1. Introduction

In this chapter, I explain sets and sets of numbers.

2. Notation

The expression "$A := B$" means that I am defining A to be equal to B. I occasionally use the logical symbol $\forall$ to mean "for all," and the logical symbol $\exists$ to mean "there exists."

I sometimes use the abbreviation "s.t." for "such that." For example, I might write, "let n be an integer s.t. $n^2 + 7n + 5$ is even."

If P and Q are mathematical statements, then I use the notation "$P \Rightarrow Q$" to mean that *P implies Q*.

3. Sets

Most (if not all) of mathematics can be described in terms of sets. Set theory is an interesting and important part of mathematics, and it would take a whole book to explain set theory in any depth. Fortunately, set theory is also very natural, and an intuitive understanding of set theory is sufficient for a good deal of mathematics.

Sets are basic objects of mathematics, so they cannot be defined in terms of other objects. However, there is some notation that goes along with sets that I would like to review. For example, sets are defined in terms of their elements. We write $x \in S$ to mean that x is an element of S, and $x \notin S$ means that x is not an element of S. If a set is finite we can give it just by listing its elements. For example, $S := \{1, 2, 3, 4\}$ is the set whose elements are the numbers $1, 2, 3$ and 4. So, $4 \in S$ but 5 is not an element of S.

Note that, when giving a set as a list of elements, the order of the elements doesn't matter. So $\{1, 2, 3, 4\}$ is the same set as $\{2, 4, 1, 3\}$. It is also convenient in proofs and definition to allow ourselves to list the same element twice. So $\{1, 2, 3, 1, 4\}$ is the same thing as $\{1, 2, 3, 4\}$.

We say that a set T is a subset of a set S if every element of T is an element of S. In this case, we write $T \subset S$. If S is a set and P is a property that element of S may or may not have, then $\{x \in S : P(x)\}$ denotes the set of all elements of S having the property P. For example, if $S = \{1, 2, 3, 4\}$ then $\{2, 4\} = \{x \in S : x \text{ is even}\}$.

There is one set which has no elements. This set is called the *empty set* and written $\emptyset$. Since a set is determined by its elements, the empty set is the unique set with no elements. A set S is called a *singleton* if it has exactly one element. So, for example, $\{3\}$ and $\{\{1, 2, 3, 4\}\}$ are singletons, but $\emptyset$ and $\{1, 2\}$ are not.

4. Sets of Numbers

We use the following notation for various sets of numbers.

(1) The set of natural numbers is $\mathbb{N} := \{0, 1, 2, 3, \ldots\}$.
(2) The set of integers is $\mathbb{Z} := \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$.
(3) The set of positive integers is $\mathbb{Z}_+ = \{1, 2, 3, \ldots\}$.
(4) The set of real numbers is $\mathbb{R}$.
(5) The set of rational numbers is $\mathbb{Q} = \{a/b \in \mathbb{R} : a, b \in \mathbb{Z}, b \neq 0\}$.
(6) If $a, b \in \mathbb{R}$, then

$$
\begin{array}{ll}
[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}, & (a, b) := \{x \in \mathbb{R} : a < x < b\}, \\
[a, b) := \{x \in \mathbb{R} : a \leq x < b\}, & (a, b] := \{x \in \mathbb{R} : a < x \leq b\}, \\
[a, \infty) := \{x \in \mathbb{R} : a \leq x\}, & (a, \infty) := \{x \in \mathbb{R} : a < x\}, \\
(-\infty, a] := \{x \in \mathbb{R} : x \leq a\}, & (-\infty, a) := \{x \in \mathbb{R} : x < a\}.
\end{array}
$$

(7) The set of complex number is $\mathbb{C} = \{x + iy : x, y \in \mathbb{R}\}$. Addition in $\mathbb{C}$ is defined by $(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)$, $(x_1 + iy_1)(x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$.

5. Arithmetic

Definition 5.1. Suppose a and b are in $\mathbb{Z}$. We say that *a divides b* and write $a \mid b$ if there exists $c \in \mathbb{Z}$ such that $ac = b$. If a does not divide b we write $a \nmid b$. If $a \mid b$ then a is said to be a *divisor* of b. A positive integer n is said to be *prime* if it has exactly two positive divisors. A positive integer n is *composite* if there exists integers a, b such that $1 < a \leq b < n$ and $ab = n$.

