CHAPTER 3

ARROWS INSTEAD OF EPSILON

> "*The world of ideas is not revealed to us in one stroke; we must both permanently and unceasingly recreate it in our consciousness*".
>
> René Thom

In this chapter we examine a number of standard set-theoretic constructions and reformulate them in the language of arrows. The general theme, as mentioned in the introduction, is that concepts defined by reference to the "internal" membership structure of a set are to be characterised "externally" by reference to connections with other sets, these connections being established by functions. The analysis will eventually lead us to the notions of *universal property* and *limit*, which encompass virtually all constructions within categories.

3.1. Monic arrows

A set function $f: A \to B$ is said to be *injective*, or *one-one* when no two distinct inputs give the same output, i.e. for inputs $x, y \in A$,

$$\text{if} \quad f(x) = f(y), \quad \text{then} \quad x = y.$$

Now let us take an injective $f: A \to B$ and two "parallel" functions $g, h: C \rightrightarrows A$ for which

$$\begin{tikzcd} C \arrow[r, "g"] \arrow[d, "h"'] & A \arrow[d, "f"] \\ A \arrow[r, "f"] & B \end{tikzcd}$$

commutes, i.e. $f \circ g = f \circ h$.

Then for $x \in C$, we have $f \circ g(x) = f \circ h(x)$, i.e. $f(g(x)) = f(h(x))$. But as f is injective, this means that $g(x) = h(x)$. Hence g and h, giving the same output for every input, are the same function, and we have shown that an injective f is "left-cancellable", i.e.

$$\textit{whenever} \quad f \circ g = f \circ h, \quad \text{then} \quad g = h.$$

On the other hand, if f has this left-cancellation property, it must be injective. To see this, take x and y in A, with $f(x) = f(y)$.

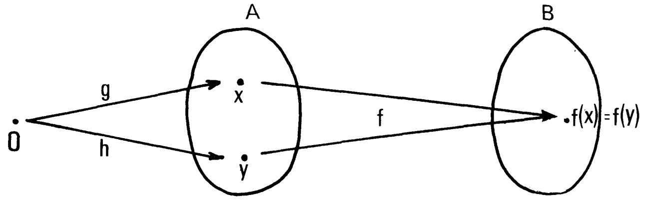

Fig. 3.1

The instructions "$g(0) = x$", "$h(0) = y$" establishes a pair of functions g, h from $\{0\}$ (i.e. the ordinal 1) to A for which we have $f \circ g = f \circ h$. By left cancellation, $g = h$, so $g(0) = h(0)$, i.e. $x = y$.

We thus see that the injective arrows in **Set** are precisely the ones that are left cancellable. The point of all this is that the latter property is formulated entirely by reference to arrows and leads to the following abstract definition:

An arrow $f : a \to b$ in a category $\mathscr{C}$ is *monic* in $\mathscr{C}$ if for any parallel pair $g, h : c \rightrightarrows a$ of $\mathscr{C}$-arrows, the equality $f \circ g = f \circ h$ implies that $g = h$. The symbolism $f : a \rightarrowtail b$ is used to indicate that f is monic. The name comes from the fact that an injective algebraic homomorphism (i.e. an arrow in a category like **Mon** or **Grp**) is called a "monomorphism".

EXAMPLE 1. In the category **N** (Example 6, Chapter 2) *every arrow is monic.* Left-cancellation here means that

$$\text{if} \quad m + n = m + p, \quad \text{then} \quad n = p$$

which is certainly a true statement about addition of numbers.

EXAMPLE 2. In a pre-order, every arrow is monic: given a pair $g, h : c \rightrightarrows a$, we must have $g = h$, as there is at most one arrow $c \to a$.

EXAMPLE 3. In **Mon, Grp, Met, Top** the monics are those arrows that are injective as set functions (see e.g. Arbib and Manes [75]).

EXAMPLE 4. In a comma category $\mathscr{C} \downarrow a$, an arrow k from (b, f) to (c, g),

$$\begin{tikzcd} b \arrow[rr, "k"] \arrow[dr, "f"'] & & c \arrow[dl, "g"] \\ & a & \end{tikzcd}$$

is monic in $\mathscr{C} \downarrow a$ iff k is monic in $\mathscr{C}$ as an arrow from b to c.

Exercises

In any category
(1). $g \circ f$ is monic if both f and g are monic.
(2) If $g \circ f$ is monic then so is f.

3.2. Epic arrows

A set function $f: A \to B$ is *onto*, or *surjective* if the codomain B is the range of f, i.e. for each $y \in B$ there is some $x \in A$ such that $y = f(x)$, i.e. every member of B is an output for f. The "arrows-only" definition of this concept comes from the definition of "monic" by simply reversing the arrows. Formally:

An arrow $f: a \to b$ is *epic* (right-cancellable) in a category $\mathscr{C}$ if for any pair of $\mathscr{C}$-arrows $g, h: b \rightrightarrows c$, the equality $g \circ f = h \circ f$ implies that $g = h$, i.e. whenever a diagram

$$\begin{tikzcd} a \arrow[r, "f"] \arrow[d, "f"'] & b \arrow[d, "g"] \\ b \arrow[r, "h"'] & c \end{tikzcd}$$

commutes, then $g = h$. The notation $f: a \twoheadrightarrow b$ is used for epic arrows.

In **Set**, the epic arrows are precisely the surjective functions (exercise for the reader, or Arbib and Manes, p. 2). A surjective homomorphism is known as an *epimorphism.*

In the category **N**, every arrow is epic, as $n + m = p + m$ implies that $n = p$. In any pre-order, all arrows are epic.

In the categories of our original list, where arrows are functions, the arrows that are surjective as functions are always epic. The converse is true in **Grp**, but not in **Mon**. The inclusion of the natural numbers into the integers is a monoid homomorphism (with respect to $+$), that is certainly not onto, but nevertheless is right cancellable in **Mon**. (Arbib and Manes p. 57).

3.3. Iso arrows

A function that is both injective and surjective is called *bijective.* If $f: A \rightarrowtail\!\!\!\!\rightarrow B$ is bijective then the passage from A to B under f can be reversed or "inverted". We can think of f as being simply a "relabelling" of A. Any $b \in B$ is the image $f(a)$ of some $a \in A$ (surjective property) and in fact is the image of only one such a (injective property). Thus the rule which assigns to b this unique a, i.e. has

$$g(b) = a \quad \text{iff} \quad f(a) = b$$

establishes a function $B \to A$ which has

$$g(f(a)) = a, \quad \text{all} \quad a \in A$$

and

$$f(g(b)) = b, \quad \text{all} \quad b \in B.$$

Hence

$$g \circ f = id_A$$

and

$$f \circ g = id_B.$$

A function that is related to f in this way is said to be an *inverse* of f. This is an essentially arrow-theoretic idea, and leads to a new definition.

A $\mathscr{C}$-arrow $f : a \to b$ is *iso*, or *invertible*, in $\mathscr{C}$ if there is a $\mathscr{C}$-arrow $g : b \to a$, such that $g \circ f = \mathbf{1}_a$ and $f \circ g = \mathbf{1}_b$.

There can in fact be at most one such g, for if $g' \circ f = \mathbf{1}_a$, and $f \circ g' = \mathbf{1}_b$, then $g' = \mathbf{1}_a \circ g' = (g \circ f) \circ g' = g \circ (f \circ g') = g \circ \mathbf{1}_b = g$. So this g, when it exists, is called *the inverse of* f, and denoted by $f^{-1} : b \to a$. It is defined by the conditions $f^{-1} \circ f = \mathbf{1}_a$, $f \circ f^{-1} = \mathbf{1}_b$. The notation $f : a \cong b$ is used for iso's.

An iso arrow is always monic. For if $f \circ g = f \circ h$, and f^{-1} exists, then $g = \mathbf{1}_a \circ g = (f^{-1} \circ f) \circ g = f^{-1} \circ (f \circ g) = f^{-1} \circ (f \circ h) = \mathbf{1}_a \circ h = h$, and so f is left-cancellable. An analogous argument shows that iso's are always epic.

Now in **Set** a function that is epic and monic has an inverse, as we saw at the beginning of this section. So in **Set**, "iso" is synonymous with "monic and epic". The same, we shall learn, goes for any topos, but is certainly not so in all categories.

In the category **N** we already know that *every* arrow is both monic and epic. But the only iso is $0 : N \to N$. For if m has inverse n, $m \circ n = \mathbf{1}_N$, i.e. $m + n = 0$. Since m and n are both natural numbers, hence both non-negative, this can only happen if $m = n = 0$.

The inclusion map mentioned at the end of the last section is in fact epic and monic, but cannot be iso, since if it had an inverse it would, as a set function, be bijective.

In a poset category $\mathbf{P}=(P,\sqsubseteq)$, if $f:p\to q$ has an inverse $f^{-1}:q\to p$, then $p\sqsubseteq q$ and $q\sqsubseteq p$, whence by antisymmetry, $p=q$. But then f must be the unique arrow $\mathbf{1}_p$ from p to p. Thus in a poset, every arrow is monic and epic, but the only iso's are the identities.

Groups

A *group* is a monoid $(M, *, e)$ in which for each $x\in M$ there is a $y\in M$ satisfying $x*y=e=y*x$. There can in fact be only one such y for a given x. It is called the inverse of x, and denoted x^{-1}. Thinking of a monoid as a category with one object, the terminology and notation is tied to its above usage: a group is essentially the same thing as a one-object category in which every arrow is iso.

EXERCISE 1. Every identity arrow is iso.

EXERCISE 2. If f is iso, so is f^{-1}.

EXERCISE 3. $f\circ g$ is iso if f, g are, with $(f\circ g)^{-1}=g^{-1}\circ f^{-1}$.

