Chapter 8  Basics of Module Theory:

Recall at the beginning of Chapter 6 we defined a vector space $V$ over a field $F$ as an additive group endowed with an action of $F$ that satisfied certain nice compatibilities between the operations. It is natural to ask if it is really necessary that $F$ be a field in order to get a nice theory. Here we relax the condition to include the case that $R$ is merely a ring. After developing the basics, we will then specialize to the case that $R$ is a PID.

Def: Let $R$ be a ring. A left $R$-module is an abelian group $(M, +)$ along with an action of $R$ on $M$, denoted by $rm$ for $r \in R, m \in M$ so that

1) $(r_1 + r_2)m = r_1m + r_2m \quad \forall r_1, r_2 \in R, m \in M$

2) $(r_1r_2)m = r_1(r_2m) \quad \forall r_1, r_2 \in R, m \in M$

If $R$ has an identity $1_R$, then

3) $1_Rm = m \quad \forall m \in M$

One can easily see that if $R$ is a field, this is precisely the definition of a vector space. Thus, a vector space is just a module defined over a field.
Write \( \text{End}_{\text{grp}}(M) \) for the set of group homomorphisms from \( M \) to \( M \). This set is a ring where addition is given by pointwise addition and multiplication is given by composition of functions: 
\[
(f + g)(m) = f(m) + g(m) \\
(f \cdot g)(m) = f(g(m)).
\]

One should check this is actually a ring. Suppose we are given a ring \( R \) and a ring homomorphism
\[
\varphi : R \to \text{End}_{\text{grp}}(M).
\]

Let \( r \cdot m = \varphi(r)(m) \). We claim this makes \( M \) into an \( R \)-module. Let \( r_1, r_2, r \in R \), with \( m, m_1, m_2 \in M \). We have

1) \[
(r_1 + r_2) \cdot m = \varphi(r_1 + r_2)(m) \\
\hspace{2cm} = (\varphi(r_1) + \varphi(r_2))(m) \quad (\text{cp ring homom}) \\
\hspace{2cm} = \varphi(r_1)(m) + \varphi(r_2)(m) \quad (\text{def of ring structure}) \\
\hspace{2cm} = r_1 \cdot m + r_2 \cdot m.
\]

2) \[
\varphi(r)(r_1 \cdot m + r_2 \cdot m) = \varphi(r)(r_1) \cdot m + \varphi(r)(r_2) \cdot m_2 \quad (\text{cp } \varphi(r) \in \text{End}_{\text{grp}}(M)) \\
\hspace{2cm} = \varphi(r)(r_1) \cdot m_1 + \varphi(r)(r_2) \cdot m_2 \\
\hspace{2cm} = r_1 \cdot m_1 + r_2 \cdot m_2.
\]

3) \[
(r_1 \cdot r_2) \cdot m = \varphi(r_1 \cdot r_2)(m) \\
\hspace{2cm} = \varphi(r_1)(\varphi(r_2)(m)) \quad (\text{by def of mult in } \text{End}_{\text{grp}}(M) \text{ and that } \varphi \text{ is ring hom}) \\
\hspace{2cm} = \varphi(r_1)(r_2 \cdot m).
\]
If \( R \) has 1 and \( \text{End}_{\mathbb{Z}}(M) \) has 1, then if we assume \( \phi \) sends \( 1_r \) to \( 1_{\text{End}_{\mathbb{Z}}(M)} \), we have the result in this case as well. As we see that given a ring homomorphism \( \phi \) from \( R \) to \( \text{End}_{\mathbb{Z}}(M) \), we obtain an \( R \)-module structure on \( M \).

Now suppose that \( M \) is an \( R \)-module. We obtain a ring homomorphism \( \phi : R \rightarrow \text{End}_{\mathbb{Z}}(M) \) by setting

\[
\phi(r)(m) = rm.
\]

It is an exercise to check this is a ring homomorphism.

This shows that \( M \) is an \( R \)-module if and only if there is a ring homomorphism from \( R \) to \( \text{End}_{\mathbb{Z}}(M) \).

**Def:** Let \( R \) be a ring and \( M \) an \( R \)-module. An \( R \)-submodule of \( M \) is a subgroup \( N \) of \( M \) that is closed under the action of \( R \), i.e., \( \forall n \in N \land \forall r \in R, nr \in N \).

**Examples:**

1) Let \( R \) be a ring and set \( M = R^n \). Then \( M \) is an \( R \)-module via componentwise addition and

\[
(r_1, \ldots, r_n) + (s_1, \ldots, s_n) = (r_1 + s_1, \ldots, r_n + s_n), \quad (r \in R) \land (r_1, \ldots, r_n) \in R^n
\]

2) Let \( M = \mathbb{Z} \langle e \rangle \). Then we have \( M \) is a \( \mathbb{Z} \)-module via
usual addition and

\[ n \cdot (a + bi) = na + nbi \]

for \( n \in \mathbb{Z}, a, b \in \mathbb{R} \).

3) More generally, let \( G \) be any abelian group. Then, \( G \) is a \( \mathbb{Z} \)-module via

\[ ng = g + \ldots + g \quad (n \text{-times}) \]

if \( n \geq 0 \). For \( n = 0 \), set \( ng = 0_G \), which we will write as \( 0 \) since \( G \) is an abelian group written additively, and

for \( n < 0 \), write

\[ ng = -g - \ldots - g \quad (|n| \text{-times}) \]

Conversely, any \( \mathbb{Z} \)-module \( G \) is clearly an abelian group. Thus, we have that abelian groups and \( \mathbb{Z} \)-modules are the same thing. We will use this later to prove structure theorems such as the Fundamental Theorem of Finitely Generated Abelian Groups by deducing it as a special case of a basic structure theorem for modules over a PID.

**Exercise:** Show the notion of submodule of a \( \mathbb{Z} \)-module is coincident with subgroup of an abelian group.
Let $K/Q$ be a Galois extension of fields and let $G = \text{Gal}(K/Q)$.

Let $E$ be an elliptic curve defined over $K$, namely, $E$ is given by an equation of the form

$$y^2 = x^3 + ax + b \quad a, b \in K$$

along with a "point at oo". Let $E(K)$ be the set defined as

$$E(K) = \{ (\alpha, \beta) \in K^2 : \beta^2 = \alpha^3 + a\alpha + b \} \cup \{oo\}.$$

We can represent this graphically as follows:

![Graph of an elliptic curve](image)

(Where we have really drawn $E(K)$ to get a nice picture.)

The set $E(K)$ can be given the structure of a group as follows: draw a line through $P$ and $Q$. Counting multiplicity, it hits $E(K)$ in a third spot. Now draw a vertical line. The point at oo is what the line goes through off the page. The third intersection point on the page is $Poo$. 

The group $E(k)$ is a $\mathbb{Z}[G]$-module, where the action of $G$ is given by $\sigma \in G$, $\tau((a, \rho)) = (\sigma a, \sigma \rho)$. One can check this gives a module structure on $E(k)$.

5) Let $F$ be a field and $V$ a vector space over $F$. Let $R = F[x]$.

Let $T \in \text{Hom}_F(V, V)$. We can use $T$ to make $V$ into a $R$-module. The module structure depends on $T$ of course.

Let $T^n = T \cdots T$ for $n \in \mathbb{Z}_{\geq 0}$ and $T^0 = \text{id}$.

Let $f(x) \in F[x]$ and $v \in V$. Write

$$f(x) = a_n x^n + \cdots + a_1 x + a_0, \quad a_i \in F.$$  

Define

$$f(x) \cdot v = (a_n T^n + \cdots + a_1 T + a_0)(v).$$

One can now check that this makes $V$ into a $F[x]$-module.

The choice of $T$ is important and for different choices one obtains different modules. One should pair a vector space $V$ and a few different $T$'s and work out the details to get a feel for this.

This shows that a vector space $V$ over $F$ and $T \in \text{Hom}_F(V, V)$ gives rise to a $F[x]$-module. Conversely, given a $F[x]$-module $V$, we obtain a vector space $V$ over $F$ and a linear transformation $T \in \text{Hom}_F(V, V)$. This is given by restricting the action to $F$ to get $V$. 
vector space and defining $T$ via

$$T(v) = x \cdot v.$$ 

Thus, we have a bijection

$$\left\{ \text{V an $F$-vector space} \right\} \quad \leftrightarrow \quad \left\{ \text{V an $F[x]$-module} \right\}.$$ 

It is natural to ask about submodules in this set-up. Given a subspace $W \subseteq V$ and $T \in \text{Hom}_F(V, V)$, we say $W$ is $T$-stable if $T(W) \subseteq W$. Let $W \subseteq V$ be an $F[x]$-submodule. Then clearly we must have $x \cdot W \subseteq W,$

i.e., $W$ must be $T$-stable. Conversely, if we let $T \in \text{Hom}_F(V, V)$ and $W$ a $T$-stable subspace, then $T^n(W) \subseteq W$ for all $n \geq 0$ so fix $w \in W$ for all fixed $F[x]$. Thus, we have a bijection

$$\left\{ \text{W an $F$-module} \right\} \quad \leftrightarrow \quad \left\{ \text{W a subspace} \right\}.$$ 

**Prop. 8.1:** Let $R$ be a ring and $M$ an $R$-module. A subset $W \subseteq M$

is a submodule of $M$ iff

1) $W \neq \emptyset$

2) $x, y \in W \Rightarrow x + y \in W, r \cdot x \in W, \forall r \in R.$

**Proof:** Exercise.
Def: Let $M$ and $N$ be $R$-modules.

1) A map $\phi: M \to N$ is an $R$-module homomorphism if
   \[
   \begin{align*}
   &\phi(x+y) = \phi(x) + \phi(y) \quad \forall \ x, y \in M \\
   &\phi(rx) = r\phi(x) \quad \forall \ r \in R, \ x \in M.
   \end{align*}
   \]
   The set of all $R$-module homomorphisms is denoted $\text{Hom}_R(M, N)$.

2) We say $\phi \in \text{Hom}_R(M, N)$ is an isomorphism if $\phi$ is bijective.
   We write $M \cong N$ if there is an isomorphism $\phi$ from $M$ to $N$ and say $M$ is isomorphic to $N$.