Remark 5.2. If $a \neq 0$, then $a \mid b \Leftrightarrow b/a \in \mathbb{Z}$. If $a = 0$, then $a \mid b$ implies that $b = 0$. Obviously 1 is a divisor of every integer b, since $b = (b)(1)$.

Lemma 5.3. *Suppose n is a positive integer. Then exactly one of the following hold.*

(1) *$n = 1$;*
(2) *n is prime;*
(3) *n is composite.*

Proof. If $n = 1$, then n has exactly one divisor, so n is not prime. Clearly n is not composite. If n is composite, then $n = ab$ with $1 < a \leq b < n$. So n has at least three positive divisors. Therefore n is not prime.

Suppose n is an integer greater than 1. If n is not prime, then there exists a positive divisor d of n satisfying $1 < d < n$. Setting $a = \min(d, n/d)$, $b = \max(d, n/d)$, we see that n is composite. □

Much of arithmetics comes from the following axiom.

Axiom 5.4 (Well-ordered property of the natural numbers)**.** *Suppose S is a non-empty subset of $\mathbb{N}$. Then S has a smallest element. In other words, there is an element $n \in S$ such that, for all $m \in S$, $n \leq m$.*

Theorem 5.5 (Fundamental Theorem of Arithmetic)**.** *Suppose n is an integer strictly greater than 1. Then there exists a unique positive integer r and prime numbers $p_1, \ldots, p_r$ satisfying $p_1 < p_2 < \cdots < p_r$ and*

$$n = p_1 p_2 \cdots p_r.$$

In other words, every integer $n > 1$ can be factored uniquely as a product of primes.

Proof. We first prove the existence of a prime factorization, then we prove the uniqueness (which is, in fact, the more subtle part).

Suppose, to get a contradiction, that there exists an integer strictly greater than 1 which cannot be factored as a product of primes. Then, by Axiom 5.3, there exists a smallest such integer n. If $n = ab$ for integers a, b satisfying $1 < a \leq b < n$, then, by our assumption on n, a and b can be written as products of primes. But then obviously so can n. So we cannot write $n = ab$ with $1 < a \leq b < n$. It follows that 1 and n are the only positive divisors of n. So n is itself prime. But this contradicts are assumption on n. So, it completes the proof of existence.

To prove the uniqueness of prime factorization, we assume, to get a contradiction, that there exists an integer strictly greater than 1 with two distinct prime factorizations. Again,

by Axiom 5.3, there exists a smallest such integer n. Suppose then that

$$\begin{aligned} n &= p_1 \cdots p_r \\ &= q_1 \cdots q_s \end{aligned}$$

are two distinct factorizations with $p_1 \leq p_2 \leq \cdots \leq p_r$ and $q_1 \leq q_2 \leq \cdots \leq q_s$. It is easy to see that r and s must both be strictly greater than 1. Moreover, we can assume that $p_1 \leq q_1$.

Assume first that $p_1 = q_1$. In this case, $p_2 \cdots p_r = q_1 \cdots q_s$. So, since $p_2 \cdots p_r < n$, it follows that $r = s$ and $p_i = q_i$ for $1 < i \leq r$. In other words, the two factorizations are the same, which contradicts our assumption on n.

So assume that $p_1 < q_1$, and set $m = (q_1 - p_1)q_2 \cdots q_s$. Then $m < n$. So m must have a unique prime factorization. Factor $q_1 - p_1 = r_1 \cdots r_t$ with r_i prime. Then $m = r_1 \cdots r_t q_2 \cdots q_s$ is a factorization of m into primes. Note that p_1 does not divide any of the r_i since p_1 does not divide $q_1 - p_1$. (If it did, it would have to divide q_1 which contradicts our assumption that $p_1 < q_1$). On the other hand, $m = n - p_1 q_2 \cdots q_s = p_1 p_2 \cdots p_r - p_1 q_2 \cdots q_s = p_1(p_2 \cdots p_r - q_2 \cdots q_s)$. So, by factoring, $p_2 \cdots p_r - q_2 \cdots q_s$ into primes, we get a factorization of m where p_1 appears. So we have two distinct prime factorizations of m. This contradicts our assumption that n was the smallest positive integer with two distinct prime factorizations. So, it completes the proof.

