

## INTRODUCING TOPOI

*“This is the development on the basis of elementary (first-order) axioms of a theory of “toposes” just good enough to be applicable not only to sheaf theory, algebraic spaces, global spectrum, etc. as originally envisaged by Grothendieck, Giraud, Verdier, and Hakim but also to Kripke semantics, abstract proof theory, and the Cohen–Scott–Solovay method for obtaining independence results in set theory.”*

F. W. Lawvere

## 4.1. Subobjects

If  $A$  is a subset of  $B$ , then the inclusion function  $A \hookrightarrow B$  is injective, hence monic. On the other hand any monic function  $f: C \rightarrow B$  determines a subset of  $B$ , viz  $\text{Im } f = \{f(x) : x \in C\}$ . It is easy to see that  $f$  induces a bijection between  $C$  and  $\text{Im } f$ , so  $C \cong \text{Im } f$ .

Thus the domain of a monic function is isomorphic to a subset of the codomain. Up to isomorphism, the domain is a subset of the codomain. This leads us to the categorial versions of subsets, which are known as *subobjects*:

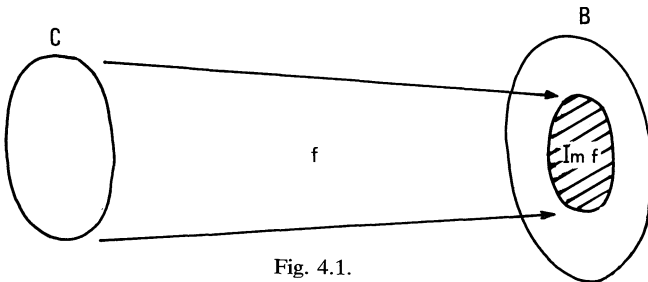


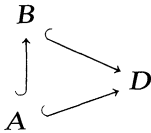
Fig. 4.1.

a *subobject* of a  $\mathcal{C}$ -object  $d$  is a monic  $\mathcal{C}$ -arrow  $f: a \rightarrow d$  with codomain  $d$ .

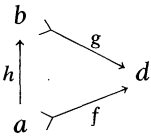
Now if  $D$  is a set, then the collection of all subsets of  $D$  is known as the *powerset* of  $D$ , denoted  $\mathcal{P}(D)$ . Thus

$$\mathcal{P}(D) = \{A : A \text{ is a subset of } D\}.$$

The relation of set inclusion is a partial ordering on the power set  $\mathcal{P}(D)$ , i.e.  $(\mathcal{P}(D), \subseteq)$  is a poset, and becomes a category in which there is an arrow  $A \rightarrow B$  iff  $A \subseteq B$ . When there is such an arrow, the diagram



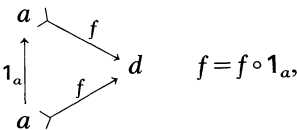
commutes. This suggests a way of defining an “inclusion” relation between subobjects of  $d$ . Given  $f: a \rightarrow d$  and  $g: b \rightarrow d$ , we put  $f \subseteq g$  iff there is a  $\mathcal{C}$ -arrow  $h: a \rightarrow b$  such that



commutes, i.e.  $f = g \circ h$ . (such an  $h$  will always be monic, by Exercise 3.1.2, so  $h$  will be a subobject of  $b$ , enhancing the analogy with the **Set** case). Thus  $f \subseteq g$  precisely when  $f$  factors through  $g$ .

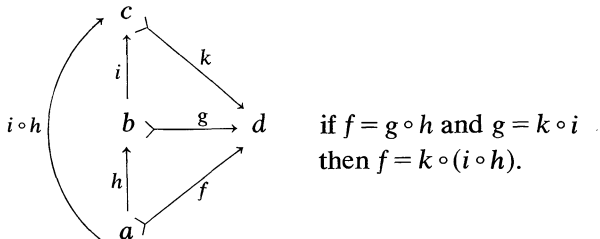
The inclusion relation on subobjects is

(i) *reflexive*;  $f \subseteq f$ , since

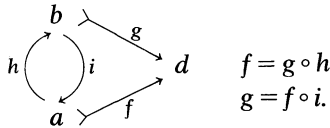


and

(ii) *transitive*; if  $f \subseteq g$  and  $g \subseteq k$ , then  $f \subseteq k$ , since



Now if  $f \subseteq g$  and  $g \subseteq f$ , then  $f$  and  $g$  each factor through each other, as in



In that case,  $h : a \rightarrow b$  is iso, with inverse  $i$  (exercise for the reader). Thus when  $f \subseteq g$  and  $g \subseteq f$ , they have isomorphic domains, and so we call them *isomorphic subobjects* and write  $f \approx g$ . Now in order for  $\subseteq$  to be antisymmetric, we require that when  $f \approx g$ , then  $f = g$ . This may not in fact be so, indeed we may have  $a \neq b$ . So  $\subseteq$  will in general be a preordering on the subobjects of  $d$  as defined, and not a partial ordering. If we left things there, we would run into difficulties later. We really do want to be able to think of  $\subseteq$  as being antisymmetric. The machinery that allows this was set up in §3.12. The relation  $\approx$  is an equivalence relation (exercise – use (i), (ii) above). Each  $f : a \rightarrow d$  determines an equivalence class

$$[f] = \{g : f \approx g\},$$

and we form the collection

$$\text{Sub}(d) = \{[f] : f \text{ is a monic with } \text{cod } f = d\}.$$

We are now going to refer to the members of  $\text{Sub}(d)$  as the subobjects, i.e. we *redefine* a subobject of  $d$  to be an equivalence class of monics with codomain  $d$ . To obtain an inclusion notion for these entities, we put (using the same symbol as before)

$$[f] \subseteq [g] \quad \text{iff} \quad f \subseteq g.$$

Here we come up against the question mentioned in §3.12. Is the definition, given via representatives of equivalence classes, independent of the choice of representative? The answer is yes. If  $[f] = [f']$  and  $[g] = [g']$ , then  $f \subseteq g$  iff  $f' \subseteq g'$ , i.e.  $\subseteq$  is stable under  $\approx$  (exercise).

The point of this construction was to make  $\subseteq$  antisymmetric. But when  $[f] \subseteq [g]$  and  $[g] \subseteq [f]$ , then  $f \subseteq g$  and  $g \subseteq f$ , so  $f \approx g$  and hence  $[f] = [g]$ . Thus the subobjects of  $d$ , as now defined, form a poset  $(\text{Sub}(d), \subseteq)$ .

This lengthy piece of methodology is not done with yet. It now starts to bite its own tail as we blur the distinction between equivalence class and representative. We shall usually say “the subobject  $f$ ” when we mean “the subobject  $[f]$ ”, and “ $f \subseteq g$ ” when strictly speaking “[ $f$ ]  $\subseteq$  [ $g$ ]” is intended, etc. All properties and constructions of subobjects used will however be stable under  $\approx$  (indeed being categorial they will only be

defined up to isomorphism anyway). So this *abus de langage* is technically justifiable and has great advantages in terms of conceptual and notational clarity. The only point on which we shall continue to be precise is the matter of identity. “ $f \simeq g$ ” will be used whenever we mean that  $f$  and  $g$  are the same subobject, i.e.  $[f] = [g]$ , while “ $f = g$ ” will be reserved for when they are the same actual arrow.

EXERCISE 1. In **Set**,  $\text{Sub}(D) \cong \mathcal{P}(D)$ . □

### Elements

Having described subsets categorially, we turn to actual elements of sets. A member  $x$  of set  $A$ , ( $x \in A$ ), can be identified with the “*singleton*” subset  $\{x\}$  of  $A$ , and hence with the arrow  $\{x\} \hookrightarrow A$ , from the terminal object  $\{x\}$  to  $A$ . In the converse direction, a function  $f: 1 \rightarrow A$  in **Set** determines an element of  $A$ , viz the  $f$ -image of the only member of the terminal object  $1$ . Thus; if category  $\mathcal{C}$  has a terminal object  $1$ , then an *element* of a  $\mathcal{C}$ -object  $a$  is defined to be a  $\mathcal{C}$ -arrow  $x: 1 \rightarrow a$ . (Note that  $x: 1 \rightarrow a$  is always monic – Exercise 3.6.3.)

Of course the question is – does this notion in general reflect the behaviour of elements in **Set**? Must a non-initial  $\mathcal{C}$ -object have elements? Can two different  $\mathcal{C}$ -objects have the same elements? Can we characterise monic and epic arrows in terms of elements of their dom and cod? These matters will be taken up in due course.

### Naming arrows

A function  $f: A \rightarrow B$  from set  $A$  to set  $B$  is an *element* of the set  $B^A$ , i.e.  $f \in B^A$ , and so determines a function  $\ulcorner f \urcorner: \{0\} \rightarrow B^A$ , with  $\ulcorner f \urcorner(0) = f$ . Then if  $x$  is an element of  $A$ , we have a categorial “element”  $\bar{x}: \{0\} \rightarrow A$ , with  $\bar{x}(0) = x$ . Since  $ev(\langle f, x \rangle) = f(x)$  we find that  $ev \circ \langle \ulcorner f \urcorner, \bar{x} \rangle(0) = ev(\ulcorner f \urcorner(0), \bar{x}(0)) = f(x) = f(\bar{x}(0))$ , and hence we have an equality of functions:

$$ev \circ \langle \ulcorner f \urcorner, \bar{x} \rangle = f \circ \bar{x}.$$

This situation can be lifted to any category  $\mathcal{C}$  that has exponentials. Given a  $\mathcal{C}$ -arrow  $f: a \rightarrow b$ , let  $f \circ pr_a: 1 \times a \rightarrow b$  be the composite  $f \circ pr_a: 1 \times a \rightarrow a \rightarrow b$ . Then the *name* of  $f$  is, by definition, the arrow  $\ulcorner f \urcorner: 1 \rightarrow b^a$  that is the exponential adjoint of  $f \circ pr_a$ . Thus  $\ulcorner f \urcorner$  is the unique arrow making

$$\begin{array}{ccc} b^a \times a & & \\ \uparrow \ulcorner f \urcorner \times 1_a & \searrow ev & \\ 1 \times a & \xrightarrow{f \circ pr_a} & b \end{array}$$

commute. Then we have that for any  $\mathcal{C}$ -element  $x : 1 \rightarrow a$  of  $a$ ,

$$ev \circ \langle [f], x \rangle = f \circ x.$$

EXERCISE 2. Prove this last statement.

