

# MATH 22

Lecture N: 10/16/2003

## FUNCTIONS: COMPOSITION & INVERSES

Mad world! mad kings! mad composition!  
—Shakespeare, *King John*, II:1

# Administrivia

- <http://denenberg.com/LectureN.pdf>
- Next Thursday's lecture will be, in part, a review for the next exam. Email questions or topics in advance to [nobodylistening@blackhole.com](mailto:nobodylistening@blackhole.com).
- Warning: Problem 18 of §5.6 continues to (e) on the next column!
- Project 4: Grader pounds me, I pound you:  
**This is your last warning!**
  - You can't manipulate non-equations.
  - “if  $p$ , then  $q$ ” is always true when  $p$  is false.

# A Formal Peculiarity

Consider the function  $f_1 : \mathbb{Z} \rightarrow \mathbb{Z}$  where  $f_1(x) = x^2$ . This function is neither one-to-one nor onto.

Now consider  $f_2 : \mathbb{Z} \rightarrow \{0, 1, 4, \dots\}$  where  $f_2(x) = x^2$ . This function also is not one-to-one, but *is* onto. (We've already noted the sensitivity of onto-ness to codomain.)

But what, formally, are  $f_1$  and  $f_2$ ? Well, they're sets of ordered pairs, like any function. Look at this set:

$$S = \{ (x, y) \mid x \in \mathbb{Z} \text{ and } y = x^2 \}$$

Here are some elements of  $S$ :

$$(3,9) \quad (0,0) \quad (-5,25) \quad (5,25) \quad (-2,4) \quad (9,81)$$

... and so forth. Which of these functions is the set  $S$ ?

**Answer:** Both are! But how can it be that  $f_1 = S = f_2$ , that is,  $f_1$  and  $f_2$  are the same identical object (a set), yet  $f_1$  *isn't* surjective and  $f_2$  *is* surjective?

**Answer:** There's no good answer. Grimaldi's formal definition of "function" *doesn't encode the codomain*.

Grimaldi would say  $f_1 \neq f_2$ , since they have different codomains, but is silent on what that means formally.

**Bottom line:** Don't be confused. Understanding is more important than formalism.

# Omission & Warning

Let  $S$  be any set. The *identity function on  $S$*  is the function  $I : S \rightarrow S$  such that  $I(x) = x$  for all  $x \in S$ . (I takes its input and spits it out unchanged as output.)

The identity function is a bijection on  $S$ ; we might also think of it as a first projection on  $S$ . Important example.

Let  $S$  be a set and let  $*$  be a binary operation on  $S$  with an identity  $e$ . Suppose that for some  $x \in S$  there is an element  $y$  in  $S$  such that  $x * y = y * x = e$ . Then we call  $y$  the *inverse of  $x$  under  $*$* , and we write  $y = x^{-1}$ .

Example: The inverse of 8 under addition is  $-8$ , since  $8 + -8 = -8 + 8 = 0$ . The inverse of 8 under multiplication is  $1/8$ . 0 has no multiplicative inverse.

To have inverses, a binary operation must have an identity (so *min* has no inverses). But *some binary operations have identity but no inverses*, e.g.  $\cap$  and  $\cup$ .

The study of inverse elements is hugely important. But we're not studying them here! We're studying inverses of functions. These *are* inverses of composition considered as a binary operation, hence not unrelated, but inverses in the abstract are not on the program.

# Preimages

Let  $f$  be a function from  $X$  to  $Y$ . Recall the following:

- If  $f(x) = y$ , then  $y$  is the *image* of  $x$
- If  $f(x) = y$ , then  $x$  is a *preimage* of  $y$
- If  $A$  is a subset of the domain of  $f$ , then  $f(A)$  is the set consisting of all  $y$  such that  $y = f(x)$  for some  $x$  in  $A$ , and we call  $f(A)$  the *image of  $A$* .

