3/17/16

Welcome back.
HW 8 : TBA.

Before the break we proved the

★ F.T.F.G.M.P.I.D., Part I :

Let R be a PID and let M be a finitely generated R-module of rank k. Then we have

$$M \approx R^{\oplus k} \oplus \mathrm{Tor}_R(M). \quad ///$$

Part II of the FTFGMPID will explain the structure of the torsion submodule $\mathrm{Tor}_R(M) \subseteq M$, and to give the correct proof of this we will transform the problem into "matrix algebra", as follows.

If M is generated by the set $A \to M$ then by definition there exists a canonical surjective homomorphism of R-modules

$$R^{\oplus A} \xrightarrow{\pi} \!\!\!\rightarrow M$$

Now consider $\ker \pi \subseteq R^{\oplus A}$. By the "key lemma for PIDs" we know that $\ker \pi$ is a free R-module of rank $\leq |A|$. That is, we have a basis $B \to \ker \pi$ and a canonical isomorphism

$$R^{\oplus B} \twoheadrightarrow \ker \pi,$$

which we can extend to an injective map

$$R^{\oplus B} \xrightarrow{\varphi} R^{\oplus A}$$

with $\operatorname{im} \varphi = \ker \pi$. In summary, we have obtained a (short) exact sequence

$$(*) \qquad 0 \longrightarrow R^{\oplus B} \xrightarrow{\varphi} R^{\oplus A} \xrightarrow{\pi} M \longrightarrow 0,$$

called a "free resolution of length 1". ///

Remarks:

- One can show (but we won't) that R is a PID if and only if every R-module has a resolution of this form.

This is another good motivation for the PID concept.

- If the short exact sequence (*) splits then M is isomorphic to a submodule of $R^{\oplus A}$, hence M is free. In particular, this implies that $\mathrm{Tor}_R(M) = 0$. Conversely, if $\mathrm{Tor}_R(M) = 0$ and if $|A| < \infty$ then we proved during the FTFGMPID that M is free, and hence (*) splits.

In summary, the splitting of (*) is controlled by $\mathrm{Tor}_R(M)$ and whether M is finitely generated.

- Using other language [see the Midterm Exam Problem 1] we can say that $\pi : R^{\oplus A} \to M$ is the cokernel of $\varphi : R^{\oplus B} \to R^{\oplus A}$. Thus the module M is determined up to isomorphism by the map φ.

///

This leads to a great idea:

★ We can replace the study of the module M by the study of the homomorphism

$$R^{\oplus B} \xrightarrow{\varphi} R^{\oplus A}.$$

For this reason we will now turn our attention to the study of homomorphisms between free modules.

Q: What is a "matrix"?

A: A homomorphism between finite rank free modules, expressed "in coordinates".

Let R be a commutative ring and consider two R-modules M, N. We saw previously that the set

$$\mathrm{Hom}_R(N, M)$$

carries a natural R-module structure defined by

$$(\varphi_1 + r\varphi_2)(n) := \varphi_1(n) + r\varphi_2(n)$$

for all $\varphi_1, \varphi_2 \in \mathrm{Hom}_R(N, M)$, $r \in R$, $n \in N$.
If M & N are both free then we have isomorphisms

$$M \approx R^{\oplus \mathrm{rk}(M)} \quad \& \quad N \approx R^{\oplus \mathrm{rk}(N)},$$

but I want to emphasize that these isomorphisms are not canonical; the choice of specific isomorphisms is equivalent to to choice of specific bases for M & N. Indeed, let

$$A = \{m_a\} \subseteq M \quad \& \quad B = \{n_b\} \subseteq N$$

be bases. Then the corresponding isomorphisms are given by

$$R^{\oplus A} \xrightarrow{\sim} M \quad \& \quad R^{\oplus B} \xrightarrow{\sim} N$$

$$\sum_a r_a i_a \longmapsto \sum_a r_a m_a \qquad \sum_b r_b i_b \longmapsto \sum_b r_b n_b$$

Now consider any R-linear map

$$\varphi : N \longrightarrow M.$$

Composing with the isomorphisms above gives a homomorphism from $R^{\oplus B}$ to $R^{\oplus A}$:

$$R^{\oplus B} \xrightarrow{\sim} N \xrightarrow{\varphi} M \xrightarrow{\sim} R^{\oplus A}$$

$$\sum_b r_b i_b \longmapsto \sum_b r_b n_b \longmapsto \sum_b r_b \varphi(n_b) \longmapsto \quad ?$$

To compute the final step we need to express each $\varphi(n_b) \in M$ in terms of the basis $A = \{m_a\} \subseteq M$. Let's just say

$$\varphi(n_b) = \sum_a r_{ab} m_a .$$

Then the final step of the map is given by

$$\sum_b r_b \varphi(n_b) = \sum_b r_b \left(\sum_a r_{ab} m_a \right)$$

$$= \sum_a \left(\sum_b r_{ab} r_b \right) m_a$$

$$\longmapsto \sum_a \left(\sum_b r_{ab} r_b \right) i_a .$$

and the full map is given by

$$\sum_b r_b i_b \longmapsto \sum_a \left(\sum_b r_{ab} r_b \right) i_a .$$

Now we desperately need a good notation to work with this mess. The notation of "matrices" & "matrix multiplication" was invented by Cayley & Sylvester in the 1850s for just this purpose.

★ **Definition:** Suppose that $|A| < \infty$, $|B| < \infty$. Then we will write

$$A = \{m_1, m_2, \cdots, m_{|A|}\} \quad \& \quad B = \{n_1, n_2, \cdots, n_{|B|}\}.$$

We will express elements of M & N as "column vectors"

$$[\Sigma r_i m_i]_A := \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_{|A|} \end{pmatrix} \quad \& \quad [\Sigma r_j n_j]_B := \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_{|B|} \end{pmatrix}$$

and we will express a homomorphism $\varphi : N \longrightarrow M$ as a "matrix"

$$[\varphi]_{BA} := \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1|B|} \\ r_{21} & r_{22} & & r_{2|B|} \\ \vdots & \vdots & & \vdots \\ r_{|A|1} & r_{|A|2} & \cdots & r_{|A||B|} \end{pmatrix}$$

where the entries $r_{ij} \in R$ are defined by

$$\varphi(n_j) = \sum_i r_{ij} m_i .$$

In other words, the j^{th} column of the matrix $[\varphi]_{BA}$ is given by

$$[\varphi(n_j)]_A = \begin{pmatrix} r_{1j} \\ r_{2j} \\ \vdots \\ r_{|A|j} \end{pmatrix}$$

Now we want to define an operation on matrices so that for all $n \in N$ we have

$$[\varphi(n)]_A = [\varphi]_{BA} \uparrow [n]_B .$$

mystery operation.

The correct definition is forced on us by the requirement of "linearity". If

$$[n]_B = [\sum r_j n_j]_B = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_{|B|} \end{pmatrix}$$

Then we must have

$$[\varphi(n)]_A = \left[\varphi\left(\sum r_j n_j\right)\right]_A$$

$$= \left[\sum r_j \varphi(n_j)\right]_A$$

$$= \sum r_j [\varphi(n_j)]_A$$

↓ "linearity"

and it follows that

$$[\varphi]_{BA}[n]_B := [\varphi(n)]_A$$

$$= \sum_j r_j [\varphi(n_j)]_A$$

$$= \sum_j (j\text{th entry of } [n]_B)(j\text{th col of } [\varphi]_{BA}).$$

[Well, OK. If that's the correct definition then we'll learn to live with it.]

Based on this definition, we will also define the product of two "non-vector" matrices.

∫

Let P be another free R-module with basis $C = \{p_1, p_2, \dots, p_{|C|}\} \subseteq P$ and consider a homomorphism

$$\psi : P \to N$$

Then we will define the matrix $[\varphi]_{BA}[\psi]_{CB}$ such that for all $p \in P$ we have.

$$\left([\varphi]_{BA}[\psi]_{CB}\right)[p]_C = [\varphi]_{BA}\left([\psi]_{CB}[p]_C\right).$$

In other words, we will define

$$[\varphi]_{BA}[\psi]_{CB} := [\varphi \circ \psi]_{CA}.$$

Then to compute this matrix, we note that

$$\begin{aligned}
&(j\text{th col of } [\varphi]_{BA}[\psi]_{CB}) \\
&\quad = (j\text{th col of } [\varphi \circ \psi]_{CA}) \\
&\quad = [\varphi(\psi(p_j))]_A \\
&\quad = [\varphi]_{BA}[\psi(p_j)]_B \\
&\quad = [\varphi]_{BA}(j\text{th col of } [\psi]_{CB}).
\end{aligned}$$

//

Remarks:

- Wait a minute! Wasn't this notation supposed to clean up a mess? It looks like we just made the mess bigger.

