The Category of Sets

Hans Halvorson

October 16, 2016

[[Note to students: this is a first draft. Please report typos. A revised version will be posted within the next couple of weeks.]]

1 Introduction

The aim of **metatheory** is to theorize about theories. For simplicity, let's use M to denote this hypothetical theory about theories. Thus, M is not the *object* of our study, it is the *tool* we will use to study other other theories. And yet, it might be helpful to give you a sort of user's manual for M. That's the aim of this chapter.

Let's think about what we hope to do with M. We want to be able to talk about theories, which are supposed to be "collections of things," or better, "structured collections of things." In the first instance, we will think of theories as structured collections of sentences.[1] What's more, sentences themselves are structured collections of symbols. Fortunately, we won't need to press the inquiry further into the question of the nature of symbols. It will suffice to assume that there are enough symbols, and that there is some primitive notion of identity of symbols. For example, I assume that you understand that "p" is the same symbol as "p", and is different from "q".

Fortunately, much of the theory we need was worked out in previous generations. At the beginning of the 20th century, much effort was spent on formulating a theory of abstract collections (i.e. sets) that would be adequate to serve as a foundation for all of mathematics. Amazingly, the theory of sets can be formalized in predicate logic with only one symbol of non-logical vocabulary, a binary relation symbol "$\in$". In the resulting first-order theory — usually called Zermelo-Frankel set theory — the quantifiers can be thought of as ranging over sets, and the relation symbol $\in$ can be used to define further notions such as subset, Cartesian products of sets, functions from one set to another, etc..

[1] I don't mean to be begging the question here about what a theory is. We could just as well think of a theory as a structured collection of models. And just as sentences can be broken down into smaller components until we reach undefined primitives, so models can be broken down into smaller components until we reach undefined primitives. In both cases, metatheory bottoms out in undefined primitives. We can call these primitives "symbols," or "sets," or anything else we want. But the name we choose doesn't affect the inferences we're permitted to draw.

Set theory can be presented informally (sometimes called "naive set theory"), or formally ("axiomatic set theory"). In both cases, the relation $\in$ is primitive. However, we're going to approach things from a different angle. We're not concerned as much with what sets *are*, but with what we can *do* with them. Thus, I'll present a version of ETCS, the elementary theory of the category of sets.[2] Here "elementary theory" indicates that this theory can be formalized in elementary (i.e. first-order) logic. The phrase "category of sets" indicates that this theory treats the collection of sets as a structured object — a category consisting of sets and functions between them.

Axiom 1: Sets is a category

Sets is a **category**, i.e. it consists of two kinds of things: objects, which we call **sets**, and arrows, which we call **functions**. To say that **Sets** is a category means that:

1. Every function f has a domain set $d_0 f$ and a codomain set $d_1 f$. We write $f : X \to Y$ to indicate that $X = d_0 f$ and $Y = d_1 f$.
2. Compatible functions can be composed. For example, if $f : X \to Y$ and $g : Y \to Z$ are functions, then $g \circ f : X \to Z$ is a function. (We frequently abbreviate $g \circ f$ as gf.)
3. Composition of functions is associative:

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

when all these compositions are defined.

4. For each set X, there is a function $1_X : X \to X$ that acts as a left and right identity relative to composition.

Discussion. If our goal was to formalize ETCS rigorously in first-order logic, we might use two-sorted logic, with one sort for sets, and one sort for functions. The primitive vocabulary of this theory would include symbols $\circ, d_0, d_1, 1$, but it would *not* include the symbol $\in$. In other words, containment is *not* a primitive notion of ETCS.

Set theory makes frequent use of bracket notation, such as:

$$\{n \in N \mid n > 17\}.$$

These symbols should be read as, "the set of n in N such that $n > 17$." Similarly, $\{x, y\}$ designates a set consisting of elements x and y. But so far, we have no

[2] All credit to William Lawvere for introducing this approach to set theory. For a gentle introduction, see his *Sets for Mathematics*.

rules for reasoning about such sets. In the following sections, we will gradually add axioms until it becomes clear which rules of inference are permitted vis-a-vis sets.

Suppose for a moment that we understand the bracket notation, and suppose that X and Y are sets. Then given an element $x \in X$, and an element $y \in Y$, we can take the set $\{x, \{x, y\}\}$ as an "ordered pair" consisting of x and y. The pair is ordered because x and y play asymmetric roles: the element x occurs by itself, as well as with the element y. If we could then gather together these ordered pairs into a single set, we would designate it by $X \times Y$, which we call the **Cartesian product** of X and Y. The Cartesian product construction should be familiar from high school mathematics. For example, the plane (with x and y coordinates) is the Cartesian product of two copies of the real number line.

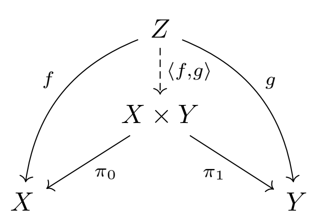
In typical presentations of set theory, the existence of product sets is derived from other axioms. Here we will proceed in the opposite direction: we will take the notion of a product set as primitive.

Axiom 2: Cartesian products

For any two sets X and Y, there is a set $X \times Y$, and functions $\pi_0 : X \times Y \to X$ and $\pi_1 : X \times Y \to Y$, such that: for any other set Z and functions $f : Z \to X$ and $g : Z \to Y$, there is a unique function $\langle f, g\rangle : Z \to X \times Y$ such that $\pi_0\langle f, g\rangle = f$ and $\pi_1\langle f, g\rangle = g$.

Here the angle brackets $\langle f, g\rangle$ are not intended to indicate anything about the internal structure of the denoted function. This notation is chosen merely to indicate that $\langle f, g\rangle$ is uniquely determined by f and g.

The defining conditions of a product set can be visualized by means of an arrow diagram.



Here each node represents a set, and arrows between nodes represent functions. The dashed arrow is meant to indicate that the axiom asserts the existence of such an arrow (dependent on the existence of the other arrows in the diagram).

Discussion. There is a close analogy between the defining conditions of a Cartesian product and the introduction and elimination rules for conjunction. If $\phi \wedge \psi$ is a conjunction, then there are arrows (i.e. derivations) $\phi \wedge \psi \to \phi$ and $\phi \wedge \psi \to \psi$. That's the $\wedge$ elimination rule.[3] Moreover, for any sentence θ, if there are deriva-

[3]Here I'm intentionally being ambiguous between the relation $\vdash$ and the connective $\to$.

tions $\theta \to \phi$ and $\theta \to \psi$, then there is a unique derivation $\theta \to \phi \wedge \psi$. That's the $\wedge$ introduction rule.

Definition. Let γ and γ' be paths of arrows in a diagram that begin and end at the same node. We say that γ and γ' **commute** just in case the composition of the functions along γ is equal to the composition of the functions along γ'. We say that the diagram as a whole **commutes** just in case any two paths between nodes are equal. Thus, for example, the product diagram above commutes.

The functions $\pi_0 : X \times Y \to X$ and $\pi_1 : X \times Y \to Y$ are typically called **projections** of the product. What features do these projections have? Before we say more on that score, let's pause to talk about features of functions.

You will recall from secondary school that functions can be one-to-one, or onto, or continuous, etc.. For bare sets, there is no notion of continuity of functions, per se. And, with only the first two axioms in place, we do not yet have the means to define what it means for a function to be one-to-one or onto. Indeed, recall that a function $f : X \to Y$ is typically said to be one-to-one just in case $f(x) = f(y)$ implies $x = y$ for any two "points" x and y of X. But we don't yet have a notion of points!

Nonetheless, there are point-free surrogates for the notions of being one-to-one and onto.

Definition. A function $f : X \to Y$ is said to be a **monomorphism** just in case for any two functions $g, h : Z \rightrightarrows X$, if $fg = fh$ then $g = h$.

Definition. A function $f : X \to Y$ is said to be a **epimorphism** just in case for any two functions $g, h : Y \to Z$, if $gf = hf$ then $g = h$.

We will frequently say, "…is monic" as shorthand for "…is a monomorphism," and "…is epi" for "…is an epimorphism."

Definition. A function $f : X \to Y$ is said to be an **isomorphism** just in case there is a function $g : Y \to X$ such that $gf = 1_X$ and $fg = 1_Y$. If there is an isomorphism $f : X \to Y$, we say that X and Y are **isomorphic**, and we write $X \cong Y$.

Exercise 1.1. Show the following:

1. If gf is monic, then f is monic.
2. If fg is epi, then f is an epi.
3. If f and g are monic, then gf is monic.
4. If f and g are epi, then gf is epi.
5. If f is an isomorphism, then f is epi and monic.

The analogy is more clear if we use the symbol $\to$ the latter.

Proposition 1.2. *Suppose that both* (W, p_0, p_1) *and* (Z, q_0, q_1) *are Cartesian products of* X *and* Y*. Then there is an isomorphism* $f : W \to Z$ *such that* $q_0 f = p_0$ *and* $q_1 f = p_1$*.*

Proof. Since (Z, q_0, q_1) is a Cartesian product, there is a unique function $\langle p_0, p_1\rangle : W \to Z$ such that $q_0\langle p_0, p_1\rangle = p_0$ and $q_1\langle p_0, p_1\rangle = p_1$. There is a similar function $\langle q_0, q_1\rangle : Z \to W$. We claim that these functions are inverse to each other. Indeed,

$$q_0 \circ \langle p_0, p_1\rangle \circ \langle q_0, q_1\rangle \;=\; p_0\langle q_0, q_1\rangle \;=\; q_0,$$

and similarly, $q_1 \circ \langle p_0, p_1\rangle \circ \langle q_0, q_1\rangle = q_1$. Thus, by the uniqueness clause in the definition of Cartesian products,

$$\langle p_0, p_1\rangle \circ \langle q_0, q_1\rangle \;=\; 1_Z.$$

A similar argument shows that

$$\langle q_0, q_1\rangle \circ \langle p_0, p_1\rangle \;=\; 1_W.$$

Therefore, $f = \langle p_0, p_1\rangle$ is the requisite isomorphism. □

Definition. If X is a set, we let $\delta : X \to X \times X$ denote the unique arrow $\langle 1_X, 1_X\rangle$ given by the definition of $X \times X$. We call δ the **diagonal** of X, or the **equality relation** on X. Note that δ is monic, since $\pi_0\delta = 1_X$ is monic.

