Category Theory

Notes by Peter Johnstone

Michaelmas 2020

4 Limits

In this chapter we study limits and colimits, which are constructions that we perform on diagrams in a category. We have already been drawing diagrams in categories in an informal way, but at this point we need to introduce a formal definition of diagram. Perhaps surprisingly, a diagram is just another name for a functor; but it is (usually) a functor of a rather special kind, which we think of in a special way, and we introduce some new notation and terminology for it.

Definition 4.1 Let J be a category (in practice J will almost always be small, and often finite). By a *diagram of shape* J in $\mathcal{C}$, we mean a functor $D\colon J \to \mathcal{C}$. We refer to the objects $D(j)$, $j \in \text{ob } J$, as *vertices* of the diagram, and the morphisms $D(\alpha)$, $\alpha \in \text{mor } J$, as its *edges*.

A usful example to keep in mind is that when J is the finite category with four objects and five non-identity morphisms represented by the picture

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

Then a diagram of shape J is just a commutative square of objects and morphisms of $\mathcal{C}$. Thus we think of J as a 'template' which indicates the shape of the diagram we are trying to draw, and the act of drawing a diagram as labelling the vertices and edges of the template with the names of objects and morphisms of $\mathcal{C}$, in a way which respects domains, codomains and composition. (Of course, when we draw a commutative square in an actual category, we usually don't draw in the diagonal arrow, because it is assumed that we know how to compose morphisms of $\mathcal{C}$.) Note also that we can think of not-necessarily-commutative squares in $\mathcal{C}$ as diagrams of shape J', where J' differs from J in having two diagonal edges (one being the composite of the top and right edges, and the other the composite of the left and bottom).

Definition 4.2 Let $D\colon J \to \mathcal{C}$ be a diagram. By a *cone over* D, we mean an object A of $\mathcal{C}$ (the *apex* of the cone) together with a family of morphisms $(\lambda_j : A \to D(j) \mid j \in \text{ob } J)$ (its *legs*), which are compatible with the edges of the diagram in the sense that $D(\alpha)\lambda_j = \lambda_{j'}$ for all $\alpha\colon j \to j'$ in J. A *morphism of cones* from $(A, (\lambda_j))$ to $(B, (\mu_j))$ is a morphism $f\colon A \to B$ satisfying $\mu_j f = \lambda_j$ for all j; thus we have a category $\text{Cone}(D)$ of cones over D. A *limit* for D is a terminal object of $\text{Cone}(D)$, if it exists. Dually we have the notion of a cone *under* D (sometimes called a *cocone*), and a *colimit* for D is an initial cone under D.

Clearly, limits are unique up to canonical isomorphism if they exist; so, as with adjoints, we often speak loosely of the (rather than a) limit of a particular diagram. Note that if $\Delta\colon \mathcal{C} \to [J, \mathcal{C}]$ denotes the functor which sends an object A to the *constant diagram* all of whose vertices are A and whose edges are 1_A, then a cone over D with apex A is the same thing as a natural transformation $\Delta A \to D$; and in fact $\mathrm{Cone}(D)$ is just another name for the arrow category $(\Delta \downarrow D)$, defined as in 3.3. Thus 3.3 tells us that assigning limits to all diagrams of shape J in $\mathcal{C}$ is equivalent to specifying a right adjoint for Δ; dually $\mathcal{C}$ has colimits for all diagrams of shape J iff Δ has a left adjoint.

To understand this concept, we need to see some examples.

Examples 4.3 (*a*) The simplest example of a shape category J is the empty category. Of course there is just one empty diagram in any category: it is very easy to draw (!), but not so easy to see since there's nothing there. But note that you *can* see a cone over the empty diagram, since it has an apex (although there are no legs); so if D is the empty diagram then $\mathrm{Cone}(D)$ is simply isomorphic to $\mathcal{C}$. Thus, although we defined a limit as a certain terminal object, we can also think of terminal objects as a special case of limits, namely limits for the empty diagram. (Combining this with 3.3 and the remark after 4.2, we see that each of the three concepts 'limit', 'right adjoint' and 'terminal object' is definable in terms of either of the others.)

(*b*) Let J be the discrete category with two objects. Then a diagram of shape J is just a pair of objects (A, B), and a cone over it is what is sometimes called a *span*, i.e. a pair of morphisms $(f : C \to A, g : C \to B)$ with common domain. It is easy to see that a limit for this diagram is exactly a categorical product $A \times B$, as we defined it in 2.6(*e*); and a colimit for it is a coproduct $A + B$.

(*c*) More generally, if J is a (small) discrete category, then a diagram of shape J is just a J-indexed family of objects $(A_j \mid j \in J)$; and a limit (respectively colimit) for this diagram is just a product $\prod_{j \in J} A_j$ (respectively a coproduct $\sum_{j \in J} A_j$), defined as the obvious generalization of the case when J has two objects.

(*d*) Let J be the category

$$\begin{tikzcd} \bullet \arrow[r, shift left] \arrow[r, shift right] & \bullet \end{tikzcd} .$$

Then a diagram of shape J is a parallel pair of morphisms $f, g\colon A \rightrightarrows B$; a cone over it consists of a span $(h\colon C \to A, k\colon C \to B)$ satisfying $fh = k = gh$. Equivalently, we can think of it as a single morphism h satisfying $fh = gh$, since k is uniquely determined by h; so a limit for the diagram is just an equalizer of (f, g) as we defined it in 2.6(*f*).