3.4. Isomorphic objects

Objects a and b are *isomorphic* in $\mathscr{C}$, denoted $a\cong b$, if there is a $\mathscr{C}$-arrow $f:a\to b$ that is iso in $\mathscr{C}$, i.e. $f:a\cong b$.

In **Set**, $A\cong B$ when there is a bijection between A and B, in which case each set can be thought of as being a "relabelling" of the other. As a specific example take a set A and put

$$B=A\times\{0\}=\{\langle x, 0\rangle : x\in A\}.$$

In effect B is just A with the label "0" attached to each of its elements. The rule $f(x)=\langle x, 0\rangle$ gives the bijection $f:A\to B$ making $A\cong B$.

In **Grp**, two groups are isomorphic if there is a group homomorphism (function that "preserves" group structure) from one to the other whose set-theoretic inverse exists and is a group homomorphism (hence is present in **Grp** as an inverse). Such an arrow is called a *group isomorphism.*

In **Top**, isomorphic topological spaces are usually called *homeomorphic.* This means there is a homeomorphism between them, i.e. a continuous bijection whose inverse is also continuous.

In these examples, isomorphic objects "look the same". One can pass freely from one to the other by an iso arrow and its inverse. Moreover these arrows, which establish a "one-one correspondence" or "matching" between the elements of the two objects, preserve any relevant structure. This means that we can replace some or all of the members of one object by their counterparts in the other object without making any difference to the structure of the object, to its appearance. Thus isomorphic groups look exactly the same, *as groups;* homeomorphic topological spaces are indistinguishable by any topological property, and so on. Within any mathematical theory, isomorphic objects are indistinguishable in terms of that theory. The aim of that theory is to identify and study constructions and properties that are "invariant" under the isomorphisms of the theory (thus topology studies properties that are not altered or destroyed when a space is replaced by another one homeomorphic to it). An object will be said to be "unique up to isomorphism" in possession of a particular attribute if the only other objects possessing that attribute are isomorphic to it. A concept will be "defined up to isomorphism" if its description specifies a particular entity, not uniquely, but only uniquely up to isomorphism.

Category theory then is the subject that provides an abstract formulation of the idea of mathematical isomorphism and studies notions that are invariant under all forms of isomorphism. In category theory, "is isomorphic to" is virtually synonymous with "is". Indeed most of the basic definitions and constructions that one can perform in a category do not specify things uniquely at all, but only, as we shall see, "up to isomorphism".

Skeletal categories

A skeletal category is one in which "isomorphic" does actually mean the same as "is", i.e. in which whenever $a \cong b$, then $a = b$. We saw in the last section that in a poset, the only iso arrows are the identities. This then gives us a categorial account of antisymmetry in pre-orders. A poset is precisely a *skeletal pre-order category.*

Exercise 1. For any $\mathscr{C}$-objects

(i) $a \cong a$;

(ii) if $a \cong b$ then $b \cong a$;

(iii) If $a \cong b$ and $b \cong c$, then $a \cong c$.

Exercise 2. **Finord** is a skeletal category.

3.5. Initial objects

What arrow properties distinguish $\emptyset$, the null set, in **Set**? Given a set A, can we find any function $\emptyset \to A$? Recalling our formulation of a function as a triple $\langle A, B, X\rangle$ with $X \subseteq A \times B$ (§2.1), we find by checking the details of that definition that $f = \langle \emptyset, A, \emptyset\rangle$ is a function from $\emptyset \to A$. The graph of f is empty, and f is known as the *empty function* for A. Since $\emptyset \times A$ is empty, $\emptyset$ is the only subset of $\emptyset \times A$, and hence f is the only function from $\emptyset$ to A. This observation leads us to the following:

DEFINITION. An object 0 is *initial* in category $\mathscr{C}$ if for every $\mathscr{C}$-object a there is one and only one arrow from 0 to a in $\mathscr{C}$.

Any two initial $\mathscr{C}$-objects must be isomorphic in $\mathscr{C}$. For if 0, $0'$ are such objects there are unique arrows $f: 0' \to 0$, $g: 0 \to 0'$. But then $f \circ g: 0 \to 0$ must be 1_0, as 1_0 is the *only* arrow $0 \to 0$, 0 being initial. Similarly, as $0'$ is initial, $g \circ f: 0' \to 0'$ is $1_{0'}$. Thus f has an inverse (g), and $f: 0' \cong 0$.

The symbol 0 of course is used because in **Set** it is a name for $\emptyset$, and $\emptyset$ is initial in **Set**. In fact $\emptyset$ is the only initial object in **Set**, so whereas the initial $\mathscr{C}$-object may only be "unique up to isomorphism", when $\mathscr{C} = \textbf{Set}$ it is actually unique.

In a pre-order $(P, \sqsubseteq)$ an initial object is an element $0 \in P$ with $0 \sqsubseteq p$ for all $p \in P$ (i.e. a *minimal* element). In a poset, where "isomorphic" means "equal", then there can be at most one initial object (the *minimum*, or *zero* element). Thus in the poset $\{0, \ldots, n-1\}$, 0 is *the* initial object, whereas in the two-object category with diagram

$$\begin{tikzcd} \bullet \arrow[loop left] \arrow[r, bend left] & \bullet \arrow[l, bend left] \arrow[loop right] \end{tikzcd}$$

both objects are initial.

In **Grp**, and **Mon**, an initial object is any one element algebra $(M, *, e)$, i.e. $M = \{e\}$, and $e * e = e$. Each of these categories has infinitely many initial objects.

In $\textbf{Set}^2$, the category of pairs of sets, the initial object is $\langle \emptyset, \emptyset\rangle$, while in $\textbf{Set}^{\to}$, the category of functions, it is $\langle \emptyset, \emptyset, \emptyset\rangle$, the empty function from $\emptyset$ to $\emptyset$. In $\textbf{Set}{\downarrow}\mathbb{R}$, the category of real valued functions, it is $f = \langle \emptyset, \mathbb{R}, \emptyset\rangle$. Given $g: A \to \mathbb{R}$, the only way to make the diagram

$$\begin{tikzcd} \emptyset \arrow[rr, "k"] \arrow[dr, "f"'] & & A \arrow[dl, "g"] \\ & \mathbb{R} & \end{tikzcd}$$

commute is to put $k = \langle \emptyset, A, \emptyset \rangle$, the empty map from $\emptyset$ to A.

Notation. The exclamation mark "!" is often used to denote a uniquely existing arrow. We put $! : 0 \to a$ for the unique arrow from 0 to a. It is also denoted 0_a, i.e. $0_a : 0 \to a$.

3.6. Terminal objects

By reversing the direction of the arrows in the definition of initial object, we have the following idea:

Definition. An object 1 is *terminal* in a category $\mathscr{C}$ if for every $\mathscr{C}$-object a there is one and only one arrow from a to 1 in $\mathscr{C}$.

In **Set**, the terminal objects are the singletons, i.e. the one-element sets $\{e\}$. Given set A, the rule $f(x) = e$ gives a function $f : A \to \{e\}$. Since e is the only possible output, this is the only possible such function. Thus **Set** has many terminal objects. They are all isomorphic (terminal objects in any category are isomorphic) and the paradigm is the ordinal $1 = \{0\}$, whence the notation.

Again we may write $! : a \to 1$ to denote the unique arrow from a to 1, or alternatively $!_a : a \to 1$.

In a pre-order a terminal object satisfies $p \sqsubseteq 1$, all p (a maximal element). In a poset, 1 is unique (the *maximum*), when it exists, and is also called the *unit* of **P**.

In **Grp** and **Mon**, terminal objects are again the one element monoids. Hence the initial objects are the same as the terminal ones (and so the equation $0 = 1$ is "true up to isomorphism"). An object that is both initial and terminal is called a *zero* object. **Set** has no zero's. The fact that **Grp** and **Mon** have zeros precludes them, as we shall see, from being topoi.

In $\mathbf{Set}\downarrow\mathbb{R}$, $(\mathbb{R}, \mathrm{id}_{\mathbb{R}})$ is a terminal object. Given (A, f), the only way to make

$$\begin{tikzcd} A \arrow[rr, "k"] \arrow[dr, "f"'] & & \mathbb{R} \arrow[dl, "\mathrm{id}_{\mathbb{R}}"] \\ & \mathbb{R} & \end{tikzcd}$$

commute is to put $k = f$.

Exercise 1. Prove that all terminal $\mathscr{C}$-objects are isomorphic.

EXERCISE 2. Find terminals in $\mathbf{Set}^2$, $\mathbf{Set}^{\rightarrow}$, and the poset $\mathbf{n}$.

EXERCISE 3. Show that an arrow $1 \rightarrow a$ whose domain is a terminal object must be monic.

3.7. Duality

We have already observed that the notion of epic arrow arises from that of monic by "reversing the arrows". The same applies to the concepts of terminal and initial objects. These are two examples of the notion of *duality* in category theory, which we will now describe a little more precisely.

If Σ is a statement in the basic language of categories, the *dual* of Σ, Σ^{op}, is the statement obtained by replacing "dom" by "cod", "cod" by "dom", and "$h = g \circ f$" by "$h = f \circ g$". Thus all arrows and composites referred to by Σ are reversed in Σ^{op}. The notion or construction described by Σ^{op} is said to be *dual* to that described by Σ. Thus the notion of epic arrow is dual to that of monic arrow. The dual of "initial object" is "terminal object", and so on.

