that we know experientially to be false.

Proposition 7.1. All horses are the same color.

"Proof". Let n denote the number of horses in a group. If $n = 1$, then there is only one horse, so all n horses have the same color. Now suppose the statement holds for $n = k$ horses, and consider a group of $k+1$ horses. Choosing a horse A arbitrarily and removing it from the group leaves k horses, which by the inductive hypothesis must have the same color. Returning A and removing a different horse B from the group then yields another group of size k, so all remaining horses must have the same color. Because A and B have the same color as all of the rest of the horses, all $k+1$ horses must have the same color. We conclude that any group of n horses must have just one color, so all horses are the same color. □

Clearly this this is false; you likely have seen multiple different colors of horses before. Broadly, the argument fails because the required base case is $n = 2$, not $n = 1$. To see this, consider the line "because A and B have the same color as all of the rest of the horses". For this to work, the set 'all of the rest of the horses' must be nonempty, for which we need $k + 1 \geq 3$. The inductive step therefore depends on the $n = 2$ case (which is clearly false) rather than just the $n = 1$ case.

Of course, our experiences with different colors of horses in the past make us doubt this argument from the beginning. Sometimes, however, the same flaw is not so detectable. Recall the Fibonacci numbers $\{1, 1, 2, 3, 5, 8, \ldots\}$, where $F_n = F_{n-1} + F_{n-2}$ for $n \geq 3$.

Proposition 7.2. Let ϕ denote the Golden Ratio $(1+\sqrt{5})/2$ and let $\psi = -\phi^{-1}$. Then the n-th Fibonacci number F_n is given by $\phi^n + \psi^n$.

"Proof". Observe that ϕ and ψ are roots of the polynomial $p(x) = x^2 - x - 1$. Because ϕ satisfies the equation $\phi^2 - 1 = \phi$, when $n = 1$, we have $F_1 = 1 = \phi - \phi^{-1}$ as desired.

Now suppose $F_k = \phi^k + \psi^k$ for all $k < n$. Then

$$F_n = F_{n-1} + F_{n-2} = \phi^{n-1} + \phi^{n-2} + \psi^{n-1} + \psi^{n-2} = \phi^n + \psi^n,$$

so we are done. □

The issue here is that to prove the n-th case, both the $n-1$ and $n-2$ cases are required. We therefore need $n - 1 \geq 1$ and $n - 2 \geq 1$ for the above calculations to make sense, which is to say $n \geq 3$. This shows that the case $n = 2$ also merits special attention. However, at the outset, we only checked a single base case. The formula given would yield $F_2 = \phi^2 + \phi^{-2} = 3 \neq 1$ and is therefore incorrect.[1]

Our next example is a computation gone awry.

Proposition 7.3. $-1 = 1$.

Proof. Let $i = \sqrt{-1}$. We begin with

$$\frac{-1}{1} = \frac{1}{-1}.$$

Taking the square root of both sides, we obtain

$$\sqrt{\frac{-1}{1}} = \sqrt{\frac{1}{-1}},$$

which we can simplify to reveal

$$i = \frac{\sqrt{-1}}{1} = \frac{1}{\sqrt{-1}} = \frac{1}{i}.$$

Multiplying through by i then reveals $-1 = 1$, as claimed. □

It may be tempting to suggest that the issue arose from taking the square root, but since we established our choice of $\sqrt{-1}$ at the outset, we avoided any inconsistencies in what we meant by $\sqrt{\ }$. Instead, we lazily applied to negative numbers the fact that for *positive* real numbers a and b, $\sqrt{a/b} = \sqrt{a}/\sqrt{b}$. In doing so, we created a choice between putting the negative sign in the numerator or the denominator, which affected where the factor of i appeared. More specifically, the meaning of $\sqrt{a}/\sqrt{b} = \sqrt{a/b}$ in $\mathbb{C}$ is that for *choices* of $\sqrt{a}$ and $\sqrt{b}$, $\sqrt{a}/\sqrt{b}$ is one of the square roots of a/b, but it is not the only one.

