The primary text for Algebra II is Dummit and Foote's *Abstract Algebra*, and will cover the following topics:

- modules and advanced linear algebra;
- representation theory of finite groups;
- field theory and Galois theory.

Contents

Modules and Advanced Linear Algebra

Tensor Products of Modules

The first major topic in Modules and Advanced Linear Algebra is tensor products.

Introduction and Basic Definitions

To motivate tensor products, we recall a basic fact from linear algebra. If we assume that R is a field, and M, N are finite-dimensional R-vector spaces, then the following equation necessarily holds:

$$\dim(M \oplus N) = \dim(M) + \dim(N).$$

We want to construct a similar operation on vector spaces, $M \otimes N$, that satisfies

$$\dim(M \otimes N) = \dim(M)\dim(N).$$

For now, we will label this by $M \bar{\otimes} N$, where we use the $\bar{\otimes}$ to refer to the fact that this is a temporary definition. Naively, we might seek to define $M \bar{\otimes} N$ as follows. If we let $\{x_1, \dots, x_k\}$ be a basis for M and $\{y_1, \dots, y_\ell\}$ a basis for N, then we will define $M \bar{\otimes} N$ to be all the formal R-linear combinations over the basis

$$B = \{x_i \otimes y_j \mid 1 \leq i \leq k, 1 \leq j \leq \ell\}.$$

While this is technically correct — as in, this does yield a vector space with

$$\dim(M \bar{\otimes} N) = \dim(M)\dim(N),$$

the issue is that this definition is not canonical, in that it depends on chosen bases for M and N. Furthermore, it is not clear how one may generalize from this definition to modules over arbitrary rings, which do not necessarily have bases. To resolve this issue, we will go about defining a construction that "extends," in a sense, this definition of $M \bar{\otimes} N$.

To start, we define the simple tensor $m \otimes n$ for any $m \in M$ and $n \in N$. If we let

$$m = \sum_{i=1}^{k} \lambda_i x_i$$

$$n = \sum_{j=1}^{\ell} \mu_j y_j,$$

then we will define

$$m \otimes n = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \lambda_i \mu_j (x_i \otimes y_j).$$

We observe that every element of $M \bar{\otimes} N$ is a sum (i.e., an *integral* linear combination) of simple tensors, as by regrouping we may take

$$\sum_{i=1}^{k} \sum_{j=1}^{\ell} \lambda_{ij} (x_i \otimes y_j) = \sum_{i=1}^{k} \sum_{j=1}^{\ell} (\lambda_{ij} x_i) \otimes y_j.$$

The simple tensors satisfy the following relations:

(R1) $(m_1 + m_2) \otimes n = m_1 \otimes n + m_2 \otimes n$;

(R2) $m \otimes (n_1 + n_2) = m \otimes n_1 + m \otimes n_2$;

(R3) $(\alpha m) \otimes n = m \otimes (\alpha n)$

for $m, m_1, m_2 \in M$, $n, n_1, n_2 \in N$, and $\alpha \in R$.

Proposition: These are the defining relations for $M \bar{\otimes} N$ in the category of abelian groups.

We will simply take this proposition as fact.

Now, let

$$\begin{aligned} Q &= M \times N \\ &= \{(m, n) \mid m \in M, n \in N\} \end{aligned}$$

be the Cartesian product of M and N as sets. We will then take $\mathbb{Z}[Q]$ to be the standard free $\mathbb{Z}$-module (i.e., free abelian group) on Q. That is, $\mathbb{Z}[Q]$ is the set of formal linear combinations

$$v = \sum_{q \in Q} \lambda_q q,$$

where $\lambda_q \in \mathbb{Z}$ and only finitely many coefficients are nonzero. By the universal property of free abelian groups, the map $(m, n) \mapsto m \otimes n$ descends to a unique homomorphism $\varphi \colon \mathbb{Z}[Q] \to M \bar{\otimes} N$. Such a homomorphism is necessarily surjective as every element of $M \bar{\otimes} N$ is an integral linear combination of simple tensors, meaning that we have

$$M \bar{\otimes} N \cong \mathbb{Z}[Q] / \ker(\varphi)$$

as abelian groups.

