INTRODUCTION TO PROOFS

formerly called Foundation and Proofs

Notes by Dr. Lynne H. Walling

1. Introduction: Sets and Functions

A set is a collection considered as a unit. We are familiar with many sets, such as the set of integers, the set of rational numbers, and so on. In mathematics we use certain sets so often that we have abbreviated notation for them:

$\mathbb{Z}$ is the set of integers; so $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \ldots\}$.

$\mathbb{Q}$ is the set of rational numbers; so $\mathbb{Q} = \left\{ \frac{a}{b} : \ a, b \in \mathbb{Z}, \ b \neq 0 \right\}$, meaning that $\mathbb{Q}$ is the set of all objects of the form $\frac{a}{b}$ that meet the conditions that $a, b \in \mathbb{Z}$ and $b \neq 0$ (recall that $a, b \in \mathbb{Z}$ means that a, b are elements of the set $\mathbb{Z}$).

$\mathbb{R}$ is the set of real numbers.

$\mathbb{C}$ is the set of complex numbers, so $\mathbb{C} = \{a + b\sqrt{-1} : \ a, b \in \mathbb{R} \}$.

$\{\}$ is the empty set (i.e. the set with no elements), which is also denoted by $\emptyset$.

Note: Suppose X is a set. We cannot say "choose $x \in X$" unless we know $X \neq \emptyset$; however, we can say "suppose $x \in X$", even when we don't know whether X is nonempty.

We write $\mathbb{Z}_+$ to denote the set of positive integers, $\mathbb{Q}_+$ the set of positive rational numbers, and $\mathbb{R}_+$ the set of positive real numbers. (**Note:** 0 is neither positive nor negative.) For $x \in \mathbb{R}$, recall that

$$|x| = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0. \end{cases}$$

The notation

$$A = \{x \in \mathbb{R} : \ x > \sqrt{2} \}$$

means that A is the set of all real numbers x that meet the condition $x > \sqrt{2}$. We write $A \subseteq X$ when X is a set and A is a subset of X, meaning that every element of A is also an element of X. We write $A \subsetneq X$ when A is a a proper subset of the set X, meaning that A is a subset of X but A is not equal to X. (The use of the notation $A \subset X$ is not consistent throughout mathematical literature, so we will avoid using this notation.) We write $A \not\subseteq B$ when A is not a subset of B.

Note that $\emptyset$ is the only subset of $\emptyset$.

Example: $\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$.

Note: Suppose A, X are sets. Showing $A = X$ is equivalent to showing $A \subseteq X$ and $X \subseteq A$.

There are various ways of constructing $\mathbb{R}$ from $\mathbb{Q}$ (see, for instance, S. Krantz, *Real Analysis and Foundations*, CRC Press, 1991; or K.A. Ross, *Elementary Analysis: The Theory of Calculus*, Springer, 1980).

Recall that we know that with $X = \mathbb{Z}$ or $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{C}$, for any $a, b, c \in X$ we have $a + b, -a, ab \in X$, $a + b = b + a$, $ab = ba$, and $c(a + b) = ca + cb$. Further, for $X = \mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{C}$ and $a \in X$ with $a \neq 0$, we have $\frac{1}{a} \in X$. We

also know that for $a, b \in \mathbb{Z}_+$, we have $a \leq ab$, and $a = ab$ only when $b = 1$. In addition, we know that for any $a, b \in \mathbb{C}$, we have $ab = 0$ only when $a = 0$ or $b = 0$; this means that for $a, b, c \in \mathbb{C}$ with $ab = ac$ and $a \neq 0$, we have $a(b - c) = 0$ and hence $b - c = 0$ so $b = c$.

A noteable property of $\mathbb{R}$ is that it is linearly ordered, meaning that for every $x, y \in \mathbb{R}$, either $x \leq y$ or $y \leq x$, and if $x, y \in \mathbb{R}$ with $x \leq y$ and $y \leq x$ then $x = y$. Note that any subset of $\mathbb{R}$ is also linearly ordered. We say a (nonempty) subset A of $\mathbb{R}$ is bounded above if there is some $M \in \mathbb{R}$ so that $M \geq a$ for all $a \in A$, and we say A is bounded below if there is some $m \in \mathbb{R}$ so that $m \leq a$ for all $a \in A$.