Definition. Suppose that $f : W \to Y$ and $g : X \to Z$ are functions. Consider the following diagram:

$$\begin{array}{ccccc} W & \xleftarrow[q_0]{} & W \times X & \xrightarrow[q_1]{} & X \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f \times g} & & \downarrow{\scriptstyle g} \\ Y & \xleftarrow[\pi_0]{} & Y \times Z & \xrightarrow[\pi_1]{} & Z \end{array}$$

We let $f \times g = \langle fq_0, gq_1\rangle$ be the unique function from $W \times X$ to $Y \times Z$ such that

$$\pi_0(f \times g) = fq_0, \qquad \pi_1(f \times g) = gq_1.$$

Proposition 1.3. *Suppose that* $f : A \to B$ *and* $g : B \to C$ *are functions. Then* $1_X \times (g \circ f) = (1_X \times g) \circ (1_X \times f)$*.*

Proof. The following diagram commutes:

$$\begin{array}{ccccc} X & \longleftarrow & X \times A & \longrightarrow & A \\ {\scriptstyle 1_X}\downarrow & & \downarrow{\scriptstyle 1_X \times f} & & \downarrow{\scriptstyle f} \\ X & \longleftarrow & X \times B & \longrightarrow & B \\ {\scriptstyle 1_X}\downarrow & & \downarrow{\scriptstyle 1_X \times g} & & \downarrow{\scriptstyle g} \\ X & \longleftarrow & X \times C & \longrightarrow & C \end{array}$$

Thus, $(1_X \times g) \circ (1_X \times f)$ has the defining properties of $1_X \times (g \circ f)$. □

Exercise 1.4. Show that $1_X \times 1_Y = 1_{X \times Y}$.

Definition. Let X be a fixed set. Then X induces two mappings, as follows:

1. A mapping $Y \mapsto X \times Y$ of sets to sets.
2. A mapping $f \mapsto 1_X \times f$ of functions to functions. That is, if $f : Y \to Z$ is a function, then $1_X \times f : X \times Y \to X \times Z$ is a function.

By the previous results, the second mapping is compatible with the composition structure on arrows. In this case, we call the pair of mappings a **functor** from **Sets** to **Sets**.

Exercise 1.5. Suppose that $f : X \to Y$ is a function. Show that the following diagram commutes.

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle \delta_X}\downarrow & & \downarrow{\scriptstyle \delta_Y} \\ X \times X & \xrightarrow[f \times f]{} & Y \times Y \end{array}$$

We will now recover the idea that sets consist of points by requiring the existence of a single-point set 1, which plays the privileged role of determining identity of functions.

Axiom 3: Terminal Object

There is a set 1 with the following two features:

1. For any set X, there is a unique function
$$X \xrightarrow{\beta_X} 1$$
In this case, we say that 1 is a **terminal object** for **Sets**.
2. For any sets X and Y, and functions $f, g : X \rightrightarrows Y$, if $f \circ x = g \circ x$ for all functions $x : 1 \to X$, then $f = g$. In this case, we say that 1 is a **separator** for **Sets**.

Exercise 1.6. Show that if X and Y are terminal objects in a category, then $X \cong Y$.

Definition. We write $x \in X$ to indicate that $x : 1 \to X$ is a function. In this case, we say that x is an **element** of X, and we write $x \in X$. If $f : X \to Y$ is a function, we sometimes write $f(x)$ for $f \circ x$. With this notation, the statement that 1 is a separator says: $f = g$ is and only if $f(x) = g(x)$, for all $x \in X$.

Discussion. In ZF set theory, equality between functions is completely determined by equality between sets. Indeed, in ZF, functions $f, g : X \rightrightarrows Y$ are defined to be certain subsets of $X \times Y$; and subsets of $X \times Y$ are defined to be equal just in case they contain the same elements. In the ETCS approach to set theory, equality between functions is primitive, and Axiom 3 stipulates that this equality can be detected by checking elements.

Some might see this difference as arguing in favor of ZF: it is more parsimonious, because it derives $f = g$ from something more fundamental. However, the defender of ETCS might claim in reply that her theory defines $x \in y$ from something more fundamental. Which is *really* more fundamental, equality between arrows (functions), or containment of objects (sets)? We'll leave that for other philosophers to think about.

Exercise 1.7. Show that any function $x : 1 \to X$ is monic.

Proposition 1.8. *A set X has exactly one element if and only if $X \cong 1$.*

Proof. The terminal object 1 has exactly one element, since there is a unique function $1 \to 1$.

Suppose now that X has exactly one element $x : 1 \to X$. We will show that X is a terminal object. First, for any set Y, there is a function $x \circ \beta_Y$ from Y to X. Now suppose that f, g are functions from Y to X such that $f \neq g$. By Axiom 3, there is an element $y \in Y$ such that $fy \neq gy$. But then X has more than one element, a contradiction. Therefore, there is a unique function from Y to X, and X is a terminal object. □

Proposition 1.9. *In any category with Cartesian products, if 1 is a terminal object then $X \times 1 \cong X$, for any object X.*

Proof. Let $\pi_0 : X \times 1 \to X$ and $\pi_1 : X \times 1 \to 1$ be the projections. Then $\langle 1_X, \beta_X \rangle$ is a function from X to $X \times 1$. We claim that π_0 is a left and right inverse for $\langle 1_X, \beta_X \rangle$. Since

$$\pi_0 \circ \langle 1_X, \beta_X \rangle \;=\; 1_X,$$

π_0 is a left inverse for $\langle 1_X, \beta_X \rangle$. To show that π_0 is a right inverse for $\langle 1_X, \beta_X \rangle$, we use the fact that $1_{X \times 1}$ is the unique function from $X \times 1$ to itself such that $\pi_0 \circ 1_{X \times 1} = \pi_0$, and $\pi_1 \circ 1_{X \times 1} = \pi_1$. But we also have,

$$\pi_0 \circ (\langle 1_X, \beta_X \rangle \circ \pi_0) \;=\; \pi_0,$$

and since π_1 is the unique function from $X \times 1$ to 1,

$$\pi_1 \circ (\langle 1_X, \beta_X \rangle \circ \pi_0) \;=\; \pi_1.$$

Thus, $\langle 1_X, \beta_X \rangle \circ \pi_0 = 1_{X \times 1}$, and π_0 is right inverse to $\langle 1_X, \beta_X \rangle$. □

Proposition 1.10. *Let a and b be elements of $X \times Y$. Then $a = b$ if and only if $\pi_0(a) = \pi_0(b)$ and $\pi_1(a) = \pi_1(b)$.*

Proof. Suppose that $\pi_0(a) = \pi_0(b)$ and $\pi_1(a) = \pi_1(b)$. By the uniqueness property of the product, there is a unique function $c : 1 \to X \times Y$ such that $\pi_0(c) = \pi_0(a)$ and $\pi_1(c) = \pi_1(a)$. Since a and b both satisfy this property, $a = b$. □

Note. The previous proposition justifies the use of the notation

$$X \times Y \;=\; \{\langle x, y\rangle \mid x \in X, y \in Y\}.$$

Here the identity condition for ordered pairs is given by

$$\langle x, y\rangle = \langle x', y'\rangle \qquad \text{iff} \qquad x = x' \text{ and } y = y'.$$

Proposition 1.11. *Let $(X \times Y, \pi_0, \pi_1)$ be the Cartesian product of X and Y. If Y is non-empty, then π_0 is an epimorphism.*

Proof. Suppose that Y is non-empty, and that $y : 1 \to Y$ is an element. Let $\beta_X : X \to 1$ be the unique map, and let $f = y \circ \beta_X$. Then $\langle 1_X, f\rangle : X \to X \times Y$ such that $\pi_0\langle 1_X, f\rangle = 1_X$. Since 1_X is epi, π_0 is epi. □

Definition. We say that $f : X \to Y$ is **injective** just in case: for any $x, y \in X$ if $f(x) = f(y)$, then $x = y$. Written more formally:

$$\forall x \forall y[f(x) = f(y) \to x = y]$$

Note. "Injective" is synonymous with "one-to-one".

Exercise 1.12. Let $f : X \to Y$ be a function. Show that if f is monic, then f is injective.

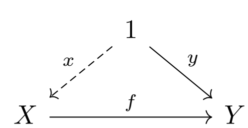
Proposition 1.13. *Let $f : X \to Y$ be a function. If f is injective then f is monic.*

Proof. Suppose that f is injective, and let $g, h : A \to X$ be functions such that $f \circ g = f \circ h$. Then for any $a \in A$, we have $f(g(a)) = f(h(a))$. Since f is injective, $g(a) = h(a)$. Since a was an arbitrary element of A, Axiom 3 entails that $g = h$. Therefore, f is monic. □

Definition. Let $f : X \to Y$ be a function. We say that f is **surjective** just in case: for each $y \in Y$, there is an $x \in X$ such that $f(x) = y$. Written formally:

$$\forall y \exists x[f(x) = y]$$

And in diagrammatic form:



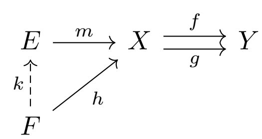
Note. "Surjective" is synonymous with "onto".

Exercise 1.14. Show that if $f : X \to Y$ is surjective then f is an epimorphism.