We complete this duality with the following definition: Suppose that  $B$  is a subset of the codomain of  $f$ . Then the *preimage of  $B$*  is the set of all  $x$  such that  $f(x) = y$  for some  $y \in B$ . We write  $f^{-1}(B)$  for the preimage of  $B$ . (If  $B$  consists of a single point  $B = \{y\}$ , we write  $f^{-1}(y)$ .)  
[blackboard picture of preimage]

**Example:** Suppose  $g : \{\text{cities}\} \rightarrow \{\text{states}\}$  is the “state-located-in” function. Then  $g^{-1}(\{\text{MA, NE}\}) = \{\text{Natick, Newton, Omaha, Grand Island, } \dots\}$ . Also,  $g^{-1}(\text{IA}) = \{\text{Council Bluffs, Des Moines, } \dots\}$

**Example:** Suppose  $f$  is the floor function  $\lfloor x \rfloor$ . Then

$$\begin{aligned} f^{-1}(\{0, 1, 2\}) &= [0, 3) \\ f^{-1}(\{5, 7\}) &= [5, 6) \cup [7, 8) \end{aligned}$$

# Examples & Theorems

**Example:** Let  $h : \{\text{states}\} \rightarrow \{\text{cities}\}$  be the “capital-of” function. Then

$$\begin{aligned}h^{-1}(\{\text{Montpelier, Helena}\}) &= \{\text{Montana, Vermont}\} \\h^{-1}(\text{Omaha}) &= \emptyset\end{aligned}$$

**Example:** Let  $i : \underline{\mathbf{R}} \rightarrow \underline{\mathbf{R}}$  be the function  $i(x) = x^2 + 10$ . Then  $i^{-1}(\underline{\mathbf{R}}) = i^{-1}(\underline{\mathbf{R}}^+) = \underline{\mathbf{R}}$  and  $i^{-1}([0,10]) = \{0\}$ .

As with images, the preimage of a set is *always* a set, even if some of these sets have only one element.

E.g., if  $f$  is our familiar  $+1$  function, then  $f^{-1}(7)$  really means  $f^{-1}(\{7\})$  and equals  $\{6\}$ , not 6.

Here are a few simple results on preimages, similar to but simpler than the corresponding theorem on images. Let  $f : X \rightarrow Y$  and let  $B_1$  and  $B_2$  be subsets of  $Y$ . Then

$$\begin{aligned}f^{-1}(B_1 \cap B_2) &= f^{-1}(B_1) \cap f^{-1}(B_2) \\f^{-1}(B_1 \cup B_2) &= f^{-1}(B_1) \cup f^{-1}(B_2) \\f^{-1}(-B_1) &= -f^{-1}(B_1)\end{aligned}$$

Proofs in the text, where G also gives a zillion examples of preimages, mostly with numerical functions. (But he does it in a way sure to confuse you, as we’ll soon see.)

# Functional Composition

We think of a function as a box that takes an input and produces an output. To *compose* two functions means to connect the output of the first to the input of the second!  
[Blackboard picture]

Consider again the squaring function  $f_1$  (which takes any  $x$  to  $x^2$ ) and the +1 function  $f_2$  (which takes  $x$  to  $x+1$ ). Suppose I take 5, shove it through  $f_1$ , then take the output and put it through  $f_2$ . The result is 26. If I wrap these two functions together and consider it as a single function, I get a new function that takes any input  $x$  and produces output  $x^2+1$ .

In symbols,  $f_1(x) = x^2$  and  $f_2(x) = x+1$ . So the new composite function is  $f_2(f_1(x)) = x^2 + 1$ . We call this new function “ $f_2$  after  $f_1$ ” and write it  $f_2 \circ f_1$ . Given two functions  $f$  and  $g$  the new function  $f \circ g$  is defined as follows:  $(f \circ g)(x) = f(g(x))$

[The notation can be a little confusing: seeing  $f \circ g$  you might think that  $f$  operates first. Read the symbol  $\circ$  as “*after*” and you won’t get confused. The other thing that you MUST keep in mind is that  $f \circ g$  is a new function that stands on its own, just like 8 in  $5+3 = 8$ .]

# Appropriateness

We said: Given two functions  $f$  and  $g$  we can make a new function  $f \circ g$ . But this isn't true for *any* two functions; the output of  $g$  must be connectible to the input of  $f$ !

For the function  $f \circ g$  to be defined, the codomain of  $g$  must be the domain of  $f$ . (Actually, it suffices that the range of  $g$  be a subset of the domain of  $f$ .)

**Formally:** Suppose  $f : S \rightarrow T$  and  $g : T \rightarrow R$  are functions. Then  $g \circ f : S \rightarrow R$  is **the function that takes any  $x \in S$  to  $g(f(x))$** . (But  $f \circ g$  is meaningless;  $(f \circ g)(x)$  would be  $f(g(x))$ , so  $x$  must be in  $T$ , but then  $g(x) \in R$  and we can't take  $f(\text{something in } R)$ !)