Proposition 1.1. *Suppose A is a nonempty subset of $\mathbb{Z}$.*

(1) *If A is bounded above then A contains a maximal element, meaning A is bounded above by an element of A.*
(2) *If A is bounded below then A contains a minimal element, meaning A is bounded below by an element of A.*

Proof. (1) Suppose A is bounded above by $N \in \mathbb{R}$. Choose any $c \in A$. Then there are finitely many integers between c and N, so there are finitely many $a \in A$ so that $c \leq a \leq N$; A is bounded above by the largest of these elements a.

(2) Suppose A is bounded below by $n \in \mathbb{R}$. Choose any $c \in A$. Then there are finitely many elements of a so that $n \leq a \leq c$; A is bounded below by the smallest of these elements a. □

Corollary 1.2. *Any nonempty subset of $\mathbb{Z}_+$ has a minimal element.*

Proof. If $A \subset \mathbb{Z}_+$ and A is nonempty, then A is a subset of $\mathbb{Z}$ that is bounded below by 1, and hence by the above theorem A has a minimal element. □

A cautionary tale regarding sets. Consider the following situation: "The barber is a man in town who shaves all those, and only those, men in town who do not shave themselves." Who shaves the barber?

In 1901 Bertrand Russell presented a version of this paradox to the mathematical community; this resulted in widespread fear that the foundations of mathematics were "built on quicksand". This paradox shows that a condition that contains an inherent contradiction does not determine a set.

There are many sources that discuss Russell's Paradox (easily found by searching the internet); students are encouraged to peruse these.

In mathematics, we are very often concerned with functions (also called maps). Some functions model the behaviour of complex systems, while other functions allow us to compare two sets. We are accustomed to functions that are given by a formula, as when studying Calculus. Here we develop a formal definition of a function.

Definitions. Given sets X, Y, we define the Cartesian product of X and Y as

$$\{(x, y) : \ x \in X, \ y \in Y \ \},$$

and we denote this set by $X \times Y$. (So $X \times Y$ is the set of all ordered pairs (x, y) that meet the conditions that $x \in X$ and $y \in Y$.) Note that if X or Y is the empty set, then so is $X \times Y$. With X, Y nonempty sets, a function f

from X into Y is a set of ordered pairs $f \subseteq X \times Y$ with the property that for each element $x \in X$ there is exactly one $y \in Y$ so that $(x, y) \in f$. When f is a function with $(x, y) \in f$, we write $f(x)$ to denote y. Thus using this notation, when f is a function from X to Y,

$$f = \{(x, f(x)) : \ x \in X \ \}.$$

(So a function f from X into Y pairs each element of X with exactly one element of Y, which we denote by $f(x)$.) We write $f : X \to Y$ to denote that f is a function from X into Y (so implicit in the notation $f : X \to Y$ is that X, Y are nonempty sets). Suppose $f : X \to Y$. We say X is the domain of f and Y is the codomain of f. The range of f, denoted $f(X)$, is the set

$$f(X) = \{f(x) : \ x \in X \ \},$$

i.e. the set of all values $f(x)$ where x meets the condition that $x \in X$. Since $f(x) \in Y$ for any $x \in X$, we also have

$$f(X) = \{y \in Y : \text{ for some } x \in X, \ f(x) = y \ \}.$$

More generally, for any $A \subseteq X$,

$$f(A) = \{f(x) : \ x \in A \ \}.$$

Example: Let $X = \{1, 2, 3\}$, $Y = \{4, 5, 6\}$. Then

$$X \times Y = \{(1, 4), (1, 5), (1, 6), (2, 4), (2, 5), (2, 6), (3, 4), (3, 5), (3, 6)\}.$$