□

6. Interesection and Union

Suppose A and B are sets. The *union* of A and B is the set $A \cup B$ with the property that $x \in A \cup B$ if and only if $x \in A$ or $x \in B$. The *intersection* of A and B is the set $A \cap B$ with the property that $x \in A \cap B$ if and only if $x \in A$ and $x \in B$.

More generally, suppose that I is a set and, for each $i \in I$, A_i is a set. Then the *union* of the A_i is the set $\cup_{i \in I} A_i$ with the property that $x \in \cup_{i \in I} A_i$ if and only if $x \in A_i$ for some $i \in I$. The *intersection* of the A_i is the set $\cap_{i \in I} A_i$ with the property that $x \in \cap_{i \in I} A_i$ if and only if, for all $i \in I$, $x \in A_i$.

Example 6.1. Let $I = (0, 1)$ and for each $i \in I$, set $F_i = [0, i]$. Then $\cup_{i \in I} F_i = [0, 1)$, $\cap_{i \in I} F_i = \{0\}$.

7. Power Sets

Suppose X is a set. The *power set* of X is the set $\mathcal{P}(X)$ of all subsets of X. Explicitly, we have $x \in \mathcal{P}(X)$ if and only if $x \subset X$.

If X is a finite set with n elements, then $\mathcal{P}(X)$ has 2^n elements.

Example 7.1. For each $n \in \mathbb{N}$, write $S_n := \{x \in \mathbb{N} : x < n\}$. Then

(1) $S_0 = \emptyset$, and $\mathcal{P}(S_0) = \{\emptyset\}$,
(2) $S_1 = \{1\}$, and $\mathcal{P}(S_1) = \{\emptyset, \{1\}\}$,
(3) $S_2 = \{1, 2\}$, and $\mathcal{P}(S_2) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}$.

Remark 7.2. In some situtation power sets can be a little confusing. I have two pieces of advice regarding this. Firstly, in most mathematical situations, what the elements of a set actually are as sets, doesn't really matter. So, for example, the most importatn thing about $\mathcal{S}$ is usually that it has 4 elements, not that its elements are $\emptyset, \{1\}, \{2\}$ and $\{1, 2\}$. On the other hand, there are of course some situations in which what the elements are really does matter. In these rare situations, my advice is to think carefully about what each element really is. There is really nothing that complicated about elementary set theory: it just requires careful thought.

8. Ordered Pairs and Cartesian Products

Suppose X and Y are sets. Then the *Cartsesian product* of X and Y is the set $X \times Y$ of all ordered pairs (x, y) where $x \in X$ and $y \in Y$. More generally, if $X_1, \ldots, X_n$ are sets, then $X_1 \times X_2 \cdots \times X_n$ is the set of ordered n-tuples $(x_1, x_2, \ldots, x_n)$ with $x_i \in X_i$ for all i. If n is a positive integer and X is a set then X^n is the set of all ordered n-tuples $(x_1, x_2, \ldots, x_n)$ with $x_i \in X$ for all i.

Example 8.1. The set of ordered pairs of real numbers is $\mathbb{R}^2$, the set of ordered triples is $\mathbb{R}^3$. The set $[0, 1] \times [0, 2]$ is a rectangle in $\mathbb{R}^2$ with width 1 and height 2.