So, e.g., if  $f : \{\text{cities}\} \rightarrow \{\text{states}\}$  is “located-in”, and  $g : \{\text{states}\} \rightarrow \mathbb{Z}$  is “population of”, then the function  $g \circ f : \{\text{cities}\} \rightarrow \mathbb{Z}$  takes any city  $c$  to the population of the state of  $c$ , e.g.  $(g \circ f)(\text{Omaha}) \approx 1.7\text{M}$ . But  $f \circ g$  is meaningless: what would  $(f \circ g)(\text{NE})$  be?

Of course, if  $f$  and  $g$  are both functions from  $S$  to  $S$  then both  $f \circ g$  and  $g \circ f$  are well-defined. Are they the same? If  $f$  is squaring, and  $g$  is increment, is  $f(g(x))$  the same as  $g(f(x))$ ?



# More on Composition

**Example:** Suppose  $f$  is “state-located-in” as above, and  $g : \{\text{states}\} \rightarrow \{\text{cities}\}$  is the “capital city of” function. Then  $g \circ f : \{\text{cities}\} \rightarrow \{\text{cities}\}$  maps any city to the city that’s the capital of its state, e.g.,  
 $(g \circ f)(\text{Medford}) = \text{Boston}$      $(g \circ f)(\text{Augusta}) = \text{Augusta}$

**Example:** Suppose  $g : \mathbf{R} \rightarrow \mathbf{R}$  is  $g(x) = x^2 + 0.5$ . Then  $\text{floor} \circ g$  is the function  $\lfloor x^2 + 0.5 \rfloor$  but  $g \circ \text{floor}$  is the function  $\lfloor x \rfloor^2 + 0.5$ . Are these the same?

**Example:** Suppose  $g : \mathbf{R} \rightarrow \mathbf{R}$  is  $g(x) = x^2 + 1$ . Then  $g \circ g : \mathbf{R} \rightarrow \mathbf{R}$  is the function

$$(g \circ g)(x) = g(g(x)) = x^4 + 2x^2 + 2$$

We call this function  $g^2$ ; by definition,  $g^2 = g \circ g$ . (Note that this definition makes sense only when the domain and codomain of  $g$  are the same.) We have  $g^2(x) = g(g(x))$ . And  $g^3 = g^2 \circ g$ , so  $g^3(x) = g(g(g(x)))$ , and so forth recursively: for any  $n > 1$ ,  $g^n$  is defined as  $g^{n-1} \circ g$ . These functions are the *powers of  $g$* .

**Quirky point:** The base case of the above recursive definition is implicitly  $g^1 = g$ . But what should  $g^0$  be?

# Associativity

Theorem: Let  $h : A \rightarrow B$ ,  $g : B \rightarrow C$ , and  $f : C \rightarrow D$  be functions. Then  $(f \circ g) \circ h = f \circ (g \circ h)$ .

[explanation and intuitive proof by blackboard diagram]

**Proof:** Let  $x$  be an element of  $A$ . We need to show that

$$((f \circ g) \circ h)(x) = (f \circ (g \circ h))(x)$$

The left-hand side is, by definition,  $(f \circ g)(h(x))$ , which in turn is  $f(g(h(x)))$ . The right-hand side is  $f((g \circ h)(x))$  which is also  $f(g(h(x)))$ . So the LHS and RHS are equal.

Keep in mind that  $\circ$  is a closed binary operation, like addition or multiplication, since it takes two things and spits out another thing of the same type (which is what a closed binary operation is supposed to do). We've just proven that  $\circ$  is an *associative* binary operation. And we saw earlier that  $\circ$  is *not* a commutative operation:

In general, it's not true that  $f \circ g \neq g \circ f$ .

Question: Is  $\circ$  idempotent?

Problem: Are there any functions such that  $f \circ g = g \circ f$ ?

# Simple Results

If  $f$  and  $g$  are injective, then  $g \circ f$  is injective.

**Informal proof:** If  $x$  and  $y$  are distinct, then  $f(x)$  and  $f(y)$  are distinct since  $f$  is injective. From this and the injectivity of  $g$  it follows that  $g(f(x))$  and  $g(f(y))$  are distinct. That is, if  $x$  and  $y$  are distinct then  $g(f(x))$  and  $g(f(y))$  are distinct, which is to say that  $g \circ f$  is injective.