Let $f = \{(1, 4), (2, 5), (3, 4)\}$, $g = \{(1, 4), (1, 5), (3, 6)\}$. Then f is a function from X into Y, since for each $x \in X$, there is exactly one $y \in Y$ so that $(x, y) \in f$. However, g is not a function from X into Y: We have $1 \in X$, but there are two values of $y \in Y$ (namely $y = 4$ and $y = 5$) so that $(1, y) \in g$; further, $2 \in X$, but there is no value of $y \in Y$ so that $(2, y) \in g$. We also have

$$f(X) = \{f(1), \ f(2), \ f(3)\} = \{4, \ 5\}.$$

Example: Define $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ by $f((m, n)) = n^2$. Let $A = \{(m, n) \in \mathbb{Z} \times \mathbb{Z} : \ n = 2m \ \}$. So $A = \{(m, 2m) : \ m \in \mathbb{Z} \ \}$. Then

$$\begin{aligned} f(A) &= \{f((m, n)) : \ (m, n) \in A \ \} \\ &= \{f((m, 2m)) : \ m \in \mathbb{Z} \ \} \\ &= \{(2m)^2 : \ m \in \mathbb{Z} \ \} \\ &= \{4m^2 : \ m \in \mathbb{Z} \ \}. \end{aligned}$$

We often need to quantify objects in mathematics, meaning we need to distinguish between a condition always being met, or the existence of a case where a condition is met. Sometimes we also need to distinguish whether there is a unique case where a condition is met. For instance, suppose we have a function $f : X \to Y$. This means that for every $x \in X$ there exists a unique $y \in Y$ so that $(x, y) \in f$. Notice that the order of the quantifying phrases is important:

"For every $x \in X$, there is a unique $y \in Y$ so that $(x, y) \in f$" means that the choice of x determines the value of y (in this particular situation, we

have $y = f(x)$). Contrastingly, suppose there exists a unique $y \in Y$ so that for every $x \in X$, $(x, y) \in f$; this means there is a unique $y \in Y$ so that

$$f = \{(x, y) : \ x \in X \ \},$$

meaning f is a constant function (as in the case $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 5$).

Notation: We use the symbol $\forall$ to denote "for all", or equivalently, "for every". We use the symbol $\exists$ to denote "there exists", and we use $\exists!$ to denote "there exists a unique", or equivalently "there exists one and only one".

Notes: (1) To show that something is unique, a standard technique is to show first that one such thing exists, and to show then that if another exists, it is equal to the first. For example, suppose X, Y are sets and $f \subseteq X \times Y$. Then f is a function from X to Y if, $\forall x \in X$, $\exists y \in Y$ so that $(x, y) \in f$, and $\forall y' \in Y$, if $(x, y') \in f$ then $y' = y$.

(2) When we write "Suppose $c \in X$" or "Choose $c \in X$" or "Take $c \in X$" without stating further assumptions on c, we mean that we are choosing c arbitrarily from X; thus anything we then conclude about c applies to every element of X.

(3) We will sometimes write "for $c \in X$" to mean "$\forall c \in X$", as the only condition being imposed on c is that it is in X. Somewhat similarly, we will sometimes write "for some $c \in X$" to mean "$\exists c \in X$".

Theorem 1.3. *Suppose $f : X \to Y$, $g : X \to Y$. Then $f = g$ if and only if $\forall x \in X$, $f(x) = g(x)$.*

Proof. First suppose that $\forall x \in X$, $f(x) = g(x)$. Thus

$$f = \{(x, f(x)) : \ x \in X \ \} = \{(x, g(x)) : \ x \in X \ \} = g.$$

Now suppose that $f = g$. Thus $(x, y) \in f$ if and only if $(x, y) \in g$. Take [arbitrary] $x \in X$, and then choose [the unique] $y \in Y$ so that $(x, y) \in f = g$; thus $y = f(x)$, and also $y = g(x)$. Hence for every $x \in X$, we have $f(x) = g(x)$. □

Definitions. We say a function $f : X \to Y$ is injective (or one-to-one, or an injection) if, $\forall x_1, x_2 \in X$, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$. (Here $\forall x_1, x_2 \in X$ means $\forall x_1 \in X$, $\forall x_2 \in X$.) We say a function $f : X \to Y$ is surjective (or onto, or a surjection) if, $\forall\ y \in Y$, $\exists\ x \in X$ so that $f(x) = y$. (Thus $f : X \to Y$ is surjective if the range of f is Y.) A function is called bijective if it is both injective and surjective.

Note: We have defined (for example) a map $f : X \to Y$ to be injective if, $\forall x_1, x_2 \in X$, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$. According to standard English usage, a definition is a precise statement of what a word or expression means. Thus saying that a map $f : X \to Y$ is injective is *equivalent* to saying that $\forall x_1, x_2 \in X$, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$.