(*e*) Let J be the category

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

so that a diagram of shape J is a *cospan*

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

A cone over this diagram has three legs; but the middle one with codomain C is redundant, since it is determined by either of the others, so we may think of it as a span $(h: D \to A, k: D \to B)$ completing the diagram to a commutative square. A limit for this diagram is called a *pullback* of the pair (f, g); we think of k as the morphism obtained by 'pulling back' f along g, although this way of thinking disguises the symmetry of the notion betwen f and g. In any category with binary products and equalizers, we may construct pullbacks by first forming the product $A \times B$ and then taking the equalizer of $(f\pi_1, g\pi_2): A \times B \rightrightarrows C$; this is a particular case of something we shall prove in 4.4 below. Dually, colimits of shape J^{op} (that is, universal completions of spans to commutative squares) are known as *pushouts*; and they may be constructed from binary coproducts and coequalizers.

(*f*) Let M be the two-element monoid $\{1, e\}$ with $e^2 = e$. A diagram of shape M in $\mathcal{C}$ is just an object equipped with an idempotent endomorphism; and a limit (respectively, a colimit) for this diagram is just the monic (respectively epic) part of a splitting of the idempotent (cf. Exercise 2.13 and Example 3.2(*d*)). Thus $\mathcal{C}$ has limits of shape M iff it has colimits of shape M, iff all its idempotents split.

(*g*) The final example we shall mention will not play a very important role in this course, but it is included for historical reasons, because it helps to explain the origin of the name 'limit'. You may well have been wondering why category-theorists chose to call these things limits, since they don't look very like the limits with which you are familiar in analysis and topology; but here is one which does look a bit more like a topological limit. Let $J = \mathbb{N}$ be the ordered set of natural numbers, so that a diagram of shape $\mathbb{N}$ is what is sometimes called a *direct sequence* of objects and morphisms

$$A_0 \longrightarrow A_1 \longrightarrow A_2 \longrightarrow \cdots \quad ;$$

a colimit for this diagram consists of an object A_∞ equipped with morphisms $A_n \to A_\infty$ for all n, which are compatible with the transition maps $A_n \to A_{n+1}$, and which is 'best possible' among such. Such things were studied in algebra and topology, and given the name 'direct limits', before category theory existed; we may think of the A_n as successive 'approximations' to the object A_∞ which we are trying to build. (For example, in topology an infinite-dimensional CW-complex X is thought of as the direct limit of its *n-skeletons* X_n, i.e. the subspaces which are the union of all Euclidean cells of dimension $\leq n$ used in the construction of X.) Dually, we have the notion of *inverse sequence*, which is a diagram of shape $\mathbb{N}^{\mathrm{op}}$, and of *inverse limit* which is a limit for such a diagram; for example, the ring of p-adic integers is the limit of the inverse sequence in **Rng** whose nth term is $\mathbb{Z}/p^n\mathbb{Z}$. So when category-theorists became aware of the need to study this construction for more general diagrams, they simply borrowed the name 'limit' from these examples. (Actually, since direct limits look more like topological limits than inverse ones, it might have been more sensible to use the name 'limit' for values of the left adjoint to $\Delta : \mathcal{C} \to [J, \mathcal{C}]$, and 'colimit' for the right adjoint; but it's too late to change that now. Incidentally, there was a period in the 1960s when some people tried using the terms 'adjoint' and 'coadjoint' instead of 'left adjoint' and 'right adjoint'; but since different authors adopted different conventions about which was which, this terminology soon dropped out of use.)

We remarked earlier that pullbacks may be constructed from products and equalizers. This is a particular case of a general phenomenon:

Lemma 4.4 *Let $\mathcal{C}$ be a category.*

(i) *If $\mathcal{C}$ has equalizers and all small products, then $\mathcal{C}$ has all small limits (that is, limits for all diagrams whose shapes are small categories).*

(ii) *If $\mathcal{C}$ has equalizers and all finite products, then $\mathcal{C}$ has all finite limits.*

(iii) *If $\mathcal{C}$ has pullbacks and a terminal object, then $\mathcal{C}$ has all finite limits.*

Proof The proofs of (i) and (ii) are identical, so we give them together. Given $D: J \to \mathcal{C}$ where J is small or finite as appropriate, form the products $P = \prod_{j \in \text{ob } J} D(j)$ and $Q = \prod_{\alpha \in \text{mor } J} D(\text{cod } \alpha)$. We have two morphisms $f, g \colon P \rightrightarrows Q$ defined by $\pi_\alpha f = \pi_{\text{cod } \alpha}$ and $\pi_\alpha g = (D\alpha)\pi_{\text{dom } \alpha}$ respectively; let $e \colon L \to P$ be their equalizer, and let $\lambda_j = \pi_j e \colon L \to D(j)$. We claim that $(L, (\lambda_j \mid j \in \text{ob } J))$ is a limit for D. Clearly, we have $(D\alpha)\lambda_j = \pi_\alpha g e = \pi_\alpha f e = \lambda_{j'}$ for any $\alpha \colon j \to j'$ in J, so the λ_j do indeed form a cone over D. But any cone over D is in particular a cone over the discrete diagram with vertices $D(j)$, $j \in \text{ob } J$, so it factors uniquely through the product P. And this factorization has equal composites with f and g, so it in turn factors uniquely through e.