If  $f$  and  $g$  are onto, then  $g \circ f$  is onto.

**Even more informal proof:** If  $f$  is onto then its domain maps to its entire codomain; same for  $g$ . Therefore  $g \circ f$  maps the domain of  $f$  first to the whole domain of  $g$  and then to the whole codomain of  $g$ , i.e.,  $g \circ f$  is surjective.

If  $f$  and  $g$  are bijections, then  $g \circ f$  is a bijection.

**Completely formal proof:** Trivial, given the two results above.

[To understand this completely you might find examples of functions such that  $f$  is injective but neither  $g$  nor  $f \circ g$  is injective. And the same thing in three other cases.]

# Invertibility

Once again, let  $f : \underline{\mathbb{Z}} \rightarrow \underline{\mathbb{Z}}$  be the +1 function, and now consider the function  $g : \underline{\mathbb{Z}} \rightarrow \underline{\mathbb{Z}}$  defined by  $g(x) = x - 1$ . Notice that  $f(g(x)) = f(x-1) = (x-1)+1 = x$  for any  $x$ , that is,  $f \circ g$  is the identity function. Similarly,  $g \circ f$  is the identity function. In such a situation we say that  $g$  is the *inverse* of  $f$ ; as a box,  $g$  undoes whatever  $f$  does, and  $f$  undoes whatever  $g$  does.

In the example above the domain and codomain are the same. But we can be more general: Suppose  $f : X \rightarrow Y$  is a function, and suppose  $g : Y \rightarrow X$  is a function such that  $f(g(y)) = y$  for every  $y \in Y$ , and  $g(f(x)) = x$  for every  $x \in X$ . Said another way,  $f \circ g$  is the identity function on  $Y$ , and  $g \circ f$  is the identity function on  $X$ . Then we say that  $f$  is *invertible* and  $g$  is the *inverse* of  $f$ , and we write  $g = f^{-1}$ . (Since the definition is symmetric, we can also say  $g$  is invertible,  $f$  is the inverse of  $g$ , and  $f = g^{-1}$ .)

(Technical point: I said “suppose  $g$  is a function such that...” but I didn’t prove that there can be only one such function; maybe there there are several such functions  $g$ ! But in fact if there is a  $g$  it must be unique; proof in G.)

# Examples / Formalism

**Example:** Suppose  $f : \underline{\mathbf{R}} \rightarrow \underline{\mathbf{R}}$  is  $f(x) = 3x + 7$ .  
Then  $f^{-1} : \underline{\mathbf{R}} \rightarrow \underline{\mathbf{R}}$  is the function  $f^{-1}(x) = (x-7)/3$ ;  
note that  $f(f^{-1}(x)) = f^{-1}(f(x)) = x$ . [Both must hold!]

**Example:** Suppose  $f : \underline{\mathbf{Z}} \rightarrow \underline{\mathbf{E}}$  is the function  $f(x) = 2x$ .  
Then  $f^{-1} : \underline{\mathbf{E}} \rightarrow \underline{\mathbf{Z}}$  is  $f^{-1}(x) = x/2$ . But  $f_1 : \underline{\mathbf{Z}} \rightarrow \underline{\mathbf{Z}}$   
with  $f_1(x) = 2x$  is *not* invertible; there is no  $g : \underline{\mathbf{Z}} \rightarrow \underline{\mathbf{Z}}$   
such that  $g(f_1(x)) = x$  and  $f_1(g(x)) = x$ . (Try  $x = 3$ !)

Formally, recall that a function is a **set of ordered pairs**.  
Then we form the *converse* of that set of ordered pairs by  
reversing each pair, i.e., we replace each  $(x,y)$  with  $(y,x)$ .  
Of course the result may not be a function (why?). But if  
it is, then the original function is invertible, and the  
converse is the inverse function. Totally unilluminating.

**Caution:** Do not confuse “inverse” with “preimage”,  
even though they use the same notation! The preimage  
of a set is defined for any function, but not all functions  
are invertible. We use the same notation for the  
following reason: **If  $f$  is invertible, then for any set  $A$  the  
preimage of  $A$  under  $f$  is equal to the image of  $A$  under  
the inverse of  $f$ . So both are written  $f^{-1}(A)$ .**

# Invertible = Bijective

Suppose  $f : X \rightarrow Y$  is invertible and that  $g : Y \rightarrow X$  is its inverse. What can we say about  $f$  and  $g$ ?