3) Let $\phi \in \text{Hom}_R(M, M)$. The kernel of $\phi$ is given by
   \[\ker \phi = \{x \in M : \phi(x) = 0\}\]
   and the image is given by
   \[\text{Im} \phi = \{\phi(x) : x \in M\}\]

Prop. 8.2: Let $\phi \in \text{Hom}_R(M, N)$. The kernel and image of $\phi$ are

submodules of $M$ and $N$ respectively.

Proof: Exercise.

Example: Let $M$ and $N$ be $\mathbb{Z}$-modules, i.e., $M$ and $N$ are abelian

groups. Let $\phi \in \text{Hom}_{\mathbb{Z}}(M, N)$. Clearly we have $\text{Coker} \phi \in \text{Hom}_{\mathbb{Z}}(M, N)$

by definition. Now suppose $\psi \in \text{Hom}_{\mathbb{Z}}(M, N)$. Then we have

$\psi(x+y) = \psi(x) + \psi(y)$

by definition. Furthermore, we have

$\psi(nx) = \psi(x + \cdots + x) = \psi(x) + \cdots + \psi(x) = n \psi(x)\]
for $n > 0$, 
\[ \psi(x) = 0 \]
and 
\[ \psi(-nx) = \psi(-x - \cdots - x) = -\psi(x) - \cdots - \psi(x) = -n\psi(x) \]

for $n > 0$, so $\psi \in \text{Hom}_\mathbb{Z}(M, N)$. Thus, we have 
\[ \text{Hom}_\mathbb{Z}(M, N) = \text{Hom}_\text{grp}(M, N). \]

**Prop. 3.3:** Let $M, N,$ and $L$ be $R$-modules.

1) A map $\phi : M \rightarrow N$ is in $\text{Hom}_R(M, N)$ iff 
\[ \phi(rx + y) = r \phi(x) + \phi(y) \]
for every $r \in R$, $x, y \in M$.

2) Let $\phi, \psi \in \text{Hom}_R(M, N)$. Define $\phi + \psi$ by 
\[ (\phi + \psi)(x) = \phi(x) + \psi(x). \]
Then $\phi + \psi \in \text{Hom}_R(M, N)$ and as $\text{Hom}_R(M, N)$ is an
abelian group. If $R$ is a commutative then we
define 
\[ (r\phi)(x) = r \phi(x). \]
We have 
\[ r\phi \in \text{Hom}_R(M, N) \]
and as if $R$ is commutative then $\text{Hom}_R(M, N)$
is an $R$-module.

3) Let $\phi \in \text{Hom}_R(M, N)$, $\psi \in \text{Hom}_R(N, L)$. Then $\psi \phi \in \text{Hom}_R(M, L)$.

4) If we define multipl. as function composition, we have 
\[ \text{Hom}_R(M, M) \]
is a ring with identity.

**Proof:** Most of these results are straightforward so we only hit
the highlights.

1) This is an easy exercise.

2) This is indeed straightforward. The only point to be made is that we require \( R \) to be commutative for \( r \circ p \) to still be in \( \text{Hom}_R(M,N) \) because we must have
\[
(r \circ p)(sx) = r((cp)(sx)) = r(scp(s)) = r(s \circ p(x)) = s \circ r(p(x)),
\]
so we must be able to interchange \( r \) with \( s \).

3) Exercise.

4) Exercise.

Recall that when we studied vector spaces, we were able to use the group structure of the underlying vector space to define quotient of vector spaces. We can do the same thing here. In particular, given an \( R \)-module \( M \) and a sub-module \( N \), we set \( M/N \) to be the quotient group. Since \( M \) is abelian, \( N \subseteq M \) as a group so \( M/N \) is again an abelian group under addition. We now see it is in fact
Prop. 8.4: Let $R$ be a ring, $M$ an $R$-module, and $N$ an $R$-submodule. The quotient group $M/N$ is an $R$-module with scalar multiplication given by

$$r(x+N) = rx + N \quad \text{for } r \in R, x \in M.$$ 

The natural projection

$$\pi : M \to M/N$$

$$x \mapsto x+N$$

is a module homomorphism with $\ker(\pi) = N$.

Proof: We already have $M/N$ is an abelian group under addition. First we show the action is well-defined. Let $x+N = y+N$ so $x-y \in N$. Then $r(x-y) \in N$ since $N$ is a submodule, i.e., $rx+N = ry+N$ and so the action is well-defined.

It is now straightforward to check this is an $R$-module, for example, let $r_1, r_2 \in R$, $x+N \in M/N$, then

$$(r_1 + r_2)(x+N) = (r_1 + r_2)x + N$$

$$= (r_1x + r_2x) + N$$

$$= r_1x + N + r_2x + N$$

$$= r_1(x+N) + r_2(x+N).$$

It remains to show $\pi$ is a surjection.
homom. with kernel $N$. We know it is a group homom. from our work in group theory. We have

$$
\pi (rx) = r x + N
$$

$$
= r (x + N)
$$

$$
= r \pi (x),
$$

so $\pi \in \text{Hom}_R(M, M/N)$. We also have $\ker(\pi) = N$ because if we consider $\pi \in \text{Hom}_{\text{grp}}(M, M/N)$, then $\ker(\pi) = N$ and by the definitions the kernels are the same when one considers a map as a group homom. or a module homom.

As one might expect, we have the analogous isomorphism theorem as were given for groups. The proofs are similar to as above: one has the result for groups so we need only check the maps used respect scalar multipl.

**Theorem 8.5:** 1) (1st Isom. Thm) Let $M, N$ be $R$-modules,

$\Phi \in \text{Hom}_R(M, N)$. Then $\text{im}(\Phi) \cong \text{ker}(\Phi)$.

2) (2nd Isom. Thm) Let $A, B$ be submodules of $M$.

Then $\text{(A+B)}_B \cong \text{A}/(A \cap B)$. 

3) (3rd Isom. Thm) Let $M$ be a $R$-module, $A, B$ $R$-submodules of $M$ with $A \subseteq B$. Then

$$\frac{(M/A)}{(M/B)} \cong M/B.$$ 

4) (Lattice Isom. Thm) Let $N \subseteq M$ be a submodule. Then there is a bijection between the submodules of $M$ that contain $N$ and the submodules of $M/N$ given by

$$A \leftrightarrow A/N.$$ 

This correspondence commutes with the processes of taking sums and intersections.

One of the nicest properties of a finite dimensional vector space is that it has a basis of finitely many elements and it is isomorphic to a power of the field. For modules, this leads us into the concepts of finitely generated modules and free modules.

**Def.** Let $M$ be a $R$-module and $N_1, \ldots, N_n$ submodule.

1) The sum of $N_1, \ldots, N_n$ is defined by

$$N_1 + \cdots + N_n = \gen{\sum x_i \in N_i : x_i \in N_i}.$$ 

2) Let $A \subseteq M$ be a subset. Let

$$RA = \gen{r_1a_1 + \cdots + r_ma_m : r_i \in R, a_i \in A, m \in \mathbb{N}}.$$ 

If $A$ is finite, say $A = \{a_1, \ldots, a_n\}$, then we write $Ra_1 + \cdots + Ra_n$ for $RA$. The set $RA$ is called the
submodule generated by \( A \). (You should check it is a submodule!)

If \( N \) is a submodule of \( M \) and \( N = RA \), we say \( N \) is generated by \( A \) and the elements of \( A \) are referred to as generators for \( N \).

3) A submodule \( N \) of \( M \) (\( N \subseteq M \) is included) is said to be finitely generated if there is a finite subset \( A \) of \( M \) so that \( N = RA \).

4) A submodule is said to be cyclic if it is generated by one element.

Examples:

1) Let \( R = \mathbb{Z} \) be any ring and consider \( R \) as a module over itself. Assume \( R \) has identity \( 1_R \). Then \( R \) is cyclic since \( R \) is generated by \( 1_R \). The submodules of \( R \) that are cyclic are precisely the principal ideals.

For example, if \( R = \mathbb{Z} \), then every submodule is cyclic since \( \mathbb{Z} \) is a PID.

2) Let \( V \) be a finite dimensional vector space over \( \mathbb{F} \). Then \( V \) has a basis \( x_1, \ldots, x_n \) over \( \mathbb{F} \), and so \( V \) is finitely generated over \( \mathbb{F} \), i.e.

\[ V = \mathbb{F} x_1 + \cdots + \mathbb{F} x_n. \]

We know submodules in this case correspond to vector subspaces. These are all finitely generated as
well since subspaces of a vector space have finite dimension if the vector space is finite dimensional.

3) Let \( R \) be a ring and let \( S = R[x_1, x_2, \ldots] \) be the polynomial ring in infinitely many variables. As we noted in the first example, \( S \) is cyclic over \( S \) so certainly is finitely generated. However, \( N = (x_1, x_2, \ldots) \) is a submodule of \( S \) that is not finitely generated.

This shows that submodules of a finitely generated module need not be finitely generated.

4) Let \( M = F[x] \). This is a module over \( F \). It is not finitely generated since given any finite set \( \{f_1(x), \ldots, f_m(x)\} \), if we set \( n = \max \deg f_i(x) \), then \( x^m \) will not be in \( F f_1(x) + \ldots + F f_m(x) \). We have \( M \) is generated by \( \{x^i\}_{i=0}^{\infty} \) over \( F \).

---

**Def.:** Let \( \{M_j : j \in I\} \) be an arbitrary collection of \( R \)-modules. The direct product of the \( M_j \) is defined to be their direct product as abelian groups with standard multiplication defined componentwise, i.e., if \( f_j : M_j \to R \) and \( g_j : M_j \to R \)

\[
\prod_{j \in I} M_j = \left\{ (f_j)_{j \in I} : f_j \in M_j \right\}
\]

with \( (f_{j_1}, \ldots, f_{j_m}) + (g_{j_1}, \ldots, g_{j_m}) = \left( f_{j_1} + g_{j_1}, \ldots, f_{j_m} + g_{j_m} \right) \).
The direct sum of $\prod M_j$ is defined to be the direct sum as abelian groups with componentwise scalar multiplication, i.e.,

$$\bigoplus M_j = \bigoplus m_j : m_j \in M_j, \ m_j = 0 \text{ for all but finitely many } j?$$

with $r \otimes m_j = \otimes rm_j$ and $\otimes m_j \otimes m_j = \otimes (m_j + m_j')$.

