

Chapter 7. Universal constructions in category-theoretic terms.

The language of category theory has enabled us to give general definitions of “free object”, “product”, “coproduct”, “equalizer” and various other universal constructions. It is clear that these different constructions have many properties in common. Let us now look for ways to unify them, so that we will be able to prove results about them by general arguments, rather than piecemeal.

7.1. Universality in terms of initial and terminal objects. In all the above constructions, we deal with mathematical entities with certain “extra” structure, and seek one entity E with such structure that is “universal”. This suggests that we make the class of entities with such extra structure into a category, and examine the universal property of E there.

For instance, the free group on three generators is universal among systems (G, a, b, c) where G is a group, and $a, b, c \in |G|$. If we define a category whose objects are these systems (G, a, b, c) , and where a morphism $(G, a, b, c) \rightarrow (G', a', b', c')$ means a group homomorphism $f: G \rightarrow G'$ such that $f(a) = a'$, $f(b) = b'$, $f(c) = c'$, we see that the universal property of the free group (F, x, y, z) says that it has a unique morphism into every object of the category – in other words, that it is an initial object.

Similarly, given a group G , the *abelianization* of G is universal among pairs (A, f) where A is an abelian group, and f a group homomorphism $G \rightarrow A$. If we define a morphism from one such pair (A, f) to another such pair (B, g) to mean a group homomorphism $m: A \rightarrow B$ such that $mf = g$, we see that the definition of the abelianization of G says that it is initial in *this* category.

Finally, a group, a ring, a lattice, etc., with a presentation $\langle X \mid R \rangle$ clearly means an initial object in the category whose objects are groups, etc., with specified X -tuples of elements satisfying the system of equations R , and whose morphisms are homomorphisms respecting these distinguished X -tuples of elements.

The above were examples of what we named “left universal” properties in §3.8. Let us look at one “right universal” property, that of a *product* of two objects A and B in a category \mathbf{C} . We see that the relevant auxiliary category should have for objects all 3-tuples (X, a, b) , where $X \in \text{Ob}(\mathbf{C})$, $a \in \mathbf{C}(X, A)$, $b \in \mathbf{C}(X, B)$, and for morphisms $(X, a, b) \rightarrow (Y, a', b')$ all morphisms $X \rightarrow Y$ in \mathbf{C} making commuting triangles with the maps into A and B . A direct product of A and B in \mathbf{C} is seen to be a *terminal* object (P, p_1, p_2) in this category.

You can likewise easily translate the universal properties of *pushouts*, *pullbacks* and *coproducts* in arbitrary categories to those of initial or terminal objects in appropriately defined auxiliary categories.

So all the universal properties we have considered reduce to those of being an initial or a terminal object in an appropriate category. This view of universal constructions is emphasized by Lang [31, p.57 et seq.], who gives these two types of objects the more poetic designations “universally repelling” and “universally attracting”. Since a terminal object in \mathbf{C} is an initial object in \mathbf{C}^{op} , all these universal properties ultimately reduce to that of initial objects!

Lemma 6.8.2 tells us that initial (and hence terminal) objects are *unique* up to unique isomorphism. This gives us, in one fell swoop, uniqueness up to canonical isomorphism for free groups, abelianizations, products, coproducts, pushouts, pullbacks, objects presented by generators and relations, and all the other universal constructions we have considered. The canonical

isomorphisms that these constructions are “unique up to” correspond to the unique morphisms between any two initial objects of a category. I.e., given two realizations of one of our universal constructions, these isomorphisms will be the unique morphisms from each to the other that preserve the extra structure.

We will look at questions of *existence* of initial objects in §7.10.

7.2. Representable functors, and Yoneda’s Lemma. The above approach to universal constructions is impressive for its simplicity; but we would also like to relate these universal objects to the original categories in question: Though the free group on an S -tuple of generators is initial in the category of groups given with S -tuples of elements, and the kernel of a group homomorphism $f: G \rightarrow H$ is terminal in the category of groups L given with homomorphisms $L \rightarrow G$ having trivial composite with f , we also want to understand these constructions in relation to the category **Group**.

Note that the objects of the various auxiliary categories we have used can be written as pairs (X, a) , where X is an object of the original category \mathbf{C} , and a is some additional structure on X . If we write $F(X)$ for the set of *all possible values* of this additional structure (e.g., in the case that leads to the free group on a set S , the set of all S -tuples of elements of X), we find that F is in general a functor, covariant or contravariant, from \mathbf{C} to **Set**. The condition characterizing a *left* universal pair (R, u) is that for every $X \in \text{Ob}(\mathbf{C})$ and $x \in F(X)$, there should be a unique morphism $f: R \rightarrow X$ such that $F(f)(u) = x$. This condition – which we see requires a covariant F so that the latter equation will make sense – is equivalent to saying that for each object X , the set of morphisms $f \in \mathbf{C}(R, X)$ is sent bijectively to the set of elements of $F(X)$ by the map $f \mapsto F(f)(u)$. The bijectivity of this correspondence for each X leads to an isomorphism between the functor $\mathbf{C}(R, -)$, i.e., $h_R: \mathbf{C} \rightarrow \mathbf{Set}$, and the given functor $F: \mathbf{C} \rightarrow \mathbf{Set}$. Thus, the universal property of R can be formulated as a statement of this isomorphism:

Theorem 7.2.1. *Let \mathbf{C} be a category, and $F: \mathbf{C} \rightarrow \mathbf{Set}$ a functor. Then the following data are equivalent:*

- (i) *An object $R \in \text{Ob}(\mathbf{C})$ and an element $u \in F(R)$ having the universal property that for all $X \in \text{Ob}(\mathbf{C})$ and all $x \in F(X)$, there exists a unique $f \in \mathbf{C}(R, X)$ such that $F(f)(u) = x$.*
- (ii) *An initial object (R, u) in the category whose objects are all ordered pairs (X, x) with $X \in \text{Ob}(\mathbf{C})$ and $x \in F(X)$, and whose morphisms are morphisms among the first components of these pairs which respect the second components.*
- (iii) *An object $R \in \text{Ob}(\mathbf{C})$ and an isomorphism of functors $i: h_R \cong F$ in $\mathbf{Set}^{\mathbf{C}}$.*

Namely, given (R, u) as in (i) or (ii), one obtains an isomorphism i as in (iii) by letting $i(X)$ take $f \in h_R(X)$ to $F(f)(u) \in F(X)$, while in the reverse direction, one obtains u from i as $i(R)(\text{id}_R)$.

Sketch of Proof. The equivalence of the structures described in (i) and (ii) is immediate.

Concerning our description of how to pass from these structures to that of (iii), it is a straightforward verification that for any $u \in F(R)$, the map i described there gives a *morphism of functors* $h_R \rightarrow F$. That this is an isomorphism is then the content of the universal property of (i). In the opposite direction, given an isomorphism i as in (iii), if u is defined as indicated, then the universal property of (i) is just a restatement of the bijectivity of the maps $i(X): h_R(X) \rightarrow F(X)$.

Finally, it is easy to check that if one goes as above from universal element to isomorphism of functors and back, one recovers the original element, and if one goes from isomorphism to

universal element and back, one recovers the original isomorphism. \square

Exercise 7.2:1. Write out the “straightforward verifications” referred to in the second sentence of the above proof, and those implied in the phrases “is then the content of” and “is just a restatement of” in the next two sentences.

Dualizing (i.e., applying Theorem 7.2.1 to \mathbf{C}^{op} and stating the resulting assertion in terms of \mathbf{C}), we get

Theorem 7.2.2. *Let \mathbf{C} be a category, and F a contravariant functor from \mathbf{C} to \mathbf{Set} (i.e., a functor $\mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}$). Then the following data are equivalent:*

- (i) *An object $R \in \text{Ob}(\mathbf{C})$ and an element $u \in F(R)$ with the universal property that for any $X \in \text{Ob}(\mathbf{C})$ and $x \in F(X)$, there exists a unique $f \in \mathbf{C}(X, R)$ such that $F(f)(u) = x$.*
- (ii) *A terminal object (R, u) in the category whose objects are all ordered pairs (X, x) with $X \in \text{Ob}(\mathbf{C})$ and $x \in F(X)$, and whose morphisms are morphisms among the first components of these pairs which respect the second components.*
- (iii) *An object $R \in \text{Ob}(\mathbf{C})$ and an isomorphism of contravariant functors $i: h^R \cong F$ in $\mathbf{Set}^{\mathbf{C}^{\text{op}}}$.*

Namely, given (R, u) as in (i) or (ii), one obtains an isomorphism i as in (iii) by letting $i(X)$ take $f \in h^R(X)$ to $F(f)(u) \in F(X)$, while in the reverse direction, one obtains u from i as $i(R)(\text{id}_R)$. \square

Note that in Theorem 7.2.1(ii), the last phrase, “which respect second components”, meant that for a morphism $f: X \rightarrow Y$ to be considered a morphism $(X, x) \rightarrow (Y, y)$, we required $F(f)(x) = y$, while in Theorem 7.2.2(ii), the corresponding condition is $F(f)(y) = x$.

We remark that the auxiliary categories used in point (ii) of the above two theorems are comma categories, $(1 \downarrow F)$ (Exercise 6.8:26(iii)).

The properties described above have names:

Definition 7.2.3. *Let \mathbf{C} be a category.*

A covariant functor $F: \mathbf{C} \rightarrow \mathbf{Set}$ is said to be representable if it is isomorphic to a covariant hom-functor h_R for some $R \in \text{Ob}(\mathbf{C})$.

A contravariant functor $F: \mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}$ is likewise said to be representable if it is isomorphic to a contravariant hom-functor h^R for some $R \in \text{Ob}(\mathbf{C})$.

In each case, R is called the representing object for F , and if i is the given isomorphism of functors, then $i(R)(\text{id}_R)$ is called the associated universal element of $F(R)$.

So from this point of view, universal problems of the sort considered above in a category \mathbf{C} are questions of the *representability* of certain set-valued functors on \mathbf{C} . Let us examine a few set-valued functors, and see which of them are representable.

If U is the underlying-set functor on **Group**, a representing object for U should be a group with a universal element of its underlying set. The object with this property is the free group on one generator. More generally, if a category has free objects with respect to a concretization U , then U will be represented by the free object on one generator, while the free object on a general set I can be characterized as representing the functor U^I (Definition 6.8.5).

The functor associating to every group the set of its elements of exponent 2 is represented by the group \mathbb{Z}_2 . More generally, the group with presentation by generators and relations $\langle X \mid R \rangle$ represents the functor associating to every group G the set of X -tuples of members of G which

satisfy the relations R .

Is the functor associating to every commutative ring K the set $|K[t]|$ of all polynomials over K in one indeterminate t representable? A representing object would be a ring R with a universal polynomial $u(t) \in R[t]$. The universal property would say that given any polynomial $p(t)$ over any ring K , there should exist a unique homomorphism $R \rightarrow K$ which, applied coefficient-wise to polynomials, carries $u(t)$ to $p(t)$. But clearly there is a problem here: The polynomial u will have some degree n , and if we choose a polynomial p of degree $> n$, it cannot be obtained from u in this way. So the set-of-polynomials functor is not representable.

However, there is a concept close to that of polynomial but not subject to the restriction that only finitely many of the coefficients be nonzero, that of a *formal power series* $a_0 + a_1t + a_2t^2 + \dots$. If K is a ring, then the ring of formal power series over K is denoted $K[[t]]$; its underlying set $|K[[t]]| = \{a_0 + a_1t + a_2t^2 + \dots\}$ can be identified with the set of all sequences (a_0, a_1, \dots) of elements of K , i.e., with $|K|^\omega = U^\omega(K)$. We know that the functor U^ω is represented by the free commutative ring on an ω -tuple of generators, that is, the polynomial ring $\mathbb{Z}[A_0, A_1, \dots]$. And indeed, the formal power series ring over this polynomial ring contains the element $A_0 + A_1t + A_2t^2 + \dots$, which clearly has the property of a universal power series.

Exercise 7.2:2. (i) Show that the functor associating to every monoid S the set of its invertible elements is representable, but that the functor associating to S the set of its right-invertible elements is not.

(ii) What about the functor associating to every monoid S the set of pairs (x, y) such that $xy = e$ and $yx = e$? The set of pairs (x, y) merely satisfying $xy = e$? The set of 3-tuples (x, y, z) such that $xy = xz = e$?

(iii) Determine which, if any, of the functors mentioned in (i) and (ii) are isomorphic to one another.

Exercise 7.2:3. Let P denote the contravariant power-set functor, associating to every set X the set $\mathbf{P}(X)$ of its subsets, and E the contravariant functor associating to every set X the set $\mathbf{E}(X)$ of equivalence relations on X . Determine whether each of these is representable.

Exercise 7.2:4. Let A, B be objects of a category \mathbf{C} . Describe a set-valued functor F on \mathbf{C} such that a *product* of A and B , if it exists in \mathbf{C} , means a representing object for F , and likewise a functor G such that a *coproduct* of A and B in \mathbf{C} means a representing object for G . (One of these will be covariant and the other contravariant.)

Exercise 7.2:5. Let (\mathbf{C}, U) be a concrete category. Show that the following conditions are equivalent. (a) The concretization functor U is *representable*. (b) \mathbf{C} has a free object on one generator. Moreover, show that if \mathbf{C} has coproducts, then these are also equivalent to (b') \mathbf{C} has free objects on all sets.

Students who know some Lie group theory might try

Exercise 7.2:6. Let \mathbf{LieGp} denote the category of Lie groups and continuous group homomorphisms. Let $T: \mathbf{LieGp} \rightarrow \mathbf{Set}$ denote the functor associating to a Lie group L the set of tangent vectors to L at the neutral element. Which of the following covariant functors $\mathbf{LieGp} \rightarrow \mathbf{Set}$ are representable? (a) the functor T , (b) the functor $T^2: L \mapsto T(L) \times T(L)$, (c) the functor $L \mapsto \{(x, y) \in T(L) \times T(L) \mid [x, y] = x\}$.