## 4.2. Classifying subobjects

In set theory, the powerset  $\mathcal{P}(D)$  is often denoted  $2^D$ . The later symbol, according to our earlier definition, in fact denotes the collection of all functions from  $D$  to  $2 = \{0, 1\}$ . The justification for the usage is that  $\mathcal{P}(D) \cong 2^D$ , i.e. there is a bijective correspondence between subsets of  $D$  and functions  $D \rightarrow 2$ . This isomorphism is established as follows: given a subset  $A \subseteq D$ , we define the function  $\chi_A : D \rightarrow 2$ , called the *characteristic function of  $A$* , by the rule, “for those elements of  $D$  in  $A$ , give output 1 and for those not in  $A$ , give output 0”. i.e.

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

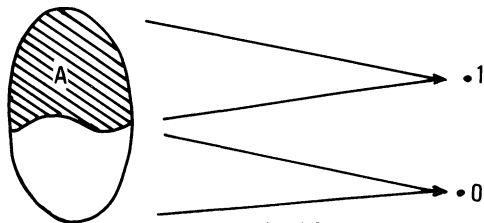


Fig. 4.2.

The assignment of  $\chi_A$  to  $A$  is injective from  $\mathcal{P}(D)$  to  $2^D$ , i.e. if  $\chi_A = \chi_B$  then  $A = B$  (why?). It is also surjective, for if  $f \in 2^D$ , then  $f = \chi_{A_f}$ , where

$$A_f = \{x : x \in D \text{ and } f(x) = 1\}.$$

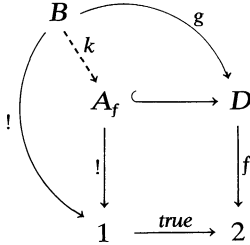
This correspondence between subset and characteristic function can be “captured” by a pullback diagram. The set  $A_f$  just defined is the inverse image under  $f$  of the subset  $\{1\}$  of  $\{0, 1\}$ , i.e.

$$A_f = f^{-1}(\{1\}),$$

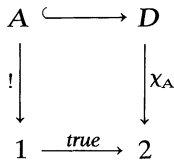
and so according to §3.13

$$\begin{array}{ccc} A_f & \hookrightarrow & D \\ \downarrow ! & & \downarrow f \\ \{1\} & \hookrightarrow & 2 \end{array}$$

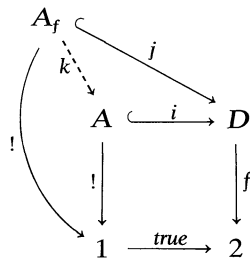
is a pullback square, i.e.  $A_f$  arises by pulling back  $\{1\} \hookrightarrow 2$  along  $f$ . We are going to modify this picture slightly. The bottom arrow, which outputs the element 1 of  $\{0, 1\}$  is replaced by the function from  $1 = \{0\}$  to  $2 = \{0, 1\}$  that outputs 1. We give this function the name *true*, for reasons that will emerge in Chapter 6. It has the rule;  $true(0) = 1$ . Then the inner square of



is a pullback. To see this, suppose the “outer square” commutes for some  $g$ . Then if  $b \in B$ ,  $f(g(b)) = true(! (b)) = 1$ , so  $g(b) \in A_f$ . Hence  $k : B \rightarrow A_f$  can be defined by the rule  $k(b) = g(b)$ . This  $k$  makes the whole diagram commute, and is clearly the only one that could do so. It follows that if  $A \subseteq D$ , then



is a pullback, since pulling *true* back along  $\chi_A$  yields the set  $\{x : \chi_A(x) = 1\}$ , which is just  $A$ . But more than this follows –  $\chi_A$  can be identified as the *one and only* function from  $D$  to  $2$  that makes the above diagram a pullback, i.e. the only function along which *true* pulls back to yield  $A$ . If, for some  $f$ , the inner square of



is a pullback, then for  $x \in A$ ,  $f(x) = 1$ , so  $x \in A_f$ . Hence  $A \subseteq A_f$ . But the outer square commutes – indeed it is a pullback as we saw above – and so

the unique  $k$  exists with  $i \circ k = j$ . Since  $i$  and  $j$  are inclusions,  $k$  must be as well. Thus  $A_f \subseteq A$ , and altogether  $A = A_f$ . But  $f$  is the characteristic function of  $A_f$ , and so,  $f = \chi_A$ .

So the set  $2$  together with the function  $true : 1 \rightarrow 2$  play a special role in the transfer from subset to characteristic function, a role that has been cast in the language of categories, in such a way as to lead to an abstract definition:

**DEFINITION.** If  $\mathcal{C}$  is a category with a terminal object  $1$ , then a *subobject classifier* for  $\mathcal{C}$  is a  $\mathcal{C}$ -object  $\Omega$  together with a  $\mathcal{C}$ -arrow  $true : 1 \rightarrow \Omega$  that satisfies the following axiom.

**$\Omega$ -AXIOM.** For each monic  $f : a \rightarrow d$  there is one and only one  $\mathcal{C}$ -arrow  $\chi_f : d \rightarrow \Omega$  such that

$$\begin{array}{ccc}
 a & \xrightarrow{f} & d \\
 \downarrow ! & & \downarrow \chi_f \\
 1 & \xrightarrow{true} & \Omega
 \end{array}$$

is a pullback square.

The arrow  $\chi_f$  is called the *characteristic arrow*, or the *character*, of the monic  $f$  (subobject of  $d$ ). The arrow  $true$  will often be denoted by the letter “ $\top$ ”.

A *subobject classifier*, when it exists in a category, is *unique up to isomorphism*. If  $\top : 1 \rightarrow \Omega$  and  $\top' : 1 \rightarrow \Omega'$  are both subobject classifiers we have the diagram

$$\begin{array}{ccc}
 1 & \xrightarrow{\top} & \Omega \\
 \downarrow & & \downarrow \chi'_{\top} \\
 1 & \xrightarrow{\top'} & \Omega' \\
 \downarrow & & \downarrow \chi_{\top'} \\
 1 & \xrightarrow{\top} & \Omega
 \end{array}$$

The top square is the pullback that gives the character  $\chi'_{\top}$  of  $\top$  using  $\top'$  as classifier (remember any arrow with  $\text{dom}=1$  is monic). The bottom

square is the pullback that gives the character of  $\top'$ , when  $\top$  is used as classifier.

Hence by the PBL (§3.13, Example 8) the outer rectangle

$$\begin{array}{ccc}
 1 & \xrightarrow{\top} & \Omega \\
 \downarrow & & \downarrow \chi_{\top'} \circ \chi'_{\top} \\
 1 & \xrightarrow{\top} & \Omega
 \end{array}$$

is a pullback. But by the  $\Omega$ -axiom there is only one arrow  $\Omega \rightarrow \Omega$  making this square a pullback, and  $1_{\Omega}$  would do that job (why?) Thus  $\chi_{\top'} \circ \chi'_{\top} = 1_{\Omega}$ . Interchanging  $\top$  and  $\top'$  in this argument gives

$$\chi'_{\top} \circ \chi_{\top'} = 1_{\Omega'}$$

and so  $\chi_{\top'}: \Omega' \cong \Omega$ .

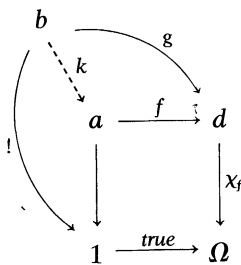
Since  $\top' = \chi'_{\top} \circ \top$  we have that any two subobject classifiers may be obtained from each other by composing with an iso arrow between their codomains.

The assignment of  $\chi_f$  to  $f$  establishes a one-one correspondence between subobjects of an object  $d$ , and arrows  $d \rightarrow \Omega$ , as shown by:

**THEOREM.** For  $f: a \rightarrow d$  and  $g: b \rightarrow d$ ,

$$f \simeq g \quad \text{iff} \quad \chi_f = \chi_g.$$

**PROOF.** Suppose first that  $\chi_f = \chi_g$ . Consider



Since  $\chi_f = \chi_g$ , the outer square commutes (indeed is a pullback) and so as the inner square is a pullback there exists  $k$  factoring  $g$  through  $f$ , hence  $g \subseteq f$ . Interchanging  $f$  and  $g$  on the diagram leads to  $f \subseteq g$  and altogether  $f \simeq g$ .

Conversely if  $f \simeq g$ , then the arrow  $k$  in the above diagram does exist and is iso with an inverse  $k^{-1}: a \cong b$ . Using this one can show that the



outer square is a pullback, which can only be so if  $\chi_f$  is the unique character of  $g$ ,  $\chi_f = \chi_g$ . □

Thus the assignment of  $\chi_f$  to  $f$  (more exactly to  $[f]$ ) injects  $\text{Sub}(d)$  into  $\mathcal{E}(d, \Omega)$ . But given any  $h : d \rightarrow \Omega$ , if we pull *true* back along  $h$ ,

$$\begin{array}{ccc} a & \xrightarrow{f} & d \\ \downarrow & & \downarrow h \\ 1 & \xrightarrow{\text{true}} & \Omega \end{array}$$

the resulting arrow  $f$  will be monic (since *true* is monic and the pullback of a monic is always itself monic – Exercise, §3.13). Hence  $h$  must be  $\chi_f$ . So in a category where these constructions are possible we get

$$\text{Sub}(d) \cong \mathcal{E}(d, \Omega).$$

NOTATION. For any  $\mathcal{C}$ -object  $a$ , the composite  $\text{true} \circ !_a$ , of arrows  $! : a \rightarrow 1$  and *true*, will be denoted  $\text{true}_a$ , or  $\top_a$ , or sometimes *true*!

$$\begin{array}{ccc} a & \longrightarrow & 1 \\ & \searrow \text{true}_a & \downarrow \text{true} \\ & & \Omega \end{array}$$

EXERCISE 1. Show that the character of  $\text{true} : 1 \rightarrow \Omega$  is  $1_\Omega$

$$\begin{array}{ccc} 1 & \xrightarrow{\text{true}} & \Omega \\ \downarrow & & \downarrow 1_\Omega \\ 1 & \xrightarrow{\text{true}} & \Omega \end{array}$$

i.e.  $\chi_{\text{true}} = 1_\Omega$ .

