LIMITS, COLIMITS AND HOW TO CALCULATE THEM IN THE CATEGORY OF MODULES OVER A PID.

KAIRUI WANG

Abstract. The goal of this paper is to introduce methods that allow us to calculate certain limits and colimits in the category of modules over a principal ideal domain. We start with a quick review of basic categorical language and duality. Then we develop the concept of universal morphisms and derive limits and colimits as special cases. The completeness and cocompleteness theorems give us methods to calculate the morphisms associated with limits and colimits in general. To use these methods, we then specialize to the case of finitely generated modules over a PID. We first develop the Smith normal form as a computational tool and prove the structure theorem for finitely generated modules over a PID. Lastly, we discuss how to use the abstract methods suggested by the completeness and cocompleteness theorems in the context of ${}_R\mathbf{Mod}$.

Contents

1. Basic Definitions and Concepts 1
2. Universality 3
3. Limits and Colimits 5
4. Smith Normal Form and Finitely Generated Modules over a PID 9
5. Finding Limits and Colimits of Special Modules 13
Acknowledgments 15
References 15

1. Basic Definitions and Concepts

We begin with a quick review of the definitions of category, functor and natural transformation.

Definition 1.1. A **category** $\mathcal{C}$ consists of a class $\mathrm{Ob}(\mathcal{C})$ whose elements are called **objects**, where for each pair of objects $X, Y \in \mathrm{Ob}(\mathcal{C})$, there is a set $\mathrm{Hom}(X, Y)$ whose elements are called **morphisms**; [1] and for each object X, there is a morphism $1_X : X \to X$ called the **identity morphism** on X. Also, between any three objects X, Y, Z, there is a composition operation $\cdot : \mathrm{Hom}(X, Y) \times \mathrm{Hom}(Y, Z) \to \mathrm{Hom}(X, Z)$ that is closed; *i.e.*, if $f \in \mathrm{Hom}(X, Y)$ and $g \in \mathrm{Hom}(Y, Z)$, then there exists a morphism $f \cdot g \in \mathrm{Hom}(X, Z)$. In addition, morphisms must satisfy the following properties:

(1) Composition of compatible morphisms is associative.

Date: August 26, 2011.

[1] Conventionally, each element $f \in \mathrm{Hom}(X, Y)$ is written as $f : X \to Y$.

(2) Composition involving identity morphisms behave as expected, *i.e.*, for any morphism $f : X \to Y$, $f \cdot 1_Y = f$ and $1_X \cdot f = f$. We call this property the unit axioms.

Notation 1.2. From now on, we will drop the $\cdot$ when writing the composition of morphisms. This means if X, Y, Z are some objects, and $f : X \to Y$ and $g : Y \to Z$ are morphisms between them, then we write the composition $f \cdot g$ as fg from now on.

Using this notation, associativity in definition 1.1 may be stated as: if f, g, h are three composable morphisms, then $(fg)h = f(gh)$.

We can think of a functor as a morphism between two categories.

Definitions 1.3. A **covariant functor** F, from a category $\mathcal{C}$ to a category $\mathcal{D}$ assigns for each object $X \in \mathrm{Ob}(\mathcal{C})$, an object $FX \in \mathrm{Ob}(\mathcal{D})$, and assigns for each morphism $f : X \to Y$ in $\mathrm{Hom}(X, Y)$, a morphism $Ff : FX \to FY$ in $\mathrm{Hom}(FX, FY)$, such that the following properties are satisfied:

(1) Identity morphisms are taken to identity morphisms, so that $F(1_X) = 1_{FX}$.
(2) F respects the composition of morphisms, so that if f, g are composable morphisms in $\mathcal{C}$, then $F(fg) = (Ff)(Fg)$ in $\mathcal{D}$.

We distinguish between a covariant and a **contravariant functor**. Both types of functors act on objects in the same way, but for every morphism $f : X \to Y$, a contravariant functor assigns to it a morphism $Ff : FY \to FX$. So correspondingly, property (2) for a contravariant functor ought to read:

(2') F respects the composition of morphisms, so that if f, g are composable morphisms in $\mathcal{C}$, then $F(fg) = (Fg)(Ff)$ in $\mathcal{D}$.

When we speak of a "functor" without modifications, we mean a covariant functor.

We can think of a natural transformation as a morphism between two functors.

Definition 1.4. Let F and G be two functors from category $\mathcal{C}$ to category $\mathcal{D}$. A **natural transformation** η from F to G is a class of morphisms in $\mathcal{D}$ indexed by the objects in $\mathcal{C}$, such that for any two objects $X, Y \in \mathrm{Ob}(\mathcal{C})$ and any morphism $f : X \to Y$, there is a commutative diagram:

$$\begin{array}{ccc} FX & \xrightarrow{Ff} & FY \\ {\scriptstyle \eta_X}\downarrow & & \downarrow{\scriptstyle \eta_Y} \\ GX & \xrightarrow[Gf]{} & GY \end{array} .$$

A natural transformation between two contravariant functors from category $\mathcal{C}$ to category $\mathcal{D}$ is defined analogously. The horizontal arrows of the commutative diagram above are reversed.

There is a symmetry in the definition of a category. Whenever we have two objects X, Y in a category $\mathcal{C}$, there could exist a morphism $f : X \to Y$, but we could just as well reverse this arrow and consider a morphism $f' : Y \to X$. A new category arises when we "reverse the source and target" of every morphism.

Definition 1.5. Let $\mathcal{C}$ be a category. We define its **opposite category**, $\mathcal{C}^{\mathrm{op}}$ to be the target category of a contravariant functor Op, which takes objects in $\mathcal{C}$ to themselves, and morphisms $f : X \to Y$ in $\mathcal{C}$ to $f : Y \to X$.

The idea of a category and its opposite category form a duality in category theory. When considering any idea in category theory, it is a good idea to also think about it in the dual category. Often, a theorem that holds in a category has a dual theorem that holds in its opposite category. The duality principle states that proving a categorical statement also implies its dual statement. It is useful to keep duality in mind, so we will point out situations where duality arises. For example, the notion of limit and colimit are dual. We will see in our applications that even though they are dual in a categorical sense, the limit and colimit in specific categories may look very different. Since we will not be proving the duality principle, we give the proof for dual statements as well (this usually amounts to reversing all arrows in a proof).