Example: Define $f : \mathbb{Z}_+ \to \mathbb{Z}_+$ by $f(x) = x^2$. This function is injective but not surjective.

Example: Let $\mathbb{R}_{\geq 0} = \{y \in \mathbb{R} : \ y \geq 0 \ \}$. Define $g : \mathbb{R} \to \mathbb{R}_{\geq 0}$ by $g(x) = x^2$; this function is surjective, but not injective.

Example: Define $h : \mathbb{R} \to \mathbb{R}$ by $h(x) = x^3$; this function is bijective.

Example: Let $(-1, 1) = \{x \in \mathbb{R} : \ -1 < x < 1 \ \}$. Define $g : (-1, 1) \to \mathbb{R}$ by

$$g(x) = \frac{x}{1 - |x|}.$$

We claim that g is surjective. To see this, suppose $y \in \mathbb{R}$.

[**Scratch work:** We want to find $x \in (-1, 1)$ so that $g(x) = y$. Thus we want $\frac{x}{1-|x|} = y$, or equivalently, $x + y|x| = y$. Suppose $x \geq 0$; then $|x| = x$, so we want $x(1 + y) = y$, or equivalently, $x = \frac{y}{1+y}$. Suppose $x < 0$; then $|x| = -x$, so we want $x(1 - y) = y$, or equivalently, $x = \frac{y}{1-y}$. Notice that for $x \in (-1, 1)$, $1 - |x| > 0$; thus when $y = \frac{x}{1-|x|}$, we have $y \geq 0$ exactly when $x \geq 0$. So this suggests we should take $x = \frac{y}{1+|y|}$. **Note:** We will need to verify that this choice of x actually is an element of $(-1, 1)$.]

Let $x = \frac{y}{1+|y|}$. We first show that $x \in (-1, 1)$, or equivalently, that $|x| < 1$. Note that $\big|1 + |y|\big| = 1 + |y|$, so $|x| = \frac{|y|}{1+|y|}$. Certainly $|y| < 1 + |y|$, so $|x| = \frac{|y|}{1+|y|} < 1$. Hence we indeed have $x \in (-1, 1)$.

Now we show that with this choice of x, we have $g(x) = y$. When $y \geq 0$ we have $x \geq 0$, so $x = \frac{y}{1+y}$ and

$$g(x) = \frac{x}{1 - |x|} = \frac{\frac{y}{1+y}}{1 - \frac{y}{1+y}} = \frac{y}{1 + y - y} = y.$$

When $y < 0$ we have $x < 0$, so $x = \frac{y}{1-y}$ and

$$g(x) = \frac{x}{1 + x} = \frac{\frac{y}{1-y}}{1 + \frac{y}{1-y}} = \frac{y}{1 - y + y} = y.$$

Hence $g(x) = y$. Since this argument holds for all $y \in \mathbb{R}$, g is surjective.

Warning: Do not confuse the definition of injective with the definition of a function. For example, consider $f = \{(y^2, y) : \ y \in \mathbb{Z} \ \} \subseteq \mathbb{Z} \times \mathbb{Z}$. So for each $y \in \mathbb{Z}$, $\exists!\ x \in \mathbb{Z}$ so that $(x, y) \in f$ (namely $x = y^2$). But f is not a function, as, for example, $(4, 2), (4, -2) \in f$.

Proposition 1.4. *Suppose $f : X \to Y$.*

(a) *f is injective if and only if $\forall y \in f(X)$, $\exists! x \in X$ so that $f(x) = y$.*

(b) *f is bijective if and only if $\forall\ y \in Y$, $\exists!\ x \in X$ so that $f(x) = y$.*

(So when f is bijective, f gives us a one-to-one correspondence between the elements of X and the elements of Y.)

Proof. (a) Suppose first that f is injective, and suppose $y \in f(X)$. Thus $\exists x \in X$ so that $f(x) = y$. Now suppose $x' \in X$ so that $x' \neq x$. Then since f is injective, $f(x') \neq f(x) = y$. Thus x is the only element of X so that $f(x) = y$; in other words, x is the unique element of X so that $f(x) = y$. In summary, for $y \in f(X)$, $\exists! x \in X$ so that $f(x) = y$.

