Introduction to Proofs

Notes by Dr. Lynne H. Walling and Dr. Steffi Zegowitz

September 2018

The Introduction to Proofs course is organised into the following nine sections.

1. **Introduction: sets and functions**
   Notation and discussion of sets; Cartesian products; definition of a function; injective, surjective, and bijective functions; composition of functions; invertible functions; proof by contradiction.

2. **Truth tables, equivalences and contrapositive**
   Notation used in truth tables; logical equivalence; the contrapositive of a statement.

3. **Negations and contrapositives of statements with quantifiers**
   Notation for a statements dependent on a variable; an algorithmic approach for negating complex statements; an equivalent definition for a function being injective.

4. **Set operations**
   Union, intersection, complement of a set and difference of two sets; De Morgan's Laws and similar statements; inverse image of a set under a function; relations between inverse images.

5. **Partitioning sets, equivalence relations and congruences**
   Relations on a set; reflexive, symmetric, and transitive relations; a correspondence between a partition and an equivalence relation; congruences.

6. **Algorithms and recursion**
   Division algorithm; greatest common factors; least common multiples; Euclid's algorithm; the Chinese Remainder Theorem.

7. **Mathematical induction and the Fundamental Theorem of Arithmetic**
   Proof by induction and strong induction; definition of a prime number; proof that there are infinitely many prime numbers; an application of the Fundamental Theorem of Arithmetic.

8. **Cardinality**
   Definition of a countable set; Cantor-Schröder-Bernstein Theorem; results regarding the cardinality of subsets of the positive integers; proof that the Cartesian product $\mathbb{N} \times \mathbb{N}$ is countable; proof that the union of a countable number of pairwise disjoint countable sets is a countable set.

9. **Uncountable sets and power sets**
   Cantor's diagonalisation proof that the unit interval $(0, 1)$ is uncountable; proof of Cantor's Theorem that the cardinality of the power set of a set A is strictly larger than the cardinality of A.

References for the course:

- Larry Gerstein, *Discrete Mathematics and Algebraic Structures*, W.H. Freeman and Company, 1987.
- D.J. Velleman, *How to Prove It: A Structured Approach*, Cambridge University Press, 2006.
- P.J. Eccles, *An Introduction to Mathematical Reasoning: Numbers, Sets and Functions*, Cambridge University Press, 1997.

1 Introduction: Sets and Functions

A mathematical theory is not to be considered complete until you have made it so clear that you can explain it to the first man whom you meet on the street.
- David Hilbert, mathematician -

This course is heavily based on

(i) definitions that are used to capture mathematical concepts, and

(ii) using these definitions to solve mathematical problems.

1.1 Sets

We will now start by defining some terms and introducing some notation that will be used frequently.

Definition 1.1. *A **set** is a collection of distinct objects, 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, there are certain sets which are used so often that we have abbreviated notations for them:

Notation.

- $\mathbb{Z}$ *is the **set of integers**, that is* $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$.
- $\mathbb{N}$ *is the **set of natural numbers**, that is* $\mathbb{N} = \{1, 2, 3, \ldots\}$.
- $\mathbb{N}_0$ *is the **set of natural numbers including zero**, that is* $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$.
- $\mathbb{Q}$ *is the **set of rational numbers**, that is* $\mathbb{Q} = \left\{ \frac{a}{b} : \ a, b \in \mathbb{Z},\ b \neq 0 \ \right\}$.
- $\mathbb{Q}_+$ *is the **set of positive rational numbers**, that is* $\mathbb{Q}_+ = \left\{ \frac{a}{b} : \ a, b \in \mathbb{N}, \right\}$.
- $\mathbb{R}$ *is the **set of real numbers**.*
- $\mathbb{R}_+$ *is the **set of positive real numbers**.*
- $\mathbb{C}$ *is the **set of complex numbers**, that is* $\mathbb{C} = \{a + ib : \ a, b \in \mathbb{R} \ \}$.
- $\{\}$ *is the **empty set** (i.e. the set with no elements), which is also denoted by* $\emptyset$.