One can easily see that these notions coincide in the case where the indexing set is finite.

**Prop. 8.6:** Let $M$ be an $R$-module and $N_1, \ldots, N_n$ submodules. T.F.A.E.

1. The map $\pi : N_1 \times N_2 \times \cdots \times N_n \rightarrow N_1 + N_2 + \cdots + N_n$

   $$(x_1, x_2, \ldots, x_n) \mapsto x_1 + x_2 + \cdots + x_n$$

   is an isomorphism of $R$-modules.

2. $N_1 \cap (N_1 + N_2 + \cdots + N_n) = 0$ for all $j$ except $j$.

3. Every $x \in N_1 + \cdots + N_n$ can be written uniquely in the form $x_1 + x_2 + \cdots + x_n$ for some $x_i \in N_i$.

**Proof:**

1) $\Rightarrow$ 2) Suppose there is a $j$ so that $N_j \cap (N_1 + N_2 + \cdots + N_n) \neq 0$, and let $x_j \neq 0$ be in this intersection. Then $\exists x_1, \ldots, x_n \in \ker \pi$ with $x_i \in N_i$ so that

$$x_j = x_1 + \cdots + x_j + x_{j+1} + \cdots + x_n.$$

Thus,

$$x_1, x_2, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n \in \ker \pi$$

and is nonzero, a contradiction to 2). Thus, 1) $\Rightarrow$ 2).

2) $\Rightarrow$ 3) Suppose 2 holds. Suppose $x, y \in N_i$ so that

$$x_1 + \cdots + x_n = y_1 + \cdots + y_n.$$
Suppose for some \( j \) we have \( x_j \neq y_j \). Then

\[
0 \neq x_j - y_j = (y_j, x_1) + \cdots + (y_j, x_n) + (y_j, x_1) + \cdots + (y_n, x_n),
\]

so

\[
x_j - y_j \in N_j \cap (N_1 + \cdots + N_j + \cdots + N_n),
\]

which contradicts 2), so \( 1 \Rightarrow 3 \).

3) = 1) Finally, assume 3) holds. Note \( \pi \) is clearly surjective and a homomorphism. The fact that 3) holds gives that \( \pi \)
is injective, so we have the result.

If there are submodules \( N_1, \ldots, N_n \) of \( M \) so that

\[
M = N_1 + \cdots + N_n
\]

and any of the equivalent conditions of Prop. 8.6 are satisfied, we say \( M \) is the \underline{internal direct sum} of the \( N_i \) and write \( M = N_1 \oplus \cdots \oplus N_n \). Note that this means every element of \( M \) can be written uniquely in the form \( x_1 + \cdots + x_n \) with \( x_i \in N_i \).

We now extend this notion to free modules.

**Def.** An \( R \)-module \( F \) is said to be \underline{free} on the set \( A \) if for every nonzero \( x \in F \), there exist unique \( r_1, \ldots, r_n \in R \) and \( a_1, \ldots, a_n \in A \) so that

\[
x = r_1 a_1 + \cdots + r_n a_n
\]

for some \( n \in \mathbb{Z}^+ \). We say \( A \) is a \underline{basis} or \underline{set of free generators} for \( F \) and the \underline{cardinality} of \( A \) is said to be \underline{the rank} of \( F \). We remark on page 20 that a definition.
This notion can be confusing at first and we need to distinguish it from the internal direct sum. We illustrate with a couple of typical examples.

**Example:** 1) Let $V$ be a finite dimensional vector space over a field $F$. Let $e_1, \ldots, e_n$ be a basis for $V$. Then $V$ is free on the set $\{e_1, \ldots, e_n\}$. Certainly every element in $V$ can be written as a sum

$$c_1 e_1 + \cdots + c_n e_n$$

for some $c_i \in F$. Suppose we have

$$x = c_1 e_1 + \cdots + c_n e_n$$

$$= d_1 e_1 + \cdots + d_n e_n.$$

Then

$$(c_1 - d_1)e_1 + \cdots + (c_n - d_n)e_n = 0,$$

so from vector space theory we conclude $c_i = d_i$.

2) Now consider $R = \mathbb{Z}$ and $M = \mathbb{F}_p \oplus \cdots \oplus \mathbb{F}_p \cong \bigoplus_{i=1}^m \mathbb{F}_p \cong \mathbb{F}_p^n$.

This will not be a free module over $R$, though from 1) we see it is free over $\mathbb{F}_p$. To see it is not free over $R = \mathbb{Z}$, consider a set $\{x_1, \ldots, x_m\}$, suppose we have $x \in M$ can be written as

$$x = c_1 x_1 + \cdots + c_m x_m$$

for some $c_i \in \mathbb{Z}$. We see immediately that

$$x = (c_1 + p) x_1 + \cdots + (c_m + p) x_m.$$
so the coefficients are not unique. This illustrates the main difference between free and a direct sum.

For free, one requires coefficients to be unique as well. Our second representation $x$ above does not violate the uniqueness for the direct sum because in $\mathbb{F}_p$ we have $(c+p)x_i = c_i x_i$, so each term is unique as an element of $\mathbb{F}_p$.

3) Let $R = \mathbb{Z}$ and $M = \mathbb{Z}^n$. Let $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $e_n = (0, \ldots, 0, 1)$. Then $M$ is free on the set $\{ e_1, \ldots, e_n \}$.

Write $x = (x_1, \ldots, x_n) \in M$. Then

$$x = x_1 e_1 + \cdots + x_n e_n$$

and one has that the $x_i$ are unique here.

Given a set $A$, we can form a free module on the set.

and have a nice universal property characterizing this module.

**Theorem 8.7:** Let $A$ be any set and $R$ a ring. Then there is a free $R$-module $F(A)$ on $A$ that satisfies the following universal property: if $M$ is any $R$-module and $\phi: A \to M$ is any map of sets, then there is a unique $\tilde{\phi} \in \text{Hom}_R(F(A), M)$ so that the following diagram commutes:

```latex
\begin{array}{ccc}
A & \xrightarrow{\phi} & M \\
\downarrow{\tilde{\phi}} & & \downarrow{id} \\
F(A) & \xrightarrow{=} & M
\end{array}
```
If $A$ is a finite set $\{x_1, \ldots, x_n\}$, then $F(A) = \mathbb{R} x_1 \oplus \cdots \oplus \mathbb{R} x_n \cong \mathbb{R}^n$.

**Proof:** If $A = \emptyset$, we set $F(A) = \emptyset$ and there is nothing to show so assume $A \neq \emptyset$. Let $F(A)$ be the collection of set maps $f : A \to \mathbb{R}$ so that $f(a) = 0$ for all but finitely many $a \in A$. For $f, g \in F(A)$ and $r \in \mathbb{R}$, define $f + g$ and $rf$ via
\[
(f + g)(a) = f(a) + g(a) \\
(rf)(a) = rf(a).
\]
We claim this makes $F(A)$ into an $\mathbb{R}$-module. We have associativity of addition because $(\mathbb{R}, +)$ is an abelian group. The identity is the function that sends $a$ to 0 for every $a \in A$. The inverse of $f$ is $-f$ where $(-f)(a) = -f(a)$. Thus, we have $F(A)$ is an abelian group. The module conditions are just as easy to check, for example,
\[
((r+s)f)(a) = (r+s)f(a) \\
= r f(a) + s f(a) \\
= (rf)(a) + (sf)(a).
\]
Thus, we have an $\mathbb{R}$-module $F(A)$.

We can realize $A$ as a subset of $F(A)$ by sending $a$ to the function $f_a$ that sends $a$ to 1 and everything else in $A$ to 0. This allows us to
View \( f \in F(A) \) as a finite sum
\[ r_1 a_1 + \ldots + r_n a_n \]
where \( f(a_i) = r_i \) and \( f(a) = 0 \) if \( a \notin \{a_1, \ldots, a_n\} \).

Clearly each element in \( F(A) \) can be expressed uniquely in this way.

It remains to establish the universal property of \( F(A) \). Let \( \phi \) be a map of \( A \) into an \( R \)-module \( M \). Define

\[ \Phi : F(A) \rightarrow M \]

\[ \sum_{i=1}^{n} r_i a_i \mapsto \sum_{i=1}^{n} r_i \phi(a_i) \]

where we use the above identification of \( F(A) \) with finite sums \( \sum_{i=1}^{n} r_i a_i \). This map is well-defined by the uniqueness of such representations. Moreover, if we write \( f = \sum_{a \in A} r_a a \), where \( r_a = 0 \) for all but finitely many \( a \) and \( g = \sum_{a \in A} s_a a \), then \( rf \) is given by \( \sum_{a \in A} r_a s_a a \), and

\[ f + g = \sum_{a \in A} (r_a + s_a) a \]  

This allows us to immediately see \( \Phi \) is an \( R \)-module homomorphism. If we restrict \( \Phi \) to \( A \), we have \( \Phi(a) = \phi(a) \) and so the diagram

\[
\begin{array}{ccc}
A & \xrightarrow{\text{incl}} & F(A) \\
\downarrow{\Phi} & & \downarrow{\Phi} \\
M & \xrightarrow{\phi} & M
\end{array}
\]

commutes. Finally, if \( \Phi' : F(A) \rightarrow M \) is another extension
\[ \varphi(a) = \phi(a) = \varphi(a), \]

and since any \( R \)-module homomorphism on \( F(A) \) is determined by its values on \( A \), we have \( \varphi = \phi \). So \( \varphi \) is the unique extension of \( \phi \).

If \( A = \mathbb{F} \times \cdots \times \mathbb{F} \), then Prop. 8.6 gives \( F(A) \cong \mathbb{F} \times \cdots \times \mathbb{F} \). Now one has \( \mathbb{F} \cong \mathbb{F} \times \cdots \mathbb{F} \times \mathbb{F} \). So \( \mathbb{F} \)-modules via the map \( r \mapsto r \mathbb{F} \), so \( F(A) \cong \mathbb{F}^n \).

**Corollary 8.8:** 1. If \( F_1 \) and \( F_2 \) are free modules on the same set \( A \), there is a unique isomorphism between \( F_1 \) and \( F_2 \).