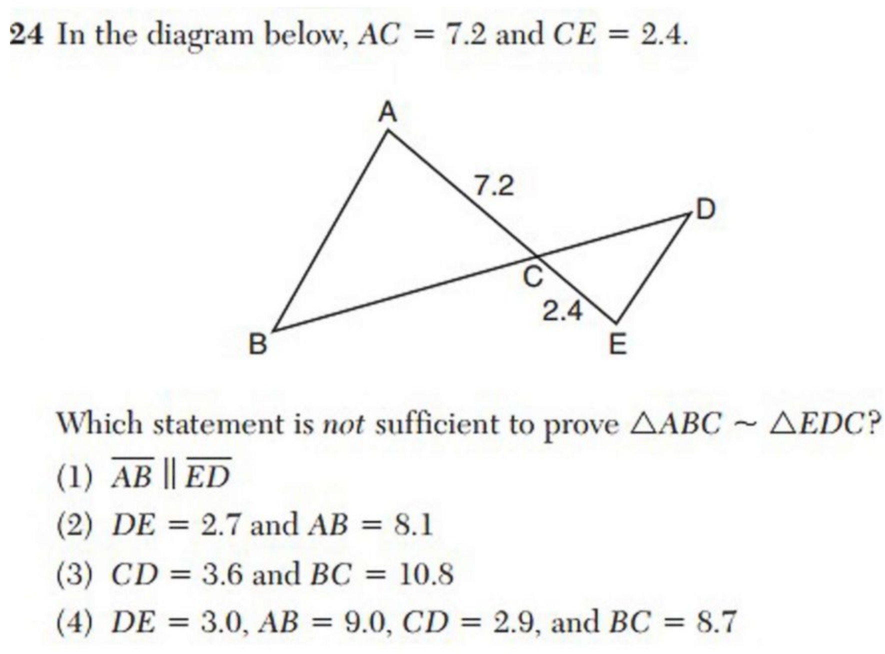
Sloppy use of the $\sqrt[n]{\ }$ symbol resulted in an error in the original proof of an important result in number theory concerning how to detect if a number is an n-th power in modular arithmetic, now known as the Grunwald-Wang theorem. First "proven" by Willhelm

[1] The formula given is actually the closed form of the Lucas numbers L_n, which follow the Fibonacci recursion $L_n = L_{n-1} - L_{n-2}$ but satisfy $L_0 = 2, L_1 = 1$. There is interesting interplay between the Fibonacci numbers and the Lucas numbers, and they often appear in combinatorial formulas for each other.

Grunwald in 1933, Shianghao Wang found a counterexample in 1948. The counterexample helped Wang identify the error in the original proof rather quickly, and within 2 years, Wang released a corrected version of the theorem handling all cases. Most of the reasoning in the original statement was correct, but modifications were needed to the formulation and proof of the theorem to make it correct.

The next example from the New York Regents Exam in 2017 can perhaps be seen as the opposite of being too sloppy.

Problem 7.4. *See Figure 1.*


24 In the diagram below, $AC = 7.2$ and $CE = 2.4$.

Which statement is *not* sufficient to prove $\triangle ABC \sim \triangle EDC$?

(1) $\overline{AB} \parallel \overline{ED}$

(2) $DE = 2.7$ and $AB = 8.1$

(3) $CD = 3.6$ and $BC = 10.8$

(4) $DE = 3.0$, $AB = 9.0$, $CD = 2.9$, and $BC = 8.7$

Figure 1: Problem 24 from the 2017 New York Regents Examination in mathematics.

The intention of the problem was to test students' abilities to recognize that the Side-Side-Angle (SSA) property is insufficient to prove that two triangles are congruent (or similar if the side lengths are merely proportionate). Hence, answer choice (2) was the "correct" answer according to the Regents board. However, with the side lengths given in choice (2) (specifically that $AB = 8.1 > AC$), it is possible to prove that angles B and D are both *acute* as drawn in the picture; this extra bit of information resolves the issue with concluding congruence from SSA presented in Figure 1 of section 3. In the case of acute triangles, SSA is enough to deduce similarity, so no answer choice was actually correct.[2]

An example of an incorrect proof that gives the correct answer is sometimes seen to justify the chain rule in first year calculus classes.

