HW3 Clarification

Mason Rogers

April 26, 2018

This week, I'd like to focus on exercise 3.3.2 and unpack a few difficulties that people had with this problem. I'll begin with a bullet list of some common errors, then walk through the process of thinking about how to solve the problem and writing a proper solution.

Problem. The following statements were proven for functions $f : X \to Y$ and $g : Y \to Z$ as a part of the previous section's exercises. Write their converses and prove or disprove those.

a. If f and g are injective functions, then $g \circ f$ is an injective function.

b. If f and g are surjective functions, then $g \circ f$ is a surjective function.

Moreover, consider the related statement 'if f is injective and g is surjective, then $g \circ f$ is bijective'. Prove or disprove it, then write its converse and prove or disprove that.

Common Errors

1. Writing the contrapositives of (a) and (b) instead of their converses and thus proving the wrong things
2. Using imprecise language instead of definitions and thus falling short of a rigorous argument, leading to error 3
3. Insisting on proof rather than disproof and thus making logical leaps to prove false statements

These mistakes are possible to make in almost any proof, so we will point out where they pop up to learn from them.

Step 1: Correctly identify what needs to be proven or disproven.

The first step of writing any proof, regardless of strategy, is writing out exactly what we need to prove or disprove. The problem asks us to prove the *converses* of statements (a) and (b) rather than the statements themselves.[1] Recall that for a statement of form 'if P then Q', the converse is the statement 'if Q then P'. This is *not* equivalent the contrapositive 'if

[1] If you never proved the statements themselves, it might be a good exercise. You can use the strategy outlined in this document to guide you.

not Q then not P'; several examples are given in section 3.1 if you want to see more detail as to why.

Taking the converses of statements (a) and (b) gives the following for functions $f : X \to Y$ and $g : Y \to Z$.

a'. (*Converse of (a).*) If $g \circ f$ is injective, then f and g are injective.

b'. (*Converse of (b).*) If $g \circ f$ is surjective, then f and g are surjective.

Note that we have now avoided common error 1.

Step 2: Use precise definitions to clarify what you need to show.

If we want to prove something involving injectivity and surjectivity, we had better *use the definitions* of injectivity and surjectivity, which are mentioned in exercise 2.3.1 and 2.3.3 respectively. Otherwise, we will use vague language like "$g \circ f$ maps only one input to each output" or "$g \circ f$ hits every element of the range". Although such intuitions are helpful for thinking about our argument, they *are not rigorous* and must be translated into precise statements in our proof. We thus reformulate:

a'. (*Precise, manipulable restatement of (a').*) If $g(f(x_1)) = g(f(x_2))$ implies that $x_1 = x_2$ for all $x_1, x_2 \in X$, then $f(x_1) = f(x_2)$ implies that $x_1 = x_2$ and $g(y_1) = g(y_2)$ implies that $y_1 = y_2$ for all $x_1, x_2 \in X$ and $y_1, y_2 \in Y$.

b'. (*Precise, manipulable restatement of (b').*) If for any $z \in Z$ there exists $x \in X$ such that $g(f(x)) = z$, then for any $y \in Y$ there exists $x \in X$ such that $f(x) = y$ and for any $z \in Z$ there exists $y \in Y$ such that $f(y) = Z$.

We are now in a position to avoid common error 2.

Step 3: Embark on the proof or disproof attempt.

Now that we know what we need to prove or disprove, we have decisions to make: Should we try for a proof or a counterexample? If proof, should we proceed directly or by contrapositive? Suppose we fail to find a counterexample to (a') after thinking for a minute or two, so we decide to try a direct proof.

Attempted proof of (a'). We need to show separately that f and g are injective, so we begin with f. Suppose that $f(x_1) = f(x_2)$ for some $x_1, x_2 \in X$. We thus have $g(f(x_1)) = g(f(x_2))$.[2] Since $g \circ f$ is injective, *by the definition of injectivity*, we have that $x_1 = x_2$. We conclude that if $f(x_1) = f(x_2)$ then $x_1 = x_2$, which *is by definition the statement* that f is injective.

Now we aim to show that g is injective. (Spoiler alert: This is where the proof attempt is going to go off the rails, but watch carefully to see what temptations to avoid.) Suppose $g(y_1) = g(y_2)$ for some $y_1, y_2 \in Y$. We want to show that $y_1 = y_2$, but it seems impossible to pull the y_1 and y_2 out from g. We thus switch to the contrapositive approach, whereby we

[2] Note that we can say this here because g is a bona-fide function, unlike $\sqrt[3]{\ }$ from last week.

aim to prove that if g is *not* injective, $g \circ f$ *cannot* be injective. *By the definition of injectivity* and the assumption that g is not injective, we can choose $y_1, y_2 \in Y$ such that $g(y_1) = g(y_2)$ and $y_1 \neq y_2$. It may be tempting to write $y_1 = f(x_1)$ and $y_2 = f(x_2)$, conclude by the injectivity of f that $x_1 \neq x_2$, and declare that $g(f(x_1)) = g(f(x_2))$ but $x_1 \neq x_2$. However, we are not guaranteed the existence of x_1 and x_2 such that $f(x_1) = y_1$ and $f(x_2) = y_2$; this would require f to be surjective, which we cannot assume.

