LECTURE NOTES ON RELATIONS AND FUNCTIONS

PETE L. CLARK

Contents

1. Relations

1.1. **The idea of a relation.** Let X and Y be two sets. We would like to formalize the idea of a **relation** between X and Y. Intuitively speaking, this is a well-defined "property" R such that given any $x \in X$ and $y \in Y$, either x bears the property R to y, or it doesn't (and not both!). Some important examples:

Example 1.1. *Let X be a set of objects and let Y be a set of sets. Then "membership" is a relation R from X to Y: i.e., we have xRy if $x \in y$.*

Example 1.2. *Let S be a set, and let $X = Y = 2^S$, the power set of S (recall that this is the set of all subsets of S. Then containment, $A \subseteq B$ is a relation between X and Y. (Proper containment, $A \subsetneq B$, is also a relation.)*

Example 1.3. *Let $X = Y$. Then equality is a relation from X to Y: we say xRy iff $x = y$. Also inequality is a relation between X and Y: we say xRy iff $x \neq y$.*

Example 1.4. *Let $X = Y = \mathbb{R}$. Then $\leq, <, \geq, >$ are relations between $\mathbb{R}$ and $\mathbb{R}$.*

Example 1.5. *Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Then we can define a relation from $\mathbb{R}$ to $\mathbb{R}$, by xRy if and only if $y = f(x)$.*

Date: April 15, 2016.

Example 1.6. *Let $X = Y = \mathbb{Z}$. Then divisibility is a relation between $\mathbb{Z}$ and $\mathbb{Z}$: we say xRy if $x \mid y$.*

Example 1.7. *Let $X = Y = \mathbb{Z}$. Then "having the same parity" is a relation between $\mathbb{Z}$ and $\mathbb{Z}$.*

In many of the above examples we have $X = Y$. This will often (but certainly not always!) be the case, and when it is we may speak of relations **on X**.

1.2. The formal definition of a relation.

We still have not given a formal definition of a relation between sets X and Y. In fact the above way of thinking about relations is easily formalized, as was suggested in class by Adam Osborne: namely, we can think of a relation R as a function from $X \times Y$ to the two-element set $\{\mathrm{TRUE}, \mathrm{FALSE}\}$. In other words, for $(x, y) \in X \times Y$, we say that xRy if and only if $f((x, y)) = \mathrm{TRUE}$.

This is a great way of thinking about relations. It has however one foundational drawback: it makes the definition of a relation depend on that of a function, whereas the standard practice for about one hundred years is the reverse: we want to define a function as a special kind of relation (c.f. Example 5 above). The familiar correspondence between logic and set theory leads us to the official definition:

Definition: A relation R between two sets X and Y is simply a **subset** of the Cartesian product $X \times Y$, i.e., a collection of ordered pairs (x, y).

(Thus we have replaced the basic logical dichotomy "TRUE/FALSE" with the basic set-theoretic dichotomy "is a member of/ is not a member of".) Note that this new definition has some geometric appeal: we are essentially identifying a relation R with its *graph* in the sense of precalculus mathematics.

We take advantage of the definition to adjust the terminology: rather than speaking (slightly awkwardly) of relations "from X to Y" we will now speak of relations **on X × Y**. When $X = Y$ we may (but need not!) speak of relations **on X**.

Example 1.8. *Any curve in $\mathbb{R}^2$ defines a relation on $\mathbb{R} \times \mathbb{R}$. E.g. the unit circle*

$$x^2 + y^2 = 1$$

is a relation in the plane: it is just a set of ordered pairs.

1.3. Basic terminology and further examples.

Let X,Y be sets. We consider the set of all relations on $X \times Y$ and denote it by $\mathcal{R}(X, Y)$. According to our formal definition we have

$$\mathcal{R}(X, Y) = 2^{X \times Y},$$

i.e., the set of all subsets of the Cartesian product $X \times Y$.

Example 1.9. *a) Suppose $X = \varnothing$. Then $X \times Y = \varnothing$ and $\mathcal{R}(X \times Y) = 2^{\varnothing} = \{\varnothing\}$. That is: if X is empty, then the set of ordered pairs (x, y) for $x \in X$ and $y \in Y$ is empty, so there is only one relation: the empty relation.*
b) Suppose $Y = \varnothing$. Again $X \times Y = \varnothing$ and the discussion is the same as above.

Example 1.10. *a) Suppose $X = \{\bullet\}$ consists of a single element. Then $X \times Y = \{(\bullet, y) \mid y \in Y\}$; in other words, $X \times Y$ is essentially just Y itself, since the first coordinate is always the same. Thus a relation R on $X \times Y$ corresponds to a subset of Y: formally, the set of all $y \in Y$ such that $\bullet R y$.*
b) Suppose $Y = \{\bullet\}$ consists of a single element. The discussion is analogus to that of part a), and relations on $X \times Y$ correspond to subsets of X.

Example 1.11. *Suppose X and Y are finite sets, with $\#X = m$ and $\#Y = n$. Then $\mathcal{R}(X, Y) = 2^{X \times Y}$ is finite, of cardinality*

$$\# 2^{X \times Y} = 2^{\# X \times Y} = 2^{\# X \cdot \# Y} = 2^{mn}.$$

The function 2^{mn} grows rapidly with both m and n, and the upshot is that if X and Y are even moderately large finite sets, the set of all relations on $X \times Y$ is very large. For instance if $X = \{a, b\}$ and $Y = \{1, 2\}$ then there are $2^{2 \cdot 2} = 16$ relations on $X \times Y$. It is probably a good exercise for you to write them all down. However, if $X = \{a, b, c\}$ and $Y = \{1, 2, 3\}$ then there are $2^{3 \cdot 3} = 512$ relations on $X \times Y$, and – with apologies to the Jackson 5? – it is less easy to write them all down.

Exercise 1.1. *Let X and Y be nonempty sets, at least one of which is infinite. Show: $\mathcal{R}(X, Y)$ is infnite.*

Given two relations R_1 and R_2 between X and Y, it makes sense to say that $R_1 \subseteq R_2$: this means that R_1 is "stricter" than R_2 or that R_2 is "more permissive" than R_1. This is a very natural idea: for instance, if X is the set of people in the world, R_1 is the brotherhood relation – i.e., $(x, y) \in R_1$ iff x and y are brothers – and R_2 is the sibling relation – i.e., $(x, y) \in R_2$ iff x and y are siblings – then $R_1 \subsetneq R_2$: if x and y are brothers then they are also siblings, but not conversely.

