Notes on Category Theory

Robert L. Knighten

November 9, 2007

Preface

There are many fine articles, notes, and books on category theory, so what is the excuse for publishing yet another tome on the subject. My initial excuse was altruistic, a student asked for help in learning the subject and none of the available sources was quite appropriate. But ultimately I recognized the personal and selfish desire to produce my own exposition of the subject. Despite that I have some hope that other students of the subject will find these notes useful.

Target Audience & Prerequisites

Category theory can sensibly be studied at many levels. Lawvere and Schanuel in their book *Conceptual Mathematics* [47] have provided an introduction to categories assuming very little background in mathematics, while Mac Lane's *Categories for the Working Mathematician* is an introduction to categories for those who already have a substantial knowledge of other parts of mathematics. These notes are targeted to a student with significant "mathematical sophistication" and a modest amount of specific knowledge. The sophistication is primarily an ease with the definition-theorem-proof style of mathematical exposition, being comfortable with an axiomatic approach, and finding particular pleasure in exploring unexpected connections even with unfamiliar parts of mathematics

Assumed Background: The critical specific knowledge assumed is a basic understanding of set theory. This includes such notions as subsets, unions and intersections of sets, ordered pairs, Cartesian products, relations, and functions as relations. An understanding of particular types of functions, particularly bijections, injections, surjections and the associated notions of direct and inverse images of subsets is also important. Other kinds of relations are important as well, particularly equivalence relations and order relations. The basic ideas regarding finite and infinite sets, cardinal and ordinal numbers and induction will also be used.

All of this material is outlined in Appendix A on informal axiomatic set theory, but this is not likely to be useful as a first exposure to set theory.

Although not strictly required some minimal understanding of elementary group theory or basic linear algebra will certainly make parts of the text much easier to understand.

There are many examples scattered through the text which require some knowledge of other and occasionally quite advanced parts of mathematics. In particular Appendix B (Catalog of Categories) contains a discussion of a large variety of specific categories. These typically assume some detailed knowledge of some parts of mathematics. None of these examples are required for understanding the body of the notes, but are included primarily for those readers who do have such knowledge and secondarily to encourage readers to explore other areas of mathematics.

Notation: The rigorous development of axiomatic set theory requires a very precise specification of the language and logic that is used. As part of that there is some concise notation that has become common in much of mathematics and which will be used throughout these notes. Occasionally, often in descriptions of sets, we will use various symbols from sentential logic particularly logical conjunction $\wedge$ for "and", logical disjunction $\vee$ for "or", implication $\Rightarrow$ for "implies" and logical equivalence $\iff$ for "if and only if". We also use $\forall$ and $\exists$ from existential logic with $\forall$ meaning "for all" and $\exists$ meaning "there exists".

Here is an example of the usage: For any sets A and B

$$\forall A\, \forall B,\ A + B = \{x : (x \in A \wedge x \notin B) \vee (x \notin A \wedge x \in B)\}$$

from which we conclude

$$\forall A\, \forall B,\ A + B = A \Rightarrow A \cap B = \emptyset$$

We have adopted two of Halmos' fine notational conventions here as well: the use of "iff" when precision demands "if and only if" while felicity asks for less; and the end (or absence) of a proof is marked with ∎.

Note on the Exercises

There are 170 exercises in these notes, freely interspersed in the text. A list of the exercises is included in the front matter, just after the list of definitions. Although the main purpose of the exercises is to develop your skill working with the concepts and techniques of category theory, the results in the exercises are also an integral part of our development. Solutions to all of the exercises are provided in Appendix C, and you should understand them. If you have any doubt about your own solution, you should read the solution in the Appendix before continuing on with the text. If you find an error in the text, in the solutions, or just have a better solution, please send your comments to the author at notes@knighten.org. They will be much appreciated.

Alternative Sources

There are many useful accounts of the material in these notes, and the study of category theory benefits from this variety of perspectives. In Appendix D are included brief reviews of the various books and notes, along with an indication of their contents.

Throughout these Notes specific references are included for alternative discussions of the material being treated, but no attempt has been made to provide attribution to original sources.

Contents

List of Definitions

List of Exercises

Introduction

Over the past century much of the progress in mathematics has been due to generalization and abstraction. Groups arose largely from the study of symmetries in various contexts, and group theory came when it was realized that there was a general abstraction that captured ideas that were being developed separately. Similarly linear algebra arose initially by recognizing the common ground under the development of linear equations, matrices, determinants and other notions. Then as linear algebra was codified it was recognized that it applied to very different situations and so, for example, its relevance to functional analysis was recognized and powerfully shaped the development of that field.

Topology, as the study of topological spaces, began around the middle of the 19th century. What we now call Algebraic Topology largely began with the work in Poincaré's series of papers called *Compléents à l'Analysis Situs* which he began publishing in 1895. Over the next 30 years Algebraic Topology developed rather slowly, but this was the same time that abstract algebra as was coming into being (as exemplified by van der Waerden's still well-named *Moderne Algebra* published in 1931.) About 1925 homology groups began to appear in all of their glory, and over the next twenty years much of the basics of modern algebraic topology appeared. But a basic insight was still missing – the recognition that in algebraic topology the important operations not only assign groups to topological spaces *but* also assign group homomorphisms to the continuous maps between the spaces. Indeed in order to axiomatize homology and cohomology theory the notion of equivalence between such operations was also needed. That was provided by Eilenberg and Mac Lane in their ground breaking paper *"General theory of natural equivalences"* [20], where the definitions of categories, functors and natural equivalences were first given. (A very good and extensive book on all of this and a very great deal more is Dieudonné's *A History of Algebraic and Differential Topology 1900-1960* [15].)

X

Chapter I

Mathematics in Categories

I.1 What is a Category?

Definition I.1: A **category** $\mathscr{C}$ has **objects** $A, B, C, \cdots, P, Q, \ldots$, and **morphisms** $f, g, h, i, \cdots, x, y, \cdots$. To each morphism is associated two objects, its **domain** and **codomain**. If f is a morphism with domain A and codomain B, this is indicated by $f : A \longrightarrow B$. Each object, A, has an associated **identity morphism** written $1_A : A \longrightarrow A$. Finally if $f : A \longrightarrow B$ and $g : B \longrightarrow C$, there is a **composition** $gf : A \longrightarrow C$, and these all satisfy the following relations:

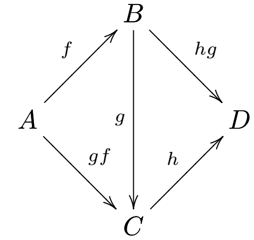
1. (**Associativity**) If $f : A \longrightarrow B$, $g : B \longrightarrow C$ and $h : C \longrightarrow D$, then $h(gf) = (hg)f : A \longrightarrow D$.

2. (**Identity Morphisms**) If $f : A \longrightarrow B$, then $f1_A = f = 1_B f$.

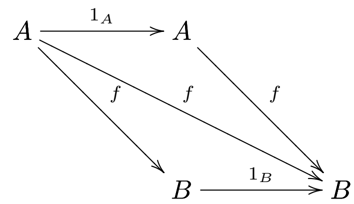
Sometimes gf is unclear and $g \circ f$ will be used instead. These are both read as "f composed with g" or as "g following f".

Relations such as these are indicated by saying diagrams of the following sort *commute*, meaning that any sequence of compositions of morphisms in the diagram that start and end at the same nodes in the diagram are equal.

1.



2.



For example in diagram (1.) commutativity of the left triangle says that gf is the g following f and the left triangle says that hg is h following g, which is certainly true but just the meaning of gf and hg. More interesting the top composite, $(hg)f$, is equal to the bottom composite, $h(gf)$, which is exactly Associativity. There is also h following g following f, and what Associativity allows us to say is that this is equal to both $(hg)f$ and $h(gf)$, i.e., the order matters, but parenthesis are unneeded.

In diagram (2.), commutativity of the top triangle says $f1_A$ is equal to f, while commutativity of the lower triangle say f if equal to 1_Bf. And these are exactly the requirements on the identify morphisms.

Note: Because category theory is applicable to so many diverse areas which have their own terminology, often well established before categories intruded, it is common even when discussing category theory to use a variety of terminology. For example while morphism is the most commonly used term, these elements of a category are also called "maps",and sometimes "arrows".Indeed we will occasionally use the word "map" as a synonym for morphism. Similarly what we called the domain of a morphism is sometimes called the "source",while "target"is an alternative for codomain.

Note: Throughout these notes script capital letters such as $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$, ..., $\mathscr{X}$, $\mathscr{Y}$, $\mathscr{Z}$ will be used without further comment to denote categories.

Examples of categories, familiar and unfamiliar, are readily at hand, but rather than listing them here Appendix B provides a Catalog of Categories where many examples are listed, together with detailed information about each of them. Each time a new concept or theorem appears it will be worthwhile to browse that Appendix for relevant examples.

There are two extreme examples of categories that are worthy of mention here. The first is the category **Set** of sets. **Set** has as objects all sets, and as morphisms all functions between sets. (For details see Section B.1.1 in the Catalog of Categories.) It is common in the development of the theory of sets to identify a function with its graph as a subset of the Cartesian product of its domain and its codomain. (See definition A.30 in Section A.8.) One result of this is that from the function (considered just as a set) it is possible to recover the domain of the function, but not in general its codomain. So when we say the morphisms are "all functions between sets" we actually consider a function as including a specification of its codomain. This same remark applies to many of the other "familiar" categories that we consider such as the categories of groups, Abelian groups, vector spaces, topological spaces, manifolds, etc.

By contrast suppose that (M, μ) is a monoid, that is a *set* M together an associative binary operation on M that has identity element (which we call 1.) Using M we define a category $\mathscr{M}$ which has only one object, which has M as its set of morphisms, all with the one object as both domain and codomain, with the identity of the monoid as the identity morphism on the unique object, and composition defined by the multiplication μ. We will usually refer to the category $\mathscr{M}$ by writing "consider the monoid M as a category with one object" without using any special name.

Conversely if $\mathscr{C}$ is any category with only one object, the morphisms with composition as multiplication form a monoid – save for one significant caveat, the definition of a monoid stipulates that M is a set.

For more information on monoids, look at the material in Section B.2.3 of the Catalog of Categories.

Our basic reference for topics in algebra is Mac Lane and Birkhoff's *Algebra* [55]. In particular for the definition and basic properties of a monoid see [55, I.11].

As an algebraic gadget the expected definition of a category is probably something like: A category, $\mathscr{C}$, consists of two sets Objects and Morphisms and functions:

$$\text{domain} : \text{Morphisms} \longrightarrow \text{Objects},$$
$$\text{codomain} : \text{Morphisms} \longrightarrow \text{Objects},$$
$$\text{id} : \text{Objects} \longrightarrow \text{Morphisms}$$

and a partial function

$$\text{composition} : \text{Morphisms} \times \text{Morphisms} \longrightarrow \text{Morphisms}$$

such that ....

The reason the definition we've given makes no mention of sets at all is because the most familiar categories, such as **Set**, do not have either a set of objects or a set of morphisms.

The connection between set theory and category theory is an odd one. Exactly how category theory should be explained in terms of set theory is still a topic of controversy, while at the same time most writers on either set theory or category theory give the subject scant attention. And that is what we will do as well. For those who are interested in more information about these issues, consult [50], [51] and [24] as a start.

There is also an active effort to use category theory as alternative foundation for set theory or even all of mathematics. Some references for these topics include Lawvere's "An Elementary Theory of the Category of Sets" [41, 44] and "The category of categories as a foundation for mathematics" [42], a couple of efforts to correct some errors, "Lawvere's basic theory of the category of categories" [7, 6]. A later discussion of axiomatizing the category of categories is McLarty's "Axiomatizing a Category of Categories" [57] , and a later discussion of axiomatizing the category of sets is Osius' "Categorical Set Theory: a Characterisation of the Category of Sets" [60].

With the arrival of the theory of topoi, that became the most important tool for discussing category theory and foundations. See Joyal and Moerdijk, [34], Mac Lane and Moerdijk, [56] and Lawvere and Rosebrugh [46] .

After that digression, we make the following definitions.

Definition I.2: A **small category** is one in which the collection of morphisms is a set. Note that as a consequence the collection of objects is a set as well.

Definition I.3: A **large category** is one in which the collection of morphisms is *not* a set.

Definition I.4: A **finite category** is one in which the collection of morphisms is a finite set. Note that as a consequence the collection of objects is finite as well.

So now going back to the connection between categories and monoids, we see that there is a natural correspondence between monoids and small categories with a single object. This is sufficiently strong that we will usually just write something like "consider the monoid as a category with one object."

I.1.1 Hom and Related Notation

Definition I.5: For any two objects in a category $\mathrm{Hom}(A, B)$ is the collection of all morphisms from A to B. If the morphisms are in the category $\mathscr{C}$ and we need to emphasize this, we will write $\mathscr{C}(A, B)$.

In a small category, $\mathrm{Hom}(A, B)$ is a set. In some large categories $\mathrm{Hom}(A, B)$ will not be a set, but in the familiar ones it is a set, so we make that a definition and a convention.

Definition I.6: A **locally small category** is one in which $\mathrm{Hom}(A, B)$ is a set for all objects A and B.

CONVENTION: Unless explicitly mentioned to the contrary, all categories considered in these notes are locally small.

With this convention every category with one object is a small category and so "is" a monoid.

"Hom" comes from "homomorphism", as does "morphism". Other notations that are sometimes used in place of Hom include "Map", "Mor", "Arr" and just parentheses – $[A, B]$ or (A, B).

Notation: It is often convenient to have anonymous functions, i.e., ones to which we give no special name. One common way of doing this is by writing something like $x \mapsto x^2$ in place of say $sq(x) = x^2$. If you are familiar with the elementary notation of the λ-calculus, this is equivalent to writing $\lambda x.x^2$. More generally if $\phi(x)$ is some formula involving x, writing $x \mapsto \phi(x)$, $\lambda x \,.\, \phi(x)$, and $f(x) = \phi(x)$ all have essentially the same effect except the expression $f(x) = \phi(x)$ requires providing the name f.

Now to use this notation, the composition of morphisms gives us a function:

$$\begin{aligned} \operatorname{Hom}(B,C) \times \operatorname{Hom}(A,B) &\longrightarrow \operatorname{Hom}(A,C) \\ (f,g) &\longmapsto fg \end{aligned}$$

And that in turn allows the application of the following simple but important observation. [**Warning:** For convenience we have $f : B \longrightarrow C$ and $g : A \longrightarrow B$ rather than the other way round.]

For any sets X, Y and Z, if we have a function $X \times Y \longrightarrow Z$ (which we will write (anonymously) as $(x,y) \mapsto xy$), then each element $x \in X$ defines a function $x_* : Y \longrightarrow Z$ by $x_*(y) = xy$. Similarly each element $y \in Y$ defines a function $y^* : X \longrightarrow Z$ by $y^*(x) = xy$. We can go even further: write Z^Y for the set of all functions from Y to Z, then from the function $X \times Y \longrightarrow Z$ we get a function $\lambda : X \longrightarrow Z^Y$ which is defined by $\lambda(x) = x_*$. There is a similar function $Y \longrightarrow Z^X$ which we will leave for the reader to actually name and describe.

As we noted above, the composition of morphisms gives us a function

$$\begin{aligned} \operatorname{Hom}(B,C) \times \operatorname{Hom}(A,B) &\longrightarrow \operatorname{Hom}(A,C) \\ (f,g) &\longmapsto fg \end{aligned}$$

to which we can apply this observation. So for each $f : B \longrightarrow C$, i.e., $f \in \operatorname{Hom}(B,C)$, we get a function $f_* : \operatorname{Hom}(A,B) \longrightarrow \operatorname{Hom}(A,C)$ defined by $f_*(g) = fg$. Equally for each $g : A \longrightarrow B$ we get $g^* : \operatorname{Hom}(B,C) \longrightarrow \operatorname{Hom}(A,C)$ with $g^*(f) = fg$.

Notice the subscript $*$ on f_* and the superscript $*$ on g^*. This use of subscript and superscript $*$ has historical roots dating at least as far back as tensor calculus and the use of subscripts and superscripts for covariant and contravariant tensors.

Notation: It occasionally happens that we want to discuss $\operatorname{Hom}(A,B)$ where we fix A but vary B. Usually the name B just confuses the issue, so in this situation we will often write $\operatorname{Hom}(A,\bullet)$ instead. This same thing applies for many things besides Hom, and the meaning should be clear in all cases. Other sources sometimes write $\operatorname{Hom}(A,\_)$ with the same meaning as $\operatorname{Hom}(A,\bullet)$ here.

In the language of computer science $\operatorname{Hom}(A,\bullet)$ is *polymorphic*, i.e., $\operatorname{Hom}(A,B)$ is defined for all objects B and is a set, while $\operatorname{Hom}(A,f)$ is defined for suitable morphisms and is a function. As we will see in detail in Chapter III (Functors), $\operatorname{Hom}(A,\bullet)$ is a *covariant* functor, while $\operatorname{Hom}(\bullet,C)$ is a *contravariant* functor, with the words covariant and contravariant having historical roots in tensor calculus.

As we'll see when we discuss functors in general, the more general notation for f_* is the rather cumbersome $\operatorname{Hom}(A,f)$ and for g^* it is $\operatorname{Hom}(g,C)$. We will commonly use f_* and g^*, but sometimes f_*^A if we need to keep track of the object A and sometimes the full $\mathrm{Hom}(A, f)$ or even $\mathscr{C}(A, f)$ if all the information is needed. Similarly we will usually write g^*, but sometimes g_C^*, $\mathrm{Hom}(g, C)$ or $\mathscr{C}(g, C)$.

We sometimes want to describe the functions f_*, and g^* without using the notation or even necessarily mentioning f or g specifically, so then we will write of "the induced functions" on the Hom sets.

For the record, here are the formal definitions.

Definition I.7: For each morphism $f : B \longrightarrow C$, the formula $f_*(g) = fg$ defines the induced function $f_* : \mathrm{Hom}(A, B) \longrightarrow \mathrm{Hom}(A, C)$.

Definition I.8: For each morphism $g : A \longrightarrow B$, the formula $g^*(f) = fg$ defines the induced function $g^* : \mathrm{Hom}(B, C) \longrightarrow \mathrm{Hom}(A, C)$.

When we come to discuss functors in Chapter III (Functors), these will be primary examples. In that context we want to note a few simple facts that we will use often: $(fg)_* = f_*g_*$ and $(fg)^* = g^*f^*$; $(1_B)_* = Hom(A, 1_B) = 1_{\mathrm{Hom}(A,B)}$ and $(1_A)^* = Hom(1_A, B) = 1_{\mathrm{Hom}(A,B)}$.

We also want to note another connection between categories and monoids. If $\mathscr{C}$ is any category and C is any object in $\mathscr{C}$, then $\mathscr{C}(C, C)$ is a monoid with composition as the binary operation and 1_C as the identity.

This is sufficiently important that we have a couple of definitions.

Definition I.9: In any category, a morphism in which the domain and codomain are equal is called an **endomorphism**.

Definition I.10: When C is an object of $\mathscr{C}$ as above, $\mathscr{C}(C, C)$ is the **monoid of endomorphisms**, or the **endomorphism monoid**, of C.

I.1.2 Subcategories

The most convenient source of additional categories is through the notion of a subcategory.

Definition I.11: A category $\mathscr{S}$ is a **subcategory** of category $\mathscr{C}$ provided:

1. Every object of $\mathscr{S}$ is an object of $\mathscr{C}$.
2. If $f \in \mathscr{S}(A, B)$, then $f \in \mathscr{C}(A, B)$.
3. If $f : A \longrightarrow B$ and $g : B \longrightarrow C$ in $\mathscr{S}$, then gf is also the composition of g following f in $\mathscr{C}$.
4. If 1_A is the identity morphism for A in $\mathscr{S}$, then 1_A is also the identity morphism for A in $\mathscr{C}$.

Examples of subcategories abound. A few of the examples that are discussed in the Catalog of Categories (Appendix B) are: The category of finite sets, **FiniteSet**, cf. Section B.1.2, as a subcategory of the category of sets; the category of Abelian groups, **Ab** cf. Section B.2.7, as a subcategory of the category of groups, **Group** (cf. Section B.2.5;) the category of lattices, **Lattice** cf. Section B.6.3, as a subcategory of the category of partially ordered sets, **Poset** cf. Section B.6.2; the category of compact Hausdorff spaces, **Comp** cf. Section B.9.4, as a subcategory of the category of topological spaces, **Top** cf. Section B.9.3; and the category of Hilbert spaces, **Hilbert** cf. Section B.13.4, as a subcategory of the category of Banach spaces, cf. Section B.13.1.

In addition the notion of subcategory generalizes the notion of submonoid, subgroup, etc. For if S is a submonoid of the monoid M, then, considered as categories, S is a subcategory of the category M, etc.

For the moment our interest in subcategories will be entirely as a way of specifying additional categories. To date all we have mentioned are just gotten specifying some collection of objects in the containing category. That is sufficiently common and important that we record it as:

Definition I.12: A subcategory $\mathscr{S}$ of $\mathscr{C}$ is **full** when for all objects A and B of $\mathscr{S}$ we have $\mathscr{S}(A,B) = \mathscr{C}(A,B)$.

Of course subcategories need not be full subcategories. As an example, starting with the category of sets we define a new category Iso(**Set**) with the same objects, i.e., sets, but as morphisms only the bijections. As every identity function is a bijections and the composition of bijections is a bijection, Iso(**Set**) is clearly a subcategory of **Set**. Not every function is a bijection, so Iso(**Set**) is *not* a full subcategory. The notation Iso(**Set**) may appear odd, but this is part of a general situation as we will see in Section I.2.1 (see page 11.)

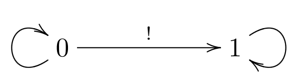
I.1.3 Recognizing Categories

To practice recognizing categories, we start with some very small and somewhat artificial examples.

The empty category **0** is the category with no objects and no morphisms. All the requirements in the definition of a category are vacuously satisfied. It is interesting and useful in much the same way the empty set is useful.

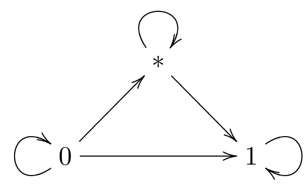
The one element category **1** is an essentially unique category with one object and one morphism which must be the identity morphism on the one object.

The category we call **2** or the **arrow category** is illustrated by the following diagram.



where all of the arrows represent distinct morphisms, and there are no other morphisms. For this to be a category, the two circular arrows must be the identity morphisms, and that completely determines composition, which is easily seen to be associative.

Next consider the following diagram:

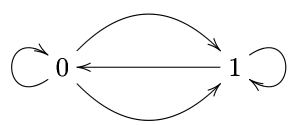


with each node representing an object and each arrow representing a distinct morphism.

Exercise I.1. Verify that there is a unique fashion in which this is a category with three objects and six morphisms. This category is named **3**.

Although these examples may seem strange they will actually recur in various application later in these notes.

As the last little example look at the following diagram:



Exercise I.2. If each arrow in the above diagram represents a distinct morphism, can this be a category with two objects and five morphisms?

In the previous section we saw the example of Iso(**Set**) as a somewhat unusual category. Let's look at a few more.

Define a subcategory $\mathscr{M}$ of **Set** to have as objects all sets, but the only morphisms are the injective functions, i.e., $f : X \longrightarrow Y$ where $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$. Clearly the identity function on any set is in $\mathscr{M}$, and the composition of any two injective functions is again injective, so we see $\mathscr{M}$ is indeed a subcategory, but not a full subcategory, of **Set**.

Exercise I.3. Define a subcategory $\mathscr{E}$ of **Set** to have as objects all sets, but the only morphisms are the surjective functions, i.e., $f : X \longrightarrow Y$ where for every $y \in Y$ there exists some $x \in X$ with $f(x) = y$. Verify that $\mathscr{E}$ is a subcategory, but not a full subcategory, of **Set**.

I.2 Special Morphisms

CONVENTION: In this section all the objects and morphisms are in the one fixed category $\mathscr{C}$ Unless explicitly mentioned to the contrary.

I.2.1 Isomorphisms

Besides the identity morphisms that exist in every category, there are other morphisms that are important and interesting parts of mathematics in a category. The first are isomorphisms.

Definition I.13: An **isomorphism**, $f : A \longrightarrow B$, is a morphism with an **inverse**, $f^{-1} : B \longrightarrow A$, satisfying $f^{-1}f = 1_A$ and $ff^{-1} = 1_B$.

Note that if f has an inverse, then it is unique, justifying the notation f^{-1}. To see the uniqueness note that if g and h are both inverses, then $g = 1_A g = hfg = h1_B = h$.

Notation: As is often done in algebra, we will use the symbol $\cong$ to indicate an isomorphism. So we will write $f : A \xrightarrow{\cong} B$ to indicate that f is an isomorphism, and we will write $A \cong B$ and say "A is isomorphic to B" when there is an isomorphism from A to B.

Just a bit earlier we introduced the endomorphism monoid of an object. Now in any monoid M with identity element 1, there is the submonoid of invertible elements: $G = \{m \in M \mid \exists m^{-1} \text{ such that } mm^{-1} = 1 = m^{-1}m\}$. As every element of G has an inverse, G is actually a group called, unoriginally, the **group of invertible elements** of M.

Applying this to the monoid of endomorphisms of any object in a category, we get the **group of automorphisms** or **automorphism group** of the object.

Here are the formal definitions.

Definition I.14: An **automorphism** is an endomorphism that is also an isomorphism.

Definition I.15: The **automorphism group** of an object C is the group of all automorphisms of C. This is usually denoted by $\mathrm{Aut}(C)$.

So $\mathrm{Aut}(C)$ is the group of invertible elements of the monoid of endomorphisms of C. In particular the identity morphism for any object C is an automorphism of C and is the identity in the group $\mathrm{Aut}(C)$.

Back in Section I.1.2 we met Iso(**Set**), a subcategory of the categories of sets that is not a full subcategory. As mentioned there this is an example of a general construction. For any category $\mathscr{C}$ there is the subcategory $\mathrm{Iso}(\mathscr{C})$ of $\mathscr{C}$ which has the same objects, but with the morphisms of $\mathrm{Iso}(\mathscr{C})$ being the isomorphisms of $\mathscr{C}$ (and that is the origin of the notation.) This is a full subcategory of $\mathscr{C}$ only when every morphism of $\mathscr{C}$ is an isomorphism (and then, of course, the two categories are the same.)

Categories such as this are actually of sufficient interest to deserve special study in their own right – for example see Higgins [32]. As a start we note the following definition.

Definition I.16: A **groupoid** is a category in which every morphism is an isomorphism.

Now just as a monoid is a category with just one object, a group is a groupoid with just one object. More information, including examples and applications, are in Appendix B.18 (Catalog of Categories).

A frequently useful technique in category theory is to connect properties of morphisms with properties of the functions they induce on Hom sets. Isomorphisms provide our first example in the next two exercises.

Exercise I.4. For any category $\mathscr{C}$, prove that if $f : A \longrightarrow B$ is an isomorphism, then for every object C the functions

$$f_*^C : \mathrm{Hom}(C, A) \longrightarrow \mathrm{Hom}(C, B) \quad \text{and} \quad f_C^* : \mathrm{Hom}(B, C) \longrightarrow \mathrm{Hom}(A, C)$$

are isomorphisms (i.e., bijections) as well.

Exercise I.5. Suppose $f : A \longrightarrow B$ is a morphism where for every object C the functions

$$f_*^C : \mathrm{Hom}(C, A) \longrightarrow \mathrm{Hom}(C, B) \quad \text{and} \quad f_C^* : \mathrm{Hom}(B, C) \longrightarrow \mathrm{Hom}(A, C)$$

are bijections. Prove that f is an isomorphism.

Exercise I.6. Suppose that $f : A \longrightarrow B$ is an isomorphism in $\mathscr{C}$. Define a function $\mathscr{C}(A, A) \longrightarrow \mathscr{C}(B, B)$ by $e \in \mathscr{C}(A, A) \mapsto fef^{-1} \in \mathscr{C}(B, B)$. Show that this function is a monoid homomorphism, and indeed an isomorphism.

In general an isomorphism of monoids is easily seen to also give rise to an isomorphism of the group of invertible elements in the monoids, so in particular whenever $f : A \longrightarrow B$ is an isomorphism the function defined above also gives an isomorphism between $\mathrm{Aut}(A)$ and $\mathrm{Aut}(B)$.

I.2.2 Sections and Retracts

The definition of an isomorphism has two parts which are really separable, and that leads to the notions of sections and retracts which we define here.

Definition I.17: For any morphism $f : A \longrightarrow B$, a **section** of f is a morphism $s : B \longrightarrow A$ such that $fs = 1_B$.

Definition I.18: For any morphism $f : A \longrightarrow B$, a **retract** (or **retract**) of f is a morphism $r : B \longrightarrow A$ such that $rf = 1_A$.

A section is also called a **right inverse**,while a retract is alternatively called a **left inverse**.

Exercise I.7. Consider any morphism f. Verify the following:

1. f has a section iff f_* always has a section.
2. f has a retract iff f^* always has a section.
3. f has a retract implies f_* always has a retract.
4. f has a section implies f^* always has a retract.

It is a trivial observation that s is a section for r iff r is a retract for s.

Note the usage of "**iff**"as an abbreviation for "if and only if". As noted in the preface (p. iv) we will use that from time to time throughout these notes.

It is not an accident that we have iff in the first two parts of this exercise, but only implication in the last two. Can you find examples where f_* has a retract, but f does not? What about where f^* has a retract but f does not have a section? [*Warning*: You won't find such examples in the category of sets.] We will return to this in exercises I.12 and I.18.

Exercise I.8. Clearly if f is an isomorphism, then f^{-1} is both a section and a retract for f. Show that if s is a section for f and r is a retract for f, then $r = s$ and so f is an isomorphism.

Exercise I.9. In the category of sets, give examples of functions that have a section but not a retract, and that have a retract but not a section. Also give examples of functions that have neither a section nor a retract.

I.2.3 Epimorphisms and Monomorphisms

In the category of Sets, the categorical notions of having a section or having a retract capture the essence of surjections and injections without any explicit mention of the elements of the sets. But these notions do not work nearly so well in other categories.

As one simple example consider the homomorphism $q : \mathbf{Z} \longrightarrow \mathbf{Z_2}$ (in the category of Abelian groups)defined by $q(n) = n \bmod 2$ ($\mathbf{Z}$ being the group of integers, and $\mathbf{Z_2}$ the group of integers modulo 2.) Now q is a surjection, but it certainly does not have a section - indeed the only homomorphism of $\mathbf{Z_2}$ to $\mathbf{Z}$ takes both elements of $\mathbf{Z_2}$ to 0. [For more information see the section on the category of groups see Section B.2.7 of the Catalog of Categories. The general reference for information about Abelian groups, and other general topics in abstract algebra, is Mac Lane and Birkhoff's *Algebra* [55].]

Fortunately there is a weaker property than having a right inverse, that captures the notion of a surjection in a categorical fashion for a great many categories. Moreover it turns out to be important and useful quite generally. Note that in many "algebraic" categories such as the categories of groups, Abelian groups, rings, etc, a surjective homomorphism is usually called an epimorphism, and that is the name used throughout these notes.

Definition I.19: In any category, a morphism $e : A \longrightarrow B$ is an **epimorphism** iff $fe = ge$ implies $f = g$.

The equation $fe = ge$ means that f and g are two morphisms with domain B and the same codomain. The codomain wasn't explicitly mentioned because it's name is irrelevant.

Note: As is common in mathematical writing, there is an implicit universal quantifier in the definition of an epimorphism. We have an epimorphism e only if for **all** morphisms f and g, $ef = eg$ implies e=f. Even one exception and e is **not** an epimorphism.

This is called a *cancellation law*, and we say that e is an epimorphism iff it can be canceled on the right or has right cancellation. That "right cancellation" is a weakened form of "having a right inverse" is the content of this next exercise.

Exercise I.10. Prove that if a morphism f has a section, then f is an epimorphism.

This result can be restated as "every retract is an epimorphism" which partly explains why retracts are also called *split epimorphisms* .

Notation: We often use the special arrow $\twoheadrightarrow$ to indicate an epimorphism.

When writing about epimorphisms, other words are sometimes used – we sometimes use the abbreviated form *epi*, and *epic*, particularly as an adjective ("the map f is epic".)

As with sections and retracts, let's also connect this with the Hom sets.

Exercise I.11. Show that $f : A \longrightarrow B$ is an epimorphism iff for every object C, the functions f^*_C are all injective.

Exercise I.12. Find an example where $f : A \longrightarrow B$ is an epimorphism, but for some object C the function f^C_* is not surjective. [*Warning:* You need to use something other than the category of sets.]

The next three exercises ask you to work out the meaning of epimorphisms in a few special cases where the categories are monoids.

There are two familiar binary operation on the set of natural numbers, addition and multiplication. Each of them gives us a monoid. The first is called the *additive monoid of natural numbers*, while the second is called the *multiplicative monoid of natural numbers*.

Exercise I.13. Consider $\mathbb{N}$, the monoid of natural numbers with the binary operation of addition as a category with one object. Show that every morphism in $\mathbb{N}$is an epimorphism.

The next exercise asks you to consider epimorphisms in a more complicated monoid.

Let $A = \{a, b\}$ be a set with two elements. Define A^* to be the set of all finite sequences from A, i.e.,

$$A^* = \{(), (a), (b), (a, a), (a, b), (b, a), (b, b), (a, a, a), \cdots\}.$$

Define a binary operation on A^* to be concatenation, i.e., if $s = (s_1, s_2, \cdots, s_m)$ and $t = (t_1, t_2, \cdots, t_n)$ are in A^*, then $st = (s_1, s_2, \cdots, s_m, t_1, t_2, \cdots, t_n)$. This makes A^* into a monoid with (), the empty sequence as the identity. [This is the **free monoid** on A and is one example of many "free" structures that will be discussed in these notes. More information on free monoids can be found in Sections B.19.3 and III.2.12, and in Bourbaki's *Algebra (Part I)* [10, I, §7].]

We will usually write just a or b rather than (a) and (b). And with that notation $(s_1 s_2 \cdots s_m) = s_1 s_2 \cdots s_m$ and we can think of that as either the product of the elements $s_1, s_2, \ldots$, or as composition of the morphisms $s_1, s_2, \ldots$.

[In mathematical logic and theoretical computer science the monoid A^* is called the *Kleene closure* of A after the logician Stephen Kleene who used it in his study of regular expressions. The construction actually make sense for sets with any number of elements, and this is the more general context of the Kleene closure.]

Exercise I.14. Show that every morphism in A^* is an epimorphism.

And finally we look at a monoid where not every morphism is an epimorphism.

Let R be the compatible equivalence relation on A^* generated by $\{a^2, ab\}$, and consider the quotient monoid $B = A^*/R$. This means that in B we have $a^2 = ab$, $a^2 b = ab^2 = a^3$, and all other relations that follow from $a^2 = ab$. By contrast ba is not equal to anything else.

[More information about generated equivalence relations and the associated quotients can be found in the Appendix on Set Theory (see Section A.7), and in the material on the categories **Monoid** and **Semigroup** in the Catalog of Categories (see Sections B.2.3 and B.2.2.) There is also a detailed treatment in Bourbaki [10, I, §1.6].]

In B every element has a unique canonical form that is one of a^m, b^n, or $b^n a^m$. There is nothing more complicated because any element of the form

$a^m b^n = a^{m+n}$ and that in turn will simplify any expression that has an ab in it somewhere.

> **Exercise I.15.** Using the above show that b is an epimorphism in B considered as a monoid with one object, but a is not an epimorphism.

Now the term epimorphism already has a meaning in a number of familiar categories, so we would certainly like to know that the new definition we have just given is actually the same as the usual one. In just a moment we'll see this is indeed true in some important examples, but there are other cases where it is not at all easy to verify this and many more where it is false. More information on this topic for the specific categories discussed can be found in the Catalog of Categories in Appendix B.

In the category **Ab** of Abelian groups and group homomorphisms (see Section B.2.7,) it is easy to see that every epimorphism in the usual sense, i.e., a surjective group homomorphism, is an epimorphism in the the sense of definition I.19. Just note that if e is a surjective homomorphism and $fe = ge$, then for each $b \in B$ there is some $a \in A$ so that $b = e(a)$. But then $f(b) = f(e(a)) = g(e(a)) = g(b)$, and so $f = g$.

The converse uses some more information about Abelian groups. Recall that for any homomorphism $f : A \longrightarrow B$ of Abelian groups the image of f, $\mathrm{Im}(f) = \{f(a) | a \in A\}$ is a (normal) subgroup and there is a quotient group $B/\mathrm{Im}(f)$ and a quotient map $q : B \longrightarrow B/\mathrm{Im}(f)$.

Now suppose that $e : A \longrightarrow B$ has the cancellation property that is the definition of an epimorphism in a category, and consider

$$A \xrightarrow{e} B \underset{q}{\overset{0}{\rightrightarrows}} B/\mathrm{Im}(e)$$

where q is the quotient map, and 0 maps every element to the zero element. Surely $0e = 0$, but also $qe = 0$. So $q = 0$! But that says $B/\mathrm{Im}(e)$ is the zero group, i.e., that $B = \mathrm{Im}(e)$ and e is surjective.

Very much the same thing is true in the category **Vect** of vector spaces and linear transformations (see Section B.4.3.) Every surjective linear transformation is an epimorphism as defined above with exactly the same argument as above.

The proof of the converse is much the same as well. Every linear transformation $f : A \longrightarrow B$ has an image $\mathrm{Im}(f) = \{f(a) | a \in A\}$ which is a subspace of B, and there is a quotient space $B/\mathrm{Im}(f)$ and a quotient map $q : B \longrightarrow B/\mathrm{Im}(f)$. Just as with Abelian groups, if $e : A \longrightarrow B$ has right cancellation, then consider

$$A \xrightarrow{e} B \underset{q}{\overset{0}{\rightrightarrows}} B/\mathrm{Im}(e)$$

where q is the quotient map, and 0 maps every element to the zero element. Surely $0e = 0$, but also $qe = 0$. So $q = 0$! But that says $B/\mathrm{Im}(e)$ is the zero vector space which only happens when $B = \mathrm{Im}(e)$, so e is surjective.

In Chapter XIII we'll see that there are a large variety of additional categories where very similar arguments apply.

By contrast, proving that a homomorphism in the category of groups (i.e., not necessarily Abelian) which has the epimorphism cancellation property is actually surjective is not nearly so easy. See section B.2.5 on page 227 for more detail.

Just as "having a right inverse" can be weakened to "right cancellation", "having a left inverse" can be weakened to "left cancellation".

Definition I.20: In any category, a morphism $m : A \longrightarrow B$ is a **monomorphism** iff $mf = mg$ implies $f = g$.

This is another "cancellation law",so we say that m is a monomorphism iff it can be canceled on the left or has left cancellation.

That "left cancellation" is a weakened form of "having a left inverse" is the content of this next exercise.

Exercise I.16. Prove that if a morphism f has a retract, then f is a monomorphism.

This result can be restated as "every section is a monomorphism" which partly explains why sections are also called *split monomorphisms* .

Just as with epimorphisms, other words are sometimes used – we sometimes say a morphism is *monic* or *mono* or is a monic.

Notation: We use the special arrow $\rightarrowtail$ to indicate a monomorphism.

Again let's connect this with the Hom sets.

Exercise I.17. Show that $f : A \longrightarrow B$ is a monomorphism iff for all objects C the functions f_*^C are always injective.

Exercise I.18. Find an example where $f : A \longrightarrow B$ is an monomorphism, but for some object C the function f^*_C is not surjective. [*Warning:* You need to use something other than the category of sets.]

Just as with epimorphism, monomorphism already has a meaning in a number of familiar categories and we would certainly like to know that the new definition we have just given is actually the same as the usual one. In just a moment we'll see this is indeed true in some important examples, but there are other cases where it is not at all easy to verify this. More information on this topic for the specific categories discussed can be found in the Catalog of Categories in Appendix B.

In the category **Ab** of Abelian groups and group homomorphism, it is easy to see that every monomorphism in the usual sense, i.e., a group homomorphism that is injective or one-to-one, is a monomorphism in the the sense of definition I.20. Just note that if e is an injective homomorphism and $mf = mg$, then for each $a \in A$, $mf(a) = mg(a)$ and so $f(a) = g(a)$, i.e., $f = g$.

The converse uses some more information about Abelian groups. Recall that for each homomorphism $f : A \longrightarrow B$ of Abelian groups we have the kernel of f, $\mathrm{Ker}(f) = \{a \in A | f(a) = 0\}$, a subgroup of A with inclusion map $i : \mathrm{Ker}(f) \longrightarrow A$.

Now suppose that $m : A \rightarrowtail B$ has the left cancellation property that is the definition of an monomorphism in a category, and consider

$$\mathrm{Ker}(m) \underset{i}{\overset{0}{\rightrightarrows}} A \xrightarrow{f} B$$

where i is the inclusion map, and 0 maps every element to the zero element. Now $m0 = 0 = mi$, so $0 = i$! But that says $\mathrm{Ker}(m) = \{0\}$ and so m is injective.

This argument works equally well in the category of groups as in the category of Abelian groups.

Essentially the same argument works in the category **Vect** of vector spaces and linear transformation as well. The first part of the argument is exactly the same, while there is just a small change of terminology in the second part. Associated to a linear transformation $f : A \longrightarrow B$ is its null space, $\mathrm{N}(f) = \{a \in A | f(a) = 0\}$, with inclusion map $i : \mathrm{N}(f) \longrightarrow A$. The rest of the argument is the same: if $m : A \rightarrowtail B$ has left cancellation then consider

$$\mathrm{Ker}(m) \underset{i}{\overset{0}{\rightrightarrows}} A \xrightarrow{f} B$$

where i is the inclusion map, and 0 maps every element to the zero element. Now $m0 = 0 = mi$, so $0 = i$! But that says $\mathrm{N}(m) = \{0\}$ and so m is injective.

Before continuing we record as exercises some simple observations that will be useful as we go along.

Exercise I.19. Show that in any category the composition of retracts is a retract.

Exercise I.20. Show that in any category if gf is a retract, then g is a retract.

Exercise I.21. Show that in any category the composition of epimorphisms is an epimorphism.

Exercise I.22. Show that in any category if gf is an epimorphism, then g is an epimorphism.

Exercise I.23. Show that in any category if s is an epimorphism and a section, then s is an isomorphism.

Exercise I.24. Show that in any category the composition of sections is a section.

Exercise I.25. Show that in any category if gf is a section, then f is a section.

Exercise I.26. Show that in any category the composition of monomorphisms is a monomorphism.

Exercise I.27. Show that in any category if gf is an monomorphism, then f is a monomorphism.

Exercise I.28. Show that in any category if r is a monomorphism and a retract, then r is an isomorphism.

In many, though far from all, of the familiar categories discussed in the Catalog of Categories (Appendix B) the epimorphisms are the surjective functions, and the monomorphisms are the injective functions. As a result it is often the case that morphisms that are both monic and epic are isomorphisms. Often, but not always!

Exercise I.29. Give an example of a morphism in a category of "sets with structure" that is an epimorphism, but not surjective. (Hint: Look in the category of monoids or in the category of topological spaces.)

For reasons explained in it is considerably harder to give examples of monomorphisms which are not injective. The most common example is in the category of divisible Abelian groups where the quotient homomorphism $\mathbb{Q} \longrightarrow \mathbb{Q}/\mathbb{Z}$ is a monomorphism but clearly not an injection.

Exercise I.30. Give an example of a category and a morphism in the category which is both a monomorphism and an epimorphism, but not an isomorphism.

This is sufficiently common that it deserves a name.

Definition I.21: A morphism that is both a monomorphism and an epimorphism is called a **bimorphism**.

Of course every isomorphism is a bimorphism, while the last exercises exhibits a bimorphism that is not an isomorphism. The interesting question is "When is a bimorphism an isomorphism?". There are categories where this is always true, and others where the bimorphisms that are not isomorphisms are of particular interest. This is a minor theme that will recur from time to time.

Again the situation where every bimorphism is an isomorphism has a name.

Definition I.22: A category is which every bimorphism is an isomorphism is called a **balanced category**.

I.2.4 Subobjects and Quotient Objects

Although "surjective" and "injective" are defined in terms of elements and so don't quite fit our "arrows only" motto, we will investigate the relation between these pairs of concepts (and sections and retracts as well) repeatedly. In particular the relation between these concepts not just in a single category, but between related categories.

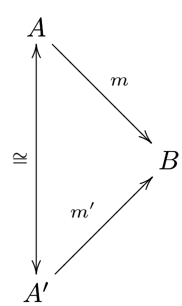
As a start notice that in most familiar categories we have notions of subobjects and quotient objects. These are usually defined in terms of elements, but we'll do it with morphisms.

Definition I.23: A **subobject** of an object B is a monomorphism $A \rightarrowtail B$.

Before continuing on, let's look at how meaningful this is in the category of sets. We know that a monomorphism $A \rightarrowtail B$ in **Set** is an injective function, so certainly if A is a subset of B, then the inclusion function is a monomorphism and so we have a subobject of B. But just because $A \rightarrowtail B$ is an injection, this does *not* mean that A is a subset of B. But any injection does factor as

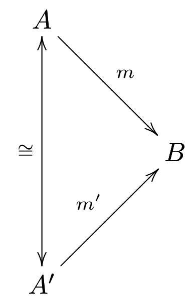
$$A \xrightarrow{\cong} S \hookrightarrow B$$

with S the image of the monomorphism being an actual subset of B and $S \hookrightarrow B$ the inclusion function. Moreover two injections $m : A \rightarrowtail B$ and $m' : A' \rightarrowtail B$ have the same image in B exactly when there is a bijection between A and A' so that



commutes. The same is true in many categories of "structured sets". Based on this we make the following definition.

Definition I.24: Two subobjects $m : A \rightarrowtail B$ and $m' : A' \rightarrowtail B$ of B are **equivalent** when there is an isomorphism between A and A' with



commuting.

So in the category of sets a subset determines a subobject and a subobject determines a subset with two different subobjects determining the same subset iff they are equivalent. In many other familiar categories such as the categories of groups, rings, vector spaces, etc., there is the same correspondence between equivalent subobjects and the familiar subgroups, subrings, etc. But, as just one example, subobjects in the category **Top** does not well correspond to subspaces of topological spaces. For the details see the various entries in the Catalog of Categories (Appendix B).

It is tempting to define a subobject of an object to be an equivalence class of equivalent monomorphisms into the object, and it is common to do so. But the difficulty with this is that such an equivalence class need not exist! In Zermelo-Fraenkel set theory as discussed in Appendix A, the "equivalence class" of all injections into a one element set, say $\{0\}$, cannot exist as a set – it is too large. Remember that for every set X there is the one element set $\{X\}$ and and a unique injection $\{X\} \rightarrowtail \{0\}$. Any two of these are equivalent, so if this equivalence class were to exist as a set it would be equinumerous with the (non-existent) set of all sets!

This issue will be dealt in a much more satisfying fashion when we discuss subobject classifiers (see definition IX.1 and the following surrounding material).

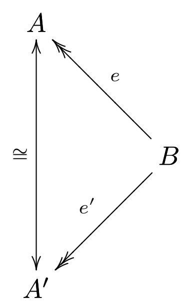
There is the dual notion of a quotient object.

Definition I.25: A **quotient object** of an object B is an epimorphism

$$B \twoheadrightarrow A.$$

Just as with subobjects we define equivalent quotient objects.

Definition I.26: Two quotient objects $e : B \twoheadrightarrow A$ and $e' : B \twoheadrightarrow A'$ are *equivalent* where there is an isomorphism between A and A' with



commuting.

Again let's look at this in the category of sets. First what *is* a quotient set? As described in the Appendix on Set Theory (see A.7 and also Mac Lane and Birkhoff [55, Sec. 1.9]) a quotient set of B is the set of equivalence classes of some equivalence relation on B. To connect that to the notion of quotient object, any function $f : B \longrightarrow A$ defines an equivalence relation on B via $b \equiv_f b'$ iff $f(b) = f(b')$. This equivalence relation is written as $b \equiv b' \mod f$.

> **Exercise I.31.** Verify that $b \equiv_f b'$ iff $f(b) = f(b')$ is an equivalence relation.

We write $B/\equiv_f$ for the quotient set of B by this equivalence relation, and $p : B \longrightarrow B/\equiv_f$ for the projection that sends each element of B into the equivalence class containing it. The projection function, p, is a surjection, and $p(b) = p(b')$ iff $b \equiv_f b'$, i.e., iff $f(b) = f(b')$. The key result is in the following exercise.

> **Exercise I.32.** Show that in the above situation there is a unique function $\overline{f} : B/\equiv_f \longrightarrow A$ with $f = \overline{f}p$. Moreover if f is a surjection, then $\overline{f}$ is a bijection.

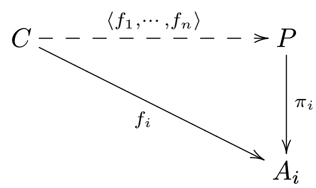
So in the category of sets quotient sets correspond precisely to equivalent quotient objects. As with subobjects this is also true in many other familiar categories such as the categories of groups, rings, vector spaces, etc., but not always. For the details see the various entries in the Catalog of Categories (Appendix B).

We will revisit quotient objects and subobjects from time to time in later sections. For the moment we do not even have the notion of an equivalence relation for general categories. This is one of the topics we will address in Section VI.1. In particular see the definition VI.6.

I.3 Special Objects

I.3.1 Products and Sums

Definition I.27: The **product** of a finite family $C_1, \cdots, C_n$ of objects in $\mathscr{C}$is an object, P, together with a family of morphisms $\pi_i : P \longrightarrow C_i$ so that for *every* family of morphisms $f_i : C \longrightarrow C_i$ there exists a *unique* morphism $\langle f_1, \cdots, f_n \rangle : C \longrightarrow P$ with that $\pi_i \langle f_1, \cdots, f_n \rangle = f_i$.



[We read this diagram as saying that $\langle f_1, \cdots, f_n \rangle$ is the *unique* morphism making the diagram commute. Note that implicitly there are n triangles in this diagram featuring (π_1, f_1), ..., (π_n, f_n) and all with the common edge $\langle f_1, \cdots, f_n \rangle$.]

Similar situations occur constantly in the study of categories, under the name *Universal Mapping Property*. The Universal Mapping Property for a product of $A_1, \cdots, A_n$ is that *every* family of morphisms $f_i : C \longrightarrow A_i$ *uniquely* factors through the family $\pi_i : P \longrightarrow A_i$ of projections. We will mention examples of other Universal Mapping Properties as they occur, and then discuss the many ramifications in Chapter V (Universal Mapping Properties).

Note: The definition of a product is a template for universal mapping property definitions throughout category theory, so it is important to understand just what is required to prove that something is a product.

The important first part is that a product is **not** just an object. In set theory there is *the* product of two sets, and it is a certain unique set. In the category of sets by contrast, a product of two sets is a set together with two projection functions. And while products in the category of sets are in a important certain sense unique (see the next proposition), it is definitely not the object that is unique.

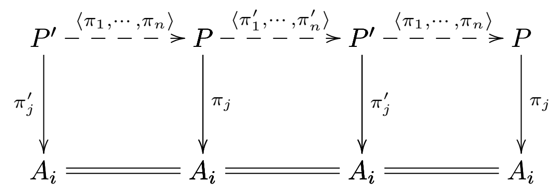
The next crucial part of the definition is the requirement that for **every** object C in the category, and for **every** family of morphisms $f_i : C \longrightarrow A_i$ a morphism $f : C \longrightarrow P$ exists such that for all i we have $\pi_i f = f_i$.

And the final critical requirement is that the morphism f which is asserted to exist in the previous paragraph is the **unique** morphism which satisfies those equations, i.e., if we have both f and g with $\pi_i f = f_i = \pi_i g$, then it must follow that $f = g$.

And equally, these are exactly the properties that can be used when we have the hypothesis that (P, $\pi_i : P \longrightarrow A_i$) is a product of the family (A_i) of objects.

Proposition I.1 *If P with $\pi_i : P \longrightarrow A_i$ and P' with $\pi'_i : P' \longrightarrow A_i$ are both products of $A_1, \cdots, A_n$, then $\langle \pi_1, \cdots, \pi_n \rangle : P \longrightarrow P'$ is an isomorphism with $\langle \pi'_1, \cdots, \pi'_n \rangle : P' \longrightarrow P$ as inverse.*

Proof: Consider the commutative diagram



But $1'_P$ is the *unique* morphism with $\pi'_j 1_{P'} = \pi'_j$, so $\langle \pi_1, \cdots, \pi_n \rangle \langle \pi'_1, \cdots, \pi'_n \rangle = 1'_P$. And 1_P is the *unique* morphism with $\pi_j 1_P = \pi_j$, so $\langle \pi'_1, \cdots, \pi'_n \rangle \langle \pi_1, \cdots, \pi_n \rangle = 1_P$. ■

Note the use of ■ to mark the end (or omission) of a proof. As noted in the preface (p. iv) We will use it in this way from time to time throughout these notes.

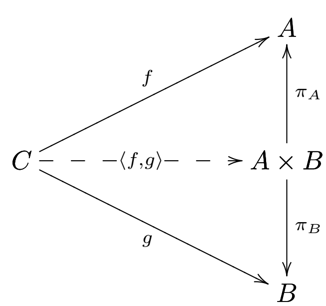
Notation: We will speak of "the" product of $A_1, \cdots, A_n$ and will denote the object as $\Pi_{i=1}^n A_i$ with the *projection morphisms* $\pi_1, \cdots, \pi_n$, but it is important to keep in mind that the product is only unique up to a unique isomorphism.

When there is no danger of confusion, we will for each morphism

$$f : C \longrightarrow \Pi_{i=1}^n A_i$$

write f_i for $\pi_i f$, so that $f = \langle f_1, \cdots, f_n \rangle$. Notice the frequently useful observation that the identity morphism on the product $\Pi_{i=1}^n A_i$ is $\langle \pi_1, \cdots, \pi_n \rangle$.

For a product of two objects we will usually write $A \times B$ for the (object of the) product of the two objects A and B and will write the projection morphisms as $\pi_A : A \times B \longrightarrow A$ and $\pi_B : A \times B \longrightarrow B$. The diagram for the definition is



In particular note that $1_{A \times B} = \langle \pi_A, \pi_B \rangle$.

When $A = B$ the notation π_A is ambiguous, so we must use the π_i notation. The special case where all the objects are the same is common and important, so we have the special notation A^2 for $A \times A$, and generally A^n for $\Pi_{i=1}^n A$.

The Catalog of Categories (Appendix B) discusses products for all of the categories there, but it seems worthwhile to at least note the situation in a few familiar categories. In **Set**, $A \times B = \{(a,b) | a \in A, b \in B\}$ with $\pi_A(a,b) = a$ and $\pi_B(a,b) = b$. And given functions $f : C \longrightarrow A$ and $g : C \longrightarrow B$, $\langle f, g\rangle(c) = (f(c), g(c))$

For most of the familiar categories of "structured sets", e.g., categories of groups, rings, vector spaces, topological spaces, etc., the same construction works equally well.

Recalling that Hom(C, A)stands for the collection of morphisms from C to A, we can also state the definition as: $(\Pi_{i=1}^n A_i, \pi_1, \cdots, \pi_n)$ is a product iff

$$\text{Hom}(C, \Pi_{i=1}^n A_i) \longrightarrow \Pi_{i=1}^n \text{Hom}(C, A_i)$$
$$f \longmapsto (\pi_1 f, \cdots, \pi_n f)$$

is a bijection with inverse

$$\Pi_{i=1}^n \text{Hom}(C, A_i) \longrightarrow \text{Hom}(C, \Pi_{i=1}^n A_i)$$
$$(f_1, \cdots, f_n) \longmapsto \langle f_1, \cdots, f_n\rangle$$

Yet another way of saying this is that for every object C, the morphism $\langle \pi_{1*}, \cdots, \pi_{n*}\rangle : \text{Hom}(C, \Pi_{i=1}^n A_i) \longrightarrow \Pi_{i=1}^n \text{Hom}(C, A_i)$ is an isomorphism in the category of sets.

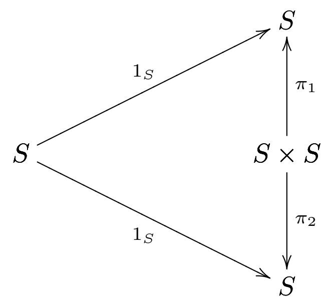
Note that $(f_1, \cdots, f_n)$ is an n-tuple of morphisms, while $\langle f_1, \cdots, f_n\rangle$ is just one morphism, the single unique morphism with $\pi_i\langle f_1, \cdots, f_n\rangle = f_i$ for $i = 1, \cdots, n$.

Now we want to look at an unfamiliar category to illustrate just what is involved in showing we do or do not have a product.

Exercise I.33. From exercise I.3 recall the category $\mathscr{E}$ with objects all sets, but with morphisms only the surjections.
Let S be the set $\{0, 1\}$ considered as object of $\mathscr{E}$. Certainly the product $(S \times S, \pi_1, \pi_2)$ exists in **Set**. Moreover $S \times S$ is an object of $\mathscr{E}$ and π_1 and π_2 are morphisms in $\mathscr{E}$.

(a) Show that if a product of S with itself exists in $\mathscr{E}$, then it must be $(S \times S, \pi_1, \pi_2)$.

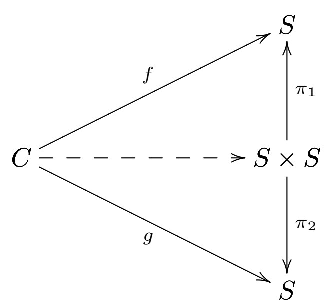
(b) Consider the diagram



and confirm that there is *no* surjection $S \longrightarrow S \times S$ which makes this diagram commute.

Why does this tell us that there is *no* product of S with itself in $\mathscr{E}$?

There are many objects and maps C, f and g in $\mathscr{E}$ where there is a morphism from $C \longrightarrow S \times S$ in $\mathscr{E}$ making the following diagram commute:



Verify that whenever such a morphism exists, it is unique.

The following exercises are designed both to exercise your facility in working with the definition, and also to develop the extent to which products do indeed work as our intuition suggests.

Exercise I.34. For any family of two or more objects, $A_1, \cdots, A_n$, in $\mathscr{C}$ prove that $\Pi_{i=1}^{n} A_i$ is isomorphic to $(\Pi_{i=1}^{n-1} A_i) \times A_n$.

Exercise I.35. Suppose $f : P \longrightarrow A \times B$is an isomorphism. Prove that $(P, \pi_1 f, \pi_2 f)$ is also a product of A and B.

Exercise I.36. Define $t : A\times B \longrightarrow B\times A$by $t_1 = \pi_2, t_2 = \pi_1$, i.e., $t = \langle \pi_2, \pi_1 \rangle$. Prove that t is an isomorphism.

The morphism t of the above exercise will be used from time to time, usually in connection with commutative operations of some sort. Because of that we want the following definition.

Definition I.28: The morphism t as defined in exercise I.36 is the **transposition isomorphism**.

This exercise and definition is a special case of a more general one which is harder only because of the bookkeeping involved in dealing with the indices.

Exercise I.37. With n a positive integer, let $p : \{1, \cdots, n\} \xrightarrow{\cong} \{1, \cdots, n\}$ be a permutation. Define $s : \Pi_{i=1}^n A_i \longrightarrow \Pi_{i=1}^n A_{p(i)}$ by $s_{p(i)} = \pi_i$. Show that s is an isomorphism with $s^{-1} : \Pi_{i=1}^n A_{p(i)} \longrightarrow \Pi_{i=1}^n A_i$ defined by $s_i^{-1} = \pi_i$

These two exercises are interesting, but can also be misleading. Notice that the first exercise *does not* say that $B \times A$ is a productof A and B. Indeed it cannot be overemphasized that a product of two objects is not just an object, but rather an object together with two projection morphisms satisfying the *Universal Mapping Property* for a product. The object that is part of the product $(B \times A,\ \pi_B,\ \pi_A)$ is isomorphic to the object $A \times B$ and so, following exercise I.35, there are suitable projections morphisms from $B \times A$ to A and B that will give a product. But those projections are not the projections that are part of $(B \times A,\ \pi_B,\ \pi_A)$, even in the special case when $A = B$!

In many of the familiar categories the projection morphisms from a product to each of its factors is surjective. That suggests that perhaps in categories they are always surjective, so we ask the following question.

Exercise I.38. Does the projection $\pi_1 : A \times B \longrightarrow A$ have to be an epimorphism? Prove or give a counter-example. (Hint: Carefully consider the category of sets.)

There are many familiar constructions for sets that carry over quite readily to products in arbitrary categories. We start with the following definition.

Definition I.29: Whenever we have $f_i : A_i \longrightarrow B_i$, we define the **product of**

the morphisms to be the unique morphism $f = \Pi_{i=1}^n f_i : \Pi_{i=1}^n A_i \longrightarrow \Pi_{i=1}^n B_i$ such that $\pi_i f = f_i \pi_i$

Just as with the product of objects, if we have just two morphisms, say f and g, we will write $f \times g$ for the product of the two. And sometimes we will write $f \times g \times h$, etc.

These definitions make for a number of exercises.

Exercise I.39. In the category of sets, if $f_1 : A_1 \longrightarrow B_1$ and $f_2 : A_2 \longrightarrow B_2$ are two functions and $(a_1, a_2) \in A_1 \times A_2$, then what is $(f_1 \times f_2)(a_1, a_2)$?

Exercise I.40. Show that $\Pi_{i=1}^n 1_{A_i} = 1_{\Pi A_i}$.

Exercise I.41. Consider the families of morphisms $f_i : A_i \longrightarrow B_i$ and $g_i : B_i \longrightarrow C_i$. Verify that $(\Pi_i g_i)(\Pi_i f_i) = \Pi_i (g_i f_i)$

Exercise I.42. Show that if f and g have retracts f' and g' respectively, then $f' \times g'$ is a retract for $f \times g$.

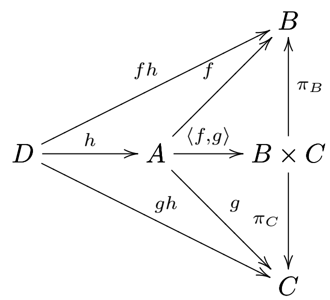
Exercise I.43. Show that if f and g are monomorphisms, then $f \times g$ is a monomorphism.

Exercise I.44. Show that if f and g have sections f' and g' respectively, then $f' \times g'$ is a section for $f \times g$

In most familiar categories the product of epimorphismsis an epimorphism as well, but that is *not* universally true. But neither is it very easy to give an example. See if you can find one!

Recall the induced function $h^* : \mathrm{Hom}(A, B) \longrightarrow \mathrm{Hom}(D, B)$ given by $h * (f) = fh$ when $h : D \longrightarrow A$ and $f : A \longrightarrow B$. Look at what happens with products:

Exercise I.45. Suppose that $f : A \longrightarrow B$, $g : A \longrightarrow C$ and $h : D \longrightarrow A$. Verify that $\langle f, g\rangle h = \langle fh, gh\rangle$.
Hint: Here is the relevant diagram.



Here is a simple but important example which is worthy of a formal definition.

Definition I.30: For every object A this is the **diagonal morphism** $\Delta = \langle 1_A, 1_A\rangle : A \longrightarrow A \times A$.

This will be written as Δ_A if we need to emphasize the particular object.

Exercise I.46. Verify that for any morphisms $f, g : A \longrightarrow B$, we have $\langle f, g\rangle = (f \times g)\Delta$.

Exercise I.47. Let X be any object of **Set**. What is $\Delta(x)$ for $x \in X$?

Exercise I.48. Let A be any Abelian group in **Ab**. What is $\Delta(a)$ for any $a \in A$?

The answers for the two previous exercises are the same, and for general reasons we will explain when we discuss Algebraic Categories in Chapter VII.

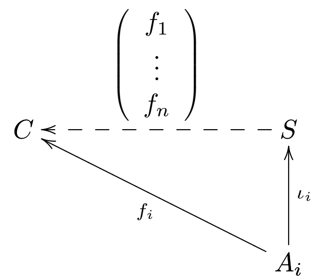
For every concept defined in a general category, there is a dual concept that is gotten by "reversing all the arrows". This notion of duality is itself a very general and important concept that will be discussed at length in Section II.1. Leading up to that we will give many example of dual definitions, theorems and proofs, starting with the dual of products.

Definition I.31: For any finite family $A_1, \cdots, A_n$ of objects in $\mathscr{C}$, a **coproduct** or **sum** of these objects is an object, S together with a family of morphism $\iota_j : A_j \longrightarrow S$ where for every family of morphisms $f_j : A_j \longrightarrow C$ there exists a *unique* morphism

$$\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} : S \longrightarrow C$$

such that

$$\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} \iota_j = f_j$$



[We read this diagram as saying that

$$\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$$

is the unique morphism that makes the diagram commute. Note that implicitly there are n triangles in this diagram featuring (ι_1, f_1), ..., (ι_n, f_n) and all with the common edge $\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$.]

Note: In Section I.3.3 we'll see the good reasons why we would like to write the morphism from a sum into another object as a column vector, but space considerations demands a more compact notation, so we will use $[f_1, \cdots, f_n]$ as a synonym for

$$\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$$

For reasons that we will explore more fully when we discuss dual categories and duality in Section II.1, for each result about products, there is a corresponding result about sums. All of the following results in this section, including the exercises, are examples.

Proposition I.2 *If S with $\iota_j : A_i \longrightarrow S$ and S' with $\iota'_j : A_i \longrightarrow S'$ are both sums of $A_1, \cdots, A_n$, then $[\iota_1, \cdots, \iota_n] : S' \longrightarrow S$ is an isomorphism with $[\iota'_1, \cdots, \iota'_n] : S \longrightarrow S'$ as inverse.*

Proof: Consider the commutative diagram

$$\begin{array}{ccccccc} A_i & = & A_i & = & A_i & = & A_i \\ {\scriptstyle \iota'_j}\downarrow & & \downarrow{\scriptstyle \iota_j} & & \downarrow{\scriptstyle \iota'_j} & & \downarrow{\scriptstyle \iota_j} \\ S' & \xrightarrow[\begin{pmatrix}\iota_1\\ \vdots\\ \iota_n\end{pmatrix}]{} & S & \xrightarrow[\begin{pmatrix}\iota'_1\\ \vdots\\ \iota'_n\end{pmatrix}]{} & S' & \xrightarrow[\begin{pmatrix}\iota_1\\ \vdots\\ \iota_n\end{pmatrix}]{} & S \end{array}$$

But $1_{S'}$ is the *unique* morphism with $1_{S'}\iota'_j = \iota'_j$, so $[\iota'_1, \cdots, \iota'_n][\iota_1, \cdots, \iota_n] = 1'_S$. And 1_S is the *unique* morphism with $1_S\iota_j = \iota_j$, so $[\iota_1, \cdots, \iota_n][\iota'_1, \cdots, \iota'_n] = 1_S$. ∎

Compare this solution with the proof of Proposition I.1. The two were written carefully to make it clear that each can be transformed into the other by "reversing the arrows" and exchanging sums and products. This is an example of duality which is discussed formally in Section II.1.

Notation: We will speak of "the" sum or coproduct of $A_1, \cdots, A_n$ and will denote the object as $\Sigma_{i=1}^n A_i$ with the *injection morphisms*

$$\iota_j : A_j \longrightarrow \Sigma_{i=1}^n A_i, j = 1, \cdots, n,$$

but it is important to keep in mind that the sum is only unique up to a unique isomorphism.

In the special case of two objects A and B we will usually write $A + B$ for the sum and write $\iota_A : A \longrightarrow A + B, \iota_B : B \longrightarrow A + B$ for the injection morphisms.

When $A = B$ the notation ι_A is ambiguous, so we must use the ι_i notation. The special case where all the objects are the same is common and important, so we have the special notation $2 \bullet A$ for $A + A$, and generally $n \bullet A$ for $\Sigma_{i=1}^n A$.

When there is no danger of confusion, we will for each morphism

$$f : \Sigma_{i=1}^n A_i \longrightarrow C$$

write f_i for $f\iota_j$, so that $f = [f_1, \cdots, f_n]$. Remember the frequently useful observation that the identity morphism on the sum $\Sigma_{i=1}^n A_i$ is $[\iota_1, \cdots, \iota_n]$.

The Catalog of Categories (Appendix B) discusses sums for all of the categories there, and for some of them (such as the categories of Abelian groups and of vector spaces) the sum is a familiar construction. But more commonly, even in the category of sets, sums are not nearly so well known as products. An initial discussion follows on page 34.

Recalling that $\mathrm{Hom}(A, C)$stands for the set of morphisms from A to C, we can also state the definition as: $(\Sigma_{i=1}^n A_i, \iota_1, \cdots, \iota_n)$ is a sum iff

$$\begin{aligned} \mathrm{Hom}(\Sigma_{i=1}^n A_i, C) &\longrightarrow \Pi_{i=1}^n \mathrm{Hom}(A_i, C) \\ f &\longmapsto (f\iota_1, \cdots, f\iota_n) \end{aligned}$$

is a bijection with inverse

$$\Pi_{i=1}^n \mathrm{Hom}(A_i, C) \longrightarrow \mathrm{Hom}(\Sigma_{i=1}^n A_i, C)$$
$$(f\iota_1, \cdots, f\iota_n) \longmapsto [f_1, \cdots, f_n]$$

Exercise I.49. For any family of two or more objects $A_1, \cdots, A_n$ in $\mathscr{C}$ prove that $\Sigma_{i=1}^n A_i$ is isomorphic to $(\Sigma_{i=1}^{n-1} A_i) + A_n$. (Compare to exercise I.34.)

Exercise I.50. Suppose $f : A + B \longrightarrow S$ is an isomorphism. Prove that $\langle S, f\iota_1, f\iota_2\rangle$ is also a sum of A and B. (Compare to exercise I.35.)

Exercise I.51. Define $t : A+B \longrightarrow B+A$ by $t_1 = \iota_2$, $t_2 = \iota_1$, i.e., $t = [\iota_2, \iota_1]$. Prove that t is an isomorphism. (Compare to exercise I.36.)

Definition I.32: Write t for $[\iota_2, \iota_1] : A + B \longrightarrow B + A$. By the immediately preceding exercise this is an isomorphism. Just as with the morphism defined in exercise I.36, we call this the **transposition isomorphism**.

Just as with products, this last exercise (and the next) is interesting, but can also be misleading. In particular the previous exercises *does not* say that $B + A$ is a sumof A and B. Indeed it cannot be overemphasized that a sum of two objects is not just an object, but rather an object together with two inclusion morphisms satisfying the *Universal Mapping Property* for a sum. The object that is part of the product $(B + A, \iota_B, \iota_A)$ is isomorphic to the object $A + B$ and so, following exercise I.50, there are suitable injection morphisms from A and B to $B + A$ that will give a sum. But those injections are not the injections that are part of $(B + A, \iota_B, \iota_A)$, even in the special case when $A = B$!

Exercise I.52. With n a positive integer, let $p : \{1, \cdots, n\} \xrightarrow{\cong} \{1, \cdots, n\}$ be a permutation. Define $s : \Sigma_{i=1}^n A_{p(i)} \longrightarrow \Sigma_{i=1}^n A_i$ by $s_{p(i)} = \iota_i$. Show that s is an isomorphism with $s^{-1} : \Sigma_{i=1}^n A_i \longrightarrow \Sigma_{i=1}^n A_{p(i)}$ defined by ${s_i}^{-1} = \iota_i$. (Compare to exercise I.37.)

Exercise I.53. Does the injection $\iota_1 : A \longrightarrow A + B$ have to be a monomorphism? Prove or give a counter-example. (Compare to exercise I.38.)

Definition I.33: Whenever we have $f_i : A_i \longrightarrow B_i$, we define the **sum of the morphisms** $f = \Sigma_{j=1}^n f_j : \Sigma_{j=1}^n A_j \longrightarrow \Sigma_{j=1}^n B_j$ by the $f\iota_j = \iota_j f_j$

Just as with the sum of objects, if we have just two morphisms, say f and g, we will write $f + g$ for the sum of the two. And sometimes we will write $f + g + h$, etc.

And just as for products this makes for a number of exercises.

Exercise I.54. Show that $\Sigma_{i=1}^{n} 1_{A_i} = 1_{\Sigma A_i}$. (Compare to exercise I.40.)

Exercise I.55. Consider the families of morphisms $f_i : A_i \longrightarrow B_i$ and $g_i : B_i \longrightarrow C_i$. Verify that $(\Sigma_i g_i)(\Sigma_i f_i) = \Sigma_i (g_i f_i)$. (Compare to exercise I.41.)

Exercise I.56. Show that if f and g have sections f' and g' respectively, then $f' + g'$ is a section for $f + g$. (Compare to exercise I.42.)

Exercise I.57. Show that if f and g are epimorphisms, then $f + g$ is an epimorphism. (Compare to exercise I.43.)

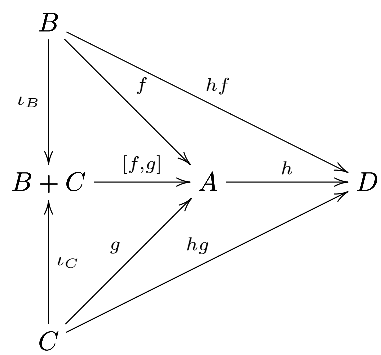
Exercise I.58. Show that if f and g have retracts f' and g' respectively, then $f' + g'$ is a retract for $f + g$. (Compare to exercise I.44.)

Just as with products of epimorphisms, in most familiar categories the sum of monomorphism is a monomorphism as well, but that is *not* universally true. (Indeed that follows from the discussion on duality in Section II.1.) But neither is it very easy to give an example. See if you can find one!

Exercise I.59. Give an example where f and g are monomorphisms, but $f+g$ is not a monomorphism. (Compare with the remark on page 28.)

Recall the induced function $h_* : \mathrm{Hom}(B, A) \longrightarrow \mathrm{Hom}(B, D)$ given by $h_*(f) = hf$ when $h : A \longrightarrow D$ and $f : B \longrightarrow A$. Look at what happens with sums.

Exercise I.60. Suppose that $f : B \longrightarrow A$, $g : C \longrightarrow A$ and $h : A \longrightarrow D$. Verify that $h[f, g] = [hf, hg]$.
Hint: Here is the relevant diagram.



(Compare to exercise I.45.)

Here is another simple but important example which is worthy of a formal definition. This is the dual of the diagonal morphism from page 29.

Definition I.34: For every object A the **codiagonal morphism** or **folding morphism** is $\nabla = [1_A, 1_A] : A + A \longrightarrow A$.

This will be written as ∇_A if we need to emphasize the particular object.

Exercise I.61. Verify that for any morphisms $f, g : B \longrightarrow A$, we have $[f, g] = \nabla(f + g)$. (Compare to exercise I.46.)

So we have a number of simple results about sums that are all dual to the corresponding properties for products, but what categories actually have sums, and what *are* sums in the categories where they exist? Most of the familiar categories do have sums, and in many cases these are very well known, but in the most basic category of sets the sum is unfamiliar.

If X and Y are disjoint sets, i.e., $X \cap Y = \emptyset$ and $f : X \longrightarrow Z$ and $g : Y \longrightarrow Z$ are any two functions, then we can define $[f, g] : X \cup Y \longrightarrow Z$ by

$$[f, g](t) = \begin{cases} f(t) & \text{if } t \in X \\ g(t) & \text{if } t \in Y \end{cases}$$

Defining $\iota_X : X \longrightarrow X \cup Y$ and $\iota_Y : Y \longrightarrow X \cup Y$ to be the inclusion maps, it is clear that $(X \cup Y,\ \iota_X,\ \iota_Y)$ is a sum of X and Y. If X and Y are not disjoint, there are nonetheless isomorphic sets X' and Y' which are disjoint. (One common construction is $X' = X \times \{0\}$ and $Y' = Y \times \{1\}$,) Then with ι_X and ι_Y as the compositions of the isomorphisms to X and Y with the inclusions into $X' \cup Y'$, we have $(X' \cup Y',\ \iota_X,\ \iota_Y)$ as a sum of X and Y.

Definition I.35: This above construction defines the **disjoint union** of two sets.

It should be clear how to extend this to a arbitrary (finite) sums, and we will leave the details for the reader.

Exercise I.62. If $f_0 : A_0 \longrightarrow B_0$ and $f_1 : A_1 \longrightarrow B_1$ are two functions, then describe $(f_0 + f_1)$ on the elements of $A_0 + A_1$.

The situation in both the category of Abelian groups and in the category of vector spaces is very simple. Define $(A + B = A \times B,\ \iota_A(a) = (a, 0),\ \iota_B(b) = (0, b))$. Now for any two homomorphisms $f : A \longrightarrow C$ and $g : B \longrightarrow C$, define $[f, g]$ by $[f, g](a, b) = f(a) + g(b)$.

Exercise I.63. Verify that $(A+B,\ \iota_A,\ \iota_B)$ is indeed a sum of A and B in **Ab**.

Exercise I.64. Let A be any Abelian group in **Ab**. What is $\nabla(a)$for any $a \in A$?

Exercise I.65. If $f_1 : A_1 \longrightarrow B_1$ and $f_2 : A_2 \longrightarrow B_2$ are two homomorphisms of Abelian groups, describe $(f_1 + f_2)$ on the elements of $A_1 + A_2$.

Exercise I.66. Verify that $(A + B,\ \iota_A,\ \iota_B)$ is indeed a sum of A and B in **Vect**.

Exercise I.67. Let A be any vector space in **Vect**. What is $\nabla(a)$for any $a \in A$?

Exercise I.68. If $f_1 : A_1 \longrightarrow B_1$ and $f_2 : A_2 \longrightarrow B_2$ are two linear transformations of vector spaces, describe $(f_1 + f_2)$ on the elements of $A_1 + A_2$.

The same results hold true more generally in the categories of modules over a ring. These are all examples of Abelian categories which are the topic of Chapter XIII. In other well-known categories, such at the category of groups, there are commonly sums, but the construction is less familiar. These are discussed in the Catalog of Categories (Appendix B).

I.3.2 Final, Initial and Zero Objects

In the category of sets there are several types of sets that play an important role. The first of these are the one element sets!

Definition I.36: An object 1 is a **final object** or *terminal object* iff each object A has exactly one morphism from A to 1. We will write this as $! : A \longrightarrow 1$ when we need to name the morphism.

> **Exercise I.69.** Prove that any two final objects in $\mathscr{C}$ are isomorphic, and the isomorphism is unique.

Notation: We will speak of "the" final object in a category and and will write it as 1 and will write the unique morphism from any object to 1 as !, but it is important to keep in mind that the final object is only unique up to a unique isomorphism.

In **Set** the final objects are all the singletons, i.e., the sets with exactly one element. These are useful objects in the category because it means there is a bijection between functions $f : 1 \longrightarrow A$ and the elements of A. Inspired by this (and other examples we will see later) we make the following definition.

Definition I.37: For any category with a final object, a **point** in an object A is a morphism $1 \longrightarrow A$.

To see some of the value of these definitions, do the following exercise.

> **Exercise I.70.** Show there is a bijection between the points of $A \times B$ and pairs (a, b) where a is a point of A and b is a point of B.

Final objects capture some other properties of singletons as well.

> **Exercise I.71.** If 1 is the final object in a category and A is any object in the category, prove that $A \xleftarrow{1_A} A \xrightarrow{!} 1$ exhibits A as the product of 1 and A. So in any category with a final object, $1 \times A \cong A$.

> **Exercise I.72.** Verify that every point is a monomorphism.

Unfortunately this is not as general as we would like. In the category **Ab** of Abelian groups, the final object is the zero group, and for every Abelian group A there is only one homomorphism from the zero group to any Abelian group. So the set of "points" in an Abelian group has no correspondence with the set of elements of the group. Although the final object in a category is in one sense very trivial, it is also another example of a *Universal Mapping Property.* Even more, as we will see in section V.1, it is the fundamental example of a Universal Mapping Property. We will also see that how it relates to each particular category is important for distinguishing various types of categories. (See Chapter IX.)

Definition I.38: A **Cartesian category** or **category with finite products** is one with a final object where every finite number of objects $C_1, \ldots,$ C_n has a product.

A final object is sensibly considered to be the product of no objects: it has no projection morphisms (there are no factors on which to project) but there is the unique morphism from any object to the final object which has suitable composition with all of those non-existent projections!

There is a dual notion (in a sense that will be explained in Section II.1) which is also important.

Definition I.39: An object 0 is an **initial object** iff there is exactly one morphism from 0 to each object A.

Exercise I.73. Prove that any two initial objects in $\mathscr{C}$ are isomorphic, and the isomorphism is unique. (Compare to exercise I.69.)

In the category of sets there is exactly one initial object, the empty set, and the unique function from the empty set to any set is the empty function. And while functions from a singleton final object to other sets are quite useful, functions from other sets to the empty set are very dull – there aren't any! But exercise I.71 does have a suitable dual.

Exercise I.74. If 0 is the initial object in a category and A is any object in the category, prove that $A \xrightarrow{1_A} A \xleftarrow{1} 0$ exhibits A as the sum of 0 and A. So in any category with a initial object, $0 + A \cong A$. (Compare to exercise I.71.)

Definition I.40: A **co-Cartesian category** or **category with finite sums** is one with an initial object where every finite collection of objects $C_1, \ldots, C_n$ has a sum.

Just as a final object is a product of no factors, an initial object is sensibly considered to be the sum of no objects: it has no injection morphisms (there is nothing to inject) but there is the unique morphism from the initial object to any object which composes properly with all of the non-existent injections!

Initial objects are just as trivial and interesting as final objects. They, too, provide another example of a *Universal Mapping Property.* Indeed in Section V.1.) we will see they are a fundamental example.

Sometimes initial objects and final objects are the same, and this deserves a special name.

Definition I.41: In a category that has both an initial object, 0, and a final object, 1, there is a unique morphism $! : 0 \longrightarrow 1$. If this is an isomorphism, then we speak of a **zero object** and write it as 0.

Of course in a category with a zero object all initial objects, final objects and zero objects are then uniquely isomorphic.

Definition I.42: If a category has a zero object, then for any objects A and B we have $A \xrightarrow{!} 0 \xrightarrow{!} B$, i.e., there is a *unique* morphism which "factors through" the zero object. All such morphisms are called **zero morphisms** and are denoted by 0.

Although the category of sets does not have a zero object, there are very many familiar categories that do. In the categories of monoids, of groups, of Abelian groups, of vector spaces, and many others, there is a trivial object that is the zero object in the category.

There are also interesting and natural categories that have a zero object but which are not "algebraic". One good example is the category of pointed sets.

Definition I.43: A **pointed set** is a pair, (X, x_0), consisting of a non-empty set, X, together with an element (called the **base point**,) x_0, of X.

Definition I.44: A morphism between pointed sets $f : (X, x_0) \longrightarrow (Y, y_0)$ is a function $f : X \longrightarrow Y$ with $f(x_0) = y_0$.

Just as we have the category **Set** of sets and functions, we have the category $\mathbf{Set}_*$ of pointed sets together with their morphisms. For more details see B.1.7.

The category $\mathbf{Set}_*$ has initial objects, final objects and zero objects: any one element set with the single element being (necessarily) the base point element.

$\mathbf{Set}_*$ has products and sums. The product of two pointed sets, (X, x_0) and (Y, y_0) "is" $(X \times Y, (x_0, y_0))$ with the usual projection maps.

The sum of two pointed sets is similar to but a bit simpler than the disjoint union of arbitrary sets. Define $(X, x_0) + (Y, y_0) = ((X \times \{y_0\}) \cup (\{x_0\} \times Y), (x_0, y_0))$, $\iota_X : (X, x_0) \longrightarrow (X, x_0) + (Y, y_0)$ by $\iota_X(x) = (x, y_0)$ and $\iota_Y : (Y, y_0) \longrightarrow (X, x_0) + (Y, y_0)$ by $\iota_Y(y) = (x_0, y)$

Definition I.45: The above construction of the sum of two pointed sets is usually called the **join** of the two pointed sets.

Now for any two morphisms $f : (X, x_0) \longrightarrow (Z, z_0)$ and $g : (Y, y_0) \longrightarrow (Z, z_0)$

we can define $[f, g] : (X, x_0) + (Y, y_0) \longrightarrow (Z, z_0)$ by

$$[f, g](t) = \begin{cases} f(t) & \text{if } t \in X \\ g(t) & \text{if } t \in Y \end{cases}$$

It is very easy to check the details to verify that this does indeed exhibit a sum of (X, x_0) and (Y, y_0), so it will be left for the reader.

It should be clear how to extend this to a arbitrary (finite) sums, and we will also leave those details for the diligent reader.

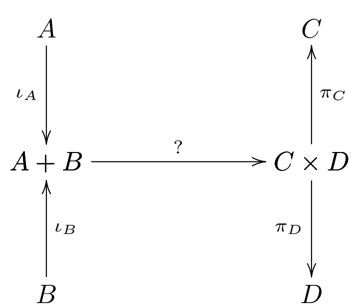
I.3.3 Direct Sums and Matrices

As an example of how this abstraction begins to connect back to more familiar mathematics, and at the same time provides a bridge to new ideas we want to look at morphisms from sums to products.

The recognition of the connections among direct sums, matrices of morphisms and addition of morphisms first appeared in Mac Lane [49](though much of the terminology has changed since then.) Most of the material in this section in much this form appeared in Eckmann and Hilton [17].Treatments can also be found in Chapter 5 of Blyth [8],section 1.591 of Freyd and Scedrov [25]and Session 26 of Lawvere and Schanuel [47]

Throughout this section we suppose that $\mathscr{C}$ is a category that has finite sums, finite products and a zero object.

Let A, B, C, and D be arbitrary objects in $\mathscr{C}$ and consider the question: "What are the morphisms from $A + B$ to $C \times D$?"



To specify a morphism from any object into $C \times D$ requires giving two morphisms from the object into C and D respectively. Specifying a morphism from $A+B$ to any object requires giving two morphisms from A and B, respectively, into the object. So to give a morphism $A + B \longrightarrow C \times D$ is equivalent to giving four morphisms

$$f : A \longrightarrow C$$
$$g : A \longrightarrow D$$
$$h : B \longrightarrow C$$
$$k : B \longrightarrow D$$

and the morphism is

$$\langle \begin{pmatrix} f \\ h \end{pmatrix}, \begin{pmatrix} g \\ k \end{pmatrix} \rangle$$

or, equivalently,

$$\begin{pmatrix} \langle f, g \rangle \\ \langle h, k \rangle \end{pmatrix}$$

As you probably suspect from the title of this sections, we are going to write this morphism as

$$\begin{pmatrix} f & g \\ h & k \end{pmatrix}$$

We have done this for $A + B \longrightarrow C \times D$ just to avoid the additional complication of indices, but that is really just a matter of bookkeeping. The real result is the following proposition.

Proposition I.3 *The morphisms* $\Sigma_{i=1}^{n} A_i \longrightarrow \Pi_{j=1}^{m} B_j$ *are exactly the "matrices"*

$$M = \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1m} \\ f_{21} & f_{22} & \cdots & f_{2m} \\ & & \ddots & \\ f_{n1} & f_{n2} & \cdots & f_{nm} \end{pmatrix}$$

where $\pi_j M \iota_i = f_{ij}$.

Exercise I.75. Prove Proposition I.3.

Just as in our initial discussion of morphisms $A + B \longrightarrow C \times D$, it is very well worth noting that this also says $M\iota_i = \langle f_{i1}, \cdots, f_{im} \rangle$ and $\pi_j M = [f_{1j}, \cdots, f_{nj}]$.

Another way of looking at this same information to to recall (see pages 25 and 32)that

$$\text{Hom}(C, \Pi_{j=1}^{m} B_j) \cong \Pi_{j=1}^{m} \text{Hom}(C, B_j)$$

and

$$\text{Hom}(\Sigma_{i=1}^{n} A_i, C) \cong \Pi_{i=1}^{n} \text{Hom}(A_i, C)$$

Put those together and look at

$$\text{Hom}(\Sigma_{i=1}^{n} A_i, \Pi_{j=1}^{m} B_j) \cong \Pi_{j=1}^{m} \text{Hom}(\Sigma_{i=1}^{n} A_i, B_j) \cong \Pi_{i=1}^{n} \Pi_{j=1}^{m} \text{Hom}(A_i, B_j)$$

We started this section with the assumption that the category under consideration has finite sums, finite products and a zero object. But the zero object has as yet made no appearance. That changes when we consider morphisms

$$\Sigma_{i=1}^{n} A_i \longrightarrow \Pi_{j=1}^{n} A_j$$

i.e., when we have the same objects in both the sum and product. Now in general there is no distinguished morphism from A_i to A_j when i and j are different, but when the category has a zero object this changes and we have the zero morphism $0 : A_i \longrightarrow 0 \longrightarrow A_j$. And as a result we have the morphism

$$I = \begin{pmatrix} 1_{A_1} & 0 & \cdots & 0 \\ 0 & 1_{A_2} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & 1_{A_n} \end{pmatrix} : \Sigma_{i=1}^{n} A_i \longrightarrow \Pi_{j=1}^{n} A_j$$

Definition I.46: This matrix of morphisms is called the **identity matrix** and is denoted by I.

This notation is deliberately very suggestive, and in some important cases it is exactly right. For example consider the category $\mathbf{FDVect}_K$ of finite dimensional vector spaces over a field K. In this category every vector space of dimension n is isomorphic to $K^n = \Pi_{i=1}^{n} K$. Moreover the morphisms, i.e. linear transformations, $L : K \longrightarrow K$ have the form $L(x) = lx$ for some $l \in K$. For good measure the morphism $I : \Sigma_{i=1}^{n} K \longrightarrow K^n$ that we constructed above is an isomorphism. Even more every morphism $K^m \longrightarrow K^n$ can now be identified with an ordinary matrix

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ & & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

of elements of K.

For more details, see Section B.4.4.

Of course in some cases the suggestion is misleading. For example consider

$$I = \begin{pmatrix} 1_X & 0 \\ 0 & 1_Y \end{pmatrix} : X + Y \longrightarrow X \times Y$$

in the category of pointed sets.

If X and Y are not singletons, then this is not an isomorphism, instead:

Exercise I.76. Show that $I : X + Y \longrightarrow X \times Y$ corresponds exactly to the inclusion $(X \times \{y_0\} \cup \{x_0\} \times Y) \subseteq X \times Y$.
Use this to show that if X and Y are not singletons, then I is not surjective.

Categories where the morphism $I : \Sigma_{i=1}^n A_i \longrightarrow \Pi_{j=1}^n A_j$ is an isomorphism are both interesting and common. We will discuss this considerably more in Chapter XIII on Additive and Abelian Categories, but there are interesting consequences we can see right now.

The first simple observation is that when I is an isomorphism, then following exercises I.35 and I.50 we can equip any sum or product with projections and injections that make it both a sum and a product. Such an object equipped with both injections and projections such as to make it a sum and product is called a *direct sum* of the objects. We write this as

$$A \underset{\iota_A}{\overset{\pi_A}{\rightleftarrows}} A \oplus B \underset{\iota_B}{\overset{\pi_B}{\rightleftarrows}} B$$

for two objects, and as

$$\bigoplus A_i \underset{\iota_i}{\overset{\pi_i}{\rightleftarrows}} A_i$$

in the general case.

Once we do this the morphism

$$I : \bigoplus A_i \xrightarrow{\begin{pmatrix} 1_{A_1} & 0 & \cdots & 0 \\ 0 & 1_{A_2} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & 1_{A_n} \end{pmatrix}} \bigoplus A_i$$

really is the identity morphism.

This is sufficiently important that we make a formal definition.

Definition I.47: In a category with finite products, finite sums and a zero object, a **direct sum** of objects $A_1, \ldots, A_n$ is an object, $\bigoplus_{i=1}^n A_i$, and morphisms $\pi_i : \bigoplus_{i=1}^n A_i \longrightarrow A_i$ and $\iota_i : A_i \longrightarrow \bigoplus_{i=1}^n A_i$ for $i = 1$ to n such that

i $(\bigoplus_{i=1}^n A_i, \pi_i : i = 1, \cdots, n)$ is a product of the A_i;

ii $(\bigoplus_{i=1}^n A_i, \iota_i : i = 1, \cdots, n)$ is a sum of the A_i;

iii For each i, $\pi_i \iota_i = 1_{A_i}$; and

iv For each i and each $j \neq i$, $\pi_j \iota_i = 0$.

The name "direct sum" comes from its use for Abelian groups, modules and vector spaces. Another term that is commonly used in the category theory literature is *biproduct*.

Definition I.48: A *category with direct sums* is a category with a zero object where every finite family of objects has a direct sum.