Exercise 7.2:7. Given a set X , let $\mathbf{GpStruct}(X)$ denote the set of all group-structures on X (consisting of a composition operation μ , an inverse operation ι , and a neutral element e). A group can be considered as a set X given with an element $s \in \mathbf{GpStruct}(X)$, and the category \mathbf{Group} has an initial object. This looks as though it should mean the underlying set of this group is a representing object for $\mathbf{GpStruct}$; but something is clearly wrong, since a map from

this set into a set X does not determine a group structure on X . Resolve this paradox.

The equivalence, in each of Theorems 7.2.1 and 7.2.2, of parts (ii) and (iii) shows that the concept of representable functor can be characterized in terms of initial and terminal objects. The reverse is also true:

Exercise 7.2:8. Let \mathbf{C} be any category. Display a covariant functor F and a contravariant functor G from \mathbf{C} to \mathbf{Set} such that an initial, respectively a terminal object of \mathbf{C} is equivalent to a representing object for F , respectively G .

The implication (i) \Rightarrow (iii) in Theorem 7.2.1 shows that an isomorphism between the hom-functor h_R associated with an object R , and an arbitrary functor F , is equivalent to a specification of an element of $F(R)$ with the universal property given in (i). In fact, *every* morphism, invertible or not, from a hom-functor h_R to a functor F corresponds to a choice of *some* element of $F(R)$. Though utterly simple to prove, this is an important tool. We give both this result and its contravariant dual in

Theorem 7.2.4 (Yoneda's Lemma). *Let \mathbf{C} be a category, and R an object of \mathbf{C} .*

If $F: \mathbf{C} \rightarrow \mathbf{Set}$ is a covariant functor, then morphisms $f: h_R \rightarrow F$ are in one-to-one correspondence with elements of $F(R)$, under the map $f \mapsto f(R)(\text{id}_R)$.

Likewise, if $F: \mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}$ is a contravariant functor, morphisms $f: h^R \rightarrow F$ are in one-to-one correspondence with elements of $F(R)$, again under the map $f \mapsto f(R)(\text{id}_R)$.

Proof. In the covariant case, we must describe how to get from an element $x \in F(R)$ an appropriate morphism $f_x: h_R \rightarrow F$. We define f_x to carry $a \in h_R(X) = \mathbf{C}(R, X)$ to $F(a)(x) \in F(X)$. The verification that this is a morphism of functors, and that this construction is inverse to the indicated map from morphisms of functors to elements of $F(R)$, is immediate.

The contravariant case follows by duality (or by the dualized argument). \square

Again –

Exercise 7.2:9. Show the verifications omitted in the proof of the above result.

The following line of thought yields some intuition on Yoneda's Lemma. Recall that if G is a group, then a G -set, i.e., a functor from the category G_{cat} to \mathbf{Set} , can be looked at as a (possibly non-faithful) representation of G by permutations. In the same way, for any category \mathbf{C} , a functor $F: \mathbf{C} \rightarrow \mathbf{Set}$ can be thought of as a (possibly non-faithful) representation of \mathbf{C} by sets and set maps. Like a G -set, such a representation F can be regarded as a mathematical ‘‘object’’; in this case the ‘‘elements’’ of that object are the members of the sets $F(X)$ ($X \in \text{Ob}(\mathbf{C})$). This was the point of view of our development of Cayley's Theorem for small categories. In proving that result, we constructed such an object by introducing one generator in each set $F(X)$, and no relations; in the discussion that followed we observed that if one introduced only a generator in the set $F(X)$ for a *particular* $X \in \text{Ob}(\mathbf{C})$, and again no relations, the resulting ‘‘freely generated’’ object would be essentially the hom-functor which we named h_X . Yoneda's Lemma is the statement of the universal property of this ‘‘free’’ construction – that a morphism from this ‘‘representation of \mathbf{C} by sets’’ to any other ‘‘representation of \mathbf{C} by sets’’ is uniquely determined by specifying where the one generator, the identity element $\text{id}_X \in h_X(X)$, is to be sent. We make this formulation explicit in

Corollary 7.2.5. *Let \mathbf{C} be a category and R an object of \mathbf{C} .*

In the (large) category whose objects are pairs (F, x) where F is a functor $\mathbf{C} \rightarrow \mathbf{Set}$ and x an element of $F(R)$, the pair (h_R, id_R) is the initial object. Equivalently, the object $h_R \in \text{Ob}(\mathbf{Set}^{\mathbf{C}})$ is a representing object for the “evaluation at R ” functor $\mathbf{Set}^{\mathbf{C}} \rightarrow \mathbf{Set}$, the universal element being $\text{id}_R \in h_R(R)$.

Likewise, in the category whose objects are pairs (F, x) where F is a functor $\mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}$ and x an element of $F(R)$, the pair (h^R, id_R) is the initial object; equivalently, the object $h^R \in \text{Ob}(\mathbf{Set}^{\mathbf{C}^{\text{op}}})$ represents the (again covariant!) “evaluation at R ” functor $\mathbf{Set}^{\mathbf{C}^{\text{op}}} \rightarrow \mathbf{Set}$. \square

This points to a general principle worth keeping in mind: when dealing with a morphism from a hom-functor to an arbitrary set-valued functor, look at its value on the identity map!

What if we apply Yoneda’s Lemma (covariant or contravariant) to the case where the arbitrary functor F is another hom-functor h_S , respectively h^S ? We get

Corollary 7.2.6. *Let \mathbf{C} be a category.*

Then for any two objects $R, S \in \text{Ob}(\mathbf{C})$, the morphisms from h_R to h_S as functors $\mathbf{C} \rightarrow \mathbf{Set}$ are in one-to-one correspondence with morphisms $S \rightarrow R$. Thus, the mapping $R \mapsto h_R$ gives a contravariant full embedding of \mathbf{C} in $\mathbf{Set}^{\mathbf{C}}$, the “Yoneda embedding”.

Likewise, morphisms from h^R to h^S as functors $\mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}$ correspond to morphisms $R \rightarrow S$, giving a covariant full “Yoneda embedding” of \mathbf{C} in $\mathbf{Set}^{\mathbf{C}^{\text{op}}}$.

These two embeddings may both be obtained from the bivariate hom-functor $\mathbf{C}^{\text{op}} \times \mathbf{C} \rightarrow \mathbf{Set}$ by distinguishing one or the other argument, i.e., regarding this bifunctor in one case as a functor $\mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}^{\mathbf{C}}$, and in the other as a functor $\mathbf{C} \rightarrow \mathbf{Set}^{\mathbf{C}^{\text{op}}}$.

Sketch of Proof. By Lemma 6.10.1 the bivariate hom functor does indeed yield functors $\mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}^{\mathbf{C}}$ and $\mathbf{C} \rightarrow \mathbf{Set}^{\mathbf{C}^{\text{op}}}$ on distinguishing one or the other argument, and we see that the object R is sent to h_R , respectively h^R . Given a morphism $f: S \rightarrow R$ in \mathbf{C} , one verifies that the induced morphism of functors $h_f: h_R \rightarrow h_S$ takes id_R to $f \in h_S(R)$. Yoneda’s Lemma with $F = h_S$ tells us that the map $f \mapsto h_f$ is one-to-one and onto, so our functor $\mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}^{\mathbf{C}}$ is full and faithful. The contravariant case follows by duality. \square

Exercise 7.2:10. Verify the above characterization of the morphism of functors induced by a morphism $f: S \rightarrow R$.

Exercise 7.2:11. Show how to answer most of the parts of Exercise 6.9:5, and also Exercise 6.9:7(i), using Yoneda’s Lemma.

Remark 7.2.7. It may seem paradoxical that we get the *contravariant* Yoneda embedding using *covariant* hom-functors, and the *covariant* Yoneda embedding using *contravariant* hom-functors, but there is a simple explanation. When we write the hom bifunctor $\mathbf{C}^{\text{op}} \times \mathbf{C} \rightarrow \mathbf{Set}$ as a functor to a functor category, $\mathbf{C} \rightarrow \mathbf{Set}^{\mathbf{C}^{\text{op}}}$ or $\mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}^{\mathbf{C}}$, by distinguishing one variable, the variance in that variable determines the variance of the resulting Yoneda embedding, while the variance in the other variable determines the variance of the hom-functors that the embedding takes on as its values. Whichever way we slice it, we get covariance in one, and contravariance in the other.

What is the value of the Yoneda embedding? First, note that categories of the form $\mathbf{Set}^{\mathbf{C}}$ have very good properties; e.g., they have small limits and colimits. Hence Yoneda embeddings

embed arbitrary categories into “good” categories. Moreover, if one wishes to extend a category \mathbf{C} by adjoining additional objects with particular properties, one can often do this by identifying \mathbf{C} with the category of representable contravariant functors on \mathbf{C} , or the opposite of the category of representable covariant functors, and then taking for the additional objects certain other functors that are not quite representable.

In §6.5 we saw that systems of universal constructions could frequently be linked together, by natural morphisms among the constructed objects, to give functors. From the above corollary, we see that this should happen in situations where the functors that these universal objects are constructed to *represent* are linked by a corresponding system of morphisms of functors, in other words (by Lemma 6.10.1) where they form the components of a *bifunctor*. There is a slight complication in formulating this precisely, because the given representable functors are not themselves the hom-functors h_R or h^R , but only isomorphic to these, and the choice of representing objects R is likewise determined only up to isomorphism. To prepare ourselves for this complication, let us prove a lemma showing that a system of objects separately isomorphic to the values of a functor in fact form the values of an isomorphic functor.

Lemma 7.2.8. *Let $F: \mathbf{C} \rightarrow \mathbf{D}$ be a functor, and for each $X \in \text{Ob}(\mathbf{C})$, let $i(X)$ be an isomorphism of $F(X)$ with another object $G(X) \in \text{Ob}(\mathbf{D})$.*

Then there is a unique way to assign to each morphism of \mathbf{C} , $f \in \mathbf{C}(X, Y)$ a morphism $G(f) \in \mathbf{D}(G(X), G(Y))$ so that the objects $G(X)$ and morphisms $G(f)$ constitute a functor $G: \mathbf{C} \rightarrow \mathbf{D}$, and i constitutes an isomorphism of functors, $F \cong G$.

Proof. If G is to be a functor and i a morphism of functors, then for each $f \in \mathbf{C}(X, Y)$ we must have $G(f) i(X) = i(Y) F(f)$. Since $i(X)$ is an isomorphism, we can rewrite this as $G(f) = i(Y) F(f) i(X)^{-1}$. It is straightforward to verify that G , so defined on morphisms, is indeed a functor. This definition of $G(f)$ insures that i is a morphism of functors $F \rightarrow G$, and it clearly has an inverse, defined by $i^{-1}(X) = i(X)^{-1}$. \square

Exercise 7.2:12. Write out the verification that G , constructed as above, is a functor.

We can now get our desired result about tying representing objects together into a functor. In thinking about results such as the next lemma, I find it useful to keep in mind the case where $\mathbf{C} = \mathbf{Set}$, $\mathbf{D} = \mathbf{Group}$, and A is the bifunctor associating to every set X and group G the set $|G|^X$ of X -tuples of elements of G .

Lemma 7.2.9. *Suppose that \mathbf{C} and \mathbf{D} are categories, and that for each $X \in \text{Ob}(\mathbf{C})$ we are given a functor $A(X, -): \mathbf{D} \rightarrow \mathbf{Set}$ and an object $F(X) \in \text{Ob}(\mathbf{D})$ representing this functor, via an isomorphism $i(X): A(X, -) \cong h_{F(X)}$.*

Then if the given functors $A(X, -)$ are in fact the values of a bifunctor $A: \mathbf{C}^{\text{op}} \times \mathbf{D} \rightarrow \mathbf{Set}$ at the objects of \mathbf{C} , then the objects $F(X)$ of \mathbf{D} can be made the values of a functor $F: \mathbf{C} \rightarrow \mathbf{D}$ in a unique way so that the isomorphisms $i(X)$ comprise an isomorphism of bifunctors

$$(7.2.10) \quad i: A(-, -) \cong \mathbf{D}(F(-), -).$$

Conversely, if the objects $F(X)$ are the values at the objects X of a functor $F: \mathbf{C} \rightarrow \mathbf{D}$, we can make the family of functors $A(X, -)$ into a bifunctor $A: \mathbf{C}^{\text{op}} \times \mathbf{D} \rightarrow \mathbf{Set}$ in a unique way so that the isomorphisms $i(X)$ again give an isomorphism (7.2.10) of bifunctors.

Proof. On the one hand, if $A: \mathbf{C}^{\text{op}} \times \mathbf{D} \rightarrow \mathbf{Set}$ is a bifunctor, the induced system of functors $A(X, -): \mathbf{D} \rightarrow \mathbf{Set}$ will together constitute a single functor which we may call $B: \mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}^{\mathbf{D}}$ (Lemma 6.10.1). For each $X \in \text{Ob}(\mathbf{C})$ we have an isomorphism $i(X)$ of $B(X)$ with a hom-functor $h_{F(X)}$, so by the preceding lemma we get an isomorphic functor $C: \mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}^{\mathbf{D}}$, such that $C(X) = h_{F(X)}$, and the isomorphism $i: B \cong C$ is made up of the $i(X)$'s. Now by Corollary 7.2.6, the covariant hom-functors h_Y ($Y \in \text{Ob}(\mathbf{D})$) form a full subcategory of $\mathbf{Set}^{\mathbf{D}}$ isomorphic to \mathbf{D} via the Yoneda embedding $Y \mapsto h_Y$. Hence the functor $C: \mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}^{\mathbf{D}}$ is induced by precomposing this embedding $\mathbf{D}^{\text{op}} \rightarrow \mathbf{Set}^{\mathbf{D}}$ with a unique functor $\mathbf{C}^{\text{op}} \rightarrow \mathbf{D}^{\text{op}}$, which is equivalent to a functor $F: \mathbf{C} \rightarrow \mathbf{D}$, and this F is the functor of the statement of the lemma.