Among all elements of $\mathcal{R}(X, Y)$, there is one relation $R_\varnothing$ which is the strictest of all, namely $R_\varnothing = \emptyset$:[1] that is, for *no* $(x, y) \in X \times Y$ do we have $(x, y) \in R_\varnothing$. Indeed $R_\varnothing \subset R$ for any $R \in \mathcal{R}(X, Y)$. At the other extreme, there is a relation which is the most permissive, namely $R_{X \times Y} = X \times Y$ itself: that is, for *all* $(x, y) \in X \times Y$ we have $(x, y) \in R_{X \times Y}$. And indeed $R \subset R_{X \times Y}$ for any $R \in \mathcal{R}(X, Y)$.

Example 1.12. *Let $X = Y$. The equality relation $R = \{(x, x) \mid x \in X\}$ can be thought of geometrically as the diagonal of $X \times Y$.*

The **domain**[2] of a relation $R \subseteq X \times Y$ is the set of $x \in X$ such that there exists $y \in Y$ with $(x, y) \in R$. In other words, it is the set of all elements in x which relate to at least one element of Y.

Example 1.13. *The circle relation $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ has domain $[-1, 1]$.*

Given a relation $R \subset X \times Y$, we can define the **inverse relation** $R^{-1} \subset Y \times X$ by interchanging the order of the coordinates. Formally, we put

$$R^{-1} = \{(y, x) \in Y \times X \mid (x, y) \in R\}.$$

Geometrically, this corresponds to reflecting across the line $y = x$.

[1] The notation here is just to emphasize that we are viewing $\varnothing$ as a relation on $X \times Y$.

[2] I don't like this terminology. But it is used in the course text, and it would be confusing to change it.

Example 1.14. *Consider the relation $R \subset \mathbb{R} \times \mathbb{R}$ attached to the function $f(x) = x^2$:*

$$R = \{(x, x^2) \mid x \in \mathbb{R}\}.$$

The graph of this relation is an upward-opening parabola: it can also be described by the equation $y = x^2$. The inverse relation R^{-1} is $\{(x^2, x) \mid x \in \mathbb{R}\}$, which corresponds to the equation $x = y^2$ and geometrically is a parabola opening rightward. Note that the domain of the original relation R is $\mathbb{R}$, whereas the domain of R^{-1} is $[0, \infty)$. Moreover, R^{-1} is not a function, since some values of x relate to more than one y-value: e.g. $(1, 1)$ and $(1, -1)$ are both in R^{-1}.

Example 1.15. *Consider the relation attached to the function $f(x) = x^3$: namely*

$$R = \{(x, x^3) \mid x \in \mathbb{R}\}.$$

This relation is described by the equation $y = x^3$; certainly it is a function, and its domain is $\mathbb{R}$. Consider the inverse relation

$$R^{-1} = \{(x^3, x) \mid x \in \mathbb{R}\},$$

which is described by the equation $x = y^3$. Since every real number has a unique real cube root, this is equivalent to $y = x^{\frac{1}{3}}$. Thus this time R^{-1} is again a function, and its domain is $\mathbb{R}$.

Later we will study functions in detail and one of our main goals will be to understand the difference between Examples 1.14 and 1.15.

1.4. **Properties of relations.**

Let X be a set. We now consider various properties that a relation R on X – i.e., $R \subset X \times X$ may or may not possess.

Reflexivity: For all $x \in X$, $(x, x) \in R$.

In other words, each element of X bears relation R to itself. Another way to say this is that the relation R contains the equality relation on X.

Exercise 1.2. *Which of the relations in Examples 1.1 through 1.15 are reflexive?*

Anti-reflexivity: For all $x \in X$, $(x, x) \not inR$.

Certainly no relation on X is both reflexive and anti-reflexive (except in the silly case $X = \varnothing$ when both properties hold vacuously) . However, notice that a relation need not be either reflexive or anti-reflexive: if there are $x, y \in X$ such that $(x, x) \in R$ and $(y, y) \notin R$, then neither property holds.

Symmetry: For all $x, y \in X$, if $(x, y) \in R$, then $(y, x) \in R$.

Again, this has a geometric interpretation in terms of symmetry across the diagonal $y = x$. For instance, the relation associated to the function $y = \frac{1}{x}$ is symmetric since interchanging x and y changes nothing, whereas the relation associated to the function $y = x^2$ is not. (Looking ahead a bit, a function $y = f(x)$ is symmetric iff it coincides with its own inverse function.)

Exercise 1.3. *Which of the relations in Examples 1.1 through 1.15 are symmetric?*

Example 1.16. *Let V be a set. A **(simple, loopless, undirected) graph** – in the sense of graph theory, not graphs of functions! – is given by a relation E on V which is irreflexive and symmetric. Thus: for $x, y \in V$, we say that x and y are **adjacent** if $(x, y) \in E$. Moreover x is never adjacent to itself, and the adjacency of x and y is a property of the unordered pair $\{x, y\}$: if x is adjacent to y then y is adjacent to x.*

Anti-Symmetry: for all $x, y \in X$, if $(x, y) \in R$ and $(y, x) \in R$, then $x = y$.

Exercise 1.4. *Which of the relations in Examples 1.1 through 1.16 are anti-symmetric?*

Transitivity: for all $x, y, z \in X$, if $(x, y) \in R$ and $(y, z) \in R$, then $(x, z) \in R$.

"Being a parent of" is not transitive, but "being an ancestor of" is transitive.

Exercise 1.5. *Which of the relations in Examples 1.1 through 1.15 are transitive?*

Worked Exercise 1.6.
Let R be a relation on X. Show the following are equivalent:
(i) R is both symmetric and anti-symmetric.
(ii) R is a subrelation of the equality relation.
***Solution**: Suppose that we have a relation R on X which is both symmetric and anti-symmetric. Then, for all $x, y \in R$, if $(x, y) \in R$, then by symmetry we have also $(y, x) \in R$, and then by anti-symmetry we have $x = y$. Thus we've shown that if (i) holds, the only possible elements $(x, y) \in R$ are those of the form (x, x), which means that R is a subrelation of the equality relation. Conversely, if R is a subrelation of equality and $(x, y) \in R$, then $y = x$, so $(y, x) \in R$. Similarly, if $(x, y) \in R$ and $(y, x) \in R$ then $x = y$. So R is both symmetric and anti-symmetric.*

Now we makes two further defintions of relations with possess certain combinations of these basic properties. The first is the most important definition in this section.