EXERCISE 2. Show that  $\chi_{1_\Omega} = \text{true}_\Omega = \text{true} \circ !_\Omega$ .

$$\begin{array}{ccc} \Omega & \xrightarrow{!_\Omega} & \Omega \\ \downarrow !_\Omega & & \downarrow \text{true} \circ !_\Omega \\ 1 & \xrightarrow{\text{true}} & \Omega \end{array}$$

EXERCISE 3. Show that for any  $f: a \rightarrow b$ ,

$$\begin{array}{ccc}
 a & \xrightarrow{f} & b \\
 \searrow \text{true}_a & & \swarrow \text{true}_b \\
 & \Omega & 
 \end{array}$$

$$\text{true}_b \circ f = \text{true}_a.$$

□

### 4.3. Definition of topos

DEFINITION. An *elementary topos* is a category  $\mathcal{E}$  such that

- (1)  $\mathcal{E}$  is finitely complete,
- (2)  $\mathcal{E}$  is finitely co-complete,
- (3)  $\mathcal{E}$  has exponentiation,
- (4)  $\mathcal{E}$  has a subobject classifier.

As observed in Chapter 3, (1) and (3) constitute the definition of “Cartesian closed”, while (1) can be replaced by

- (1')  $\mathcal{E}$  has a terminal object and pullbacks,

and dually (2) replaced by

- (2')  $\mathcal{E}$  has an initial object  $0$ , and pushouts.

The definition just given is the one originally proposed by Lawvere and Tierney, in terms of which they started topos theory in 1969. Subsequently C. Juul Mikkelsen discovered that condition (2) is implied by the combination of (1), (3) and (4) (cf. Paré [74]). Thus a topos can be defined as a Cartesian closed category with a subobject classifier. In §4.7 we shall consider a different definition, based on a categorial characterisation of power sets.

The word “elementary” (which from now on will be understood) has a special technical meaning to do with the nature of the definition of topos. This usage will be explained in Chapter 11.

The list of topoi that follows in this chapter is intended to illustrate the generality of the concept. By no means all of the detail is given – for the most part we concentrate on the structure of the subobject classifier.

### 4.4. First examples

EXAMPLE 1. **Set** is a topos – the prime example and the motivation for the concept in the first place.

EXAMPLE 2. **Finset** is a topos, with limits, exponentials, and  $\top:1 \rightarrow \Omega$  exactly as in **Set**.

EXAMPLE 3. **Finord** is a topos. Every finite set is isomorphic to some finite ordinal ( $A \cong n$  if  $A$  has  $n$  elements). Hence all categorial constructions in **Finset** “transfer” into **Finord** (as we have already observed for product, exponentials). The subobject classifier in **Finord** is the same function  $true: \{0\} \rightarrow \{0, 1\}$  as in **Finset** and **Set**.

EXAMPLE 4. **Set**<sup>2</sup>, the category of pairs of sets is a topos. All constructions are obtained by “doubling up” the corresponding constructions in **Set** (cf. Example 10, §2.5).

A terminal object is a pair  $\langle \{0\}, \{0\} \rangle$  of singleton sets. Given two arrows  $\langle f, g \rangle: \langle A, B \rangle \rightarrow \langle E, F \rangle, \langle h, k \rangle: \langle C, D \rangle \rightarrow \langle E, F \rangle$  with common codomain in **Set**<sup>2</sup>, form the pullbacks

$$\begin{array}{ccc}
 P & \xrightarrow{j} & C \\
 i \downarrow & & \downarrow h \\
 A & \xrightarrow{f} & E
 \end{array}
 \qquad
 \begin{array}{ccc}
 Q & \xrightarrow{v} & D \\
 u \downarrow & & \downarrow k \\
 B & \xrightarrow{g} & F
 \end{array}$$

in **Set**. Then

$$\begin{array}{ccc}
 \langle P, Q \rangle & \xrightarrow{\langle j, v \rangle} & \langle C, D \rangle \\
 \langle i, u \rangle \downarrow & & \downarrow \langle h, k \rangle \\
 \langle A, B \rangle & \xrightarrow{\langle f, g \rangle} & \langle E, F \rangle
 \end{array}$$

will be a pullback in **Set**<sup>2</sup>.

The exponential has

$$\langle C, D \rangle^{\langle A, B \rangle} = \langle C^A, D^B \rangle$$

with evaluation arrow from

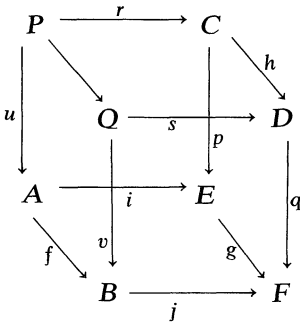
$$\langle C, D \rangle^{\langle A, B \rangle} \times \langle A, B \rangle = \langle C^A \times A, D^B \times B \rangle$$

to  $\langle C, D \rangle$  as the pair  $\langle e, f \rangle$  where  $e: C^A \times A \rightarrow C$  and  $f: D^B \times B \rightarrow D$  are the appropriate evaluation arrows in **Set**.

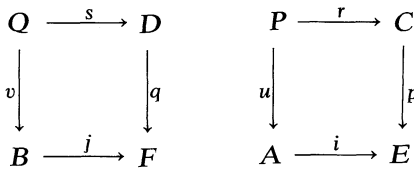
The subobject classifier is  $\langle \top, \top \rangle: \langle \{0\}, \{0\} \rangle \rightarrow \langle 2, 2 \rangle$ . The category **Set** plays no special role here. If  $\mathcal{E}_1$  and  $\mathcal{E}_2$  are any topoi, then the product category  $\mathcal{E}_1 \times \mathcal{E}_2$  is a topos.

**EXAMPLE 5.** **Set** $^\rightarrow$ , the category of functions. The terminal object is the identity function  $\text{id}_{\{0\}}$  from  $\{0\}$  to  $\{0\}$ .

*Pullback:* Consider the “cube”

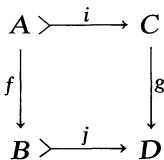


$f, g, h$  are given as **Set** $^\rightarrow$ -objects with  $\langle i, j \rangle$  an arrow from  $f$  to  $g$ ,  $\langle p, q \rangle$  an arrow from  $h$  to  $g$ . The rest of the diagram obtains by forming the pullbacks



in **Set**. The arrow  $k$  exists by the universal property of the pullback of  $j$  and  $q$ . Then in **Set** $^\rightarrow$  the arrows  $\langle u, v \rangle$  and  $\langle r, s \rangle$  are the pullbacks of  $\langle i, j \rangle$  and  $\langle p, q \rangle$ .

*Classifier:* If  $f: A \rightarrow B$  is a subobject of  $g: C \rightarrow D$  in **Set** $^\rightarrow$  then there is a commutative **Set** diagram



We will take the monics to be actual inclusions, so that  $A \subseteq C$ ,  $B \subseteq D$  and  $f$  is the restriction of  $g$ , i.e.  $f(x) = g(x)$  for  $x \in A$ . The picture is

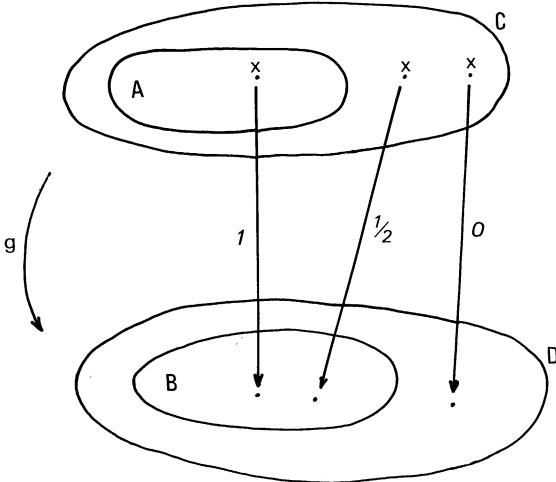


Fig. 4.3.

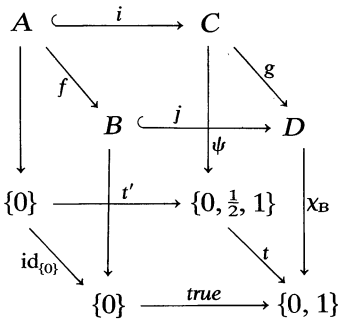
An element  $x$  of  $C$  can be classified now in three ways. Either

- (i)  $x \in A$ , or
- (ii)  $x \notin A$ , but  $g(x) \in B$ , or
- (iii)  $x \notin A$ , and  $g(x) \notin B$ .

So we introduce a 3-element set  $\{0, \frac{1}{2}, 1\}$  and define  $\Psi: C \rightarrow \{0, \frac{1}{2}, 1\}$  by

$$\psi(x) = \begin{cases} 1 & \text{if (i) holds} \\ \frac{1}{2} & \text{if (ii) holds} \\ 0 & \text{if (iii) holds} \end{cases}$$

We can now form the cube



where  $true(0) = t'(0) = 1$ ,  $t: \{0, \frac{1}{2}, 1\} \rightarrow \{0, 1\}$  has  $t(0) = 0$ , and  $t(1) = t(\frac{1}{2}) = 1$ .  $\chi_B$  is the characteristic function of  $B$ .

The base of the cube displays the subobject classifier  $\top: 1 \rightarrow \Omega$  for  $\mathbf{Set}^{\rightarrow}$ .  $\top$  is the pair  $\langle t', true \rangle$  from  $1 = id_{\{0\}}$  to  $\Omega = t: \{0, \frac{1}{2}, 1\} \rightarrow \{0, 1\}$ .

The front and back faces of the cube are each pullbacks in **Set**. The whole diagram exhibits  $\langle \psi, \chi_B \rangle$  as the character in  $\mathbf{Set}^\rightarrow$  of the monic  $\langle i, j \rangle$ .