Inversely, if F is given as a functor, let us consider each functor $A(X, -)$ as an object $B(X)$ of $\mathbf{Set}^{\mathbf{D}}$. Then for each X we have an isomorphism $i(X): B(X) \cong h_{F(X)}$, and applying the preceding lemma to the isomorphisms $i(X)^{-1}$, we conclude that the objects $B(X)$ are the values of a functor $B: \mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}^{\mathbf{D}}$, which we may regard as a bifunctor $A: \mathbf{C}^{\text{op}} \times \mathbf{D} \rightarrow \mathbf{Set}$, and again the values of i become an isomorphism of bifunctors. \square

The above lemma concerns systems of objects representing covariant hom-functors; let us state the corresponding result for contravariant hom-functors. A priori, this means replacing \mathbf{D} by \mathbf{D}^{op} . But it is then natural to replace the ‘‘parametrizing’’ category \mathbf{C}^{op} by \mathbf{C} so as to keep the parametrization of the constructed objects of \mathbf{D} covariant. And having done that much, why not interchange the names of \mathbf{C} and \mathbf{D} so as to get a set-up parallel to that of the preceding case? Doing so, we get

Lemma 7.2.11. *Suppose that \mathbf{C} and \mathbf{D} are categories, and that for each $Y \in \text{Ob}(\mathbf{D})$ we are given a functor $A(-, Y): \mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}$ and an object $U(Y) \in \text{Ob}(\mathbf{C})$ representing this contravariant functor, via an isomorphism $j(Y): A(-, Y) \cong h^{U(Y)}$.*

Then if the given functors $A(-, Y)$ are the values of a bifunctor $A: \mathbf{C}^{\text{op}} \times \mathbf{D} \rightarrow \mathbf{Set}$ at the objects of \mathbf{D} , the family of objects $U(Y)$ of \mathbf{C} can be made the values of a functor $U: \mathbf{D} \rightarrow \mathbf{C}$ in a unique way so that the isomorphisms $j(Y)$ constitute an isomorphism of bifunctors

$$(7.2.12) \quad j: A(-, -) \cong \mathbf{C}(-, U(-)).$$

Conversely, if the objects $U(Y)$ are the images of the objects Y under a functor $U: \mathbf{D} \rightarrow \mathbf{C}$, we can make the family of functors $A(-, Y)$ into a bifunctor $A: \mathbf{C}^{\text{op}} \times \mathbf{D} \rightarrow \mathbf{Set}$ in a unique way so that the isomorphisms $j(Y)$ together give an isomorphism (7.2.12) of bifunctors. \square

7.3. Adjoint functors. Let us look at some examples of the situation of the two preceding lemmas – families of objects that we characterized individually as the representing objects for certain naturally occurring functors, but that turned out, themselves, to fit together into a functor. By those lemmas, this means that the system of functors that these objects represented fit together into a bifunctor. We shall see that in each of these cases, this structure of bifunctor was actually present in the original situation, providing an explanation of why our constructions yielded functors.

The free group on each set X is the object of **Group** representing the functor $G \mapsto |G|^X = \mathbf{Set}(X, U(G))$. So the free group functor arises by representing the family of functors **Group** \rightarrow **Set** obtained by inserting all sets as the first argument of the bifunctor

$$\mathbf{Set}(-, U(-)): \mathbf{Set}^{\text{op}} \times \mathbf{Group} \rightarrow \mathbf{Set}.$$

The analogous description obviously applies in any category \mathbf{C} having free objects with respect to

a concretization $U: \mathbf{C} \rightarrow \mathbf{Set}$.

If G is a group, the *abelianization* of G is the object of \mathbf{Ab} representing the functor $\mathbf{Ab} \rightarrow \mathbf{Set}$ given by $A \mapsto \mathbf{Group}(G, A)$. The symbol $\mathbf{Group}(G, A)$ makes sense because \mathbf{Ab} is a subcategory of \mathbf{Group} , but to put this example in the context of the general pattern, let us write V for the inclusion functor of \mathbf{Ab} in \mathbf{Group} . We then see that the abelianization functor arises by representing the family of set-valued functors obtained by inserting values in the first argument of the bifunctor

$$\mathbf{Group}(-, V(-)): \mathbf{Group}^{\text{op}} \times \mathbf{Ab} \rightarrow \mathbf{Set}.$$

In the same way, if W denotes the forgetful functor $\mathbf{Group} \rightarrow \mathbf{Monoid}$, then the functor taking a monoid to its universal enveloping group arises by representing the family of set-valued functors obtained by inserting values in the first argument of the bifunctor

$$\mathbf{Monoid}(-, W(-)): \mathbf{Monoid}^{\text{op}} \times \mathbf{Group} \rightarrow \mathbf{Set}.$$

The above were left universal examples, that is, constructions $F: \mathbf{C} \rightarrow \mathbf{D}$ such that each object $F(X)$ represented a covariant functor $\mathbf{D} \rightarrow \mathbf{Set}$. We see that in each such case, the bifunctor from which these covariant functors were extracted had the form

$$(7.3.1) \quad \mathbf{C}(-, U(-)): \mathbf{C}^{\text{op}} \times \mathbf{D} \rightarrow \mathbf{Set},$$

for some functor $U: \mathbf{D} \rightarrow \mathbf{C}$. Taking (7.3.1) to be the A in Lemma 7.2.9, we see that the universal properties of the objects $F(X)$ in terms of U can be formulated in each of these cases as

$$\mathbf{C}(-, U(-)) \cong \mathbf{D}(F(-), -)$$

– a strikingly symmetrical condition!

Let us consider one right universal example. Given a monoid S , we considered above the construction of the universal group G with a homomorphism of S into G_{md} ; but there is also a universal group G with a homomorphism of G_{md} into S , namely the group $G = S_{\text{inv}}$ of invertible elements (“units”) of S . If we write $F: \mathbf{Group} \rightarrow \mathbf{Monoid}$ for the forgetful functor $G \mapsto G_{\text{md}}$, and call the above group-of-units functor $U: \mathbf{Monoid} \rightarrow \mathbf{Group}$, we see that $U(S)$ represents the contravariant functor associating to each group G the set $\mathbf{Monoid}(F(G), S)$. If we write \mathbf{C} and \mathbf{D} for \mathbf{Group} and \mathbf{Monoid} , then on taking $\mathbf{D}(F(-), -)$ for the bifunctor A in the last formulation of Lemma 7.2.11, we get an isomorphism characterizing this right universal construction U :

$$\mathbf{D}(F(-), -) \cong \mathbf{C}(-, U(-)).$$

This is exactly the same as the isomorphism characterizing our examples of left universal constructions – but written in reverse order, and looked at as characterizing U in terms of F , rather than F in terms of U ! The fact that these two situations are characterized by the same isomorphism means that a functor F gives objects representing the covariant functors $\mathbf{C}(X, U(-))$ if and only if U gives objects representing the contravariant functors $\mathbf{D}(F(-), Y)$.

Let us test this conclusion, by turning our characterization of the free group construction upside down. Since the free group $F(X)$ on a set X is left universal among groups G with set maps of X into their underlying sets $U(G)$, the *underlying set* $U(G)$ of a group G should be right-universal among all sets X with group homomorphisms from the free group $F(X)$ into G . And indeed, though it may seem bizarre to treat the free-group construction as given and the

underlying-set construction as something to be characterized, the universal property certainly holds: For any group G , $U(G)$ is a set with a homomorphism $u: F(U(G)) \rightarrow G$, such that given any homomorphism f from a free group $F(X)$ on any set into G , there is a unique set map $h: X \rightarrow U(G)$ (which you should be able to describe) such that $f = uF(h)$. This property of underlying sets is sometimes even useful. For instance, in showing that every group can be presented by generators and relations, one wishes to write an arbitrary group G as a homomorphic image of a free group on some set X . The above property says that there is a universal choice of such X , namely the underlying set $U(G)$ of G .

Before setting out to tie together all our ways of describing these universal constructions, let us prove a lemma that will allow us to relate isomorphisms of bifunctors as above to systems of maps $X \rightarrow U(F(X))$ and $F(U(Y)) \rightarrow Y$. (This observation is an instance of the general principle noted following Corollary 7.2.5.)

Lemma 7.3.2. *Let \mathbf{C} and \mathbf{D} be categories and $U: \mathbf{D} \rightarrow \mathbf{C}$, $F: \mathbf{C} \rightarrow \mathbf{D}$ functors, and consider the two bifunctors $\mathbf{C}^{\text{op}} \times \mathbf{D} \rightarrow \mathbf{Set}$,*

$$\mathbf{C}(-, U(-)), \quad \mathbf{D}(F(-), -).$$

Then a morphism of bifunctors

$$(7.3.3) \quad a: \mathbf{C}(-, U(-)) \rightarrow \mathbf{D}(F(-), -)$$

is determined by its values on identity morphisms $\text{id}_{U(D)} \in \mathbf{C}(U(D), U(D))$ ($D \in \text{Ob}(\mathbf{D})$). In fact, given a as above, if we write $\alpha(D) = a(U(D), D)(\text{id}_{U(D)}) \in \mathbf{D}(F(U(D)), D)$, then this family of morphisms comprises a morphism of functors,

$$(7.3.4) \quad \alpha: FU \rightarrow \text{Id}_{\mathbf{D}}$$

and this construction yields a bijection between morphisms (7.3.3) and morphisms (7.3.4). Given a morphism (7.3.4), the corresponding morphism (7.3.3) can be described as acting on $f \in \mathbf{C}(C, U(D))$ by first applying F to get $F(f): F(C) \rightarrow FU(D)$, then composing this with $\alpha(D): FU(D) \rightarrow D$, getting $a(f) = \alpha(D)F(f): F(C) \rightarrow D$.

Likewise, a morphism of bifunctors in the opposite direction to (7.3.3),

$$(7.3.5) \quad b: \mathbf{D}(F(-), -) \rightarrow \mathbf{C}(-, U(-))$$

is determined by its values on identity morphisms, in this case morphisms $\text{id}_{F(C)} \in \mathbf{D}(F(C), F(C))$ ($C \in \text{Ob}(\mathbf{C})$), and writing $\beta(C) = b(C, F(C))(\text{id}_{F(C)}) \in \mathbf{C}(C, U(F(C)))$, we get a bijection between morphisms (7.3.5) and morphisms

$$(7.3.6) \quad \beta: \text{Id}_{\mathbf{C}} \rightarrow UF.$$

Given β , the corresponding morphism b can be described as taking $f \in \mathbf{D}(F(C), D)$ to $U(f)\beta(C) \in \mathbf{C}(C, U(D))$.

Sketch of Proof. Consider a morphism a as in (7.3.3). For each $D \in \text{Ob}(\mathbf{D})$ this gives a morphism of functors $\mathbf{C}(-, U(D)) \rightarrow \mathbf{D}(F(-), D)$. Since the first of these functors is $h^{U(D)}$, the Yoneda Lemma says this morphism is determined by its value on the identity morphism of $U(D)$. It is straightforward to verify that the condition that these morphisms of functors $\mathbf{C}(-, U(D)) \rightarrow \mathbf{D}(F(-), D)$ should comprise a single morphism of bifunctors (7.3.3) is equivalent to the condition that the values of these morphisms on identities should comprise a morphism of functors (7.3.4).

The reader can easily check that the description of how to recover (7.3.3) from (7.3.4) also leads to a morphism of functors, and that this construction is inverse to the first.

The second paragraph follows by duality. \square

Exercise 7.3:1. Write out the ‘‘straightforward’’ verification and the ‘‘easy check’’ referred to in the first paragraph of the proof of the lemma.

To get a feel for the above construction, you might start with the morphism of bifunctors a that associates to every set map from a set X to the underlying set $U(G)$ of a group G the induced group homomorphism from the free group $F(X)$ into G . Determine the morphism of functors α that the above construction yields, and check explicitly that the ‘‘inverse’’ construction described does indeed recover a . In this example, one finds that a is invertible; calling its inverse b , you might similarly work out for this b the constructions of the second assertion of the lemma.

With the help of Lemmas 7.2.9, 7.2.11 and 7.3.2, we can now give several descriptions of the type of universal construction discussed at the beginning of this section.

Theorem 7.3.7. *Let \mathbf{C} and \mathbf{D} be categories. Then the following data are equivalent:*

(i) *A pair of functors $U: \mathbf{D} \rightarrow \mathbf{C}$, $F: \mathbf{C} \rightarrow \mathbf{D}$, and an isomorphism*

$$i: \mathbf{C}(-, U(-)) \cong \mathbf{D}(F(-), -)$$

of functors $\mathbf{C}^{\text{op}} \times \mathbf{D} \rightarrow \mathbf{Set}$.

(ii) *A functor $U: \mathbf{D} \rightarrow \mathbf{C}$, and for every $C \in \text{Ob}(\mathbf{C})$, an object $R_C \in \text{Ob}(\mathbf{D})$ and an element $u_C \in \mathbf{C}(C, U(R_C))$ which are universal among such object-element pairs, i.e., which represent the covariant functor $\mathbf{C}(C, U(-)): \mathbf{D} \rightarrow \mathbf{Set}$ (cf. Theorem 7.2.1 and Definition 7.2.3).*

(ii*) *A functor $F: \mathbf{C} \rightarrow \mathbf{D}$, and for every $D \in \text{Ob}(\mathbf{D})$, an object $R_D \in \text{Ob}(\mathbf{C})$ and an element $v_D \in \mathbf{D}(F(R_D), D)$ which are universal among such object-element pairs, i.e., which represent the contravariant functor $\mathbf{D}(F(-), D): \mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}$.*

(iii) *A pair of functors $U: \mathbf{D} \rightarrow \mathbf{C}$, $F: \mathbf{C} \rightarrow \mathbf{D}$, and a pair of morphisms of functors*

$$\eta: \text{Id}_{\mathbf{C}} \rightarrow UF, \quad \varepsilon: FU \rightarrow \text{Id}_{\mathbf{D}},$$

such that the two composites

$$U \xrightarrow{\eta \circ U} UFU \xrightarrow{U \circ \varepsilon} U, \quad F \xrightarrow{F \circ \eta} FUF \xrightarrow{\varepsilon \circ F} F,$$

are the identity morphisms of U and F respectively. (For the ‘‘ \circ ’’ notation see Lemma 6.10.2.)