Well, be patient. Linear algebra was the "category theory" of its day. At first it looks like a mess of abstract nonsense, until you've internalized the technology. Then you'll see that it's an amazingly efficient language.

- I defined matrices and matrix multiplication in terms of columns. However, the strength of the notation is that we can also think of it in terms of rows.

By convention we think of columns as "vectors" (i.e. lines) and rows as "covectors" (i.e. hyperplanes). The matrix notation balances these concepts in a beautiful way.

3/22/16

HW3 : TBA soon!

Midterm Exam Stats :

Total	=	32
Average	=	24.5
Median	=	26
St. Dev.	=	6.1

Last time I tried to discover the definition of matrix multiplication from scratch. The idea is as follows. Let R be a commutative ring consider two natural numbers $m, n \in \mathbb{N}$. We know that the hom set

$$\mathrm{Hom}_R(R^{\oplus n}, R^{\oplus m})$$

has a natural R-module structure, and if we choose specific bases for $R^{\oplus m}$ & $R^{\oplus n}$ then each $\varphi : R^{\oplus n} \to R^{\oplus m}$ is determined uniquely by $m \times n$ elements of R. How is the structure of $\mathrm{Hom}_R(R^{\oplus n}, R^{\oplus m})$ reflected combinatorially by these $m \times n$ elements ?

After playing around we were led to the following definition.

★ Definition: Let $\mathrm{Mat}_{m\times n}(R)$ denote the set of "$m\times n$ matrices over R"

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} =: (a_{ij}).$$

The operations

$$(a_{ij}) + (b_{ij}) := (a_{ij} + b_{ij}) \quad \& \quad r(a_{ij}) := (ra_{ij})$$

make $\mathrm{Mat}_{m\times n}(R)$ into a free R-module of rank $m\times n$, which is canonically isomorphic to $(R^{\oplus m})^{\oplus n}$ or $(R^{\oplus n})^{\oplus m}$ [see HW 8].

Given $A = (a_{ij}) \in \mathrm{Mat}_{\ell\times m}(R)$ and $B = (b_{ij}) \in \mathrm{Mat}_{m\times n}(R)$, we define the product matrix $AB = (c_{ij}) \in \mathrm{Mat}_{\ell\times n}(R)$ by

$$\boxed{c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}}$$

In particular, we have

$(i,j)^{th}$ entry $AB = (i^{th}\text{ row } A)(j^{th}\text{ col } B)$

i^{th} row $AB = (i^{th}\text{ row } A)\,B$

j^{th} col $AB = A(j^{th}\text{ col } B)$.

One can check that the operation $(A,B) \mapsto AB$ is R-bilinear and associative (when defined), and there is a very good reason for this.

★ Theorem justifying the definition:

Choosing bases for $R^{\oplus m}$ & $R^{\oplus n}$ determines an isomorphism of R-modules

$$\mathrm{Hom}_R(R^{\oplus n}, R^{\oplus m}) \xrightarrow{\sim} \mathrm{Mat}_{m\times n}(R).$$

Furthermore, choosing a basis for $R^{\oplus l}$ determines a commutative square

$$\begin{array}{ccc} \mathrm{Hom}_R(R^{\oplus m}, R^{\oplus l}) \times \mathrm{Hom}_R(R^{\oplus n}, R^{\oplus m}) & \longrightarrow & \mathrm{Mat}_{l\times m}(R) \times \mathrm{Mat}_{m\times n}(R) \\ \downarrow & & \downarrow \\ \mathrm{Hom}_R(R^{\oplus n}, R^{\oplus l}) & \longrightarrow & \mathrm{Mat}_{l\times n}(R), \end{array}$$

where the left arrow is composition of morphisms and the right arrow is multiplication of matrices.

Special Cases:

- Let M be an R-module. Since R is a cyclic module over itself (generated by $1 \in R$) we have an isomorphism

$$\begin{aligned} \mathrm{Hom}_R(R, M) &\xrightarrow{\sim} M \\ \varphi &\longmapsto \varphi(1). \end{aligned}$$

If $M \approx R^{\oplus m}$ is free then any choice of basis determines an isomorphism

$$R^{\oplus m} \approx \mathrm{Hom}_R(R, R^{\oplus m}) \xrightarrow{\sim} \mathrm{Mat}_{m \times 1}(R).$$

Thus we can identify $R^{\oplus m}$ with the module of vectors (i.e. column vectors).

- If M is an R-module then it is not generally true that

$$\mathrm{Hom}_R(M, R) \approx M.$$

[Example: $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2\mathbb{Z}, \mathbb{Z}) = 0 \not\cong \mathbb{Z}/2\mathbb{Z}$.]

However, if $M \approx R^{\oplus m}$ is a free module of finite rank then choosing a basis gives an isomorphism

$$\mathrm{Hom}_R(R^{\oplus m}, R) \xrightarrow{\sim} \mathrm{Mat}_{1\times m}(R) \approx R^{\oplus m}.$$

Here we are thinking of $R^{\oplus m}$ as a module of <u>covectors</u> (i.e. row vectors).

[For a general R-module M we define the <u>dual</u> R-module

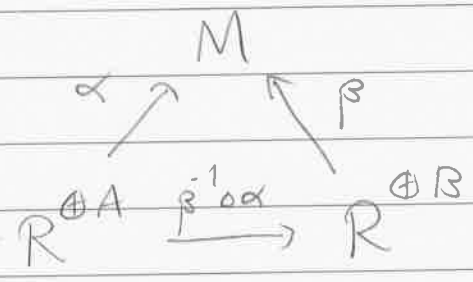
$$M^\vee := \mathrm{Hom}_R(M, R).$$

More on this later.]

Matrix notation is nice because it is concrete and allows us to work with vectors & covectors simultaneously. However, we must pay for this convenience with the fact that <u>coordinates are not canonical</u>.

Change of Basis :

Let M be a free R-module. If $A, B \subseteq M$ are two bases for M then we obtain canonical isomorphisms



The composition $\beta^{-1}\circ\alpha : R^{\oplus A} \longrightarrow R^{\oplus B}$ is called the change-of-basis homomorphism. If $A = \{a_1, \cdots, a_m\}$ & $B = \{b_1, \cdots, b_m\}$ then we can express it as a matrix

$$\left[\beta^{-1}\circ\alpha\right]_{AB} = \left(\left[\beta^{-1}\circ\alpha(a_1)\right]_B \cdots \left[\beta^{-1}\circ\alpha(a_m)\right]_B \right)$$

$$= \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mm} \end{pmatrix}$$

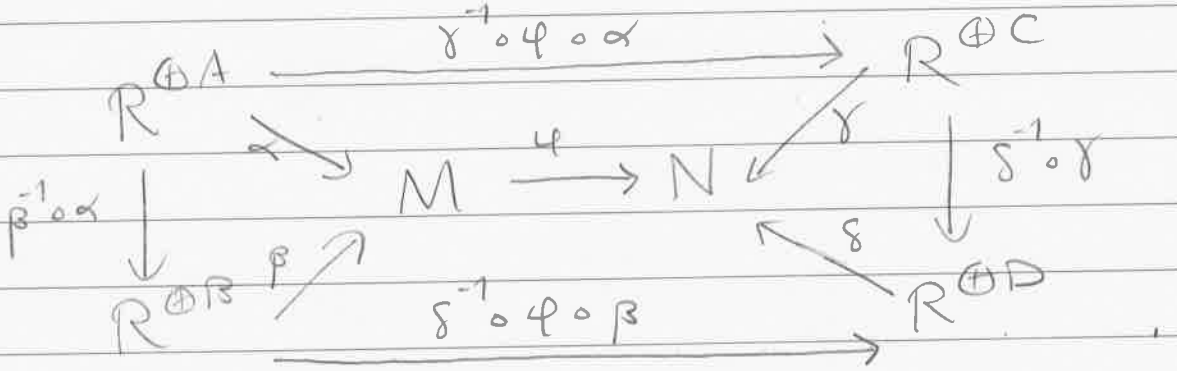
where $\beta^{-1}\circ\alpha(a_j) = \sum_i r_{ij} b_i$.