Remark 8.2. In mathematics, everything is supposed to be a set. So we should be able to say what an ordered pair is. Here's the standard definition: Suppose x and y are sets. Then

$$(x, y) := \{\{x\}, \{x, y\}\}$$

To see that the definition makes sense, we have to check that, for sets x, y, z, w, $(x, y) = (z, w)$ only if $x = z$ and $y = w$. To see this, first note that, if $x = y$ then (x, y) has one element. Otherwise (x, y) has two elements. So, if $x = y$ then we must have $z = w$ and $\{x\} = \{z\}$. This implies that $x = z$ since a set is determined by its elements. So $(x, y) = (x, x) = (z, z) = (z, w)$. On the other hand, if $x \neq y$, then $\{x\}$ is the only element of (x, y) with one element. Similarly, $\{z\}$ is the only element of (z, w) with one element. So $\{x\} = \{z\}$. Therefore, $x = z$. Therefore, $\{x, y\} = \{z, w\} = \{x, w\}$. This implies that $z = w$.

9. Relations and Functions

Definition 9.1. Suppose X and Y are sets. A *relation* from X to Y is a subset $R \subset X \times Y$. We write $\mathrm{Rel}(X, Y)$ for the set of all relations from X to Y. If $R \in \mathrm{Rel}(X, Y)$ and $S \subset X$, then we set

$$R[S] := \{y \in Y : \text{there exists } x \in S : (x, y) \in R\}.$$

We write R^{op} for the relation from Y to X given by

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

There are two types of relations which are especially useful: functions and equivalence relation.

Definition 9.2. Suppose X and Y are sets. A relation $f \in \mathrm{Rel}(X, Y)$ is a function if, for each $x \in X$, there is a unique $y \in Y$ such that $(x, y) \in f$. In this case, we write $f(x) := y$. We write $\mathrm{F}(X, Y)$ for the set of all functions from X to Y, and we write $f : X \to Y$ to indicate that f is a function from X to Y.

Remark 9.3. With our defintions, if $f : X \to Y$ is a function and $x \in X$, then $f[\{x\}] = \{f(x)\}$. Note that, if $f : X \to Y$ is a function then $f = \cup_{x \in X}\{(x, f(x))\}$. In particular, if f and g are functions from X to Y and $f(x) = g(x)$ for every $x \in X$, then $f = g$.

Remark 9.4. We usually write down a function $f : X \to Y$ by specifying a rule for computing an element $f(x) \in Y$ given an element $x \in X$. Sometimes this rule is given in the form $x \mapsto f(x)$. For example, we can define a function $f : \mathbb{R} \to \mathbb{R}$ by saying that f is the function given by $x \mapsto x^3 + 3x + 5$.

Definition 9.5. Suppose $f : X \to Y$ and $g : Y \to Z$ are functions. Then the *composition* of g with f is the function $g \circ f : X \to Z$ given by $(g \circ f)(x) = g(f(x))$ for $x \in X$.

Definition 9.6. Suppose Y is a set. The *identity function* on Y is the function $\mathrm{id}_Y : Y \to Y$ given by $y \mapsto y$.

Proposition 9.7. *Suppose X, Y, Z and U are sets. Let $f : X \to Y$, $g : Y \to Z$ and $h : Z \to U$ be functions. Then*

(1) $\mathrm{id}_Z \circ g = g = g \circ \mathrm{id}_Y$;
(2) $h \circ (g \circ f) = (h \circ g) \circ f$.

Proof. (1): Suppose $y \in Y$. Then $(\mathrm{id}_Z \circ g)(y) = \mathrm{id}_Z(g(y)) = g(y) = g(\mathrm{id}_Y(y)) = (g \circ \mathrm{id}_Y)(y)$.

(2): Suppose $x \in X$. Then $(h \circ (g \circ f))(x) = h((g \circ f)(x)) = h(g(f(x))) = (h \circ g)(f(x)) = ((h \circ g) \circ f)(x)$. □

Definition 9.8. A function $f : X \to Y$ is said to be

(1) *one-one* if, for every $x_1, x_2 \in X$, $f(x_1) = f(x_2)$ only if $x_1 = x_2$;
(2) *onto* if, for every $y \in Y$, there exists an $x \in X$ such that $f(x) = y$.

We say that f is an *isomorphism of sets* if it is one-one and onto.