There is no common name for this type of category, though Lawvere [43] has argued for "linear category" largely because of results such as those expounded in this section. That name is not adopted here because it is not commonly used, and because linear category is also used in the literature for at least two other types of categories.

Just as with products (I.38) and sums (I.40), the zero object in such a category is sensibly considered to be the direct sum of zero objects. Contrast this with exercises I.38 and I.53.

In direct analogy with products and sums we will write the direct sum of two objects as $(A \oplus B, \pi_A, \pi_B, \iota_A, \iota_B)$.

Each projection $\pi_i : \bigoplus_{i=1}^n A_i \longrightarrow A_i$ has the corresponding injection $\iota_i : A_i \longrightarrow \bigoplus_{i=1}^n A_i$ as a section (and the injection has the projection as a retract), so the projections are always epimorphisms, and the injections are always monomorphisms.

Now for any two morphisms $f : A \longrightarrow B$ and $g : C \longrightarrow D$, we have already defined $f \times g : A \times B \longrightarrow C \times C$ and $f + g : A + B \longrightarrow C + D$. But when we have direct sums, that means they both are morphisms from $A \oplus B$ to $C \oplus D$ and so both are given by 2×2-matrices. So what is the matrix for $f \times g$? Well we know that $f \times g$ is the unique morphism with $\pi_B(f \times g) = f\pi_A$ and $\pi_D(f \times g) = g\pi_B$. To compute the matrix we need $\pi_C(f \times g)\iota_A$, $\pi_C(f \times g)\iota_B$, $\pi_D(f \times g)\iota_A$, and $\pi_D(f \times g)\iota_B$. But $\pi_C(f \times g)\iota_A = f\pi_A\iota_A = f1_A = f$ and $\pi_C(f \times g)\iota_B = f\pi_A\iota_B = f0 = 0$. A similar computation for the other two shows that the matrix is

$$\begin{pmatrix} f & 0 \\ 0 & g \end{pmatrix}$$

Exercise I.77. By means of a similar computation show that the matrix of $f + g$ is the same as the matrix of $f \times g$, so $f + g = f \times g$.

And so from now on we will write $f \oplus g : A \oplus B \longrightarrow C \oplus D$ in this situation. Also we want to write $\bigoplus_{i=1}^n f_i : \bigoplus_{i=1}^n A_i \longrightarrow \bigoplus_{i=1}^n B_i$ so we need the more general result of the next exercise.

Exercise I.78. In a category with direct sums, verify that the matrix of both $\Pi_{i=1}^n f_i : \bigoplus_{i=1}^n A_i \longrightarrow \bigoplus_{i=1}^n B_i$ and $\Sigma_{i=1}^n f_i : \bigoplus_{i=1}^n A_i \longrightarrow \bigoplus_{i=1}^n B_i$ is

$$\begin{pmatrix} f_1 & 0 & \cdots & 0 \\ 0 & f_2 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & f_n \end{pmatrix}$$

What makes ordinary matrices interesting is primarily that they can be multiplied, and that multiplication of matrices corresponds to composition of the corresponding maps, as is the case for the category of vector spaces. In order to define multiplication of these new matrices, we may be able to use composition of morphisms in place of multiplication, but what is to take the place of sum? We actually have a pair of candidates which we will see are actually the same. As we will see in a bit this has several important consequences, and is not at all just a happy accident.

The two "sums" we define on $\mathrm{Hom}(A, B)$ are:

For $f, g \in \mathrm{Hom}(A, B)$, define

Definition I.49:

$$\mathbf{f} \vartriangle \mathbf{g} = A \xrightarrow{\Delta} A \oplus A \xrightarrow{[f,g]} B$$

and

Definition I.50: $\mathbf{f} \triangledown \mathbf{g} = A \xrightarrow{\langle f,g\rangle} B \oplus B \xrightarrow{\nabla} B$

> **Exercise I.79.** In the category **Ab** of Abelian groups, what are $(f \vartriangle g)(a)$ and $(f \triangledown g)(a)$?

Proposition I.4 *In any category with direct sums* $\vartriangle$ *and* $\triangledown$ *are the same binary operations on* $\mathrm{Hom}(A, B)$

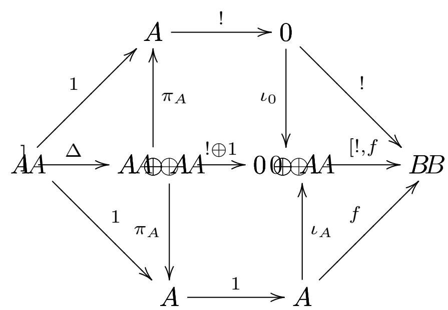
Proof: For f, g in $\mathrm{Hom}(A, B)$,

$$\begin{aligned} f \vartriangle g &= [f, g]\Delta && \text{by definition I.49} \\ &= \nabla(f + g)\Delta && \text{by exercise I.61} \\ &= \nabla(f \times g)\Delta && \text{by exercise I.77} \\ &= \nabla\langle f, g\rangle && \text{by exercise I.46} \\ &= f \triangledown g && \text{by definition I.50} \end{aligned}$$

▮

Definition I.51: For any category with direct sums and any two morphisms $f, g : A \longrightarrow B$ in the category we define $\mathbf{f} + \mathbf{g} = \nabla(f \oplus g)\Delta$. By the preceding proposition $f + g = f \vartriangle g = f \triangledown g$.

Exercise I.80. In any category with direct sums, show that for any morphism f, $f + 0 = f = 0 + f$.
Hint: Explain and use the following commutative diagram.



Exercise I.81. In any category with direct sums, consider morphisms where hf, hg, $f+g$, fe and ge are all defined. Show that $h(f+g) = hf + hg$ and $(f+g)e = fe + ge$.

Using this exercise you can readily verify a result that is quite familiar for Abelian groups, vector spaces, etc.

Exercise I.82. In any category with direct sums, show that for any direct sum $A \oplus B$ we have $\iota_A \pi_A + \iota_B \pi_B = 1_{A \oplus B}$

This exercise also allows us to get back to the motivation for the development of the sum of morphisms: composition of matrices of morphisms.

Proposition I.5 *In any category with direct sums, the composition*

$$\bigoplus_{i=i}^{n} A_i \xrightarrow{\begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1m} \\ f_{21} & f_{22} & \cdots & f_{2m} \\ & & \ddots & \\ f_{n1} & f_{n2} & \cdots & f_{nm} \end{pmatrix}} \bigoplus_{j=1}^{m} B_j \xrightarrow{\begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1p} \\ g_{21} & g_{22} & \cdots & g_{2p} \\ & & \ddots & \\ g_{m1} & g_{m2} & \cdots & g_{mp} \end{pmatrix}} \bigoplus_{k=1}^{p} C_k$$

is the $n \times p$*-matrix* (h_{ik}) *with* $h_{ik} =$

$$\begin{pmatrix} g_{1k} \\ g_{2k} \\ \vdots \\ g_{mk} \end{pmatrix} \quad \langle f_{i1}, f_{i2}, \cdots, f_{im} \rangle$$

Proof: To simplify the notation, we will write F for the matrix (f_{ij}), G for the matrix (g_{jk}) and H for the matrix (h_{ik}). Remember that $h_{ik} = \pi_k H \iota_i = \pi_k G F \iota_i$. But $\pi_k G$ *is* $[g_{1k}, \cdots, g_{nk}]$ and $F\iota_i$ *is* $\langle f_{i1}, f_{i2}, \cdots, f_{im}\rangle$. ∎

Now one more proposition will take us to our target.

Proposition I.6 *In any category with direct sums, the composition*

$$A \xrightarrow{\langle f_1, \cdots, f_m\rangle} \bigoplus_{j=1}^{m} B_j \xrightarrow{[g_1, \cdots, g_m]} C$$

is $g_1 f_1 + \cdots + g_m f_m$

Unfortunately there is a problem. We have not defined the sum of more than two morphisms – well so what, we certainly know how to iterate sums as $(((f+g)+h)+k)$, etc. Unfortunately we don't yet know that our "sum" is associative! So we need a bit of a detour to fix these details, and the first step is some convenient notation – for any object, A, we write A^n for $\bigoplus_1^n A$. (Compare the notation A^n on page 24 and $n \bullet A$ on page 31.)

Definition I.52: In any category with direct sums and any finite number of morphisms $f_1, \cdots, f_n$, all from A to B we define $f_1 + \cdots + f_n$ to be

$$A \xrightarrow{\Delta} A^n \xrightarrow{f_1 \oplus \cdots \oplus f_n} B^n \xrightarrow{\nabla} B$$

Now we have a definition, but we really want to know that this sum is associative. The proof of that is tricky, and we will actually defer it to the discussion preceding corollary 1 on page 54.

Proof of Proposition I.6 Now of course we have things arranged so that the proof is trivial.

$$\begin{aligned}
[g_1, \cdots, g_m]\langle f_1, \cdots, f_m\rangle &= \nabla(g_1 \oplus \cdots \oplus g_m)(f_1 \oplus \cdots \oplus f_m)\Delta \\
&= \nabla(g_1 f_1 \oplus \cdots \oplus g_m f_m)\Delta \\
&= g_1 f_1 + \cdots + g_m f_m
\end{aligned}$$

∎

With the first aspects of "matrix multiplication" in hand, there are many additional developments possible, but this is a good point to change our point of view a bit and consider doing various types of algebra in other categories. Once we have some of that in hand we'll revisit matrices of morphisms as an application and see how it connects to what we have just done.

I.4 Algebraic Objects

A recurring theme in these notes is moving back and forth between investigating what we can do using familiar concepts in general categories as a way to study categories (as in the previous section where we found that arbitrary categories with finite direct sums have many familiar properties), and using categories to investigate familiar concepts (as we are about to do in this section.) Many familiar algebraic structures, such as groups and rings, can be defined in very general categories. Specialized to familiar categories they then give us new insights and connections.

We will actually return to this theme several times, particularly in Chapter VII (Algebraic Categories). This first discussion is largely based on the treatment given by Eckmann and Hilton [17, 18, 19].

I.4.1 Magmas in a Category

We start with the simplest of algebraic gadgets because it will serve as a useful base for most of what follows, and because it usefully simplifies what we need to do. Unfortunately the simplest algebraic gadget is also an unfamiliar one. It is so simple that it really has little interesting theory of its own.

Following Bourbaki [10, I.1] we have the most basic definition.

Definition I.53: A **magma** is a set, M, together with a *binary operation* or *law of composition*, $\mu : M \times M \longrightarrow M$.

Most commonly the binary operation in a magma is written as $\mu(m, n) = mn$ though in particular examples the operation may be written as $m + n$, $m * n$, m^n or in some quite different fashion.

Here we make **NO** assumptions about the operation – it need not be associative, commutative, nor have any sort of identities. Stipulating that the operation is associative, commutative or satisfies some other identities results in other, often more familiar, objects.

Now as we are discussing categories we immediately want to define a suitable morphism.

Definition I.54: A **magma homomorphism** is a function $f : M \longrightarrow N$ such that $f(xy) = f(x)f(y)$.

Formally, if $(M,\ \mu_M)$ and $(N,\ \mu_N)$ are magmas, then a magma homomorphism $f : (M, \mu_M) \longrightarrow (N, \mu_N)$ is a function $f : M \longrightarrow N$ such that $f(\mu_M(x, y)) = \mu_N(f(x), f(y))$ for $x, y \in M$.

Now we have the category **Magma**of magmas which has as objects all magmas and as morphisms the magma homomorphisms. (You should convince yourself that this is indeed a category.) And look at Section B.2.1 in the Catalog of Categories (Appendix B) for more information about this category.

We could now go forward and define commutative magmas, and topological magmas, smooth magmas and many other variations on this theme, thereby getting a plethora of additional categories. But there is a better way.

Let $\mathscr{C}$ be any category with finite products.

Definition I.55: A **magma** in $\mathscr{C}$ is an object, M, together with a **binary**

operation $\mu : M \times M \longrightarrow M$.

And to go along we define the corresponding morphisms.

Definition I.56: A **magma morphism**, or **morphism of magmas** in $\mathscr{C}$ is a morphism $h : M \longrightarrow N$ such that

$$\begin{array}{ccc} M \times M & \xrightarrow{h \times h} & N \times N \\ \downarrow{\scriptstyle \mu_M} & & \downarrow{\scriptstyle \mu_N} \\ M & \xrightarrow[h]{} & N \end{array}$$

> **Exercise I.83.** Show that there is a category, $\mathbf{Magma}_{\mathscr{C}}$, with objects the magmas in $\mathscr{C}$ and as morphisms the magma morphisms.

Of course $\mathbf{Magma}_{\mathbf{Set}} = \mathbf{Magma}$, as defined in Section B.2.1.

Now we have such categories as $\mathbf{Magma}_{\mathbf{Top}}$, the category of continuous magmas, and $\mathbf{Magma}_{\mathbf{Manifold}}$, the category of smooth magmas.

The category $\mathbf{Magma}_{\mathscr{C}}$ directly inherits some properties from the mother category $\mathscr{C}$. First note that if 1 is a final object in $\mathscr{C}$ then the unique morphism $! : 1 \times 1 \longrightarrow 1$ exhibits 1 as an object in $\mathbf{Magma}_{\mathscr{C}}$. Moreover for any other magma, (M, μ) in $\mathscr{C}$ clearly

$$\begin{array}{ccc} M \times M & \xrightarrow{! \times !} & 1 \times 1 \\ \downarrow{\scriptstyle \mu_M} & & \downarrow{\scriptstyle !} \\ M & \xrightarrow[!]{} & 1 \end{array}$$

commutes, so ! is the unique magma morphism from M to 1. Thus 1 (with its unique binary operation) is also a final object in $\mathbf{Magma}_{\mathscr{C}}$.

Also if $M_1, \ldots, M_n$ are magmas in $\mathscr{C}$ (with binary operations $\mu_1, \ldots, \mu_n$), then we can use the μ_i to define a binary operation, μ, on $\Pi_1^n M_i$. This is done using the isomorphism between $\Pi_1^n M_i \times \Pi_1^n M_i$ and $\Pi_1^n (M_i \times M_i)$ which is given by exercise I.37. Because we will use this particular isomorphism repeatedly it is worthwhile to make it explicit. It is

$$\langle\langle \pi_1, \pi_1 \rangle, \cdots, \langle \pi_n, \pi_n \rangle\rangle : \Pi_{i=1}^n M_i \times \Pi_{i=1}^n M_i \longrightarrow \Pi_{i=1}^n (M_i \times M_i)$$

While the inverse is

$$\langle\langle \pi_1, \cdots, \pi_n \rangle, \langle \pi_1, \cdots, \pi_n \rangle\rangle$$

.

Now we define μ to be the composition of $\Pi_1^n M_i \times \Pi_1^n M_i \xrightarrow{\cong} \Pi_1^n(M_i \times M_i)$ and $\Pi_1^n \mu_i : \Pi_1^n(M_i \times M_i) \longrightarrow \Pi_1^n M_i$. Some diagram chasing leads to the conclusion that $(\Pi_1^n M_i, \mu)$ is the product of the (M_i, μ_i) in $\mathbf{Magma}_{\mathscr{C}}$. The details are in the following exercise.

Exercise I.84. Prove that if $\mathscr{C}$ is any category with finite products, then $\mathbf{Magma}_{\mathscr{C}}$ is a category with finite products.

Now we want to preview a bit of what we will start doing once we introduce functors, the morphisms between categories (cf. Chapter III (Functors).)

Recall that we write $\mathrm{Hom}(A, B)$ for the set of morphisms from A to B(or $\mathscr{C}(A, B)$ if we want to emphasize which category is being discussed.) Moreover we have $\mathrm{Hom}(C, A \times B) \cong \mathrm{Hom}(C, A) \times \mathrm{Hom}(C, B)$ (look back at page 25.) Lets apply that to a magma $(M,\ \mu)$ and notice that for any object C we get a binary operation, $\triangledown$, on $\mathrm{Hom}(C, M)$ defined by $f \triangledown g = \mu\langle f, g\rangle$ Even more, if $h : M \longrightarrow N$ is a magma morphism, then the induced function $h_* : \mathrm{Hom}(C, M) \longrightarrow \mathrm{Hom}(C, N)$is a magma homomorphism.

Exercise I.85. Verify the above assertion that if $h : M \longrightarrow N$ is a magma morphism, then h_* is a magma homomorphism.

Now if $h : D \longrightarrow C$ is any morphism in the category, we also have the induced function $h^* : \mathrm{Hom}(C, M) \longrightarrow \mathrm{Hom}(D, M)$.Recalling exercise I.45 we have the following closely related exercise.

Exercise I.86. Verify that if M is a magma in $\mathscr{C}$, and $h : D \longrightarrow C$ is any morphism, then $h^* : \mathrm{Hom}(C, M) \longrightarrow \mathrm{Hom}(D, M)$ is a magma homomorphism.

One of the motivating examples for this discussion of magmas is the codiagonal $\nabla : A \oplus A \longrightarrow A$ in any category with direct sums. In particular the definition of $\triangledown$ in $\mathrm{Hom}(A, B)$ given in Definition I.50 and several of the results following that definition are special cases of the material in this section. And pretty much all of them will be included once we combine the material here with that in the next section.

Before we move on to that, we want to record the following theorem.

Theorem I.1 *Let M be an object in $\mathscr{C}$ and suppose that for every object C in $\mathscr{C}$ a binary operation, $\triangledown$, is defined on $\mathrm{Hom}(C, M)$ in such a way that for every morphism $h : C \longrightarrow D$ the function $h^* : \mathrm{Hom}(D, M) \longrightarrow \mathrm{Hom}(C, M)$ is a magma homomorphism. Then there is a unique binary operation $\mu : M \times M \longrightarrow M$ so $f \triangledown g = \mu\langle f, g\rangle$ for all f and g in $\mathrm{Hom}(C, M)$*

Proof: Consider $\mu = \pi_1 \triangledown \pi_2 : M \times M \longrightarrow M$. Then for f and g in $\mathrm{Hom}(C, M)$

we have $\langle f, g\rangle : C \times C \longrightarrow M$ and

$$\begin{aligned}\mu\langle f, g\rangle &= (\pi_1 \triangledown \pi_2)\langle f, g\rangle \\ &= \langle f, g\rangle^*(\pi_1 \triangledown \pi_2) \\ &= \langle f, g\rangle^*(\pi_1) \triangledown \langle f, g\rangle^*(\pi_2) \\ &= \pi_1\langle f, g\rangle \triangledown \pi_2\langle f, g\rangle \\ &= f \triangledown g\end{aligned}$$

And if ν is some binary operation on M that induces $\triangledown$, then

$$\begin{aligned}\mu &= \pi_1 \triangledown \pi_2 \\ &= \nu\langle \pi_1, \pi_2\rangle \\ &= \nu 1_{M \times M} \\ &= \nu\end{aligned}$$

∎

Many interesting categories have the property that the Hom sets can be naturally considered as objects in some category other than **Set**. The study of such categories and the consequences of the additional structure is the subject of Enriched Category Theory which is the topic of Chapter XII.

I.4.2 Comagmas in a Category

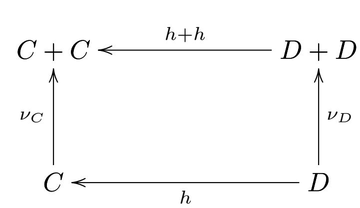
As with just about everything in category theory (and as we will explore more fully in Section II.1)there is a dual to the concept of magma. We are going straight to the "categorical" definition for reasons that should be clear very quickly.

Let $\mathscr{C}$ be any category with finite sums. The dual of the magmas are comagmas. Here is the definition.

Definition I.57: A **comagma** in $\mathscr{C}$ is an object, C, together with a **co-operation**$\nu : C + C \longrightarrow C$.

And to go with it, here is the definition of the appropriate morphisms.

Definition I.58: A **comagma morphism**, or **morphism of comagmas** in $\mathscr{C}$ is a morphism $h : D \longrightarrow C$ such that



> **Exercise I.87.** Show that there is a category, $\mathbf{Comagma}_{\mathscr{C}}$, with objects the comagmas in $\mathscr{C}$ and as morphisms the comagma morphisms. (Compare to exercise I.83.)

We now have, in particular, the category $\mathbf{Comagma} = \mathbf{Comagma}_{\mathbf{Set}}$ as well as many others. Of course the reason that comagmas are actually interesting is probably obscure at this point. For instance there are no interesting comagmas from the category of sets in these notes! Indeed the first interesting specialization of comagmas is to comonoids, the dual of monoids, But the only comonoid in the category of sets is the empty set. (See page 62.) But in time we will see interesting examples on other categories.

The primary reason that we are discussing comagmas here is because of what happens with Hom.Recall that $\mathrm{Hom}(A+B, C) \cong \mathrm{Hom}(A, C) \times \mathrm{Hom}(B, C)$ (look back at page 32.) Lets apply that to a comagma (C, ν) and notice that for any object B we get a binary operation on $\mathrm{Hom}(C, B)$ defined by $f \triangledown g = [f, g]\nu$ Even more, if $h : C \longrightarrow D$ is a comagma morphism, then the induced function $h^* : \mathrm{Hom}(D, B) \longrightarrow \mathrm{Hom}(C, B)$is a magma homomorphism.

> **Exercise I.88.** Verify the above assertion that if $h : C \longrightarrow D$ is a comagma morphism, then h^* is a magma homomorphism. (Compare to exercise I.89.)

Now if $h : A \longrightarrow B$ is any morphism in the category, we also have $h_* : \mathrm{Hom}(C, A) \longrightarrow \mathrm{Hom}(C, B)$.Recalling exercise I.60 we have:

> **Exercise I.89.** Verify that if C is a comagma in $\mathscr{C}$, and $h : A \longrightarrow B$ is any morphism, then $h_* : \mathrm{Hom}(C, A) \longrightarrow \mathrm{Hom}(D, B)$is a magma homomorphism.

The category of comagmas in a category directly inherits some properties from the mother category. First note that if 0 is a initial object in $\mathscr{C}$ the unique morphism $0 \xrightarrow{!} 0 + 0$ exhibits 0 as an object in $\mathbf{Comagma}_{\mathscr{C}}$. Moreover for any other comagma, (C, ν) in $\mathscr{C}$ clearly

$$\begin{array}{ccc} C + C & \xleftarrow{\;!+!\;} & 0 + 0 \\ {\scriptstyle \nu_C}\uparrow & & \uparrow{\scriptstyle !} \\ C & \xleftarrow[\;!\;]{} & 0 \end{array}$$

commutes, so ! is the unique comagma morphism from 0 to C. Thus 0 (with its unique co-operation) is also a initial object in $\mathbf{Comagma}_{\mathscr{C}}$.

Also if $C_1, \ldots, C_n$ are comagmas in $\mathscr{C}$ (with co-operations $\nu_1, \ldots, \nu_n$), then we can use the ν_i to define a co-operation, ν, on $\Sigma_1^n C_i$ by taking ν to be the composition of

$$\Sigma_{i=1}^n \nu_i : \Sigma_1^n C_i \longrightarrow \Sigma_1^n (C_i + C_i)$$

and

$$\Sigma_1^n(C_i + C_i) \xrightarrow{\cong} \Sigma_1^n C_i + \Sigma_1^n C_i$$

Some diagram chasing leads to the conclusion that $(\Sigma_1^n C_i, \nu)$ is the sum of the (C_i, ν_i) in $\mathbf{Comagma}_{\mathscr{C}}$. The details are in the next exercise.

Exercise I.90. If $\mathscr{C}$ is any category with finite sums, then $\mathbf{Comagma}_{\mathscr{C}}$ is a category with finite sums. (Compare to exercise I.84.)

The motivating example for this discussion of comagmas is the diagonal morphism $\Delta : C \longrightarrow C \oplus C$ in any category with direct sums. In particular the definition of $\triangle$ in $\mathrm{Hom}(A, C)$ given in I.50 and several of the results following that definition are special cases of the material in this section. And pretty much all of them will be included once we combine the material here with that in the previous section on magmas.

Before we move on to that, we want to note that if $\mathrm{Hom}(\bullet, M)$ is always a magma and gives magma homomorphisms, then M "is" a magma. This easy theorem has a multitude of ramifications.

Theorem I.2 *Let C be an object in $\mathscr{C}$ and suppose that for every object A in $\mathscr{C}$ a binary operation, $\triangle$, is defined on $\mathrm{Hom}(C, A)$ in such a way that for every morphism $h : A \longrightarrow B$ the function $h_* : \mathrm{Hom}(C, A) \longrightarrow \mathrm{Hom}(C, B)$ is a magma homomorphism. Then there is a unique co-operation $\nu : C \longrightarrow C + C$ so $f \triangle g = [f, g]\nu$ for all f and g in $\mathrm{Hom}(C, A)$. (Compare to Theorem I.1.)*

Proof: Consider $\nu = \iota_1 \triangle \iota_2 : C \longrightarrow C + C$. Then for f and g in $\mathrm{Hom}(C, A)$ we have $[f, g] : C + C \longrightarrow A$ and

$$\begin{aligned}
[f, g]\nu &= [f, g](\iota_1 \triangle \iota_2) \\
&= [f, g]_*(\iota_1 \triangle \iota_2) \\
&= [f, g]_*(\iota_1) \triangle [f, g]_*(\iota_2) \\
&= [f, g]\iota_1 \triangle [f, g]\iota_2 \\
&= f \triangle g
\end{aligned}$$

And if μ is some co-operation on M that induces $\triangle$, then

$$\begin{aligned}
\nu &= \iota_1 \triangle \iota_2 \\
&= [\iota_1, \iota_2]\mu \\
&= 1_{M+M}\mu \\
&= \mu
\end{aligned}$$

∎

Again note the great similarity (or more precisely the duality) of everything we've done to this point in this section with what was done in the preceding section on magmas. This is yet another example of what we will formalize in Section II.1.

I.4.2.1 Comagmas and Magmas Together

In a category with direct sums, every object naturally has both a magma and a comagma structure, and in Section I.3.3 we saw that the two magma structures that were induced on the Hom sets are equal. Here we want to revisit that material in the context of magmas and comagmas. So suppose that A is a comagma in $\mathscr{C}$ with co-operation $\vartriangle$ (deliberately intended to remind us of the diagonal Δ) and B is a magma in $\mathscr{C}$ with binary operation $\triangledown$ (deliberately intended to remind us of the codiagonal ∇). Then, as above, $\mathrm{Hom}(A, B)$ has two binary operations $\vartriangle$ and $\triangledown$ defined by

$$\begin{aligned} f \vartriangle g &= [f, g] \vartriangle \\ h \triangledown k &= \triangledown \langle h, k \rangle \end{aligned}$$

for f, g, h, and k all in $\mathrm{Hom}(A, B)$. And these are connected by this **master identity**:

Proposition I.7 $\boxed{(f \vartriangle h) \triangledown (g \vartriangle k) = (f \triangledown g) \vartriangle (h \triangledown k)}$

Proof:

$$\begin{array}{rcll} (f \vartriangle g) \triangledown (h \vartriangle k) & = & \triangledown \langle f \vartriangle g, h \vartriangle k \rangle & \text{by definition of } \triangledown \\ & = & \triangledown \langle [f, g] \vartriangle, [h, k] \vartriangle \rangle & \text{by definition of } \vartriangle \\ & = & \triangledown \langle [f, g], [h, k] \rangle \vartriangle & \text{by exercise I.89} \\ & = & \triangledown [\langle f, h \rangle, \langle g, k \rangle] \vartriangle & \text{by Proposition I.3} \\ & = & [\triangledown \langle f, h \rangle, \triangledown \langle g, k \rangle] \vartriangle & \text{by exercise I.86} \\ & = & [f \triangledown h, g \triangledown k] \vartriangle & \text{by definition of } \triangledown \\ & = & (f \triangledown h) \vartriangle (g \triangledown k) & \text{by definition of } \vartriangle \end{array}$$

∎

Now assume that $\mathscr{C}$ has a zero object and that the zero morphism in $\mathrm{Hom}(A, B)$ is the identity for both $\vartriangle$ and $\triangledown$, i.e., $f \vartriangle 0 = f = 0 \vartriangle f$ and $g \triangledown 0 = g = 0 \triangledown g$ for all f and g in $\mathrm{Hom}(A, B)$. This assumption combined with the master identity has amazing consequences.

First, taking $g = 0$ and $h = 0$, we get

$$\begin{array}{rcl} f \triangledown k & = & (f \vartriangle 0) \triangledown (0 \vartriangle k) \\ & = & (f \triangledown 0) \vartriangle (0 \triangledown k) \\ & = & f \vartriangle k \end{array}$$

i.e., the two binary operations $\vartriangle$ and $\triangledown$ on $\mathrm{Hom}(A, B)$. And this means the master identity can be rewritten as $\boxed{(f + h) + (g + k) = (f + g) + (h + k)}$

where $+$ stands for either/both of $\vartriangle$ and $\triangledown$.

Part of the reason for using $+$ as the name of the binary operation is because taking $f = 0$ and $k = 0$ gives

$$\begin{aligned} h + g &= (0 + h) + (g + 0) \\ &= (0 + g) + (h + 0) \\ &= g + h \end{aligned}$$

i.e., the binary operation is commutative.

And finally, taking just $h = 0$ we get

$$\begin{aligned} f + (g + k) &= (f + 0) + (g + k) \\ &= (f + g) + (0 + k) \\ &= (f + g) + k \end{aligned}$$

Which is to say that the two (equal) binary operations on $\mathrm{Hom}(A, B)$ are also associative.

Recall that a set together with a binary operation that is associative, commutative and has an identity is a commutative monoid, so the above is summarized in the following theorem.

Theorem I.3 *Let $\mathscr{C}$ be a category with a zero object, $(A, \triangle)$ a comagma in $\mathscr{C}$ and $(B, \triangledown)$ a magma in $\mathscr{C}$ such that the zero morphism in $\mathrm{Hom}(A, B)$ is the identity for the binary operations induced by $\triangle$ and $\triangledown$, then the two binary operations are the same and make $\mathrm{Hom}(A, B)$ a commutative monoid.*

As a first consequence we have the following corollary.

Corollary 1 *If $\mathscr{C}$ is a category with finite direct sums, then for any two objects A and B in the category, the binary operation on $\mathrm{Hom}(A, B)$ defined by the diagonal morphism Δ and the folding morphism ∇ makes it a commutative monoid with the zero morphism as the identity.*

∎

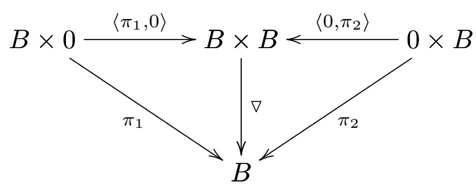
As another interesting consequence notice that if every object in $\mathscr{C}$ admits a comagma structure and a magma structure such that the 0 morphism is the identity in the magma structure induced on the Hom sets, then these structures are necessarily unique. This follows by combining the above proposition with Theorems I.1 and I.2 which say that the magma structure on the Hom sets determine the magma and comagmas in $\mathscr{C}$.

In particular this applies to every category with finite direct sums – the comagma and magma structures defined by Δ and ∇ are unique!

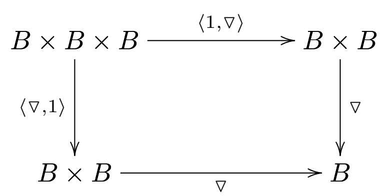
As a direct lead-in to the next section we have the following summary result.

Proposition I.8 *Let $\mathscr{C}$ be a category with a zero object, $(A, \triangle)$ a comagma in $\mathscr{C}$ and $(B, \triangledown)$ a magma in $\mathscr{C}$ such that the zero morphism in $\mathrm{Hom}(A, B)$ is the identity for the common binary operation on $\mathrm{Hom}(A, B)$, then the following diagrams commute:*

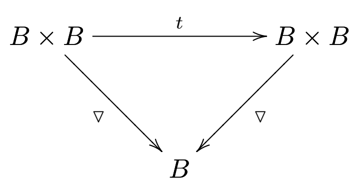
1. $(B, \triangledown)$ has an identity:



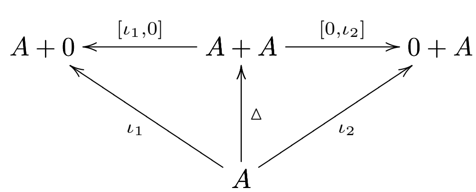
2. $(B, \triangledown)$ is associative:



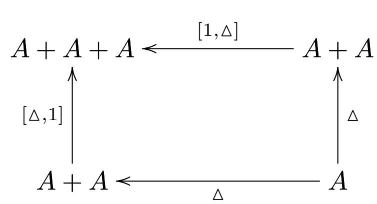
3. $(B, \triangledown)$ is commutative:



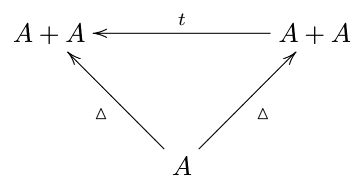
4. $(A, \vartriangle)$ has a co-identity:



5. $(A, \vartriangle)$ is co-associative:



6. (A, $\triangle$) is co-commutative:



Notation: It occasionally happens that we want to discuss $\mathrm{Hom}(A, B)$ where we fix B but vary A. Usually the name A just confuses the issue, so in this situation – the following proof is an example – we will often write $\mathrm{Hom}(\bullet, B)$instead. (The notation $\mathrm{Hom}(-, B)$ is also sometimes used for the same purpose.) This same thing applies for many things besides Hom, and the meaning should be clear in all cases.

Proof: For (1.) note that $\nabla\langle\pi_1, 0\rangle = \pi_1 + 0 = \pi_1$ and $\nabla\langle 0, \pi_2\rangle = 0 + \pi_2 = \pi_2$ as 0 is the identity for $+$, the induced binary operation on $\mathrm{Hom}(\bullet, B)$

For (2.) we see that $\nabla\langle\nabla, 1\rangle = \nabla + 1$, while $\nabla\langle 1, \nabla\rangle = 1 + \nabla$. But we know that $+$ is commutative, so $\nabla + 1 = 1 + \nabla$.

For (3.) we observe that ∇t is also a binary operation on B for which 0 is the identity in the monoid $\mathrm{Hom}(\bullet, B)$. But such a binary operation on B is unique, so $\nabla t = \nabla$.

We will leave parts (4.) – (6.) as the next exercise. ∎

Exercise I.91. Complete the proof of Proposition I.8

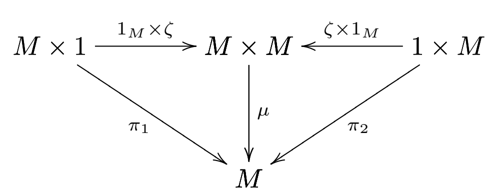
In case they are not clear, the titles for each of the six parts of this proposition will be explained in more detail in the next two sections.

I.4.3 Monoids in a Category

Throughout this section all categories are assumed to have finite products.

Definition I.59: A **monoid object**, or usually just a **monoid**, in $\mathscr{C}$ is a magma in $\mathscr{C}$ $(M, \mu : M \times M \longrightarrow M)$ plus an **identity** or **unit** $\zeta : 1 \longrightarrow M$ for which the following diagrams commute:

identity



associativity

$$\begin{array}{ccc} M \times M \times M & \xrightarrow{\langle 1,\mu\rangle} & M \times M \\ \Big\downarrow{\scriptstyle\langle\mu,1\rangle} & & \Big\downarrow{\scriptstyle\mu} \\ M \times M & \xrightarrow[\mu]{} & M \end{array}$$

This definition will likely make more sense when we see that a monoid in **Set** is just a monoid in the usual sense. Write $\mu(m,n) = m\nabla n$ and $\zeta(*) = 1 \in M$ where $*$ is the unique element in the singleton set 1. (The use of ∇ for the product is intended to remind of the codiagonal, see page 34.) Now the above diagrams become:

1. identity

$$\begin{array}{ccccc} (m,*) \longmapsto (m,1) & & (1,m) \longleftarrow\!\shortmid (*,m) \\ \searrow \quad \Big\downarrow & & \Big\downarrow \quad \swarrow \\ m = m\nabla 1 & & 1\nabla m = m \end{array}$$

2. associativity

$$\begin{array}{ccc} (m,n,p) & \longmapsto & (m,n\nabla p) \\ \Big\downarrow & & \Big\downarrow \\ (m\nabla n,p) & \longmapsto & m\nabla(n\nabla p) = (m\nabla n)\nabla p \end{array}$$

So the first says that 1 is an identity for the multiplication while the second is just the associative law for multiplication, i.e., we have a monoid in the usual sense.

In light of the last section we also want to define commutative monoids in a category.

Definition I.60: Let $t : M \times M \longrightarrow M \times M$ be the transposition isomorphism defined in exercise I.36. A monoid (M, μ, ζ) is **commutative** iff

$$\begin{array}{ccc} M \times M & \xrightarrow{t} & M \times M \\ & {\scriptstyle\mu}\searrow \quad \swarrow{\scriptstyle\mu} & \\ & M & \end{array}$$

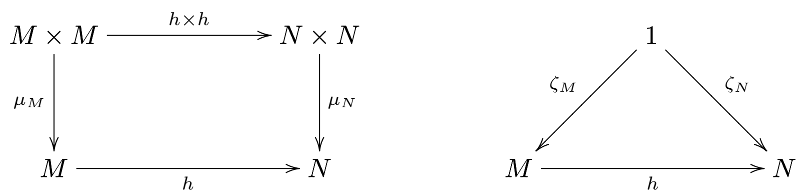
is commutative.

Note that in **Set** this just says $m \nabla n = n \nabla m$, i.e., the multiplication is commutative in the ordinary sense.

Now we can restate the first three parts of Proposition I.8 as saying that the object B is a commutative monoid in the category. The last three parts say that A is a co-commutative comonoid, and that will be explained and justified in the next section.

As we are studying categories, we are, of course, looking for a category of monoids, so we want the following definition.

Definition I.61: If $(M,\, \mu_M,\, \zeta_M)$ and $(N,\, \mu_N,\, \zeta_N)$ are two monoids in $\mathscr{C}$, then a **monoid morphism** from M to N is a morphism $h : M \longrightarrow N$ in $\mathscr{C}$ for which the following diagrams commute:



The subscripts on μ and ζ will usually be omitted as the intended subscript is usually clear from the context.

Again if we look at what this means in the category of sets, then the first square is the familiar $h(m \vartriangle m') = h(m) \vartriangle h(m')$, while the second is $h(1) = 1$, exactly what it means for h to be a monoid homomorphism.

Exercise I.92. Show that there is a category, $\mathbf{Monoid}_{\mathscr{C}}$, with objects the monoids in $\mathscr{C}$ and with morphisms the monoid morphisms.

Of course $\mathbf{Monoid}_{\mathbf{Set}} = \mathbf{Monoid}$, as defined in Section B.2.3. Well, there is a tiny issue: the definition of a monoid in **Set** isn't exactly the usual definition of a monoid which just asserts that some identity element exists. Fortunately it is an easy theorem that a monoid has a unique identity and so the two concepts really do coincide. For more details, see the material on monoids in Section B.2.3 of the Catalog of Categories. Also look at Mac Lane and Birkhoff [55, I.11].

Exercise I.93. Show that there is a category with objects the commutative monoids in $\mathscr{C}$ and with morphisms the monoid morphisms between commutative monoids.

Of course you've already done much of the work for these last two exercises in exercise I.83.

This exercise is really about showing that the category of commutative monoids in $\mathscr{C}$ is a subcategory of the category of monoids in $\mathscr{C}$.

Just as with the category of magmas in a category, the category $\mathbf{Monoid}_{\mathscr{C}}$ directly inherits various properties from $\mathscr{C}$. Again you've already done most of the next exercise in exercise I.84.

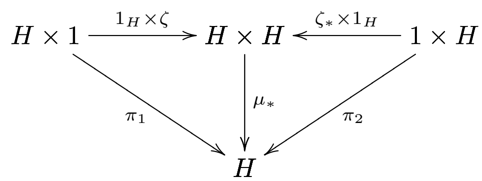
Exercise I.94. Prove that if $\mathscr{C}$ is any category with finite products, then $\mathbf{Monoid}_{\mathscr{C}}$ is a category with finite products.

Similarly for the next two exercises, you did most of the work back in exercises I.85 and I.86.

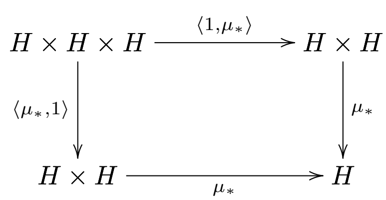
To start, for any monoid (M, μ, ζ) and any object C, the Hom set $\mathrm{Hom}(C, M)$ is naturally a monoid with binary operation $\triangledown$ defined by $f \triangledown g = \mu\langle f, g\rangle$. The identity is $e = \zeta_*(!) \in \mathrm{Hom}(C, M)$. [Remember that $\mathrm{Hom}(C, 1)$ has just the one element ! as 1 is a final object in $\mathscr{C}$.]

Of course we need to verify the identity and associativity relations for $\mathrm{Hom}(C, M)$ and that is easily done by noting that the $\mathrm{Hom}(C, \bullet)$ operation applied to the diagrams for M become (writing H for $\mathrm{Hom}(C, M)$):

1. identity



2. associativity



and, as we saw on page 57 these really are the same as the ordinary definition of a monoid.

Even more, if $h : M \longrightarrow N$ is a monoid morphism, then the induced function $h_* : \mathrm{Hom}(C, M) \longrightarrow \mathrm{Hom}(C, N)$is a monoid homomorphism.

Exercise I.95. Verify the above assertion that if $h : M \longrightarrow N$ is a monoid morphism, then $\mathrm{Hom}(C, h)$ is a monoid homomorphism.

Exercise I.96. Verify that if M is a monoid in $\mathscr{C}$, and $h : D \longrightarrow C$ is any morphism, then $h^* : \mathrm{Hom}(C, M) \longrightarrow \mathrm{Hom}(D, M)$ is a monoid homomorphism.

Finally just as for magmas we have the reverse: If $\mathrm{Hom}(\bullet, M)$ is always a monoid and gives monoid homomorphisms, then M is a monoid.

Theorem I.4 *Let M be an object in $\mathscr{C}$. Suppose that for every object C in $\mathscr{C}$ the Hom set* $\mathrm{Hom}(C, M)$ *is a monoid in such a fashion that for every morphism $h : C \longrightarrow D$ the induced function $h^* : \mathrm{Hom}(D, M) \longrightarrow \mathrm{Hom}(C, M)$ is a monoid homomorphism. Then there are unique morphisms μ and ζ so that (M, μ, ζ) is a monoid in $\mathscr{C}$ inducing the monoid structure in* $\mathrm{Hom}(C, M)$

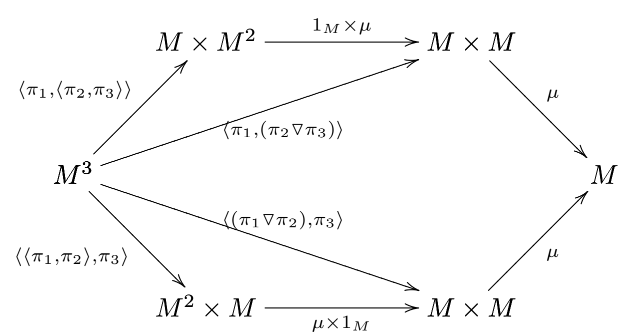
Proof: The first part of this is Theorem I.1 which gives us the unique binary operation μ on M inducing the binary operation on $\mathrm{Hom}(\bullet, M)$.

To get the identity, $\zeta : 1 \longrightarrow M$, note that if it exists it is in $\mathrm{Hom}(1, M)$ and it must be the identity element in that monoid. So let us define ζ to be the identity element in the monoid $\mathrm{Hom}(1, M)$ and prove that it is also the identity for the binary operation μ. The next thing to notice is that as 1 is the final object, for every object N the unique morphism $! : N \longrightarrow 1$ induces a monoid homomorphism $!^* : \mathrm{Hom}(1, M) \longrightarrow \mathrm{Hom}(N, M)$ which in particular takes the identity element in $\mathrm{Hom}(1, M)$ to the identity element in $\mathrm{Hom}(N, M)$, i.e., $\zeta!$ is the identity element in $\mathrm{Hom}(N, M)$.

Applying this to $\pi_2 : M \times 1 \longrightarrow 1$ we first notice that π_2 is really !, so $\zeta\pi_2$ is the identity element in $\mathrm{Hom}(M \times 1, M)$. And $\mu(1_M \times \zeta) = \pi_1 \nabla \zeta\pi_2 = \pi_1$.

The argument for the other half of ζ being an identity is essentially the same and is left to the reader.

To verify the associativity of μ we need to make the diagram in the definition a bit more fulsome:



And finally $\mu\langle \pi_1, \pi_2 \nabla \pi_3 \rangle = \pi_1 \nabla (\pi_2 \nabla \pi_3)$, while $\mu\langle (\pi_1 \nabla \pi_2), \pi_3 \rangle = (\pi_1 \nabla \pi_2) \nabla \pi_3$. But these two are equal because the binary operation ∇ on $\mathrm{Hom}(M^3, M)$ is associative. ▮

In this same situation, if $\mathrm{Hom}(\bullet, M)$ is a commutative monoid, then the corresponding monoid structure on M is commutative.

Theorem I.5 *Let M be an object in $\mathscr{C}$. Suppose that for every object C in $\mathscr{C}$ the Hom set* $\mathrm{Hom}(C, M)$ *has a natural* commutative *monoid structure in such a fashion that for every morphism $h : C \longrightarrow D$ the induced function*

$h^* : \mathrm{Hom}(D, M) \longrightarrow \mathrm{Hom}(C, M)$ is a monoid homomorphism. Then there are unique morphisms μ and ζ so that (M, μ, ζ) is a commutative *monoid in $\mathscr{C}$ inducing the monoid structure in* $\mathrm{Hom}(C, M)$

Exercise I.97. Prove Theorem I.5

As promised we now have an alternative proof of the first three parts of Proposition I.8 on page 54. Indeed this can be recast as saying that when $\mathscr{C}$ is a category with a zero object, $(A, \triangle)$ is a comagma in $\mathscr{C}$, $(B, \triangledown)$ a magma in $\mathscr{C}$, and the zero morphism in $\mathrm{Hom}(A, B)$ is the identity for the common binary operation on $\mathrm{Hom}(A, B)$, then $(B, \triangledown, 0)$ is a commutative monoid.

The last three parts say that $(A, \triangle, 0)$ is a co-commutative comonoid, and the next section explains and justifies this claim.

I.4.4 Comonoids in a Category

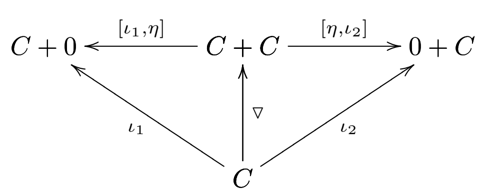
Even though the definition of duality in categories is still to come in Section II.1, the many examples to date should make it clear that every thing in the previous section can be "dualized". We will take advantage of that to leave all the results in this section as exercises.

To start we make assume the dual of the presumption of section I.4.3.

Throughout this section all categories are assumed to have finite sums.

If you find any of this confusing, you should be able to read ahead in Section II.1 as well as reading the solutions.

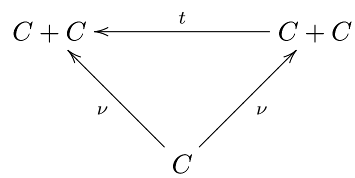
Definition I.62: Suppose $\mathscr{C}$ is a category with finite sums. A **comonoid object** or usually just a **comonoid** in $\mathscr{C}$ consists of an object C and morphisms $\nu : C \longrightarrow C+C$ and $\eta : C \longrightarrow 0$ so that (C, ν) is a comagma and the following diagram commutes:



Terminology: We will often just speak of a comonoid or a comonoid object if the category is clear. The morphism ν is, of course, called the **comultiplication**,while η is called the **co-unit**of the comonoid.

While for monoids we were able to explain the definition of a monoid object by pointing out that it was familiar in the category of sets, we don't have that advantage with comonoids. Comonoids are not familiar, and in the category of sets there is only one comonoid, namely the empty set! So for the moment this is largely an exercise in formal manipulation, with a promise that this will be more interesting later on

Definition I.63: Let $t : C + C \longrightarrow C + C$ be the transposition isomorphismdefined in definition I.32 on 32. A co-monoid $(C,\, \nu,\, \eta)$ is **co-commutative** iff

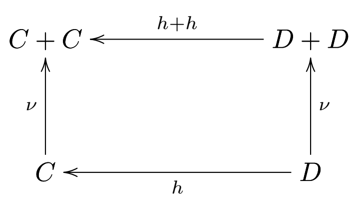


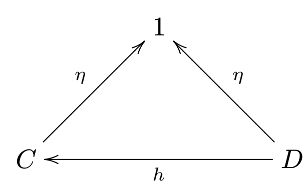
is commutative.

Now see that the second three parts of Proposition I.8 does just say that A is a co-commutative comonoid.

As we are studying categories, we are, of course, looking for a category of comonoids, so we want the following definition.

Definition I.64: If $(C,\, \nu,\, \eta)$ and $(D,\, \nu,\, \eta)$ are two comonoids in $\mathscr{C}$, then a **comonoid morphism** from C to D is a morphism $h : C \longrightarrow D$ in $\mathscr{C}$ for which the following diagrams commute:





Exercise I.98. Show that there is a category, $\mathbf{Comonoid}_{\mathscr{C}}$, with objects the comonoids in $\mathscr{C}$ and as morphisms the comonoid morphisms.

Exercise I.99. Show that there is a category with objects the co-commutative comonoids in $\mathscr{C}$ and as morphisms the comonoid morphisms between co-commutative comonoids.

Of course you've already done most of the work for these two exercises in exercise I.87.

Just as with the category of comagmas in a category, the category $\mathbf{Comonoid}_{\mathscr{C}}$ directly inherits various properties from $\mathscr{C}$. Again you've already done most of the next exercise in exercise I.90.

> **Exercise I.100.** Prove that if $\mathscr{C}$ is any category with finite sums, then $\mathbf{Comonoid}_{\mathscr{C}}$ is a category with finite sums.

Back on page 59 we saw that for any monoid M and any object C the Hom set $\mathrm{Hom}(C, M)$ is naturally a monoid. There is a similar dual result for comonoids leading to the next few exercises.

For any comonoid $(C,\ \nu,\ \eta)$ and any object A the Hom set $\mathrm{Hom}(C, A)$ is naturally a monoid with binary operation, $\triangledown$, on $\mathrm{Hom}(C, M)$ defined by $f \triangledown g = [f, g]\nu$, with identity $\zeta \in \mathrm{Hom}(C, A)$ as the image of $\eta^* : \mathrm{Hom}(0, A) \longrightarrow \mathrm{Hom}(C, A)$, i.e., $\zeta = \eta^*(!) = !\eta$ where $!$ is the unique morphism from an initial object 0 to A.

Of course we need to verify the identity and associativity relations for $\mathrm{Hom}(C, A)$ and that is easily done by noting that the $\mathrm{Hom}(\bullet, A)$ operation applied to the diagrams for C gives (writing H for $\mathrm{Hom}(C, A)$):

1. identity

$$\begin{array}{ccccc} H \times 1 & \xrightarrow{1_H \times \eta} & H \times H & \xleftarrow{\eta_* \times 1_H} & 1 \times H \\ & \searrow^{\iota_1} & \downarrow{\scriptstyle \nu_*} & \swarrow_{\iota_2} & \\ & & H & & \end{array}$$

2. associativity

$$\begin{array}{ccc} H \times H \times H & \xrightarrow{\langle 1, \nu^* \rangle} & H \times H \\ {\scriptstyle \langle \nu^*, 1 \rangle}\downarrow & & \downarrow{\scriptstyle \nu^*} \\ H \times H & \xrightarrow[\nu^*]{} & H \end{array}$$

and, as we saw on page 57 these really are the same as the ordinary definition of a monoid.

Even more, if $h : C \longrightarrow D$ is a comonoid morphism,then the induced function $h^* : \mathrm{Hom}(D, A) \longrightarrow \mathrm{Hom}(C, A)$is a monoid homomorphism as you will verify in the next exercise. Note that most of the work in the next two exercises was done in exercises I.88 and I.89.

> **Exercise I.101.** Verify that if $h : C \longrightarrow D$ is a comonoid morphism, then for every object A the induced function $h^* : \mathrm{Hom}(D, A) \longrightarrow \mathrm{Hom}(C, A)$ is a monoid homomorphism.

> **Exercise I.102.** Verify that if C is a comonoid in $\mathscr{C}$, and $h : A \longrightarrow B$ is any morphism, then $h_* : \mathrm{Hom}(C, A) \longrightarrow \mathrm{Hom}(C, B)$ is a monoid homomorphism.

And finally we have the is dual of Theorem I.4 on page 60, the result for Hom and monoids: If $\mathrm{Hom}(C, \bullet)$ is always a monoid and gives monoid homomorphisms, then C is a comonoid.

Theorem I.6 *Let C be an object in $\mathscr{C}$. Suppose that for every object A in $\mathscr{C}$ the Hom set $\mathrm{Hom}(C, A)$ is a monoid in such a way that for every morphism $h : A \longrightarrow B$ the induced function $h_* : \mathrm{Hom}(C, A) \longrightarrow \mathrm{Hom}(C, B)$ is a monoid homomorphism. Then there are unique morphisms ν and η so that $(C,\, \nu,\, \eta)$ is a comonoid in $\mathscr{C}$ inducing the monoid structure in $\mathrm{Hom}(C, A)$*

Exercise I.103. Prove Theorem I.6

In this same situation, if $\mathrm{Hom}(C, \bullet)$ is a commutative monoid, then the corresponding comonoid structure on C is co-commutative.

Theorem I.7 *Let C be an object in $\mathscr{C}$. Suppose that for every object A in $\mathscr{C}$ the Hom set $\mathrm{Hom}(C, A)$ is a* commutative *monoid in such a way that for every morphism $h : A \longrightarrow B$ the induced function $h_* : \mathrm{Hom}(C, A) \longrightarrow \mathrm{Hom}(C, B)$ is a monoid homomorphism. Then there are unique morphisms ν and η so that $(C,\, \nu,\, \eta)$ is a* co-commutative *comonoid in $\mathscr{C}$ inducing the monoid structure in $\mathrm{Hom}(C, A)$*

Exercise I.104. Prove Theorem I.7

As promised we now have an alternative proof of the second part of Proposition I.8. Indeed this can now be recast as saying that when $\mathscr{C}$ is a category with a zero object, $(A,\, \triangle)$ is a comagma in $\mathscr{C}$ and $(B,\, \nu)$ a magma in $\mathscr{C}$ such that the zero morphism in $\mathrm{Hom}(A, B)$ is the identity for the common binary operation on $\mathrm{Hom}(A, B)$, then $(A,\, \triangle,\, 0)$ is a co-commutative comonoid.

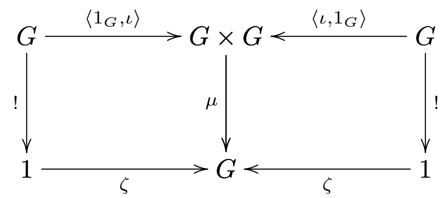
Now as you may have guessed we could continue on to define monoids acting on other objects, or rings or other algebraic gadgets in quite general categories in much the same way we done with magmas and monoids. But we'll postpone that until we have more machinery that will simplify the process considerably (see Chapter VII, Algebraic Categories.) But we do want to look at the categorical generalization of one of the most familiar concepts: groups.

I.4.5 Groups in a Category

A group is a monoid in which every element has an inverse. The corresponding definition in a category is the following.

Definition I.65: Suppose $\mathscr{C}$ is a category with finite products. A **group object** or usually just a **group** in $\mathscr{C}$ consists of an object G and morphisms

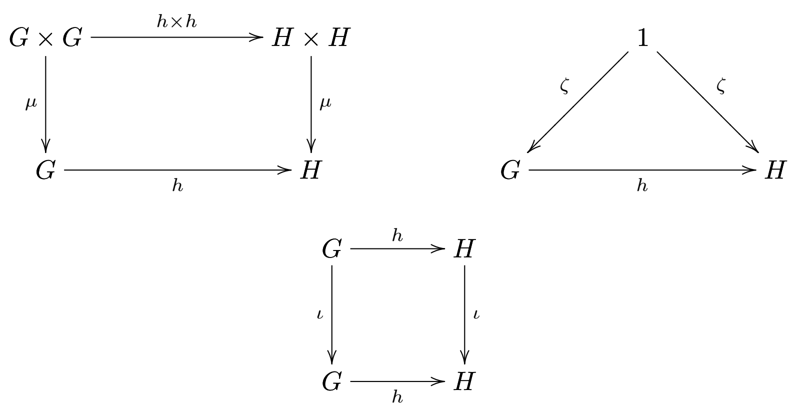
$\mu : G \times G \longrightarrow G$, $\zeta : 1 \longrightarrow G$, and $\iota : G \longrightarrow G$ so that (G, μ, ζ) is a monoid in $\mathscr{C}$ and ι is the inverse. This mean that the following diagrams commute:



Of course this is really just saying that a group object is a monoid object together with an inverse, in particular in the category of sets $\iota : G \longrightarrow G$ is just $\iota(g) = g^{-1}$.

Now we can go one step further and define the corresponding morphisms.

Definition I.66: If (G, μ, ζ, ι) and (H, μ, ζ, ι) are two groups in $\mathscr{C}$, then a **group morphism** from G to H is a morphism $h : G \longrightarrow H$ in $\mathscr{C}$ for which the following diagrams commute:



As you should expect this is a monoid homomorphism that also "carries inverses to inverses", the natural generalization from **Set**.

We call also easily define Abelian, i.e., commutative, groups.

Definition I.67: A group (G, μ, ζ, ι) is **commutative** or **Abelian** iff the monoid (G, μ, ζ) is commutative.

In the next exercise you will show that there is a category of groups in $\mathscr{C}$. Of course most of the work was already done in exercise I.92

> **Exercise I.105.** Show that there is a category, $\mathbf{Group}_{\mathscr{C}}$, with objects the groups in $\mathscr{C}$ and as morphisms the group morphisms.

Of course $\mathbf{Group}_{\mathbf{Set}} = \mathbf{Group}$, as defined in Section B.2.5, right? Well, there is actually a bit to be said here. The issue is that the definition of a group in **Set** does not immediately translate into the usual definition of a group. For example in their book Algebra [55], Mac Lane and Birkhoff give the definition: "A *group* G is a set G together with a binary operation $G \times G \longrightarrow G$, written $(a, b) \mapsto ab$ such that

i This operation is associative.

ii There is an element $u \in G$ with $ua = a = au$ for all $a \in G$.

iii For this element u, there is to each element $a \in G$ an element $a' \in G$ with $aa' = u = a'a$.

Now even though this was written by Saunders Mac Lane, one of the fathers of category theory, there is no mention of the functions ζ and ι that are part of our definition of a group in the category **Set**! Of course the existence of the element u guaranteed by the above definition allows us to define the function $\zeta : 1 \longrightarrow G$ by $\zeta(*) = u$. But a priori it appears that a group in the above sense might allow several different such elements u and therefor several different group objects in **Set**. Of course this is not the case as it is an easy theorem that there is only one such element u.

Similarly the existence for each $a \in G$ of a' with $aa' = u = a'a$ allows us (perhaps with the aid of the axiom of choice!) to define an $\iota : G \longrightarrow G$ so that $a\iota(a) = u = \iota(a)a$. Again there is an a priori possibility that there might be many such functions leading to many different group objects in **Set** associated to a particular group. But again this is not in fact the case as another easy theorem shows that such an element a' is unique for each a and so there is only one such function ι (and the axiom of choice is not needed.)

For more details, see the material on groups in Section B.2.5 of the Catalog of Categories. Also look at Mac Lane and Birkhoff [55, II].

This kind of issue comes up rather regularly when moving between the standard set based definitions and those reformulated to fit category theory. In these early stages we will usually write a bit about the issues, but then we will leave it to the reader without comment unless there are particularly interesting issues.

There is also the category of Abelian groups in $\mathscr{C}$ which is a subcategory of the category of groups in $\mathscr{C}$. That's the content of the next exercise where you've already done the work for in the last exercise and in exercise I.93.

Exercise I.106. Show that there is a category with objects the commutative groups in $\mathscr{C}$ and as morphisms the group morphisms between commutative groups.

This is an opportunity to point out a subtlety that occasionally confuses newcomers to category theory. With the usual definition of groups and monoids (that is as sets with a binary operation), it is quite true that the category of groups is a subcategory of the category of monoids. But if we look at the definition we have given of monoid and group in the category **Set**, it is *not true* that every group is a monoid! The reason for this is that each group has the inverse function (ι) in addition to binary operation (μ) and the identity (ζ). In principal it might be the case that we could have two group objects (G, μ, ζ, ι) and (G, μ, ζ, γ) where the *only* difference is the inverse morphism. In the category of sets this cannot happen – we know that if an element in a monoid has an inverse, the inverse is unique – and that is why the category of groups is a subcategory of the category of monoids. In a general category with finite products there are no elements to have inverses, unique or no. Indeed the role of objects in categories is a very secondary one, so the notion of a subcategory is interesting primarily for the examples arising as subcategories of familiar categories. Even more interesting is the situation illustrated by the relations between magmas, monoids and groups in a category. Every group object does have an underlying monoid object, and each monoid object has an underlying magma object – in both cases the underlying object is gotten by simply "forgetting" the extra structure, i.e., the underlying monoid of a group is gotten by forgetting the inverse, while the underlying magma of a monoid is gotten by forgetting the identity. Moreover the group morphisms between two group objects is a subset of the monoid homomorphisms between the two underlying monoids, and equally well a subset of the magma homomorphism between the two underlying magmas. This is a very common situation and is codified in the discussion of based categories and forgetful functors coming up in Sections II.5 and III.2.6.

Just as with the categories of magmas and monoids in a category, the category $\mathbf{Group}_{\mathscr{C}}$ directly inherits various properties from $\mathscr{C}$. Again you've already done most of the next exercise in exercise I.94.

Exercise I.107. Prove that if $\mathscr{C}$ is any category with finite products, then $\mathbf{Group}_{\mathscr{C}}$ is a category with finite products.

For any object C and any monoid object M, the set of morphisms $\mathrm{Hom}(C, M)$ is naturally a monoid (see page 59.) Similarly for any group object G, the set of morphisms $\mathrm{Hom}(C, G)$ is naturally a group – the multiplication and identity come from the monoid structure that G gives as a monoid. The inverse on $\mathrm{Hom}(C, G)$ is just ι_* where ι is the inverse on G.

Exercise I.108. Verify that ι_* is indeed the inverse on $\mathrm{Hom}(C, G)$ as claimed.

For the next two exercises, most of the work was back in exercises I.95 and I.96.

Exercise I.109. Verify the above assertion that if $h : G \longrightarrow H$ is a group morphism, then for any C $\mathrm{Hom}(\bullet, h)$ is a group homomorphism.

Exercise I.110. Verify that if G is a group in $\mathscr{C}$, and $h : D \longrightarrow C$ is any morphism, then $h^* : \mathrm{Hom}(C, G) \longrightarrow \mathrm{Hom}(D, G)$ is a group homomorphism.

Finally we see that, just as for magmas and monoids, if $\mathrm{Hom}(\bullet, G)$ is always a group and gives group homomorphisms, then G is a group.

Theorem I.8 *Let G be an object in $\mathscr{C}$. Suppose that for every object C in $\mathscr{C}$ the Hom set* $\mathrm{Hom}(C, G)$ *is a group in such a way that for every morphism $h : C \longrightarrow D$ the function $h^* : \mathrm{Hom}(D, G) \longrightarrow \mathrm{Hom}(C, G)$ is a group homomorphism. Then there are unique morphisms μ, ζ, and ι so that (G, μ, ζ, ι) is a group in $\mathscr{C}$ inducing the group structure in* $\mathrm{Hom}(C, G)$

Exercise I.111. Prove Theorem I.8.

Combining this proposition with Theorem I.5 gives us

Theorem I.9 *Let G be an object in $\mathscr{C}$. Suppose that for every object C in $\mathscr{C}$ the Hom set* $\mathrm{Hom}(C, G)$ *is an* Abelian *group in such a way that for every morphism $h : C \longrightarrow D$ the function $h^* : \mathrm{Hom}(D, G) \longrightarrow \mathrm{Hom}(C, G)$ is a group homomorphism. Then there are unique morphisms μ, ζ, and ι so that (G, μ, ζ, ι) is an* Abelian *group in $\mathscr{C}$ inducing the group structure in* $\mathrm{Hom}(C, G)$

To date we have some of the elementary definitions regarding monoids and groups, but no actual development of any interesting information about these categories themselves, not even such simple results as the isomorphism theorems for groups. We could in fact develop that material in a fairly direct manner, but it is much more worth while to first develop some of the machinery of category theory which we can then apply.

Chapter II

Constructing Categories

Examples of categories abound, and we saw that familiar notions in the category of sets generalize to other bases categories to give even more categories. But category theory also suggests other ways of getting more categories – products and sums of categories, subcategories, quotient categories, functor categories, etc. These are particularly important as a way of organizing complexity, especially when we study functors and functor categories will come in the next chapter.

II.1 Duality and Dual Category

Two points determine a line and, in the projective plane, two lines always intersect in a point. This is an ancient example of the duality between lines and points with Pappus' Theorem (see [68]) on nine points and nine lines being perhaps the first duality theorem. Today there are dozens if not hundreds of duality theorems in mathematics, many of them key results.

Duality pairs up objects and relations in a complementary fashion so that there are dual theorems with dual proofs. The simplest examples of duality include negation of propositions in logic (where for example the "and" operation is dual to "or"), dual polyhedra (see [67]) where faces and vertices are swapped (the dual of the cube is the octahedron; the dual of the dodecahedron is the icosahedron; while the tetrahedron is its own dual), and pairing regions and vertices in planar graphs.

More sophisticated examples include Poincaré Duality (see [4]) in algebraic topology and Pontryagin Duality (see [69]) for topological groups.

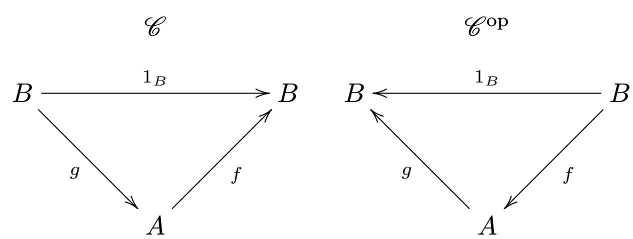
Category theory is another area where duality plays an important role. What makes it particularly interesting is that a great many of the other notions of duality can be expressed by duality for categories. The starting point is the definition of the dual of a category.

Definition II.1: Associated to any category $\mathscr{C}$ is a **dual** or **opposite** category denoted $\mathscr{C}^{\mathrm{op}}$. The objects and morphisms of $\mathscr{C}^{\mathrm{op}}$ are exactly the objects and morphisms of $\mathscr{C}$, but f is a morphism from A to B in $\mathscr{C}$ iff f is a morphism from B to A in $\mathscr{C}^{\mathrm{op}}$. Finally composition is reversed in $\mathscr{C}^{\mathrm{op}}$ from what it is in $\mathscr{C}$. If $h = gf$ in $\mathscr{C}$, then $h = fg$ in $\mathscr{C}^{\mathrm{op}}$.

Clearly $(\mathscr{C}^{\mathrm{op}})^{\mathrm{op}}$ is just $\mathscr{C}$.

Although the definition of the dual category is a very simple formality, it has a great deal of force. The first import is that it reduces the number of proofs in category theory by a factor of 2! The reason for this is that when we prove a theorem for all categories, it of course applies not only to each particular category, but also to its dual category. But the theorem applied in the dual category is (usually) another different theorem in the original category. Here are some simple examples that where we have already done twice the work that was needed.

First we note that section and retract are dual concepts. If $f : A \longrightarrow B$ is a section of $g : B \longrightarrow A$ in $\mathscr{C}$, then $f : B \longrightarrow A$ is a retract of $g : A \longrightarrow B$ in $\mathscr{C}^{\mathrm{op}}$.



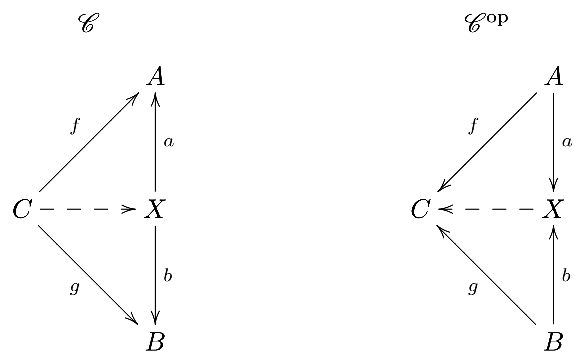
Similarly epimorphism and monomorphism are dual concepts. If the morphism $f : A \twoheadrightarrow B$ is an epimorphism in $\mathscr{C}$, then $f : B \rightarrowtail A$ is a monomorphism in $\mathscr{C}^{\mathrm{op}}$. If $gf = hf \Rightarrow g = h$ in $\mathscr{C}$, then $fg = fh \Rightarrow g = h$ in $\mathscr{C}^{\mathrm{op}}$ (and conversely.)

Next note that final object and initial object are dual concepts, i.e., if X is a final object in the category $\mathscr{C}$, then X is an initial object in the category $\mathscr{C}^{\mathrm{op}}$:

For each object C there is the unique morphism $C \longrightarrow X$ in $\mathscr{C}$ and the unique morphism $X \longrightarrow C$ in $\mathscr{C}^{\mathrm{op}}$.

As our final examples for the moment we see that product and sum are dual concepts. Suppose $\langle X, a : X \longrightarrow A, b : X \longrightarrow B \rangle$ is a product of A and

B in $\mathscr{C}$, then $\langle X, a : A \longrightarrow X, b : B \longrightarrow X\rangle$ is a sum of A and B in $\mathscr{C}^{\text{op}}$:



For each pair of morphisms $f : C \longrightarrow A$ and $g : C \longrightarrow B$ in $\mathscr{C}$ there is a unique morphism $C \longrightarrow X$ in $\mathscr{C}$ making the left diagram commute, and in $\mathscr{C}^{\text{op}}$ for each pair of morphisms $f : A \longrightarrow C$ and $g : B \longrightarrow C$ there is a unique morphism $X \longrightarrow C$ making the right diagram commute. We've only indicated this for pairs of objects, but you should easily see that it is true for arbitrary products and sums as well.

Now consider the following pairs of propositions:

1. a) If a morphism f has a section, then f is an epimorphism. (See exercise I.10.)

   b) If a morphism f has a retract, then f is a monomorphism. (See exercise I.16.)

2. a) (Proposition I.1) If P with $\pi_i : P \longrightarrow A_i$ and P' with $\pi'_i : P' \longrightarrow A_i$ are both products of $A_1, \cdots, A_n$, then $\langle \pi_1, \cdots, \pi_n\rangle : P \longrightarrow P'$ is an isomorphism with $\langle \pi'_1, \cdots, \pi'_n\rangle : P' \longrightarrow P$ as inverse.

   b) (Proposition I.2) If S with $\iota_j : A_i \longrightarrow S$ and S' with $\iota'_j : A_i \longrightarrow S'$ are both sums of $A_1, \cdots, A_n$, then $[\iota_1, \cdots, \iota_n] : S' \longrightarrow S$ is an isomorphism with $[\iota'_1, \cdots, \iota'_n] : S \longrightarrow S'$ as inverse.

3. a) For any family of two or more objects, $A_1, \cdots, A_n$, in $\mathscr{C}$ prove that $\Pi_{i=1}^{n} A_i$ is isomorphic to $\Pi_{i=1}^{n-1} A_i \times A_n$. (See exercise I.34.)

   b) For any family of two or more objects, $A_1, \cdots, A_n$, in $\mathscr{C}$ prove that $\Sigma_{i=1}^{n} A_i$ is isomorphic to $\Sigma_{i=1}^{n-1} A_i + A_n$. (See exercise I.49.)

4. a) Suppose $f : P \longrightarrow A \times B$ is an isomorphism. Prove that $\langle P, \pi_1 f, \pi_2 f\rangle$ is also a product of A and B. (See exercise I.35.)

   b) Suppose $f : A + B \longrightarrow S$ is an isomorphism. Prove that $\langle S, f\iota_1, f\iota_2\rangle$ is also a sum of A and B. (See exercise I.50.)

5. a) Define $t : A \times B \longrightarrow B \times A$ by $t_1 = \pi_2, t_2 = \pi_1$. Prove that t is an isomorphism. (See exercise I.36.)

b) Define $f : A + B \longrightarrow B + A$ by $f_1 = \iota_2, f_2 = \iota_1$. Prove that f is an isomorphism. (See exercise I.51.)

6. a) Any two final objects in $\mathscr{C}$ are isomorphic, and the isomorphism is unique. (See exercise I.69.)

   b) Any two initial objects in $\mathscr{C}$ are isomorphic, and the isomorphism is unique. (See exercise I.73.)

7. a) In any category with a final object, $1 \times A = A$. (See exercise I.71.)

   b) In any category with an initial object, $0 + A = A$. (See exercise I.74.)

These are pairs of dual propositions. In all cases each proposition is a logical consequence of the other because each applies equally well to dual categories. In our proofs of these results (most being in the Solutions (Appendix C)), we have tried to write the proofs to emphasize the duality.Going forward we will prove only one of the dual theorems of interest, and indeed may use the dual theorem without even explicitly stating it.

There are a number of well known results in mathematics that are truly about dual categories. One of the simplest, and one of the motivating examples in the development of the original definitions of category theory, is that the category of finite dimensional vector spaces over a fixed field is self-dual, i.e., it is "equivalent" to its own dual category. To justify that will be one of the early goals of Chapter III (Functors).

II.2 Quotient Categories

Besides subgroups, subrings, subspaces, etc., there are also quotient groups, quotient rings, quotient spaces, etc. Similarly there is the notion of a quotient category.

The initial motivating examples of quotient categories are the various homotopy categories in algebraic topology. For more details look at in the Catalog of Categories at the homotopy category of topological space (B.9.6), the category of H-Spaces (B.9.7), the homotopy category of Kan complexes (B.10.2), and the homotopy category of chain complexes (B.11.4).

As a preliminary to the definition of a quotient category, we need to specify the kind of equivalence relation that is relevant.

Definition II.2: A **congruence** $\sim$on a category is an equivalence relation on the morphisms in the category such that

i If $f \sim g$, then f and g have the same domain and codomain.

ii If $f \sim g$ and $h \sim k$ and hf is defined, then $hf \sim kg$.

If $\sim$ is a congruence on the category $\mathscr{C}$, define $\mathscr{C}/\sim$ as the category with the same objects as $\mathscr{C}$, and with the morphisms from A to B in $\mathscr{C}/\sim$ the equivalence classes of $\mathscr{C}(A,B)$ with respect to the equivalence relation $\sim$. The identity morphism on A is the equivalence class of 1_A, and the composition of equivalence classes is the equivalence class of the composition of any two morphisms in the equivalence classes.

Exercise II.1. Verify that $\mathscr{C}/\sim$ as described above is in fact a category.

Definition II.3: $\mathscr{C}/\sim$ is called the *quotient category* of $\mathscr{C}$ by the congruence $\sim$.

II.3 Product of Categories

Along with subcategories and quotient categories, there are also products of categories. Here is the definition.

Definition II.4: $\mathscr{B}$ are two categories, the **product category** $\mathscr{A} \times \mathscr{B}$ has as objects the pairs (A, B) with A an object of $\mathscr{A}$and B an object of $\mathscr{B}$and as morphisms the pairs (f, g) with f a morphism in $\mathscr{A}$and g a morphism in $\mathscr{B}$. The domains, codomains, composition, and identities are exactly as expected: if $f : A_0 \longrightarrow A_1$ and $f' : A_1 \longrightarrow A_2$ in $\mathscr{A}$while $g : B_0 \longrightarrow B_1$ and $g' : B_1 \longrightarrow B_2$ in $\mathscr{B}$, then $(f, g) : (A_0, B_0) \longrightarrow (A_1, B_1)$ and $(f', g') : (A_1, B_1) \longrightarrow (A_2, B_2)$ in $\mathscr{A} \times \mathscr{B}$. Moreover $(f', g')(f, g) = (f'f, g'g)$ and $1_{(A,B)} = (1_A, 1_B)$.

The domain of (f, g) is the pair (domain(f), domain(g)) and the cod while the of objects from $\mathscr{A}$ and $\mathscr{B}$ respectively, while the morphisms are similar pairs of morphisms.

Exercise II.2. Verify that $\mathscr{A} \times \mathscr{B}$ is indeed a category.

We've only defined the product of two categories, but the definition actually extends to the product of any indexed family of categories as follows:

Definition II.5: If $(\mathscr{C}_i : i \in I)$ is an indexed family of categories with I any set, then we define the **product category** $\Pi_{i \in I} \mathscr{C}_i$ with the objects being indexed families of objects $(C_i : i \in I)$ and the morphisms being indexed families of of morphisms $(f_i : i \in I)$.

The only (apparent) issue in this construction is that we are treading close to foundational issues when the objects in these categories do not form a set. This can be addressed, but we will just refer to the discussions mentioned earlier, cf. page 5.

II.4 Sum of Categories

Besides products of categories there are also sums of categories. Here is the definition.

Definition II.6: If $(\mathscr{C}_i : i \in I)$ is an indexed family of categories with I any set, then we define a category $\Sigma_{i\in I}\mathscr{C}_i$ with the objects being pairs (C, i) with $i \in I$ and C an object of $\mathscr{C}_i$. There are no morphisms from (C, i) to (D, j) unless $i = j$ in which case every morphism $f : C \longrightarrow D$ in $\mathscr{C}_i$ gives a morphism $(f, i) : (C, i) \longrightarrow (D, i)$.

Exercise II.3. Verify that $\Sigma_{i\in I}\mathscr{C}_i$, as defined above, is indeed a category.

II.5 Concrete and Based Categories

The overwhelming majority of the familiar categories (see the Catalog of Categories (Appendix B)) come from the category of sets by defining the objects of the new category as sets with some additional structure (e.g., a binary operation [as with **Monoid**] or a family of subsets [as with **Top**]), while the new morphisms are defined to be functions between the sets which in some fashion preserve the structure.

Rather than trying to specify just what is meant by a structure we will abstract this situation to get the following (tentative) definition.

Definition II.7: A **concrete category** is a category $\mathscr{C}$ and a specification of the underlying sets and functions in **Set**. This last means that for each object C of $\mathscr{C}$ there is specified a set $U(C)$ and each morphism $f : C \longrightarrow \mathbf{D}$ of $\mathscr{C}$ there is a function $U(f) : U(C) \longrightarrow U(D)$. Moreover U has to satisfy some consistency conditions:

1. $U(1_C) = 1_{U(C)}$
2. $U(gf) = U(g)U(f)$
3. if $f, f' : C \longrightarrow D$ and $U(f) = U(f')$, then $f = f'$.

The image here is that $U(C)$ is the underlying set of the object, i.e., the set gotten by forgetting about the structure. Similarly $U(f)$ is the function that preserves the structure. The three consistency conditions say that the identity function preserves the structure, that composition of the morphisms in the new category is just composition of functions, and that the only morphisms in the new category are indeed just suitable functions.

This is labeled a tentative definition because as soon as we have the definition of a functor we will restate this definition and see that U is just a faithful functor from $\mathscr{C}$ to **Set**, called the underlying or forgetful functor (details in definition III.8 and Section III.2.6.)

Exercise II.4. There are many categories listed in the Catalog of Categories (Appendix B) that are quite naturally concrete categories. The first and simplest is the category of sets itself. There U is just the identity on objects and morphisms.
Decide which of the categories in the Appendix are naturally concrete categories and for those describe a forgetful functor. [This is a trivial exercise, particularly as the answer in each case is in the article for the given category. It is just an excuse to get you to review these categories and reflect on concrete categories. Not all of the categories listed are concrete categories, but proving that there is no forgetful functor can be non-trivial.]

This is also a special case of a more general phenomenon that we saw repeatedly in Section I.4. Much as with concrete categories, the various categories of algebraic objects were constructed from some base category by defining the objects of the new category as objects in the base category with some additional structure (e.g., a binary operation as with magmas in a category, $\mathbf{Magma}_{\mathscr{C}}$), while the new morphisms were defined to be morphisms in the base category which respected the structure.

Again rather than trying to specify just what is meant by a structure we will abstract this situation to get the following (tentative) definition.

Definition II.8: A **based category** on a base category $\mathscr{B}$ is a category $\mathscr{C}$ and a specification of the objects and morphisms in $\mathscr{B}$. This last means that for each object C of $\mathscr{C}$ there is specified an object $U(C)$ in $\mathscr{B}$ and each morphism $f : C \longrightarrow \mathsf{D}$ of $\mathscr{C}$ there is a morphism $U(f) : U(C) \longrightarrow U(D)$. Moreover U has to satisfy the consistency conditions:

1. $U(1_C) = 1_{U(C)}$
2. $U(gf) = U(g)U(f)$
3. if $f, f' : C \longrightarrow D$ and $U(f) = U(f')$, then $f = f'$.

This too is labeled a tentative definition because it also will be superseded as soon as we have the definition of a functor and can restate this definition as saying that U is just a faithful functor from $\mathscr{C}$ to $\mathscr{B}$ (details in definition III.8 and Section III.2.6.)

The next exercise is just looking at a very few of the multitude of examples of based categories that will naturally occur in these notes.

Exercise II.5. For each of the following situations, confirm that we do have a based category as asserted. [Again this exercise is trivial and just intended to get you to reflect on based categories.]

1. **Group** based on **Set**
2. **Group** based on **Monoid**
3. **Monoid** based on **Magma**
4. $\mathbf{Group}_{\mathscr{C}}$ based on $\mathscr{C}$
5. $\mathbf{Group}_{\mathscr{C}}$ based on $\mathbf{Monoid}_{\mathscr{C}}$
6. $\mathbf{Monoid}_{\mathscr{C}}$ based on $\mathbf{Magma}_{\mathscr{C}}$.
7. $\mathbf{Magma}_{\mathscr{C}}$ based on $\mathscr{C}$ (see Section I.4.1, in particular exercise I.83.)
8. **LieGroup** based on **Manifold**
9. $\mathbf{Module}_{\mathbf{R}}$ based on **Ab**

II.6 Morphism Categories

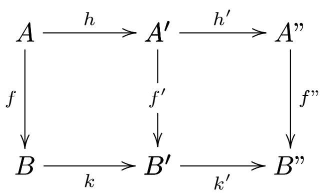
For any category there are a variety of associated categories where the *objects* of the new category are certain *morphisms* from the original category. Throughout this section let $\mathscr{C}$ be a fixed category.

- The first and simplest example is the case where we consider all the morphisms as the objects of our new category. The category $\mathscr{C}^{\mathbf{2}}$, called the **morphism category of** $\mathscr{C}$, has as objects the morphisms of $\mathscr{C}$, while a morphism in $\mathscr{C}^{\mathbf{2}}$ from $f : A \longrightarrow B$ to $f' : A' \longrightarrow B'$ is a pair (h, k) of $\mathscr{C}$-morphisms so that

$$\begin{array}{ccc} A & \xrightarrow{h} & A' \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f'} \\ B & \xrightarrow[k]{} & B' \end{array}$$

commutes.

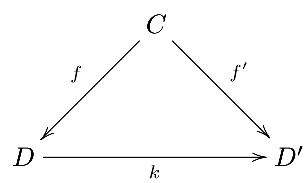
The identity morphism on f is the pair $(1_A, 1_B)$, while the composition of $(h, k) : f \longrightarrow f'$ with $(h', k') : f' \longrightarrow f$" is $(h'h, k'k)$ as in



That $\mathscr{C}^{\mathbf{2}}$ is a category follows immediately from the fact that $\mathscr{C}$ is a category.

The explanation for the notation $\mathscr{C}^{\mathbf{2}}$ and the connection with the arrow category **2** is in Section III.6.7 below.

- The next interesting case arises when we restrict our morphisms to have a fixed domain. For C a fixed object of $\mathscr{C}$, the **Category of objects under** C, written $(C \downarrow \mathscr{C})$ has as objects morphisms in $\mathscr{C}$ with domain C, while a morphism in $(C \downarrow \mathscr{C})$ is a $\mathscr{C}$-morphism between the codomains that makes the following triangle commute:

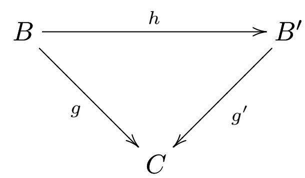
  So here k is a morphism from f to f' in $(C \downarrow \mathscr{C})$.

  $(C \downarrow \mathscr{C})$ is immediately seen to be the subcategory of $\mathscr{C}^{\mathbf{2}}$ consisting of just those "objects" with domain C and just the "morphisms" of the form $(1_C, k)$.

  Both objects under C and the following definition of objects over C are examples of the more general and useful construction of comma categories which will be discussed in Section IV.1 where the notation will be expanded and discussed.

  As an example, considering **Set**, the category of sets, and C to be a final object 1, i.e., any fixed one element set, we see that $(\mathbf{Set} \downarrow 1)$ is essentially the same as the category $\mathbf{Set}_*$ of pointed sets. (Cf. Section I.3.2, especially page 38, and Section B.1.7.) For a function from 1 to any set S is completely determined the point $s_0 \in S$ (the base point) that is the image of the function, and a morphism in $(\mathbf{Set} \downarrow 1)$ is exactly a function that takes base point to base point. Formally what we see is that the category $(\mathbf{Set} \downarrow 1)$ is isomorphic to the category $\mathbf{Set}_*$, and this will essentially be a proof of that fact as soon as we actually define "isomorphism of category" which we will do in definition III.6 on page 81.

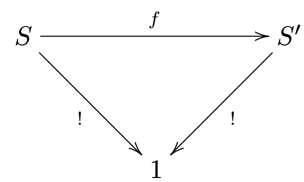
- The dual to the notion of the category of objects under C is of the **category of objects over** C, written $(\mathscr{C} \downarrow C)$. Here the objects are morphisms in $\mathscr{C}$ with codomain C, and a morphism is a $\mathscr{C}$-morphism between the codomains that making the triangle commute:

  Here h is a morphism (in $(\mathscr{C} \downarrow C)$) from g to g'.

  Again $(\mathscr{C} \downarrow C)$ is a subcategory of $\mathscr{C}^{\mathbf{2}}$ consisting of just those "objects" with codomain C and just the "morphisms" of the form $(h, 1_C)$.

  By contrast with $(\mathbf{Set} \downarrow 1)$, the category $(1 \downarrow \mathbf{Set})$ is really the same as **Set** itself. For any set S there is a unique function from S to 1, and for any function $f : S \longrightarrow S'$, the triangle

  commutes.

  This argument really is just using the fact that 1 is a final object in **Set**. The same argument shows that for any category with final object the category $(\mathscr{C} \downarrow 1)$ is isomorphic to $\mathscr{C}$.

  Dually for any category with initial object 0, the category $(0 \downarrow \mathscr{C})$ is isomorphic to $\mathscr{C}$.

Chapter III

Functors and Natural Transformations

III.1 What is a Functor?

A functor is a morphism from one category to another. Here is the actual definition.

Definition III.1: A **functor** F from $\mathscr{C}$ to $\mathscr{D}$ assigns to each object A in $\mathscr{C}$ an object $F(A)$ in $\mathscr{D}$ and to each morphism $f : A \longrightarrow B$ in $\mathscr{C}$ a morphism $F(f) : F(A) \longrightarrow F(B)$. Moreover $F(1_A) = 1_{F(A)}$ for every object in A in $\mathscr{C}$, and whenever $A \xrightarrow{f} B \xrightarrow{g} C$ in $\mathscr{C}$, $F(gf) = F(g)F(f)$ in $\mathscr{D}$.

The canonical examples of functors are $\mathscr{C}(C, \bullet) : \mathscr{C} \longrightarrow \mathbf{Set}$, the Hom functors.For each object A of $\mathscr{C}$ we have the set $\mathscr{C}(C, A)$ and for each morphism $f : A \longrightarrow B$ we have the function $f_* = \mathrm{Hom}(C, f) : \mathrm{Hom}(C, A) \longrightarrow \mathrm{Hom}(C, B)$ from Definition I.7. Clearly $(1_A)_*$ is the identity function on $\mathscr{C}(A, A)$, and $(gf)_* = g_* f_*$ is exactly the associative law in $\mathscr{C}$.

But we have another canonical example: $\mathscr{C}(\bullet, C) : \mathscr{C} \longrightarrow \mathbf{Set}$. For each object A of $\mathscr{C}$ we have the set $\mathscr{C}(A, C)$ and for each morphism $f : A \longrightarrow B$ we have the function $f^* = \mathrm{Hom}(f, C) : \mathrm{Hom}(B, C) \longrightarrow Hom(A, C)$ from Definition I.8. And this is *not* a functor because $(gf)^* = f^* g^*$ rather than $(gf)^* = g^* f^*$.

This is sufficiently important that it too gets a definition.

Definition III.2: A *contravariant functor* F from $\mathscr{C}$ to $\mathscr{D}$ assigns to each object A in $\mathscr{C}$ an object $F(A)$ in $\mathscr{D}$ and to each morphism $f : A \longrightarrow B$ in $\mathscr{C}$ a morphism $F(f) : F(B) \longrightarrow F(A)$. Moreover $F(1_A) = 1_{F(A)}$ for every object in A in $\mathscr{C}$, and whenever $A \xrightarrow{f} B \xrightarrow{g} C$ in $\mathscr{C}$, $F(gf) = F(f)F(g)$ in $\mathscr{D}$.

Sometimes for emphasis a functor is called a covariant functor. To confuse things a bit, a few authors have used the term *cofunctor.* for contravariant functor.

So we also have the contravariant Hom functors $\mathscr{C}(\bullet, C) : \mathscr{C} \longrightarrow \mathbf{Set}$.

As we would expect for "morphisms of categories", there are identity functors and composition of functors.

Definition III.3: The **identity functor** $1_{\mathscr{C}} : \mathscr{C} \longrightarrow \mathscr{C}$ has $1_{\mathscr{C}}(C) = C$ and $1_{\mathscr{C}}(f) = f$ for every object C and morphism f in $\mathscr{C}$. This is clearly a functor.

Definition III.4: For two functors $F : \mathscr{A} \longrightarrow \mathscr{B}$ and $G : \mathscr{B} \longrightarrow \mathscr{C}$, the **composition** $GF : \mathscr{A} \longrightarrow \mathscr{C}$ has $GF(A) = G(F(A))$ and $GF(f) = G(F(f))$ for every object A and morphism f in $\mathscr{A}$. And again it is clear that the composition is a functor.

The same definition applies equally well if either F, G or both are contravariant functors. Note that if just one is a contravariant functor, then the composition is a contravariant functor, while if both are covariant or both are contravariant, then the composition is covariant.

For every category we have the special, but unnamed, contravariant functor $\mathscr{C} \longrightarrow \mathscr{C}^{\mathrm{op}}$ which takes every object and every morphism to itself, but in the dual category. More, every contravariant functor $F : \mathscr{C} \longrightarrow \mathscr{D}$ uniquely factors into $\mathscr{C} \longrightarrow \mathscr{C}^{\mathrm{op}} \longrightarrow \mathscr{D}$ with the first functor here being the special contravariant functor and the second being a covariant functor. So every contravariant functor $F : \mathscr{C} \longrightarrow \mathscr{D}$ can be considered as a functor from $\mathscr{C}^{\mathrm{op}} \longrightarrow \mathscr{D}$ which we will also call F.

Of course it is also the case that every contravariant functor $F : \mathscr{C} \longrightarrow \mathscr{D}$ when composed with the special canonical contravariant functor from $\mathscr{D} \longrightarrow \mathscr{D}^{\mathrm{op}}$ gives a covariant functor from $\mathscr{C}$ to $\mathscr{D}^{\mathrm{op}}$. Again as no confusion should result, this composite will also be called F.

Just as we have functions of several variables, we naturally have functors of several variables as well. Our very first example is Hom : $\mathscr{C} \times \mathscr{C} \longrightarrow \mathbf{Set}$. We'll leave it as an exercise in choosing clear notation to verify that Hom is indeed a functor of two variables, or, as it is often called, a **bifunctor**. Here is the actual definition.

Definition III.5: A **bifunctor** from categories $\mathscr{C}$ and $\mathscr{B}$ to $\mathscr{D}$ is a functor from $\mathscr{C} \times \mathscr{D}$ to $\mathscr{D}$. A bifunctor from categories $\mathscr{C}$ and $\mathscr{B}$ to $\mathscr{D}$ which is contravariant in the first argument and covariant in the second is a functor from $\mathscr{C}^{\mathrm{op}} \times \mathscr{D}$ to $\mathscr{D}$.