**First:**  $f$  must be one-to-one; it's not possible that there are distinct  $x_1$  and  $x_2$  in  $X$  such that  $f(x_1) = f(x_2)$ .

**Intuitively:** If there were  $x_1$  and  $x_2$  that collapsed into the same  $y$ , how could  $g$ , given  $y$ , produce both of them?

**Formal proof:** Suppose we are given  $x_1$  and  $x_2$  in  $X$  as above. So  $g(f(x_1)) = g(f(x_2))$ . But  $g(f(x_1)) = x_1$  and  $g(f(x_2)) = x_2$  by definition of inverse. So  $x_1 = x_2$ . QED

**Second:**  $f$  must be onto; for each  $y \in Y$  there must be some  $x \in X$  such that  $f(x) = y$ .

**Proof:** Since  $g$  is a function, for each  $y \in Y$  we have  $g(y) \in X$ . But then  $f(g(y)) = y$  by definition of inverse, and we've found the  $x$  we need.

We've proved that every invertible function is bijective.

It's also true that every bijective function is invertible!

**Sketch:** Let  $f : X \rightarrow Y$  be bijective. For each  $y \in Y$  there is an  $x \in X$  such that  $f(x) = y$  [since  $f$  is onto] and there is in fact exactly one such  $x$  [since  $f$  is 1-1]. So define a function  $g$  as follows: for each  $y \in Y$ , let  $g(y)$  be that unique  $x$ . It's easy to prove that this  $g$  is the inverse of  $f$ .

# An Important Theorem

Suppose  $f : X \rightarrow Y$  and  $g : Y \rightarrow Z$  are both invertible functions. Then the function  $g \circ f : X \rightarrow Z$  is also invertible, and in fact  $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ .

What does this theorem say? It says that if each step you go forward can be reversed, then you can also reverse the effect of going two steps forward. *But you must take the reverse steps in backwards order!* [blackboard picture]

I strongly recommend that you prove this theorem yourself; it's not very hard, and it's a great exercise. We'll do one example: Let  $f : \mathbb{Z} \rightarrow \mathbb{Z}$  be our old friend the +1 function, and let  $g : \mathbb{Z} \rightarrow \mathbb{E}$  be the function that doubles its argument:  $g(x) = 2x$ , where  $\mathbb{E}$  is the set of even integers. (Why did I put  $\mathbb{E}$  here instead of  $\mathbb{Z}$ ?) Then  $g \circ f : \mathbb{Z} \rightarrow \mathbb{E}$  is the function  $(g \circ f)(x) = 2x + 2$  since “g after f” means “add one then double”.

How do we invert this? Do we subtract one and halve?

No indeed—we must halve, *then* subtract one.  $f^{-1}$  is “subtract one” and  $g^{-1}$  is “halve”, so it must be that  $(g \circ f)^{-1}$ , the inverse of “g after f”, is “ $f^{-1}$  after  $g^{-1}$ ”.

# A Counting Theorem

Suppose  $f : X \rightarrow Y$  for *finite* sets  $X$  and  $Y$  such that  $|X| = |Y|$ . Then the following statements are equivalent:

- (a)  $f$  is invertible
- (b)  $f$  is bijective
- (c)  $f$  is injective
- (d)  $f$  is surjective

(Remember what this means: These statements are either all true or all false. Said another way, if  $f$  possesses any of these properties, it possesses them all.)

Recall how we prove a theorem of equivalent statements: we prove that (a)  $\iff$  (b), that (b)  $\iff$  (c), that (c)  $\iff$  (d), and finally that (d)  $\iff$  (a). Here things are simpler since we've already proved (a) and (b) equivalent, and (b) implies both (c) and (d) by definition. So if we prove that (c)  $\iff$  (b) and (d)  $\iff$  (b) we're done. Details in G, using the Pigeonhole Principle.

**But here's the intuition:** When two sets are of the same size, anything injective must be surjective since there can't be any extra elements in the codomain. And any surjection must be one-to-one since otherwise there aren't enough elements in the domain to cover the codomain. [proof by blackboard picture]