(iii) We shall show that the hypotheses of (iii) imply those of (ii). Suppose 1 is a terminal object of $\mathcal{C}$. Then, for any two objects A and B, the pullback of

$$\begin{tikzcd} & A \arrow[d] \\ B \arrow[r] & 1 \end{tikzcd}$$

has the universal property of a product $A \times B$; and we can then construct the product of any finite family $(A_1, A_2, \ldots, A_n)$ as $(\cdots((A_1 \times A_2) \times A_3) \cdots) \times A_n$. So $\mathcal{C}$ has all finite products; to obtain an equalizer for $f, g \colon A \rightrightarrows B$, consider the pullback of

$$\begin{tikzcd} & A \arrow[d, "{(1_A, f)}"] \\ A \arrow[r, "{(1_A, g)}"] & A \times B \end{tikzcd}$$

where we have adopted the convention of denoting morphisms into a product by vectors of their composites with the projections. It is clear that any cone over this diagram with apex C has its two legs $C \rightrightarrows A$ equal; so a pullback for it has the same universal property as an equalizer of (f, g). $\square$

We say a category is *complete* (respectively *cocomplete*) if it has all small limits (respectively colimits). Using 4.4(i), it is easy to see that categories like **Set**, **Gp** and **Top** are all complete and cocomplete. Next, we consider the ways in which functors may interact with limits.

Definition 4.5 Let $G \colon \mathcal{D} \to \mathcal{C}$ be a functor.

(a) We say G *preserves* limits of shape J if, whenever $D \colon J \to \mathcal{D}$ and $(L, (\lambda_j))$ is a limit for D, then $(GL, (G\lambda_j))$ is a limit for GD.

(b) We say G *reflects* limits of shape J if, whenever $D \colon J \to \mathcal{D}$, any cone over D which is mapped by G to a limit for GD is itself a limit for D.

(c) We say G *creates* limits of shape J if, given $D \colon J \to \mathcal{D}$ and a limit $(L, (\lambda_j))$ for GD, there exists a cone over D whose image under G is isomorphic to $(L, (\lambda_j))$, and any such cone is a limit for D.

Part (c) of this definition is slightly nonstandard; most textbooks give a stricter definition of creation, in which each limit cone over GD is required to lift uniquely 'on the nose' to a cone over D, and that lifting is a limit. We have preferred to avoid this because it goes against the general principle that categorical properties should be invariant under equivalence; in particular, an equivalence functor (that is, one satisfying the hypotheses of 1.12) need not create limits in this strict sense, simply because the 'on the nose' liftings of limit cones may not exist (or may not be unique). Our weaker definition avoids this problem. However, it has to be admitted that most of the instances of limit-creation we shall encounter will satisfy the strict condition.

The notions of 4.5 may behave rather oddly if we do not assume that all limits of the given shape exist; so normally in (a) we assume that $\mathcal{D}$ has limits of shape J, and in (b) and (c) we assume that $\mathcal{C}$ does. With this hypothesis, creation is equivalent to the conjunction of preservation and reflection.

Corollary 4.6 *In any of the three statements of* 4.4*, we may replace the two occurrences of '$\mathcal{C}$ has' by either '$\mathcal{D}$ has and $G\colon \mathcal{D} \to \mathcal{C}$ preserves' or '$\mathcal{C}$ has and $G\colon \mathcal{D} \to \mathcal{C}$ creates'.*

Proof All the constructions in the proof of 4.4 are preserved (respectively created) by G. $\square$

Examples 4.7 (a) The forgetful functor $\mathbf{Gp} \to \mathbf{Set}$ creates all small limits (in the strict sense): given a family of groups $(G_j \mid G \in J)$ there is a unique group structure on $\prod_{j\in J} G_j$ which makes the projections into homomorphisms, and it makes $\prod_{j\in J} G_j$ into a product in $\mathbf{Gp}$. (Similarly, equalizers in $\mathbf{Set}$ of parallel pairs of group homomorphisms are subgroups.) It does not preserve or reflect colimits in general: for example, the coproduct in $\mathbf{Gp}$ of G and H is their free product $G * H$, which is not the same as their disjoint union. (But see 5.13(a) below for a particular class of colimits which it does create.)

(b) The forgetful functor $U\colon \mathbf{Top} \to \mathbf{Set}$ preserves all small limits and colimits, but does not reflect them: given a diagram $D : J \to \mathbf{Top}$, it is generally possible to impose a topology on the limit of UD which is strictly finer than that which makes it into a limit in $\mathbf{Top}$, and so we obtain a non-limit cone in $\mathbf{Top}$ which is mapped by U to a limit cone.

(c) Examples of functors which reflect (co)limits without preserving them are rather hard to come by, but here is one: the inclusion $\mathbf{AbGp} \to \mathbf{Gp}$ reflects binary coproducts but does not preserve them. The reason is that a free product of two nontrivial groups is always nonabelian; so the only cones in $\mathbf{AbGp}$ which can map to coproduct cones in $\mathbf{Gp}$ are those where at least one of the two summands is trivial. But if A is the trivial group, then $A + B \cong B$ in either $\mathbf{Gp}$ or $\mathbf{AbGp}$.