Remark. *We have that* 0 *is neither positive nor negative.*

The notation

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

means that A is the set of all real numbers x that satisfy the condition that $x > \sqrt{2}$.

Definition 1.2. *A set A is a **subset** of a set X, denoted by $A \subseteq X$, if every element of A is also an element of X. We write $A \nsubseteq B$ when A is <u>not</u> a subset of B.*

Definition 1.3. *A set A is a **proper subset** of a set X, denoted by $A \subset X$, if A is a subset of X but $A \neq X$.*

Example. *Let $A = \{4n : n \in \mathbb{N}\}$ and $B = \{2n : n \in \mathbb{N}\}$.*

(i) Let $x \in A$. Then $x = 4n$ for some $n \in \mathbb{N}$. Further, $x = 4n = 2(2n) = 2m$, for some $m \in \mathbb{N}$. Hence, $x \in B$. Therefore, $A \subseteq B$.

(ii) Take $x = 10$. Then $x \in B$ but $x \notin A$. Therefore, $B \not\subseteq A$.

Combining (i) and (ii), it follows that $A \subset B$.

Remark. *We have that $\emptyset$ is the only subset of $\emptyset$.*

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

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

Remark. *Suppose A and X are sets. Showing that $A = X$ is equivalent to showing that $A \subseteq X$ and $X \subseteq A$.*

Recall that we have the following properties of sets:

- Let $X = \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$ and let $a, b, c \in X$. Then $a + b, -a, ab \in X$ and

$$a + b = b + a,$$

$$ab = ba,$$

$$c(a + b) = ca + cb.$$

- Let $X = \mathbb{Q}, \mathbb{R}$, or $\mathbb{C}$ and let $a \in X$ with $a \neq 0$. Then $\frac{1}{a} \in X$.
- For $a, b \in \mathbb{N}$, we have that $a \leq ab$, and $a = ab$ if and only if $b = 1$.
- For any $a, b \in \mathbb{C}$, we have that $ab = 0$ if and only if $a = 0$ or $b = 0$.
- $\mathbb{R}$ (and any subset of $\mathbb{R}$) is **linearly ordered**: Let $x, y, z \in \mathbb{R}$. Then
  - (i) either $x \leq y$ or $y \leq x$.
  - (ii) if $x \leq y$ and $y \leq x$, then $x = y$.
  - (iii) if $x \leq y$ and $y \leq z$, then $x \leq z$.

Remark. *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." But who shaves the barber? In 1901, Bertrand Russell presented a version of this paradox to the mathematical community. This resulted in a 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 which discuss Russell's Paradox (easily found via Google) and students are encouraged to peruse these.*

1.2 Functions

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 will now develop a formal definition for a function.

Definition 1.4. *Given two sets X, Y, we define the **Cartesian product** of X and Y, denoted by $X \times Y$, by*

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

Remark. *Note that if $X = \emptyset$ or $Y = \emptyset$, then $X \times Y = \emptyset$.*

Example. *We have that*

$$\mathbb{R} \times \mathbb{R} = \{(x, y) : \ x, y \in \mathbb{R} \ \}.$$

So $\mathbb{R} \times \mathbb{R}$ is the Cartesian plane.

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

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

Definition 1.5. *Let X, Y be non-empty sets. A **function** f from X into Y, denoted by $f : X \to Y$, is a set $f \subseteq X \times Y$ which satisfies that for each element $x \in X$ there exists exactly one element $f(x) \in Y$ such that $(x, f(x)) \in f$. That is,*

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

Definition 1.6. *Suppose that $f : X \to Y$. We say that X is the **domain** of f and Y is the **co-domain** of f. The **range** (or **image**) of f, denoted by $f[X]$, is the set*

$$\begin{aligned} f[X] &= \{f(x) : \ x \in X \ \}. \\ &= \{y \in Y : y = f(x), \ \text{for } x \in X\}. \end{aligned}$$

Example. *Let $X = \{1, 2, 3\}$ and $Y = \{4, 5, 6\}$. 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 $y \in Y$ such that $(x, y) \in f$. However, g is not a function from X into Y. This follows since $1 \in X$, but there are two values of $y \in Y$ (namely $y = 4$ and $y = 5$) such that $(1, y) \in g$. In addition, $2 \in X$, but there is no value of $y \in Y$ such that $(2, y) \in g$.