An **equivalence relation** on a set X is a relation on X which is reflexive, symmetric and transitive.

A **partial ordering** on a set X is a relation on X which is reflexive, anti-symmetric and transitive.

Exercise 1.7. *Which of the relations in Examples 1.1 through 1.16 are equivalence relations? Which are partial orderings?*

We often denote equivalence relations by a tilde – $x \sim y$ – and read $x \sim y$ as "x is equivalent to y". For instance, the relation "having the same parity" on $\mathbb{Z}$ is an equivalence relation, and $x \sim y$ means that x and y are both even or both odd. Thus it serves to group the elements of $\mathbb{Z}$ into subsets which share some common property. In this case, all the even numbers are being grouped together and all the odd numbers are being grouped together. We will see shortly that this is a characteristic property of equivalence relations: every equivalence relation on a set X determines a **partition** on X and conversely, given any partition on X we can define an equivalence relation.

The concept of a partial ordering should be regarded as a "generalized less than or equal to" relation. Perhaps the best example is the containment relation $\subseteq$ on the power set $\mathcal{P}(S)$ of a set S. This is a very natural way of regarding one set as "bigger" or "smaller" than another set. Thus the insight here is that containment satisfies many of the formal properties of the more familiar $\leq$ on numbers. However there is one property of $\leq$ on numbers that does not generalize to $\subseteq$ (and hence not to an arbitrary partial ordering): namely, given any two real numbers x, y we must have either $x \leq y$ or $y \leq x$. However for sets this does not need to be the case (unless S has at most one element). For instance, in the power set of the positive integers, we have $A = \{1\}$ and $B = \{2\}$, so neither is it true that $A \subseteq B$ or that $B \subseteq A$. This is a much stronger property of a relation:

Totality: For all $x, y \in X$, either $(x, y) \in R$ or $(y, x) \in R$.

A **total ordering** (or **linear ordering**) on a set X is a partial ordering satisfying dichotomy.

Example 1.17. *The relation $\leq$ on $\mathbb{R}$ is a total ordering.*

There is an entire branch of mathematics – **order theory** – devoted to the study of partial orderings.[3] In my opinion order theory gets short shrift in the standard mathematics curriculum (especially at the advanced undergraduate and graduate levels): most students learn only a few isolated results which they apply frequently but with little context or insight. Unfortunately we are not in a position to combat this trend: partial and total orderings will get short shrift here as well!

1.5. Partitions and Equivalence Relations.

Let X be a set, and let $\sim$ be an equivalence relation on X.

For $x \in X$, we define the **equivalence class of x** as

$$[x] = \{y \in X \mid y \sim x\}.$$

For example, if $\sim$ is the relation "having the same parity" on $\mathbb{Z}$, then

$$[2] = \{\ldots, -4, -2, 0, 2, 4, \ldots\},$$

i.e., the set of all even integers. Similarly

$$[1] = \{\ldots -3, -1, 1, 3, \ldots\}$$

is the set of all odd integers. But an equivalence class in general has many "representatives". For instance, the equivalence class $[4]$ is the set of all integers having the same parity as 4, so is again the set of all even integers: $[4] = [2]$. More generally, for any even integer n, we have $[n] = [0]$ and for any odd integer n we have $[n] = [1]$. Thus in this case we have partitioned the integers into two subsets: the even integers and the odd integers.

We claim that given any equivalence relation $\sim$ on a set X, the set $\{[x] \mid x \in X\}$ forms a partition of X. Before we proceed to demonstrate this, observe that we are now strongly using our convention that there is no "multiplicity" associated to membership in a set: e.g. the sets $\{4, 2+2, 1^1 + 3^0 + 2^1\}$ and $\{4\}$ are equal. The

[3] For instance, there is a journal called **Order**, in which a paper of mine appears.

above representation $\{[x] \mid x \in X\}$ is highly redundant: for instance in the above example we are writing down the set of even integers and the set of odd integers infinitely many times, but it only "counts once" in order to build the set of subsets which gives the partition.

With this disposed of, the verification that $\mathcal{P} = \{[x] \mid x \in X\}$ gives a partition of X comes down to recalling the definition of a partition and then following our noses. There are three properties to verify:

(i) That every element of $\mathcal{P}$ is nonempty. Indeed, the element $[x]$ is nonempty because it contains x! This is by reflexivity: $x \sim x$, so $x \in \{y \in X \mid y \sim x\}$.
(ii) That the union of all the elements of $\mathcal{P}$ is all of X. But again, the union is indexed by the elements x of X, and we just saw that $x \in [x]$, so every x in X is indeed in at least one element of $\mathcal{P}$.
(iii) Finally, we must show that if $[x] \cap [y] \neq \varnothing$, then $[x] = [y]$: i.e., any two elements of $\mathcal{P}$ which have a common element must be the same element. So suppose that there exists $z \in [x] \cap [y]$. Writing this out, we have $z \sim x$ and $z \sim y$. By symmetry, we have $y \sim z$; from this and $z \sim x$, we deduce by transitivity that $y \sim x$, i.e., $y \in [x]$. We claim that it follows from this that $[y] \subset [x]$. To see this, take any $w \in [y]$, so that $w \sim y$. Since $w \sim x$, we conclude $w \sim x$, so $w \in [x]$. Rerunning the above argument with the roles of x and y interchanged we get also that $[y] \subset [x]$, so $[x] = [y]$. This completes the verification.

Note that the key fact underlying the proof was that any two equivalence classes $[x]$ and $[y]$ are either disjoint or coincident. Note also that we did indeed use all three properties of an equivalence relation.

Now we wish to go in the other direction. Suppose X is a set and $\mathcal{P} = \{U_i\}_{i \in I}$ is a partition of X(here I is just an index set). We can define an equivalence relation $\sim$ on X as follows: we say that $x \sim y$ if there exists $i \in I$ such that $x, y \in U_i$. In other words, we are decreeing x and y to be equivalent exactly when they lie in the same "piece" of the partition. Let us verify that this is an equivalence relation. First, let $x \in X$. Then, since $\mathcal{P}$ is a partition, there exists some $i \in I$ such that $x \in U_i$, and then x and x are both in U_i, so $x \sim x$. Next, suppose that $x \sim y$: this means that there exists $i \in I$ such that x and y are both in U_i; but then sure enough y and x are both in U_i ("and" is commutative!), so $y \sim x$. Similarly, if we have x, y, z such that $x \sim y$ and $y \sim z$, then there exists i such that x and y are both in U_i and a possibly different index j such that y and z are both in U_j. But since $y \in U_i \cap U_j$, we must have $U_i = U_j$ so that x and z are both in $U_i = U_j$ and $x \sim z$.