[2] The error with the problem was originally discovered by high school student and geometry tutor Benjamin Calfato, who submitted a proof that the information in answer choice (2) was enough to prove $\triangle ABC \sim \triangle EDC$. The board conceded that he was correct but initially refused to rescore the exam, no doubt in part because the exam had already been regraded due to another correction. This led to a back-and-forth between Calfato and the Regents board about whether the scoring should be changed, which Calfato ultimately won.

Proposition 7.5 (*Chain rule*)**.** If $f, g : \mathbb{R} \to \mathbb{R}$ are differentiable functions, then $(g \circ f)' = (g' \circ f) \cdot f'$.

"Proof". We write

$$\frac{dg}{dx} = \frac{dg}{\cancel{df}} \frac{\cancel{df}}{dx} = \frac{dg}{df}\frac{df}{dx}.$$

□

It should seem alarming that the definition of the derivative of $g \circ f$ is never invoked in the given argument. While the 'fraction' notation for differentiation is well-motivated, differentials do not have an honest definition in single variable calculus (and the framework in differential geometry that *does* give a useful meaning of differentials *uses* the chain rule!). The proof ignores the actual mathematical content of the statement by wrapping it up into notation, then abusing the notation.

A similar example gives a slick "proof" of the Cayley-Hamilton theorem.

Theorem 7.6 (*Cayley-Hamilton*)**.** For an $n \times n$ matrix A, define $p(t)$ to be the characteristic polynomial $\det(tI_n - A)$. Then $p(A) = 0$.

"Proof". We write

$$p(A) = \det(AI - A) = \det(0) = 0.$$

□

We define p as a polynomial in the *scalar* unknown t, then want to say something about p as a polynomial in the *matrix* unknown A. But without being clear about what we mean by allowing p to take a matrix unknown, it is tempting to assume that we can substitute A in for t in the formula given to us for $p(t)$. By definition, $p(A)$ is the polynomial whose A^i coefficient matches the t^i coefficient of $p(t)$ (e.g. if $p(t) = 3t^2+t+4$ then $p(A) = 3A^2+A+4I_n$). However, $\det(AI - A)$ is just the number 0 with no content.

We have now seen examples of sloppy proofs that give wrong answers and sloppy proofs that give right answers. It is also worth examining instances in which the proof given is the correct proof of some mathematical statement, but not the intended one.

Proposition 7.7. For natural numbers a and b, ab is odd if and only if a and b are both odd.

"Proof". First, suppose a and b are odd, so $a = 2k + 1$ and $b = 2\ell + 1$ for some $k, \ell \in \mathbb{Z}$. Then $ab = (2k + 1)(2\ell + 1) = 2(2k\ell + k + \ell) + 1$, which is odd.

Now suppose ab is even, so $2 \mid ab$. Then by Euclid's lemma, $2 \mid a$ or $2 \mid b$. Thus, either a or b is even, and we are done. □

Both proofs, while a bit terse, are correct proofs of the implication 'if a and b are odd then ab is odd'. However, the other direction 'if ab is odd then a and b are odd' is never addressed. Thus, the proof of the proposition is incomplete.

We conclude with one more example that is more subtle than the previous examples. The statement we will "prove" is true in many easily imagined cases, and coming up with am explicit counterexample is hard if you do not know where to look.

Proposition 7.8. Let $f_n : \mathbb{R} \to \mathbb{R}$ be a sequence of differentiable functions such that for each $x \in \mathbb{R}$, $f_n(x)$ is a convergent sequence. Define $f(x) = \lim_{n\to\infty} f_n(x)$. Then for all $x \in \mathbb{R}$, $f'(x) = \lim_{n\to\infty} f_n'(x)$.