There are similar definitions of a bifunctor which is covariant in the first and contravariant in the second arguments, and of a bifunctor which is contravariant in both arguments.

More generally a **multifunctor**from categories $\mathscr{C}_1, \ldots, \mathscr{C}_n$ to $\mathscr{D}$ which is covariant in some arguments and contravariant in the others is a functor from a suitable product of the categories corresponding to covariant arguments and the duals of the categories corresponding to the contravariant arguments to $\mathscr{D}$.

The full formal definition of a multifunctor would be long and tedious definition without providing additional insight, so we settle for this informal statement.

The most basic example of a bifunctor is Hom.

Exercise III.1. Verify that Hom is a bifunctor from $\mathscr{C}$ and $\mathscr{C}$ to **Set** which is contravariant in the first argument.

It is also possible to compose multifunctors in much the same way we can compose functions of multiple variables, but the actual complicated definition will not be spelled out but just used as needed.

Now that we have functors, composition of functors and identity functors, we immediately get the notion of an isomorphism of categories.

Definition III.6: A functor $F : \mathscr{A} \longrightarrow \mathscr{B}$ is an **isomorphism**of categories iff there is an inverse functor $F^{-1} : \mathscr{B} \longrightarrow \mathscr{A}$ so that $F^{-1}F = 1_{\mathscr{A}}$ and $FF^{-1} = 1_{\mathscr{B}}$.

Applications of category theory within particular categories systematically ignore the difference between isomorphic objects. Correspondingly two isomorphic categories are "the same" from the standpoint of category theory. But isomorphism of categories is more stringent than is truly interesting. There are many categories that are not isomorphic, perhaps because one has more isomorphic copies of some object than the other, but are truly equivalent. The interesting notion for functors is not that they be inverses, but only that they be "naturally equivalent", one of the original concepts introduced when Eilenberg and Mac Lane defined categories and functors. The definition and further discussion is in section III.5.

Before we get there we need to record a couple of other definitions, and see a small sampling of the diversity of functors in our universe.

Part of a functor, F, is a function

$$F : \mathrm{Hom}(A, B) \longrightarrow \mathrm{Hom}(F(A), F(B))$$

Definition III.7: A functor F is **full** iff the function between the Hom sets is surjective for all objects.

This is, of course, closely connected to the notion of a full subcategory (see definition I.12 and also the next section.)

Definition III.8: A functor F is **faithful** iff the function between the Hom sets is injective for all objects.

III.2 Examples of Functors

III.2.1 Subcategories and Inclusion Functors

Way back in Section I.1.2 we defined subcategories. Just as with other subobjects, there are inclusion morphisms associated with subcategories, but we weren't prepared to explain them because functors were not yet defined. Now we can state the obvious definition.

Definition III.9: For any subcategory $\mathscr{S}$ of $\mathscr{C}$ the inclusion functor$\mathscr{S} \longrightarrow \mathscr{C}$ takes each object and morphism in $\mathscr{S}$ into the same object and morphism in $\mathscr{C}$.

The definition of a subcategory is exactly what is needed to ensure that the inclusion functor is in fact a functor.

Every inclusion functor is faithful,and $\mathscr{S}$ is a full subcategory exactly if the inclusion functor is full.

III.2.2 Quotient Categories and Quotient Functors

Definition III.10: If $\mathscr{C}/\sim$ is the quotient category of $\mathscr{C}$ by the congruence $\sim$, there is the **quotient functor**which takes each object in $\mathscr{C}$ into the same object in $\mathscr{C}/\sim$ and each morphism into its equivalence class of morphisms.

The quotient functor is always full. It is only faithful when the quotient functor is an isomorphism, that is when $\sim$ is the equality equivalence relation.

III.2.3 Product of Categories and Projection Functors

As this is category theory this definition is quite unsound – where is the Universal Mapping Property? That lacunae is rectified with the following definition.

Definition III.11: The **projection functors** $\pi_j : \Pi_{i\in I}\mathscr{C}_i \longrightarrow \mathscr{C}_j$ are $\pi_j(C_i : i \in I) = C_j$ and $\pi_j(f_i : i \in I) = f_j$.

Now for any family of functors $F_i : \mathscr{C} \longrightarrow \mathscr{C}_i$ we have the functor $F : \mathscr{C} \longrightarrow \pi_j : \Pi_{i\in I}\mathscr{C}_i$ defined on objects by $F(C) = (F_i(C) : i \in I)$ and on morphism by $F(f) = (F_i(f) : i \in I)$. And the Universal Mapping Property , F is the unique functor such that $\pi_i F = F_i$ for all $i \in I$, is easily checked:

1. Each π_i is a functor. The initial part of that, π_i taking objects to objects and morphisms to morphisms is clear. The other parts are:

   a) That π_i takes identity morphisms to identity morphisms is the observation that the identity morphism on $(C_i : i \in I)$ is $(1_{C_i} : i \in I)$ and π_i applied to this gives 1_{C_i};

   b) That π_i preserves composition of morphisms is noting that $(g_i : i \in I)(f_i : i \in I) = (g_i f_i : i \in I)$ and $\pi_i((g_i : i \in I)(f_i : i \in I)) = g_i f_i$.

2. F is a functor. Again the initial part is clear. The rest is:

   a) F takes identity morphisms to identify morphisms: The identity morphism on $(C_i : i \in I)$ is $(1_{C_i} : i \in I)$ and F applied to this is $(F_i(1_{C_i}) : i \in I) = (1_{F_i(C_i)} : i \in I)$ (because each F_i is a functor.) And this is the identity on $(F_i(C_i) : i \in I)$ which is $F(C_i : i \in I)$.

   b) F preserves composition of morphism:

$$\begin{aligned} F((g_i : i \in I)(f_i : i \in I)) &= F(g_i f_i : i \in I) \\ &= (F_i(g_i f_i) : i \in I) \\ &= (F_i(g_i)F_i(f_i) : i \in I) \\ &= (F_i(g_i) : i \in I)(F_i(f_i) : i \in I) \\ &= F(g_i : i \in I)F(f_i : i \in I) \end{aligned}$$

   c) That F is the *unique* functor whose composition with the π_i is F_i is immediate from the definition of $\Pi_{i\in I}\mathscr{C}_i$ and the π_i.

▮

Exercise III.2. Give examples to show that projection functors need not be either faithful or full.
Remember this just means there is some situation where $\mathscr{C}(C_1, C_2) \times \mathscr{D}(D_1, D_2) \longrightarrow \mathscr{C}(C_1, C_2)$, $(f, g) \mapsto f$, is not injective, and some possibly different situation where this function is not surjective.

III.2.4 Sum of Categories and Injection Functors

And here is the appropriate Universal Mapping Property.

Definition III.12: The **injection functors** $\iota_j : \mathscr{C}_j \longrightarrow \Sigma_{i\in I}\mathscr{C}_i$ are $\iota_j(C) = (C, j)$ and $\iota_j(f) = (f, j)$.

Now for any family of functors $F_i : \mathscr{C} \longrightarrow \mathscr{C}_i$ we have the functor $F : \mathscr{C} \longrightarrow \pi_j : \Pi_{i\in I}\mathscr{C}_i$ defined on objects by $F(C) = (F_i(C) : i \in I)$ and on morphism by $F(f) = (F_i(f) : i \in I)$. And the Universal Mapping Property, that F is the unique functor such that $\pi_i F = F_i$ for all $i \in I$, is easily checked:

For any family of functors $F_i : \mathscr{C}_i \longrightarrow \mathscr{C}$ we have the functor $F : \Sigma_{i\in I}\mathscr{C}_i \longrightarrow \mathscr{C}$ defined on objects by $F(C, i) = F_i(C)$ and on morphisms $(f, i) : (C, i) \longrightarrow (D, i)$ is $F_i(f)$. The Universal Mapping Property, that F is the unique functor such that $F\iota_i = F_i$ for all $i \in I$ is easily checked:

1. Each ι_i is a functor. The initial part of that, ι_i taking objects to objects and morphisms to morphisms is clear. The other parts are:

a) That ι_i takes identity morphisms to identity morphisms for $\iota_i(1_C) = (1_C, i)$ which is the identity morphism on $\iota_i(C) = (C, i)$.

b) ι_i preserves composition of morphisms for the only compositions to be preserved are gf where g and f and in $\mathscr{C}_i$, and then $\iota_i(gf) = (gf, i) = (g, i)(f, i) = \iota_i(g)\iota_i(f)$.

2. F is a functor. Again the initial part is clear. Here is the rest.

a) F takes identity morphisms to identify morphisms: The identity morphism on $(C,\ i)$ is 1_C and F applied to this is $(F_i(1_C), i) = (1_{F_i(C)}, i)$ (because each F_i is a functor.) And this is the identity on $F(C, i)$.

b) F preserves composition of morphism:

$$\begin{aligned} F((g,i)(f,i)) &= F(gf,i) \\ &= F_i(gf) \\ &= F_i(g)F_i(f) \\ &= F(g,i)F(f,i) \end{aligned}$$

c) That F is the *unique* functor with $F\iota_i = F_i$ for all i in I is immediate from the definition of $\Sigma_{i\in I}\mathscr{C}_i$ and the ι_i.

▮

III.2.5 Constant Functors

The very simplest functors are the constant functors.

Definition III.13: Let D be any object in the category $\mathscr{D}$. Then for any category $\mathscr{C}$ the **constant functor** from $\mathscr{C}$ to $\mathscr{D}$determined by D is the functor of the same name $D : \mathscr{C} \longrightarrow \mathscr{D}$ where $D(C) = D$ and $D(f) = 1_D$ for all objects and morphisms in $\mathscr{C}$.

When $S : \mathscr{C} \longrightarrow \mathscr{D}$ is a constant functor there is a unique object D in $\mathscr{D}$ with $S = D$ and we say that D is the object *selected* by S.

There are constant contravariant functors as well, indeed every constant functor is both covariant and contravariant!

Constant functors are so trivial that they may not seem worthy of consideration, but in Sections IV.1 and V.1 we will see they are quite useful.

Recall from Section I.1.3 the category **1** with just one object and one morphism. Constant functors are characterized by the fact that they factor through **1**, i.e., $F : \mathscr{C} \longrightarrow \mathscr{D}$ is constant iff there is a commutative triangle

$$\begin{array}{ccccc} & & \mathbf{1} & & \\ & \overset{!}{\nearrow} & & \overset{D}{\searrow} & \\ \mathscr{C} & & \xrightarrow[F]{} & & \mathscr{D} \end{array}$$

with D some object in $\mathscr{D}$ (the one selected by F.)

III.2.6 Forgetful Functors

Back in Section II.5 when discussing Concrete and Based Categories are tentative definitions of concrete and based categories. We can now give the final definitions.

Definition III.14: A **based category** on a base category $\mathscr{B}$ is a category $\mathscr{C}$ together with a faithful functor $U : \mathscr{C} \longrightarrow \mathscr{B}$.

Definition III.15: A **concrete category** is a category based on **Set**.

In all cases the functor U to the base category is called the **underlying** or **forgetful** functor. The commonly used phrase is "forgetful functor" rather than "underlying functor" which would suggest using the letter F rather than the nearly universal use of the letter U. The reason for using U even when calling it the forgetful functor is the close association of forgetful functors with free functors (cf. Sections III.2.12, III.4, V.4.1) for which the letter F will be reserved.

If $\mathscr{A}$ is a subcategory of $\mathscr{B}$, then the inclusion functor $\mathscr{A} \hookrightarrow \mathscr{B}$ (see Section III.2.1) is faithful and so $\mathscr{A}$ together with the inclusion functor exhibits it as a category based on $\mathscr{B}$. While most inclusion functors do not fit our intuitive notion of a forgetful functor, there are examples that do – **Ab** $\hookrightarrow$ **Group**, just forget that an Abelian group is commutative; **Group** $\hookrightarrow$ **Monoid**, just forget that a group has inverses; **Monoid** $\hookrightarrow$ **Semigroup**, just forget that a monoid has an identity element; **Semigroup** $\hookrightarrow$ **Magma**, just forget that a semigroup is associative – are all examples, as are all the of possible compositions.

III.2.7 The Product Functor

If $\mathscr{C}$ is a category with finite products, fix an object C of $\mathscr{C}$ and consider $\times C : \mathscr{C} \longrightarrow \mathscr{C}$ where for each object A we have $\times C(A) = A \times C$ and for each morphism $f : A \longrightarrow B$ we have $\times C(f) = f \times 1_C$.

Exercise III.3. Show that $\times C$ as defined in the previous paragraph is a functor.

Duality applies here to give a sum functor with everything left as an exercise.

Exercise III.4. For $\mathscr{C}$ a category with finite sums and C an object of $\mathscr{C}$, show there is a functor $+C : \mathscr{C} \longrightarrow \mathscr{C}$ which takes each object to its sum with C.

III.2.8 The Sum Bifunctor

If $\mathscr{C}$ is a category with finite sums, define $+ : \mathscr{C} \times \mathscr{C} \longrightarrow \mathscr{C}$ by $+(C_1, C_2) = C_1 + C_2$ and $+([f_1, f_2]) = f_1 + f_2$.

Exercise III.5. Confirm that + is a functor.

Duality applies here to give a product bifunctor with everything left as an exercise.

Exercise III.6. For $\mathscr{C}$ a category with finite products, show there is a bifunctor from $\mathscr{C}$ to itself which takes each pair of objects and each pair of morphisms to their respective products.

III.2.9 Power Set Functor

In the category of sets, let $\mathcal{P}(X)$ denote the set of all subsets of X (cf. 185.) And for any function $f : X \longrightarrow Y$ and any $S \subseteq X$ define $\mathcal{P}(f)$ to be the *direct image* of f.

$$\mathcal{P}(f)(S) = f(S) = \{y | \exists x \in S , y = f(x)\}$$

$\mathcal{P}(X)$ is called the **power set** of X.

Notation: As mentioned in the Introduction, we will occasionally, as above, use the two symbols $\exists$ and $\forall$ from mathematical logic. The symbol $\exists$ is shorthand for "there exists", while $\forall$ is used in place of "for all".

Exercise III.7. Verify that $\mathcal{P}$: **Set** $\longrightarrow$ **Set** is a functor.

The power set actually gives rise to a contravariant functor as well. Let $P(X)$ denote the set of all subsets of X, but for any function $f : X \longrightarrow Y$ and any $T \subseteq Y$ define $P(f)(T) = f^{-1}(T) = \{x | f(x) \in T\}$,

Exercise III.8. Verify that P : **Set** $\longrightarrow$ **Set** is a contravariant functor.

III.2.10 Monoid Homomorphisms are Functors

Recall (see p. 5) that a monoid can be considered as a small category with one object. If M and N are two monoids, then a monoid homomorphism $h : M \longrightarrow N$ takes the identity in M to the identity in N and satisfies $h(mm') = h(m)h(m')$. But that exactly says that h "is" a functor between these two one object categories.

III.2.11 Forgetful Functor for Monoid

Consider the category $\mathbf{Monoid}_{\mathscr{C}}$ of monoids in category $\mathscr{C}$ andrecall that a monoid (M, μ, ζ) in $\mathscr{C}$ consists of an object, M, a binary operation, μ and an identity, ζ. Define $U : \mathbf{Monoid}_{\mathscr{C}} \longrightarrow \mathscr{C}$ by $U(M, \mu, \zeta) = M$ and $U(f) = f$ for f any monoid morphism.

Exercise III.9. Verify that U is a functor.

U is called the **underlying** or **forgetful** functor.

There are obvious variations on this for categories of magmas, categories of Abelian monoids, categories of groups and categories of Abelian groups. You should also see immediately how to define $U : \mathbf{Group}_{\mathscr{C}} \longrightarrow \mathbf{Monoid}_c atC$ and $U : \mathbf{Monoid}_{\mathscr{C}} \longrightarrow \mathbf{Magma}_{\mathscr{C}}$. Indeed when we consider the category of sets, there are also obvious forgetful functors from the categories of rings, topological spaces, posets, etc.

III.2.12 Free Monoid Functor

Closely related to the forgetful functor $U : \mathbf{Monoid} \longrightarrow \mathbf{Set}$ is the functor $F : \mathbf{Set} \longrightarrow \mathbf{Monoid}$ defined as follows: $F(A) = A^*$,the monoid of strings on the alphabet A. A^* consists of all finite sequences of elements of A, including the empty sequence. (See exercise I.14 on page 15.) The binary operation in A^* is concatenation – if $(a_1, a_2, \cdots, a_n)$ and $(b_1, \cdots, b_m)$ are two such strings, their product is $(a_1, a_2, \cdots, a_n, b_1, \cdots, b_m)$. This product is clearly associative, and the empty string is the identity.

For a function $f : A \longrightarrow B$, define $F(f) : A^* \longrightarrow B^*$ by $F(f)(a_1, a_2, \cdots, a_n) = f(a_1)f(a_2)\cdots f(a_n)$ [NOTE: Writing $g(x_1, x_2, \cdots, x_n)$ rather than $g((x_1, x_2, \cdots, x_n))$ is an extremely common abuse of notation that we will happily adopt without further comment.]

F as so defined is clearly a functor.

Now suppose that $f : A \longrightarrow U(M)$ is any function from a set A to (the underlying set of) a monoid M. Then we can define a monoid homomorphism $f^* : A^* \longrightarrow M$ by $f^*() = 1$ and $f^*(a_1, a_2, \cdots, a_n) = f(a_1)f(a_2)\cdots f(a_n)$. [Note that () denotes the empty sequence and $(a_1, a_2, \cdots, a_n)$ is a sequence of n elements, while $f(a_1)f(a_2)\cdots f(a_n)$ is the product of n elements in the monoid M.]

Exercise III.10. Verify that f^* is indeed a monoid homomorphism.

Clearly f^* is the *unique* monoid homomorphism extending f, i.e., if $h : A^* \longrightarrow M$ is a monoid homomorphism such that $h(a) = f(a)$ for every $a \in A$, then $h = f^*$.

This is called the *Universal Mapping Property* for the **free monoid**A^* generated by A. This is just one of a multitude of "free" gadgets that we will encounter and systematize in Section V.4.1.

Another way of stating this result is that we have a function

$$\mathbf{Set}(A, U(M)) \longrightarrow \mathbf{Monoid}(A^*, M)$$

which is a bijection.

Exercise III.11. Show that this function is in fact an isomorphism (i.e., a bijection.)

We will see that this is an example, and a very important one, where U and F are a pair of adjoint functors. (See section V.6, page 136)

III.2.13 Polynomial Ring as Functor

Consider the category, **CommutativeRing**, of commutative rings with identity. If R is a commutative ring, then we get a new commutative ring $R[X]$, the polynomial ring in one variable with coefficients in R. Moreover if $h : R \longrightarrow S$ is a homomorphism in **CommutativeRing**, then h can be extended to a homomorphism, H, from $R[X]$ to $S[X]$ by defining $H(\Sigma_{i=1}^n a_i X^i) = \Sigma_{i=1}^n h(a_i) X^i$. This allows us to define a functor F : **CommutativeRing** $\longrightarrow$ **CommutativeRing** by $F(R) = R[X]$ and $F(h) = H$.

Exercise III.12. Verify that F is indeed a functor.

III.2.14 Commutator Functor

For any group G, the commutatorof two elements $g_1, g_2 \in G$ is $[g_1, g_2] = g_1 g_2 g_1^{-1} g_2^{-1}$. The commutator, $[G, G]$, of G is $\{[g_1, g_2] | g_1, g_2 \in G\}$. Moreover if $f : G \longrightarrow H$ is any group homomorphism, then $f([G, G]) \subseteq [H, H]$, so $f|[G, G] : [G, G] \longrightarrow [H, H]$ is a group homomorphism. (For more information, see Mac Lane and Birkhoff [55, III.10,XII.4].)

Define C : **Group** $\longrightarrow$ **Group** by $C(G) = [G, G]$ and $C(f) = f|[G, G]$ for any group G and any group homomorphism $f : G \longrightarrow H$.

Exercise III.13. Prove that C as above is in fact a functor.

III.2.15 Abelianizer: Groups Made Abelian

Continuing where with the material in the previous section, $[G, G]$ is a normal subgroup of G, and $G/[G, G]$ is an Abelian group. As noted, if $f : G \longrightarrow H$ is any group homomorphism, then $f([G, G]) \subseteq [H, H]$, so there is a unique homomorphism $\overline{f} : G/[G, G] \longrightarrow H/[H, H]$ with $p_H f = \overline{f} p_G$ where p_G and

p_H are the projection homomorphisms to $G/[G,G]$ and $H/[H,H]$ respectively. (For more information, see Mac Lane and Birkhoff [55, III.10, XII.4].)

Define $A : \mathbf{Group} \longrightarrow \mathbf{Ab}$ as follows: for each group G, $A(G)$ is $G/[G,G]$, while if $f : G \longrightarrow H$ is a group homomorphism, then $A(f) = \overline{f}$.

Exercise III.14. Prove that A as above is in fact a functor.

The functor A is called the **Abelianizer**, and $A(G) = G/[G,G]$ is often called *"G made Abelian."*

Next let $I : \mathbf{Ab} \longrightarrow \mathbf{Group}$ denote the inclusion functor of the subcategory of Abelian groups into the category of groups. If $f : G \longrightarrow I(H)$ is a group homomorphism into an Abelian group H, then $h([G,G]) \subseteq [H,H] = 0$ as H is commutative. So $\overline{f} : G/[G,G] \longrightarrow H/0 \cong H$. Thus we have a function from $\mathbf{Group}(G, I(H)) \longrightarrow \mathbf{Ab}(A(G), H)$.

Exercise III.15. Show that this function is in fact an isomorphism.

This is another example where A and I are a pair of adjoint functors. (See Section V.6.) We'll see that adjoint functorsare very common indeed.

III.2.16 Discrete Topological Space Functor

Recall that a topological space consists of a set X, and a topology $\mathsf{T} \subseteq \mathcal{P}(X)$, whose elements are called the open sets of the topology, satisfying certain axioms. (For details see sections B.9.3 and B.19.5.) In particular $SP(X)$ is a topology on X, called the **discrete topology**.

The category **Top** has as objects topological spaces and as morphism continuous functions, i.e., functions for which the inverse of every open set is an open set. If $(X, \mathcal{P}(X))$ is a discrete topological space, then *all* functions from X to any other topological space are continuous. So if we define $F : \mathbf{Set} \longrightarrow \mathbf{Top}$ by $F(X) = (X, \mathcal{P}(X))$ and $F(f) = f$, then F is clearly a functor. Moreover $\mathbf{Top}(F(X), Y) \cong \mathbf{Set}(X, U(Y))$ where $U : \mathbf{Top} \longrightarrow \mathbf{Set}$ is the forgetful functor. And this is yet another example of an adjoint pair of functors.

Exercise III.16. Verify all of the details in the above two paragraphs.

The last few examples are going to be lacking crucial details, and are included just to indicate a very few of the great many advanced topics where functors arise.

III.2.17 The Lie Algebra of a Lie Group

To every Lie group,i.e., a group object in the category of smooth manifolds, we associate a Lie algebraas follows. (Details will be found in most modern books on differential geometry, for example Lang's Differential and Riemannian Manifolds [40],though the way we introduce the Lie product is different but equivalent.)

As with any manifold,there is at each point $g \in G$ a tangent space which is a vector space ofthe same dimension as the group which we will denote by T_g. For each element $g \in G$ the function $L_g : G \longrightarrow G$ given by $L_g(x) = gx$ is a smooth function with inverse $L_{g^{-1}}$, and so the differential of L_g is an isomorphism between T_e, the tangent space at the identity, and T_g.

Considering T_e, the commutatorfunction $G \times G \longrightarrow G$ given by $(g,h) \mapsto ghg^{-1}h^{-1}$sends (e,e) to e and so induces a bilinear function $T_e \times T_e \longrightarrow T_e$ which we denote by $[u,v]$. This makes T_e into a Lie algebra, which is usually denoted by $\mathfrak{g}$. Beyond that if $f : G \longrightarrow H$ is any Lie group homomorphism, then of course $f(e) = e$ and $df : T_e(G) \longrightarrow T_e(H)$ is not only a linear transformation, it is also a Lie algebra homomorphism as well. The result is that we have the basic functor **LieGroup** $\longrightarrow$ **LieAlgebra** which is fundamental in the study of Lie groups.

III.2.18 Homology Theory

As we mentioned back in the introduction a key impetus for the introduction of categories and functors was the realization that a good description of the homology and cohomology groups requires them. In particular Eilenberg and Steenrod [22] define a homology theory by giving axioms for a sequence of functors from an "admissible" topological category to the category of Abelian groups. The admissible categories consists of certain pairs (X, A) with $A \subseteq X$ of topological spaces, while the morphisms $f : (X, A) \longrightarrow (Y, B)$ (homotopy classes) are continuous functions $f : X \longrightarrow Y$ with $f(A) \subseteq B$. There is more, involving natural transformations (topic of the next section) and exact sequences which we will finally discuss in Chapter XIV.

III.3 Categories of Categories

We now have the makings of a category: objects, namely categories, and morphisms, namely functors, but as we noted near the beginning (cf. page 5) there is no category of all categories, for much the same reason there is no set of all sets. We can make the following definition.

Definition III.16: Cat, the **category of small categories**, has as objects all small categories, and as morphisms all functors between small categories.

Recall that monoids can be regarded as small categories with one object (cf. page 5), and as we note below monoid homomorphisms are exactly functors between such categories. This tells us that **Monoid**, the category of monoids, is a subcategory of **Cat**. A better way of saying this is that we have a functor **Monoid** $\longrightarrow$ **Cat** which takes each monoid to the corresponding category with a single element, and each monoid homomorphism to the corresponding functor. This functor is faithful, and produces an isomorphism of categories between **Monoid** and the full subcategory of all small categories with a single object.

The use of monoids was essentially arbitrary. For any of the other structures which can be regarded as categories and the morphisms between them as functors we have the same situation – a faithful functor from the category of structures into **Cat** which gives an isomorphism with a full subcategory of **Cat**. In particular **Set** "is" a subcategory of **Cat**, the full subcategory of all small discrete categories.

For more details look at the examples in Section B.19 in the Catalog of Categories (Appendix B).

The discussion of product of categories back in Section II.3 shows that **Cat** has products and we include that proposition here for the record.

Proposition III.9 *The category* **Cat** *of small categories has products.*

The result just proved is actually stronger. It says that products (and not just finite products) of arbitrary families of small categories have the Universal Mapping Property for products.

In the same way, the discussion of sum of categories back in Section II.4 shows that **Cat** has sums. Which we formally note here.

Proposition III.10 *The category* **Cat** *of small categories has sums.*

The result just proved is actually stronger. It says that sums (and not just finite sums) of arbitrary families of small categories have the Universal Mapping Property for sums.

The best known difficulty in naive set theory is Russell's Paradox: If there were a set of **all** sets, Ω, then there would be a set of all ordinary sets $\mathbf{O} = \{S | S \in \Omega \& S \notin S\}$. But this is not possible as $\mathbf{O} \notin \mathbf{O}$ implies $\mathbf{O} \in \mathbf{O}$ and $\mathbf{O} \in \mathbf{O}$ implies $\mathbf{O} \notin \mathbf{O}$. So from this we conclude that Ω cannot exist.

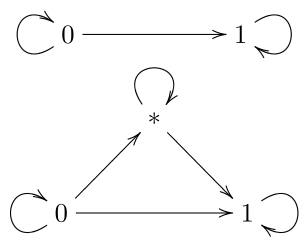
There is an analogous argument that can be made about a category of all categories: If there were a category of **all** categories, $\mathscr{C}$, then consider the full subcategory $\mathscr{O}$ of all ordinary categories where an ordinary category is one which is **not** an object in itself. Just as in Russell's Paradox, the category $\mathscr{O}$ is ordinary iff it is not ordinary. So we conclude that $\mathscr{O}$ cannot exist and so $\mathscr{C}$ cannot exist.

Reflection on this analogy reveals a significant issue. The most fundamental principle of set theory is the Axiom of Specification: If A is any set and $P(x)$ is a boolean predicate on the elements of A (i.e., $P(x)$ is either true or false for every element of A), then there exists a set B whose elements are exactly those elements x of A for which $P(x)$ is true. (For more detail see section A.2 in the

Appendix on Set Theory. Also read the discussion in Halmos' Naive Set Theory [29, Section 2].) By contrast it is not clear we can select a full subcategory from the objects of an arbitrary category using an arbitrary boolean predicate. Tacitly we have been freely assuming that we can do this, and we will continue to do so. It appears that we can always find a way to deal with this issue using a suitable choice for our foundations, but in these notes we have been and will continue to blithely ignore all these foundational details.

III.4 Digraphs and Free Categories

Back in Section I.1.3 we considered two small categories by looking at the following diagrams.



But these are also two just two small digraphs (see Section B.8.2.) There is a close connection between digraphs and categories that will be explored here. Indeed one approach to defining categories is to start with directed graphs as precursors and define categories as graphs with additional structure – see Barr and Wells [3, Secs. 1.3, 1.4, 2.1] and Mac Lane [53, Sec. II.7]

From Section B.8.2 in the Catalog of Categories we recall the two basic definitions.

Definition III.17: A **directed graph** or **digraph**G consists of a pair (V, E) of sets together with a two functions init:$E \longrightarrow V$ and ter:$E \longrightarrow V$.

The elements of V are the **vertices** or **nodes** of G, while the elements of E are the **edges** or **arcs**. The vertex init(e) is the **initial vertex** of the edge, and ter(e) is the terminal vertex. The edge e is said to be directed from init(e) to ter(e).

Definition III.18: Let G and H be digraphs. A **homomorphism** from G to H, $f : G \longrightarrow H$ is a pair of functions $f_V : V(G) \longrightarrow V(H)$ and $f_E : E(G) \longrightarrow E(H)$ $\text{init}(f_E(e)) = f_V(\text{init}(e))$ and $\text{ter}(f_E(e)) = f_V(\text{ter}(e))$.

And with them we also have the category **Digraph** of digraphs

This is sufficiently reminiscent of the domain/codomain aspect of categories that it it is no surprise to have the forgetful functor from U : **Cat** $\longrightarrow$ **Digraph** that just forgets the identities and composition of a category. In more detail, U sends each small category $\mathscr{C}$ to the digraph with vertices the objects of $\mathscr{C}$, edges the morphisms of $\mathscr{C}$, init(e) the domain of e, and ter(e)

the codomain of e. Each functor F takes objects to objects and morphisms to morphisms, thereby producing a digraph homomorphism.

Definition III.19: The graph $U(\mathscr{C})$ is called the **(underlying) graph of the category** $\mathscr{C}$.

Of course for consistency this should be called the *digraph of the category*, but the phrase "graph of the category" is the one commonly used in the category theory literature. The reason for the discrepancy is that there are many varieties of graphs and usually the most common type under discussion is blessed with the name "graph", while other get various adjectives attached. So many books on graph theory would use the phrase "directed multigraph with loops" for what is here called a digraph, while books on category theory such as Barr and Wells [3] and MacLane [53] use just "graph", but don't discuss other types of graphs. Here both digraphs and ordinary graphs will be considered, whence both names.

The view of categories by diagrams of arrows as at the beginning of this section or in Section I.1.3 is actually just looking at the graph of the category. In particular we saw that there are some digraphs which are uniquely the graph of a category, while there are other digraphs which are not the graph of any category.

There is no composition of edges in a digraph, but there is a closely related notion in graph theory.

Definition III.20: A **path** in a digraph G is a finite sequence $(e_1, \cdots, e_n)$ of edges with $\text{ter}(e_j) = \text{init}(e_{j+1})$ for $j = 1, \ldots, n-1$. This is called a path of length n from $\text{init}(e_1)$ to $\text{ter}(e_n)$. In addition, for each vertex v there is a path of length 0 from v to itself. For convenience we will write this as 1_v.

And paths lead from the forgetful functor $U : \mathbf{Cat} \longrightarrow \mathbf{Digraph}$ to the free category functor $F : \mathbf{Digraph} \longrightarrow \mathbf{Cat}$ defined as follows: For each digraph G, $F(G)$ is the category with objects the vertices of G, while the morphisms are are all the paths in G. The identity morphism on the vertex v is the zero length path 1_v introduced above. If $p = (e_1, \cdots, e_n)$ is a path from $v_0(= \text{init}(e_1))$ to $v_n(= \text{ter}(e_n))$, then we define the domain of p to be v_0 and the codomain to be v_n. In particular the domain and codomain of 1_v is v. Composition is concatenation of sequences, i.e., if p is a path from v_0 to v_n and $q = (e_{n+1}, \cdots, e_m)$ is a path from v_n to v_{n+m}, then $qp = (e_1, \cdots, e_n, e_{n+1}, \cdots, e_m)$. Clearly composition is associative and the zero length paths are identities for composition, so $F(G)$ is a category as desired.

Now if $f : G \longrightarrow H$ is a digraph homomorphism, then $F(f)$ must be a *functor* from $F(G)$ to $F(H)$. It is defined to be f_V on the objects, i.e., the vertices, while on the morphisms, i.e., the paths, $F(f)(1_v)$ is defined to be $1_{f_V(v)}$ while for $p = (e_1, \cdots, e_n)$ a path from v_0 to v_n we define $F(f)(p) tobethepath (f_E(e_1), \cdots, f_E(e_n))$. The definition of a digraph homomorphism ensures that $F(f)(p)$ is indeed a path that goes from $F(f)(v_0)$ to $F(f)(v_n)$. Clearly $F(f)(qp) = F(f)(q)F(f)(p)$, so $F(f)$ is a functor as claimed.

For any digraph homomorphism $f : G \longrightarrow U(\mathscr{C})$ there is an associated functor $\hat{f} : F(G) \longrightarrow \mathscr{C}$. On objects, i.e., vertices of G, it is $\hat{f}(v) = f_V(v)$, while on morphisms, i.e., paths in G, it is $\hat{f}(e_1, \cdots, e_n) = f_E(e_n) \cdots f_E(e_1)$ where this last is composition of morphisms in $\mathscr{C}$.

Exercise III.17. Verify that $\hat{f}$ is indeed a functor.

Clearly f^* is the *unique* functor extending f, i.e., if $h : F(G) \longrightarrow \mathscr{C}$ is any functor with $h(v) = f_V(v)$ for every vertex and $h(e) = f_E(e)$ for every edge, then $h = \hat{f}$.

This is very similar to the Universal Mapping Property for the free monoid generated by a set (see Section III.2.12), and leads to the following definition.

Definition III.21: A **free category** generated by a digraph G is a category $\mathscr{F}$ together with a digraph homomorphism $i : G \longrightarrow U(\mathscr{F})$ with the following Universal Mapping Property: For any digraph homomorphism $f : G \longrightarrow U(\mathscr{C})$ there is a unique functor $\hat{f} : \mathscr{F} \longrightarrow \mathscr{C}$ such that $U(\hat{f})i = f$.

The definition of $F(G)$ given above is the explicit construction of a free category. To see that we need to define $i : G \longrightarrow UF(G)$ which is just $i_V(v) = v$ and $i_E(e) = (e)$, i.e., each edge goes into a path of length one. And the above discussion gives us the following proposition.

Proposition III.11 *For each digraph G, the category $F(G)$ together with the digraph homomorphism $i : G \longrightarrow UF(G)$ defined above is a free category generated by G.*

∎

This is yet another of the multitude of "free" gadgets that we will encounter and systematize in Section V.4.1.

Just as with free monoids, another way of stating this result is that we have a function $\mathbf{Digraph}(G, U(\mathscr{C})) \longrightarrow \mathbf{Cat}(F(G), \mathscr{C})$ which is a bijection.

Exercise III.18. Show that this function is in fact an isomorphism.)

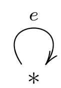
We will see shortly that this is another example where U and F are a pair of adjoint functors. (See section V.6.)

Here are some simple examples of digraphs and the associated free category. First note that there is an empty digraph with no vertices and no edges. The associated free category in this case is the empty category **0**. In this case the graph of the category **0** is just the empty graph as well.

Next consider the digraph with one vertex and no edges. The associated free category is the category **1** with the one object and just the single identity morphism on that object.

Slightly more generally, there is the discrete digraph with an arbitrary set V of vertices, but no edges. Then the associated free category is the category with V as the set of objects and with no morphisms other than the identity morphism for each object, i.e., the discrete categoryassociated to the set V (cf. Section B.19.1 in the Catalog of Categories.)

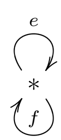
Now look at the digraph



with one vertex and one edge from $*$ to itself (called a loop in the digraph.) This is the graph of the category **1** which we just saw is the free category on the digraph with one vertex and no edges. In this case the free category on this digraph, which we will call $\mathscr{F}$, has, of course, just the one object $*$, but an infinity of morphisms, namely $1, (e), (e, e), (e, e, e), \cdots$. For convenience we will rewrite these as $1, e, e^2, e^3, \cdots$. Of course the exponent is the number of terms in the composition. On reflection this is seen to be the free monoid on one generator (considered as a category with one object.)

Note the large difference between the free category generated by that digraph and the category **1** that produced the digraph. This free category is actually $F(U(\mathbf{1}))$, i.e., the free category generated by the underlying graph of **1**. Exercise III.18 shows there is a bijection $\mathbf{Digraph}(G, U(\mathscr{C})) \cong \mathbf{Cat}(F(G), \mathscr{C})$ for any digraph G and category $\mathscr{C}$. Taking $\mathscr{C} = \mathbf{1}$ and $G = U(\mathbf{1})$ we see there is a distinguished functor $\varepsilon : F(U(\mathbf{1})) \longrightarrow \mathbf{1}$ corresponding to the identity homomorphism on $U(\mathbf{1})$. In this particular case ε is clear: it takes the unique object in $\mathscr{F}$ into the unique object in **1** and takes all the morphisms in $\mathscr{F}$ into the unique morphism in **1**.

There is an easy generalization which is worth doing in stages. Look at the digraph



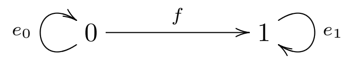
with one vertex and two distinct loops on that edge. In this case the free category has a much greater variety of morphisms: $1, (e), (f), (e, e), (e, f), (f, e), (f, f), (e, e, e), (e, e, f), \cdots$, but this is just the free monoid generated by the set $\{e, f\}$ (cf. Section III.2.12.) More generally we have the following proposition.

Proposition III.12 *The free category generated by the digraph having just one vertex, v, and as edges a set E of loops is the free monoid generated by E (considered as a category with one object.)*

∎

There is another way of viewing this which is helpful. Recall that in any category $\mathscr{C}$ we always have the monoid of endomorphisms $\mathscr{C}(C, C)$ for each object C. For the free category $\mathscr{F}$ generated by the digraph having just one vertex, v, and a set E of loops on that vertex we see that $\mathscr{F}(v,v)$ is the free monoid generated by E.

Next look at the following digraph.



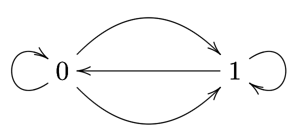
which is the graph of the category **2** (cf. Sections I.1.3 and B.20.3.)

Writing $\mathscr{F}$ for the associated free category, we immediately see from the description of the construction of a free category that $\mathscr{F}(0,0)$ is the free monoid generated by e_0 (i.e., the elements of $\mathscr{F}(0,0)$ are the morphisms e_0^m for some natural number m), while $\mathscr{F}(1,1)$ is the free monoid generated by e_1, $\mathscr{F}(1,0)$ is empty, and the morphisms in $\mathscr{F}(0,1)$ all have the form $e_0^m f e_1^n$ for natural numbers m and n.

Here $\mathscr{F}$ is $F(U(\mathbf{2}))$, the free category generated by the underlying graph of **2**. Again exercise III.18 shows there is a bijection $\mathbf{Digraph}(U(\mathbf{2}),U(\mathbf{2})) \cong \mathbf{Cat}(F(U(\mathbf{2})),\mathbf{2})$, so there is a distinguished functor $\varepsilon : F(U(\mathbf{2})) \longrightarrow \mathbf{2}$ corresponding to the identity homomorphism on $U(\mathbf{2})$. The next exercise asks you to describe this functor.

Exercise III.19. Describe the canonical functor $\varepsilon : F(U(\mathbf{2})) \longrightarrow \mathbf{2}$ in detail.

As a last example, in exercise I.2 the following diagram



was seen, in the terminology of this section, not to be the graph of any category.

Exercise III.20. Describe in detail the free category generated by the above digraph.

III.5 Natural Transformations

> As Eilenberg-Mac Lane first observed, "category" has been defined in order to define "functor" and "functor" has been defined in order to define "natural transformation".
> *Saunders Mac Lane* [53, p. 18]

Definition III.22: For two functors $F, G : \mathscr{A} \longrightarrow \mathscr{B}$. A **natural transformation** $\tau : F \longrightarrow G$, assigns to each object A of $\mathscr{A}$ a morphism $\tau_A : F(A) \longrightarrow G(A)$

in $\mathscr{B}$ such that the diagrams

$$\begin{CD} F(A) @>{\tau_A}>> G(A) \\ @V{F(f)}VV @VV{G(f)}V \\ F(B) @>>{\tau_B}> G(B) \end{CD}$$

A natural transformation is more naturally called a **morphism of functors** or a **functor morphism**. We will use all of these on various occasions. As the subscript on the name of a natural transformation, e.g., the A on τ_A, is usually redundant, we will often omit it.

Natural transformations compose: If $F, G, H : \mathscr{A} \longrightarrow \mathscr{B}$ are functors with $\tau : F \longrightarrow G$ and $\sigma : G \longrightarrow H$ natural transformations, then $\sigma\tau : F \longrightarrow H$ has $(\sigma\tau)_A = \sigma_A \tau_A$. And for any functor there is, of course, an identity morphism $1_F : F \longrightarrow F$ with $1_{FA} = 1_{F(A)}$ and $\tau 1_G = \tau$, $1_G \sigma = \sigma$.

Definition III.23: A natural transformation $\tau : F \longrightarrow G$ is a **natural equivalence** or **natural isomorphism** iff there is a natural transformation $\tau^{-1} : G \longrightarrow F$ with $\tau\tau^{-1} = 1_G$ and $\tau^{-1}\tau = 1_F$.

It is equivalent simply to require that for each object A of $\mathscr{A}$ the morphism τ_A be an isomorphism, for then the morphisms τ_A^{-1} are the components of the required τ^{-1}.

As mentioned earlier isomorphism of categories (see definition III.6) is too strict a notion. The useful alternative is the weaker notion of equivalence of categories.

Definition III.24: The functor $F : \mathscr{A} \longrightarrow \mathscr{B}$ is an **equivalence of categories** iff there is a functor $G : \mathscr{B} \longrightarrow \mathscr{A}$ with GF naturally isomorphic to $1_{\mathscr{A}}$ and FG naturally isomorphic to $1_{\mathscr{B}}$.

Examples of natural transformations abound. The simplest are those between constant functors.

Exercise III.21. Let D and D' objects of $\mathscr{D}$. Show that natural transformations between the constant functors D and D' correspond to morphisms between the objects.

The most common home for natural transformations is functor categories which we will define and produce shortly, but first let's see some stand alone examples of natural transformations.

III.6 Examples of Natural Transformations

III.6.1 Dual Vector Spaces

The first example of a natural equivalence to appear in print was that of finite dimensional vector spaces and their double duals. It is hard to do better, so we will continue the tradition.

Consider the category $\mathscr{V}$ of vector spaces over a field K. We have the contravariant functor $D(\bullet) = \mathscr{V}(\bullet, \mathsf{K})$. For any vector space V, $D(V)$ is called the dual space of V and is usually written as V^*. If $f : V \longrightarrow W$ is a linear transformation, then we have the dual linear transformation $D(f) : D(W) \longrightarrow D(V)$ which we will write as $f^* : W^* \longrightarrow V^*$. (Notice this is consistent with definition I.8 and is part of the reason for that definition.) The elements of V^* are called linear functionals on V.

The **second dual** or **double dual** is just gotten by iterating the dual space construction $DD(V) = D(D(V))$, and DD is a functor from $\mathscr{V}$ to itself. $\tau : 1_{\mathscr{V}} \longrightarrow DD$ is the familiar construction: each $v \in V$ defines $\tau(v) \in V^{**}$ by $\tau(v)(v^*) = v^*(v)$ for every $v^* \in V^*$.

All of the difficulty here is getting notation straight. To verify that τ is a natural transformation we must show that if $f : V \longrightarrow W$ is a linear transformation and $f^{**} : V^{**} \longrightarrow W^{**}$ is its double dual, then $f^{**}\tau = \tau f$. And that is easy once you understand that if $v^{**} \in V^{**}$, then $L^{**}(v^{**})(w*) = v^{**}(w * f)$ for $w* : W \longrightarrow \mathsf{K}$, an element of W^*.

Warning: Here we have v, v^*, v^{**}, w, w^*, and w^{**}. The only relations here are that $v \in V$, $v^* \in V^*$, etc. In particular v^* is *not* the result of applying some * operator to v.

Exercise III.22. Verify that $\tau : 1_{\mathscr{V}} \longrightarrow DD$ as defined above is a natural transformation. Also verify that τ_V is injective.

From this we see that when restricting to $\mathscr{V}_0$, the full subcategory of finite dimensional vector spaces, we find τ is a natural isomorphism. And this in turn shows that $D : \mathscr{V}_0^{\text{op}} \longrightarrow \mathscr{V}_0$ is an equivalence of categories. *The category of finite dimensional vector spaces is self-dual!*

III.6.2 Free Monoid Functor

In Section III.2.12 we introduced the forgetful functor U : **Monoid** $\longrightarrow$ **Set** and the related free monoid functor F : **Set** $\longrightarrow$ **Monoid**. Here we will meet a related natural transformation. This is leading up to the material on adjoint functors in section V.6

First, there is a natural transformation $\eta : 1_{\mathbf{Set}} \longrightarrow UF$ with $\eta_A(a) = (a)$. [Here (a) is the sequence with just one term, a.]

Exercise III.23. Verify that $\eta : 1_{\mathbf{Set}} \longrightarrow UF$ is indeed a natural transformation.

Second, there is a natural transformation $\varepsilon : FU \longrightarrow 1_{\mathbf{Monoid}}$ with $\varepsilon(m_1, m_2, \cdots, m_n) = m_1 m_2 \cdots m_n$. [Here $(m_1, m_2, \cdots, m_n)$ is a sequence of n elements from the monoid M (considered as a set), while $m_1 m_2 \cdots m_n$ is the product of those elements in the monoid.]

Exercise III.24. Verify that $\varepsilon : FU \longrightarrow 1_{\mathbf{Monoid}}$ is indeed a natural transformation.

In that earlier discussion we presented a function $\mathbf{Set}(A, U(M)) \longrightarrow \mathbf{Monoid}(F(A), M)$. Let's get that in another way. If $f : A \longrightarrow U(M)$ is any function from the set A into the underlying set of the monoid M, then we have the homomorphism $F(f) : F(A) \longrightarrow FU(M)$. We compose that with ε_M to get $\varepsilon_M F(f) : F(A) \longrightarrow M$.

Exercise III.25. Show that the function $\mathbf{Set}(A, U(M)) \longrightarrow \mathbf{Monoid}(F(A), M)$ given by $f \mapsto \varepsilon F(f)$ is the same function specified in Section III.2.12.

Finally, we can get the inverse function explicitly. If $h : F(A) \longrightarrow M$ is a monoid homomorphism, then we can forget that and consider it just as a function between sets, i.e., $U(h) : UF(A) \longrightarrow U(M)$. But then $U(h)$ composed with η_A gives the function $U(h)\eta_A : A \longrightarrow U(M)$, and this is the desired inverse.

Exercise III.26. Show that the function $\mathbf{Monoid}(F(A), M) \longrightarrow \mathbf{Set}(A, U(M))$ given by $h \mapsto U(h)\eta$ is the inverse of the function of the preceding exercise.

Exercise III.27. Show that $\mathbf{Monoid}(F(\bullet), \bullet)$ is naturally equivalent to $\mathbf{Set}(\bullet, U(\bullet))$ with both considered as functors from $\mathbf{Set}^{\mathrm{op}} \times \mathbf{Monoid}$ to $\mathbf{Monoid}$.

This is the basic mojo for adjoint functors, but we will go through it a couple more times before we finally make state the theorem.

III.6.3 Commutator and Abelianizer

Refer back to sections III.2.14 and III.2.15 for the basic information used here. Recall that we have the commutator functor $C : \mathbf{Group} \longrightarrow \mathbf{Group}$, the inclusion functor $I : \mathbf{Ab} \longrightarrow \mathbf{Group}$ and the Abelianizer $A : \mathbf{Group} \longrightarrow \mathbf{Ab}$. There are also a number of interesting natural transformations among them.

First define $\iota : C \longrightarrow 1_{\mathbf{Group}}$ by taking $\iota_G : [G, G] \longrightarrow G$ to be the inclusion of the subgroup.

Exercise III.28. Verify that ι as defined above is a natural transformation.

Next define $\pi : 1_{\mathbf{Group}} \longrightarrow IA$ by $\pi_G : G \longrightarrow G/[G,G]$ is the canonical homomorphism from G to the quotient group $G/[G,G]$.

Exercise III.29. Verify that π as defined above is a natural transformation.

Using this we define a function $\mathbf{Ab}(A(G), A) \longrightarrow \mathbf{Group}(G, I(A))$ as follows. If $f : A(G) \longrightarrow A$ is a homomorphism of Abelian groups, then we consider it as a group homomorphism and get $I(f) : IA(G) \longrightarrow I(A)$, so we get $\pi A(f) : G \longrightarrow I(A)$.

Now observe that the functor $AI : \mathbf{Ab} \longrightarrow \mathbf{Ab}$ is "take an Abelian group and produce its quotient group modulo the commutator subgroup (which is 0)", i.e., it is the identity functor. So of course we have the identity natural transformation $1 : AI \longrightarrow 1_{\mathbf{Ab}}$.

And that allows us to define a function from $\mathbf{Group}(G, I(A)) \longrightarrow \mathbf{Ab}(A(G), A)$ via: For $g : G \longrightarrow I(A)$ we get

$$A(G) \xrightarrow{A(g)} AI(A) \xrightarrow{\cong} A$$

which is in $\mathbf{Ab}(A(G), A)$.

Exercise III.30. Show that the two functions just defined are inverse to one another and give bijections between $\mathbf{Ab}(A(G), A)$ and $\mathbf{Group}(G, I(A))$.

That and a tiny bit more work gives the following result.

Exercise III.31. Show that $\mathbf{Ab}(A(\bullet), \bullet)$ is naturally equivalent to $\mathbf{Group}(\bullet, I(\bullet))$ with both considered as functors from $\mathbf{Group}^{\mathrm{op}} \times \mathbf{Ab}$ to $\mathbf{Set}$.

R

This is another, simpler, example of the basic mojo for adjoint functors. Also this factorization of $G \longrightarrow A$ (G a group, A an Abelian group) through $G/[G,G]$ was understood and named "natural" before natural transformations were defined. Indeed it was one of the examples that inspired Eilenberg and Mac Lane to give the name to natural transformations.

III.6.4 The Discrete Topology and the Forgetful Functor

In Section III.2.16 on the Discrete Topology Functor we introduced the forgetful functor $U : \mathbf{Top} \longrightarrow \mathbf{Set}$ and the discrete topological space functor $F : \mathbf{Set} \longrightarrow \mathbf{Top}$.

In direct analogy to what we did the the last two sections we first define a natural transformation $\eta : 1_{\mathbf{Set}} \longrightarrow UF$. Note that F equips a set with the discrete topology, while U just throws it away. So we take $\eta_S = 1_S$. Clearly that is a natural transformation.

From this trivial natural transformation we get a function $\mathbf{Top}(F(S), X) \longrightarrow \mathbf{Set}(S, U(X))$: If $f : F(S) \longrightarrow X$ is a continuous function, then we get the plain old function $U(f) : UF(S) \longrightarrow U(X)$. Composing that with $\eta_S : S \longrightarrow UF(S)$ we get $U(f)\eta_S : S \longrightarrow U(X)$.

[Yes, this really is a lot of formalism for a trivial observation, but its best to start with the easy cases!]

Second, there is a natural transformation $\varepsilon : FU \longrightarrow 1_{\mathbf{Top}}$ with $\varepsilon_T = 1_T$. [ε is more interesting than η as FU takes a topological space, (X, T), and produces the topological space $(X, \mathcal{P}(X))$. So it is key to observe that $1_X : (X, \mathcal{P}(X)) \longrightarrow (X, \mathsf{T})$ is always continuous.]

Exercise III.32. Verify that $\varepsilon : FU \longrightarrow 1_{\mathbf{Top}}$ is indeed a natural transformation.

Again we can use this natural transformation to define a function $\mathbf{Set}(S, U(X)) \longrightarrow \mathbf{Top}(F(S), X)$: If $f : S \longrightarrow U(X)$ is any function, then we get $F(f) : F(S) \longrightarrow FU(X)$, and $\varepsilon_X F(f) : F(S) \longrightarrow X$ is in $\mathbf{Top}(F(S), X)$.

Exercise III.33. Show that the two functions just defined are inverse to one another and give bijections between $\mathbf{Top}(F(S), X)$ and $\mathbf{Set}(S, U(X))$.

That and a tiny bit more work gives

Exercise III.34. Show that $\mathbf{Top}(F(\bullet), \bullet)$ is naturally isomorphic to $\mathbf{Set}(\bullet, I(\bullet))$ with both considered as functors from $\mathbf{Set}^{\mathrm{op}} \times \mathbf{Top}$ to $\mathbf{Set}$.

Yes, this is the basic mojo for adjoint functors once again. So the proof of the theorem of which this is an example (which is coming in Section V.6) should be old hat.

III.6.5 The Godement Calculus

We've noted that composition of functors, and also the composition of natural transformations. But we can also form other composites of functors and natural transformations. The Godementcalculus extends and codifies this. Although The primary use of this material is in Chapter XV (*2-Categories*), definitions III.25 (of βF) and III.26 (of $G\alpha$) are used extensively not only throughout these notes, but in the literature of category theory as a whole.

Godement's 1958 book *Topologie Algébrique et Théorie des Faisceaux* [26]was very influential in introducing the mathematical world to the importance of sheaves which had been introduced by Leray and developed by Cartan, Lazard and others a few years earlier. Part of this was also demonstrating that category theory was more than just a convenient language. As an incidental aspect of this Godement presented "cinq règles de calcul fonctoriel" and these are now commonly known as the Godement calculus of natural transformations.

Other discussions can be found in Barr and Wells [3]and Arbib and Manes [1].Much of this can also be found in other sources, without mention of the Godement calculus, under the names "vertical composition" and "horizontal composition", for example see Mac Lane [53, Sec. II.5].

Definition III.25: Suppose $F : \mathscr{B} \longrightarrow \mathscr{C}$ and $G, G' : \mathscr{C} \longrightarrow \mathscr{D}$ are functors and $\beta : G \longrightarrow G'$ is a natural transformation. Then $\beta\mathbf{F}$ is the natural transformation from GF to $G'F$ defined by $\beta F_B = \beta_{F(B)}$

Exercise III.35. Verify that βF as just defined is indeed a natural transformation from GF to $G'F$.

Definition III.26: Suppose $F, F' : \mathscr{B} \longrightarrow \mathscr{C}$ and $G : \mathscr{C} \longrightarrow \mathscr{D}$ are functors and $\alpha : F \longrightarrow F'$ is a natural transformation. Then $\mathbf{G}\alpha$ is the natural transformation from GF to GF' defined by $G\alpha_B = G(\alpha_B)$.

Exercise III.36. Verify that $G\alpha$ as just defined is indeed a natural transformation from GF to GF'.

III.6.6 Functor Categories

Once again We have the makings of new categories: objects, now functors, and morphisms, this time natural transformations. And again there are set theoretic foundations place some limitations – we are not able to prove the existence of the category of all functors between two arbitrary categories. But if $\mathscr{S}$ is a small category, while $\mathscr{C}$ is any category, then we can make the

Definition III.27: The **functor category** $\mathscr{C}^{\mathscr{S}}$ has as objects all functors from $\mathscr{S}$ to $\mathscr{C}$, and as morphisms the natural transformations between them.

If $\mathscr{C}$ is a small category, then $\mathscr{C}^{\mathscr{S}}$ is also a small category. For details about these set theoretic issues, see Mac Lane [53, Sec. II.4]

III.6.7 Examples of Functor Categories

The first example of a functor category is both universal and trivial. *Every category is a functor category!* More precisely, let $\mathbf{1}$ be the category that has exactly one object and one morphism (cf. Section B.20.2 in the Catalog of Categories (Appendix B).)

Exercise III.37. Define the *diagonal functor* $\Delta : \mathscr{C} \longrightarrow \mathscr{C}^{\mathbf{1}}$ so on objects $\Delta(C) = C$, the constant functor selecting C and on morphisms Δ just selects the corresponding natural transformation between the relevant constant functors. (See Section III.2.5 and exercise III.21.) Verify that Δ is an isomorphism.

The category $\mathbf{1}$ is almost the simplest case of a discrete category, i.e., a category that has only identity morphisms. (Cf. Section B.19.1 in the Catalog of Categories (Appendix B) for more discussion.) And the above exercise is just a special case of

Exercise III.38. If $\mathscr{D}$ is a discrete finite category with n objects, show that $\mathscr{C}^{\mathscr{D}} \cong \mathscr{C}^n$ where $\mathscr{C}^n$ is the n-fold product of $\mathscr{C}$ with itself.

The very simplest case of a discrete category is the empty category $\mathbf{0}$. The functor category $\mathscr{C}^{\mathbf{0}}$ is isomorphic to $\mathbf{1}$ as the only functor is the empty functor and the only natural transformation of the empty functor is the identity natural transformation.

Another simple example of a functor category is $\mathscr{C}^{\mathbf{2}}$ where $\mathbf{2}$ is the morphism category discussed in Sections I.1.3 and B.20.3. The category $\mathbf{2}$ is illustrated by

$$\circlearrowright\, 0 \xrightarrow{\ \ !\ \ } 1 \,\circlearrowleft$$

where the two circular arrows are the identity maps.

We introduced $\mathscr{C}^{\mathbf{2}}$ back in Section II.6 with a quite different definition: the **morphism category of** $\mathscr{C}$ has as objects the morphisms of $\mathscr{C}$, while a morphism in $\mathscr{C}^{\mathbf{2}}$ from $f : A \longrightarrow B$ to $f' : A' \longrightarrow B'$ is a pair (h, k) of $\mathscr{C}$-morphisms so that

$$\begin{array}{ccc} A & \xrightarrow{\ h\ } & A' \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f'} \\ B & \xrightarrow[\ k\ ]{} & B' \end{array}$$

commutes.

But these are really the "same" or, more formally, the two categories are isomorphic. To see this let us for the moment write $\mathscr{M}$ for the morphism category of $\mathscr{C}$ as above and $\mathscr{C}^{\mathbf{2}}$ for the functor category.

Define a functor from $\mathscr{M}$ to $\mathscr{C}^{\mathbf{2}}$ as follows: An object f in $\mathscr{M}$ (i.e., a morphism $f : A \longrightarrow B$ in $\mathscr{C}$) goes to the object F in $\mathscr{C}^{\mathbf{2}}$ which is the functor $F : \mathbf{2} \longrightarrow \mathscr{C}$ with $F(0) = A$, $F(1) = B$ and $F(!) = f$. And a morphism $(h, k) : f \longrightarrow f'$ in $\mathscr{M}$ to the morphism $\tau : F \longrightarrow F'$ in $\mathscr{C}^{\mathbf{2}}$ which is the natural transformation $\tau_0 : F(0) \longrightarrow F'(0) = h : A \longrightarrow A'$ and $\tau_1 : F(1) \longrightarrow F'(1) = k : B \longrightarrow B'$. The commuting square we must have to complete the verification that this is a natural transformation is

$$\begin{array}{ccc} F(0) & \xrightarrow{\tau_0} & F'(0) \\ {\scriptstyle F(!)}\downarrow & & \downarrow{\scriptstyle F'(!)} \\ F(1) & \xrightarrow[\tau_1]{} & F'(1) \end{array}$$

which is exactly the square

$$\begin{array}{ccc} A & \xrightarrow{h} & A' \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f'} \\ B & \xrightarrow[k]{} & B' \end{array}$$

which is guaranteed to commute by the definition of a morphism in $\mathscr{M}$. This is clearly a functor.

And the inverse functor from $\mathscr{C}^{\mathbf{2}}$ to $\mathscr{M}$ takes each object F of $\mathscr{C}^{\mathbf{2}}$ to the object $F(!)$ of $\mathscr{M}$, and each morphism $\tau : F \longrightarrow F'$ goes to the morphism (τ_0, τ_1) in $\mathscr{M}$. This is also clearly a functor and equally clearly the inverse of the previous functor.

From now on we will always write $\mathscr{C}^{\mathbf{2}}$ for these categories and use whichever description is convenient for the purpose at hand.

III.6.8 Discrete Dynamical Systems

The simplest type of discrete dynamical system consists of a set, S – the state space, and a transition function $t : S \longrightarrow S$. The image is that this models a system which starts in some initial state and then moves from state to state by application of the function t. The items of interest are the "flows" in the dynamical system which are the sequences $(s_0, s_1, s_2, \cdots)$ from S where $s_0 \in S$ is some initial state and then $s_1 = t(s_0)$, $s_2 = t(s_1)$, etc.

If (R, t_R) and (S, t_S) are two discrete dynamical systems, then a morphism $f : (R, t_R) \longrightarrow (S, t_S)$ is a function $f : R \longrightarrow S$ with $t_S f = f t_R$. Clearly such a morphism produces a function taking the flows on (R, t_R) to those on (S, t_S) by $f(r_0, r_1, r_2, \cdots) = (f(r_0), f(r_1), f(r_2), \cdots)$. We will use $\mathscr{D}$ for the category of discrete dynamical systems.

Next consider the category $\mathbf{Set}^{\mathbb{N}}$ where $\mathbb{N}$is the additive monoid of natural numbers considered as a small category with one object. 0 is the identity morphism of $\mathbb{N}$, so we will also use it for the unique object as well.

We are going to define a functor from $\mathbf{Set}^{\mathbb{N}}$ to $\mathscr{D}$, the category of discrete dynamical systems. Observe that an element of $\mathbf{Set}^{\mathbb{N}}$, i.e., a functor from $\mathbb{N}$to $\mathbf{Set}$, assigns to the object 0 some set S and to each n in $\mathbb{N}$some function $S \longrightarrow S$. Writing $t : S \longrightarrow S$ for the function assigned to $1 \in \mathbb{N}$ by the functor we note that the function assigned to an integer $n \geq 1$ is f^n, the composition of f with itself n times. So an object in $\mathbf{Set}^{\mathbb{N}}$ is completely determined by the pair (S, t) which is exactly an object in $\mathscr{D}$! And every such pair defines a functor from $\mathbb{N}$to **Set**.

For good measure, if (R, t_R) and (S, t_S) are two such pairs defining functors in $\mathbf{Set}^{\mathbb{N}}$, a natural transformation between them is a function $\tau : R \longrightarrow S$ such that $\tau t_S^n = t_R^n \tau$ which certainly includes the requirement that $\tau t_S = t_R \tau$ and so is a morphism in $\mathscr{D}$. Actually $\tau t_S = t_R \tau$ implies $\tau t_S^n = t_R^n \tau$ so in fact we have an isomorphism of categories $\mathbf{Set}^{\mathbb{N}} \cong \mathscr{D}$.

Chapter IV

Constructing Categories - Part II

IV.1 Comma Categories

An extremely versatile and useful construction of categories has the unenlightening name "comma category". In his thesis [45], Lawvere began a program to develop category theory as a foundation for mathematics separate and independent of set theory. As part of that he associated with any two functors S and T with common codomain another category (F, T). Unfortunately he gave this construction no name, but it's value was soon recognized and, for lack of anything better, the name comma category was soon attached. To this day no good name has appeared even though the notation has changed. (The origin of the original notation is explained in example 5 below. The use of the notation (S, T) is just so common that some alternative notation is essential.)

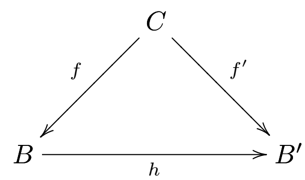
Definition IV.1: For any two functors $S : \mathscr{A} \longrightarrow \mathscr{C}$ and $T : \mathscr{B} \longrightarrow \mathscr{C}$ with common codomain, define the category $(S \downarrow T)$ whose objects are triples (A, f, B) where A is an object of $\mathscr{A}$, B is an object of $\mathscr{B}$, and $f : S(A) \longrightarrow T(B)$ is a morphism in $\mathscr{C}$. A morphism from (A, f, B) to (A', f', B') is a pair of morphisms (a, b) with $a : A \longrightarrow A'$, $b : B \longrightarrow B'$ and $T(b)f = f'S(a)$. This is summarized in the following diagram:

$$\text{object } (A,f,B): \quad \begin{array}{c} S(A) \\ \Big\downarrow f \\ T(B) \end{array} \qquad\qquad \text{morphism } (a,b): \quad \begin{array}{ccc} S(A) & \xrightarrow{S(a)} & S(A') \\ \Big\downarrow f & & \Big\downarrow f' \\ T(B) & \xrightarrow[T(b)]{} & T(B') \end{array}$$

Examples and Special Cases

1. $(C \downarrow \mathscr{C})$ – Category of objects under C

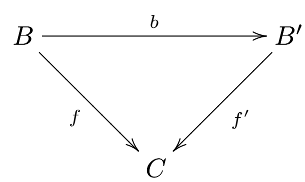
The first special case occurs with S being a constant functor C for some object C in $\mathscr{C}$, and T the identity functor on $\mathscr{C}$. We then have the **category of objects under C** and written $(C \downarrow \mathscr{C})$. This is also known as the **coslice category with respect to C**. The objects (C, f, B) can be simplified to (B, f) as C is constant; and $f : C \longrightarrow T(B)$ simplifies to $f : C \longrightarrow B$. As well a morphism $(c, b) : (B, f) \longrightarrow (B', f')$ simplifies to $b : B \longrightarrow B'$, (as c must be identity morphism on C) with the following diagram commuting.



This example, as well as the next two, were presaged in Section II.6

2. $(\mathscr{C} \downarrow C)$ – Category of objects over C

   Similarly, S might be the identity functor and T a constant functor: this is the **category of objects over** C (where C is the object of $\mathscr{C}$ selected by T), written $(\mathscr{C} \downarrow C)$. This is also known as the **slice category over** C. It is the dual concept to objects-under-C. The objects are pairs (B, f) with $f : B \longrightarrow C$, while the morphisms are just morphisms $b : B \longrightarrow B'$ in $\mathscr{C}$ with the following diagram commuting.



3. $(\mathscr{C} \downarrow \mathscr{C})$ – Morphism Category of $\mathscr{C}$

   When both S and T are the identity functor $1_{\mathscr{C}}$, the comma category $(\mathscr{C} \downarrow \mathscr{C})$ is immediately seen to be the same as the category $\mathscr{C}^{\mathbf{2}}$ as discussed in Sections II.6 and III.6.7.

4. $(A \downarrow T)$ and $(S \downarrow B)$

   In either of the above two cases, the identity functor may be replaced with some other functor; this yields a family of categories particularly useful in the study of Universal Mapping Properties, the topic of the next chapter. For example, if T is the forgetful functor carrying a monoid to its underlying set, and S is a constant functor selecting the set A, then $(A \downarrow T)$ is the comma category whose objects are pairs $(A, f : A \longrightarrow T(M))$ with f a function from the set A to the underlying set of some monoid M.

5. $(A \downarrow B)$

   Another special case occurs when both S and T are constant functors with domain the category **1**. If S selects A and T selects B, then the comma category produced is equivalent to the set of morphisms between A and B. (Strictly, it is a discrete category – all the morphisms are identity morphisms – which may be identified with the set of its objects.)

6. $(\bullet \downarrow \mathbf{Set})$

   The category of pointed sets is a comma category $(\bullet \downarrow \mathbf{Set})$, with $\bullet$ being (a functor selecting) any singleton set, and **Set** (the identity functor of) the category of sets. Each object of this category is a set, together with a function selecting some element of the set: the "base point". Morphisms are functions on sets which map base points to base points. Similarly there is the category of pointed spaces $(\bullet \downarrow \mathbf{Top})$.

7. $(\mathbf{Set} \downarrow D)$ – the category of graphs

   The category of graphs is $(\mathbf{Set} \downarrow D)$, with the functor D taking a set s to $s \times s$. The objects (a, b, f) then consist of two sets and a function; a is an indexing set, b is a set of nodes, and $f : a \longrightarrow b \times b$ chooses pairs of elements of b for each input from a. That is, f picks out certain edges from the set of possible edges. A morphism in this category is made up of two functions, one on the indexing set and one on the node set. They must "agree" according to the general definition above, meaning that $(g, h) : (a, b, f) \longrightarrow (a', b', f')$ must satisfy $f'g = S(h)f$. In other words, the edge corresponding to a certain element of the indexing set, when translated, must be the same as the edge for the translated index.

In addition to providing a very general method of constructing additional categories, comma categories are closely related to universal mapping properties which are the topic of the next chapter. In addition comma categories themselves arise from a universal mapping property. The starting point for this discussion is to note that there are "projection" functors π_S and π_T from $(S \downarrow T)$ to $\mathscr{A}$ and to $\mathscr{B}$. On objects the functors are $\pi_S(A, f, B) = A$ and $\pi_T(A, f, B) = B$, while on morphisms they are $\pi_S(a, b) = a$ and $\pi_T(a, b) = b$. Of course this is equivalent to specifying the functor $< \pi_S, \pi_T >: (S \downarrow T) \longrightarrow \mathscr{A} \times \mathscr{B}$ and we will use this in Proposition V.14 below.

There is also a canonical natural transformation $\alpha : S\pi_{\mathscr{A}} \longrightarrow T\pi_{\mathscr{B}}$ defined by $\alpha_{(A,f,B)} = f$. Verification that $\pi_{\mathscr{A}}$ and $pi_{\mathscr{B}}$ are functors and that α is a natural transformation consists of a simple use of the definition of $(S \downarrow T)$ and is left for the reader.

The more interesting result is that $(S \downarrow T)$ is universal in this regard as captured in the next proposition.

Proposition IV.13 *Starting with functors $S : \mathscr{A} \longrightarrow \mathscr{C}$ and $T : \mathscr{B} \longrightarrow \mathscr{C}$, suppose there is a category $\mathscr{D}$ together with functors $P_{\mathscr{A}} : \mathscr{D} \longrightarrow \mathscr{A}$ and $P_{\mathscr{B}} : \mathscr{D} \longrightarrow \mathscr{B}$ and a natural transformation $\beta : SP_{\mathscr{A}} \longrightarrow TP_{\mathscr{B}}$. Then there* *is a* unique *functor* $P : \mathscr{D} \longrightarrow (S \downarrow T)$ *such that* $P_{\mathscr{A}} = \Pi_{\mathscr{A}} P$, $P_{\mathscr{B}} = \Pi_{\mathscr{B}} P$ *and* $\beta = \alpha P$

Exercise IV.1. Unwind the conditions in the above proposition to see what P must be, and thereby prove the proposition.

Chapter V

Universal Mapping Properties

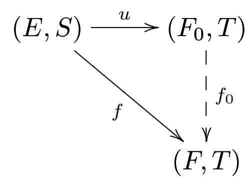
Beyond the foundational concepts of category, functor and natural transformation, the most important idea in category theory is that of a "universal mapping property". We've mentioned various universal mapping properties along the way to this point, without ever making the term precise. In this chapter we will finally give precision to that term, but in many different ways and still without giving a precise meaning to that phrase! Rather we will define universal elements, universal arrows, representable functors, adjoint functors, Kan extensions and a number of other concepts, and show that in some sense that they are all equivalent. But they are also all useful in different ways and important ways. This chapter is fundamental for everything that follows in these notes.

The notion of universal mappings probably started with Poincaré's study of universal covering surfaces (see Dieudonné [15, Part 3, Sec. I.2]), and topology is the source of a very large collection of important examples both many that predate category theory and some that drove the development of category theory itself.

Another important motivating source of universal mappings was the notion of free structures in algebra. Indeed Pierre Samuel in "On universal mappings and free topological groups" [64] was probably the first to use the phrase "universal mapping" in print. The approach that he used – in terms of sets with structures – was then substantially developed by Bourbaki[9, IV.3] in his *Elements of Mathematics: Theory of Sets*

Neither Samuel nor Bourbaki discussed categories, rather they wrote of structures and species of structures, and in this context the problem of "universal mappings" becomes: Given a set E with a structure S and appropriate mappings $f : E \longrightarrow F$ into sets with a compatible structure T, find a "universal set" F_0 and suitable mapping $u : E \longrightarrow F_0$ so that every $f : E \longrightarrow F$

factors uniquely as $f = f_0 u$.



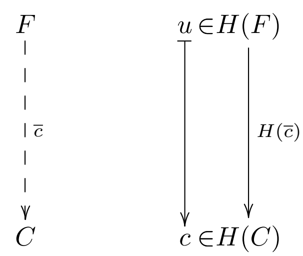
Nicely presented example applications of this approach are to be found in Bourbaki[9, IV.3.] and include free algebraic structures, rings and fields of fractions, tensor product of modules, completion of a uniform space, Stone-Čech compactification, free topological groups, almost periodic functions on a topological group and the Albanese variety of an algebraic variety.

The approach via category theory encompasses all of this and much more, so all of these examples will be treated at some point in these notes, mostly in Chapter VII. Refer to the index for specific locations

V.1 Universal Elements

The simplest of the many formalizations of the notion of "universal mapping property" is the notion of a universal element for a functor to the category of sets.

Definition V.1: A **universal element** for a functor $H : \mathscr{C} \longrightarrow \mathbf{Set}$ is a pair (F, u) with F an object of $\mathscr{C}$ and $u \in H(F)$ satisfying the following Universal Mapping Property: For each pair (C, c) with C an object of $\mathscr{C}$ and $c \in H(C)$ there is a unique morphism $\overline{c} : F \longrightarrow C$ so that $H(\overline{c})(u) = c$.



This definition applies equally well to a contravariant functor considered as a functor from $\mathscr{C}^{\mathrm{op}}$ to **Set**. Writing it out directly in terms of $\mathscr{C}$, a universal element for a contravariant functor H is a pair (F, u) with F an object of $\mathscr{C}$ and $u \in H(F)$ satisfying the following Universal Mapping Property: for each pair (C, c) with Cq an object of $\mathscr{C}$ and $c \in H(C)$ there is a unique morphism $\overline{c} : C \longrightarrow F$ so that $H(\overline{c})(u) = c$. The only change from the definition for a functor is that the morphism $\overline{c}$ is reversed (and the functor reverses morphisms.)

Example: For any category $\mathscr{C}$ define the constant functor $H : \mathscr{C} \longrightarrow \mathbf{Set}$ to take every object to the final object $1 = \{0\}$ and every morphism to the unique function from 1 to itself. Then H has a universal element iff $\mathscr{C}$ has an initial object,and the universal element consists of the initial object and the unique element in one point set.

Equally well we have the constant contravariant functor $H : \mathscr{C}^{\mathrm{op}} \longrightarrow \mathbf{Set}$, and it has a universal element iff $\mathscr{C}$ has a final object with the object in the universal element being the final object.

Compare this discussion with the original discussion of the universal mapping properties of initial (37) and final (37) objects.

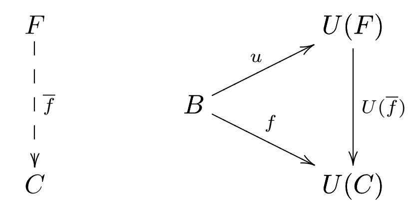
Example: The first example of a universal mapping property was for products, and that fits here very well. Fixing C_1 and C_2, two objects of the category $\mathscr{C}$, define the functor $H : \mathscr{C} \longrightarrow \mathbf{Set}$ on an object C by $H(C) = \{(f_1, f_2) : f_1 : C_1 \longrightarrow C, \wedge f_2 : C_2 \longrightarrow C\}$ and on a morphism $f : C \longrightarrow D$ by $H(f)(f_1, f_2) = (ff_1, ff_2)$. (The easy verification that H is a functor is left to the reader.) Then a universal element for H consists of an object $C_1 \times C_2$ of $\mathscr{C}$ and an element $(\pi_1, \pi_2) \in H(C_1 \times C_2)$ such that for every object C of $\mathscr{C}$ and every $(f_1, f_2) \in H(C)$ there is a unique morphism $\langle f_1, f_2\rangle : C_1 \times C_2 \longrightarrow C$ (using the notation for products) with $H(\langle f_1, f_2\rangle)(\pi_1, \pi_2) = (f_1, f_2)$. Of course this is just another way of writing that $\langle f_1, f_2\rangle$ is the unique morphism such that $\pi_1\langle f_1, f_2\rangle = f_1$ and $\pi_2\langle f_1, f_2\rangle = f_2$. We summarize thus by saying that a product of C_1 and C_2 is a universal element for the functor H. This again emphasizes very well that a product of two objects is *not* just an object.

Sums fit here equally well. For fixed objects C_1 and C_2 there is the contravariant functor $K : \mathscr{C} \longrightarrow \mathbf{Set}$ defined by $K(C) = \{(f_1, f_2) : f_1 : C \longrightarrow C_1, \wedge f_2 : C \longrightarrow C_2\}$ and on a morphism $f : D \longrightarrow C$ by $H(f)(f_1, f_2) = (f_1 f, f_2 f)$. Then a universal element for K consists of an object $C_1 + C_2$ and an element $(\iota_1, \iota_2) \in K(C_1 + C_2)$ such that for every object C of $\mathscr{C}$ and every $(f_1, f_2) \in K(C)$ there is a unique morphism $[f_1, f_2] : C \longrightarrow C_1 + C_2$ (using the notation for sums) with $H([f_1, f_2])(\iota_1, \iota_2) = (f_1, f_2)$. Of course this is just another way of writing that $[f_1, f_2]$ is the unique morphism such that $[f_1, f_2]\iota_1 = f_1$ and $[f_1, f_2]\iota_2 = f_2$. We summarize that by saying that a sum of C_1 and C_2 is a universal element for the functor K. This again emphasizes very well that a sum of two objects is *not* just an object.

Example: The other early example of a universal mapping property was the free monoid (see 87.) It fits here by taking a fixed set A and defining the functor H : **Monoid** $\longrightarrow$ *Set* on a monoid M to be $H(M) = \{f : A \longrightarrow U(M)\}$ and on a homomorphism $h : M \longrightarrow N$ to be $H(h)(f) = fU(h)$. (Here U is the forgetful functor from **Monoid** to **Set**. Again verification that H is indeed a functor is left to the reader.) Now a universal element for H is a monoid A^* and an element of $H(A^*)$, i.e., a function $\varepsilon : A \longrightarrow U(A^*)$ such that every monoid M and each function $f : A \longrightarrow U(M)$ there is a unique monoid homomorphism $f^* : A^* \longrightarrow M$ extending f, i.e., the universal mapping property of the free monoid generated by A.

V.2 Universal Arrows

Definition V.2: For a functor $U : \mathscr{C} \longrightarrow \mathscr{B}$ and an object B of $\mathscr{B}$, a **universal arrow** from B to U is a pair (F, u) where F is an object of $\mathscr{C}$ and $u : B \longrightarrow U(F)$ satisfying the following Universal Mapping Property: For each pair (C, f) with $f : B \longrightarrow U(C)$ in $\mathscr{B}$ there is a unique $\overline{f} : F \longrightarrow C$ so that $f = U(\overline{f})u$.



Again, this definition applies equally well to a contravariant functor considered as a functor from $\mathscr{C}^{\text{op}}$ to $\mathscr{B}$. Writing it out directly in terms of $\mathscr{C}$, a universal arrow for a contravariant functor U is a pair (F, u) with F an object of $\mathscr{C}$ and $u : D \longrightarrow U(F)$ satisfying the Universal Mapping Property: for each pair (C, c) with C an object of $\mathscr{C}$ and $c : B \longrightarrow U(C)$ there is a unique morphism $\overline{c} : C \longrightarrow F$ so that $H(\overline{c})(u) = c$. And again the only change from the definition for a functor is that the morphism $\overline{c}$ is reversed (and the functor reverses morphisms.)

Universal elements are really examples of universal arrows – there is a natural correspondence between the elements of a set S and the morphisms from a final object, 1, and S. So a universal element for a functor $H : \mathscr{C} \longrightarrow \mathbf{Set}$ is "the same" as a universal arrow from a final object 1 in **Set** to H.

Universal arrows are equally well examples of universal elements! Starting with a functor $U : \mathscr{C} \longrightarrow \mathscr{B}$ and an object B of $\mathscr{B}$, define a new functor $H : \mathscr{C} \longrightarrow \mathbf{Set}$ as $H(\bullet) = \mathscr{C}(B, U(\bullet))$. Then a universal element for H is a pair (F, u) where F is an object of $\mathscr{C}$ and $u \in \mathscr{C}(B, U(F))$, i.e., a morphism $u : B \longrightarrow U(F)$, and the universal mapping property for the universal element is just another way of writing the universal mapping property of the universal arrow.

V.3 Representable Functors

Definition V.3: A **representable functor** $F : \mathscr{C} \longrightarrow \mathbf{Set}$ is one naturally equivalent to a Hom functor $\text{Hom}(C, \bullet) : \mathscr{C} \longrightarrow \mathbf{Set}$. Explicitly, F is representable iff there is an object C and a natural equivalence $\eta : \text{Hom}(C, \bullet) \xrightarrow{\cong} F$.

Proposition V.14 *The functor $F : \mathscr{C} \longrightarrow$ **Set** is representable iff it has a universal element. If (U, u) is a universal element for F, then*

$$\eta : \mathrm{Hom}(U, \bullet) \xrightarrow{\cong} F$$

is defined by $\eta_C(f) = f_*(u)$. And if

$$\eta : \mathrm{Hom}(U, \bullet) \xrightarrow{\cong} F$$

is a natural isomorphism, then $(U, \eta_U(1_U))$ is a universal element for F.

Proof: With (U, u) a universal element for F we verify that $\eta_C(f) = f_*(u)$ does indeed define a natural transformation, i.e., that for any $g : C \longrightarrow D$ this square commutes:

$$\begin{array}{ccc}
\mathrm{Hom}(U, C) & \xrightarrow{\eta_C} & F(C) \\
{\scriptstyle g_*}\downarrow & & \downarrow{\scriptstyle F(g)} \\
\mathrm{Hom}(U, D) & \xrightarrow[\eta_D]{} & F(D)
\end{array}$$

Which, for each $f \in \mathrm{Hom}(U, C)$, requires

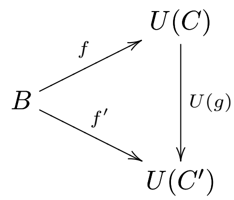
$$\begin{aligned}
F(g)\eta_C(f) &= F(g)F(f)(u) \\
&= F(gf)(u) \\
&= \eta_D(gf) \\
&= \eta_D(g_*(f)) \\
&= \eta_D g_*(f)
\end{aligned}$$

Going the other way, to show that for each object C and each $c \in F(C)$, there is $f : U \longrightarrow C$ with $f(u) = c$ it suffices to show that for each morphism $c \in \mathrm{Hom}(U, C) \cong F(C)$ there is an f with $f_*(1_U) = c$, and for that we can take $f = c$. ▮

V.4 Initial and Final Objects

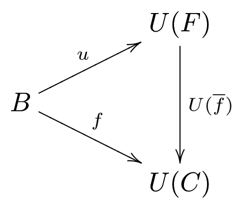
We've exhibited initial and final objects as universal elements, but we can equally well see that universal arrows are special cases of initial and final objects! Given the datum for a universal arrow, i.e., a functor $U : \mathscr{C} \longrightarrow \mathscr{B}$ and an object B of $\mathscr{B}$, we have the comma category (see Section IV.1) $(B \downarrow U)$ which has as objects the pairs $(C, f : B \longrightarrow U(C))$ while a morphism from $(C, f : B \longrightarrow U(C))$ to $(C', f' : B \longrightarrow U(C'))$ is a morphism $g : C \longrightarrow C'$

in $\mathscr{C}$ so that



commutes.

In this category an initial object is an object $(F, u : B \longrightarrow U(F))$ where for any other object $(C, f : B \longrightarrow U(C))$ there is a unique morphism $\overline{f} : F \longrightarrow C$ in $\mathscr{C}$ so that



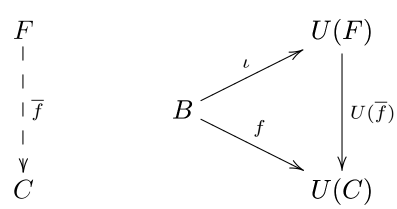
commutes. And that is exactly the definition of a universal arrow from B to U. It is worthwhile to summarize this in a proposition.

Proposition V.15 *Let U be a functor from $\mathscr{C}$ to $\mathscr{B}$ and B an object of $\mathscr{B}$, then a **universal arrow** from B to U is an initial object of $(B \downarrow U)$, and conversely.*

V.4.1 Free Objects

As mentioned above the motivating source of the name "universal mappings" was the notion of free structures in algebra. The context of based categories seems to give the best home for this notion. So let $U : \mathscr{C} \longrightarrow \mathscr{B}$ be a category based on $\mathscr{B}$ and B and object of $\mathscr{B}$.

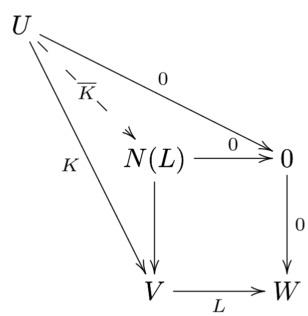
Definition V.4: A **free object** in $\mathscr{C}$ generated by an object B of $\mathscr{B}$ is an object F of $\mathscr{C}$ together with a morphism $\iota : B \longrightarrow U(F)$ in $\mathscr{B}$ satisfying the following Universal Mapping Property: For every object C in $\mathscr{C}$ and every morphism $f : B \longrightarrow U(C)$ there is a unique morphism $\overline{f} : F \longrightarrow C$



V.5 Limits and Colimits

The most familiar and widely used universal mapping properties are limits and colimits. Our very first examples of limits and colimits were products and coproducts, with initial and final objects close behind. As with these, other limits and colimits are most immediately associated with the internal structure of a category. But the preservation and creation of limits is also an important structural property of functors.

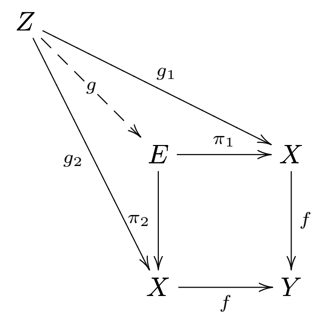
In various guises, and with various names, limits (and colimits) well predate the introduction of categories and functors. In most cases the universal mapping property was not made explicit, often because it was essentially trivial. For example the null space of a linear transformation has the following universal mapping property: For any linear transformation $L : V \longrightarrow W$ between two vector spaces over a field K there is the inclusion $N(L) \hookrightarrow V$ of the null space of L, i.e., the subspace of V consisting of all vectors v with $L(v) = 0$. If $K : U \longrightarrow V$ is any other linear transformation with $LK = 0$, then there is a *unique* linear transformation $\overline{K} : U \longrightarrow N(L)$ such that the following diagram commutes:



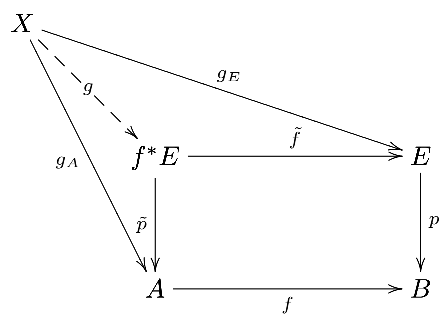
This is representative of a number of similar situations. For example L might be a group homomorphism whereupon we similarly have the kernel of L, a normal subgroup, in place of $N(L)$. In the category of sets the situation is more complicated as there is no direct analog of the kernel for functions, but rather the equivalence relation

$$E = \{(x_1, x_2) : x_1, x_2 \in X \ \text{ with } f(x_1) = f(x_2)\}$$

defined by the function (see exercise I.31 on page 22.) Here for any *pair* of functions $g_1, g_2 : Z \longrightarrow X$ with $fg_1 = fg_2$ there is a *unique* function $g : Z \longrightarrow E$ such that $g_1 = \pi_1 g$ and $g_2 = \pi_2 g$. This is summarized in the following commutative diagram.

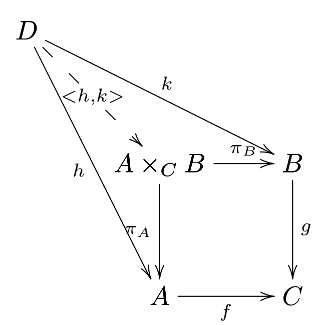


Another similar construction occurs when discussing fiber spaces in topology. A *fiber space* is a continuous surjection $p : E \longrightarrow B$. (Of course E and B must be topological spaces, and there are usually additional conditions that are imposed such as requiring that E is "locally a product".) If $f : A \longrightarrow B$ is any continuous function, then there is a *pullback of* p *along* f which is a fiber space $\tilde{p} : f^*E \longrightarrow A$ with the universal mapping property captured in the following commutative diagram:



This last universal mapping property is suitable and sufficiently important to state in an arbitrary category.

Definition V.5: Given morphisms $f : A \longrightarrow C$ and $g : B \longrightarrow C$, a **pullback** of f and g is a pair of morphisms $\pi_A : A \times_C B \longrightarrow A$ and $\pi_B : A \times_C B \longrightarrow B$ satisfying the universal mapping property:



This is indeed a generalization of the previous examples – they are all special cases for suitably chosen categories and morphisms.

Just as with products and final objects and universal mapping objects in general, pullbacks need not exist, but when they do they are unique up to a unique isomorphism. The statement and proof of this result for arbitrary limits, not just pullbacks, is in Proposition V.16.

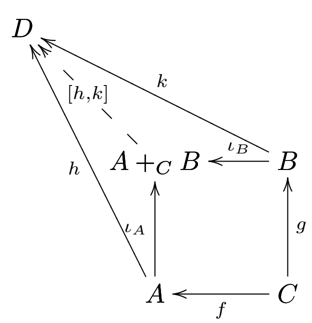
Exercise V.1. Verify that pullbacks always exist in **Set**. For functions $f : A \longrightarrow C$ and $g : B \longrightarrow C$, define

$$A \times_C B = \{(a, b) : a \in A, b \in B \wedge f(a) = g(b)\} \subseteq A \times B$$

and define the two projection functions to be the restrictions of the projections on $A \times B$, then show that this has the requisite universal mapping property.

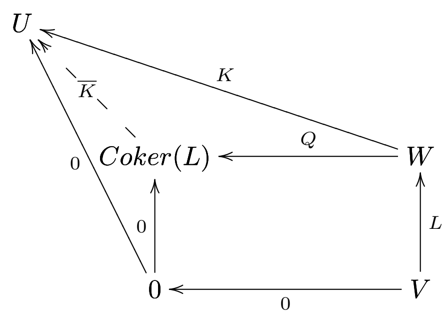
Products, final objects and all the versions of pullbacks are examples of limits. Just as for products and final objects there are the dual notions of coproducts and initial objects, there is a dual to the notion of pullbacks. These are called *pushouts* and are a fundamental examples of colimits. We could just say that a pushout is a pullback in the dual category, but it is worthwhile to spell out the definition explicitly in the category itself.

Definition V.6: Given morphisms $f : C \longrightarrow A$ and $g : C \longrightarrow B$, a **pushout** of f and g is a pair of morphisms $\iota_A : A \longrightarrow A +_C B$ and $\iota_B : B \longrightarrow A +_C B$ satisfying the universal mapping property:



So again pushouts need not exist, but when they do they are unique up to a unique isomorphism.

Various familiar constructions in various parts of mathematics are examples of pushouts. For any linear transformation $L : V \longrightarrow W$ between two vector spaces over a field K there is the cokernel of L which is $W/Im(L)$. This is the pushout of L and 0 as shown here.



The universal mapping property here is "$KL = 0$ implies there exists a unique linear transformation $\overline{K}$ with $K = \overline{K}Q$".