Moreover, the processes of passing from an equivalence relation to a partition and from a partition to an equivalence relation are mutually inverse: if we start with an equivalence relation R, form the associated partition $\mathcal{P}(R)$, and then form the associated equivalence relation $\sim (\mathcal{P}(R))$, then we get the equivalence relation R that we started with, and similarly in the other direction.

1.6. **Examples of equivalence relations.**

Example 1.18. *(Congruence modulo n) Let $n \in \mathbb{Z}^+$. There is a natural partition of $\mathbb{Z}$ into n parts which generalizes the partition into even and odd. Namely, we put*

$$Y_1 = \{\ldots, -2n, -n, 0, n, 2n, \ldots\} = \{kn \mid k \in \mathbb{Z}\}$$

the set of all multiples of n,

$$Y_2 = \{\ldots, -2n+1, -n+1, 1, n+1, 2n+1 \ldots\} = \{kn+1 \mid k \in \mathbb{Z}\},$$

and similarly, for any $0 \leq d \leq n-1$, we put

$$Y_d = \{\ldots, -2n+d, -n+d, d, n+d, 2n+d \ldots\} = \{kn+d \mid kin\mathbb{Z}\}.$$

That is, Y_d is the set of all integers which, upon division by n, leave a remainder of d. Earlier we showed that the remainder upon division by n is a well-defined integer in the range $0 \leq d < n$. Here by "well-defined", I mean that for $0 \leq d_1 \neq d_2 < n$, the sets Y_{d_1} and Y_{d_2} are disjoint. Recall why this is true: if not, there exist k_1, k_2 such that $k_1 n + d_1 = k_2 n + d_2$, so $d_1 - d_2 = (k_2 - k_1)n$, so $d_1 - d_2$ is a multiple of n. But $-n < d_1 - d_2 < n$, so the only multiple of n it could possibly be is 0, i.e., $d_1 = d_2$. It is clear that each Y_d is nonempty and that their union is all of $\mathbb{Z}$, so $\{Y_d\}_{d=0}^{n-1}$ gives a partition of $\mathbb{Z}$.

The corresponding equivalence relation is called **congruence modulo** n, and written as follows:

$$x \equiv y \pmod{n}.$$

What this means is that x and y leave the same remainder upon division by n.

Proposition 1.19. *For integers x, y, the following are equivalent:*
(i) $x \equiv y \pmod{n}$.
(ii) $n \mid x - y$.

Proof. Suppose that $x \equiv y \pmod{n}$. Then they leave the same remainder, say d, upon division by n: there exist $k_1,\ k_2 \in \mathbb{Z}$ such that $x = k_1 n + d$, $y = k_2 n + d$, so $x - y = (k_1 - k_2)n$ and indeed $n \mid x - y$. Conversely, suppose that $x = k_1 n + d_1$, $y = k_2 n + d_2$, with d_1 and d_2 distinct integers both in the interval $[0, n-1]$. Then, if n divides $x - y = (k_1 - k_2)n + (d_1 - d_2)$, then it also divides $d_1 - d_2$, which as above is impossible since $-n < d_1 - d_2 < n$. □

Example 1.20. *(Fibers of a function) Let $f : X \to Y$ be a function. We define a relation R on X by $(x_1, x_2) \in R$ iff $f(x_1) = f(x_2)$. This is an equivalence relation. The equivalence class of $[x]$ is called the **fiber over f(x)**.*

1.7. Extra: composition of relations.

Suppose we have a relation $R \subset X \times Y$ and a relation $S \subset Y \times Z$. We can define a **composite relation** $S \circ R \subset X \times Z$ in a way which will generalize compositions of functions. Compared to composition of functions, composition of relations is much less well-known, although as with many abstract concepts, once it is pointed out to you, you begin to see it "in nature'. This section is certainly optional reading.

The definition is simply this:

$$S \circ R = \{(x, z) \in X \times Z \mid \exists y \in Y \text{such that } (x, y) \in R \text{ and } (y, z) \in S\}.$$

In other words, we say that x in the first set X relates to z in the third set Z if there exists at least one intermediate element y in the second set such that x relates to y and y relates to z.

In particular, we can always compose relations on a single set X. As a special case, given a relation R, we can compose it with itself: say

$$R^{(2)} = R \circ R = \{(x,z) \in X \times X \mid \exists y \in X \text{ such that } xRy \text{ and } yRz\}.$$

Proposition 1.21. *For a relation R on X, the following are equivalent:*
(i) R is transitive.
(ii) $R^{(2)} \subseteq R$.

Exercise 1.8. *Show that the composition of relations is associative.*

Exercise 1.9. *Show: $(S \circ R)^{-1} = R^{-1} \circ S^{-1}$.*

Exercise 1.10. *Let $X = \{1, \ldots, N\}$. To a relation R on X we associate its **adjacency matrix** $M = M(R)$: if $(i,j) \in R$, we put $M(i,j) = 1$; otherwise we put $M(i,j) = 0$. Show that the adjacency matrix of the composite relation R^2 is the product matrix $M(R) \cdot M(R)$ in the sense of linear algebra.*

2. Functions

Let X and Y be sets. A **function** $f : X \to Y$ is a special kind of relation between X and Y. Namely, it is a relation $R \subset X \times Y$ satisfying the following condition: for all $x \in X$ there exists exactly one $y \in Y$ such that $(x,y) \in R$. Because element of y attached to a given element x of X is unique, we may denote it by $f(x)$.

Geometrically, a function is a relation which passes the **vertical line test**: every vertical line $x = c$ intersects the graph of the function in exactly one point. In particular, the domain of any function is all of X.