This matrix is clearly invertible with

$$[\beta^{-1}\circ\alpha]_{AB}^{-1} = [\alpha^{-1}\circ\beta]_{BA}.$$

Now let N be another free R-module with two bases $C, D \subseteq N$, and let $\varphi: M \to N$ be any homomorphism.

We obtain a diagram:



The matrix of φ in the bases A & C is defined by the morphism

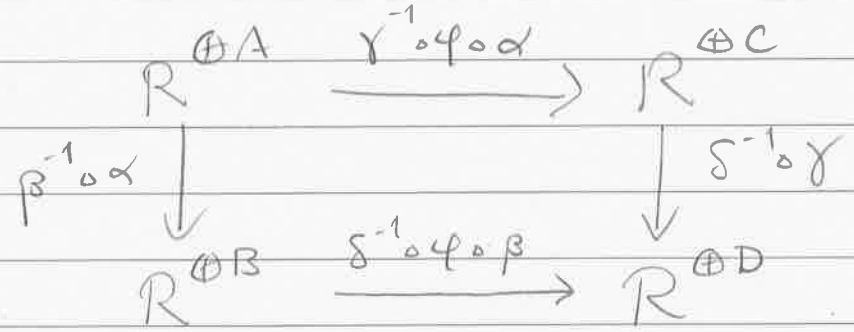
$$[\varphi]_{AC} = [\gamma^{-1}\circ\varphi\circ\alpha]_{AC}$$

and the matrix of φ in B & D is defined by

$$[\varphi]_{BD} = [\delta^{-1}\circ\varphi\circ\beta]_{BD}$$

$\}$

To see how these matrices are related we just erase the middle part of the diagram:



We obtain

$$[\varphi]_{AC} = [\gamma^{-1}\circ\varphi\circ\alpha]_{AC}$$

$$= [(\gamma^{-1}\circ\delta)\circ(\delta^{-1}\circ\varphi\circ\beta)\circ(\beta^{-1}\circ\alpha)]_{AC}$$

$$= [\gamma^{-1}\circ\delta]_{DC}\,[\varphi]_{BD}\,[\beta^{-1}\circ\alpha]_{AB}\,.$$

In summary,

★ We can think of a matrix $Z \in \mathrm{Mat}_{m\times n}(R)$ as a homomorphism $R^{\oplus n} \to R^{\oplus m}$ expressed with respect to some choice of bases for $R^{\oplus m}$ & $R^{\oplus n}$.

Furthermore, two such matrices Z_1 & Z_2 represent the same homomorphism with respect to different bases if and only if there exist invertible matrices $X \in GL_m(R)$ and $Y \in GL_n(R)$ such that

$$Z_1 = X Z_2 Y .$$

//

We can also phrase this in the language of last semester (MTH 761).

We can define an action of the group $GL_m(R) \times GL_n(R)$ on the set $Mat_{m\times n}(R)$ by

$$(A, B) \cdot C := ACB^{-1} .$$

Recall from MTH 761 HW3.4(b) that the corresponding semi-direct product

$$Mat_{m\times n}(R) \rtimes \left(GL_m(R) \times GL_n(R)\right)$$

is isomorphic to the group of invertible block upper-triangular matrices:

$$\left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right) \begin{matrix} \} m \\ \} n \end{matrix}$$

$\underbrace{\quad}_{m} \quad \underbrace{\quad}_{n}$

Our next goal is to classify the orbits of the action

$$\left(GL_m(R) \times GL_n(R)\right) \circlearrowright Mat_{m\times n}(R),$$

i.e., to determine when two $m \times n$ matrices represent the same homomorphism in different coordinates.

One way to do this is to define a "complete normal form" $NF : Mat_{m\times n}(R) \to Mat_{m\times n}(R)$ such that

$NF(C_1) = NF(C_2) \iff C_1$ & C_2 are in the same orbit.

3/24/16

HW 3 due Thurs Apr 7.

FYI: The Final Exam is Wed May 4 at 11:00am – 1:30pm. There will also be a Qual (I don't know when).

Let R be a commutative ring. Last time we defined the R-module of $m \times n$ matrices $\mathrm{Mat}_{m\times n}(R)$ and discussed how "choosing bases" determines an isomorphism

$$\mathrm{Hom}_R(R^{\oplus n}, R^{\oplus m}) \xrightarrow{\sim} \mathrm{Mat}_{m\times n}(R).$$

In the special case that $m = n$ we will write $\mathrm{Mat}_n(R) := \mathrm{Mat}_{n\times n}(R)$. Then from the R-bilinearity of matrix multiplication we obtain an isomorphism of R-algebras

$$\mathrm{End}_R(R^{\oplus n}) \xrightarrow{\sim} \mathrm{Mat}_n(R).$$

What?

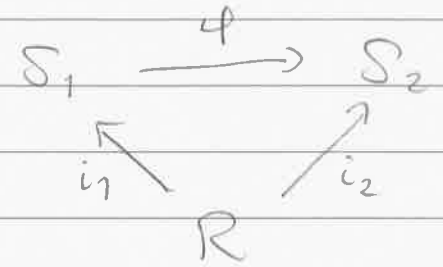
☆ Definition: Let R be a commutative ring. We define an R-algebra as a pair (i, S) where

{

- S is an arbitrary ring.
- $i : R \to S$ is a ring homomorphism such that

$$\text{im } i \subseteq Z(S) := \{ a \in S : ab = ba \ \forall \ b \in S \}$$

We define a morphism of R-algebras $\varphi : (i_1, S_1) \to (i_2, S_2)$ as a ring homomorphism $\varphi : S_1 \to S_2$ satisfying



We'll call the resulting category R-Alg. ///

Remark: Let $i : R \to S$ be an R-algebra and for all $r \in R$ define a function

$$\lambda_r : S \to S \quad \text{by} \quad \lambda_r(s) := i(r) \circ s .$$

You checked on HW2.5(b) that this defines a ring homomorphism $\lambda : R \to \text{End}_{Ab}(S)$ making S into an R-module.

Then we can say that the monoid structure on S is R-bilinear in the sense that for all $a, b, c \in S$ and $r \in R$ we have

$$(a + \lambda_r(b)) \circ c = a \circ c + \lambda_r(b \circ c) \quad \text{and}$$
$$a \circ (b + \lambda_r(c)) = a \circ b + \lambda_r(a \circ c).$$

This observation leads to an alternate definition:

An (associative) R-algebra is an R-module together with an (associative) R-bilinear operation.

Whatever the definition, the prototype of an R-algebra is $\mathrm{Mat}_n(R)$.

Indeed, we know that matrix multiplication is R-bilinear. To see that $\mathrm{Mat}_n(R)$ satisfies the more abstract definition of R-algebra, consider the ring homomorphism

$$R \longrightarrow \mathrm{Mat}_n(R)$$

$$r \longmapsto \begin{pmatrix} r & & 0 \\ & \ddots & \\ 0 & & r \end{pmatrix} = rI_n.$$

and recall from the end of MTH 761 (specifically, from 12/1/15) that

(*) $Z(Mat_n(R)) = \{ rI_n : r \in R \} \approx R$.