*Exponentiation:* Let  $f: A \rightarrow B$ ,  $g: C \rightarrow D$  be two  $\mathbf{Set}^\rightarrow$ -objects. Then  $g^f$  is the  $\mathbf{Set}^\rightarrow$ -object (function)  $g^f: E \rightarrow F$ , where

$$F = D^B \quad (\text{exponential in } \mathbf{Set})$$

$E$  is the collection of all  $\mathbf{Set}^\rightarrow$ -arrows from  $f$  to  $g$  i.e.

$$E = \{ \langle h, k \rangle : \begin{array}{ccc} A & \xrightarrow{h} & C \\ \downarrow f & & \downarrow g \\ B & \xrightarrow{k} & D \end{array} \text{ commutes} \}$$

and

$$g^f(\langle h, k \rangle) = k.$$

The product object of  $g^f$  and  $f$  in  $\mathbf{Set}^\rightarrow$  is the product map

$$g^f \times f: E \times A \rightarrow F \times B \quad (\text{cf. §3.8})$$

and the evaluation arrow from  $g^f \times f$  to  $g$  is the pair  $\langle u, v \rangle$

$$\begin{array}{ccc} E \times A & \xrightarrow{u} & C \\ \downarrow g^f \times f & & \downarrow g \\ F \times B & \xrightarrow{v} & D \end{array}$$

where  $v$  is the usual evaluation arrow in **Set**, and  $u$  takes input  $\langle \langle h, k \rangle, x \rangle$  to output  $h(x)$ .

The constructions just given for  $\top: 1 \rightarrow \Omega$  and  $g^f$  will be seen in Chapter 9 to be instances of a more general definition that yields a whole family of topoi.

## 4.5. Bundles and sheaves

One of the primary sources of topos theory is algebraic geometry, in particular the study of *sheaves*. To understand what a sheaf is requires some knowledge of topology and the full story about sheaves and their relation to topoi would take us beyond our present scope. The idea is closely tied up with models of intuitionistic logic, but is much more general than that. Indeed, sheaf theory constitutes a whole conceptual framework and language of its own, and to ignore it completely, even at

this stage, would be to distort the overall significance and point of view of topos theory.

For the benefit of the reader unfamiliar with topology we shall delay its introduction and first consider the underlying set-theoretic structure of the sheaf concept, to be called a *bundle*.

Let us assume we have a collection  $\mathcal{A}$  of sets, no two of which have any elements in common. That is, any two members of  $\mathcal{A}$  are sets that are disjoint. We need a convenient notation for referring to these sets so we presume we have a set  $I$  of *labels*, or *indices*, for them. For each index  $i \in I$ , there is a set  $A_i$  that belongs to our collection, and each member of  $\mathcal{A}$  is labelled in this way, so we write  $\mathcal{A}$  as the collection of all these  $A_i$ 's,

$$\mathcal{A} = \{A_i : i \in I\}.$$

The fact that the members of  $\mathcal{A}$  are pairwise disjoint is expressed by saying that for *distinct* indices  $i, j \in I$

$$A_i \cap A_j = \emptyset$$

We visualise the  $A_i$ 's as "sitting over" the index set  $I$  thus:

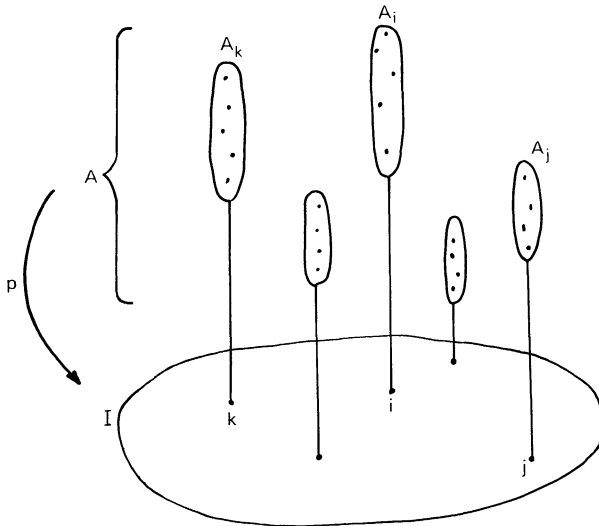


Fig. 4.4.

If we let  $A$  be the union of all the  $A_i$ 's, i.e.

$$A = \{x : \text{for some } i, x \in A_i\}$$

then there is an obvious map  $p : A \rightarrow I$ . If  $x \in A$  then there is exactly one  $A_i$  such that  $x \in A_i$ , by the disjointness condition. We put  $p(x) = i$ . Thus

all the members of  $A_i$  get mapped to  $i$ , all the members of  $A_j$  to  $j$ , etc. We can then re-capture  $A_i$  as the inverse image under  $p$  of  $\{i\}$ , for

$$p^{-1}(\{i\}) = \{x: p(x) = i\} = A_i.$$

The set  $A_i$  is called the *stalk*, or *fibre* over  $i$ . The members of  $A_i$  are called the *germs* at  $i$ . The whole structure is called a *bundle* of sets over the *base space*  $I$ . The set  $A$  is called the *stalk space* (l'espace étalé) of the bundle. The reason for the botanical terminology is evident – what we have is a bundle of stalks, each with its own head of germs (think of a bunch of asparagus spears).

This construction looks rather special, but it is to be found whenever there are functions. We have just seen that a bundle has an associated map  $p$  from its stalk space to the base. (If in fact every stalk is nonempty then  $p$  will be surjective, but in general we will allow the possibility that  $A_i = \emptyset$ ). Conversely, if  $p: A \rightarrow I$  is an arbitrary function from some set  $A$  to  $I$ , then we can *define*  $A_i$  to be  $p^{-1}(\{i\})$ , for each  $i \in I$ , and *define*

$$\mathcal{A} = \{p^{-1}(\{i\}): i \in I\} = \{A_i: i \in I\}.$$

Then  $\mathcal{A}$  is a bundle of sets over  $I$  whose stalk space is the original  $A$ , and induced map  $A \rightarrow I$  the original  $p$  (the stalks are disjoint, as no  $x \in A$  can have two different  $p$ -outputs).

So a bundle of sets over  $I$  is “essentially just” a function with codomain  $I$ . The two are not of course identical conceptually. To construe a function as a bundle is to offer a new, and provocative, perspective. To emphasise that, we will introduce a new name  $\mathbf{Bn}(I)$  for the category of bundles over  $I$ , although we have already described it in Example 12 of Chapter 2 as the Comma category  $\mathbf{Set} \downarrow I$  of functions with codomain  $I$ . Thus the  $\mathbf{Bn}(I)$ -objects are the pairs  $(A, f)$ , where  $f: A \rightarrow I$  is a set function and the arrows  $k: (A, f) \rightarrow (B, g)$  have  $k: A \rightarrow B$  such that

$$\begin{array}{ccc} A & \xrightarrow{k} & B \\ f \searrow & & \swarrow g \\ & I & \end{array}$$

commutes, i.e.  $g \circ k = f$ . This means that if  $f(x) = i$ , for  $x \in A$ , then  $g(k(x)) = i$ , i.e. if  $x \in A_i$ , then  $k(x) \in B_i$ . Thus  $k$  maps germs at  $i$  in  $(A, f)$  to germs at  $i$  in  $(B, g)$ .

Now a topos is to be thought of as a generalisation of the category  $\mathbf{Set}$ . An object in a topos is a “generalised set”. A “set” in the topos  $\mathbf{Bn}(I)$  is a bundle of ordinary sets. Many categorial notions when applied to  $\mathbf{Bn}(I)$



prove to be bundles of the corresponding entities in **Set**, as we shall now see.

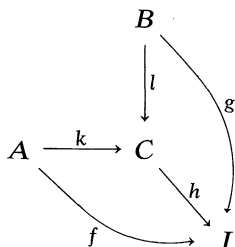
The terminal object  $1$  for  $\mathbf{Bn}(I)$  is  $\text{id}_I : I \rightarrow I$ , and for any bundle  $(A, f)$ , the unique arrow  $(A, f) \rightarrow (I, \text{id}_I)$  is  $f : A \rightarrow I$  itself (cf. §3.6). Now the stalk of  $\text{id}_I$  over  $i$  is  $\text{id}^{-1}(\{i\}) = \{i\}$ , which is terminal in **Set**. Thus the  $\mathbf{Bn}(I)$  terminal is a bundle of **Set**-terminals over  $I$ , and the unique arrow  $f : (A, f) \rightarrow (I, \text{id}_I)$  can be construed as a bundle

$$\{f_i : i \in I\}$$

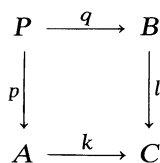
of unique **Set**-arrows, where

$$f_i = ! : f^{-1}(\{i\}) \rightarrow \{i\}.$$

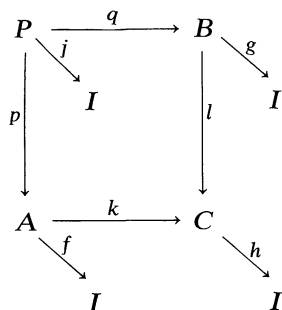
*Pullback:* Given  $\mathbf{Bn}(I)$ -arrows  $k : \langle A, f \rangle \rightarrow \langle C, h \rangle$  and  $l : \langle B, g \rangle \rightarrow \langle C, h \rangle$ , so that



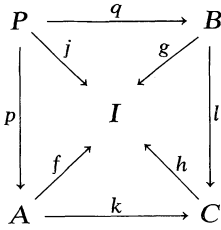
commutes, form the pullback



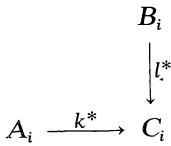
in **Set** of  $k$  and  $l$ . Then



is a pullback of  $k$  and  $l$  in  $\mathbf{Bn}(I)$ , where  $j = f \circ p = h \circ k \circ p = h \circ l \circ q = g \circ q$ . The diagram is probably more usefully given as the commutative



Now if  $A_i, B_i, C_i$  are the stalks over  $i$  for the bundles  $f, g, h$ , then the pullback of



has domain  $\{(x, y) : x \in A_i, y \in B_i, \text{ and } k(x) = l(y)\}$  which can be seen to be the same as

$$\{(x, y) : x \in A, y \in B \text{ and } j\langle x, y \rangle = i\} = j^{-1}(\{i\}),$$

which is the stalk over  $i$  of  $j : P \rightarrow I$ .

Thus the pullback object  $(P, j)$  is a bundle of pullbacks from  $\mathbf{Set}$ .

*Subobject classifier:* The classifier for  $\mathbf{Bn}(I)$  is a bundle of two-element sets, i.e. a bundle of  $\mathbf{Set}$ -classifiers.