Example 2.1. *The equality relation $\{(x,x) \mid x \in X\}$ on X is a function: $f(x) = x$ for all x. We call this the **identity function** and denote it by 1_X.*

Example 2.2. *a) Let Y be a set. Then $\varnothing \times Y = \varnothing$, so there is a unique relation on $\varnothing \times Y$. This relation is – vacuously – a function.*
b) Let X be a set. Then $X \times \varnothing = \varnothing$, so there is a unique relation on $X \times \varnothing$, with domain $\varnothing$. If $X = \varnothing$, then we get the empty function $f : \varnothing \to \varnothing$. If $X \neq \varnothing$ then the domain is not all of X so we do not get a function.

If $f : X \to Y$ is a function, the second set Y is called the **codomain** of f. Note the asymmetry in the definition of a function: although every element x of the domain X is required to be associated to a unique element y of Y, the same is not required of elements y of the codomain: there may be multiple elements x in X such that $f(x) = y$, or there may be none at all.

The **image** of $f : X \to Y$ is $\{y \in Y$ such that $y = f(x)$ for some $x \in X.\}$[4]

In calculus one discusses functions with domain some subset of $\mathbb{R}$ and codomain $\mathbb{R}$. Moreover in calculus a function is usually (but not always...) given by some relatively simple algebraic/analytic expression, and the convention is that the domain is the largest subset of $\mathbb{R}$ on which the given expression makes sense.

[4] Some people call this the **range**, but also some people call the set Y (what we called the codomain) the range, so the term is ambiguous and perhaps best avoided.

Example 2.3.
a) The function $y = 3x$ is a function from $\mathbb{R}$ to $\mathbb{R}$. Its range is all of $\mathbb{R}$.
b)The function $y = x^2$ is a function from $\mathbb{R}$ to $\mathbb{R}$. Its range is $[0, \infty)$.
c) The function $y = x^3$ is a function from $\mathbb{R}$ to $\mathbb{R}$. Its range is all of $\mathbb{R}$.
d) The function $y = \sqrt{x}$ is a function from $[0, \infty)$ to $\mathbb{R}$. Its range is $[0, \infty)$.
e) The arctangent $y = \arctan x$ is a function from $\mathbb{R}$ to $\mathbb{R}$. Its range is $(\frac{-\pi}{2}, \frac{\pi}{2})$.

2.1. The set of all functions from X to Y.

Let X and Y be sets. We denote the set of all functions $f : X \to Y$ by Y^X. Why such a strange notation? The following simple and useful result gives the motivation. Recall that for $n \in \mathbb{Z}^+$, we put $[n] = \{1, 2, \ldots, n\}$, and we also put $[0] = \varnothing$. Thus $\#[n] = n$ for all $n \in \mathbb{N}$.

Proposition 2.4. *Let $m, n \in \mathbb{N}$. Then we have*

$$\#[m]^{[n]} = m^n.$$

In words: the set of all functions from $\{1, \ldots, n\}$ to $\{1, \ldots, m\}$ has cardinality m^n.

Proof. To define a function $f : \{1, \ldots, n\} \to \{1, \ldots, m\}$, we must specify a sequence of elements $f(1), \ldots, f(n)$ in $\{1, \ldots, m\}$. There are m possible choices for $f(1)$, also m possible choices for $f(2)$, and so forth, up to m possible choices for $f(n)$, and these choices are independent. Thus we have $m \cdots m$ n times $= m^n$ choices overall. □

2.2. Injective functions.

From the perspective of our course, the most important material on functions are the concepts *injectivity, surjectivity and bijectivity* and the relation of these properties with the existence of inverse functions.

A function $f : X \to Y$ is **injective** if every element y of the codomain is associated to at most one element $x \in X$. That is, f is injective if for all $x_1, x_2 \in X$, $f(x_1) = f(x_2)$ implies $x_1 = x_2$.

Let us meditate a bit on the property of injectivity. One way to think about it is via a horizontal line test: a function is injective if and only if each horizontal line $y = c$ intersects the graph of f in **at most** one point. Another way to think about an injective function is as a function which entails no loss of information. That is, for an injective function, if your friend tells you $x \in X$ and you tell me $f(x) \in Y$, then I can, in principle, figure out what x is because it is uniquely determined.

Consider for instance the two functions $f(x) = x^2$ and $f(x) = x^3$. The first function $f(x) = x^2$ is not injective: if y is any positive real number then there are two x-values such that $f(x) = y$, $x = \sqrt{y}$ and $x = -\sqrt{y}$. Or, in other words, if $f(x) = x^2$ and I tell you that $f(x) = 1$, then you are in doubt as to what x is: it could be either $+1$ or -1. On the other hand, $f(x) = x^3$ is injective, so if I tell you that $f(x) = x^3 = 1$, then we can conclude that $x = 1$.

How can we verify in practice that a function is injective? One way is to construct an inverse function, which we will discuss further later. But in the special case when $f : \mathbb{R} \to \mathbb{R}$ is a continuous function, the methods of calculus give useful criteria for injectivity.

Before stating the result, let us first recall the definitions of increasing and decreasing functions. A function $f : \mathbb{R} \to \mathbb{R}$ is **(strictly) increasing** if for all $x_1, x_2 \in \mathbb{R}$, $x_1 < x_2 \implies f(x_1) < f(x_2)$. Similarly, f is **(strictly) decreasing** if for all $x_1, x_2 \in \mathbb{R}$, $x_1 < x_2 \implies f(x_1) > f(x_2)$. Notice that a function which is increasing or decreasing is injective. The "problem" is that a function need not be either increasing or decreasing, although "well-behaved" functions of the sort one encounters in calculus have the property that their domain can be broken up into intervals on which the function is either increasing or decreasing. For instance, the function $f(x) = x^2$ is decreasing on $(-\infty, 0)$ and increasing on $(0, \infty)$.

Theorem 2.5. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function.*
a) If f is injective, then f is either increasing or decreasing.
b) If f is differentiable and either $f'(x) > 0$ for all $x \in \mathbb{R}$ or $f'(x) < 0$ for all $x \in \mathbb{R}$, then f is injective.

It is something of a sad reflection on our calculus curriculum that useful and basic facts like this are not established in a standard calculus course. However, the full details are somewhat intricate. We sketch a proof below.