Now let me recall the proof of (*):

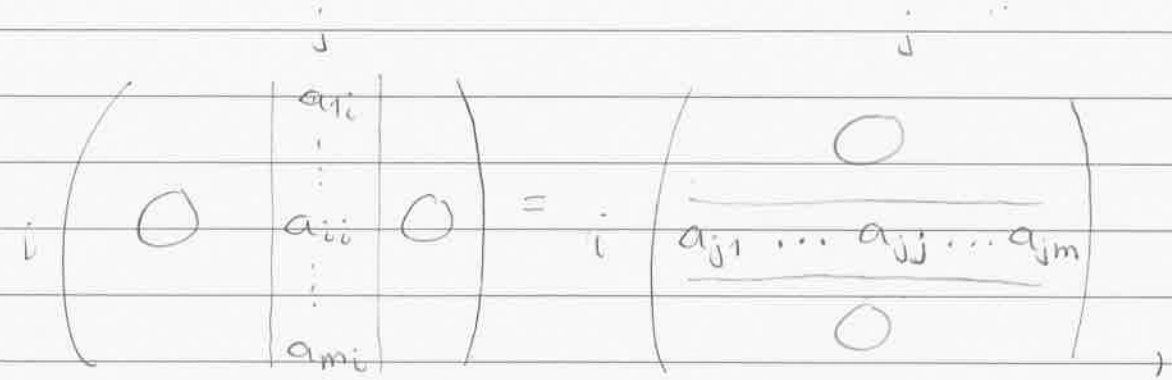
For all $1 \leq i,j \leq n$ we define the matrix unit

$$e_{ij} := \begin{array}{c} \\ i \\ \\ \end{array} \overset{\begin{array}{c} \\ j \\ \\ \end{array}}{\begin{pmatrix} 0 & \vdots & 0 \\ \cdots & 1 & \cdots \\ 0 & \vdots & 0 \end{pmatrix}} \in Mat_n(R).$$

Note that $\{e_{ij}\}$ is a basis for $Mat_n(R)$ as a free R-module. Then consider any matrix $A \in Mat_n(R)$. Since matrix multiplication is R-bilinear we conclude that

$A \in Z(Mat_n(R)) \iff Ae_{ij} = e_{ij}A$ for all $1 \leq i,j \leq n$.

The equation $Ae_{ij} = e_{ij}A$ can be written as



which implies that

- $a_{ii} = a_{jj}$
- $a_{kl} = 0$ for $k \neq i$ and $l = i$
- $a_{kl} = 0$ for $k = j$ and $l \neq j$.

Thus if $Ae_{ij} = e_{ij}A$ holds for all $1 \leq i, j \leq n$ we conclude that $A = rI_n$ for some $r \in R$, as desired. //

Actually, I lied; this is not the result that I proved on 12/1/15. To see that result we need two more definitions.

★ Definition: Let R be a commutative ring. Since $\mathrm{Mat}_n(R)$ is a ring we can consider its group of units

$$GL_n(R) := \mathrm{Mat}_n(R)^\times \approx \mathrm{Aut}_R(R^{\oplus n})$$

called the general linear group.

★ Definition: For all $1 \leq i, j \leq n$ we define the elementary matrices $\in \mathrm{Mat}_n(R)$ as follows.

(1) $E_{ij}(r) := I_n + r e_{ij}$ for $i \neq j$

(2) $E_{ii}(r) := I_n + (r-1) e_{ii}$

(3) $P_{ij} := I_n + e_{ij} + e_{ji} - e_{ii} - e_{jj}$ for $i \neq j$.

These matrices act on the left/right of $\mathrm{Mat}_n(R)$ by elementary row/column operations.

(1) Replace i^{th} row/column by itself plus r times the j^{th} row/column

(2) Multiply i^{th} row/column by r.

(3) Swap i^{th} & j^{th} rows/columns. //

Then here's the result from 12/1/15.

★ Theorem: $Z(GL_n(R)) \approx R^\times$.

Proof: Suppose that $A \in Z(GL_n(R))$. Then since $E_{ij}(1) \in GL_n(R)$ we must have

$$\begin{aligned} A E_{ij}(1) &= E_{ij}(1) A \\ A(I_n + e_{ij}) &= (I_n + e_{ij}) A \\ A + Ae_{ij} &= A + e_{ij} A \\ Ae_{ij} &= e_{ij} A . \end{aligned}$$

Since this is true for all $1 \le i, j \le n$ we conclude as before that $A = r I_n$ for some $r \in R$. Then since $A \in GL_n(R)$ we must have $r \in R^\times$. //

For $i \neq j$ we have $P_{ij}, E_{ij}(r) \in GL_n(R)$ because $(P_{ij})^{-1} = P_{ji}$ & $E_{ij}(r)^{-1} = E_{ij}(-r)$.

If $r \in R^\times$ then we also have $E_{ii}(r) \in GL_n(R)$ because $E_{ii}(r)^{-1} = E_{ii}(r^{-1})$.

↓

It is interesting to investigate the subgroups of $GL_n(R)$ generated by certain kinds of elementary matrices.

Examples:

- The subgroup $W := \langle P_{ij} \rangle \leq GL_n(R)$, called a Weyl subgroup, is isomorphic to the symmetric group S_n.

- If R is a field then we know from Gaussian elimination that every invertible matrix is a product of elementary matrices.

- We will see that this is also true when R is a Euclidean domain, but it fails in general. For example, let $R = \mathbb{Z}[\theta]$ where $\theta = (1+\sqrt{-19})/2$. Then the matrix

$$\begin{pmatrix} 3-\theta & 2+\theta \\ -3-2\theta & 5-2\theta \end{pmatrix}$$

is invertible but it is not a product of elementary matrices. [Remark: This $\mathbb{Z}[\theta]$ is a PID but it is not Euclidean.]

- If R is a Euclidean domain then one can also show that the special linear group $SL_n(R) := \ker(\det : GL_n(R) \longrightarrow R^\times)$ is generated by the matrices

$$E_{ij}(r) \quad \text{for} \quad i \neq j,$$

sometimes called "transvections" or "shear mappings". Again, this is not true in general.

//

==

Next time we will use elementary row & column operations (and, more generally, "quasi-elementary" operations) to reduce a matrix over a PID to diagonal form.

This will lead to the FTFGMPID, Part II.

3/29/16

HW 3 due Thurs Apr 7
[several small typos corrected]

Let R be a commutative ring. Last time we defined the elementary $m \times m$ matrices as follows:

(1) $E_{ij}(r) = I_m + re_{ij}$, $r \in R$, $i \neq j$

(2) $E_{ii}(r) = I_m + (r-1)e_{ii}$, $r \in R$

(3) $P_{ij} = I_m + e_{ij} + e_{ji} - e_{ii} - e_{jj}$, $i \neq j$.

These matrices act on the left of $Mat_{m \times n}(R)$ by row operations and act on the right of $Mat_{l \times m}(R)$ by column operations.

If we restrict ourselves to invertible operations (i.e. we only use $E_{ii}(r)$ when $r \in R^\times$) then we can define group actions. More generally, we define an R-linear action of the group $GL_m(R) \times GL_n(R)$ on the R-module $Mat_{m \times n}(R)$ by

$$(g,h)\cdot A := gAh^{-1}.$$

[I say this action is "more general" because the group $GL(R)$ is not necessarily generated by elementary matrices. It will be when R is a Euclidean domain.]

We have also seen that the induced action

$$(GL_m(R) \times GL_n(R)) \curvearrowright \mathrm{Hom}_R(R^{\oplus n}, R^{\oplus m})$$

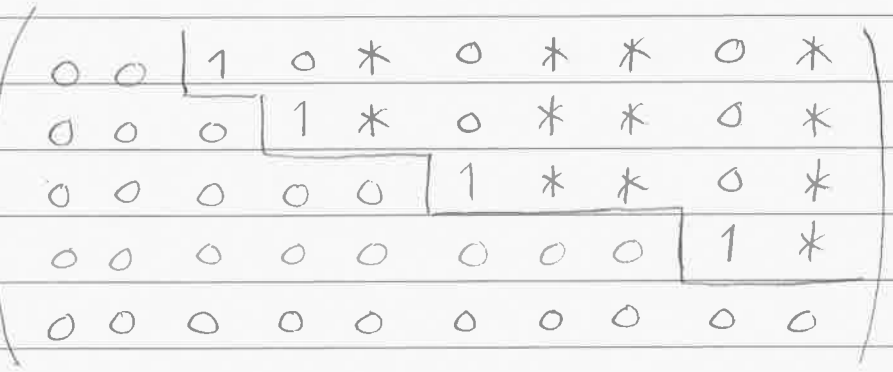
corresponds to "change of basis" on the modules $R^{\oplus m}$ & $R^{\oplus n}$. We are thus interested in classinging the orbits of this action. I'm sure you're familiar with the following special case.