2. Universality

One type of problem in category theory is the universal mapping problem. Informally, these problems look for a morphism (called the universal morphism) that satisfies some desired property, such that any other morphism that satisfies the property "factors through" it in the sense that it is the universal morphism composed with some other morphism. We can put this idea in formal terms. We start by defining a comma category.

Definition 2.1. Let $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ be three categories, and $F : \mathcal{A} \to \mathcal{C}$ and $G : \mathcal{B} \to \mathcal{C}$ be two functors with the same target category. The **comma category** $(F \downarrow G)$ has objects that are triples (α, β, f) where α is an object in $\mathcal{A}$, β is an object in $\mathcal{B}$ and $f : F\alpha \to G\beta$ is a morphism in $\mathcal{C}$.

The morphisms between two objects (α, β, f) and (α', β', f') are pairs (g, h) where $g : \alpha \to \alpha'$ and $h : \beta \to \beta'$ are morphisms in $\mathcal{A}$ and $\mathcal{B}$, respectively, such that the following diagram commutes:

$$\begin{array}{ccc} F\alpha & \xrightarrow{Fg} & F\alpha' \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f'} \\ G\beta & \xrightarrow[Gh]{} & G\beta' \end{array} .$$

The identity morphism for any object (α, β, f) is $(1_\alpha, 1_\beta)$, which exists because 1_α and 1_β exist in $\mathcal{A}$ and $\mathcal{B}$, respectively. The composition rule is that if $g' : \alpha' \to \alpha''$ is composable with g in $\mathcal{A}$ and $h' : \beta' \to \beta''$ is composable with h in $\mathcal{B}$, then (g, h) and (g', h') are composable, with

$$(g, h)(g', h') = (gg', hh')$$

This composition has source (α, β, f) and target (α'', β'', f''). The two commutative diagrams for (g, h) and (g', h') make the overall diagram commute:

$$\begin{array}{ccccc} F\alpha & \xrightarrow{Fg} & F\alpha' & \xrightarrow{Fg'} & F\alpha'' \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f'} & & \downarrow{\scriptstyle f''} \\ G\beta & \xrightarrow[Gh]{} & G\beta' & \xrightarrow[Gh']{} & G\beta'' \end{array}$$

Notice that the horizontal arrows composed to the required $F(gg')$ and $G(hh')$ by functoriality. It is easy to verify our morphisms and composition rule satisfy the two axioms of a category in defnition 1.1.

Associativity of morphism composition follows from associativity in each component morphism. The unit axioms follow from the unit axioms for each component morphism. That is, the morphism (g, h), composing with the identity morphism $(1_{\alpha'}, 1_{\beta'})$ gives

$$(g, h)(1_{\alpha'}, 1_{\beta'}) = (g1_{\alpha'}, h1_{\beta'}) = (g, h).$$

Precomposing with $(1_\alpha, 1_\beta)$ also leaves (g, h) unchanged and can be shown analogously.

We will be considering comma categories where one functor has as source the category with one object and the identity morphism. This functor simply maps to an object $c \in \mathrm{Ob}(C)$. We will call this functor c.

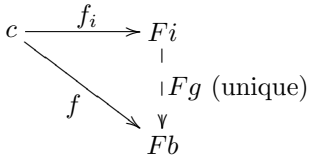
Definitions 2.2. Let $\mathcal{C}$ be a category. An **initial object** in $\mathcal{C}$ is an object such that for any object X in $\mathcal{C}$, there exists a unique morphism from the initial object to X. Dually, a **terminal object** in $\mathcal{C}$ is an object such that for any object Y in $\mathcal{C}$, there exists a unique morphism from Y to the terminal object.

Proposition 2.3. *Initial and terminal objects, if they exist, are unique up to unique isomorphism.*

Proof. Say I, I' are initial objects in a category $\mathcal{C}$. By definition, there exist unique morphisms $f : I \to I'$ and $g : I' \to I$. Composing f and g, and noting that the unique morphisms from I to itself and from I' to itself are their respective identity morphisms, we have $fg = 1_I$ and $gf = 1_{I'}$, so f is the unique isomorphism between I and I'. The proof for terminal objects is analogous. □

Definition 2.4. Let c be an object in a category $\mathcal{C}$ and $F : \mathcal{B} \to \mathcal{C}$ be a functor. If we consider c as a functor from the category with one object $(*)$ and its identity morphism to $\mathcal{C}$, we define the **universal morphism** *from c to F* as the initial object in the comma category $(c \downarrow F)$.

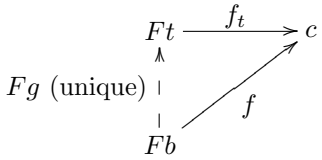
Let's unpack this definition. Objects in $(c \downarrow F)$ are of the form $(*, b, f)$ for some object b in $\mathcal{B}$. This gives us the morphism in $\mathcal{C}$: $f : c \to Fb$. Suppose $(*, i, f_i)$ is the initial object in the comma category. This means there is a unique morphism $(1_*, g)$ from $(*, i, f_i)$ to another object $(*, b, f)$. Morphisms in this comma category make the following diagram commute:



Definition 2.5. The dual to definition 2.4 is that the **universal morphism** *from F to c* is the terminal object in the comma category $(F \downarrow c)$.

Unpacking this definition, if we call the terminal object in $(F \downarrow c)$ $(t, *, f_t)$, then for any object $(b, *, f)$ there exists a unique morphism $(g, 1_*)$ to the terminal object

that makes the following diagram commute:



Corollary 2.6. *Universal morphisms are unique up to unique isomorphism.*

3. Limits and Colimits

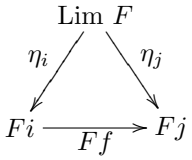
Limits and their dual, colimits, are solutions to universal mapping problems, that generalize many other universal constructions.