Sketch of Proof. The equivalence of (i) with (ii) and with (ii*) is given by Lemma 7.2.9 with $A(-, -) = \mathbf{C}(-, U(-))$, and Lemma 7.2.11 with $A(-, -) = \mathbf{D}(F(-), -)$, respectively. By Lemma 7.3.2, an isomorphism of bifunctors as in (i) must correspond to a pair of morphisms of functors $\eta: \text{Id}_{\mathbf{C}} \rightarrow UF$, $\varepsilon: FU \rightarrow \text{Id}_{\mathbf{D}}$ which induce mutually inverse morphisms of bifunctors. I claim that the conditions needed for these induced morphisms to be mutually inverse are those shown diagrammatically in (iii).

In the verification of this statement (made an exercise below), one assumes α and β given as in Lemma 7.3.2, uses the formulas for a and b in terms of these to express the composites ab and ba , and must prove that these composites are the identity morphisms. By Yoneda’s Lemma, it suffices to check these equalities on appropriate identity morphisms. (With what objects of \mathbf{C} and \mathbf{D} in the slots of $\mathbf{D}(F(-), -)$, respectively $\mathbf{C}(-, U(-))$?) This approach quickly gives the

desired statements. However, if one prefers to see directly that these statements are equivalent to ab and ba fixing *all* morphisms $f \in \mathbf{D}(F(C), D)$, respectively $g \in \mathbf{C}(C, U(D))$, then one may combine the equations saying that the latter conclusions hold with the commutativity of the diagram expressing the functoriality of a , respectively b , applied to the morphism f , respectively g . \square

- Exercise 7.3:2.** (i) Write out the verification sketched in the last paragraph of the above proof.
 (ii) Show that η will be composed of the “universal morphisms” u_C of point (ii) of the theorem, and ε will be composed of the universal morphisms v_D of point (ii*).*
 (iii) Take one universal construction, e.g., that of free groups, write down the equalities expressed diagrammatically in part (iii) of the above theorem for this construction in terms of maps of set- and group-elements, and explain why they hold in *this case*.

Definition 7.3.8. Given categories \mathbf{C} and \mathbf{D} and functors $U: \mathbf{D} \rightarrow \mathbf{C}$, $F: \mathbf{C} \rightarrow \mathbf{D}$, an isomorphism

$$i: \mathbf{C}(-, U(-)) \cong \mathbf{D}(F(-), -)$$

of bifunctors $\mathbf{C}^{\text{op}} \times \mathbf{D} \rightarrow \mathbf{Set}$, or equivalently, a pair of morphisms of functors ε, η satisfying the condition of point (iii) of the above theorem, is called an adjunction between U and F .

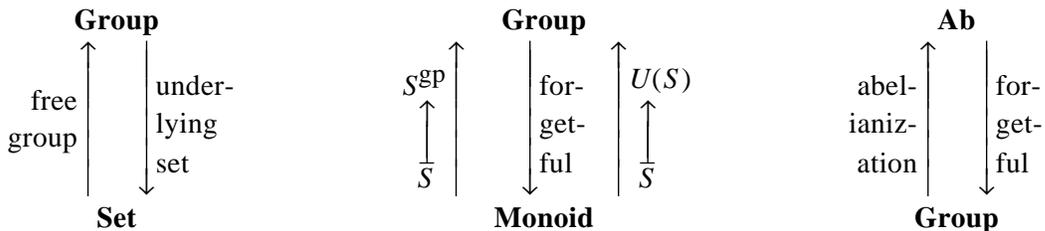
In this situation, U is called the “right adjoint” of F , and F the “left adjoint” of U (referring to their occurrence in the right and left slots of the hom-bifunctors in the above isomorphism). The morphisms of functors η and ε are called, respectively, the unit and counit of the adjunction.

Historical note: The term “adjoint” was borrowed from analysis, where the adjoint of a bounded operator between Hilbert spaces, $A: X \rightarrow Y$, is the operator $B: Y \rightarrow X$ characterized by the condition on inner products $(x, By) = (Ax, y)$.

The student who finds condition (iii) of Theorem 7.3.7 hard to grasp will be happy to know that we will not make much use of it in the next few chapters. (I have trouble with it myself.) But we will use the morphisms η and ε named in that condition, so you should get a clear idea of how these act. (What we will seldom use is the fact that the indicated compositional condition on a pair of morphisms η, ε is equivalent to their being the unit and counit of an adjunction. Nevertheless, I recommend working Exercise 7.3:2 this once.)

The terms “unit” and “counit” will be easier to explain when we consider the concepts of *monad* and *comonad* in Chapter 10 (not yet written).

We can now characterize as right or left adjoints many of the universal constructions we are familiar with. The three diagrams below show the cases we used above to motivate the concept. In each of these, a pair of successive vertical arrows between two categories represents a pair of mutually adjoint functors, the right adjoint being shown on the right and the left adjoint on the left.



The middle diagram is interesting in that the forgetful functor there (in the notation of §3.11, $G \mapsto G_{\text{md}}$) has both a left and a right adjoint. In the first diagram, we can, as mentioned, replace **Group** with any category \mathbf{C} having free objects with respect to a concretization U . A still wider generalization is noted in the next exercise.

Exercise 7.3:3. If you did not do Exercise 7.2:5, prove that if \mathbf{C} is a category with small coproducts and $U: \mathbf{C} \rightarrow \mathbf{Set}$ a functor, then U has a left adjoint if and only if it is *representable*.

(Exercise 7.2:5 was essentially the case of this result where U was faithful, so that it could be called a “concretization” and its left adjoint a “free object” construction; but faithfulness played no part in the proof. In Chapter 9 we shall extend the concept of “representable functor” from set-valued functors to algebra-valued functors, and generalize the above result to the resulting much wider context.)

Exercise 7.3:4. Show that the left (or right) adjoint of a functor, if one exists, is unique up to canonical isomorphism, and conversely, that if A and B are isomorphic functors, then any functor which can be made a left (or right) adjoint of A can also be made a left (or right) adjoint of B .

Exercise 7.3:5. Show that if $A: \mathbf{C} \rightarrow \mathbf{D}$, $B: \mathbf{D} \rightarrow \mathbf{C}$ give an equivalence of categories, then B is both a right and a left adjoint to A .

The next exercise is a familiar example in disguise.

Exercise 7.3:6. Let \mathbf{C} be the category with $\text{Ob}(\mathbf{C}) = \text{Ob}(\mathbf{Group})$, but with morphisms defined so that for groups G and H , $\mathbf{C}(G, H) = \mathbf{Set}(|G|, |H|)$. Thus **Group** is a subcategory of \mathbf{C} , with the same object set but smaller morphism sets. Does the inclusion functor $\mathbf{Group} \rightarrow \mathbf{C}$ have a left and/or a right adjoint?

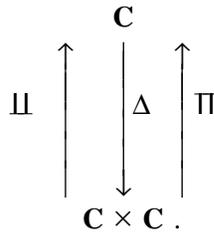
There are many other constructions whose universal properties translate into adjointness statements: The forgetful functor $\mathbf{Ring}^1 \rightarrow \mathbf{Monoid}$ that remembers only the multiplicative structure has as left adjoint the *monoid ring* construction. The forgetful functor $\mathbf{Ring}^1 \rightarrow \mathbf{Ab}$ that remembers only the additive structure has for left adjoint the *tensor ring* construction. (These two constructions were discussed briefly toward the end of §3.12.) The inclusion of the category of compact Hausdorff spaces in that of arbitrary topological spaces has for left adjoint the Stone-Ćech compactification functor (§3.17). The functor associating to every commutative ring its Boolean ring of idempotent elements has as left adjoint the construction asked for in Exercise 3.14:3(iv). The forgetful functors going from **Lattice** to \vee -**Semilattice** and \wedge -**Semilattice**, and from these in turn to **POSet**, have left adjoints which you were asked to construct in Exercise 5.1:8.

The student familiar with Lie algebras (§8.7 below) will note that the functor associating to an associative algebra A the Lie algebra A_{Lie} with the same underlying vector space as A , and with the commutator operation of A for Lie bracket, has for left adjoint the *universal enveloping algebra* construction. (The Poincaré-Birkhoff-Witt Theorem gives a normal form for this universal object; I hope to treat such normal form results in a much later chapter. Cf. [40])

Suppose \mathbf{C} is a category having *products* and *coproducts* of all pairs of objects. We know that each of these constructions will give a functor $\mathbf{C} \times \mathbf{C} \rightarrow \mathbf{C}$. Can these functors be characterized as adjoints of some functors $\mathbf{C} \rightarrow \mathbf{C} \times \mathbf{C}$? Similarly, can the *tensor product* functor $\mathbf{Ab} \times \mathbf{Ab} \rightarrow \mathbf{Ab}$ be characterized as an adjoint of some functor $\mathbf{Ab} \rightarrow \mathbf{Ab} \times \mathbf{Ab}$?

The universal property of the product functor $\mathbf{C} \times \mathbf{C} \rightarrow \mathbf{C}$ is a right universal one, so if it arises as an adjoint, it should be a right adjoint to some functor $A: \mathbf{C} \rightarrow \mathbf{C} \times \mathbf{C}$. No such functor was evident in our definition of products. However, we can search for such an A by posing the universal problem whose solution would be a *left* adjoint to the product functor: Given

$X \in \text{Ob}(\mathbf{C})$, does there exist $(Y, Z) \in \text{Ob}(\mathbf{C} \times \mathbf{C})$ with a universal example of a morphism $X \rightarrow Y \times Z$? Since a morphism $X \rightarrow Y \times Z$ corresponds to a morphism $X \rightarrow Y$ and a morphism $X \rightarrow Z$, this asks whether there exists a pair (Y, Z) of objects of \mathbf{C} universal for having a morphism from X to each member of this pair. In fact, the pair (X, X) is easily seen to have the desired universal property. This leads us to define the “diagonal functor” $\Delta: \mathbf{C} \rightarrow \mathbf{C} \times \mathbf{C}$ taking each object X to (X, X) , and each morphism f to (f, f) . It is now easy to check that the universal property of the direct product construction is that of a right adjoint to Δ . Moreover, similar reasoning shows that the universal property of the coproduct is that of a left adjoint of Δ . So in a category \mathbf{C} having both products and coproducts, we have the diagram of adjoint functors



We recall that if \mathbf{C} is \mathbf{Ab} , the constructions of pairwise products and coproducts (“direct products and direct sums”) coincide. So in that case we get a “cyclic” diagram of adjoints.

Exercise 7.3:7. Does the direct product construction on \mathbf{Set} have a *right* adjoint? Does the coproduct construction have a *left* adjoint?

The next exercise is one of my favorites:

Exercise 7.3:8. Recall that $\mathbf{2}$ denotes the category with two objects, 0 and 1 , and exactly one nonidentity morphism, $0 \rightarrow 1$, so that for any category \mathbf{C} , an object of $\mathbf{C}^{\mathbf{2}}$ corresponds to a choice of two objects $A_0, A_1 \in \text{Ob}(\mathbf{C})$ and a morphism $f: A_0 \rightarrow A_1$.

Let $p_0: \mathbf{Group}^{\mathbf{2}} \rightarrow \mathbf{Group}$ denote the functor taking each object (A_0, A_1, f) to its first component A_0 , and likewise every morphism $(a_0, a_1): (A_0, A_1, f) \rightarrow (B_0, B_1, g)$ of $\mathbf{Group}^{\mathbf{2}}$ to its first component a_0 .

Investigate whether p_0 has a left adjoint, and whether it has a right adjoint. If a left adjoint is found, investigate whether this in turn has a left adjoint (clearly it has a right adjoint – namely p_0); likewise if p_0 has a right adjoint, investigate whether this in turn has a right adjoint; and so on, as long as further adjoints on either side can be found.

Exercise 7.3:9. Let G be a group, and $G\text{-Set}$ the category of all G -sets.

You can probably think of one or more very easily described functors from \mathbf{Set} to $G\text{-Set}$, or vice versa. Choose one of them, and apply the idea of the preceding exercise; i.e., look for a left adjoint and/or a right adjoint, and for further adjoints of these, as long as you can find any.

When you are finished, does the chain of functors you have gotten contain all the “easily described functors” between these two categories that you were able to think of? If not, take one that was missed, and do the same with it.

Exercise 7.3:10. Translate the idea indicated in observation (a) following Exercise 3.8:1 into questions of the existence of adjoints to certain functors between categories $G_1\text{-Set}$ and $G_2\text{-Set}$, determine whether these adjoints do in fact exist, and if they do, describe them as well as you can.

Let us now consider the case of the tensor product construction, $\otimes: \mathbf{Ab} \times \mathbf{Ab} \rightarrow \mathbf{Ab}$. It is the solution to a left universal problem, and we can characterize this problem as arising, as described in Lemma 7.2.9, from the bifunctor $\text{Bil}: (\mathbf{Ab} \times \mathbf{Ab})^{\text{op}} \times \mathbf{Ab} \rightarrow \mathbf{Set}$, where for abelian groups A, B, C we let $\text{Bil}((A, B), C)$ denote the set of bilinear maps $(A, B) \rightarrow C$. From the preceding

examples, we might expect $\text{Bil}((A, B), C)$ to be expressible in the form $(\mathbf{Ab} \times \mathbf{Ab})((A, B), U(C))$ for some functor $U: \mathbf{Ab} \rightarrow \mathbf{Ab} \times \mathbf{Ab}$.