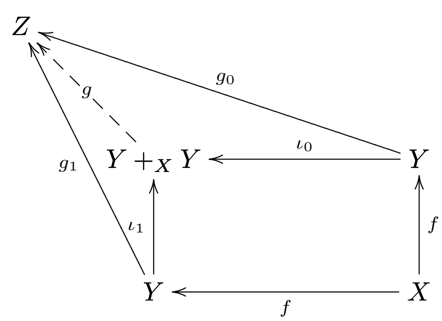
Just as with pullbacks, this is representative of a number of similar situations. For example L might be a group homomorphism whereupon we similarly have the cokernel of L, which is the quotient of W by the normal subgroup generated by the image of L in place of $N(L)$.

In the category of sets the situation is considerably less familiar, just as coproducts of sets are less familiar than products. Also there is no direct analog of the cokernel for functions, rather for $f : X \longrightarrow Y$ there is the "amalgamated sum" $Y +_X Y$ which is the quotient of the disjoint union of Y with itself $(Y + Y)$ by the equivalence relation that identifies elements coming from a common element of X. In detail,

$$Y + Y = \{(y_i, i) : i \in \{0, 1\} \wedge y_i \in Y\}$$

(see definition I.35) and $Y +_X Y$ is $(Y + Y)/\sim$ where $\sim$ is the equivalence relation generated by the requirement that $(f(x), 0) \sim (f(x), 1)$ for all $x \in X$ (see the discussion on page 192.)

Here for any *pair* of functions $g_0, g_1 : Y \longrightarrow Z$ with $g_0 f = g_1 f$ there is a *unique* function $g : Y +_X Y \longrightarrow Z$ such that $g_0 = g\iota_0$ and $g_1 = g\iota_1$. This is summarized in the following commutative diagram.



This construction is most commonly used in topology under the name "adjunction space". When X and Y are topological spaces, A is a subspace of Y

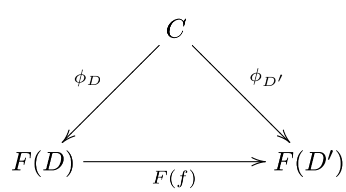
(with the inclusion function $\iota : A \hookrightarrow Y$) and $f : A \longrightarrow X$ is a continuous function (which is called the *attaching map.*) Then the pushout $X +_A Y$ is called an adjunction space and is constructed by taking the disjoint union of X and Y modulo the equivalence relation generated by requiring that a is equivalent to $f(a)$ for all $a \in A$. Intuitively we think of the spaces X and Y being glued together along the subspace A, with the function f providing the glue.

A simple and important example arises when X and Y are each a closed disk, A is the bounding circle and f is the inclusion. Then the adjunction space is (homeomorphic to) the 2-sphere. This works equally well when X and Y are both closed n-balls (or n-cells) and A is the bounding n^{-1}-sphere to give the (n+1)-sphere. Iteratively attaching cells of various dimensions leads to the definition of CW-complexes, an important class of spaces much used in algebraic topology. For information on CW-complexes look at Lundell [48] or Hatcher [30].

V.5.1 Cones and Limits

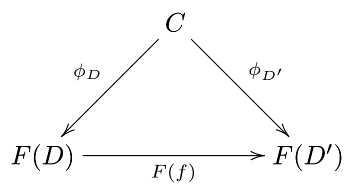
All of the examples take the form of some common "diagram" of objects and morphisms from the category $\mathscr{C}$ and then considering an object with a universal family of mappings into the diagram. A number of different ways have been developed for specifying what is meant by a suitable diagram, using graphs, diagram schemes, free categories and, the method used here, arbitrary categories. (For discussions of the various approaches see Barr and Wells [3], Popescu and Popescu [63], and Mac Lane [54].)

For any small category $\mathbf{D}$ and arbitrary category $\mathscr{C}$ we have the **diagonal functor** $\Delta : \mathscr{C} \longrightarrow \mathscr{C}^{\mathbf{D}}$ which on objects has $\Delta(C) = C$, the constant functor selecting C, and on morphisms just selects the corresponding natural transformation between the relevant constant functors. For any functor $F : \mathbf{D} \longrightarrow \mathscr{C}$ we also have the constant functor $\mathbf{1} \longrightarrow \mathscr{C}^{\mathbf{D}}$ which selects F, and from these we can form the comma category $(\Delta \downarrow F)$. It is helpful to describe the objects and morphisms somewhat more concretely. From the general definition, an object $(C, \phi : C \longrightarrow F, 1)$ in $(\Delta \downarrow F)$ consists of an object C from $\mathscr{C}$ and the unique object 1 in $\mathbf{1}$ together with a natural transformation ϕ from the constant functor selecting C to the functor F. The object 1 being unvarying, this is the same as saying an object C together with a natural transformation $\phi : C \longrightarrow F$. And this is, for each object D in $\mathbf{D}$, a morphism $\phi_D : C \longrightarrow F(D)$ (in $\mathscr{C}$) such that for every morphism $f : D \longrightarrow D'$ of $\mathbf{D}$ the diagram



is commutative.

Definition V.7: Let $F : \mathbf{D} \longrightarrow \mathscr{C}$ be a functor. A **cone** over F is an object C of $\mathscr{C}$ and a natural transformation $\phi : C \longrightarrow F$ from the constant functor C to F. This is the same as saying that for every morphism $f : D \longrightarrow D'$ of D the triangle

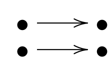


commutes.

A morphism from the cone (C, ϕ) to the cone (C', ϕ') is a morphism $f : C \longrightarrow C'$ with $\phi' f = \phi$. (NB the morphism $f : C \longrightarrow C'$ can be considered equally well as a morphism in $\mathscr{C}$ and as a natural transformation between the two constant functors.)

Clearly there is a category of cones over F with objects and morphism as in the preceding definition. We write this category as $\mathscr{C}/F$.

When the category D is very small a functor can be conveniently specified by giving the objects and morphisms that are the values of the functor. For example we will speak of "a pair of morphisms" instead of talking about "the functor that maps the two non-trivial morphisms in D to the pair of morphisms of interest". We will be equally casual when we actually need to talk about that category D and just describe it as

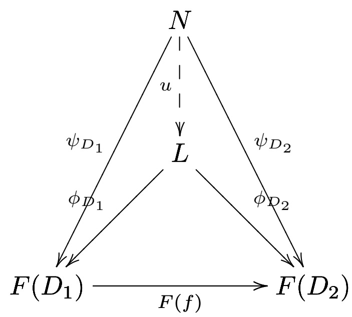


This category has four distinct objects with the corresponding four identity morphisms, the two morphisms indicated and no other morphisms.

This is all preliminary to giving the actual definition of a limit.

Definition V.8: Let $F : \mathbf{D} \longrightarrow \mathscr{C}$ be a functor. A **limit** for F is a final object (L, ϕ) in $\mathscr{C}/F$. This is also called a *universal cone* over F and is a cone $\phi : L \longrightarrow F$ with the universal mapping property that for any cone $\psi : N \longrightarrow F$ there is a unique morphism $u : N \longrightarrow L$ such that for every object D of $\mathbf{D}$ $\phi_D o u = \psi_D$. Explicitly, this says for for every morphism $f : D_1 \longrightarrow D_2$

in **D** this diagram is commutative:



As with all universal mapping objects limits are essentially unique.

Proposition V.16 *Let $F : \mathbf{D} \longrightarrow \mathscr{C}$ be a functor. If $\phi : L \longrightarrow F$ and $\phi' : L' \longrightarrow F$ are both limits of F, then there is a unique isomorphism $u : L \longrightarrow L'$ such that $\phi' u = \phi$.*

Proof: The proof is much the same as every other uniqueness proof for universal mapping properties. As L' is a limit of F and ϕ is a cone over F there is a unique morphism $u : L' \longrightarrow L$ such that $\phi' u = \phi$. And as L is a limit of F and ϕ' is a cone over F there is a unique morphism $v : L' \longrightarrow L$ such that $\phi v = \phi'$. Next we note that $\phi' uv = \phi v = \phi'$ whence uv must be 1_L as that is the unique morphism which composed with ϕ gives ϕ. Finally as $\phi vu = \phi' u = \phi$ we equally well conclude that $vu = 1_{L'}$. ∎

As with other universal mapping objects we will use this to justify referring to *the* limit of the functor F, and we will write $\varprojlim F$ for the object of the cone over F that is its limit. For each object D of **D** we will write $\pi_D : \varprojlim F \longrightarrow F(D)$ for the component of the natural transformation $\pi : \varprojlim F \longrightarrow F$ that is the remaining part of the limit.

Examples

1. **Final Object**

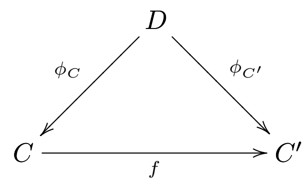
   The **0** empty category is the initial category, i.e., there is a unique (empty!) functor from **0** to any other category $\mathscr{C}$. A cone over the empty functor is nothing but an object of $\mathscr{C}$ as a natural transformation from C to that empty functor is vacuous. So a universal cone in this case is just an object 1 where there is a unique morphism from every object to 1, i.e., a final object in $\mathscr{C}$.

2. **Initial Object**

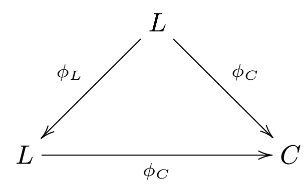
   Curiously initial objects are also limits in an even more extreme manner, as the limit of the identity functor $1_{\mathscr{C}}$.

Proposition V.17 *If* $\varprojlim 1_{\mathscr{C}}$ *exists in* $\mathscr{C}$*, then* $\varprojlim 1_{\mathscr{C}}$ *is an initial object in* $\mathscr{C}$*, and if* $\mathscr{C}$ *has an initial object then it is a limit for* $1_{\mathscr{C}}$*.*

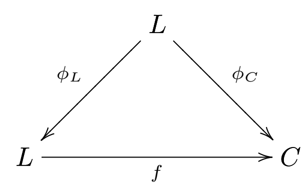
Proof: Recall that a cone over the identity functor $1_{\mathscr{C}}$ is a natural transformation $\psi : D \longrightarrow 1_{\mathscr{C}}$ from some constant functor D selecting the corresponding object in $\mathscr{C}$. This is a family of morphisms $\psi_C : D \longrightarrow C$ having the property that for every morphism $f : C \longrightarrow C'$ the triangle



commutes. A limit for $1_{\mathscr{C}}$ is a universal cone (L, ϕ) which is final in $\mathscr{C}/1_{\mathscr{C}}$. Consider in particular the triangle



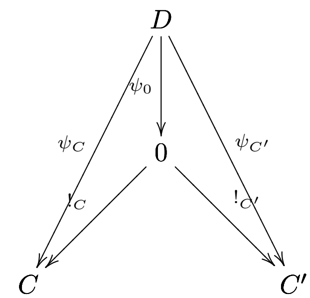
For every object C this commutes just because ϕ is a cone over $1_{\mathscr{C}}$. But this says that $\phi\phi_L$ is not only a cone over $1_{\mathscr{C}}$ but is the same cone as ϕ from which we conclude that ϕ_L must be 1_L. Now considering



with f an arbitrary morphism from L to C we see that $\phi_C = f\phi_L = f$ as $\phi_L = 1_L$. So for each object C there is exactly one morphism (ϕ_C) from L to C which is just what it means for L to be an initial object of $\mathscr{C}$.

For the other direction, if 0 is an initial object of $\mathscr{C}$, then we certainly have a cone over $1_{\mathscr{C}}$ given by $! : 0 \longrightarrow 1_{\mathscr{C}}$ with $!_C$ being the unique morphism from 0 to C. We also see that $!_0$ must be 1_0 as there is exactly one morphism from 0 to 0.

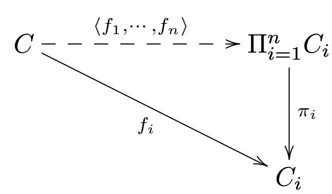
Now suppose $\psi : D \longrightarrow 1_{\mathscr{C}}$ is a cone over $1_{\mathscr{C}}$, then in particular we have $\psi_0 : D \longrightarrow 0$ which is a cone morphism from ψ to $!$, i.e., the following diagram commutes:



In order to show that 0 is the limit of $1_{\mathscr{C}}$ we have to verify that ψ_0 is the *unique* such morphism. But if $g : D \longrightarrow 0$ is such a morphism then in particular we have $\psi_0 = !_0 g = g$. $\blacksquare$

3. **Product**

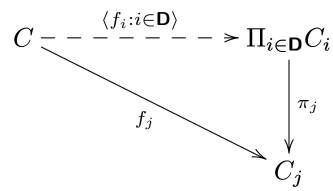
When $\mathbf{D}$ is a finite discrete category with objects $1, \cdots, n$, a functor from $\mathbf{D}$ to $\mathscr{C}$ simply selects n objects $C_1, \cdots, C_n$. A cone over the functor selecting $(C_1, \cdots, C_n)$ is just a family of morphisms $f_i : C \longrightarrow C_i$ and a limit (universal cone) is an object $\Pi_{i=1}^n C_i$ and a family of morphisms $\pi_1, \cdots, \pi_n$ satisfying the universal mapping property that for any family of morphisms $f_i : C \longrightarrow C_i$ there exists a *unique* morphism $\langle f_1, \cdots, f_n \rangle : C \longrightarrow \Pi_{i=1}^n C_i$ with $\pi_i \langle f_1, \cdots, f_n \rangle = f_i$.



Of course as advertised this is exactly the universal mapping property for the product of $C_1, \cdots, C_n$ in $\mathscr{C}$.

The empty category is actually the discrete category with zero objects, so this is yet again reinforcing the remark that the final object is just the empty product!

While the above discussion assumed that $\mathbf{D}$ was a finite discrete category, that was relevant only to connecting the limit of a functor from $\mathbf{D}$ to $\mathscr{C}$ with the prior discussion of products. This allows us to *define* the product of an arbitrary family of objects in $\mathscr{C}$ to be the limit of a functor from some discrete category which selects those objects. From this point on we will freely use diagrams such as the following:



Of course just because we can define products of arbitrary families of objects, that says noting about existence of such product. In particular while assuming all products of two objects exists guarantees that all finite products exists, it says nothing about existence of products of infinite families. This will be discussed further in section V.5.3

Proposition V.18 *Suppose that in the category $\mathscr{C}$ all products exist, i.e., for every functor $F : \mathscr{D} \longrightarrow \mathscr{C}$ with $\mathscr{D}$* any *discrete category, the product $\varprojlim F$ exists, then $\mathscr{C}$ is preordered.*

Proof: Suppose that $f, g : B \longrightarrow C$ in $\mathscr{C}$, and consider product $P = \Pi_{\mathscr{D}} C$ where $\mathscr{D}$ is a proper class. For each D in $\mathscr{D}$ take $f_D : B \longrightarrow P$ to be the unique morphism such that $\pi_D f_D = f$ and $\pi_{D'} f_D = g$ for $D' \neq D$. If g is not equal to f, then each of the f_D is distinct and so $\mathrm{Hom}(B, P)$ is a proper class contrary to the definition of a category. Whence we conclude that any parallel morphisms are equal, which is the definition of a preordered category! ∎

4. **Kernel**

Whenever $\mathscr{C}$ is a category with a zero object, we can define the kernel of a morphism.

Definition V.9: A **kernel** of $f : C \longrightarrow D$ is a morphism $k : K \longrightarrow C$ with

a) $fk = 0$, and
b) If g is any morphism with $fg = 0$, then there is a unique morphism $\overline{g}$ so that $g = k\overline{g}$.

As always with universal mapping properties the item of interest is the morphism k (which determines K), but it is K which is named as $\mathrm{Ker}(f)$. Following the usual practice we will often write of "the kernel $\mathrm{Ker}(f)$ of f", but it must be understood that this tacitly includes the specific morphism $k : \mathrm{Ker}(f) \longrightarrow C$. Of course the reason for this is that in the most familiar categories the kernel of a morphism is a subset of the domain and the morphism k is just the inclusion

In general a morphism may not have a kernel, but if it does it is unique in the strong sense that when $k : K \longrightarrow C$ and $k' : K' \longrightarrow C$ are two kernels of $f : C \longrightarrow D$, there is a unique isomorphism $\overline{k} : K \longrightarrow K'$ with $k'\overline{k} = k$. The inverse being $\overline{k'} : K' \longrightarrow K$.

An easy but important observation is that every kernel is a monomorphism.

Proposition V.19 *In any category with zero object, if* $k : \mathrm{Ker}(f) \longrightarrow C$ *is a kernel, then it is a monomorphism.*

Proof: Suppose g_1 and g_2 are morphisms to $\mathrm{Ker}(f)$ and $kg_1 = kg_2$. Then fkg_1 and fkg_2 are both zero morphisms, so there is a unique morphism g so that $kg = kg_1$. But both g_1 and g_2 have that property so they must be equal. ∎

While every kernel is a monomorphism, the converse is not true.

Exercise V.2. Give an example of a monomorphism that is not a kernel. (Hint: Consider the inclusion of $\mathbb{N}$ into $\mathbb{Z}$ in the category of monoids.)

We've previously noted that the category theory notion of monomorphism does not capture all the meaning associated with injective homomorphisms. This is part of that and leads to the following definitions.

Definition V.10: In a category with zero, a **normal monomorphism** is a morphism that is the kernel of some morphism.

Definition V.11: A **normal category** is a category with zero in which every monomorphism is normal.

The most familiar examples of normal categories are the categories of modules, including the category of Abelian groups and the category of vector spaces over a particular field. These are all example of Abelian categories and will be discussed at length in Chapter XIII.

Examples of kernels abound. In the category of groups there is the usual notion of the kernel of a group homomorphism $h : G \longrightarrow G'$, namely $\mathrm{Ker}(h) = \{g \in G : h(g) = 1\}$ where 1 is the identity element in G'. Then the inclusion morphism $i : \mathrm{Ker}(h) \longrightarrow G$ is a kernel of h according to the above definition.

Similarly in the category of modules over a particular ring there is again the usual notion of the kernel of a module homomorphism $h : M \longrightarrow M'$ as $\mathrm{Ker}(h) = \{m \in M : h(m) = 0\}$ where 0 is the zero element in M'. Again the inclusion morphism $i : \mathrm{Ker}(h) \longrightarrow M$ is a kernel of h as above.

This last includes kernels for homomorphisms of Abelian groups ($\mathbb{Z}$-modules) and null spaces for linear transformations (K-modules for the base field K.) At the same time it is really a special case of the kernel of a group homomorphism as all that is being used here is the kernel of the module homomorphism considered as a homomorphism of (Abelian) groups. In Chapter XIII on Abelian Categories, we will see that in a certain sense these are all the interesting cases, but we will also see the value of the abstraction.

A more general case is the category of monoids. Again conventionally the kernel of a monoid homomorphism $h : M \longrightarrow M'$ is $\text{Ker}(h) = \{m \in M : h(m) = 1\}$ where 1 is the identity element in M'. Again the inclusion morphism $i : \text{Ker}(h) \longrightarrow M$ is a kernel of h as above. This includes the kernel of group homomorphisms as a special case, but we will see later that this illustrates that the notion of a kernel is only useful in special cases.

There are numerous other cases in algebra where it is common to speak of kernels that are not captured by this definition. For example in the category of commutative rings it is usual to define the kernel of a ring homomorphism $f : R \longrightarrow S$ as $\text{Ker}(f) = \{r \in R : f(r) = 0\}$. But this is an *ideal* of R rather than a subring. Indeed the category **CommutativeRing** has no zero object (as the rings in **CommutativeRing** all have 1 different from 0), so the definition of a kernel given in this section simply does not apply.

5. **Equalizer**

Kernels are specializations of the more general notion of equalizers which concerns two morphisms between the same objects.

Definition V.12: Two morphisms are **parallel** if they have the same domain and the same codomain. This is usually written symbolically as $\bullet \underset{g}{\overset{f}{\rightrightarrows}} \bullet$ or just $\underset{g}{\overset{f}{\rightrightarrows}}$.

Definition V.13: An **equalizer** of parallel morphisms f and g is a morphism k with

a) $fk = gk$, and
b) If h is any morphism with $fh = gh$, then there is a unique morphism $\overline{h}$ so that $h = k\overline{h}$.

Kernels are special cases of equalizers where one of the two parallel morphisms is a zero morphism. Just as with kernels, equalizers need not exist, but if they do exists any two equalizers of the same parallel pair are isomorphic via the unique morphisms between guaranteed by the definition.

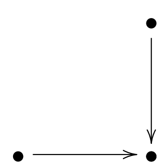
Besides the examples of kernels, equalizers also arise in the category of sets and related categories. In **Set** if $f, g : X \longrightarrow Y$ are two parallel functions, then taking $K \longrightarrow \{x \in X : f(x) = g(x)\}$ and $k : K \longrightarrow X$ the inclusion function exhibits k as an equalizer of f and g. This same construction serves in the category **Top** of topological spaces and continuous maps where K is simply considered as a subspace of X.

Equalizers are also called *difference kernels* because in certain situations they arise as the kernel of the difference of two morphisms. For example in all categories of modules the Hom-set natural structure of an Abelian group, so for any pair of morphisms $f, g : M \longrightarrow N$ there is the morphism $f - g : M \longrightarrow N$ with $f - g(m) = f(m) - g(m)$. And in this case the kernel of $f - g$ is an equalizer of f and g. This situation prevails in all preadditive categories as will be discussed in Chapter XIII.

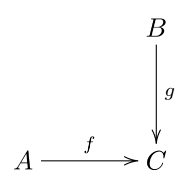
Just as every kernel is a monomorphism, so is every equalizer and with essentially the same proof. Suppose that $k : K \longrightarrow C$ is an equalizer of $f, g : C \longrightarrow D$ and h_1, h_2 are morphisms to $\mathrm{Ker}(f)$ with $kh_1 = kh_2$. Then $fkh_1 = gkh_1$, so there is a *unique* morphism h so that $kh = kh_1$. But both h_1 and h_2 have that property so they must be equal. ∎

6. **Pullback**

Pullbacks were defined back on page 118 as examples of limits, and here is that remark made precise. When $\mathbf{D}$ is the category



a functor F from $\mathbf{D}$ to $\mathscr{C}$ selects three objects A, B and C in $\mathscr{C}$ together with two morphisms $f : A \longrightarrow C$ and $g : B \longrightarrow C$, i.e., the diagram



and the limit of F is exactly the pullback of f and g as previously defined.

Equalizers are just pullbacks where the morphisms are parallel, i.e., they have the same domain as well as the same codomain.

In the category of sets the canonical pullback of f and g is the subset

$$A \times_C B = (a, b) \in A \times B such that f(a) = g(b),$$

of $A \times B$ together with the restrictions of the projection maps π_1 and π_2 to $A \times_C B$.

This example motivates another way of characterizing the pullback: as the equalizer of the morphisms $f\pi_1, g\pi_2 : A \times_C B \rightrightarrows C$.

> **Exercise V.3.** Verify that in any category with finite products, pullbacks and equalizers, the pullback of $f : A \longrightarrow C$ and $g : B \longrightarrow C$ is the equalizer of $f\pi_1$ and $g\pi_2$.

This shows that pullbacks exist in any category with binary products and equalizers.

Clearly in any category with a terminal object 1, the pullback $A \times_1 B$ is just the product $A \times B$.

Every square

$$\begin{array}{ccc} P & \xrightarrow{\;p\;} & X \\ {\scriptstyle q}\downarrow & & \downarrow{\scriptstyle f} \\ Y & \xrightarrow[\;g\;]{} & Z \end{array}$$

where (P, p, q) is a pullback of (f, g) is called a *Cartesian square.*

Proposition V.20 *Any morphism $f : X \longrightarrow Y$ is a monomorphism iff the commutative square*

$$\begin{array}{ccc} X & \xrightarrow{\;1_X\;} & X \\ {\scriptstyle 1_X}\downarrow & & \downarrow{\scriptstyle f} \\ X & \xrightarrow[\;f\;]{} & Y \end{array}$$

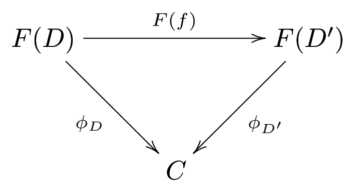
is Cartesian.

V.5.2 Cocones and Colimits

Everything in the previous section has a dual, but these dual notions are sufficiently common and important that this section provides all of the definitions and results explicitly, though the proofs follow from the dual results so nothing further need be said.

The dual of a cone is a cocone, again a family of commuting triangles but now pointing down.

Definition V.14: Let $F : \mathbf{D} \longrightarrow \mathscr{C}$ be a functor. A **cocone** over F is an object C of $\mathscr{C}$ and a natural transformation $\phi : F \longrightarrow C$ from F to the constant functor C. This is the same as saying that for every morphism $f : D \longrightarrow D'$ of $\mathscr{D}$ this triangle commutes:

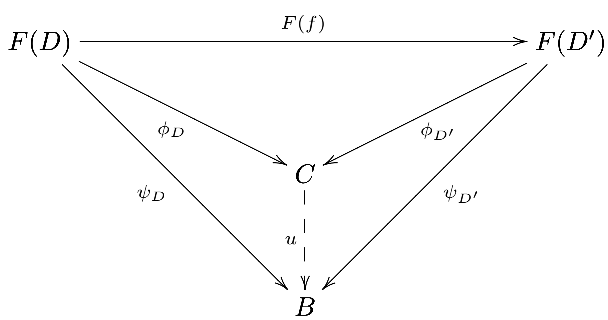


A morphism from the cocone (C, ϕ) to the cocone (C', ϕ') is a morphism $f : C \longrightarrow C'$ with $\phi f = \phi'$.

Clearly there is a category of cocones over F with objects and morphism as in the preceding definition. We will denote this category by $F/\mathscr{C}$.

Dual to the notion of a limit is a colimit which is a universal cocone just as a limit is a universal cocone..

Definition V.15: Let $F : \mathbf{D} \longrightarrow \mathscr{C}$ be a functor. A **colimit** for F is an initial object (C, ϕ) in $F/\mathscr{C}$. This is also called a *universal cocone* over F and is a cocone $\phi : F \longrightarrow C$ with the universal mapping property that for any cocone $\psi : F \longrightarrow B$ there is a unique morphism $u : C \longrightarrow B$ so $u\phi = \psi$. Explicitly, this says for for every morphism $f : D \longrightarrow D'$ in $\mathbf{D}$ this diagram is commutative:



Proposition V.21 *Let $F : \mathbf{D} \longrightarrow \mathscr{C}$ be a functor. If $\phi : F \longrightarrow C$ and $\psi : F \longrightarrow B$ are both colimits of F, then there is a unique isomorphism $u : B \longrightarrow C$ such that $u\phi = \psi$.*

As with other universal mapping objects we will use this to justify referring to *the* colimit of the functor F, and we will write $\varinjlim F$ for the object of the cocone over F that is its limit. We will write $\iota : F \longrightarrow \varinjlim F$ for the natural transformation that is the remaining part of the colimit.

Examples

1. **Initial Object**

   The **0** empty category is the initial category, i.e., there is a unique (empty!) functor from **0** to any other category $\mathscr{C}$. A cocone over the empty functor is nothing but an object of $\mathscr{C}$ as a natural transformation from C to that empty functor is vacuous. So a universal cocone in this case is just an object 0 where there is a unique morphism from every object to 0, i.e., an initial object in $\mathscr{C}$.

2. **Final Object**

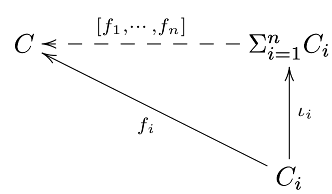
   A final objects is the colimit of the identity functor $1_{\mathscr{C}}$, and conversely.

   Proposition V.22 *If* $\varinjlim 1_{\mathscr{C}}$ *exists in* $\mathscr{C}$*, then* $\varinjlim 1_{\mathscr{C}}$ *is a final object in* $\mathscr{C}$*, and if* $\mathscr{C}$ *has a final object 1 then it is the colimit of* $1_{\mathscr{C}}$*.*

3. **Sum**

   Sums are the duals of products, and they occur as colimits of functors on discrete categories.

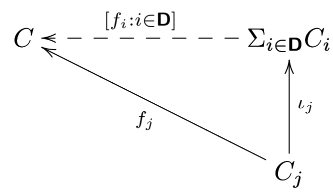
   When D is a finite discrete category with objects $1, 2, \cdots, n$, a functor from D to $\mathscr{C}$ simply selects n objects $C_1, \ldots, C_n$. A cocone over the functor selecting $(C_1, \cdots, C_n)$ is just a family of morphisms $f_i : C_i \longrightarrow C$ and a colimit (universal cocone) is an object $\Sigma_{i=1}^n C_i$ and a family of morphisms $\iota_1, \cdots, \iota_n$ satisfying the universal mapping property that for any family of morphisms $f_i : C_i \longrightarrow C$ there exists a *unique* morphism $[f_1, \cdots, f_n] : \Sigma_{i=1}^n C_i \longrightarrow C$ with $[f_1, \cdots, f_n]\iota_i = f_i$.



   Of course as advertised this is exactly the universal mapping property for the sum of $C_1, \cdots, C_n$ in $\mathscr{C}$.

   The empty category is actually the discrete category with zero objects, so this is yet again reinforcing the remark that the initial object is just the empty sum!

   While the above discussion assumed that D was a *finite* discrete category, that was relevant only to connecting the colimit of a functor from D to $\mathscr{C}$ with the prior discussion of sums. This allows us to *define* the sum of an arbitrary family of objects in $\mathscr{C}$ to be the colimit of a functor from some discrete category which selects those objects. From this point on we will freely use diagrams such as the following:



Of course just because we can define sums of arbitrary families of objects, that says noting about existence of such product. In particular while assuming all sums of two objects exists guarantees that all finite sums exists, it says nothing about existence of sums of infinite families. This will be discussed further in section V.5.3.

4. **Cokernel**

Dual to kernels are cokernels, and just as with kernels they can be defined in any category that has a zero object (a self-dual notion.)

Definition V.16: In a category with a zero object, a **cokernel** of $f : E \longrightarrow D$ is a morphism $c : D \longrightarrow C$ with

a) $cf = 0$, and
b) If g is any morphism with $gf = 0$, then there is a unique morphism $\overline{g}$ so that $g = \overline{g}c$.

As always with universal mapping properties the item of interest is the morphism c (which determines C), but it is C which is named as Coker(f). Following the usual practice we will often speak of "the cokernel Coker(f) of f", but it must be understood that this tacitly includes the specific morphism $c : D \longrightarrow \text{Coker}(f)$.

In general a morphism may not have a cokernel, but if it does it is unique in the strong sense that when $c : D \longrightarrow C$ and $c' : D \longrightarrow C'$ are two cokernels of $f : E \longrightarrow D$, there is a unique isomorphism $\overline{c} : C \longrightarrow C'$ with $\overline{c}c' = c$. The inverse being $\overline{c'} : C' \longrightarrow C$.

Of course just as every every kernel is a monomorphism, every cokernel is an epimorphism, a fact we record as a proposition.

Proposition V.23 *In any category with zero object, if* $c : B \longrightarrow \text{Coker}(f)$ *is a cokernel, then it is an epimorphism.*

While every cokernel is an epimorphism, the converse is not true.

Exercise V.4. Give an example of an epimorphism that is not a cokernel. (Hint: Consider the inclusion of $\mathbb{N}$ into $\mathbb{Z}$ in the category of monoids.)

We've previously noted that the category theory notion of epimorphism does not capture all the meaning associated with surjective homomorphisms. This is part of that and leads to the following definitions.

Definition V.17: A morphism in a category with zero is a **normal epimorphism** provided it is the cokernel of some morphism.

Definition V.18: A **conormal category** is a category in which every epimorphism is conormal.

The most familiar examples of conormal categories are all categories of modules, including the category of Abelian groups and the category of vector spaces over a particular field. These are all example of Abelian categories and will be discussed at length in Chapter XIII.

In the category of modules over a ring there is again the notion of the cokernel of a module homomorphism being the quotient of the codomain by the image. This includes cokernels for homomorphisms of Abelian groups ($\mathbb{Z}$-modules) and cokernel for linear transformations (K-modules for the base field K.) At the same time it is really a special case of the cokernel of a group homomorphism as all that is being used here is the cokernel of the module homomorphism considered as a homomorphism of (Abelian) groups. In Chapter XIII on Abelian Categories, we will see that in a certain sense these are all the interesting cases.

5. **Coequalizer**

Cokernels are specializations of the more general notion of coequalizers which concerns two parallel morphisms.

Definition V.19: An **coequalizer** of parallel morphisms f and g is a morphism c with

a) $cf = cg$, and
b) If h is any morphism with $hf = hg$, then there is a unique morphism $\overline{h}$ so that $h = \overline{h}c$.

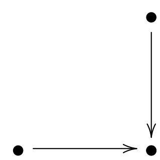
Cokernels are special cases of equalizers where one of the two parallel morphisms is a zero morphism. Just as with cokernels, equalizers need not exist, but if they do exists any two coequalizers of the same parallel pair are isomorphic via the unique morphisms between guaranteed by the definition.

Coequalizers are also called *difference cokernels* because in certain situations they arise as the kernel of the difference of two morphisms. For example in all categories of modules the Hom-set is naturally an Abelian group, so for any pair of morphisms $f, g : M \longrightarrow N$ there is the morphism $f - g : M \longrightarrow N$ with $f - g(m) = f(m) - g(m)$. And in this case the cokernel of $f - g$ is an coequalizer of f and g. This situation prevails in all preadditive categories as will be discussed in Chapter XIII.

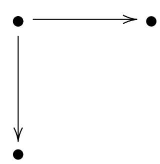
Just as every cokernel is an epimorphism, so is every coequalizer.

6. **Pushout**

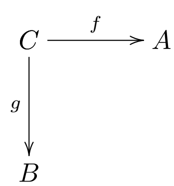
Pushout are the duals of pullbacks and provide a good example of the value of alternative views of duality. The diagram category for a pullback is



so a pushout in the category $\mathscr{C}$ can be described as the limit of a functor from this diagram category to $\mathscr{C}^{\text{op}}$. But that diagram can equally well be considered as a functor from the dual of the diagram category to $\mathscr{C}$, and usually we consider the diagram category **D**



and specific diagrams like



So the pushout $A \times_C B$ of f and g is the colimit of the functor from **D** to $\mathscr{C}$ that selects f and g.

By duality, exercise V.3 shows that in any category with finite sums and coequalizers, the pushout $A +_C B$ of $f : C \longrightarrow A$ and $g : C \longrightarrow B$ is the coequalizer of $\iota_1 f$ and $\iota_2 g$.

This is the abstract justification of the construction of pushouts in **Set** given on page 119.

Dual to the results for pullbacks, this shows that pushouts exist in any category with binary sums and coequalizers, and that in a category with an initial object 0, the pushout $X +_0 Y$ is just the sum $X + Y$.

Every square

$$\begin{array}{ccc} Y & \xrightarrow{i} & P \\ {\scriptstyle f}\uparrow & & \uparrow{\scriptstyle j} \\ X & \xrightarrow[g]{} & Z \end{array}$$

where (P, p, q) is a pushout of (f, g) is called a *coCartesian square.*

Proposition V.24 *Any morphism* $f : X \longrightarrow Y$ *is an epimorphism iff the commutative square*

$$\begin{array}{ccc} Y & \xrightarrow{1_Y} & Y \\ {\scriptstyle f}\uparrow & & \uparrow{\scriptstyle 1_Y} \\ X & \xrightarrow[f]{} & Y \end{array}$$

is coCartesian.

Of course this is the dual of proposition V.20, so no additional proof is needed.

V.5.3 Complete Categories

Definition V.20: A category is *complete* if it has all small limits.

Theorem V.10 *small products and kernels implies complete*

Definition V.21: A category is *cocomplete* if it has all small colimits.

Theorem V.11 *small sums and cokernels implies cocomplete*

V.6 Adjoint Functors

“Adjoints are everywhere”

That slogan, useful for anyone studying categories, has already been reinforced by the examples of adjoint functors provided in a number of earlier exercises. See exercises III.11, III.15, III.16, III.18, III.25, III.27, III.31, and III.34.

No definition has been provided to this point, so this section will provide the definition, including a number of alternative formulations, and some basic theory about adjoint functors.

The basic context for adjoints consists of two functors $U : \mathscr{M} \longrightarrow \mathscr{S}$ and $F : \mathscr{S} \longrightarrow \mathscr{M}$. [Throughout this section we will generally use these names as they are intended to suggest the underlying and free functors between the category of modules (or monoids) and the category of sets.]

The basic definition is in terms of natural transformations $\varepsilon : 1_{\mathscr{S}} \longrightarrow UF$ and $\eta : FU \longrightarrow 1_{\mathscr{M}}$, while the memorable formulation is in terms of the bifunctors $\mathscr{M}(F(\bullet), \bullet)$ and $\mathscr{S}(\bullet, U(\bullet))$ both from $\mathscr{S}^{\mathrm{op}} \times \mathscr{M}$ to **Set**.

As bifunctors these are both functors from $\mathscr{S}^{\mathrm{op}} \times \mathscr{M}$ to **Set**. So for each pair of morphisms $(f : T \longrightarrow S, h : M \longrightarrow N)$ we have the induced functions $\mathscr{M}(F(f), h)$ and $\mathscr{S}(f, U(h))$. This notation is so cumbersome that we will usually use the notation $(F(f), h)_*$ and $(f, U(h))_*$ in its place. Notice that for example $(f, U(h))_*(g) = U(h)gf$ so that this is a direct extension of the notation first introduced in definitions I.7 and I.8. In particular in the case that f is an identity morphism, we will write h_* instead of $\mathscr{M}(F(S), h)$ and $U(h)_*$ instead of $\mathscr{S}(S, U(h))$. Similarly when h is an identity morphism, we will write $F(f)^*$ instead of $\mathscr{M}(F(f), M)$ and f^* instead of $\mathscr{S}(f, U(M))$.

The first observation is that the natural transformations ε and η induce natural transformations $\phi : \mathscr{M}(F(\bullet), \bullet) \longrightarrow \mathscr{S}(\bullet, U(\bullet))$ and $\psi : \mathscr{S}(\bullet, U(\bullet)) \longrightarrow \mathscr{M}(F(\bullet), \bullet)$ defined as follows.

If $s \in \mathscr{S}(S, U(M))$, then $\phi_{S,M}(s) = \eta_M F(s)$

$$\begin{array}{ccc} F(S) & \xrightarrow{F(s)} & FU(M) \\ & \underset{\phi_{S,M}(s)}{\searrow} & \downarrow{\scriptstyle \eta_M} \\ & & M \end{array}$$

while for $m \in \mathscr{M}(F(S), M)$, the definition is $\psi_{S,M}(m) = U(m)\varepsilon$

$$\begin{array}{ccc} UF(S) & \xrightarrow{U(m)} & U(M) \\ {\scriptstyle \varepsilon_S}\uparrow & \overset{\psi_{S,M}(m)}{\nearrow} & \\ S & & \end{array}$$

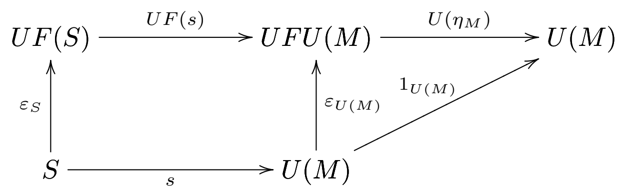
To verify that ϕ and ψ are natural transformations requires checking that for all morphisms $t : T \longrightarrow S$ in $\mathscr{S}$ and $n : M \longrightarrow N$ in $\mathscr{M}$ the following are commutative squares.

$$\begin{array}{ccc} \mathscr{M}(F(S), M) & \xrightarrow{\phi_{S,M}} & \mathscr{S}(S, U(M)) \\ {\scriptstyle (F(t),n)_*}\downarrow & & \downarrow{\scriptstyle (t,U(n))_*} \\ \mathscr{M}(F(T), N) & \xrightarrow[\phi_{T,N}]{} & \mathscr{S}(T, U(N)) \end{array} \qquad \begin{array}{ccc} \mathscr{S}(S, U(M)) & \xrightarrow{\psi_{S,M}} & \mathscr{M}(F(S), M) \\ {\scriptstyle (t,U(n))_*}\downarrow & & \downarrow{\scriptstyle (F(t),n)_*} \\ \mathscr{S}(T, U(N)) & \xrightarrow[\psi_{T,N}]{} & \mathscr{M}(F(T), N) \end{array}$$

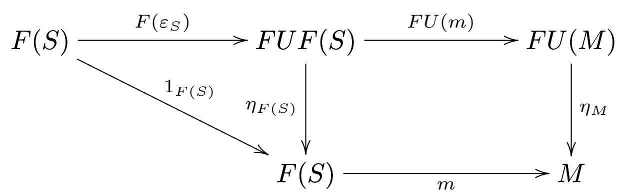
This is easily done by explicit calculation. In the first square $(t, U(n))_*\phi_{S,M}(m) = U(n)U(m)\varepsilon_S t$ and $\phi_{T,N}(F(t), n)_*(m) = U(n)U(m)UF(t)\varepsilon_T$. But as ε is a natural transformation we have $\varepsilon_S t = UF(t)\varepsilon_T$ and so the square is commutative.

Similarly in the second square $(F(t), n)_*\psi_{S,M}(s) = n\eta_M F(s)F(t)$ and $\psi_{T,N}(t, U(n))_*(s) = \eta_N FU(n)F(s)F(t)$. But as η is a natural transformation we have $\eta_N FU(n) = n\eta_N$ and so the square is commutative.

The second result is that when $(U\eta)(\varepsilon U) = 1_U$ and $(F\varepsilon)(\eta F) = 1_F$ the above natural transformations ϕ and ψ are natural isomorphisms that are inverse to one another. The first assumption gives the commutative diagram



and shows that $\phi_{S,M}\psi_{S,M}(s) = U(\eta_M)UF(s)\varepsilon_S = s$, while the second assumption gives the commutative diagram



and shows that $\psi_{S,M}\phi_{S,M}(m) = \eta_M FU(m)F(\varepsilon_S) = m$.

This result leads to the following definition and theorem.

Definition V.22: An **adjunction** between two functors $U : \mathscr{M} \longrightarrow \mathscr{S}$ and $F : \mathscr{S} \longrightarrow \mathscr{M}$ consists of two natural transformations $\eta : FU \longrightarrow 1_{\mathscr{M}}$ and $\varepsilon : 1_{\mathscr{S}} \longrightarrow UF$ satisfying

$$U \xrightarrow{\varepsilon U} UFU \xrightarrow{U\eta} U = 1_U$$

$$F \xrightarrow{F\varepsilon} FUF \xrightarrow{\eta F} F = 1_F$$

We write $F \dashv U$ when there is an adjunction between F and U. This is also described by saying that F is the **left adjoint** of U and U is the **right adjoint** of F. The natural transformation η is the **unit** of the adjunction, while ε is the **counit** of the adjunction.

The notation adjoint and coadjoint has also been used, but it was never generally agreed which was which, so the right/left terminology is dominant.

The discussion of the second result above is a detailed statement and proof of the following theorem.

Theorem V.12 *An adjunction $F \dashv U$ induces a natural equivalence between the bifunctors $\mathscr{M}(F(\bullet), \bullet)$ and $\mathscr{S}(\bullet, U(\bullet))$.*

The converse of this theorem is true as well.

Theorem V.13 *If the bifunctors $\mathscr{M}(F(\bullet), \bullet)$ and $\mathscr{S}(\bullet, U(\bullet))$. are naturally equivalent, then $F \dashv U$.*

Proof: This is our third result and just as with the first two most of the interest is in the details. Suppose that $\phi : \mathscr{M}(F(\bullet), \bullet) \longrightarrow \mathscr{S}(\bullet, U(\bullet))$ is a natural equivalence with $\psi : \mathscr{S}(\bullet, U(\bullet)) \longrightarrow \mathscr{M}(F(\bullet), \bullet)$ its inverse. Then for each object S of $\mathscr{S}$ we can define $\varepsilon_S = \phi_{S,F(S)}(1_{F(S)}) \in \mathscr{S}(S, UF(S))$, and for each object M of $\mathscr{M}$ we can define $\eta_M = \psi_{U(M),M}(1_U(M)) \in \mathscr{M}(FU(M), M)$.

Starting with $h : F(S) \longrightarrow M$, the commutative diagram

$$\begin{array}{ccc}
\mathscr{M}(F(S), F(S)) & \xrightarrow{\phi_{S,F(S)}} & \mathscr{S}(S, UF(S)) \\
\downarrow{\scriptstyle h_*} & & \downarrow{\scriptstyle U(h)_*} \\
\mathscr{M}(F(S), M) & \xrightarrow[\phi_{S,M}]{} & \mathscr{S}(S, U(M))
\end{array}$$

gives us that $\phi_{S,M}(h) = U(h)\varepsilon_S$ by considering the image of $1_{F(S)} \in \mathscr{M}(F(S), F(S))$ along the two different paths from $\mathscr{M}(F(S), F(S))$ to $\mathscr{S}(S, U(M))$.

In the same way, starting with $f : S \longrightarrow U(M)$, the commutative diagram

$$\begin{array}{ccc}
\mathscr{S}(U(M), U(M)) & \xrightarrow{\psi_{U(M),M}} & \mathscr{M}(FU(M), M) \\
\downarrow{\scriptstyle f^*} & & \downarrow{\scriptstyle F(f)^*} \\
\mathscr{S}(S, U(M)) & \xrightarrow[\psi_{S,M}]{} & \mathscr{M}(F(S), M)
\end{array}$$

gives us that $\psi_{S,M}(f) = \eta_M F(f)$ by considering the image of $1_{U(M)} \in \mathscr{S}(U(M), U(M))$ along the two different paths from $\mathscr{S}(U(M), U(M))$ to $\mathscr{M}(F(S), M)$.

That ε is a natural transformation from $1_{\mathscr{S}}$ to UF and η is a natural transformation from FU to $1_{\mathscr{M}}$ follows from the commutative square:

$$\begin{array}{ccc}
\mathscr{M}(F(T), F(T)) & \xrightarrow{\phi_{T,F(T)}} & \mathscr{S}(T, UF(T)) \\
\downarrow{\scriptstyle F(f)^*} & & \downarrow{\scriptstyle f^*} \\
\mathscr{M}(F(S), F(T)) & \xrightarrow[\phi_{S,F(T)}]{} & \mathscr{S}(S, UF(T))
\end{array}$$

Just note that the path along the top and right takes $1_{F(T)}$ to ε_T and then to $f\varepsilon_T$, while the path down and left take $1_{F(T)}$ to F(f) and then, using the above characterization of ϕ, to $UF(f)\varepsilon_S$. The equality of these two is exactly what is needed to verify the naturality of ε.

The naturality of $\eta : FU \longrightarrow 1_{\mathscr{M}}$ follows from the commutativity of the following square in essentially the same fashion.

$$\begin{array}{ccc}
\mathscr{S}(U(M), U(M)) & \xrightarrow{\psi_{U(M),M}} & \mathscr{M}(FU(M), M) \\
{\scriptstyle U(h)_*}\big\downarrow & & \big\downarrow{\scriptstyle h_*} \\
\mathscr{S}(U(M), U(N)) & \xrightarrow[\psi_{U(M),N}]{} & \mathscr{M}(FU(M), N)
\end{array}$$

■

While the natural equivalence "$\mathscr{M}(F(S), M) = \mathscr{S}(S, M)$" is much easier to remember than the definition in terms of ε and η, the definition of an adjunction has the advantage of applying to arbitrary categories, not just locally small categories.

Lawvere introduced comma categories (see section IV.1) as a way of capturing the "$\mathscr{M}(F(S), M) = \mathscr{S}(S, M)$" specification of adjoints without the restriction to locally small categories. His treatment was done entirely inside a suitable category of categories. A tainted version (because elements appear here) follows.

For any functor $F : \mathscr{S} \longrightarrow \mathscr{M}$ we also have the functor $1_{\mathscr{M}} : \mathscr{M} \longrightarrow \mathscr{M}$ and so the comma category $(F \downarrow \mathscr{M})$ (i.e., $(F \downarrow 1_{\mathscr{M}})$. Recall that the objects of the comma category are triples (S, f, M) where S is an object of $\mathscr{S}$, M is an object of $\mathscr{M}$, and $f : F(S) \longrightarrow M$ is a morphism in $\mathscr{M}$. A morphism from (S, f, M) to (S', f', M') is a pair of morphisms (s, m) with $s : S \longrightarrow S'$, $m : M \longrightarrow M'$ and $mf = f'F(s)$. This is summarized in the following diagram:

$$\text{object } (S, f, M) : \quad \begin{array}{c} F(S) \\ {\scriptstyle f}\big\downarrow \\ M \end{array} \qquad\qquad \text{morphism } (s, m) : \quad \begin{array}{ccc} F(S) & \xrightarrow{F(s)} & F(S') \\ {\scriptstyle f}\big\downarrow & & \big\downarrow{\scriptstyle f'} \\ M & \xrightarrow[m]{} & M' \end{array}$$

The comma category $(F \downarrow \mathscr{M})$ is in many ways the bifunctor $\mathscr{M}(F(\bullet), \bullet)$ made into a category.

Similarly for any functor $U : \mathscr{M} \longrightarrow \mathscr{S}$ we also have the comma category $(\mathscr{S} \downarrow U)$ where the objects of the comma category are triples (S, g, M) where S is an object of $\mathscr{S}$, M is an object of $\mathscr{M}$, and $g : S \longrightarrow U(M)$ is a morphism in $\mathscr{S}$. A morphism from (S, g, M) to (S', g', M') is a pair of morphisms (s, m) with $s : S \longrightarrow S'$, $m : M \longrightarrow M'$ and $U(m)g = g's$. This is summarized in the following diagram:

$$\text{object } (S,g,M):\quad \begin{CD} S \\ @VgVV \\ U(M) \end{CD} \qquad\qquad \text{morphism } (s,m):\quad \begin{CD} S @>s>> S' \\ @VgVV @VV{g'}V \\ U(M) @>>{U(m)}> U(M') \end{CD}$$

Lawvere's insight was that $F \dashv U$ is almost equivalent to $(F \downarrow \mathscr{M}) \cong (\mathscr{S} \downarrow U)$. There is one additional detail – requiring that the following diagram is commutative

$$\begin{array}{ccc} (F \downarrow \mathscr{M}) & \xrightarrow{\theta} & (\mathscr{S} \downarrow U) \\ & \searrow \quad \swarrow & \\ & \mathscr{S} \times \mathscr{M} & \end{array}$$

where the two functors to $\mathscr{S} \times \mathscr{M}$ are the canonical functors discussed in connection with the Universal Mapping Property of comma categories (page 109).

In more detail, we have projection functors from $(F \downarrow \mathscr{M})$ to $\mathscr{S}$ and to $\mathscr{M}$. Taken together they define a canonical functor from $(F \downarrow \mathscr{M})$ to $\mathscr{S} \times \mathscr{M}$ which is given by taking each morphism (s,m) in $(F \downarrow \mathscr{M})$ to (s,m) as a morphism in $\mathscr{S} \times \mathscr{M}$. Equally well there is again a canonical functor from $(\mathscr{S} \downarrow U)$ to $\mathscr{S} \times \mathscr{M}$ and again it takes each morphism (s,m) in $(\mathscr{S} \downarrow U)$ to the same pair considered as a morphism in $\mathscr{S} \times \mathscr{M}$.

Here is the theorem.

Theorem V.14 *An isomorphism $\theta : (F \downarrow \mathscr{M}) \longrightarrow (\mathscr{S} \downarrow U)$ for which*

$$\begin{array}{ccc} (F \downarrow \mathscr{M}) & \xrightarrow{\theta} & (\mathscr{S} \downarrow U) \\ & \searrow \quad \swarrow & \\ & \mathscr{S} \times \mathscr{M} & \end{array}$$

commutes is equivalent to an adjunction between F and U. (The two functors to $\mathscr{S} \times \mathscr{M}$ are the canonical functors just defined.)

Proof: The proof of this theorem is gotten by adapting the proofs of Theorems V.12 and V.13.

A natural transformation $\eta : FU \longrightarrow 1_{\mathscr{M}}$ gives rise to the functor $\theta : (F \downarrow \mathscr{M}) \longrightarrow (\mathscr{S} \downarrow U)$ defined as taking the morphism

$$\begin{array}{ccc} F(S) & \xrightarrow{F(s)} & F(S') \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f'} \\ M & \xrightarrow[m]{} & M' \end{array}$$

in $(F \downarrow \mathscr{M})$ to the morphism

$$\begin{array}{ccc} S & \xrightarrow{s} & S' \\ {\scriptstyle \eta_S}\downarrow & & \downarrow{\scriptstyle \eta'_S} \\ UF(S) & \xrightarrow[UF(s)]{} & UF(S') \\ {\scriptstyle U(f)}\downarrow & & \downarrow{\scriptstyle U(f')} \\ U(M) & \xrightarrow[U(m)]{} & U(M') \end{array}$$

in $(\mathscr{S} \downarrow U)$.

To check that θ is a functor, note that applying θ to $(1_S, 1_M)$, the identity morphism on the object (S, f, M) of $(F \downarrow \mathscr{M})$, produces $(1_S, 1_M)$, the identity morphism on the object $\theta(S, f, M) = (S, U(f)\varepsilon_S, M)$ of $(\mathscr{S} \downarrow U)$, while verification that $\theta((s', m')(s, m)) = \theta(s', m')\theta(s, m)$ is easily seen by inspecting the following commutative diagram which comes directly from the definitions:

$$\begin{array}{ccccc} S & \xrightarrow{s} & S' & \xrightarrow{s'} & S'' \\ {\scriptstyle \varepsilon_S}\downarrow & & \downarrow{\scriptstyle \varepsilon_{S'}} & & \downarrow{\scriptstyle \varepsilon_{S''}} \\ UF(S) & \xrightarrow{UF(s)} & UF(S') & \xrightarrow{UF(s')} & UF(S'') \\ {\scriptstyle U(f)}\downarrow & & \downarrow{\scriptstyle U(f')} & & \downarrow{\scriptstyle U(f'')} \\ U(M) & \xrightarrow[U(m)]{} & U(M') & \xrightarrow[U(m')]{} & U(M'') \end{array}$$

Clearly θ commutes with the projections to $\mathscr{S} \times \mathscr{M}$.

Similarly a natural transformation $\varepsilon : 1_{\mathscr{S}} \longrightarrow UF$ gives rise to the functor $\xi : (\mathscr{S} \downarrow U) \longrightarrow (F \downarrow \mathscr{M})$ which takes the morphism

$$\begin{array}{ccc} S & \xrightarrow{s} & S' \\ {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle g'} \\ U(M) & \xrightarrow[U(m)]{} & U(M') \end{array}$$

in $(\mathscr{S} \downarrow U)$ to the morphism

$$\begin{array}{ccc} F(S) & \xrightarrow{F(s)} & F(S') \\ {\scriptstyle F(g)}\downarrow & & \downarrow{\scriptstyle F(g')} \\ FU(M) & \xrightarrow[FU(m)]{} & FU(M') \\ {\scriptstyle \varepsilon_M}\downarrow & & \downarrow{\scriptstyle \varepsilon_{M'}} \\ M & \xrightarrow[m]{} & M' \end{array}$$

in $(F \downarrow \mathscr{M})$. Checking that ξ is a functor is much the same as for θ, as is verifying that ξ commutes with the projections to $\mathscr{S} \times \mathscr{M}$.

To see that θ and ξ are inverse to one another, we first show that the equation $(\eta F)(F\varepsilon) = 1_F$ implies $\xi\theta$ is the identity on $(F \downarrow \mathscr{M})$, and then that the equation $(U\eta)(\varepsilon U) = 1_U$ implies $\theta\xi$ is the identity on $(\mathscr{S} \downarrow U)$.

Applying $\xi\theta$ to the morphism $(s, m) : (S, f, M) \longrightarrow (S', f', M')$ gives $(s, m) : (S, \xi\theta(f), M) \longrightarrow (S', \xi\theta(f'), M')$ where $\xi\theta(f) = \eta_M FU(f)F(\varepsilon_S)$ As η is a natural transformation $\eta_M FU(f)F(\varepsilon_S) = f\eta_{F(S)}F(\varepsilon_S)$ and that together with the hypothesis $(\eta F)(F\varepsilon) = 1_F$ gives $\xi\theta(f) = f$, which is the desired result.

The similar verification that the equation $(U\eta)(\varepsilon U) = 1_U$ implies $\theta\xi = 1_{(\mathscr{S}\downarrow U)}$ is left as the next exercise.

The proof of the converse is in steps much as above. Starting with a functor $\theta : (F \downarrow \mathscr{M}) \longrightarrow (\mathscr{S} \downarrow U)$ commuting with the projections to $\mathscr{S} \times \mathscr{M}$ consider for each object S of $\mathscr{S}$ the object $\theta(F(S), 1_{F(S)}, F(S))$ of $(\mathscr{S} \downarrow U)$. The hypothesis that θ commutes with the two projections to $\mathscr{S} \times \mathscr{M}$ guarantees this has the form $(S, \varepsilon_S, F(S))$, and this defines the morphism $\varepsilon_S : S \longrightarrow UF(S)$.

This actually defines a natural transformation $\varepsilon : 1_{\mathscr{S}} \longrightarrow UF$ with the naturality coming by noting that

$$\begin{array}{ccc} S & \xrightarrow{\varepsilon_S} & UF(S) \\ {\scriptstyle s}\downarrow & & \downarrow{\scriptstyle UF(s)} \\ S' & \xrightarrow[\varepsilon_{S'}]{} & UF(S') \end{array}$$

is commutative because it is the image under θ of the commutative square

$$\begin{CD} F(S) @>{1_{F(S)}}>> F(S) \\ @V{F(s)}VV @VV{F(s)}V \\ F(S') @>>{1_{F(S')}}> F(S') \end{CD}$$

Next we notice that θ takes the morphism $(1_S, f) : (S, 1_{F(S)}, F(S)) \longrightarrow (S, f, M)$ in $(F \downarrow \mathscr{M})$ to the morphism $(1_S, f) : (S, \varepsilon_S, F(S)) \longrightarrow (S, \theta(f), M)$ in $(\mathscr{S} \downarrow U)$, and from that we see that $\theta(f) = U(f)\varepsilon_S$.

Similarly starting with a functor $\xi : (U \downarrow \mathscr{M}) \longrightarrow (F \downarrow \mathscr{M}$ commuting with the projections to $\mathscr{S} \times \mathscr{M}$, define, for each object M of $\mathscr{M}$, the morphism $\eta_M : FU(M) \longrightarrow M$ by the formula $\xi(U(M), 1_{U(M)}, U(M)) = (U(M), \eta_M, M)$

Again this defines the natural transformation $\eta : FU \longrightarrow 1_{\mathscr{M}}$ with the square

$$\begin{CD} FU(M) @>{\eta_M}>> M \\ @V{FU(m)}VV @VV{m}V \\ FU(M') @>>{\eta_{M'}}> M' \end{CD}$$

being commutative because it is the image under ξ of the commutative square

$$\begin{CD} U(M) @>{1_{U(M)}}>> U(M) \\ @V{U(m)}VV @VV{U(m)}V \\ U(M') @>>{1_{U(M')}}> U(M') \end{CD}$$

And we see that ξ takes the morphism $(g, 1_M) : (S, g, M) \longrightarrow (U(M), 1_M, M)$ in $(\mathscr{S} \downarrow U)$ to the morphism $(g, 1_M) : (S, \xi(g), M) \longrightarrow (S, \eta_M, M)$ in $(F \downarrow \mathscr{M}$, and from that we see that $\xi(g) = \eta_M F(g)$.

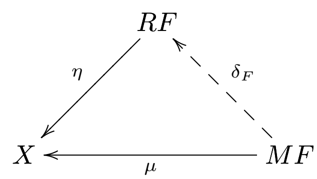
Now $\xi\theta(f) = \eta_M FU(f)F(\varepsilon_S) = f\eta_{F(S)}\varepsilon_{F(S)}$, so $\xi\theta = 1_{(F\downarrow\mathscr{M})}$ implies $(\eta F)(F\varepsilon) = 1_F$.

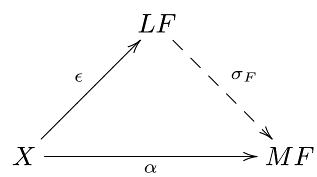
Equally $\theta\xi(g) = U(\eta_M)UF(g)\varepsilon_S = U(\eta_M)\varepsilon_{U(M)}g$, so $\theta\xi = 1_{(\mathscr{S}\downarrow)U}$ implies $(U\eta)(\varepsilon U) = 1_U$.

∎

Exercise V.5. Complete the proof of the above theorem by showing that $(U\eta)(\varepsilon U) = 1_U$ implies $\theta\xi = 1_{(\mathscr{S}\downarrow U)}$.

V.7 Kan Extensions

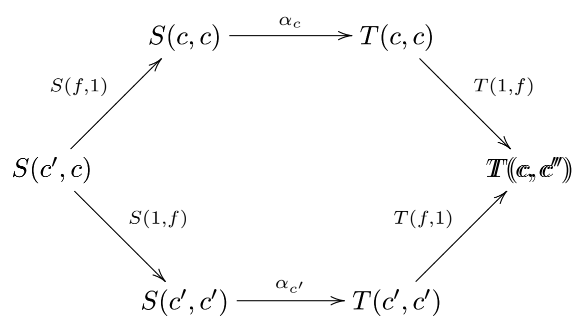




V.7.1 Ends and Coends

From wikipedia

a dinatural transformation α between two functors $S, T : \mathscr{C}^{\mathrm{op}} \times \mathscr{C} \longrightarrow \mathscr{X}$, written $\alpha : S \overset{..}{\longrightarrow} T$, is a function which to every object c of $\mathscr{C}$ associates an arrow $\alpha_c : S(c,c) \longrightarrow T(c,c)$ of $\mathscr{X}$ and satisfies the following coherence property: for every morphism $f : c \longrightarrow c'$ of $\mathscr{C}$ the diagram



commutes.

From wikipedia

an end of a functor $S : \mathscr{C}^{\mathrm{op}} \times \mathscr{C} \longrightarrow \mathscr{X}$ is a universal dinatural transformation from an object e of $\mathscr{X}$ to S.

More explicitly, this is a pair (e, ω), where e is an object of $\mathscr{X}$ $\omega : e \overset{..}{\longrightarrow} S$ is a dinatural transformation, such that for every dinatural transformation $\beta : x \overset{..}{\longrightarrow} S$ there exists a unique morphism $h : x \longrightarrow e$ of X with $\beta_a = \omega_a \circ h$ for every object a of $\mathscr{C}$.

By abuse of language the object e is often called the end of the functor S (forgetting ω) and is written $e = \int_c S(c,c)$ or just $\int_{\mathscr{C}} S$.

Coend

The definition of the coend of a functor $S : \mathscr{C}^{\mathrm{op}} \times \mathscr{C} \longrightarrow \mathscr{X}$ is the dual of the definition of an end. Thus, a coend of S consists of a pair (d, ζ), where

d is an object of $\mathscr{X}$ and $\zeta : S \overset{\cdot\cdot}{\longrightarrow} d$ is a dinatural transformation, such that for every dinatural transformation $\gamma : S \overset{\cdot\cdot}{\longrightarrow} x$ there exists a unique morphism $g : d \longrightarrow x$ of $\mathscr{X}$ with $\gamma_a = g \circ \zeta_a$ for every object a of $\mathscr{C}$.

The coend d of the functor S is written $d = \int^c S(c,c)$ or $\int^{\mathscr{C}} S$.

Chapter VI

More Mathematics in a Category

VI.1 Relations In Categories

When we discussed subobjects and quotient objects back in Section I.2.4 we saw that in **Set**, the category of sets, equivalent quotient objects corresponds exactly to the quotient sets with respect to an equivalence relation. We also noted we did not even have the notion of an equivalence relation on an object for general categories. As promised we will here give the needed definition and relate it to quotient objects.

In **Set** a relation on a set S is defined as a subset of $S \times S$. We could follow that definition and define a relation on an object C as a subobject of $C \times C$, but that restricts the definition to those categories where $C \times C$ exists. Instead we note that specifying a subobject R of $C \times C$, i.e., a monomorphism $d : R \rightarrowtail C \times C$, is equivalent to giving a pair of morphisms (d_0, d_1) from R to C where $fd_0 = fd_1$ and $gd_0 = gd_1$, implies $f = g$. That gives rise to the following definitions.

[For convenience we will write $(d_0, d_1) : R \longrightarrow C$ as shorthand for $d_0 : R \longrightarrow C$ and $d_1 : R \longrightarrow C$.]

Definition VI.1: A pair of morphisms $(d_0, d_1) : R \longrightarrow C$ is **jointly monic**,or is a **monic pair** provided $d_0 f = d_1 f$ and $d_0 g = d_1 g$, implies $f = g$

Definition VI.2: A **relation** on an object C is a pair of morphisms $(d_0, d_1) : R \longrightarrow C$ which are **jointly monic**.

Exercise VI.1. Verify that if the product $C \times C$ exists, then d_0, d_1 are jointly monic iff $\langle d_0, d_1 \rangle : R \longrightarrow C \times C$ is monic.

Recall that an equivalence relation on a set satisfies the three conditions of being *reflexive*, *symmetric*, and *transitive*. (See Mac Lane and Birkhoff [55, Sec. 1.11] for details.) So we want to translate these notions to the setting of category theory.

For sets, a relation, R is reflexive iff (c, c) is always in R. That's the same as saying that the diagonal Δ is a subset of R, and that in turn is the same as requiring the diagonal morphism to factor through $\langle d_0, d_1 \rangle$, i.e., there is a function $r : C \longrightarrow R$ with $\langle d_0, d_1 \rangle r = \Delta$ which is the same as saying that $d_0 r = 1_C$ and $d_1 r = 1_C$. And that gives us a definition of reflexive relation which is valid in any category.

Definition VI.3: A relation $d_0, d_1 : R \longrightarrow C$ is **reflexive** iff there is a morphism $r : C \longrightarrow R$ such that $d_0 d = 1_C = d_1 d$.

Note this says d_0 and d_1 have r as a common retract.

As often, the Hom functors gives us a way of reflecting our knowledge of **Set** to a general category. If $d_0, d_1 : R \longrightarrow C$ is a relation, then for any object X, the pair $\mathrm{Hom}(X, d_0), \mathrm{Hom}(X, d_1) : \mathrm{Hom}(X, R) \longrightarrow \mathrm{Hom}(X, C)$ is a relation on $\mathrm{Hom}(X, C)$ in **Set**. The next exercise shows that reflexive relations are preserved (d_0, d_1 a reflexive relation implies $\mathrm{Hom}(\bullet, d_0)$, $\mathrm{Hom}(\bullet, d_1)$ is a reflexive relation) and reflected ($\mathrm{Hom}(\bullet, d_0)$, $\mathrm{Hom}(\bullet, d_1)$ a reflexive relation implies d_0, d_1 a reflexive relation.)

Exercise VI.2. Suppose that in $\mathscr{C}$ the pair d_0, d_1 (both morphisms from R to C) is a relation on C. Show that it is a reflexive relation iff for every object X, the pair $\mathrm{Hom}(X, d_0)$ and $\mathrm{Hom}(X, d_1)$ is a reflexive relation on $Hom(X, C)$ in **Set**.

Similarly, for sets a symmetric relation is one where $(c, c') \in R \iff (c', c) \in R$. That's the same as saying that the pair (d_1, d_0) defines the same relation as the pair (d_0, d_1) which in turn is the same as saying there is an isomorphism $\tau : R \longrightarrow R$ with $d_1 = d_0 \tau$, which gives us the general definition we want.

Definition VI.4: A relation $d_0, d_1 : R \longrightarrow C$ is **symmetric** iff there is an isomorphism $\tau : R \longrightarrow R$ with $d_1 = d_0 \tau$.

Exercise VI.3. Suppose that in $\mathscr{C}$ the pair $(d_0, d_1) : R \longrightarrow C$ is a relation on C. Show that it is a symmetric relation iff for every object X, the pair $(\mathrm{Hom}(X, d_0), \mathrm{Hom}(X, d_1))$ is a symmetric relation on $Hom(X, C)$ in **Set**.

Definition VI.5: A relation $d_0, d_1 : R \longrightarrow C$ is **transitive** iff

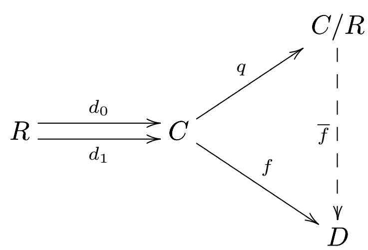
Exercise VI.4. Suppose that in $\mathscr{C}$ the pair $(d_0, d_1) : R \longrightarrow C$ is a relation on C. Show that it is a transitive relation iff for every object X, the pair $(\mathrm{Hom}(X, d_0), \mathrm{Hom}(X, d_1))$ is a transitive relation on $Hom(X, C)$ in **Set**.

Definition VI.6: An **equivalence relation** on an object C is a relation on C that is reflexive, symmetric, and transitive.

Of course we are usually interested in equivalence relations because they allow us to define quotient objects, so our next definition makes that connection.

Definition VI.7: If $(d_0, d_1) : R \longrightarrow C$ is an equivalence relation on C, then a **quotient** of C by R is an object C/R and a morphism $q : C \longrightarrow C/R$ satisfying the following Universal Mapping Property:

1. $qd_0 = qd_1$, and
2. if f is any morphism with $fd_0 = fd_1$, then f factors uniquely through q, i.e., there is a unique morphism $\overline{f}$ with $\overline{f}q = f$.



The morphism q is called "the" *quotient morphism* and the object C/R is called "the" *quotient object* of C modulo R. This common terminology is misleading in that the equivalence relation is actually the pair of morphisms (d_0, d_1), not their common domain R, while as with all universal mapping properties, the object C/R and the morphism q are only determined up to a unique isomorphism.

Of course we already have a definition of a quotient object, and in the category of sets we know quotient objects (i.e., equivalence classes of epimorphisms, see definition I.25) correspond to quotients modulo an equivalence relation. So the next step is to see the extent that extends to general categories. The first stage is observing that the quotient morphism q is always an epimorphism.

Treated in Popescu and Popescu [63, pp. 54–56].

Definition VI.8: A **regular monomorphism** is an equalizer of a pair.

Definition VI.9: A **extremal monomorphism** is a monomorphism that cannot be factored by an epimorphism that is not an isomorphism.

Chapter VII

Algebraic Categories

Rhetorical algebra, as instructions written entirely with words and numbers for solving various concrete arithmetic problems, appeared around four millenia ago in Babylonia and Egypt.

How many apples are needed if four persons of six receive one-third, one-eighth, one-fourth, and one-fifth, respectively, of the total number, while the fifth receives ten apples, and one apple remains left for the sixth person?

(Eves, Great Moments in Mathematics (Before 1650), 1980, pp. 127-128.)

Algebra slowly changed to the geometric algebra of the classical Greek mathematicians and the Vedic Indian mathematicians, and still later evolved into syncopated algebra using abbreviations and some symbols. Finally in the 16th century symbolic algebra emerged with the use of symbols for addition, subtraction, multiplication, exponents, coefficients, unknowns, radicals, etc. By this time the primary preoccupation of algebra had become understanding roots of polynomial equations, and this led to the acceptance of negative and then complex numbers. After the invention of analytic geometry by Descartes and Pascal, it also led increasingly to the intertwinning of algebra and geometry.

During the 17th and 18th centuries the Fundamental Theorem of Algebra (FTA) slowly emerged along with an understanding of the complex numbers. This culminated with the proof of the FTA at the end of the 18th century. By this time abstract algebra was emerging as attention shifted from solving polynomial equations to studying the number systems and other mathematical structures that arose while understanding the theory of equations. Euler, Gauss, Cauchy, Abel and others all used groups in their work throughout the 18th and 19th centuries but without the name. Early in the 19th century Galois explicitly recognized groups in his invention of Galois theory, and they have been an important unifying concept every since. In addition Hamilton (with the invention of Quaternions), Grassmann (in his study of vectors and Grassmann algebras), Cayley (studying the Octonions) and Boole (with his algebraic treatment of basic logic) developed the immediate precursors of abstract algebra as we know it today.

Although the axiomatic approach to mathematics was one of the important contributions of classical Greek mathematics, its applicability to algebra did not come to the fore until the latter part of the 19th century. At the turn of the century the impact of non-Euclidean geometry, the antimonies of axiomatic set theory, and the general recognition of the general value of the axiomatic approach to algebra lead to the recognition of the subject called Universal Algebra. Indeed the first book on this subject was published by A. N. Whitehead [70] in 1898.

Whitehead clearly recognized and stated the basic goal of a comparative study of various systems of finitary operations on a set, with the intention of understanding the common properties of such varied algebraic structures as monoids, groups, rings, Boolean algebra and lattices. When Whitehead published his book, the knowledge of these structures was insufficiently developed to provide the experience needed to develop a general theory. It was not until the 1930s that Birkhoff began publishing results in the area that is the subject of this chapter.

The first section here develops the basics of the theory in much the fashion expounded by Birkhoff [5], while later sections discuss the approach through category theory that was developed by Lawvere [45].

VII.1 Universal Algebra

As a first step into universal algebra consider a fixed set Ω (the set of operator symbols) and a function arity : $\Omega \longrightarrow \mathbb{N}$ giving the arity of the operator. An Ω-*algebra* A is a set $|A|$ (the *carrier* of A) and for each $\omega \in \Omega$ of arity $n = \text{arity}(\omega)$ an operator $o_\omega : |A|^n \longrightarrow |A|$

For general information on universal algebra consult the books of that title by P. M. Cohn [13] and G. Grätzer [27].

VII.2 Algebraic Theories

The original source for this material is Lawvere's thesis [45] which is available in the Theory and Application of Categories Reprint Series.

See the article "Algebraic Categories" by Pedicchio and Rovatti [61].

VII.3 Internal Categories

One of the most fascinating aspects of the theory of categories is its reflective nature. One aspect of that is that it is quite natural to define categories and functors within a category, at least one that has finite limits. Throughout this section we will take $\mathscr{S}$ to be a Cartesian category.

Definition VII.1: An **internal category** in $\mathscr{S}$ consists of

Definition VII.2:

Definition VII.3:

Chapter VIII

Cartesian Closed Categories

Suppose that $\mathscr{V}$ is a Cartesian category, i.e., one with all finite products. Then for each object C of $\mathscr{V}$ there is the functor $\bullet \times C : \mathscr{V} \longrightarrow \mathscr{V}$. When this functor has a right adjoint it is written as $\bullet^C$ A category where all such functors have right adjoints is called a **Cartesian closed category**.

VIII.1 Partial Equivalence Relations and Modest Sets

Definition VIII.1: A **partial equivalence relation** $\approx$ on a set A is a binary relation on A that is symmetric ($a \approx a' \Rightarrow a' \approx a$) and transitive ($a \approx a'$ and $a' \approx a'' \Rightarrow a \approx a''$).

Note that the difference between a partial equivalence relation and an equivalence relation (see definition A.24) is that a partial equivalence relation need not be reflexive, i.e., $a \approx a$ is not guaranteed. So every equivalence relation is a partial equivalence relation but the converse is not true with an example being the relation $\{(0, 0)\}$ on $\{0, 1\}$. In other words we have $0 \approx 0$ and nothing more. Clearly this is symmetric and transitive, but it is not reflexive as $1 \approx 1$ is false.

The phrase "partial equivalence relation" will be used sufficiently often that we will frequently use the abbreviation PER.

The reason for the word "partial" is because all PERs are similar to the example. When $\approx$ is a PER on the set A, consider the subset $D = \{a \in A : a \approx a\}$. Notice that if a and a' are elements of A and $a \approx a'$, then by symmetry $a' \approx a$ and by transitivity both $a \approx a$ and $a' \approx a'$ so both a and a' are in D. That means $\approx$ can be considered as a relation on D, and there it is an equivalence relation. D is called the *domain of definition* of the PER $\approx$ and we may write $D(\approx)$ to emphasize the relation, particularly when considering more than one PER.

The converse is equally well true — if D is a subset of A and $\approx$ is an *equivalence relation* on D, the $\approx$ may equally well be considered as a relation on A and it is a partial equivalence relation on A. So an alternative description of a PER on A is an equivalence relation on some subset of A.

Equivalence relations on sets are closely related to quotient sets (see the discussions in sections I.26 and A.27), so in a certain sense a PER on A is much the same thing as a quotient of a subset of A, a perspective we will explore further.

For $\approx_A$ a PER on A and $\approx_B$ is a PER on B, a function $f : A \longrightarrow B$ is *compatible* when $a \approx_A a' \Rightarrow f(a) \approx_B f(a')$. This tells us that

$$f(D(\approx_A)) \subseteq D(\approx_B)$$

and so we can consider

$$f|_{D(\approx_A)} : D(\approx_A) \longrightarrow D(\approx_B)$$

Going one step further we also have the quotient maps $q_A : D(\approx_A) \longrightarrow D(\approx_A)/\approx_A$ and $q_B : D(\approx_B) \longrightarrow D(\approx_B)/\approx_B$) (see page 196) and they all fit into the commutative square:

$$\begin{array}{ccc} D(\approx_A) & \xrightarrow{\quad f \quad} & D(\approx_B) \\ {\scriptstyle q_A}\downarrow & & \downarrow{\scriptstyle q_B} \\ D(\approx_A)/\approx_A & \underset{\overline{f}}{\dashrightarrow} & D(\approx_B)/\approx_B \end{array}$$

where $\overline{f}$ is the unique function with $q_B f = \overline{f} q_A$ (see exercise I.32.)

Writing $f : (A, \approx_A) \longrightarrow (B, \approx_B)$ for a compatible function from A to B, we can summarize this by saying that every compatible function f induces a function $\overline{f} : D(\approx_A)/\approx_A \longrightarrow D(\approx_B)/\approx_B$.

As an intermediate step to defining the **PER** category we define the category of compatible functions to have as objects pairs $(A, \approx_A)$ with A a set and $\approx_A$ a PER on A, and as morphisms the compatible functions between such pairs. The identity morphisms are just the identity functions, and composition is just composition of functions which works because the composition of two compatible functions is again a compatible function.

The penultimate step here is to observe that we can define a congruence on the category of compatible functions (recall definition II.2) by saying that two compatible functions f and f' from $(A, \approx_A)$ to $(B, \approx_B)$ are congruent ($f \sim f'$) when for all $a \in A$

$$a \approx_A a \Rightarrow f(a) \approx_B f'(a)$$

This is an equivalence relation on compatible functions from $(A, \approx_A)$ to $(B, \approx_B)$ exactly because these are compatible functions and $\approx_B$ is a PER. The other part of the congruence condition is:

$$f \sim g \wedge h \sim k \wedge hf \text{ defined } \Rightarrow hf \sim kg.$$

Translated to this situation we must show

$$(\forall a \in A,\ a \approx_A a \Rightarrow f(a) \approx_B g(a) \wedge \forall b \in B,\ b \approx_B b \Rightarrow h(b) \approx_C k(b)) \\ \Rightarrow \forall a \in A,\ a \approx_A a \Rightarrow hf(a) \approx_C kg(a)$$

which comes from $hf(a) \approx_C hg(a)$ (because h is compatible) and $h(g(a)) \approx_C k(g(a)$ (which follows from the second assumption because $g(a) \approx_B g(a)$.)

The point of this equivalence relation is exactly that two compatible morphisms f and f' are equivalent iff they induce the same function from $D(\approx_A)/\approx_A$ to $D(\approx_B)/\approx_B$.

With all of this out of the way we are ready to define the **PER** category.

Definition VIII.2: The category **PER** is the quotient category (see definition II.3) of the category of compatible functions by the congruence described above. So **PER** has as objects pairs $(A, \approx_A)$ with A a set and $\approx_A$ a PER on A, while a morphism $f : (A, \approx_A) \longrightarrow (B, \approx_B)$ is an equivalence class of compatible functions from $(A, \approx_A)$ to $(B, \approx_B)$.

As for all quotient categories, the identity morphisms are the equivalence classes of identity functions and composition is defined by the equivalence class of the composition of representative functions. As common in such situations we will often abuse the notation and deliberately confuse compatible functions with the equivalence class they represent.

The category **PER** was introduced as the inspiration and example of a whole family of others. For $\mathscr{C}$ any Cartesian closed category we are going to introduce the category $\mathbf{PER}(\mathscr{C})$ which is based on following the construction of **PER** except that the objects will be PERs on the points of an object in $\mathscr{C}$. We start with the following definition.

Definition VIII.3: When $\mathscr{C}$ is a CCC and A is an object of $\mathscr{C}$, a **partial equivalence relation** (PER) on A is a PER on the set of points of A, i.e., the set $\mathscr{C}(1, A)$.

If $\approx_A$ is a PER on A (i.e., on $\mathscr{C}(1, A)$) and $\approx_B$ is a PER on B, then a morphism $f : A \longrightarrow B$ in $\mathscr{C}$ is *compatible* when the induced function $f_* : \mathscr{C}(1, A) \longrightarrow \mathscr{C}(1, B)$ is a compatible function. In full this means that $\forall a, a' \in \mathscr{C}(1, A), a \approx_A a' \Rightarrow fa \approx_B fa'$.

As above we define the (anonymous) category of compatible morphisms on $\mathscr{C}$ to have as objects pairs $(A, \approx_A)$ with A an object of $\mathscr{C}$ and $\approx_A$ a PER on A, and as morphisms the compatible morphisms between such pairs. The identity morphisms are just the identity functions, and composition is just composition of morphisms which works because the composition of two compatible morphisms is again a compatible morphism.

Of course the next step is to define a congruence on the category of compatible morphisms on $\mathscr{C}$ by saying that two compatible morphisms f and f'

from $(A, \approx_A)$ to $(B, \approx_B)$ are congruent $(f \sim f')$ when for all $a \in \mathscr{C}(1, A)$

$$a \approx_A a \Rightarrow fa \approx_B f'a$$

This is an equivalence relation on compatible morphisms from $(A, \approx_A)$ to $(B, \approx_B)$ exactly because these are compatible morphisms and $\approx_B$ is a PER. The other part of the congruence condition is:

$$f \sim g \wedge h \sim k \wedge hf \text{ defined } \Rightarrow hf \sim kg.$$

Translated to this situation we must show

$$(\forall a \in \mathscr{C}(1, A),\ a \approx_A a \Rightarrow fa \approx_B ga \wedge \forall b \in \mathscr{C}(1, B),\ b \approx_B b \Rightarrow hb \approx_C kb) \\ \Rightarrow \forall a \in \mathscr{C}(1, A),\ a \approx_A a \Rightarrow hfa \approx_C kga$$

which comes from $hfa \approx_C hga$ (because h is compatible) and $hga \approx_C kga$ (which follows from the second assumption because $ga \approx_B ga$.)

And now we are ready to define the **PER** category.

Definition VIII.4: The category $\mathbf{PER}(\mathscr{C})$ is the quotient category of the category of compatible functions by the congruence described above. So $\mathbf{PER}(\mathscr{C})$ has as objects pairs $(A, \approx_A)$ with A an object of $\mathscr{C}$ and $\approx_A$ a PER on the set $\mathscr{C}(1, A)$, while a morphism $f : (A, \approx_A) \longrightarrow (B, \approx_B)$ is an equivalence class of compatible morphisms from $(A, \approx_A)$ to $(B, \approx_B)$, this last meaning that the induced function

$$f_* : (\mathscr{C}(1, A), \approx_{\mathscr{C}} (1, A)) \longrightarrow (\mathscr{C}(1, B), \approx_{\mathscr{C}} (1, B))$$

is a compatible function.

Just as with the original **PER**, the identity morphisms are the equivalence classes of identity functions and composition is defined by the equivalence class of the composition of representative functions. As common in such situations we will often abuse the notation and deliberately confuse compatible functions with the equivalence class they represent.

Now we have both **PER** and $\mathbf{PER}(\mathbf{Set})$, but these are really the same thing as the next exercise will have you verify.

Exercise VIII.1. Starting with the observation that the natural transformation

$$\epsilon : \mathbf{Set}(1, \bullet) \longrightarrow 1_{\mathbf{Set}}(\bullet)$$

given by $\epsilon(x) = x(0)$ is a natural equivalence, unwind all the details to verify that **PER** and $\mathbf{PER}(\mathbf{Set})$ are isomorphic categories.

Definition VIII.5: A **modest set** is a triple (I, A, e) where I and A are sets and $e : I \longrightarrow \mathcal{P}(A)$ defines an indexed family of subsets of A subject to:

1. For all i in I, $e(i) \neq \emptyset$;
2. If i and j are distinct elements of I, then $e(i) \cap e(j) = \emptyset$.

When (I, A, e) is a modest set, there is the subset $\bigcup_{i \in I} e(i)$ of A and the partition $\{e(i) : i \in I\}$ of this subset. So a modest set is equally well defined as an indexed partition of some subset of X. Now partition and equivalence relations are intimately intertwined (see Section A.7), so specifying a partition of a subset of A is equivalent to specifying an equivalence relation on the subset and so is equivalent to specifying a PER on A. The reason for the apparently more complicated definition is two-fold. First it provides flexibility when constructing new modest sets, and second it is a form that will conveniently generalize to other categories.

Definition VIII.6: A **morphism of modest sets**

$$f = (f_i, f_e) : (I, A, e_A) \longrightarrow (J, B, e_B)$$

is a pair of functions $f_i : I \longrightarrow J$ and $f_e : A \longrightarrow B$ with

$$\forall i \in I, a \in A,\ a \in e_A(i) \Rightarrow f_e(a) \in e_B(f_i(i))$$

Chapter IX

Topoi

Definition IX.1: If $\mathscr{C}$ is a category with a terminal object 1, the **subobject classifier** in $\mathscr{C}$ is a morphism $t : 1 \longrightarrow \Omega$ with the following universal mapping property: for each monomorphism $f : A \longrightarrow B$ in $\mathscr{C}$ there is a unique morphism $\chi_f : A \longrightarrow \Omega$ making

$$\begin{array}{ccc} A & \stackrel{f}{\rightarrowtail} & B \\ {\scriptstyle !}\downarrow & & \downarrow{\scriptstyle \chi_f} \\ 1 & \xrightarrow[t]{} & \Omega \end{array}$$

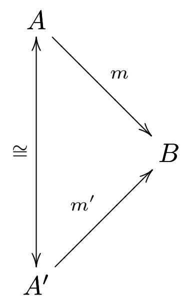
a pullback square.

The monomorphism χ_f is called the *characteristic morphism* of the monomorphism f (which is here considered as a subobject of B.) This is in analogy to the *characteristic function* $\chi_A : B \longrightarrow 2$ of a subset A of a set B where χ_A is the constant function with value 1.

Exercise IX.1. As in Appendix A (Set Theory) we write $1 = \{0\}$ and $2 = \{0, 1\}$. As noted previously 1 is a terminal object in **Set**. Defining $t : 1 \longrightarrow 2$ by $t(0) = 1$, show that t is a subobject classifier in **Set** by verifying that if $A \subseteq B$ and $i : A \longrightarrow B$ is the inclusion function, then the characteristic function χ_A is the characteristic morphism for i.

Recall from definition I.24 that two subobjects $m : A \rightarrowtail B$ and $m' : A' \rightarrowtail B$ of B are equivalent when there is an isomorphism between A and A'

with



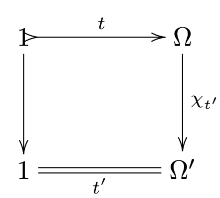
commuting.

This is captured by subobject classifiers as noted in the next exercise.

Exercise IX.2. If the category $\mathscr{C}$ has a subobject classifier $t : 1 \longrightarrow \Omega$, then $m : A \rightarrowtail B$ and $m' : A' \rightarrowtail B$ are equivalent subobjects of the object B in $\mathscr{C}$ iff $\chi_m = \chi_{m'}$.

As always with universal mapping properties if a subobject classifier exists it is unique up to a unique isomorphism, and this is the result of the next exercise.

Exercise IX.3. Verify that if $t : 1 \longrightarrow \Omega$ and $t' : 1 to \Omega'$ are both subobject classifiers in $\mathscr{C}$, then there is a unique isomorphism $\chi_{t'} : \Omega' \longrightarrow \Omega$ so that



commutes.

Chapter X

The Category of Sets Reconsidered

This is largely a retelling of the story in *Sets for Mathematics* by Lawvere and Rosebrugh [46].

Definition X.1: 1 is a *separator* in $\mathscr{C}$ if $f, g : C \longrightarrow D$, then f and g equal on elements of C implies $f = g$.

In the following definition we consider only categories with a terminal object.

Definition X.2: A morphism $f : C \longrightarrow D$ is a *surjection* when for each element $d : \mathbf{1} \longrightarrow D$ of D there is at least one element $c : \mathbf{1} \longrightarrow C$ with $d = fc$.

The category of sets **Set** has the following properties:

1. **Set** has a final object, **1**.
2. **Set** has an initial object, **0**.
3. **1** separates morphisms in **Set**.
4. **0** is *not* isomorphic to **1**.
5. **Set** has sums.
6. **Set** has products.
7. **Set** has finite (inverse) limits.
8. **Set** has finite colimits.
9. [Exponentiation] For all objects X and Y in **Set** there is a mapping object Y^X.

10. There is a truth value object $t : 1 \longrightarrow \Omega$ in **Set**, i.e., there a natural correspondence between parts of an object X and morphisms $X \longrightarrow \Omega$.

11. [**Set** is Boolean] $\langle t, f\rangle : 1 \longrightarrow \Omega$ is an isomorphism.

12. [Axiom of Choice] Every surjection in **Set** is a retract (and so an epimorphism.)

13. [Axiom of Choice] Every epimorphism is a retract (splits).

Chapter XI

Monoidal Categories

References: Kelly [37],

Chapter XII

Enriched Category Theory

See Kelly [37] and Dubuc [16].

Chapter XIII

Additive and Abelian Categories

Chapter XIV

Homological Algebra

References: Cartan and Eilenberg [11] ; Weibel [65]; Kashiwara and Schapira [35]
.

History in Weibel article [66].

XIV.1 Introduction

XIV.2 Additive Categories

XIV.3 Abelian Categories

First appearance in Mac Lane [49]. Connection with homological algebra as Buchsbaum's exact categories which were describe in the appendix to Cartan and Eilenberg Homological Algebra [11]. Definitive appearance in Grothendieck's "Sur Quelques Points d'Algbre Homologique" [28].

XIV.4 Ext and Tor

XIV.5 Category of Complexes

XIV.6 Triangulated Categories

XIV.7 Derived Categories

XIV.8 Derived Functors

Chapter XV

2-Categories

Chapter XVI

Fibered Categories

Appendices

Appendix A

Set Theory

The category of sets is used as the basic example of a category throughout theses notes. Understanding that category requires knowledge of the rudiments of set theory. For convenient reference the necessary information is outlined here.

This appendix is not intended to be an introduction to set theory, so although all the concepts that are discussed will be formally defined, they will on occasion be discussed informally before the actual definition is given.

Because the connection between category theory and the theory of sets occasionally skirts the edge of the foundation of set theory, the treatment is an informal axiomatic treatment which was inspired by Halmos' *Naive Set Theory*, [29]. though for reasons explained below the actual theory is different.

When set theory was first emerging as a part of mathematics there was the belief that for any boolean predicate $P(x)$ (i.e., where for every x the predicate $P(x)$ is either true or false) there is the set $\{x : P(x)\}$ of all elements x for which the predicate was true. That idea was demolished by Bertrand Russell when he propounded his famous paradox: which expressed If $S = \{x : x \text{ is a set and } x \notin x\}$, then $S \notin S \Rightarrow S \in S$, while $S \in S \Rightarrow S \notin S$.

Around 1900 a number of such paradoxes were found, all in some fashion involving very "big" sets. Beginning with Zermelo in 1908 and continuing down to the present day, a number of inequivalent approaches to set theory have been advanced which avoid these problems and are strong enough to serve as a foundation for almost all of mathematics.

The best known of these is Zermelo-Fraenkel set theory which is presented very attractively in Halmos' *Naive Set Theory* [29] . Bourbaki's famous series of books published under the general title *Éléments de Mathématique* developed a large part of modern mathematics based on Zermelo-Fraenkel set theory which is exposed in great detail in Book I – *Theory of Sets* [9] . This book also contains a useful history.

Zermelo-Fraenkel set theory is commonly abbreviated ZFC (actually standing for Zermelo-Fraenkel set theory with the Axiom of Choice.) That notation will occasionally occur in these notes.