Definitions 3.1. By a **small category**, we mean a category where the class of objects is a set. Let $F : \mathcal{B} \to \mathcal{C}$ be a functor. If c is an object of $\mathcal{C}$, we say F is the **constant functor** at c if F maps every object in $\mathcal{B}$ to c, and if F maps every morphism of $\mathcal{B}$ to 1_c, the identity morphism of c.

Let J be a small category. The **diagonal functor** $\Delta : \mathcal{C} \to \mathcal{C}^J$ is a functor from $\mathcal{C}$ to $\mathcal{C}^J$, the category of functors from J to $\mathcal{C}$. Δ takes objects c in $\mathcal{C}$ to the constant functor Δc. Δ takes morphisms $f : c \to c'$ to natural transformations $\Delta f = \Delta c \to \Delta c'$. Because Δc and $\Delta c'$ are constant functors, Δf is just the morphism $f : c \to c'$ for every object in J.

Definitions 3.2. We call a functor F from a small category J to a category $\mathcal{C}$ a **diagram** over J in $\mathcal{C}$. Clearly, F is an object in $\mathcal{C}^J$. The **limit** of diagram F, denoted Lim F, is the universal morphism from Δ to F.

This means the limit of diagram F is the terminal object in the comma category $(\Delta \downarrow F)$. Objects in this comma category are $(c, *, f)$, where c is an object of $\mathcal{C}$. The triple represents the natural transformation $f : \Delta c \to F$. The limit is the natural transformation $\eta : \Delta(\text{Lim } F) \to F$, where Lim F here denotes some object in $\mathcal{C}$. Since $\Delta(\text{Lim } F)$ is a constant functor, naturality of η gives the following commutative diagram for every morphism $f : i \to j$ in J:

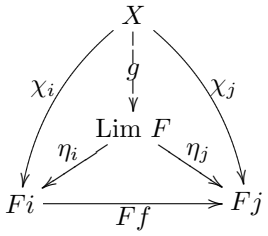


Definition 3.3. We call a natural transformation from a constant functor at $c \in \text{Ob}(\mathcal{C})$ to a diagram $F : J \to \mathcal{C}$ a **cone** over F.

Remark 3.4. This is called "cone" because when we draw the naturality diagram implied by a cone, we have one object (the vertex) along with morphisms to all other objects in the image of F. The relationships in F resemble the base of a geometric cone.

Remark 3.5. With the definition of cones, we can now interpret the comma category $(\Delta \downarrow F)$ as the category of cones over diagram F, thus the limit of F is the terminal object in the category of cones over F. We denote this category by $\mathbf{Cone}_F$.

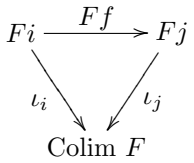
The cone interpretation makes it clear exactly what universal mapping problem the limit solves: given a diagram F with source category J and target category $\mathcal{C}$, is there an object Lim F with associated morphisms from Lim F to each object Fj where $j \in \mathrm{Ob}(J)$, such that Lim F and its associated morphisms commute with all morphisms Ff where f is a morphism in J, and satisfy the universal property that if another object X and morphisms from X to all Fj commute with all morphisms Ff, then there exists a unique morphism $g : X \to \mathrm{Lim}\ F$, such that for every $f : i \to j$ in J, the following diagram commutes



Dually, we can define the colimit of a diagram.

Definitions 3.6. The **colimit** of diagram F, denoted Colim F, is the universal morphism from F to Δ.

Analogous to limits, the colimit of F is the initial object in the comma category $(F \downarrow \Delta)$. It is a natural transformation $\iota : F \to \Delta(\mathrm{Colim}\ F)$, where Colim F is some object in $\mathcal{C}$. Naturality of ι gives the following commutative diagram for every morphism $f : i \to j$ in J:

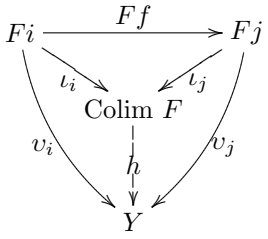


Definition 3.7. A natural transformation from a diagram $F : J \to \mathcal{C}$ to a constant functor at $c \in \mathrm{Ob}(\mathcal{C})$ is a **cocone** over F.

Remark 3.8. This allows us to interpret the comma category $(F \downarrow \Delta)$ as the category of cocones over F, thus the colimit is the initial object in the category of cocones over F, we denote this category by $\mathbf{Cocone}_F$.

The universal mapping problem solved is: given a diagram $F : J \to \mathcal{C}$, is there an object Colim F with associated morphisms from each Fj, where $j \in \mathrm{Ob}(J)$, to Colim F such that Colim F and its associated morphisms commute with all morphisms Ff where f is a morphism in J, and is universal in that if an object Y and morphisms from each Fj to Y also commute with all morphisms Ff, then there exists a unique morphism $h : \mathrm{Colim}\ F \to Y$, such that for any $f : i \to j$ in J,

the following diagram commutes:



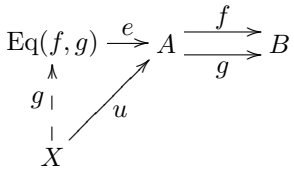
Now we give two specific, and very important examples of limits (dually colimits): the product and the equalizer (dually the coproduct and the coequalizer). They are fundamental in that every other limit and colimit can be constructed with them.

Definitions 3.9. Let $\mathscr{S}$ be a set of objects in category $\mathcal{C}$. The **product** of objects in $\mathscr{S}$ is an object in $\mathcal{C}$, denoted $\prod_{S\in\mathscr{S}} S$, along with morphisms $\eta_S : \prod_{S\in\mathscr{S}} S \to S$ for each object $S \in \mathscr{S}$, which is universal in the sense that if X is an object in $\mathcal{C}$ with morphisms $\chi_S : X \to S$ for every $S \in \mathscr{S}$, then there exists a unique morphism $g : X \to \prod_{S\in\mathscr{S}} S$, such that for every object $S \in \mathscr{S}$, $\chi_S = g\eta_S$.