✡ Theorem on RREF:

Let K be a field and consider the action

$$GL_m(K) \curvearrowright \mathrm{Mat}_{m\times n}(K), \quad g\cdot A = gA$$

By applying invertible row operations we can bring any matrix into reduced row echelon form (RREF):



One can show that this defines a function

$$RREF : Mat_{m\times n}(R) \longrightarrow Mat_{m\times n}(R)$$

and that we have

$$RREF(A) = RREF(B) \iff \exists\, g \in GL_m(R) \text{ with } gA = B.$$

As a corollary of this, if we consider the action $GL_m(K) \circlearrowright Mat_{m\times m}(K)$ then for all $g \in GL_m(K) \subseteq Mat_{m\times m}(K)$ we must have

$$RREF(g) = RREF(g^{-1}g) = I_m.$$

This means that there exist elementary matrices $E_1, E_2, \cdots, E_k$ such that

$$E_k E_{k-1} \cdots E_2 E_1 g = I_m.$$

Then inverting gives

$$g = E_1^{-1} E_2^{-1} \cdots E_{k-1}^{-1} E_k^{-1}$$

and we conclude that every $g \in GL_m(K)$ is a product of elementary matrices. ///

In summary, the RREF is a "complete invariant" for the action $GL_m(K) \curvearrowright Mat_{m\times n}(K)$.

Remark: This version of RREF was described independently by Wilhelm Jordan & B.-I. Clasen in 1887. The algorithm of "row reduction" is attributed to Gauss who invented it in order to solve the problem of "least-squares approximation" (and to find the missing dwarf planet Ceres in 1801). However the method already appeared in China (~300 BC) in the "Nine Chapters on the Mathematical Art". Actually the Chinese were interested in systems of linear "Diophantine" equations so they probably had something like the following result.

☆ Theorem on HNF (Hermite, 1851):

Consider the action

$$GL_m(\mathbb{Z}) \circlearrowright Mat_{m\times n}(\mathbb{Z}), \quad g \circ A = gA.$$

By applying invertible row operations (i.e. we only consider $E_{ij}(r)$ for $r \in \mathbb{Z}^\times = \{\pm 1\}$) we can bring any matrix into Hermite normal form (HNF). For example

$$\begin{pmatrix} 3 & 1 & 4 & 2 \\ 1 & 1 & 0 & 1 \\ 2 & 0 & 4 & 0 \end{pmatrix} \xrightarrow{\text{HNF}} \begin{pmatrix} \boxed{1} & 1 & 0 & 0 \\ 0 & \boxed{2} & -4 & 0 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}$$

I won't define the HNF in detail but I'll just mention that it is a complete invariant for the action. ///

More generally, we might try to find complete invariants for the two-sided action

$$(GL_m(R) \times GL_n(R)) \circlearrowright Mat_{m\times n}(R).$$

★ **Theorem**: Let K be a field. Then every matrix $A \in Mat_{m\times n}(K)$ is equivalent to a unique matrix of the form

$$\left(\begin{array}{c|c} I_k & 0 \\ \hline 0 & 0 \end{array}\right)$$

where $k \leq \min(m,n)$ is called the rank of the matrix A.

Proof: Let $A = (a_{ij})$. If $A = 0$ then we're done. Otherwise by swapping rows and columns we can assume that $a_{11} \neq 0$. Since K is a field we can multiply the first row (or column) by a_{11}^{-1} and then we can use row and column operations to obtain a matrix of the form

$$\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right).$$

Then by induction we obtain a matrix of the desired form.

To show uniqueness, consider the corresponding linear function

$$\varphi_A : K^{\oplus n} \longrightarrow K^{\oplus m}$$

We have seen that by choosing bases for $R^{\oplus m}$ & $R^{\oplus n}$ we can write

$$[\varphi_A] = \left(\begin{array}{c|c} I_k & 0 \\ \hline 0 & 0 \end{array}\right).$$

But then observe that k is the dimension of the image in $\varphi_A \subseteq K^{\oplus m}$ as a K-vector space, which does not depend on the choice of bases. ///

If we try to do the same for abelian groups ($\mathbb{Z}$-modules) we get something called the Smith Normal Form (SNF), which was described by H. J. S. Smith in 1861. To pave the way let's consider an example.

Example: Suppose we want to study the submodule of $\mathbb{Z}$ generated by the two element set

$$\{24, 62\} \subseteq \mathbb{Z}.$$

Then we have a canonical $\mathbb{Z}$-module homomorphism from the free module $\mathbb{Z}^{\oplus 2}$,

$$\varphi : \mathbb{Z}^{\oplus 2} \longrightarrow \mathbb{Z}$$
$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto 24x + 62y$$

and we want to study the image $\text{im}\varphi \subseteq \mathbb{Z}$. Note that $\mathbb{Z}$ is represented by the matrix

$$[\varphi] = (24 \quad 62)$$

since for all elements $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{Z}^{\oplus 2}$ we have

$$[\varphi]\begin{pmatrix} x \\ y \end{pmatrix} = (24 \quad 62)\begin{pmatrix} x \\ y \end{pmatrix} = 24x + 62y.$$

Now our goal is to find a new basis for $\mathbb{Z}^{\oplus 2}$ in which the matrix $[\varphi]$ is as nice as possible.

And to do this we are allowed to multiply $(24\ 62)$ on the right by 2×2 $\mathbb{Z}$-invertible elementary matrices. We can proceed as follows:

$$(24\ 62)\begin{pmatrix}1 & -2\\ 0 & 1\end{pmatrix} = (24\ 14)$$

$$(24\ 14)\begin{pmatrix}1 & 0\\ -1 & 1\end{pmatrix} = (10\ 14)$$

$$(10\ 14)\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix} = (10\ 4)$$

$$(10\ 4)\begin{pmatrix}1 & 0\\ -2 & 1\end{pmatrix} = (2\ 4)$$

$$(2\ 4)\begin{pmatrix}1 & -2\\ 0 & 1\end{pmatrix} = (2\ 0).$$

Since we're not allowed to divide by 2, this is as simple as it gets.
In summary, we have

$$(24\ 62)\begin{pmatrix}13 & -31\\ -5 & 12\end{pmatrix} = (2\ 0)$$

[Note that we have

$$\det \begin{pmatrix} 13 & -31 \\ -5 & 12 \end{pmatrix} = 12 \cdot 13 - 31 \cdot 5 = 1 \in \{\pm 1\}$$

as expected because this matrix is invertible over $\mathbb{Z}$.]

In summary, if we make the change of basis

$$\mathbb{Z}^{\oplus 2} \longrightarrow \mathbb{Z}^{\oplus 2}$$

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} 13 \\ -5 \end{pmatrix}, \begin{pmatrix} -31 \\ 12 \end{pmatrix}$$

Then our homomorphism $\varphi : \mathbb{Z}^{\oplus 2} \to \mathbb{Z}$ is represented by the matrix $[\varphi] = (2 \ 0)$ and we conclude that

$$\begin{aligned} \operatorname{im} \varphi &= \left\{ [\varphi] \begin{pmatrix} x \\ y \end{pmatrix} : \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{Z}^{\oplus 2} \right\} \\ &= \left\{ (2 \ 0) \begin{pmatrix} x \\ y \end{pmatrix} : \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{Z}^{\oplus 2} \right\} \\ &= \{ 2x : x \in \mathbb{Z} \} \\ &= 2\mathbb{Z} \subseteq \mathbb{Z}. \end{aligned}$$

///

3/31/16

HW 3 due next Thurs Apr 7.
Final: Wed May 4, 11:00 – 1:30.

Last time we considered the subgroup (i.e. $\mathbb{Z}$-submodule) $M \subseteq \mathbb{Z}$ generated by the set $\{24, 62\} \subseteq \mathbb{Z}$. This generating set defines a canonical morphism

$$\varphi : \mathbb{Z}^{\oplus 2} \twoheadrightarrow M$$
$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto 24x + 62y$$

We can extend this to a morphism of free modules

$$\varphi : \mathbb{Z}^{\oplus 2} \longrightarrow \mathbb{Z}$$

which is represented by the matrix

$$[\varphi] = (24 \;\; 62).$$

Then to compute the isomorphism type of $M = \operatorname{im} \varphi$ we tried to simplify the matrix by changing the basis on $\mathbb{Z}^{\oplus 2}$.