Remark 9.9. Suppose $f : X \to Y$ is a function. If f is one-one we also say that f is *injective*. If f is onto we also say that f is *surjective*. If $S \subset X$, then the subset $f[S] \subset Y$ is called the *image* of S under S. The *image* of f is $f[X]$. Note that f is onto if and only if $f[X] = Y$. It is common to abuse notation and write $f(S)$ for $f[S]$ and I will do that as long as it doesn't lead to confusion. Similary, if $T \subset Y$, then the subset $f^{\mathrm{op}}[T]$ of X is called the *inverse image of* T. It is common to abuse notation and write $f^{-1}(T)$ instead of $f^{\mathrm{op}}[T]$. The relation f^{op} is an element of $\mathrm{Rel}(f(X), X)$. Explicitly, for $T \subset Y$, $f^{-1}(T) = \{x \in X : f(x) \in T\}$.

Proposition 9.10. *Suppose $f : X \to Y$ and $g : Y \to Z$ are functions. Prove the following:*

(1) *If f and g are one-one, then so is $g \circ f$.*
(2) *If f and g are onto, so is $g \circ f$.*
(3) *If $g \circ f$ is one-one, then so is f.*
(4) *If $g \circ f$ is onto, then so is g.*

Proof. (1): Suppose $x_1, x_2 \in X$. Then $(g \circ f)(x_1) = (g \circ f)(x_2) \Leftrightarrow g(f(x_1)) = g(f(x_2))$. Since g is one-one, this implies $f(x_1) = f(x_2)$. Since f is one-one, this implies $x_1 = x_2$.

(2): Suppose $z \in Z$. Since g is onto, there exists $y \in Y$ such that $g(y) = z$. Since f is onto, there exists $x \in X$ such that $f(x) = y$. So $z = g(f(x)) = (g \circ f)(x)$.

(3): Suppose $f(x_1) = f(x_2)$ for $x_1 \neq x_2$ in X. Then $g(f(x_1)) = g(f(x_2))$. So if f is not one-one, neither is $g \circ f$.

(4): Suppose $g \circ f$ is onto. Pick $z \in Z$. Then there exists $x \in X$ such that $g(f(x)) = z$. Setting $y = f(x)$, we see that $g(y) = z$. So g is onto as well. □

Proposition 9.11. *Suppose $f : X \to Y$ is a function. Then the following are equivalent.*

(1) *The function f is a an isomorphism of sets from X to Y;*
(2) *The relation $h := f^{\mathrm{op}}$ in $\mathrm{Rel}(Y, X)$ is in $\mathrm{F}(Y, X)$. Moreover, $f \circ h = \mathrm{id}_Y$ and $h \circ f = \mathrm{id}_X$.*
(3) *There exists a function $g \in \mathrm{F}(Y, X)$ such that $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$.*

Moreover if (1)-(3) hold, $f^{\mathrm{op}} = g$.

Proof. (1)⇒ (2): Suppose $f : X \to Y$ is a an isomorphism from X to Y. Then, for every $y \in Y$, there is a unique $x \in X$ such that $y = f(x)$. In other words, for each $y \in Y$, there is a unique $x \in X$ such that $(x, y) \in f$. Therefore, for each $y \in Y$, there is a uniqe $x \in X$ such that $(y, x) \in f^{\mathrm{op}}$. Therefore, by definition, $f^{\mathrm{op}} \in \mathrm{F}(Y, X)$.

Suppose $x \in X$ and $y = f(x)$. By definition, $(y, x) \in h$. So $x = h(y)$. In other words, $h(f(x)) = x$ So $h \circ f = \mathrm{id}_X$. Therefore $f^{\mathrm{op}} \circ f = \mathrm{id}_X$. Since $f = h^{\mathrm{op}}$ is a fucntion from X to Y, it follows that $f \circ h = h^{\mathrm{op}} \circ h = \mathrm{id}_Y$.

(2)⇒ (3): Obvious.

(3)⇒ (1): Suppose $x_1, x_2 \in X$. Then $f(x_1) = f(x_2) \Rightarrow x_1 = g(f(x_1)) = g(f(x_2)) = x_2$. So f is one-one. Suppose $y \in Y$. Then $y = f(g(y))$. So f is onto.