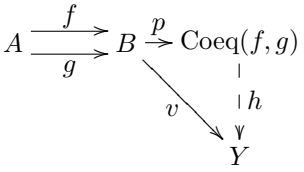
Dually, the **coproduct** of objects in $\mathscr{S}$ is an object in $\mathcal{C}$ denoted $\coprod_{S\in\mathscr{S}} S$, along with morphisms $\iota_S : S \to \coprod_{S\in\mathscr{S}} S$ for every $S \in \mathscr{S}$, which is universal in the sense that if Y is an object with morphisms $\upsilon_S : S \to Y$ for every $S \in \mathscr{S}$, then there exists a unique morphism $h : \coprod_{S\in\mathscr{S}} S \to Y$ such that for every object $S \in \mathscr{S}$, $\upsilon_S = \iota_S h$.

We can see that the product of objects in $\mathscr{S}$ is the limit of a diagram F from a category with only identity morphisms to $\mathcal{C}$, where each object is taken to an object in $\mathscr{S}$; the **coproduct** of objects in $\mathscr{S}$ is the colimit of F.

Definitions 3.10. Given two objects A, B in category $\mathcal{C}$ with two morphisms $f, g : A \to B$, the **equalizer** of f and g is an object, denoted $\mathrm{Eq}(f, g)$, and a morphism $e : \mathrm{Eq}(f, g) \to A$ such that $ef = eg$, and is universal in that if X is an object with morphism $u : X \to A$ such that $uf = ug$, then there exists a unique morphism $g : X \to \mathrm{Eq}(f, g)$ such that the following diagram commutes



Dually, the **coequalizer** of f and g is an object, denoted $\mathrm{Coeq}(f, g)$, and a morphism $p : B \to \mathrm{Coeq}(f, g)$ such that $fp = gp$, and is universal in that if Y is an object with morphism $v : B \to Y$ such that $fv = gv$, then there exists a unique morphism $h : \mathrm{Coeq}(f, g) \to Y$ such that the following diagram commutes



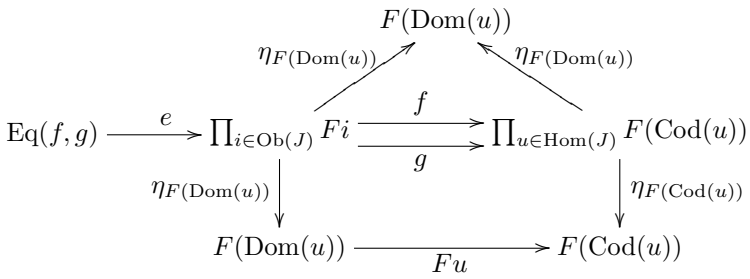
We can see that the equalizer is the limit of a diagram F from a category with two objects i, j and two morphisms $i \to j$, where $Fi = A$, $Fj = B$ and the two morphisms are taken to f and g; the coequalizer is the colimit of F.

Definition 3.11. By a **complete category**, we mean a category where limits over diagrams with a small source category exist; a **cocomplete category** is a category where colimits over diagrams with a small source category exist.

Theorem 3.12 (Completeness Theorem)**.** *A category $\mathcal{C}$ is complete if and only if the product of any set of objects in $\mathcal{C}$ exists and the equalizer between any two morphisms with the same source and target exists.*

Proof. The forward direction is obvious: products of a set of objects and equalizers are special cases of limits over diagrams from a small source category.

The proof of the converse can be summarized in a single diagram. We will explain the diagram after we draw it:



For any diagram $F : J \to \mathcal{C}$, the product of the image under F of all objects in J, $A = \prod_{i\in \mathrm{Ob}(J)} Fi$, and the product of the image under F of all objects that are the codomain of some morphism in J, $B = \prod_{u\in \mathrm{Hom}(J)} F(\mathrm{Cod}(u))$, both exist by hypothesis.

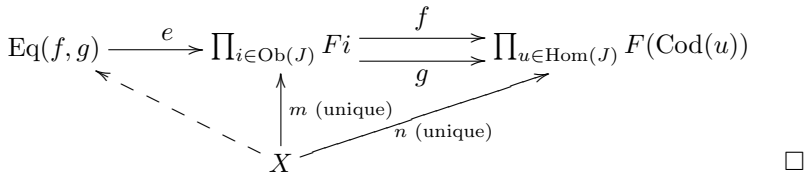
For any $u \in \mathrm{Hom}(J)$, $F(\mathrm{Cod}(u))$ is a component of both products A and B, so there are morphisms $\eta_{F(\mathrm{Cod}(u))} : A \to F(\mathrm{Cod}(u))$ for each u. The universality of B gives a unique morphism $f : A \to B$. (See the top triangle.)

There is another way to construct morphisms from A to each $F(\mathrm{Cod}(u))$. Since the domain of each u under F is a component of the product A, there is a morphism from A to each $F(\mathrm{Dom}(u))$. The morphism Fu then takes $F(\mathrm{Dom}(u))$ to $F(\mathrm{Cod}(u))$. The composition $\eta_{F(\mathrm{Dom}(u))}Fu$ gives another morphism from A to each $F(\mathrm{Cod}(u))$. Using the morphisms $\eta_{F(\mathrm{Dom}(u))}Fu$ for each u gives another cone over objects in J that are the target of morphisms, so universality of B gives another morphism $g : A \to B$. (See the bottom rectangle of the diagram).

We claim the object Lim F is the object of the equalizer $\mathrm{Eq}(f, g)$. For any morphism $u : i \to j$ in J, there are morphisms from $\mathrm{Eq}(f, g)$ to Fi and Fj, namely $e\eta_{Fi}$ and $ef\eta_{Fj}$, respectively such that these morphisms commute with Fu.