We will eventually establish that all epimorphisms are surjective. However, first we need a couple more axioms. Given a set X, and some definable condition ϕ on X, we would like to be able to construct a subset consisting of those elements in X that satisfy ϕ. The usual notation here is $\{x \in X \mid \phi(x)\}$, which we read as, "the x in X such that $\phi(x)$." But the important question is: which features ϕ do we allow? As an example of a definable condition ϕ, consider the condition of, "having the same value under the functions f and g," that is, $\phi(x)$ just in case $f(x) = g(x)$. We call the subset $\{x \in X \mid f(x) = g(x)\}$ the **equalizer** of f and g.

Axiom 4: Equalizers

Suppose that $f, g : X \rightrightarrows Y$ are functions. Then there is a set E and a function $m : E \to X$ with the following property: $fm = gm$, and for any other set F and function $h : F \to X$, if $fh = gh$, then there is a unique function $k : F \to E$ such that $mk = h$.



We call (E, m) an **equalizer** of f and g.

Exercise 1.15. Suppose that (E, m) and (E', m') are both equalizers of f and g. Show that there is an isomorphism $k : E \to E'$.

Definition. Let A, B, C be sets, and let $f : A \to C$ and $g : B \to C$ be functions. We say that g **factors through** f just in case there is a function $h : B \to A$ such that $fh = g$.

Exercise 1.16. Let $f, g : X \rightrightarrows Y$, and let $m : E \to X$ be the equalizer of f and g. Let $x \in X$. Show that x factors through m if and only if $f(x) = g(x)$.

Proposition 1.17. *In any category, if (E, m) is the equalizer of f and g. Then m is a monomorphism.*

Proof. Let $x, y : Z \to E$ such that $mx = my$. Since $fmx = gmx$, there is a unique arrow $z : Z \to E$ such that $mz = mx$. Since both $mx = mx$ and $my = mx$, it follows that $x = y$. Therefore, m is monic. $\square$

Definition. Let $f : X \to Y$ be a function. We say that f is a **regular monomorphism** just in case f is the equalizer of a pair of arrows $g, h : Y \rightrightarrows Z$.

Exercise 1.18. Show that if f is an epimorphism and a regular monomorphism, then f is an isomorphism.

In other approaches to set theory, one uses $\in$ to define a relation of inclusion between sets:

$$X \subseteq Y \iff \forall x(x \in X \to x \in Y).$$

We cannot define this exact notion in our approach since, for us, elements are attached to some particular set. However, for typical applications, every set under consideration will come equipped with a canonical monomorphism $m : X \to U$, where U is some fixed set. Thus, it will suffice to consider a relativized notion.

Definition. A **subobject** or **subset** of a set X is a set B and a monomorphism $m : B \to X$, called the **inclusion** of B in X. Given two subsets $m : B \to X$ and $n : A \to X$, we say that B is a subset of A (relative to X), written $B \subseteq_X A$ just in case there is a function $k : B \to A$ such that $nk = m$. When no confusion can result, we omit X and write $B \subseteq A$.

Let $m : B \to Y$ be monic, and let $f : X \to Y$. Consider the following diagram:

$$f^{-1}(B) \xrightarrow{\;k\;} X \times B \underset{mp_1}{\overset{fp_0}{\rightrightarrows}} Y$$

where $f^{-1}(B)$ is defined as the equalizer of $f\pi_0$ and mp_1. Intuitively, we have

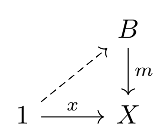
$$\begin{aligned} f^{-1}(B) &= \{\langle x, y\rangle \in X \times B \mid f(x) = y\} \\ &= \{\langle x, y\rangle \in X \times Y \mid f(x) = y \text{ and } y \in B\} \\ &= \{x \in X \mid f(x) \in B\}. \end{aligned}$$

Now we verify that $f^{-1}(B)$ is a subset of X.

Proposition 1.19. *The function $p_0k : f^{-1}(B) \to X$ is monic.*

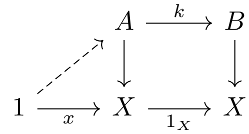
Proof. To simplify notation, let $E = f^{-1}(B)$. Let $x, y : Z \to E$ such that $p_0kx = p_0ky$. Then $fp_0kx = fp_0ky$, and hence $mp_1kx = mp_1ky$. Since m is monic, $p_1kx = p_1ky$. Thus, $kx = ky$. (The identity of a function into $X \times B$ is determined by the identity of its projections onto X and B.) Since k is monic, $x = y$. Therefore, p_0k is monic. $\square$

Definition. Let $m : B \to X$ be a subobject, and let $x : 1 \to X$. We say that $x \in B$ just in case x factors through m as follows:



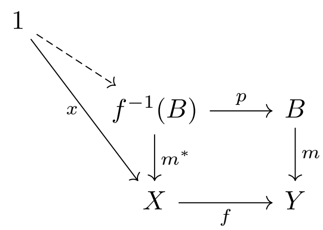
Proposition 1.20. *Let $A \subseteq B \subseteq X$. If $x \in A$ then $x \in B$.*

Proof.



□

Recall that $x \in f^{-1}(B)$ means: $x : 1 \to X$ factors through the inclusion of $f^{-1}(B)$ in X. Consider the following diagram:



First look just at the lower-right square. This square commutes, in the sense that following the arrows from $f^{-1}(B)$ clockwise gives the same answer as following the arrows from $f^{-1}(B)$ counterclockwise. The square has another property: for any set Z, and functions $g : Z \to X$ and $h : Z \to B$, there is a unique function $k : Z \to f^{-1}(B)$ such that $m^*k = g$ and $pk = h$. [To understand better, draw a picture!] When a commuting square has this property, then it's said to be a **pullback**.

Proposition 1.21. *Let $f : X \to Y$, and let $B \subseteq Y$. Then $x \in f^{-1}(B)$ if and only if $f(x) \in B$.*

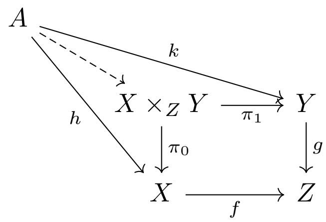
Proof. If $x \in f^{-1}(B)$, then there is an arrow $\hat{x} : 1 \to f^{-1}(B)$ such that $m^*\hat{x} = x$. Thus, $fx = mp\hat{x}$, which entails that the element $f(x) \in Y$ factors through B, i.e. $f(x) \in B$. Conversely, if $f(x) \in B$, then since the square is a pullback, $x : 1 \to X$ factors through $f^{-1}(B)$, i.e. $x \in f^{-1}(B)$. □

Definition. Given functions $f : X \to Z$ and $g : Y \to Z$, we define

$$X \times_Z Y \ = \ \{\langle x, y\rangle \in X \times Y \mid f(x) = g(y)\}.$$

In other words, $X \times_Z Y$ is the equalizer of $f\pi_0$ and $g\pi_1$. The set $X \times_Z Y$, together with the functions $\pi_0 : X \times_Z Y \to X$ and $\pi_1 : X \times_Z Y \to Y$ is called the **pullback** of f and g, alternatively the **fibered product** of f and g.

The pullback of f and g has the following distinguishing property: for any set A, and functions $h : A \to X$ and $k : A \to Y$ such that $fh = gk$, there is a unique function $j : A \to X \times_Z Y$ such that $\pi_0 j = h$ and $\pi_1 j = k$.



The following is an interesting special case of a pullback.

Definition. Let $f : X \to Y$ be a function. Then the **kernel pair** of f is the pullback $X \times_Y X$, with projections $p_0 : X \times_Y X \to X$ and $p_1 : X \times_Y X \to X$. Intuitively, $X \times_Y X$ is the relation, "having the same image under f." Written in terms of braces,

$$X \times_Y X \;=\; \{\langle x, x' \rangle \in X \times X \mid f(x) = f(x')\}.$$

In particular, f is injective if and only if, "having the same image under f" is coextensive with the equality relation on X. That is, $X \times_Y X = \{\langle x, x \rangle \mid x \in X\}$, which is the diagonal of X.

Exercise 1.22. Let $f : X \to Y$ be a function, and let $p_0, p_1 : X \times_Y X \rightrightarrows X$ be the kernel pair of f. Show that the following are equivalent:

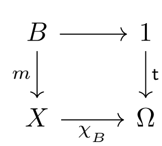
1. f is a monomorphism.
2. p_0 and p_1 are isomorphisms.
3. $p_0 = p_1$.

2 Truth values and subsets

Axiom 5: Truth-value object

There is a set Ω with the following features:

1. Ω has exactly two elements, which we denote by $\mathtt{t} : 1 \to \Omega$ and $\mathsf{f} : 1 \to \Omega$.
2. For any set X, and subobject $m : B \to X$, there is a unique function $\chi_B : X \to \Omega$ such that the following diagram is a pullback:



In other words, $B = \{x \in X \mid \chi_B(x) = \mathsf{t}\}$.

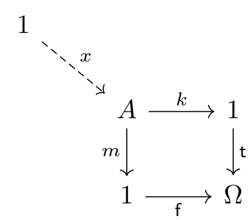
Intuitively speaking, the first part of Axiom 5 says that Ω is a two-element set, say $\Omega = \{\mathsf{f}, \mathsf{t}\}$. The second part of Axiom 5 says that Ω classifies the subobjects of a set X. That is, each subobject $m : B \to X$ corresponds to a unique **characteristic function** $\chi_B : X \to \{\mathsf{f}, \mathsf{t}\}$ such that $\chi_B(x) = \mathsf{t}$ if and only if $x \in B$.

The terminal object 1 is a set with one element. Thus, it should be the case that 1 has two subsets, the empty set and 1 itself.