Computing the range of f, we get

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

Example. *Let $X = \{x \in \mathbb{R} : \ -2 \leq x \leq 2 \ \}$, $Y = \mathbb{R}$, and*

$$f = \{(x, x^2) : x \in \mathbb{R} \text{ and } -2 \leq x \leq 2\}.$$

Then $f : X \to \mathbb{R}$ is defined by $f(x) = x^2$, for all $x \in X$.

Example. *Define $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ by $f((m,n)) = n^2$, for all $m, n \in \mathbb{Z}$. Then the range range of f is $\{n^2 : n \in \mathbb{Z}\}$.*

Now, let $A = \{(m,n) : m, n \in \mathbb{Z}, n = 2m\}$. Then $A = \{(m, 2m) : m \in \mathbb{Z}\}$. Hence, we have

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

We often need to quantify objects in mathematics, that is we need to distinguish between a criteria always being met, or the existence of a case where a criteria is met. Sometimes we also need to distinguish whether there is a unique case where a criteria is met.

Notation.

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

Remark. *To show that something is unique, we first show that one such case exists, and then proceed to show that if another case exists, it is equal to the first case.*

Example. *Let X, Y be two non-empty sets and $f \subseteq X \times Y$. Then f is a function from X to Y if*

(i) $\forall\, x \in X$, $\exists\, y \in Y$ such that $(x,y) \in f$, and

(ii) $\forall\, y' \in Y$, if $(x,y') \in f$ then $y' = y$.

We have the following theorem.

Theorem 1.7. *Suppose $f : X \to Y$ and $g : X \to Y$. Then $f = g$ if and only if $f(x) = g(x)$, for all $x \in X$.*

Proof.

($\Rightarrow$) First, suppose that $f = g$. Take $x \in X$ and choose $y \in Y$ such that $(x,y) \in f$. Then $(x,y) \in g$ (since $f = g$) and $y = f(x) = g(x)$. Similarly, if $(x',y') \in g$, then $(x',y') \in f$ and $y' = g(x') = f(x')$. It follows that $f(x) = g(x)$ for every $x \in X$.

($\Leftarrow$) Second, suppose that $f(x) = g(x)$ for all $x \in X$. Then

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

It follows that $f = g$ if and only if $f(x) = g(x)$, for all $x \in X$. □

Remark. *Let A and B be two statements. A statement can be true or false, but it cannot be both. To prove that 'A if and only if B' is true, we must prove that **both** 'If A, then B' and 'If B, then A' are true. Note that if we want to prove that 'A if and only if B' is false, it is sufficient to show that either 'If A, then B' is false or 'If B, then A' is false.*

Definition 1.8. *We say a function $f : X \to Y$ is **injective** (or **one-to-one**, or an **injection**) if $\forall\, x, x' \in X$, we have that $x \neq x'$ implies that $f(x) \neq f(x')$.*

Definition 1.9. *We say a function $f : X \to Y$ is **surjective** (or **onto**, or a **surjection**) if $\forall\, y \in Y$, $\exists\, x \in X$ such that $f(x) = y$.*

Remark. *If a function $f : X \to Y$ is surjective, then $f[X] = Y$.*

Definition 1.10. *A function is called **bijective** if it is both injective and surjective.*

Example. *Define $f : \mathbb{N} \to \mathbb{N}$ by $f(x) = x^2$, for all $x \in \mathbb{N}$. The 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 $g(x) = x^2$, for all $x \in \mathbb{R}$. The function is surjective but not injective.*