Now say X is an object such that for each morphism $u : i \to j$ in J, there are morphisms $p_{Fi} : X \to Fi$ and $p_{Fj} : X \to Fj$ that commute with Fu. Since every object has an identity morphism, X has a morphism to every object Fi where $i \in \mathrm{Ob}(j)$. Then by the universality of the products A and B, there are unique morphisms $m : X \to A$ and $n : X \to B$. Since morphisms are closed under composition, we know then mf and mg are morphisms from X to B. But n is the unique morphism from X to B, so $mf = mg$. By the universality of the equalizer $\mathrm{Eq}(f, g)$, there is a unique morphism $X \to \mathrm{Eq}(f, g)$, so we see $\mathrm{Eq}(f, g) = \mathrm{Lim}\ F$,

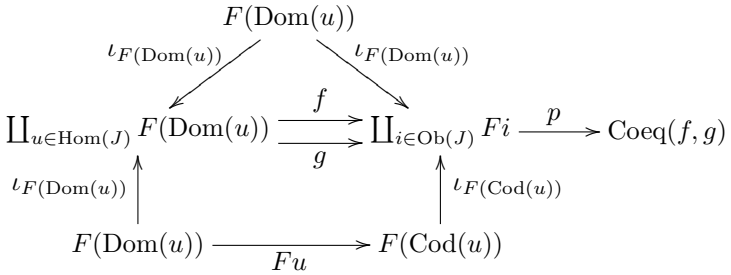
and for each $u : i \to j$ in J, the $e\eta_{F_i}$ and $ef\eta_{F_j}$ give the cone morphisms over F_i and F_j, respectively.



□

Theorem 3.13 (Cocompleteness Theorem)**.** *A category $\mathcal{C}$ is cocomplete if and only if the coproduct of any set of objects in $\mathcal{C}$ exists and the coequalizer between any two morphisms with the same source and target exists.*

Proof. The proof is dual to the proof of theorem 3.12. Forward direction is obvious. The following diagram summarizes the proof of the reverse direction (notice how this is precisely the first diagram in the proof of theorem 3.12 with all arrows reversed):



For any morphism u in J, there are two morphisms from $F(\mathrm{Dom}(u))$ to $\coprod_{i\in\mathrm{Ob}(J)} Fi$: $\iota_{F(\mathrm{Dom}(u))}$ and $Fu\iota_{F(\mathrm{Cod}(u))}$. They give two cocones over the image under F of objects that are domains in J, so by universality we get two morphisms $f, g : \coprod_{u\in\mathrm{Hom}(J)} F(\mathrm{Dom}(u))$. $\mathrm{Coeq}(f, g)$ is the object Colim F. For any $u : i \to j$, in J, the cocone are the morphisms $\iota_{Fi}fp$, $\iota_{Fj}p$ from Fi, Fj respectively to $\mathrm{Coeq}(f, g)$.

To see the universality of $\mathrm{Coeq}(f, g)$, any object X with a cocone over F, has cocone over the image under F of all objects in J, and cocone over the image under F of all objects that are domains in J, so there are unique morphisms $m : \coprod_{u\in\mathrm{Hom}(J)} F(\mathrm{Dom}(u)) \to X$ and $n : \coprod_{i\in\mathrm{Ob}(J)} Fi \to X$. Uniqueness of m means $m = fn = gn$, so the universality of $\mathrm{Coeq}(f, g)$ gives a unique morphism $\mathrm{Coeq}(f, g) \to X$. □

4. Smith Normal Form and Finitely Generated Modules over a PID

Before we look at finding limits and colimits of modules, we need to develop some computational tools (the Smith normal form of a matrix) and know how to extract information about modules from homomorphisms between them.

Definitions 4.1. If I, J are two ideals in a ring, then by their **sum**, we mean the set

$$I + J = \{s + t \mid s \in I, t \in J\}.$$

A **Bézout domain** is an integral domain where the sum of two principal ideals is another principal ideal.

It is not hard to see that the sum of two ideals is also an ideal. Then it follows that every PID is a Bézout domain since every ideal in a PID is principal. In particular, the sum of two ideals (a), (b) in a PID is generated by $\gcd(a, b)$. The definition of a Bézout domain tells us is that in a PID, for any elements a, b, there are elements s, t such that $sa + tb = \gcd(a, b)$. This is the Bézout identity.

Definition 4.2. Let R be a PID. A matrix T with coefficients in R is said to be in a **Smith normal form** if T is diagonal, and has the following divisibility relation on its diagonal elements: $a_1|a_2|\cdots|a_k$, where a_i is the ith row and column entry.

Theorem 4.3 (Existence of a Smith normal form)**.** *Every nonzero $m \times n$ matrix T with entries in a PID, R, can be put in Smith normal form;* i.e.*, there exist invertible matrices A, B such that ATB is in Smith normal form.*

Remark 4.4. We want to find invertible matrices A and B, which basically limits us to the elementary row and column operations, which are: 1). multiplying a row/column by a unit; 2). adding two rows/columns together; and 3). interchanging two rows or columns. 1). limits the operations we can perform over PIDs compared to matrices over fields, because every nonzero element is a unit in a field, but far less elements are units in a PID.

Proof. Without loss of generality, we can assume the entry $a_{1,1} \neq 0$, interchanging rows and columns to achieve this if necessary. Consider first all the elements in the first column $a_{j,1}$, where $1 < j \leq m$. For each j, there exists $s, t \in R$ such that $sa_{1,1} + ta_{j,1} = u$, where $u = \gcd(a_{1,1}, a_{j,1})$ by Bézout's identity. This means

$$s(a_{1,1}/u) + t(a_{j,1}/u) = \pm 1. \tag{4.5}$$

Therefore, we can multiply on the left by the $m \times m$ identity matrix with the (j, j)th entry replaced by s, (j, k)th replaced by t, (k, j)th replaced by $a_{j,1}/u$ and (k, k)th replaced by $-a_{1,1}/u$. Notice how we used the coefficients in equation 4.5 to guarantee that the determiniant of the matrix multiplied on the left is a unit (hence the matrix itself is invertible). This replaces $a_{1,1}$ with u, and $a_{j,1}$ with 0 in the original matrix T.