Proposition 2.1. *The terminal object* 1 *has exactly two subobjects.*

Proof. By Axiom 5, subobjects of 1 correspond to functions $1 \to \Omega$, that is, to elements of Ω. By Axiom 5, Ω has exactly two elements. Therefore, 1 has exactly two subobjects. □

Obviously the function $\mathsf{t} : 1 \to \Omega$ corresponds to the subobject $\mathrm{id}_1 : 1 \to 1$. Can we say more about the subobject $m : A \to 1$ corresponding to the function $\mathsf{f} : 1 \to \Omega$? Intuitively, we should have $A = \{x \in 1 \mid \mathsf{t} = \mathsf{f}\}$, in other words, the empty set. To confirm this intuition, consider the pullback diagram:



Note that m and k must both be the unique function from A to 1, that is $m = k = \beta_A$. Suppose that A is nonempty, i.e. there is a function $x : 1 \to A$. Then $\beta_A \circ x$ is the identity $1 \to 1$, and since the square commutes, $\mathsf{t} = \mathsf{f}$, a contradiction. Therefore, A has no elements.

Exercise 2.2. Show that $\Omega \times \Omega$ has exactly four elements.

We now use the existence of a truth-value object in **Sets** to demonstrate further properties of functions.

Definition. Let $f : X \to Y$ be a function. Then f is said to be a **regular monomorphism** just in case there are functions $g, h : Y \rightrightarrows Z$ such that f is the equalizer of g and h.

Exercise 2.3. Show that in any category, if $f : X \to Y$ is a regular monomorphism, then f is monic.

Proposition 2.4. *Every monomorphism between sets is regular, i.e. an equalizer of a pair of parallel arrows.*

Proof. Let $m : B \to X$ be monic. By Axiom 5, the following is a pullback diagram:

$$\begin{array}{ccc} B & \longrightarrow & 1 \\ {\scriptstyle m}\downarrow & & \downarrow{\scriptstyle t} \\ X & \underset{\chi_B}{\longrightarrow} & \Omega \end{array}$$

A straightforward verification shows that m is the equalizer of $X \overset{\beta_X}{\to} 1 \overset{\mathsf{t}}{\to} \Omega$ and $\chi_B : X \to \Omega$. Therefore, m is regular monic. □

Students with some background in mathematics might assume that if a function $f : X \to Y$ is both a monomorphism and an epimorphism, then it is an isomorphism. However, that isn't true in all categories! [For example, it's not true in the category of monoids.] Nonetheless, **Sets** is a special category, and in this case we have the result:

Proposition 2.5. *In* **Sets***, if a function is both a monomorphism and an epimorphism, then it is an isomorphism.*

Proof. In any category, if m is regular monic and epi, then m is an isomorphism (Exercise 1.18). □

Definition. Let $f : X \to Y$ be a function, and let $y \in Y$. The **fiber** over y is the subset $f^{-1}\{y\}$ of X given by the following pullback:

$$\begin{array}{ccc} f^{-1}\{y\} & \longrightarrow & 1 \\ \downarrow & & \downarrow{\scriptstyle y} \\ X & \overset{f}{\longrightarrow} & Y \end{array}$$

Proposition 2.6. *Let $p : X \to Y$. If p is not a surjection, then there is a $y_0 \in Y$ such that the fiber $p^{-1}\{y_0\}$ is empty.*

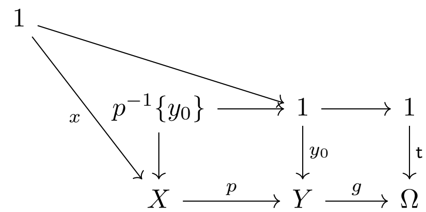
Proof. Since p is not a surjection, there is a $y_0 \in Y$ such that for all $x \in X$, $p(x) \neq y_0$. Now consider the pullback:

$$\begin{array}{cccc} 1 & & & \\ & \overset{z}{\searrow} & & \\ & p^{-1}\{y_0\} & \longrightarrow & 1 \\ & \downarrow{\scriptstyle m} & & \downarrow{\scriptstyle y_0} \\ & X & \overset{p}{\longrightarrow} & Y \end{array}$$

If there were a morphism $z : 1 \to p^{-1}\{y_0\}$, then we would have $p(m(z)) = y_0$, a contradiction. Therefore, $p^{-1}\{y_0\}$ is empty. □

Proposition 2.7. *In* **Sets***, epimorphisms are surjective.*

Proof. Suppose that $p : X \to Y$ is not a surjection. Then there is a $y_0 \in Y$ such that for all $x \in X$, $p(x) \neq y_0$. Since 1 is terminal, the morphism $y_0 : 1 \to Y$ is monic. Consider the following diagram:

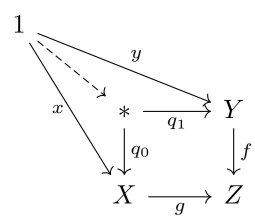


Here g is the characteristic function of $\{y_0\}$; by Axiom 5, g is the unique function that makes the right hand square a pullback. Let $x \in X$ be arbitrary. If we had $g(p(x)) = \mathsf{t}$, then there would be an element $x' \in p^{-1}\{y_0\}$, in contradiction with the fact that the latter is empty (Proposition 2.6). By Axiom 5, either $g(p(x)) = \mathsf{t}$ or $g(p(x)) = \mathsf{f}$; therefore, $g(p(x)) = \mathsf{f}$. Now let h be the composite $Y \to 1 \xrightarrow{\mathsf{f}} \Omega$. Then, for any $x \in X$, we have $h(p(x)) = \mathsf{f}$. Since $g \circ p$ and $h \circ p$ agree on arbitrary $x \in X$, we have $g \circ p = h \circ p$. Since $g \neq h$, it follows that p is not an epimorphism. $\square$

In a general category, there is no guarantee that an epimorphism pulls back to an epimorphism. However, in **Sets**, we have the following:

Proposition 2.8. *In* **Sets***, the pullback of an epimorphism is an epimorphism.*

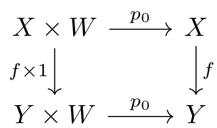
Proof. Suppose that $f : Y \to Z$ is epi, and let $x \in X$. Consider the pullback diagram:



By Proposition 2.7, f is surjective. In particular, there is a $y \in Y$ such that $f(y) = g(x)$. Since the diagram is a pullback, there is a unique $\langle x, y\rangle : 1 \to *$ such that $q_0\langle x, y\rangle = x$ and $q_1\langle x, y\rangle = y$. Therefore, q_0 is surjective, and hence epi. $\square$

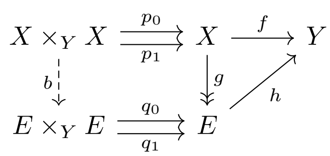
Proposition 2.9. *If $f : X \to Y$ and $g : W \to Z$ are epimorphisms, then so is $f \times g : X \times W \to Y \times Z$.*

Proof. Since $f \times g = (f \times 1) \circ (1 \times g)$, it will suffice to show that $f \times 1$ is epi when f is epi. Now, the following diagram is a pullback:



By Proposition 2.8, if f is epi, then $f \times 1$ is epi. $\square$

Suppose that $f : X \to Y$ is a function, and that $p_0, p_1 : X \times_Y X \rightrightarrows X$ is the kernel pair of f. Suppose also that $h : E \to Y$ is a function, that $q_0, q_1 : E \times_Y E \rightrightarrows E$ is the kernel pair of h, and that $g : X \twoheadrightarrow E$ is an epimorphism. Then there is a unique function $b : X \times_Y X \to E \times_Y E$, such that $q_0 b = g p_0$ and $q_1 b = g p_1$.



An argument similar to the one above shows that b is an epimorphism. We will use this fact below to describe the properties of epimorphisms in **Sets**.

3 Relations

3.1 Equivalence relations and equivalence classes

A relation R on a set X is a subset of $X \times X$; i.e. a set of ordered-pairs. A relation is said to be an equivalence relation just in case it is reflexive, symmetric, and transitive. One particular way that equivalence relations on X arise is from functions with X as domain: given a function $f : X \to Y$, let say that $\langle x, y \rangle \in R$ just in case $f(x) = f(y)$. [Sometimes we say that, "x and y lie in the same fiber over Y."] Then R is an equivalence relation on X.

Given an equivalence relation R on X, and some element $x \in X$, let $[x] = \{y \in X \mid \langle x, y \rangle \in R\}$ denote the set of all elements of X that are equivalent to X. We say that $[x]$ is the **equivalence class** of x. It's straightforward to show that for any $x, y \in X$, either $[x] = [y]$ or $[x] \cap [y] = \emptyset$. Moreover, for any $x \in X$, we have $x \in [x]$. Thus the equivalence classes form a **partition** of X into disjoint subsets.

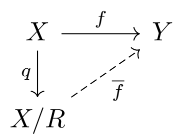
We'd like now to be able to talk about the set of these equivalence classes, i.e. something that might intuitively be written as $\{[x] \mid x \in X\}$. The following axiom guarantees the existence of such a set, called X/R, and a canonical mapping $q : X \to X/R$ that takes each element $x \in X$ to its equivalence class $[x] \in X/R$.

Our next axiom guarantees the existence of the set of equivalence classes.

Axiom 6: Equivalence classes

Let R be an equivalence relation on X. Then there is a set X/R, and a function $q : X \to X/R$ with the properties:

1. $\langle x, y\rangle \in R$ if and only if $q(x) = q(y)$.
2. For any set Y and function $f : X \to Y$ that is constant on equivalence classes, there is a unique function $\overline{f} : X/R \to Y$ such that $\overline{f} \circ q = f$.



Here f is constant on equivalence classes just in case $f(x) = f(y)$ whenever $\langle x, y\rangle \in R$.

An equivalence relation R can be thought of as a subobject of $X \times X$, i.e. a subset of ordered pairs. Accordingly, there are two functions $p_0 : R \to X$ and $p_1 : R \to X$ given by: $p_0\langle x, y\rangle = x$ and $p_1\langle x, y\rangle = y$. Then condition (1) in the above axiom says that $q \circ p_0 = q \circ p_1$. And condition (2) says that for any function $f : X \to Y$ such that $f \circ p_0 = f \circ p_1$, there is a unique function $\overline{f} : X/R \to Y$ such that $\overline{f} \circ q = f$. In this case, we say that q is a **coequalizer** of p_0 and p_1.