Finally, suppose $g \in \mathrm{F}(Y, X)$ is a function such that $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$. Then $g = g \circ \mathrm{id}_Y = g \circ (f \circ h) = (g \circ f) \circ h = h$. □

Definition 9.12. Suppose $f : Y \to Z$ is a function and X and W are sets. I write $f_* : \mathrm{F}(X, Y) \to \mathrm{F}(X, Z)$ for the function which sends $g \in \mathrm{F}(X, Y)$ to $f \circ g \in \mathrm{F}(X, Z)$. I write $f^* : \mathrm{F}(Z, W) \to \mathrm{F}(Y, W)$ for the funciton which sends $h \in \mathrm{F}(Z, W)$ to $h \circ f \in \mathrm{F}(Y, W)$.

Proposition 9.13. *Suppose $f : Y \to Z$ is a function. Then*

(1) *The function f is injective if and only if, for every set X, the map $f_* : \mathrm{F}(X, Y) \to \mathrm{F}(X, Z)$ is injective.*
(2) *the function f is surjective if and only if, for every set W, the map $f^* : \mathrm{F}(Z, W) \to \mathrm{F}(Y, W)$ is injecitive.*

Proof. (1): ⇒: Suppose $f : Y \to Z$ is injective, and suppose $g_1, g_2 : X \to Y$ are two functions such that $f \circ g_1 = f \circ g_2$. Then, if $x \in X$, we have $f(g_1(x)) = f(g_2(x))$. So $g_1(x) = g_2(x)$. Since this holds for any $x \in X$, it follows that $g_1 = g_2$.

⇐: Set $X = \{0\}$. Then the map $\mathrm{F}(X, Y) \to Y$ given by $g \mapsto g(0)$ is an isomorphism of sets. It follows easily that, if $f_* : \mathrm{F}(X, Y) \to \mathrm{F}(X, Z)$ is injective, so is $f : Y \to Z$.

(2): ⇒: Suppose $f : Y \to Z$ is surjective, and suppose $h_1, h_2 : Z \to W$ are two functions such that $h_1 \circ f = h_2 \circ f$. Pick $z \in Z$. Since f is surjective, we can find a $y \in Y$ such that $f(y) = z$. Then $h_1(z) = h_1(f(y)) = h_2(f(y)) = h_2(z)$. Since this holds for any $z \in Z$, it follows that $h_1 = h_2$.

⇐: Suppose the map $f : Y \to Z$ is not surjective. Pick $\zeta \in Z \setminus f(Y)$. Set $W = \{0, 1\}$. Define a function $h_0 : Z \to W$ by setting $h_0(z) = 0$ for all $z \in Z$, define a function $h_1 : Z \to W$ by

$$h_1(z) = \begin{cases} 0, & z \neq \zeta; \\ 1, & z = \zeta. \end{cases}$$

Then $h_0 \circ f = h_1 \circ f$ because $h_0(f(y)) = h_1(f(y)) = 0$ for all $y \in Y$. But $h_0 \neq h_1$. So $f^* : \mathrm{F}(Z, W) \to \mathrm{F}(Y, W)$ is not injective. □

10. Equivalence Relations, Partitions, Quotients and Kernels

Definition 10.1. Let X be a set. A *relation* on X is a relation $R \subset X \times X$. If R is a relation on X, we write $x \sim_R y$ to mean that $(x, y) \in R$. When the relation R is obvious, we simply write $x \sim y$ for $x \sim_R y$.

Definition 10.2. A relation R on X is called an *equivalence relation* if it satisfies the following three axioms:

(R) For all $x \in X$, $(x, x) \in R$.
(S) If $(x, y) \in R$ then (y, x) in R.
(T) If (x, y) and (y, z) are in R, then (x, z) is in R.

Remark 10.3. In Definition 10.2, R stands for "reflexive," S stands for "symmetric" and T stands for "transitive."

Example 10.4. For any set X, let $\Delta_X = \{(x, x) \in X \times X\}$. This set is called the *diagonal* in $X \times X$. It is obviously an equivalence relation and we have $x \sim y \Leftrightarrow x = y$.