   That is the identity map on \( A \). Note that this shows the

   second map \( \varphi_1 \) is free, and \( \varphi_2 \) is free, respectively.

   Since \( A \) is connected, \( \psi_1 \) and \( \psi_2 \) are the same.

   2. If \( F_1 \) is any free \( \mathbb{F} \)-module on \( A \), then \( F(A) \cong \mathbb{F} \).

**Proof:** Since \( F_2 \) is free on \( A \), we have a natural map \( \varphi_2 : A \to F_2 \) as inclusion. Thus, Theorem 8.7 gives the commutative diagram

\[
\begin{array}{cc}
A & \xrightarrow{\varphi_1} & F_1 \\
\downarrow{\varphi_2} & & \downarrow{\psi_1} \\
F_2 & & \\
\end{array}
\]

However, applying the same property to \( F_1 \) gives an

inverse map \( \psi_1 : F_2 \to F_1 \). Composing \( \psi_1 \) and \( \psi_2 \)
gives a map \( F_1 \xrightarrow{\psi_1} F_2 \xrightarrow{\psi_1} F_1 \). Since each map

is unique, the composition is unique and so must be
Def: 1) The left $R$-module $M$ is said to be a **Noetherian $R$-module** or to satisfy the ascending chain condition (a.c.c) on submodules if given any sequence of submodules

\[ M_1 \subseteq M_2 \subseteq M_3 \subseteq \ldots \]

there is an integer $m$ so that for all $k \geq m$, we have $M_k = M_m$; i.e., the chain stabilizes.

2) A ring $R$ is said to be **Noetherian** if it is Noetherian as a left $R$-module over itself.

**Theorem 8.9:** Let $R$ be a ring and $M$ a left $R$-module. T.F.A.E.:

1) $M$ is Noetherian

2) Every nonempty set of submodules of $M$ contains a maximal element under inclusion.

3) Every submodule of $M$ is finitely generated.

**Proof:** Suppose $\mathcal{N} = \{N_i\}$ be a nonempty collection of submodules. If $N_1$ is maximal we are done; if not we have $N \subseteq N_1 \subseteq N_2$ with $N \subseteq N_2$. If $N_2$ is maximal we are done, and if not then $N_1 \subseteq N_3 \subseteq \cdots \subseteq N$, continuing in this pattern, if we do not have a maximal element then we can construct an infinite strictly increasing sequence of submodules, contradicting 2).
2. Suppose \( N \) is not finitely generated. Let \( N \) be the collection of finitely generated submodules of \( N \). Then there is a maximal element \( N' \) of \( N \). We claim \( N' = N \). If not, there is a \( x \in N - \langle N' \rangle \). But then \( N'' = N' + \langle x \rangle \) is a finitely generated submodule of \( N \) that properly contains \( N' \). This contradicts the maximality of \( N' \) as we must have \( N = N' \), so \( N \) is finitely generated.

3. Suppose \( N_1 \subseteq N_2 \subseteq N_3 \subseteq \ldots \) is a sequence of submodules of \( M \). Then set

\[
N = \bigcup_{j \geq 1} N_j.
\]

It is easy to check this is again a submodule of \( M \), and so by assumption is finitely generated. Let \( x_1, \ldots, x_n \) be generators for \( N \). Then there exist some \( m \geq 1 \) so that \( x_1, \ldots, x_n \in N_m \). Thus, \( N = N_m \) and so the sequence stabilizes and \( M \) is Noetherian. \( \blacksquare \)

As an immediate consequence we have the following corollary.

**Cor. 5.10:** If \( R \) is a PID then every nonempty set of ideals in \( R \) has a maximal element and \( R \) is a Noetherian ring.

Our next major goal is a classification of finitely generated modules over PIDs. This will give the Fundamental Theorem for Finitely Generated Abelian Groups as an immediate
Prop. 5.11: Let $R$ be an integral domain and $M$ a free $R$-module of finite rank $n$. Then given any $x_1, \ldots, x_m \in M$ there are elements $r_1, \ldots, r_m \in R$, not all zero, so that

$$r_1 x_1 + \cdots + r_m x_m = 0,$$

i.e., any collection of $m$ elements is linearly dependent.

Proof: Let $F = \text{Fr}_n(R)$. Since $M$ is free with $n$ generators we have $M \cong R \oplus \cdots \oplus R \cong F \oplus \cdots \oplus F$. We have $F^n$ is an $n$-dimensional vector space over $F$, so considering $x_1, \ldots, x_m$ as elements of $F^n$, we must have $a_1, \ldots, a_m \in F$, not all zero, so that

$$a_1 x_1 + \cdots + a_m x_m = 0.$$

Since $a_i \in F = \text{Fr}_n(R)$, we can write $a_i = b_i/c_i$ for some $b_i, c_i \in R$. If we set $C = c_1 \cdots c_m$, then

$$C a_1 x_1 + \cdots + C a_m x_m = 0$$

and $C a_i \in R$, which gives the result.

Def.: Let $R$ be an integral domain. The rank of an $R$-module $M$ is the maximal number of linearly independent elements in $M$.
In the case that \( R \) is an integral domain and \( M \) is a free module, we must show the two notions of rank agree. First we show the original definition of rank is well-defined if \( F \) is free of finite rank. Let \( R \) be commutative and assume \( R^n \cong R^m \).

Let \( m \) be a maximal ideal of \( R \). Then from your homework you have seen that

\[
\frac{R^n}{m \cdot R^n} \cong \frac{R^n}{m} \times \cdots \times \frac{R^n}{m} \cong F \times \cdots \times F = F^n
\]

for \( F = R/m \). However, if \( R^n \cong R^m \), then we get \( \frac{R^n}{m \cdot R^n} \cong \frac{R^n}{m \cdot R^n} \)

and so \( F^n \cong F^m \). We may use matrix space theory to conclude \( m = n \).

We may assume \( R \) is an integral domain. Let \( F \) be a free \( R \)-module of finite rank \( n \). Then \( F \cong R^n \).

Let \( \{x_1, \ldots, x_n\} \) be a set of linearly independent elements in \( F \). Then \( n = \text{rank } F \).

Thus, \( F \) has a basis with \( n \) elements, so unless \( m \) is maximal, this contradicts the maximality of \( n \). Thus, the two notions of rank coincide. This, combined with the above argument, shows rank defined in the case of integral domains is unique as well.

We define the torsion submodule of \( M \) to be

\[
\text{Tor}_R(M) = \{ x \in M : rx = 0 \text{ for some } r \in R - 0 \}\]
We say $M$ is a torsion module if

$$\operatorname{Tor}_R(M) = M$$

and $M$ is said to be torsion free if

$$\operatorname{Tor}_R(M) = 0.$$

Let $N \subseteq M$ be a submodule. The annihilator of $N$ is defined by

$$\operatorname{Ann}_R(N) = \{ r \in R : r \cdot n = 0 \quad \forall n \in N \}.$$

This is clearly an ideal in $R$.

**Exercise:**
1) Show $\operatorname{Ann}_R(N) = 0$ if $\operatorname{Tor}_R(N) = N$.
2) If $L \subseteq N$ are submodules, show $\operatorname{Ann}_R(N) = \operatorname{Ann}_R(L)$.

**Example:**
1) Let $V$ be a vector space. Then $\operatorname{Tor}_R(V) = 0$.
2) Let $M = \mathbb{Z}/n\mathbb{Z} \oplus \mathbb{Z}^r$ as a $\mathbb{Z}$-module. Then

$$\operatorname{Tor}_R(M) \cong \mathbb{Z}/n\mathbb{Z}.$$

Let $N_1 = (\mathbb{Z}/n\mathbb{Z}) \oplus 0$ and $N_2 = 0 \oplus \mathbb{Z}^r$. Then

$$\operatorname{Ann}_R(N_1) = n\mathbb{Z},$$

$$\operatorname{Ann}_R(N_2) = 0.$$
of a free module is bounded above by the rank of the module.

One must exercise care here, though, as a given \( R \)-module \( M \) of rank \( n \) need not have a basis, i.e., it may not be free. In fact, even if \( \text{Tor}_1^R(M) = 0 \) and \( \text{rank}_R(M) = n \), \( M \) need not be a free module.

**Example:** Let \( R = \mathbb{Z}[x] \) and \( M = (2, x) \). We have seen before that \( M \) is not a principal ideal, so as a module \( M \) cannot be free of rank 1. Let \( f(x), g(x) \in (2, x) \). Then we can write

\[
0 = g(x) \cdot f(x) + (-f(x))g(x)
\]

so the elements \( f(x) \) and \( g(x) \) cannot be linearly independent over \( R \). Thus, \( \text{rank}_R(M) = 1 \). Thus, \( M \) has rank 1, is clearly torsion free, and is not a free module of rank 1.

In the case where \( M \) is free of finite rank over a P.I.D., and \( N \) is a submodule of \( M \), we have the following main result.

**Theorem 8.12:** Let \( R \) be a P.I.D. and \( M \) a free \( R \)-module of finite rank \( n \), \( N \subseteq M \) a submodule. Then
1) \( N \) is free of rank \( m \) \( \in \mathbb{N} \).

2) there is a basis \( y_1, \ldots, y_n \) of \( M \) and matrices \( a_1, \ldots, a_n \in R \) with \( a_1 y_1 + \cdots + a_n y_n = 0 \) so that \( a, y_1, \ldots, a, y_n \) is a basis of \( N \).

\textbf{Proof:} The result holds trivially if \( N = 0 \), so assume \( N \neq 0 \).

Let \( \psi \in \text{Hom}_R(M, R) \). Then \( \psi(N) \) is a submodule of \( R \), i.e., it is an ideal of \( R \). We are assuming that \( R \) is a PID, so \( \psi(N) = (a_\psi) \) for some \( a_\psi \in R \). Let

\[ \Sigma = \{ a_\psi : \psi \in \text{Hom}_R(M, R) \} . \]

We have \( \Sigma \neq \emptyset \) because \( (0) \in \Sigma \) via the map sending everything in \( M \) to \( 0 \). Thus, we have a nonempty collection of ideals, so \( \Sigma \) contains a maximal element, i.e., there is a homomorphism \( \psi \) so that \( \psi(N) = (a_\psi) \) is not properly contained in any other element in \( \Sigma \). Let \( a = a_\psi \) and let \( y \in N \) be the element so that \( \psi(y) = a \).