Exercise 3.1. Show that in any category, coequalizers are unique up to isomorphism.

Exercise 3.2. Show that in any category, a coequalizer is an epimorphism.

Exercise 3.3. For a function $f : X \to Y$, let $R = \{\langle x, y\rangle \in X \times X \mid f(x) = f(y)\}$. That is, R is the kernel pair of f. Show that R is an equivalence relation.

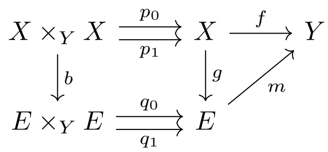
Definition. A function $f : X \to Y$ is said to be a **regular epimorphism** just in case f is a coequalizer.

Exercise 3.4. Show that in any category, if $f : X \to Y$ is both a monomorphism and a regular epimorphism, then f is an isomorphism.

Proposition 3.5. *Every epimorphism in* **Sets** *is regular. In particular, every epimorphism is the coequalizer of its kernel pair.*

Proof. Let $f : X \to Y$ be an epimorphism. Let $p_0, p_1 : X \times_Y X \rightrightarrows X$ be the kernel pair of f. By Axiom 6, the coequalizer $g : X \to E$ of p_0 and p_1 exists; and since f also coequalizes p_0 and p_1, there is a unique function $m : E \to Y$

such that $f = mg$.



Here $E \times_Y E$ is the kernel pair of m. Since $mgp_0 = fp_0 = fp_1 = mgp_1$, there is a unique function $b : X \times_Y X \to E \times_Y E$ such that $gp_0 = q_0 b$ and $gp_1 = q_1 b$. By the considerations at the end of the previous section, b is an epimorphism. Furthermore,

$$q_0 b = gp_0 = gp_1 = q_1 b,$$

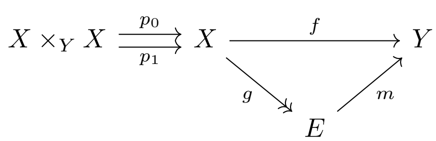
and therefore $q_0 = q_1$. By Exercise 1.22, m is a monomorphism. Since $f = mg$, and f is epi, m is also epi. Therefore, by Proposition 2.5, m is an isomorphism. □

This last proposition actually shows that **Sets** is what is known as a **regular category**. In general, a category **C** is said to be **regular** just in case it has all finite limits, if coequalizers of kernel pairs exist, and if regular epimorphisms are stable under pullback. Now, it's known that if a category has products and equalizers, then it has all finite limits. Thus **Sets** has all finite limits. Our most recent axiom says that **Sets** has coequalizers of kernel pairs. And finally, all epimorphisms in **Sets** are regular, and epimorphisms in **Sets** are stable under pullback; therefore, regular epimorphisms are stable under pullback.

Regular categories have several nice features that will prove quite useful. In the remainder of this section, we will discuss one such features: factorization of functions into a regular epimorphism followed by a monomorphism.

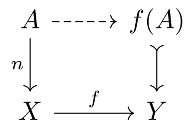
3.2 The epi-monic factorization

Let $f : X \to Y$ be a function, and let $p_0, p_1 : X \times_Y X \rightrightarrows X$ be the kernel pair of f. By Axiom 6, the kernel pair has a coequalizer $g : X \twoheadrightarrow E$. Since f also coequalizes p_0 and p_1, there is a unique function $m : E \to Y$ such that $f = mg$.



An argument similar to the one in Proposition 3.5 shows that m is a monomorphism. Thus, (E, m) is a subobject of Y, which we call the **image** of X under f, and we write $E = f(X)$. The pair (g, m) is called the **epi-monic factorization** of f. Since epis are surjections, and monics are injections, (g, m) can also be called the surjective-injective factorization.

Definition. Suppose that A is a subset of X, in particular, $n : A \to X$ is monic. Then $f \circ n : A \to Y$, and we let $f(A)$ denote the image of A under $f \circ n$.

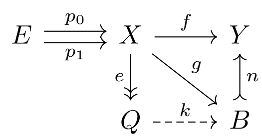


We also use the suggestive notation

$$f(A) \;=\; \exists_f(A) \;=\; \{y \in Y \mid \exists x \in A. f(x) = y\}.$$

Proposition 3.6. *Let $f : X \to Y$ be a function, and let A be a subobject of X. The image $f(A)$ is the smallest subobject of Y through which f factors.*

Proof. Let $e : X \to Q$ and $m : Q \to Y$ be the epi-monic factorization of f. Suppose that $n : B \to Y$ is a subobject, and that f factors through n, say $f = ng$. Consider the following diagram.

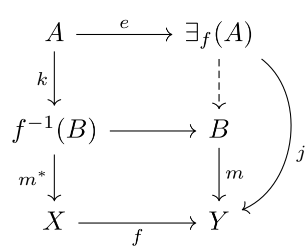


Then $ngp_0 = fp_0 = fp_1 = ngp_1$, since p_0, p_1 is the kernel pair of f. Since n is monic, $gp_0 = gp_1$, i.e. g coequalizes p_0 and p_1. Since $e : X \to Q$ is the coequalizer of p_0 and p_1, there is a unique function $k : Q \to B$ such that $ke = g$. By uniqueness of the epi-monic factorization, $nk = m$. Therefore, $Q \subseteq B$. □

Proposition 3.7. *For any $A \subseteq X$ and $B \subseteq Y$, we have*

$$A \subseteq f^{-1}(B) \quad \textit{if and only if} \quad \exists_f(A) \subseteq B.$$

Proof. Suppose first that $A \subseteq f^{-1}(B)$, in particular that $k : A \to f^{-1}(B)$. Consider the following diagram:



By definition, je is the epi-mono factorization of fm^*k. Since fm^*k also factors through $m : B \to Y$, we have $\exists_f(A) \subseteq B$, by Proposition 3.6.

Suppose now that $\exists_f(A) \subseteq B$. Using the fact that the lower square in the diagram is a pullback, we see that there is an arrow $k : A \to f^{-1}(B)$ such that m^*k is the inclusion of A in X. That is, $A \subseteq f^{-1}(B)$. □

Exercise 3.8. Use the previous result to show that $A \subseteq f^{-1}(\exists_f(A))$, for any subset A of X.

3.3 Functional relations

Definition. A relation $R \subseteq X \times Y$ is said to be **functional** just in case for each $x \in X$ there is a unique $y \in Y$ such that $\langle x, y\rangle \in R$.

Definition. Suppose that $f : X \to Y$ is a function. We let $\text{graph}(f) = \{\langle x, y\rangle \mid f(x) = y\}$.

Exercise 3.9. Show that graph(f) is a functional relation.

The following result is helpful for establishing the existence of arrows $f : X \to Y$.

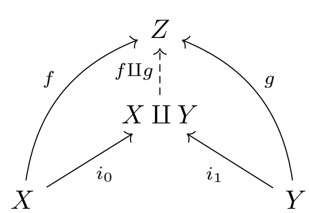
Proposition 3.10. *Let $R \subseteq X \times Y$ be a functional relation. Then there is a unique function $f : X \to Y$ such that $R = \text{graph}(f)$.*

The proof of this result is somewhat complicated, and we omit it (for the time being).

4 Colimits

Axiom 7: Coproducts

For any two sets X, Y, there is a set $X \amalg Y$ and functions $i_0 : X \to X \amalg Y$ and $i_1 : Y \to X \amalg Y$ with the feature that for any set Z, and functions $f : X \to Z$ and $g : Y \to Z$, there is a unique function $f \amalg g : X \amalg Y \to Z$ such that $(f \amalg g) \circ i_0 = f$ and $(f \amalg g) \circ i_1 = g$.



We call $X \amalg Y$ the **coproduct** of X and Y. We call i_0 and i_1 the coprojections of the coproduct.

Intuitively speaking, the coproduct $X \amalg Y$ is the disjoint union of the sets X and Y. What we mean by "disjoint" here is that if X and Y share elements in common (which doesn't make sense in our framework, but does in some frameworks), then these elements are dis-identified before the union is taken. For example, in terms of elements, we could think of $X \amalg Y$ as consisting of elements of the form $\langle x, 0\rangle$, with $x \in X$, and elements of the form $\langle y, 1\rangle$, with $y \in Y$. Thus, if x is contained in both X and Y, then $X \amalg Y$ contains two separate copies of x, namely $\langle x, 0\rangle$ and $\langle x, 1\rangle$.

We now show that that the inclusions $i_0 : X \to X \amalg Y$ and $i_1 : Y \to X \amalg Y$ do, in fact, have disjoint images.