As an example, we now show how $a_{2,1}$ is eliminated using the above method in a 3×3 matrix. Let $\sigma = a_{2,1}/u$ and $\tau = a_{1,1}/u$:

$$\begin{pmatrix} s & t & 0 \\ \sigma & -\tau & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \to \begin{pmatrix} u & * & * \\ 0 & * & * \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}.$$

After every $a_{j,1}$ $1 < j \leq m$ is 0, we can apply an analogous procedure to the row $a_{1,k}$ where $1 < k \leq n$. This time we will be using the analogous column operations. It is possible that after cancelling all the row elements $a_{1,k}$, we would have applied some column operations that make some elements of $a_{j,1}$ nonzero again. We may choose to cancel elements in the column again, but that leaves the possibility of getting back nonzero rows elements. We claim that by alternating between cancelling row elements and column elements, after a finite number of steps, every $a_{j,1}$ and $a_{1,k}$ will be zero.

Every time we cancel rows or columns, we replace the $a_{1,1}$ by u which divides $a_{1,1}$, so successive applications give a divisibility chain where each successive $(1, 1)$th entry divides the previous ones. This corresponds to an ascending chain of ideals, which must stabilize after finitely many steps because every PID is Noetherian.

The stabilized $(1,1)$th entry divides every other entry $a_{j,1}$ and $a_{1,k}$. In this case, Bézout's identity is something like

$$1(a_{1,1}/a_{1,1}) + 0(a_{j,1}/a_{1,1}) = 1(1) + 0(a_{j,1}/a_{1,1}) = \pm 1.$$

Using these coefficients the left (or right) multiplication simply cancels the row (or column) element by replacing that row (or column) by $a_{j,1}/a_{1,1}$ times the first row (or column). And since the first row (or column) is not altered here, elements $a_{j,1}$ and $a_{1,k}$ that have already been canceled will stay zero.

Repeat this cancellation procedure for successive entries along the diagonal, $a_{t,t}$, where we interchange rows and columns greater than t to try to get $a_{t,t} \neq 0$, and if we have $a_{t,t} \neq 0$, use the alternating procedure to cancel elements $a_{j,t}$ and $a_{t,k}$ where $t < j < m$ and $t < k < n$. This will be sure to leave what has already been canceled in its place. If it is not possible to perform row and column interchanges to get $a_{t,t} \neq 0$, then that means every $a_{j,k}$ where $t \leq j \leq m$ and $t \leq k \leq n$ is zero (everything to the right and below $a_{t,t}$ is zero, including itself). In this case, we are done. Otherwise, the process terminates when t reaches m or n, whichever one comes first.

At this point, we have a diagonalized matrix, but it may not satisfy the required divisibility relation. If $j < k$, we can make $a_j | a_k$ by first adding the kth column to the jth column (here, we only show the relevant j, kth rows and columns as a 2×2 submatrix)

$$\begin{pmatrix} a_j & 0 \\ a_k & a_k \end{pmatrix},$$

then from the Bézout identity $sa_j + ta_k = \gcd(a_j, a_k) = u$ we get the left multiplication

$$\begin{pmatrix} s & t \\ a_k/u & -a_j/u \end{pmatrix} \begin{pmatrix} a_j & 0 \\ a_k & a_k \end{pmatrix} \to \begin{pmatrix} u & ta_k \\ 0 & -a_j a_k/u \end{pmatrix}.$$

$u | ta_k$, so adding an appropriate number of copies of column j to column k gives

$$\begin{pmatrix} u & 0 \\ 0 & -a_j a_k/u \end{pmatrix}.$$

Since $pu = a_j$ and $qu = a_k$ for some $p, q \in R$, $-a_j a_k/u = -pqu$ is certainly divisible by u, hence $u | -a_j a_k/u$, the new kth entry along the diagonal. Applying this procedure starting from between a_1 and a_2, a_1 and $a_3 \ldots$, then a_2 and $a_3, \ldots$, etc. makes $a_1 | a_2, a_3 \ldots$, $a_2 | a_3, \ldots$, etc. which gives the desired divisibility conditions. □

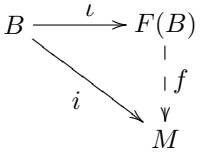
We now give a useful fact about modules.

Proposition 4.6. *Every R-module M is a quotient of a free R-module F.*

Proof. Let F be the free R-module on M, $F(M)$. $F(M)$ is the direct sum of copies of the ring R, indexed by M. For $i : M \to M$ the identity map, the universality of $F(M)$ gives an R-module homomorphism $f : F(M) \to M$ that agrees with i on all $m \in M$. [4, p. 471] Since i is surjective, so is f. Then by the first isomorphism theorem, $M \cong F(M)/\ker f$. □

For a finitely generated module over a PID, R, let B be the finite set of generators. The universal property on free R-modules which we used in proposition 4.6

gives the following commutative diagram.



Where ι is the usual inclusion from a summand to a direct sum. i is the inclusion that takes each generator to itself in M. f is unique; the f that makes this diagram commutes sends $(r_e)_{e\in B} \mapsto \sum_{e\in B} r_e e$. (where $r_e \in R$). $\ker f$ is the submodule of sequences $(r_e)_{e\in B}$ that make the sum $\sum_{e\in B} r_e e = 0$, we call these sums *relations* in M. $\ker f$ is also finitely generated, as a submodule of a finitely generated free R-module. Say $\ker f$ is generated by $\{k_1, \ldots, k_m\}$, where each $k_i = (a_{ij})_{j\in B}$ is a sequence in R. If $|B| = n$, then the $m \times n$ matrix where row i has the entries (a_{ij}) is called a **relation matrix** for M.