(d) If $\mathcal{D}$ has limits of shape J, then so does any functor category $[\mathcal{C}, \mathcal{D}]$, and the forgetful functor $[\mathcal{C}, \mathcal{D}] \to \mathcal{D}^{\mathrm{ob}\, \mathcal{C}}$ creates them (strictly). We may regard a diagram of shape J in $[\mathcal{C}, \mathcal{D}]$ as a functor $D\colon J \times \mathcal{C} \to \mathcal{D}$; if A is any object of $\mathcal{C}$, then $D(-, A)$ is a diagram of shape J in $\mathcal{D}$, and so has a limit $(LA, (\lambda_{j,A} \mid j \in \mathrm{ob}\, J))$. There is a unique way of making $A \mapsto LA$ into a functor so that the $\lambda_{j,-}$ become natural transformations: namely, given $f\colon A \to B$ in $\mathcal{C}$ we define Lf to be the unique factorization of the cone whose legs are the composites $D(j, f)\lambda_{j,A} : LA \to D(j, B)$ through the limit of $D(-, B)$. As usual, uniqueness implies functoriality of $f \mapsto Lf$; and it also implies that $(L, (\lambda_{j,-}))$ is a limit cone in $[\mathcal{C}, \mathcal{D}]$.

Remark 4.8 It is easy to see that a morphism $f\colon A \to B$ is monic iff the square

$$\begin{tikzcd} A \arrow[r, "1_A"] \arrow[d, "1_A"'] & A \arrow[d, "f"] \\ A \arrow[r, "f"] & B \end{tikzcd}$$

is a pullback. Thus we may conclude from 4.7(d) that *if $\mathcal{D}$ has pullbacks, then any monomorphism in a functor category $[\mathcal{C},\mathcal{D}]$ is a pointwise monomorphism*, since this pullback must be constructed pointwise. (The converse is trivially true.) This (or rather its dual) fulfils a promise we made before 2.11, since **Set** has pushouts.

Lemma 4.9 *If $G\colon \mathcal{D} \to \mathcal{C}$ has a left adjoint, then G preserves all limits which exist in $\mathcal{D}$.*

Proof We shall give two proofs: the first is a more conceptual proof which contains the real reason why the result holds, but which requires the assumption that both $\mathcal{C}$ and $\mathcal{D}$ have all limits of some particular shape, whereas the second, more pedestrian, proof requires only the existence of a limit for a particular diagram in $\mathcal{D}$. For the first proof, suppose $(F \dashv G)$; then it is easy to see that the square

$$\begin{tikzcd} \mathcal{C} \arrow[r, "F"] \arrow[d, "\Delta"] & \mathcal{D} \arrow[d, "\Delta"] \\ {[J,\mathcal{C}]} \arrow[r, "{[J,F]}"] & {[J,\mathcal{D}]} \end{tikzcd}$$

commutes, where $[J,F]$ is the functor which applies F to diagrams of shape J. But all the functors in this diagram have right adjoints (it is easy to see that we have $([J,F] \dashv [J,G])$), so by 3.6 the diagram

$$\begin{tikzcd} {[J,\mathcal{D}]} \arrow[r, "{[J,G]}"] \arrow[d, "\lim_J"] & {[J,\mathcal{C}]} \arrow[d, "\lim_J"] \\ \mathcal{D} \arrow[r, "G"] & \mathcal{C} \end{tikzcd}$$

commutes up to isomorphism. But this is exactly the assertion that G preserves limits of shape J.

For the second proof, suppose given $D\colon J \to \mathcal{D}$ with limit $(L,(\lambda_j))$. Given a cone $(A,(\alpha_j))$ over GD with apex A, the naturality of the adjunction ensures that the transposes $\overline{\alpha_j} : FA \to D(j)$ of the α_j form a cone over D. So we have a unique morphism of cones $\overline{f}\colon FA \to L$; and its transpose $f\colon A \to GL$ yields the unique factorization of $(A,(\alpha_j))$ through $(GL,(G\lambda_j))$. □

The Adjoint Functor Theorem asserts that, *morally*, the converse of 4.9 is true: if $\mathcal{D}$ has and $G : \mathcal{D} \to \mathcal{C}$ preserves all limits, then G has a left adjoint. This is a true theorem (and we shall prove it shortly); but in this 'primitive' form it applies only to functors between preorders, since it requires limits which are 'too big' to exist in more general categories (see Exercise 4.20 below). However, there are versions of the theorem which, by imposing suitable set-theoretic restrictions on $\mathcal{C}$, $\mathcal{D}$ and/or G, allow us to get away with using only small limits; and we shall also prove two of these. We begin with a lemma which, along with Theorem 3.3, is central to all versions of the theorem.