We claim \( a \neq 0 \). Let \( x_1, \ldots, x_n \) be a basis of \( M \) and define \( \pi : \text{Hom}_R(M, R) \to R \) via

\[ \pi(\psi)(x_1, \ldots, x_n) = a \psi . \]

Since \( N \neq 0 \), there is an \( i \) so that \( \pi(\psi)(N) \neq 0 \). Thus, using that \( \{ a \} \) is maximal in \( \Sigma \), we see \( a \neq 0 \).

Our next goal is to show \( (a, \psi(y)) \) for every \( \psi \in \text{Hom}_R(M, R) \). Let \( (d) = (a_1, \psi(y)) \). Then \( d \mid a_1 \) and \( d \mid \psi(y) \) in \( R \) and \( f \in R \) so that
\[ d = r_1 \xi_1 + r_2 \phi(y). \]

Consider \( \eta = r_1 \psi + r_2 \phi \in \text{Hom}_R(M, R). \) Then \( \eta(y) = r_1 \xi_1 + r_2 \phi(y) = d. \)

Thus, \( d + \eta(N) \) and so \( (d) \subseteq \eta(N). \) However, \( d \) is a divisor of \( a_1, \) so \( (a_1) \subseteq (d) \subseteq \eta(N). \) Since \( (a_1) \) is maximal, we must have \( (a_1) = \eta(N) \) and so \( a_1 \mid d. \) Thus, we must have \( a_1 \mid \phi(y) \) as claimed since \( a_1, d \mid \psi(y). \)

We can apply the maps \( \pi_i \) to see that \( a_1 \mid \pi_i(y) \) for all \( i. \) Write \( \pi_i(y) = a_1 b_i \) for some \( b_i \in R, \) \( 1 \leq i \leq n. \)

Let
\[ y_i = \sum_{i=1}^{n} b_i \xi_i. \]

We have \( a_1 y_i = y_i. \) Observe that \( \psi(y_i) = \psi(a_1 y_i) = a_1 \psi(y_i). \)

Since \( a_1 \neq 0 \) and \( R \) is an integral domain we have \( \psi(y_i) = 1. \)

Our next step is to show that \( y_i \) can be taken as a basis element for \( M \) and \( a_1 y_i \) can be taken as a basis element for \( N. \) This is equivalent to checking that

1) \( M = R y_i \otimes \ker \psi \)
2) \( N = R a_1 y_i \otimes (\ker \psi) M. \)

We begin with 1), i.e. \( x \in M \) and write \( x = \psi(x) y_i + (x - \psi(x) y_i). \)

We have
\[ \psi(x - \psi(x) y_i) = \psi(x) - \psi(x) \psi(y_i) = \psi(x) - \psi(x) \cdot 1 = 0. \]
As \( x - \psi(x)y_1 \in \ker \psi \). Thus \( M = Ry_1 + \ker \psi \). Suppose \( ry_1 \in \ker \psi \). Then \( \psi(ry_1) = \psi(ry_1) = rz = r \) and \( \psi(ry_1) = 0 \), so we have \( r = 0 \). Thus, the sum is actually direct. This gives 1).

We now prove 2). Since \( a_1 \) is a generator for \( \psi(N) \), we have \( a_1 | \psi(x') \) for all \( x' \in N \). Write \( \psi(x') = ba_1 \) for \( b \in R \). Then as above we have

\[
\psi(x') = \psi(x')y_1 + (x' - \psi(x')y_1) = ba_1y_1 + (x' - ba_1y_1)
\]

where \( x' = ba_1y_1 + \ker \psi \cap N \). Thus, \( N = Ra_1y_1 + (N \cap \ker \psi) \).

The fact that the sum is direct follows because the sum in the proof of 1) is direct and this is a special case of that.

We can now prove the fact that \( N \) is free of rank \( m \).

By induction on the rank of \( N \). If \( m = 0 \) then \( N \) is a torsion module. However, free modules are torsion free and since \( M \) is free, \( \text{Tor}_1(N) = 0 \) as \( N = 0 \). Now assume \( m > 0 \). We have that

\[
N = Ra_1y_1 + (N \cap \ker \psi)
\]

so the rank of \( N \cap \ker \psi \) is \( m - 1 \). Thus, we apply induction to conclude that \( N \cap \ker \psi \) is free of rank \( m - 1 \).

Thus, if we adjoin \( a_1y_1 \) to any basis of \( N \cap \ker \psi \), we have a basis of \( N \) as \( N \) is free of rank \( M \).

It remains to prove 2) in the statement of the
thenem. We proceed by induction on $n = \text{rank}_R(M)$. Apply the previous paragraph to $ker \psi$ to see this submodule is free and rank $n-1$ since

$$M = Ra_0 \oplus ker \psi.$$ 

We now apply the induction hypothesis to $ker \psi$ and its submodule $ker \psi \cap N$, we have a basis $y_1, \ldots, y_n$ of $ker \psi$ and $a_1, \ldots, a_m \in R$ with $a_1y_1, \ldots, a_my_m$ so that $a_1y_1, \ldots, a_my_m$ is a basis of $ker \psi \cap N$. We may use that the sums are direct to get that $y_1, \ldots, y_n$ is a basis of $M$ and $a_1y_1, \ldots, a_my_m$ is a basis of $N$.

It only remains to show $a_1a_2$. Define $\phi : \text{Hom}_R(M, R)$ by $\phi(y_1) = \phi(y_2) = 1$ and $\phi(y_i) = 0$ for $i > 2$. Then $\phi(a_1y_1) = \phi(a_2y_1)$ so $a_1 \in \phi(N)$; thus $(a_1) \leq (N)$. But $(a_1)$ is maximal in $\Sigma$ so $(a_1) = (N)$. However, $a_2 = \phi(a_1y_1) \in \phi(N)$, so $a_2 \in (a_1)$, i.e., $a_1a_2$ and we have the result.

We would also like to give a matrix version of this result as it will help illuminate the result as well as make it easier to do computations with. Before doing this we set aside a couple...
of important results that come directly from the proof we just finished.

**Corol. 8.13:** Let \( M \) be an \( R \)-module, \( F \) a free \( R \)-module of finite rank, and let \( \phi \in \text{Hom}_R(M, F) \) be surjective. Then \( M \) has a submodule \( F' \cong F \) so that
\[
M = F' \oplus \ker \phi.
\]

**Proof:** Exercise. This follows from any given in the proof of Thm 8.12. Note we only proved
there it was assumed that \( R \) was a PID, but the same any.
gives the result in general.

**Corol. 8.14:** Let \( R \) be a PID and \( F \) a free \( R \)-module of finite rank \( n \). Then every submodule of \( F \) is free of rank \( \leq n \).

Let \( \phi \in \text{Hom}_R(F, F) \). Then for each \( i \), we can write
\[
\phi(x_i) = \sum_{j=1}^{n} a_{ij} x_j
\]
for some \( a_{ij} \in R \). Thus, \( \phi \) determines a matrix
\[
A_\phi = (a_{ij}) \in \text{Mat}_n(R).
\]
Conversely, given a matrix \( A \in \text{Mat}_n(\mathbb{R}) \), we obtain a \( \Phi_A \in \text{Hom}_R(F,F) \) by setting
\[
\Phi_A(x_j) = \sum_{i=1}^n a_{ij} x_i
\]
for \( A = (a_{ij}) \). The fact that \( F \) is free tells us \( \Phi_A \) is uniquely determined by \( A \). Thus, this correspondence between \( \text{Hom}_R(F,F) \) and \( \text{Mat}_n(\mathbb{R}) \) is bijective. In fact, one can easily check that this map is a ring isomorphism if we define addition on \( \text{Hom}_R(F,F) \) as \( (\Phi + \Psi)(a) = \Phi(a) + \Psi(a) \) and multiplication as composition. The isomorphism is non-canonical as it depends on the choice of basis for \( F \).

If we restrict our attention to \( \text{Aut}_R(F) \), the \( R \)-module automorphisms of \( F \), then we have an isomorphism between \( \text{Aut}_R(F) \) and \( \text{GL}_n(\mathbb{R}) \). We say an element in \( \text{GL}_n(\mathbb{R}) \) is invertible or non-singular.

Suppose we have a set \( \{y_1, \ldots, y_n\} \) of elements of \( F \). We can write
\[
y_j = \sum_{i=1}^n a_{ij} x_i
\]
for some \( A = (a_{ij}) \in \text{Mat}_n(\mathbb{R}) \). We have an associated map \( \Phi_A \in \text{Hom}_R(F,F) \).
Lemma 8.15: The following are equivalent:

1) \( \mathbf{y}_1, \ldots, \mathbf{y}_n \) is a basis of \( F \)

2) \( \mathbf{y}_i \in \text{Ann}_R(\mathbf{F}) \)

3) \( \mathbf{A} \in \text{GL}_n(\mathbb{R}) \).

Proof: We have already seen (1) \( \Rightarrow \) (3).

Suppose \( \mathbf{A} \) is invertible. Since \( \mathbf{y}_i \in \text{Ann}_R(\mathbf{F}) \), there exists \( \mathbf{c} \in \mathbb{R}^n \) such that \( \mathbf{A} \mathbf{c} = \mathbf{0} \).

Write \( \mathbf{x} = \mathbf{r}_1 \mathbf{x}_1 + \cdots + \mathbf{r}_n \mathbf{x}_n \). Then

\[
\mathbf{y}_i = \mathbf{A}^{-1} \mathbf{y}_i = \mathbf{r}_1 \mathbf{A}^{-1} \mathbf{x}_1 + \cdots + \mathbf{r}_n \mathbf{A}^{-1} \mathbf{x}_n
\]

Thus, \( \mathbf{y}_1, \ldots, \mathbf{y}_n \) spans \( F \). We have uniqueness because if

\[
\mathbf{y} = \mathbf{r}_1 \mathbf{y}_1 + \cdots + \mathbf{r}_n \mathbf{y}_n = \mathbf{r}_1 \mathbf{y}_1 + \cdots + \mathbf{r}_n \mathbf{y}_n,
\]

then

\[
\mathbf{x} = \mathbf{A}^{-1} \mathbf{y} = \mathbf{r}_1 \mathbf{x}_1 + \cdots + \mathbf{r}_n \mathbf{x}_n = \mathbf{r}_1 \mathbf{x}_1 + \cdots + \mathbf{r}_n \mathbf{x}_n,
\]

so \( \mathbf{r} = \mathbf{s} \).