But, in fact, it cannot be so expressed; in other words, the tensor product construction $\mathbf{Ab} \times \mathbf{Ab} \rightarrow \mathbf{Ab}$, though it is a left universal construction, is *not* a left adjoint. The details (and a different sense in which the tensor product *is* a left adjoint functor construction) are something you can work out:

Exercise 7.3:11. (i) Show that the functor $\otimes: \mathbf{Ab} \times \mathbf{Ab} \rightarrow \mathbf{Ab}$ has no left or right adjoint.

(ii) On the other hand, show that for any fixed abelian group A , the functor $A \otimes -: \mathbf{Ab} \rightarrow \mathbf{Ab}$ is left adjoint to the functor $\text{Hom}(A, -): \mathbf{Ab} \rightarrow \mathbf{Ab}$. (I am writing $\text{Hom}(A, B)$ for the *abelian group* of homomorphisms from A to B , in contrast to $\mathbf{Ab}(A, B)$ the *set* of such homomorphisms – an arbitrary and ad hoc notational choice.)

(iii) Investigate whether the functor $A \otimes -: \mathbf{Ab} \rightarrow \mathbf{Ab}$ has a left adjoint, and whether $\text{Hom}(A, -): \mathbf{Ab} \rightarrow \mathbf{Ab}$ has a right adjoint. If such adjoints do not *always* exist, do they exist for *some* choices of A ?

If you are familiar enough with ring theory, generalize the above problems to modules over a fixed commutative ring k , or to bimodules over pairs of noncommutative rings.

Exercise 7.3:12. For a fixed set A , does the functor $\mathbf{Set} \rightarrow \mathbf{Set}$ given by $S \mapsto S \times A$ have a left or right adjoint?

A situation which is similar to that of the tensor product, in that the question of whether a construction is an adjoint depends on what we take as the variable, is considered in

Exercise 7.3:13. In this exercise “ring” will mean commutative ring with 1; recall that we denote the category of such rings $\mathbf{CommRing}^1$.

If R is a ring and X any set, $R[X]$ will denote the polynomial ring over R in an X -tuple of indeterminates.

(i) Show that for X a nonempty set, the functor $P_X: \mathbf{CommRing}^1 \rightarrow \mathbf{CommRing}^1$ taking each ring R to $R[X]$ has neither a right nor a left adjoint, and similarly that for R a ring, the functor $Q_R: \mathbf{Set} \rightarrow \mathbf{CommRing}^1$ taking each set X to $R[X]$ has neither a right nor a left adjoint.

(ii) On the other hand, show that the functor $\mathbf{CommRing}^1 \times \mathbf{Set} \rightarrow \mathbf{CommRing}^1$ taking a pair (R, X) to $R[X]$ is an adjoint (on the appropriate side) of an easily described functor.

(iii) For any ring R , let $\mathbf{CommRing}_R^1$ denote the category of commutative R -algebras (rings S given with homomorphisms $R \rightarrow S$), and R -algebra homomorphisms (ring homomorphisms making commuting triangles with R . In the notation of Exercise 6.8:26(ii), this is the comma category $(R \downarrow \mathbf{CommRing}^1)$.)

Similarly, for any set X , let $\mathbf{CommRing}_X^1$ denote the category of rings S given with set maps $X \rightarrow |S|$, and again having for morphisms the ring homomorphisms making commuting triangles. (This is the comma category $(X \downarrow U)$, where U is the underlying set functor of $\mathbf{CommRing}^1$. Note that to keep the symbols $\mathbf{CommRing}_R^1$ and $\mathbf{CommRing}_X^1$ unambiguous, we must remember to use distinct symbols for rings and sets.)

Show that for any R , the functor $\mathbf{Set} \rightarrow \mathbf{CommRing}_R^1$ taking X to $R[X]$ can be characterized as an adjoint, and that for any X , the functor $\mathbf{CommRing}^1 \rightarrow \mathbf{CommRing}_X^1$ taking R to $R[X]$ can also be characterized as an adjoint.

(iv) Investigate similar questions for the formal power series construction, $R[[X]]$; in particular, whether the analog of (i) is true.

Here is still another way to make the tensor product construction into an adjoint functor:

Exercise 7.3:14. (i) Let **Bil** be the category whose objects are all 4-tuples (A, B, β, C) where A, B, C are abelian groups, and $\beta: (A, B) \rightarrow C$ is a bilinear map, and with morphisms defined in the natural way. (Say what this natural way is!) Show that the forgetful functor $\mathbf{Bil} \rightarrow \mathbf{Ab} \times \mathbf{Ab}$, taking each such 4-tuple to its first two components, has a left adjoint, which is “essentially” the tensor-product construction.

(ii) Show that an analogous trick can be used to convert any isomorphism of bifunctors as in the Lemma 7.2.9 into an adjunction. (Between what categories?) Do the same for the situation of Lemma 7.2.11.

Exercise 7.3:15. Describe all pairs of adjoint functors at least one member of which is a *constant* functor, i.e., a functor taking all objects of its domain category to a single object X of its codomain category, and all morphisms of its domain category to id_X .

What happens when we compose two functors arising from adjunctions?

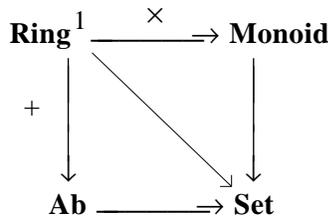
Note that the *abelianization* of the *free* group on a set X is a *free abelian* group on X . That is, when we compose these two functors, each of which is a left adjoint, we get another functor with that property. The general statement is simple, and is delightfully easy to prove.

Theorem 7.3.9. Suppose $\mathbf{E} \begin{matrix} \xrightarrow{U} \\ \xleftarrow{F} \end{matrix} \mathbf{D} \begin{matrix} \xrightarrow{V} \\ \xleftarrow{G} \end{matrix} \mathbf{C}$ are pairs of adjoint functors, with U and V the right adjoints, F and G the left adjoints. Then $\mathbf{E} \begin{matrix} \xrightarrow{VU} \\ \xleftarrow{FG} \end{matrix} \mathbf{C}$ are also adjoint, with VU the right adjoint and FG the left adjoint.

Proof. $\mathbf{C}(-, VU(-)) \cong \mathbf{D}(G(-), U(-)) \cong \mathbf{E}(FG(-), -)$. \square

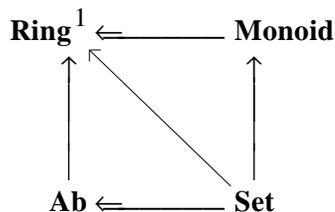
Exercise 7.3:16. Suppose U, V, F and G are as above, η and ε are the unit and counit of the adjoint pair U, F , and η' and ε' are the unit and counit of the adjoint pair V, G . Describe the unit and counit of the adjoint pair VU, FG .

For further examples of the above theorem, consider two ways we can factor the forgetful functor from \mathbf{Ring}^1 to \mathbf{Set} . We can first pass from a ring to its multiplicative monoid, then go to the underlying set thereof, or we can first pass from the ring to its additive group, and then to the underlying set:



Taking left adjoints, we get the two decompositions of the *free ring* construction noted in §3.12: as the free-monoid functor followed by the monoid-ring functor, and as the free abelian group functor

followed by the tensor algebra functor:



7.4. Number-theoretic interlude: the p -adic numbers, and related constructions. While you digest the concept of adjunction (fundamentally simple, yet daunting in its multiple facets), let us look at some constructions of a different sort, which we did not meet any examples of in the ‘‘Cook’s tour’’ of Chapter 3. In this section we will develop a particular case important in number theory; the general category-theoretic concept will be defined in the next section. A much broader generalization, which also embraces several constructions we *have* studied, will be developed in the section after that.

Suppose we are interested in solving the equation

$$(7.4.1) \quad x^2 = -1$$

in the integers, \mathbb{Z} . Of course, we know it has no solution in the real numbers, let alone the integers, but we will ignore that dreary fact for the moment.

We may observe that the above equation does have a solution in the finite ring \mathbb{Z}_5 , in fact, two solutions, 2 and 3. Up to sign, these are the same, so let us look for a solution of (7.4.1) in \mathbb{Z} satisfying

$$x \equiv 2 \pmod{5}.$$

An integer x which is $\equiv 2 \pmod{5}$ has the form $5y+2$, so we may rewrite (7.4.1) as

$$(5y+2)^2 = -1$$

and expand, to see what information we can learn about y . We get $25y^2+20y = -5$. Hence $20y \equiv -5 \pmod{25}$, and dividing by 5, we get $4y \equiv -1 \pmod{5}$. This has the unique solution

$$y \equiv 1 \pmod{5},$$

which, substituted back, determines x modulo 25:

$$x = 5y+2 \equiv 5 \cdot 1+2 = 7 \pmod{25}.$$

We continue in the same fashion: At the next stage, putting $x = 25z+7$ we have $(25z+7)^2 = -1$. You should verify that this implies

$$z \equiv 2 \pmod{5},$$

which leads to

$$x \equiv 57 \pmod{125}.$$

Can we go on indefinitely? This is answered in

Exercise 7.4:1. (i) Show that given $i > 0$, and $c \in \mathbb{Z}$ such that $c^2 \equiv -1 \pmod{5^i}$, there exists $c' \in \mathbb{Z}$ such that $c'^2 \equiv -1 \pmod{5^{i+1}}$, and $c' \equiv c \pmod{5^i}$.

(ii) Show that any integer is uniquely determined by its residues modulo $5, 5^2, 5^3, \dots, 5^i, \dots$.

Part (ii) of the above exercise shows that if there *were* an integer satisfying (7.4.1), the sequence of residues arising by repeated application of the step of part (i) would determine it. But now let us return to our senses, and remember that (7.4.1) has no real solution, and ask what, if anything, we *have* found.

Clearly, we have shown that there exists a sequence of residues, $x_1 \in |\mathbb{Z}_5|$, $x_2 \in |\mathbb{Z}_{5^2}|$, \dots , $x_i \in |\mathbb{Z}_{5^i}|$, \dots , each of which satisfies (7.4.1) in the ring in which it lives, and which are “consistent”, in the sense that each x_{i+1} is a “lifting” of x_i , under the series of natural ring homomorphisms

$$\dots \rightarrow \mathbb{Z}_{5^{i+1}} \rightarrow \mathbb{Z}_{5^i} \rightarrow \dots \rightarrow \mathbb{Z}_{5^2} \rightarrow \mathbb{Z}_5.$$

Let us name the i th homomorphism in the above sequence $f_i: \mathbb{Z}_{5^{i+1}} \rightarrow \mathbb{Z}_{5^i}$; thus, f_i takes the residue of any integer n modulo 5^{i+1} to the residue of n modulo 5^i . Now note that the set of all strings

$$(7.4.2) \quad (\dots, x_i, \dots, x_2, x_1) \text{ such that } x_i \in |\mathbb{Z}_{5^i}| \text{ and } f_i(x_{i+1}) = x_i \text{ (} i = 1, 2, \dots \text{)}$$

forms a ring under componentwise operations. What we have shown is that *this ring* contains a square root of -1 . Since, as we have noted, an integer n is determined by its residues modulo the powers of 5 , the ring \mathbb{Z} is *embedded* in this ring, though of course the square root, in this ring, of $-1 \in |\mathbb{Z}|$ does not lie in the embedded copy of the ring \mathbb{Z} . (“If the fool would persist in his folly, he would become wise,” William Blake [52].)

The ring of sequences (7.4.2) is called the *ring of 5-adic integers*. The corresponding object constructed for any prime p , using the system of maps

$$(7.4.3) \quad \dots \rightarrow \mathbb{Z}_p^{i+1} \rightarrow \mathbb{Z}_p^i \rightarrow \dots \rightarrow \mathbb{Z}_p^2 \rightarrow \mathbb{Z}_p,$$

is called the ring of *p-adic integers*. These rings are of fundamental importance in modern number theory, and come up in many other areas as well. The notation for them is not uniform; the symbol we will use here is $\hat{\mathbb{Z}}_{(p)}$. (The (p) in parenthesis denotes the *ideal* of the ring \mathbb{Z} generated by the element p . What is meant by putting it as a subscript of \mathbb{Z} and adding a hat will be seen a little later. Many number-theorists simply write \mathbb{Z}_p for the p -adic integers, denoting the field of p elements by $\mathbb{Z}/p\mathbb{Z}$ or \mathbb{F}_p ; cf. [24, p.272], [31, p.162, Example].)

The construction of \mathbb{Z}_p is in some ways analogous to the construction of the real numbers from the rationals. Real numbers are entities that can be approximated by rational numbers under the *distance metric*; p -adic integers are entities that can be approximated by integers via *congruences* modulo arbitrarily high powers of p . This analogy is made stronger in

Exercise 7.4:2. Let p be a fixed prime number. If n is any integer, let $v_p(n)$ denote the greatest integer e such that p^e divides n , or the symbol $+\infty$ if $n = 0$. The *p-adic metric* on \mathbb{Z} is defined by $d_p(m, n) = p^{-v_p(m-n)}$. Thus, it makes m and n “close” if they are congruent modulo a high power of p .

(i) Verify that d_p is a metric on \mathbb{Z} , and that the ring operations are continuous in this metric. Deduce that the *completion* of \mathbb{Z} with respect to this metric (the set of Cauchy sequences modulo the usual equivalence relation) can be made a ring containing \mathbb{Z} .

- (ii) Show that this completion is isomorphic to $\hat{\mathbb{Z}}_{(p)}$.
- (iii) Show that every element x of this completion has a unique “left-facing base- p expression” $x = \sum_{0 \leq i < \infty} c_i p^i$, where each $c_i \in \{0, 1, \dots, p-1\}$. In particular, show that any such infinite sum is convergent in the p -adic metric. What is the expression for -1 in this form?