Lemma 4.7. *If a relation matrix T for a finitely generated module M over PID, R, is similar to a diagonal matrix S with diagonal entries $\{a_1, \ldots, a_k\}$, then*

$$M \cong R/(a_1) \oplus R/(a_2) \oplus \cdots \oplus R/(a_k).$$

Proof. S is another relation matrix for M. The kernel K is generated by each row, which is a sequence zero everywhere except a_i at the ith position for each $1 \le i \le k$. Consider also the map $\varphi : R^k \to R/(a_1) \oplus R/(a_2) \oplus \cdots \oplus R/(a_k)$ which sends $(r_1, \ldots, r_k) \mapsto (r_1 + (a_1), r_2 + (a_2), \ldots, r_k + (a_k))$. K is also the kernel of φ. φ is a projection in every component $R/(a_i)$ so is itself a surjection. Therefore, by the first isomorphism theorem, we have the following

$$M \cong R^k/K \cong R/(a_1) \oplus R/(a_2) \oplus \cdots \oplus R/(a_k). \quad \square$$

Theorem 4.8 (Structure theorem for finitely generated modules over a PID)**.** *Let R be a PID. Any finitely generated R-module M is a direct sum of a free part and a torsion part,* i.e.*,*

$$M \cong R^n \oplus R/(d_1) \oplus R/(d_2) \oplus \cdots \oplus R/(d_k),$$

where n is called the **free rank** *of M, and d_i are the principal generators of ideals in R, where $d_1|d_2|\cdots|d_k$. Furthermore, the free rank is unique, and the divisibility condition on the generators d_i make the entire decomposition unique.*

Proof. A relation matrix T for M is a matrix with entries in the PID, R. It is similar to a Smith normal matrix S, which by lemma 4.7 gives the isomorphism

$$M \cong R/(a_1) \oplus \cdots \oplus R/(a_n).$$

There are possibly zeros at the end of the diagonal of S. A quotient ring $R/(0) \cong R$, so

$$M \cong R/(a_1) \oplus \cdots \oplus R/(a_k) \oplus R^l.$$

Divisibility is built into the Smith normal form. $\square$

5. Finding Limits and Colimits of Special Modules

We now look at the category of left R-modules, ${}_R\mathbf{Mod}$. (Our discussion applies equally to the category of right R-modules $\mathbf{Mod}_R$). First, we give some definitions that will become useful.

Definitions 5.1. Let $\mathcal{M} = \{M_i\}_{i\in I}$ be a set of left R-modules indexed by set I. The **direct product** of modules in $\mathcal{M}$, denoted $\prod_{i\in I} M_i$, is the left R-module whose elements are sequences $(a_i)_{i\in I}$ where $a_i \in M_i$, and whose addition and ring multiplication are defined componentwise. The **direct sum** of modules in $\mathcal{M}$, denoted $\bigoplus_{i\in I} M_i$, is the left R-module whose elements are sequences $(a_i)_{i\in I}$ where $a_i \in M_i$ and only finitely many $a_i \neq 0$.

Definition 5.2. Let $T : M \to N$ be an R-module homomorphism. The **cokernel** of T, denoted coker$T$, is defined to be the quotient module $N/\mathrm{Im}T$. [2]

Proposition 5.3. *In ${}_R\mathbf{Mod}$, the product of a set $\mathcal{M} = \{M_i\}_{i\in I}$ of left R-modules is the direct product of modules in $\mathcal{M}$.*

Proof. The R-module homomorphisms from $\prod_i M_i$ to each module $M_k \in \mathcal{M}$ are the projection maps $p_{M_k} : \prod_i M_i \to M_k$, where $(a_i)_{i\in I} \mapsto a_k$; *i.e.*, p_{M_k} picks out the kth component of the sequence (a_i). Now suppose a left R-module X also has homomorphisms φ_{M_i} to every $M_i \in \mathcal{M}$. Define a homomorphism $f : X \to \prod_i M_i$ by $f = (\varphi_{M_i})_{i\in I}$; *i.e.*, each $x \in X$ is mapped to a sequence $(\varphi_{M_i}(x))_{i\in I}$. For each module $M_i \in \mathcal{M}$, it is clear that f composed with a projection to M_i gives the homomorphism from X to M_i, so $fp_{M_i} = \varphi_{M_i}$. Now suppose g is another homomorphism from X to $\prod_i M_i$ and $gp_{M_i} = \varphi_{M_i}$ for all $i \in I$. But this precisely means g takes x to a sequence $(g_i(x))_{i\in I}$ where projection in each M_i shows every $g_i = \varphi_{M_i}$, hence $f = g$. So $f = (\varphi_{M_i})$ is the unique homomorphism from X to $\prod_i M_i$. □

Proposition 5.4. *In ${}_R\mathbf{Mod}$, the coproduct of a set $\mathcal{M} = \{M_i\}_{i\in I}$ of left R-modules is the direct sum of modules in $\mathcal{M}$.*

Proof. The R-module homomorphism from each module $M_k \in \mathcal{M}$ to $\bigoplus_i M_i$ is the inclusion map $i_{M_k} : M_k \to \bigoplus_i M_i$, where $a_k \in M_k$ is mapped to the sequence with 0 everywhere except a_k in the kth component.

Suppose a left R-module Y also has homomorphisms $\tau_{M_i} : M_i \to \bigoplus_i M_i$ for each $M_i \in \mathcal{M}$. Define a homomorphism $f : \bigoplus_i M_i \to Y$ by

$$(a_i)_{i\in I} \mapsto \sum_{i\in I} \tau_{M_i}(a_i).$$

It is clear that for any $M_i \in \mathcal{M}$, $i_{M_i} f = \tau_{M_i}$. Now suppose g is another homomorphism from $\bigoplus_i M_i \to Y$ with $i_{M_i} g = \tau_{M_i}$ for every $M_i \in \mathcal{M}$. But every sequence $(a_i)_{i\in I}$ is equal to the sum $\sum_{i\in I}(b_{ik})_{k\in I}$, where (b_{ik}) is 0 everywhere except at $b_{ii} = a_i$. Each sequence $(b_{ik})_{i\in I}$ is the image of $a_i \in M_i$ under i_{M_i}, so

[2] As the name cokernel suggests, the kernel and cokernel are dual notions. Generally speaking, the kernel of a morphism $f : X \to Y$, if it exists, is the equalizer of f and the zero morphism from X to Y; the cokernel of f, if it exists, is the coequalizer of f and the zero morphism from X to Y. It is straightforward to check that our definitions for the kernel and cokernel in ${}_R\mathbf{Mod}$ satisfy the general definitions.