(1) \( \Rightarrow \) (3). If \( \mathbf{y}_1, \ldots, \mathbf{y}_n \) is a basis, we can write

\[
\mathbf{x} = \sum_{j=1}^n b_j \mathbf{y}_j
\]

for some \( b_j \in \mathbb{R} \). Then \( B = (b_{ij}) \in \text{Mat}_n(\mathbb{R}) \) and

\[
AB = BA = I,
\]

so \( B \in \text{GL}_n(\mathbb{R}) \).

As in linear algebra over a field, it may be convenient at times to choose a new basis. Thus, if we write \( \mathbf{F} \in \text{Hom}_R(F,F) \)

as a matrix, we will want to be able to relate the matrix to the one we would get for a different choice of basis. As above,
let $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ be bases of $F$. Let $\phi \in \text{End}_F(F)$ be defined so that $\phi(x_i) = y_i$. Let $\psi \in \text{End}_F(F)$. Then we want to relate the matrix of $\psi$ with respect to $x_1, \ldots, x_n$ to the matrix of $\psi$ with respect to $y_1, \ldots, y_n$. Write

$$\psi(y_j) = \sum_{i=1}^n b_{ij} y_i$$

and $B = (b_{ij})$. Then

$$\psi(\phi(x_j)) = \sum_{i=1}^n b_{ij} \phi(x_i).$$

Thus,

$$\phi^{-1} \psi \phi(x_j) = \sum_{i=1}^n b_{ij} x_i,$$

so $B = A^{-1} \psi A$, i.e., it is the matrix of $\phi^{-1} \psi \phi$ with respect to $x_1, \ldots, x_n$. In particular, if

$$\phi(x_j) = \sum_{i=1}^n \alpha_{ij} x_i, \quad A = (\alpha_{ij})$$

and

$$\psi(x_j) = \sum_{i=1}^n \beta_{ij} x_i, \quad C = (\beta_{ij})$$

then

$$B = A^{-1} C A.$$

One can define determinants the same as for a matrix over a field. The same algebraic properties hold here.
Now return to the case that $F$ is a free module and $\mathbf{x} = \langle x_1, \ldots, x_n \rangle$ is a basis of $F$ and $N$ a submodule of $F$ with basis $\mathcal{N} = \{y_1, \ldots, y_m\}$.

We can write

$$y_j = \sum_{i=1}^{n} a_{ij} x_i$$

for some $a_{ij} \in R$. Let $A$ be the matrix $(a_{ij}) \in \text{Mat}_{m \times n}(R)$, as it is the matrix associated to the choice of bases of $F$ and $N$. Now what happens when we choose different bases of $F$ and $N'$?

Let $F' = \langle x'_1, \ldots, x'_n \rangle$ and $N' = \{y'_1, \ldots, y'_m\}$.

Write

$$x'_j = \sum_{i=1}^{n} s_{ij} x_i$$

and

$$y'_j = \sum_{i=1}^{m} t_{ij} y_i.$$ 

where $S = (s_{ij})$ and $T = (t_{ij})$ are in $\text{GL}_n(R)$ and $\text{GL}_m(R)$ respectively. We can express the $x'_i$'s in terms of the $x_i$'s via the matrix $S^{-1}$: set $S^{-1} = (\tilde{s}_{ij})$, then

$$\sum_{i=1}^{n} \tilde{s}_{ij} x'_i = \sum_{i=1}^{n} \sum_{k=1}^{n} s_{ij} t_{ki} x_k$$

$$= \sum_{k=1}^{n} s_{kj} x_k$$

$$= x_j$$

where we have used that $\tilde{s}_{ij} s_{kj} = k_{ij}$, entry in the identity matrix.
Thus, we can write

\[ y_j' = \sum_{i=1}^{m} t_{ij} y_i = \sum_{i=1}^{m} t_{ij} \sum_{k=1}^{n} a_{ki} x_k \]

\[ = \sum_{i=1}^{m} \sum_{k=1}^{n} t_{ij} a_{ki} x_k \]

\[ = \sum_{i=1}^{m} \sum_{k=1}^{n} t_{ij} a_{ki} \sum_{l=1}^{n} \delta_{lk} x_{l'} \]

\[ = \sum_{i=1}^{m} \sum_{k=1}^{n} \delta_{ik} a_{ki} t_{ij} x_{l'} \]

\[ = \sum (S^{-1}AT)_{ij} x_{l'} . \]

Thus, the matrix of \( N' \) with respect to \( \phi' \) is given by

\[ A' = S^{-1}AT. \]

As by a suitable change of basis in \( F \) and \( N \) we can replace the matrix \( A \) by any matrix \( SAT \) for \( S \in \text{GL}_n(\mathbb{R}) \), \( T \in \text{GL}_m(\mathbb{R}) \). We say \( A \) is equivalent to \( B \) if there are matrices in \( \text{GL}_n(\mathbb{R}) \) and \( \text{GL}_m(\mathbb{R}) \) so that \( A = SBT \).

Another way we can rewrite the Theorem 8.12 in terms of matrices is by the same double-count.

**Theorem:** Let \( A \in \text{Mat}_{n \times n}(\mathbb{R}) \) where \( R \) is a PID. Then \( A \) is equivalent to a matrix

\[ \text{diag}(d_1, d_2, \ldots, d_n) \]

with \( d_1 \leq d_2 \leq \cdots \leq d_n \) where \( u = \min\{n, m\} \).
We have just seen how to go from Theorem 8.12 to this theorem. Now suppose we have this theorem. Let \( F \) and \( N \) be bases of \( F \) and \( N \) as above. Let \( A \) be the matrix giving \( N \) in terms of \( F \). Then the theorem gives \( S \in \text{GL}_n(\mathbb{R}), T \in \text{GL}_m(\mathbb{R}) \) so that

\[
S^{-1}AT = \text{diag}(d_1, \ldots, d_u)
\]

with \( d_1, d_2, \ldots, d_u \). The matrices \( S \) and \( T \) determine new basis \( F' \) and \( N' \). The matrix \( A' \) of \( N' \) with respect to \( F' \) is

\[
\text{diag}(d_1, d_2, \ldots, d_u).
\]

Thus,

\[
y'_1 = d_1 x'_1, \ldots, y'_u = d_u x'_u.
\]

Now set \( d_{u+1} = \ldots = d_n = 0 \) we get Theorem 8.12.

We can now sketch an alternate proof of Theorem 8.12 by proving this theorem. We won't be super rigorous, but it will be helpful for computational purposes. First we recall elementary row and column operations on a matrix. Define

\[
F_{ij} = \text{matrix that results from interchanging the } \ i^{th} \text{ and } j^{th} \text{ rows of the identity matrix.}
\]

\[
G_i(u) = \text{diag}(1, \ldots, 1, u, 1, \ldots, 1)
\]

with its place.

\[
H_{ij}(n) = \left( h_{ij} \right) \quad \text{when } h_{ii} = 1, h_{ij} = 0, \quad h_{ii,x} = 0 \ldots \text{etc.}
\]

\[
\tilde{H}_{ij}(n) = (h_{ij}).
\]
The first thing to note is these are all elements in $GL_n(R)$.

**Lemma 8.16:** The effect of multiplying a matrix on the left is:

1. $F_{ij} A = \text{interchange row } i \text{ and } j \text{ of } A$
2. $G_{i \cdot} A = \text{multiply row } i \text{ by } i \text{ in } A$
3. $H_{ij}(r) A = \text{add } r \text{ times row } j \text{ in } A \text{ to row } i$

The effect of multiplying on the right is:

4. $A F_{ij} = \text{interchange column } i \text{ and } j$
5. $A G_{\cdot i} = \text{multiply column } i \text{ by } i$
6. $A H_{j i}(r) = \text{add } r \text{ times column } j \text{ to } i$

**Proof:** Exercise. 

Let $R$ be a PID, so a UFD. We define a length function on $R^2$ as follows. Since $R$ is a UFD, given $x \in R$, we can write $x = u p_1 \cdots p_n$ for $u \in R^\times$ and $p_i$ primes.

Define $\lambda : R^2 \to \mathbb{Z}$. $\lambda(x) = r$. We have

$$\lambda(xy) = \lambda(x) + \lambda(y)$$

so this is a group homom. We can now see how to reduce a matrix over $R$ to the form we want. Let $A = (a_{ij})$ be the matrix we are working with. The goal is to transform
it into a matrix of the form
\[
C = \begin{pmatrix}
d_1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & C \end{pmatrix}
\]
(4)

where \(d_1\) divides each entry of \(C\). One then applies the same method to \(C\) starting to get the result. We can assume \(a_{11} \neq 0\) for otherwise we can use elementary row and column operations to put it in the form so \(a_{11} = 0\). If \(a_{11}\) divides all \(a_{i,j}\), then we can transform all the \(a_{i,j}\) to be 0 by appropriate column operations.

Assume then in \(a_{i,j}\) so that \(a_{11} \mid a_{i,j}\). We can assume \(j = 2\) by switching columns. Let \(d = \gcd(a_{11}, a_{22})\) and write
\[
a_{11} = dy_1, \\
a_{12} = dy_2.
\]

Since \(a_{11} \mid a_{12}\), \(y_1\) is not a unit. Thus, \(\lambda(y_1) \geq 1\) and so \(\lambda(d) < \lambda(a_{11})\).