From a given category $\mathscr{C}$ we construct its *dual* or *opposite* category $\mathscr{C}^{\mathrm{op}}$ as follows:

$\mathscr{C}$ and $\mathscr{C}^{\mathrm{op}}$ have the same objects. For each $\mathscr{C}$-arrow $f: a \rightarrow b$ we introduce an arrow $f^{\mathrm{op}}: b \rightarrow a$ in $\mathscr{C}^{\mathrm{op}}$, these being all and only the arrows in $\mathscr{C}^{\mathrm{op}}$. The composite $f^{\mathrm{op}} \circ g^{\mathrm{op}}$ is defined precisely when $g \circ f$ is defined in $\mathscr{C}$ and has

$$\begin{tikzcd} a \arrow[r, "f"] & b \arrow[l, bend left, "f^{\mathrm{op}}"] \arrow[r, "g"] & c \arrow[l, bend left, "g^{\mathrm{op}}"] \end{tikzcd}$$

$f^{\mathrm{op}} \circ g^{\mathrm{op}} = (g \circ f)^{\mathrm{op}}$. Note that $\mathrm{dom}\, f^{\mathrm{op}} = \mathrm{cod}\, f$, and $\mathrm{cod}(f^{\mathrm{op}}) = \mathrm{dom}\, f$.

EXAMPLE 1. If $\mathscr{C}$ is discrete, $\mathscr{C}^{\mathrm{op}} = \mathscr{C}$.

EXAMPLE 2. If $\mathscr{C}$ is a pre-order (P, R), with $R \subseteq P \times P$, then $\mathscr{C}^{\mathrm{op}}$ is the pre-order (P, R^{-1}), where $pR^{-1}q$ iff qRp, i.e. R^{-1} is the inverse relation to R.

EXAMPLE 3. For any $\mathscr{C}$, $(\mathscr{C}^{\mathrm{op}})^{\mathrm{op}} = \mathscr{C}$.

The dual of a construction expressed by Σ can be interpreted as the original construction applied to the opposite category. If Σ is true of $\mathscr{C}$, Σ^{op} will be true of $\mathscr{C}^{\mathrm{op}}$. Thus the initial object $\emptyset$ in **Set** is the terminal object of $\mathbf{Set}^{\mathrm{op}}$. Now if Σ is a theorem of category theory, i.e. derivable from the category axioms, then Σ will be true in all categories. Hence Σ^{op} will hold in all categories of the form $\mathscr{C}^{\mathrm{op}}$. But any category $\mathscr{D}$ has this form (put $\mathscr{C} = \mathscr{D}^{\mathrm{op}}$), and so Σ^{op} holds in all categories. Thus from any true statement Σ of category theory we immediately obtain another true statement Σ^{op} by this *Duality Principle.*

The Duality Principle cuts the number of things to be proven in half. For example, we note first that the concept of iso arrow is self-dual. The dual of an invertible arrow is again an invertible arrow – indeed $(f^{\mathrm{op}})^{-1} = (f^{-1})^{\mathrm{op}}$. So having proven

any two initial $\mathscr{C}$-objects are isomorphic

we can conclude without further ado, the dual fact that

any two terminal $\mathscr{C}$-objects are isomorphic.

The Duality Principle comes from the domain of logic. It is discussed in a more rigorous fashion in Hatcher [68] §8.2.

3.8. Products

We come now to the problem of giving a characterisation, using arrows, of the product set

$$A \times B = \{\langle x, y\rangle : x \in A \text{ and } y \in B\}$$

of two sets A and B. The uninitiated may find it hard to believe that this can be achieved without any reference to ordered pairs. But in fact it can be, up to isomorphism, and the way it is done will lead us to a general description of what a "construction" in a category is.

Associated with $A \times B$ are two special maps, the *projections*

$$p_A : A \times B \to A$$

and

$$p_B : A \times B \to B$$

given by the rules

$$p_A(\langle x, y\rangle) = x$$

$$p_B(\langle x, y\rangle) = y.$$

Now suppose we are given some other set C with a pair of maps $f: C \to A$, $g: C \to B$, Then we define $p: C \to A \times B$

$$\begin{tikzcd} & C \arrow[dl, "f"'] \arrow[d, dashed, "p"] \arrow[dr, "g"] & \\ A & A \times B \arrow[l, "p_A"] \arrow[r, "p_B"'] & B \end{tikzcd}$$

by the rule $p(x) = \langle f(x), g(x)\rangle$. Then we have $p_A(p(x)) = f(x)$, and $p_B(p(x)) = g(x)$ for all $x \in C$, so $p_A \circ p = f$ and $p_B \circ p = g$, i.e. the above diagram commutes. Moreover, p as defined is the only arrow that can make the diagram commute. For if $p(x) = \langle y, z\rangle$ then simply knowing that $p_A \circ p = f$ tells us that $p_A(p(x)) = f(x)$, i.e. $y = f(x)$. Similarly if $p_B \circ p = g$, we must have $z = g(x)$.

The map p associated with f and g is usually denoted $\langle f, g\rangle$, the *product map* of f and g. Its definition in **Set** is $\langle f, g\rangle(x) = \langle f(x), g(x)\rangle$.

The observations just made motivate the following:

DEFINITION. A *product* in a category $\mathscr{C}$ of two objects a and b is a $\mathscr{C}$-object $a \times b$ together with a pair $(pr_a: a \times b \to a,\ pr_b: a \times b \to b)$ of $\mathscr{C}$-arrows such that for *any* pair of $\mathscr{C}$-arrows of the form $(f: c \to a, g: c \to b)$ there is exactly one arrow $\langle f, g\rangle: c \to a \times b$ making

$$\begin{tikzcd} & c \arrow[dl, "f"'] \arrow[d, dashed, "\langle f, g\rangle"] \arrow[dr, "g"] & \\ a & a \times b \arrow[l, "pr_a"] \arrow[r, "pr_b"'] & b \end{tikzcd}$$

commute, i.e. such that $pr_a \circ \langle f, g\rangle = f$ and $pr_b \circ \langle f, g\rangle = g$. $\langle f, g\rangle$ is the *product arrow of f and g with* respect to the *projections* pr_a, pr_b.

Notice that we said *a* product of a and b, not *the product.* This is because $a \times b$ *is only defined up to isomorphism.* For suppose $(p: d \to a, q: d \to b)$ also satisfies the definition of "a product of $a \times b$" and consider the diagram

$$\begin{tikzcd}
 & d \arrow[dl, "p"'] \arrow[dr, "q"] \arrow[d, dashed, "\langle p, q\rangle"] & \\
a & a \times b \arrow[l, "pr_a"'] \arrow[r, "pr_b"] \arrow[d, dashed, "\langle pr_a, pr_b\rangle"] & b \\
 & d \arrow[ul, "p"] \arrow[ur, "q"'] &
\end{tikzcd}$$

$\langle p, q\rangle$ is the unique product arrow of p and q with respect to "the" product $a \times b$. $\langle pr_a, pr_b\rangle$ is the unique product arrow of pr_a and pr_b with respect to "the" product d.

Now, since d is a product of a and b there can be only one arrow $s : d \to d$ such that

$$\begin{tikzcd}
 & d \arrow[dl, "p"'] \arrow[dr, "q"] \arrow[d, dashed, "s"] & \\
a & d \arrow[l, "p"] \arrow[r, "q"'] & b
\end{tikzcd}$$

commutes. But putting $s = 1_d$ makes this diagram commute, while the commutativity of the previous diagram implies that putting $s = \langle pr_a, pr_b\rangle \circ \langle p, q\rangle$ also works (more fully – $p \circ \langle pr_a, pr_b\rangle \circ \langle p, q\rangle = pr_a \circ \langle p, q\rangle = p$ etc.). By the uniqueness of s we must conclude

$$\langle pr_a, pr_b\rangle \circ \langle p, q\rangle = 1_d.$$

Interchanging the roles of d and $a \times b$ in this argument leads to $\langle p, q\rangle \circ \langle pr_a, pr_b\rangle = 1_{a\times b}$. Thus $\langle p, q\rangle : d \cong a \times b$, so the two products are isomorphic and furthermore the iso $\langle p, q\rangle$ when composed with the projections for $a \times b$ produces the projections for d, as the last diagram but one indicates. Indeed, $\langle p, q\rangle$ is the only arrow with this property.

In summary then our definition characterises the product of a and b "uniquely up to a unique commuting isomorphism", which is enough from the categorial viewpoint.

Example 1. In **Set**, **Finset**, **Nonset**, the product of A and B is the Cartesian product set $A \times B$.

EXAMPLE 2. In **Grp** the product of two objects is the standard direct product of groups, with the binary operation defined "component-wise" on the product set of the two groups.

EXAMPLE 3. In **Top**, the product is the standard notion of product space.

EXAMPLE 4. In a pre-order $(P, \sqsubseteq)$ a product of p and q when it exists is defined by the properties

(i) $p \times q \sqsubseteq p$, $p \times q \sqsubseteq q$, i.e. $p \times q$ is a "lower bound" of p and q;

(ii) if $c \sqsubseteq p$ and $c \sqsubseteq q$, then $c \sqsubseteq p \times q$, i.e. $p \times q$ is "greater" than any other lower bound of p and q.

In other words $p \times q$ is a *greatest lower bound* (g.l.b.) of p and q. In a poset, being skeletal, the g.l.b. is unique, when it exists, and will be denoted $p \sqcap q$. A poset in which every two elements have a g.l.b. is called a *lower semilattice.* Categorially a lower semilattice is a skeletal pre-order category in which any two objects have a product.

EXAMPLE 5. If A and B are finite sets, with say m and n elements respectively, then the product set $A \times B$ has $m \times n$ elements (where the last "$\times$" denotes multiplication). This has an interesting manifestation in the skeletal category **Finord**. There the product of the ordinal numbers m and n exists and is quite literally the ordinal $m \times n$.

EXERCISE 1. $\langle pr_a, pr_b \rangle = 1_{a \times b}$

$$\begin{tikzcd}
 & b \\
a \times b \arrow[ur, "pr_b"] \arrow[dr, "pr_a"'] \arrow[r, dashed, "\langle pr_a, pr_b \rangle"] & a \times b \arrow[u, "pr_b"'] \arrow[d, "pr_a"] \\
 & a
\end{tikzcd}$$

EXERCISE 2. If $\langle f, g \rangle = \langle k, h \rangle$, then $f = k$ and $g = h$.