We showed above that one could find a solution to the equation $x^2 = -1$ in $\hat{\mathbb{Z}}_{(5)}$. Let us note some simpler equations one can solve:

- Exercise 7.4.3.** (i) Show that every integer n not divisible by p is invertible in $\hat{\mathbb{Z}}_{(p)}$.
- (ii) Are the “base- p expressions” (in the sense of the preceding exercise) for the elements n^{-1} eventually periodic?

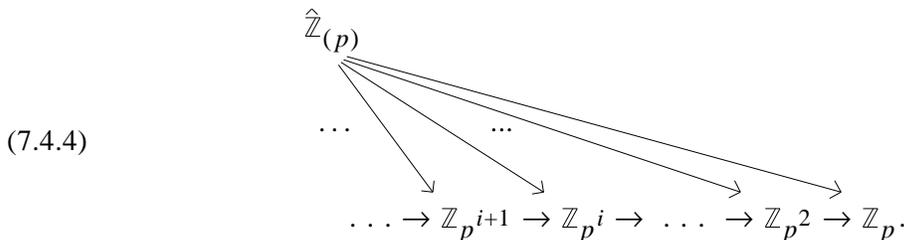
It follows from point (i) of the above exercise that we can embed into the p -adic integers not only \mathbb{Z} , but the subring of \mathbb{Q} consisting of all fractions with denominators not divisible by p . Now when one adjoins to a commutative ring R inverses of all elements not lying in some prime ideal P , the resulting ring (which, if R is an integral domain, is a subring of the field of fractions of R) is denoted R_P , so what we have embedded in the p -adic integers is the ring $\mathbb{Z}_{(p)}$. In $\mathbb{Z}_{(p)}$, every nonzero element is clearly an invertible element times a power of p , from which it follows that the nonzero ideals are precisely the ideals (p^i) . It is easy to verify that the factor-ring $\mathbb{Z}_{(p)}/(p^i)$ is isomorphic to \mathbb{Z}_p^i ; hence the system of finite rings and homomorphisms (7.4.3) can be described as consisting of *all* the proper factor-rings of $\mathbb{Z}_{(p)}$, together with the canonical maps among them. Hence the p -adic integers can be thought of as elements which can be approximated by members of $\mathbb{Z}_{(p)}$ modulo all *nonzero ideals* of that ring. Ring-theorists call the ring of such elements the *completion* of $\mathbb{Z}_{(p)}$ with respect to the system of its nonzero ideals, hence the symbol $\hat{\mathbb{Z}}_{(p)}$.

We will not go into a general study of what algebraic equations have solutions in the ring of p -adic integers. A result applicable to a large class of rings including the p -adics is *Hensel’s Lemma*; see [24, Theorem 8.5.6] or [22, §III.4.3] for the statement.

Let us characterize abstractly the relation between the diagram (7.4.3) and the ring of p -adic integers which we have constructed from it. Since a p -adic integer is by definition a sequence $(\dots, x_i, \dots, x_2, x_1)$ with each $x_i \in \mathbb{Z}_p^i$, the ring of p -adic integers has *projection* homomorphisms p_i onto each ring \mathbb{Z}_p^i . (Apologies for the double use of the letter “ p ”!) Since the components x_i of each element satisfy the compatibility conditions $f_i(x_{i+1}) = x_i$, these projection maps satisfy

$$f_i p_{i+1} = p_i,$$

i.e., they make a commuting diagram



I claim that $\hat{\mathbb{Z}}_{(p)}$ is right universal for these properties. Indeed, given any ring R with homomorphisms $r_i: R \rightarrow \mathbb{Z}_p^i$ which are “compatible”, i.e., satisfy $f_i r_{i+1} = r_i$, we see that for

any $a \in R$, the system of images $(\dots, r_i(a), \dots, r_2(a), r_1(a))$ defines an element $r(a) \in \hat{\mathbb{Z}}_{(p)}$. The resulting map $r: R \rightarrow \hat{\mathbb{Z}}_{(p)}$ will be a homomorphism such that $r_i = p_i r$ for each i , and will be uniquely determined by these equations.

This universal property is expressed by saying that $\hat{\mathbb{Z}}_{(p)}$ is the *inverse limit* of the system (7.4.3); one writes

$$\hat{\mathbb{Z}}_{(p)} = \varprojlim_i \mathbb{Z}_p^i.$$

We will give the formal definition of this concept in the next section.

A very similar example of an inverse limit is that of the system

$$(7.4.5) \quad \dots \rightarrow k[x]/(x^{i+1}) \rightarrow k[x]/(x^i) \rightarrow \dots \rightarrow k[x]/(x^2) \rightarrow k[x]/(x),$$

where $k[x]$ is the ring of polynomials in x over a field k , and (x^i) the ideal of all multiples of x^i . A member of $k[x]/(x^i)$ can be thought of as a polynomial in x specified modulo terms of degree $\geq i$. If we take a sequence of such partially specified polynomials, each extending the next, these determine a *formal power series* in x . So the inverse limit of the above system is the formal power series ring $k[[x]]$. This ring is well known as a place where one can solve various sorts of equations. Some of these results are instances of Hensel's Lemma, referred to above; others, such as the existence of formal-power-series solutions to differential equations, fall outside the scope of that lemma.

We constructed the p -adic integers using the canonical *surjections* $\mathbb{Z}_p^{i+1} \rightarrow \mathbb{Z}_p^i$. Now there are also canonical *embeddings* $\mathbb{Z}_p^i \rightarrow \mathbb{Z}_p^{i+1}$, sending the residue of n modulo p^i to the residue of pn modulo p^{i+1} . These respect addition but not multiplication, i.e., they are homomorphisms of abelian groups but not of rings. If we write out this system of groups and embeddings,

$$(7.4.6) \quad \mathbb{Z}_p \rightarrow \mathbb{Z}_p^2 \rightarrow \dots \rightarrow \mathbb{Z}_p^i \rightarrow \mathbb{Z}_p^{i+1} \rightarrow \dots$$

it is natural to think of each group as a subgroup of the next, and to try to take their "union" G . But they are not literally subgroups of one another, so we need to think further about what we want this G to be.

Clearly, for every element x of each group in the above system, we want there to be an element of G representing the image of x . Furthermore, if an element x of one of the above groups is mapped to an element y of another by some composite of the maps shown in (7.4.6), then these two elements should have the same image in G . Hence to get our G , let us form a disjoint union of the underlying sets of the given groups, and divide out by the equivalence relation that equates two elements if the image of one under a composite of the given maps is the other. It is straightforward to verify that this *is* an equivalence relation on the disjoint union, and that because the maps in the above diagram are group homomorphisms, the quotient by this relation inherits a group structure. If we call the maps in (7.4.6) $e_i: \mathbb{Z}_p^i \rightarrow \mathbb{Z}_p^{i+1}$, and the maps to the group we have constructed $q_i: \mathbb{Z}_p^i \rightarrow G$, then the identifications we have made have the effect that for each i ,

$$q_{i+1}e_i = q_i,$$

i.e., that the diagram

Definition 7.5.1. Let P be a partially ordered set.

P is said to be directed (or upward directed) if it is nonempty, and for any two elements x, y of P , there exists an element z majorizing both x and y .

P is said to be inversely directed (or downward directed) if it is nonempty and for any two elements x, y of P , there exists an element z which is \leq both x and y ; equivalently, if P^{op} is directed.

(The word “filtered” is sometimes used instead of “directed” in these definitions.)

(If you did Exercise 5.2:9, you will find that these conditions are certain of the “interpolation” properties of that exercise.)

We can now give the general definitions of direct and inverse limits. The formulations we give below assume that the morphisms of our given systems go in the “upward” direction with respect to the ordering on the indexing set. It happens that in our initial example of $\hat{\mathbb{Z}}_{(p)}$, the standard ordering on the positive integers is such that the morphisms went the *opposite* way; in our construction of \mathbb{Z}_p^∞ they went the “right” way; while in the case of germs of analytic functions, if one orders neighborhoods of z by inclusion, the morphisms again go the “wrong” way (namely, from the set of functions on a larger neighborhood to the set of functions on a smaller neighborhood). This can be corrected formally by using, when necessary, the opposite partial ordering on the index set. Informally, in discussing direct and inverse limits one often just specifies the system of *objects and maps*, and understands that to apply the formal definition, one should partially order the set indexing the objects so as to make maps among them go “upward”.

Definition 7.5.2. Let \mathbf{C} be a category, and suppose we are given a family of objects $X_i \in \text{Ob}(\mathbf{C})$ ($i \in I$), a partial ordering on the index set I , and a system (f_{ij}) of morphisms, $f_{ij} \in \mathbf{C}(X_i, X_j)$ ($i < j, i, j \in I$) such that for $i < j < k$, one has $f_{jk}f_{ij} = f_{ik}$. (In brief, suppose we are given a partially ordered set I , and a functor $F: I_{\text{cat}} \rightarrow \mathbf{C}$.)

If I is inversely directed, then $(X_i, f_{ij})_I$ is called an inversely directed system of objects and maps in \mathbf{C} . An inverse limit of this system means an object L given with morphisms $p_i: L \rightarrow X_i$ which are compatible, in the sense that for all $i < j \in I$, $p_j = f_{ij}p_i$, and which is universal for this property, in the sense that given any object W and morphisms $w_i: W \rightarrow X_i$ such that for all $i < j \in I$, $w_j = f_{ij}w_i$, there exists a unique morphism $w: W \rightarrow L$ such that $w_i = p_i w$ for all $i \in I$.

Likewise, if I is directed, then $(X_i, f_{ij})_I$ is called a directed system in \mathbf{C} ; and a direct limit of this system means an object L given with morphisms $q_i: X_i \rightarrow L$ such that for all $i < j \in I$, $q_i = q_j f_{ij}$, and which is universal in the sense that given any object Y and morphisms $y_i: X_i \rightarrow Y$ such that for all $i < j \in I$, $y_i = y_j f_{ij}$, there exists a unique morphism $y: L \rightarrow Y$ such that $y_i = y q_i$ for all $i \in I$.

(Synonyms sometimes used for inverse and direct limit are projective and inductive limit, respectively.)

Loosely, one often writes the inverse limit object $\varprojlim_i X_i$, and the direct limit object $\varinjlim_i X_i$. More precisely, letting F denote the functor $I_{\text{cat}} \rightarrow \mathbf{C}$ corresponding to the inversely directed or directed system (X_i, f_{ij}) , one writes these objects as $\varprojlim F$ and $\varinjlim F$ respectively.

The morphisms $p_j: \varprojlim_i X_i \rightarrow X_j$ are called the projection maps associated with this inverse limit, and the $q_j: X_j \rightarrow \varinjlim_i X_i$ the coprojection maps associated with the direct limit.

In the next-to-last paragraph of the above definition, by the “functor ... corresponding to the ...

system (X_i, f_{ij}) we understand the functor which takes on the value X_i at the object i , the value f_{ij} at the morphism (i, j) ($i < j$ in I), and the value id_{X_i} at the morphism (i, i) . In the case where the indexing partially ordered set consists of the positive or negative integers, note that the full system of morphisms is determined by the morphisms $f_{i, i+1}$ (which can be arbitrary), hence in such cases one generally specifies only these morphisms in describing the system.

Direct and inverse limits in **Set** may be constructed by the techniques we illustrated earlier:

Lemma 7.5.3. *Every inversely directed system (X_i, f_{ij}) of sets and set maps has an inverse limit, given by*

$$(7.5.4) \quad \varprojlim X_i = \{(x_i) \in \prod_I X_i \mid x_j = f_{ij}(x_i) \text{ for } i < j \in I\}, \text{ with} \\ \text{the } p_j \text{ given by projection maps, } \varprojlim X_i \subseteq \prod X_i \rightarrow X_j.$$

Likewise, every directed system (X_i, f_{ij}) of sets and set maps has a direct limit, gotten by forming the disjoint union of the X_i and dividing out by the equivalence relation under which $x \in X_i$ and $x' \in X_{i'}$ are equivalent if and only if they have the same image in some X_j ($j > i, i'$). \square

One may ask what the point is, in our definitions of direct and inverse limit, of requiring that the partially ordered set I be directed or inversely directed. One could set up the definitions without that restriction, and in most familiar categories one can, in fact, construct objects which satisfy the resulting condition. But the behavior of these constructions tends to be quite different from those we have discussed *unless* these directedness assumptions are made. (For instance, the explicit description in Lemma 7.5.3 of the equivalence relation in the construction of a direct limit of sets is no longer correct.) In any case, such a generalized definition would be subsumed by a still more general definition to be made in the next section! So the value of the definition in the form given above is that it singles out a situation in which the limit objects can be studied by certain techniques.

Exercise 7.5:1. (i) If $(X_i, f_{ij})_I$ is a directed system in a category **C**, and J a *cofinal* subset of I , show that $\varinjlim_J X_j \cong \varinjlim_I X_i$; precisely, that J will also be a directed partially ordered set, and that any object with the universal property of the direct limit of the given system can be made into a direct limit of the subsystem in a natural way, and vice versa.

(ii) Show that the isomorphism of (i) is an instance of a morphism (in one direction or the other) between $\varinjlim_J X_j$ and $\varinjlim_I X_i$ which can be defined whenever $J \subseteq I$ are both directed and both limits exist, whether or not J is cofinal.

(iii) State the result corresponding to (i) for inverse limits. (For this we need a term for a subset of a partially ordered set which has the property of being cofinal under the opposite ordering; let us use “downward cofinal”. When speaking of inverse systems, one sometimes just says “cofinal”, with the understanding that this is meant in the only sense that is relevant to such systems.)

(iv) What can you deduce from (i) and (iii) about direct limits over directed partially ordered sets having a greatest element, and inverse limits over inversely directed partially ordered sets having a least element?

(v) Given any directed partially ordered set I and any *noncofinal* directed subset J of I , show that there exists a directed system of sets, (X_i, f_{ij}) , indexed by I , such that $\varinjlim_I X_i \not\cong \varinjlim_J X_j$.