Since \(Ra_{11} + Ra_{12} = R1\), we can write
\[
d = r_1a_{11} + r_2a_{12}
\]
for some \(r_1, r_2 \in \mathbb{R}\). Then
\[
d = r_1dy_1 + r_2dy_2,
\]
i.e.,
\[
1 = r_1y_1 + r_2y_2.
\]

Thus, the matrix
\[
\begin{pmatrix}
x_1 & -y_2 & 0 \\
x_2 & y_1 & 0 \\
0 & 0 & 1
\end{pmatrix}
\]
has determinant 1 and so is in $GL_n(\mathbb{R})$. Now consider the matrix $A$, which is equivalent to $A$ and has a (1,1) entry of $x_1 a_{11} + x_2 a_{12} = 1$. Now we repeat the process until $d$ being our new $a_{11}$. We check if $d$ divides the rest of the first row. If it does, we subtract multiples of the first column to eliminate other entries in the first row. If there is a first row entry of does not divide, we repeat the above process. Since $\det(A) > \chi(a_{11})$, this process must eventually terminate. We do the same process with the first column. Thus, we will end up with a matrix of the form

$$
\begin{pmatrix}
1 & 0 & \ldots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \lambda
\end{pmatrix}
$$

after finitely many steps. If $d$ divides all the entries of $D^*$ we have reached our goal. If not, say $d \nmid d_{ij}$, then add the $i$th row to the top row and use the above process. This shows how to get the desired matrix. We now work on example.

**Example:** Let $F$ be a free $\mathbb{Z}$-module with some basis $x_1, x_2, x_3, x_4$ and $N$ a submodule with basis $y_1, y_2, y_3$, which can be written as

$$
y_1 = 2x_1 + 2x_2 + 3x_3 + x_4,$n_2 = 7x_1 + x_2 - x_3,$n_3 = -x_1 - x_2 + 2x_3 + 3x_4.$

Then the matrix expressing $N$ in terms of $F$ is given by
\[
A = \begin{pmatrix}
2 & 7 & -1 \\
2 & 1 & -1 \\
3 & -1 & 2 \\
1 & 0 & 3
\end{pmatrix} \in \text{Mat}_{4 \times 3}(\mathbb{Z}).
\]

Note, in terms of matrices this says \( y_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \), then

\[
Ay_1 = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} =
\]

which, in terms of the basis \( \mathcal{F} \), is \( 2x_1 + 2x_2 + 3x_3 + x_4 \), as desired. As \( A \) is the matrix that write the basis \( N \) in terms of the basis \( \mathcal{F} \).

We now want to write

\[
A' = S^{-1}AS
\]

where \( S \in \text{GL}_4(\mathbb{Z}) \), \( T \in \text{GL}_3(\mathbb{Z}) \) and \( A' \) is \( \text{diag}(d_1, d_2, d_3) \) with \( d_1 \mid d_2 \mid d_3 \). We do this by elementary row and column operations as outlined above.

\[
A \rightarrow \left\{ R_1 \leftrightarrow R_4 \right\} = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & -1 \\ 3 & -1 & 2 \\ 2 & 7 & -1 \end{pmatrix} = F_{41}A
\]

\[
\rightarrow \left\{ -2R_3 + R_2, -3R_2 + R_3, -2R_2 + R_4 \right\} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -7 \\ 0 & 1 & -7 \\ 0 & 1 & -7 \end{pmatrix} = H_{31}(-2)H_{21}(-3)F_{41}A
\]

\[
\rightarrow \left\{ -3c_1 + c_5 \right\} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -7 \\ 0 & 1 & -7 \\ 0 & 1 & -7 \end{pmatrix} = H_{41}(-2)H_{31}(-3)H_{21}(-2)F_{41}AH_{31}(-3) = A^*
\]

Since \( d_1 - 1 \) divides every entry in the submatrix, we may work...
on the automorphs?

\[
\begin{align*}
\{ R_2 + R_3 \} & \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & -7 \\ 0 & 0 & 14 \end{array} \right) = H_{y_2}(-7) H_{y_2}(-14) A^k \\
\{ R_2 + R_4 \} & \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & -7 \\ 0 & 0 & 11 \end{array} \right) = H_{y_2}(-7) H_{y_2}(-11) A^k \\
\{ 3 R_3 + R_4 \} & \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 14 \end{array} \right) = H_{y_2}(-7) H_{y_2}(-11) A^k \\
\{ 7 C_2 + C_3 \} & \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -14 \end{array} \right) = H_{y_2}(-7) H_{y_2}(-11) A^k H_{y_2}(7) \\
\{ G_3(-11) \} & \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 14 \end{array} \right) = G_3(-11) H_{y_2}(-7) H_{y_2}(-11) A^k H_{y_2}(7).
\end{align*}
\]

Thus we have \( d_1 = 1 = d_2, \ d_2 = 14 \) and

\[
S^{-1} = G_3(-11) H_{y_2}(3) H_{y_2}(-7) H_{y_2}(11) H_{y_2}(-2) H_{y_2}(-3) H_{y_2}(-2) F_{14}
\]

\[
= \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right) \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 14 \end{array} \right) \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right) \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right) \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right) \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right) \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right)
\]

\[
= \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right)
\]

\[
S = \left( \begin{array}{ccc} 2 & 7 & 2 & 1 \\ 2 & 1 & 0 & 0 \\ 3 & -1 & -1 & 0 \\ 1 & 0 & 0 & 0 \end{array} \right)
\]

and

\[
T = \left( \begin{array}{ccc} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right) \left( \begin{array}{ccc} 10 & 0 & 0 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{array} \right) = \left( \begin{array}{ccc} 1 & 0 & -3 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{array} \right).
\]

With these matrices we can write down the new bases \( N' \) and \( F' \) in
terms of $\sigma_j$. We know

$$X'_j = \sum_{i=1}^{n} \sigma_{ij} x_i,$$

so

$$X'_1 = \sum_{i=1}^{n} \sigma_{1i} x_i = 2x_1 + 2x_2 + 3x_3 + x_4$$

$$X'_2 = 7x_1 + x_2 - x_3$$

$$X'_3 = 3x_1 - x_3$$

$$X'_4 = x_4$$

and

$$y'_1 = x'_1$$

$$y' = x'_2$$

$$y'_3 = 14x'_3.$$

We can also use these convenient bases to easily see the structure

of the quotient $F/N$. We have

$$F = \mathbb{Z}x'_1 \oplus \mathbb{Z}x'_2 \oplus \mathbb{Z}x'_3 \oplus \mathbb{Z}x'_4$$

$$N = \mathbb{Z}x'_1 \oplus \mathbb{Z}x'_2 + 14\mathbb{Z}x'_3 + 0$$

so

$$F/N \cong \mathbb{Z}/14\mathbb{Z} \oplus \mathbb{Z}.$$
and relations. Note we are only interested in abelian groups here.

Recall a group is given by generators and relations

if it can be written as

\[ G = \langle x_1, \ldots, x_n : \sum_{i=1}^{m} r_{ij} x_i = 0 \text{ for } j = 1, \ldots, m, \ r_{ij} \in \mathbb{Z} \rangle. \]

The elements \( x_1, \ldots, x_n \) are called the generators and the expressions \( \sum_{i=1}^{m} r_{ij} x_i \) are referred to as the relations. We would like to be able to determine the structure of such a \( G \) as a finitely generated abelian group and to be able to compare two groups defined in terms of generators and relations. We accomplish this by realizing \( G \) as a quotient of free modules.

Let \( F = \mathbb{Z}e_1 \oplus \ldots \oplus \mathbb{Z}e_n \) be the free \( \mathbb{Z} \)-module on \( n \) generators and let \( N \) be the submodule defined by

\[ N = \mathbb{Z} \left( \sum_{i=1}^{n} r_{i1} e_i \right) \oplus \ldots \oplus \mathbb{Z} \left( \sum_{i=1}^{n} r_{im} e_i \right). \]

Consider the \( \mathbb{Z} \)-module given by \( F/N \). Define \( \varphi : F \to G \)

\[ e_i \mapsto x_i. \]

This is clearly a surjective \( \mathbb{Z} \)-linear map and the fact that \( G \) is given via the relations \( \sum_{i=1}^{m} r_{ij} x_i = 0 \) for \( j = 1, \ldots, m \) says exactly that \( N = \ker \varphi \). Thus, it is enough to study \( F/N \).
in order to determine the structure of $G$. However, we have just seen how to do this! Let $y = \sum_{i=0}^{n} y_i x_i$ so that $y = \sum_{i=1}^{m} y_i$ is a basis of $N$. Then following the above procedure, there exists $F' = \{ e', \ldots, e_n' \}$ a basis of $F'$ and $a_1, a_2, \ldots, a_m$ so that $N' = \langle d_1, e', \ldots, d_m, e'_m \rangle$ is a basis of $N$. Then

$$F/N \cong \mathbb{Z}^{-m} \otimes \mathbb{Z}/a_1 \mathbb{Z} \otimes \cdots \otimes \mathbb{Z}/a_m \mathbb{Z}.$$ 

**Example:** Let $G = \langle x_1, x_2, x_3, x_4 : 2x_1 + 2x_2 + 3x_3 + x_4 = 0, 7x_1 + x_2 - x_3 = 0, -x_1 - x_2 + 2x_3 + 3x_4 = 0 \rangle$.

Then the corresponding $F$ and $N$ are exactly the ones in the previous example, so

$$G \cong F/N \cong \mathbb{Z} \otimes \mathbb{Z}/m \mathbb{Z}.$$ 

One should note such problems have occurred in the problem, so this gives an easy way to answer such questions.

We now resume our goal of a structure theorem for finitely generated modules over a PID. We begin by establishing existence theorems for the decomposition in an invariant factor and then an elementary divisor form. We then prove uniqueness.
Theorem 8.10: (Fundamental Theorem of Finitely Generated $R$-modules) Let $R$ be a P ID and $M$ a finitely generated $R$-module.

1) Then $M$ is isomorphic to a direct sum of finitely many cyclic modules. More precisely,
$$M \cong R^n \oplus R/(d_1) \oplus R/(d_2) \oplus \cdots \oplus R/(d_m)$$
for some $n \geq 0$ and monic elements $d_1, \ldots, d_m \in R^* - R^*$ that satisfy $d_1|d_2|\cdots|d_m$.

2) $M$ is torsion free iff $M$ is free.

3) If $M$ is a torsion module, then $R^m/(d_1) \oplus \cdots \oplus R^m/(d_m) = 0$ in particular, $M$ is a torsion module iff $r = 0$ and in this case $\text{Ann}_R(M) = (d_m)$.