Now, consider the subgroup of $\mathbb{Z}[Q]$, which we denote $\langle K \rangle$, that is generated by the following elements:

(I) $(m_1 + m_2, n) - (m_1, n) - (m_2, n)$;

(II) $(m, n_1 + n_2) - (m, n_1) - (m, n_2)$;

(III) $(\alpha m, n) - (m, \alpha n)$

for $m_1, m_2, m \in M$, $n_1, n_2, n \in N$, and $\alpha \in R$. Then, from proposition that the relations (R1) through (R3) define $M \bar{\otimes} N$, it follows that $\langle K \rangle = \ker(\varphi)$. Thus, we may define the tensor product canonically as follows.

Definition: Letting M, N, Q, K be as above, we define

$$M \otimes N \coloneqq \mathbb{Z}[Q]/\langle K \rangle, \tag{$\dagger$}$$

and define $m \otimes n = (m, n) + K$.

So far, this has only given us an abelian group. We may ask how to define $\mathbb{Z}[Q]/\langle K \rangle$ as an R-vector space, which naturally seems to be defined by

$$r\left(\sum_{i=1}^{n} m_i \otimes n_i\right) = \sum_{i=1}^{n} (rm_i) \otimes n_i \tag{$*$}$$

To show that the right-hand side of $(*)$ is well-defined is a very difficult task. We will not do it here.

Now, we can actually quite easily generalize $(\dagger)$ to modules over non-fields.

- If R is a commutative ring with 1, and M and N are left R-modules, the definition in $(\dagger)$ copies over exactly.
- If R is non-commutative with 1, then the definition in $(\dagger)$ makes sense, but the scalar multiplication in $(*)$ does *not* hold.

  In fact, we need to change the assumptions for M and N as R-modules. In particular, we need M to be a *right* R-module, and N to be a left R-module, and take the generators of type (III) for K to be defined by

  (III') $(mr, n) - (m, rn)$

  for $m \in M$, $n \in N$, and $r \in R$. This gives the tensor product $M \otimes_R N$ an abelian group structure, but does not endow it with a R-module structure.

We may now consider some simple examples computing tensor products.

Example: Let $R = \mathbb{Z}$. We will show that $\mathbb{Z}/n\mathbb{Z} \otimes_R \mathbb{Q} = 0$.

As a general strategy, in order to show that a tensor product is the zero module, it suffices to show for every simple tensor. Observe that $0 \otimes y = 0$ for any tensor product, since we may take

$$\begin{aligned} 0 \otimes y &= (0 + 0) \otimes y \\ &= 0 \otimes y + 0 \otimes y. \end{aligned}$$

Therefore, we may write

$$\begin{aligned} [a] \otimes b &= (n[a]) \otimes \left(\frac{b}{n}\right) \\ &= [na] \otimes \frac{b}{n} \\ &= 0 \otimes \frac{b}{n} \\ &= 0. \end{aligned}$$

Universal Property

We may now work towards understanding one of the defining properties of tensor products in general. This requires a discussion of a weakened version of R-bilinear maps.

Definition: Let R be a ring, M a right R-module, N a left R-module, and L an abelian group written additively. A map $\varphi\colon M \times N \to L$ is called *R-balanced* if

(BM1) $\varphi(m_1 + m_2, n) = \varphi(m_1, n) + \varphi(m_2, n)$

(BM2) $\varphi(m, n_1 + n_2) = \varphi(m, n_1) + \varphi(m, n_2)$

(BM3) $\varphi(mr, n) = \varphi(m, rn)$

for all $r \in R$, $m, m_1, m_2 \in M$, and $n, n_1, n_2 \in N$.

Theorem: Let R, M, N, L be as above. Let

$$\Omega = \{\Phi \colon M \otimes N \to L \mid \Phi \text{ a group homomorphism}\}$$
$$\Delta = \{\varphi \colon M \times N \to L \mid \varphi \ R\text{-balanced}\}.$$

Define the map $J \colon \Omega \to \Delta$ by

$$(J\Phi)(m, n) = \Phi(m \otimes n).$$

Then, J is bijective.

Proof. We have that J is injective since $J\Phi$ captures the value of Φ on simple tensors, and Φ is completely determined by its value on simple tensors since Φ is a group homomorphism, and elements of $M \otimes N$ are sums of simple tensors.