We're stuck... $\square$

Step 4: Listen to your difficulties.

This is an insidious difference between computation-based math and proof-based math. In calculus class, if you feel stuck, you can usually do anything remotely reasonable to get an answer. If a step feels suspicious or unjustified, you can try it anyway and see whether it allows you to proceed. But in Math 79SI, if your gut is telling you that you cannot prove why g is injective without doing anything fishy, *you need to listen to it* rather than creatively (by which I mean wrongly) sidestepping the problem.

We've proven that f is injective, but we are having trouble showing that g is injective. We now realize that it might be a good idea to search for a counterexample, *and* from what we learned in our proof attempt, we know that in our counterexample, f still has to be injective. We thus can ponder, "how can an injective function be followed by a non-injective function such that the composition is still injective?" Our attempt at proving the injectivity of g also gives us another clue: We could finish the proof of (a') if f were assumed to be surjective, so we should look for a counterexample where f is not surjective.

Step 5: When looking for a counterexample, think small.

It may be tempting to try to think of examples of continuous functions $f, g : \mathbb{R} \to \mathbb{R}$ because these are the functions with which we are most familiar. But injectivity and surjectivity are very bare-bones properties of functions, and thus it often suffices to work with bare-bones examples of functions when looking for a counterexample. We suggest a counterexample in Figure 1. Additionally, we have successfully avoided common error 3, and we are done with the hard work.

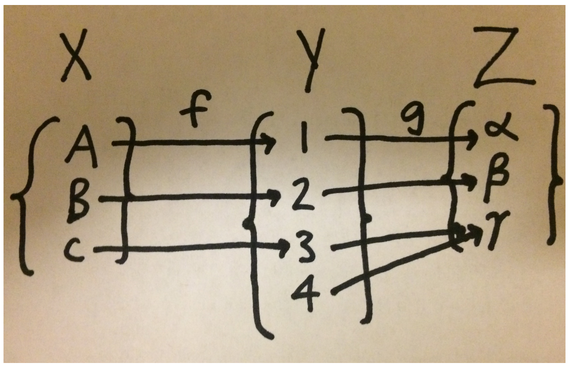

Figure 1: Observe that $g \circ f$ is the bijection $A \mapsto \alpha, B \mapsto \beta, C \mapsto \gamma$, but g is not injective and f is not surjective.

We could then repeat this whole process with statement (b'):

Attempted proof of (b'). We need to show separately that f and g are surjective given the fact that $g \circ f$ is surjective, and we begin with g. Choose $z \in Z$. *By the definition of surjectivity*, we can choose $x \in X$ such that $g(f(x)) = z$. Then $y = f(x) \in Y$ satisfies $g(y) = z$. We can therefore choose some $y \in Y$ for all $z \in Z$ such that $g(y) = z$, which *by definition is the statement* that g is surjective.

We now aim to prove that f is surjective. (Spoiler alert: Not going to happen.) Choosing $y \in Y$, let $z = g(y)$. From *the definition of surjectivity* and the assumption that $g \circ f$ is surjective, we can write $g(y) = z = g(f(x))$ for some $x \in X$. Here, it may be tempting to conclude that $y = f(x)$, but this would assume that g is injective, which is not necessarily true.

We're stuck again. □

If we wanted to come up with a new counterexample, we could ask ourselves "How can a surjective function follow a non-surjective function in such a way that the composition is still surjective?" We also have the other clue that letting g be injective would allow us to finish the proof of (b'), so we look for a counterexample where g is not injective.

Fortunately, Figure 1 is a counterexample to (b') as well.

We conclude that statements (a') and (b') are both false, but we learned something from trying to prove them: If $g \circ f$ is injective, f is injective, and if $g \circ f$ is surjective, g is surjective. This helps us with the last part of the problem:

Problem (Last part)**.** Consider the related statement 'if f is injective and g is surjective, then $g \circ f$ is bijective'. Prove or disprove it, then write its converse and prove or disprove that.

I leave disproving the original statement as an exercise along with the reminder to think small. The converse in question is 'if $g \circ f$ is bijective then f is injective and g is surjective'. I claim we're already done:

Proof. Since $g \circ f$ is bijective, $g \circ f$ is both injective and surjective. We proved above that $g \circ f$ being injective implies that f is injective and that $g \circ f$ being surjective implies that g is surjective. We thus conclude that f is injective and g is surjective, as desired. □

This time, we did not have to do any definitional unpacking *since we already did all of the heavy lifting earlier.*

A final thought

This example is illustrative of what doing math should feel like. We try and often "fail", but it is precisely this "failure" that informs how we may succeed in our next attempt. It is easy to write up a clean argument once you know the solution, and given that we are constantly presented with thoroughly refined proofs in classes and textbooks, it is easy to feel like doing math should be a linear process. However, this is seldom the case.

In general, I think it is a good idea to think through the intuition and the formalism in parallel. Drawing pictures and thinking non-rigorously might give you a strategy, and attempting to formalize that strategy with precise, mathematical language will help you realize how your strategy needs to be modified. At the end of the day, though, make sure that your rigorous argument relies on appropriate definitions and previous results. Then you will have both valuable intuition *and* a correct proof.