After a sequence of column operations (equivalent to the Euclidean algorithm) we found that

$$\begin{pmatrix} 24 & 62 \end{pmatrix} \begin{pmatrix} 13 & -31 \\ -5 & 12 \end{pmatrix} = \begin{pmatrix} 2 & 0 \end{pmatrix}.$$

[Note that $\det \begin{pmatrix} 13 & -31 \\ -5 & 12 \end{pmatrix} = 1 \in \{\pm 1\} = \mathbb{Z}^\times$ as expected because this matrix is invertible over $\mathbb{Z}$.]

In other words by changing the basis

$$\mathbb{Z}^{\oplus 2} \longrightarrow \mathbb{Z}^{\oplus 2}$$

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} 13 \\ -5 \end{pmatrix}, \begin{pmatrix} -31 \\ 12 \end{pmatrix}$$

we find that φ is represented by the matrix $[\varphi] = \begin{pmatrix} 2 & 0 \end{pmatrix}$ and we conclude that

$$M = \operatorname{im} \varphi = \left\{ \begin{pmatrix} 2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} : \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{Z}^{\oplus 2} \right\}$$

$$= \{ 2x : x \in \mathbb{Z} \}$$

$$= 2\mathbb{Z}.$$

We also conclude that

$$\mathbb{Z}/M \approx \mathbb{Z}/2\mathbb{Z},$$

which was not obvious from the original definition of M.

For a slightly larger example, consider the submodule $M \subseteq \mathbb{Z}^{\oplus 2}$ generated by the two vectors

$$\left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\} \subseteq \mathbb{Z}^{\oplus 2}.$$

We want to compute the isomorphism type of the quotient $\mathbb{Z}^{\oplus 2}/M$. To do this we consider the canonical morphism

$$\begin{aligned} \varphi : \mathbb{Z}^{\oplus 2} &\longrightarrow \mathbb{Z}^{\oplus 2} \\ \begin{pmatrix} x \\ y \end{pmatrix} &\longmapsto x\begin{pmatrix} 2 \\ 1 \end{pmatrix} + y\begin{pmatrix} -1 \\ 2 \end{pmatrix} \end{aligned}$$

represented by the matrix $[\varphi] = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$

and then we try to simplify the matrix as much as possible by changing the basis on the domain and codomain:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & -5 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 0 & -5 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -5 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & -5 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}.$$

In summary we have

$$\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$$

$$\left(\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}\right)^{-1}\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}\left(\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}\right) = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}.$$

This tells us that we have performed

The change of basis

$$a, b \longrightarrow a', b'$$
$$\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} \qquad \begin{pmatrix}2\\1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}$$

on the codomain and the change of basis

$$c, d \longrightarrow c', d'$$
$$\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} \qquad \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}2\\-1\end{pmatrix}$$

on the domain. With respect to these new bases we have

$$\mathbb{Z}^{\oplus 2} \xrightarrow{\varphi} \mathbb{Z}^{\oplus 2}$$

$$\begin{pmatrix}x\\y\end{pmatrix} \longmapsto x\begin{pmatrix}1\\0\end{pmatrix} + y\begin{pmatrix}0\\5\end{pmatrix}$$

and now it becomes clear that

$$M = \operatorname{im}\varphi = \mathbb{Z} \oplus 5\mathbb{Z} \approx \mathbb{Z} \oplus \mathbb{Z} \approx \mathbb{Z}^{\oplus 2}$$

[we say M is a full rank subgroup of $\mathbb{Z}^{\oplus 2}$].

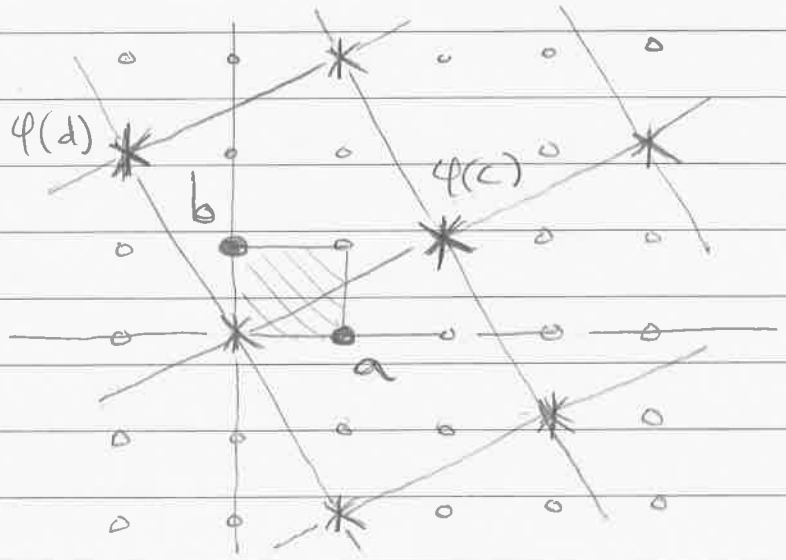
Also, since $[\varphi] = \begin{pmatrix}1&0\\0&5\end{pmatrix}$ is diagonal the quotient $\mathbb{Z}^{\oplus 2}/M$ respects the direct sum in the numerator,

So we obtain

$$\mathbb{Z}^{\oplus 2}/M \approx \left(\frac{\mathbb{Z}}{\mathbb{Z}}\right) \oplus \left(\frac{\mathbb{Z}}{5\mathbb{Z}}\right)$$

$$\approx 0 \oplus \mathbb{Z}/5\mathbb{Z} \approx \mathbb{Z}/5\mathbb{Z}.$$

It's easier to see what this means in terms of pictures. Note that $\mathbb{Z}^{\oplus 2}$ is a free module with basis a, b and that $M = \text{im}\,\varphi$ is a free submodule with basis $\varphi(c), \varphi(d)$ as shown below:



We can interpret our calculation as finding new bases for $\mathbb{Z}^{\oplus 2}$ & M that are more obviously related.

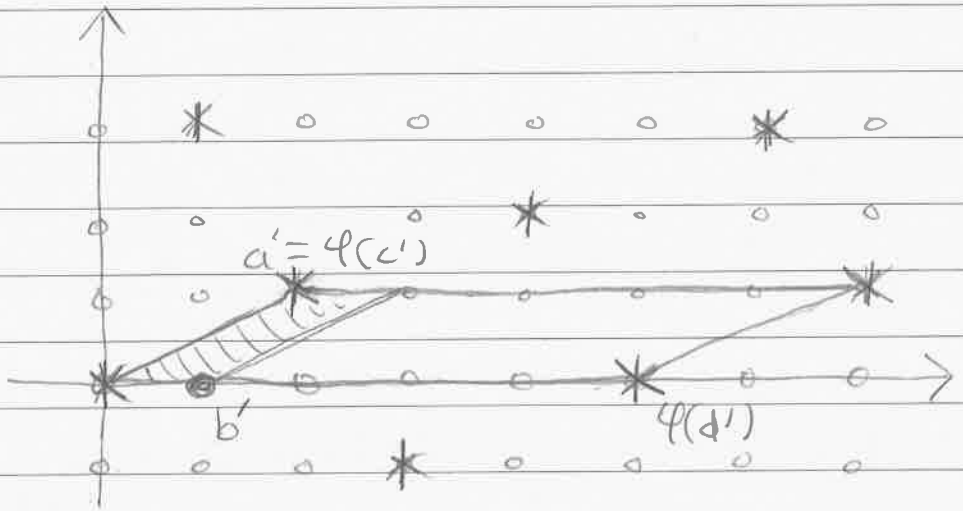
The new basis for $\mathbb{Z}^{\oplus 2}$ is

$$a' = 2a + b \quad , \quad b' = a$$

and the new basis for M is

$$\varphi(c') = \varphi(c) = 2a + b \quad , \quad \varphi(d') = \varphi(2c - d) = 2\varphi(c) - \varphi(d) = 2(2a+b) - (-a+2b) = 5a$$

as shown below:



Now it's obvious why the quotient is isomorphic to $\mathbb{Z}/5\mathbb{Z}$ [because the parallelogram $\varphi(c') \times \varphi(d')$ is 1×5 times the parallelogram $a' \times b'$].

Now I will present the general theorem in this subject. It was first described by Henry John Stephen Smith in 1861.