Proof. We prove part a) by contraposition: that is, we assume that f is continuous and *neither* increasing nor decreasing, and we wish to show that it is not injective. Since f is not decreasing, there exist $x_1 < x_2$ such that $f(x_1) \leq f(x_2)$. Since f is not increasing, there exist $x_3 < x_4$ such that $f(x_3) \geq f(x_4)$. If $f(x_3) = f(x_4)$. We claim that it follows that there exist $a < b < c$ such that either
Case 1:$f(b) \geq f(a)$ and $f(b) \geq f(c)$, or
Case 2: $f(b) \leq f(a)$ and $f(b) \leq f(c)$.
This follows from a somewhat tedious consideration of cases as to in which order the four points x_1, x_2, x_3, x_4 occur, which we omit here. Now we apply the Intermediate Value Theorem to f on the intervals $[a, b]$ and $[b, c]$. In Case 1, every number smaller than $f(b)$ but sufficiently close to it is assumed both on the interval $[a, b]$ and again on the interval $[b, c]$, so f is not injective. In Case 2, every number larger than $f(b)$ but sufficiently close to it is assumed both on the interval $[a, b]$ and again on $[b, c]$, so again f is not injective.
As for part b), we again go by contraposition and assume that f is not injective: that is, we suppose that there exist $a < b$ such that $f(a) = f(b)$. Applying the Mean Value Theorem to f on $[a, b]$, we get that there exists c, $a < c < b$, such that

$$f'(c) = \frac{f(b) - f(a)}{b - a} = 0,$$

contradicting the assumption that $f'(x)$ is always positive or always negative. □

Remark: The proof shows that we could have replaced part b) with the apparently weaker hypothesis that for all $x \in \mathbb{R}$, $f'(x) \neq 0$. However, it can be shown that this is equivalent to f' always being positive or always being negative, a consequence of the **Intermediate Value Theorem For Derivatives**.

Example 2.6. *a) Let $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = \arctan x$. We claim f is injective. Indeed, it is differentiable and its derivative is $f'(x) = \frac{1}{1+x^2} > 0$ for all $x \in \mathbb{R}$. Therefore f is strictly increasing, hence injective.*
b) Let $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = -x^3 - x$. We claim f is injective. Indeed, it is

differentiable and its derivative is $f'(x) = -3x^2 - 1 = -(3x^2 + 1) < 0$ *for all* $x \in \mathbb{R}$. *Therefore* f *is strictly decreasing, hence injective.*

Example 2.7. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be given by* $f(x) = x^3$. *One meets this function in precalculus and calculus mathematics, and one certainly expects it to be injective. Unfortunately the criterion of Theorem 2.5 falls a bit short here: the derivative is* $f'(x) = 3x^2$, *which is always non-negative but is* 0 *at* $x = 0$.

We will show "by hand" that f *is indeed injective. Namely, let* $x_1, x_2 \in \mathbb{R}$ *and suppose* $x_1^3 = x_2^3$. *Then*

$$0 = x_1^3 - x_2^3 = (x_1 - x_2)(x_1^2 + x_1x_2 + x_2^2).$$

Seeking a contradiction, we suppose that $x_1 \neq x_2$. *Then* $x_1 - x_2 \neq 0$, *so we can divide through by it, getting*

$$0 = x_1^2 + x_1x_2 + x_2^2 = (x_1 + \frac{x_2}{2})^2 + \frac{3}{4}x_2^2.$$

Because each of the two terms in the sum is always non-negative, the only way the sum can be zero is if

$$(x_1 + \frac{x_2}{2})^2 = \frac{3}{4}x_2^2 = 0.$$

The second equality implies $x_2 = 0$, *and plugging this into the first inequality gives* $x_1^2 = 0$ *and thus* $x_1 = 0$. *So* $x_1 = 0 = x_2$*: contradiction.*

We gave a proof of the injectivity of $f : x \mapsto x^3$ to nail down the fact that Theorem 2.5 gives a sufficient but not necessary criterion for a differentiable function to be injective. But we would really like to able to *improve* Theorem 2.5 so as to handle this example via the methods of caclulus. For instance, let n be a positive integer. Then we equally well believe that the function $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^{2n+1}$ should be injective. It is possible to show this using the above factorization method....but it is real work to do so. The following criterion comes to the rescue to do this and many other examples easily.

Theorem 2.8. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a differentiable function.*
a) Suppose that $f'(x) \geq 0$ *for all* x *and that there is no* $a < b$ *such that* $f'(x) = 0$ *for all* $x \in (a, b)$. *Then* f *is strictly increasing (hence injective).*
b) Suppose that $f'(x) \leq 0$ *for all* x *and that there is no* $a < b$ *such that* $f'(x) = 0$ *for all* $x \in (a, b)$. *Then* f *is strictly decreasing (hence injective).*

Proof. We prove part a); the proof of part b) is identical. Again we go by contrapositive: suppose that f is not strictly increasing, so that there exists $a < b$ such that $f(a) \leq f(b)$. If $f(a) < f(b)$, then applying the Mean Value Theorem, we get a c in between a and b such that $f'(c) < 0$, contradiction. So we may assume that $f(a) = f(b)$. Then, by exactly the same MVT argument, $f'(x) \geq 0$ for all x implies that f is at least weakly increasing, i.e., $x_1 \leq x_2 \implies f(x_1) \leq f(x_2)$. But a weakly increasing function f with $f(a) = f(b)$ must be constant on the entire interval $[a, b]$, hence $f'(x) = 0$ for all x in (a, b), contradicting the hypothesis. □

Worked Exercise 2.1. *We will show that for any* $n \in \mathbb{Z}^+$, *the function* $f : \mathbb{R} \to \mathbb{R}$ *given by* $x \mapsto x^{2n+1}$ *is injective. Indeed we have* $f'(x) = (2n+1)x^{2n}$, *which is non-negative for all* $x \in \mathbb{R}$ *and is* 0 *only at* $x = 0$. *So Theorem 2.8a) applies to show that* f *is strictly increasing, hence injective.*

2.3. **Surjective functions.** A function $f : X \to Y$ if its image $f(X)$ is equal to the codomain Y. More plainly, for all $y \in Y$, there is $x \in X$ such that $f(x) = y$.

In many ways surjectivity is the "dual property" to injectivity. For instance, it can also be verified by a horizontal line test: a function f is surjective if and only if each horizontal line $y = c$ intersects the graph of f in **at least one point**.

Worked Exercise 2.2. *Let m and b be real numbers. Is $f(x) = mx+b$ surjective?*

Solution: It is surjective if and only if $m \neq 0$. First, if $m = 0$, then $f(x) = b$ is a constant function: it maps all of $\mathbb{R}$ to the single point b and therefore is at the opposite extreme from being surjective. Conversely, if $m \neq 0$, write $y = mx + b$ and solve for x: $x = \frac{y-b}{m}$. Note that this argument also shows that if $m \neq 0$, f is injective: given an arbitary y, we have solved for a unique value of x.