We define  $\Omega = (2 \times I, p_I)$ , where  $p_I : 2 \times I \rightarrow I$  is the projection  $p_I(\langle x, y \rangle) = y$  onto the “second factor”. Now the product set  $2 \times I$  is in fact the (disjoint) union of the sets

$$\{0\} \times I = \{\langle 0, i \rangle : i \in I\}$$

and

$$\{1\} \times I = \{\langle 1, i \rangle : i \in I\},$$

each isomorphic to  $I$ , and we visualise  $\Omega$  as shown in Fig. 4.5. The stalk over a particular  $i$  is the two-element set

$$\Omega_i = \{\langle 0, i \rangle, \langle 1, i \rangle\} = 2 \times \{i\}.$$

The classifier arrow  $\top : 1 \rightarrow \Omega$  can be thought of as a bundle of copies of the set function *true*. We define  $\top : I \rightarrow 2 \times I$  by

$$\top(i) = \langle 1, i \rangle.$$

In terms of the limit approach to products,  $\top$  is the product map  $\langle \text{true}!, \text{id}_I \rangle$  of  $\text{true} \circ ! : I \rightarrow \{0\} \rightarrow \{0, 1\}$  and  $\text{id}_I$ .

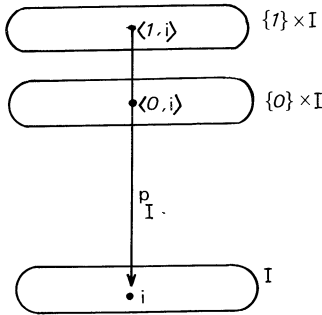
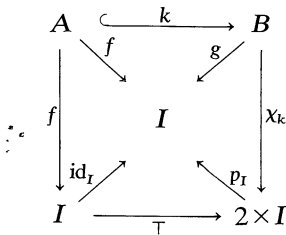


Fig. 4.5.

To see how  $\top$  classifies subobjects we take a monic  $k : (A, f) \rightarrow (B, g)$  in  $\mathbf{Bn}(I)$ , and in fact suppose that  $k$  is an inclusion, i.e.  $A \subseteq B$  and  $f(x) = g(x)$ , all  $x \in A$ . We wish to define the character  $\chi_k : (B, g) \rightarrow \Omega = (2 \times I, p_I)$  so that



commutes and gives a pullback in  $\mathbf{Bn}(I)$ . Now any  $x \in B$  is classified according to whether  $x \in A$  or  $x \notin A$ .

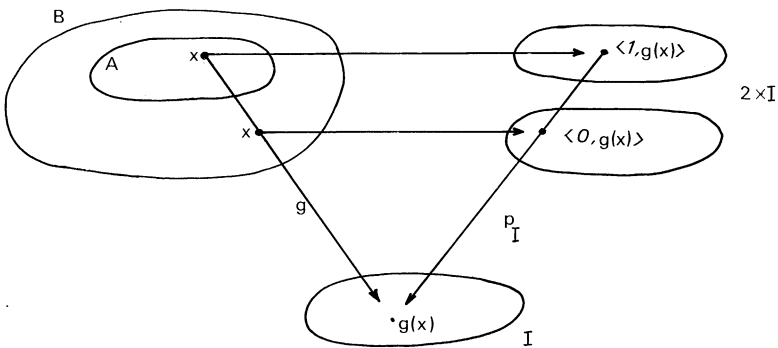


Fig. 4.6.

We make  $\chi_k$  assign as “1” or “0” accordingly, and also make these choices in the right stalks, so that  $p_I \circ \chi_k = g$ . Formally,  $\chi_k : B \rightarrow 2 \times I$  is the product map  $\langle \chi_A, g \rangle : B \rightarrow 2 \times I$ , where  $\chi_A : B \rightarrow 2$  is the usual characteristic function of  $A$ , i.e.

$$\chi_k(x) = \begin{cases} \langle 1, g(x) \rangle & \text{if } x \in A \\ \langle 0, g(x) \rangle & \text{if } x \notin A. \end{cases}$$

EXERCISE 1. Verify that this construction satisfies the  $\Omega$ -axiom. □

*Sections:* The function  $\top: I \rightarrow 2 \times I$  has an interesting property – for input  $i$  the output  $\top(i) = \langle 1, i \rangle$  is a germ at  $i$ . Such a function from the base set  $I$  to the stalk space that picks one germ out of each stalk is called a *section* of the bundle. In general  $s: I \rightarrow A$  is a section of bundle  $f: A \rightarrow I$  if  $s(i) \in A_i = f^{-1}(\{i\})$ , for all  $i \in I$ . This means precisely that  $f(s(i)) = i$ , all  $i$ , and hence that

$$\begin{array}{ccc}
 I & \xrightarrow{s} & A \\
 \searrow & & \swarrow \\
 & \text{id}_I & f \\
 & & I
 \end{array}$$

commutes. So another way of looking at a section is to say that it is a **Bn**( $I$ )-arrow from the terminal  $(I, \text{id}_I)$  to  $(A, f)$ . Thus a section of the bundle  $(A, f)$  is an *element* of the **Bn**( $I$ )-object  $(A, f)$  in the sense of the definition at the end of §4.1. But our initial picture of a section is a bundle of germs, one from each stalk. So an “element” in **Bn**( $I$ ) is a bundle of ordinary elements.

Elements of  $\Omega$ , i.e. arrows  $1 \rightarrow \Omega$ , in any topos  $\mathcal{E}$  are known as the *truth-values* of  $\mathcal{E}$ , and have a special role in the logical structure of  $\mathcal{E}$  (See Chapter 6). We know (§4.2) that there is a bijective correspondence  $\text{Sub}(1) \cong \mathcal{E}(1, \Omega)$  between elements of  $\Omega$  and subobjects of 1. Now in **Bn**( $I$ ) a subobject  $k: (A, f) \twoheadrightarrow 1$  of 1 must have

$$\begin{array}{ccc}
 A & \twoheadrightarrow k & I \\
 \searrow & & \swarrow \\
 & f & \text{id}_I \\
 & & I
 \end{array}$$

commuting, so  $k = f$ . Thus a subobject of 1 can be identified with an injective function  $f: A \twoheadrightarrow I$ , i.e. with a subobject of  $I$  in **Set**. The latter of course is essentially a subset of  $I$ , and we conclude that there is a bijection

$$\mathcal{P}(I) \cong \mathbf{Bn}(I)(1, \Omega)$$

i.e. we may identify truth-values (elements of  $\Omega$ ) in **Bn**( $I$ ) with subsets of  $I$ . It is instructive to spell this out fully:

Given  $A \subseteq I$ , let  $S_A: I \rightarrow 2 \times I$  be the product map  $\langle \chi_A, \text{id}_I \rangle$ , i.e.

$$S_A(i) = \begin{cases} \langle 1, i \rangle & \text{if } i \in A \\ \langle 0, i \rangle & \text{if } i \notin A \end{cases}$$

then  $S_A$  is a section of  $\Omega$ , whose image is shown shaded in the picture.

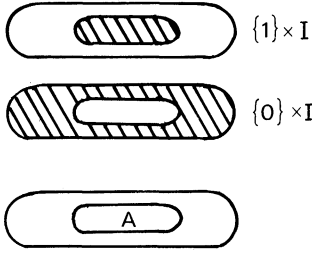


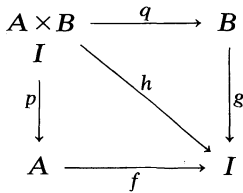
Fig. 4.7.

The assignment of  $S_A : 1 \rightarrow \Omega$  to  $A$  is injective (exercise). Moreover if  $S : 1 \rightarrow \Omega$  is any section, and  $A = \{i : S(i) = \langle 1, i \rangle\}$ , then  $S = S_A$ , so the assignment is also surjective.

Note that whereas **Set** has two truth values,  $\mathcal{P}(I)$  may well be infinite (it certainly will be if  $I$  is infinite).

EXERCISE 2. What are the truth-values in **Set**<sup>2</sup> and in **Set**<sup>→</sup>? □

*Products.* Let  $(A, f)$  and  $(B, g)$  be bundles over  $I$  and form the pullback

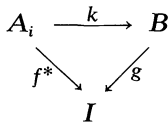


Then  $(A \times_I B, h)$  is the product of  $(A, f)$  and  $(B, g)$  in **Bn**( $I$ ), where  $h = f \circ p = g \circ q$ , and has projection arrows  $p$  and  $q$ . Note that the stalk (fibre) over  $i$  is

$$\{(x, y) : f(x) = g(y) = i\} = A_i \times B_i,$$

the product of the fibres over  $i$  in  $(A, f)$  and  $(B, g)$ . Hence the name “fibred product” that is sometimes used for “pullback”.

*Exponentials.* Given bundles  $f : A \rightarrow I$  and  $g : B \rightarrow I$  we form their exponential as a bundle of the exponentials  $B_i^{A_i}$  of the stalks of  $A$  and  $B$ . More precisely let  $D_i$  be the collection of functions  $k : A_i \rightarrow B$  such that



commutes and so  $k$  carries  $A_i$  into the stalk  $B_i$  of  $g$  over  $i$  (where, as previously,  $f^*$  denotes a function that has the same rule as  $f$  but may vary as to domain or codomain). Now the  $D_i$ 's may not be pairwise disjoint, so we define  $E_i = \{i\} \times D_i$ , for each  $i$ , and then  $\{E_i : i \in I\}$  is a bundle. The induced function  $p : E \rightarrow I$  where  $E$  is the union of the  $E_i$ 's has  $p(\langle i, k \rangle) = i$ .  $(E, p)$  is the exponential

$$(B, g)^{(A, f)}.$$

The evaluation arrow  $ev : (E, p) \times (A, f) \rightarrow (B, g)$  is the function  $ev : E \times_I A \rightarrow B$ , where

$$ev(\langle \langle i, k \rangle, x \rangle) = k(x).$$

The reader who has the patience to wade through the details of checking that this construction is well defined and satisfies the definition of exponentiation will no doubt get his reward in heaven. For the present he will perhaps appreciate the advantages of the categorial viewpoint, wherein all we need to say about the exponential, to know what it is, is that it satisfies the universal property described in §3.16. (We shall return to this example in Chapter 15).