Proof: This result is essentially what we have just shown. Let $x_1, \ldots, x_n$ be generators for $M$ and let $F$ be a free $R$-module generated by $e_1, \ldots, e_n$. Define
$$\varphi: F \to M$$
$$e_i \mapsto x_i.$$

This is surjective and so the 1st isomorphism theorem gives
$$F/\ker \varphi \cong M.$$

We apply Theorem 8.12 to $F$ and $N = \ker \varphi$ and so
$$M \cong (R^{x_1} \oplus \cdots \oplus R^{x_n})/(Rd_1 x_1 \oplus \cdots \oplus Rd_m x_m)$$
$$\cong R^{n-m} \oplus R/(d_1) \oplus \cdots \oplus R/(d_m).$$
If $d_i$ is a unit then $(d_i) = \mathbb{R}$, so we can remove these terms.

This gives 1).

We clearly have $M$ is torsion free iff $M \cong \mathbb{R}^{\oplus n}$ via $(*)$, so this gives 2). Finally, 3) follows immediately from the definitions.

Recall that since $\mathbb{R}$ is a PID, it is necessarily a UFD as well. Given $d \in \mathbb{R}$, we can write

$$d = u \cdot p_1^{\alpha_1} \cdots p_n^{\alpha_n}$$

for $u \in \mathbb{R}^\times$, $p_i$ distinct primes, and $\alpha_i \geq 0$. This is unique up to units, so the ideals $(p_i^{\alpha_i})$ are uniquely defined. We can apply the Chinese Remainder Theorem to obtain

$$\mathbb{R}/(d) \cong \mathbb{R}/(p_1^{\alpha_1}) \oplus \cdots \oplus \mathbb{R}/(p_n^{\alpha_n})$$

as rings. This certainly gives that they are isomorphic as $\mathbb{R}$-modules. If we apply this to each factor $\mathbb{R}/(d_i)$ in Theorem 8.17, we obtain the elementary division form of the theorem.

**Theorem 8.18:** (Fundamental Theorem existence of elementary divisions) Let $\mathbb{R}$ be a PID and $M$ a finitely generated $\mathbb{R}$-module. Then $\mathbb{R}$ is the direct sum of a finite number of cyclic modules.
whose annihilator is either (0) or generated by powers of primes in \( R \), i.e.,

\[
M \cong R^e \oplus R/(p_1^{k_1}) \oplus \cdots \oplus R/(p_n^{k_n})
\]

where \( e \in \mathbb{Z}_{\geq 0} \) and the \( p_i \) are primes, not necessarily distinct, and \( e \in \mathbb{Z}_{\geq 0} \).

Another version of this theorem is given as follows.

**Theorem 8.19:** Let \( R \) be a PID and \( M \) a non-zero torsion \( R \)-module with \( \text{Ann}_R(M) = (a) \). Write

\[
a = u_1 p_1^{k_1} \cdots p_n^{k_n}
\]

for the prime factorization of \( a \) into distinct prime factors. Let

\[
N_i = \left\{ x \in M : p_i^{k_i} x = 0 \right\}
\]

for \( 1 \leq i \leq n \). Then \( N_i \) is a submodule of \( M \) with \( \text{Ann}_R(N_i) = (p_i^{k_i}) \) and \( N_i = M[p_i^{k_i}] \) when we write \( M[p_i^{k_i}] \) for the set of elements annihilated by some power of \( p_i \). We have

\[
M = N_1 \oplus N_2 \oplus \cdots \oplus N_n.
\]

In particular, we can write

\[
M = \bigoplus_{p \text{ prime}} M[p_i^{k_i}].
\]

**Proof:** If \( M \) is finitely generated this is immediate by reordering
the terms in Theorem 8.18 so all the terms involving $p_i$ are together. If $M$ is not finitely generated, one can still use the Chinese Remainder Theorem as before Theorem 8.18 to get this result. One should write out the details as an exercise.

**Def.** The terms in $M[p^m]$ are referred to as the $p$-primary components of $M$. One should think of them as analogous to the $p$-Sylow subgroups in finite group theory.

Our next goal is to prove the uniqueness part of the decomposition of a finitely generated module over a PID. Namely, we have the following theorem.

**Theorem 8.20 (Fundamental Theorem Uniqueness):** Let $R$ be a PID.

1) Two finitely generated $R$-modules $M_1$ and $M_2$ are isomorphic iff they have the same free rank and the same list of invariant factors.

2) Two finitely generated $R$-modules $M_1$ and $M_2$ are isomorphic iff they have the same free rank and the same elementary divisors.

Before we prove this, we need the following lemma.
Lemma 8.21: Let $R$ be a PID and $p$ a prime in $R$. Let $K$ denote the field $R/(p)$. (Recall primes are maximal in a PID.)

1) Let $M = R^n$. Then $M/pM \cong R^n$.

2) Let $M = R/(d)$ where $d \in R - 0$. Then

$$M/pM \cong \begin{cases} K & \text{if } p \not\mid d \text{ in } R \\ 0 & \text{if } p \mid d \text{ in } R. \end{cases}$$

3) Let $M = R/(d) \oplus \cdots \oplus R/(d_i)$ where $p \mid d_i$ for each $i$.

Then $M/pM \cong K^n$.

Proof: Exercise. 1) was homework. 2) follows from considering the fact that $(p)+(d) = (\gcd(p,d))$, and 3) is 1) and 2) combined.

Proof (Thm. 8.20): It is clear that if $M_1$ and $M_2$ have the same rank and same invariant factors in elementary divisors that they are isomorphic so we only need to prove the other direction.

Suppose $M_1 \cong M_2$. Then if $\phi$ is the isomorphism, we must have $\phi(\text{Tor}_R(M_1)) = \text{Tor}_R(M_2)$, so $\text{Tor}_R(M_1) \cong \text{Tor}_R(M_2)$.

Thus, $M_1/\text{Tor}_R(M_1) \cong M_2/\text{Tor}_R(M_2)$. But we have

$$R^n \cong M_1/\text{Tor}_R(M_1) \cong M_2/\text{Tor}_R(M_2) \cong R^n,$$

when $n = \text{rank}_R(M_i)$. Thus, $R^n \cong R^n$. Let $(p)$ be any maximal ideal in $R$. Then

$$K^n \cong (R/(p))^n \cong (R/(p))^n \cong K^n.$$
and we can use results from linear algebra on fields to conclude that \( n_1 = n_2 \). Thus, if \( M_1 \cong M_2 \) they must have the same rank.

We can now restrict our attention to \( \text{Tor}_R(M_1) \) and \( \text{Tor}_R(M_2) \). Thus, we may as well \( M_1 \) and \( M_2 \) are torsion modules to simplify the notation. Our first step is to show they have the same elementary divisors. In particular, it is enough to show for any fixed prime that the elementary divisors that are a power of \( p \) match up. One can check if \( M_1 \cong M_2 \), then necessarily we have \( M_1[p^{\infty}] \cong M_2[p^{\infty}] \), so it is enough to restrict to the cases that \( M_1 \) and \( M_2 \) are annihilated by a power of \( p \).

Suppose \( \text{Ann}_R(M_1) = (p^n) = \text{Ann}_R(M_2) \). We proceed by induction on \( n \).

1. If \( n = 0 \), then \( M_1 = 0 = M_2 \) and we are done.
2. If \( n > 0 \), then \( M_1 \) and \( M_2 \) necessarily have nontrivial elementary divisors. Let the elementary divisors of \( M_1 \) be

\[
\begin{align*}
&\ p, p, \ldots, p, p^{\alpha_1}, p^{\alpha_2}, \ldots, p^{\alpha_m} \\
&\text{m - coprime}
\end{align*}
\]

with \( 2 \leq \alpha_1 < \alpha_2 < \ldots < \alpha_m \). Thus, \( M_1 \) is the direct sum of cyclic modules with generators \( X_1, \ldots, X_m, X_{m+1}, \ldots, X_{m+s} \) with annihilations \( (p), (p), (p^{\alpha_1}), (p^{\alpha_2}), \ldots, (p^{\alpha_m}) \).

The submodule \( pM_1 \) then has elementary divisors...
\[ p_{r_1}^{a_{r_1}}, \ldots, p_{r_s}^{a_{r_s}}. \]

Similarly, if the elementary divisors of \( M_2 \) are
\[
\underbrace{p_{r_1}, \ldots, p_{r_t}}_{\text{r copies}}
\]
then \( p_{M_2} \) has elementary divisors
\[
p_{r_1}^{a_{r_1}} \cdots p_{r_t}^{a_{r_t}}.
\]

Since \( M_1 \cong M_2 \), we have \( p_{M_1} \cong p_{M_2} \) and \( \Ann_{\mathbb{Z}}(p_{M_1}) = (q^{e-1}) \).

We apply the induction hypothesis to get the elementary divisors of \( p_{M_1} \) are the same as for \( p_{M_2} \), i.e., \( s \leq t \) and \( a_{r_s} = \beta_s \). Now we use that
\[
\frac{M_1}{p_{M_1}} \cong \frac{M_2}{p_{M_2}}
\]
to obtain
\[
K^{m \cdot s} \cong K^{r \cdot t}.
\]

Since \( s \leq t \), this gives \( m \cdot r \) and \( s \cdot M_1 \) and \( M_2 \) have the same elementary divisors.

We now must deal with the invariant factors.

Let \( d_1, d_2, \ldots, d_m \) be invariant factors for \( M_1 \). We get the elementary divisors by taking the prime power factors of these elements. Since \( d_1, d_2, \ldots, d_m \), we have that \( d_m \) is the product of the largest prime power among the elementary divisors, \( d_m-1 \) is the largest of the product
of the largest prime powers of the elementary divisors once the factors of \( dm \) have been removed, and so on.

If \( e_1, e_2, \ldots, e_n \) are the invariant factors for \( M_2 \), they can be characterized in terms of the elementary divisors of \( M_2 \) in the same way. However, since the elementary divisors are, in the same, the invariant factors must be as well and we have the result. \( \square \)

**Addition:** By setting \( R = \mathbb{Z} \) in the results of this chapter we obtain the Fundamental Theorem of Finitely Generated Abelian Groups.

At this point one can easily study Rational canonical forms and Jordan canonical forms of matrices as special cases of these results. While these are very important, we omit them for now due to time constraints.