To prove surjectivity, we recall that

$$M \otimes N = \mathbb{Z}[M \times N]/\langle K \rangle.$$

Let $\varphi \colon M \times N \to L$ be an R-balanced map. By the universal property for free modules, there is a homomorphism $\widetilde{\varphi} \colon \mathbb{Z}[M \times N] \to L$ taking $(m, n) \mapsto \varphi(m, n)$.

We only need to show now that $\widetilde{\varphi}$ kills the elements of K that generate $\langle K \rangle$, but this follows from the fact that φ is R-balanced. Therefore, we get an induced map

$$\begin{aligned} \Phi \colon M \otimes N &\to L \\ m \otimes n &\mapsto \varphi(m, n), \end{aligned}$$

so we are done. □

Definition: Let R be a commutative ring, M, N, L left R-modules. A map $\varphi \colon M \times N \to L$ is called R-bilinear if it satisfies (BM1), (BM2), and

(BM3') $\varphi(m, rn) = \varphi(rm, n) = r\varphi(m, n)$

Theorem: If R, M, N, L are as above, then there exists a natural bijection between $\hom_R(M \otimes N, L)$ and $\hom_R(M \times N, L)$.

The proof is the same as the proof in the case of R-balanced maps, mutatis mutandis.

Proposition: Let R be a commutative ring, and M, N free left R-modules with respective bases X and Y. Then, $M \otimes N$ is a free module with basis

$$Z = \{x \otimes y \mid x \in X, y \in Y\}.$$

Proof. We have that Z generates $M \otimes N$ as a R-module, so we only need to show that Z is linearly independent.

Let

$$v = \sum_{i=1}^{t} r_i x_i \otimes y_i.$$

Without loss of generality, we assume that $r_1 \neq 0$. It is enough to find a homomorphism $\varphi \colon M \otimes N \to R$ such that $\varphi(v) \neq 0$.

Toward this end, we construct an R-bilinear map, which we only need to specify on the basis. Define

$$\alpha \colon M \to R$$

$$x_i \mapsto \begin{cases} 0 & x_i \neq x_1 \\ 1 & x_i = x_1 \end{cases}$$

$$\beta \colon N \to R$$

$$y_i \mapsto \begin{cases} 0 & y_i \neq y_1 \\ 1 & y_i = y_1 \end{cases}.$$

The map $\varphi \colon M \times N \to R$ given by $\varphi(x_i, y_i) = \alpha(x_i)\beta(y_i)$ is thus R-bilinear and induces a map on the tensor product that is nonzero at v. Thus, v is not the zero vector. □

Module Structure on Tensor Products

Thus far, we have only shown that there is a module structure on the tensor product whenever we are considering modules over commutative rings. Else, we only have an abelian group. We will specify the case when there is a module structure on the tensor product.

Definition: Let R and S be unital rings. An (S, R)-bimodule is an abelian group $(M, +)$ that is both a left S-module and right R-module satisfying the compatibility condition

$$(sm)r = s(mr)$$

for all $m \in M$, $r \in R$, and $s \in S$.

Example:

(i) If R and S are both subrings of the same ring T with $1_R = 1_S = 1_T$, then T is an (S, R) bimodule, where S acts by left-multiplication and R acts by right-multiplication.

(ii) If R is commutative, M a left R-module, then M can be considered as an (R, R)-bimodule by setting $m.r = rm$.

Proposition: Let R and S be rings, M an (S, R)-bimodule, and N a right R-module. Then, $M \otimes N$ can be endowed with a unique S-module structure by taking $s(m \otimes n) = sm \otimes n$.

Proof. Fix $s \in S$. Define the map $\varphi_s \colon M \times N \to M \otimes_R N$ by taking $\varphi_s(m, n) = sm \otimes n$.

This map is R-balanced; we will verify (BM3) for this purpose:

$$\begin{aligned} \varphi_s(mr, n) &= s(mr) \otimes n \\ &= (sm)r \otimes n \\ &= sm \otimes rn \\ &= \varphi_s(m, rn). \end{aligned}$$

Thus, by the universal property, there is a homomorphism of abelian groups $\overline{\varphi_s} \colon M \otimes N \to M \otimes N$ such that $\overline{\varphi_s}(m \otimes n) = sm \otimes n$. We will then define the action of s on $M \otimes N$ by taking $s.u = \overline{\varphi_s}(u)$ for any $u \in M \otimes N$. □

The most useful case for this proposition is the *extension of scalars*. If $R \subseteq S$ is a unital subring, then we may view S as an (S, R)-bimodule; for any left R-module N, we have that $S \otimes_R N$ may be endowed with the structure of an S-module. We call this module the extension of scalars of N from R to S.