Exercise 3. $\langle f \circ h,\ g \circ h\rangle = \langle f, g\rangle \circ h$

$$\begin{tikzcd}
 & & b \\
d \arrow[r, "h"] \arrow[urr, bend left, "g \circ h"] \arrow[drr, bend right, "f \circ h"'] & c \arrow[r, dashed, "\langle f, g\rangle"] \arrow[ur, "g"] \arrow[dr, "f"'] & a \times b \arrow[u, "pr_b"'] \arrow[d, "pr_a"] \\
 & & a
\end{tikzcd}$$

Exercise 4. We saw earlier that in **Set**, $A \cong A \times \{0\}$. Show that if category $\mathscr{C}$ has a terminal object 1 and products, then for any $\mathscr{C}$-object a, $a \cong a \times 1$ and indeed $\langle 1_a, !_a\rangle$ is iso

$$\begin{tikzcd}
 & a \arrow[dl, "1_a"'] \arrow[d, "\langle 1_a, !_a\rangle"] \arrow[dr, "!_a"] & \\
a & a \times 1 \arrow[l, "pr_a"] \arrow[r, "pr_1"'] & 1
\end{tikzcd}$$

Product maps

Given set functions $f: A \to B$, $g: C \to D$ we obtain a function from $A \times C$ to $B \times D$ that outputs $\langle f(x), g(y)\rangle$ for input $\langle x, y\rangle$. This map is denoted $f \times g$, and we have

$$f \times g(\langle x, y\rangle) = \langle f(x), g(y)\rangle,$$

It is not hard to see that $f \times g$ is just the product map of the two composites $f \circ p_A : A \times C \to A \to B$ and $g \circ p_C : A \times C \to C \to D$, so we can define the following.

Definition If $f: a \to b$ and $g: c \to d$ are $\mathscr{C}$-arrows then $f \times g: a \times b \to c \times d$ is the $\mathscr{C}$-arrow $\langle f \circ pr_a, g \circ pr_b\rangle$

$$\begin{tikzcd}
 & c \arrow[r, "g"] & d \\
a \times c \arrow[ur, "pr_c"] \arrow[dr, "pr_a"'] \arrow[rr, dashed, "\langle f \circ pr_a, g \circ pr_c\rangle"] & & b \times d \arrow[u] \arrow[d] \\
 & a \arrow[r, "f"] & b
\end{tikzcd}$$

(Of course $f \times g$ is only defined when $a \times c$ and $b \times d$ exist in $\mathscr{C}$).

EXERCISE 5. $1_a \times 1_b = 1_{a \times b}$

$$\begin{tikzcd}
 & b \arrow[r, "1_b"] & b \\
a \times b \arrow[ur] \arrow[dr] \arrow[rr, dashed, "1_a \times 1_b"] & & a \times b \arrow[u] \arrow[d] \\
 & a \arrow[r, "1_a"'] & a
\end{tikzcd}$$

EXERCISE 6. $a \times b \cong b \times a$.

EXERCISE 7. Show that $(a \times b) \times c \cong a \times (b \times c)$

$$\begin{tikzcd}
(a \times b) \times c \arrow[d] \arrow[ddrr] \arrow[rr, dash, dashed] \arrow[drr, dashed] & & a \times (b \times c) \arrow[d] \arrow[ddll] \arrow[dll, dashed] \\
(a \times b) \arrow[d] \arrow[dr] & & (b \times c) \arrow[d] \arrow[dl] \\
a & b & c
\end{tikzcd}$$

EXERCISE 8. Show that (i)

$$\begin{tikzcd}
 & c \arrow[r, "h"] & d \\
e \arrow[ur, "k"] \arrow[r, "\langle g, k \rangle"] \arrow[dr, "g"'] & a \times c \arrow[u] \arrow[d] \arrow[r, "f \times h"] & b \times d \arrow[u] \arrow[d] \\
 & a \arrow[r, "f"] & b
\end{tikzcd}$$

$(f \times h) \circ \langle g, k \rangle = \langle f \circ g, h \circ k \rangle$ and

(ii)

$$\begin{tikzcd}
e' \arrow[r, "k"] & c \arrow[r, "h"] & d \\
e \times e' \arrow[u] \arrow[d] \arrow[r, "g \times k"] & a \times c \arrow[u] \arrow[d] \arrow[r, "f \times h"] & b \times d \arrow[u] \arrow[d] \\
e \arrow[r, "g"] & a \arrow[r, "f"] & b
\end{tikzcd}$$

$(f\times h)\circ(g\times k)=(f\circ g)\times(h\circ k)$. □

The use we have been making of the broken arrow symbol $\dashrightarrow$ is a standard one in category theory. When present in any diagram it indicates that there is one and only one arrow that can occupy that position and allow the diagram to commute.

Finite products

Given sets A, B, C we extend the notion of product to define $A\times B\times C$ as the set of ordered *triples* $\langle x, y, z\rangle$. First elements come from A, second from B, and third from C. Thus $A\times B\times C=\{\langle x, y, z\rangle : x\in A, y\in B$, and $z\in C\}$. This idea can be extended to form the product of any finite sequence of sets $A_1, A_2, \ldots, A_m$. We define $A_1\times A_2\times\ldots\times A_m$ to be the set

$$\{\langle x_1, \ldots, x_m\rangle : x_1\in A_1, x_2\in A_2, \ldots, x_m\in A_m\}$$

of all "m-tuples", or "m-length sequences", whose "i-th" members come from A_i.

As a special case of this concept we have the *m-fold product* of a set A, as the set

$$A^m=\{\langle x_1, \ldots, x_m\rangle : x_1, x_2, \ldots, x_m\in A\}$$

of all m-tuples whose members all come from A. Associated with A^m are m different projection maps $pr_1^m, pr_2^m, \ldots, pr_m^m$ from A^m to A, given by the rules

$$pr_1^m(\langle x_1, \ldots, x_m \rangle) = x_1$$
$$pr_2^m(\langle x_1, \ldots, x_m \rangle) = x_2$$
$$\vdots$$
$$pr_m^m(\langle x_1, \ldots, x_m \rangle) = x_m$$

Given a set C and m maps $f_1 : C \to A, \ldots, f_m : C \to A$, we can then form a product map $\langle f_1, \ldots, f_m \rangle$ from C to A^m by stipulating, for input $c \in C$, that

$$\langle f_1, \ldots, f_m \rangle(c) = \langle f_1(c), f_2(c), \ldots, f_m(c) \rangle.$$

The construction just outlined can be developed in any category $\mathscr{C}$ that has products of any two $\mathscr{C}$-objects. For a given $\mathscr{C}$-object a, we define the m-fold product of a (with itself) to be

$$a^m = \underbrace{a \times a \times \ldots \times a}_{m\text{-copies}}$$

There is an ambiguity here. Should, for example, a^3 be taken as $(a \times a) \times a$ or $a \times (a \times a)$? However, Exercise 7 above allows us to gloss over this point, since these last two objects are isomorphic.

By applying the definition of products of pairs objects to the formation of a^m we may show that a^m has associated with it m projection arrows $pr_1^m : a^m \to a, \ldots, pr_m^m : a^m \to a$, with the universal property that for any $\mathscr{C}$-arrows $f_1 : c \to a, \ldots, f_m : c \to a$ with common domain, there is exactly one (product) arrow $\langle f_1, \ldots, f_m \rangle : c \to a^m$ making

$$\begin{tikzcd}
 & & c \arrow[dll, "f_1"'] \arrow[dl, "f_2"] \arrow[dd, dashed] \arrow[drr, "f_m"] & & \\
a & a & \cdots & & a \\
 & & a^m \arrow[ull, "pr_1^m"] \arrow[ul, "pr_2^m"'] \arrow[urr, "pr_m^m"'] & &
\end{tikzcd}$$

commute. For $m = 1$, we take a^1 to be just a, and $pr_1 : a \to a$ to be $\mathbf{1}_a$.

Finite products will play an important role in the "first-order" semantics of Chapter 11.

Exercise 9. Analyse in detail the formation of the projection arrows $pr_1^m, \ldots, pr_m^m$, and verify all assertions relating to the last diagram. Show that for any product arrow

$$c \xrightarrow{\langle f_1, \ldots, f_m \rangle} a^m,$$

we have $pr_j^m \circ \langle f_1, \ldots, f_m \rangle = f_j$, for $1 \leqslant j \leqslant m$.

EXERCISE 10. Develop the notion of the product $a_1 \times a_2 \times \ldots \times a_m$ of m objects (possibly different) and the product $f_1 \times f_2 \times \ldots \times f_m$ of m arrows (possibly different).

3.9. Co-products

The dual notion to "product" is the *co-product,* or *sum,* of objects, which by the duality principle we directly define as follows.

DEFINITION A *co-product* of $\mathscr{C}$-objects a and b is a $\mathscr{C}$-object $a+b$ together with a pair $i_a : a \to a+b$, $i_b : b \to a+b$) of $\mathscr{C}$-arrows such that for any pair of $\mathscr{C}$-arrows of the form $(f : a \to c,\ g : b \to c)$ there is exactly one arrow $[f, g] : a+b \to c$ making

$$\begin{tikzcd} a \arrow[r, "i_a"] \arrow[dr, "f"'] & a+b \arrow[d, dashed, "{[f, g]}"] & b \arrow[l, "i_b"'] \arrow[dl, "g"] \\ & c & \end{tikzcd}$$

commute, i.e. such that $[f, g] \circ i_a = f$ and $[f, g] \circ i_b = g$.

$[f, g]$ is called the *co-product* arrow of f and g with respect to the *injections* i_a and i_b.