★ Theorem on Smith Normal Form (SNF):

Let R be a PID. By applying a sequence of invertible row & column operations we can bring any $A \in Mat_{m\times n}(R)$ into diagonal form

$$\left(\begin{array}{c|c} d_1 & \\ & d_2 \\ & \ddots \\ & d_k \\ \hline 0 & 0 \end{array}\right)$$

where $k \leq \min(m,n)$ is the rank of A. Furthermore, we can obtain that

$$d_1 \mid d_2 \mid \cdots \mid d_k$$

and in this case the matrix is a complete invariant for the action

$$\left(GL_m(R) \times GL_n(R)\right) \curvearrowright Mat_{m\times n}(R).$$

Proof: I'll give the proof for Euclidean domains and then mumble some words about how to extend it to PIDs.

So let R be a Euclidean domain with size function $\delta : R \setminus \{0\} \to \mathbb{N}$. Recall the definition: for all $a, b \in R$ with $b \neq 0$ there exist $q, r \in R$ such that

- $a = qb + r$
- $r = 0$ or $\delta(r) < \delta(b)$.

Now consider a matrix $A = (a_{ij}) \in \mathrm{Mat}_{m \times n}(R)$. I'll describe a (possibly very inefficient) algorithm to put A into SNF.

Step 1: If $A = 0$ then stop. Otherwise we can permute rows & columns so that a_{11} has minimal size.

Step 2: Now we try to clear the 1st row & 1st column. If at any point we create an entry smaller than a_{11}, go back to Step 1 (by well-ordering of $\mathbb{N}$ this can only happen finitely many times).

- If there exists $a_{i1} \neq 0$ with $i > 1$ then we divide by a_{11} to obtain

$$a_{i1} = q a_{11} + r$$

Subtract q(row 1) from (row i) to replace the entry a_{i1} by r. If $r = 0$ then we're happy. Otherwise we have $\delta(r) < \delta(a_{11})$ so we go back to Step 1.

- If there exists $a_{1j} \neq 0$ with $j > 1$ we perform a similar procedure to replace a_{1j} with 0, or with an element smaller than a_{11} in which case we go back to Step 1.

At the end of Step 2 we obtain a matrix of the form

$$\left(\begin{array}{c|c} a_{11} & 0 \cdots 0 \\ \hline 0 & \\ \vdots & A' \\ 0 & \end{array}\right)$$

where $a_{11} \neq 0$.

If a_{11} divides every entry of A' then we can apply recursion to obtain the SNF:

$$\left(\begin{array}{c|ccc} a_{11} & 0 & 0 & \cdots & 0 \\ \hline 0 & a_{22} & 0 & \cdots & 0 \\ \hline 0 & 0 & & & \\ \vdots & \vdots & & A'' & \\ 0 & 0 & & & \end{array}\right), \text{ etc.}$$

But <u>does</u> a_{11} divide every entry of A'? We need one more step.

Step 3: If there exists an entry b in A' such that a_{11} does not divide b then we replace (row 1) by itself plus the row containing b:

$$\left(\begin{array}{c|ccccc} a_{11} & * & \cdots & * & b & * \cdots * \\ \hline 0 & & & & & \\ \vdots & & & A' & & \\ 0 & & & & & \end{array}\right)$$

Then we go back to Step 2. Then dividing b by a_{11} produces $b = qa_{11} + r$ with $r \neq 0$, hence $\delta(r) < \delta(a_{11})$.

And now we're back to Step 1!

///

I emphasize that this process will stop because if not then we obtain an infinite subset of $\mathbb{N}$ with no smallest element [i.e. the sizes of the various entries a_{11}].

This completes the proof for Euclidean domains. Now what about PIDs?

There are two issues to consider.

1. Elementary row/column operations are no longer enough. Instead we will define the pseudo-elementary matrices for $i \neq j$:

$$E_{ij}\begin{pmatrix} a & b \\ c & d \end{pmatrix} := I + (a-1)e_{ii} + be_{ij} + ce_{ji} + (d-1)e_{jj}.$$

where $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in Mat_{2\times 2}(R)$ is invertible.

If x_i, x_j are the i^{th}, j^{th} rows of A the i^{th}, j^{th} rows of $E_{ij}\begin{pmatrix} a & b \\ c & d \end{pmatrix} A$ are

$$ax_i + bx_j \;,\; cx_i + dx_j.$$

{

and if y_i, y_j are the i^{th}, j^{th} columns of A then the i^{th}, j^{th} columns of $AE_{ij}\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ are

$$ay_i + cy_j \;,\; by_i + dy_j .$$

Now consider any two elements $x, y \in R$ (not both zero). Since R is a PID we have $(x, y) = (d)$ for some nonzero $d \in R$ and hence there exist $a, b \in R$ such that $ax + by = d$. It follows that

$$\begin{pmatrix} a & b \\ -y & x \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} d \\ 0 \end{pmatrix} \;\&\; (x\ y)\begin{pmatrix} a & -y \\ b & x \end{pmatrix} = (d\ 0)$$

where the matrices $\begin{pmatrix} a & b \\ -y & x \end{pmatrix}$ & $\begin{pmatrix} a & -y \\ b & x \end{pmatrix}$ have determinant $d \neq 0$ and hence are invertible.

In this way we can perform all of the necessary reductions in the algorithm by pseudo-elementary operations.

2. To guarantee that the new algorithm will stop we can no longer rely on the well-ordering of $\mathbb{N}$.

Instead we want to use the fact that a PID is "Noetherian" [i.e. has no infinite increasing chain of ideals. We already used this property when proving that PID $\Rightarrow$ UFD].

At each step of the algorithm when the "size" of a_{11} decreased we can instead arrange to replace a_{11} by a "proper divisor" [indeed if a_{11} divides an element b then we can get rid of b the old-fashioned way, and otherwise we replace a_{11} by d where

$$(a_{11}) \subsetneq (a_{11}, b) = (d).\ ]$$

If the algorithm doesn't terminate then we will obtain an infinite chain

$$(a_{11}) \subsetneq (a_{11}') \subsetneq (a_{11}'') \subsetneq \cdots,$$

which is impossible for the usual reason [see HW 2.1(c)].

QED.

Remark:

You might ask, is there a more general kind of ring that has Smith Normal Form? The answer is yes: they are called "elementary divisor rings" and they are basically defined by the property of having SNF.

It has been my experience that in ring theory there is a bijection between theorems and definitions. That is, to the question

"What is the most general kind of ring that satisfies this theorem?"

the answer is always

"We don't know, so let's give that kind of ring a special name."

11

4/5/16

HW 3 due Thurs.
HW 2 stats:

Total	=	25
Average	=	20.2
Median	=	19.5
St. Dev.	=	3.4

Last time I proved the existence of SNF over a Euclidean domain and sketched how to extend the proof to PIDs. In summary,

☆ Let R be a PID. Then for any matrix $A \in Mat_{m\times n}(R)$ there exist invertible matrices $P \in GL_m(R)$ & $Q \in GL_n(R)$ such that

$$PAQ^{-1} = \left(\begin{array}{ccc|c} d_1 & & & \\ & d_2 & & \\ & & \ddots & 0 \\ & & d_k & \\ \hline & 0 & & 0 \end{array}\right)$$

for some nonzero elements

$$d_1 \mid d_2 \mid d_3 \mid \cdots \mid d_k.$$

I also claimed that this matrix is unique, hence it defines a complete invariant for the action

$$\left(GL_m(R) \times GL_n(R)\right) \curvearrowright Mat_{m\times n}(R).$$

We will call it the Smith Normal form of A.

Before proving the uniquess of SNF, let's examine its implications for the structure theory of modules over a PID.

==

Let R be a PID and consider a finitely generated R-module M. By definition this means that there exists a finite set A and a canonical surjective morphism from the free module $R^{\oplus A}$:

$$R^{\oplus A} \xrightarrow{\pi} \!\!\!\rightarrow M.$$

Now consider the submodule $\ker \pi \subseteq R^{\oplus A}$. Since $R^{\oplus A}$ is free we know from the "Key Lemma for PIDs" that $\ker \pi$ is a free module of rank $\leq |A|$.