Lemma 4.10 *Suppose $\mathcal{D}$ has and $G\colon \mathcal{D} \to \mathcal{C}$ preserves limits of shape J. Then, for each $A \in \mathrm{ob}\,\mathcal{C}$, the category $(A \downarrow G)$ has limits of shape J and the forgetful functor $(A \downarrow G) \to \mathcal{D}$ creates them.*

Proof The second statement tells you how to prove the first. Suppose given $D\colon J \to (A \downarrow G)$; write $D(j) = (UD(j), f_j)$ where U is the forgetful functor, and let $(L,(\lambda_j \mid j \in \mathrm{ob}\,J))$ be a limit for UD. The f_j form a cone over GUD, since the edges of UD are morphisms in $(A \downarrow G)$, and so induce a unique $h\colon A \to GL$ such that $(G\lambda_j)h = f_j$ for all j. But then (L,h) is an object of $(A \downarrow G)$, and it is the unique lifting of L to an object of this category which makes the λ_j into morphisms of $(A \downarrow G)$. The fact that they form a limit cone in $(A \downarrow G)$ is straightforward. □

By 3.3, we know that we need to find an initial object of $(A \downarrow G)$ for each A. We know that an initial object can be represented as a colimit (of the empty diagram); can we also represent it as a limit? It's helpful to think about posets at this point: the least element of a poset, if it has one, is the join of the empty family, but it is also the meet of the family of all elements of the poset. This suggests that we should look for the limit of the 'full diagram' containing all the objects and morphisms of our category. And indeed this works:

Lemma 4.11 *Let $\mathcal{C}$ be a category. Specifying an initial object of $\mathcal{C}$ is equivalent to specifying a limit for the identity functor $1_{\mathcal{C}}$, considered as a diagram of shape $\mathcal{C}$ in $\mathcal{C}$.*

Proof First suppose given an initial object I. The unique morphisms $I \to A$, $A \in \mathrm{ob}\,\mathcal{C}$, clearly form a cone over the identity diagram, and it is a limit cone, since if $(B, (f_A \mid A \in \mathrm{ob}\,\mathcal{C}))$ is any cone over the identity functor then f_I is its unique factorization through the one with apex I.

Conversely, suppose given a limit cone $(I, (f_A \mid A \in \mathrm{ob}\,\mathcal{C}))$ for the identity functor. Then, for any $g\colon I \to A$, we have $gf_I = f_A$, so we need to prove that $f_I = 1_I$. But specializing to the case when $g = f_A$, we see that f_I is a factorization of the limit cone through itself, so it must be the identity. Hence f_A is the unique morphism $I \to A$, for each $A \in \mathrm{ob}\,\mathcal{C}$. □

The 'primitive' form of the Adjoint Functor Theorem follows immediately from 4.10 and 4.11. But it requires $\mathcal{D}$ to have limits of shape $(A \downarrow G)$ for all $A \in \mathrm{ob}\,\mathcal{C}$, and in general categories which are not preorders do not have limits for diagrams which are 'as big as themselves'. The most we can hope for is that $\mathcal{D}$ is complete; so we need to impose set-theoretic conditions which ensure that we can represent the initial objects of the $(A \downarrow G)$ as small limits. We give two versions of the theorem: the first, known as the General Adjoint Functor Theorem, imposes minimal restrictions on $\mathcal{C}$ and $\mathcal{D}$, but requires a set-theoretic condition on the functor G. The second (Special) theorem imposes stronger conditions on $\mathcal{C}$ and $\mathcal{D}$, which ensure that this set-theoretic condition automatically holds for any functor preserving small limits, so we do not have to verify it. Both versions are due to Peter Freyd, and first appeared as exercises for the reader (!) in his book *Abelian Categories*.

Theorem 4.12 *(General Adjoint Functor Theorem) Let $G\colon \mathcal{D} \to \mathcal{C}$ be a functor where $\mathcal{D}$ is locally small and complete. Then G has a left adjoint iff it preserves all small limits and satisfies the 'solution-set condition' that, for any $A \in \mathrm{ob}\,\mathcal{C}$, there is a set of objects of $(A \downarrow G)$ which is collectively weakly initial, i.e. every object admits a morphism from at least one member of this set.*

Proof The left-to right implication is immediate from 4.9 and 3.3, since the initial object (FA, η_A) of $(A \downarrow G)$ forms a singleton solution-set.

For the converse, by 4.10, $(A \downarrow G)$ inherits completeness from $\mathcal{D}$; and it also inherits local smallness since the morphisms $(B, f) \to (B', f')$ in $(A \downarrow G)$ are a subset of the morphisms $B \to B'$ in $\mathcal{D}$. So we are reduced to proving that if $\mathcal{A} = (A \downarrow G)$ is complete and locally small, and has a weakly initial set of objects $\{S_i \mid i \in I\}$, then it has an initial object.