In Zermelo-Fraenkel set theory the problem of "big" sets is avoided by changing the specification of sets to allow only $\{x : P(x) \wedge x \in S\}$ with S a pre-existing set and providing other "safe" constructions of sets with the result that "big" sets are simply excluded from discussion. (For clarification, see Halmos [29], particularly p. 11.) Unfortunately for the purpose of these notes that means that most categories of interest cannot be considered within the realm of set theory. Other approaches allow these "big" sets (formally called *classes*). The version outlined here is due to A. P. Morse and John Kelley . It is sketched in the Appendix to Kelley's *General Topology* [36] and presented in detail in Morse [59]. While an introductory book length treatment is in Monk's *Introduction to Set Theory* [58] . The organization of this appendix is inspired by Halmos' book, while the theory presented is largely an amalgam of the presentation by Monk and Kelley. Details, i.e., proofs, are almost completely omitted.

Necessarily a key part of axiomatic set theory is a detailed specification of the formal language and logic used in the theory, but we will leave out those details. A discussion of what is needed is in the Introduction and Appendix of Monk's book.

A.1 Extension Axiom

Intuitively sets include a pack of wolves, a bunch of grapes or a flock of pigeons, but in developing mathematics it is a wondrous fact that sets of sets are all that is needed. So in set theory (as developed here) we have just two primitive notions – **class** and **membership**. (Sets will be *defined* to be certain classes.) Membership is a relationship between some, definitely not all, classes. The membership relation is written as $A \in B$ when the class A is a **member** of the class B. Other phrases used include "A is an **element** of B" and "A belongs to B".

The negation, i.e., A is *not* an element of B, is written as $A \notin B$.

Synonyms for class include *collection*, *family* and *aggregate*, but not *set* which has the following special meaning.

Definition A.1: The class A is a **set** iff there is a class B with $A \in B$.

Definition A.2: A **proper class** is a class which is not a set, i.e., it does not belong to any other class.

Lower case letters will usually be used to indicate sets, so the trivial theorem:

$$\forall a\ \exists B,\ a \in B$$

says that for every *set* a there exists a class B with a an element of B.

The theory of sets begins with the simple notion of sets as collections of elements with two sets being equal exactly when they have the same elements. This is the first of our axioms, but for classes rather than sets.

Axiom of Extension: Two classes are equal iff they have the same elements. Symbolically this is:

$$\forall A\, \forall B\, [\forall C\, (C \in A \iff C \in B) \iff A = B]$$

A.2 Axiom of Specification

At the moment this theory may be vacuous as the axiom of extension doesn't guarantee that there are any classes at all! While the simplistic use of predicates to specify arbitrary sets leads to paradox, we can safely use them to specify classes as in the following axiom schema.

Axiom of Specification: For every set theoretic boolean predicate $\varphi(x)$ (i.e., $\varphi(x)$ is either true or false for every *set* x) in which A is not mentioned, the following is an axiom:
There exists a class A so that for every x, $x \in A \iff x$ is a set and $\varphi(x)$ or, symbolically,

$$\exists A\, \forall X,\ X \in A \iff X \text{ is a set } \wedge \varphi(X)$$

This is an *axiom schema* as there is an additional axiom for every predicate $\varphi(x)$ and every symbol A. Here we use capital letters to refer to classes, but there is no limit on the actual symbols that can be used as long as they are all understood to refer only to classes.

Using the axiom of extension it is easy to see that the class A where

$$\forall X,\ X \in A \iff X \text{ is a set and } \varphi(X)$$

is unique; it will be written as $\{x : \varphi(x)\}$

A lack remains – what is a "set theoretic boolean predicate in which A is not mentioned"? We will continue without giving a precise formulation, but more detail and precision is required and for that we will refer to Monk [58, page 15] for a informal discussion of the appropriate language.

Here are some interesting examples illustrating some simple choices for $\varphi(x)$:

$$R = \{x : x \notin x\}, \quad V = \{x : x = x\}, \quad \emptyset = \{x : x \neq x\}$$

The first example is the starting point for Russell's Paradox (see p. 179), but now if and we ask if $R \in R$ then we conclude that $R \notin R$. While from $R \notin R$ we simply conclude that R is a proper class and there is no contradiction.

The second example is simply the assertion that there is the *class* V of *all* sets. (We will consistently use V for the universe of all sets in honor of Godël who used it in his ground breaking work *The Consistency of The Axiom of Choice and of The Continuum Hypothesis With The Axioms of Set Theory.* In the next section we will see that V is a proper class, not a set.

The third example has no elements and is called the **empty class** and is denoted by $\emptyset$. Although it is part of the axiom of infinity (see p. 200), it is convenient to postulate at this point that this is a set.

Axiom of Existence: The empty class is a set. Symbolically this is

$$\exists C(\emptyset \in C)$$

The empty set is a subset of every set, but at this stage $\emptyset$ is the only set guaranteed to exist. Curiously, and usefully, for the purposes of mathematics it seems to be enough to have just sets that can be built from the empty set, but we do need axioms that guarantee the existence of more than just the empty set. The axiom in next section is a start.

A.2.1 Boolean Algebra of Classes

The first two axioms are enough to prove some of the basic algebraic manipulation of classes.

Definition A.3: $A \cup B = \{x : x \in A \vee x \in B\}$
$A \cup B$ is called the **union** of A and B.

Definition A.4: $A \cap B = \{x : x \in A \wedge x \in B\}$
$A \cap B$ is called the **intersection** of A and B.

Notice that even if A and B are sets, the two axioms do not guarantee that either $A \cup B$ or $A \cap B$ is a set.

Here are some basic facts about the union of classes which are easily proved:

- $A \cup \emptyset = A$
- $A \cup V = V$, with V the class of all sets
- $A \cup B = B \cup A$
- $A \cup (B \cup C) = (A \cup B) \cup C$
- $A \cup A = A$

And here are corresponding basic facts about the intersection of classes:

- $A \cap \emptyset = \emptyset$
- $A \cap V = A$, with V the class of all sets
- $A \cap B = B \cap A$
- $A \cap (B \cap C) = (A \cap B) \cap C$
- $A \cap A = A$

As algebraic operations we see that $\cup$ and $\cap$ are commutative, associative and have identities. There are distributive laws as well:

- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Besides union and intersection, there is a third basic operation, complement or difference.

Definition A.5: $A' = \{x : x \notin A\}$

A' is called the complement of A. The notation $\sim A$ is also often used for the complement.

Definition A.6:

$$A - B = \{x \in A : x \notin B\}$$

$A - B$ is called the **difference** between A and B or the **complement of** B **in** A. Writings that use $\sim A$ for the complement of A usually write $A \sim B$ for the complement of B in A.

The connection between complement and difference is direct: $A' = V - A$ and $A - B = A \cap B'$

The basic facts about complements include the following easy results:

- $(A')' = A$
- $\emptyset' = V$; $V' = \emptyset$
- $V' = \emptyset$; $\emptyset' = V$
- $A \cup A' = V$; $A \cap A' = \emptyset$
- De Morgan laws: $(A \cup B)' = A' \cap B'$; $(A \cap B)' = A' \cup B'$
- $A - B = A \cap B'$
- $A - (A - B) = A \cap B$
- $A \cap (B - C) = (A \cap B) - (A \cap C)$

An important thrust in the development of set theory was to show that it was adequate for the development of *all* of mathematics. In axiomatic set theory such as outlined here a key point is that there is *nothing* in the theory other than sets (well, here, nothing but classes.) So in particular every set is itself a collection of sets, i.e., a set whose elements are sets. But most of the time in other parts of mathematics we write of sets whose elements are numbers, or points in a space, or other irreducible items. The fact that these can themselves be constructed as sets of some sort is irrelevant to the mathematics being developed. As a result it is common to specifically single out sets of sets (more commonly called collections of sets,) and to connect with this convention we will follow suit. In particular we will often use special notation such as $\mathcal{C}$ when we want to note that we are paying special attention to the fact that the elements of the class $\mathcal{C}$ are themselves sets. Two very common example are unions and intersections of collections of sets.

Definition A.7: $\bigcup \mathcal{C} = \{x : \exists c(x \in c \wedge c \in \mathcal{C}\}$
$\bigcup \mathcal{C}$ is the **union** of the collection of sets $\mathcal{C}$.

The notation $\bigcup \mathcal{C}$ is actually less common than more convoluted notation such as

$$\bigcup \{x : x \in \mathcal{C}\} \quad \text{or} \quad \bigcup_{x \in \mathcal{C}} x. \tag{A.1}$$

Definition A.8: $\bigcap \mathcal{C} = \{x : \forall c,\ c \in \mathcal{C} \Rightarrow x \in c\}$
$\bigcap \mathcal{C}$ is the **intersection** of the members of $\mathcal{C}$.

As with the union of collections, the notation $\bigcap \mathcal{C}$ is actually less common than more convoluted notation such as

$$\bigcap \{x : x \in \mathcal{C}\} \quad \text{or} \quad \bigcap_{x \in \mathcal{C}} x.$$

Note that the elements of $\bigcap \mathcal{C}$ belong to the members of $\mathcal{C}$ and need not (but may) belong to $\mathcal{C}$ itself. Similarly $\bigcup \mathcal{C}$ is the union of the members of $\mathcal{C}$. In particular x belongs to $\bigcap \mathcal{C}$ iff x belongs to every member of $\mathcal{C}$, while x belongs to $\bigcup \mathcal{C}$ iff x belongs to at least one member of $\mathcal{C}$.

Two extreme cases are $\bigcup \emptyset$ which is $\emptyset$, and $\bigcap \emptyset$ which is the universe V of all sets.

Besides equality of classes, there is also the fundamental notion of one class being a subclass of another.

Definition A.9: $A \subseteq B$ iff $x \in A \Rightarrow x \in B$
$A \subseteq B$ is read as A is a **subclass** of B, or A *is contained in*, or B **includes** A.

In line with the notation in this definition, $A \subset B$ means A is a subclass of B but A is not equal to B. Be warned that in much of the set theory literature $\subset$ is used where we use $\subseteq$.

Basic facts about inclusion of classes include:

- $\emptyset \subseteq A$ and $A \subseteq V$.
- $A = B$ iff $A \subseteq B$ and $B \subseteq A$.
- If $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.
- $A \subseteq B$ iff $A \cup B = B$.
- $A \subseteq B$ iff $A \cap B = A$.
- If $A \subseteq B$, then $\bigcup A \subseteq \bigcup B$ and $\bigcap B \subseteq \bigcap A$.
- If $a \in A$, then $a \subseteq \bigcup A$ and $\bigcap A \subseteq a$.

- $A \subseteq B$ iff $B' \subseteq A'$.
- $A \subseteq B$ iff $A - B = \emptyset$
- $A \cap B \subseteq (A \cap C) \cup (B \cap C')$
- $(A \cup C) \cap (B \cup C) \subseteq A \cup B$

The results in this subsection are used freely throughout these notes, usually without overt mention.

A.3 Power-set Axiom

A very basic axiom is that every subclass of a set is actually a set, indeed this axiom goes further and postulates that they are all elements of a common set.

Power-set Axiom: If a is a set, then there is a set b with every subclass of a an element of b. Symbolically this is expressed as:

$$\forall a \, \exists b \, \forall C \; (C \subseteq a \Rightarrow C \in b)$$

This axiom doesn't specify the set consisting of *exactly* the subsets of a given set, but the axiom of specification rectifies this and so we can define the power set of any set.

Definition A.10: $\mathcal{P}(a) = \{s : s \subseteq a\}$
This is called **power set** of the set a.

Actually $\mathcal{P}(A) = \{s : s \subseteq A\}$ is perfectly well-defined for an arbitrary class A but is of only passing interest when A is a proper class. In particular notice that $\mathcal{P}(V) = V$ $(SP(V) \subseteq V$ because $\mathcal{P}(V)$ is a class; every element of V is a set, therefore a subset of V and so an element of $\mathcal{P}(V)$. $\blacksquare$)

There are at least a couple of reasons for the name "power set". The first is the observation that if a is a finite set with n elements, then the number of elements in $\mathcal{P}(a)$ is 2^n. The second will be explained in the next section where we see the connection between $\mathcal{P}(a)$ and the set $\mathbf{2}^a$ of all functions from the set a to a two element set.

As every subclass of a set is itself a set, and we will emphasize this by calling them *subsets*.

In the previous section we noted that the intersection of two sets was not guaranteed to be a set, but now we see that

$$a \cap b = \{x \in a : x \in b\}$$

which is also equal to

$$\{x \in b : x \in a\}$$

and so equal to

$$\{x : x \in a \wedge x \in b\}$$

is a subclass of the set a (and a subclass of the set b) and so is a set.

Essentially this same argument applies to any *non-empty* collection of sets. If $c \in \mathcal{C}$, then $\bigcap \mathcal{C} \subseteq c$ and so $\mathcal{C}$ is a set.

By contrast, as noted in the preceding section, $\bigcap \emptyset = V$.

We also have $\bigcap V = \emptyset$ and $\bigcup V = V$. The first comes from noting that as $\emptyset$ is a set it is an element of V and so $\bigcap V \subseteq \emptyset$ and so is equal to $\emptyset$. While the second comes from noting that if $x \in V$, then $x \in \mathcal{P}(x)$ and $\mathcal{P}(x) \in V$ so $x \in \bigcup V$. But as with every class $\bigcup V \subseteq V$ and so $\bigcup V = V$.

At this point in our development we still have a paucity of sets. Besides the empty set we now have $\mathcal{P}(\emptyset)$ which is the one element set $\{\emptyset\}$, and we can iterate to get $\mathcal{P}(\mathcal{P}(\emptyset))$, etc. But, for example, we still have no guarantee that the union of two sets is a set.

We can however exhibit various proper classes. In the previous section (see p. 181) we saw that the Russell class $R = \{x : x \notin x\}$ is a proper class. But now from that we can conclude that V is a proper class as well, for $R \subseteq V$ and if V were a set then R would be a subset, not a proper class.

A.4 Axiom of Pairs

For any set x we have the *singleton* class $\{x\}$ that has just x as a member. Following Kelley [36, p. 258], it is convenient to extend the definition of a singleton to allow x to be a proper classes as follows.

Definition A.11: $\{X\} = \{y : X \in V \Rightarrow y = X\}$

So when X is a set, $\{X\}$ is still the singleton class with just the one member, but when X is a proper class $\{X\}$ is the universe V.

Notice that when x is a set we have the set $\mathcal{P}(x)$ with $\{x\} \subseteq \mathcal{P}(x)$ and so $\{x\}$ is also a set. The next axiom extends this to pairs of sets.

Axiom of Pairs: If x and y are sets, then $x \cup y$ is a set.

Definition A.12: $\{X, Y\} = \{X\} \cup \{Y\}$

And $\{X, Y\}$ is called a **pair** (or, for emphasis, an **unordered pair**).

Note that if either X or Y is a proper class, then $\{X, Y\} = V$, while for sets x and y we have $\{x, y\} = \{z : z = x \vee z = y\}$.

This is another place where there is an inversion between the development of set theory without mention of classes (e.g. ZFC) and the class based versions. In ZFC pairs are introduced via an axiom of pairs and then singletons are defined as pairs in which the two elements are the same, viz $\{a\} = \{a, a\}$, as contrasted to the definition here of a pair as a union of singletons.

With three sets a, b, and c we can define: $\{a, b, c\} = \{a, b\} \cup \{c\}$. The set $\{a, b, c\}$ is also equal to $\{a\} \cup \{b, c\}$, and precisely because these are equal we will freely write $\{a\} \cup \{b\} \cup \{c\}$ for any and all of them. The definition and notation for more terms is obvious and we will freely write $\{a_1, a_2, \cdots, a_n\}$

for the set which has as members exactly a_1, a_2, ..., a_n. Also we will follow common convention and write $a_1 \cup a_2 \cup \cdots \cup a_n$ and

$$\bigcup_{i=1}^{n} a_i$$

for all of these.

A.5 Union Axiom

Even though we have not yet formally defined "finite", we know from the previous section that finite unions of sets, i.e., $\bigcup_{i=1}^{n} a_i$, are also sets. The Union Axiom extends this to unions of sets of sets.

Union Axiom: If x is a set, then so is $\bigcup x$.

In studying the algebraic properties of operations on sets, intersection behaves like a product, but we don't yet have the appropriate "sum" of two sets. It is usually called the "symmetric difference".

Definition A.13: The **symmetric difference** (or **Boolean sum**) of two classes A and B is

$$A + B = (A - B) \cup (B - A)$$

When A and B are sets, $A + B$ is a set as well.

Other common notation for $A + B$ includes $A \triangle B$ and $A \oplus B$.

This operation is commutative ($A+B = B+A$), associative ($A+(B+C) = (A + B) + C$), has an identity ($A + \emptyset = A$), and every class is its own inverse ($A + A = \emptyset$). Also intersection distributes ($A \cap (B + C) = (A \cap B) + (A \cap C)$). Finally $A + (B - A) = B$. When a and b are subsets of a common set x, then $a + b$ is also a subset of x.

Suitably translated this says that $(\mathcal{P}(x), +, \cap)$ is a Boolean ring with $\emptyset$ as the zero and x as the multiplicative identity. For the details see Section B.6.5.

When dealing with subsets of a fixed set x it is convenient and common to use the notation a' to mean the complement of a in x. Then the basic facts become the following:

- $(a')' = a$
- $\emptyset' = x$; $x' = \emptyset$
- $x' = \emptyset$; $\emptyset' = x$
- $a \cup a' = x$; $a \cap a' = \emptyset$
- De Morgan laws: $(a \cup b)' = a' \cap b'$; $(a \cap b)' = a' \cup b'$
- $a - b = a \cap b'$
- $a - (a - b) = a \cap b$

- $a \cap (b - c) = (a \cap b) - (a \cap c)$

where all of the sets are subsets of x.

With a little extension of the notation the De Morgan laws can be generalized to general unions and intersections.

For any sub-collection $\mathcal{C}$of $\mathcal{P}(x)$, define $\mathcal{D} = \{a \in \mathcal{P}(x) : a' \in \mathcal{C}\}$, and write

$$\bigcup_{a \in \mathcal{C}} a' \quad \text{and} \quad \bigcap_{a \in \mathcal{C}} a'$$

for the union and intersection of $\mathcal{D}$. Then the generalized De Morgan laws are:

$$(\bigcup_{a \in \mathcal{C}} a)' = \bigcap_{a \in \mathcal{C}} a'$$
$$(\bigcap_{a \in \mathcal{C}} a)' = \bigcup_{a \in \mathcal{C}} a'$$

and they are immediate from the definitions.

A few other simple facts include:

- $\bigcap_{a \in \mathcal{C}} \mathcal{P}(a) = \mathcal{P}(\bigcap_{a \in \mathcal{C}} a)$
- $\bigcup_{a \in \mathcal{C}} \mathcal{P}(a) \subseteq \mathcal{P}(\bigcup_{a \in \mathcal{C}} a)$
- if $x \subseteq y$, then $\mathcal{P}(x) \subseteq \mathcal{P}(y)$
- if $\emptyset \in \mathcal{C}$, then $\bigcap_{a \in \mathcal{C}} = \emptyset$
- if $x \in \mathcal{C} \subseteq \mathcal{P}(x)$, then $\bigcup_{a \in \mathcal{C}} = x$

A.6 Ordered Pairs and Cartesian Products

In notable contrast to a set, which is entirely determined by its elements, a list depends on the order of its elements. So the set $\{a, b, a, c\}$ is exactly the same as the set $\{c, a, c, b, b\}$, both being the three element set $\{a, b, c\}$, while the lists (a, b, a, c) and (c, a, c, b, b) are entirely different, even having a different number of elements. The special case of a list with two elements, usually called an *ordered pair*, is particularly important. The characteristic property of an ordered pair (a, b) is that it is equal to another ordered pair (c, d) iff $a = c$ and $b = d$. It is possible to enrich the theory of sets with an additional primitive notion of an ordered pair, or even the more elaborate notion of a list, but fortunately that is unnecessary – it is possible to define an ordered pair as a particular kind of set.

Definition A.14: $(a, b) = \{\{a\}, \{a, b\}\}$

(a, b) is called the **ordered pair** of the sets a and b. From the Axiom of Pairs it is certainly a set.

Proposition A.25 $(a,b) = (c,d)$ *iff* $a = c$ *and* $b = d$

$\blacksquare$

Halmos [29, Section 6] has a nice explanation as to how this definition arises, and also proves the proposition. This proposition is also proved as Theorem 1.33 in Monk [58] and on page 259 of Kelley [36]. The proofs are all somewhat different in interesting ways.

The primary reason for introducing ordered pairs is to allow the following definition.

Definition A.15: $A \times B = \{(a,b) : a \in \mathbf{A}, b \in \mathbf{B}\}$

$A \times B$ is called the **Cartesian product** of the classes A and B.

For any *sets* $\mathcal{A}$ and $\mathcal{B}$, if $a \in \mathcal{A}$ and $b \in \mathcal{B}$, then both $\{a\}$ and $\{a,b\}$ are subsets of $\mathcal{A} \cup \mathcal{B}$ and so elements of $\mathcal{P}(\mathcal{A} \cup \mathcal{B})$. Thus $(a,b) = \{\{a\},\{a,b\}\}$ is a subset of $\mathcal{P}(\mathcal{A} \cup \mathcal{B})$ and so an element of $\mathcal{P}(\mathcal{P}(\mathcal{A} \cup \mathcal{B}))$. This makes it clear that $\mathcal{A} \times \mathcal{B} = \{(a,b) : a \in \mathcal{A}, b \in \mathcal{B}\}$ is a *subset* of $\mathcal{P}(\mathcal{P}(\mathcal{A} \cup \mathcal{B}))$.

The sole purpose of introducing the rather odd *definition* of an ordered pair was to allow the construction of Cartesian products without introducing an independent concept and additional axioms. The definition itself will not be used going forward, but only the basic property that $(a,b) = (c,d)$ iff $a = c$ and $b = d$. Thus, for example, the strange fact that $\{a,b\}$ is an element of (a,b) will **not** be used, and is a somewhat embarrassing artifact. It is one of the reasons that fans of category theory are often annoyed by axiomatic set theory as *the* modern foundation for mathematics. A good explanation of this perspective is given in Barr [2].

As you should suspect this is not the only possible construction that will serve as an ordered pair. For example defining $[a,b] = \{\{\emptyset, a\}, \{\{\emptyset\}, b\}\}$ you can readily prove that $[a,b]$ is a set and that $[a,b] = [c,d]$ iff $a = c$ and $b = d$. The choice of what to use as the actual definition of an ordered pair is largely arbitrary.

As with previous constructions, we record some basic facts about Cartesian products which will be used freely when needed: For $\mathcal{A}$, $\mathcal{B}$, $\mathcal{X}$ and $\mathcal{Y}$ any sets we have

- $\mathcal{A} \times \mathcal{B} \neq \emptyset$ iff $\mathcal{A} \neq \emptyset$ and $\mathcal{B} \neq \emptyset$.
- If $\mathcal{A} \subseteq \mathcal{X}$ and $\mathcal{B} \subseteq \mathcal{Y}$, then $\mathcal{A} \times \mathcal{B} \subseteq \mathcal{X} \times \mathcal{Y}$.
- If $\mathcal{A} \times \mathcal{B} \neq \emptyset$ and $\mathcal{A} \times \mathcal{B} \subseteq \mathcal{X} \times \mathcal{Y}$, then $\mathcal{A} \subseteq \mathcal{X}$ and $\mathcal{B} \subseteq \mathcal{Y}$.
- $(\mathcal{A} \cup \mathcal{B}) \times \mathcal{X} = (\mathcal{A} \times \mathcal{X}) \cup (\mathcal{B} \times \mathcal{X})$.
- $(\mathcal{A} \cap \mathcal{B}) \times \mathcal{X} = (\mathcal{A} \times \mathcal{X}) \cap (\mathcal{B} \times \mathcal{X})$.
- $(\mathcal{A} - \mathcal{B}) \times \mathcal{X} = (\mathcal{A} \times \mathcal{X}) - (\mathcal{B} \times \mathcal{X})$.

Products are a fundamental notion in category theory with the Cartesian product of sets as the basic example, but we can not yet discuss the category of sets using axiomatic set theory as we do not yet have functions! Functions will actually be defined via Cartesian products, and first we discuss the more general notion of relations.

Even more, here in set theory the fundamental notion is really the ordered pair and Cartesian products *ARE* particular sets of ordered pairs, while in the category **Set** the particular *object* is not important, only the Universal Mapping Property of the object *together with its projection morphisms* (which we have not yet even defined!)

A.7 Relations

Throughout this section script letters such as $\mathcal{A}$, $\mathcal{B}$, $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{R}$, as well as small letters, will denote sets so that we can conveniently write $a \in \mathcal{A}$, etc. with the implication that $\mathcal{A}$ is a set of sets.

Definition A.16: A **relation** from X to Y is a subclass of $X \times Y$.

This is often called a *binary relation*, and ternary and even general *n-ary* relations are defined as subsets of $X \times Y \times Z$ and of $X_1 \times \cdots \times X_n$ respectively. In this appendix we will restrict attention to binary relation and just refer to them as relations.

There are other closely related definitions as well.

Definition A.17: If R is a relation from X to Y, the **domain of definition** of R is the class

$$\text{domain}(R) = \{x \in X : \exists y \in Y\ [(x, y) \in R]\}$$

Definition A.18: The **range** of R is the class

$$\text{range}(R) = \{y \in Y : \exists x \in X\ [(x, y) \in R]\}$$

For every relation R from X to Y we have $\text{domain}(R) \subseteq X$ and $\text{range}(R) \subseteq Y$, and both of these inclusions may be strict.

When X and Y are sets, any relation from X to Y is a set, as are the domain of definition and range of the relation.

By contrast to the definition here it is usual to define a relation to be any class of ordered pairs. It is then easy to see that a relation R will be a subclass of $\text{domain}(R) \times \text{range}(R)$. The reason for the different definition here is for convenience in discussing the category of relations (see Section B.1.3.) A notational confusion arises as a result – usually the domain of definition of a relation is simply called the domain, but in the category **Rel** a morphism from the set X to the set Y is a relation from X to Y so X is the domain (and Y is the codomain of the relation.) Hence our use of the term *domain of definition* even though we still write domain(R). In practice this seldom causes confusion.

Examples of relations include:

- The empty set is a relation from any class to any other class.
- $X \times Y$ is a relation from X to Y.
- The **equality relation** or **diagonal** on any class X is the class $\{(x, x) : x \in X\}$. We usually denote this by E_X or Δ_X. The subscript X will be omitted if it is clear. In particular there is E_V with V the universal class. This is usually denoted by just E or Δ. Observe that for any class X we have $\Delta_X = \Delta \cap X$.
- For any set $\mathcal{X}$ there is the relation $\mathcal{E}$ from $\mathcal{X}$ to $\mathcal{P}(\mathcal{X})$ with $\mathcal{E} = \{(x, \mathcal{A}) : x \in \mathcal{A}\}$.

There is a variety of notation, terminology and definitions that we need to record. If R is a relation, we usually write xRy instead of $(x, y) \in R$. If R is a relation from X to itself we say that R is a *relation on* X.

In all of the following definitions, R is a relation on X.

Definition A.19: R is **reflexive** iff xRx for every $x \in X$.

Definition A.20: R **symmetric** iff xRy implies yRx.

Definition A.21: R is **antisymmetric** iff xRy and yRx implies $x = y$.

Definition A.22: R is **connected** iff always xRy or yRx.

Definition A.23: R is **transitive** iff xRy and yRz implies xRz.

Definition A.24: R is an **equivalence relation** iff it is reflexive, symmetric and transitive.

The fundamental equivalence relation on any class is the equality relation.

As classes, relations from X to Y can be compared to see if they are subclasses of one another. A relation R is **smaller** or **finer** than a relation S when $R \subseteq S$. Equally well we say that S is **larger** or **coarser** than R. With this terminology, a relation R on X is reflexive iff it is coarser than the equality relation.

Other important special types of relations, partial orders, are discussed at length in Section A.12.

Associated to each relation from X to Y is the *opposite* or *inverse* relation gotten by reversing the coordinates in each ordered pair in the relation. Symbolically

$$\mathcal{R}^{\text{op}} = \mathcal{R}^{-1} = \{(y,x) \in \mathcal{Y} \times \mathcal{X} : (x,y) \in \mathcal{R}\}$$

Observe that a relation R on X is symmetric iff $R = R^{-1}$.

Relations can also be composed as in the next definition.

Definition A.25: If R is a relation from X to Y and $S$is a relation from Y to Z then the **composition** of R followed by S is

$$\mathcal{S} \circ \mathcal{R} = \{(x,z) \in \mathcal{X} \times \mathcal{Z} : \exists y \in \mathcal{Y} \text{ with } (x,y) \in \mathcal{R} \wedge (y,z) \in \mathcal{S}\}$$

As a trivial example, if R is any relation from X to Y and E_X is the equality relation on X, then $R \circ E_X = R$. Of course if E_Y is the equality relation on Y, then equally well $E_Y \circ R = R$.

A relation R on X is transitive iff $R \circ R = R$.

The intersection of any family of transitive relations on a *set* $\mathcal{X}$ is again a transitive relation on $\mathcal{X}$. For a particular relation $\mathcal{R}$ on $\mathcal{X}$, the intersection of the family of all transitive relations on $\mathcal{X}$ that include $\mathcal{R}$gives a smallest transitive relation that is coarser than $\mathcal{R}$. This is called the **transitive closure** of $\mathcal{R}$. The same thing is true about reflexive relations and symmetric relations, so we also have the **reflexive closure** and **symmetric closure** of any relation $\mathcal{R}$ on $\mathcal{X}$.

These closures can be described in other ways as well. The reflexive closure of a relation $\mathcal{R}$ is $E \cup \mathcal{R}$, with E the identity relation, while the symmetric closure of $\mathcal{R}$ is $\mathcal{R} \cup \mathcal{R}^{-1}$. The transitive closure of $\mathcal{R}$ can be described in terms of $\mathcal{R} \circ \mathcal{R}$, $\mathcal{R} \circ (\mathcal{R} \circ \mathcal{R})$ and so on, but the details will have to wait until Section A.10. This alternative description actually allows us to define the transitive closure on an arbitrary relation, not just a relation on a set.

In the same vein, the intersection of any family of equivalence relations on $\mathcal{X}$ is an equivalence relation on $\mathcal{X}$, so the intersection of the family of all equivalence relations on $\mathcal{X}$ that are larger than $\mathcal{R}$ is the smallest equivalence relation on $\mathcal{X}$ including $\mathcal{R}$. This is called the **equivalence relation generated by** $\mathcal{R}$. The alternative description of the transitive closure also allows us to define the equivalence relation generated by an arbitrary relation, not just a relation on a set.

Equivalence relations are tightly intertwined with two other basic notions – partitions and quotient sets.

Definition A.26: When R is an equivalence relation on X and $x \in X$, the **equivalence class** of x with respect to R is the class $\{y \in X : xRy\}$.

The basic property of equivalence classes is the following proposition.

Proposition A.26 *With R an equivalence relation on X, $x/R = y/R \iff xRy$.*

There is no really common notation for equivalence classes, but we will usually write x/R. When X is a set the equivalence classes for any relation on X are subsets of X. When X is a proper class, the status of the equivalence classes is varied. For example on any class X and any element x, the equivalence class for the equality relation is the set $x/E = \{x\}$. At the other extreme taking R to be the equivalence relation $X \times X$ on X, all elements are equivalent and so the one equivalence class is X itself. So if X is a proper class, then so is this equivalence class. All kinds of mixtures of sets and proper classes are possible.

Our primary interest in equivalence relations is because of their connection with quotient sets (see I.26.)

Definition A.27: $\mathcal{X}/\mathcal{R} = \{y \in \mathcal{P}(\mathcal{X}) : \exists x \in \mathcal{X}(y = x/\mathcal{R})\}$

With $\mathcal{R}$ an equivalence relation on the set $\mathcal{X}$, the set $\mathcal{X}/\mathcal{R}$ is called the "quotient of $\mathcal{X}$ by $\mathcal{R}$." It is also called "$\mathcal{X}$ modulo $\mathcal{R}$" or "$\mathcal{X}$ mod $\mathcal{R}$". Such sets in general are called *quotient sets.*

As a collection of sets $\mathcal{X}/\mathcal{R}$ has three characteristic properties. First it is disjoint.

Definition A.28: Two classes A and B are **disjoint** when $A \cap B = \emptyset$. A class X is **disjoint** provided any two distinct members of X are disjoint.

$\mathcal{X}/\mathcal{R}$ is disjoint, for if $y \in x/\mathcal{R} \cap z/\mathcal{R}$, then $x\mathcal{R}y$ and $z\mathcal{R}y$ so also $y\mathcal{R}z$ and then $x\mathcal{R}y$. Whereupon $x/\mathcal{R} = y/\mathcal{R}$.

Second every element of $\mathcal{X}/\mathcal{R}$ is non-empty for (x, x) is always in $x/\mathcal{R}$.

Third the union of $\mathcal{X}/\mathcal{R}$ is $\mathcal{X}$ for every $x \in \mathcal{X}$ belongs to exactly one equivalence class, namely $x/\mathcal{R}$.

The process can be reversed, and for that we start with the relevant definition.

Definition A.29: A **partition** of a set $\mathcal{X}$ is disjoint collection of non-empty subsets of $\mathcal{X}$ whose union is $\mathcal{X}$.

Given a partition $\mathcal{P}$ of $\mathcal{X}$, define a relation $\mathcal{R}$ by saying

$$x\mathcal{R}y \text{ iff } \exists s \in \mathcal{R} \text{ such that } x \in s \wedge y \in s$$

Then $\mathcal{R}$ is an equivalence relation: $\mathcal{X}$ is the union of $\mathcal{P}$, so each $x \in \mathcal{X}$ belongs to some set in $\mathcal{P}$. As x is certainly in the same set as itself, this says that $\mathcal{R}$ is reflexive. $\mathcal{R}$ is symmetric for if x and y are in the same $s \in \mathcal{P}$, then certainly y and x are both in s. And finally $\mathcal{R}$ is transitive because this is just saying that if x and y are in the same set and y and z are in the same set, then x and z are in the same set.

So an equivalence relation $\mathcal{R}$ gives a partition $\mathcal{P}$, while $\mathcal{P}$ in turn gives an equivalence relation which is clearly seen to be $\mathcal{R}$. In the other direction, any partition $\mathcal{P}$ gives rise to an equivalence relation where the set of equivalence classes is $\mathcal{P}$.

Just as with products, the quotient of a set $\mathcal{X}$ by an equivalence relation $\mathcal{R}$really *IS* a specific subset of $\mathcal{P}(\mathcal{X})$, while in the category **Set** the particular *object* is not important, only the universal mapping property of the quotient function (which we'll finally define in the next section!)

A.8 Functions

The mathematical term "function", as a magnitude associated to a curve, was introduced by Leibniz late in the 17th century. By the mid 18th century Euler was using function to describe an analytic expression, e.g. $\sin(x^3)$. During the 19th century many and varied "monster" functions were discovered: Dirichlet's example of a function that is continuous nowhere, or worse a function that is continuous at the irrational numbers and discontinuous at the rational numbers; Weierstrass' function that is continuous everywhere but nowhere differentiable; and various examples of space-filling curves. Cantor's introduction of set theory was motivated by his efforts to understand the variety of functions, particularly those represented by Fourier series. His own first shocking result was in showing that the unit interval was in one-to-one correspondence with the unit square. Throughout this period the concept of function was gradually being clarified and finally reached is current form in the early 20th century.

Because of this long history there are many other words such as *map*, *mapping*, *transformation*, *correspondence*, and *operator* that have been used as synonyms for *function*. And even today we frequently find such definitions as "A variable so related to another that for each value assumed by one there is a value determined for the other", "A correspondence in which values of one variable determine the values of another", or "A rule for associating a member of one set with a member of another set." Although the spirit of these definitions is different, the actually utility is well captured by the following definition.

Definition A.30: A **function** from the class X to the class Y is a relation F from X to Y with the property that for every $x \in X$ there is a *unique* $y \in Y$ with $(x, y) \in F$. That unique y with $(x, y) \in F$ is denoted by $F(x)$. For functions the notation $y = F(x)$ is universally used in place of $(x, y) \in F$ and xFy

The notation $F : X \longrightarrow Y$ means "F is a function from X to Y". As with any relation, a function $F : X \longrightarrow Y$ has a domain of definition and a range. The domain of definition of F is X which is also the domain of F as a morphism in the category of sets. The range may be a proper subclass of Y. More, a function from X to Y does *not* determine Y. In the category of sets, Y is the *codomain* of F. The fact that a function does not specify its codomain is the reason that the morphisms in the category of sets are not just functions.

At times many of the words mentioned above will be used as synonyms for function. In particular we will freely use the phrase *binary operation* on X for a function from $X \times X$ to X.

Just as with relations in general, if X and Y are sets then every function from X to Y is a set, being a subset of $X \times Y$.

When the range of F is equal to Y, we say that F maps X *onto* Y or, more commonly today, that F is *surjective* or is a *surjection*. We may also say that F is an *epimorphism* though this is usually reserved for those cases where the functions of interest are called homomorphisms.

When we write $y = F(x)$ we will speak of the *argument* x and of y as the *value* of F at x.

What we here define to be a function is what is often called the *graph* of the function. For example in elementary mathematics it is common to distinguish between the function $\sin(x)$ and it's graph $\{(x, \sin(x)) \in \mathbb{R}^2 : x \in \mathbb{R}\}$. But here no distinction is made between the two, as all the relevant properties of the function are captured by its graph.

The range of F is also commonly called the *image* of F. When A is any subset of X, the subset

$$F(A) = \{y \in Y : \exists x \in A\ [y = F(x)]\}$$

of Y is called the **image of** A **under** F. Of course $F(X)$ is the image of F.

If Y is not empty, the simplest examples of functions are the *constant functions*: if $y \in Y$, then $X \times \{y\} \subseteq X \times Y$ is a constant function which has the value y at every argument x.

For X a subset of Y, the set $X \times X$ is a function i :X$\longrightarrow Y$ called the *inclusion* or *injection* of X into Y. A far more common way of describing i is to say it is the function defined by $i(x) = x$ for all $x \in X$. The phrase "the function F defined by ..." is the most common way of defining functions, and is intended to imply there is a unique function defined by specifying the value y of the function at each argument x. For example, we usually use y to name a constant function what has always the value y.

In the special case of X considered as a subset of itself, the inclusion function is called the *identity* function: $1_X : X \longrightarrow X$ is defined by $1_X(x) = x$. Of course this is just another name for the equality relation on X.

If $F : X \longrightarrow Y$ and $G : Y \longrightarrow Z$, then we have the composition $G \circ F$ as a relation from X to Z. This is again a function and is just called the composition of the functions. It is often written as just GF. As noted earlier with regard to relations $F1_X = F$ and $1_Y F = F$. Also if z is a constant function from $Y \longrightarrow Z$, then $zF = z$, the second z being the constant function from X to Z. And if y is a constant function from X to Y, then $Gy = G(y)$, where $G(y)$ is the constant function from X to Z with value $G(y)$.

When X is a subset of Y and $F : Y \longrightarrow Z$, we have the inclusion $i : X \longrightarrow Y$ and the composition Fi. This composition is called the **restriction** of F to X and is written $F|_X$. If $G : X \longrightarrow Z$ and $F|_X = G$, then F is called an **extension** of G to Y.

For any two classes X and Y we have the Cartesian product $X \times Y$. Two very important functions in this situation are the **coordinate projections** $\pi_1 : X \times Y \longrightarrow X$ and $\pi_2 : X \times Y \longrightarrow Y$ defined by $\pi_1(x, y) = x$ and $\pi_2(x, y) = y$.

Here we are deliberately violating our notational convention and using the lower case letter π for a function that may be a proper class. As in the notes proper the capital letter Π will be reserved for products.

Looking at the actual definition of π_1 as a subset of $(X \times Y) \times X$, especially taking it all the way back to the definition of ordered pairs, illustrates the importance of abstracting away the crucial properties. Getting away from any such definitional complexity to the critical universal mapping property of products in the category **Set** is another such step.

Another general example of a function comes from an equivalence relation. For any equivalence relation $\mathcal{R}$ on the *set* $\mathcal{X}$, define the *quotient map* $q : X \longrightarrow X/R$ by $q(x) = x/R$. Clearly q is surjective. [For the significance of this in the categorical context, see 22.]

As shown in exercise I.31 every function $f : \mathcal{X} \longrightarrow \mathcal{Y}$ defines an equivalence relation on $\mathcal{X}$ via $x \equiv_F x'$ iff $F(x) = F(x')$. Now the projection map $q : \mathcal{X} \longrightarrow \mathcal{X}/\equiv_f$ is always surjective, and as shown in exercise I.32 there is a unique function $\overline{f} : \mathcal{X}/\equiv_f \longrightarrow \mathcal{Y}$ with $f = \overline{f}q$.

The function $\overline{f}$ has the important property that if $\overline{f}(x/\equiv_f) = \overline{f}(x'/\equiv_f)$, then $x/\equiv_f = x'/\equiv_f$.

A function $F : X \longrightarrow Y$ is called *one-to-one* or, more commonly today, *injective* or an *injection* when $F(x) = F(x') \Rightarrow x = x'$. We may also say the function is a *monomorphism* though this is usually reserved for those cases where the functions of interest are called homomorphisms.

A function that is both an injection and a surjection (i.e., is one-to-one and onto) is called a *bijection*.

Every function $F : X \longrightarrow Y$, considered as a relation, has an inverse relation F^{-1}. If the inverse relation is also a function, then F is a bijection. Conversely if F is a bijection, then F^{-1} is a function. When F is a bijection, then $F^{-1} \circ F = 1_X$ and $F \circ F^{-1} = 1_Y$. If X and Y are two sets where there exists a bijection between them, then we will write $X \sim Y$ and say the two sets are *bijective*, *equivalent*, *equinumerous*, *equipotent* or *have the same number of elements*. Of course we haven't yet defined "number of elements", so the last phrase is a bit premature, but we will rectify that in Section A.9.

For sets $\mathcal{X}$ and $\mathcal{Y}$, the class of all functions from $\mathcal{X}$ to $\mathcal{Y}$ is a subset of $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$ and is denoted by $\mathcal{Y}^{\mathcal{X}}$. The reason this is written as an exponential should become clear in Section A.11 when we actually define an exponential function.

Some basic properties of $\mathcal{Y}^{\mathcal{X}}$ include the following:

- For every set $\mathcal{Y}$, including the empty set, $\mathcal{Y}^{\emptyset} = \{\emptyset\}$.
- If $\mathcal{X}$ is not the empty set, then $\emptyset^{\mathcal{X}} = \emptyset$.
- $\mathcal{P}(\mathcal{X}) \sim \mathbf{2}^{\mathcal{X}}$: Let $\mathbf{2}$ denote a set with exactly two distinct elements (which we will call 0 and 1). For any set $\mathcal{X}$ consider the set $\mathbf{2}^{\mathcal{X}}$. Whenever $\mathcal{A}$ is a subset of $\mathcal{X}$, define the *characteristic function* of $\mathcal{A}$ to be

$$\chi_{\mathcal{A}} : \mathcal{X} \longrightarrow \mathbf{2} \text{ with } \chi_{\mathcal{A}}(x) = 1 \text{ if } x \in \mathcal{A}, \chi_{\mathcal{A}} = 0 \text{ otherwise}$$

From this we get the function from $\mathcal{P}(\mathcal{X})$ to $\mathbf{2}^{\mathcal{X}}$ defined by $\mathcal{A} \mapsto \chi_{\mathcal{A}}$, and this function is a bijection. The inverse function is given by associating to any map $\chi : \mathcal{X} \longrightarrow \mathbf{2}$ the subset $\{x \in \mathcal{X} : \chi(x) = 1\}$.

- $(\mathcal{Z}^{\mathcal{Y}})^{\mathcal{X}} \sim \mathcal{Z}^{(\mathcal{X}\times\mathcal{Y})}$: The bijection from $(\mathcal{Z}^{\mathcal{Y}})^{\mathcal{X}}$ to $\mathcal{Z}^{(\mathcal{X}\times\mathcal{Y})}$ is given by $\phi \mapsto \Phi$ where $\Phi(x, y) = \phi(x)(y)$. The inverse is just gotten by reading this formula the other way – a function Φ of two variables defines a function of the first variable (x) whose value is a function of the second variable (y). In theoretical computer science this process of going from a function of two variables to a function of one variable with values being other functions is called *currying* in honor of the logician Haskell Curry, though this observation long predates Curry.

- There are other variations on the idea that sending a function to its value at a particular argument defines a function whose arguments are functions. The simplest version of this is the function

$$\text{eval} : \mathcal{X} \longrightarrow \mathcal{Y}^{(\mathcal{Y}^{\mathcal{X}})}$$

defined by eval(x)(f) = f(x).

A.8.1 Families and Cartesian Products

There are many situations where functions are used in a fashion that leads to very different notation and terminology. The most familiar examples are sequences such as 1, 1/2, 1/3, 1/4, ... where nearly all the details of the function are at best implicit. (Sequences will be explicitly defined in the next section.)

The functions of interest here are called **indexed families** or just **families**. A family is a function x from an **index set** I to the **indexed set** X. An element i of I is called an **index**, while the value of x at i is called the $i-$ th **term** and is denoted by x_i. In this section we will be discussing families of sets, and the common terminology is something like "a family $\{A_i\}$ of subsets of X" by which we understand a function from some index set I to $\mathcal{P}(X)$. It is even more common to use just "a family $\{A_i\}$ of sets" leaving both the index set and the indexed set unspecified!

If $\{A_i\}$ is a family of sets of X, then the union of the family, written

$$\bigcup_{i\in I} A_i \quad \text{or even} \quad \bigcup_i A_i,$$

is the union of the range of the family. From the definition of union we see that $x \in \bigcup_i A_i$ iff x belongs to A_i for at least one i in I.

Similarly the intersection of a family $\{A_i\}$ of sets, written

$$\bigcap_{i\in I} A_i \quad \text{or even} \quad \bigcap_i A_i,$$

is the intersect of the range of the family. Unlike unions this only makes sense if this is a non-empty family, i.e., the index set is non-empty. From the definition of intersection we see that $x \in \bigcap_i A_i$ iff x belongs to A_i for every index i.

The basic identities about unions and intersections that were listed for three sets in Section A.5 generalize immediately to families:

- $B \cap \bigcup_i A_i = \bigcup_i (B \cap A_i)$
- $B \cup \bigcap_i A_i = \bigcap_i (B \cup A_i)$
- $(\bigcup_i A_i) \cap (\bigcup_j B_j) = \bigcup_{i,j} (A_i \cap B_j)$
- $(\bigcap_i A_i) \cup (\bigcap_j B_j) = \bigcap_{i,j} (A_i \cup B_j)$

where $\{A_i\}$ and $\{B_j\}$ are families of sets and B is a set, and the notation $\bigcup_{i,j}$ is shorthand for $\bigcup_{(i,j) \in I \times J}$.

In discussing unions and intersections of families, this is nothing but a convenient if sloppy notation as the collections could be used equally well. And the reverse is true as well as any collection can be considered as a family indexed by itself. But when generalizing Cartesian products families, in one guise or another, are essential.

Definition A.31: If $\{X_i\}$ is a family of sets with index set I, the **Cartesian product** is the set of all families $\{x_i\}$ with $x_i \in X_i$ for every i in I. It is denoted by

$$\Pi_{i \in I} X_i \quad \text{or} \quad \Pi_i X_i$$

As $\Pi_i X_i \subseteq (\bigcup_i X_i)^I$, so it is a set.

For each $j \in I$, the **projection** on the j-th coordinate is the function $\pi_j : \Pi_{i \in I} X_i \longrightarrow X_j$ defined by $\pi_j(\{x_i\}) = x_j$

If $\{X_i\}$ is a family of sets with index set $\{1, 2\}$ with 1 and 2 being two unspecified but unequal elements, then we have both

$$X_1 \times X_2 \quad \text{and} \quad \Pi_{i \in \{1,2\}} X_i$$

As these have been defined, they are *very* different sets, nonetheless it is customary to identify the set $\Pi_{i \in \{1,2\}} X_i$ of functions with the Cartesian product $X_1 \times X_2$ a set of ordered pairs as defined earlier. The justification is a "natural" bijection between the two sets. In the category **Set**, both of these sets together with their projections are products of X_1 and X_2, and the natural bijection is just the unique function between the two that commutes with the projections.

There are a pair of equations that relate unions, intersections and Cartesian products:

- $(\bigcup_i A_i) \times (\bigcup_j B_j) = \bigcup_{i,j} (A_i \times B_j)$
- $(\bigcap_i A_i) \times (\bigcap_j B_j) = \bigcap_{i,j} (A_i \times B_j)$

where, of course, the intersections must be of non-empty families.

All of the classes in this section were assumed to be sets, but much of the discussion can be carried through for classes more generally. It is a good exercise to verify just how much can be done.

A.8.2 Images

Every relation $\mathcal{R}$ from $\mathcal{X}$ to $\mathcal{Y}$ defines two functions between the power sets of $\mathcal{X}$ and $\mathcal{Y}$. The first is $\mathcal{R} : \mathcal{P}(\mathcal{X}) \longrightarrow \mathcal{P}(\mathcal{Y})$ defined by

$$\mathcal{R}(\mathcal{A}) = \{y \in \mathcal{Y} : \exists a \in \mathcal{A}\ (a\mathcal{R}y)\}$$

for every subset $\mathcal{A}$ of $\mathcal{X}$, while the second is $\mathcal{R}^{-1} : \mathcal{P}(\mathcal{Y}) \longrightarrow \mathcal{P}(\mathcal{X})$ defined by

$$\mathcal{R}^{-1}(\mathcal{B}) = \{x \in \mathcal{X} : \exists b \in \mathcal{B}\ (x\mathcal{R}b)\}$$

for $\mathcal{B}$ any subset of $\mathcal{Y}$.

Definition A.32: $\mathcal{R}(\mathcal{A})$ is the **image** or **direct image** of $\mathcal{A}$ under $\mathcal{R}$.

In Bourbaki's useful phrase [9, p. VI] the use of $\mathcal{R}$ as the name of the function as well as the relation which produces it is an abuse of notation, but it improves readability while seldom creating confusion.

Definition A.33: $\mathcal{R}^{-1}(\mathcal{B})$ is the **inverse image** of $\mathcal{B}$ under $\mathcal{R}$.

$\mathcal{R}^{-1}(\mathcal{B})$ has another possible meaning, the image of $\mathcal{B}$ under the relation $\mathcal{R}^{-1}$. Happily the two meanings are the same.

A.9 Natural Numbers

In Section A.8 we said that two sets have the same number of elements when they are bijective, i.e., there is a bijection between them. This equinumerous relation between sets is an equivalence relation on the universe V:

1. Every set is equinumerous with itself – the identity function is a bijection from the set to itself.

2. If X is equinumerous with Y, then Y is equinumerous with X – for if there is a bijection from X to Y, then its inverse is a bijection from Y to X.

3. If X is equinumerous with Y and Y is equinumerous with Z, then X is equinumerous with Z – for if f is a bijection from X to Y and g is a bijection from Y to Z, then gf is a bijection from X to Z with inverse $f^{-1}g^{-1}$.

This suggests defining the "number of elements" in a set to be the equinumerous equivalence class of that set. But with the exception of the equivalence class of $\emptyset$ (which is $\{\emptyset\}$) all of the equivalence classes are proper classes, i.e., there is not even a class of "numbers". An alternative approach is to define numbers to be certain canonical sets, one from each such "equivalence class" and that is what will be done here. Eventually we will associate to every set a "cardinal number" as the number of elements in the set. But the first step is to define the natural numbers 0, 1, 2, ... within set theory.

The starting point is the number zero, and for that there is only one possibility.

Definition A.34: The number **zero** is the empty set. When we are speaking of it as a number it will, of course, be written as 0.

The next step is to introduce a standard method for getting the next larger number. The very natural idea is to define $n + 1 = \{0, 1, \cdots, n\}$. Notice that as n should be $\{0, 1, \cdots, n - 1\}$ we can equally write this as $n + 1 = \{n\} \cup n$ which is very fortunate as we can make that a quite general definition.

Definition A.35: For any set x, the **successor** of x, written x^+, is $x \cup \{x\}$.

So $0^+ = \{0\}$ and has a single element. Of course we are going to name it 1. Similarly $1^+ = \{0, 1\}$ is a set with two elements, and it will be called 2. We can keep going in this fashion as long as we want, but nothing we have yet guarantees the existence of a set of all the successors of 0. For that we need another axiom.

Axiom of Infinity: There exists a set containing 0 and containing the successor of each of its elements.

Solely for the purposes of the following discussion, we say that a set Ω is a *successor set* provided $0 \in \Omega$ and $x \in \Omega \Rightarrow x^+ \in \Omega$. The Axiom of Infinity guarantees that at least one such set exists, and because Ω^+ is also a successor set there are many others. More importantly there is a smallest successor set. The intersection of any family of successor sets is a successor set, so the intersection of *all* successor sets in Ω is itself a successor set which is denoted by ω. The set ω is actually a subset of *every* successor set (for if S is an successor set, then $S \cap \omega$ is a both a subset and a superset of ω and so must be ω.) With these preliminaries we make the following definition.

Definition A.36: The minimal successor set ω which is a subset of every successor set is the set of **natural numbers**, and the elements of ω are called natural numbers.

The justification for calling ω the set of natural numbers is in the next two sections where the connection with the Peano Axioms and arithmetic are explored.

The use of ω to denote the set of natural numbers is common, but not universal, in the literature of set theory, so it will be used throughout this appendix. Elsewhere in these notes the notation $\mathbb{N}$ is used consistently for the set of natural numbers.

Definition A.37: A family with index set ω is called an **infinite sequence**, while a family with index set a natural number is called a **finite sequence**.

If $\{a_i\}$ is a finite sequence of sets indexed by the natural number n^+, then the union of the sequence is denoted by

$$\bigcup_{i=0}^{n} a_i \quad \text{or} \quad a_0 \cup \cdots \cup a_n.$$

Similarly if $\{a_i\}$ is an infinite sequence of sets, then the union of the sequence is denoted by

$$\bigcup_{i=0}^{\infty} a_i \quad \text{or} \quad a_0 \cup a_1 \cup \cdots .$$

Similar notation is used for intersections and Cartesian products of sequences:

$$\bigcap_{i=0}^{n} a_i \quad \text{or} \quad a_0 \cap \cdots \cap a_n.$$

$$\Pi_{i=0}^{n} a_i \quad \text{or} \quad a_0 \times \cdots \times a_n.$$

$$\bigcap_{i=0}^{\infty} a_i \quad \text{or} \quad a_0 \cap a_1 \cap \cdots .$$

$$\Pi_{i=0}^{\infty} a_i \quad \text{or} \quad a_0 \times a_1 \times \cdots .$$

There are a number of other uses of the word "sequence" which are all variations on this. We will freely use such notation as

$$\bigcup_{i=3}^{\infty} a_i$$

with the expectation that the meaning is clear.

A.10 Peano Axioms

This section is a digression from the development of axiomatic set theory to justify calling ω the set of natural numbers, and also to indicate the initial part of the development of (nearly) all of mathematics from set theory.

The starting point are the axioms for the foundation of arithmetic which were introduced by Giuseppe Peano in his book *Arithmetices principia, nova methodo exposita* published in 1889.

Recast in more modern language Peano showed the set $\mathbb{N}$ of natural numbers could be specified by the following Peano axioms.

1. $0 \in \mathbb{N}$.
2. Every natural number n has a successor, denoted n'.
3. 0 is not the successor of any natural number.
4. If $n' = m'$, then $n = m$.
5. (Induction Principle) If $Q \subseteq \mathbb{N}$ with
   - $0 \in Q$, and
   - $q \in Q \Rightarrow q' \in Q$

   then $Q = \mathbb{N}$.

The successor on n is what we commonly call $n+1$ (but the Peano Axioms allow us to define addition and *prove* that $n' = n + 1$.)

The Induction Principle captures what is often call the Principle of Mathematical Induction and stated as:

Let $p(n)$ be a proposition depending on integer variable n. If $p(0)$ is true and $\forall k \geq 0[p(k) \Rightarrow p(k+1)]$, then $p(n)$ is true for all $n \geq 0$.

The Principle of Mathematical Induction follows from Peano's Induction Principle by taking $Q = \{n \in \mathbb{N} : p(n)\}$. Q contains 0 because $p(0)$ is true by assumption, and $k \in Q \Rightarrow k' \in Q$ exactly because $p(k) \Rightarrow p(k+1)$. Of course we are cheating a bit because we have identified k' with $k+1$ and the justification for that will actually happen in the next section.

Prior to showing that the Peano Axioms are, suitably construed, theorems of set theory, we will explore some of the basic consequences of these axioms. To do that we first formalize the following variant concept.

Definition A.38: A **Peano system** consists of a set P together with a distinguished element $p_0 \in P$ and a function $S : P \longrightarrow P$ satisfying:

1. $\forall p \in P[p_0 \neq S(p)]$.
2. S is injective, i.e., $S(p) = S(q) \Rightarrow p = q$.
3. (Induction Principle) If $Q \subseteq P$ with

- $p_0 \in Q$, and
- $p \in Q \Rightarrow S(p) \in Q$,

then $Q = P$.

As so often in mathematics, we will usually name a Peano system (P, p_0, S) by just the set P. Of course this is an abuse of notation but is not likely to cause confusion.

The most basic property of Peano systems is the ability to give recursive definitions of functions as in the following theorem.

Theorem A.15 (Recursion Theorem) *Let (P, p_0, S) is any Peano system. If x_0 is an element of a set X, and f is a function from X to X, then there is a unique function $\phi : P \longrightarrow X$ with $\phi(p_0) = x_0$ and $\forall p \in P[\phi(S(p)) = f(\phi(p))]$.*

Proof: Uniqueness: Suppose we have $\phi : P \longrightarrow X$ with $\phi(p_0) = x_0$ and $\forall p \in P[\phi(S(p)) = f(\phi(p))]$ and $\psi : P \longrightarrow X$ with $\psi(p_0) = x_0$ and $\forall p \in P[\psi(S(p)) = f(\psi(p))]$ and consider the subset $Q = \{p \in P : \phi(p) = \psi(p)\}$. Certainly $p_0 \in Q$ as $\phi(p_0) = x_0 = \psi(p_0)$. And if $p \in Q$, then $\phi(S(p)) = f(\phi(p)) = f(\psi(p)) = \psi(S(p))$, i.e., $S(p) \in Q$. So $Q = P$ and $\phi = \psi$.

This type of proof based on appeal to the Induction Principle is called *proof by induction*.

Existence: The proof of existence has two parts – construction ϕ as a relation and then verifying that it is actually a function.

Consider relations F from P to X which have the two properties that $(p_0, x_0) \in F$ and $(p, x) \in F \Rightarrow (S(p), f(x)) \in F$. The relation $P \times X$ is one such, so the set $\mathcal{F}$ of all such relations is non-empty. Define $\phi = \bigcap \mathcal{F}$. Clearly ϕ itself belongs to $\mathcal{F}$, so if ϕ is a function the proof is complete.

So it remains to verify that if both $(p, x) \in \phi$ and $(p, y) \in \phi$, then $x = y$. Define $Q = \{p \in P : \forall x \forall y[(p, x) \in \phi \wedge (p, y) \in \phi \Rightarrow x = y\}$. A proof by induction will show that $Q = P$ and complete the proof of the Recursion Theorem.

First note that $p_0 \in Q$, for if not there is some $y \in X$ different from x_0 with $(p_0, y) \in \phi$. But then (p_0, x_0) is in the relation $\phi - \{(p_0, y)\}$ because $x_0 \neq y$. Also if (p, x) is in $\phi - \{(p_0, y)\}$, then so is $(S(p), f(x))$ for $S(p)$ is guaranteed not to be p_0 and so $(S(p), f(x))$ is certainly not the removed element (p_0, y). This says that $\phi - \{(p_0, y)\}$ is in $\mathcal{F}$ contradicting the definition of ϕ as $\bigcap \mathcal{F}$.

A very similar argument shows that $p \in Q \Rightarrow S(p) \in Q$. For suppose that $p \in Q$, but $S(p) \notin Q$. Then there is a unique element x with $(p, x) \in \phi$, The pair $(S(p), f(x))$ is in ϕ but assuming that $S(p)$ is not in Q there must be an element y, different from $f(x)$, so that $(S(p), y) \in \phi$ as well. So consider the relation $\phi - \{(S(p), y)\}$. As above $(p_0, x_0) \in \phi - \{(S(p), y)\}$ because $S(p) \neq p_0$. Similarly $(q, z) \in \phi - \{(S(p), y)\} \Rightarrow (S(q), f(z)) \in \phi - \{(S(p), y)\}$. To see this consider two possibilities – $q = p$ and $q \neq p$. When $q = p$, it must be that $z = x$ and so $f(z) = f(x) \neq y$. But if $q \neq p$, then $S(q) \neq S(p)$ and so

$(S(q), f(z)) \neq (S(p), y)$. Thus $\phi - \{(S(p), y)\} \in \mathcal{F}$ contradicting the definition of $\phi = \bigcap \mathcal{F}$.

This completes the proof by induction that ϕ is a function, and so the proof of the Recursion Theorem. ∎

It is worth noting that the proof of the recursion theorem uses all three parts of the definition of a Peano system. Indeed the theorem would not be true were any of the conditions removed from the definition.

A first consequence of the Recursion Theorem is that Peano Sets are essential unique.

Theorem A.16 (Uniqueness of Peano Sets) *If (P, p_0, S) and (Q, q_0, T) are two Peano systems, then there is a unique bijection $f : P \longrightarrow Q$ such that $f(p_0) = q_0$ and $\forall p \in P(f(S(p)) = T(f(p)))$.*

Proof: (Uniqueness Theorem) When (P, p_0, S) and (Q, q_0, T) are two Peano sets, the Recursion Theorem applies to guarantee there are functions $f : P \longrightarrow Q$ and $g : Q \longrightarrow P$ satisfying $f(p_0) = q_0$, $g(q_0) = p_0$, $f(S(p)) = T(f(p))$, and $g(T(q)) = S(g(q))$. So then $gf : P \longrightarrow P$ and $fg : Q \longrightarrow Q$ have the properties $gf(p_0) = p_0$, $gf(S(p)) = S(gf(p))$, $fg(q_0) = q_0$, and $fg(T(q)) = T(fg(q))$. But 1_P satisfies same two properties as gf, while 1_Q satisfies the same two properties as fg. As the Recursion Theorem guarantees there is only one function function from P to P and one function from Q to Q having these properties we conclude that $fg = 1_P$ and $gf = 1_Q$. ∎

It is good to know that Peano systems are essentially unique, but this is not very interesting unless at least one exists. That is the basic point of the definition of the set ω in the last section.

Theorem A.17 *The set ω with 0 (i.e., the empty set) as distinguished element and $S(n) = n^+$, is a Peano system.*

Proof: To show that $(\omega, 0, S)$ is a Peano system we need to verify:

1. $\forall n \in \omega,\ 0 \neq n^+$.
2. $\forall n, m \in \omega,\ n^+ = m^+ \Rightarrow n = m$.
3. (Induction Principle) If $Q \subseteq \omega$ with
   - $0 \in Q$, and
   - $n \in Q \Rightarrow n^+ \in Q$,

   then $Q = \omega$.

The first condition is clearly true as $n \in n^+$, and so $n^+ \neq \emptyset$. The Induction Principle is essentially the definition of ω: $Q \subseteq \omega$ and Q is a successor set, while by definition ω is contained in every successor set. Ergo $Q = \omega$.

Curiously only the second condition $(n, m \in \omega(n^+ = m^+ \Rightarrow n = m)$ requires significant work, and it requires a small detour into some of the arcane points of set theory. Most applications of set theory discuss only sets and their elements. Occasionally discussion of families of sets is required, but discussion involving long $\in$-chains, i.e., $a \in b \in c \in \cdots$ just do not arise. By contrast they do occur in the internals of set theory and this is the first illustration. This discussion will recur in the discussion of well-ordering and ordinal numbers. Recall that $0 \in 1 \in 2 \in \cdots$. That relationship is captured in the next definition.

Definition A.39: A set n is **transitive** provided $\forall a \forall b(a \in b \wedge b \in n \Rightarrow a \in n)$. Equivalently n is transitive when $\forall b(b \in n \Rightarrow b \subseteq n)$.

To prove the second condition we will use the following two technical lemmas.

Lemma 1 *Every member of ω is transitive.*

Lemma 2 $\forall n \in \omega(m \subseteq n \Rightarrow m \notin n)$

Assuming the lemmas, suppose that n and m are natural numbers with $n^+ = m^+$. As $n \in n^+ = m^+$, so $n \in m^+$ and thus $n \in m$ or $n = m$. By symmetry we also have $m \in n$ or $n = m$. If $n \neq m$, then $n \in m$ and $m \in n$. The first lemma tells us that n is transitive and so $n \in n$. But as $n \subseteq n$ the second lemma guarantees that $n \notin n$. So we must have $n = m$. $\blacksquare$

Proof of Lemma 1: The proof is by induction. Consider the set S of all transitive elements of ω. Trivially $0 \in S$. If $n \in S$ and $k \in n^+$, then either $k \in n$ or $k = n$. By assumption n is transitive, so if $k \in n$, then $k \subseteq n$ and so $k \subseteq n^+$. And if $k = n$, then certainly $k \subseteq n^+$. So $n^+ \in S$. The Induction Principal now applies to say $S = \omega$ which is the lemma. $\blacksquare$

Proof of Lemma 2: Again the proof is by induction. Let $S = \{n \in \omega : n \subseteq m \Rightarrow m \notin n\}$. Certainly $0 \in S$ as $\forall m(m \notin 0)$. Next assume $n \in S$. Note first that $n \subseteq n$, so our assumption means $n \notin n$. Now consider some set m containing n^+. This means $n \subseteq m$ and $n \in m$. But then $m \notin n$ as n is in S, and $m \neq n$ because $n \notin n$, so $m \notin n^+$. The Induction Principal now applies to say $S = \omega$ thereby completing the proof. $\blacksquare$

The technicalities occurring in these two lemmas and their proofs have very little to do with applications of set theory in other parts of mathematics, but they will recur and be expanded in the development of the theory of ordinal and cardinal numbers later in this appendix.

We now have both existence and uniqueness of a Peano system, but note that even though ω provides a very elegant and natural example of a Peano system it is *not* the only one. Indeed if (P, p_0, S) is a Peano system, then so too is $(P - \{p_0\}, S(p_0), S|_{P-\{p_0\}})$. This corresponds to the observation that proof by induction can start with 1 (or any other integer) rather 0.

Here are some fundamental facts about natural numbers that are easily established directly from the Peano axioms:

- If n is any natural number, then $n \neq n^+$.

- If n is a non-zero natural number, then there is a unique natural number m so that $n = m^+$.
- The set ω of natural numbers is transitive.

Of course the familiar facts about addition and multiplication of natural numbers is still lacking, but those operations can be defined and their properties proved from the Peano axioms as will be done in the next section.

A.11 Arithmetic

The intuition behind the definition of the natural numbers is that of sequentially adding one more item. In that view natural numbers m and n are added by starting with m and counting up by n. Formalizing that is easily done using the Recursion Theorem. For each natural number m there exists a function $s_m : \omega \longrightarrow \omega$ with $s_m(0) = m$ and $s_m(n^+) = s_m(n)^+$.

Definition A.40: $\forall m \in \omega \forall n \in \omega [m + n = s_m(n)]$
Of course $m+n$ is called the **sum** of m and n, and the binary operation $+$ is called **addition**.

Proposition A.27 *For all natural numbers m, n and p,*

(i) $m + 0 = m$.

(ii) $m + 1 = m^+$.

(iii) $0 + n = n$.

(iv) $m^+ + n = (m + n)^+$.

(v) $m + n = n + m$.

(vi) $(m + n) + p = m + (n + p)$.

(vii) $m + n = m + p \Rightarrow n = p$.

(viii) $n \neq 0 \Rightarrow m + n \neq 0$.

(ix) *If $m \neq n$, then $\exists k \in \omega\ [m = n + k]$ or $\exists j \in \omega\ [n = m + j]$.*

Proof:

(i) By definition $m + 0 = s_m(0) = m$.

(ii) $m + 1 = s_m(0^+) = s_m(0)^+ = m^+$.

(iii) While s_0 is the unique function on ω with $s_0(0) = 0$ and $s_0(n^+) = s_0(n)^+$, the identity function 1_ω has the same properties. Thus $s_0 = 1_\omega$ and so $0 + n = s_0(n) = 1_\omega(n) = n$.

(iv) The function s_{m^+} is characterized by the two conditions: $s_{m^+}(0) = m^+$ and $s_{m^+}(n^+) = s_{m^+}(n)^+$. But the function t_m defined by $t_m(n) = s_m(n^+)$ satisfies those same two conditions. Therefore $s_{m^+} = t_m$ and in particular $m^+ + n = s_{m^+}(n) = t_m(n) = s_m(n^+) = s_m(n)^+ = (m+n)^+$.

(v) Consider $S = \{m \in \omega : \forall n \in \omega[m + n = n + m]\}$. Above saw that $0 + n = n = n + 0$, so 0 is an element of S. And if $m \in S$, then

$$\begin{aligned} m^+ + n &= (m+n)^+ && \text{part 2 of this proposition} \\ &= (n+m)^+ && \text{because } m \text{ in } S \\ &= n + m^+ && \text{by definition of addition} \end{aligned}$$

and so $m^+ \in S$ as well. The Induction Principal now applies to tell us that $S = \omega$ and thereby completes the proof of this part.

(vi) The proof that $(m+n)+p = m+(n+p)$ works by induction on p. First both $(m+n)+0$ and $m+(n+0)$ are equal to $m+n$ by the definition of addition. For the induction step suppose that $(m+n)+p = m+(n+p)$ for all m and n. Then

$$\begin{aligned} (m+n)+p^+ &= ((m+n)+p)^+ && \text{by definition of addition} \\ &= (m+(n+p))^+ && \text{by induction hypothesis} \\ &= m+(n+p)^+ && \text{by definition of addition} \\ &= m+(n+p^+) && \text{by definition of addition} \end{aligned}$$

and so the proof by induction is complete.

(vii) The proof that $m+n = m+p \Rightarrow n = p$ is by induction on m. The assertion is certainly true for $m = 0$ as $0+n = n$. So suppose that $m^+ + n = m^+ + p$. Then as $m^+ + n = m + n^+$, it follows that $m + n^+ = m + p^+$. The induction hypothesis gives $n^+ = p^+$ and $n = p$ follows from the fourth Peano Axiom and the proof by induction is complete.

(viii) This is also proved by induction on m. Certainly $n \neq 0 \Rightarrow 0 + n \neq 0$. Note that $m^+ + n = m + n^+$ and n^+ is always non-zero, so the induction hypothesis give $m + n^+ \neq 0$, so the proof by induction is complete.

(ix) Fix an arbitrary natural number n and consider the set

$$S = \{m \in \omega : \exists k \in \omega\ m = n + k \vee \exists j \in \omega\ n = m + j \vee \}$$

Certainly $0 \in S$, so now for the induction step suppose that $m \in S$. To see that m^+ is also in S the two cases:

(a) $m = n + k$ – then $m^+ = n + k^+$. Note this includes the case that $k = 0$ which is also the case $n = m + 0$.

(b) $n = m + j$ with $j \neq 0$ – then there exists $p \in \omega$ with $j = p^+$, and so $n = m + j = m + p^+ = m^+ + p$.

So $m^+ \in S$ and so $S = \omega$.

∎

At the same primitive level, multiplication of natural numbers can be considered as repeated addition which can again be formalized using the Recursion Theorem: for each natural number there exists a m unique function $p_m : \omega \longrightarrow \omega$ with $p_m(0) = 0$ and $p_m(n^+) = p_m(n) + m$.

Definition A.41: $\forall m \in \omega \forall n \in \omega [m \cdot n = p_m(n)]$

The natural number $m \cdot n$ is called the **product** of m and n, and the binary operation $\cdot$ is call **multiplication**.

The $\cdot$ is often omitted with multiplication being indicated by juxtaposition if no confusion will result. Other notation sometimes used in place of the $\cdot$ to indicate multiplication includes $\times$ and $*$.

Proposition A.28 *For all natural numbers m, n and p,*

(i) $m \cdot 0 = 0$.

(ii) $0 \cdot n = 0$.

(iii) $m \cdot 1 = m$.

(iv) $(m \cdot n) \cdot p = m \cdot (n \cdot p)$.

(v) $m \cdot n = n \cdot m$.

(vi) $m \cdot (n + p) = (m \cdot n) + (m \cdot p)$.

(vii) *if* $m \neq 1$, *then* $m \cdot n = m \cdot p \Rightarrow n = p$.

(viii) *if* $n \neq 1$, *then* $m \cdot n \neq 1$.

The proof of this proposition is very similar to that of the preceding proposition and is omitted. ∎

Finally the same process can be applied to exponents which arise from repeated multiplication. Using the Recursion Theorem, for each natural number m there exists a function $e_m : \omega \longrightarrow \omega$ with $e_m(0) = 1$ and $e_m(n^+) = e_m(n) \cdot m$.

Definition A.42: $\forall m \in \omega \forall n \in \omega [m^n = e_m(n)]$

The natural number m^n is called the **n-th power of m**, and the operation is called **exponentiation**.

Proposition A.29 *For all natural numbers m, n and p,*

(i) $m^0 = 1$.

(ii) $0^n = 0$ *except that* $0^0 = 1$.

(iii) $m^1 = m$.

(iv) $(m^n)^p = m^{(n^p)}$.

(v) $m^{(n+p)} = (m^n) \cdot (m^p)$.

(vi) $m^n = m^p \Rightarrow n = p$ *except when* m^n *is equal to 0 or 1.*

(vii) $m^n = 1 \Rightarrow m = 1 \vee n = 0$.

The proof of this proposition is very similar to that of the preceding two propositions and is omitted. ∎

A.12 Order

The natural arithmetic operations (addition and multiplication) on the set ω of natural numbers were developed in the previous section, but there is another familiar and important aspect of the natural numbers that hasn't yet been mentioned: order. The order relation in ω is the familiar succession $0 < 1 < 2 < \cdots$. The first step is to make the order relation precise, and that is based on the understanding that m is larger than n when m is gotten by from n by successive increments, which is the same as adding some natural number to n.

Definition A.43: For all $n, m \in \omega$, $n \leq m$ iff there is some natural number k so that $m = n + k$.

In particular every natural number is equal to or greater than 0.

Proposition A.30 *For all* m*,* n *and* p *in* ω*:*

(i) (Reflexive) $m \leq m$.

(ii) (Antisymmetric) $m \leq n \wedge n \leq m \Rightarrow m = n$.

(iii) (Transitive) $m \leq n \wedge n \leq p \Rightarrow m \leq p$.

(iv) (Connected) Either $m \leq n$ *or* $n \leq m$.

Proof: The proof is based on Proposition A.27. (i) is because $m + 0 = m$. For (ii) note that if $m = n + k$ and $n = m + l$, then $m + 0 = m = (m + l) + k = m + (l + k)$ and so l + k = 0 and so l = k = 0, i.e., m = n. (iii) comes from noting that if $n = m + k$ and $p = n + l$, then $p = (m + k) + l = m + (k + l)$. (iv) is just a restatement of the last part of Proposition A.27. ∎

Order relations are both generally important and specifically important in the study of infinite sets in set theory. The starting point here are some general definitions and simple observations.

Definition A.44: A **preorder** on X is a relation R on X which is reflexive ($\forall x \in X,\ xRx$), and transitive ($\forall x, y, z \in X,\ xRy \wedge yRz \Rightarrow xRz$).

Clearly the order relation on the natural numbers defined above is a preorder.

Every preorder defines an equivalence relation: $x \equiv_R y \iff (xRy \wedge yRx)$. [Proof: $x \equiv_R x$ because the preorder R is reflexive; $x \equiv_R y \Rightarrow y \equiv_R x$ by the symmetry in the definition of $\equiv_R$; and $\equiv_R$ is transitive because R is transitive. $\blacksquare$]

The preorder also defines a preorder on the quotient set $X/\equiv_R$ in the expected way: $x/\equiv_R\ Ry/\equiv_R$ iff xRy. [Proof: If $x \equiv_R x'$ and $y \equiv_R y'$, and xRy, then $xRy \wedge yRy' \Rightarrow xRy'$, while $x'Rx \wedge xRy' \Rightarrow x'Ry'$. $\blacksquare$]

This preorder on the quotient set now is *antisymmetric*, i.e., xRy and yRx implies $x = y$. An antisymmetric preorder is actually the usual basic object of study in order theory and is formalized in the following definition.

Definition A.45: A **partial order** on X is a relation on X which is reflexive, transitive and antisymmetric.

Commonly a partial order is denoted by $\leq$ or something reminiscent such as $\preceq$, $\subseteq$, or $\sqsubseteq$.

Proposition A.30 shows that the order relation on the natural numbers is not just a pre-order, but a partial order.

Associated to a partial order $\leq$ on X are several other closely related relations on X just as is familiar for order relation(s) on various types of numbers. For x and y in X, $y \geq x$ is the same as $x \leq y$ (i.e., $\geq$ is the inverse relation of $\leq$,) The relation $\geq$ is also a (usually different) partial order on X.

For x and y in X, $x < y$ means $x \leq y$ and $x \neq y$, and $y > x$ is the same as $x < y$. When $x < y$ we say x is *less* than y, or *smaller* than y or *precedes* y. Similarly when $y > x$ we say that y is *greater* than y, or *larger* than y or *succeeds* or *follows* y.

The relation $<$ is *not* a partial order, but gets its own definition.

Definition A.46: A **strict order** on X is a relation on X which is transitive, and anti-reflexive.

Commonly a strict order is denoted by $<$ or something reminiscent such as $\prec$, $\subset$, or $\sqsubset$.

Defining $x < y$ to mean $x \leq y$ and $x \neq y$ associates to a partial order $\leq$ a strict order $<$. Conversely starting with a strict order $<$ and defining $x \leq y$ to mean $x < y$ or $x = y$ produces a partial order.

Definition A.47: A **partially ordered set** or, more commonly, a **poset** is a set together with a partial order on it.

Note that this is really just another name for a partial order (with domain a set) because a partial order determines the set on which it is a partial order! Despite this it is nearly universal to use the name of the domain of the partial order to refer to the poset. The restriction to a partially ordered set rather than a partially ordered *class* is inessential but customary. The trivial extension of what follows is left to the interested reader.

As discussed above, every poset has associated with it two partial orders and two strict orders. This relationship is so intimate and so familiar that we will freely move among the four relations with no comment at all.

The generic example of a partially ordered set is a power set $\mathcal{P}(x)$ with the subset relation $\subseteq$ as the partial ordering. Indeed any subset y of $\mathcal{P}(x)$ is a poset under the subset relation, and all posets are like that. Here is what that means.

Definition A.48: If $(a, \leq)$ and $(b, \preceq)$ are posets, then a **monotone** function from a to b is a function $f : a \longrightarrow b$ such that $a_1 \leq a_2$ implies $f(a_1) \preceq f(a_2)$.

Monotone functions are also called *order preserving* or *isotone*.

Definition A.49: An **isomorphism** between $(a, \leq)$ and $(b, \preceq)$ is a monotone function from a to b with a monotone inverse $f^{-1} : b \longrightarrow a$.

Proposition A.31 *If $(a, \leq)$ is a poset, then there is a subset b of $\mathcal{P}(a)$ with $(a, \leq)$ isomorphic to $(b, \subseteq)$.*

Proof: Associate to each element x of a the subset $f(x) = \{y \in a : y \leq x\}$ of a, and take $b = \{f(x) : x \in a\}$. This actually defines f as the desired isomorphism. ∎.

There are various special properties of partial orders that are of particular interest.

Definition A.50: A connected partial order, i.e., for every x and y in X either $x \leq y$ or $y \leq x$, is called a **total order**. (This is also called a *linear order* or a *simple order*.)

Definition A.51: A **totally ordered set** is a partially ordered set where the partial order is a total order. A totally ordered set is often called a **chain**.

The full content of Proposition A.30 is that the order relation on the set of natural numbers is a total order.

Examples: Every equivalence relation is a partial order. Every pair of distinct elements is incomparable. The only relation that is both a partial order and an equivalence relation is equality. The Hasse diagram of an equivalence relation is a discrete graph, i.e., there are no edges.

Special Elements.

Definition A.52: A **least element**, **minimum element**, or **first element** in a partially ordered set $(X, \leq)$ is an element $x_0 \in X$ such that $\forall x \in X\ (x_0 \leq x)$.

Definition A.53: A **greatest element**, **maximum element** or **last element** in a partially ordered set $(X, \leq)$ is an element $x_1 \in X$ such that $\forall x \in X \ (x \leq x_1)$.

A greatest element $(X, \leq)$ is a least element in $(X, \geq)$, and a least element in $(X, \leq)$ is a greatest element in $(X, \geq)$. This is our first example of *duality* for order.

Least and greatest elements may fail to exist. For instance with our usual understanding of the ordering of the natural numbers, 0 is a least element while there is no greatest element. But when they do exist least and greatest elements are clearly unique.

There are also the closely related but distinct notions of minimal and maximal elements.

Definition A.54: A **minimal element** in a partially ordered set $(X, \leq)$ is an element $x_0 \in X$ such that $x \leq x_0 \Rightarrow x = x_0$.

If x_0 is a least element in X, then it is certainly a minimal element in X, but the converse is not true. Just as with least elements, minimal elements may not exist in a particular poset. But, by contrast with least elements, minimal elements may not be unique even when they do exist.

The dual notion is that of maximal elements which are the minimal elements for the poset $(X, \geq)$. Here is the direct definition.

Definition A.55: A **maximal element** in a partially ordered set $(X, \leq)$ is an element $x_0 \in X$ such that $x \geq x_0 \Rightarrow x = x_0$.

When $(X, \leq)$ is a poset and Y is a subset of X we define the relation $\leq|_Y$ on Y to be $\leq \cap Y \times Y$. Usually we will just speak of the inherited order on Y and use the same symbol for the partial order on Y as for the partial order on X.

There are several special elements associated to subsets of a poset.

Definition A.56: An **upper bound** for a subset Y of a poset X is an element $x \in X$ with $\forall y \in Y \ y \leq x$

Dually we have lower bounds.

Definition A.57: A **lower bound** for a subset Y of a poset X is an element $x \in X$ with $\forall y \in Y \ y \geq x$

Of special interest are least elements among upper bounds and greatest elements among lower bounds, so we have the following definitions.

Definition A.58: Associated to each subset Y of the poset X we have the subset U consisting of all upper bounds of Y. A least element of U is called a **least upper bound** of Y. As with least elements in general, this may not exist, but is unique when it does exist. A least upper bound is also commonly called a *supremum* or a *lub*, and denoted $\mathbf{lub}(Y)$ or $\bigvee(Y)$.

Definition A.59: Associated to each subset Y of the poset X we have the subset L consisting of all lower bounds of Y. A greatest element of L is called a **greatest lower bound** of Y. As with greatest elements in general, this may not exist, but is unique when it does exist. A greatest lower bound is also commonly called an *infimum* or a *glb*, and denoted $\mathbf{glb}(Y)$ or $\bigwedge(Y)$.

The order relations defined on the natural numbers were all defined using the properties of the set of natural numbers as a Peano system. This process can also be reversed and the Peano system structure defined in terms of the order relation. The first part is simply that 0 is the least element of the ordered set ω. The harder part is the definition of the successor function and proof of the induction principal. That will be done in Section A.16. Beyond that there is an alternative perspective on the order relation that is more directly tied to the details on how ω and its elements are defined that will be considered there as well.

Duality in Order

Several times above we have remarked on pairs of dual concepts in a posets. Examples of dual pairs include "least" and "greatest", "minimal" and "maximal", "lower bound" and "upper bound", and "lub" and "glb".

Every order theoretic definition has its dual, obtained by applying the given definition to the inverse relation. Equally every theorem in order theory also has a dual theorem which does not require a separate proof.

The connection between posets and categories (see Section B.19.4) exposes duality in order as yet another example of duality of categories.

A.13 Number Systems

There are several number systems besides the natural numbers that are commonly considered: integers, rational numbers, real numbers and complex numbers. All of these can be defined, there existence demonstrated and their properties proved within axiomatic set theory. One valuable classic on the topic is Landau's *Foundations of analysis; the arithmetic of whole, rational, irrational, and complex numbers* [39], while two more recent treatments are *The Structure of the Real Number System* [12] by Cohen and Ehrlich and *The Number Systems: Foundations of Algebra and Analysis* [23] by Feferman.

The discussion of number systems is not very close to the theory of sets, but this is a convenient point to provide a précis of the material for reference in the remainder of the notes.

From Section B.2.3 of the Catalog we recall that a *monoid* consists of a set M and an associative binary operation on M that has an identity. If the binary operation is commutative, then M is a *commutative monoid*

Proposition A.32 *The set of natural numbers with the binary operation of addition is a commutative monoid with 0 as the identity.*

This is just another way of stating parts of Proposition A.27. ∎

Proposition A.33 *The set of natural numbers with the binary operation of multiplication is a commutative monoid with 1 as the identity.*

This is just another way of starting parts of Proposition A.28. ∎

As there are two natural and common ways in which the set of natural numbers are a monoid it is sometimes necessary to distinguish them. The set of natural numbers together with the binary operation of addition is called the *additive monoid of natural numbers*, while when considering multiplication we speak of the *multiplicative monoid of natural numbers*.

Proposition A.34 *The additive monoid of natural numbers is a free monoid generated by 1.*

Proof: This means that if $(M, \cdot, 1)$ is any monoid and m_0 is an arbitrary element of M, there there is a *unique* monoid homomorphism $h : \omega \longrightarrow M$ with $h(1) = m_0$.

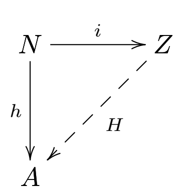
As preparation to using the Recursion Theorem, define a function $f : M \longrightarrow M$ by $f(m) = m_0 \cdot m$. Then there is a *unique* function $h : \omega \longrightarrow M$ so that $h(0) = 1$ and $h(S(n)) = m_0 \cdot h(n)$. Then $h(1) = h(S(0)) = m_0 \cdot 1 = m_0$.

Next we use induction to prove that h is a homomorphism. Consider the set $Q = \{n \in \omega : \forall n' \in \omega\ [h(n+n') = h(n) \cdot h(n')]\}$. Certainly $0 \in Q$ as $h(0+n') = h(n') = h(0) \cdot h(n')$ for $h(0) = 1$. And if $n \in Q$, then $h(S(n)+n') = h(S(n+n')) = m_0 \cdot h(n+n') = m_0 \cdot h(n) \cdot h(n') = h(S(n)) \cdot h(n')$, so $S(n) \in Q$. By the Induction Principle $Q = \omega$ and so h is a homomorphism. ∎

Proposition A.35 *The multiplicative monoid of positive natural numbers is a free monoid generated by the prime numbers.*

The Abelian group of integers is constructed from the monoid of natural numbers in a fashion that applies in the following generality.

Proposition A.36 *For every commutative monoid N there is an Abelian group Z and a monoid homomorphism $i : N \longrightarrow Z$ so that for every monoid homomorphism $h : M \longrightarrow A$ of M to an Abelian group A there is a* unique *group homomorphism $H : Z \longrightarrow M$ so that $h = Hi$.*



The pair (Z, i) is called Abelian group completion of the commutative monoid N.

For the proof see Proposition B.39.

Of course this is an example of a universal mapping property and as in all such cases the Abelian group is unique up to isomorphism.

Definition A.60: The Abelian group of integers, $\mathbb{Z}$, is the Abelian group completion of the commutative monoid of natural numbers.

This "definition" has a problem. What exactly is the *set* $\mathbb{Z}$? In particular is the set ω of natural numbers a subset of $\mathbb{Z}$? Common mathematical usage certainly assumes that every natural number is an integer! But it is also common in the development of number systems from the axioms of set theory or even from the Peano axioms to identify the set $\mathbb{Z}$ as a set of equivalence classes of pair of natural numbers as in the proof of Proposition B.39 so that ω is not a subset of $\mathbb{Z}$. This is a minor but niggling annoyance which can be addressed in several different ways.

One solution is to change the definition of the natural numbers or the integers so that the set of natural numbers *is* a subset of the set of integers. Again there are several possible ways of doing that with perhaps the simplest being to define the integers by just adjoining suitable negative integers to ω as is the common notation. Verifying that this works is a bit of a nuisance, but it certainly works.

Another solution is to provide an axiomatic characterization of the integers (for example as a certain type of ordered integral domain,) and show that the system of integers must have a compatible Peano system as a subsystem (which is, of course, isomorphic to the Peano system of natural numbers.) Then arbitrarily fix one such as "the" integers and "the" natural numbers.

Still another approach is to incorporate additional axioms that define all of the usual number systems (natural numbers, integers, and rational, real and complex numbers) and ensure the expected inclusion relations. This involves changing the approach used here so that there are elements (commonly called *urelements*) besides those that are sets. Then the constructions here and in Section A.11 serve to show that the more complex axiom systems that define the number systems are equiconsistent with the limited set theory axioms in this Appendix.

Introducing urelements seems to refute the remark that "for the purposes of mathematics it seems to be enough to have just sets that can be built from the empty set" (page 182) but the use or lack of other elements is entirely a matter of convenience. (It is left as an exercise for the reader to decide where the convenience lies!)

As remarked in Section A.6 the same thing is equally well true for ordered pairs. Rather than using the abstruse definition given in these notes it is entirely feasible to introduce them as independent entities having their characteristic property, and then allow ordered pairs of element as elements as well. In the same vein it is also possible to introduce functions as new objects subject to suitable axioms and then allow them to be elements of sets as well. The reason this is seldom done is because of the complexity associated with having many different logical entities and the large complicated collection of axioms needed to specify this complexity.

In this Appendix we will follow the same general approach we have used throughout the notes of being somewhat cavalier in accepting these number systems without completely specifying exactly the set at issue.

At this stage the set of integers as a mathematical object is defined as being a certain Abelian group, while the set of integers so familiar to us all also has multiplication and an order relation extending the same features of the natural numbers.

A.14 Axiom of Choice

Axiom of Choice: The Cartesian product of a non-empty family of non-empty sets is non-empty.

A.15 Zorn's Lemma

A.16 Well Ordered Sets

Definition A.61: A **well ordered set** is a poset in which every non-empty subset has a least element.

Definition A.62: For a well ordered set X and x an element of X, the **segment** of X determined by x is the set $\{y \in X : y < x\}$

Theorem A.18 *Let $(X, \leq)$ be a well ordered set. Suppose that S is a subset of X such that:*

1. *the smallest element of X is an element of S, and*
2. $\forall x \in X \; [\forall y \in X \; (y < x \Rightarrow y \in S) \Rightarrow x \in S]$

Then $S = X$.

A.17 Transfinite Induction

A.18 Axiom of Substitution

Axiom of Substitution: If f is a function and the domain of f is a set, then the image of f is a set.

A.19 Ordinal Numbers and Their Arithmetic

A.20 The Schröder-Bernstein Theorem

A.21 Cardinal Numbers and Their Arithmetic

A.22 Axiom of Regularity

Axiom of Regularity:

Appendix B

Catalog of Categories

The purpose of this appendix is to catalog a variety of categories, together with interesting properties and interrelations. It comes here at the end exactly because we feel free to discuss all of the notions that occur earlier in these notes. But the intention is that each section can be read only part way for information relevant to early parts of the notes. So in particular the properties of these categories are discussed in approximately the same order as those properties are discussed in the notes.

There are a substantial variety of categories discussed here, and it is unlikely that all of them will be familiar to any reader of these notes. For just that reason each section is intended to be largely independent of the other sections. Moreover we provide references for readers who want more information about the particular categories. The references are not usually about the categories per se, but usually to sources where the objects and morphisms that constitute the category and discussed.

Each of the "big three" topics in Mathematics: Algebra, Topology and Analysis contribute a large number of categories that are familiar, by which we mean widely and commonly studied. Adding to those the categories associated to the foundational topics of set theory, order and logic provides the source of most of the categories listed in this chapter. That also is the organizing principle, at least up to section B.16.

While the study of categories is fundamentally about morphisms, the familiar categories are named after their objects. This is a reflection that the first level of abstraction was such objects as groups, rings, topological spaces, etc. The study of the morphisms between suitable objects came later. Even when morphisms began to be studied, the starting point was objects as "sets with structure" and morphisms were just certain functions that "respected" the structure.

In most of the categories described below the identity function on (the underlying set of) an object is a clearly a morphism, and the function composition of two morphisms is again a morphism and with these observations it is a trivial formality to verify that the proposed category is indeed a category. So in most cases the verification will be omitted without comment.

B.1 Sets

B.1.1 Set − sets

The most fundamental category of all is the category of functions between sets. Because this is so fundamental and because it is the first example the definition will be given in pedantic detail. Such detail will be omitted in other examples.