In particular, there exists a set B such that $|B| \leq |A|$ and a canonical bijective homomorphism

$$R^{\oplus B} \xrightarrow{\varphi} \ker \pi$$

Then by extending the codomain of φ we obtain a short exact sequence of R-modules

$$0 \longrightarrow R^{\oplus B} \xrightarrow{\varphi} R^{\oplus A} \xrightarrow{\pi} M \longrightarrow 0$$

Note, in particular, that

$$M = \operatorname{im} \pi \approx R^{\oplus A} / \ker \pi = R^{\oplus A} / \operatorname{im} \varphi$$

so that the isomorphism type of M is completely determined by the homomorphism

$$R^{\oplus B} \xrightarrow{\varphi} R^{\oplus A},$$

which can be thought of in coordinates as an $|A| \times |B|$ matrix. Now the idea is to change coodinates on $R^{\oplus A}$ & $R^{\oplus B}$ so that the matrix of φ becomes as simple as possible.

This is exactly what the SNF does for us.

Suppose that $|A| = m$ & $|B| = n \leq m$. Then according to the theorem, there exist new bases

$$A' = \{a_1, \dots, a_m\} \subseteq R^{\oplus A}$$

$$B' = \{b_1, \dots, b_n\} \subseteq R^{\oplus B}$$

such that

$$[\varphi]_{B'A'} = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \\ \hline & 0 & \end{pmatrix} \begin{matrix} \} n \\ \\ \\ \\ \} m-n \end{matrix}$$

$$\underbrace{\qquad\qquad}_{n}$$

In other words, we have

$$\varphi(b_i) = d_i a_i \quad \text{for all } 1 \leq i \leq n.$$

What does this tell us about the module M?

First it tells us that the image of φ is

$$\begin{aligned}\operatorname{im}\varphi &= \left\{ \varphi\left(\sum_{j=1}^{n} r_j b_j\right) : r_j \in R \right\} \\ &= \left\{ \sum_{j=1}^{n} r_j \varphi(b_j) : r_j \in R \right\} \\ &= \left\{ \sum_{j=1}^{n} r_j d_j a_j : r_j \in R \right\}.\end{aligned}$$

Now I claim that

$$M = \frac{R^{\oplus A}}{\operatorname{im}\varphi} \approx \frac{R}{(d_1)} \oplus \cdots \oplus \frac{R}{(d_n)} \oplus R^{\oplus(m-n)}.$$

Indeed, we have an obvious homomorphism

$$\overline{\Phi} : R^{\oplus A} \twoheadrightarrow \frac{R}{(d_1)} \oplus \cdots \oplus \frac{R}{(d_n)} \oplus R^{\oplus(m-n)}$$

defined by

$$\overline{\Phi}\left(\sum_{i=1}^{m} s_i a_i\right) := \left(s_1 + (d_1), \ldots, s_n + (d_n), s_{n+1}, \ldots, s_m\right).$$

Finally, note that $\ker \overline{\Phi} = \operatorname{im}\varphi$. Indeed, we have

$$\overline{\Phi}\left(\sum_{i=1}^{m} s_i a_i\right) = \left(0 + (d_1), \ldots, 0 + (d_n), 0, \ldots, 0\right)$$

if and only if

$s_{n+1} = \cdots = s_m = 0$ and $s_j \in (d_j)$ for $1 \le j \le n$

i.e., if and only if

$$\sum_{i=1}^{m} s_i a_i = \sum_{j=1}^{n} r_{j,j} a_j$$

for some elements $r_j \in R$. Then the 1st Isomorphism Theorem gives

$$M \approx R^{\oplus A} / \text{im}\,\varphi$$

$$= R^{\oplus A} / \ker \overline{\Phi}$$

$$\approx \text{im}\,\overline{\Phi}$$

$$= R/(d_1) \oplus \cdots \oplus R/(d_n) \oplus R^{\oplus(m-n)}.$$

//

This (finally) completes the proof of the Fundamental Theorem of Finite Abelian Groups (from MTH 761, 12/3/15), and allows me to state the rest of the F.T.F.G.M.P.I.D..

✡ F.T.F.G.M.P.I.D., Part II :

Let R be a PID and let M be a finitely generated R-module of rank k. Then there exist nonzero elements $d_1, \dots, d_n \in R$ such that

$$d_1 \mid d_2 \mid \cdots \mid d_n$$

and such that

$$M \cong R^{\oplus k} \oplus \frac{R}{(d_1)} \oplus \cdots \oplus \frac{R}{(d_n)}$$

Moreover, these elements $d_1, \dots, d_n$ are unique up to associates so we can give them a special name: we call them the invariant factors of M.

[Comparing with Part I we see that

$$\mathrm{Tor}_R(M) \cong R/(d_1) \oplus \cdots \oplus R/(d_n),$$

so Part II is really a structure theorem for torsion modules over a PID.]

The existence is done; it remains only to show uniqueness. This will also complete the proof of uniqueness for the Smith Normal Form [since the two theorems are basically equivalent].

The proof goes in 4 steps.

Step 1: Since we already know that rank is well-defined [see HW 2] we can restrict our attention to the torsion submodule $T := \mathrm{Tor}_R(M)$.

Step 2: Suppose that

$$T = R/(d_1) \oplus \cdots \oplus R/(d_n)$$

where $(d_n) \subseteq (d_{n-1}) \subseteq \cdots \subseteq (d_1)$. We want to show that this chain of ideals only depends on the isomorphism type of T and not on any specific choice of generating set. To this end we define the annihilator ideal of T:

$$\mathrm{Ann}_R(T) := \{ r \in R : rt = 0 \ \forall t \in T \}.$$

and then prove that

$$\mathrm{Ann}_R(T) = (d_n).$$

Indeed, suppose that $r \in \mathrm{Ann}_R(T)$. Then, in particular, we must have

$$\begin{aligned}(0+(d_1), \cdots, 0+(d_n)) &= r(1+(d_1), \cdots, 1+(d_n)) \\ &= (r+(d_1), \cdots, r+(d_n))\end{aligned}$$

and then $r+(d_n) = 0+(d_n) \Longrightarrow r \in (d_n)$. Conversely, suppose that $r \in (d_n)$. Then since $(d_n) \subseteq \cdots \subseteq (d_1)$ we must have $r \in (d_i)$ for all i, and it follows that for all $(r_1+(d_1), \cdots, r_n+(d_n)) \in T$ we have

$$r(r_1+(d_1), \cdots, r_n+(d_n)) = (0+(d_1), \cdots, 0+(d_n)),$$

as desired. ///

We have shown that the ideal (d_n) depends only on the isomorphism type of T. What about the ideals

$$(d_{n-1}) \subseteq \cdots \subseteq (d_1) \quad ?$$

Step 3 : We know that every PID is also a UFD. So there exists a unique factorization

$$d_n = p_1^{\alpha} p_2^{a_{2n}} \cdots p_m^{a_{mn}}$$

where $p_1, \cdots, p_m$ are some distinct prime elements of R. Since $d_1 | d_2 | \cdots | d_n$ we must also have

$$d_j = p_1^{\alpha_{1j}} p_2^{\alpha_{2j}} \cdots p_m^{\alpha_{mj}}$$

for some $\alpha_{ij} \in \mathbb{N}$ with $1 \leq i \leq m$, $1 \leq j \leq n$ such that

$$\alpha_{i1} \leq \alpha_{i2} \leq \cdots \leq \alpha_{in}$$

for each i. You will show on HW4 that

$$R/(d_j) \approx \bigoplus_i R/(p_i^{\alpha_{ij}})$$

and hence that

$$T \approx \bigoplus_{i,j} R/(p_i^{\alpha_{ij}}).$$

$\S$

[Once we have proved uniquess of the elements $p_i^{\alpha_{ij}}$ they will be called the elementary divisors of T.]

Step 4: We already know that the primes $p_i \in R$ are determined uniquely (up to associates) by the module T. Indeed, we know that the ideal (d_n) is determined uniquely. Then we that observe that (p_i) are precisely the prime ideals containing (d_n). [This is just another way to say "unique prime factorization".]

Thus it remains only to show that the natural numbers

$$\alpha_{ij} \in \mathbb{N}$$

are uniquely determined by T.

You will finish the proof on HW 4 :) .

But first I need to write HW 4 :(.

QED.