In **Set**, the co-product of A and B is their *disjoint union,* $A+B$. This is the union of two sets that look the same as (i.e. are isomorphic to) A and B but are *disjoint* (have no elements in common). We put

$$A' = \{\langle a, 0 \rangle : a \in A\} = A \times \{0\}$$

and

$$B' = \{\langle b, 1 \rangle : b \in B\} = B \times \{1\}$$

(why does $A' \cap B' = \emptyset$?) and then define

$$A + B = A' \cup B'.$$

The injection $i_A : A \to A + B$ is given by the rule

$$i_A(a) = \langle a, 0 \rangle,$$

while $i_B : B \to A + B$ has $i_B(b) = \langle b, 1 \rangle$.

EXERCISE 1. Show that $A + B$, i_A, i_B as just defined satisfy the co-product definition. (First you will have to determine the rule for the function $[f, g]$ in this case.)

EXERCISE 2. If $A \cap B = \emptyset$, show $A \cup B \cong A + B$. □

In a pre-order $(P, \sqsubseteq)$, $p + q$ is defined by the properties

(i) $p \sqsubseteq p + q$, $q \sqsubseteq p + q$ (i.e. $p + q$ is an "upper bound" of p and q);

(ii) if $p \sqsubseteq c$ and $q \sqsubseteq c$, then $p + q \sqsubseteq c$, i.e. $p + q$ is "less than" any other upper bound of p and q.

In other words $p + q$ is a *least upper bound* (l.u.b.) of p and q. In a poset the l.u.b. is unique when it exists, and will be denoted $p \sqcup q$. A poset in which any two elements have a l.u.b. *and a g.l.b.* (§3.8) is called a *lattice.*

Categorially then a lattice is a skeletal pre-order having a product and a co-product for any two of its elements.

The disjoint union of two finite sets, with say m and n elements respectively is a set with (m plus n) elements. Indeed in **Finord**, the co-product of m and n is the ordinal number $m + n$ (where "+" means "plus" quite literally). With regard to the ordinals $1 = \{0\}$ and $2 = \{0, 1\}$ it is true then in the skeletal category **Finord** that

$$1 + 1 = 2,$$

while in **Finset**, or **Set** it would be more accurate to say

$$1 + 1 \cong 2$$

(Co-products being defined only up to isomorphism.)

Later in §5.4 we shall see that there are categories in which this last statement, under an appropriate interpretation, is false.

EXERCISE 3. Define the co-product arrow $f + g : a + b \to c + d$ of arrows $f : a \to c$ and $g : b \to d$ and dualise all of the Exercises in §3.8.

3.10. Equalisers

Given a pair $f,g : A \rightrightarrows B$ of parallel functions in **Set**, let E be the subset of A on which f and g agree, i.e.

$$E = \{x : x \in A \text{ and } f(x) = g(x)\}$$

Then the inclusion function $i : E \hookrightarrow A$ is called the *equaliser* of f and g. The reason for the name is that under composition with i we find that $f \circ i = g \circ i$, i.e. the two functions are "equalised" by i. Moreover, i is a "canonical" equaliser of f and g – if $h : C \to A$ is any other such equaliser of f and g, i.e. $f \circ h = g \circ h$,

$$\begin{tikzcd} E \arrow[r, hook, "i"] & A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \\ C \arrow[u, dashed, "k"] \arrow[ur, "h"'] & & \end{tikzcd}$$

then h "factors" uniquely through $i : E \hookrightarrow A$, i.e. there is exactly one function $k : C \to E$ such that $i \circ k = h$. In other words, given h, there is only one way to fill in the broken arrow to make the above diagram commute. That there can be at most one way is clear – if $i \circ k$ is to be the same as h, then for $c \in C$ we must have $i(k(c)) = h(c)$, i.e. $k(c) = h(c)$ (i being the inclusion). But this *does* work, for $f(h(c)) = g(h(c))$, and so $h(c) \in E$.

The situation just considered is now abstracted and applied to categories in general.

An arrow $i : e \to a$ in $\mathscr{C}$ is an *equaliser* of a pair $f,g : a \to b$ of $\mathscr{C}$-arrows if

(i) $f \circ i = g \circ i$, and

(ii) Whenever $h : c \to a$ has $f \circ h = g \circ h$ in $\mathscr{C}$ there is exactly one $\mathscr{C}$-arrow $k : c \to e$ such that $i \circ k = h$

$$\begin{tikzcd} e \arrow[r, "i"] & a \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & b \\ c \arrow[u, dashed, "k"] \arrow[ur, "h"'] & & \end{tikzcd}$$

An arrow will simply be called an *equaliser* in $\mathscr{C}$ if there are a pair of $\mathscr{C}$-arrows of which it is an equaliser.

THEOREM 1. *Every equaliser is monic.*

PROOF. Suppose i equalises f and g. To show i monic (left cancellable), let $i \circ j = i \circ l$, where $j, l : c \rightrightarrows e$. Then in the above diagram let $h : c \to a$ be the arrow $i \circ j$. We have $f \circ h = f \circ (i \circ j) = (f \circ i) \circ j = (g \circ i) \circ j = g \circ h$, and so there is a unique k with $i \circ k = h$. But $i \circ j = h$ (by definition), so k must be j. However, $i \circ l = i \circ j = h$, so $k = l$. Hence $j = l$. □

The converse of Theorem 1 does not hold in all categories. For instance in the category **N**, 1 is monic (all arrows are), but cannot equalise any pair (m, n) of arrows. If it did, we would have $m \circ 1 = n \circ 1$, i.e. $m + 1 = n + 1$, hence $m = n$. But then $m + 0 = n + 0$, which would imply that 0 factors uniquely through 1, i.e. there is a unique k having $1 + k = 0$. But of course there is no such *natural number* k.

Recalling that in **N** every arrow is epic, while 0 is the only iso, the next theorem gives a somewhat deeper explanation of the situation just described.

THEOREM 2. *In any category, an epic equaliser is iso.*

PROOF. If i equalises f and g, then $f \circ i = g \circ i$, so if i is epic, $f = g$. Then in the equaliser diagram, put $c = a$, and $h = \mathbf{1}_a$. We have

$$\begin{tikzcd} e \arrow[r, "i"] & a \\ a \arrow[u, dashed, "k"] \arrow[ur, "\mathbf{1}_a"'] & \end{tikzcd}$$

$f \circ \mathbf{1}_a = g \circ \mathbf{1}_a = f$, so there is a unique k with $i \circ k = \mathbf{1}_a$. Then $i \circ k \circ i = \mathbf{1}_a \circ i = i = i \circ \mathbf{1}_b$. But i is an equaliser, therefore left-cancellable, (Theorem 1), so $k \circ i = \mathbf{1}_b$. This gives k as an inverse to i, so i is iso. □

While monics may not be equalisers in all categories, they are certainly so in **Set** (and in fact in any topos). For if $f : E \rightarrowtail A$ is injective, define $h : A \to \{0, 1\}$ by the rule $h(x) = 1$, all $x \in A$, and $g : A \to \{0, 1\}$ by the rule

$$g(x) = \begin{cases} 1 & \text{if } x \in \operatorname{Im} i \\ 0 & \text{if } x \notin \operatorname{Im} i \end{cases}$$

Then f equalises g and h.

EXERCISE 1. Prove the last assertion.

EXERCISE 2. Show that in a poset, the only equalisers are the identity arrows.

3.11 Limits and co-limits

The definitions of the product of two objects and the equaliser of two arrows have the same basic form. In each case the entity in question has a certain property "canonically", in that any other object with that property "factors through" it in the manner indicated above. In the case of an equaliser the property is that of "equalising" the two original arrows. In the case of the product of a and b the property is that of being the domain of a pair of arrows whose codomains are a and b. This sort of situation is called a *universal construction.* The entity in question is *universal* amongst the things that have a certain property.

We can make this idea a little more precise (without being too pedantic, hopefully) by considering diagrams. By a diagram D in a category $\mathscr{C}$ we simply mean a collection of $\mathscr{C}$-objects $d_i, d_j, \ldots$, together with some $\mathscr{C}$-arrows $g: d_i \to d_j$ between certain of the objects in the diagram. (Possibly more than one arrow between a given pair of objects, possibly none.)

A *cone* for diagram D consists of a $\mathscr{C}$-object c together with a $\mathscr{C}$-arrow $f_i: c \to d_i$ for each object d_i in D, such that

$$\begin{tikzcd} d_i \arrow[rr, "g"] & & d_j \\ & c \arrow[ul, "f_i"] \arrow[ur, "f_j"'] & \end{tikzcd}$$

commutes, whenever g is an arrow in the diagram D. We use the symbolism $\{f_i: c \to d_i\}$ to denote a cone for D.

A *limit* for a diagram D is a D-cone $\{f_i: c \to d_i\}$ with the property that for any other D-cone $\{f'_i: c' \to d_i\}$ there is exactly one arrow $f: c' \to c$ such

$$\begin{tikzcd} & d_i & \\ c' \arrow[ur, "f'_i"] \arrow[rr, dashed, "f"'] & & c \arrow[ul, "f_i"'] \end{tikzcd}$$

commutes for every object d_i in D.

This limiting cone, when it exists, is said to have the *universal property* with respect to D-cones. It is universal amongst such cones – any other D-cone factors uniquely through it as in the last diagram. A limit for diagram D is unique up to isomorphism:- if $\{f_i: c \to d_i\}$ and $\{f'_i: c' \to d_i\}$ are both limits for D, then the unique commuting arrow $f: c' \dashrightarrow c$ above is iso (its inverse is the unique commuting arrow $c \dashrightarrow c'$ whose existence follows from the fact that $\{f'_i: c' \to d_i\}$ is a limit).

EXAMPLE 1. Given $\mathscr{C}$-objects a and b let D be the arrow-less diagram

$$a \qquad\qquad b$$

A D-cone is then an object c, together with two arrows f, and g of the form

$$\begin{tikzcd} a & & b \\ & c \arrow[ul, "f"] \arrow[ur, "g"'] & \end{tikzcd}$$

A limiting D-cone, one through which all such cones factor, is none other than a product of a and b in $\mathscr{C}$.