Definition 10.5. Suppose $f : X \to Y$ is a function. The *difference kernel* of f is the set $K_f = \{(x_1, x_2) \in X \times X : f(x_1) = f(x_2)\}$. We also occasionally write $X \times_Y X$.

Proposition 10.6. *Suppose $f : X \to Y$ is a function. Then K_f is an equivalence relation on X.*

Proof. Write $x_1 \sim x_2$ to mean that $(x_1, x_2) \in K_f$. Clearly, $x_1 = x_2 \Rightarrow f(x_1) = f(x_2) \Rightarrow x_1 \sim x_2$. So the reflexive property holds. Suppose $x_1 \sim x_2$. Then $f(x_1) = f(x_2)$. So $x_2 \sim x_1$. So the symmetric property holds. Finally, if $x_1 \sim x_2$ and $x_2 \sim x_3$, then $f(x_1) = f(x_2) = f(x_3)$. So $x_1 \sim x_3$, and this proves transitivity. □

Lemma 10.7. *Suppose $f : X \to Y$ and $g : Y \to Z$ are functions. Set $h = g \circ f$. Then $K_f \subset K_h$.*

Proof. Suppose $(x_1, x_2) \in K_f$. Then, by definition, $f(x_1) = f(x_2)$. Therefore $h(x_1) = g(f(x_1)) = g(f(x_2)) = h(x_2)$. So $(x_1, x_2) \in K_h$. □

Theorem 10.8. *Suppose $f : X \to Y$ is a surjective function and Z is a set. Suppose $g \in \mathrm{F}(X, Z)$. Then $g \in f^*(\mathrm{F}(Y, Z))$ if and only if $g(x_1) = g(x_2)$ for all $x_1, x_2 \in X$ such that $f(x_1) = f(x_2)$.*

Proof. $\Rightarrow$: Suppose $g \in f^*(\mathrm{F}(Y, Z))$. Then $g = h \circ f$ for some function $h : Y \to Z$. Then $f(x_1) = f(x_2) \Rightarrow g(x_1) = h(f(x_1)) = h(f(x_2)) = g(x_2)$.

$\Leftarrow$: Suppose $g(x_1) = g(x_2)$ for all $x_1, x_2 \in X$ such that $f(x_1) = f(x_2)$. Let $h = \{(y, z) \in Y \times Z : \exists x \in X\, y = f(x), z = g(x)\}$. Since f is surjective, for every $y \in Y$, there is an $x_1 \in X$ such that $y = f(x_1)$. So $(y, g(x_1)) \in h$. Suppose $(y, z) \in h$. Then, by definition, $z = g(x_2)$ for some $x_2 \in X$ such that $f(x_2) = y$. Then $f(x_1) = f(x_2)$, so $z = g(x_2) = g(x_1)$. It follows that $h \in \mathrm{F}(Y, Z)$. Moreover, for $x \in X$, we have $(h \circ f)(x) = h(f(x)) = g(x)$. So $f^*(h) = h \circ f = g$. □

Corollary 10.9. *Suppose $f : X \to Y$ is a surjective function and $h : X \to Z$ is function such that $K_f \subset K_h$. Then there exists a unique function $g : Y \to Z$ such that $h = g \circ f$. Moreover,*

(1) *If $K_f = K_h$, then g is injective.*
(2) *If h is surjective, then so is g.*
(3) *If $K_f = K_g$ and h is a surjective, then g is an isomorphism of sets.*

Proof. Since $K_f \subset K_h$, $h(x_1) = h(x_2)$ for every $(x_1, x_2) \in K_f$. By Theorem 10.8, $h \in f^*(\mathrm{F}(Y, Z))$, so $h = g \circ f$ for some $g : Y \to Z$. Since f is surjective, g is unique by Proposition 9.13.

(1): Suppose $g(y_1) = g(y_2)$ for $y_1, y_2 \in Y$. Since f is surjective, we can find $x_1, x_2 \in X$ such that $y_i = f(x_i)$ for $i = 1, 2$. Then $h(x_1) = h(x_2)$, so $(x_1, x_2) \in K_h$. Therefore, $(x_1, x_2) \in K_f$. So $y_1 = f(x_1) = f(x_2) = y_2$.