$g((b_{ik})_{k\in I}) = \tau_{M_i}(a_i)$. Therefore,

$$g((a_i)_{i\in I}) = g\big(\sum_{i\in I}(b_{ik})_{k\in I}\big) = \sum_{i\in I} g((b_{ik})_{k\in I}) = \sum_{i\in I}\tau_{M_i}(a_i),$$

hence $f = g$; f is the unique homomorphism from $\bigoplus_i M_i \to Y$. □

Proposition 5.5. *In ${}_R\mathbf{Mod}$, the equalizer between two homomorphisms $f, g : M \to N$ is the kernel of $f - g$, sometimes called the* **difference kernel** *of f and g.*

Proof. The homomorphism $e : \ker(f-g) \to M$ is the inclusion map. $(f-g)(x) = 0$ implies $f(x) = g(x)$, so for any $x \in \ker(f-g)$, $f(e(x)) = f(x) = g(x) = g(e(x))$ so $ef = eg$. Suppose X is another module with $\varphi : X \to M$ such that $\varphi f = \varphi g$. This means for any $x \in X$, $f(\varphi(x)) = g(\varphi(x))$ so $\varphi(x) \in \ker(f-g)$, so there is a homomorphism $h : X \to \ker(f-g)$ where $h = \varphi$, and clearly $he = \varphi$. If $l : X \to \ker(f-g)$ is another homomorphism such that $le = \varphi$, we know for any $x \in X$, e takes $l(x) \in \ker(f-g)$ to $l(x) = \varphi(x) \in M$, so $l = \varphi = h$; h is the unique homomorphism from X to $\ker(f-g)$. □

Proposition 5.6. *In ${}_R\mathbf{Mod}$, the coequalizer between two homomorphisms $f, g : M \to N$ is the cokernel of $f - g$, sometimes called the* **difference cokernel** *of f and g.*

Proof. The homomorphism $p : N \to \operatorname{coker}(f-g)$ is the projection map from N to its quotient $N/\mathrm{Im}(f-g)$. For any $m \in M$, $p(f(m)-g(m)) = 0$, so $p(f(m)) = p(g(m))$, hence $fp = gp$. Suppose Y is another module with $\varphi : N \to Y$ such that $f\varphi = g\varphi$. Then there is a homomorphism $h : \operatorname{coker}(f-g) \to Y$ where $h(x+\mathrm{Im}(f-g)) = \varphi(x)$. Say y is another representative of $x + \mathrm{Im}(f-g)$. Then we know $x - y \in \mathrm{Im}(f-g)$, and

$$h(x + \mathrm{Im}(f-g) - y + \mathrm{Im}(f-g)) = h(0 + \mathrm{Im}(f-g))$$

implies $\varphi(x) = \varphi(y)$, so h is well-defined. It is easy to see that $ph = \varphi$.

If $l : \operatorname{coker}(f-g) \to Y$ is another homomorphism such that $pl = \varphi$, then for any $n \in N$, $l(p(n)) = l(n + \mathrm{Im}(f-g)) = \varphi(n)$, so $l = h$; h is the unique homomorphism from $\operatorname{coker}(f-g)$ to Y. □

Corollary 5.7. ${}_R\mathbf{Mod}$ *is a complete and a cocomplete category.*

Proof. From theorems 3.12 and 3.13, we see that it is sufficient to show that products and coproducts exist for any set of R-modules, and the equalizer and coequalizer between any two R-module homomorphisms exist. We proved these four facts in propositions 5.3, 5.4, 5.5 and 5.6 □

Let R be a PID. Free R-modules are isomorphic to a direct sum of copies of the base ring R. Homomorphisms between finitely generated free R-modules can be conveniently expressed as a matrix with entries in R. If we have a finite set of finitely generated free R-modules, then the product and coproduct objects coincide. [3] They are isomorphic to the direct sum of copies of R, with free rank equal the sum of the free ranks of each component free R-module.

[3] More generally, finite products and finite coproducts in ${}_R\mathbf{Mod}$ coincide for *any* base ring R. This is in turn a special case of the general result that in an abelian category, finite products and finite coproducts coincide. See () Is this even true?????**********

For matrix manipulation, we will use the convention that a matrix acts on arguments on the right. (This is consistent with writing $g \circ f$ as fg). Under such a convention, The matrix that represents projection from a direct sum $\bigoplus_i M_i$ to a summand M_k is the kth column of the identity matrix from $\bigoplus_i M_i$ to itself. The matrix that represents the inclusion map $M_k \to \bigoplus_i M_i$ is a row of zeros, with the 1 at the kth column.

By the structure theorem 4.8, the kernel of a homomorphism f as a submodule of a finitely generated free R-module is also a finitely generated free R-module. After putting the matrix representing f into Smith normal form, the number of rows of zeros gives the free rank of the kernel.

The cokernel of a homomorphism $f : M \to N$ can be read directly from the Smith normal form of f as a matrix. It is the direct sum of quotients $R/(a_i)$ where the a_i are the diagonal entries of the Smith normal form. The reasoning is much like the one we used for lemma 4.7. In this case, the rows generate $\mathrm{Im}(f)$, and the codomain is N, so we have

$$N/\mathrm{Im}(f) = \mathrm{coker}(f) \cong L \cong R/(a_1) \oplus R/(a_2) \oplus \cdots \oplus R/(a_n).$$

for some R-module L.

Using the formula from the completeness and cocompleteness theorems 3.12 and 3.13, we now have an algorithm that allows us to calculate any limits and colimits over finitely generated free R-modules and homomorphisms between them.

Acknowledgments. I would like to thank my mentor John Wiltshire-Gordon for the support he has provided me throughout the REU experience. He was very helpful and enthusiastic to introduce different areas of mathematics to me. In addition, he has made me believe in the importance and utility of computation in abstract algebra.

References

[1] W. A. Adkins and S. H. Weintraub. *Algebra: An approach via module theory.* Springer. (1992). 307-18.

[2] D. S. Dummit and R. M. Foote. *Abstract algebra, 3rd ed.*. John Wiley & Sons Inc. (2004). 460-4.

[3] S. Mac Lane. *Categories for the working mathematician, 2nd ed.*. Springer. (1998). 7-19, 31-5, 55-9, 62-72, 109-15.

[4] J. J. Rotman. *Advanced modern algebra.* Prentice Hall. (2003). 471-3, 687-90.