Now suppose that $\forall y \in f(X)$, $\exists! x \in X$ so that $f(x) = y$. Suppose $x_1, x_2 \in X$ so that $x_1 \neq x_2$; let $y_1 = f(x_1)$. By assumption, x_1 is the only element of X that f maps to y_1. Hence $y_1 \neq f(x_2)$, so $f(x_1) \neq f(x_2)$. Thus we have shown that for $x_1, x_2 \in X$ with $x_1 \neq x_2$, we have $f(x_1) \neq f(x_2)$, meaning that f is injective.

(b) Say f is bijective; then $f(X) = Y$, so by (1), $\forall y \in Y$, $\exists! x \in X$ so that $f(x) = y$.

Now suppose that $\forall y \in Y$, $\exists! x \in X$ so that $f(x) = y$. Then f is surjective, since $\forall y \in Y$, $\exists x \in X$ so that $f(x) = y$. Therefore $f(X) = Y$, and by (1), f is injective. Hence f is bijective. □

Definitions. Suppose we have functions $f : X \to Y$, $g : Y \to Z$. We define the composition of g and f, denoted $g \circ f$, by

$$(g \circ f)(x) = g(f(x)) \text{ for any } x \in X.$$

Since f assigns to $x \in X$ exactly one value $f(x) \in Y$, and g assigns to $f(x) \in Y$ exactly one value in Z, we have that $g \circ f$ is a function from X to Z, i.e. $g \circ f : X \to Z$. We say a function $f : X \to Y$ is invertible if there exists a function $g : Y \to X$ so that $g \circ f$ is the identity function on X (meaning that for all $x \in X$, $(g \circ f)(x) = x$), and $f \circ g$ is the identity function on Y. Note that when g is an inverse for f, we also have that f is an inverse for g.

Example: Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = 2x + 3$ and define $g : \mathbb{R} \to \mathbb{R}$ by $g(x) = (x - 3)/2$. Then for any $x \in \mathbb{R}$,

$$g \circ f(x) = g(f(x)) = \frac{f(x) - 3}{2} = \frac{(2x + 3) - 3}{2} = x,$$

and

$$f \circ g(x) = f(g(x)) = 2g(x) + 3 = 2 \cdot \frac{x - 3}{2} + 3 = x.$$

Hence $g \circ f$ is the identity function on the domain of f, and $f \circ g$ is the identity function on the domain of g. So f is invertible with g as an inverse.

Proposition 1.5. *Suppose $f : X \to Y$, $g : Y \to Z$, $h : Z \to W$. Then $h \circ (g \circ f) = (h \circ g) \circ f$.*

Proof. To show $h \circ (g \circ f) = (h \circ g) \circ f$, we need to show that for all $x \in X$, we have $h \circ (g \circ f)(x) = (h \circ g) \circ f(x)$. So take $x \in X$; then

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

and

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

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

Theorem 1.6. *Suppose $f : X \to Y$, $g : Y \to Z$.*

(a) *If f and g are injective then so is $g \circ f$.*
(b) *If f and g are surjective then so is $g \circ f$.*

Proof. We will prove (a) and leave (b) as an exercise.

Suppose f, g are injective, and suppose $x_1, x_2 \in X$ so that $x_1 \neq x_2$. Since f is injective, this means that $f(x_1) \neq f(x_2)$. Set $y_1 = f(x_1)$, $y_2 = f(x_2)$. Thus $y_1, y_2 \in Y$ with $y_1 \neq y_2$. Since g is injective, this means $g(y_1) \neq g(y_2)$. Substituting for y_1, y_2, this means

$$g \circ f(x_1) = g(f(x_1)) = g(y_1) \neq g(y_2) = g(f(x_2)) = g \circ f(x_2).$$

Summarising, for any $x_1, x_2 \in X$, if $x_1 \neq x_2$ then $g \circ f(x_1) \neq g \circ f(x_2)$. Hence $g \circ f$ is injective. □

Note: This theorem shows that if $f : X \to Y$ and $g : Y \to Z$ are both bijective, then $g \circ f : X \to Z$ is also bijective.

As an exercise, one proves the following.