Exercise 7.5:2. (i) Suppose $(X_i, f_{ij})_I$ is a directed system in a category \mathbf{C} , and $f: J \rightarrow I$ a surjective isotone map, such that J , like I , is directed. Show that $\varinjlim_{j \in J} X_{f(j)} \cong \varinjlim_{i \in I} X_i$.

(ii) Deduce that if \mathbf{D} is a subcategory of \mathbf{C} , and L is an object of \mathbf{C} that can be written as a direct limit of objects and morphisms in \mathbf{D} , then L can be written as such a direct limit taken over a directed partially ordered set of the form $\mathbf{P}_{\text{fin}}(S)$, where S is a set, and $\mathbf{P}_{\text{fin}}(S)$ denotes the partially ordered set of all finite subsets of S , ordered by inclusion.

The next few exercises concern direct and inverse limits of *sets*. We shall see in the next chapter that direct and inverse limits of algebras have as their underlying sets the direct or inverse limits of the objects' underlying sets (assuming, in the case of *direct* limits, that the algebras have only finitary operations); hence the results obtained for sets in the exercises below will be applicable to algebras.

The construction of the p -adic integers was based on a system of *surjective* homomorphisms. The first point of the next exercise looks at inverse systems with the opposite property, and the second considers the dual situation for direct limits.

Exercise 7.5:3. (i) Let (S_i, f_{ij}) be an inversely directed system in \mathbf{Set} such that all the morphisms f_{ij} are one-to-one, and let us choose any element $i_0 \in I$. Show that $\varprojlim_i S_i$ can be identified with the intersection, in S_{i_0} , of the sets $f_{ii_0}(S_i)$ ($i < i_0$).

(ii) Let (S_i, f_{ij}) be a directed system in \mathbf{Set} such that all the morphisms f_{ij} are onto, and let us choose any element $i_0 \in I$. Show that $\varinjlim_i S_i$ can be identified with the quotient set of S_{i_0} by the union of the equivalence relations induced by the maps $f_{i_0 i}: S_{i_0} \rightarrow S_i$ ($i > i_0$).

Exercise 7.5:4. (i) Show that the inverse limit of any inverse system of *finite nonempty* sets is nonempty.

(Suggestions: Either build the description of an element of the inverse limit up “from below”, by looking at partial assignments satisfying appropriate extendibility conditions, and apply Zorn's Lemma to get a maximal such assignment, or else “narrow down on an element from above”, by looking at “subsystems” of the given inverse system, i.e., systems of nonempty subsets of the given sets carried into one another by the given mappings, and using Zorn's Lemma to get a minimal such subsystem. You might find it instructive to work out both of these proofs.)

(ii) Show that (i) can fail if the condition “finite” is removed, even for inverse limits over the totally ordered set of negative integers.

(iii) If you have some familiarity with general topology, see whether you can generalize statement (i) to a result on topological spaces, with “compact Hausdorff” replacing “finite”.

As an application of part (i) of the above exercise, suppose we are given a subdivision of the plane into regions, possibly infinitely many, and are studying the problem of coloring these regions with n colors so that no two adjacent regions are the same color. Let the set of all our regions be denoted R , the adjacency relation $A \subseteq R \times R$ (i.e., $(r_1, r_2) \in A$ if and only if r_1 and r_2 are adjacent regions), and the set of colors C . For any subset $S \subseteq R$, let X_S denote the set of all colorings of S (maps $S \rightarrow C$) under which no two adjacent regions have the same color; let us call these “permissible colorings of S ”. If $S \subseteq T$, then the restriction to S of a permissible coloring of T is a permissible coloring of S ; thus we have a restriction map $X_T \rightarrow X_S$. Now –

Exercise 7.5:5. (i) Show that in the above situation, the sets X_S , as S ranges over the *finite* subsets of R , form an inversely directed system, and that X_R may be identified with the inverse limit of this system in \mathbf{Set} .

(ii) Deduce using Exercise 7.5:4(i) that if each finite family $S \subseteq R$ can be colored, then the

whole picture R can be colored. (Note: the assumption that every finite family S can be colored does *not* say that *every* permissible coloring of a finite family S can be extended to a permissible coloring of every larger finite family T !)

- Exercise 7.5:6.** (i) Show that if (X_i, f_{ij}) is a directed system of sets, and each f_{ij} is one-to-one, then the canonical maps $q_j: X_j \rightarrow \varinjlim X_i$ are all one-to-one.
- (ii) Let (X_i, f_{ij}) be an inversely directed system of sets such that each f_{ij} is surjective. Show that if I is *countable*, then the canonical maps $p_j: \varprojlim X_i \rightarrow X_j$ are surjective. (Suggestion: First prove this in the case where I is the set of \leftarrow negative integers. Then show that any countable inversely directed partially ordered set either has a least element, or has a downward-cofinal subset order-isomorphic to the negative integers, and apply Exercise 7.5:1(iii).)
- (iii) Does this result remain true for uncountable I ? In particular, what if I is the opposite of an uncountable cardinal?

Exercise 7.5:7. Show that every group is a direct limit of finitely presented groups.

(This result is not specific to groups. We shall be able to extend it to more general algebras when we have developed the necessary language in the next chapter.)

The remaining exercises in this section develop some particular examples and applications of direct and inverse limits, including some further results concerning the p -adic integers. In these exercises you may assume the result which, as noted earlier, will be proved in the next chapter, that a direct or inverse limit of algebras whose operations are finitary can be constructed by forming the corresponding limit of underlying sets and giving this an induced algebra structure. None of these exercises, or the remarks connecting them, is needed for the subsequent sections of these notes.

One can sometimes achieve interesting constructions by taking direct limits of systems in which all objects are the same; this is illustrated in the next three exercises. The first shows a sophisticated way to get a familiar construction; in the next two, direct limits are used to get curious examples.

Exercise 7.5:8. Consider the directed system (X_i, f_{ij}) in **Ab**, where I is the set of positive integers, partially ordered by divisibility (i considered less than or equal to j if and only if i divides j), each object X_i is the additive group \mathbb{Z} , and for $j = ni$, $f_{ij}: \mathbb{Z} \rightarrow \mathbb{Z}$ is given by multiplication by n .

- (i) Show that $\varinjlim X_i$ may be identified with the additive group of the rational numbers.
- (ii) Show that if you perform the same construction starting with an arbitrary abelian group A in place of \mathbb{Z} , the result is a \mathbb{Q} -vector-space which can be characterized by a universal property relative to A .
- (iii) Can you describe the ring multiplication of \mathbb{Q} in terms of the description of its underlying abelian group in (i)?

Exercise 7.5:9. For this exercise, assume known the facts that every subgroup of a free group is free, and in particular, that in the free group on two generators x, y , the subgroup generated by the two commutators $x^{-1}y^{-1}xy$ and $x^{-2}y^{-1}x^2y$ is free on those two elements.

Let F denote the free group on x and y , and f the endomorphism of F taking x to $x^{-1}y^{-1}xy$ and y to $x^{-2}y^{-1}x^2y$. Let G denote the direct limit of the system $F \rightarrow F \rightarrow F \rightarrow \dots$, where all the arrows shown are the above morphism f .

Show that G is a nontrivial group such that every finitely generated subgroup of G is free, but that G is equal to its own commutator, $G = [G, G]$; i.e., that the abelianization of G is the trivial group. Deduce that though G is “locally free”, it is not free.

Exercise 7.5:10. Let k be a field. Let R denote the direct limit of the system of k -algebras $k[x] \rightarrow k[x] \rightarrow k[x] \rightarrow \dots$, where each arrow is the homomorphism sending x to x^2 . Show that R is an integral domain in which every finitely generated ideal is principal, but not every

ideal is finitely generated. (Thus, for each ideal, the minimum cardinality of a generating set is either 0, 1 or infinite.)

For the student familiar with the Galois theory of finite-dimensional field extensions, the next exercise shows how the Galois groups of infinite-dimensional extensions can be characterized in terms of the finite-dimensional case.

Exercise 7.5:11. Suppose E/K is a normal algebraic field extension, possibly of infinite degree. Let I be the set of subfields of E normal and of *finite* degree over K . If $F_2 \subseteq F_1$ in I , let $f_{F_1, F_2}: \text{Aut}_K F_1 \rightarrow \text{Aut}_K F_2$ denote the map which acts by *restricting* automorphisms of F_1 to the subfield F_2 .

- (i) Show that the definition of f_{F_1, F_2} makes sense, and gives a group homomorphism.
- (ii) Show that if we order I by reverse inclusion of fields, then the groups $\text{Aut}_K F$ ($F \in I$) and homomorphisms f_{F_1, F_2} ($F_1 \leq F_2$) form an inversely directed system of groups.
- (iii) Show that $\text{Aut}_K E$ is the inverse limit of this system in **Group**.
- (iv) Can you find a normal algebraic field extension whose automorphism group is isomorphic to the additive group of the p -adic integers?

Exercise 7.5:12. (i) (Open question.) Suppose a group G is the inverse limit of a system of finite groups. If G is a torsion group (i.e., if all elements of G are of finite order), must G have finite exponent (i.e., must there exist an integer n such that $x^n = e$ is an identity of G)?

Though the above question is very difficult, the next two parts are reasonable exercises, and may help render that question more tractable:

- (ii) Show that (i) is equivalent to the corresponding question in which we assume that G is the inverse limit of a system of finite groups indexed by the negative integers (under the natural ordering), with all connecting morphisms surjective.
- (iii) Translate (i) (possibly with the help of (ii)) into a question on finite groups which you could pose to a person not familiar with the concept of inverse limit. (The more natural-sounding, the better.)

Back to the p -adic integers, now. Part (i) of the next exercise seemed to me too simple to be true when I saw it described (in a footnote in a Ph.D. thesis) as “well-known”. But it is, in fact, not hard to verify

Exercise 7.5:13. (i) Show that $\mathbb{Z}[[x]]/(x-p) \cong \hat{\mathbb{Z}}_{(p)}$, where $\mathbb{Z}[[x]]$, we recall, denotes the ring of formal power series over \mathbb{Z} in one indeterminate x , and $(x-p)$ denotes the ideal of that ring generated by $x-p$.

- (ii) Examine other constructions of factor-rings of formal power series rings. For instance, can you describe $\mathbb{Z}[[x]]/(x-p^2)$? $\mathbb{Z}[[x]]/(x^2-p)$? $\mathbb{Z}[[x]]/(px^2-1)$? $R[[x]]/(f(x))$ for a general commutative ring R and a polynomial or power series $f(x)$, perhaps subject to some additional conditions? $R[[x, y]]/I$ for some fairly general class of ideals I ?

(If you consider $\mathbb{Z}[[x]]/(x-n)$ for n not a prime power, you might first look at Exercise 7.5.15 below.)

Exercise 7.5:14. (i) Show that the function v_p of Exercise 7.4:2 satisfies $v_p(xy) = v_p(x) + v_p(y)$ and $v_p(x+y) \geq \min(v_p(x), v_p(y))$ ($x, y \in \mathbb{Z}$).

- (ii) Deduce that $\hat{\mathbb{Z}}_{(p)}$ is an integral domain.
- (iii) Show that v_p can be extended in a unique manner to a $\mathbb{Z} \cup \{+\infty\}$ -valued function on \mathbb{Q} satisfying the properties noted in (i).
- (iv) Show that the completion of \mathbb{Q} with respect to the metric d_p induced by the above extended function v_p is the field of fractions of $\hat{\mathbb{Z}}_{(p)}$.

(v) Show that elements of this field have expansions $x = \sum_i c_i p^i$, where again $c_i \in \{0, 1, \dots, p-1\}$, and where i now ranges over all integer values (not necessarily positive), but subject to the condition that the set of i such that c_i is nonzero is bounded below.

This field is called the field of *p-adic rationals*, and denoted $\hat{\mathbb{Q}}_{(p)}$ (or \mathbb{Q}_p).

Is the “adic” construction limited to primes p , or can one construct, say, a ring of “10-adic integers”, $\hat{\mathbb{Z}}_{(10)}$? One encounters a trivial difficulty in that there are two ways of interpreting this symbol. But we shall see below that they lead to the same ring; so there is a well-defined object to which we can give this name. However, its properties will not be as nice as those of the p -adic integers for prime p .

Exercise 7.5:15. Let $\mathbb{Z}_{(10)}$ denote the ring of all rational numbers which can be written with denominators relatively prime to 10.

(i) Determine all nonzero ideals $I \subseteq \mathbb{Z}_{(10)}$ and the structures of the factor-rings $\mathbb{Z}_{(10)}/I$. Sketch the diagram of the inverse system of these factor-rings and the canonical maps among them.

(ii) Show that the inverse system $\dots \rightarrow \mathbb{Z}_{10^i} \rightarrow \dots \rightarrow \mathbb{Z}_{100} \rightarrow \mathbb{Z}_{10}$ constitutes a downward *cofinal* subsystem of the above inverse system.

Hence by Exercise 7.5:1 the inverse limits of these two systems are isomorphic, and we shall denote their common value $\hat{\mathbb{Z}}_{(10)}$. It is clear from the form of the second inverse system that elements of $\hat{\mathbb{Z}}_{(10)}$ can be described by “infinite decimal expressions to the left of the decimal point”.

(iii) Show that the relation $[2] \cdot [5] = [0]$ in \mathbb{Z}_{10} can be lifted to get a pair of nonzero elements which have product 0 in \mathbb{Z}_{100} , that these can be lifted to such elements in \mathbb{Z}_{1000} , and so on, and deduce that $\hat{\mathbb{Z}}_{(10)}$ is not an integral domain.

(iv) Prove, in fact, that $\hat{\mathbb{Z}}_{(10)} \cong \hat{\mathbb{Z}}_{(2)} \times \hat{\mathbb{Z}}_{(5)}$.