Definition: Let R be a commutative ring with 1. An R-algebra is a ring A which is also an R-module such that multiplication $\mu \colon A \times A \to A$, $\mu(a, b) = ab$ is an R-bilinear map.

Example (Some R-algebras)**:** The following are R-algebras:

(i) the polynomial ring $R[x_1, \ldots, x_n]$ in commuting variables;

(ii) the ring of noncommutative polynomials $R\langle x_1, \ldots, x_n \rangle$;

(iii) the matrices $\mathrm{Mat}_n(R)$ of $n \times n$ matrices over R.

Theorem: Let A and B be R-algebras, where R is commutative. Then, $A \otimes B$ has a unique R-algebra structure such that

$$(a \otimes b)(c \otimes d) = ac \otimes bd.$$

Proof. See Proposition 1.4.3 in this document. □

Lemma: If $R \subseteq S$ is a unital subring, with S commutative, and M is an R-algebra, then $S \otimes_R M$ is an S-algebra.

Lemma: With the same assumptions as above, we have isomorphisms

$$S \otimes_R \text{Mat}_n(R) \cong \text{Mat}_n(S)$$
$$S \otimes_R R[x_1, \ldots, x_n] \cong S[x_1, \ldots, x_n]$$
$$S \otimes_R R[x]/(f(x)) \cong S[x]/(f(x)).$$

The general proof approach for these types of problems is to use the universal properties to construct an isomorphism of abelian groups, then prove that the isomorphism is compatible with the S-action and products of elements.

Finitely Generated Modules over PIDs

We will now turn our focus towards proving the structure theorem for finitely generated modules over a PID. For this, we will use the Smith Normal Form.

Smith Normal Form

Definition: Let R be a PID, and $A \in \text{Mat}_n(R)$. We say A is in Smith normal form if there exists $m \leq \min(k, n)$ and elements $d_1, \ldots, d_m$ such that

$$A = \begin{pmatrix} d_1 & \cdots & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & d_m & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

where $d_i \neq 0$ for all i and $(d_1) \supseteq (d_2) \supseteq \cdots \supseteq (d_m)$.

Theorem: If R is a PID with $A \in \text{Mat}_n(R)$, then there are $U \in \text{GL}_k(R)$ and $V \in \text{GL}_n(R)$ such that UAV is in Smith normal form.

This decomposition is essentially unique, in that if we have two Smith normal forms

$$D = \begin{pmatrix} d_1' & \cdots & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & d_m & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$D' = \begin{pmatrix} d_1' & \cdots & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & d_s' & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

then $s = m$ and d_i and d_i' are associates for each i.

To show existence, we consider the elementary row (and column) operations, viewed as elementary matrices. Given $\lambda \in R$, we consider $E_{ij}(\lambda)$ with $i < j$ to be a square matrix of size to be determined size that has 1 along the diagonal, λ in position (i, j), and 0 everywhere else. Furthermore, for any $i \neq j$, we let F_{ij} be the matrix obtained from an identity matrix of a certain size by swapping rows i and j.[I] Then, we have the following table of correspondences.

Elementary Row Operation	Interpretation as Matrix Multiplication
$R_i(A) \mapsto R_i(A) + \lambda R_j(A)$	$A \mapsto A_{ij}(\lambda)A$
$C_i(A) \mapsto C_i(A) + \lambda C_j(A)$	$A \mapsto AE_{ji}(\lambda)$
$R_i(A) \leftrightarrow R_j(A)$	$A \mapsto F_{ij}A$
$C_i(A) \leftrightarrow C_j(A)$	$A \mapsto AF_{ij}$

Thus, it suffices to show that every A can be replaced via these elementary row operations to a matrix in Smith normal form.

We will prove the special case where R is a Euclidean domain, since the primary applications we will be using are for the cases of $\mathbb{Z}$ and $F[x]$ for some field F, and both of these rings are Euclidean domains. If N is the Euclidean norm for R, then we will let $N(A) = \min_{a_{ij} \neq 0} N(a_{ij})$ denote the norm of the matrix.