Definition B.1: The objects of the category **Set** are all sets. The morphisms of **Set** are triples (A, f, B) where A and B are sets and f is a function from A to B (see definition A.30) with domain A and range a subset of B. The domain of (A, f, B) is A, while the codomain is B. The identity morphism associated to a set A is the morphism $(A, 1_A, A)$ where 1_A is the identity function from A to itself. Finally the composition of $f : A \longrightarrow B$ with $g : B \longrightarrow C$ is (A, gf, C) where gf is the composition of the two functions.

To confirm that this does in fact define a category we need to check that composition of morphisms is associative:

$$A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D \Rightarrow h(gf) = (hg)f$$

and the defining properties of the identity morphisms:

$$A \xrightarrow{1_A} \xrightarrow{f} B \xrightarrow{1_B} B \Rightarrow f1_A = f \wedge 1_B f = f$$

both of which are follow from the corresponding facts about functions.

Already here we are following the nearly universal mathematical custom of naming things using a small and ambiguous part of their full name. Here it is naming a morphism such as (A, f, B) by just the function f. Notice that the function f actually does determine its domain A, but not its codomain B. The same function f is part of many different morphisms with different codomains. And from here on the full detail (A, f, B) will be written as $f : A \longrightarrow B$ or $A \xrightarrow{f} B$.

An outline of the information about sets that is used in this section and throughout these notes is collected in Appendix A (Set Theory). The Appendix is not about the category of sets, but what is needed about set theory in these notes is established there.

In **Set** the epimorphisms are exactly the surjective functions, or, more precisely, those morphisms where the range of the function is the codomain of the morphism. The monomorphisms are the injective functions, and the isomorphisms are the bijective functions. In this category every epimorphism has a section, while every monomorphism has a retract. As a consequence (cf. exercise I.8) every morphism that is both monic and epic is an isomorphism.

Set has an initial object, the empty set with the unique function ! from the initial object to any other object being the empty function. **Set** has many final objects, namely all of the singleton sets with !, the unique function from any object to the final object, being the function that sends every element to the single element.

Set has products and sums. The product of two sets X and Y "is" the usual Cartesian product $X \times Y$ (see definition A.15) consisting of ordered pairs of elements from the two sets, with $\pi_1(x, y) = x$ and $\pi_2(x, y) = y$. Of course as is true is any category, if there is an isomorphism (i.e., a bijection in this context) from Z to $X \times Y$ then composing the bijection with π_1 and π_2 gives two projections morphisms from Z to X and Y which makes Z a product of the two sets as well.

The sum of sets is the less familiar disjoint union as discussed on page 35.

By contrast, Lawvere and Rosebrugh (*Sets for Mathematics*) [46] give an informal axiomatic treatment of the category of sets which does not depend on axiomatic set theory as treated in the appendix.

Barr [2] provides a fine polemic explaining why categorical set theory is preferable to traditional axiomatic set theory.

B.1.2 FiniteSet – finite sets

The category **FiniteSet** of finite sets is the full subcategory of **Set** whose objects are the finite sets.

Just as in **Set** the epimorphisms in **FiniteSet** are exactly the surjective functions, or, more precisely, those morphisms where the range of the function is the codomain of the morphism. The monomorphisms are the injective functions, and the isomorphisms are the bijective functions. In this category every epimorphism has a section, while every monomorphism has a retract. As a consequence (cf. exercise I.8) every bimorphism, i.e., one that is both monic and epic, is an isomorphism.

The unique initial object in **Set**, namely the empty set, is also the unique initial object in **FiniteSet**. And all of the many final objects in **Set**, namely all of the singleton sets, are equally well the final objects in **FiniteSet**.

Also the products and sums in **Set** are equally well the products and sums in **FiniteSet**, for the product and sum of two finite sets are also finite sets.

B.1.3 Rel – the category of sets and relations

The objects of **Rel** are sets, while the morphisms are relations between sets. (See Section A.7 for the relevant definitions.)

B.1.4 RefRel – the category of sets and reflexive relations

B.1.5 SymRel – the category of sets and symmetric relations

B.1.6 PSet – the category of sets and partial functions

Definition B.2: A **partial function** F from a set A to a set B is a function $f : S \longrightarrow B$ with S a subset of A. S is called the **domain of definition** or **support** of F. When $S = A$, the partial function F is called a **total function**.

We will usually speak of a partial function f, confusing the partial function with the function whose domain is a subset of A. If more detail is needed we will write (f, S) or even (f, S): $A \longrightarrow B$.

Of course for any set A we have the total function $(1_A, A)$. And if $(f, S) : A \longrightarrow B$ and $(g, T) : B \longrightarrow C$ are partial functions, then we take the composition $(g, T)(f, S) : A \longrightarrow C$ to be the partial function with domain of definition $S \cap g^{-1}(T) \subseteq A$ and function

$$gf|_{S\cap g^{-1}(T)}$$

B.1.7 $\mathbf{Set}_*$ – the category of pointed sets

The objects in the category $\mathbf{Set}_*$are pointed sets (cf. definition I.43), and the morphisms are morphisms of pointed sets. As noted above, $\mathbf{Set}_*$ has epimorphisms, namely the surjective functions, while monomorphisms are the injective functions, and the isomorphisms are the bijective functions. In this category every epimorphism has a section, while every monomorphism has a retract. Basically the arguments for **Set** carry over with little change.

Again, $\mathbf{Set}_*$ has initial objects, final objects and zero objects: any one element set with the single element being (necessarily) the base point.

Also $\mathbf{Set}_*$ has both products and sums. The product of two pointed sets, $\langle X, x_0\rangle$ and $\langle Y, y_0\rangle$ "is" $\langle X \times Y, \langle x_0, y_0\rangle\rangle$ with the usual projection maps, while the sum of two pointed sets is the join defined on 38.

B.1.8 Ord – the category of ordinal numbers

B.2 Semigroups, Monoids, Groups and Their Friends

B.2.1 Magma – the category of magmas

Actually this is not likely to be a familiar category, but it is included here partly because it is the base for so many more familiar categories, but mainly because it is a useful notion when we discuss algebraic objects in categories (cf. Section I.4). A detailed treatment of magmas is to be found at the very beginning of Bourbaki's Algebra [10].

Recall the definitions from I.53 and I.54.

Definition: A **magma**is a set, M, together with a **binary operation** or **law of composition**, $\mu : M \times M \longrightarrow M$.

Most commonly the binary operation in a magma is written as $\mu(m, n) = mn$ though in particular examples the operation may be written in some quite different fashion.

We make **NO** assumptions about the operation – it may not be associative, commutative, nor have any sort of identities. All of those result in other, often more familiar objects.

Definition: A **magma homomorphism**, or **morphism of magmas** is a function $f : M \longrightarrow N$ such that $f(xy) = f(x)f(y)$. Notice that this is equivalent to saying that $\mu(f \times f) = f\mu$.

All of the categories discussed in this section are subcategories of **Magma**. The selected objects that define a particular subcategory typically are those which satisfy additional laws such as the associative, commutative or identity laws.

B.2.2 Semigroup – the category of semigroups

Definition B.3: A **semigroup** is an associative magma, i.e., one in which the binary operation satisfies the associative law: $(ab)c = a(bc)$.

The category **Semigroup** is the full subcategory of **Magma** with objects the semigroups. In particular a semigroup homomorphism is just a homomorphism of the magmas.

Definition B.4: A **commutative semigroup** or **Abelian semigroup** is a semigroup in which the binary operation is commutative: $ab = ba$.

Of course there is the corresponding full subcategory of **Semigroup** with objects the commutative semigroups. This subcategory will be denoted by **ASemigroup**.

Let S be an Abelian semigroup. The Grothendieck group of S is $K(S) = S \times S/\sim$, where $\sim$ is the equivalence relation: $(s, t) \sim (u, v)$ if there exists $r \in S$ such that $s + v + r = t + u + r$. This is indeed an Abelian group with zero element (s, s) (any $s \in S$) and inverse $-(s, t) = (t, s)$. It is common to use the suggestive notation $t - s$ for (t, s).

B.2.3 Monoid – monoids

Definition B.5: A **monoid** is a binary operation that is associative and has an identity element.

In more detail a monoid $(M, \cdot, 1)$ consists of a set M, a binary operation $\cdot$ on that set, and a distinguished element $1 \in M$ satisfying the Associative Law: $\forall m, m', m'' \in M\ [(m \cdot m') \cdot m'' = m \cdot (m' \cdot m'')]$; and the Identity Law: $\forall m \in M,\ m \cdot 1 = m = 1 \cdot m$.

As usual a monoid $(M, \cdot, 1)$ is typically named only by the set M, and no special name is given for the binary operation. Indeed, as in much of these notes, the effect of the binary operation in a monoid is often shown by juxtaposition.

Definition B.6: A **monoid homomorphism** is a magma homomorphism that carries the identity to the identity, i.e., a monoid homomorphism $f : M \longrightarrow N$ is a function between the sets satisfying:

- $\forall m, m' \in M,\ f(m \cdot m') = f(m) \cdot f(m')$.
- $f(1) = 1$.

Definition B.7: The category **Monoid** has as objects all monoids and as morphisms all monoid homomorphisms.

In **Monoid** as with all concrete categories the surjective morphisms are epimorphisms, but there is no good characterization of all epimorphisms. The example of the inclusion of the additive monoid of natural numbers into the additive monoid of integers is an epimorphism that is not surjective.

The monomorphisms are the injective homomorphisms, and the isomorphisms are the bijective homomorphisms.

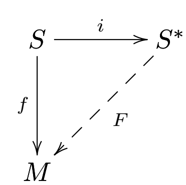
The category **Monoid** has initial, final and zero objects, all being the single element monoids consisting of just the identity element.

This category also has products. The product of $(M, \cdot, 1)$ and $(N, \cdot, 1)$ is $(M \times N, \cdot, (1, 1))$ where $(m, n) \cdot (m', n') = (m \cdot m', n \cdot n')$, and the projection homomorphisms are just the set projection maps.

Verification that $(M \times N, \cdot, (1, 1))$ is indeed a monoid, that the projection functions are homomorphism, and that the universal mapping property for the product in **Monoid** does hold are all straight forward.

Just as the sum of sets is much less familiar and more complicated than the product of sets, so it is with the sum of monoids. It is simplest to show that the sum of monoids exists after a digression on free monoids.

Proposition B.37 *Every set generates a free monoid. In detail this means that if S is any set there is a monoid S^* and a function $i : S \longrightarrow S^*$ so that for every function $f : S \longrightarrow M$ of S to a monoid M there is a* unique *monoid homomorphism $F : S^* \longrightarrow M$ so that $f = Fi$.*



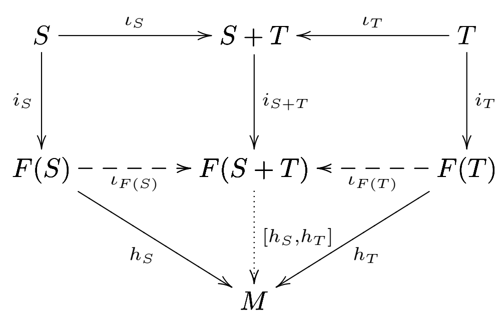
Proof: See Section III.2.12. ∎

Proposition B.38 $F(S+T) = F(S) + F(T)$.

The details of just what this means are in the following proof.

Proof: We have the injections $\iota_S : S \longrightarrow S+T$ and $\iota_T : T \longrightarrow S+T$, and the canonical function $i_{S+T} : S+T \longrightarrow F(S+T)$, which give us the functions $i_{S+T}\iota_S$ and $i_{S+T}\iota_T$ which in turn induce monoid homomorphisms from $F(S)$ and $F(T)$ to $F(S+T)$ which we will call $\iota_{F(S)}$ and $\iota_{F(T)}$ respectively. All these morphisms, as well as all the rest discussed in this proof, are shown in the diagram below.

The detailed claim is that $F(S+T)$ with the monoid homomorphisms $\iota_{F(S)}$ and $\iota_{F(T)}$ is a sum of $F(S)$ and $F(T)$. To verify this, suppose $h_S : F(S) \longrightarrow M$ and $h_T : F(T) \longrightarrow M$ are any two monoid homomorphisms. Then we have the function $[h_S i_S, h_T i_T] : S+T \longrightarrow M$ which in turn induces a unique monoid homomorphism $[h_S, h_T] : F(S+T) \longrightarrow M$ with $[h_s, h_T]i_{S+T} = [h_S i_S, h_T i_T]$ But then $h_S i_S = [h_S, h_T]\iota_{F(S)} i_S$ and $h_T i_T = [h_S, h_T]\iota_{F(T)} i_T$. From the first it follows that $[h_S, h_T]\iota_{F(S)} = h_S$ and from the second that $[h_S, h_T]\iota_{F(T)} = h_T$, and that these are the unique such morphisms. ∎



Definition B.8: An element p in a monoid M is **right cancelable** provided $mp = np \Rightarrow m = n$. It is **left cancelable** provided $pm = pn \Rightarrow m = n$, and it is **cancelable** provided it is right and left cancelable.

An element p is right cancelable iff it is an epimorphism in M considered as a category with one object, and p is left cancelable iff it is a monomorphism in that category.

Definition B.9: An element p in a monoid M is **right invertible** provided there exists an element m such that $mp = 1$. It is **left invertible** provided there exists an element n such that $pn = 1$, and it is **invertible** provided it is right and left invertible.

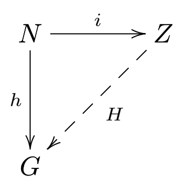
Considering M as a category, a right invertible element is one that has a section or, equivalently, is a retract, while a left invertible element is one that has a retract or, equivalently, is a section. An invertible element is an isomorphism.

Suppose have monoid M and monoid homomorphism $h : M \longrightarrow U$. Let $S = \{s \in M : h(s) \text{ is invertible }\}$. Then S is a submonoid.

Theorem B.19 *For any subset S of a monoid M there is a monoid $S^{-1}M$ and a monoid homomorphism $h : M \longrightarrow S^{-1}M$ with the following Universal Mapping Property:*

1. *$\forall s \in S$, $h(s)$ is invertible in $S^{-1}M$.*
2. *If $f : M \longrightarrow N$ is a monoid homomorphism with $f(s)$ invertible for all $s \in S$, then there exists a* unique *monoid homomorphism $\overline{f} : S^{-1}M \longrightarrow N$ with $f = \overline{f}h$.*

Corollary 2 *For every monoid N there is a group Z and a monoid homomorphism $i : N \longrightarrow Z$ with the universal mapping property that for every monoid homomorphism $h : N \longrightarrow G$ of N to a group G there is a* unique *group homomorphism $H : Z \longrightarrow G$ so that $h = Hi$.*



B.2.4 CMonoid – commutative monoids

Definition B.10: A **commutative monoid** is a monoid in which the operation is commutative.

Definition B.11: **CMonoid** is the full subcategory of **Monoid** whose objects are the commutative monoids.

In **CMonoid** the epimorphisms are the surjective homomorphisms, the monomorphisms are the injective homomorphisms, and the isomorphisms are the bijective homomorphisms.

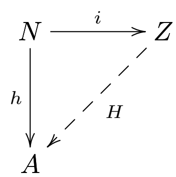
The category **CMonoid** has initial, final and zero objects, all being the single element monoids consisting of just the identity element.

This category also has products. The product of $(M, \cdot, 1)$ and $(N, \cdot, 1)$ is $(M \times N, \cdot, (1, 1)$ where $(m, n) \cdot (m', n') = (m \cdot m', n \cdot n')$, and the projection homomorphisms are just the set projection maps.

Verification that $(M \times N, \cdot, (1, 1))$ is indeed a monoid, that the projection functions are homomorphism, and that the universal mapping property for the product in **CMonoid** does hold are all straight forward.

Just as the sum of sets is much less familiar and more complicated than the product of sets, so it is with the sum of commutative monoids.

Proposition B.39 *For every* commutative *monoid N there is an Abelian group Z and a monoid homomorphism $i : N \longrightarrow Z$ so that for every monoid homomorphism $h : M \longrightarrow A$ of M to an Abelian group A there is a* unique *group homomorphism $H : Z \longrightarrow M$ so that $h = Hi$.*



Proof: The construction of Z is based on the intuition that the elements of Z arise as differences of elements of N. Moreover $n' - n = m' - m$ when $n' + m = n + m'$, so define $Z = N \times N/\sim$ where $(n', n) \sim (m', m)$ when $n' + m = n + m'$ and $i(n)$ is the equivalence class of $(n, 0)$ which we will denote for the duration of this proof as $((n', n))$. The binary operation on Z is defined in the usual way: $((n', n)) + ((m', m)) = ((n' + m', n + m))$ and this is well-defined, i.e., not dependent of the particular representatives. Straight forward computations show that Z is an Abelian group with $((0, 0))$ as identity and the inverse of $((n', n))$ is $((n, n'))$.

Finally given the homomorphism $h : N \longrightarrow A$, if there is a group homomorphism $H : Z \longrightarrow A$ with $Hi = h$, then because $((n', n)) = i(n') - i(n)$ it follows that $H((n', n)) = h(n') - h(n)$. Taking that as a definition, simple computations show that H is a well-defined group homomorphism with $Hi = h$ and this guarantees uniqueness as well as existence. ∎

Of course this is an example of a universal arrow (see Section V.2) from the commutative monoid N to the forgetful functor from the category **Ab** of Abelian groups to the category of commutative monoids.

B.2.5 Group – groups

Definition B.12: A **group** is a monoid in which every element has an inverse.

Definition B.13: A **group homomorphism** is just a monoid homomorphism.

It is natural to expect a requirement that a group homomorphism carries inverses to inverses, but that is a formal consequence.

Definition B.14: The category **Group** has as objects all groups and as morphisms the group homomorphisms between them.

B.2.6 FiniteGroup – finite groups

The category **FiniteGroup** has as objects all groups and as morphisms the group homomorphisms between them.

B.2.7 Ab – Abelian groups

Definition B.15: An **Abelian group** (also called a commutative group) is a group in which the binary operation is commutative.

The category **Ab** has as objects all *Abelian* groups and as morphisms the group homomorphisms between them.

B.2.7.1 TorAb – Torsion Abelian groups

B.2.7.2 DivAb – Divisible Abelian groups

B.2.7.3 TorsionFreeAb – Torsion Free Abelian groups

B.3 Rings

Rings are algebraic gadgets which have two binary operations, addition and multiplication, that are connected in some approximation of the familiar situation with numbers. The most common type of ring is sometimes called an associative ring with identity, but here the name ring is reserved for this type.

Definition B.16: A **ring** is a set R equipped with two binary operations $+$ and $\cdot$, called addition and multiplication, and distinguished elements 0 and 1 such that:

- $(R, +, 0)$ is an Abelian group:
  $\forall a, b, c \in R$
  - $0 + a = a = a + 0$
  - $(a + b) + c = a + (b + c)$
  - $a + b = b + a$
  - For each $a \in R$ there is an element $-a$ such that $a + -a = 0 = -a + a$
- $(R, \cdot, 1)$ is a monoid:
  - $(a \cdot 1) = a = 1 \cdot a$
  - $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
- Multiplication distributes over addition:
  - $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$
  - $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$

As with groups the symbol $\cdot$ is usually omitted and multiplication is just denoted by juxtaposition. Also the standard order of operation rules are used, so that, for example, $a + bc$ is an abbreviation for $a + (b \cdot c)$.

Although addition in a ring is commutative (i.e. $a + b = b + a$), multiplication is *not* assumed to be commutative. Rings in which multiplication is also commutative (such as the ring of integers) are called commutative rings.

The category **Ring** of rings is discussed in Section B.3.3 below while its subcategory of commutative rings is discussed in Section B.3.4.

There are a wide variety of interesting generalizations, most of them gotten by eliminating one or more of the items in the above definition. For example rngs are rings without identity. Here is the actual definition.

Definition B.17: A **rng** is a set R equipped with two binary operations $+$ and $\cdot$, called addition and multiplication, and a distinguished element 0 such that:

- $(R, +, 0)$ is an Abelian group.
- $(R, \cdot)$ is a semigroup.
- Multiplication distributes over addition.

The category **Rng** of rings without identity is discussed in Section B.3.2. Contrary to what the name suggests a rng may have a multiplicative identity, in particular every ring is a rng.

The simplest examples of rngs are gotten by defining multiplication on any Abelian group $(A, +, 0)$ by $a \cdot b = 0$ for all a, and b in A. The only case when this has an identity is the zero ring which has only one element, 0.

The system $(\mathbb{N}, +, 0, \cdot, 1)$ satisfies almost all of the condition to be a commutative ring – the only thing missing are negative numbers. Such gadgets are called *rigs* ("rings without negatives") or *semirings* and defined as follows.

Definition B.18: A **rig** is a set R equipped with two binary operations $+$ and $\cdot$, called addition and multiplication, and distinguished element 0 and 1 such that:

- $(R, +, 0)$ is a commutative monoid.
- $(R, \cdot, 1)$ is a monoid.
- Multiplication distributes over addition.

A particularly interesting class is the non-associative rings (remove the assumption that multiplication is associative) and various specific variations such as Lie rings and Jordan rings (which replace the associative law with a particular alternative).

Another direction is the class of rings called algebras. These are rings of one or another of the various types already mentioned but where the additive

Abelian group is assumed to be a vector space over some field or, more generally, a module over some base ring. A variety of categories of algebras are discussed in Section B.5.

B.3.1 Rig – rings without negatives

Recall that *rigs* were defined back in Section B.3.3 as follows.

A **rig** is a set R equipped with two binary operations $+$ and $\cdot$, called addition and multiplication, and distinguished elements 0 and 1 such that:

- $(R, +, 0)$ is an commutative monoid group.
- $(R, \cdot, 1)$ is a monoid.
- Multiplication distributes over addition.

Examples

- The most familiar example of a rig that is not a ring is $\mathbb{N}$ with the usual addition and multiplication.
- If R is any rig and X is any non-empty set, then the set R^X of all functions from X to R is a rig – the operations in R^X come from pointwise operations: $(f + g)(x) = f(x) + g(x)$, (fg)(x) = f(x)g(x), 0(x) = 0, and 1(x) - 1.

Definition B.19: A **rig homomorphism** is a function from one rig to another which is a group homomorphism between the additive monoids and a semigroup homomorphism between the multiplicative semigroups.

In detail a rig homomorphism $f : R \longrightarrow S$ satisfies:

(i) $f(0) = 0$.

(ii) $f(r + r') = f(r) + f(r')$.

(iii) $f(1) = 1$.

(iv) $f(r \cdot r') = f(r) \cdot f(r')$.

Definition B.20: The category **Rig** has as object all rigs (as defined above) and as morphisms all rig homomorphisms.

The identity function on a rig is a rig homomorphism, and the function composition of two rig homomorphism is again a rig homomorphism. With those two observations it is a now familiar formality to verify that **Rig** is indeed a category.

Theorem B.20 *Given a rig* $(R, +, 0, \cdot, 1)$*, there is a universal morphism from* R *to a ring.*

Proof: On $R \times R$ define $(a,b)\ (c,d) <=> \exists z \in R,\ a+d+z = c+b+z$. The quotient $R(R) = (RR)/$ is a ring and the map $R \longrightarrow R(R)$ given by $x \mapsto [(x,0)]/$ "(= "x - 0")" is a rig morphism. In general, the morphism $R \longrightarrow R(R)$ need not be injective. ∎

B.3.2 Rng – rings without identity

Definition B.21: A **rng homomorphism** is a function from one rng to another which is a group homomorphism between the additive groups and also a semigroup homomorphism between the multiplicative semigroups.

In detail a rng homomorphism $f : R \longrightarrow S$ satisfies:

(i) $f(0) = 0$.

(ii) $f(r + r') = f(r) + f(r')$.

(iii) $f(r \cdot r') = f(r) \cdot f(r')$.

The identity function on a rng is a rng homomorphism, and the function composition of two rng homomorphism is again a rng homomorphism. With those two observations it is a now familiar formality to verify that **Rng** is indeed a category.

B.3.3 Ring – associative rings with identity

As with groups the symbol $\times$ is usually omitted and multiplication is just denoted by juxtaposition. Also the standard order of operation rules are used, so that, for example, $a + bc$ is an abbreviation for $a + (b \times c)$.

Although ring addition is commutative (i.e. $a + b = b + a$), note that the commutativity for multiplication ($a \times b = b \times a$) is *not* among the ring axioms listed above. Rings that also satisfy commutativity for multiplication (such as the ring of integers) are called commutative rings. Not all rings are commutative.

Definition B.22: A **ring homomorphism** is just an rng homomorphism.

In detail a ring homomorphism $f : R \longrightarrow S$ satisfies:

(i) $f(0) = 0$.

(ii) $f(r + r') = f(r) + f(r')$.

(iii) $f(1) = 1$.

(iv) $f(r \cdot r') = f(r) \cdot f(r')$.

The identity function on a ring is a ring homomorphism, and the function composition of two ring homomorphism is again a ring homomorphism. With those two observations it is a now familiar formality to verify that **Ring** is indeed a category.

B.3.4 CommutativeRing – commutative rings

CommutativeRing is the category of commutative rings with identity with morphisms being the rings homomorphisms (which take 1 to 1.)

B.3.5 Field – fields

Definition B.23: A **field** is a commutative ring in which every non-zero element has a multiplicative inverse.

B.4 Modules

B.4.1 Module – modules over a commutative ring

Throughout this section R is a commutative ring.

Definition B.24: An R**-module** M is an Abelian group $(M, +, 0)$ together with a function $R \times M \longrightarrow M$ (written as $(r, m) \mapsto rm$) satisfying: For all m, n in M and for all r, s in R

(i) r(m + n) = rm + rn.

(ii) 1m = m.

(iii) (r + s)m = rm + sm.

(iv) (rs)m = r(sm).

This is actually a family of categories. For each commutative ring R there is the category $\mathbf{Module_R}$ with objects all R-modules and morphisms all R-module homomorphisms. In the case that $R = \mathbb{Z}$, the category $\mathbf{Module}_{\mathbb{Z}}$ is the same as **Ab**. As you should expect there is in general a great deal of similarity between **Ab** and other categories of modules.

Much of the time the particular ring is not interesting, in which case **Module** will be written to stand for $\mathbf{Module_R}$ with R unspecified.

B.4.2 Matrices – matrices over a commutative ring

This is another parametrised family of categories. For each commutative ring R, the category $\mathbf{Matrices}_R$ has as objects all positive integers and as morphisms matrices with entries in R. In particular an $m \times n$ matrix is a morphism from n to m with the identity morphism on m being the $m \times m$ identity matrix, and composition being the usual matrix multiplication.

This category is actually equivalent to the full subcategory of all finitely generated free R-modules in $\mathbf{Module_R}$.

B.4.3 Vect – vector spaces

For each field K we have the category $\mathbf{Vect}_K$ which has as objects all vector spaces over K and as morphisms all linear transformations between them. This is really just the category $\mathbf{Module}_K$, but it is sufficiently interesting and well studied in its own right that we will treat it separately.

Much of the time the particular field is not interesting, in which case **Vect** will be written to stand for $\mathbf{Vect}_K$ with K unspecified.

B.4.4 FDVect – finite dimensional vector spaces

The full subcategory $\mathbf{FDVect}_K$ of $\mathbf{FDVect}_K$ has as objects just the finite dimensional vector spaces in $\mathbf{Vect}_K$. This is is sufficiently interesting that it deserves separate mention.

Much of the time the particular field is not interesting, in which case **FDVect** will be written to stand for $\mathbf{FDVect}_K$ with K unspecified.

B.5 Algebras

Algebras are rings in which the additive structure is not just an Abelian group, but a module over a commutative ring. The most commonly discussed algebras are associative algebras over fields, but here we will reserve the name algebra for associative algebras over a commutative ring.

Throughout this section R will be a commutative ring.

Definition B.25: An R-algebra is both an R-module and a ring in such a way that multiplication in the ring is an R-bilinear map. In detail an R-algebra

B.5.1 Algebra – associative algebras

This is actually a family of categories. For a fixed commutative ring (with identity), R, $\mathbf{Algebra}_R$ is the category with objects all R-algebras and morphisms all the algebra homomorphisms. By an algebra we certainly mean a set with both an operation of addition and an operation of multiplication. Moreover each ring is an Abelian group with respect to addition.

And we require the algebra homomorphisms to take 1 to 1.

Much of the time the particular commutative ring is not interesting, in which case **Algebra** will be written to stand for $\mathbf{Algebra}_R$ with R unspecified.

B.5.2 LieAlgebra – Lie algebras

This is actually a family of categories. For a fixed commutative ring (with identity), R, **LieAlgebra** is the category with objects all R Lie algebras and morphisms all the Lie algebra homomorphisms.

B.6 Order

B.6.1 Preorder – preorders

The category **Preorder** has as objects all preorders and as morphisms all the order preserving functions between them.

As the identity function on any preorder is certainly order preserving and the composition of two order preserving functions is also order preserving, we see that **Preorder** is indeed a category.

Proposition B.40 *The category* **Preorder** *has products.*

Proof: For partially ordered sets X and Y define a partial order on product set $X \times Y$ by $(x, y) \leq (x', y')$ iff $x \leq x'$ and $y \leq y'$. The usual projection functions π_X and π_Y are certainly order preserving. Moreover if $f : Z \longrightarrow X$ and $g : Z \longrightarrow Y$ are preorder morphisms, i.e., monotone functions, then the function $(f, g) : Z \longrightarrow X \times Y$ is monotone and clearly has the universal mapping property required for a product. ∎

Proposition B.41 *The category* **Preorder** *has sums.*

B.6.2 Poset – partially ordered sets

Partially ordered sets, or **posets**, are those preorders where the order relation is *antisymmetric*:

$$\forall x, y \ \in X, x \leq y \wedge y \leq x \Rightarrow x = y.$$

The category **Poset** of partially ordered sets is the full subcategory of **Preorder** with objects the posets.

For more information about posets, consult Mac Lane and Birkhoff [55, II.8, IV.6, XIV].)

B.6.3 Lattice – lattices

The category **Lattice** has as objects all lattices and lattice morphisms between them.

There are two quite different ways to describe lattices – as certain types of partially ordered sets, and as certain types of algebraic structures. We start with the poset approach.

Lattices as posets In any poset L we have the notions of upper and lower bounds and from them the notions least upper bounds and greatest lower bounds, the essential notions for defining a lattice.

Recall from definition A.56 that if S is any subset of a poset L, then $u \in L$ is an *upper bound* for S provided $s \in S \Rightarrow s \leq u$.

Equally well there is the dual notion of a lower bound: l is a *lower bound* for S provided $s \in S \Rightarrow s \geq l$.

A least upper bound is a minimum among upper bounds, while a greatest lower bound is a maximum among lower bounds. The notion of glb is dual to the notion of lub.

Now a lattice is a poset in which non-empty finite sets have lubs and glbs.

Definition B.26: A **lattice** is a poset in which every pair of elements has both a lub and a glb. The lub of $\{x, y\}$ is called the **join** of x and y and is written as $x \vee y$, while the glb of $\{x, y\}$ is called the **meet** of x and y and is written as $x \wedge y$.

So every lattice, L, has binary operations $\wedge : L \times L \longrightarrow L$ and $\vee : L \times L \longrightarrow L$. The order relation can be retrieved from the meet and join: $x \leq y \iff x \wedge y = y \wedge x \vee y = x$. This will be the connection between the order and algebraic definitions.

The morphisms between lattices are not just order preserving, they are required to preserve meets and joins as well.

Definition B.27: A **lattice morphism** $f : L \longrightarrow L'$ is a function with $f(x \wedge y) = f(x) \wedge f(y)$ and $f(x \vee y) = f(x) \vee f(y)$.

While no requirement that a lattice morphism be order preserving was explicitly mentioned, the connection between the order relation and the meets and joins guarantees that a lattice morphism is a poset morphism as well.

Although the definition of a lattice only requires the existence of lubs and glbs for two element sets, induction on the number of elements in a set easily shows that lubs and glbs for all non-empty finite sets exists in every lattice.

For more information consult Mac Lane and Birkhoff [55, XIV].)

It will be stated explicitly whenever a lattice is required to have a least or greatest element. If both of these special elements do exist, then the poset is called a bounded lattice. An easy induction argument also shows the existence of all suprema and infima of non-empty finite subsets of any lattice.

the above definition in terms of the existence of suitable Galois connections between related posets – an approach that is of special interest for category theoretic investigations of the concept.

Lattices as algebraic structures

Consider an algebraic structure in the sense of universal algebra, given by $(L, \vee, \wedge)$, where $\vee$and $\wedge$are two binary operations. L is a lattice if the following identities hold for all elements a, b, and c in L:

- Idempotent laws: $a \vee a = a$, $a \wedge a = a$
- Commutative laws: $a \vee b = b \vee a$, $a \wedge b = b \wedge a$
- Associative laws: $a \vee (b \vee c) = (a \vee b) \vee c$, $a \wedge (b \wedge c) = (a \wedge b) \wedge c$

- Absorption laws: $a \vee (a \wedge b) = a$, $a \wedge (a \vee b) = a$

It turns out that the idempotency laws can be deduced from absorption and thus need not be stated separately.

Bounded lattices include neutral elements 0 and 1 for the meet and join operations in the above definition.

Connection between both definitions

Obviously, an order theoretic lattice gives rise to two binary operations $\vee$and $\wedge$. It now can be seen very easily that this operation really makes $(L, \vee, \wedge)$ a lattice in the algebraic sense. More surprisingly the converse of this result is also true: consider any algebraically defined lattice $(M, \vee, \wedge)$. Now define a partial order $\leq$ on M by setting

$x \leq y$ iff $x = x \wedge y$

or, equivalently,

$x \leq y$ iff $y = x \vee y$

for all elements x and y in M. The above laws for absorption ensure that both definitions are indeed equivalent. An easy check shows that the relation $\leq$ introduced in this way defines a partial ordering within which binary meets and joins are given through the original operations $\vee$and $\wedge$. Conversely, the order induced by the algebraically defined lattice $(L, \vee, \wedge)$ that was derived from the order theoretic formulation above coincides with the original ordering of L.

Hence, the two definitions can be used in an entirely interchangeable way, depending on which of them appears to be more convenient for a particular purpose.

Examples

- For any set A, the collection of all finite subsets of A (including the empty set) can be ordered via subset inclusion to obtain a lattice. The lattice operations are intersection (meet) and union (join) of sets, respectively. This lattice has the empty set as least element, but it will only contain a greatest element if A itself is finite. So it is not a bounded lattice in general.

- The natural numbers in their common order are a lattice, the lattice operations given by the min and max operations. The least element of this lattice is 0, but no greatest element exists.

- Any complete lattice (also see below) is a (rather specific) bounded lattice. A broad range of practically relevant examples belongs to this class.

- The set of compact elements of an arithmetic complete lattice is a lattice with a least element, where the lattice operations are given by restricting the respective operations of the arithmetic lattice. This is the specific property which distinguishes arithmetic lattices from algebraic lattices, for which the compacts do only form a join-semilattice. Both of these classes of complete lattices are studied in domain theory.

Further examples are given for each of the additional properties that are discussed below.

Morphisms of lattices

The appropriate notion of a morphism between two lattices can easily be derived from the algebraic definition above: given two lattices $(L, \vee, \wedge)$ and $(M, \cup, \cap)$, a homomorphism of lattices is a function $f : L \longrightarrow M$ with the properties that

- $f(x \vee y) = f(x) \cup f(y)$, and
- $f(x \wedge y) = f(x) \cap f(y)$.

In the order-theoretical formulation, these conditions just state that a homomorphism of lattices is a function that preserves binary meets and joins. For bounded lattices, preservation of least and greatest elements is just preservation of join and meet of the empty set.

Note that any homomorphism of lattices is necessarily monotone with respect to the associated ordering relation. For an explanation see the article on preservation of limits. The converse is of course not true: monotonicity does by no means imply the required preservation properties.

Using the standard definition of isomorphisms as invertible morphisms, one finds that an isomorphism of lattices is exactly a bijective lattice homomorphism. Lattices and their homomorphisms obviously form a category.

Properties of lattices

The definitions above already introduced the simple condition of being a bounded lattice. A number of other important properties, many of which lead to interesting special classes of lattices, will be introduced below.

B.6.4 Full subcategories of equationally defined lattices

Complete Lattices **Definition B.28:** A lattice is **complete** provided all of its subsets has both a least upper bound and a greatest lower bound.

Distributive Lattices Since any lattice comes with two binary operations, it is natural to consider distributivity laws among them.

Definition B.29: A lattice $(L, \vee, \wedge)$ is **distributive**, provided x, y and z in L implies

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$$

Surprisingly, this condition turns to be equivalent to its dual statement:

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$$

Modular Lattices Distributivity is too strong a condition for certain applications. A strictly weaker but still useful property is modularity.

Definition B.30: A lattice $(L, \vee, \wedge)$ is **modular** provided x, y, and z in L implies

$$x \vee (y \wedge (x \vee z)) = (x \vee y) \wedge (x \vee z)$$

Another equivalent statement of this condition is as follows: if $x \leq z$ then for all y

$$x \vee (y \wedge z) = (x \vee y) \wedge z$$

For example, the lattice of submodules of a module and the lattice of normal subgroups of a group have this special property. Furthermore, every distributive lattice is indeed modular.

B.6.5 Boolean – the categories of Boolean algebras and Boolean rings

There are two apparently quite different notions – Boolean rings and Boolean algebras – that have been recognized as "equivalent" for a long time. For both of these notions we have categories and the categories are naturally equivalent.

Definition B.31: A **Boolean ring** is a ring (cf. Section B.3.3 on **Ring**) in which every element is idempotent, i.e., for every element b of a Boolean ring B we have $b^2 = b$.

The canonical example of a Boolean ring is $\mathbb{Z}_2$, the integers modulo 2.

The other basic example is the ring of subsets of a given set: For any set X we have the set $\mathcal{P}(X)$ of all subsets of X (cf page 86.) Define the "sum" of two elements of $\mathcal{P}(X)$ to be their symmetric difference (cf 187)), i.e., $S + T = (S - T) \cup (T - S) = (S \cup T) - (S \cap T) = \{x : x \in S \vee x \in T \text{ but not both}\}$ and the product of two elements to be their intersection, i.e., $S \bullet T = S \cap T$. In fact if X is a one element set, then

A Boolean algebra is a lattice $(A, \wedge, \vee)$ (considered as an algebraic structure) with the following four additional properties:

1. bounded below: $\exists 0 \in A\ \forall a \in A\ a \vee 0 = a$.
2. bounded above: $\exists 1 \in A\ \forall a \in A\ a \wedge 1 = a$.
3. distributive law: $\forall a, b, c\ \in A (a \vee b) \wedge c = (a \wedge c) \vee (b \wedge c)$.
4. existence of complements: $\forall a \in A\ \exists \neg a \in A\ (a \vee \neg a = 1 \wedge a \wedge \neg a = 0)$.

From these axioms, we find that the smallest element 0, the largest element 1, and the complement $\neg a$ of any element a are uniquely determined.

Like any lattice, a Boolean algebra $(A, \wedge, \vee)$ gives rise to a partially ordered set $(A, \leq\ )$ by defining $a \leq b$ iff $a = a \wedge b$ (which is also equivalent to $b = a \vee b$).

Equally well a Boolean algebra can be defined as distributive lattice $(A, \leq)$ (considered as a partially ordered set) with least element 0 and greatest element 1, within which every element x has a complement $\neg x$ such that $x \wedge \neg x = 0$ and $x \vee \neg x = 1$.

Here $\wedge$ and $\vee$ are used to denote the infimum (meet) and supremum (join) of two elements. Again, if complements in the above sense exist, then they are uniquely determined.

The algebraic and the order theoretic perspective as usual can be used interchangeably and both are of great use to import results and concepts from both universal algebra and order theory. In many practical examples an ordering relation, conjunction, disjunction, and negation are all naturally available, so that it is straightforward to exploit this relationships.

Here are several other theorems valid in all Boolean algebras. For example, for all elements a and b of a Boolean algebra, $a \wedge 0 = 0$ and $a \vee 1 = 1$, $\neg 1 = 0$ and $\neg 0 = 1$, $\neg\neg a = a$ and that both de Morgan's laws are valid, i.e. $\neg(a \wedge b) = (\neg a) \vee (\neg b)$ and $\neg(a \vee b) = (\neg a) \wedge (\neg b)$.

General insights from duality in order theory apply to Boolean algebras. Especially, the order dual of every Boolean algebra, or, equivalently, the algebra obtained by exchanging $\wedge$ and $\vee$, is also a Boolean algebra. Thus the dual version of the distributive law, $(a \wedge b) \vee c = (a \vee c) \wedge (b \vee c)$ also holds true. In general, any law valid for Boolean algebras can be transformed into another valid dual law by exchanging 0 with 1, $\wedge$ with $\vee$, and $\leq$ with $\geq$.

A homomorphism between the Boolean algebras A and B is a function $f : A \longrightarrow B$ such that $\forall a \in A \; \forall b \in A$, $f(a \wedge b) = f(a) \wedge f(b)$ and $f(a \vee b) = f(a) \vee f(b)$ and $f(0) = 0, f(1) = 1$.

It then follows that $f(\neg a) = \neg f(a)$ for all $a \in A$ as well. The class of all Boolean algebras, together with this notion of morphism, forms a category. An isomorphism from A to B is a homomorphism from A to B which is bijective. The inverse of an isomorphism is also an isomorphism, and we call the two Boolean algebras A and B isomorphic. From the standpoint of Boolean algebra theory, they cannot be distinguished; they only differ in the notation of their elements.

Boolean rings, ideals and filters

Every Boolean algebra $(A, \wedge, \vee)$ gives rise to a ring $(A, +, *)$ by defining $a+b = (a \wedge \neg b) \vee (b \wedge \neg a)$ (this operation is called "symmetric difference" in the case of sets and XOR in the case of logic) and $a * b = a \wedge b$. The zero element of this ring coincides with the 0 of the Boolean algebra; the multiplicative identity element of the ring is the 1 of the Boolean algebra. This ring has the property that $a * a = a$ for all $a \in A$; rings with this property are called Boolean rings.

Conversely, if a Boolean ring A is given, we can turn it into a Boolean algebra by defining $x \vee y = x + y - xy$ and $x \wedge y = xy$. Since these two operations are inverses of each other, we can say that every Boolean ring arises from a Boolean algebra, and vice versa. Furthermore, a map $f : A \longrightarrow B$ is a homomorphism of Boolean algebras if and only if it is a homomorphism of Boolean rings. The categories of Boolean rings and Boolean algebras are equivalent.

An ideal of the Boolean algebra A is a subset I such that for all $x, y \in I$ we have $x \wedge y \in I$ and for all $a \in A$ we have $a \wedge x \in I$. This notion of ideal coincides with the notion of ring ideal in the Boolean ring A. An ideal I of A is called **prime** if $I \neq A$ and if $a \wedge b \in I$ always implies $a \in I$ or $b \in I$. An ideal I of A is called **maximal** if $I \neq A$ and if the only ideal properly containing I is A itself. These notions coincide with ring theoretic ones of prime ideal and maximal ideal in the Boolean ring A.

The dual of an ideal is a **filter**. A filter of the Boolean algebra A is a subset p such that for all $x, y \in p$ we have $x \wedge y \in p$ and for all $a \in A$ if $a \vee x = a$ then $a \in p$.

A Boolean algebra (Boolean algebra) is a set A together with binary operations $+$ and $\times$ and a unary operation -, and elements 0, 1 of A such that the following laws hold: commutative and associative laws for addition and multiplication, distributive laws both for multiplication over addition and for addition over multiplication, and the following special laws:

- $x + (x \times y) = x$
- $x \times (x + y) = x$
- $x + (-x) = 1$
- $x \times (-x) = 0$

These laws are better understood in terms of the basic example of a Boolean algebra, consisting of a collection A of subsets of a set X closed under the operations of union, intersection, complementation with respect to X, with members $\emptyset$ and X. Many elementary laws follow from these axioms, keeping in mind this example for motivation. Any Boolean algebra has a natural partial order $\leq$ defined upon it by saying that $x \leq y$ if and only if $x + y = y$. This corresponds in our main example to $\subseteq$. Of special importance is the two-element Boolean algebra, formed by taking the set X to have just one element. An important elementary result is that an equation holds in all Boolean algebras if and only if it holds in the two-element Boolean algebra. Next, we define $x \oplus y = (x - y) + (y - x)$. Then A together with $\oplus$ and $\times$, along with 0 and 1, forms a ring with identity in which every element is idempotent. Conversely, given such a ring, with addition $\oplus$ and multiplication $\times$, define $x + y = x \oplus y \oplus (x \times y)$ and $-x = 1 \oplus x$. This makes the ring into a Boolean algebra. These two processes are inverses of one another, and show that the theory of Boolean algebras and of rings with identity in which every element is idempotent are definitionally equivalent. This puts the theory of Boolean algebras into a standard object of research in algebra.

An atom in a Boolean algebra is a nonzero element a such that there is no element b with $0 < b < a$. A Boolean algebra is atomic if every nonzero element of the Boolean algebra is above an atom. Finite Boolean algebras are atomic, but so are many infinite Boolean algebras. Under the partial order $\leq$ above, $x + y$ is the least upper bound of x and y, and $x \times y$ is the greatest lower bound of x and y. We can generalize this: ΣX is the least upper bound of a set X of elements, and ΠX is the greatest lower bound of a set X of elements. These do not exist for all sets in all Boolean algebras; if they do always exist, the Boolean algebra is said to be **complete**.

2. The elementary algebraic theory

Several algebraic constructions have obvious definitions and simple properties for Boolean algebras: subalgebras, homomorphisms, isomorphisms, and direct products (even of infinitely many algebras). Some other standard algebraic constructions are more peculiar to Boolean algebras. An **ideal** in a Boolean algebra is a subset I closed under $+$, with 0 as a member, and such that if $a \leq b \in I$, then also $a \in I$. Although not immediately obvious, this is the same as the ring-theoretic concept. There is a dual notion of a filter (with no counterpart in rings in general). A **filter** is a subset F closed under $\times$, having 1 as a member, and such that if $a \geq b \in F$, then also $a \in F$. An **ultrafilter** on A is a filter F with the following properties: $0 \notin F$, and for any $a \in A$, either $a \in F$ or $-a \in F$. For any $a \in A$, let $S(a) = \{F : F \text{ is an ultrafilter on } A \wedge a \in F\}$. Then S is an isomorphism onto a Boolean algebra of subsets of the set X of all ultrafilters on A. This establishes the basic Stone representation theorem, and clarifies the origin of Boolean algebras as concrete algebras of sets. Moreover, the sets $S(a)$ can be declared to be a base for a topology on X, and this turns X into a totally disconnected compact Hausdorff space. This establishes a one-to-one correspondence between the class of Boolean algebras and the class of such spaces. As a consequence, used very much in the theory of Boolean algebras, many topological theorems and concepts have consequences for Boolean algebras. If x is an element of a Boolean algebra, we let $0x = -x$ and $1x = x$. If $(x(0), \ldots, x(m-1))$ is a finite sequence of elements of a Boolean algebra A, then every element of the subalgebra of A generated by $\{x(0), \ldots, x(m-1)\}$ can be written as a sum of monomials $e(0)x(0) \times \cdots \times e(m-1)x(m-1)$ for e in some set of functions mapping $m = \{0, \ldots, m-1\}$ into $2 = \{0, 1\}$. This is an algebraic expression of the disjunctive normal form theorem of sentential logic. A function f from a set X of generators of a Boolean algebra A into a Boolean algebra B can be extended to a homomorphism if and only if $e(0)x(0) \ldots e(m-1)x(m-1) = 0$ always implies that $e(0)f(x(0)) \ldots e(m-1)f(x(m-1)) = 0$. This is Sikorski's extension criterion. Every Boolean algebra A can be embedded in a complete Boolean algebra B in such a way that every element of B is the least upper bound of a set of elements of A. B is unique up to A-isomorphism, and is called the completion of A. If f is a homomorphism from a Boolean algebra A into a complete Boolean algebra B, and if A is a subalgebra of C, then f can be extended to a homomorphism of C into B. This is Sikorski's extension theorem. Another general algebraic notion which applies to Boolean algebras is the notion of a free algebra. This can be concretely constructed for Boolean algebras. Namely, the free Boolean algebra on K is the Boolean algebra of closed-open subsets of the two element discrete space raised to the K power.

B.6.5.1 Boolean – the category of complete Boolean algebras

B.7 Ordered Algebraic Structures

B.7.1 OrderedMagma – Ordered Magmas

An ordered magma is both a magma and a poset where the order is respected by the binary operation of the magma.

Definition B.32: An **ordered magma** is a magma $(M, \cdot)$ which also has a partial order $\leq$ with the property:

$$\forall m, n, p \in M \; [m \leq n \Rightarrow mp \leq np \wedge pm \leq pn]$$

The basic example of an ordered magma is the additive magma of natural numbers with the usual ordering. In this case the defining property connecting the order with addition reads:

$$\forall m, n, p \in \mathbb{N} \; [m \leq n \Rightarrow m + p \leq n + p \wedge p + m \leq p + n]$$

The multiplicative magma of natural numbers with the usual ordering is another quite different ordered magma where we have the familiar property:

$$\forall m, n, p \in \mathbb{N} \; [m \leq n \Rightarrow mp \leq np \wedge pm \leq pn]$$

B.7.2 OrderedMonoid – Ordered Monoids

Definition B.33: An **ordered monoid** is a monoid $(M, 1, \cdot)$ which also has a partial order $\leq$ with the property:

$$\forall m, n, p \in M \; [m \leq n \Rightarrow mp \leq np \wedge pm \leq pn]$$

The basic example of an ordered monoid is the additive monoid of natural numbers with the usual ordering. In this case the defining property connecting the order with addition reads:

$$\forall m, n, p \in \mathbb{N} \; [m \leq n \Rightarrow m + p \leq n + p \wedge p + m \leq p + n]$$

The multiplicative monoid of natural numbers with the usual ordering is another quite different ordered monoid.

B.7.3 OrderedGroup – Ordered Groups

Definition B.34: An **ordered group** is an ordered monoid in which every element has an inverse.

By contrast with an ordered monoid the order relation can be reduced to comparison with the identity: $g \leq h$ iff $1 \leq g^{-1}h$ which is also equivalent to $1 \leq hg^{-1}$. This can be carried a step further and the order specified in terms of an invariant subset. When G is an ordered group, define $G^{+} = \{g \in G : 1 \leq g\}$. Then $G+$ has three characteristic properties:

(i) $1 \in G^{+}$.

(ii) $g, h \in G^{+} \Rightarrow gh \in G^{+}$.

(iii) $\forall g \in G\ [h \in G^{+} \Rightarrow g^{-1}hg \in G^{+}$.

In the reverse direction, if G^{+} is a subset of a group G and has the above three properties, then defining $g \leq h$ iff $g^{-1}h \in G^{+}$ makes $\leq$ a partial order on G in such a way that it is an ordered group.

The basic example of an ordered group is the additive group of integers with the usual ordering.

B.7.4 OrderedRig – Ordered Rigs

Definition B.35: An **ordered rig** is a rig R together with a partial order, $\leq$, which is an ordered monoid under addition and $\forall r, s, t \in R,\ r \leq s \wedge 0 \leq t \Rightarrow rt \leq st \wedge tr \leq ts$.

Examples

- The most familiar example of an ordered rig is $\mathbb{N}$ with all the usual structure.
- Every rig is an ordered rig for the trivial partial order where $a \leq b$ iff $a = b$.
- If R is any rig and X is any non-empty set, then the set R^X of all functions from X to R is a rig (see Section B.3.1) and is an ordered rig for the partial order $f \leq g \iff \forall x \in X,\ f(x) \leq g(x)$.

Definition B.36: A **homomorphism of ordered rigs** is a rig homomorphism that is also a monotone function between the two partially ordered sets.

Definition B.37: The category **OrderedRig** of ordered rigs has as objects all ordered rigs and as morphisms the homomorphisms between them.

The identity function on an ordered rig is a homomorphism of ordered rigs, and the function composition of two homomorphisms of ordered rigs is again a homomorphism of ordered rigs, so it is a now familiar formality to verify that **OrderedRig** is indeed a category.

Proposition B.42 *For every ordered commutative rig R there is an ordered commutative ring S and a universal ordered rig homomorphism $i : R \longrightarrow S$, i.e., if $h : R \longrightarrow T$ is an ordered rig homomorphism to an ordered ring T, then there is a* unique *ordered ring homomorphism $\overline{h} : S \longrightarrow T$ with $\overline{h}i = h$*

Proof: There is an easy construction based on the perception that the elements of S are differences of elements of R.

Let $S' = R \times R$ and define the relation on S' by (r_1, r_2) (r'_1, r'_2) iff there is some $r'' \in R$ such that

$\blacksquare$

B.7.5 OrderedRing – Ordered Rings

Throughout this section only non-zero rings will be considered. This is the same as saying that $0 \neq 1$ in all of these rings.

Definition B.38: A **compatible order** on a ring R is a subset R^+ of R satisfying: $\forall r, s \in R$,

(i) $r, s \in R^+ \Rightarrow r + s \in R^+ \wedge rs \in R^+$.

(ii) (Trichotomy Law) If $r \neq 0$, then either $r \in R^+$ or $-r \in R^+$, but not both.

An **ordered ring** is a non-zero ring together with a compatible order.

B.8 Graphs

B.8.1 Graph – graphs

The category **Graph** has as objects all graphs and graph homomorphisms between them.

Definition B.39: A **graph** is a pair of sets $G = (V,\ E)$ with each element of E a 2-element subset of V. The elements of V are called the **vertices**of G, and the elements of E are called the **edges**of G.

A good introductory treatment of graph theory is Diestel [14]. Of course the category of graphs must have morphisms, but like most texts on graph theory Diesel makes no mention of graph homomorphism (though isomorphisms and colorings appear), and graph are assumed finite. Graph homomorphisms have actually been studied for 50 years or so, but there are only a couple of books that discuss them in any depth, for example see Hell and Nešetřil [31] which is however almost exclusively about finite graphs.

Definition B.40:

B.8.2 Digraph – directed graphs

The category **Digraph** has as objects all directed graphs (commonly called digraphs) and digraph homomorphisms between them.

Definition B.41:

Definition B.42:

B.9 Topology

B.9.1 Metric – metric spaces

The category **Metric** has as objects all metric spaces and as morphisms all contractions between them. [RIGHT CHOICE?]

B.9.2 Uniform – uniform spaces

The category **Uniform** has as objects all uniform spaces and as morphisms all uniformly continuous functions between them.

B.9.3 Top – topological spaces

The category **Top** has as objects all topological spaces and as morphisms all continuous functions between them.

B.9.4 Comp – compact Hausdorff spaces

The category **Comp** has as objects all compact Hausdoff topological spaces and as morphisms all continuous functions between them.

B.9.5 Kspace – K-spaces

The category **Kspace** has as objects all compactly generated Hausdorff topological spaces (K-spaces) and as morphisms all continuous functions between them.

B.9.6 Homotopy – the homotopy category of topological spaces

Homotopy is the quotient category of **Top** where the equivalence relation on continuous functions is homotopy, i.e., for $f_0, f_1 : X \longrightarrow Y$, $f_0 \simeq f_1$ iff there exists a continuous functions (a "homotopy") $F : X \times [0,1] \longrightarrow Y$ with $F(x,0) = f_0(x)$ and $F(x,1) = f_1(x)$ for all $x \in X$.

B.9.7 HSpace – H-Spaces

HSpace are

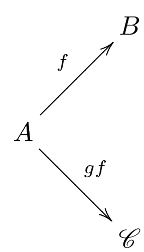
B.10 Simplicial Categories

B.10.1 Simplicial – simplicial sets

The category of simplicial sets has as objects all simplicial sets and as morphisms ...

Let $\mathscr{C}$be a small category. It is easy to define the sets $N(\mathscr{C})_k$ for small k, which leads to the general definition. In particular, there is a 0-simplex of $N(\mathscr{C})$ for each object of $\mathscr{C}$. There is a 1-simplex for each morphism $f : x \longrightarrow y$ in $\mathscr{C}$. Now suppose that $f : x \longrightarrow y$ and $g : y \longrightarrow z$ are morphisms in $\mathscr{C}$. Then we also have their composition $gf : x \longrightarrow z$.

A 2-simplex.



The diagram suggests our course of action: add a 2-simplex for this commutative triangle. Every 2-simplex of $N(\mathscr{C})$ comes from a pair of composable morphisms in this way. Note that the addition of these 2-simplices does not erase or otherwise disregard morphisms obtained by composition, it merely remembers that that is how they arise.

In general, $N(\mathscr{C})_k$ consists of the k-tuples of composable morphisms

$A_0 \longrightarrow A_1 \longrightarrow A_2 \longrightarrow \cdots \longrightarrow A_{k-1} \longrightarrow A_k$

of $\mathscr{C}$. To complete the definition of $N(\mathscr{C})$ as a simplicial set, we must also specify the face and degeneracy maps. These are also provided to us by the structure of $\mathscr{C}$as a category. The face maps

$d_i : N(\mathscr{C})_k \longrightarrow N(\mathscr{C})_{k-1}$

are given by composition of morphisms at the i-th object. This means that d_i sends the k-tuple

$A_0 \longrightarrow A_1 \longrightarrow A_2 \longrightarrow \cdots \longrightarrow A_{k-1} \longrightarrow A_k$

to the (k-1)-tuple

$A_0 \longrightarrow \cdots A_{i-1} \longrightarrow A_{i+1} \longrightarrow \cdots \longrightarrow A_k$.

That is, the map d_i composes the morphisms $A_{i-1} \longrightarrow A_i$ and $A_i \longrightarrow A_{i+1}$ into the morphism $A_{i-1} \longrightarrow A_{i+1}$, yielding a (k-1)-tuple for every k-tuple.

Similarly, the degeneracy maps

$s_i : N(\mathscr{C})_k \longrightarrow N(\mathscr{C})_{k+1}$

are given by inserting an identity morphism at the object A_i.

Recall that simplicial sets may also be regarded as functors $\Delta^{\mathrm{op}} \longrightarrow Set$, where Δ is the category of totally ordered finite sets and order-preserving morphisms. Every partially ordered set P yields a (small) category i(P) with objects the elements of P and with a unique morphism from p to q whenever $p < q$ in P. We thus obtain a functor i from the category Δ to the category of small categories. We can now describe the nerve of the category $\mathscr{C}$as the functor $\Delta^{\mathrm{op}} \longrightarrow \mathbf{Set}$

$N(\mathscr{C})(?) = Fun(i(?), \mathscr{C})$.

This description of the nerve makes functoriality quite transparent; for example, a functor between small categories $\mathscr{C}$and $\mathscr{D}$induces a map of simplicial sets $N(\mathscr{C}) \longrightarrow N(\mathscr{D})$. Moreover a natural transformation between two such functors induces a homotopy between the induced maps. This observation can be regarded as the beginning of one of the principles of higher category theory.

The *nerve* of an internal category is discussed by Johnstone [33, B2.3.2]

B.10.2 Kan – the homotopy category of Kan complexes

B.11 Differential, Graded and Filtered Algebraic Gadgets

B.11.1 Graded Category

Definition B.43: Let $\mathcal{I}$ be a category. An **$\mathcal{I}$-graded category** is a category $\mathscr{C}$ together with a functor $G : \mathscr{C} \longrightarrow \mathcal{I}$.

For each object I in $\mathcal{I}$, $G(I)$ is called the I-component of G.

B.11.2 GradedModule

Definition B.44: This is another family of categories.

B.11.3 GradedRing

Definition B.45: This is another family of categories. If N is any monoid, an N-graded ring $R_\bullet$ consists of a family $\{R_n : n \in N\}$ of Abelian groups together with a multiplication $R_n \times R_m \longrightarrow R_{m+n}$ denoted in the usual way as $(r, s) \mapsto rs$ satisfying:

(i) r(s + s') = rs + rs'; (r + r')s = rs + r's.

(ii) r(st) = (rs)t

Definition B.46: The ring **associated** to an M-graded ring $R_\bullet$

B.11.4 ChainComplex– Chain complexes

Definition B.47: A **chain complex** $C_\bullet$ is a $\mathbb{Z}$-graded module together with an endomorphism of degree -1 such that $d \circ d = 0$.

The endomorphism d is called the **boundary map** of the chain complex.

Definition B.48: A **chain map** or **morphism of chain complexes** is a 0-degree morphism of the graded modules that commutes with the boundary map.

In detail a chain map $f_\bullet : A_\bullet \longrightarrow B_\bullet$ is a family of homomorphisms $f_n : A_n \longrightarrow B_n$ such that $f_n d_n = d_n f_n$

Definition B.49: A chain complex $A_\bullet$ is **bounded below** if there is an integer N such that $A_n = 0$ for all $n < N$. It is **bounded above** if there is an integer M such that $A_n = 0$ for all $n > M$. It is **bounded** iff it is bounded above and below.

Definition B.50: The **cycles** of a chain complex $A_\bullet$ is the graded submodule $Ker(d)$ which has $Ker(d_n) \subseteq A_n$ in degree n.

Definition B.51: The **boundaries** of a chain complex $A_\bullet$ is the graded submodule $Im(d)$ which has $Im(d_{n+1}) \subseteq A_n$ in degree n.

Because $d \circ d = 0$, the boundaries are all cycles.

Definition B.52: A chain complex is **acyclic** or **exact** iff every cycle is a boundary, i.e., for every integer n we have $Ker(d_n) = Im(d_{n+1})$.

Definition B.53: The **homology** of a chain complex $A_\bullet$ is the graded module $H_\bullet(A_\bullet)$ where $H_n(A_\bullet) = Ker(d_n)/Im(d_{n+1})$.

$A_\bullet$ is acyclic iff $H_\bullet(A_\bullet) = 0$, i.e., $H_n(A_\bullet) = 0$ in every degree n.

If $f_\bullet : A_\bullet \longrightarrow B_\bullet$ is a chain map, then $f(Ker(d_n)) \subseteq Ker(d_n)$, and $f(Im(d_{n+1})) \subseteq Im(d_{n+1})$, so there is an induced homomorphism $H(f_\bullet) : H(A_\bullet) \longrightarrow H(B_\bullet)$.

Definition B.54: When $f_\bullet$ and $g_\bullet$ are two chain maps between the same chain complexes $A_\bullet$ and $B_\bullet$, a **chain homotopy** between $f_\bullet$ and $g_\bullet$ is homomorphism $H_\bullet : A_\bullet \longrightarrow B_\bullet$ of degree 1 such that $f - g = Dd + dD$

Proposition B.43 *If $f_\bullet$ and $g_\bullet$ are chain homotopic chain maps, then $H(f_\bullet) = H(g_\bullet)$.*

B.12 Topological Algebras

Most of the structures studied in Algebra have another interesting variety where the operations live on topological spaces and are continuous, while the homomorphisms are also required to be continuous. A variety of those categories are cataloged in this section.

B.12.1 TopGroup – topological groups

B.12.2 TopAb – Abelian topological groups

B.12.3 TopVect – topological vector spaces

B.12.4 HausdorffTopVect – Hausdorff topological vector spaces

B.13 Analysis

B.13.1 Banach – Banach spaces

The category **Banach** has as objects all complex Banach spaces, and as morphisms all bounded linear transformations between them.

B.13.2 FDBanach – finite dimensional Banach spaces

The category **FDBanach** has as objects all finite dimensional complex Banach spaces, and as morphisms all bounded linear transformations between them. The finite dimensional Banach spaces are sufficiently special and well studied, and the category sufficiently interesting that this category deserves its own study.

B.13.3 BanachAlgebra

Definition B.55: A **Banach algebra** A is an algebra over the real or complex numbers which is also a Banach space such that for all a and b in A we have $||ab|| \leq ||a||||b||$.

B.13.3.1 C*-algebra

B.13.4 Hilbert – Hilbert spaces

The category **Hilbert** has as objects all complex Hilbert spaces, and as morphisms all bounded linear transformations between them.

Definition B.56: An **inner product** on a real or complex vector space is a positive-definite nondegenerate symmetric sequilinear form.

A sesquilinear form on a complex vector space V is a map $< \bullet, \bullet >: VVto\mathbb{C}$ such that for all $u, v, x, y \in V$ and all $a \in \mathbb{C}$

- $< u+v, x+y > = < u, x > + < u, y > + < v, x > + < v, y >$.
- $< au, x > = \overline{a} < u, x >$.
- $< u, ax > = a < u, x >$

This means that for a fixed $v \in V$ the map $x \mapsto \; < v, x >$ is a linear functional on V (i.e. an element of the dual space V^*). Likewise, the map $x \mapsto \; < x, v >$ is a conjugate-linear functional on V.

The form is symmetric when $\forall x, y \in V \;\; < x, y > = \overline{< y, x >}$. This condition implies that $< x, x > \in \mathbb{R}$ for all $x \in V$, because $< x, x > = \overline{< x, x >}$.

The form is nondegenerate provided the induced map from V to the dual space V' is an isomorphism. This means that $\forall y \in V \;, < x, y > = 0$ iff $x = 0$, and if $\phi : V \longrightarrow \mathbb{C}$ is a continuous linear functional, then there exists $v \in V$ so that $\phi = < \bullet, v >$.

Finally the form is positive-definite provided $\forall x \in V \;, < x, x > \geq 0$. (This makes sense because $< x, x > \in \mathbb{R}$ for all $x \in V$.)

Examples

The simplest example of a real inner product space is the space of real numbers with the usual multiplication of numbers as the inner product.

The straight forward generalization of this is to any Euclidean space $\mathbb{R}^n$ with the dot product

$$< (x_1, \cdots, x_n), (y_1, \cdots, y_n) > = \Sigma_{i=1}^{n} x_i y_i = x_1 y_1 + \cdots + x_n y_n$$

as the inner product.

The general form of an inner product on $\mathbb{C}^n$ is given by:

$$< \mathbf{x}, \mathbf{y} > = \mathbf{x} * \mathbf{My}$$

with M any positive-definite matrix, and $x*$ the conjugate transpose of x.

¡x,y¿

B.13.5 FDHilb – finite dimensional Hilbert spaces

B.14 Differential Geometry

B.14.1 Manifold– smooth manifolds

The category **Manifold** has as objects all real finite dimensional smooth, i.e., C^∞, manifolds, and as morphisms all smooth functions between them.

B.14.2 LieGroup– Lie groups

The category **LieGroup** is the category of group objects in the category **Manifold**. Concretely it has as objects all real finite dimensional Lie groups, and as morphisms all smooth homomorphisms between them.

B.15 Algebraic and Analytic Geometry

B.15.1 Sheaf

Definition B.57: Let $(X, \mathcal{T})$ be a topological space, and $\mathscr{C}$ a category. A **presheaf** on X with values in $\mathscr{C}$ is a contravariant functor from $\mathcal{T}$ to $\mathscr{C}$.

Explicitly a presheaf F on the topological space $(X, \mathcal{T})$ has for each open set $U \subseteq X$ an object $F(U)$ of $\mathscr{C}$, and for each inclusion of open sets $V \subseteq U$ a morphism $\rho_{U,V} : F(U) \longrightarrow F(V)$ satisfying the two conditions: (i) $\rho_{U,U} = 1_{F(U)}$ and (ii) for $W \subseteq V \subseteq U$, $\rho_{U,W} = \rho_{V,W}\rho_{U,V}$.

When U is an open subset of X and F is a presheaf on X, then $F(U)$ is called the (object of) sections of F over U, and the morphism $\rho_{U,V}$ is called "restriction from U to V". This terminology comes from the following canonical example of a presheaf.

Let $\pi : E \longrightarrow X$ be any function. For each open set U of X, take $F(U)$ to be the sections of π over U, i.e., $F(U) = \{s : U \longrightarrow E : \pi s = 1_U\}$ and for open sets $V \subseteq U \subseteq X$ take $\rho_{U,V}(s) = s|_V$. That F is a presheaf is immediate. In this situation particularly $F(U)$ is sometimes written as $\Gamma(U, F)$, and the elements of $F(X)$ are called *global sections* of F.

ok, sheaves are in fact defined as coming

B.15.2 RingedSpace

Definition B.58: A **ringed space** is a topological space together with a sheaf of commutative rings on the space. The sheaf is called the **structure sheaf** of the ringed space.

If X is the topological space in a ringed space, the structure sheaf is commonly named $\mathcal{O}_X$ and the ringed space is written as $(X, \mathcal{O}_X)$.

B.15.3 Scheme – algebraic schemes

Schemes are the fundamental objects of modern algebraic geometry. The basic schemes are the affine schemes.

B.16 Unusual Categories

B.17 Cat – small categories

B.18 Groupoid – groupoids

B.19 Structures as Categories

B.19.1 Every set is a category

If S is any set, there is the category that has as objects the elements of S and as morphisms the elements of S with every morphism the identity morphism on itself! Inspired by this observation we say that **discrete categories**are those which have only identity morphisms. So a small discrete category is one that comes from a set as above. Of course if the category is not small, then it doesn't come from a set, but it comes from a class in the same way.

Even more a functor between small discrete categories is really just a function between the objects of the two categories. So the slogan is "sets are small discrete categories".

B.19.2 Every monoid is a category

Recall that a monoid consists of a set M, an associative binary operation on M (conventionally written $\langle m, n\rangle \longrightarrow mn$), and an identity element, 1, for the binary operation. But this is really the same as the definition of a small category with one object.

B.19.3 Monoid of Strings

An interesting example of a monoid is the *monoid of strings* over a given alphabet. If A is any set, the *alphabet*, the monoid of string, A^* consists of all finite sequences of elements of A, including the empty string. The "product" in A^* is concatenation – if $a_1 a_2 \cdots a_n$ and $b_1 \cdots b_m$ are two such strings, their product is $a_1 a_2 \cdots a_n b_1 \cdots b_m$. This product is clearly associative, and the empty string is the identity.

As with any monoid, this may be considered a category with exactly one object. As a category we can ask, does it have products?

B.19.4 Every preorder is a category

If $\langle \underline{P}, \preceq\rangle$ is a preorder, it can equally well be considered as a small category with objects the elements of $\underline{\mathrm{P}}$while there is a unique morphism from $ptoq \in \underline{P}$ iff $p \preceq q$. Note for any two objects p and q there is either exactly one morphism from p to q (if $p \preceq q$), or there are no morphisms from p to q (if $p \preceq q$ is false.)

Composition is what it must be: if $p \longrightarrow q \longrightarrow r$, then $p \preceq q \preceq r$ and so $p \preceq r$ and that gives the composition. The identity morphisms arise because $p \preceq p$.

Exercise B.1. Suppose $\mathscr{P}$ is a small category with the property that for any two objects A and B of $\mathscr{P}$ there is at most one morphism from A to B. Define a relation on the set P of objects of $\mathscr{P}$ by $A \preceq B$ iff there is a morphism from A to B. Demonstrate that P together with this relation is a preorder.

B.19.5 Every topology is a category

Recall that a topology on a set X is a collection, T, of subsets of X subject to

1. $\emptyset, X \in \mathsf{T}$
2. If $X_1, \cdots, X_n$ is any finite collection of elements of T, then $X_1 \cap \cdots \cap X_n \in \mathsf{T}$
3. If $\mathsf{S} \subseteq \mathsf{T}$, then $\bigcup_{S \in \mathsf{S}} S \in \mathsf{T}$

The objects of the category T are the elements of T, and a morphism $X_1 \longrightarrow X_2$ means $X_2 \supseteq X_1$.

This is lumped with posets exactly because we are really just considering a topology as a poset with $X_1 \preceq X_2$ iff $X_2 \supseteq X_1$.

We record this special case exactly because of the connection with sheaves, Grothendieck topologies and topoi as discussed at length in Chapter IX

B.20 Little Categories

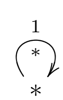
B.20.1 0 – the empty category

0 is the category with no objects and no morphisms. It is interesting and useful much as the empty set. It is the "initial category", i.e., there is a unique (empty) functor from **0** to any category.

The discussion of special morphisms and special objects in **0** is left as an exercise for those with metaphysical leanings.

B.20.2 1 – the one morphism category

1 is the category with one object and one morphism, the identity morphism on the one object, represented by the following diagram.

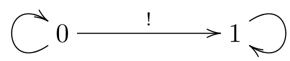


1 is the "final category", i.e., there is a unique functor from any category to **1** – it takes each object to the unique object, and each morphism to the unique morphism.

Of course the one morphism in **1** is an isomorphism, and so is monic and epic, has a retract and a section. Products and sums exists in **1** with all objects being, of course, the one object and all morphisms being the one morphism.

B.20.3 2 – the arrow category

The category **2** is illustrated by

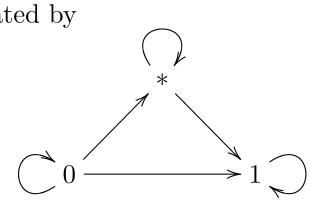


where the two circular arrows are the identity maps. Note that this category has an initial object, 0, a distinct final object, 1, the required morphisms and nothing else. This category has both sums and products: as noted before (I.71) $0 + 0 = 0$, $0 + 1 = 1$, $0 \times 1 = 0$, and $1 \times 1 = 1$.

Exercise B.2. Prove that $1 + 1 = 1$ and $0 \times 0 = 0$ in the category **2**.

B.20.4 3 – the commutative triangle category

The category **3** is illustrated by



Note that this category has an initial object, 0, a distinct final object, 1, and one other object which is not isomorphic to either of those. It has the morphisms required based on those objects, and nothing more.

B.20.5 The parallel arrows category

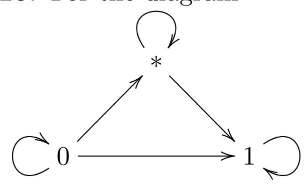
The category $\downdownarrows$ has objects 0 and 1 with their respective identity morphisms, two distinct morphisms from 0 to 1, and nothing else. The interest of this category is nothing more nor less than being the domain for functors whose limits are equalizers.

Appendix C

Solutions of Exercises

C.1 Solutions for Chapter I

Solution for I.1 on p. 10: For the diagram



the problem is to show there is a unique fashion in which this is a category with three objects and five morphisms.

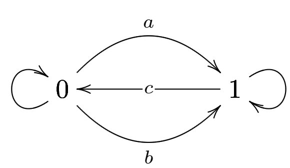
The first thing to note is that if this is a category, then there is at most one morphism between any two objects. So in particular the arrow from each object to itself must be the identity morphism. This in turn specifies the composition of those three morphisms with arrows that start or end at the corresponding node. There are a total of 9 such compositions, three for each node, and they are all consistent with this being a category. Moreover there are no other definitions possible.

There is only one other composition to be defined, that of $0 \longrightarrow * \longrightarrow 1$. Again there is only one morphism from 0 to 1, so that is what the composition must be. Thus all the definition of composition is forced by the uniqueness of the morphisms between objects.

At this stage we see that there is a composition defined respecting the domains and codomains, and there is only one possible way it can be defined. Moreover there are identity morphisms for each object, they are identities for the composition operation and associativity does trivially check whenever at least two of the three morphisms is an identity morphism.

The last detail that must be checked is associativity of compositions involving the composition $0 \longrightarrow * \longrightarrow 1$. But the third morphism is any such composition must be an identity, so all such triple compositions are just $0 \longrightarrow 1$ and so are equal. ∎

Solution for I.2 on p. 10: The arrows in the following diagram *cannot* be all the distinct morphism of a category with two objects and five morphisms.



If this did represent a category, the unique arrows that loop around to 0 and 1 would have to be the identities on the objects 0 and 1. That determines the compositions of those two with the remaining three, a, b and c. So far there are clearly no conflicts. Now ca and cb must be morphisms from 0 to itself and so must be the unique morphism, namely the identity morphism on the object

0. Similarly ac and ab must both be equal to the identity morphism on the object 1. But associativity now presents a contradiction:

$$\begin{aligned} a &= a1 \\ &= a(ca) \\ &= a(cb) \\ &= (ac)b \\ &= 1b \\ &= b \end{aligned}$$

while our basic assumption is that a and b are not equal. ∎

Solution for I.3 on p. 10: To verify that there is a subcategory $\mathscr{E}$ of **Set** consisting of all and only the surjective functions as morphisms, first note that every identity functions is a surjection and so in $\mathscr{E}$. Also if $f : A \longrightarrow B$ and $g : B \longrightarrow C$ are surjections, then so if gf: for each $c \in C$ there exists $b \in B$ with $g(b) = c$ because g is surjective, and there exists $a \in A$ with $f(a) = b$ because f is surjective. Combining these, for each $c \in C$ there exists $a \in A$ with $gf(a) = g(b) = c$ proving that gf is surjective.

There is a multitude of examples showing that $\mathscr{E}$ is not full, but one particularly interesting sample are the empty functions $\emptyset : \emptyset \longrightarrow X$ where X is *any* non-empty set. The function $\emptyset$ is not surjective in this case, i.e., not in $\mathscr{E}$, demonstrating that $\mathscr{E}$ is not full. But more, peering ahead a bit (see Section I.3.2,) this shows that $\mathscr{E}$ does *not* have an initial object.

Solution for I.4 on p. 12: If $f : A \longrightarrow B$ is an isomorphism, then for every object C the induced function $f_* : \mathrm{Hom}(C, A) \longrightarrow \mathrm{Hom}(C, B)$ is a bijection with inverse f_*^{-1}. For given any $g \in \mathrm{Hom}(C, A)$, $f_*(g) = fg$, so $f_*^{-1} f_*(g) = f_*^{-1} f_* g = g$ and for any $h \in \mathrm{Hom}(C, B)$, $f_* f_*^{-1}(h) = f_* f_*^{-1} h = h$

Similarly for every object C $f^* : \mathrm{Hom}(B, C) \longrightarrow \mathrm{Hom}(A, C)$ is a bijection with inverse $(f^{-1})^*$. For given any $h \in \mathrm{Hom}(B, C)$, $f^*(h) = hf$, so $(f^{-1})^* f^*(h) = hf^*(f^{-1})^* = h$ and for any $g \in \mathrm{Hom}(A, C)$, $f^*(f^{-1})^*(h) = h(f^{-1})^* f^* = h$. ∎

Solution for I.5 on p. 12: Here we want to show that if $f : A \longrightarrow B$ is a morphism where for every object C the induced functions

$$f_* : \mathrm{Hom}(C, A) \longrightarrow \mathrm{Hom}(C, B)$$

and

$$f^* : \mathrm{Hom}(B, C) \longrightarrow \mathrm{Hom}(A, C)$$

are bijections, then f is an isomorphism.

This converse to the last exercise is most neatly done in three pieces. First note that if $f : A \longrightarrow B$ has $f_* : \mathrm{Hom}(B, A) \longrightarrow \mathrm{Hom}(B, B)$ surjective, then there is a g in $\mathrm{Hom}(B, A)$ with $f_*(g) = 1_B$, i.e., $fg = 1_B$. (See definition I.17 for more context.)

Second note that if $f^* : \mathrm{Hom}(B,A) \longrightarrow \mathrm{Hom}(A,A)$ is surjective, then there is an $h \in \mathrm{Hom}(B,A)$ with $f^*(h) = 1_A$, i.e., $hf = 1_A$. (See definition I.18 for more context.)

Finally note that $g = 1_A g = (hf)g = h(fg) = h1_B = h$, so we have an inverse for f. ∎

Solution for I.6 on p. 12: When $f : A \longrightarrow B$ is an isomorphism in $\mathscr{C}$ we define a function $\mathscr{C}(A,A) \longrightarrow \mathscr{C}(B,B)$ by $e \in \mathscr{C}(A,A) \mapsto fef^{-1} \in \mathscr{C}(B,B)$. We will show that this function is a monoid isomorphism.

Clearly $f1_Af^{-1} = 1_B$ and $fee'f^{-1} = (fef^{-1})(fe'f^{-1})$, so the function is a monoid homomorphism. And it is an isomorphism because it has the inverse homomorphism given by $g \in \mathscr{C}(B,B)$ maps to $f^{-1}gf \in \mathscr{C}(A,A)$. ∎

Solution for I.7 on p. 13: For any morphism $f : A \longrightarrow B$ we will verify the following:
(a.) f has a section iff f_* always has a section.
(b.) f has a retract iff f^* always has a section.
(c.) f has a retract implies f_* always has a retract.
(d.) f has a section implies f^* always has a retract.

This exercise is really just an elaboration of exercise I.5 as you can easily see by looking at the solution for that exercise.

(a.) f has a section means there is a morphism $s : B \longrightarrow A$ so that $fs = 1_B$. But then $s_* : \mathrm{Hom}(C,B) \longrightarrow \mathrm{Hom}(C,A)$ is a section for f_*, for if $h \in \mathrm{Hom}(C,B)$, then $f_*s_*(h) = fsh = 1_Bh = h$.

Conversely if f_* has a section, then in particular $f_* : \mathrm{Hom}(B,A) \longrightarrow \mathrm{Hom}(B,B)$ is surjective and so there is an element s of $\mathrm{Hom}(B,A)$ with $f_*(s) = 1_B$. But that is just another way of saying that $fs = 1_B$, i.e., that s is a section for f. ∎

(b.) Similarly f has a retract when there is a morphism $r : B \longrightarrow A$ so that $rf = 1_A$. But then $r^* : \mathrm{Hom}(A,C) \longrightarrow \mathrm{Hom}(B,C)$ is a section for f^*, for if $g \in \mathrm{Hom}(A,C)$, then $f^*r^*(g) = rfg = 1_Ag = g$.

Conversely if f^* has a section, then in particular $f^* : \mathrm{Hom}(B,A) \longrightarrow \mathrm{Hom}(A,A)$ is surjective and so there is an element r of $\mathrm{Hom}(B,A)$ with $f^*(r) = 1_A$. But that is just another way of saying that $rf = 1_A$, i.e., that r is a retract for f ∎

I hope you noticed the great similarity between these two results. The two proofs look almost the same, although the order of composition is reversed. This is our first example of duality which is discussed formally in Section II.1, but with many examples like this before then.

(c.) Now if f has a retract r, then $r_* : \mathrm{Hom}(C,B) \longrightarrow \mathrm{Hom}(C,A)$ is a retract for f_*, for if $k \in \mathrm{Hom}(C,A)$, then $r_*f_*(k) = rfk = 1_Ak = k$. ∎

(d.) Similarly if f has a section s, then $s^* : \mathrm{Hom}(A, C) \longrightarrow \mathrm{Hom}(B, C)$ is a retract for f^*, for if $j \in \mathrm{Hom}(B, C)$, then $s^* f^*(j) = jfs = j1_B = j$. ∎

Note again the great similarity between these last two results. The two proofs look almost the same, although the order of composition is reversed. This is another example of duality.

Solution for I.8 on p. 13: This is just like the observation that in a monoid every element that has a right inverse and a left inverse has an inverse because they are equal: $r = r1_B = r(fs) = (rf)s = 1_A s = s$. ∎

Solution for I.9 on p. 13: It's even more interesting to observe that for functions between sets a function has a section exactly if it is surjective, while a function has a retract exactly if it is injective. To see this note that if $f : X \longrightarrow Y$ has a section $s : Y \longrightarrow X$, then for each $y \in Y$, $s(y)$ is an element in X so that $f(s(y)) = y$, i.e., f is surjective. Conversely if f is surjective, then for each $y \in Y$, $f^{-1}(\{y\}) = \{x | f(x) = y\}$ is non-empty. So the axiom of choice allows us to define a function $s : Y \longrightarrow X$ with $s(y) \in f^{-1}(\{x\})$, and that is the desired section. For this reason a section is also sometimes called a *choice map*.

Now if f has a retract and $f(x) = f(x')$, then $x = r(f(x)) = r(f(x')) = x'$ and so f is an injection. Conversely if f is an injection, select some arbitrary element $x_0 \in X$ and define $r : Y \longrightarrow X$ by

$$r(y) = \begin{cases} x & \text{if there exists a (necessarily unique) } x \in X \text{ with } f(x) = y \\ x_0 & \text{otherwise} \end{cases}$$

Then r is the required retract.

Now for an actual solution to the exercise consider the functions

$$s : \{1\} \longrightarrow \{1, 2\}$$

and

$$r : \{1, 2\} \longrightarrow \{1\}$$

with

$$\begin{aligned} s(1) &= 1, \\ r(1) &= 1, \\ r(2) &= 1. \end{aligned}$$

Then r is a retract for s and s is a section for r, but r does not have a retract and s does not have a section. Finally note that $sr : \{1, 2\} \longrightarrow \{1, 2\}$ is neither injective nor surjective and so has neither a retract nor a section. ∎

Solution for I.11 on p. 14: To show that $f : A \longrightarrow B$ is an epimorphism iff f^* is always injective, observe that the definition of f being epic ($gf = hf \Rightarrow$

$g = h$) is, by the definition of f^*, the same as saying ($f^*(g) = f^*(h) \Rightarrow g = h$), which is exactly what it means to says that f^* is injective. ∎

Solution for I.12 on p. 14: We want an example of a category and a morphism $f : A \longrightarrow B$ which is an epimorphism, but where f_* is not always surjective. We will actually see several examples from different categories, but preparatory to that let us see just what this means.

In the solution of exercise I.9 we noted that in the category of sets a function is surjective iff it has a section. But part (a.) of exercise I.7. tells us that f_* always has a sections iff f has a section. Combining those in this context we see that we are simply looking for an epimorphism that does not have a section. This also tells us we will *not* find our example in the category of sets, where morphisms are epic iff they have sections.

In the category of monoids consider the inclusion homomorphism $i : \mathbb{N} \longrightarrow \mathbb{Z}$ with $\mathbb{N}$ the additive monoid of non-negative integers and $\mathbb{Z}$ the additive monoid of all integers. Certainly i is a monomorphism, but more interesting in this context is that it is also an epimorphism! To see this we make a minor digression on monoids.

Let M be a multiplicative monoid with the operation denoted by $*$ (so the identity of M is 1.) If $h : \mathbb{Z} \longrightarrow M$ is a monoid homomorphism, then $h(0) = 1$, and if n is a positive integer, then $1 = h(0) = h(n + (-n)) = h(n) * h(-n)$. And in a monoid, if an element $m \in M$ has an inverse, it is unique: Suppose $l * m = 1 = m * r$, then

$$\begin{aligned} l &= l * 1 \\ &= l * (m * r) \\ &= (l * m) * r \\ &= 1 * r \\ &= r \end{aligned}$$

[Compare this with exercise I.8.]

Now if h and g are two monoid homomorphisms from $\mathbb{Z}$ to M with $hi = gi$, then certainly $h(n) = g(n)$ whenever n is a non-negative integer. But the digression shows that $h(-n) = h(n)^{-1} = g(n)^{-1} = g(-n)$ as well, i.e., h and g are equal. So i is an epimorphism.

But i does *not* have a section for if $s : \mathbb{Z} \longrightarrow \mathbb{N}$ were a section we would have $is(-1) = -1$, i.e., $-1 \in \mathbb{N}$. And this finally tells us that i is an epimorphism in **Monoid** where i_* is not always surjective.

Looking at this concretely can be interesting as well. There is a canonical bijection between the set $\mathbf{Monoid}(\mathbb{Z}, \mathbb{Z})$ and the set $\mathbb{Z}$ because there is for each $n \in \mathbb{Z}$ a unique monoid homomorphism h with $h(1) = n$. But the only monoid homomorphism in $\mathbf{Monoid}(\mathbb{N}, \mathbb{Z})$ take every element of $\mathbb{Z}$ to 0 as 0 is the only invertible element in $\mathbb{N}$. So the function

$$i_* : \mathbf{Monoid}(\mathbb{N}, \mathbb{Z}) \longrightarrow \mathbf{Monoid}(\mathbb{Z}, \mathbb{Z})$$

is very far indeed from being surjective.

$\blacksquare$

In the category **CommutativeRing** (cf. section B.3.4) of commutative rings consider the inclusion homomorphism $i : \mathbb{Z} \longrightarrow \mathbb{Q}$ with $\mathbb{Z}$ the ring of all integers and $\mathbb{Q}$ the field of all rational numbers. Certainly i is a monomorphism, but it is an epimorphism as well! The reason is that every homomorphism of commutative rings with domain $\mathbb{Q}$ is determined by its value on 1. For note that if $h : \mathbb{Q} \longrightarrow R$ is such a ring homomorphism, then $h(n/m) = h(n)h(m)^{-1}$ and so $h(n/m) = h(n)h(m)^{-1} = g(n)g(m)^{-1}$, i.e., $h = g$.

But certainly i has no section $s : \mathbb{Q} \longrightarrow \mathbb{Z}$ for then $s(1/2) \in \mathbb{Z}$ would be a multiplicative inverse for 2 which certainly does not exist! $\blacksquare$

For another example, in the category of topological spaces consider the inclusion $i : \mathbb{Q} \longrightarrow \mathbb{R}$ where $\mathbb{R}$ is the space of real numbers with the usual topology and $\mathbb{Q}$ is the subspace of all rational numbers. Certainly i is a monomorphism, but it is equally well an epimorphism because $\mathbb{Q}$ is dense in $\mathbb{R}$ and it is a general truth that two continuous functions which agree on a dense subspace are equal. But i does not have a section for the most blatant of reasons: the cardinality of $\mathbb{Q}$ is countable, while the cardinality of $\mathbb{R}$ is uncountable. $\blacksquare$

Solution for I.13 on p. 15: We want to verify that in $\mathbb{N}$, the additive monoid of natural numbers considered as a category with one object, every morphism is an epimorphism. So consider any natural number n and its "composition" with two others, j and k. Then we need to verify that $n + j = n + k \Rightarrow j = k$, which certainly qualifies as a well known property of the natural numbers! $\blacksquare$

Solution for I.14 on p. 15: In this exercise we want to verify that for $A = \{a, b\}$ and A^* the monoid of all all finite sequences from A, every morphism in A^* is an epimorphism. So consider any three finite sequences $(c_1, \cdots, c_p)$, $(d_1, \cdots, d_q)$, and $(e_1, \cdots, e_r)$, and suppose $(c_1, \cdots, c_p, d_1, \cdots, d_q) = (c_1, \cdots, c_p, e_1, \cdots, e_r)$. To say that two sequences are equal means there are the same number of terms, and term by term they are equal. So $q = r$ and $d_1 = e_1, \ldots, d_q = e_q$., i.e., $(d_1, \cdots, d_q) = (e_1, \cdots, e_r)$ which is exactly what is needed to show that $(c_1, \cdots, c_p)$ is an epimorphism. [Those of you with a sharp eye may have noticed that this proof is a bit slippery, as we are carefully avoiding the necessary use of induction by casually using $\cdots$ and $\ldots$. And you are correct, but we will continue to leave the details of such inductive proofs to be filled in by the ambitious reader.] $\blacksquare$

Notice that the only information we used about A^* is that it is a monoid of sequences. That A has two elements, or what those particular elements happen to be, is completely irrelevant. So in fact we've really seen that every morphism in a free monoid is an epimorphism.

Solution for I.15 on p. 16: With B the quotient monoid of A^* (as above) by the equivalence relation generated by $\{a^2, ab\}$, we want to verify that b is an epimorphism in B, while a is not.

That a is not an epimorphism is directly from the definition of B, for $aa = ab$, but $a \neq b$.

To see that b is an epimorphism is tedious but elementary; just consider the cases.

Recall that in B every element has a unique canonical form which is one of a^m, b^n, or $b^n a^m$. For $m > 0$, ba^m is certainly not equal to ba^n with $n \neq m$, nor to b^n for any n, and is only equal $b^n a^m$ for $n = 1$. Equally well $bb^n = b^{(n+1)}$ and not to any other canonical form, while $b(b^n a^m) = b^{(n+1)} a^m$ and again is not equal to any other canonical form. ∎

Solution for I.10 on p. 14: Suppose that $f : A \longrightarrow B$ has a section $s : B \longrightarrow A$ so that $fs = 1_B$. Now if $gf = hf$, then $g = g1_B = g(fs) = (gf)s = (hf)s = h(fs) = h1_B = h$ and so f is an epimorphism. ∎

Solution for I.17 on p. 17: In order to show that $f : A \longrightarrow B$ is a monomorphism iff f_* is always injective, observe that the definition of f being monic ($fg = fh \Rightarrow g = h$) is, by the definition of f_*, the same as saying $(f_*(g) = f_*(h) \Rightarrow g = h)$, which is exactly what it means to says that f_* is injective. ∎

Notice this is almost the same as the solution of exercise I.11. The difference between the two being the exchange of epimorphism and monomorphism, f_* and f^* and the order of composition. This is another example of duality which we will discuss formally in Section II.1 after seeing a great many examples.

Solution for I.18 on p. 17: We want an example of a category and a morphism $f : A \longrightarrow B$ which is a monomorphism, but where f^* is not always surjective. We will actually see several examples from different categories, but preparatory to that let us see just what this means.