By the intermediate value theorem, if a continuous function $f : \mathbb{R} \to \mathbb{R}$ takes on two values $m \leq M$, then it also takes on every value in between. In particular, if a continuous function takes on arbitrarily large values and arbitrarily small values, then it is surjective.

Theorem 2.9. *Let $a_0, \ldots, a_n \in \mathbb{R}$ and suppose $a_n \neq 0$. Let $P : \mathbb{R} \to \mathbb{R}$ by*

$$P(x) = a_n x^n + \ldots + a_1 x + a_0.$$

Thus P is a polynomial of degree n. Then: P is surjective if and only if n is odd.

Proof. Suppose that n is odd. Then, if the leading term a_n is positive, then

$$\lim_{x\to\infty} P(x) = +\infty, \ \lim_{x\to-\infty} P(x) = -\infty,$$

whereas if the leading term a_n is negative, then

$$\lim_{x\to\infty} P(x) = -\infty, \ \lim_{x\to-\infty} P(x) = +\infty,$$

so either way P takes on arbitarily large and small values. By the Intermediate Value Theorem, its range must be all of $\mathbb{R}$.
Now suppose n is even. Then if a_n is positive, we have

$$\lim_{x\to\infty} P(x) = \lim_{x\to-\infty} P(x) = +\infty.$$

It follows that there exists a non-negative real number M such that if $|x| \geq M$, $P(x) \geq 0$. On the other hand, since the restriction of P to $[-M, M]$ is a continuous function on a closed interval, it is bounded below: there exists a real number m such that $P(x) \geq m$ for all $x \in [-M, M]$. Therefore $P(x) \geq m$ for all x, so it is not surjective. Similarly, if a_n is negative, we can show that P is bounded above so is not surjective. □

2.4. **Bijective functions.**

A function $f : X \to Y$ is **bijective** if it is both injective and surjective.

Exercise 2.3. *Show: or any set X, the identity function $1_X : X \to X$ by $1_X(x) = x$ is bijective.*

Exercise 2.4. *Determine which of the functions introduced so far in this section are bijective.*

A function is bijective iff for every $y \in Y$, there exists a unique $x \in X$ such that $f(x) = y$.

The following result is easy but of the highest level of importance.

Theorem 2.10. *For a function $f : X \to Y$, the following are equivalent:*
(i) f is bijective.
(ii) The inverse relation $f^{-1} : Y \to X = \{(f(x), x) \mid x \in X\}$ is itself a function.

Proof. Indeed, we need f to be surjective so that the domain of f^{-1} is all of Y and we need it to be injective so that each y in Y is associated to no more than one x value. □

2.5. **Composition of functions.**

Probably the most important and general property of functions is that they can, under the right circumstances, be *composed*.[5] For instance, in calculus, complicated functions are built up out of simple functions by plugging one function into another, e.g. $\sqrt{x^2+1}$, or $e^{\sin x}$, and the most important differentiation rule – the Chain Rule – tells how to find the derivative of a composition of two functions in terms of the derivatives of the original functions.

Let $f : X \to Y$ and $g : Y \to Z$: that is, the codomain of f is equal to the domain of g. Then we can define a new function $g \circ f : X \to Z$ by:

$$x \mapsto g(f(x)).$$

Remark: Note that $g \circ f$ means first perform f and then perform g. Thus function composition proceeds from right to left, counterintuitively at first. There was a time when this bothered mathematicians enough to suggest writing functions *on the right*, i.e., $(x)f$ rather than $f(x)$. But that time is past.

Remark: The condition for composition can be somewhat relaxed: it is not necessary for the domain of g to equal the codomain of f. What is precisely necessary and sufficient is that for every $x \in X$, $f(x)$ lies in the domain of g, i.e.,

$$\text{Range}(f) \subseteq \text{Codomain}(g).$$

Example: The composition of functions is generally *not* commutative. In fact, if $g \circ f$ is defined, $f \circ g$ need not be defined at all. For instance, suppose $f : \mathbb{R} \to \mathbb{R}$ is the function which takes every rational number to 1 and every irrational number to 0 and $g : \{0, 1\} \to \{a, b\}$ is the function $0 \mapsto b$, $1 \mapsto a$. Then $g \circ f : \mathbb{R} \to \{a, b\}$ is defined: it takes every rational number to a and every irrational number to b. But $f \circ g$ makes no sense at all:

$$f(g(0)) = f(b) = ???.$$

Remark: Those who have taken linear algebra will notice the analogy with the multiplication of matrices: if A is an $m \times n$ matrix and B is an $n \times p$ matrix, then the product AB is defined, an $m \times p$ matrix. But if $m \neq p$, the product BA is not defined. (In fact this is more than an analogy, since an $m \times n$ matrix A can be viewed as a linear transformation $L_A : \mathbb{R}^n \to \mathbb{R}^m$. Matrix multiplication is indeed

[5] This is a special case of the composition of relations described in §$X.X$, but since that was optional material, we proceed without assuming any knowledge of that material.

a special case of composition of functions.)

Even when $g \circ f$ and $f \circ g$ are both defined – e.g. when $f, g : \mathbb{R} \to \mathbb{R}$, they need not be equal. This is again familiar from precalculus mathematics. If $f(x) = x^2$ and $g(x) = x + 1$, then

$$g(f(x)) = x^2 + 1, \text{ whereas } f(g(x)) = (x+1)^2 = x^2 + 2x + 1.$$

On the other hand, function composition is always **associative**: if $f : X \to Y$, $g : Y \to Z$ and $h : Z \to W$ are functions, then we have

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

Indeed the proof is trivial, since both sides map $x \in X$ to $h(g(f(x)))$.[6]

Exercise: Let $f : X \to Y$.
a) Show that $f \circ 1_X = f$.
b) Show that $1_Y \circ f = f$.

2.6. Basic facts about injectivity, surjectivity and composition.

Here we establish a small number of very important facts about how injectivity, surjectivity and bijectivity behave with respect to function composition. First:

Theorem 2.11. *Let $f : X \to Y$ and $g : Y \to Z$ be two functions.*
a) If f and g are injective, then so is $g \circ f$.
b) If f and g are surjective, then so is $g \circ f$.
c) If f and g are bijective, then so is $g \circ f$.