"Proof". We use the limit definition of the derivative to write

$$f_n'(x) = \lim_{h\to 0} \frac{f_n(x+h) - f_n(x)}{h},$$

and we therefore have

$$\lim_{n\to\infty} f_n'(x) = \lim_{n\to\infty} \lim_{h\to 0} \frac{f_n(x+h) - f_n(x)}{h}.$$

Because $\lim_{n\to\infty} f_n(x+h) = f(x+h)$ for any $x, h \in \mathbb{R}$, we have

$$\lim_{n\to\infty} f_n'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h},$$

and the right side is $f'(x)$ by definition. □

The issue with this "proof" is that we tacitly swapped the order of the limits, which is not allowed in general. There are many theorems stating specific cases in which different limits may be taken in any order, but they often require more hypotheses than one might intuitively expect. Even though it may not seem like there is not a reason to avoid switching the limits in question, a *proof* is necessary to justify doing so in any particular case.

We have now seen a number of different flawed proofs, where the flaws arise for many different reasons. In your own proof writing, it is important to watch out for these potential pitfalls, as well as any others that you learn to avoid through experience.

7.3 Exercises

1. Find the error in the following "proof" that $n = 0$ for all $n \in \mathbb{N}$.

   Problem: Prove that $n = 0$ for all $n \in \mathbb{Z}$.[3]

   "Proof": Choose $a \in \mathbb{Z}$ and let $b = a$. Multiplying both sides of $a = b$ by a gives $a^2 = ab$. Subtracting b^2 from both sides gives $a^2 - b^2 = ab - b^2$. The left side factors as $(a+b)(a-b)$, while the right factors as $b(a-b)$. We thus divide through by $a - b$ to reveal $a + b = b$, which simplifies to $a = 0$.

2. Find and explain any errors in the following false proof invented by Thomas Clausen:

   Problem: Prove for all $n \in \mathbb{Z}$ that $e^{-4\pi^2 n^2} = 1$.

   "Proof": For all $n \in \mathbb{Z}$, $e^{2\pi i n} = 1$, so $e^{2\pi i n + 1} = e$. We therefore have

$$e^{(2\pi i n+1)^2} = \left(e^{2\pi i n+1}\right)^{2\pi i n+1} = e^{2\pi i n+1} = e.$$

[3] The same argument also "works" for all $z \in \mathbb{C}$.

But $(2\pi in+1)^2 = 1+4\pi ni-4\pi^2n^2$, so for all n we have

$$e^{1+2\pi(2n)i-4\pi^2n^2} = e.$$

Simplifying gives

$$e^{1+2\pi(2n)i}e^{-4\pi^2n^2} = e\cdot e^{-4\pi^2n^2} = e,$$

from which we deduce that $e^{-4\pi^2n^2} = 1$ for all $n \in \mathbb{Z}$.

3. View the Numberphile video (5:19) 'All Triangles Are Equilateral' (link listed in the Further Reading Section 7.4). What is the main error in their "proof"?

4. Choose two proof-based exercises from previous sections for which your original proof was correct, but you feel that now you could articulate your reasoning better. Write up a new, easier-to-read version of the proof, and summarize any changes you made.

5. Choose two proof-based exercises from previous sections for which your original proof was incorrect. Correct it and describe what changes were required. Were you able to salvage your old intuition, or did you have to begin anew?

6. Google 'Andrew Wiles' and 'Proof of Fermat's Last Theorem', and read about the developments between the original announcement of his proof to its correction and eventual acceptance. This is an excellent example of how learning why one approach fails provides insight as to why another approach may succeed. (*Optional: View the Nova documentary on Wiles' proof of Fermat's Last Theorem, linked to in the Further Reading Section 7.4.*)

7.4 Further Reading

CPZ Chapter 0	More about good writing habits in mathematics
"Primer on Proof", Jenny Wilson	An excellent, to-the-point introduction to proof writing which contains a list of common mistakes to avoid. web.stanford.edu/~jchw/PrimerOnProof.pdf
"All Triangles Are Equilateral", Numberphile	A tricky "proof" that all triangles are equilateral, one of many fun Numberphile videos. youtube.com/watch?v=Yajonhixy4g
"The Proof", Nova documentary dir. Simon Singh	An interesting documentary (50 min) on the proof of Fermat's Last Theorem, a famous longstanding open problem in number theory answered by Andrew Wiles in 1994. http://www.dailymotion.com/video/x1btavd