(2): This follows from Proposition 9.10 (4).

(3): This follows from (1) and (2). □

Definition 10.10. Supppose X is a set. A *partition* of X is a set P of non-empty subset of X such that

(1) $X = \cup_{S \in P} S$,
(2) For $S, T \in P$, $S \neq T \Rightarrow S \cap T = \emptyset$.

Given a set X with a partition P, we define a map $\pi_P : X \to P$ sending $x \in X$ to the unique $S \in P$ such that $x \in S$.

Remark 10.11. By Definition 10.10 2, the map $\pi_P : X \to P$ is surjective.

Example 10.12. Suppose $X = \{0, 1, 2\}$. The set $P = \{\{0, 1\}, \{2\}\}$ is a partition of X. The map $\pi_P : X \to P$ is given by $\pi_P(0) = \pi_P(1) = \{0, 1\}$ and $\pi_P(2) = \{2\}$.

Definition 10.13. Suppose X is a set, R is an equivalence relation on X and $x \in X$. The *equivalence class* of x is the subset $[x]$ of X given by

$$[x] = \{y \in X : (x, y) \in R\}.$$

We write X/R for the set of equivalence relations of X.

Proposition 10.14. *Suppose X is a set and R is an equivalence relation on X. Then*

(1) *For $x, y \in X$, $x \in [y] \Leftrightarrow [y] = [x]$.*
(2) *X/R is a partition of X.*
(3) *For $x, y \in X$, we have $x \sim_R y \Leftrightarrow [x] = [y]$.*
(4) *The map $\pi_{X/R} : X \to X/R$ sends $x \in X$ to $[x]$. The difference kernel of $\pi_{X/R}$ is R.*
(5) *Suppose $f : X \to Y$ is any function such that $R \subset K_f$, then there is a unique $g : X/R \to Y$ such that $f = g \circ \pi_{X/R}$.*

Proof. (1): Since $x \in [x]$, $[y] = [x] \Rightarrow x \in [y]$. To prove the converse, we first prove that $x \in [y] \Rightarrow [y] \subset [x]$. To see this, suppose $x, z \in [y]$. Then $y \sim z$ and $y \sim x$. So $z \sim y$. Therefore $z \sim x$ and, thus, $x \sim z$. So $z \in [x]$. This shows that $[y] \subset [x]$.

Now note that $y \in [y] \subset [x]$. Therefore $[x] \subset [y]$. So $[x] = [y]$.

(2): Since $x \in [x]$, $\cup_{[x] \in X/R}[x] = X$. This proves that Definition 10.10.(1) holds. Suppose $z \in [x] \cap [y]$, for $x, y, z \in X$. Then, by (1), $[x] = [z] = [y]$. This proves that Definition 10.10.(2) holds.

(3): We have $x \sim_R y \Leftrightarrow y \in [x] \Leftrightarrow [y] = [x]$.

(4): This is just a restatment of (3).

(5): This follows from (4) and from Corollary 10.9. □

Remark 10.15. If X is a set and R is an equivalence relation, we abuse notation and write π_R or just π for the map $\pi_{X/R} : X \to X/R$.

11. Exercises

Exercise 11.1. Suppose $f : X \to Y$ is a function. Show that the relation f^{-1} is in $\mathrm{Rel}(f[X], X)$, and that f is one-one if and only if $f^{-1} \in \mathrm{F}(f[X], X)$.

Exercise 11.2. Suppose X, Y, Z and W are sets, and suppose $P \in \mathrm{Rel}(X, Y)$, $Q \in \mathrm{Rel}(Y, Z)$ and $R \in \mathrm{Rel}(Z, W)$ are relations. Define a relation $Q \circ R \in \mathrm{Rel}(X, Z)$ by setting

$$Q \circ R = \{(x, z) \in X \times Z : \text{ there exists } y \in Y \text{ such that } (x, y) \in R, (y, z) \in Q\}.$$

(1) Show that if Q and R are functions, then the composition defined above is the same as that defined in Definition 9.5.
(2) Show that $(P \circ Q) \circ R = P \circ (Q \circ R)$.