**FUNDAMENTAL THEOREM.** *Not only is  $\mathbf{Bn}(I) = \mathbf{Set} \downarrow I$  a topos, but more generally if  $\mathcal{E}$  is any topos and  $a$  an  $\mathcal{E}$ -object, then the category  $\mathcal{E} \downarrow a$  of  $\mathcal{E}$ -arrows over  $a$  (§2.5, Example 12) is also a topos.*

This fact has been called the *Fundamental Theorem of Topoi* by Freyd [72]. The reader can probably sort out many of the details from the above, e.g. if  $\tau : 1 \rightarrow \Omega$  is the classifier in  $\mathcal{E}$ , then in  $\mathcal{E} \downarrow a$  it is  $\langle \tau_a, 1_a \rangle$ , i.e.

$$\begin{array}{ccc} a & \xrightarrow{\langle \tau_a, 1_a \rangle} & \Omega \times a \\ & \searrow 1_a & \swarrow pr_a \\ & & a \end{array}$$

The definition of exponentials in  $\mathcal{E} \downarrow a$  would carry us too far afield at present. It requires the development of a categorial theory of "partial functions" and their classification, which will be considered in Chapters 11 and 15.

## Sheaves

A *sheaf* is a bundle with some additional topological structure. Let  $I$  be a topological space, with  $\mathcal{O}$  its collection of open sets. A *sheaf* over  $I$  is a

pair  $(A, p)$  where  $A$  is a topological space and  $p : A \rightarrow I$  is a continuous map that is a local homeomorphism. This means that each point  $x \in A$  has an open neighbourhood  $U$  in  $A$  that is mapped homeomorphically by  $p$  onto  $p(U) = \{p(y) : y \in U\}$ , and the latter is open in  $I$ . The category  $\mathbf{Top}(I)$  of sheaves over  $I$  has such pairs  $(A, p)$  as objects, and as arrows  $k : (A, p) \rightarrow (B, q)$  the continuous maps  $k : A \rightarrow B$  such that

$$\begin{array}{ccc} A & \xrightarrow{k} & B \\ p \searrow & & \swarrow q \\ & I & \end{array}$$

commutes. Such a  $k$  is in fact an open map (as is a local homeomorphism) and in particular  $\text{Im } k = k(A)$  will be an open subset of  $B$ .

$\mathbf{Top}(I)$  is a topos, known as a *spatial topos*. The terminal object is  $\text{id}_I : I \rightarrow I$ . The subobject classifier is the *sheaf of germs of open sets in  $I$* . Its construction illustrates a common method of building a bundle over  $I$ . There will be some ambient set  $X$  and each point  $i \in I$  will determine an equivalence relation  $\sim_i$  on  $X$ . The stalk over  $i$  will then be defined as the quotient set  $X/\sim_i$  of equivalence classes of  $X$  under  $\sim_i$ .

In the present case  $X$  is the collection  $\Theta$  of open sets in  $I$ . At  $i \in I$ , we define  $\sim_i$  by declaring, for  $U, V \in \Theta$

$$U \sim_i V \text{ iff there is some open set } W \text{ such that } i \in W \text{ and } U \cap W = V \cap W$$

Then  $\sim_i$  is an equivalence relation. The intuitive idea is that  $U \sim_i V$  when the points in  $U$  that are close to  $i$  are the same as those that are in  $V$  and close to  $i$ , i.e. “locally” around  $i$ ,  $U$  and  $V$  look the same, i.e. the statement “ $U = V$ ” is “locally true” at  $i$ .

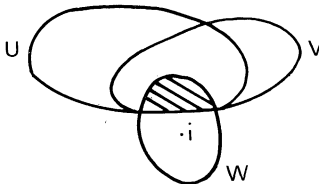


Fig. 4.8.

The equivalence class

$$[U]_i = \{V : U \sim_i V\}$$

is called the *germ of  $U$  at  $i$* . Intuitively it “represents” the collection of points in  $U$  that are “close” to  $i$ .

We then take as the stalk over  $i$ ,

$$\Omega_i = \{\langle i, [U]_i \rangle : U \text{ open in } I\}.$$

Then  $\Omega$  is the corresponding function  $p : \hat{I} \rightarrow I$ , where  $\hat{I}$  is the union of the stalks  $\Omega_i$ , and  $p$  gives output  $i$  for inputs from  $\Omega_i$ . The topology on  $\hat{I}$  has as base all sets of the form

$$[U, V] = \{\langle i, [U]_i \rangle : i \in V\}$$

where  $V$  is open and  $U \subseteq V$ . This makes  $p$  a local homeomorphism, and also makes each stalk a discrete space under the relative topology.

If we denote by  $\Theta_i$  the collection of open neighbourhoods of  $i$  then we have the following facts about germs of open sets:

(i)  $[U]_i = [I]_i$  iff  $i \in U$

(ii)  $[I]_i = \Theta_i$

(iii)  $[U]_i = [\emptyset]_i$  iff  $i$  is separated from  $U$  (i.e. there exists  $V \in \Theta_i$  such that  $U \cap V = \emptyset$ )

[The reader familiar with lattices may care to note that the open sets in  $I$  form a distributive lattice  $(\Theta, \cap, \cup)$  in which  $\Theta_i$  is a (prime) filter. The stalk  $\Omega_i$  is essentially the quotient lattice  $\Theta/\Theta_i$ , i.e.  $\sim_i$  is the standard definition of the lattice congruence determined by  $\Theta_i$ .]

Before examining  $\Omega$  as a subobject classifier we will look at truth-values  $s : 1 \rightarrow \Omega$ . Such an arrow is a *continuous* section of  $\Omega$ , generally called a *global* section of the sheaf. (We may also consider *local sections*  $s : U \rightarrow \hat{I}$  of  $\hat{I}$  defined on (open) subsets  $U$  of  $I$ ).

$$\begin{array}{ccc} I & \xrightarrow{s} & \hat{I} \\ \text{id}_I \searrow & & \swarrow p \\ & I & \end{array}$$

Now if  $U$  is open in  $I$ , define  $S_U : I \rightarrow \hat{I}$  by  $S_U(i) = \langle i, [U]_i \rangle$ . We then find  $S_U$  is a continuous global section, i.e.  $S_U : 1 \rightarrow \Omega$ . By (i) above we note that  $S_U(i) = \langle i, [I]_i \rangle$  iff  $i \in U$ . Then if  $s : 1 \rightarrow \Omega$  is any continuous section of  $\Omega$  and  $U = \{i : s(i) = \langle i, [I]_i \rangle\}$  we find that  $U$  is open ( $U = s^{-1}([I, I])$ ) and  $S_U = s$ .

We thus have that the truth values in **Top**( $I$ ) are “essentially” the *open* subsets of  $I$ , whereas in **Bn**( $I$ ) they were all the subsets of  $I$ . This will be a continuing theme. We shall later see other constructions that have a set-theoretic and a topological version, and find that the latter arise from the form by replacing “subset” by “open subset”.



The arrow  $\tau : 1 \rightarrow \Omega$  is the continuous section  $\tau : I \rightarrow \hat{I}$  that has  $\tau(i) = \langle i, [I]_i \rangle$ , all  $i \in I$ . Now if  $k$  is monic, where

$$\begin{array}{ccc}
 A & \xrightarrow{k} & B \\
 \searrow p & & \swarrow q \\
 & I &
 \end{array}$$

commutes, and  $A$  is an open subset of  $B$ , we obtain the character  $\chi_k : (B, q) \rightarrow \Omega$  as follows.

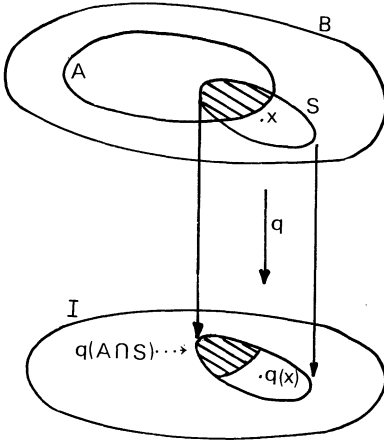


Fig. 4.9.

If  $x \in B$ , choose a neighbourhood  $S$  of  $x$  on which  $q$  is a local homeomorphism. Then  $\chi_k : B \rightarrow \hat{I}$  takes  $x$  to the germ of  $q(A \cap S)$  at  $q(x)$ , i.e.

$$\chi_k(x) = \langle q(x), [q(A \cap S)]_{q(x)} \rangle$$

Intuitively, the germ of  $q(A \cap S)$  at  $q(x)$  represents in  $I$ , under the local homeomorphism  $q$ , the set of points in  $A$  close to  $x$ . It provides a measure of the *extent* to which  $x$  is in  $A$ . Whereas in set theory classification admits of only two possibilities – either  $x \in A$  or  $x \notin A$  – in a topological context we may make more subtle distinctions by classifying according to how close  $x$  is to  $A$ . We use the germs at  $q(x)$  as a system of entities for measuring proximity of  $x$  to open subsets of  $B$ . A partial ordering on  $\Omega_{q(x)}$  is given by

$$[U]_{q(x)} \sqsubseteq [V]_{q(x)} \quad \text{iff there is some open set } W \text{ such that } q(x) \in W \text{ and } U \cap W \subseteq V \cap W,$$

i.e. iff the statement “ $U \subseteq V$ ” is locally true at  $q(x)$ .

Then the “larger” the germ of  $q(A \cap S)$  is in terms of this ordering, the closer will  $x$  be to  $A$ . If in fact  $x \in A$ , then  $q(x) \in q(A \cap S)$  and so by (i) above, the germ of  $q(A \cap S)$  is as large as it could be, i.e.  $[q(A \cap S)]_{q(x)} = [I]_{q(x)}$ . At the other extreme, if  $x$  is separated from  $A$ , then the germ of  $q(A \cap S)$  is as small as it could be, i.e.  $[q(A \cap S)] = [\emptyset]$ . Otherwise, when  $x$  is on the boundary of  $A$ ,  $[q(A \cap S)]$  is strictly between the germs of  $\emptyset$  and  $I$ ,  $[\emptyset] \subset [q(A \cap S)] \subset [I]$ .

EXERCISE 1. Verify that the definition of  $\chi_k(x)$  does not depend on the choice of neighbourhood  $S$  of  $x$  on which  $q$  is a local homeomorphism.