Theorem 1.7. *Suppose $f : X \to Y$, and $g : Y \to X$, $h : Y \to X$ are inverses of f. Then $g = h$; that is, if f has an inverse then its inverse is unique.*

Theorem 1.8. *Suppose $f : X \to Y$. Then f is invertible if and only if f is bijective.*

Proof. There are 2 statements we need to prove:

(1) f is invertible if f is bijective, or equivalently, f is bijective implies f is invertible;
(2) f is invertible only if f is bijective, or equivalently, f is invertible implies f is bijective.

To show (1): Suppose f is bijective. [So we are assuming that f is injective and surjective.] Set

$$g = \{(y, x) \in Y \times X : \ (x, y) \in f \ \}.$$

Since f is bijective, $\forall y \in Y$, $\exists! x \in X$ so that $(x, y) \in f$; thus g is a function. So $g : Y \to X$, and for any $y \in Y$, $g(y) = x$ where $f(x) = y$. Now we need to show that g is an inverse of f. For this, first take any $x \in X$. Set $y = f(x)$. Thus by the definition of g, $g(y) = x$, so $(g \circ f)(x) = x$. As x was chosen arbitrarily from X, this shows $g \circ f$ is the identity function on X. Now choose any $y \in Y$. Since f is bijective, there is a unique $x \in X$ with $f(x) = y$. Thus $g(y) = x$, and hence $(f \circ g)(y) = f(x) = y$. Since y was chosen arbitrarily from Y, this shows $f \circ g$ is the identity function on Y. Hence when f is bijective, we have that f is invertible. Thus (1) $\implies$ (2).

To show (2): Suppose f is invertible. Then there is some $g : Y \to X$ so that g is an inverse of f. To show f is surjective, take $y \in Y$. Then

$$y = (f \circ g)(y) = f(g(y))$$

and $g(y) \in X$. Hence f is surjective. To show f is injective, suppose that $x_1, x_2 \in X$ with $x_1 \neq x_2$. Since $g \circ f$ is the identity map on X, we know

$$g(f(x_1)) = g \circ f(x_1) = x_1, \ g(f(x_2)) = g \circ f(x_2) = x_2,$$

so $g(f(x_1)) \neq g(f(x_2))$. Hence $f(x_1) \neq f(x_2)$ (else $g(f(x_1)) = g(f(x_2))$). Hence f is injective, and thus (2) $\implies$ (1). □

Note: Suppose $f : X \to Y$ bijective. In proving the above results, we found a "recipe" for defining $f^{-1} : Y \to X$:

For any $y \in Y$, $f^{-1}(y) = x$ where $x \in X$ so that $f(x) = y$.

Example: Suppose $a, b, c, d \in \mathbb{R}$ so that $a < b$ and $c < d$. Let $[a, b]$ denote the closed interval from a to b; that is,

$$[a, b] = \{x \in \mathbb{R} : \ a \leq x \leq b \ \}.$$

We claim there is a bijection between $[a, b]$ and $[c, d]$. Intuitively, the idea is the we stretch or shrink the interval $[a, b]$ to be the same length as $[c, d]$, and shift this. The map $f_1(x) = x - a$ will take $[a, b]$ to $[0, b - a]$, then $f_2(x) = x \cdot \frac{d-c}{b-a}$ will take $[0, b-a]$ to $[0, d-c]$, and then $f_3(x) = c+x$ will take

$[0, d-c]$ to $[c, d]$. We set $f = f_3 \circ f_2 \circ f_1$; thus we set $f(x) = c + \frac{(x-a)(d-c)}{(b-a)}$. We want to show $f : [a, b] \to [c, d]$. To do this, take $x \in [a, b]$, we have $0 \leq x - a \leq b - a$. We know $a < b$ and $c < d$, so $b - a > 0$ and $d - c > 0$. Hence

$$0 \leq \frac{(x-a)(d-c)}{(b-a)} \leq (d-c),$$

and then

$$c \leq c + \frac{(x-a)(d-c)}{(b-a)} \leq d.$$

So we indeed have that $f : [a, b] \to [c, d]$. (**Warning:** One may be tempted to argue by first *assuming* that $c \leq f(x) \leq d$, and then *deducing* that $a \leq x \leq b$, but what we need to show is that *if* $x \in [a, b]$ *then* $f(x) \in [c, d]$. In one's scratch work one might first assume that $c \leq f(x) \leq d$ and then deduce that $a \leq x \leq b$, but then one must determine whether these steps can be reversed to obtain a proof of what is needed. More generally, to prove a statement of the form "If A then B", it is **incorrect** to begin by assuming what is to be deduced.)