Proposition 4.1. *Coproducts in* **Sets** *are disjoint. In other words, if* $i_0 : X \to X \amalg Y$ *and* $i_1 : Y \to X \amalg Y$ *are the coprojections, then* $i_0(x) \neq i_1(y)$ *for all* $x \in X$ *and* $y \in Y$*.*

Proof. Suppose for reductio ad absurdum that $i_0(x) = i_1(y)$. Let $g : X \to \Omega$ be the unique map that factors through $\mathsf{t} : 1 \to \Omega$. Let $h : Y \to \Omega$ be the unique map that factors through $\mathsf{f} : 1 \to \Omega$. By the universal property of the coproduct, there is a unique function $g \amalg h : X \amalg Y \to \Omega$ such that $(g \amalg h)i_0 = g$ and $(g \amalg h)i_1 = h$. Thus, we have

$$\mathsf{t} = g(x) = (g \amalg h)i_0 x = (g \amalg h)i_1 y = h(y) = \mathsf{f},$$

a contradiction. Therefore, $i_0(x) \neq i_1(y)$, and the ranges of i_0 and i_1 are disjoint. □

Proposition 4.2. *The coprojections* $i_0 : X \to X \amalg Y$ *and* $i_1 : Y \to X \amalg Y$ *are monomorphisms.*

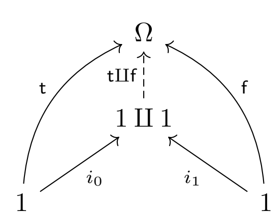
Proof. We will show that i_0 is monic; the result then follows by symmetry. Suppose first that X has no elements. Then i_0 is trivially injective, hence monic by Proposition 1.13. Suppose now that X has an element $x : 1 \to X$. Let $g = x \circ \beta_Y$, where $\beta_Y : Y \to 1$. Then $(1_X \amalg g)i_0 = 1_X$, and Exercise 1.1 entails that i_0 is monic. □

Proposition 4.3. *The coprojections are jointly surjective. That is, for each* $z \in X \amalg Y$*, either there is an* $x \in X$ *such that* $z = i_0(x)$*, or there is a* $y \in Y$ *such that* $z = i_1(y)$*.*

Proof. Suppose for reductio ad absurdum that z is neither in the image of i_0 nor in the image of i_1. Let $g : (X \amalg Y) \to \Omega$ be the characteristic function of $\{z_0\}$. Then for all $x \in X$, $g(i_0(x)) = f$. And for all $y \in Y$, $g(i_1(y)) = f$. Now let $h : (X \amalg Y) \to \Omega$ be the constant f function, i.e. $h(z) = f$ for all $z \in X \amalg Y$. Then $gi_0 = hi_0$ and $gi_1 = hi_1$. Since functions from $X \amalg Y$ are determined by their coprojections, $g = h$, a contradiction. Therefore, all $z \in X \amalg Y$ are either in the range of i_0 or in the range of i_1. □

Proposition 4.4. *The function* $\mathsf{t} \amalg \mathsf{f} : 1 \amalg 1 \to \Omega$ *is an isomorphism.*

Proof. Consider the diagram:



Then $\mathtt{t} \amalg \mathtt{f}$ is monic since every element of $1 \amalg 1$ factors through either i_0 or i_1 (Proposition 4.3), and since $\mathtt{t} \neq \mathtt{f}$. Furthermore, $\mathtt{t} \amalg \mathtt{f}$ is epi since $\mathtt{t}$ and $\mathtt{f}$ are the only elements of Ω. By Proposition 2.5, $\mathtt{t} \amalg \mathtt{f}$ is an isomorphism. □

Proposition 4.5. *Let X be a set, and let B be a subset of X. Then the inclusion $B \amalg X \backslash B \to X$ is an isomorphism.*

Proof. Using the fact that Ω is Boolean, for every $x \in X$, either $x \in B$ or $x \in X \backslash B$. Thus the inclusion $B \amalg X \backslash B \to X$ is a bijection, hence an isomorphism. □

Axiom 8: Empty set

There is a set $\emptyset$ with the following properties:

1. For any set X, there is a unique function

$$\emptyset \xrightarrow{\alpha_X} X$$

   In this case, we say that $\emptyset$ is an **initial object** in **Sets**.

2. $\emptyset$ is empty, i.e. there is no function $x : 1 \to \emptyset$.

Exercise 4.6. Show that in any category with coproducts, if A is an initial object, then $X \amalg A \cong X$, for any object X.

Proposition 4.7. *Any function $f : X \to \emptyset$ is an isomorphism.*

Proof. Since $\emptyset$ has no elements, f is trivially surjective. We now claim that X has no elements. Indeed, if $x : 1 \to X$ is an element of X, then $f(x)$ is an element of $\emptyset$. Since X has no elements, f is trivially injective. By Proposition 2.5, f is an isomorphism. □

Proposition 4.8. *A set X has no elements if and only if $X \cong \emptyset$.*

Proof. By Axiom 8, the set $\emptyset$ has no elements. Thus if $X \cong \emptyset$, then X has no elements.

Suppose now that X has no elements. Since $\emptyset$ is an initial object, there is a unique arrow $\alpha_X : \emptyset \to X$. Since X has no elements, α_X is trivially surjective. Since $\emptyset$ has no elements, α_X is trivially injective. By Proposition 2.5, f is an isomorphism. □

5 Sets of functions and sets of subsets

[[Note to students. This section is highly technical, and you are *not* required to read it. Please simply familiarize yourself with the definition of an exponential object, and the definition of the powerset of a set.]]

One distinctive feature of the category of sets is its ability to model almost any mathematical construction. One such construction is gathering together old things into a new set. For example, given two sets A and X, can we form a set X^A of all functions from A to X? Similarly, given a set X, can we form a set $\mathscr{P}X$ of all subsets of X?

As usual, we won't be interested in hard questions about what it takes to be a set. Rather, we're interested in hypothetical questions: if such a set existed, what would it be like? The crucial features of X^A seem to be captured by the following axiom:

Axiom 9: Exponential objects

Suppose that A and X are sets. Then there is a set X^A, and a function $e_X : A \times X^A \to X$ such that for any set Z, and function $f : A \times Z \to X$, there is a unique function $f^\sharp : Z \to X^A$ such that $e_X \circ (1_A \times f^\sharp) = f$.

$$\begin{array}{ccc} A \times X^A & \xrightarrow{e_X} & X \\ {\scriptstyle 1\times f^\sharp}\uparrow & \nearrow_{f} & \\ A \times Z & & \end{array}$$

The set X^A is called an **exponential object**, and the function $f^\sharp : Z \to X^A$ is called the **transpose** of $f : A \times Z \to X$.

The way to remember this axiom is to think of Y^X as the set of functions from X to Y, and to think of $e : X \times Y^X \to Y$ as a meta-function that takes an element $f \in Y^X$, and an element $x \in X$, and returns the value $e(f, x) = f(x)$. For this reason, $e : X \times Y^X \to Y$ is sometimes called the **evaluation function**.

Note further that if $f : X \times Z \to Y$ is a function, then for each $z \in Z$, $f(-, z)$ is a function from $X \to Y$. In other words, f corresponds uniquely to a function from Z to functions from Y to X. This latter function is the transpose $f^\sharp : Z \to Y^X$ of f.

We have written Axiom 9 in first-order fashion, but it might help to think of it as stating that there is a one-to-one correspondence between two sets:

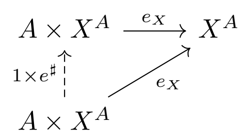
$$\mathrm{hom}(X \times Z, Y) \;\cong\; \mathrm{hom}(Z, Y^X),$$

where $\mathrm{hom}(A, B)$ is though of as the set of functions from A to B. As a particular case, when $Z = 1$, the terminal object, we have

$$\hom(X, Y) \;\cong\; \hom(1, Y^X).$$

In other words, elements of Y^X in the "internal sense" correspond to elements of $\hom(X, Y)$ in the "external sense."

Consider now the following special case of the above construction:

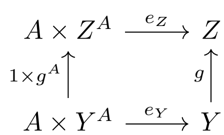


Thus, $e_X^\sharp = 1_{X^A}$.

Definition. Suppose that $g : Y \to Z$ is a function. We let $g^A : X^A \to Y^A$ denote the transpose of the function:

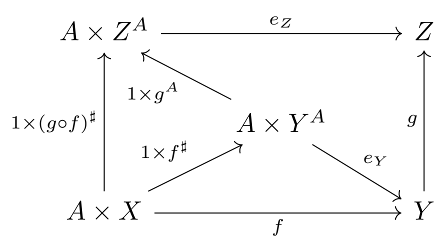
$$A \times Y^A \xrightarrow{e_Y} Y \xrightarrow{g} Z$$

That is, $g^A = (g \circ e_Y)^\sharp$, and the following diagram commutes:



Proposition 5.1. *Let $f : A \times X \to Y$ and $g : Y \to Z$ be functions. Then $(g \circ f)^\sharp = g^A \circ f^\sharp$.*

Proof. Consider the following diagram:



The bottom triangle commutes by the definition of $f^\sharp$. The upper right triangle commutes by the definition of g^A. And the outer square commutes by the definition of $(g \circ f)^\sharp$. It follows that

$$e_Z \circ (1 \times (g^A \circ f^\sharp)) \;=\; g \circ f,$$

and hence $g^A \circ f^\sharp = (g \circ f)^\sharp$. □

Consider now the following particular case:

$$\begin{array}{ccc} A \times (A \times X)^A & \xrightarrow{e} & A \times X \\ {\scriptstyle 1\times p}\uparrow & \nearrow_{1} & \\ A \times X & & \end{array}$$

Here $p = 1^\sharp$ is the unique function such that $e(1_A \times p) = 1_{A\times X}$. Intuitively, we can think of p as the function that takes an element $x \in X$, and returns the function $p_x : A \to A \times X$ such that $p_x(a) = \langle a, x\rangle$. Thus, $(1 \times p)\langle a, x\rangle = \langle a, p_x\rangle$, and $e(1 \times p)\langle a, x\rangle = p_x(a) = \langle a, x\rangle$.