First we form the product $P = \prod_{i \in I} S_i$. Clearly P is a single weakly initial object, since it admits morphisms to every S_i; but in general it will not be initial, so we need to 'cut it down'. We do this using what might be called an 'industrial-strength equalizer': that is, we form the limit of the diagram whose vertices are two copies of P and whose edges, all going from the first vertex to the second, are all the endomorphisms of P (this is the point at which we use local smallness). Let us denote this by $i\colon I \to P$; then I is again clearly weakly initial. But it is actually initial, since if we are given $f, g\colon I \rightrightarrows T$, we may form their (ordinary) equalizer $e\colon E \to I$, and since E is an object of $\mathcal{A}$ there exists $h\colon P \to E$. Now ieh is an endomorphism of P, but so is 1_P, and hence we have $iehi = 1_P i = i$. But i is monic (for the same reason as an ordinary equalizer), so we may cancel it on the left to obtain $ehi = 1_I$. So e is split epic, and hence $f = g$. □

Examples 4.13 (a) Imagine, if you will, that you have never seen the construction of free groups, and you want to prove that the forgetful functor $U\colon \mathbf{Gp} \to \mathbf{Set}$ has a left adjoint. Clearly, $\mathbf{Gp}$ is locally small and complete, and U preserves small limits (since, as we observed in 4.7(a), it creates them); so we need to verify the solution-set condition at an arbitrary set A. But any function $f : A \to UG$ factors through UG', where G' is the subgroup of G generated by the image of f; and this subgroup has cardinality bounded by $\max\{\aleph_0, \operatorname{card} A\}$. So by taking a fixed set B of this cardinality, imposing all possible group structures on all subsets of B, and considering all mappings from A into the underlying sets of these groups, we obtain a solution-set at A.

Saunders Mac Lane, who gives this example in his book *Categories for the Working Mathematician*, goes on to say that it has the advantage of avoiding the tedious construction of free groups using equivalence classes of words in the generators and their inverses. But that is slightly disingenuous, since the way we obtained the cardinality bound on G' was precisely because we knew its elements could all be represented by words; and in general it seems to be the case that, if you know enough about a functor to verify the solution-set condition for it, you know enough to give an explicit construction of its left adjoint. And the latter is more likely to be useful in practice; if you want to prove anything nontrivial about free groups (for example, the Nielsen–Schreier Theorem that every subgroup of a free group is free), you won't get very far by using the Adjoint Functor Theorem proof that they exist.

(b) We should give an example to show that the solution-set condition can fail. Consider the forgetful functor $U : \mathbf{CLat} \to \mathbf{Set}$, where $\mathbf{CLat}$ is the category of complete lattices (i.e., posets with arbitrary meets and joins, and maps preserving both meets and joins). This satisfies the other hypotheses of 4.12 for the same reasons as the forgetful functor from $\mathbf{Gp}$; but in 1965 Andrew Hales proved that, for any cardinal κ, there is a complete lattice L_κ of cardinality at least κ which is generated by a three-element subset (i.e. no proper sub-complete-lattice of L_κ contains those three elements). So the solution-set condition fails for the set $A = \{a, b, c\}$, and U has no left adjoint.

Before proving the Special Adjoint Functor Theorem, we need to introduce a new definition, and also to prove a lemma about pullbacks of monomorphisms. First the definition:

Definition 4.14 By a *subobject* of an object A in a category $\mathcal{C}$, we mean simply a monomorphism $m\colon A' \rightarrowtail A$ with codomain A. The subobjects of A form a preorder $\mathrm{Sub}(A)$ under the relation $m \leq m'$ iff m factors through m'; we say $\mathcal{C}$ is *well-powered* if $\mathrm{Sub}(A)$ is equivalent to a (small) poset for every A — equivalently, there exists a set of subobjects of A which meets every isomorphism class in $\mathrm{Sub}(A)$.

For example, $\mathbf{Set}$ is well-powered because every subobject of A is isomorphic to the inclusion of a subset (and it is the power-set axiom in set theory which says that these form a set — hence the name 'well-powered'). Incidentally, the dual notion (that is, well-poweredness of $\mathcal{C}^{\mathrm{op}}$) should properly be known as 'well-copoweredness'; but many authors wrongly write 'cowell-powered', which ought if anything to mean 'badly powered'. There is a fairly close relationship between well-poweredness and local smallness; see Exercise 4.22 below.

Lemma 4.15 *(Monomorphisms are stable under pullback) Let*

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

be a pullback square in a category where f is monic. Then k is also monic.

Proof Suppose given $l, m\colon D \rightrightarrows P$ with $kl = km$. Then $fhl = gkl = gkm = fhm$, and since f is monic we have $hl = hm$. So l and m are factorizations of the same cone through the pullback, and are therefore equal. □

Theorem 4.16 *(Special Adjoint Functor Theorem) Suppose both $\mathcal{C}$ and $\mathcal{D}$ are locally small, and additionally that $\mathcal{D}$ is complete and well-powered and has a coseparating set of objects (i.e. one satisfying the dual of 2.7(a)). Then a functor $G\colon \mathcal{D} \to \mathcal{C}$ has a left adjoint iff it preserves all small limits.*

Proof Once again, the left-to-right implication is just 4.9. Conversely, suppose $\mathcal{C}$, $\mathcal{D}$ and G satisfy the conditions, and let A be an object of $\mathcal{C}$. As before, $(A \downarrow G)$ inherits completeness and local smallness from $\mathcal{D}$; but it also inherits well-poweredness, since the subobjects of (B, f) in $(A \downarrow G)$ are just those $m : B' \rightarrowtail B$ for which f factors through Gm (note that since the forgetful functor $(A \downarrow G) \to \mathcal{D}$ preserves pullbacks, it preserves monomorphisms by 4.8). And if $\{S_i \mid i \in I\}$ is a coseparating set for $\mathcal{D}$, it is easy to see that $(A \downarrow G)$ has a coseparating set consisting of all pairs (S_i, f) with $i \in I$ and $f\colon A \to GS_i$ (this is the point where we use local smallness of $\mathcal{C}$).