EXERCISE 2. (Alternative definition of  $\chi_k(x)$ ). Let

$$U_x = \{i \in I: \text{for some local section } s \text{ of } (B, q), s(i) \in A \text{ and } s(q(x)) = x\}$$

be the set of points in  $I$  that are carried into  $A$  by some local section of  $(B, q)$  that takes  $q(x)$  to  $x$ . Show that

$$[U_x]_{q(x)} = [q(A \cap S)]_{q(x)},$$

where  $S$  is as above. □

### 4.6. Monoid actions

Let  $\mathbf{M} = (M, *, e)$  be a monoid (cf. §2.5). Then any given  $m \in M$  determines a function  $\lambda_m : M \rightarrow M$ , called *left-multiplication by  $m$* , and defined by the rule  $\lambda_m(n) = m * n$ , for all  $n \in M$ . We thus obtain a family  $\{\lambda_m : m \in M\}$  of functions, indexed by  $M$ , which satisfies

- (i)  $\lambda_e = \text{id}_M$ , since  $\lambda_e(m) = e * m = m$ , and
- (ii)  $\lambda_m \circ \lambda_p = \lambda_{m * p}$ , since  $\lambda_m(\lambda_p(n)) = m * (p * n) = (m * p) * n$ .

Condition (ii) in fact says that the collection of  $\lambda_m$ 's is closed under functional composition. Indeed, it forms a monoid under this operation with identity  $\lambda_e$ .

The notion just described can be generalised. Suppose we have a set  $X$  and a collection  $\{\lambda_m : X \rightarrow X : m \in M\}$  of functions  $\lambda_m$  from  $X$  to  $X$ , the collection being indexed by the elements of our original monoid, and satisfying

$$\begin{aligned} \lambda_e &= \text{id}_X \\ \lambda_m \circ \lambda_p &= \lambda_{m * p}. \end{aligned}$$

The collection of  $\lambda_m$ 's is called an *action* of  $\mathbf{M}$  on the set  $X$ , and can be

replaced by a single function  $\lambda : M \times X \rightarrow X$ , defined by

$$\lambda(m, x) = \lambda_m(x), \quad \text{all } m \in M, \quad x \in X.$$

The above two conditions become

$$\lambda(e, x) = x$$

and

$$\lambda(m, \lambda(p, x)) = \lambda(m * p, x).$$

An **M**-set is defined to be a pair  $(X, \lambda)$ , where  $\lambda : M \times X \rightarrow X$  is such an action of **M** on  $X$ .

**EXAMPLE 1.** **M** is the monoid  $(\mathbb{N}, +, 0)$  of natural numbers under addition.  $X$  is the set of real numbers.  $\lambda$  is *addition*:—  $\lambda(m, r) = m + r$ .

**EXAMPLE 2.**  $X$  is the set of vectors of a vector space, **M** the multiplicative monoid of its scalars,  $\lambda$  is scalar multiplication of vectors.

**EXAMPLE 3.**  $X$  is the set of points in the Euclidean plane. **M** is the group of Euclidean transformations (rotations, reflections, translations) with  $*$  as function composition.  $\lambda(m, x)$  is  $m(x)$ , i.e. the result of applying transformation  $m$  to point  $x$ .

**EXAMPLE 4.**  $X$  is the set of states of a computing device. **M** is the set of input words (strings) with  $*$  the operation of concatenation or juxtaposition of strings.  $\lambda(m, x)$  is the state the machine goes into in response to being fed input  $m$  while in state  $x$ .  $\square$

For a given monoid **M**, the **M**-sets are the objects of a category **M-Set**, which is a topos. An arrow  $f : (X, \lambda) \rightarrow (Y, \mu)$  is an *equivariant*, or *action-preserving* function  $f : X \rightarrow Y$ , i.e. one such that

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ \lambda_m \downarrow & & \downarrow \mu_m \\ X & \xrightarrow{f} & Y \end{array}$$

commutes for each  $m \in M$ . In other words,  $f(\lambda(m, x)) = \mu(m, f(x))$ , all  $m$  and  $x$ . Composition of arrows is functional composition.

The terminal object is a singleton  $\mathbf{M}$ -set. We take  $1 = (\{0\}, \lambda_0)$  where  $\lambda_0(m, 0) = 0$ , all  $m$ .

The product of  $(X, \lambda)$  and  $(Y, \mu)$  is  $(X \times Y, \delta)$ , where  $\delta_m$  is  $\lambda_m \times \mu_m : X \times Y \rightarrow X \times Y$ . The pullback of

$$\begin{array}{ccc} & (Y, \mu) & \\ & \downarrow g & \\ (X, \lambda) & \xrightarrow{f} & (Z, \gamma) \end{array}$$

is  $(X \times_Z Y, \delta)$  with  $\delta$  as above.

Now a set  $B \subseteq \mathbf{M}$  is called a *left ideal* of  $\mathbf{M}$  if it is closed under left-multiplication, i.e. if  $m * b \in B$  whenever  $b \in B$  and  $m$  is any element of  $M$ . For example,  $M$  and  $\emptyset$  are left ideals of  $\mathbf{M}$ . We put  $\Omega = (L_M, \omega)$  where  $L_M$  is the set of left ideals in  $\mathbf{M}$ , and  $\omega : M \times L_M \rightarrow L_M$  has  $\omega(m, B) = \{n : n * m \in B\}$ .  $\top : 1 \rightarrow \Omega$  is the function  $\top : \{0\} \rightarrow L_M$  with  $\top(0) = M$ . Thus  $\top$  picks out the largest left-ideal  $M$  of  $\mathbf{M}$ .

To illustrate the workings of the subobject classifier, suppose  $k : (X, \lambda) \rightarrow (Y, \mu)$  is in fact the inclusion  $X \hookrightarrow Y$  (since  $k$  is *equivariant* this means  $\mu(m, x) = \lambda(m, x)$ , all  $x \in X$ ). The character  $\chi_k : (Y, \mu) \rightarrow \Omega$  of  $k$  is  $\chi_k : Y \rightarrow L_M$  defined by

$$\chi_k(y) = \{m : \mu(m, y) \in X\}, \quad \text{all } y \in Y.$$

**EXERCISE 1.** Check all the details – that  $\omega$  is an action of  $\mathbf{M}$  on  $L_M$ , that  $\chi_k(y)$  is a left-ideal, and that  $\chi_k$  satisfies the  $\Omega$ -axiom.  $\square$

## Exponentiation

Our initial motivation showed that  $* : M \times M \rightarrow M$  is itself an action of  $\mathbf{M}$  on  $M$ , i.e. that  $(M, *)$  is an  $\mathbf{M}$ -set. Given  $(X, \lambda)$  and  $(Y, \mu)$  we define the exponential

$$(Y, \mu)^{(X, \lambda)} = (E, \sigma)$$

where  $E$  is the set of equivariant maps  $f$  of the form  $f : (M, *) \times (X, \lambda) \rightarrow (Y, \mu)$  and  $\sigma_m : E \rightarrow E$  takes such an  $f$  to the function  $g = \sigma_m(f) : M \times X \rightarrow Y$  given by

$$g(n, x) = f(m * n, x)$$

The evaluation arrow

$$ev : (E, \sigma) \times (X, \lambda) \rightarrow (Y, \mu)$$

has

$$ev(f, x) = f(e, x).$$

Then given an arrow  $f : (X, \lambda) \times (Y, \mu) \rightarrow (Z, \nu)$ , the exponential adjoint  $\hat{f} : (X, \lambda) \rightarrow (Z, \nu)^{(Y, \mu)}$  takes  $x \in X$  to the equivariant map  $\hat{f}_x : M \times Y \rightarrow Z$  having

$$\hat{f}_x(m, y) = f(\lambda_m(x), y).$$

Categories of the form **M-Set** provide a rich source of examples, particularly of topoi that have “non-classical” properties. They will be “re-created” from a different perspective in Chapter 9.

EXERCISE 2. Describe all the left-ideals in  $(N, +, 0)$ .

EXERCISE 3. Show that  $M$  is a group iff  $M$  and  $\emptyset$  are the only left-ideals of **M**, i.e. iff  $L_M = \{M, \emptyset\}$ . □

### 4.7. Power objects

The exponential  $\Omega^a$  in a topos is the analogue of  $2^A$  in **Set**. Since  $2^A \cong \mathcal{P}(A)$  it is natural to wonder whether the object  $\Omega^a$  behaves like the “powerset” of the “set”  $a$ . In fact it does, as we shall see by first developing an independent categorial description of  $\mathcal{P}(A)$  in **Set**.

Now given sets  $A$  and  $B$  there is a bijective correspondence between the functions from  $B$  to  $\mathcal{P}(A)$  and the relations from  $B$  to  $A$ . Given function  $f : B \rightarrow \mathcal{P}(A)$  define relation  $R_f \subseteq B \times A$  by stipulating  $xR_f y$  iff  $y \in f(x)$ , for  $x \in B, y \in A$ . Conversely, given  $R \subseteq B \times A$ , define  $f_R : B \rightarrow \mathcal{P}(A)$  by  $f_R(x) = \{y : y \in A \text{ and } xRy\}$ .

It is not hard to see that the assignments of  $f_R$  to  $R$  and  $R_f$  to  $f$  are inverse to each other and establish the asserted isomorphism.

In order to capture this correspondence in terms of arrows we examine a special relation  $\in_A$  from  $\mathcal{P}(A)$  to  $A$ .  $\in_A$  is the membership relation and contains all the information about which subsets of  $A$  contain which elements of  $A$ . Precisely

$$\in_A = \{(U, x) : U \subseteq A, x \in A, \text{ and } x \in U\}.$$

Passing from  $\mathcal{P}(A)$  to  $2^A$ , the condition “ $x \in U$ ” becomes “ $\chi_U(x) = 1$ ”, and we see that  $\in_A$  is isomorphic to the set

$$\in'_A = \{(\chi_U, x) : U \subseteq A, x \in A, \text{ and } \chi_U(x) = 1\} \subseteq 2^A \times A$$

What is the characteristic function of  $\in'_A$  as a subset of  $2^A \times A$ ? Well it is none other than the evaluation arrow  $ev: 2^A \times A \rightarrow 2$ , since  $ev(\chi_U, x) = \chi_U(x)$ . Thus we are lead to a characterisation of  $\in'_A$  (and hence  $\in_A$  up to isomorphism) by the pullback square

$$\begin{array}{ccc} \in'_A & \hookrightarrow & 2^A \times A \\ \downarrow ! & & \downarrow ev \\ 1 & \xrightarrow{\text{true}} & 2 \end{array}$$

Now given a relation  $R \subseteq B \times A$ , we have  $\langle x, y \rangle \in R$  iff  $y \in f_R(x)$  iff  $\langle f_R(x), y \rangle \in \in_A$ , and so  $R$  is the inverse image of  $\in_A$  under the map  $f_R \times 1_A$ , that takes  $\langle x, y \rangle$  to  $\langle f_R(x), y \rangle$ .