EXAMPLE 2. Let D be the diagram

$$\begin{tikzcd} a \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & b \end{tikzcd}$$

A D-cone is a pair $h: c \rightarrow a$, $j: c \rightarrow b$ such that

$$\begin{tikzcd} a \arrow[rr, "f"] & & b \\ & c \arrow[ul, "h"] \arrow[ur, "j"'] & \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} a \arrow[rr, "g"] & & b \\ & c \arrow[ul, "h"] \arrow[ur, "j"'] & \end{tikzcd}$$

commute. But this requires that $j = f \circ h = g \circ h$, so we can simply say that a D-cone in this case is an arrow $h: c \rightarrow a$ such that

$$\begin{tikzcd} c \arrow[r, "h"] & a \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & b \end{tikzcd}$$

commutes, i.e. $f \circ h = g \circ h$. We then see that a D-limit is an *equaliser* of f and g.

EXAMPLE 3. Let D be the *empty* diagram

i.e. no objects and no arrows. A D-cone is then simply a $\mathscr{C}$-object c (there are no f_i's as D has no d_i's). A limiting cone is then an object c

such that for any other $\mathscr{C}$-object (D-cone) c', there is exactly one arrow $c' \dashrightarrow c$. In other words, a limit for the empty diagram is a *terminal object!*

□

By *duality* we define a *co-cone* $\{f_i : d_i \to c\}$ for diagram D to consist of an object c, and arrows $f_i : d_i \to c$ for each object d_i in D. A *co-limit* for D is then a co-cone $\{f_i : d_i \to c\}$ with the *co-universal property* that for any other co-cone $\{f_i' : d_i \to c'\}$ there is exactly one arrow $f : c \to c'$ such

$$\begin{tikzcd} & d_i \arrow[dl, "f_i"'] \arrow[dr, "f_i'"] & \\ c \arrow[rr, dashed, "f"] & & c' \end{tikzcd}$$

commutes for every d_i in D.

A co-limit for the diagram of Example 1 is a co-product of a and b, while a co-limit for the empty diagram is a category $\mathscr{C}$ is an *initial object* for $\mathscr{C}$.

3.12. Co-equalisers

The co-equaliser of a pair (f, g) of parallel $\mathscr{C}$-arrows is a co-limit for the diagram

$$\begin{tikzcd} a \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & b \end{tikzcd}$$

It can be described as a $\mathscr{C}$-arrow $q : b \to e$ such that

(i) $q \circ f = q \circ g$, and

(ii) whenever $h : b \to c$ has $h \circ f = h \circ g$ in $\mathscr{C}$ there is exactly one $\mathscr{C}$-arrow $k : e \to c$ such that

$$\begin{tikzcd} a \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & b \arrow[r, "q"] \arrow[dr, "h"'] & e \arrow[d, dashed, "k"] \\ & & c \end{tikzcd}$$

commutes. The results of §3.10 immediately dualise to tell us that co-equalisers are epic, that the converse is true in **Set**, and that a monic co-equaliser is iso.

In **Set** an "$\in$-related" description of the co-equaliser comes through the very important notion of *equivalence relation.* An equivalence relation on a set A is, by definition, a relation $R \subseteq A \times A$ that is

(a) *reflexive,* i.e. aRa, for every $a \in A$;
(b) *transitive,* i.e. whenever aRb and bRc, then aRc; and
(c) *symmetric,* i.e. whenever aRb, then bRa.

Equivalence relations arise throughout mathematics (and elsewhere) in situations where one wishes to identify different things that are 'equivalent'. Typically one may be concerned with some particular property (properties) with respect to which different things may be indistinguishable. The relation that holds between two things when they are thus indistinguishable will then be an equivalence relation.

We have in fact already met this idea in the discussion in §3.4 of isomorphism. Two objects in a category that are isomorphic might just as well be the same object, as far as categorial properties are concerned, and indeed

$$\{\langle a, b\rangle : a \cong b \text{ in } \mathscr{C}\}$$

is a relation on $\mathscr{C}$-objects that is reflexive, transitive, and symmetric. (Exercise 3.4.1).

The process of "identifying equivalent things" is rendered explicit by lumping together all things that are related to each other and treating the resulting collection as a single entity. Formally, for $a \in A$ we define the *R-equivalence class* of a to be the set

$$[a] = \{b : aRb\}$$

of all members of A to which a is R-related. Different elements may have the same subset of A as their equivalence class, and the situation in general is as follows:

(1) $[a] = [b]$ iff aRb
(2) if $[a] \neq [b]$ then $[a] \cap [b] = \emptyset$
(3) $a \in [a]$

(the proof of these depends on properties (a), (b), (c) above). Statement (1) tells us that equivalent elements are related to precisely the same elements, and conversely (2) says if two equivalence classes are not the same, then they have no elements in common at all. This, together with

(3) (which holds by (a)), implies that each $a \in A$ is a member of one and only one R-equivalence class.

The actual identification process consists in passing from the original set to a new set whose elements are the R-equivalence classes, i.e. we shift from A to the set

$$A/R = \{[a] \colon a \in A\}$$

The transfer is effected by the *natural map* $f_R : A \to A/R$, where $f_R(a) = [a]$, for $a \in A$.

Thus, by (1), when aRb we have $f_R(a) = f_R(b)$, and so R-equivalent elements are identified by the application of f_R.

What has all this to do with co-equalisers? Well the point is that f_R is the co-equaliser of the pair $f, g : R \rightrightarrows A$ of projection functions from R to A, i.e. the functions

$$f(\langle a, b\rangle) = a$$

and

$$g(\langle a, b\rangle) = b.$$

The last paragraph explained in effect why $f_R \circ f = f_R \circ g$. To see why the diagram

$$\begin{tikzcd} R \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & A \arrow[r, "f_R"] \arrow[dr, "h"'] & A/R \arrow[d, dashed, "k"] \\ & & B \end{tikzcd}$$

can be "filled in" by only one k, given $h \circ f = h \circ g$, we suppose we have a k such that $k \circ f_R = h$. Then for $[a] \in A/R$ we must have $k([a]) = k(f_R(a)) = k \circ f_R(a) = h(a)$. So the only thing we can do is define k to be the function that for input $[a]$ gives output $h(a)$. There is a problem here about whether k is a *well-defined* function, for if $[a] = [b]$, our rule also tells us to output $h(b)$ for input $[a] = [b]$. In order for there to be a unique output for a given input, we would need to know in this case that $h(a) = h(b)$. But in fact if $[a] = [b]$ then $\langle a, b\rangle \in R$ and our desideratum follows, because $h \circ f = h \circ g$.

The question of "well-definedness" just dealt with occurs repeatedly in working with so called "quotient" sets of the form A/R. Operations on, and properties of an R-equivalence class are defined by reference to some selected member of the equivalence class, called its *representative*.

One must always check that the definition does not depend on which representative is chosen. In other words a well defined concept is one that is *stable* or *invariant* under R, i.e. is not altered or destroyed when certain things are replaced by others to which they are R-equivalent.

Equivalence relations can be used to construct the co-equaliser in **Set** of any pair $f,g: A \rightrightarrows B$ of parallel functions. To co-equalise f and g we have to identify $f(x)$ with $g(x)$, for $x \in A$. So we consider the relation

$$S = \{\langle f(x), g(x)\rangle : x \in A\} \subseteq B \times B.$$

S may not be an equivalence relation on B. However, it is possible to build up S until it becomes an equivalence relation, and to do this in a "minimal" way. There is an equivalence relation R on B such that

(i) $S \subseteq R$, and

(ii) if T is any other equivalence on B such that T contains S, then $R \subseteq T$

(i.e. R is the "smallest" equivalence relation on B that contains S). The co-equaliser of f and g is then the natural map $f_R : B \to B/R$. (See Arbib and Manes, p. 19, for the details of how to construct this R).

3.13. The pullback

A *pullback* of a pair $a \xrightarrow{f} c \xleftarrow{g} b$ of $\mathscr{C}$-arrows with a common codomain is a limit in $\mathscr{C}$ for the diagram

$$\begin{tikzcd} & b \arrow[d, "g"] \\ a \arrow[r, "f"] & c \end{tikzcd}$$

A cone for this diagram consists of three arrows f', h, g', such that

$$\begin{tikzcd} d \arrow[r, "f'"] \arrow[d, "g'"'] \arrow[dr, "h"] & b \arrow[d, "g"] \\ a \arrow[r, "f"'] & c \end{tikzcd}$$

commutes. But this requires that $h = g \circ f' = f \circ g'$, so we may simply say that a cone is a pair $a \xleftarrow{g'} d \xrightarrow{f'} b$ of $\mathscr{C}$-arrows such that the "square"

$$\begin{tikzcd} d \arrow[r, "f'"] \arrow[d, "g'"'] & b \arrow[d, "g"] \\ a \arrow[r, "f"'] & c \end{tikzcd}$$

commutes, i.e. $f \circ g' = g \circ f'$.

Thus we have, by the definition of universal cone, that a pullback of the pair $a \xrightarrow{f} c \xleftarrow{g} b$ in $\mathscr{C}$ is a pair of $\mathscr{C}$-arrows $a \xleftarrow{g'} d \xrightarrow{f'} b$ such that

(i) $f \circ g' = g \circ f'$, and

(ii) whenever $a \xleftarrow{h} e \xrightarrow{j} b$ are such that $f \circ h = g \circ j$, then

$$\begin{tikzcd} e \arrow[dr, dashed, "k"] \arrow[drr, bend left, "j"] \arrow[ddr, bend right, "h"'] & & \\ & d \arrow[r, "f'"'] \arrow[d, "g'"] & b \arrow[d, "g"] \\ & a \arrow[r, "f"] & c \end{tikzcd}$$

there is exactly one $\mathscr{C}$-arrow $k : e \dashrightarrow d$ such that $h = g' \circ k$ and $j = f' \circ k$. In other words when h and j are such that the outer "square", or "boundary" of the above diagram commutes, then there is only one way to fill in the broken arrow to make the whole diagram commute.