A construction often used in number theory is characterized in

Exercise 7.5:16. Show that the inverse limit of the system of all factor-rings of \mathbb{Z} by nonzero ideals is isomorphic to $\prod_p \hat{\mathbb{Z}}_{(p)}$, where the direct product is taken over all primes p . (This ring is denoted $\hat{\mathbb{Z}}$.)

A feature we have not yet mentioned, but which is important in the study of inverse limits, is topological structure. Recall that the inverse limit of a system of sets and set maps (X_i, f_{ij}) was constructed as a subset of $\prod X_i$. Let us now regard each X_i as a discrete topological space, and give $\prod X_i$ the product topology. In general, a product of discrete spaces is not discrete; however, a product of compact spaces is compact, so if our discrete spaces X_i are *finite*, their product will be compact. It is not hard to show that the subset $\varprojlim X_i \subseteq \prod X_i$ will be closed in the product topology, and hence, if the X_i are finite, will be compact in the induced topology.

Exercise 7.5:17. (i) Verify the assertion that $\varprojlim X_i \subseteq \prod X_i$ is always closed in the product topology, and is therefore compact if all X_i are *finite*.

(ii) Show that Exercise 7.5:4(i) (and hence Exercise 7.5:5(ii)) can be deduced using the compactness of $\varprojlim X_i$.

(iii) Show that the compact topology described above agrees in the case of $\hat{\mathbb{Z}}_{(p)}$ with the topology arising from the metric d_p of Exercise 7.4:2.

In fact, results like Exercise 7.5:5(ii), saying that a family of conditions can be satisfied simultaneously if all finite subfamilies of these conditions can be so satisfied, are called by logicians “compactness” results, because the proofs can generally be formulated in terms of the compactness of some topological space.

I can now say that the usual formulation of the open question of Exercise 7.5:12(i) is, “If a compact topological group is torsion, must it have finite exponent?” (Note that a topological group is by definition required to have a Hausdorff topology.) The equivalence of this with the question of that exercise follows from a deep result, that any compact group is an inverse limit of surjective maps of compact Lie groups (see [101, Theorem IV.4.6, p.175]), combined with the observation that if any of these Lie groups had positive dimension, we would get elements of infinite order. Thus, any compact torsion group is an inverse limit of 0-dimensional compact Lie groups, i.e., finite discrete groups, under the product topology.

An inverse limit of finite structures is called *profinite* (based on the synonym “projective limit” for “inverse limit”). I hope to eventually add to these notes a chapter treating profinite algebras (meanwhile, for some interesting results see [48]), and objects with related conditions, such as profinite-dimensionality. Let us look briefly at the latter condition in

Exercise 7.5:18. Let V be a vector space over a field k .

- (i) Show that the dual space V^* is the inverse limit, over all finite-dimensional subspaces $V_0 \subseteq V$, of the spaces V_0^* .
- (ii) Can you get the result of (i) as an instance of a general result describing *duals of direct limits* of vector spaces?
- (iii) If you did Exercise 5.5:5(ii)-(iii), show that the topology described there is that of the inverse limit of the finite-dimensional discrete spaces V_0^* referred to above. Show moreover that the only linear functionals $V^* \rightarrow k$ continuous in this topology are those induced by the elements of V .

The remainder of this section constitutes a digression for curiosity’s sake.

Ordinary real numbers expressed in base p have expansions going endlessly to the right, and finitely many steps to the left of the decimal point, while p -adic rationals (Exercise 7.5:14) have expansions going endlessly to the left, and finitely many steps to the right. Is it possible to define an arithmetic of elements with formal base- p expansions going endlessly in both directions?

Exercise 7.5:19. Let p be a prime. For every integer n , we have a subgroup $p^n\mathbb{Z} \subseteq \mathbb{R}$, hence we can form the quotient group $\mathbb{R}/p^n\mathbb{Z}$. Observe that these groups are each isomorphic to the circle group \mathbb{R}/\mathbb{Z} , and form an inverse system $\dots \rightarrow \mathbb{R}/p^2\mathbb{Z} \rightarrow \mathbb{R}/p\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z} \rightarrow \dots$, where the connecting maps take the residue of a real number modulo \mathbb{Z}/p^{i+1} to its residue modulo \mathbb{Z}/p^i . Let G be the *inverse limit* of this system of groups.

- (i) Show how to express elements of G as formal doubly infinite series $\sum_{i \in \mathbb{Z}} c_i p^i$, where $c_i \in \{0, 1, \dots, p-1\}$, ($i = \dots, -1, 0, 1, \dots$). Show that such a representation is unique except for the cases where for all sufficiently small i , c_i either becomes constant with value 0 or constant with value $p-1$.
- (ii) Show that $\hat{\mathbb{Q}}_{(p)}$ and \mathbb{R} both embed as dense subgroups of G .

Groups of the above sort appear in the theory of locally compact abelian groups, where they are called “solenoids”, from a term in electronics meaning “a hollow tightly wound coil of wire”. For students familiar with Pontryagin duality, the solenoid G constructed above will be seen to be the dual of the *discrete* additive group of $\mathbb{Z}[p^{-1}]$ (the ring of rational numbers of the form np^{-i}).

The above group G may also be obtained as a completion: For p a prime, let us define a function v_p on the real numbers, by letting $v_p(x)$ be the supremum of all integers n such that $x \in p^n\mathbb{Z}$. This will be $+\infty$ if $x = 0$, a nonnegative integer if x is a nonzero integer, a negative integer if x is a noninteger rational number of the form m/p^i , and $-\infty$ if none of these cases hold. (This does not agree with the definition of $v_p(x)$ we gave in Exercise 7.5:14 for rational x ,

though it does for x in the subring $\mathbb{Z}[p^{-1}]$.) Now for any two real numbers x, y , define $d_{p,|}(x, y) = \inf_{z \in \mathbb{R}} (p^{-v_p(x-z)} + |z-y|)$. Observe that although $p^{-v_p(x-z)}$ takes on the value $+\infty$ for most z , there exist values of z for which it is finite, so the infimum shown will be finite for all x and y .

- Exercise 7.5:20.** (i) Show that $d_{p,|}$ is a metric on the real line \mathbb{R} , and is bounded above.
 (ii) Show how to obtain from a doubly infinite series $\sum_{i \in \mathbb{Z}} c_i p^i$ a Cauchy sequence in \mathbb{R} under this metric, and show that all elements of the completion of \mathbb{R} in the metric $d_{p,|}$ can be represented by such series.
 (iii) Deduce that this completion is isomorphic to the solenoid G of the preceding exercise.

- Exercise 7.5:21.** (i) Show that the topology on G arising from the above metric agrees with that obtained by regarding G as an inverse limit of compact groups $\mathbb{R}/p^n\mathbb{Z}$. Deduce that the additive group operations of \mathbb{R} extend continuously to this completion.
 (ii) Let r be a real number, and $\bar{r}: \mathbb{R} \rightarrow \mathbb{R}$ the operation of multiplication by r . Show that \bar{r} is continuous in the metric $d_{p,|}$ if and only if $r \in \mathbb{Z}[p^{-1}]$. Deduce that multiplication as a map $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is not bicontinuous in this metric. Hence the ring structure on \mathbb{R} does not extend to the solenoid.
 (iii) Can addition of elements of the solenoid be performed by the same operations on digits that one uses to add ordinary real numbers in base p ? What goes wrong if we try to apply the ordinary procedure for *multiplying* numbers in base p ?
 (iv) If n is a positive integer not a power of p , show that the elements “ n^{-1} ” of \mathbb{R} and of $\hat{\mathbb{Q}}_{(p)}$ have distinct images under the embeddings of Exercise 7.5:19(ii). Deduce that the additive group of the solenoid has nonzero elements of finite order. Can you characterize such elements in terms of their “base p ” expansions?
 (v) Show that the solenoid described above is isomorphic to the group $\mathbf{Ab}(\mathbb{Z}[p^{-1}], \mathbb{R}/\mathbb{Z})$. (This is equivalent to the assertion of in the paragraph following Exercise 7.5:19(ii)).

7.6. Limits and colimits. Direct and inverse limits are similar in their universal properties to several other constructions we have seen. Let us recall these.

Given two objects X_1, X_2 of a category \mathbf{C} , a *product* of X_1 and X_2 in \mathbf{C} is an object P given with morphisms p_1 and p_2 into X_1 and X_2 , and universal for this property.

Given a pair of parallel morphisms $X_1 \rightrightarrows X_2$ in \mathbf{C} , an *equalizer* of this system is an object K given with a morphism k into X_1 having equal composites with those two morphisms, and again universal. To improve the parallelism with similar constructions, let us rename the morphism k as k_1 , and let $k_2: K \rightarrow X_2$ denote the common value of the composites of k_1 with the two morphisms $X_1 \rightrightarrows X_2$. Then we can describe K as having a morphism into *each* of X_1, X_2 , such that the composite of $k_1: K \rightarrow X_1$ with each of the two given morphisms $X_1 \rightarrow X_2$ is the morphism $k_2: K \rightarrow X_2$, and such that (K, k_1, k_2) is universal for these properties. We see that this is exactly like the universal property of an inverse limit, except that the indexing category $\cdot \rightrightarrows \cdot$ is not of the form $I_{\mathbf{cat}}$ for a partially ordered set I .

In the same way, a *pullback* of a pair of morphisms $f_1: X_1 \rightarrow X_3, f_2: X_2 \rightarrow X_3$ can be redefined as an object P given with morphisms p_1, p_2, p_3 into X_1, X_2, X_3 respectively, satisfying $f_1 p_1 = p_3$ and $f_2 p_2 = p_3$, and universal for this property.

Let us look at a case we haven't discussed yet. If G is a group and S a G -set, then the *fixed-point set* of the action of G on S means $\{x \in |S| \mid (\forall g \in |G|) gx = x\}$. If we denote the action of each $g \in |G|$ on S by $g_S: |S| \rightarrow |S|$, then the fixed-point set is universal among sets A with maps $i: A \rightarrow |S|$ such that for all $g \in |G|, i = g_S i$. Given an object X of any category

\mathbf{C} , and an action of a group G on X , we can look for an object with the same universal property, and, if it exists, call it the “fixed object” of the action.

We have seen constructions dual to those of product, equalizer and pullback. A construction dual to that of fixed object should take an object X of \mathbf{C} with an action of G on it to an object B of \mathbf{C} with a map $j: X \rightarrow B$ unchanged under composition on the right with the actions of elements of G , and universal for this property. Examples of this concept are examined in

Exercise 7.6:1. Let G be a group.

- (i) If X is a set on which G acts by permutations, and x an element of X , one defines the *orbit* of x under G to be the set $Gx = \{gx \mid g \in |G|\}$. Let B be the set of such orbits Gx , called the *orbit space* of X . Show that this set B , together with the map $X \rightarrow B$ taking x to Gx , has the universal property discussed above.
- (ii) Show that if G acts by automorphisms on (say) a ring R , then there is an object S in the category of rings with this same universal property, but that its underlying set will not in general be the orbit space of the action of G on the underlying set of R .
- (iii) If G acts by automorphisms on an object X of **POSet**, again show the existence of an object B with the above universal property. Show moreover that if G is finite, the underlying set of B will be the orbit space of the underlying set of X , and the universal map $X \rightarrow B$ will be *strictly* isotone; but that if G is infinite, neither statement need be true.
- (iv) Do the assertions of (iii) about the case where G is finite remain true if we replace **POSet** by **Lattice**?

As noted above, the universal properties we have been examining have statements formally identical with those of direct and inverse limits, except that the partially ordered set I of that definition is replaced by other categories \mathbf{D} (for example the two-object category $\cdot \rightrightarrows \cdot$ or the one-object category G_{cat}). As names for the more general concepts embracing such cases, one uses modified versions of the terms “direct limit” and “inverse limit”.

Definition 7.6.1. Let \mathbf{C} and \mathbf{D} be categories, and $F: \mathbf{D} \rightarrow \mathbf{C}$ a functor.

Then a limit of F , written $\varprojlim F$ or $\varprojlim_{\mathbf{D}} F(X)$, means an object $L \in \text{Ob}(\mathbf{C})$ given with morphisms $p(X): L \rightarrow F(X)$ for all $X \in \text{Ob}(\mathbf{D})$, such that for $f \in \mathbf{D}(X, Y)$ one has $p(Y) = F(f)p(X)$, and universal for this property, in the sense that given any object $M \in \text{Ob}(\mathbf{C})$ and family of morphisms $m(X): M \rightarrow F(X)$ ($X \in \text{Ob}(\mathbf{D})$) which similarly make commuting triangles with the morphisms $F(f)$, there exists a unique morphism $h: M \rightarrow L$ such that for all X , $m(X) = p(X)h$.

Likewise, a colimit of F , written $\varinjlim F$ or $\varinjlim_{\mathbf{D}} F(X)$, means an object $L \in \text{Ob}(\mathbf{C})$ given with morphisms $q(X): F(X) \rightarrow L$ for all $X \in \text{Ob}(\mathbf{D})$ such that for $f \in \mathbf{D}(X, Y)$ one has $q(Y) = q(X)F(f)$, and universal for this property, in the sense that given $M \in \text{Ob}(\mathbf{C})$ and morphisms $m(X): F(X) \rightarrow M$ ($X \in \text{Ob}(\mathbf{D})$) making commuting triangles with the morphisms $F(f)$, there exists a unique morphism $h: L \rightarrow M$ such that for all X , $m(X) = hq(X)$.

The morphisms $p(X)$ in the definition of a limit may be called the projection morphisms, and the $q(X)$ in the definition of colimit may be called the coprojection morphisms.

One says that a category \mathbf{C} “has small limits” if all functors from small categories \mathbf{D} into \mathbf{C} have limits; likewise \mathbf{C} “has small colimits” if all functors from small categories into \mathbf{C} have colimits.

Remarks on terminology. Since the above concepts generalize not only direct and inverse limits, but also a large number of other pairs of constructions, they might just as well have been given names suggestive of one of the other pairs. I think that the reason “limit” and “colimit”