Example. *Define $h : \mathbb{R} \to \mathbb{R}$ by $h(x) = x^3$, for all $x \in \mathbb{R}$. The function is bijective.*

Remark. *Caution! Do not confuse the definition of injectivity with the definition of a function. For example, consider $f = \{(y^2, y) : \ y \in \mathbb{Z} \ \} \subseteq \mathbb{Z} \times \mathbb{Z}$. We have that for each $y \in \mathbb{Z}$, there exists a unique $x \in \mathbb{Z}$ such that $(x, y) \in f$ (namely $x = y^2$). But f is not a function since, for example, $(4, 2), (4, -2) \in f$.*

Example. *Define $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \times \mathbb{R}$ by*

$$f((m, n)) = (m + n, m - n),$$

for all $m, n \in \mathbb{R}$. We want to show that f is surjective. We begin by choosing $(u, v) \in \mathbb{R} \times \mathbb{R}$ (the co-domain).

[Scratch work: We want to find $(m, n) \in \mathbb{R} \times \mathbb{R}$ (the domain of f) such that $f(m, n) = (u, v)$. So we need to find $(m, n) \in \mathbb{R} \times \mathbb{R}$ such that $m + n = u$ and $m - n = v$. Hence, we must have $m = u - n$ and $m = v + n$. Then $u - n = v + n$, or equivalently, $u - v = 2n$, or equivalently, $\frac{u-v}{2} = n$. If we have $\frac{u-v}{2} = n$ and $m = u - n$, then $m = u - \frac{u-v}{2} = \frac{u+v}{2}$. Thus, having worked backwards to find m and n, we take these values for m and n and, with hope, can show that $f(m, n) = (u, v)$.]

We set $m = \frac{u+v}{2}$ and $n = \frac{u-v}{2}$. Then $(m, n) \in \mathbb{R} \times \mathbb{R}$ (the domain), and

$$f(m, n) = (m + n, m - n) = \left(\frac{u + v}{2} + \frac{u - v}{2}, \frac{u + v}{2} - \frac{u - v}{2} \right) = (u, v).$$

It follows that f is surjective.

We have the following proposition.

Proposition 1.11. *Suppose that $f : X \to Y$. Then*

(i) we have that f is injective if and only if $\forall\, y \in f[X]$, $\exists!\, x \in X$ such that $f(x) = y$.

(ii) we have that f is bijective if and only if $\forall\, y \in Y$, $\exists!\, x \in X$ such that $f(x) = y$.

Proof.

(i) ($\Rightarrow$) First, suppose that f is injective, and take $y \in f[X]$. Hence, $\exists\, x \in X$ such that $f(x) = y$. Now, suppose $\exists\, x' \in X$ such that $x' \neq x$. Since f is injective, we have that $f(x') \neq f(x) = y$. Hence, $\forall\, y \in f[X]$, $\exists!\, x \in X$ such that $f(x) = y$.

($\Leftarrow$) Second, suppose that $\forall\, y \in f[X]$, $\exists!\, x \in X$ such that $f(x) = y$. Take $x, x' \in X$ such that $x \neq x'$ and let $y = f(x)$. By assumption, x is the only element of X that f maps to y. Then $y \neq f(x')$, so $f(x) \neq f(x')$. Hence, we have shown that for $x, x' \in X$ with $x \neq x'$, we have that $f(x) \neq f(x')$. Therefore, f is injective.

It follows that f is injective if and only if $\forall\, y \in f[X]$, $\exists!\, x \in X$ such that $f(x) = y$.

(ii) Exercise 1.3.

□

Definition 1.12. *Let $f : X \to Y$ and $g : Y \to Z$. We define the **composition** of g and f, denoted by $g \circ f$, by*

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

Remark. *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, that is $g \circ f : X \to Z$.*

Definition 1.13. *We say that a function $f : X \to Y$ is **invertible** if there exists a function $g : Y \to X$ such that $g \circ f$ is the **identity function on** X (that is, $\forall\, x \in X$, $(g \circ f)(x) = x$) and $f \circ g$ is the **identity function on** Y (that is, $\forall\, y \in Y$, $(f \circ g)(y) = y$).*

Remark. *If g is an inverse for f, then 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 $g(x) = \frac{x-3}{2}$. Then $\forall\, x \in \mathbb{R}$, we have*

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

and

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

Hence $g \circ f$ is the identity function on $\mathbb{R}$, and $f \circ g$ is the identity function on $\mathbb{R}$. It follows that f is invertible where g is the inverse of f.