In the solution of exercise I.9 we noted that in the category of sets a function is surjective iff it has a section. But part (b.) of exercise I.7. tells us that f^* has a section iff f has a retract. Combining those in this context we see that we are simply looking for a monomorphism that does not have a retract. The solution of exercise I.9 also showed that in the category of sets a function is injective iff it has a retract, so this tells us we will not find our example in the category of sets, nor in any other category where morphisms are monic iff they have retracts.

In the category of monoids consider the inclusion homomorphism $i : \mathbb{N} \longrightarrow \mathbb{Z}$ with $\mathbb{N}$ the additive monoid of non-negative integers and $\mathbb{Z}$ the additive monoid of all integers. Certainly i is a monomorphism, but if it has a retract $r : \mathbb{Z} \longrightarrow \mathbb{N}$ with $ri = 1_{\mathbb{N}}$, then $1 = ri(1) = r(1)$ and also $0 = r(0) = r(1 + (-1)) = r(1) + r(-1) = 1 + r(-1)$. But there is no element of $\mathbb{N}$ which added to 1 gives 0, so no such r can exist. ∎

In the category of Abelian groups consider the inclusion homomorphism $i : \mathbb{Z} \longrightarrow \mathbb{Q}$ with $\mathbb{Z}$ the additive group of all integers and $\mathbb{Q}$ the additive group of all rational numbers. Certainly i is a monomorphism, but if it has a retract $r : \mathbb{Q} \longrightarrow \mathbb{Z}$ with $ri = 1_{\mathbb{Z}}$, then again $r(1) = 1$ but also $1 = r(1/2 + 1/2) = r(1/2) + r(1/2)$. But there is no integer which added to itself gives 1, so no such r can exist. ∎

In the category of topological spaces consider the inclusion $i : \mathbb{Q} \longrightarrow \mathbb{R}$ where $\mathbb{R}$ is the space of real numbers with the usual topology and $\mathbb{Q}$ is the subspace of all rational numbers. Certainly i is a monomorphism. But consider a hypothetical retract $r : \mathbb{R} \longrightarrow \mathbb{Q}$ for i. The space $\mathbb{R}$ is connected and the continuous image of a connected space is connected, so the existence of r would imply that $\mathbb{Q}$ is connected which is false indeed. ∎

Solution for I.16 on p. 17: Suppose that $f : B \longrightarrow A$ has a retract $r : A \longrightarrow B$ so that $rf = 1_B$. Now if $fg = fh$, then $g = 1_A g = (rf)g = r(fg) = r(fh) = (rf)h = 1_B h = h$ and so f is a monomorphism. ∎

Notice the great similarity of this solution to that of exercise I.10. The two proofs look almost the same, with the exchanges of retract and section, monic and epi, and the the order of composition. This is another example of duality which we will discuss formally in Section II.1, but we will see many more examples first.

Solution for I.19 on p. 18: Suppose that $f : A \longrightarrow B$ has a retract r and $g : B \longrightarrow C$ has a retract q, then rq is a retract for gf because

$$\begin{aligned} (rq)(gf) &= r(qg)f \\ &= r1_B f \\ &= rf \\ &= 1_A \end{aligned}$$

∎

Solution for I.20 on p. 18: If gf is a retract (of s), then g is a retract of fs because $1 = (gf)h = g(fh)$. ∎

Solution for I.21 on p. 18: Suppose $f : A \twoheadrightarrow B$ and $g : B \twoheadrightarrow C$ are epimorphisms, then gf is an epimorphism as well. To see this just note that if $h_1 gf = h_2 fg$, then $h_1 g = h_2 g$ because f is an epimorphism and then $h_1 = h_2$ because g is an epimorphism. ∎

Solution for I.22 on p. 18: Suppose $f : A \longrightarrow B$ and $g : B \longrightarrow C$ and that gf is an epimorphism, then g is an epimorphism as well. To see this note that if $h_1 g = h_2 g$, then $h_1 gf = h_2 gf$. But as gf is an epimorphism it follows that $h_1 = h_2$ which is what is needed to show that g is an epimorphism. ∎

Solution for I.23 on p. 19: Suppose that $s : A \twoheadrightarrow B$ is an epimorphism and also a section for $r : B \longrightarrow A$. Then $rs = 1_A$, but also $(sr)s = s(rs) = s1_A = s = 1_B s$ and so it follows that $sr = 1_B$, i.e., s is an isomorphism. ∎

Solution for I.24 on p. 19: Suppose that $s_1 : A \longrightarrow B$ and $s_2 : B \longrightarrow C$ are sections of r_1 and r_2 respectively. Then $s_2 s_1$ is a section of $r_1 r_2$: $r_1 r_2 s_2 s_1 = r_1 1_B s_1 = r_1 s_1 = 1_A$. ∎

Solution for I.25 on p. 19: If gf is a section (of r), then f is a section of rg because $1 = r(gf) = (rg)f$. ■

Notice this is almost the same as the solution of exercise I.20. The difference between the two being the exchange of retract and section and the the order of composition. This is another example of duality which we will discuss formally in Section II.1.

Solution for I.26 on p. 19: If gf is defined and both g and f are monomorphisms then so is gf, for suppose that $(gf)e_1 = (gf)e_2$, then $g(fe_1) = g(fe_2)$ and so $fe_1 = fe_2$ (because g is monic) whence $e_1 = e_2$ (because f is monic). And that is just what we need to know that gf is monic. ■

Solution for I.27 on p. 19: Suppose that gf is an monomorphism, then f is a monomorphism as well. for suppose that $fe_1 = fe_2$, then $gfe_1 = gfe_2$ and so $e_1 = e_2$ (because gf is monic), and that is just what we need to see that f is monic. ■

Solution for I.28 on p. 19: Suppose that r is a retract, i.e., there is s with $rs = 1$, and r is also a monomorphism, then also $sr = 1$. For $rs = 1 \Rightarrow srs = s = 1s$ and so, because s is monic, sr = 1. ■

Solution for I.29 on p. 19: This exercise asks for examples of epimorphisms which are not surjective. Of course this only makes sense if the objects of the category under discussion are "sets with structure" and the morphisms are functions of some sort. Without explicit mention we have already provided examples in the solutions given for exercise I.12. Much later we will even be able to explain why this happened.

The three examples are:

1. In **Monoid** (cf. section B.2.3), the inclusion homomorphism $i : \mathbb{N} \longrightarrow \mathbb{Z}$ is an epimorphism, but it is certainly not surjective. ■

2. In **CommutativeRing** (cf. section B.3.4) consider the inclusion homomorphism $i : \mathbb{Z} \longrightarrow \mathbb{Q}$ is an epimorphism, but it is certainly not surjective. ■

3. And in **Top** (cf. section B.9.3) the inclusion $i : \mathbb{Q} \longrightarrow \mathbb{R}$ is an epimorphism, but again it is certainly not surjective. ■

Solution for I.30 on p. 19: Here we are asked for examples of morphisms that are both monic and epic, but not iso. Again previous the solutions provided for earlier exercises have, without explicit mention, given us several. From exercise I.12 we have the following three examples.

1. In **Monoid** (cf. section B.2.3) the inclusion homomorphism $i : \mathbb{N} \longrightarrow \mathbb{Z}$ is both a monomorphism and an epimorphism, but it is certainly not an isomorphism (as i has neither a section nor a retract.) ■

2. In **CommutativeRing** (cf. section B.3.4) the inclusion homomorphism $i : \mathbb{Z} \longrightarrow \mathbb{Q}$ is a monomorphism and an epimorphism, but it is certainly not an isomorphism (as i has neither a section nor a retract.) ∎

3. And in **Top** (cf. section B.9.3) the inclusion $i : \mathbb{Q} \longrightarrow \mathbb{R}$ is both a monomorphism and an epimorphism, but again it is certainly not an isomorphism (as i has neither a section nor a retract.) ∎

We get another family of examples from exercisesI.13 and I.14 for in the monoids $\mathbb{N}$ and $\mathbf{A}^*$ (both considered as categories with one object) all of the elements considered as morphisms are both monomorphisms and epimorphisms, but only the identity is an isomorphism. This same thing is true for the free monoid on any number of generators. ∎

Solution for I.31 on p. 22: Recall that an equivalence relation on a set B is a binary relation $\equiv$ satisfying the three conditions;

1. [reflexive] For all $b \in B$, $b \equiv b$.
2. [symmetry] For all $b, b' \in B$, $b \equiv b' \iff b' \equiv b$.
3. [transitivity] For all $b, b', b'' \in B$, $b \equiv b'$ and $b' \equiv b'' \Rightarrow b \equiv b''$.

The most basic example of an equivalence relation is equality, i.e., $b \equiv b'$ means $b = b'$. This exercise is about a close relative: With $b \equiv_f b'$ defined to mean $f(b) = f(b')$, we want to see that $\equiv_f$ is an equivalence relation. The [identity] follows from identity for equality for surely $f(b)$ is equal to $f(b)$. And [symmetry] follows from symmetry for equality:

$$f(b) = f(b') \iff f(b') = f(b).$$

Finally [transitivity] follows from transitivity for equality:

$$f(b) = f(b') \text{ and } f(b') = f(b'') \Rightarrow f(b) = f(b'').$$

∎

Solution for I.32 on p. 22: Just as with any equivalence relation, for the equivalence relation $\equiv_f$ on the set B we have the quotient $p : B \longrightarrow B/\equiv_f$ where $B/\equiv_f$ is the set of all equivalence classes. In this case we also get a unique induced function $\overline{f} : B/\equiv_f \longrightarrow A$ with $f = \overline{f}p$. Indeed the definition of $\overline{f}$ is exactly that $\overline{f}(\overline{b}) = f(b)$ whenever $\overline{b}$ is one of the equivalence classes in $B/\equiv_f$, i.e., $\overline{f}(p(b)) = f(b)$. And what must be checked is that this is actually a definition – does $p(b) = p(b')$ imply $f(b) = f(b')$? But $p(b) = p(b')$ exactly if $b \equiv_f b'$ which is just the same as $f(b) = f(b')$.

Of course this also says that $\overline{f}$ is injective. And if f is surjective, then $\overline{f}$ is equally well surjective, and so is is a bijection. ∎

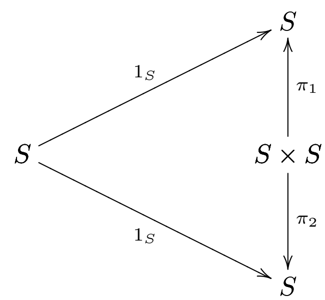
Solution for I.33 on p. 26: This exercise is about products in the category, $\mathscr{E}$, of sets and surjections.

The first thing we want to note is that if E_1 and E_2 have a product in $\mathscr{E}$, it must actually be a product in **Set**. To see this let $(E_1 \times E_2,\ \pi_1,\ \pi_2)$ be a product in **Set**, and let $(P,\ p_1,\ p_2)$ be a product of the same two objects in $\mathscr{E}$. If E_1 and E_2 are not empty, then π_1 and π_2 are surjections and so morphisms in $\mathscr{E}$. So there is a unique morphism $\langle \pi_1, pi_2 \rangle : E_1 \times E_2 \longrightarrow P$ in $\mathscr{E}$. There is also a unique morphism $\langle p_1, p_2 \rangle : P \longrightarrow E_1 \times E_2$ in **Set**. But then $\langle \pi_1, \pi_2 \rangle \langle p_1, p_2 \rangle : P \longrightarrow P$ is equal to 1_P and

$$\langle p_1, p_2 \rangle \langle \pi_1, \pi_2 \rangle : E_1 \times E_2 \longrightarrow E_1 \times E_2$$

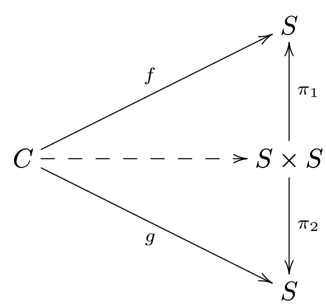
is equal to $1_{E_1 \times E_2}$, the proof being essentially the same as the proof of Proposition I.1. So $(E_1 \times E_2,\ \pi_1,\ \pi_2)$ is a product in $\mathscr{E}$.

Now consider $S = \{0, 1\}$ as an object of $\mathscr{E}$. If S has a product with itself in $\mathscr{E}$, the product must be $(S \times S,\ \pi_1,\ \pi_2)$, the usual product in **Set**. But look at the diagram



and note that the *unique* function $\Delta : S \longrightarrow S \times S$ which makes this diagram commute is *not* a surjection. So there is *NO* morphism from S to $S \times S$ in $\mathscr{E}$ that will make this diagram commute, so $(S \times S,\ \pi_1,\ \pi_2)$ is *not* a product of S with itself in $\mathscr{E}$, and so, as we just saw above, there is no product of S with itself in $\mathscr{E}$.

Considering



notice that are many objects and maps C, f and g in $\mathscr{E}$ where there is a morphism $C \longrightarrow S \times S$ in $\mathscr{E}$ making the diagram commute. And of course this morphism is necessarily the unique morphism in **Set** which makes the diagram commutes because $(S \times S,\ \pi_1,\ \pi_2)$ is a product in **Set**. ∎

Solution for I.34 on p. 26: The method is to show that $\Pi_{i=1}^{n-1} A_i \times A_n$ (with suitable morphisms) is a product of $A_1, \cdots, A_n$. The family of morphisms is

$$\pi_1' = \Pi_{i=1}^{n-1} A_i \times A_n \xrightarrow{\pi_1} \Pi_{i=1}^{n-1} A_i \xrightarrow{\pi_1} A_1,$$
$$\vdots$$
$$\pi_j' = \Pi_{i=1}^{n-1} A_i \times A_n \xrightarrow{\pi_1} \Pi_{i=1}^{n-1} A_j \xrightarrow{\pi_j} A_j,$$
$$\vdots$$
$$\pi_{n-1}' = \Pi_{i=1}^{n-1} A_i \times A_n \xrightarrow{\pi_1} \Pi_{i=1}^{n-1} A_{n-1} \xrightarrow{\pi_{n-1}} A_{n-1}$$

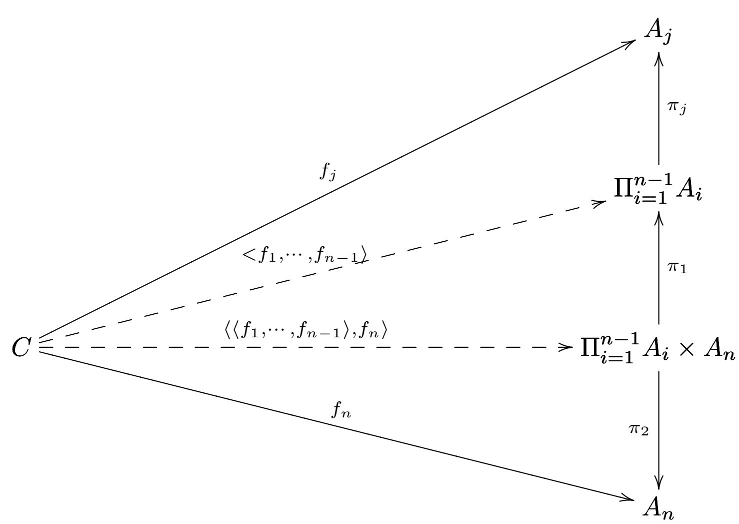
and $\pi_n' = \Pi_{i=1}^{n-1} A_i \times A_n \xrightarrow{\pi_2} A_n$.

Now given a family $f_i : C \longrightarrow A_i$ of morphisms we get

$$\langle f_1, \cdots, f_{n-1} \rangle : C \longrightarrow \Pi_{i=1}^{n-1} A_i$$

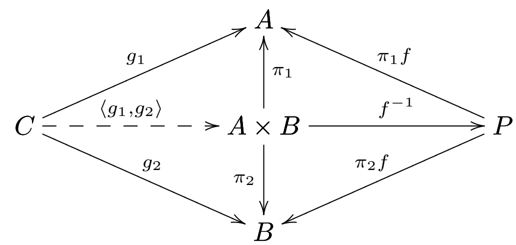
This together with f_n gives us $f = \langle \langle f_1, \cdots, f_{n-1} \rangle, f_n \rangle$ as the unique morphism from C to $\Pi_{i=1}^{n-1} A_i \times A_n$ such that $\pi_j' f = f_j$ for $j = 1, \cdots, n$. Proposition I.1 now applies to say that $\Pi_{i=1}^{n-1} A_i \times A_n$ is canonically isomorphic to $\Pi_{i=1}^{n} A_i$. ▮

This is well summarized in the commutative diagram:



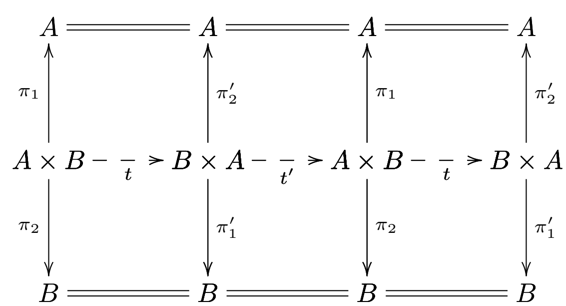
where j is in the range $1, \cdots, n-1$ and the dotted arrows are those whose *unique* existence make the diagram commute.

Solution for I.35 on p. 27: The proof that $\langle P, \pi_1 f, \pi_2 f \rangle$ is also a product is neatly captured in the following commutative diagram:



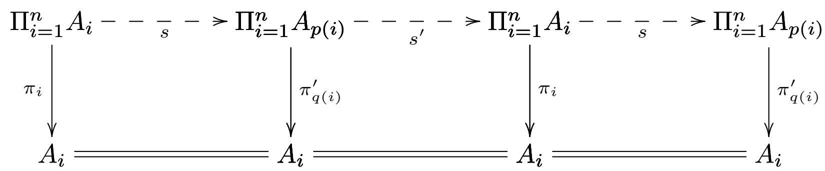
If g_1 and g_2 are arbitrary morphisms from some object C to A and B respectively, then $\langle g_1, g_2\rangle : C \longrightarrow A \times B$ is the unique morphism making the right side of the diagram commute. As f is an isomorphism, there is it's unique inverse f^{-1} as shown. Now $f^{-1}\langle g_1, g_2\rangle : C \longrightarrow P$ with $\pi_1 f f^{-1}\langle g_1, g_2\rangle = \pi_1\langle g_1, g_2\rangle = g_1$ and $\pi_2 f f^{-1}\langle g_1, g_2\rangle = \pi_2\langle g_1, g_2\rangle = g_2$. Further if $g : C \longrightarrow P$ is any morphism with $g\pi_1 f = g_1$ and $g\pi_2 f = g_2$, then $fg : C \longrightarrow A \times B$ must be equal to $\langle g_1, g_2\rangle$ and so $g = (f^{-1}f)g = f^{-1}(fg) = f^{-1}\langle g_1, g_2\rangle$. So $f^{-1}\langle g_1, g_2\rangle$ is the *unique* such morphism and so $\langle P, \pi_1 f, \pi_2 f\rangle$ is a product of A and B.
∎

Solution for I.36 on p. 27: To prove that $t : A \times B \xrightarrow{\cong} B \times A$ is just a matter of looking carefully at the following commutative diagram.



Because $A \times B$ and $B \times A$ are products, t and t' exist and are uniquely defined by the requirement that this diagram commutes. But then $\pi_1 t' t = \pi_1$ and $\pi_2 t' t = \pi_2$. As $1_{A\times B}$ is the unique morphism satisfying that pair of relations, we see that $t't = 1_{A\times B}$. The same argument using the four squares on the right shows $tt' = 1_{B\times A}$. ∎

Solution for I.37 on p. 27: With p an n-permutation with inverse q, look at the following family of commutative diagrams (with i varying from 1 to n):



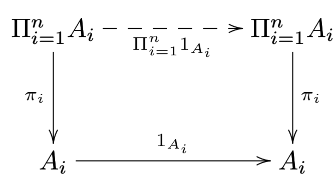
Note that s and s' as morphisms into products are *defined* by the fact that these diagrams commute. But then $s's$ must be the identity on $\Pi_{i=1}^n A_i$ and ss' must be the identity on $\Pi_{i=1}^n A_{p(i)}$ as the identities also make the relevant diagrams commute. ∎

Solution for I.38 on p. 27: No! Projections need not be epimorphisms. The simplest example is $\emptyset \times \{*\} \longrightarrow \{*\}$ in the category of sets. $\emptyset \times \{*\} = \emptyset$ as indeed $\emptyset \times A = \emptyset$ for every set A, and certainly the map from the empty set to a one element set is not surjective. ∎

Solution for I.39 on p. 28: In the category of sets, if $f_1 : A_1 \longrightarrow B_1$ and $f_2 : A_2 \longrightarrow B_2$ are two functions, then what is $(f_1 \times f_2)(a_1, a_2)$?

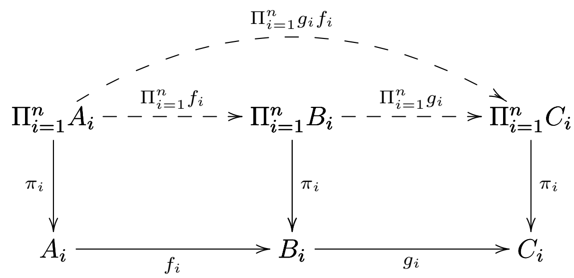
Writing $(f_1 \times f_2)(a_1, a_2) = (b_1, b_2)$, we have $b_1 = \pi_1(f_1 \times f_2)(a_1, a_2)$. But $\pi_1(f_1, f_2) = f_1\pi_1$ by the definition of $f_1 \times f_2$. So $b_1 = f_1\pi_1(a_1, a_2) = f_1(a_1)$. The same argument with 2 replacing 1 shows that $b_2 = f_2(a_2)$. So $(f_1 \times f_2)(a_1, a_2) = (f_1(a_1), f_2(a_2))$. ∎

Solution for I.40 on p. 28: To show that $\Pi_{i=1}^n 1_{A_i} = 1_{\Pi A_i}$, note that, by definition, $\Pi_{i=1}^n 1_{A_i}$ is the unique morphism that fills in this family of commutative diagrams:



But $1_{\Pi A_i}$ fills it in equally well, ergo, they are equal. ∎

Solution for I.41 on p. 28: With the possible exception of the equality of the top arc with the composition beneath it, the following family of diagrams commute.



But by the definition of a product that immediately implies the desired equality:

$$\Pi_{i=1}^n g_i \Pi_{i=1}^n f_i = \Pi_{i=1}^n g_i f_i.$$

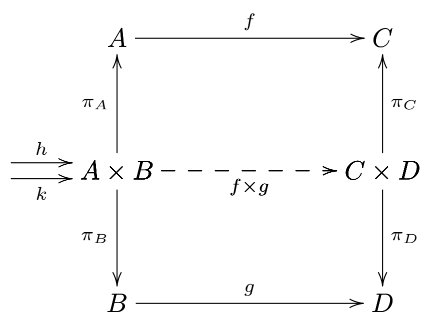
∎

Solution for I.42 on p. 28: If $f : A \longrightarrow C$ and $g : B \longrightarrow D$ have retracts f' and g' respectively, then $ff' = 1_C$ and $gg' = 1_D$ and so

$$\begin{aligned} (f \times g)(f' \times g') &= (ff' \times gg') \\ &= (1_C \times 1_D) \\ &= 1_{C \times D} \end{aligned}$$

where we are using exercises I.40 and I.41. ∎

Solution for I.43 on p. 28: Suppose f and g are monomorphisms, and we have the following commutative diagram:



In particular $(f \times g)h = (f \times g)k$. Then

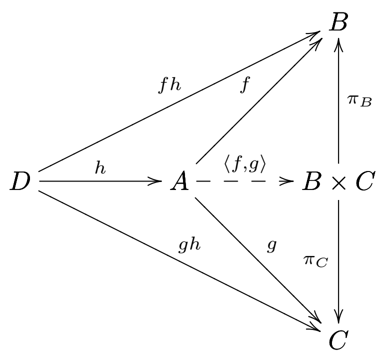
$$f\pi_A h = \pi_C(f \times g)h = \pi_C(f \times g)k = f\pi_A k$$
$$g\pi_B h = \pi_D(f \times g)h = \pi_D(f \times g)k = g\pi_B k$$

As f and g are monomorphisms, $\pi_A h = \pi_A k$ and $\pi_B h = \pi_B k$ whence we see that $h = k$. So $f \times g$ is a monomorphism. ∎

Solution for I.44 on p. 28: If f and g have sections f' and g' respectively, then $f' \times g'$ is a section for $f \times g$. This follows by combining exercises I.41 and I.40. For $(f \times g)(f' \times g') = ff' \times gf' = 1 \times 1 = 1$. ∎

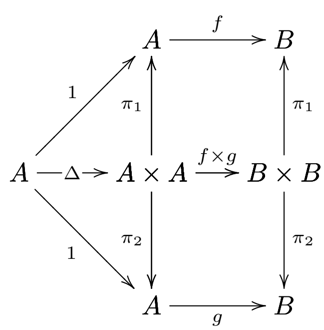
Compare this with the solution to exercise I.42. The two proofs look almost the same, although the order of composition is reversed. Notice that this is *not* an example of duality, because both involve products rather than interchanging products and sums. This proof is dual to that in the solution of exercise I.58 because that is the dual theorem.

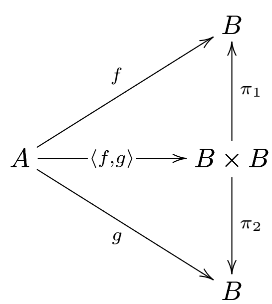
Solution for I.45 on p. 29: The method to verify that $\langle f, g\rangle h = \langle fh, gh\rangle$ is what should by now be very familiar. $\langle fh, gh\rangle$ is the unique morphism such that $\pi_B\langle fh, gh\rangle = fh$ and $\pi_C\langle fh, gh\rangle = gh$, but $\langle f, g\rangle h$ has that same property and so they are equal.



∎

Solution for I.46 on p. 29: To verify that $\langle f, g\rangle = (f \times g)\Delta$ stare at these two commutative diagrams and note that $\pi_i(f \times g)\Delta = f1 = f$ (for $i = 1, 2$) which is exactly the defining property of $\langle f, g\rangle$.





∎

Solution for I.47 on p. 29: To compute $\Delta(x)$ for any x in a set X we just note that $\pi_1(\Delta(x)) = 1_X(x) = x$ and $\pi_2(\Delta(x)) = 1_X(x) = x$, so $\Delta(x) = (x, x)$. ∎

Solution for I.48 on p. 29: To compute $\Delta(a)$ for any $a \in A$ with A an object in **Ab** we just note that $\pi_1(\Delta(a)) = 1_A(a) = a$ and $\pi_2(\Delta(a)) = 1_A(a) = a$, so $\Delta(a) = (a, a)$. ∎

Solution for I.49 on p. 32: For any family of two or more objects, $A_1, \cdots, A_n$, in $\mathscr{C}$ prove that $\Sigma_{i=1}^{n} A_i$ is isomorphic to $\Sigma_{i=1}^{n-1} A_i + A_n$.

The method is to show that $\Sigma_{i=1}^{n-1} A_i + A_n$ (with suitable morphisms) is a sum of $A_1, \cdots, A_n$. The family of morphisms is

$$\iota'_1 = \Sigma_{i=1}^{n-1} A_i + A_n \xleftarrow{\iota_1} \Sigma_{i=1}^{n-1} A_i \xleftarrow{\iota_1} A_1,$$

$$\vdots$$

$$\iota'_j = \Sigma_{i=1}^{n-1} A_i + A_n \xleftarrow{\iota_1} \Sigma_{i=1}^{n-1} A_j \xleftarrow{\iota_j} A_j,$$

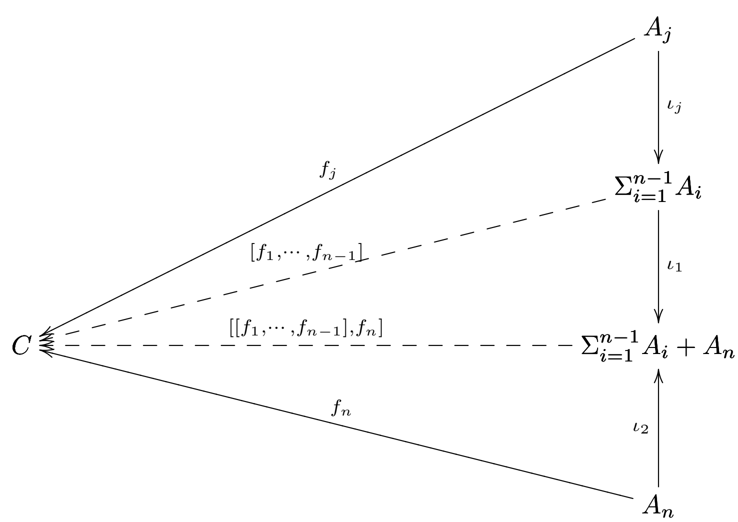
$$\vdots$$

$$\iota'_{n-1} = \Sigma_{i=1}^{n-1} A_i + A_n \xleftarrow{\iota_1} \Sigma_{i=1}^{n-1} A_{n-1} \xleftarrow{\iota_{n-1}} A_{n-1}$$

and $\iota'_n = \Sigma_{i=1}^{n-1} A_i + A_n \xleftarrow{\iota_2} A_n$.

Now given a family $f_i : A_i \longrightarrow C$ of morphisms we get $[f_1, \cdots, f_{n-1}] : \Sigma_{i=1}^{n-1} A_i \longrightarrow C$. This together with f_n gives us $f = [[f_1, \cdots, f_{n-1}], f_n]$ as the unique morphism from $\Sigma_{i=1}^{n-1} A_i + A_n$ to C such that $f\iota'_j = f_j$ for $j = 1, \cdots, n$. Proposition I.2 now applies to say that $\Sigma_{i=1}^{n-1} A_i + A_n$ is canonically isomorphic to $\Sigma_{i=1}^{n} A_i$. ∎

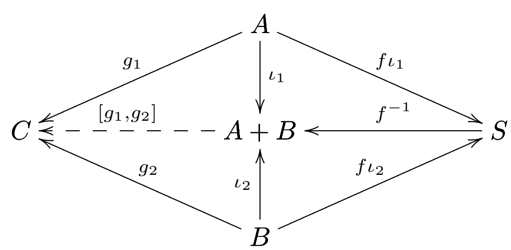
This is well summarized in the commutative diagram:



where j is in the range $1, \cdots, n-1$ and the dotted arrows are those whose *unique* existence make the diagram commute.

Compare this solution with the solution for exercise I.34 on page 26. The two were written carefully to make it clear that each can be transformed into the other by "reversing the arrows" and exchanging sums and products. This is an example of duality which is discussed formally in Section II.1.

Solution for I.50 on p. 32: The proof that $\langle S,\, f\iota_1,\, f\iota_2\rangle$ is also a sum is neatly captured in the following commutative diagram:



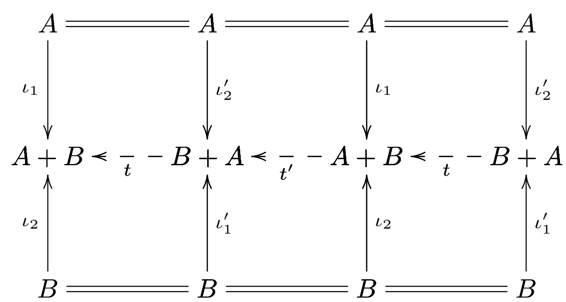
If g_1 and g_2 are arbitrary morphisms to some object C, then $[g_1, g_2] : A + B \longrightarrow C$ is the unique morphism making the left side of the diagram commute. As f is an isomorphism, there is it's unique inverse f^{-1} as shown. Now $[g_1, g_2]f^{-1} : S \longrightarrow C$ with $\iota_1 f f^{-1}[g_1, g_2] = \iota_1[g_1, g_2] = g_1$ and $\iota_2 f f^{-1}[g_1, g_2] = \iota_2[g_1, g_2] = g_2$. Further if $g : S \longrightarrow C$ is any morphism with $gf\iota_1 = g_1$ and $gf\iota_2 = g_2$, then $gf : A + B \longrightarrow$ must be equal to $[g_1, g_2]$ and so

$$\begin{aligned} g &= g(ff^{-1}) \\ &= gff^{-1} \\ &= [g_1, g_2]f^{-1}. \end{aligned}$$

Thus $[g_1, g_2]f^{-1}$ is the *unique* such morphism, and so $\langle S,\, \iota_1 f,\, \iota_2 f\rangle$ is a sum of A and B.

Compare this solution with the solution for exercise I.35 on page 27. The two were written carefully to make it clear that each can be transformed into the other by "reversing the arrows" and exchanging sums and products. This is an example of duality which is discussed formally in Section II.1.

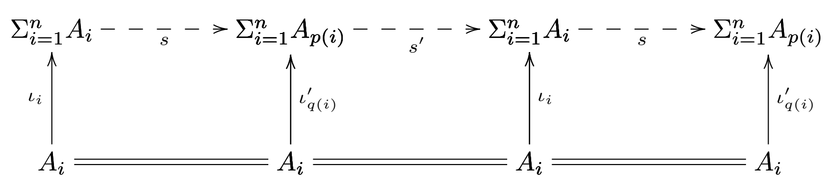
Solution for I.51 on p. 32: To prove that $A + B \cong B + A$ is just a matter of looking carefully at the following commutative diagram.



Because $A + B$ and $B + A$ are sums, t and t' exist and are uniquely defined by the requirement that this diagram commutes. But then $tt'\iota_1 = \iota_1$ and $tt'\iota_2 = \iota_2$. As 1_{A+B} is the unique morphism satisfying that pair of relations, we see that $tt' = 1_{A+B}$. The same argument using the four squares on the right shows $t't = 1_{B+A}$. ∎

Compare this solution with the solution for exercise I.36 on page 27. The two were written carefully to make it clear that each can be transformed into the other by "reversing the arrows" and exchanging sums and products. This is an example of duality which is discussed formally in Section II.1.

Solution for I.52 on p. 32: With p an n-permutation with inverse q, look at the following family of commutative diagrams (with i varying from 1 to n):



Note that s and s' as morphisms from sums are *defined* by the fact that these diagrams commute. But then ss' must be the identity on $\Sigma_{i=1}^n A_i$ and $s's$ must be the identity on $\Sigma_{i=1}^n A_{p(i)}$ as the identities also make the relevant diagrams commute. ∎

Compare this solution with the solution for exercise I.37 on page 27. The two were written carefully to make it clear that each can be transformed into the other by "reversing the arrows" and exchanging sums and products. This is an example of duality which is discussed formally in Section II.1.

Solution for I.53 on p. 32: In a great many of the familiar categories, for example **Set**, **Group**, **Ab**, **Top**, etc., every injection into a sum is a monomorphism, but this is not universally true. Perhaps the simplest familiar example where this is *not* true is **CommutativeRing**, the category of commutative rings. As discussed in section B.3.4 of the Catalog of Categories, the sum in this category is the tensor product. In particular then we have the odd feature that the sum of two non-zero rings may be zero with the simplest example being $\mathbb{Z}_2 \otimes \mathbb{Z}_3$.

In any tensor product $A \otimes B$ we have

$$(a_1 + a_2) \otimes (b_1 + b_2) = a_1 \otimes b_1 + a_1 \otimes b_2 + a_2 \otimes b_1 + a_2 \otimes b_2.$$

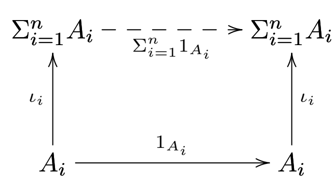
So in $\mathbb{Z}_2 \otimes \mathbb{Z}_3$ we have

$$\begin{aligned} 1 &= 1 \otimes 1 \\ &= (1+1+1) \otimes 1 \\ &= 1 \otimes 1 + 1 \otimes 1 + 1 \otimes 1 \\ &= 1 \otimes (1+1+1) \\ &= 1 \otimes 0 \\ &= 0 \end{aligned}$$

Now look at the two ring homomorphisms E_0 and E_1 from $\mathbb{Z}[X]$ to $\mathbb{Z}_2$ with $E_0(p(X)) = p(0) \bmod 2$ and $E_1(p(X)) = p(1) \bmod 2$. They are certainly not equal, but $\iota_1 E_0 = \iota_1 E_1$. So ι_1 is not a monomorphism. ∎

Looking ahead to dual categories (see Section II.1) we also see that the easiest counter-example is the counter-example in exercise I.38 interpreted in the dual category of the category of sets where the projection that is not an epimorphism in **Set** gives an injection that is not a monomorphism in **Set**$^{\text{op}}$. ∎

Solution for I.54 on p. 33: To show that $\Sigma_{i=1}^n 1_{A_i} = 1_{\Sigma A_i}$, note that, by definition, $\Sigma_{i=1}^n 1_{A_i}$ is the unique morphism that fills in this family of commutative diagrams:

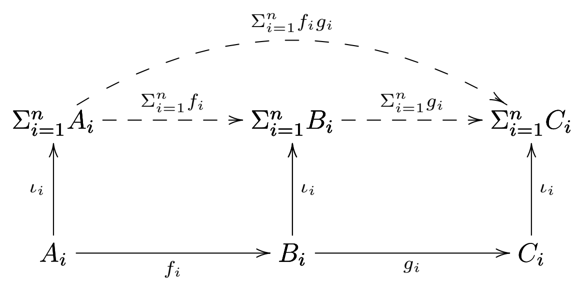


But $1_{\Sigma A_i}$ fills it in equally well, ergo, they are equal. ∎

Compare this solution with the solution for exercise I.40 on page 28. The two were written carefully to make it clear that each can be transformed into the other by "reversing the arrows" and exchanging sums and products. This is an example of duality which is discussed formally in Section II.1.

Solution for I.55 on p. 33: Consider the families of morphisms $f_i : A_i \longrightarrow B_i$ and $g_i : B_i \longrightarrow C_i$. Verify that $\Sigma_i g_i \Sigma_i f_i = \Sigma_i g_i f_i$

With the possible exception of the equality of the top arc with the composition beneath it, the following family of diagrams commute.

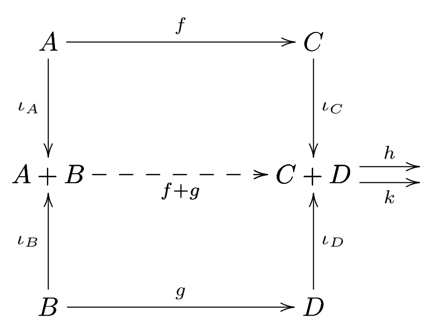


But by the definition of a sum that immediately implies the desired equality. ∎

Compare this solution with the solution for exercise I.41 on page 28. The two were written carefully to make it clear that each can be transformed into the other by "reversing the arrows", including exchanging sums and products. This is an example of duality which is discussed formally in Section II.1.

Solution for I.56 on p. 33: If $f : A \longrightarrow C$ and $g : B \longrightarrow D$ have sections f' and g' respectively, then $ff' = 1_A$ and $gg' = 1_B$, so $(f+g)(f'+g') = (ff' + gg') = 1_A + 1_C = 1_(A+B)$ and so $f' + g'$ is a section for $f + g$. Notice we are using exercises I.54 and I.55. ∎

Solution for I.57 on p. 33: Suppose f and g are epimorphisms, and we have the following commutative diagram:



In particular $h(f + g) = k(f + g)$. Then

$$h\iota_C f = h(f+g)\iota_A = k(f+g)\iota_A = k\iota_C f$$
$$h\iota_D g = h(f+g)\iota_B = k(f+g)\iota_B = k\iota_D g$$

As f and g are epimorphisms, $h\iota_C = k\iota_C$ and $h\iota_D = k\iota_D$ whence we see that $h = k$. So $f + g$ is an epimorphism. ∎

Compare this solution with the solution for exercise I.43 on page 28. The two were written carefully to make it clear that each can be transformed into the other by "reversing the arrows", exchanging sums and products and interchanging "monomorphism" and "epimorphism". This is an example of duality which is discussed formally in Section II.1.

Solution for I.58 on p. 33: Suppose that $f : A \longrightarrow C$ and $g : B \longrightarrow D$ have retracts f' and g' respectively. Then $f'f = 1_C$ and $g'g = 1_D$, so $(f' + g')(f + g) = (f'f + g'g) = 1_C + 1_D = 1_(C + D)$ and so $f' + g'$ is a retract for $f + g$. Notice we are using exercises I.54 and I.55. ∎

Solution for I.59 on p. 33: In most of the familiar categories, for example the **Set**, **Group**, **Ab**, **Top**, etc., the sum of two monomorphisms is always a monomorphism. Perhaps the simplest example where this is *not* true is **CommutativeRing**, the category of commutative rings.

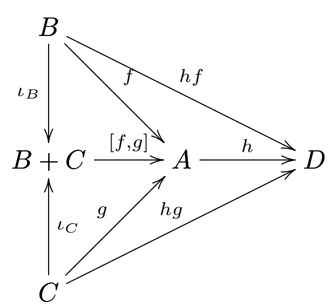
As discussed in section B.3.4 of the Catalog of Categories, the sum in this category is the tensor product. In particular then we have the odd feature that the sum of two non-zero rings may be zero with one simple example being $\mathbb{Z}_2 \otimes \mathbb{Q}$.

In any tensor product $A \otimes B$ we have $(a_1 + a_2) \otimes (b_1 + b_2) = a_1 \otimes \mathbf{b}_1 + a_1 \otimes \mathbf{b}_2 + a_2 \otimes \mathbf{b}_1 + a_2 \otimes \mathbf{b}_2$. So in $\mathbb{Z}_2 \otimes \mathbb{Q}$ we have

$$\begin{aligned} 1 &= 1 \otimes 1 \\ &= 1 \otimes (1/2 + 1/2) \\ &= 1 \otimes 1/2 + 1 \otimes 1/2 \\ &= (1 + 1) \otimes 1/2 \\ &= 0 \otimes 1/2 \\ &= 0 \end{aligned}$$

Now look at $1_{\mathbb{Z}_2} \otimes i$ with $i : \mathbb{Z} \longrightarrow \mathbb{Q}$ the inclusion. Both of these are monomorphisms, but $\mathbb{Z}_2 \otimes \mathbb{Z} \cong \mathbb{Z}_2$, while $\mathbb{Z}_2 \otimes \mathbb{Q} \cong 0$, so $1_{\mathbb{Z}_2} \otimes i$ is not a monomorphism. ∎

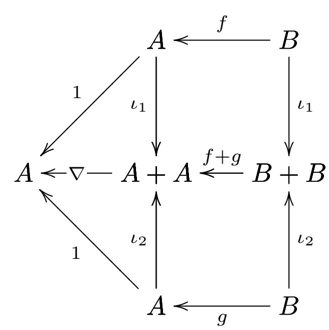
Solution for I.60 on p. 34: The method to verify that $h[f, g] = [hf, hg]$ will be boringly familiar by now. $[hf, hg]$ is the unique morphism such that $[hf, hg]\iota_B = hf$ and $[fh, gh]\iota_C = hg$, but $h[f, g]$ has that same property and so they are equal.

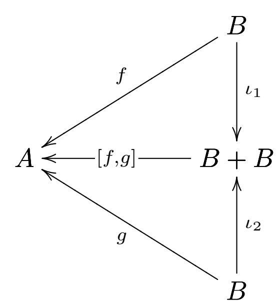


∎

Again, compare this solution with the solution for exercise I.45 on page 29. The two were written carefully to make it clear that each can be transformed into the other by “reversing the arrows”, including exchanging sums and products. And again this is an example of duality which is discussed formally in Section II.1.

Solution for I.61 on p. 34: To verify that $[f, g] = \nabla(f + g)$ stare at these two commutative diagrams and note that $\nabla(f + g)\iota_i = 1f = f$ (for $i = 1, 2$) which is exactly the defining property of $[f, g]$.





∎

Compare this solution with the solution for exercise I.46 on page 29. The two were written carefully to make it clear that each can be transformed into the other by “reversing the arrows”, including exchanging sums and products. And yet again this is an example of duality which is discussed formally in Section II.1.

Solution for I.62 on p. 35: From the definition, $(f_0 + f_1)\iota_1(a_0) = \iota_0 f_0(a_0)$ and $(f_0 + f_1)\iota_1(a_1) = \iota_1 f_1(a_1)$, i.e., $(f_1 + f_2)(a_0, 0) = (f_0(a_0), 0)$ and $(f_0 + f_1)(a_1, 1) = (f_1(a_1), 1)$. But in the category of sets $A_0 + A_1$ can be taken as $A_0 \times \{0\} \cup A_1 \times \{1\}$ ∎

Solution for I.63 on p. 35: Verification that $(A \times B,\ \iota_A(a) = (a, 0),\ \iota_B(b) = (0, b))$ is indeed a sum of A and B in **Ab** has several easy parts. The first is the trivial observation that ι_A and ι_B are actually group homomorphisms. The next is to check that for any two homomorphisms $f : A \longrightarrow C$ and $g : B \longrightarrow C$, $[f, g]$ defined by $[f, g](a, b) = f(a) + g(b)$ is indeed a group homomorphism:

$$\begin{aligned}
[f,g]((a_1,b_1)+(a_2,b_2)) &= [f,g](a_1+a_2,b_1+b_2)\\
&= f(a_1+a_2)+g(b_1+b_2)\\
&= f(a_1)+f(a_2)+g(b_1)+g(b_2)\\
&= f(a_1)+g(b_1)+f(a_2)+g(b_2)\\
&= [f,g](a_1,b_1)+[f,g](a_2,b_2)
\end{aligned}$$

and $[f,g](0,0) = f(0) + g(0) = 0 + 0 = 0$.

Finally we must check that if $h : A \times B \longrightarrow C$ is any other group homomorphism such that $h\iota_A = f$ and $h\iota_B = g$, then $h = [f,g]$:

$$\begin{aligned}
h(a,b) &= h((a,0)+(0,b))\\
&= h(a,0)+h(0,b)\\
&= h\iota_A(a)+h\iota_B(b)\\
&= f(a)+g(b)\\
&= [f.g](a,b)
\end{aligned}$$

∎

Solution for I.64 on p. 35: To compute $\nabla(a_1,a_2)$ for any $(a_1,a_2) \in A \oplus A$ with A any object in **Ab**, note that $(a_1,a_2) = (a_1,0)+(0,a_2) = \iota_1(a_1)+\iota_2(a_2)$ and so $\nabla(a_1,a_2) = \nabla(\iota_1(a_1)+\iota_2(a_2)) = \nabla(\iota_1(a_1))+\nabla(\iota_2(a_2)) = 1_A(a_1)+1_A(a_2) = a_1+a_2$.

Solution for I.65 on p. 35: From the definition, $(f_1+f_2)\iota_1(a_1) = \iota_1 f_1(a_1)$ and $(f_1+f_2)\iota_2(a_2) = \iota_2 f_2(a_2)$, i.e., $(f_1+f_2)(a_1,0) = (f_1(a_1),0)$ and $(f_1+f_2)(0,a_2) = (0,f_2(a_2))$. And as $(a_1,a_2) = (a_1,0)+(0,a_2)$ and f_1+f_2 is a homomorphism we have $(f_1+f_2)(a_1,a_2) = (f_1(a_1),f_2(a_2))$. ∎

Solution for I.66 on p. 35: Verification that $(A \times B,\ \iota_A(a) = (a,0),\ \iota_B(b) = (0,b))$ is indeed a sum of A and B in **Vect** has several easy parts. The first is the trivial observation that ι_A and ι_B are actually linear transformations. The next is to check that for any two linear transformations $f : A \longrightarrow C$ and $g : B \longrightarrow C$, $[f,g]$ defined by $[f,g](a,b) = f(a)+g(b)$ is indeed a linear transformation:

$$\begin{aligned}
[f,g]((a_1,b_1)+(a_2,b_2)) &= [f,g](a_1+a_2,b_1+b_2)\\
&= f(a_1+a_2)+g(b_1+b_2)\\
&= f(a_1)+f(a_2)+g(b_1)+g(b_2)\\
&= f(a_1)+g(b_1)+f(a_2)+g(b_2)\\
&= [f,g](a_1,b_1)+[f,g](a_2,b_2)
\end{aligned}$$

$$\begin{aligned}[f,g](\lambda(a,b)) &= [f,g](\lambda a, \lambda b) \\ &= f(\lambda a) + f(\lambda b) \\ &= \lambda f(a) + \lambda g(b) \\ &= \lambda(f(a) + g(b)) \\ &= \lambda[f,g](a,b)\end{aligned}$$

and $[f,g](0,0) = f(0) + g(0) = 0 + 0 = 0$.

Finally we must check that if $h : A \times B \longrightarrow C$ is any other linear transformation such that $h\iota_A = f$ and $h\iota_B = g$, then $h = [f,g]$:

$$\begin{aligned}h(a,b) &= h((a,0) + (0,b)) \\ &= h(a,0) + h(0,b) \\ &= h\iota_A(a) + h\iota_B(b) \\ &= f(a) + g(b) \\ &= [f.g](a,b)\end{aligned}$$

∎

Solution for I.67 on p. 35: To compute $\nabla(a_1, a_2)$ for any $(a_1, a_2) \in A \oplus A$ with A any vector space in **Vect**, note that $(a_1, a_2) = (a_1, 0) + (0, a_2) = \iota_1(a_1) + \iota_2(a_2)$ and so $\nabla(a_1, a_2) = \nabla(\iota_1(a_1) + \iota_2(a_2)) = \nabla(\iota_1(a_1)) + \nabla(\iota_2(a_2)) = 1_A(a_1) + 1_A(a_2) = a_1 + a_2$. ∎

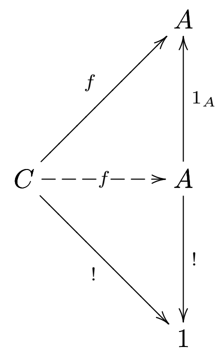
Solution for I.68 on p. 35: From the definition, $(f_1 + f_2)\iota_1(a_1) = \iota_1 f_1(a_1)$ and $(f_1 + f_2)\iota_2(a_2) = \iota_2 f_2(a_2)$, i.e., $(f_1 + f_2)(a_1, 0) = (f_1(a_1), 0)$ and $(f_1 + f_2)(0, a_2) = (0, f_2(a_2))$. And as $(a_1, a_2) = (a_1, 0) + (0, a_2)$ and $f_1 + f_2$ is a linear transformation we have $(f_1 + f_2)(a_1, a_2) = (f_1(a_1), f_2(a_2))$. ∎

Solution for I.69 on p. 36: Prove that any two final objects in $\mathscr{C}$ are isomorphism, and the isomorphism is unique.

If F and F' are both final objects, then there are *unique* morphisms $F \xrightarrow{!} F'$ and $F' \xrightarrow{F}$. But then the composition $F \xrightarrow{!} F' \xrightarrow{!} F$ must be the unique morphism from $F \longrightarrow F$, which is 1_F. Equally well the composition in the other order must be $1_{F'}$. ∎

Solution for I.70 on p. 36: There is hardly anything to say. For each and every object C, not just 1, there is a bijection between pairs of morphisms $(f : C \longrightarrow A, g : C \longrightarrow B)$ and morphisms $\langle f, g\rangle : C \longrightarrow A \times B$ of C to $A \times B$ – that is just a slightly different way of stating the definition of the product. ∎

Solution for I.71 on p. 36: Here is the diagram that exhibits $(A, 1_A, !)$ as a product of A and 1:



∎

Solution for I.72 on p. 36: To verify that every point $p : 1 \longrightarrow C$ is a monomorphism, consider the diagram

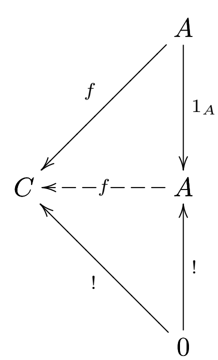
$$B \underset{g_2}{\overset{g_1}{\rightrightarrows}} 1 \xrightarrow{p} C$$

By the definition of a final object *any* two morphisms $g_1, g_2 : C \longrightarrow 1$ are equal, so certainly $pg_1 = pg_2 \Rightarrow g_1 = g_2$. ∎

Solution for I.73 on p. 37: If 0 and 0' are both initial objects, then there are unique morphisms $0 \longrightarrow 0'$ and $0' \longrightarrow 0$. But then the composition $0 \longrightarrow 0' \longrightarrow 0$ must be the unique morphism from $0 \longrightarrow 0$, i.e., 1_0. Equally well the composition in the other order must be $1_{0'}$. ∎

Compare this solution with the solution for exercise I.69 on page 37. The two were written carefully to make it clear that each can be transformed into the other by "reversing the arrows", including exchanging initial and final objects. In this particular case the proofs look identical, but this is yet another example of duality which is discussed formally in Section II.1.

Solution for I.74 on p. 37: Here is the diagram that exhibits $(A, 1_A, !)$ as a sum of A and 0:



∎

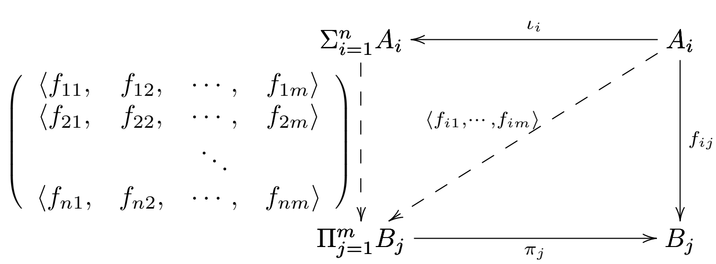
Compare this solution with the solution for exercise I.71 on page 36. The two were written carefully to make it clear that each can be transformed into the other by "reversing the arrows", including exchanging sums and products and initial and final objects. Again this is yet another example of duality which is discussed formally in Section II.1.

Solution for I.75 on p. 40: Verifying that the morphisms $\Sigma_{i=1}^{n} A_i \longrightarrow \Pi_{j=1}^{m} B_j$ are exactly the "matrices"

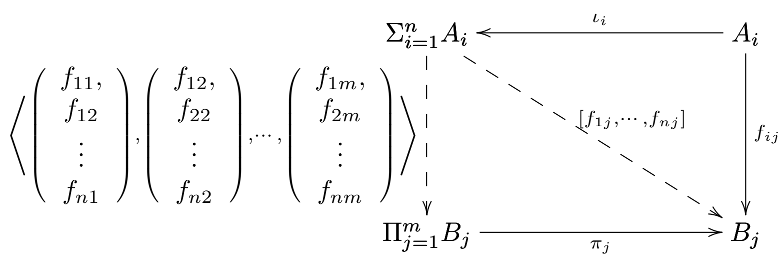
$$M = \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1m} \\ f_{21} & f_{22} & \cdots & f_{2m} \\ & & \ddots & \\ f_{n1} & f_{n2} & \cdots & f_{nm} \end{pmatrix}$$

where $\pi_j M \iota_i = f_{ij}$ is entirely a matter of unwinding the definitions. Given $f : \Sigma_{i=1}^{n} A_i \longrightarrow \Pi_{j=1}^{m} B_j$ we certainly get the matrix $M = (f_{ij})$ with $f_{ij} = \pi_j f \iota_i$.

In the reverse direction if we are given an $n \times m$-matrix (f_{ij}) of morphisms $f_{ij} : A_i \longrightarrow B_j$, then for each fixed i we get the *unique* morphism $\langle f_{i1}, \cdots, f_{im} \rangle : A_i \longrightarrow \Pi_{j=1}^{m} B_j$ whose composition with π_j is f_{ij}. And from all of those we get the *unique* morphism f (shown in "matrix" form in the following diagram) whose composition with ι_i is $\langle f_{i1}, \cdots, f_{im} \rangle$. ∎



Equally well, the $n \times m$-matrix (f_{ij}) of morphisms $f_{ij} : A_i \longrightarrow B_j$, then for each fixed j we get the *unique* morphism $[f_{1j}, \cdots, f_{nj}] : \Sigma_{i=1}^{n} A_i \longrightarrow B_j$ whose composition with ι_i is f_{ij}. And from all of those we get the *unique* morphism f (shown in "matrix" form in the following diagram) whose composition with π_j is $[f_{1j}, \cdots, f_{nj}]$. Notice as a result that the two morphisms from $\Sigma A_i \longrightarrow \Pi B_j$ defined here and above are equal.



Solution for I.76 on p. 41: In order to show that $I : X + Y \longrightarrow X \times Y$ corresponds exactly to the inclusion $(X \times \{y_0\} \cup \{x_0\} \times Y) \subset X \times Y$ we recall the construction of the sum of two pointed sets (see I.45): $\langle X, x_0\rangle + \langle Y, y_0\rangle = \langle X \times \{y_0\} \cup \{x_0\} \times Y, \langle x_0, Y_0\rangle\rangle$, $\iota_X : \langle X, x_0\rangle \longrightarrow \langle X, x_0\rangle + \langle Y, y_0\rangle$ by $\iota_X(x) = \langle x, y_0\rangle$ and $\iota_Y : \langle Y, y_0\rangle \longrightarrow \langle X, x_0\rangle + \langle Y, y_0\rangle$ by $\iota_X(y) = \langle x_0, y\rangle$

And what is I?

$$I(t) = \begin{cases} \langle t, y_0\rangle & \text{if } t \in X \\ \langle x_0, t\rangle & \text{if } t \in Y \end{cases}$$

which is exactly the desired inclusion.

Now if X and Y are not singletons, then there is $x_1 \in X$ with $x_1 \neq x_0$ and $y_1 \in Y$ with $y_1 \neq y_0$. But then $(x_1, y_1) \in X \times Y$, but it is not in the image of I. ∎

Solution for I.77 on p. 43: Starting with $f : A \longrightarrow C$ and $g : B \longrightarrow D$ we compute the the matrix $f + g$ by starting with the fact that $(f + g)\iota_A = \iota_C f$ and $(f + g)\iota_B = \iota_D g$. We need $\pi_C(f + g)\iota_A$, $\pi_C(f + g)\iota_B$, $\pi_D(f + g)\iota_A$, and $\pi_D(f + g)\iota_B$. But $\pi_C(f + g)\iota_A = \pi_C\iota_C f = 1_C f = f$ and $\pi_C(f + g)\iota_B = \pi_C\iota_D f = 0f = 0$. A similar computation for the other two shows that the matrix is

$$\begin{pmatrix} f & 0 \\ 0 & g \end{pmatrix}$$

And this is what we previously computed as the matrix of $f \times g$

Solution for I.78 on p. 43: To verify that the matrices of $\Pi_{i=1}^n f_i : \bigoplus_{i=1}^n A_i \longrightarrow \bigoplus_{i=1}^n B_i$ and $\Sigma_{i=1}^n f_i : \bigoplus_{i=1}^n A_i \longrightarrow \bigoplus_{i=1}^n B_i$ is

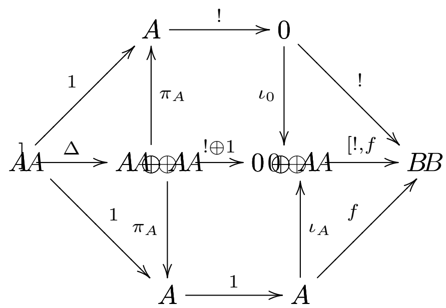
$$\begin{pmatrix} f_1 & 0 & \cdots & 0 \\ 0 & f_2 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & f_n \end{pmatrix}$$

is one of those cases where the use of indices makes the computation a little easier. The ij-entry of the matrix for the product is $\pi_i(\Pi_{k=1}^n f_k)\iota_j$ which is, by definition, equal to $f_i\pi_k\iota_j$ which is f_i when $i = j$ and 0 otherwise.

Equally well the ij-entry of the matrix for the sum is $\pi_i(\Sigma_{k=1}^n f_k)\iota_j$ which is, by definition, equal to $\pi_i\iota_j f_j$ which is f_i when $i = j$ and 0 otherwise.

Solution for I.79 on p. 44: In the category of Abelian groups, $\Delta : A \longrightarrow A \times A$ is $\Delta(a) = (a, a)$, so $(f \vartriangle g)(a) = [f, g](a, a) = f(a) + g(a)$. (Recall exercise I.63.) And $(f \triangledown g)(a) = \nabla(f(a), g(a)) = f(a) + g(a)$. ∎

Solution for I.80 on p. 45: To show that $f = 0+f$ we will use the following commutative diagram:



Notice that the left two triangle are just exhibiting $A \oplus A$ as a product, and showing the definition of Δ. The sequence $A \xrightarrow{!} 0 \xrightarrow{!} B$ is factoring $0 : A \longrightarrow B$. The middle squares exhibit the definition of $! \oplus 1_A$, and the right triangle on the right exhibit the definition of $[!, f]$. Recall from exercise I.74 that here $\iota_A : A \longrightarrow 0 \oplus A$ is an isomorphism.

Now the application of this diagram here is first that $[!, f](! \oplus 1_A)\Delta = f1_A1_A = f$, and second that $[!, f](! \oplus 1_A)\Delta = [0, f]\Delta = 0 + f$. So $f = 0 + f$. ∎

Now to get $f + 0 = f$ just interchange the top and bottom of the diagram. ∎

Solution for I.81 on p. 45: For $h(f + g)$ we have:

$$\begin{array}{rcll} h(f+g) & = & h(f \vartriangle g) & \text{by Proposition I.4} \\ & = & h[f,g]\Delta & \text{by definition I.49} \\ & = & [hf,hg]\Delta & \text{by exercise I.60} \\ & = & hf \vartriangle hg & \text{by definition I.49} \\ & = & hf + hg & \text{by Proposition I.4} \end{array}$$

And for $(f + g)e$ we have:

$$\begin{array}{rcll} (f+g)e & = & (f \triangledown g)e & \text{by Proposition I.4} \\ & = & \nabla\langle f,g\rangle e & \text{by definition I.50} \\ & = & \nabla\langle fe,ge\rangle & \text{by exercise I.45} \\ & = & fe \triangledown ge & \text{by definition I.50} \\ & = & fe + ge & \text{by Proposition I.4} \end{array}$$

∎

Solution for I.82 on p. 45: In order to verify that $\iota_A\pi_A+\iota_B\pi_B = 1_{A\oplus B}$ let's look at the matrix of $\iota_A\pi_A + \iota_B\pi_B$. For that we look at the four components:

$$\begin{aligned}\pi_A(\iota_A\pi_A + \iota_B\pi_B)\iota_A &= \pi_A\iota_A\pi_A\iota_A + \pi_A\iota_B\pi_B\iota_A\\ &= 1_A1_A + 00\\ &= 1_A + 0\\ &= 1_A\end{aligned}$$

$$\begin{aligned}\pi_A(\iota_A\pi_A + \iota_B\pi_B)\iota_B &= \pi_A\iota_A\pi_A\iota_B + \pi_A\iota_B\pi_B\iota_B\\ &= 1_A0 + 01_B\\ &= 0 + 0\\ &= 0\end{aligned}$$

$$\begin{aligned}\pi_B(\iota_A\pi_A + \iota_B\pi_B)\iota_A &= \pi_B\iota_A\pi_A\iota_A + \pi_B\iota_B\pi_B\iota_A\\ &= 01_A + 1_B0\\ &= 0 + 0\\ &= 0\end{aligned}$$

$$\begin{aligned}\pi_B(\iota_A\pi_A + \iota_B\pi_B)\iota_B &= \pi_B\iota_A\pi_A\iota_B + \pi_B\iota_B\pi_B\iota_B\\ &= 00 + 1_B1_B\\ &= 0 + 1_B\\ &= 1_B\end{aligned}$$

So the matrix of $\iota_A\pi_A + \iota_B\pi_B$ is the identity matrix. But only the identity morphism has that matrix! ∎

Solution for I.83 on p. 48: To show that there is a category, $\mathbf{Magma}_{\mathscr{C}}$, with objects the magmas in $\mathscr{C}$ and as morphisms the magma morphisms, we have to specify the identity morphisms, the composition and verify the required identities. Of course the identity morphism for a magma (M, μ) will just be the identity morphism on the object M which is clearly a magma morphism. And composition of magma morphisms will just be composition of the morphisms in $\mathscr{C}$. That composition of two magma morphisms is again a magma morphism is clear from

$$\begin{array}{ccc} M \times M & \xrightarrow{\mu} & M \\ \downarrow{\scriptstyle f\times f} & & \downarrow{\scriptstyle f} \\ N \times N & \xrightarrow[\mu]{\mu} & N \\ \downarrow{\scriptstyle g\times g} & & \downarrow{\scriptstyle g} \\ P \times P & \xrightarrow[\mu]{} & P \end{array}$$

and the fact (cf. exercise I.41) that $(g \times g)(f \times f) = (gf \times gf)$. Finally associativity of composition of magma morphisms and the identities for identity morphisms follows from associativity of composition of morphisms in $\mathscr{C}$. ∎

Solution for I.84 on p. 49:

If $\mathscr{C}$ is any category with finite products, then $\mathbf{Magma}_{\mathscr{C}}$ is also category with finite products. As noted in the text, when $M_1, \ldots, M_n$ are magmas in $\mathscr{C}$ (with binary operations $\mu_1, \ldots, \mu_n$), we define a binary operation, μ, on $\Pi_1^n M_i$ as the composite:

$$\Pi_1^n M_i \times \Pi_1^n M_i \xrightarrow{\cong} \Pi_1^n (M_i \times M_i) \xrightarrow{\Pi_1^n \mu_i} \Pi_1^n M_i$$

where the isomorphism is a special case of the result in exercise I.37.

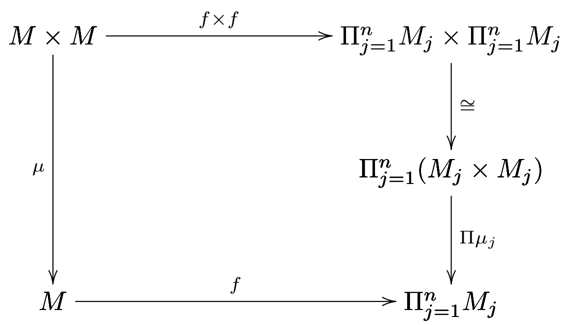
To verify this is a product in $\mathbf{Magma}_{\mathscr{C}}$ we need to confirm that the projection morphisms $\pi_i : \Pi_{i=1}^n M_i \longrightarrow M_i$ are magma morphisms, and that for every family of magma morphisms $f_i : M \longrightarrow M_i$, the unique morphism $f = \langle f_1, \cdots, f_n \rangle : M \longrightarrow \Pi_{i=1}^n M_i$ making the product diagram commute (which exists because we have the product in $\mathscr{C}$) is actually a magma morphism.

That each projection π_i is a magma homomorphism is seen by inspecting the following diagram:

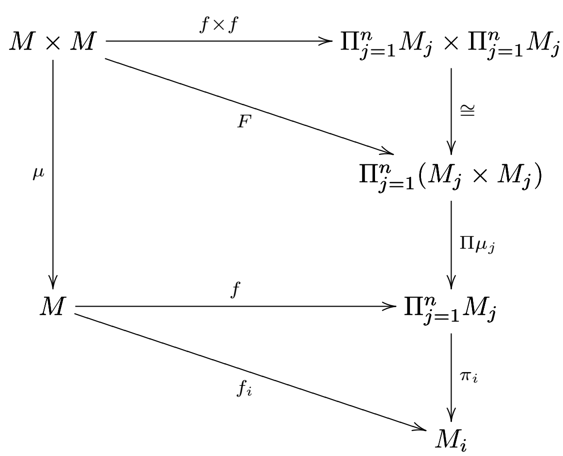
$$\begin{array}{ccc} \Pi_{j=1}^n M_j \times \Pi_{j=1}^n M_j & \overset{\pi_i \times \pi_i}{\searrow} & \\ \downarrow{\scriptstyle \cong} & & \\ \Pi_{j=1}^n (M_j \times M_j) & \xrightarrow{\pi_i} & M_i \times M_i \\ \downarrow{\scriptstyle \Pi\mu_j} & & \downarrow{\scriptstyle \mu_i} \\ \Pi_{j=1}^n M_j & \xrightarrow[\pi_i]{} & M_i \end{array}$$

The top triangle is commutative because the isomorphism is exactly the unique morphism that makes it commute. While the bottom square is commutative because this is exactly the definition of $\Pi\mu_j$.

To see that $f =< f_1, \cdots, f_n >$ is a magma morphism we need to show the following diagram commutes:

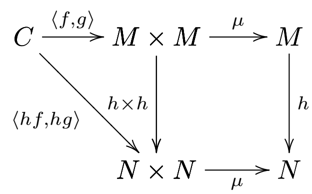


To show that two morphisms into the product $\Pi_{j=1}^n M_j$ are equal we need to show their compositions with the projections into each factor are equal. To see that let's add a few more morphisms:



(where we've written F for $\langle f_1 \times f_1, \cdots, f_n \times f_n \rangle$ to make the diagram less cluttered.) The top triangle commutes by the definition of the isomorphism, while the bottom triangle commutes by the definition of f. So we are left to verify that $\pi_i(\Pi\mu_j)F = f_i\mu$. But $\pi_i(\Pi\mu_j)F = \mu_i\pi_i F$ by definition of $\Pi\mu_j$, and this in turn is equal to $\mu_i(f_i \times f_i)$ by definition of F. And finally this is equal to $f_i\mu$ because f_i is a magma morphism by hypothesis. ∎

Solution for I.85 on p. 49: As with most everything here, verification that for M and N magmas and $h : M \longrightarrow N$ a magma morphism, then $\mathrm{Hom}(C, h)$ is a magma homomorphism is just a matter of unwinding the definitions. And as usual a commutative diagram helps:



$$
\begin{aligned}
\mathrm{Hom}(C,h)(f \vartriangle g) &= h(f \vartriangle g) \\
&= h\mu\langle f,g\rangle \\
&= \langle f,g\rangle(h \times h)\mu \\
&= \langle hf,hg\rangle\mu \\
&= hf \vartriangle hg \\
&= \mathrm{Hom}(C,h)(f) \vartriangle \mathrm{Hom}(C,h)(g)
\end{aligned}
$$

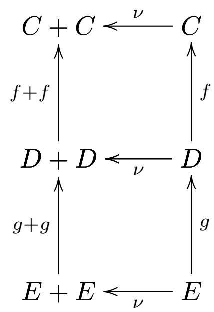
▮

Solution for I.86 on p. 49: Verifying that for any magma M and h any morphism at all $\mathrm{Hom}(h,M)$ is a magma homomorphism is even simpler than the last exercise:

$$
\begin{aligned}
\mathrm{Hom}(h,M)(f \vartriangle g) &= (f \vartriangle g)h \\
&= \mu\langle f,g\rangle h \\
&= \mu\langle fh,gh\rangle \\
&= fh \vartriangle gh \\
&= \mathrm{Hom}(h,M)(f) \vartriangle \mathrm{Hom}(h,M)(g)
\end{aligned}
$$

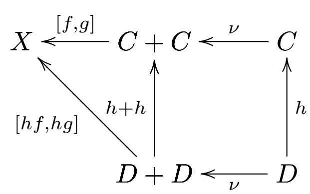
▮

Solution for I.87 on p. 51: To show that there is a category, $\mathbf{Comagma}_{\mathscr{C}}$, with objects the comagmas in $\mathscr{C}$ and as morphisms the comagma morphisms, we have to specify the identity morphisms, the composition and verify the required identities. Of course the identity morphism for a comagma (C,ν) will just be the identity morphism on the object C which is clearly a comagma morphism. And composition of comagma morphisms will just be composition of the morphisms in $\mathscr{C}$. That composition of two comagma morphisms is again a comagma morphism is clear from



and the fact (cf. exercise I.55) that $(f+f)(g+g) = (fg+fg)$. Finally associativity of composition of comagma morphisms and the identities for identity morphisms follows from associativity of composition of morphisms in $\mathscr{C}$. ∎

Solution for I.88 on p. 51: To verify that $h : D \longrightarrow C$ a comagma morphism, implies $\mathrm{Hom}(h, X)$ is a magma homomorphism is just a matter of unwinding the definitions. And as usual a commutative diagram helps:



$$
\begin{aligned}
\mathrm{Hom}(h,X)(f \triangledown g) &= (f \triangledown g)h \\
&= [f,g]\nu h \\
&= \nu(h+h)[f,g] \\
&= \nu[hf,hg] \\
&= hf \triangledown hg \\
&= \mathrm{Hom}(h,X)(f) \triangledown \mathrm{Hom}(h,X)(g)
\end{aligned}
$$

∎

Compare this solution with that of I.85. This is yet another example of duality, though now “reversing arrows” results not only in interchanging sums and products, but also interchanging magmas and comagmas.

Solution for I.89 on p. 51: Verifying that for any comagma C and h any morphism at all $\mathrm{Hom}(C, h)$ is a magma homomorphism is even simpler than the last exercise:

$$\begin{aligned}
\mathrm{Hom}(C,h)(f\triangledown g) &= h(f\triangledown g)\\
&= h[f,g]\nu\\
&= [hf,hg]\nu\\
&= hf\triangledown hg\\
&= \mathrm{Hom}(C,h)(f)\triangledown\mathrm{Hom}(C,h)(g)
\end{aligned}$$

∎

Compare this solution with that of exercise I.86. This is yet another example of duality, though now "reversing arrows" results not only in interchanging sums and products, but also interchanging magmas and comagmas.

Solution for I.90 on p. 52: The exercise here is to show that if $\mathscr{C}$ is any category with finite sums, then $\mathbf{Comagma}_{\mathscr{C}}$ is also category with finite sums.

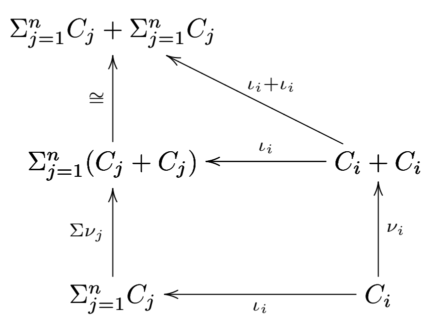
As noted in the text, when $C_1, \ldots, C_n$ are comagmas in $\mathscr{C}$ (with co-operations $\nu_1, \ldots, \nu_n$), we can define a co-operation, ν, on $\Sigma_1^n C_i$ as the composite:

$$\Sigma_1^n C_i \xrightarrow{\Sigma_1^n \nu_i} \Sigma_1^n (C_i + C_i) \xrightarrow{\cong} \Sigma_1^n C_i + \Sigma_1^n C_i$$

where the isomorphism is a special case of the result in exercise I.52.

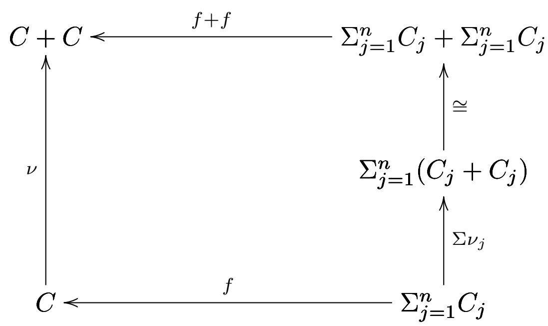
To verify this is a sum in $\mathbf{Comagma}_{\mathscr{C}}$ we need to confirm that the injection morphisms $\iota_i : C_i \longrightarrow \Sigma_{i=1}^n C_i$ are comagma morphisms, and that for every family of comagma morphisms $f_i : C \longrightarrow C_i$, the unique morphism $f = [f_1, \cdots, f_n] : C \longrightarrow \Sigma_{i=1}^n C_i$ making the sum diagram commute (which exists because we have the sum in $\mathscr{C}$) is actually a comagma morphism.

That each injection ι_i is a comagma homomorphism is seen by inspecting the following diagram:

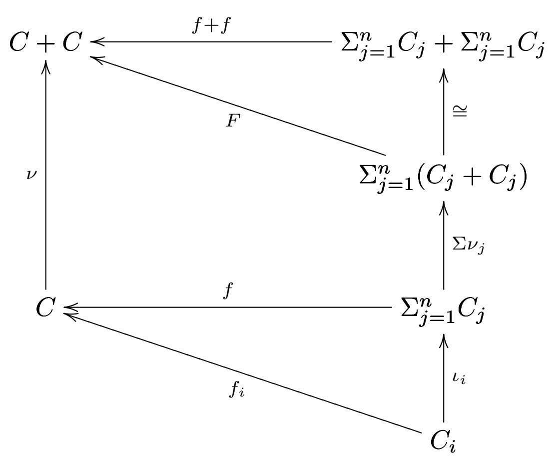


The top triangle is commutative because the isomorphism is exactly the unique morphism that makes it commute. While the bottom square is commutative because this is exactly the definition of $\Sigma\nu_j$.

To see that $f = [f_1, \cdots, f_n]$ is a comagma morphism we need to show the following diagram commutes:



To show that two morphisms from the sum $\Sigma_{j=1}^n C_j$ are equal we need to show their compositions with the injections from each summand are equal. To see that let's add a few more morphisms:



(where we've written F for $[f_1 + f_1, \cdots, f_n + f_n]$ to make the diagram less cluttered.) The top triangle commutes by the definition of the isomorphism, while the bottom triangle commutes by the definition of f. So we are left to verify that $F(\Sigma\nu_j)\iota_i = \nu f_i$.

But $F(\Sigma\nu_j)\iota_i = F\iota_i\nu_i$ by definition of $\Sigma\nu_j$, and this in turn is equal to $(f_i + f_i)\nu_i$ by definition of F. And finally this is equal to νf_i because f_i is a comagma morphism by hypothesis. ∎

Compare this solution with that of exercise I.84. This is yet another example of dual proofs of dual theorems. In particular this solution was carefully produced by taking the solution of exercise I.84 and making exactly the changes necessary to produce the dual proof.

Solution for I.91 on p. 56: The proof of parts (4.) – (6.) of Proposition I.8 were left as an exercise because they are dual to parts (1.) – (3.). Here are the dual proofs:

For (4.) note that $[\iota_1, 0] \vartriangle = \iota_1 + 0 = \iota_1$ and $[0, \iota_2] \vartriangle = 0 + \iota_2 = \iota_2$ as 0 is the identity for $+$, the induced binary operation on $\mathrm{Hom}(A, \bullet)$

For (5.) we see that $[\vartriangle, 1] \vartriangle = \vartriangle +1$, while $[1, \vartriangle] \vartriangle = 1+ \vartriangle$. But we know that $+$ is commutative, so $\vartriangle +1 = 1+ \vartriangle$.

For (6.) we observe that $\vartriangle t$ is also a binary operation on A for which 0 is the identity in the monoid $\mathrm{Hom}(A, \bullet)$. But such a binary operation on A is unique, so $\vartriangle t = \vartriangle$. ∎

Solution for I.92 on p. 58: To show that there is a category, $\mathbf{Monoid}_{\mathscr{C}}$, with objects the monoids in $\mathscr{C}$ and as morphisms the monoid morphisms, we have to specify the identity morphisms, the composition and verify the required identities. Of course the identity morphism for a monoid (M, μ) will just be the identity morphism on the object M which is clearly a monoid morphism. And composition of monoid morphisms will just be composition of the morphisms in $\mathscr{C}$. The only work to be done is verification that composition of two monoid morphisms is again a monoid morphism. The first part of that was done in verifying the existence of $\mathbf{Magma}_{\mathscr{C}}$ in the solution to exercise I.83. For the second part suppose $f : M \longrightarrow N$ and $g : N \longrightarrow P$ are monoid morphisms, so in particular $f\zeta^M = \zeta^N$ and $g\zeta^N = \zeta^P$, then $gf\zeta^M = \zeta^P$ and so gf is a monoid morphism as well.

That identity morphisms behave as required and that composition is associative is true in $\mathbf{Monoid}_{\mathscr{C}}$ just because it is true in $\mathscr{C}$. ∎

Solution for I.93 on p. 58: In exercise I.92 you displayed the category of monoids in $\mathscr{C}$, $\mathbf{Monoid}_{\mathscr{C}}$. For the category of commutative monoids in $\mathscr{C}$ there is really nothing more to say – the identities, law of composition and various identities all come from those same items in $\mathbf{Monoid}_{\mathscr{C}}$. Indeed this is just one of many examples of a full subcategory defined by a class of objects in some category (see section I.1.2, especially definition I.12.) ∎

Solution for I.94 on p. 59: If $\mathscr{C}$ is any category with finite products, then $\mathbf{Monoid}_{\mathscr{C}}$ is also a category with finite products. Much of the work of proving this has already been done in exercise I.84. What remains to be done is (1.) check that when $M_1, \ldots, M_n$ are monoids, the product binary operation on the magma $\Pi_{i=1}^n M_i$ also satisfies the associative law (see definition I.59); (2.) define an identity $\zeta : 1 \longrightarrow \Pi_{i=1}^n M_i$ (including verifying the identity law holds); (3.) verify that each of the projection morphisms $\pi_i : \Pi_{i=1}^n M_i \longrightarrow M_i$ preserves the identity; and (4.) show that for every family of monoid morphisms $f_i : M \longrightarrow M_i$, the unique magma morphism $f = \langle f_1, \cdots, f_n \rangle : M \longrightarrow \Pi_{i=1}^n M_i$ making the product diagram commute respect the identity.

(1.) Taking $M = \Pi_{i=1}^n M_i$ the associative law asserts the following diagram is commutative:

$$\begin{array}{ccc} M\times M\times M & \xrightarrow{\langle 1,\mu\rangle} & M\times M \\ {\scriptstyle \langle\mu,1\rangle}\downarrow & & \downarrow{\scriptstyle\mu} \\ M\times M & \xrightarrow[\mu]{} & M\end{array}$$

As always, to verify that two morphisms into a product are equal it is enough to verify that their composites with all of the projection morphisms are equal. So we will verify that the following diagrams are commutative:

$$\begin{array}{ccccc} M\times M\times M & \xrightarrow{1\times\mu} & M\times M & \xrightarrow{\mu} & M \\ \downarrow{\scriptstyle \pi_i\times\pi_i\times\pi_i} & & \downarrow{\scriptstyle \pi_i\times\pi_i} & & \downarrow{\scriptstyle \pi_i} \\ M_i\times M_i\times M_i & \xrightarrow[1\times\mu_i]{} & M_i\times M_i & \xrightarrow[\mu_i]{} & M_i \end{array}$$

$$\begin{array}{ccccc} M\times M\times M & \xrightarrow{\mu\times 1} & M\times M & \xrightarrow{\mu} & M \\ \downarrow{\scriptstyle \pi_i\times\pi_i\times\pi_i} & & \downarrow{\scriptstyle \pi_i\times\pi_i} & & \downarrow{\scriptstyle \pi_i} \\ M_i\times M_i\times M_i & \xrightarrow[\mu_i\times 1]{} & M_i\times M_i & \xrightarrow[\mu_i]{} & M_i \end{array}$$

Notice that the only differences in the diagrams are the $1 \times \mu$ and $1 \times \mu_i$ in the top diagram versus $\mu \times 1$ and $\mu_i \times 1$ in the bottom. But we know that for all i the bottom rows of the two diagrams are equal because the binary operations μ_i are associative. So if the diagrams are commutative, it follows that the top rows are equal and so μ is also associative.

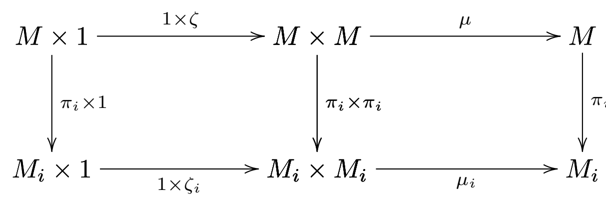
In the top diagram, the right hand square is commutative by the definition of $\mu = \Pi_{i=1}^{n}\mu_i$. For the left hand square note that $(\pi_i \times \pi_i)(1 \times \mu) = \pi_i \times (\pi_i\mu) = \pi_i \times (\mu_i(\pi_i \times \pi_i)) = (1 \times \mu_i)(\pi_i \times \pi_i \times \pi_i)$

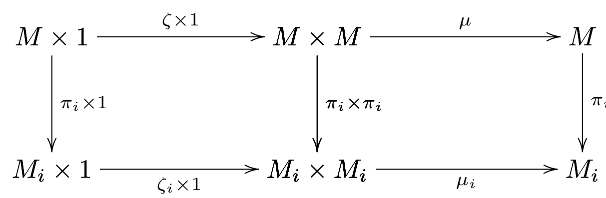
The lower diagram is commutative by essentially the same argument. ∎

(2.) With the identity on M defined as $\zeta = \langle \zeta_1, \cdots, \zeta_n\rangle : 1 \longrightarrow \Pi_{i=1}^{n} M_i$, we want to verify that the following two triangles commute.

$$\begin{array}{ccccc} M\times 1 & \xrightarrow{1_M\times\zeta} & M\times M & \xleftarrow{\zeta\times 1_M} & 1\times M \\ & {\scriptstyle\pi_1}\searrow & \downarrow{\scriptstyle\mu} & \swarrow{\scriptstyle\pi_2} & \\ & & M & & \end{array}$$

Again it is enough to verify that their composites with all of the projection morphisms are equal. So we first note that the following diagrams are commutative:





and observe that the bottom row in both cases is $\pi_1 : M_i \times 1 \longrightarrow M_i$ because ζ_i is the identity for μ_i. But that says for every i, $\pi_i\mu(1\times\zeta) = \pi_1(\pi_i \times 1)$ which in turn is equal to $\pi_i\pi_1$ and so $\mu(1 \times \zeta) = \pi_1$. ∎

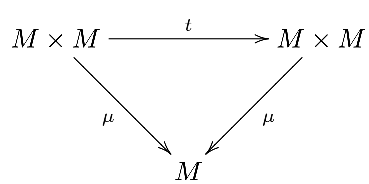
(3.) To verify that each of the projection morphisms $\pi_i : \Pi_{i=1}^n M_i \longrightarrow M_i$ preserves the identity is just to note the definition of ζ: $\pi_i\zeta = \zeta_i$. ∎

(4.) And finally verifying that $f = \langle f_1, \cdots, f_n\rangle$ always respects the identity is just noting $f\zeta = \langle f_1\zeta_1, \cdots, f_n\zeta_n\rangle = \langle \zeta_1, \cdots, \zeta_n\rangle = \zeta$. ∎

Solution for I.95 on p. 59: Suppose $h : M \longrightarrow N$ is a monoid morphism in $\mathscr{C}$ and C is any object, then in exercise I.85 you showed that h_* from $\mathrm{Hom}(C, M)$ to $\mathrm{Hom}(C, N)$ is a magma morphism. It is a monoid homomorphism because $h\zeta = \zeta$ implies that $h_*\zeta_* = \zeta_*$, i.e., $h_*(\zeta) = \zeta$ where these last ζ's are the identities in $\mathrm{Hom}(C, M)$ and $\mathrm{Hom}(C, N)$. ∎

Solution for I.96 on p. 59: Considering $h : D \longrightarrow C$ and the induced function $h^* : \mathrm{Hom}(C, M) \longrightarrow \mathrm{Hom}(D, M)$ for a monoid M in $\mathscr{C}$, half the work to show h^* is a monoid homomorphism was done in exercise I.86. What remains is to see that $h^*(\zeta) = \zeta$ and that is just because $!h =!$ where the !'s are the unique morphisms to the final object. ∎

Solution for I.97 on p. 61: Most of the work to prove Theorem I.5 was done in proving Theorem I.4. From that we have the unique morphisms so that (M, μ, ζ) is a monoid inducing the given monoid structure in $\mathrm{Hom}(C, M)$. All that remains is to prove the binary operation commutative. which means that the diagram



is commutative. Recall that $\mu = \pi_1 \nabla \pi_2$ and $t = \langle \pi_2, \pi_1 \rangle$ so $\mu t = \pi_2 \nabla \pi_1$. By hypothesis ∇ is commutative, so $\pi_1 \nabla \pi_2 = \pi_2 \nabla \pi_1$ and $\mu t = \mu$. ∎

Solution for I.98 on p. 62: To show that there is a category, $\mathbf{Comonoid}_{\mathscr{C}}$, with objects the comonoids in $\mathscr{C}$ and as morphisms the comonoid morphisms, we have to specify the identity morphisms, the composition and verify the required identities. Of course the identity morphism for a comonoid (C, ν) will just be the identity morphism on the object C which is clearly a comonoid morphism. And composition of comonoid morphisms will just be composition of the morphisms in $\mathscr{C}$. The only work to be done is verification that composition of two comonoid morphisms is again a comonoid morphism. The first part of that was done in verifying the existence of $\mathbf{Comagma}_{\mathscr{C}}$ in the solution to exercise I.87. For the second part suppose $f : C \longrightarrow D$ and $g : D \longrightarrow E$ are comonoid morphisms, so in particular $\eta^C f = \eta^D$ and $\eta^D g = \eta^E$, then $\eta^C gf = \eta^E$ and so gf is a comonoid morphism as well.

That identity morphisms behave as required and that composition is associative is true in $\mathbf{Comonoid}_{\mathscr{C}}$ just because it is true in $\mathscr{C}$. ∎

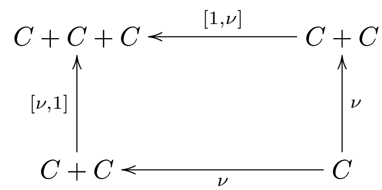
Compare this solution to the solution of exercise I.92. This is still another example of dual proofs of dual theorems. In particular this solution was carefully produced by taking the solution of exercise I.92 and making exactly the changes necessary to produce the dual proof.

Solution for I.99 on p. 62: In exercise I.98 you displayed the category of comonoids in $\mathscr{C}$, $\mathbf{Comonoid}_{\mathscr{C}}$. For the category of co-commutative comonoids in $\mathscr{C}$ there is really nothing more to say – the identities, law of composition and various identities all come from those same items in $\mathbf{Comonoid}_{\mathscr{C}}$. Indeed this is just one of many examples of a full subcategory defined by a class of objects in some category (see section I.1.2, especially definition I.12.) ∎

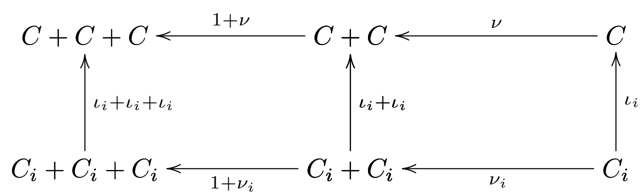
Compare this solution to the solution of exercise I.93. This time the two proofs are nearly identical, but this is still an example of dual proofs of dual theorems. In particular this solution was produced by taking the solution of exercise I.92 and following using the meaning of "dual" to make exactly the changes necessary to produce the dual proof, which in this case was just changing a few words and symbols.

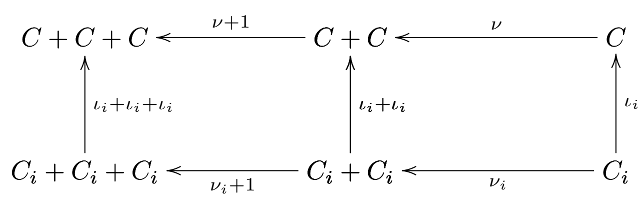
Solution for I.100 on p. 63: If $\mathscr{C}$is any category with finite sums, then $\mathbf{Comonoid}_{\mathscr{C}}$ is also a category with finite sums. Much of the work of proving this has already been done in exercise I.90. What remains to be done is (1.) check that when $C_1, \ldots, C_n$ are monoids, the co-operation on the magma $\Sigma_{i=1}^n C_i$ also satisfies the associative law (see definition I.62); (2.) define an co-unit $\eta : \Sigma_{i=1}^n C_i \longrightarrow 1$ (including verifying the co-unit law holds); (3.) verify that each of the inclusion morphisms $\iota_i : C_i \longrightarrow \Sigma_{i=1}^n C_i$ preserves the co-unit; and (4.) show that for every family of comonoid morphisms $f_i : C_i \longrightarrow C$, the unique co-magma morphism $f = [f_1, \cdots, f_n] : \Sigma_{i=1}^n C_i \longrightarrow C$ making the sum diagram commute respect the identity.

1. Taking $C = \Sigma_{i=1}^n C_i$ the co-associative law asserts the following diagram is commutative:



As always, to verify that two morphisms from a sum are equal it is enough to verify that their composites with all of the inclusion morphisms are equal. So we will verify that the following diagrams are commutative:



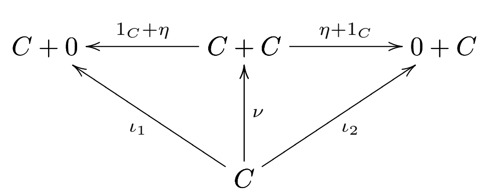


Notice that the only differences in the diagrams are the $1+\nu$ and $1+\nu_i$ in the top diagram versus $\nu+1$ and ν_i+1 in the bottom. But we know that for all i the bottom rows of the two diagrams are equal because the co-operations ν_i are co-associative. So if the diagrams are commutative, it follows that the top rows are equal as well so ν is also co-associative.