Proof. a) We must show that for all $x_1,\ x_2 \in X$, if $g(f(x_1)) = g(f(x_2))$, then $x_1 = x_2$. But put $y_1 = f(x_1)$ and $y_2 = f(x_2)$. Then $g(y_1) = g(y_2)$. Since g is assumed to be injective, this implies $f(x_1) = y_1 = y_2 = f(x_2)$. Since f is also assumed to be injective, this implies $x_1 = x_2$.
b) We must show that for all $z \in Z$, there exists at least one x in X such that $g(f(x)) = z$. Since $g : Y \to Z$ is surjective, there exists $y \in Y$ such that $g(y) = z$. Since $f : X \to Y$ is surjective, there exists $x \in X$ such that $f(x) = y$. Then $g(f(x)) = g(y) = z$.
c) Finally, if f and g are bijective, then f and g are both injective, so by part a) $g \circ f$ is injective. Similarly, f and g are both surjective, so by part b) $g \circ f$ is surjective. Thus $g \circ f$ is injective and surjective, i.e., bijective, qed. □

Now we wish to explore the other direction: suppose we know that $g \circ f$ is injective, surjective or bijective? What can we conclude about the "factor" functions f and g?

The following example shows that we need to be careful.

Example: Let $X = Z = \{0\}$, let $Y = \mathbb{R}$. Define $f : X \to Y$ be $f(0) = \pi$ (or your favorite real number; it would not change the outcome), and let f be the constant function which takes every real number y to 0: note that this is the unique function from $\mathbb{R}$ to $\{0\}$. We compute $g \circ f$: $g(f(0)) = g(\pi) = 0$. Thus $g \circ f$ is the identity function on X: in particular it is bijective. However, both f and g are far

[6] As above, this provides a conceptual reason behind the associativity of matrix multiplication.

from being bijective: the range of f is only a single point $\{\pi\}$, so f is not surjective, whereas g maps every real number to 0, so is not injective.

On the other hand, something is true: namely the "inside function" f is injective, and the outside function g is surjective. This is in fact a general phenomenon.

Theorem 2.12. *(Green and Brown Fact) Let $f : X \to Y$ and $g : Y \to Z$ be functions.*
a) If $g \circ f$ is injective, then f is injective.
b) If $g \circ f$ is surjecitve, then g is surjective.
c) If $g \circ f$ is bijective, then f is injective and g is surjective.

Proof. a) We proceed by contraposition: suppose that f is not injective: then there exist $x_1 \neq x_2$ in X such that $f(x_1) = f(x_2)$. But then $g(f(x_1)) = g(f(x_2))$, so that the distinct points x_1 and x_2 become equal under $g \circ f$: that is, $g \circ f$ is not injective.
b) Again by contraposition: suppose that g is not surjective: then there exists $z \in Z$ such that for no y in Y do we have $z = g(y)$. But then we certainly cannot have an $x \in X$ such that $z = g(f(x))$, because if so taking $y = f(x)$ shows that z is in the range of g, contradiction.
c) If $g \circ f$ is bijective, it is injective and surjective, so we apply parts a) and b). □

Remark: The name of Theorem 2.12 comes from the Spring 2009 version of Math 3200, when I presented this result using green and brown chalk, decided it was important enough to have a name, and was completely lacking in inspiration.

2.7. **Inverse Functions.**

Finally we come to the last piece of the puzzle: let $f : X \to Y$ be a function. We know that the inverse relation f^{-1} is a function if and only if f is injective and surjective. But there is another (very important) necessary and sufficient condition for invertibility in terms of function composition. Before stating it, recall that for a set X, the identity function 1_X is the function from X to X such that $1_X(x) = x$ for all $x \in X$. (Similarly $1_Y(y) = y$ for all $y \in Y$.)

We say that a function $g : Y \to X$ is the **inverse function** to $f : X \to Y$ if both of the following hold:

(IF1) $g \circ f = 1_X$: i.e., for all $x \in X$, $g(f(x)) = x$.
(IF2) $f \circ g = 1_Y$: i.e., for all $y \in Y$, $f(g(y)) = y$.

In other words, g is the inverse function to f if applying one function and then the other – in either order! – brings us back where we started.

The point here is that g is supposed to be related to f^{-1}, the inverse relation. Here is the precise result:

Theorem 2.13. *Let $f : X \to Y$.*
a) The following are equivalent:
(i) f is bijective.
(ii) The inverse relation $f^{-1} : Y \to X$ is a function.

(iii) f has an inverse function g.
b) When the equivalent conditions of part a) hold, then the inverse function g is uniquely determined: it is the function f^{-1}.

Proof. a) We already know the equivalence of (i) and (ii): this is Theorem 2.10 above.
(ii) $\implies$ (iii): Assume (ii), i.e., that the inverse relation f^{-1} is a function. We claim that it is then the inverse function to f in the sense that $f^{-1} \circ f = 1_X$ and $f \circ f^{-1} = 1_Y$. We just do it: for $x \in X$, $f^{-1}(f(x))$ is the unique element of X which gets mapped under f to $f(x)$: since x is such an element and the uniqueness is assumed, we must have $f^{-1}(f(x)) = x$. Similarly, for $y \in Y$, $f^{-1}(y)$ is the unique element x of X such that $f(x) = y$, so $f(f^{-1}(y)) = f(x) = y$.
(iii) $\implies$ (i): We have $g \circ f = 1_X$, and the identity function is bijective. By the Green and Brown Fact, this implies that f is injective. Similarly, we have $f \circ g = 1_Y$ is bijective, so by the Green and Brown Fact, this implies that f is surjective. Therefore f is bijective.[7]
b) Suppose that we have any function $g : Y \to X$ such that $g \circ f = 1_X$ and $f \circ g = 1_Y$. By the proof of part a), we know that f is bijective and thus the inverse relation f^{-1} is a function such that $f^{-1} \circ f = 1_X$, $f \circ f^{-1} = 1_Y$. Thus

$$g = g \circ 1_Y = g \circ (f \circ f^{-1}) = (g \circ f) \circ f^{-1} = 1_X \circ f^{-1} = f^{-1}.$$

□

In summary, for a function f, being bijective, having the inverse relation (obtained by "reversing all the arrows") be a function, and having another function g which undoes f by composition in either order, are all equivalent.

[7] A very similar argument shows that g is bijective as well.