We have the following proposition.

Proposition 1.14. *Suppose that $f : X \to Y$, $g : Y \to Z$ and $h : Z \to W$. Then $h \circ (g \circ f) = (h \circ g) \circ f$.*

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

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

$\forall\, x \in X$, and

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

$\forall\, x \in X$. Thus, $h \circ (g \circ f) = (h \circ g) \circ f$. □

We have the following theorem.

Theorem 1.15. *Suppose that $f : X \to Y$ and $g : Y \to Z$. Then*

(i) if f and g are injective, we have that $g \circ f$ is injective.

(ii) if f and g are surjective, we have that $g \circ f$ is surjective.

Proof.

(i) Suppose that f, g are injective, and suppose that $x, x' \in X$ are such that $x \neq x'$. Since f is injective, we have that $f(x) \neq f(x')$. Now, set $y = f(x)$ and $y' = f(x')$. Then $y, y' \in Y$ with $y \neq y'$. Further, since g is injective, we have $g(y) \neq g(y')$. It follows that

$$(g \circ f)(x) = g(f(x)) = g(y) \neq g(y') = g(f(x')) = (g \circ f)(x').$$

To conclude, for any $x, x' \in X$ with $x \neq x'$, we have $(g \circ f)(x) \neq (g \circ f)(x')$. Hence $g \circ f$ is injective.

(ii) Exercise 1.8.

□

Remark. *The above theorem shows that if both $f : X \to Y$ and $g : Y \to Z$ are bijective, then $g \circ f : X \to Z$ is also bijective.*

Theorem 1.16. *Suppose $f : X \to Y$, and suppose that $g : Y \to X$ and $h : Y \to X$ are inverses of f. Then $g = h$, that is, if f has an inverse then its inverse is unique.*

Proof. Exercise 1.9. □

We will now use a *proof by contradiction*. For a proof by contradiction, we proceed as follows. We want to prove that a certain statement P is true. To the contrary, we assume that P is false, and we use this false statement to deduce that something (which we know to be false) is true. Hence, we conclude that it is impossible that P is false, and therefore we must have that P is true.

Theorem 1.17. *Suppose that $f : X \to Y$. Then f is invertible if and only if f is bijective.*

Proof.

(⇒) First, we will show that if f is invertible, then f is bijective. So suppose f is invertible, and let $g : Y \to X$ denote the inverse of f. Then $g \circ f$ is the identity function on X, and $f \circ g$ is the identity function on Y. To show that f is injective, suppose that $x, x' \in X$ with $x \neq x'$. Then

$$x = (g \circ f)(x) = g(f(x)) \qquad \text{and} \qquad x' = (g \circ f)(x') = g(f(x')).$$

To the contrary, suppose that $f(x) = f(x')$. Then we must have $g(f(x)) = g(f(x'))$. But this contradicts the fact that $g(f(x)) = x \neq x' = g(f(x'))$. Hence, we must have that $f(x) \neq f(x')$. So if $x, x' \in X$ with $x \neq x'$ then $f(x) \neq f(x')$. That is, f is injective.

We now continue to show that f is surjective. We begin by choosing $y \in Y$. We know that $f \circ g$ is the identity function on Y, so

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

for all $y \in Y$. Now, set $x = g(y)$. Then $x \in X$, and

$$f(x) = f(g(y)) = y,$$

for all $y \in Y$. This shows that f is surjective. Therefore we have shown that when f is invertible, then f is injective and surjective. Therefore, f is bijective.