Definition. Suppose that $f : Z \to X^A$ is a function. We define $f^\flat : Z \times A \to X$ to be the following composite function:

$$A \times Z \xrightarrow{1\times f} A \times X^A \xrightarrow{e_X} X$$

Proposition 5.2. *Let $f : X \to Y$ and $g : Y \to Z^A$ be functions. Then $(g \circ f)^\flat = g^\flat \circ (1_A \times f)$.*

Proof. By definition,

$$(g \circ f)^\flat = e_X \circ (1 \times (g \circ f)) = e_X \circ (1 \times g) \circ (1 \times f) = g^\flat \circ (1 \times f).$$

□

Proposition 5.3. *For any function $f : A \times Z \to X$, we have $(f^\sharp)^\flat = f$.*

Proof. By the definitions, we have

$$(f^\sharp)^\flat \ = \ e_X \circ (1 \times f^\sharp) = f.$$

□

Proposition 5.4. *For any function $f : Z \to X^A$, we have $(f^\flat)^\sharp = f$.*

Proof. By definition, $(f^\flat)^\sharp$ is the unique function such that $e_X \circ (1 \times (f^\flat)^\sharp) = f^\flat$. But also $e_X \circ (1 \times f) = f^\flat$. Therefore, $(f^\flat)^\sharp = f$. □

Proposition 5.5. *For any set X, we have $X^1 \cong X$.*

Proof. Let $e : 1 \times X^1 \to X$ be the evaluation function from Axiom 9. We claim that e is a bijection. Recall that there is a natural isomorphism $i : 1 \times 1 \to 1$. Consider the following diagram:

$$\begin{array}{ccc} 1 \times X^1 & \xrightarrow{e} & X \\ {\scriptstyle 1\times x^\sharp}\uparrow & & \uparrow{\scriptstyle x} \\ 1 \times 1 & \xrightarrow{i} & 1 \end{array}$$

That is, for any element $x : 1 \to X$, there is a unique element $x^\sharp$ of X^1 such that $e(1 \times x^\sharp) = x$. Thus, e is a bijection, and $X \cong 1 \times X^1$ is isomorphic to X. □

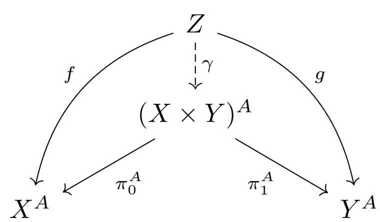
Proposition 5.6. *For any set X, we have $X^{\emptyset} \cong 1$.*

Proof. Elements of $X^{\emptyset}$ correspond functions $\emptyset \to X$. There is exactly one such function, hence $X^{\emptyset}$ has exactly one element $x : 1 \to X^{\emptyset}$. Thus, x is a bijection, and $X^{\emptyset} \cong 1$. □

Proposition 5.7. *For any sets A, X, Y, we have $(X \times Y)^A \cong X^A \times Y^A$.*

Proof. An elegant proof of this proposition would note that $(-)^A$ is a functor, and is right adjoint to the functor $A \times (-)$. Since right adjoints preserve products, $(X \times Y)^A \cong X^A \times Y^A$. Nonetheless, we will go into further detail.

By uniqueness of Cartesian products, it will suffice to show that $(X \times Y)^A$ is a Cartesian product of X^A and Y^A, with projections π_0^A and π_1^A. Let Z be an arbitrary set, and let $f : Z \to X^A$ and $g : Z \to Y^A$ be functions. Now take $\gamma = \langle f^\flat, g^\flat \rangle^\sharp$, where $f^\flat : A \times Z \to X$ and $g^\flat : A \times Z \to Y$.



We claim that $\pi_0^A \gamma = f$ and $\pi_1^A \gamma = g$. Indeed,

$$\pi_0^A \circ \gamma = \pi_0^A \circ \langle f^\flat, g^\flat \rangle^\sharp = (\pi_0 \circ \langle f^\flat, g^\flat \rangle)^\sharp = (f^\flat)^\sharp = f.$$

Thus, $\pi_0^A \gamma = f$, and similarly, $\pi_1^A \gamma = g$.

Suppose now that $h : Z \to (X \times Y)^A$ such that $\pi_0^A h = f$ and $\pi_1^A h = g$. Then

$$f = \pi_0^A \circ (h^\flat)^\sharp = (\pi_0 \circ h^\flat)^\sharp.$$

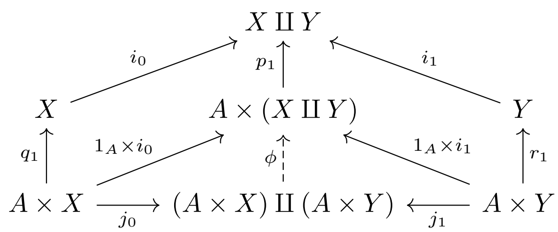
Hence, $\pi_0 \circ h^\flat = f^\flat$, and similarly, $\pi_1 \circ h^\flat = g^\flat$. That is, $h^\flat = \langle f^\flat, g^\flat \rangle$, and $h = \langle f^\flat, g^\flat \rangle^\sharp = \gamma$. □

Proposition 5.8. *For any sets A, X, Y, we have $A \times (X \amalg Y) \cong (A \times X) \amalg (A \times Y)$.*

Proof. Even without Axiom 9, there is always a canonical function from $(A \times X) \amalg (A \times Y)$ to $A \times (X \amalg Y)$, namely $\phi := (1_A \times i_0) \amalg (1_A \times i_1)$, where i_0 and i_1 are the coproduct inclusions of $X \amalg Y$. That is,

$$\phi \circ j_0 = 1_A \times i_0, \qquad \text{and} \qquad \phi \circ j_1 = 1_A \times i_1,$$

where j_0 and j_1 are the coproduct inclusions of $(A \times X) \amalg (A \times Y)$.



We will show that Axiom 9 entails that ϕ is invertible.

Let $g : A\times(X\amalg Y) \to A\times(X\amalg Y)$ be the identity, i.e. $g = 1_{A\times(X\amalg Y)}$. Then $g^\sharp : X \amalg Y \to (A \times (X \amalg Y))^A$ is the unique function such that $e(1_A \times g^\sharp) = g$. By Proposition 5.2,

$$(g^\sharp \circ i_0)^\flat = g \circ (1_A \times i_0) = 1_A \times i_0.$$

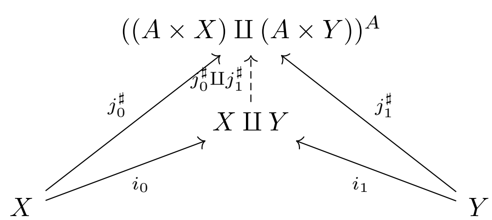
Similarly, $(g^\sharp \circ i_1)^\flat = 1_A \times i_1$. Thus,

$$g^\sharp \ = \ (1_A \times i_0)^\sharp \amalg (1_A \times i_1)^\sharp.$$

We also have $(1_A \times i_0)^\sharp = (\phi \circ j_0)^\sharp = \phi^A \circ j_0^\sharp$, and $(1^A \times i_1)^\sharp = \phi^A \circ j_1^\sharp$. Hence

$$g^\sharp \ = \ (\phi^A \circ j_0^\sharp) \amalg (\phi^A \circ j_1^\sharp) \ = \ \phi^A \circ (j_0^\sharp \amalg j_1^\sharp).$$

Now for the inverse of ϕ, we take $\psi = (j_0^\sharp \amalg j_1^\sharp)^\flat$.



It then follows that

$$(\phi \circ \psi)^\sharp \ = \ \phi^A \circ (j_0^\sharp \amalg j_1^\sharp) \ = \ g^\sharp,$$

and therefore $\phi \circ \psi = 1_{A\times(X\amalg Y)}$. Similarly,

$$(\psi \circ \phi \circ j_0)^\sharp = \psi^A \circ (\phi \circ j_0)^\sharp = \psi^A \circ g^\sharp \circ i_0 = \psi^\sharp \circ i_0 = j_0^\sharp.$$

Thus, $\psi \circ \phi \circ j_0 = j_0$, and a similar calculation shows that $\psi \circ \phi \circ j_1 = j_1$. It follows that $\psi \circ \phi = 1_{(A\times X)\amalg(A\times Y)}$. Thus, ψ is a two-sided inverse for ϕ, and $A \times (X \amalg Y)$ is isomorphic to $(A \times X) \amalg (A \times Y)$. □

Definition (Powerset)**.** If X is a set, we let $\mathscr{P}X = \Omega^X$.

Intuitively speaking, $\mathscr{P}X$ is the set of all subsets of X. For example, if $X = \{a, b\}$, then $\mathscr{P}X = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$. More rigorously, each element of Ω^X corresponds to a function $1 \to \Omega^X$, which in turn corresponds to a function $X \cong 1 \times X \to \Omega$, which corresponds to a subobject of X. Thus, we can think of $\mathscr{P}X$ as another name for Sub(X), although Sub(X) is not really an object in **Sets**.

6 Cardinality

Summary: When mathematics was rigorized in the 19th century, one of the important advances was a rigorous definition of "infinite set." It came as something of a suprise that there are different sizes of infinity, and that some infinite sets (e.g. the real numbers) are strictly larger than the natural numbers. In this section, we define "finite" and "infinite." We then add an axiom which says there is a specific set N that behaves like the natural numbers; in particular, N is infinite. Finally, we show that the powerset $\mathscr{P}X$ of a set X is always larger than X.

Definition. A set X is said to be **finite** if and only if for any function $m : X \to X$, if m is monic, then m is an isomorphism. A set X is said to be **infinite** if and only if there is a function $m : X \to X$ that is monic and not surjective.

We are already guaranteed the existence of finite sets: for example, the terminal object 1 is finite, as is the subobject classifier Ω. But the axioms we have stated thus far do not guarantee the existence of any infinite sets. We won't know that there are infinite sets until we add the "natural number object" axiom below.

Definition. We say that Y is at least as large as X, written $|X| \leq |Y|$, just in case there is a monomorphism $m : X \to Y$.

Proposition 6.1. $|X| \leq |X \amalg Y|$.

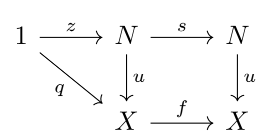
Proof. Proposition 4.2 shows that $i_0 : X \to X \amalg Y$ is monic. □

Proposition 6.2. *If Y is non-empty, then* $|X| \leq |X \times Y|$.