The inner square (f, g, f', g') of the diagram is called a *pullback square*, or *Cartesian square*. We also say that f' arises by *pulling back f along g*, and g' arises by *pulling back g along f*.

The pullback is a very important and fundamental mathematical notion, that incorporates a number of well known constructions. It is certainly the most important limit concept to be used in the study (and definition) of topoi. The following examples, illustrating its workings and generality, are commended as worthy of detailed examination.

EXAMPLE 1. In **Set**, the pullback

$$\begin{tikzcd} D \arrow[r, "f'"] \arrow[d, "g'"'] & B \arrow[d, "g"] \\ A \arrow[r, "f"'] & C \end{tikzcd}$$

of two set function f and g is defined by putting

$$D = \{\langle x, y\rangle : x \in A,\ y \in B, \text{ and } f(x) = g(y)\}$$

with f' and g' as the projections:

$$f'(\langle x, y\rangle) = y$$
$$g'(\langle x, y\rangle) = x.$$

D is then a subset of the product set $A \times B$. It is sometimes denoted $A \underset{C}{\times} B$, the product of A and B *over* C. Pullbacks are also known as "fibred products" (the use of the word "fibred" is explained in Chapter 4).

EXAMPLE 2. **Inverse images.** If $f: A \to B$ is a function, and C a subset of B, then the *inverse image of C under f*, denoted $f^{-1}(C)$, is that subset of A consisting of all the f-inputs whose corresponding outputs lie in C, i.e.

$$f^{-1}(C) = \{x : x \in A \text{ and } f(x) \in C\}$$

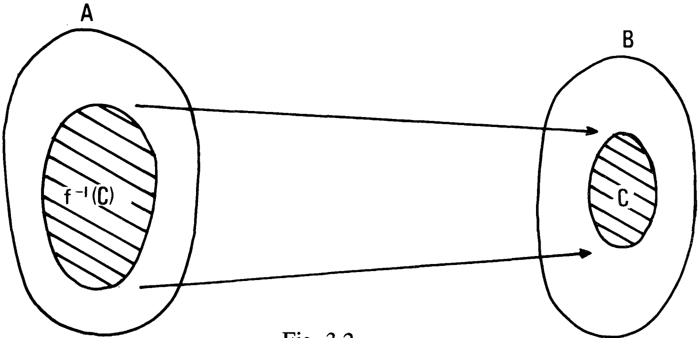

Fig. 3.2.

The diagram

$$\begin{tikzcd} f^{-1}(C) \arrow[r, "f^*"] \arrow[d, hook] & C \arrow[d, hook] \\ A \arrow[r, "f"] & B \end{tikzcd}$$

is a pullback square in **Set**, where the arrows with curved tails denote inclusions as usual, and $f^*(x) = f(x)$ for $x \in f^{-1}(C)$ (i.e. f^* is the *restriction* of f to $f^{-1}(C)$). Thus *the inverse image of C under f arises by pulling C back along f.*

The dynamical quality inherent in the notion of function (cf. §2.1) is quite forcefully present in this example of "pulling back". It would be quite unconvincing to suggest we were just dealing with sets of ordered pairs.

EXAMPLE 3. **Kernel relation.** Associated with any function $f: A \to B$ is a special equivalence relation on A, denoted R_f, and called the *kernel relation* of f (the kernel "congruence" in universal algebra, where it lies at the heart of the First Isomorphism Theorem). As a set of ordered pairs we have

$$R_f = \{\langle x, y\rangle : x \in A \text{ and } y \in A \text{ and } f(x) = f(y)\}$$

or

$$xR_f y \quad \text{iff} \quad f(x) = f(y).$$

In the light of our first example we see that

$$\begin{tikzcd} R_f \arrow[r, "p_2"] \arrow[d, "p_1"'] & A \arrow[d, "f"] \\ A \arrow[r, "f"'] & B \end{tikzcd}$$

is a pullback square, where $p_1(\langle x, y\rangle) = x$ and $p_2(\langle x, y\rangle) = y$, i.e. R_f arises as the pullback of f along itself. This observation will provide the key to some work in Chapter 5 on the "epi-monic factorisation" of arrows in a topos.

EXAMPLE 4. **Kernels** (for algebraists). Let $f: \mathbf{M} \to \mathbf{N}$ be a monoid homomorphism and

$$K = \{x : f(x) = e\}$$

the kernel of f.

Then

$$\begin{tikzcd} \mathbf{K} \arrow[r, hook] \arrow[d, "!"'] & \mathbf{M} \arrow[d, "f"] \\ \mathbf{O} \arrow[r, "!"] & \mathbf{N} \end{tikzcd}$$

is a pullback square in **Mon**, where **O** is the one-element monoid (which is initial and terminal).

This characterisation of kernels applies also to the categories **Grp** and **Vect**.

Example 5. In a pre-order $(P, \sqsubseteq)$,

$$\begin{tikzcd} s \arrow[r] \arrow[d] & q \arrow[d] \\ p \arrow[r] & r \end{tikzcd}$$

is a pullback square iff s is a product of p and q.

Example 6. In any category with a terminal object, if

$$\begin{tikzcd} d \arrow[r, "g"] \arrow[d, "f"'] & b \arrow[d, "!"] \\ a \arrow[r, "!"] & 1 \end{tikzcd}$$

is a pullback, then $\langle f, g \rangle$ is a product (g.l.b.) of a and b.

Example 7. In any category, if

$$\begin{tikzcd} e \arrow[r, "i"] \arrow[d, "i"'] & a \arrow[d, "g"] \\ a \arrow[r, "f"'] & b \end{tikzcd}$$

is a pullback, then i is an equaliser of f and g.

Example 8. The Pullback Lemma (PBL). *If a diagram of the form*

$$\begin{tikzcd} \bullet \arrow[r] \arrow[d] & \bullet \arrow[r] \arrow[d] & \bullet \arrow[d] \\ \bullet \arrow[r] & \bullet \arrow[r] & \bullet \end{tikzcd}$$

commutes, then

(i) *if the two small squares are pullbacks, then the outer "rectangle"* (with top and bottom edges the evident composites) *is a pullback;*

(ii) *if the outer rectangle and the right hand square are pullbacks then so is the left hand square.*

The PBL is a key fact, and will be used repeatedly in what follows. Its proof, though rather tedious, will certainly familiarise the reader with how a pullback works.

The PBL will often be used for a diagram of the form,

$$\begin{tikzcd} \bullet \arrow[r] \arrow[d] & \bullet \arrow[d] \\ \bullet \arrow[r] \arrow[d] & \bullet \arrow[d] \\ \bullet \arrow[r] & \bullet \end{tikzcd}$$

in which case when the outer rectangle and *bottom* square are pullbacks, we will conclude that the *top* square is a pullback.

EXAMPLE 9. In any category, an arrow $f: a \to b$ is monic iff

$$\begin{tikzcd} a \arrow[r, "1_a"] \arrow[d, "1_a"'] & a \arrow[d, "f"] \\ a \arrow[r, "f"'] & b \end{tikzcd}$$

is a pullback square.

EXERCISE. Show that if

$$\begin{tikzcd} a \arrow[r, "g"] \arrow[d] & b \arrow[d] \\ c \arrow[r, "f"] & d \end{tikzcd}$$

is a pullback square, and f is monic, then g is also monic.

3.14. Pushouts

The dual of a pullback of a pair of arrows with common codomain is a *pushout* of the two arrows with common domain:

a pushout of $b \xleftarrow{f} a \xrightarrow{g} c$ is a co-limit for the diagram

$$\begin{tikzcd} a \arrow[r, "g"] \arrow[d, "f"'] & c \\ b & \end{tikzcd}$$

In **Set** it obtained by forming the disjoint union $b+c$ and then identifying $f(x)$ with $g(x)$, for each $x \in a$ (by a co-equaliser).

EXERCISE. Dualise §3.13.

3.15. Completeness

A category $\mathscr{C}$ is *complete* if every diagram in $\mathscr{C}$ has a limit in $\mathscr{C}$. Dually $\mathscr{C}$ is *co-complete* when every $\mathscr{C}$-diagram has a co-limit. A *bi-complete* category is one that is complete and co-complete.

A *finite* diagram is one that has a finite number of objects, and a finite number of arrows between them.

A category is *finitely complete* if it has a limit for every *finite* diagram. Finite co-completeness and finite bi-completeness are defined similarly.

THEOREM 1. *If $\mathscr{C}$ has a terminal object, and a pullback for each pair of $\mathscr{C}$-arrows with common codomain, then $\mathscr{C}$ is finitely complete.* □

A proof of this theorem is beyond our present scope (and outside our major concerns). The details may be found in Herrlich and Strecker [73], Theorem 23.7, along with a number of other characterisations of finite completeness.

To illustrate the Theorem, we observe that

(A) given a terminal object and pullbacks, the product of a and b is got from the pullback of $a \to 1 \leftarrow b$ (cf. §3.13, Example 6);

(B) given pullbacks and products, from a parallel pair $f, g : a \rightrightarrows b$ we first form the product arrows

$$a \xrightarrow{\langle 1_a, f\rangle} a \times b \quad \text{and} \quad a \xrightarrow{\langle 1_a, g\rangle} a \times b$$

and then their pullback

$$\begin{tikzcd} d \arrow[r, "q"] \arrow[d, "p"'] & a \arrow[d, "\langle 1_a, g\rangle"] \\ a \arrow[r, "\langle 1_a, f\rangle"] & a \times b \end{tikzcd}$$

It follows readily (§3.8) that $p = q$, and that this arrow is an equaliser of f and g.