($\Leftarrow$) Second, we will show that if f is bijective, then f is invertible. So suppose that f is bijective. We set

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

Since f is bijective, we have that $\forall\, y \in Y$, $\exists!\, x \in X$ such that $f(x) = y$. Hence, g is a function, that is $g : Y \to X$.

We will show that g is the inverse of f. For this, we first take $x \in X$ and set $y \in Y$ to be $y = f(x)$. Then by the definition of g, we have that $g(y) = x$, so $(g \circ f)(x) = g(y) = x$. Since x was arbitrary, this shows that $g \circ f$ is the identity function on X. Now, choose any $y \in Y$. Since f is bijective then $\forall\, y \in Y$, $\exists!\, x \in X$ such that $f(x) = y$. Further, since g is a function, we must have that $g(y) = x$, and hence $(f \circ g)(y) = f(x) = y$. Since y was arbitrary, this shows that $f \circ g$ is the identity function on Y. Hence if f is bijective, we have that f is invertible.

It follows that f is invertible if and only if f is bijective. □

Remark. *Suppose that $f : X \to Y$ is bijective. In the above proof, we found a way of defining $f^{-1} : Y \to X$. That is, $\forall\, y \in Y$, we can write $f^{-1}(y) = x$ where $x \in X$ such that $f(x) = y$.*

Example. *Suppose that $a, b, c, d \in \mathbb{R}$ such 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 the closed intervals $[a, b]$ and $[c, d]$. So we need to find a bijective function which first stretches or shrinks the interval $[a, b]$ to be the same length as $[c, d]$, and then shift this. The map

$$f_1(x) = x - a$$

takes the interval $[a, b]$ to $[0, b - a]$. The map

$$f_2(x) = x \cdot \frac{d - c}{b - a}$$

takes the interval $[0, b - a]$ to $[0, d - c]$, and then

$$f_3(x) = c + x$$

takes the interval to $[0, d - c]$ to $[c, d]$. Hence, we set $f = f_3 \circ f_2 \circ f_1$, that is

$$f(x) = c + \frac{(x - a)(d - c)}{(b - a)},$$

for all $x \in [a, b]$. We need to verify that f indeed maps the interval $[a, b]$ to $[c, d]$, that is $f : [a, b] \to [c, d]$. To do this, we take $x \in [a, b]$. Then $0 \leq x - a \leq b - a$. Since $a < b$ and $c < d$, we have that $b - a > 0$ 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 $f : [a, b] \to [c, d]$.

Now we want to show that f *is bijective. So we could argue that* f *is injective and surjective. However, in this example, we will argue that* f *is bijective by finding* $g : [c, d] \to [a, b]$ *such that* $g \circ f$ *is the identity map on* $[a, b]$ *and* $f \circ g$ *is the identity map on* $[c, d]$. *Using the same technique we used to construct* f, *we define*

$$g(x) = a + \frac{(x - c)(b - a)}{(d - c)},$$

for all $x \in [c, d]$. *Then for* $x \in [c, d]$, *we have* $c \leq x \leq d$ *and hence*

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

Then $g : [c, d] \to [a, b]$. *Further, we have*

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

for all $x \in [a, b]$. *Similarly,*

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

for all $x \in [c, d]$. *It follows that* $g : [c, d] \to [a, b]$ *is the inverse of* f. *Therefore,* f *is bijective.*

We have the following results.

Proposition 1.18. *Suppose that* $f : X \to Y$ *and* $g : Y \to X$ *are such that* $g \circ f$ *is the identity function on* X.

(i) Let g *be injective. Then* $f \circ g$ *is the identity map on* Y *(and hence* $g = f^{-1}$*).*

(ii) Let f *be surjective. Then* $f \circ g$ *is the identity map on* Y *(and hence* $g = f^{-1}$*).*

Proof.

(i) Exercise 2.10.

(ii) Exercise 1.10.

□

Theorem 1.19. *Suppose* $f : X \to Y$ *and* $g : Y \to Z$ *are bijective. Then* $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Proof. Exercise 1.11. □