Thus we are reduced to proving that if a category $\mathcal{A}$ is complete, locally small and well-powered and has a coseparating set $\mathcal{S}$, then it has an initial object. As before, we begin by constructing a product P, namely the product of all members of $\mathcal{S}$. Then, instead of an industrial-strength equalizer, we use an 'industrial-strength pullback' (also sometimes called a wide pullback); that is, a limit for the diagram whose edges are all the members of a representative set of subobjects of P. We denote the apex of the limit cone by I; an easy generalization of 4.15 shows that all the legs $I \to P'$ of the limit cone are monic, so in particular $I \to P$ is monic, and it is a least subobject of P. It is then clear that I satisfies the uniqueness property of an initial object: given $f, g\colon I \rightrightarrows A$, their equalizer would yield a subobject of I and hence a subobject of P smaller than I, so it would have to be an isomorphism. But we need to do some work to show existence.

Given an object A of $\mathcal{A}$, form the product $Q = \prod_{(S,f)} S$ indexed by all pairs (S, f) with $S \in \mathcal{S}$ and $f\colon A \to S$. We have a canonical morphism $g\colon A \to Q$ defined by $\pi_{(S,f)} g = f$; the assertion that $\mathcal{S}$ is a coseparating set says precisely that g is monic. And we have $h\colon P \to Q$ defined by $\pi_{(S,f)} h = \pi_S$, so if we form the pullback

$$\begin{tikzcd} B \arrow[r] \arrow[d, tail, "k"] & A \arrow[d, tail, "g"] \\ P \arrow[r, "h"] & Q \end{tikzcd}$$

we see that k is monic by 4.15; so it is a subobject of P, and hence $I \rightarrowtail P$ factors through it. So we have a morphism $I \to B$, and hence by composition a morphism $I \to A$. □

Example 4.17 As promised, we use 4.16 to prove the existence of the Stone–Čech compactification β, i.e. the left adjoint for the inclusion $I\colon \mathbf{KHaus} \to \mathbf{Top}$. Tychonoff's Theorem says that **KHaus** is closed under products in **Top**; so it has them, and I preserves them. It is also closed under equalizers, since the equalizer of two continuous maps into a Hausdorff space is a closed subspace of their domain, and closed subspaces of compact spaces are compact. Thus **KHaus** is complete, and I preserves all small limits. **KHaus** is also clearly locally small; it is well-powered since every subobject of a compact Hausdorff space is isomorphic to the inclusion of a closed subspace; and it has a single coseparator, namely the closed unit interval $[0, 1]$, by Uryson's Lemma (which says that if X is a normal space — in particular a compact Hausdorff space — and x and y are distinct points of X, then there is a continuous $f\colon X \to [0, 1]$ with $f(x) = 0$ and $f(y) = 1$). So all the hypotheses of 4.16 are satisfied.

Remarks 4.18 (a) As we mentioned earlier, Marshall Stone and Eduard Čech gave substantially different constructions of βX when they introduced it; but Čech's construction was strikingly similar to that produced by the Special Adjoint Functor Theorem, and although Peter Freyd has never (as far as I know) publicly admitted this, it seems very likely that his formulation of SAFT was inspired directly by Čech's construction. What Čech did, given an arbitrary space X, was to form the product P of copies of $[0,1]$ indexed by all continuous maps $f : X \to [0,1]$, and the map $g : X \to P$ defined by $\pi_f g = f$; that is, he formed the product (P, g) of the members of a coseparating set for $(X \downarrow I)$. He then defined βX to be the closure of the image of g; but that is of course the smallest subobject of (P, g) in $(X \downarrow I)$. Thus the hypotheses of 4.16 are exactly the categorical translations of the topological facts that Čech used to construct β.

(b) We should mention that we could also have deduced the existence of β from 4.12; to obtain a solution set for I at an object X, we use a cardinality bound much as in 4.13(a). Any continuous map $f : X \to IY$ factors through IY' where Y' is the closure of the image of f; and one can show that if a Hausdorff space Y' contains a dense subspace X' of cardinality at most κ, then card $Y' \leq 2^{2^\kappa}$, since the mapping sending y to $\{U \cap X' \mid y \in U, U \text{ open in } Y'\}$ is an injection from Y' to the second power-set PPX'.

Exercises on chapter 4

Exercise 4.19 (i) A category J is said to be *connected* if it has just one connected component, i.e. (it is nonempty and) any two objects of J may be linked by a 'zig-zag' of morphisms. Show that a category has all finite connected limits (that is, limits of diagrams whose shapes are finite connected categories) iff it has pullbacks and equalizers. [*Hint*: given $D : J \to \mathcal{C}$ where J is finite and connected, first use pullbacks to enlarge D to a diagram D' of shape J' where J' has a weakly initial object, in such a way that every cone over D extends uniquely to a cone over D'. Then use equalizers to get a limit for D'.]

(ii) Give examples of a category which has pullbacks but not equalizers, and of one which has equalizers but not pullbacks.