Exercises

(1) Verify (B), and consider the details of that construction in **Set**.
(2) Show how to construct pullbacks from products and equalisers. A hint is given by the description (Example 1, §3.13) of pullbacks in **Set**. A co-hint appears in §3.14.
(3) Dualise the Theorem of this section.

3.16. Exponentiation

Given sets A and B we can form in **Set** the collection B^A of all functions that have domain A and codomain B, i.e.

$$B^A = \{f : f \text{ is a function from } A \text{ to } B\}$$

To characterise B^A by arrows we observe that associated with B^A is a special arrow

$$ev : B^A \times A \to B,$$

given by the rule

$$ev(\langle f, x\rangle) = f(x).$$

ev is the *evaluation* function. Its inputs are pairs of the form $\langle f, x\rangle$ where $f: A \to B$ and $x \in A$. The action of ev for such as input is to apply f to x, to evaluate f at x, yielding the output $f(x) \in B$. The categorial description of B^A comes from the fact that ev enjoys a universal property amongst all set functions of the form

$$C \times A \xrightarrow{g} B.$$

Given any such g, there is one and only one function $\hat{g}: C \to B^A$ such that

$$\begin{tikzcd} B^A \times A \arrow[dr, "ev"] & \\ & B \\ C \times A \arrow[uu, dashed, "\hat{g} \times \mathrm{id}_A"] \arrow[ur, "g"'] & \end{tikzcd}$$

commutes where $\hat{g} \times \mathrm{id}_A$ is the product function described in §3.8. For input $\langle c, a \rangle \in C \times A$ it gives output $\langle \hat{g}(c), \mathrm{id}_A(a) \rangle = \langle \hat{g}(c), a \rangle$.

The idea behind the definition of $\hat{g}$ is that the action of g causes any particular c to determine a function $A \to B$ by fixing the first elements of arguments of g at c, and allowing the second elements to range over A. In other words for a given $c \in C$ we define $g_c : A \to B$ by the rule

$$g_c(a) = g(\langle c, a \rangle), \quad \text{for each} \quad a \in A.$$

$\hat{g}: C \to B^A$ can now be defined by $\hat{g}(c) = g_c$, all $c \in C$. For any $\langle c, a \rangle \in C \times A$ we then get

$$ev(\langle \hat{g}(c), a \rangle) = g_c(a) = g(\langle c, a \rangle)$$

and so the above diagram commutes. But the requirement that the diagram commutes, i.e. that $ev(\langle \hat{g}(c), a \rangle) = g(\langle c, a \rangle)$, means that $\hat{g}(c)$ must be the function that for input a gives output $g(\langle c, a \rangle)$, i.e. $\hat{g}(c)$ must be g_c as above.

By abstraction then we say that a category $\mathscr{C}$ has *exponentiation* if it has a product for any two $\mathscr{C}$-objects, and if for any given $\mathscr{C}$-objects a and b there is a $\mathscr{C}$-object b^a and a $\mathscr{C}$-arrow $ev: b^a \times a \to b$, called an *evaluation* arrow, such that for any $\mathscr{C}$-object c and $\mathscr{C}$-arrow $g: c \times a \to b$, there is a unique $\mathscr{C}$-arrow $\hat{g}: c \to b^a$ making

$$\begin{tikzcd} b^a \times a \arrow[dr, "ev"] & \\ & b \\ c \times a \arrow[uu, dashed, "\hat{g} \times 1_a"] \arrow[ur, "g"'] & \end{tikzcd}$$

commute, i.e. a unique $\hat{g}$ such that $ev \circ (\hat{g} \times 1_a) = g$. The assignment of $\hat{g}$ to g establishes a bijection

$$\mathscr{C}(c \times a, b) \cong \mathscr{C}(c, b^a)$$

between the collection of $\mathscr{C}$-arrows from $c \times a$ to b, and the collection of those from c to b^a. For if $\hat{g} = \hat{h}$, then $ev \circ (\hat{g} \times 1_a) = ev \circ (\hat{h} \times 1_a)$, i.e. $g = h$, and so the assignment is injective. To see that it is surjective, take $h : c \to b^a$ and *define* $g = ev \circ (h \times 1_a)$. By the uniqueness of $\hat{g}$ we must have $h = \hat{g}$.

Two arrows (g and $\hat{g}$) that correspond to each other under this bijection will be called *exponential adjoints* of each other. The origin of this terminology may be found in Chapter 15.

A finitely complete category with exponentiation is said to be *Cartesian closed.*

EXAMPLE 1. If A and B are finite sets with say m and n elements, then B^A is finite and has n^m ("n to the power m") elements. In the expression n^m, the "m" is called an *exponent*, hence the above terminology. **Finord** is Cartesian closed, and indeed the exponential is literally the number n^m.

EXAMPLE 2. A *chain* is a poset $\mathbf{P} = (P, \sqsubseteq)$ that is linearly ordered, i.e. has $p \sqsubseteq q$ or $q \sqsubseteq p$ for any $p, q \in P$. If $\mathbf{P}$ is a chain with a terminal object 1, then we put

$$q^p = \begin{cases} 1 & \text{if} \quad p \sqsubseteq q \\ q & \text{if} \quad q \sqsubset p \quad (\text{i.e. } q \sqsubseteq p \quad \text{and} \quad q \neq p) \end{cases}$$

A chain always has products:

$$p \times q = \text{g.l.b. of } p \text{ and } q = \begin{cases} p & \text{if} \quad p \sqsubseteq q \\ q & \text{if} \quad q \sqsubseteq p. \end{cases}$$

We thus have two cases to consider for ev.

(i) $p \sqsubseteq q$. Then $q^p \times p = 1 \times p = p \sqsubseteq q$;

(ii) $q \sqsubseteq p$. Then $q^p \times p = q \times p = q$.

In either case $q^p \times p \sqsubseteq q$ and so ev is the unique arrow $q^p \times p \to q$ in $\mathbf{P}$. We leave it to the reader to verify that this definition gives $\mathbf{P}$ exponentiation. An explanation of why it works, and an account of exponentiation in posets in general will be forthcoming in Chapter 8. □

THEOREM 1. *Let $\mathscr{C}$ be a Cartesian closed category with an initial object* 0. *Then in $\mathscr{C}$,*

(1) $0 \cong 0 \times a$, *for any object a;*

(2) *if there exists an arrow $a \to 0$, then $a \cong 0$;*

(3) *if $0 \cong 1$, then the category $\mathscr{C}$ is degenerate, i.e. all $\mathscr{C}$-objects are isomorphic;*

(4) *any arrow* $0 \to a$ *with dom* 0 *is monic;*
(5) $a^1 \cong a$, $a^0 \cong 1$, $1^a \cong 1$.

PROOF. (1) For any $\mathscr{C}$-object b, $\mathscr{C}(0, b^a)$ has only one member (as 0 is initial). By definition of exponentiation, $\mathscr{C}(0, b^a) \cong \mathscr{C}(0 \times a, b)$. Hence the latter collection has only one member. Thus there is only one arrow $0 \times a \to b$, for any b. Hence $0 \times a$ is an initial $\mathscr{C}$-object, and since the latter are unique up to isomorphism, $0 \cong 0 \times a$.

(2) Given $f: a \to 0$, we show that $a \cong 0 \times a$, and hence by (1), $a \cong 0$. From the universal definition of product

$$\begin{tikzcd} & a \arrow[dl, "f"'] \arrow[d, dashed, "\langle f, 1_a \rangle"] \arrow[dr, "1_a"] & \\ 0 & 0 \times a \arrow[l, "pr_0"] \arrow[r, "pr_a"'] & a \end{tikzcd}$$

$pr_a \circ \langle f, 1_a \rangle = 1_a$. But $\langle f, 1_a \rangle \circ pr_a$ is an arrow from $0 \times a$ to $0 \times a$, and there is only one such, $0 \times a$ being initial. thus $\langle f, 1_a \rangle \circ pr_a = 1_{0 \times a}$, giving $\langle f, 1_a \rangle = pr_a^{-1}$ and $pr_a : 0 \times a \cong a$.

(3) If $0 \cong 1$, then for any a, since there is an arrow from a to 1, there will be one from a to 0 whence, by (2), $a \cong 0$. Thus all objects are isomorphic to 0. Ergo they are all isomorphic to each other.

(4) Given $f: 0 \to a$, suppose $f \circ g = f \circ h$, i.e.

$$b \underset{h}{\overset{g}{\rightrightarrows}} 0 \xrightarrow{f} a$$

commutes. But then by (2), $b \cong 0$, so b is an initial object and there is only one arrow $b \to 0$. Thus $g = h$, and f is left-cancellable. □

EXERCISE. Prove part (5) of the Theorem, and interpret (1)–(5) as they apply to **Set**. □

Having reached the end of this chapter, we can look back on an extensive catalogue of categorial versions of mathematical concepts and constructions. We now have some idea of how category theory has recreated the world of mathematical ideas, and indeed expanded the horizons of mathematical thought. And we have seen a number of features that distinguish **Set** from other categories. In **Set**, monic epics are iso, a property not enjoyed by **Mon**. It is however, enjoyed by **Grp** – but then **Grp** is not Cartesian closed (this follows from the above Theorem – **Grp** is not degenerate, but does have $0 \cong 1$). On the other hand the Cartesian-closed categories are not all "**Set**-like". The poset $\mathbf{n} = \{0, \ldots, n-1\}$ is Cartesian-closed (being a chain with terminal object), but has monic epics that are not iso. It would appear then that to develop a categorial set theory we will have to work in categories that have some other special features in common with **Set**, something at least that is not possessed by **Mon**, **n**, etc. In fact what we need is one more construction, a conceptually straightforward but very powerful one whose nature will be revealed in the next chapter.