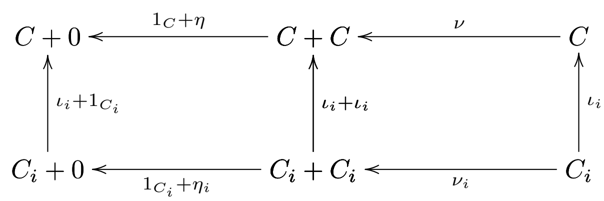
In the top diagram, the right hand square is commutative by the definition of $\nu = \Sigma_{i=1}^n \nu_i$. For the left hand square note that $(\iota_i + \iota_i)(1+\nu) = \iota_i + (\iota_i \nu) = \iota_i + (\nu_i(\iota_i + \iota_i)) = (1+\nu_i)(\iota_i + \iota_i + \iota_i)$

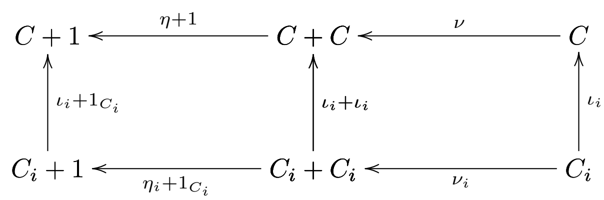
The lower diagram is commutative by essentially the same argument. ∎

2. With the co-unit on C defined as $\eta = [\eta_1, \cdots, \eta_n] : \Sigma_{i=1}^n C_i \longrightarrow 0$, we want to verify that the following two triangles commute.



Again it is enough to verify that their composites with all of the injection morphisms are equal. So we first note that the following diagrams are commutative:





and observe that the bottom row in both cases is $\iota_1 : C_i \longrightarrow C_i + 0$ because η_i is the co-unit for ν_i. But that says for every i, $\iota_i\nu(1_C + \eta) = \iota_1(\iota_i + 1_{C_i})$ which in turn is equal to $\iota_i\iota_1$ and so $\nu(1 + \eta) = \iota_1$. ∎

3. To verify that each of the inclusion morphisms $\iota_i : C_i \longrightarrow \Sigma_{i=1}^n C_i$ preserves the co-unit is just to note the definition of η as the unique morphism with $\iota_i\eta = \eta_i$. ∎

4. And finally verifying that $f = [f_1, \cdots, f_n]$ always respects the co-unit is just noting $f\eta = [f_1\eta_1, \cdots, f_n\eta_n] = [\eta_1, \cdots, \eta_n] = \eta$. ∎

Compare this solution to the solution of exercise I.94. This is still another example of dual proofs of dual theorems. In particular this solution was carefully produced by taking the solution of exercise I.94 and making exactly the changes necessary to produce the dual proof. In particular products were replaced by sums, product projections by sum injections, monoids by comonoids, and – the root of all the changes – the direction of all morphisms was reversed. This shows most clearly in all of the diagrams.

Solution for I.101 on p. 63: Suppose $h : C \longrightarrow D$ is a comonoid morphism in $\mathscr{C}$ and A is any object, then in exercise I.88 you showed that $h^* : \mathrm{Hom}(D, A) \longrightarrow \mathrm{Hom}(C, A)$ is a magma morphism. To show that it is a monoid homomorphism we just note that $\eta h = \eta$ implies that $\eta^* h^* = \eta^*$, i.e., $h^*(\zeta) = \zeta$ where these last ζ's are the identities in $\mathrm{Hom}(D, A)$ and $\mathrm{Hom}(C, A)$. ∎

Compare this to the solution of exercise I.96. Once again this is a dual proof of a dual theorem.

Solution for I.102 on p. 63: To verify that for every comonoid C in $\mathscr{C}$ and every morphism $h : A \longrightarrow B$ the induced function $h_* : \mathrm{Hom}(C, A) \longrightarrow \mathrm{Hom}(C, B)$

is actually a monoid homomorphism was half done in exercise I.89. What remains is to see that $h_*(\zeta) = \zeta$ and that is just because $h! =!$ where the !'s are the unique morphisms from an initial object. ▮

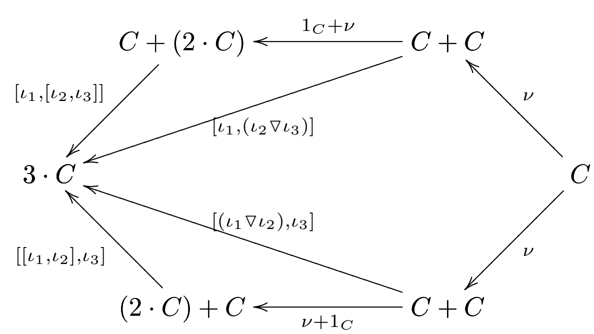
Solution for I.103 on p. 64: Proof: In order to prove Theorem I.6 note the first part of this is Theorem I.2 which gives us the unique co-operation ν on C inducing the binary operation on $\mathrm{Hom}(C, \bullet)$.

To get the co-unit, $\eta : C \longrightarrow 0$, note that if it exists it is in $\mathrm{Hom}(C, 0)$ and it must be the identity element in that monoid. So let us define η to be the identity element in $\mathrm{Hom}(C, 0)$ and prove that it is also the co-unit for the co-operation ν. The next thing to notice is that as 0 is an initial object, for every object D the unique morphism $! : 0 \longrightarrow D$ induces a monoid homomorphism $!_* : \mathrm{Hom}(C, 0) \longrightarrow \mathrm{Hom}(C, D)$ which in particular takes the identity element in $\mathrm{Hom}(C, 0)$ to the identity element in $\mathrm{Hom}(C, D)$, i.e., $!\eta$ is the identity element in $\mathrm{Hom}(C, D)$.

Applying this to $\iota_2 : 0 \longrightarrow C + 0$ we first notice that ι_2 is really !, so $\iota_2\eta$ is the identity element in $\mathrm{Hom}(C, C + 0)$. And $(1_C + \eta)\nu = \iota_2\eta\nabla\iota_1 = \iota_1$.

The argument for the other half of η being a co-unit is essentially the same and is left to the reader.

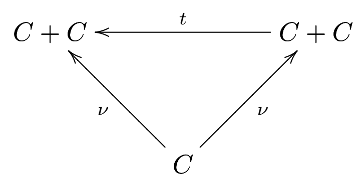
To verify the co-associativity of ν we need to make the diagram in the definition a bit more fulsome:



And finally $\nu[\iota_1, \iota_2\nabla\iota_3] = \iota_1\nabla(\iota_2\nabla = (\iota_1\nabla\iota_2)\iota_3)$, while $\nu[(\iota_1\nabla\iota_2), \iota_3]\nabla\iota_3$. But these two are equal because the binary operation ∇ on $\mathrm{Hom}(C^3, C)$ is associative. ▮

Compare this solution to the proof of Theorem I.4. This is yet one more example of a dual proofs of a dual theorem. In particular this solution was carefully produced by taking the proof of Theorem I.4 and making exactly the changes necessary to produce the dual proof. In particular products were replaced by sums, product projections by sum injections, monoids by comonoids, and – the root of all the changes – the direction of all morphisms was reversed. This shows most clearly in the rather complicated diagram.

Solution for I.104 on p. 64: Most of the work to prove Theorem I.7 was done in proving Theorem I.6. From that we have the unique morphisms so that (C, ν, η) is a comonoid inducing the given monoid structure in $\mathrm{Hom}(C, \bullet)$. All that remains is to prove the co-operation is co-commutative. which means that the diagram



is commutative. Recall that $\nu = \iota_1 \vartriangle \iota_2$ and $t = [\iota_2, \iota_1]$ so $\nu t = \iota_2 \vartriangle \iota_1$. By hypothesis $\vartriangle$ is commutative, so $\iota_1 \vartriangle \iota_2 = \iota_2 \vartriangle \iota_1$ and $\nu t = \nu$. ∎

Compare this to the solution of exercise I.97. We have yet another example of a dual proof of a dual theorem.

Solution for I.105 on p. 65: To show that there is a category, $\mathbf{Group}_{\mathscr{C}}$, with objects the groups in $\mathscr{C}$and as morphisms the group morphisms, we have to specify the identity morphisms, the composition and verify the required identities. Of course the identity morphism for a group (G, μ, ζ, ι) will just be the identity morphism on the object M which is clearly a group morphism. And composition of group morphisms will just be composition of the morphisms in $\mathscr{C}$. The only work to be done is verification that composition of two group morphisms is again a group morphism. Most of that was done in verifying the existence of $\mathbf{Monoid}_{\mathscr{C}}$ in the solution to exercise I.92. The only part remaining is to show that when $f : G \longrightarrow H$ and $g : H \longrightarrow K$ are group morphisms, so in particular $f\iota^G = \iota^H F$ and $g\iota^H = \iota^K g$, then $gf\iota^G = \iota^K$ and so gf is a group morphism as well.

That identity morphisms behave as required and that composition is associative is true in $\mathbf{Group}_{\mathscr{C}}$ just because it is true in $\mathscr{C}$. ∎

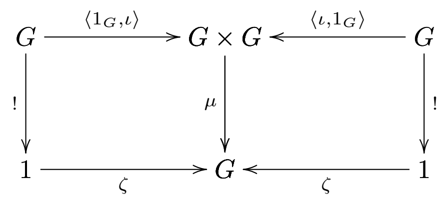
Solution for I.106 on p. 66: To show that there is a category, $\mathbf{Ab}_{\mathscr{C}}$, with objects the commutative (or Abelian) groups in $\mathscr{C}$and as morphisms the group morphisms, is essentially trivial following exercise I.105, just as exercise I.93 was trivial after exercise I.92.

The identities, law of composition and various identities in $\mathbf{Ab}_{\mathscr{C}}$ all come from those same items in $\mathbf{Group}_{\mathscr{C}}$. Indeed this is just one of many examples of a full subcategory defined by a class of objects in some category (see section I.1.2, especially definition I.12.) ∎

Solution for I.107 on p. 67: If $\mathscr{C}$is any category with finite products, then $\mathbf{Monoid}_{\mathscr{C}}$ is also a category with finite products as was confirmed in exercise I.94. So if $(G_i, \mu_i, \zeta_i, \iota_i)$ is a family of objects in $\mathbf{Group}_{\mathscr{C}}$, then (G_i, μ_i, ζ_i) is a family of objects in $\mathbf{Monoid}_{\mathscr{C}}$ with a product (G, μ, ζ) (including projections $\pi_i : G \longrightarrow G_i$ in $\mathbf{Monoid}_{\mathscr{C}}$) To get the product in $\mathbf{Group}_{\mathscr{C}}$ we need to do the following:

- Define an inverse $\iota : G \longrightarrow G$.

Which We do by defining $\iota = \Pi_{i=1}^{n}\iota_i$. (Remember this is the same as saying $\pi_i\iota = \iota_i\pi_i$ which is what we will use below.) Now we must verify that the following diagrams commute.



As always, to verify that two morphisms into a product are equal it is enough to verify that their composites with all of the projection morphisms are equal. But for every i from 1 to n we have

$$\begin{aligned}
\pi_i mu\langle 1_G, \iota\rangle &= \mu_i\pi_i\langle\langle\pi_1,\pi_1\rangle,\cdots,\langle\pi_n,\pi_n\rangle\rangle\langle 1_G,\iota\rangle \\
&= \mu_i\langle\pi_i,\pi_i\rangle\langle 1_G,\iota\rangle \\
&= \mu_i\langle\pi_i 1_G,\pi_i\iota\rangle \\
&= \mu_i\langle 1_{G_i}\pi_i,\iota_i\pi_i\rangle \\
&= \mu_i\langle 1_{G_i},\iota_i\rangle\langle\pi_i,\pi_i\rangle \\
\pi_i\zeta! &= \zeta_i! < \pi_i,\pi_i\rangle
\end{aligned}$$

and so $\mu\langle 1_G, \iota\rangle = \zeta!$ which shows the left square commutes. The proof that the right square commutes is a slight rearrangement of this. ▮

- To verify that each of the projection morphisms $\pi_i : \Pi_{i=1}^{n} G_i \longrightarrow G_i$ commutes with the inverse morphism is just to note the definition of ι: $\pi_i\iota = \iota_i\pi_i$. ▮
- And finally verifying that $f = \langle f_1, \cdots, f_n\rangle$ always commutes with the inverse is just noting $f\iota = \langle f_1\iota_1, \cdots, f_n\iota_n\rangle = \langle\iota_1 f_1, \cdots, \iota_n\, f_n\rangle = \iota f$. ▮

Solution for I.108 on p. 67: To verify that ι_* is indeed the inverse on $\mathrm{Hom}(C, G)$ we check that for each $f \in \mathrm{Hom}(C, G)$ we have

$$\begin{aligned}
f\triangledown\iota_*(f) &= \mu\langle f, \iota_*(f)\rangle \\
&= \mu_*(\langle(1_G)_*(f), \iota_*(f)\rangle \\
&= \mu_*\langle(1_G)_*, \iota_*\rangle(f) \\
&= \zeta_*!_*(f) \\
&= \zeta_*(*) \\
&= e \text{ the identity in } \mathrm{Hom}(C, G)
\end{aligned}$$

The check that $\iota_*(f)\triangledown f = e$ is essentially the same. ▮

Solution for I.109 on p. 67: Suppose $h : G \longrightarrow H$ is a group morphism in $\mathscr{C}$ and C is any object, then, forgetting the inverse, h is a monoid morphism or even, forgetting the identity, just a magma morphism. And in exercise I.95 you showed that h_* from $\mathrm{Hom}(C, G)$ to $\mathrm{Hom}(C, H)$ is a monoid morphism (and that in turn was based on exercise I.85 which shows h_* is a magma homomorphism.) Now to complete the proof that h_* is a group homomorphism just note that $h\iota = \iota h$ implies $h_*\iota_* = \iota_* h_*$. $\blacksquare$

Solution for I.110 on p. 68: Considering $h : D \longrightarrow C$ and the induced function $h^* : \mathrm{Hom}(C, G) \longrightarrow \mathrm{Hom}(D, G)$ for a group G in $\mathscr{C}$, most of the work to show h^* is a group homomorphism was done in exercise I.96 showing h^* is a monoid homomorphism (and which in turn was based on exercise I.86 showing h^* is a magma homomorphism.) What remains is to see that $h^*\iota^* = \iota^* h^*$ which follows immediately from $h\iota = \iota h$. $\blacksquare$

Solution for I.111 on p. 68: To prove Theorem I.8 we first note that Theorem I.8 gives most of the result, namely the monoid (G, μ, ζ) inducing the monoid structure in $\mathrm{Hom}(\bullet, G)$. What remains is to define the inverse morphism $\iota : G \longrightarrow G$ and verify that it induces the inverse operation on $\mathrm{Hom}(\bullet, G)$.

By hypothesis, $\mathrm{Hom}(G, G)$ is a group so in particular 1_G has an inverse with regard to this group operation and we will take this to be the inverse on G and so write it as ι. So we have $1_G \nabla \iota = e = \iota \nabla 1_G$ (e being the identity for the group $\mathrm{Hom}(G, G)$.)

For the induced operation note that for $f : C \longrightarrow G$ we know that $f^*(\iota)$ is the inverse of $f^*(1_G) = f$. But $f^*(\iota) = \iota f = \iota_*(f)$, so ι does induce the inverse operation on $\mathrm{Hom}(C, G)$. $\blacksquare$

C.2 Solutions for Chapter II

Solution for II.1 on p. 73: To verify that $\mathscr{C}/\sim$ is a category we note that we have identified the objects, the morphisms, the identity morphisms and composition of morphisms. The requirement that $f \sim g$ implies f and g have the same domain and codomain guarantees that the domain and codomain of an equivalence class of morphisms is well defined. And the requirement that $f \sim g$ and $h \sim k$ implies $hf \sim kg$. guarantees both the proper behavior of identity morphisms and associativity of composition follow from those facts in $\mathscr{C}$. $\blacksquare$

Solution for II.2 on p. 73: If $\mathscr{A}$ and $\mathscr{B}$ are two categories, the objects of $\mathscr{A} \times \mathscr{B}$ are pairs of objects from $\mathscr{A}$ and $\mathscr{B}$ respectively, while the morphisms are similar pairs of morphisms.

To verify that $\mathscr{A} \times \mathscr{B}$ is indeed a category we note that we while we have objects and morphisms, we need to identify domains and codomains, identity morphisms and compositions, then check that the appropriate identities hold.

We define the domain of $(f : A_1 \longrightarrow A_2,\ g : B_1 \longrightarrow B_2)$ to be (A_1, B_1), and the codomain to be (A_2, B_2), so $(f, g) : (A_1, B_1) \longrightarrow (A_2, B_2)$. The identity morphism on (A, B) is taken as $(1_A, 1_B)$, and composition of $(f, g) : (A_1, B_1) \longrightarrow (A_2, B_2)$ with $(h, k) : (A_2, B_2) \longrightarrow (A_3, B_3)$ is defined to be $(hf, kg) : (A_1, B_1) \longrightarrow (A_3, B_3)$.

The identities $(1_{A_1}, 1_{B_1})(f, g) = (f, g) = (f, g)(1_{A_2}, 1_{B_2})$ and $(p, q)((h, k)(f, g)) = ((p, q)(h, k))(f, g)$ follow from the corresponding identities in $\mathscr{A}$ and $\mathscr{B}$. ∎

Solution for II.3 on p. 74: To verify that $\Sigma_{i \in I}\mathscr{C}_i$, as defined in II.6, is actually a category observe that when $(f, i) : (C, i) \longrightarrow (D, i)$ and $(g, i) : (D, i) \longrightarrow (E, i)$ are morphisms in $\Sigma_{i \in I}\mathscr{C}_i$ their composition is $(gf, i) : (C, i) \longrightarrow (E, i)$ and this is associative because composition is associative in each $\mathscr{C}_i$. The identity morphism on (C, i) is $(1_C, i)$ and this has the properties of an identity morphism because 1_C is the identity morphism on C.

Solution for II.4 on p. 75: No solution is given here. As mentioned in the exercise "the answer in each case is in the article for the given category".

Solution for II.5 on p. 76:

1. **Group** is based on **Set** – A group is a set together with a binary operation that is associative, has an identity and where every element has an inverse, so the underlying set is just gotten by ignoring the binary operation. Similarly a group homomorphism is a function between the corresponding sets that behaves properly with respect to the binary operation, to the underlying function is just that function. Checking the consistency conditions is immediate.

2. **Group** is based on **Monoid** – A group is equally well a monoid where every element has an inverse, so the underlying monoid of a group is just the group, forgetting the extra properties. Additionally a group homomorphism is just a monoid homomorphism between groups, so the underlying homomorphism is just itself. Again checking the consistency condition is immediate.

3. **Monoid** based on **Magma** – A monoid is a magma where the binary operation is associative and has an identity, so the underlying magma is just the monoid, forgetting the extra properties. Additionally a monoid homomorphism is just a magma homomorphism between monoids that takes identity to identity, so the underlying homomorphism is just itself. Again checking the consistency condition is immediate.

4. When $\mathscr{C}$ is a category with finite products we have the category $\mathbf{Group}_{\mathscr{C}}$ which is based on $\mathscr{C}$: For each object $(G,\ \mu,\ \zeta,\ \iota)$ in $\mathbf{Group}_{\mathscr{C}}$, the underlying object $U(G, \mu, \zeta, \iota)$ in $\mathscr{C}$ is G and each group morphism is just a morphism in $\mathscr{C}$ anyway (for which certain diagrams commute). Checking the consistency conditions is clear and immediate.

5. Pretty much the same, when $\mathscr{C}$ is a category with finite products, $\mathbf{Magma}_{\mathscr{C}}$ is based on $\mathscr{C}$: For each object (M, μ) in $\mathbf{Magma}_{\mathscr{C}}$, the underlying object in $\mathscr{C}$ is, of course, M and for each magma morphism f the corresponding morphism $U(f)$ is just f in $\mathscr{C}$. Checking the consistency conditions is again immediate.

6. Again in just the same way, when $\mathscr{C}$ is a category with finite products, $\mathbf{Monoid}_{\mathscr{C}}$ is based on $\mathscr{C}$: For each object (M, μ, ζ) in $\mathbf{Monoid}_{\mathscr{C}}$, the underlying object in $\mathscr{C}$ is, of course, M and for each monoid morphism f the corresponding morphism $U(f)$ is just f in $\mathscr{C}$. Checking the consistency conditions is immediate.

7. And when $\mathscr{C}$ is a category with finite products so that there is both $\mathbf{Monoid}_{\mathscr{C}}$ and $\mathbf{Magma}_{\mathscr{C}}$ and the first is based on the second: For each object (M, μ, ζ) in $\mathbf{Monoid}_{\mathscr{C}}$, the underlying object in $\mathbf{Magma}_{\mathscr{C}}$ is (M, μ) and each monoid morphism is equally well a magma morphism. Checking the consistency conditions is immediate.

8. The same remarks apply in comparing $\mathbf{Group}_{\mathscr{C}}$ and $\mathbf{Monoid}_{\mathscr{C}}$: For each object $(G,\ \mu,\ \zeta,\ \iota)$ in $\mathbf{Group}_{\mathscr{C}}$, the underlying object in $\mathbf{Monoid}_{\mathscr{C}}$ is (G, μ, ζ) and each group morphism is equally well a monoid morphism. And again checking the consistency conditions is immediate.

9. **LieGroup** is based on **Manifold**, and this is just a special case of $\mathbf{Group}_{\mathscr{C}}$ being based on $\mathscr{C}$ by taking $\mathscr{C}$ to be **Manifold** whereupon $\mathbf{Group}_{\mathscr{C}}$ is the category **LieGroup**.

10. $\mathbf{Module_R}$ based on **Ab** for every R-module is just an Abelian group M together with a suitable pairing $R \times M \longrightarrow M$, so the underlying Abelian group is just M. And every module homomorphism is an Abelian group homomorphism that respects the pairing, so the underlying
consistency conditions is clear and immediate.

C.3 Solutions for Chapter III

Solution for III.1 on p. 81: To say that Hom is a bifunctor from $\mathscr{C}$ and $\mathscr{C}$ to **Set** which is contravariant in the first position means just that Hom may be regarded as a functor from $\mathscr{C}^{\text{op}} \times \mathscr{C}$ to **Set**. To verify this we must show what Hom does on objects and morphisms, and confirm the relevant identities. On an object $(C_1,\ C_2)$ in $\mathscr{C}^{\text{op}} \times \mathscr{C}$, we have $\text{Hom}(C_1, C_2)$, the set of morphisms in $\mathscr{C}$ from C_1 to C_2. Now a morphism $(f, g) : (D_1, C_1) \longrightarrow (D_2, C_2)$ in $\mathscr{C}^{\text{op}} \times \mathscr{C}$ is a pair consisting $f : D_1 \longrightarrow D_2$ in $\mathscr{C}^{\text{op}}$ and $g : C_1 \longrightarrow C_2$ in $\mathscr{C}$. Of course the morphism f in $\mathscr{C}^{\text{op}}$ is the same as a morphism $f : D_2 \longrightarrow D_1$ in $\mathscr{C}$.

The Hom bifunctor applied to (f, g) is $\text{Hom}(f, g) : Hom(D_1, C_1) \longrightarrow \text{Hom}(D_2, C_2)$ with $\text{Hom}(f, g)(h) = ghf$. So clearly $\text{Hom}(1_D, 1_C)$ is the identity function on $\text{Hom}(D, C)$, while for $(p, q) : (D_2, C_2) \longrightarrow (D_3, C_3)$ we have $\text{Hom}(fp, qg)(h) = (qg)h(fp) = q(ghf)p = \text{Hom}(p, q)\text{Hom}(f, g)(h)$. ∎

Solution for III.2 on p. 83: To find example of projection functors that are not faithful or not full, recall that

$$\mathscr{A} \times \mathscr{B}((A, B), (A', B')) = \mathscr{A}(A, A') \times \mathscr{B}(B, B')$$

and the projection functor $\pi_1 \colon \mathscr{A} \times \mathscr{A} \longrightarrow \mathscr{A}$ induces the projection function $\mathscr{A}(A, A') \times \mathscr{B}(B, B') \longrightarrow \mathscr{A}(A, A')$. So we need sets S and T where $\pi_1 \colon S \times T \longrightarrow S$ is not injective (or surjective, respectively) and then we need categories where the Hom sets are the appropriate S and T. The first part we get by taking T to be the empty set $\emptyset$ and S any non-empty set (a one element set as a specific example,) for then $S \times T = \emptyset$ as well, so $\pi_1 \colon S \times T \longrightarrow S$ is the empty function from $\emptyset$ to the non-empty set S and is neither injective nor surjective. For the categories we may take both $\mathscr{A}$ and $\mathscr{B}$ to be the category of sets **Set** as for any set S we have $\mathbf{Set}(1, S) \cong S$, whence

$$\pi_1 \colon \mathbf{Set} \times \mathbf{Set}((1, 1), (1, 0)) \longrightarrow \mathbf{Set}(1, 1)$$

is an example showing the function π_1 is neither faithful nor full. [Here 0 is the initial object $\emptyset$ and 1 is the final object $\{0\}$.]

$\blacksquare$

Solution for III.3 on p. 85: For a fixed object C in a category $\mathscr{C}$ with finite products, we want to verify that taking $F(A) = A \times C$ for every object A and $F(f) = f \times 1_C$ for every morphism $f : A \longrightarrow B$ defines a functor from $\mathscr{C}$ to itself. Certainly $F(1_A) = 1_{F(A)}$ by exercise I.40, while $F(gf) = F(g)F(f)$ by exercise I.41. $\blacksquare$

Solution for III.4 on p. 85: The functor $+C : \mathscr{C} \longrightarrow \mathscr{C}$ takes an object B to the object $B + C$ and a morphism $f : B \longrightarrow B'$ to the morphism $f + 1_C : B + C \longrightarrow B' + C$. [This morphism is also often written $f + C$.] Certainly $1_B + 1_C = 1_{B+C}$ by exercise I.54, while $f'f + 1_C = (f' + 1_C)(f + 1_C)$ by exercise I.55 and those are the two parts needed to verify that $+C$ is a functor.

$\blacksquare$

Compare these last two solutions. This is yet another example of duality, where "reversing arrows" also means interchanging sums and products.

Solution for III.5 on p. 86: In a category $\mathscr{C}$ with finite sums, define $F : \mathscr{C} \times \mathscr{C} \longrightarrow \mathscr{C}$ by $F(C_1, C_2) = C_1 + C_2$ and $F(f_1, f_2) = f_1 + f_2$, then F is a functor from $\mathscr{C} \times \mathscr{C}$ to itself. The fact that $F(1_C, 1_c) = 1_{F(C)}$ is the content of exercise I.54, while $F((g_1, g_2)(f_1, f_2)) = F(g_1, g_2)F(f_1, f_2)$ is immediate from exercise I.55. $\blacksquare$

Solution for III.6 on p. 86: In a category $\mathscr{C}$ with finite products, define $F : \mathscr{C} \times \mathscr{C} \longrightarrow \mathscr{C}$ by $F(C_1, C_2) = C_1 \times C_2$ and $F(f_1, f_2) = f_1 \times f_2$, then F is a functor from $\mathscr{C} \times \mathscr{C}$ to itself. The fact that $F(1_{C_1}, 1_{C_2}) = 1_{F(C_1, C_2)}$ is the content of exercise I.40, while $F((g_1, g_2)(f_1, f_2)) = F(g_1, g_2)F(f_1, f_2)$ is immediate from exercise I.41.

▮

Compare these last two solutions. This is yet another example of duality, where in this case "reversing arrows" shows up as just interchanging sums and products.

Solution for III.7 on p. 86: We verify that the power set $\mathcal{P} : \mathbf{Set} \longrightarrow \mathbf{Set}$ is a functor by noting that $\mathcal{P}(1_X)(S) = 1_X(S) = S$ for every $S \subseteq X$, so $\mathcal{P}(1_X) = 1_{\mathcal{P}(X)}$. While for $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ and $S \subseteq X$ we have:

$$\begin{aligned}\mathcal{P}(gf)(S) &= gf(S)\\ &= \{z \mid \exists x \in S, z = gf(x)\}\\ &= \{z \mid \exists y \in f(S), z = f(y)\}\\ &= \mathcal{P}(g)\mathcal{P}(f)(S)\end{aligned}$$

so $\mathcal{P}(gf) = \mathcal{P}(g)\mathcal{P}(f)$. ▮

Solution for III.8 on p. 86: We verify that the power set $P : \mathbf{Set} \longrightarrow \mathbf{Set}$ is a contravariant functor by noting that $\mathsf{P}(1_X)(S) = 1_X \inf(S) = S$ for every $S \subseteq X$, so $\mathsf{P}(1_X) = 1_{\mathsf{P}(X)}$. While for $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ and $T \subseteq Z$ we have:

$$\begin{aligned}\mathsf{P}(gf)(T) &= (gf)^{-1}(T)\\ &= \{x \mid gf(x) \in T\}\\ &= \{x \mid f(x) \in g^{-1}(T)\}\\ &= \mathsf{P}(f)\mathsf{P}(g)(T)\end{aligned}$$

so $\mathsf{P}(gf) = \mathsf{P}(f)\mathsf{P}(g)$. ▮

Solution for III.9 on p. 87: For the category $\mathbf{Monoid}_{\mathscr{C}}$ of monoids in the category $\mathscr{C}$ (with finite products) the forgetful functor $U : \mathbf{Monoid}_{\mathscr{C}} \longrightarrow \mathscr{C}$ is defined as $U(M, \mu, \zeta) = M$ for any monoid in $\mathscr{C}$ and $U(f) = f$ for f any monoid morphism. Verification that U is indeed a functor is completely trivial: $U(1_M) = 1_M = 1_{U(M)}$ and $U(gf) = gf = U(g)U(f)$ by the definition of the identity morphisms and composition in $\mathbf{Monoid}_{\mathscr{C}}$. ▮

Solution for III.10 on p. 87: The first part of verifying that $f^* : A^* \longrightarrow M$ is a monoid homomorphism is just the part of the definition which says $f^*(()) = 1$ as the empty sequence is the identity in A^*. For the second part we need to check that $f^*(a_1, a_2, \cdots, a_n, a_{n+1}, \cdots, a_{n+m}) = f^*(a_1, a_2, \cdots, a_n)f^*(a_{n+1}, \cdots, a_{n+m})$. But the left hand side of this is $f(a_1)f(a_2)\cdots f(a_n)f(a_{n+1})\cdots f(a_{n+m})$ while the right hand side is $f(a_1)f(a_2)\cdots f(a_n)f(a_{n+1})\cdots f(a_{n+m})$ and the equality of the two is a direct consequence of associativity of the multiplication in M.

Solution for III.11 on p. 88: The function $\mathbf{Set}(A, U(M)) \longrightarrow \mathbf{Monoid}(A^*, M)$ assigns to each function $f : AtoU(M)$ the homomorphism $f^* : A^* \longrightarrow M$. The inverse function assigns to each homomorphism $h : A^* \longrightarrow M$ the function

$$\begin{aligned} A &\longrightarrow U(M) \\ a &\mapsto h(a) \end{aligned}$$

where the a in $h(a)$ refers to the sequence consisting just of the one element a. Going from $\mathbf{Set}(A, U(M))$ to $\mathbf{Monoid}(A^*, M)$ and back to $\mathbf{Set}(A, U(M))$ clearly takes any function from A to $U(M)$ to itself. In the other direction note that if $h : A^* \longrightarrow M$ is a homomorphism, then $h(a_1\, a_2 \cdots a_n) = h(a_1)h(a_2)\cdots h(a_n)$ and this is exactly the homomorphism gotten by starting with h and taking it to $\mathbf{Set}(A, U(M))$ and then back to $\mathbf{Monoid}(A, M)$ again. $\blacksquare$

Solution for III.12 on p. 88: Verification that assigning to a commutative ring R its polynomial ring $R[X]$ and to a ring homomorphism $h : R \longrightarrow S$ its extension to the polynomial rings defines a functor from **CommutativeRing** to **CommutativeRing** is just directly checking the definition. First this certainly takes objects to object and morphisms to morphisms. Second $F(1_R) = 1_{R[X]} = 1_{F(R)}$, and third for homomorphisms $h : R \longrightarrow S$ and $k : S \longrightarrow T$ we have

$$\begin{aligned} F(kh)(\Sigma_{i=1}^n r_i X^i) &= \Sigma_{i=1}^n kh(r_i)X^i \\ &= K(\Sigma_{i=1}^n h(r_i)X^i) \\ &= K(H(\Sigma_{i=1}^n r_i X^i)) \\ &= F(k)F(h)(\Sigma_{i=1}^n r_i X^i) \end{aligned}$$

so F(kh) = F(k)F(h). $\blacksquare$

Solution for III.13 on p. 88: To prove that the commutator functor C : **Group** $\longrightarrow$ **Group** is indeed a functor is just a matter of directly checking the definition. First C certainly takes objects to object and morphisms. Second $C(1_G) = 1_G|_{[G,G]} = 1_{[G,G]} = 1_{C(G)}$, and third for $f : G \longrightarrow H$ and $f' : H \longrightarrow K$ we have $C(f'f) = C(f')C(f)$ because $f'f|_{[G,G]} = f'|_{[H,H]}f|_{[G,G]}$ and that follows from the fact that $f([G, G]) \subseteq [H, H]$. $\blacksquare$

Solution for III.14 on p. 89: We want to verify the Abelianizer A : **Group** $\longrightarrow$ **Ab** is a functor. A is given on objects by $A(G) = G/[G, G]$ and on a group homomorphism $f : G \longrightarrow H$ by $A(f) = \overline{f}$ with $\overline{f}$ being the unique homomorphism with $p_H f = \overline{f} p_G$. Certainly $A(1_G)$ is $1_{A(G)}$. While if we have homomorphism $f : G \longrightarrow H$ and $e : H \longrightarrow K$, then certainly $p_K ef = \overline{e} p_H f = \overline{e}\overline{f} p_G$ so $\overline{e}\overline{f} = \overline{ef}$ which is the same as saying that $A(ef) = A(e)A(f)$ and so A is indeed a functor.

$\blacksquare$

Solution for III.15 on p. 89: The function φ : $\mathbf{Group}(G, I(H)) \longrightarrow \mathbf{Ab}(A(G), H)$ was defined in the text, and here we wish to show it is a bijection. The easiest way is to explicitly exhibit its inverse ψ : $\mathbf{Ab}(A(G), H)$

$\longrightarrow \mathbf{Group}(G, I(H))$. So starting with a homomorphism $e \in \mathbf{Ab}(A(G), H)$ notice that $ep_G \in \mathbf{Group}(G, I(H))$ and $\varphi(ep_G) = e$ because $p_G([G,G]) = \{0\}$. So we define $\psi(e) = ep_G$ and we have already verified that $\varphi\psi$ is the identity on $\mathbf{Ab}(A(G), H)$. For the other direction recall that the defining property of $\varphi(f) = \overline{f}$ is that it is the *unique* homomorphism with $\overline{f}p_G = f$. But then $\psi(\overline{f})$ is, by definition, $\overline{f}p_G$ and that is f. Thus we have also verified that $\psi\varphi$ is the identity on $\mathbf{Group}(G, I(H))$.

∎

Solution for III.16 on p. 89: For the category **Top** we have the forgetful functor $U : \mathbf{Top} \longrightarrow \mathbf{Set}$ which takes a topological space $(X, \mathcal{T}(X))$ (where $\mathcal{T}(X)$ is the topology) to the underlying set X and a continuous function $f : (X, \mathcal{T}(X)) \longrightarrow (Y, \mathcal{T}(Y))$ to the function $f : X \longrightarrow Y$ where we are simply forgetting it is continuous. The forgetful functor patently takes identity functions (which are always continuous no matter the topology) to identity functions, and it also preserves composition because composition of continuous functions is after all just composition of functions.

We also have the discrete topology functor $F : \mathbf{Set} \longrightarrow \mathbf{Top}$ where $F(X) = (X, \mathcal{P}(X))$ and $F(f) = f$. This makes sense as for any set X the set $\mathcal{P}(X)$ of all its subsets is closed under arbitrary unions and finite intersections, i.e., it is a topology. Moreover any function $f : X \longrightarrow Y$ is continuous for any topology on Y as the inverse image of *every* subset of Y is in $\mathcal{P}(X)$. And it is a functor for clearly $F(1_X) = 1_{F(X)}$ and $F(gf) = F(g)F(f)$.

Finally the bijection between $\mathbf{Top}(F(X), Y)$ and $\mathbf{Set}(X, U(Y))$ is nothing more than reading the previous two paragraphs — every function from a set X to the underlying set of some topological space Y *is* a continuous function from X with the discrete topology to the topological space Y, and conversely.

∎

Solution for III.17 on p. 94: The verification that $\hat{f} : F(G) \longrightarrow \mathscr{C}$ is always a functor has the usual steps. The definition of $\hat{f}$ shows it assigns to each object of $F(G)$ an object of $\mathscr{C}$, and to each morphism of $F(G)$ a morphism in $\mathscr{C}$. Additionally if $p : v_0 \longrightarrow v_n$ in $F(G)$ then clearly $\hat{f}(p) : \hat{v_0} \longrightarrow \hat{v_n}$. Finally for paths $p = (e_1, \cdots, e_n)$ from $v_0 (= \mathrm{init}(e_1))$ to $v_n (= \mathrm{ter}(e_n))$ and $q = (e_{n+1}, \cdots, e_m)$ from $v_n (= \mathrm{init}(e_{n+1}))$ to $v_m (= \mathrm{ter}(e_m))$ the composition qp is the path $(e_1, \cdots, e_n, e_{n+1}, \cdots, e_m)$ and

$$\begin{aligned} \hat{f}(qp) &= (f_E(e_1) \cdots f_E(e_n) f_E(e_{n+1}) \cdots, f_E(e_m)) \\ &= (f_E(e_1), \cdots, f_E(e_n))(f_E(e_{n+1}) \cdots f_E(e_m)) \\ &= \hat{f}(q)\hat{f}(p) \end{aligned}$$

∎

Solution for III.18 on p. 94: The exercise here is to prove that the function

$$\begin{array}{ccc} \mathbf{Digraph}(G, U(\mathscr{C})) & \longrightarrow & \mathbf{Cat}(F(G), \mathscr{C}) \\ f & \mapsto & \hat{f} \end{array}$$

is a bijection.

The inverse function takes a functor $k : F(G) \longrightarrow \mathscr{C}$ to the digraph homomorphism $\check{k} : G \longrightarrow \mathscr{C}$ that just forgets about composition, i.e., $\check{k}_V(v) = k(v)$ and $\check{k}_E(e) = k(e)$. That $\check{k}$ is a digraph homomorphism is an immediate consequence of the properties in the definition of a functor not involving composition. The computation of $\check{\hat{f}}$ is also immediate: $\check{\hat{f}}_V(v) = \hat{f}(v) = f_V(v)$ and $\check{\hat{f}}_E(e)\hat{f}(e) = f_E(e)$, so $\check{\hat{f}} = f$. Also if $g : C \longrightarrow D$ is any morphism in $\mathscr{C}$, then $\hat{\check{k}}(C) = \check{k}_V(C) = k(C)$ and $\hat{\check{k}}(g) = \check{k}_E(f) = k(g)$. so $\hat{\check{k}} = k$.
∎

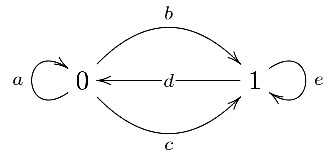
Solution for III.19 on p. 96: In the terminology of the preceding exercise, the canonical functor $\varepsilon : F(U(\mathbf{2})) \longrightarrow \mathbf{2}$ is $\widehat{1_{U(\mathbf{2})}}$, and we will use the description found above.

The category **2** has the two objects 0 and 1, the two identity morphisms 0, 1 (with the same names as their corresponding objects) and the one non-identity morphism $! : 0 \longrightarrow 1$ (so named because it is the unique morphism from the initial object 0 to the final object 1.) The underlying graph $U(\mathbf{2})$ has the two nodes 0 and 1 and the three edges: 0 from 0 to 0; 1 from 1 to 1 and ! from 0 to 1. So $F(U(\mathbf{2}))$ has the two objects 0 and 1 and as morphisms all sequences of the following forms:

1. an empty path on 0
2. ! as a path from 0 to 1
3. an empty path on 1
4. any path of positive length starting and ending at 0: $(0, \cdots, 0)$
5. any path of positive length starting and ending at 1: $(1, \cdots, 1)$
6. the other paths from 0 to 1: $(0, \cdots, 0, !, 1, \cdots, 1)$

From the above description we know that $\varepsilon = \widehat{1_{U(\mathbf{2})}}$ takes the objects 0 to 0 and 1 to 1, while on the morphisms it takes every path starting and ending at 0 to the identity morphism 0, every path starting and ending at 1 to the identity morphism 1, and every path from 0 to 1 to the morphism !.
∎

Solution for III.20 on p. 96: In order to describe in more detail the free category generated by the following digraph we have added names for the five edges.



In the free category $\mathscr{F}$ there are, of course, just the two objects and 1. So it is convenient to describe the morphisms of $\mathscr{F}$ by partitioning them into the four sets $\mathscr{F}(0,0)$, $\mathscr{F}(1,1)$, $\mathscr{F}(0,1)$, and $\mathscr{F}(1,1)$. The constructive specification of $\mathscr{F}$ shows that $\mathscr{F}(0,0)$ is the free monoid on the three generators a, db and dc. Similarly $\mathscr{F}(1,1)$ is the free monoid on the three generators e, bd and cd.

The two sets $\mathscr{F}(0,1)$ and $\mathscr{F}(1,0)$ have somewhat more complicated descriptions, but they are very similar to one another. A morphism in $\mathscr{F}(0,1)$ can be described by the following "regular expression": $\{A(b|c)Ed\}^+$ where A denotes any "word" from $\mathscr{F}(0,0)$, E is any "word" from $\mathscr{F}(1,1)$, $(b|c)$ means one of b or c and $\{\cdots\}^+$ signifies that the expression inside the braces is repeated one or more times.

The morphisms in $\mathscr{F}(1,0)$ are described by this regular expression: $\{EdA(b|c)\}^+$ ∎

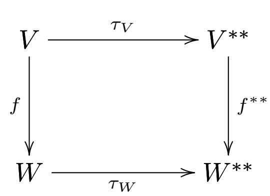
Solution for III.21 on p. 97: When $f : D \longrightarrow D'$ is a morphism in the category $\mathscr{D}$ it induces in the obvious way a natural transformation, which we will also denote by f, between the associated constant functors – for each object C of $\mathscr{C}$, $f_C : D(C) \longrightarrow D'(C)$ is just f. And if $\eta : D \longrightarrow D'$ is a natural transformation between the constant functors selecting D and D', then η_C must be the same morphism from $D(C) = D$ to $D'(C) = D'$ for every object C of $\mathscr{C}$ and this correspondence is clearly bijective. ∎

Solution for III.22 on p. 98: If $f : V \longrightarrow W$ is a linear transformation, then we have the dual linear transformation $D(f) : D(W) \longrightarrow D(V)$ which we will write as $f^* : W^* \longrightarrow V^*$. (Notice this is consistent with definition I.8 and the notation we have used elsewhere.)

Writing v^* and w^* for arbitrary elements of V^* and W^* respectively, we see that $f^*(w*)$ is in V^* and is defined by $f^*(w^*)(v) = w^*(f(v))$

Continuing in this vein and writing v^{**} and w^{**} for arbitrary elements of V^{**} and W^{**} respectively, we see that $f^{**}(v^{**})(w^*) = v^{**}(f^*(w^*)) = v^{**}(w^* f)$

Now $\tau : 1_{\mathscr{V}} \longrightarrow DD$ was defined by $\tau_V(v)(v^*) = v^*(v)$ and we want to verify the following square commutes:



Checking we see

$$\begin{aligned}
(f^{**}\tau_V)(v)(w^*) &= f^{**}(\tau_V(v))(w^*) \\
&= \tau_V(v)(f^*(w^*)) \\
&= \tau_V(v)(w^*f) \\
&= (w^*f)(v) \\
&= w^*(f(v)) \\
&= \tau_W(f(v))(w^*) \\
&= (\tau_W f)(v)(w^*)
\end{aligned}$$

The result that $\tau_V : V \longrightarrow V^{**}$ is always injective requires a bit of special knowledge about vector spaces, to wit the result that if v is any non-zero element of V, then there exists an element v^* of V^* with $v^*(v) \neq 0$. This is usually proved using the observation that every non-zero v is part of a basis for V, for then if B is such a basis we can define $v^*(\Sigma_{b\in B} c_b b) = c_v$ with c_v the coefficient of $v \in V$. For that we see that if $\tau_V(v) = 0$, then $\tau_V(v)(v^*) = 0$ for all $v^* \in V^*$ which by the above observation implies $v = 0$.

∎

Solution for III.23 on p. 98: Recall (cf. Section III.2.12) the construction of the free monoid $F(S)$ on the set S as the set of finite sequence of elements of S with the binary operation of concatenation. For $f : S \longrightarrow T$ the homomorphism $F(f) : F(S) \longrightarrow F(T)$ is given by $F(f)(s_0, \cdots, s_n) = (f(s_0), \cdots, f(s_n))$.

The natural transformation $\eta : 1_{\mathbf{Set}} \longrightarrow UF$ is defined as $\eta_S(s) = (s)$ with (s) the sequence with just one term, s. To verify that η is a natural transformation we must show that the following diagram always commutes.

$$\begin{array}{ccc}
S & \xrightarrow{\eta_S} & UF(S) \\
{\scriptstyle f}\downarrow & & \downarrow{\scriptstyle UF(f)} \\
T & \xrightarrow[\eta_T]{} & UF(T)
\end{array}$$

With all of this the actual verification is immediate:

$$UF(f)(\eta_S(s)) = UF(f)((s)) = (f(s))$$
$$\eta_T(f(s)) = (f(s))$$

∎

Solution for III.24 on p. 99: The natural transformation $\varepsilon : FU \longrightarrow 1_{\mathbf{Monoid}}$ is defined by $\varepsilon_M(m_1, m_2, \cdots, m_n) = m_1 m_2 \cdots m_n$ To verify that ε is a natural transformation we must show that the following diagram always commutes:

$$\begin{array}{ccc} FU(M) & \xrightarrow{\varepsilon_M} & M \\ {\scriptstyle FU(f)}\downarrow & & \downarrow{\scriptstyle f} \\ FU(N) & \xrightarrow[\varepsilon_N]{} & N \end{array}$$

That is just the following simple calculation:

$$f(\varepsilon_M(m_1, m_2, \cdots, m_n)) = f(m_1 m_2 \cdots m_n) = f(m_1)f(m_2)\cdots f(m_n)$$
$$\varepsilon_N(FU(f)((m_1 m_2 \cdots m_n))) = \varepsilon_N((f(m_1)f(m_2)\cdots f(m_n))) = f(m_1)f(m_2)\cdots f(m_n)$$

∎

Solution for III.25 on p. 99: In Section III.2.12 we defined a function from $\mathbf{Set}(A, U(M))$ to $\mathbf{Monoid}(F(A), M)$ by assigning to each $f : A \longrightarrow U(M)$ the monoid homomorphism $f^* : A^* \longrightarrow M$ with $f^*(()) = 1$ and $f^*(a_1, a_2, \cdots, a_n) = f(a_1)f(a_2)\cdots f(a_n)$

Now we also have the function $\mathbf{Set}(A, U(M)) \longrightarrow \mathbf{Monoid}(F(A), M)$ given by $f \mapsto \varepsilon_M F(f)$ where $\varepsilon : FU \longrightarrow 1_{\mathbf{Monoid}}$ is defined as $\varepsilon_M(m_1, m_2, \cdots, m_n) = m_1 m_2 \cdots m_n$, and this exercise is to verify that these are the same functions. The verification consists of explicitly evaluating $\varepsilon_M F(f)$: $\varepsilon_M(F(f)(a_1, a_2, \cdots, a_n)) = \varepsilon_M(f(a_1), f(a_2), \cdots, f(a_n)) = f(a_1)f(a_2)\cdots f(a_n)$ ∎

Solution for III.26 on p. 99: As a preliminary step, if we consider the function $\mathbf{Monoid}(F(A), M) \longrightarrow \mathbf{Set}(A, U(M))$ given by $h \mapsto U(h)\eta$ and write $\hat{h}$ for $U(h)\eta$ we immediately see that $\hat{h}(a) = h(a)$. [Beware of the change of type between the two occurrences of a – the first is a as an element of the set A, while the second is actually a as the single element in the one element sequence (a).]

Combining this observation with the immediately preceding exercise we see that the composition

$$\mathbf{Set}(A, U(M)) \longrightarrow \mathbf{Monoid}(F(A), M) \longrightarrow \mathbf{Set}(A, U(M))$$

assigns to a function f the function $\hat{f^*}$ and $\hat{f^*}(a)$ is just $f(a)$, i.e., $\hat{f^*} = f$.

Similarly the composition

$$\mathbf{Monoid}(F(A), M) \longrightarrow \mathbf{Set}(A, U(M)) \longrightarrow \mathbf{Monoid}(F(A), M)$$

assigns to a homomorphism h the homomorphism $\hat{h}^*$ and $\hat{h}^*(s)$ is just $h(s)$, i.e., $\hat{h}^* = h$. So the function $\mathbf{Monoid}(F(A), M) \longrightarrow \mathbf{Set}(A, U(M))$ given by $h \mapsto U(h)\eta$ is the inverse of the function of the preceding exercise. ∎

The above verification is as simple and direct as can be imagined, but there is a little more complex approach which has the advantage of connecting directly with the general study of adjunctions and adjoint functors, so we offer it as well.

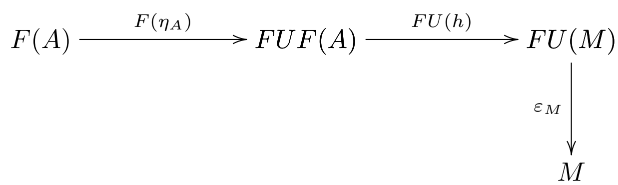
First observe that

$$F(A) \xrightarrow{F(\eta_A)} FUF(A) \xrightarrow{\varepsilon_{F(A)}} F(A)$$
$$U(M) \xrightarrow[\eta_{U(M)}]{} UFU(M) \xrightarrow[U(\varepsilon_M)]{} U(M)$$

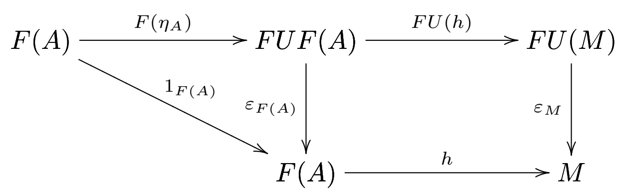
are immediately verified to be identity morphisms. Next note that the composition:

$$\mathbf{Set}(A, U(M)) \longrightarrow \mathbf{Monoid}(F(A), M) \longrightarrow \mathbf{Set}(A, U(M))$$

assigns to a function f the following composite function.

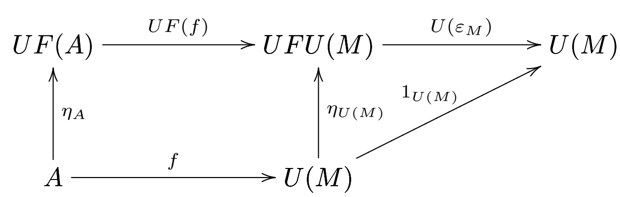


Adding a few more morphisms gives the following diagram as ε is a natural transformation.



Of course this says exactly that going from $\mathbf{Set}(A, U(M))$ to $\mathbf{Monoid}(F(A), M)$ and back to $\mathbf{Set}(A, U(M))$ is the identity.

In much the same way the following commutative diagram shows that going from $\mathbf{Monoid}(F(A), M)$ to $\mathbf{Set}(A, U(M))$ and back to $\mathbf{Monoid}(F(A), M)$ is the identity.



▮

Solution for III.27 on p. 99: The natural transformation $\phi_{A,M} : \mathbf{Set}(A, U(M)) \longrightarrow \mathbf{Monoid}(F(A), M)$ is defined by $\phi_{A,M}(f) = \varepsilon_M F(f)$, while the natural transformation $\psi_{A,M} : \mathbf{Monoid}(F(A), M) \longrightarrow \mathbf{Set}(A, U(M))$ is defined by $\psi_{A,M}(h) = U(h)\eta_A$. By the last exercise these are inverse functions, so the only work required to verify that they give a natural isomorphism is to verify that they are in fact natural transformations, which means verifying that for all morphisms $f : A \longrightarrow B$ and $h : M \longrightarrow N$ the following diagrams are commutative.

$$\begin{array}{ccc}
\mathbf{Set}(A, U(M)) & \xrightarrow{\phi_{A,M}} & \mathbf{Monoid}(F(A), M) \\
\Big\downarrow{\scriptstyle (f, U(h))_*} & & \Big\downarrow{\scriptstyle (F(f), h)_*} \\
\mathbf{Set}(B, U(N)) & \xrightarrow[\phi_{B,N}]{} & \mathbf{Monoid}(F(B), N)
\end{array}$$

$$\begin{array}{ccc}
\mathbf{Monoid}(F(A), M) & \xrightarrow{\psi_{A,M}} & Set(A, U(M)) \\
\Big\downarrow{\scriptstyle (F(f), h)_*} & & \Big\downarrow{\scriptstyle (f, U(h))_*} \\
\mathbf{Monoid}(F(B), N) & \xrightarrow[\psi_{B,N}]{} & \mathbf{Set}(B, U(N))
\end{array}$$

For the first diagram, notice that the result of starting with $b \in \mathbf{Set}(A, U(M))$ and following it across the top and down on the right is $h\varepsilon_M F(b)F(f)$ while the result of going down on the left and across the bottom is $\varepsilon_N FU(h)F(b)F(f)$. As ε is a natural transformation we know that $\varepsilon_N FU(h) = h\varepsilon_M$ and so the diagram is commutative.

For the second diagram, notice that the result of starting with $k \in \mathbf{Monoid}(F(A), M)$ and following it across the top and down on the right is $U(h)U(k)\eta_A f$ while the result of going down on the left and across the bottom is $U(h)U(k)UF(f)\eta_A$. As η is a natural transformation we know that $\eta_A f = UF(f)\eta_A$ and so the diagram is commutative.

∎

Solution for III.28 on p. 100: Verification that inclusion of the commutator subgroup into the containing group is a natural transformation is just a matter of displaying the relevant diagram:

$$\begin{array}{ccc}
[G, G] & \xrightarrow{\iota_G} & G \\
\Big\downarrow{\scriptstyle f|_{[H,H]}} & & \Big\downarrow{\scriptstyle f} \\
[H, H] & \xrightarrow[\iota_h]{} & H
\end{array}$$

∎

Solution for III.29 on p. 100: Verification that projection onto the abelianizer of a group is a natural transformation is just a matter of displaying the relevant diagram:

$$\begin{array}{ccc} G & \xrightarrow{\pi_G} & G/[G,G] \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle \overline{f}} \\ H & \xrightarrow[\pi_H]{} & H/[H,H] \end{array}$$

∎

Solution for III.30 on p. 100: This is one of those cases where the formalism complicates a very simple situation. The bijection between $\mathbf{Ab}(A(G), A)$ and $\mathbf{Group}(G, I(A))$ is just the simple observation that any group homomorphism from an arbitrary group G into an Abelian group factors uniquely through $A(G)$ as group homomorphisms preserve commutators and every commutator in a Abelian group is the identity. ∎

Solution for III.31 on p. 100: Verification that $\mathbf{Ab}(A(\bullet), \bullet)$ is naturally equivalent to $\mathbf{Group}(\bullet, I(\bullet))$ just consists of showing that the functions discussed in the last exercise are actually natural transformations, which means verifying that for all morphisms $f : H \longrightarrow F$ and $h : A \longrightarrow B$ the following squares are commutative.

$$\begin{array}{ccc} \mathbf{Ab}(A(G), A) & \xrightarrow{\phi_{G,A}} & \mathbf{Group}(G, I(A)) \\ {\scriptstyle (A(f),h)_*}\downarrow & & \downarrow{\scriptstyle (f,I(h))_*} \\ \mathbf{Ab}(A(H), B) & \xrightarrow[\phi_{H,B}]{} & \mathbf{Group}(H, I(B)) \end{array}$$

$$\begin{array}{ccc} \mathbf{Group}(G, I(A)) & \xrightarrow{\psi_{G,A}} & \mathbf{Ab}(A(G), A) \\ {\scriptstyle (f,I(h))_*}\downarrow & & \downarrow{\scriptstyle (A(f),h)_*} \\ \mathbf{Group}(H, I(B)) & \xrightarrow[\psi_{H,B}]{} & \mathbf{Ab}(A(H), B) \end{array}$$

$\phi_{G,A}$

Show that $\mathbf{Ab}(A(\bullet), \bullet)$ is naturally equivalent to $\mathbf{Group}(\bullet, I(\bullet))$ with both considered as functors from $\mathbf{Group}^{\mathrm{op}} \times \mathbf{Ab}$ to $\mathbf{Set}$.

Solution for III.32 on p. 101: Verify that $\varepsilon : FU \longrightarrow 1_{\mathbf{Top}}$ is indeed a natural transformation.

Solution for III.33 on p. 101: Show that the two functions just defined are inverse to one another and give bijections between $\mathbf{Top}(F(S), X)$ and $\mathbf{Set}(S, U(X))$.

Solution for III.34 on p. 101: Show that $\mathbf{Top}(F(\bullet), \bullet)$ is naturally isomorphic to $\mathbf{Set}(\bullet, I(\bullet))$ with both considered as functors from $\mathbf{Set}^{\mathrm{op}} \times \mathbf{Top}$ to $\mathbf{Set}$.

Solution for III.35 on p. 102: Verify that βF as just defined is indeed a natural transformation from GF to $G'F$.

Solution for III.36 on p. 102: Verify that $G\alpha$ as just defined is indeed a natural transformation from GF to GF'.

Solution for III.37 on p. 103: Show that for any category $\mathscr{C}$ the functor category $\mathscr{C}^{\mathbf{1}}$ is isomorphic to $\mathscr{C}$

Hint: Every functor $\mathbf{1} \longrightarrow \mathscr{C}$ takes the unique object in $\mathbf{1}$ to an object in $\mathscr{C}$, and that completely determines the functor, so all that is left is to note that a natural transformation between such functors is the same as a morphism between the objects.

Solution for III.38 on p. 103: If $\mathscr{D}$ is a discrete finite category with n objects, show that $\mathscr{C}^{\mathscr{D}} \cong \mathscr{C}^n$ where $\mathscr{C}^n$ is the n-fold product of $\mathscr{C}$ with itself.

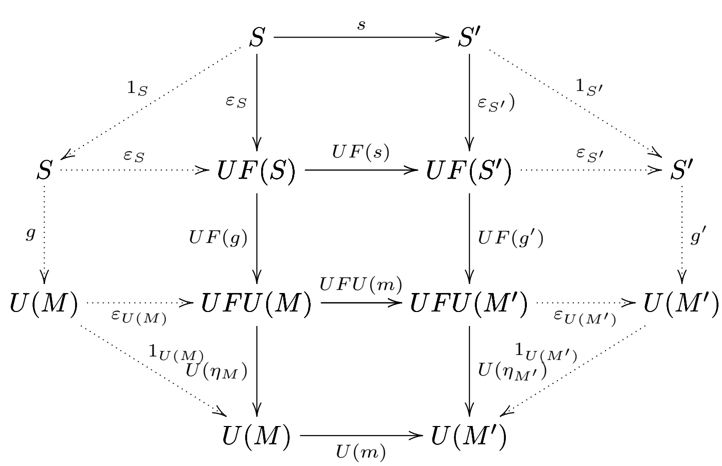
C.4 Solutions for Chapter IV

C.5 Solutions for Chapter V

Solution for V.5 on p. 144:

To see that the equation $(U\eta)(\varepsilon U) = 1_U$ implies $\theta\xi$ is the identity on $(\mathscr{S} \downarrow U)$, observe that the result of applying $\theta\xi$ to

Extend this to the following larger commutative diagram by adding the dotted arrows shown using the naturality of η and the assumption that $(\eta F)(F\varepsilon) = 1_F$



and this gives the result.

C.6 Solutions for Chapter VII

C.7 Solutions for Chapter VIII

C.8 Solutions for Chapter XIII

C.9 Solutions for Chapter XIV

C.10 Solutions for Chapter IX

C.11 Solutions for Chapter VI

Solution for VI.1 on p. 147: Verify that if the product $C \times C$ exists, then d_0, d_1 are jointly monic iff $\langle d_0, d_1 \rangle : R \longrightarrow C \times C$ is monic.

Solution for VI.2 on p. 148: Suppose that in $\mathscr{C}$ the pair $d_0, d_1 : R \longrightarrow C$ is a relation on C. Show that it is a reflexive relation iff $\text{Hom}(X, d_0), \text{Hom}(X, d_1) : \text{Hom}(X, R) \longrightarrow \text{Hom}(X, C)$ is a reflexive relation in **Set**.

C.12 Solutions for Appendix B

Solution for B.1 on p. 253: Suppose $\mathscr{P}$ is a small category with the property that for any two objects A and B of $\mathscr{P}$ there is at most one morphism from A to B. Define a relation on the objects of $\mathscr{P}$ by $A \preceq B$ iff there is a morphism from A to B. Demonstrate that the set of objects of $\mathscr{P}$ with the relation $\preceq$ is a partially ordered set. Similarly if

Solution for B.2 on p. 254: To prove that $1 + 1 = 1$ in **2**, note that there is only one morphism with domain 1, namely the identity, and it serves as the unique morphism from $1 + 1$ to 1. In the same fashion there is only one morphism with codomain 0, namely the identity, and it serves as the unique morphism from 0 to $0 \times 0 = 0$. ∎

Appendix D

Other Sources

Even though category theory is a relatively young part of mathematics, it has a huge literature. In particular there are a many accounts of the material in these notes, and the study of category theory benefits from this variety of perspectives. Here are listed a number of generally available books and notes, together with some brief comments on them. In particular all of these have been useful in preparing these notes.

D.1 Introductions

D.1.1 Abstract and Concrete Categories: The Joy of Cats (Adámek, Herrlich, and Strecker [?])

D.1.2 Category Theory (Awodey [?])

D.1.3 Category Theory: An Introduction (Herrlich and Strecker [?])

D.1.4 Arrows, Structures, and Functors: The Categorical Imperative (Arbib and Manes [1])

D.1.5 Category Theory: Lecture Notes for ESSLLI (Barr and Wells [?])

D.1.6 Category Theory for Computing Science (Barr and Wells [3])

D.1.7 Categories (Blyth [8])

D.1.8 Handbook of Categorical Algebra I: Basic Category Theory (Borceux [?])

D.1.9 Handbook of Categorical Algebra II: (Borceux [?])

D.1.10 Handbook of Categorical Algebra III: (Borceux [?])

D.1.11 Introduction to the Theory of Categories and Functors (Bucur and Deleanu [?])

D.1.12 Lecture Notes in Category Theory (Cáccamo [?])

D.1.13 Abelian Categories: An Introduction to the Theory of Functors (Freyd [?])

D.1.14 A Categorical Primer (Hillman [?])

D.1.15 An Introduction to Category Theory (Krishnan [?])

D.1.16 Conceptual Mathematics: A First Introduction to Categories (Lawvere and Schanuel [47])

D.1.17 Theory of Categories (Mitchell [?])

D.1.18 Categories and Functors (Pareigis [?])

D.1.19 Theory of Categories (Popescu and Popescu [63])

D.1.20 Categories (Schubert [?])

D.1.21 Introduction to Category Theory and Categorical Logic (Streichert [?])

D.2 For Computer Science

D.2.1 Categories, Types, and Structures: An Introduction to Category Theory for the Working Computer Scientist (Asperti and Longo [?]))

D.2.2 Categories for Types (Crole [?])

D.2.3 A Gentle Introduction to Category Theory — The Calculational Approach — (Fokkinga [?])

D.2.4 Basic Category Theory for Computer Scientists (Pierce [?])

D.2.5 Category Theory and Computer Programming (Pitt et. al. [?])

This is the proceedings of the "Category Theory and Computer Programming, Tutorial and Workshop" held in Guildford, UK, September 16-20, 1985. The primary interest here are the tutorials which present basic category theory from a computer science perspective.

- Andrzej Tarlecki: Bits and Pieces of the Theory of Institutions.
- Donald Sannella, Andrzej Tarlecki: Extended ML: an Institution - Independent Framework for Formal Program Development.
- Horst Reichel: Behavioral Program Specification.
- Hans-Dieter Ehrich: Key Extensions of Abstract Data Types, Final Algebras, and Database Semantics.

Section 3 : Categorical Logic
- Michael P. Fourman, Steven Vickers: Theories as Categories.
- Paul Taylor: Internal Completeness of Categories of Domains.
- John Cartmell: Formalizing the Network and Hierarchical Data Models - an Application of Categorical Logic.

Section 4 : Categorical Programming
- David E. Rydeheard, Rod M. Burstall: A Categorical Unification Algorithm.
- David E. Rydeheard, Rod M. Burstall: Computing with Categories.

D.2.6 Categories and Computer Science (Walters [?])

Introduction
1 The Algebra of Functions
  a) Categories
  b) General Examples
  c) Free Categories; Generators and Relations
  d) Some Large Categories
  e) The Dual of a Category
2 Products and Sums
  a) Initial and Terminal Objects
  b) Categories with Products; Circuits
  c) Products of Families
  d) Sums
  e) Categories with Sums; Flow Charts
3 Distributive Categories
  a) The Distributive Law
  b) Examples
  c) Imperative Programs
4 Data Types
  a) Arithmetic
  b) Stacks
  c) Arrays
  d) Binary Trees
  e) Queues
  f) Pointers
  g) Turing Machines
5 Categories of Functors
  a) Functors
  b) Functor Categories
  c) Directed Graphs and Regular Grammars
  d) Automata and Imperative Programs with Input
  e) The Specification of Functions
  f) What Does *Free* Mean?
  g) Adjoint Functors
6 More About Products
  a) The Free Category with Products
  b) Functional Specification with Products
  c) Context-free Languages
  d) Natural Numbers and Cartesian Closed Categories
7 Computational Category Theory
  a) The Knuth-Bendix Procedure
  b) Computing Left Kan Extensions

D.3 Advanced - Specialized

D.3.1 Theory of Mathematical Structures (Adámek [?])

Part I: CONSTRUCTS

D.3.2 Toposes, Triples, and Theories (Barr and Wells [?])

D.3.3 An Invitation to General Algebra and Universal Constructions (Bergman [?])

D.3.4 Toposes and Local Set Theories (Bell [?])

D.3.5 Categories, Allegories (Freyd and Scedrov [25])

1.6.8 GENERATING SET, BASIS
1.6.9 PRE-FILTER, FILTER
1.6.10 REPRESENTATION THEOREM FOR PRE-LOGOI, BOOLEAN ALGEBRA, ULTRA-FILTER
1.6.11 special pre-logos
1.6.12 well-joined category
1.6.13 BOOLEAN PRE-LOGOS
1.6.14 ULTRA-PRODUCT FUNCTOR, ULTRA-POWER FUNCTOR
1.6.15 properness of a subobject
1.6.16 COMPLETE MEASURE, ATOMIC MEASURE
1.6.17 PRE-TOPOS
1.6.18 AMALGAMATION LEMMA
1.6.19 DECIDABLE OBJECT
1.6.20 DIACONESCU BOOLEAN THEOREM

1.7 Logoi
1.7.1 LOGOS
1.7.2 LOCALLY COMPLETE CATEGORY
1.7.3 HEYTING ALGEBRA
1.7.4 LOCALE, category of complete Heyting algebras, category of locales
1.7.5 NEGATION
1.7.6 LAW OF EXCLUDED MIDDLE
1.7.7 scone of a Heyting algebra
1.7.8 free Heyting algebra, RETRACT
1.7.9 slice lemma for logoi
1.7.10 COPRIME OBJECT, CONNECTED OBJECT, FOCAL LOGOS
1.7.11 FOCAL REPRESENTATION
1.7.12 GEOMETRIC REPRESENTATION THEOREM FOR LOGOI
1.7.13 DOMINATES, LEFT-FULL
1.7.14 FREYD CURVE
1.7.15 STONE REPRESENTATION THEOREM FOR LOGOI
1.7.16 ATOM, ATOMICALLY BASED, ATOMLESS, periodic power
1.7.17 STONE SPACE, CLOPEN
1.7.18 MICRO-SHEAF
1.7.19 TRANSITIVE CLOSURE, TRANSITIVE-REFLEXIVE CLOSURE, TRANSITIVE (PRE-)LOGOS
1.7.20 σ-TRANSITIVE LOGOS, σ-TRANSITIVE PRE-LOGOS
1.7.21 EQUIVALENCE CLOSURE, E-STANDARD PRE-LOGOS
1.7.22 representation theorem for countable σ-transitive (pre-)logoi

1.8 Adjoint functors, Grothendieck topoi, and exponential categories
1.8.1 ADJOINT PAIR OF FUNCTORS, LEFT ADJOINT, RIGHT ADJOINT
1.8.2 REFLECTIVE SUBCATEGORY, REFLECTION
1.8.3 REFLECTIVE SUBCATEGORY, REFLECTION
1.8.4 CLOSURE OPERATION
1.8.5 COREFLECTIVE INCLUSION
1.8.6 ADJOINT ON THE RIGHT (LEFT), Galois connection
1.8.7 DIAGONAL FUNCTOR
1.8.8 diagram in one category modeled on another, lower bound, compatibility condition, greatest lower bound
1.8.9 LIMIT, COLIMIT
1.8.10 COMPLETE, COCOMPLETE (category)
1.8.11 CONTINUOUS, COCONTINUOUS (functor)
1.8.12 weak-, WEAK-LIMIT, WEAKLY-COMPLETE
1.8.13 PRE-LIMIT, PRE-COMPLETE
1.8.14 PRE-ADJOINT, PRE-REFLECTION, PRE-ADJOINT FUNCTOR, GENERAL ADJOINT FUNCTOR THEOREM
1.8.15 UNIFORMLY CONTINUOUS (functor), MORE GENERAL ADJOINT FUNCTOR THEOREM
1.8.16 ADJOINT FUNCTOR THEOREM
1.8.17 POINTWISE CONTINUOUS (functor)
1.8.18 functor generated by the elements, PETTY-FUNCTOR
1.8.19 GENERAL REPRESENTABILITY THEOREM, category of elements
1.8.20 WELL-POWERED CATEGORY, minimal object
1.8.21 cardinality function, generated by A through G
1.8.22 COGENERATING SET, SPECIAL ADJOINT FUNCTOR THEOREM
1.8.23 GIRAUD DEFINITION OF A GROTHENDIECK TOPOS
1.8.24 EXPONENTIAL CATEGORY [cartesian-closed], EVALUATION MAP
1.8.25 bifunctor
1.8.26 EXPONENTIAL IDEAL, REPLETE SUBCATEGORY

D.3.6 Topoi. The Categorical Analysis of Logic (Goldblatt [?])

D.3.7 Notes on Categories and Groupoids (Higgins[32])

Chapters

D.3.8 Topos Theory (Johnstone [?])

D.3.9 Sketches of an Elephant: a Topos Theory Compendium (Johnstone [33, ?])

VOLUME 1

VOLUME 2

D.3.10 Categories and Sheaves (Kashiwara and Schapira [35])

D.3.11 Basic Concepts of Enriched Category Theory (Kelly [37])

D.3.12 Introduction to Higher Order Categorical Logic (Lambek and Scott [?])

D.3.13 Sets for Mathematics (Lawvere and Rosebrugh [46])

D.3.14 Categories for the Working Mathematician (Mac Lane [53])

4. Preservation of Limits
5. Adjoints on Limits
6. Freyd's Adjoint Functor Theorem
7. Subobjects and Generators
8. The Special Adjoint Functor Theorem
9. Adjoints in Topology

VI. Monads and Algebras
1. Monads in a Category
2. Algebras for a Monad
3. The Comparison with Algebras
4. Words and Free Semigroups
5. Free Algebras for a Monad
6. Split Coequalizers
7. Beck's Theorem
8. Algebras are T-algebras
9. Compact Hausdorff Spaces

VII. Monoids
1. Monoidal Categories
2. Coherence
3. Monoids
4. Actions
5. The Simplicial Category
6. Monads and Homology
7. Closed Categories
8. Compactly Generated Spaces
9. Loops and Suspensions

VIII. Abelian Categories
1. Kernels and Cokernels
2. Additive Categories
3. Abelian Categories
4. Diagram Lemmas

IX. Special Limits
1. Filtered Limits
2. Interchange of Limits
3. Final Functors
4. Diagonal Naturality
5. Ends
6. Coends
7. Ends with Parameters
8. Iterated Ends and Limits

X. Kan Extensions
1. Adjoints and Limits
2. Weak Universality
3. The Kan Extension
4. Kan Extensions as Coends
5. Pointwise Kan Extensions
6. Density
7. All Concepts are Kan Extensions

D.3.15 Sheaves in Geometry and Logic (Mac Lane and Moerdijk [56])

Categorical Preliminaries

I. Categories of Functors
1. The Categories at Issue
2. Pullbacks
3. Characteristic Functions of Subobjects
4. Typical Subobject Classifiers
5. Colimits
6. Exponentials
7. Propositional Calculus
8. Heyting Algebras
9. Quantifiers as Adjoints

II. Sheaves of Sets
1. Sheaves
2. Sieves and Sheaves

D.3.16 First Order Categorical Logic (Makkai and Reyes [?])

D.3.17 Algebraic Theories (Manes [?])

D.3.18 Elementary Categories, Elementary Toposes (McLarty [?])

D.3.19 Categorical Foundations (Pedicchio and Tholen [62]

D.3.20 Introduction to Categories, Homological Algebra, and Sheaf Cohomology (Strooker [?])

D.3.21 Practical Foundations of Mathematics (Taylor [?])

Here are the categories that Taylor names:

1.1 **Set**
1.2 **PSet**
1.3 binary endorelations
1.4 well founded relations
1.5 **Preorder**
1.6 **Poset**
1.7 posets with left adjoints
1.8 complete semilattices
1.9 meet semilattices
1.10 distributive lattices
1.11 Heyting semilattices
1.12 Heyting lattices
1.13 frames
1.14 locales
1.15 directed complete posets
1.16 inductive partial orders
1.17 **Top**
1.18 **Monoid**
1.19 **Group**
1.20 commutative monoids
1.21 **Ab**
1.22 fields
1.23 **Vect**
1.24 **Ring**
1.25 **CommutativeRing**

D.3.22 An Introduction to Homological Algebra (Weibel [65])

Bibliography

[1] Jiří Adámek. *Theory of Mathematical Structures*. D. Reidel Publishing Co., Dordrecht, 1983.

[2] Jiří Adámek, Horst Herrlich, and George E. Strecker. *Abstract Concrete Categories: The Joy of Cats*. Pure and applied mathematics. John Wiley & Sons, 1990.

[3] Michael A. Arbib and Ernest G. Manes. *Arrows, Structures, and Functors: The Categorical Imperative*. Academic Press, New York, 1975.

[4] Andrea Asperti and Giuseppe Longo. *Categories, Types, and Structures: An Introduction to Category, Theory for the Working Computer Scientist*. Foundations of Computing. MIT Press, Cambridge, MA, 1991. Available at `ftp://ftp.ens.fr/pub/dmi/users/longo/CategTypesStructures`.

[5] Steve Awodey. *Category Theory*, volume 49 of *Oxford Logic Guides*. oup, 2006.

[6] Michael Barr. Functional Set Theory. Available on-line at `ftp://ftp.math.mcgill.ca/pub/barr/ttt.ps`, 1990?

[7] Michael Barr and Charles Wells. *Toposes, Triples, and Theories*. Number 278 in Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1985. Available at `http://www.cwru.edu/math/wells/pub/ttt.html`.

[8] Michael Barr and Charles Wells. *Category Theory for Computer Science*. Les Publications CRM, Montréal, 3rd edition, 1999. Centre de Recherches Mathematiques.

[9] Michael Barr and Charles Wells. Category Theory Lecture Notes for ESSLLI. Course given at the 11th European Summer School in Logic, Language and Information, Utrecht University, 9–20 August 1999, August 1999.

[10] James C. Becker and Daniel H. Gottlieb. A history of duality in algebraic topology. In I. M. James, editor, *History of Topology*, chapter 25, pages

725–745. elsevier, 1999. Available at `www.math.purdue.edu/~gottlieb/Bibliography/53.pdf`.

[11] J. L. Bell. *Toposes and Local Set Theories: An Introduction*. Number 14 in Oxford Logic Guides. Oxford University Press, 1988.

[12] George M. Bergman. *An Invitation to General Algebra and Universal Constructions*. Henry Helson, 15 the Crescent, Berkeley CA, 94708, 1998. available from `http://math.berkeley.edu/~gbergman/245/`.

[13] G. Birkhoff. *Lattice Theory, revised edition*. Number 25 in AMS Colloquium Publication. American Mathematical Society, 1948.

[14] G. Blanc and M. R. Donnadieu. Axiomatisation de la catégorie des catégories. *Cahiers Topologie Géom. Différentielle*, 17(2):135–170, 1976.

[15] Georges Blanc and Anne Preller. Lawvere's basic theory of the category of categories. *Journal of Symbolic Logic*, 40(1):14–18, 1975.

[16] T. S. Blyth. *Categories*. Longman, London, 1986.

[17] Francis Borceux. *Handbook of Categorical Algebra 1: Basic Category Theory*. Number 50 in Encyclopedia of Mathematics and Its Applications. Cambridge University Press, 1994.

[18] Francis Borceux. *Handbook of Categorical Algebra 2: Categories and Structures*. Number 51 in Encyclopedia of Mathematics and Its Applications. Cambridge University Press, 1994.

[19] Francis Borceux. *Handbook of Categorical Algebra 3: Categories of Sheaves*. Number 52 in Encyclopedia of Mathematics and Its Applications. Cambridge University Press, 1994.

[20] Nicolas Bourbaki. *Elements of Mathematics: Theory of Sets*. Addison-Wesley Publishing Co., 1968. Translated from Eléments de Mathématique, Théories des Ensembles published by Hermann, Paris.

[21] Nicolas Bourbaki. *Elements of Mathematics: Algebra I*. Addison-Wesley Publishing Co., 1974. Translated from Elements de Mathematique, Algebra. Reprinted springer, 1999.

[22] Ionel Bucur and Aristide Deleanu; Peter J. Hilton. *Introduction to the Theory of Categories and Functors*. Number 19 in Pure and applied mathematics. John Wiley & Sons, 1968.

[23] Mario José Cáccamo and Glynn Winskel. Lecture notes in category theory [draft]. Unpublished manuscript, June 2002.

[24] Henri Cartan and Samuel Eilenberg. *Homological Algebra*. Princeton University Press, 1956.

[25] Leon W. Cohen and Gertrude Ehrlich. *The Structure of the Real Number System.* The University Series in Undergraduate Mathematics. D. Van Nostrand Co., 1963.

[26] P. M. Cohn. *Universal Algebra.* Harper's Series in Modern Mathematics. Harper & Row, 1965.

[27] Roy Crole. *Categories for Types.* Cambridge Mathematical Textbooks. Cambridge University Press, 1993.

[28] Reinhard Diestel. *Graph Theory.* Graduate Texts in Mathematics. Springer-Verlag, 1997. 2nd edition, Springer (2000); Translation of Graphentheorie, Springer-Verlag (1996).

[29] Jean Alexandre Dieudonné. *A History of Algebraic and Differential Topology 1900-1960.* Birkhäuser, 1989.

[30] Eduardo Dubuc. *Kan Extensions in Enriched Category Theory.* Number 145 in Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1970.

[31] B. Eckmann and P. J. Hilton. Group-like structures in general categories I. *Math. Annalen*, 145:227–255, 1962.

[32] B. Eckmann and P. J. Hilton. Group-like structures in general categories. II. Equalizers, limits, lengths. *Math. Ann.*, 151:150–186, 1963.

[33] B. Eckmann and P. J. Hilton. Group-like structures in general categories. III. Primitive categories. *Math. Ann.*, 150:165–187, 1963.

[34] S. Eilenberg and Saunders Mac Lane. General theory of natural equivalences. *Transactions of the American Mathematical Society*, 58:231–244, 1945.

[35] Samuel Eilenberg and Saunders Mac Lane. *Eilenberg–Mac Lane: Collected Works.* Academic Press, 1986.

[36] Samuel Eilenberg and Norman Steenrod. *Foundations of Algebraic Topology.* Princeton University Press, Princeton, N.J., 1952.

[37] Solomon Feferman. *The Number Systems: Foundations of Algebra and Analysis.* Addison-Wesley Publishing Co., 1964. Reprinted by American Mathematics Society, 2003.

[38] Solomon Feferman. Set-theoretical foundations of category theory. In Mac Lane [52], pages 201–247.

[39] Maarten M. Fokkinga. A gentle introduction to category theory — the calculational approach —. Revised version of Lecture Notes of the STOP 1992 Summerschool on Constructive Algorithmics at the University of Utrecht. Available on-line at `http://wwwhome.cs.utwente.nl/~fokkinga/`, June 1994.

[40] Peter J. Freyd. *Abelian Categories: An Introduction to the Theory of Functors*. Harper and Row, 1964. Available with updated forward at `http://www.tac.mta.ca/tac/reprints`.

[41] Peter J. Freyd and Andrej Ščedrov. *Categories, Allegories*. Number 39 in North-Holland Mathematical Library. North Holland, 1990.

[42] Roger Godement. *Topologie algébrique et théorie des faisceaux*. Actualit'es Sci. Ind. No. 1252. Publ. Math. Univ. Strasbourg. No. 13. Hermann, Paris, 1958.

[43] Robert I. Goldblatt. *Topoi: The Categorical Analysis of Logic*. Number 98 in Studies in Logic and the Foundations of Mathematics. North Holland, 2nd, revised edition, 1984.

[44] George Grätzer. *Universal Algebra*. The University Series in Higher Mathematics. D. Van Nostrand Co., 1968.

[45] Alexander Grothendieck. Sur quelques points d'algèbre homologique. *Tohoku Math. J. (2)*, 9:119–221, 1957.

[46] Paul R. Halmos. *Naive Set Theory*. The University Series in Undergraduate Mathematics. D. Van Nostrand Co., 1960. Reprinted Springer-Verlag, 1974.

[47] Allen Hatcher. *Algebraic Topology*. CUP, 2002. Available on-line at `http://www.math.cornell.edu/~hatcher/AT/ATpage.html`.

[48] Pavol Hell and Jaroslav Nešetřil. *Graphs And Homomorphisms*. Oxford Lecture Series in Mathematics and Its Applications. Oxford University Press, 2004.

[49] Horst Herrlich and George E. Strecker. *Category Theory: An Introduction*. Number 1 in Sigma series in pure mathematics. Heldermann, Berlin, 2nd edition, 1979.

[50] Philip J. Higgins. *Notes on Categories and Groupoids*. D. Van Nostrand Co., 1971. Reprints in Theory and Applications of Categories, No. 7, 1971, pp. 1–195. Available on-line at `http://www.tac.mta.ca/tac/reprints`.

[51] Chris Hillman. A categorical primer. Available on-line at `http://books.pdox.net/Math/A%20Categorical%20Primer.ps`, Was evolving for several years., 2001.

[52] P. T. Johnstone. *Sketches of an Elephant: A Topos Theory Compendium: Volume 1*. Number 43 in Oxford Logic Guides. Oxford University Press, 2002.

[53] P. T. Johnstone. *Sketches of an Elephant: A Topos Theory Compendium: Volume 2.* Number 44 in Oxford Logic Guides. Oxford University Press, 2003.

[54] Peter Johnstone. *Topos Theory.* Number 10 in London Mathematical Society Monographs. Academic Press, 1977.

[55] André Joyal and Ieke Moerdijk. *Algebraic Set Theory.* Number 220 in London Mathematical Society Lecture Notes. Cambridge University Press, 1995.

[56] Masaki Kashiwara and Pierre Schapira. *Categories and Sheaves*, volume 332 of *Grundlehren der mathematischen Wissenschaften.* Springer-Verlag, 2006.

[57] John L. Kelley. *General Topology.* University Series in Higher Mathematics. D. Van Nostrand Co., 1955. Reprinted by Springer-Verlag, 1975.

[58] G. M. Kelly. *Basic Concepts of Enriched Category Theory.* Number 64 in London Mathematical Society Lecture Notes. Cambridge University Press, Cambridge, 1982. Reprinted as [38].

[59] G. M. Kelly. *Basic Concepts of Enriched Category Theory.* Number 10 in Reprints in Theory and Applications of Categories. Theory and Applications of Categories, 2005. Reprint of [37].

[60] Viakalathur Sankrithi Krishnan. *An introduction to category theory.* North Holland, New York, 1981.

[61] Joachim Lambek and P. J. Scott. *Introduction to Higher Order Categorical Logic.* Number 7 in Cambridge studies in advanced mathematics. Cambridge University Press, 1986. First paperback edition (with corrections) 1988.

[62] Edmund Landau. *Foundations of analysis; the arithmetic of whole, rational, irrational, and complex numbers.* Chelsea Publishing Co., 1951. Translated from the German by F. Steinhardt.

[63] Serge Lang. *Differential and Riemannian Manifolds.* Springer-Verlag, 3rd edition, 1995.

[64] F. W. Lawvere. An elementary theory of the category of sets. *Proceedings of the National Academy of Sciences U.S.A.*, 52:1506–1511., 1964.

[65] F. William Lawvere. The category of categories as a foundation for mathematics. In Samuel Eilenberg, D. K. Harrison, Saunders Mac Lane, and Helmut Röhrl, editors, *Categorical Algebra (La Jolla, 1965)*, pages 1–20. Springer-Verlag, 1965. For clarification see the review by J. R. Isbell in Math. Rev. 34, no. 7332.

[66] F. William Lawvere. Categories of space and of quantity. In *The space of mathematics (San Sebastiàn, 1990)*, Found. Comm. Cogn., pages 14–30. de Gruyter, Berlin, 1992.

[67] F. William Lawvere. An elementary theory of the category of sets (long version) with commentary. *Reprints in Theory and Applications of Categories*, 11:1–35, 2005. An expanded version of Lawvere [41].

[68] F. William Lawvere. *Functorial Semantics of Algebraic Theories and Some Algebraic Properties in the Context of Functorial Semantics of Algebraic Theories*. Number 5 in Reprints in Theory and Applications of Categories. TAC, 2005. Publication of Columbia Ph. D. thesis (1963) and reprint of paper (1968). Available from `http://www.tac.mta.ca/tac/reprints`.

[69] F. William Lawvere and Robert Rosebrugh. *Sets for Mathematics*. Cambridge University Press, 2003.

[70] F. William Lawvere and Stephen J. Schanuel. *Conceptual Mathematics: A First Introduction to Categories*. Cambridge University Press, 1997.

[71] Albert T. Lundell. *The Topology of CW Complexes*. The University Series in Higher Mathematics. VNR, 1969.

[72] Saunders Mac Lane. Dualities for groups. *Bulletin of the American Mathematical Society*, 56:485–516, 1950. Reprinted in [21].

[73] Saunders Mac Lane. Foundations for categories and sets. In Peter Hilton, editor, *Category theory, homology theory and their applications II*, number 92 in Lecture Notes in Mathematics, pages 146–164. Springer-Verlag, 1969. Proceedings of the conference held at the Seattle Research Center of the Battelle Memorial Institute, June 24-July 19, 1968.

[74] Saunders Mac Lane. One universe as a foundation for category theory. In *Reports of the Midwest Category Seminar III* [52], pages 192–200.

[75] Saunders Mac Lane, editor. *Reports of the Midwest Category Seminar III*, number 106 in Lecture Notes in Mathematics, Berlin, 1969. Springer-Verlag.

[76] Saunders Mac Lane. *Categories for the Working Mathematician*. Number 5 in Graduate Texts in Mathematics. Springer-Verlag, 1971.

[77] Saunders Mac Lane. *Categories for the Working Mathematician*. Number 5 in Graduate Texts in Mathematics. Springer-Verlag, 2nd edition, 1997.

[78] Saunders Mac Lane and Garrett Birkhoff. *Algebra, Third Edition*. AMS Chelsea Publishing, Providence, RI, 1999.

[79] Saunders Mac Lane and Ieke Moerdijk. *Sheaves in Geometry and Logic: a First Introduction to Topos Theory.* Universitext. Springer-Verlag, 1992.

[80] Michael Makkai and Gonzalo Reyes. *First Order Categorical Logic: Model-Theoretical Methods in the Theory of Topoi and Related Categories.* Number 611 in Lecture Notes in Mathematics. Springer-Verlag, 1977.

[81] Ernest G. Manes. *Algebraic Theories.* Number 26 in Graduate Texts in Mathematics. Springer-Verlag, 1976.

[82] Colin McLarty. Axiomatizing a category of categories. *Journal of Symbolic Logic*, 56(4):1243–1260, 1991.

[83] Colin McLarty. *Elementary Categories, Elementary Toposes.* Number 21 in Oxford logic guides. Oxford University Press, 1992.

[84] Barry Mitchell. *Theory of Categories.* Number 17 in Pure and applied mathematics. Academic Press, 1965.

[85] J. Donald Monk. *Introduction to Set Theory.* International Series in Pure and Applied Mathematics. mcgh, 1969.

[86] Anthony P. Morse. *A Theory of Sets.* Number 108 in Pure and Applied Mathematics. Academic Press, 2nd edition, 1986.

[87] Gerhard Osius. Categorical set theory: a characterisation of the category of sets. *Journal of Pure and Applied Algebra*, 4:79–119, 1974. Habilitationsschrift, University Bielefeld, May 1973.

[88] Bodo Pareigis. *Categories and Functors.* Number 39 in Pure and applied mathematics. Academic Press, 1970.

[89] Maria Cristina Pedicchio and Fabrizio Rovatti. Algebraic categories. In Pedicchio and Tholen [62], pages 269–309.

[90] Maria Cristina Pedicchio and Walter Tholen, editors. *Categorical Foundations: Special Topics in Order, Topology, Algebra, and Sheaf Theory.* Number 97 in Encyclopedia of Mathematics and Its Applications. Cambridge University Press, 2004.

[91] Benjamin C. Pierce. *Basic Category Theory for Computer Scientists.* Foundations of computing. MIT Press, Cambridge, Mass., 1991.

[92] David Pitt, Samson Abramsky, Axel Poignè, and David Ryderheard, editors. *Category Theory and Computer Programming: Tutorial and Workshop, Guildford, U.K., September 16-20, 1985: Proceedings*, number 240 in Lecture Notes in Computer Science, Berlin, 1985. Springer-Verlag.

[93] Nicolae Popescu and Liliana Popescu. *Theory of Categories.* Sijthoff & Noordhoff, 1979.

[94] P. Samuel. On universal mappings and free topological groups. *Bull. Amer. Math. Soc.*, 54:591–598, 1948.

[95] Horst Schubert. *Categories.* Springer-Verlag, 1972.

[96] Thomas Streicher. Introduction to category theory and categorical logic. Notes from courses in Summer semester '03 and Winter semester '03/'04 at Technische Universität Darmstadt. Available at `http://www.mathematik.tu-darmstadt.de/~streicher/`.

[97] Jan R. Strooker. *Introduction to Categories, homological algebra, And sheaf cohomology.* Cambridge University Press, 1978.

[98] Paul Taylor. *Practical Foundations of Mathematics.* Cambridge University Press, 1999.

[99] R. F. C. Walters. *Categories and Computer Science.* Number 28 in Cambridge Computer Science Texts. Cambridge University Press, 1992.

[100] Charles A. Weibel. *An Introduction to Homological Algebra.* Cambridge studies in advanced mathematics 38. CUP, 1994.

[101] Charles A. Weibel. History of homological algebra. In *History of topology*, pages 797–836. North-Holland, Amsterdam, 1999.

[102] Eric W. Weisstein. Dual polyhedron. `http://mathworld.wolfram.com/DualPolyhedron.html`. From MathWorld–A Wolfram Web Resource.

[103] Eric W. Weisstein. Pappus's hexagon theorem. `http://mathworld.wolfram.com/PappussHexagonTheorem.html`. From MathWorld–A Wolfram Web Resource.

[104] Eric W. Weisstein. Pontryagin duality. `http://mathworld.wolfram.com/PontryaginDuality.html`. From MathWorld–A Wolfram Web Resource.

[105] A. N. Whitehead. *A Treatise on Universal Algebra.* Cambridge at the University Press, 1898.

Index