So we see that (§3.13) the diagram

$$\begin{array}{ccc} R & \hookrightarrow & B \times A \\ g \downarrow & & \downarrow f_R \times \text{id}_A \\ \in_A & \hookrightarrow & \mathcal{P}(A) \times A \end{array}$$

is a pullback, where  $g$  is the restriction of  $f_R \times \text{id}_A$  to  $R$ . But something stronger than this can be said – given  $R$ , then without considering what  $g$  is,  $f_R$  is the *only* function  $B \rightarrow \mathcal{P}(A)$  that will give a pullback of the form of the diagram.

EXERCISE 1. Prove this last assertion. □

We are therefore lead to the following definition:

DEFINITION. A category  $\mathcal{C}$  with products is said to have *power objects* if to each  $\mathcal{C}$ -object  $a$  there are  $\mathcal{C}$ -objects  $\mathcal{P}(a)$  and  $\in_a$ , and a monic  $\in : \in_a \rightarrow \mathcal{P}(a) \times a$ , such that for any  $\mathcal{C}$ -object  $b$ , and “relation”,  $r: R \rightarrow b \times a$  there is exactly one  $\mathcal{C}$ -arrow  $f_r: b \rightarrow \mathcal{P}(a)$  for which there is a pullback in  $\mathcal{C}$  of the form

$$\begin{array}{ccc} R & \xrightarrow{r} & b \times a \\ \downarrow & & \downarrow f_r \times 1_a \\ \in_a & \xrightarrow{\in} & \mathcal{P}(a) \times a \end{array}$$

THEOREM 1. Any topos  $\mathcal{E}$  has power objects.

PROOF. For given  $\mathcal{E}$ -object  $a$ , let  $\mathcal{P}(a) = \Omega^a$  and let  $\epsilon : \epsilon_a \twoheadrightarrow \Omega^a \times a$  be the subobject of  $\Omega^a \times a$  whose character is  $ev_a : \Omega^a \times a \rightarrow \Omega$ , i.e.

$$\begin{array}{ccc} \epsilon_a & \xrightarrow{\epsilon} & \Omega^a \times a \\ \downarrow ! & & \downarrow ev_a \\ 1 & \xrightarrow{\top} & \Omega \end{array}$$

is a pullback, where  $ev_a$  is the evaluation arrow from  $\Omega^a \times a$  to  $\Omega$ . To show that this construction gives power objects take any monic  $r : R \hookrightarrow b \times a$  and let  $\chi_r : b \times a \rightarrow \Omega$  be its character. Then let  $f_r : b \rightarrow \Omega^a$  be the exponential adjoint to  $\chi_r$ , i.e. the unique arrow that makes

$$\begin{array}{ccc} \Omega^a \times a & & \\ f_r \times 1_a \uparrow & \searrow ev_a & \\ b \times a & \xrightarrow{\chi_r} & \Omega \end{array}$$

commute. Now consider the diagram

$$\begin{array}{ccc} R & \xrightarrow{r} & b \times a \\ \downarrow \text{!} & & \downarrow f_r \times 1_a \\ \epsilon_a & \xrightarrow{\epsilon} & \Omega^a \times a \\ \downarrow & & \downarrow ev_a \\ 1 & \xrightarrow{\top} & \Omega \end{array}$$

Since  $ev_a \circ (f_r \times 1_a) = \chi_r$ , the “perimeter” of this diagram is a pullback, by the  $\Omega$ -axiom. In particular it commutes, so as the bottom square is a pullback, the unique arrow  $R \twoheadrightarrow \epsilon_a$  does exist to make the whole diagram commute. But then by the PBL the top square is a pullback, as required by the definition of power objects. Moreover simply knowing that  $f_r$  is some arrow making the top square a pullback gives both squares as pullbacks and hence (PBL) the outer rectangle is a pullback. The  $\Omega$ -axiom then implies that  $ev_a \circ (f_r \times 1_a) = \chi_r$  and thus from the previous diagram  $f_r$  is uniquely determined as the exponential adjoint of  $\chi_r$ .  $\square$

Now given power objects we can recover  $\Omega$ , as  $\Omega \cong \Omega^1 = \mathcal{P}(1)$ . The monic  $\epsilon_1 \twoheadrightarrow \Omega^1 \times 1 \cong \Omega^1$  proves to be a subobject classifier. Anders Kock and C. Juul Mikkelsen have shown that power objects can also be used to

construct exponentials, and that

*a category  $\mathcal{C}$  is a topos iff  $\mathcal{C}$  is finitely complete and has power objects*

(for details consult Wraith [75]).

Currently this characterisation is being used as the definition of a topos, it being the best in terms of brevity. Paedogogically it is not however the best, for a number of reasons. Historically the idea of an elementary topos arose through examination of subobject classifiers, and this path provides the most suitable motivation. As will be evident it is the  $\Omega$ -axiom that is the key to the basic structure of a topos and it would have to be introduced anyway for the theory to get off the ground. Moreover each of the  $\Omega$ -axiom, and the notion of exponentiation, is conceptually simpler than the description of power objects.

There is another more remote matter, due to the recent development of weak set theories relating to recursion theory (admissible sets – cf. Barwise [75]). These theories produce categories of sets without general powerset formation. It therefore becomes of interest to study the ramifications of the  $\Omega$ -axiom without having to relate it to the notion of power-object.

EXERCISE 2. Examine the structure of power objects in the various topoi described in this chapter.

EXERCISE 3. Deduce from the discussion of this section, including the proof of the Theorem, that a category  $\mathcal{C}$  is a topos iff

- (i)  $\mathcal{C}$  has a terminal object and pullbacks of appropriate pairs of arrows,
- (ii)  $\mathcal{C}$  has a subobject classifier  $true : 1 \rightarrow \Omega$
- (iii) For each  $\mathcal{C}$ -object  $a$  there is a  $\mathcal{C}$ -object  $\Omega^a$  and an arrow  $ev_a : \Omega^a \times a \rightarrow \Omega$  such that for each  $\mathcal{C}$ -object  $b$  and “relation”  $r : R \rightarrow b \times a$  there is exactly one  $\mathcal{C}$ -arrow  $f_r : b \rightarrow \Omega^a$  making

$$\begin{array}{ccc}
 \Omega^a \times a & \xrightarrow{ev_a} & \Omega \\
 \uparrow f_r \times 1_a & \nearrow \chi_r & \\
 b \times a & & 
 \end{array}$$

commute.

EXERCISE 4. Show that the unique arrow  $\Omega^a \rightarrow \Omega^a$  corresponding to the relation  $\in_a \rightarrow \Omega^a \times a$  is  $1_{\Omega^a}$ . □



### 4.8. $\Omega$ and comprehension

In Lawvere [72] it is suggested that the  $\Omega$ -axiom is a form of the ZF Comprehension principle. To see this, suppose that  $B$  is a set and  $\varphi$  a property that applies to members of  $B$ . We represent  $\varphi$  in **Set** as a function  $\varphi : B \rightarrow 2$  given by

$$\varphi(x) = \begin{cases} 1 & \text{if } x \text{ has property } \varphi \\ 0 & \text{otherwise.} \end{cases}$$

Now the comprehension (separation) principle allows us to form the subset  $\{x : x \in B \text{ and } \varphi(x)\}$  of all elements of  $B$  satisfying  $\varphi$ . This set is determined by  $\varphi$  qua function as what we earlier called  $A_\varphi = \{x : \varphi(x) = 1\}$ . We have  $y \in \{x : \varphi(x)\}$  iff  $\varphi(y) = 1$ , and

$$\begin{array}{ccc} A_\varphi & \hookrightarrow & B \\ \downarrow ! & & \downarrow \varphi \\ 1 & \xrightarrow{\text{true}} & 2 \end{array}$$

is a pullback. By analogy, in a topos  $\mathcal{E}$ , if  $\varphi : b \rightarrow \Omega$  is an arrow with  $\text{cod} = \Omega$ , we let  $\{x : \varphi\} : a \rightarrow b$  be the subobject of  $b$  obtained by pulling *true* back along  $\varphi$ , as in

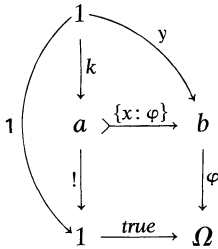
$$\begin{array}{ccc} a & \xrightarrow{\{x : \varphi\}} & b \\ \downarrow ! & & \downarrow \varphi \\ 1 & \xrightarrow{\text{true}} & \Omega \end{array}$$

Now in a general category, if  $x : 1 \rightarrow b$  is an element of object  $b$ , and  $f : a \rightarrow b$  a subobject, we define  $x$  to be a member of  $f$ ,  $x \in f$ , when  $x$  factors through  $f$ , i.e. there exists  $k : 1 \rightarrow a$  making

$$\begin{array}{ccc} & 1 & \\ k \swarrow & & \searrow x \\ a & \xrightarrow{f} & b \end{array}$$

commute. This naturally generalises the situation in **Set**.

Applying this notion of membership to the above pullback we see that if  $y : 1 \rightarrow b$  is a  $b$ -element then



$y \in \{x: \varphi\}$  iff the arrow  $k$  exists to make the whole diagram commute. But as the inner square is a pullback,  $k$  will exist (uniquely) iff the perimeter of the diagram commutes. Hence

$$y \in \{x: \varphi\} \text{ iff } \varphi \circ y = true,$$

giving us an analogue of the set-theoretic situation.

EXERCISE 1. Take  $f: a \twoheadrightarrow b$ ,  $g: c \twoheadrightarrow b$  with  $f \subseteq g$ . If  $x \in b$  (i.e.  $x: 1 \rightarrow b$ , or  $x \in 1_b$  as above) has  $x \in f$ , show  $x \in g$ .

EXERCISE 2. For any  $f: a \twoheadrightarrow d$  and  $x: 1 \rightarrow d$ ,  $x \in f$  iff  $\chi_f \circ x = true$ .