Proof. Consider the function $\langle 1_X, f\rangle : X \to X \times Y$, where $f : X \to 1 \to Y$. □

Axiom 10: Natural Number Object

There is an object N, and functions $z : 1 \to N$ and $s : N \to N$ such that for any other set X with functions $q : 1 \to X$ and $f : X \to X$, there is a unique function $u : N \to X$ such that the following diagram commutes:

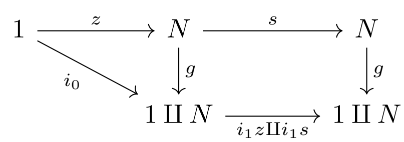


The set N is called a **natural number object**.

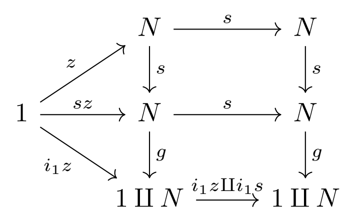
Exercise 6.3. Let N' be a set, and let $z' : 1 \to N'$ and $s' : N' \to N'$ be functions that satisfy the conditions in the axiom above. Show that N' is isomorphic to N.

Proposition 6.4. *$z \amalg s : 1 \amalg N \to N$ is an isomorphism.*

Proof. Let $i_0 : 1 \to 1 \amalg N$ and $i_1 : N \to 1 \amalg N$ be the coproduct inclusions. Using the NNO axiom, there is a unique function $g : N \to 1 \amalg N$ such that the following diagram commutes:



We will show that g is a two-sided inverse of $z \amalg s$. To this end, we first establish that $g \circ s = i_1$. Consider the following diagram:



The lower triangle commutes because of the commutativity of the previous diagram. Thus, the entire diagram commutes. The outer triangle and square would also commute with i_1 in place of $g \circ s$. By the NNO axiom, $g \circ s = i_1$. Now, to see that $(z \amalg s) \circ g = 1_N$, note first that

$$(z \amalg s) \circ g \circ z = (z \amalg s) \circ i_0 = z.$$

Furthermore,

$$(z \amalg s) \circ g \circ s = (z \amalg s) \circ i_1 = s.$$

Thus the NNO axiom entails that $(z \amalg s) \circ g = \mathrm{id}_N$. Finally, to see that $g \circ (z \amalg s) = \mathrm{id}_{1 \amalg N}$, we calculate:

$$g \circ (z \amalg s) \circ i_0 = g \circ z = i_0.$$

Furthermore,

$$g \circ (z \amalg s) \circ i_1 = g \circ s = i_1.$$

Therefore, $g \circ (z \amalg s) = \mathrm{id}_{1 \amalg N}$. This establishes that g is a two-sided inverse of $z \amalg s$, and $1 \amalg N$ is isomorphic to N. □

Proposition 6.5. *The function $s : N \to N$ is injective, but not surjective. Thus, N is infinite.*

Proof. By Proposition 4.2, the function $i_1 : N \to 1 \amalg N$ is monic. Since the images of i_0 and i_1 are disjoint, i_0 is not surjective. Since $z \amalg s$ is an isomorphism, $(z \amalg s) \circ i_1 = s$ is monic, but not surjective. Therefore, N is infinite. $\square$

Proposition 6.6. *If $m : B \to X$ is a nonempty subobject, then there is an epimorphism $f : X \to B$.*

Proof. Since B is nonempty, there is a function $g : X \backslash B \to B$. By Proposition 4.5, $B \cong B \amalg X \backslash B$. Finally, $1_B \amalg g : B \amalg X \backslash B \to B$ is an epimorphism, since 1_B is an epimorphism. $\square$

Definition. We say that a set X is **countable** just in case there is an epimorphism $f : N \to X$, where N is the natural numbers.

Proposition 6.7. *$N \times N$ is countably infinite.*

Sketch of proof. We will give two arguments: one quick, and one slow (but hopefully more illuminating). For the quick argument, define a function $g : N \times N \to N$ by $g(x, y) = 2^x 3^y$. If $\langle x, y\rangle \neq \langle x', y'\rangle$, then either $x \neq x'$ or $y \neq y'$. In either case, unique factorizability of integers gives $2^x 3^y \neq 2^{x'} 3^{y'}$. Therefore, $g : N \times N \to N$ is monic. Since $N \times N$ is nonempty, Proposition 6.6 entails that there is an epimorphism $f : N \twoheadrightarrow N \times N$. Therefore, $N \times N$ is countable.

Now for the slow argument. Imagine writing down all elements in $N \times N$ in an infinite table, whose first few elements look like this:

$$\left(\begin{array}{cccc} \langle 0,0\rangle & \langle 1,0\rangle & \langle 2,0\rangle & \cdots \\ \langle 0,1\rangle & \langle 1,1\rangle & \langle 2,1\rangle & \cdots \\ \langle 0,2\rangle & \langle 1,2\rangle & \langle 2,2\rangle & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{array} \right.$$

Now imagine running a thread diagonally through the numbers: begin with $\langle 0,0\rangle$, then move down to $\langle 1,0\rangle$ and up to $\langle 0,1\rangle$, then over to $\langle 2,0\rangle$ and down its diagonal, etc.. This process defines a function $f : N \to N \times N$ whose first few values are

$$\begin{array}{l} f(0) = \langle 0,0\rangle \\ f(1) = \langle 0,1\rangle \\ f(2) = \langle 1,0\rangle \\ \vdots \end{array}$$

It is not difficult to show that f is surjective, and so $N \times N$ is countable. $\square$

Exercise 6.8. Show that if A and B are countable then $A \cup B$ is countable.

We're now going to show that exponentiation allows us to construct sets of larger and larger size. In the case of finite sets A and X, it's easy to see that the following equation holds:

$$|A^X| \;=\; |A|^{|X|},$$

where $|X|$ denotes the number of elements in X. In particular, Ω^X can be thought of as the set of binary sequences indexed by X. We're now going to show that for any set X, the set Ω^X is larger than X.

Definition. Let $g : A \to A$ be a function. We say that $a \in A$ is a **fixed point** of g just in case $g(a) = a$. We say that A has the **fixed point property** just in case any function $g : A \to A$ has a fixed point.

Proposition 6.9. *Let A and X be sets. If there is a surjective function $p : X \to A^X$, then A has the fixed point property.*

Proof. Suppose that $p : X \to A^X$ is surjective. That is, for any function $f : X \to A$, there is an $x_f \in X$ such that $f = p(x_f)$. Let $\varphi = p^\flat$, so that $f = \varphi(x_f, -)$. Now let $g : A \to A$ be any function. We need to show that g has a fixed point. Consider the function $f : X \to A$ defined by $f = g \circ \varphi \circ \delta_X$, where $\delta_X : X \to X \times X$ is the diagonal map. Then we have

$$g\varphi(x, x) \;=\; f(x) \;=\; \varphi(x_f, x),$$

for all $x \in X$. In particular, $g\varphi(x_f, x_f) = \varphi(x_f, x_f)$, which means that $a = \varphi(x_f, x_f)$ is a fixed point of g. Since $g : A \to A$ was arbitrary, it follows that A has the fixed point property. □

Proposition 6.10. *There is no surjective function $X \to \Omega^X$.*

Proof. The function $\Omega \to \Omega$ that permutes $\mathtt{t}$ and $\mathtt{f}$ has no fixed points. The result then follows from Proposition 6.9. □

Exercise 6.11. Show that there is an injective function $X \to \Omega^X$. [The proof is easy if you simply think of Ω^X as functions from X to $\{\mathtt{t}, \mathtt{f}\}$. For a bigger challenge, try to prove that it's true using the definition of the exponential set Ω^X.]

Corollary 6.12. *For any set X, the set $\mathscr{P}X$ of its subsets is strictly larger than X.*

There are several other facts about cardinality that are important for certain parts of mathematics — in our case, they will be important for the study of topology. For example, if X is an infinite set, then the set $\mathscr{F}X$ of all finite subsets of a set X has the same cardinality as X. Similarly, a countable coproduct of countable sets is countable. However, these facts — well known from ZF set theory — are not obviously provable in ETCS.

Discussion. Intuitively speaking, X^N is the set of all sequences with values in X. Thus, we should have something like

$$X^N \;\cong\; X \times X \times \cdots$$

However, we don't have any axiom telling us that **Sets** has infinite products such as the one on the right hand side above. Can it be proven that X^N satisfies the definition of an infinite product? In other words, are there projections $\pi_i : X^N \to X$ which satisfy an appropriate universal property?

7 The Axiom of Choice

Definition. Let $f : X \to Y$ be a function. We say that f is a **split epimorphism** just in case there is a function $s : Y \to X$ such that $fs = 1_Y$. In this case, we say that s is a **section** of f.

Exercise 7.1. Prove that if f is a split epimorphism, then f is a regular epimorphism.

Exercise 7.2. Prove that if s is a section, then s is a regular monomorphism.

Axiom 11: Axiom of choice

Every epimorphism in **Sets** has a section.

The name "axiom of choice" comes from a different formulation of this axiom, which says (roughly speaking) that for any set-indexed collection of sets, say $\{X_i \mid i \in I\}$, we can choose one element from each set, say $x_i \in X_i$, and form a new set with these elements.

To translate that version of the axiom of choice into our version, suppose that the sets X_i are stacked side by side, and that f is the map that projects each $x \in X_i$ to the value i. Then a section s of f would be a map with domain I that returns an element $s(i) \in X_i$ for each $i \in I$.

Further reading

- C. Butz, "Regular categories and regular logic" `http://www.brics.dk/LS/98/2/BRICS-LS-98-2.pdf`
- W. Lawvere, *Sets for Mathematics*. Cambridge University Press, 2003.
- W. Lawvere, "An elementary theory of the category of sets" `http://www.tac.mta.ca/tac/reprints/articles/11/tr11.pdf`
- T. Leinster, "Rethinking set theory" `https://arxiv.org/abs/1212.6543`
- ETCS in nlab `https://ncatlab.org/nlab/show/ETCS`