Now we want to show that f is bijective. So we could argue that f is injective and surjective. Using the definition of injective that we have given, it is awkward to show that f is injective; in §3 we will use a result of §2 to produce an equivalent definition of injective, using the "contrapositive" of the definition we have given. (The contrapositive of a statement of the form "If A holds then B holds" is "If B does not hold then A does not hold"; in §2 we will see that the contrapositive of a statement is equivalent to the statement.) In arguing that this particular function is surjective, we would actually produce the inverse of f, so here we will argue that f is bijective by finding $g : [c, d] \to [a, b]$ so that $g \circ f$ is the identity map on $[a, b]$ and $f \circ g$ is the identity map on $[c, d]$.

Using the strategy we used to construct f, reversing the roles of a and c and the roles of b and d, we define $g(x) = a + \frac{(x-c)(b-a)}{(d-c)}$. [Alternatively, we could set $y = f(x)$ and solve for x, finding that $x = a + \frac{(y-c)(b-a)}{(d-c)}$, and then setting $g(y) = a + \frac{(y-c)(b-a)}{(d-c)}$.] Then for $x \in [c, d]$, we have $c \leq x \leq d$ and hence $a \leq a + \frac{(x-c)(b-a)}{(d-c)} \leq b$; so $g : [c, d] \to [a, b]$. Also, for $x \in [a, b]$,

$$\begin{aligned} g \circ f(x) &= g(f(x)) \\ &= a + (f(x) - c)\frac{(b-a)}{(d-c)} \\ &= a + \left(c + (x-a)\frac{(d-c)}{(b-a)} - c\right)\frac{(b-a)}{(d-c)} \\ &= a + (x-a) \\ &= x. \end{aligned}$$

Similarly,

$$\begin{aligned} f \circ g(x) &= f(g(x)) \\ &= c + (g(x) - a)\frac{(d-c)}{(b-a)} \\ &= c + \left(a + (x-c)\frac{(b-a)}{(d-c)} \right) \frac{(d-c)}{(b-a)} \\ &= x. \end{aligned}$$

Thus $g : [c,d] \to [a,b]$ is the inverse of f.

Note: Given the above definitions of f and g, it is necessary to ensure that $f([a,b]) \subseteq [c,d]$ and that $g([c,d]) \subseteq [a,b]$, else we cannot claim that $f : [a,b] \to [c,d]$ and $g : [c,d] \to [a,b]$, and knowing the domains and codomains of f and g is necessary to apply the preceding theorem. We could define $f : [a,b] \to \mathbb{R}$ by $f(x) = c + \frac{(x-a)(d-c)}{(b-a)}$ and $g : [0,2] \to \mathbb{R}$ by $g(x) = a + \frac{(c-x)(b-a)}{(d-c)}$, and then proceed mechanically to argue that $g \circ f(x) = x$, $f \circ g(x) = x$; this will work because we could have extended the domains of f and g to $\mathbb{R}$, but unless $c = 0$ and $d = 2$, this does not *prove* that there is a bijection between $[a,b]$ and $[0,1]$.

In the exercises, one proves the following. (Part (a) of this theorem is an exercise for §3, and part (b) is an exercise for this section.)

Proposition 1.9. *Suppose $f : X \to Y$, $g : Y \to X$ so that $g \circ f$ is the identity map on X, meaning that for all $x \in X$, we have $g \circ f(x) = x$.*

(a) *Suppose g is injective; then $f \circ g$ is the identity map on Y (and hence $g = f^{-1}$).*
(b) *Suppose f is surjective; then $f \circ g$ is the identity map on Y (and hence $g = f^{-1}$).*

As an exercise, one also proves the following.

Theorem 1.10. *Suppose $f : X \to Y$ and $g : Y \to Z$ are bijective (and hence we know $g \circ f$ is bijective). Then $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

One also proves this useful result.

Proposition 1.11. *Suppose $f : X \to Y$ is bijective, and $A \subseteq X$. Set $B = \{ x \in X : \ x \notin A \}$. (Standard notation for B is $X \setminus A$.) Then $f(A) \cap f(B) = \emptyset$.*