(iii) Let $\mathcal{C}$ be a category, and A an object of $\mathcal{C}$. Show that the forgetful functor $\mathcal{C}/A \to \mathcal{C}$ creates connected limits. [Here $\mathcal{C}/A$ denotes the arrow category $(1_{\mathcal{C}} \downarrow A)$, i.e. the category whose objects are the morphisms of $\mathcal{C}$ with codomain A, and whose morphisms $f \to g$ are the morphisms $h : \operatorname{dom} f \to \operatorname{dom} g$ in $\mathcal{C}$ satisfying $gh = f$.]

(iv) Let $\mathcal{C}$ be a category with all finite limits, and $F : \mathcal{C} \to \mathcal{D}$ a functor. By considering the factorization of F as

$$\mathcal{C} \xrightarrow{\widehat{F}} \mathcal{D}/F1 \xrightarrow{U} \mathcal{D}\,,$$

where 1 is the terminal object of $\mathcal{C}$, $\widehat{F}(A) = F(A \to 1)$ and U is the forgetful functor, show that F preserves all finite connected limits iff it preserves pullbacks.

Exercise 4.20 Let κ be a cardinal; let $\mathcal{C}$ and $\mathcal{D}$ be categories such that $\mathcal{C}$ has products of families of objects of cardinality κ, and each $\mathcal{D}(A, B)$ has at most κ elements. Show that any functor $F : \mathcal{C} \to \mathcal{D}$ factors through a preorder (i.e. F is constant on each $\mathcal{C}(A, B)$). Deduce that a complete small category must be a preorder.

Exercise 4.21 A *complete semilattice* is a partially ordered set A in which every subset has a least upper bound (i.e. A is cocomplete when regarded as a category); a complete semilattice homomorphism is a mapping preserving (order and) arbitrary least upper bounds. Use the Adjoint Functor Theorem to show

(i) that a poset A is a complete semilattice iff A^{op} is; and

(ii) that the category **CSLat** of complete semilattices and their homomorphisms is isomorphic to its opposite.

Exercise 4.22 (i) Let $\mathcal{C}$ be a well-powered category with finite products. Show that $\mathcal{C}$ is locally small. [*Hint*: show that two morphisms $f, g\colon A \rightrightarrows B$ are equal iff their graphs $(1_A, f)$ and $(1_A, g)$ are isomorphic as subobjects of $A \times B$.]

(ii) Let $\mathcal{C}$ be a locally small category with pullbacks and a detecting set of objects. Show that $\mathcal{C}$ is well-powered. [Show that a subobject of A is determined up to isomorphism by the set of morphisms from members of the detecting set to A which factor through it.]

(iii) Find examples of a well-powered category which is not locally small, and a locally small category which is not well-powered. [Think about 'large' groups for the first example, and 'large' preorders for the second.]

Exercise 4.23 Let $F : \mathcal{C} \to \mathcal{D}$ be a functor between small categories, and write F^* for the functor $[\mathcal{D}, \mathbf{Set}] \to [\mathcal{C}, \mathbf{Set}]$ induced by composition with F.

(i) Use the General Adjoint Functor Theorem to show that F^* has a left adjoint. [To obtain a solution set at a particular functor $G : \mathcal{C} \to \mathbf{Set}$, let κ be the cardinality of $\coprod_{A \in \mathrm{ob}\, \mathcal{C}} GA$, and consider functors $\mathcal{D} \to \mathbf{Set}$ which are epimorphic images of coproducts of at most κ representables.]

(ii) Use the Special Adjoint Functor Theorem to show that F^* has a right adjoint.

[The adjoints of F^* are called the *left* and *right Kan extension functors* along F. They can be constructed explicitly by forming colimits and limits, respectively, of diagrams whose shapes are the arrow categories $(F \downarrow B)$ and $(B \downarrow F)$, $B \in \mathrm{ob}\, \mathcal{D}$. It was the desire to study (particular cases of) these adjoints that led Daniel Kan to formulate the notion of adjoint functor.]

Exercise 4.24 (i) A continuous map $f\colon X \to Y$ of topological spaces is said to be a *local homeomorphism* if, for every $x \in X$, there exist open neighbourhoods U, V of x and $f(x)$ respectively such that f restricts to a homeomorphism $U \to V$. Show that a composite of two local homeomorphisms is a local homeomorphism.

(ii) We write **LH** for the category of topological spaces and local homeomorphisms between them. Show that **LH** has pullbacks, and that the inclusion functor $\mathbf{LH} \to \mathbf{Top}$ preserves them. [*Hint*: first show that a pullback in **Top** of a local homeomorphism is a local homeomorphism. But note that this alone is not enough, since the inclusion is not full.]

(iii) Show that **LH** has equalizers, but that the inclusion $\mathbf{LH} \to \mathbf{Top}$ does not preserve them. [*Hint*: the image of any local homeomorphism is an open subset of its codomain.] Deduce that the hypothesis '$\mathcal{C}$ has all finite limits' in Exercise 4.19(iv) cannot be weakened to '$\mathcal{C}$ has all finite connected limits'.

(iv) Show that, for every cardinal κ, there exists a topological space in which every nonempty open set has cardinality at least κ. Use this to give a direct proof that **LH** does not have a terminal object.

(v) Does **LH** have binary products? [Your first guess is probably wrong — mine was!]