Proof of Existence. By induction, we may assume that existence holds for any matrices with smaller values of $k+n$. Starting with an arbitrary matrix with dimensions $k \times n$, we claim that we can reduce any nonzero matrix to either

(i) a matrix A' with smaller norm

(ii) a matrix of the block diagonal form

$$\begin{pmatrix} d_1 & 0 \\ 0 & B \end{pmatrix}, \tag{$\dagger$}$$

where d_1 divides all the entries of B.

To see that this suffices to prove existence of Smith normal form, we start by seeing that if case 1 occurs, we apply the reduction until case 2 occurs, which necessarily happens by the well-ordering principle. Then, if A is reduced to a matrix of the block diagonal form ($\dagger$), then by the induction hypothesis we may reduce B to Smith normal form using the elementary operations. We may then apply these same operations to A to yield A to have $d_2 | d_3 | \cdots | d_m$ and the block d_1 on the upper left hand corner, so since these elementary row operations yield R-linear combinations of the original element, with d_1 dividing all the entries of B, we have that $d_1 | d_2$.

Now, we will show that this is satisfactory. Using column and row flips, we may assume that $N(A) = N(a_{11})$. Write

$$\begin{aligned} a_{1j} &= a_{11}q_{1j} + r_{1j} \\ a_{i1} &= a_{11}q_{i1} + r_{i1} \end{aligned}$$

for all $2 \leq i, j \leq n$. Then, for any i and j, we have $N(r_{i1}) < N(a_{11})$ and $N(r_{1j}) \leq N(a_{11})$, or we have $N(r_{i1}) = 0$ or $N(r_{1j}) = 0$. By applying the operations $C_j(A) \mapsto C_j(A) - q_{1j}C_1(A)$ and $R_i(A) \mapsto R_i(A) - q_{i1}R_1(A)$, we may perform the replacements $a_{i1} \mapsto r_{i1}$ and $a_{1j} \mapsto r_{1j}$. If there is at least one such nonzero remainder, then we have $N(A') < N(A)$, and this shows case (i).

If all the remainders are 0, then we have the block diagonal form ($\dagger$), and split into two subcases. If we had $a_{11} | b_{ij}$ for all $b_{ij} \in B$, then we would be done. Else, suppose there were some $s, t \geq 2$ with $a_{11} \nmid a_{st}$. In this case, we may first perform the operation $R_1(A') \mapsto R_1(A') + R_s(A')$, write $a_{st} = a_{11}q + r$ with $r \neq 0$, and then perform $C_t(A) = C_t(A) - qC_1(A)$ to find that $N(A')$ is not minimal, so we may perform the process further. This gives existence. □

[I] One may recognize this as the matrix corresponding to the permutation $\tau = (i, j)$.

Structure of Finitely-Generated Modules over PIDs

Theorem (Uniqueness of Rank)**:** Let $R \neq \{0\}$ be a commutative ring. If $R^n \cong R^m$, then $n = m$.

Proof. The theorem holds for the case of R as a field, and otherwise, pass to a quotient by a maximal ideal. □

Corollary: If M is a finitely generated free module over a commutative ring R, then if we define $\text{rk}(M)$ to be the value of $n \in \mathbb{Z}_{\geq 0}$ such that $M \cong R^n$, such a value is then unique.

Theorem (Compatible Basis Theorem)**:** Let R be a PID, and M a finitely generated free module with rank n. Let N be a submodule of M. Then, N is free with rank $m \leq n$, and there is a basis $\{y_1, \ldots, y_n\}$ for M and elements $(d_1) \supseteq \cdots \supseteq (d_m)$ in R such that $\{d_1 y_1, \ldots, d_m y_m\}$ is a basis for N.

Moreover, m is unique, and $d_1, \ldots, d_m$ are unique up to associates.

Proof. Let $\{x_1, \ldots, x_n\}$ be a basis for M, and let N be generated by $\{u_1, \ldots, u_k\}$. We may write

$$u_i = \sum_{j=1}^{n} a_{ij} x_j$$

for some $a_{ij} \in R$. Writing the system of equations as a matrix

$$\begin{pmatrix} u_1 \\ \vdots \\ u_k \end{pmatrix} = (a_{ij})_{i,j} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

we may then put the matrix $A = (a_{ij})_{i,j}$ in Smith normal form, giving $A = PDQ$ for some diagonal matrix $D = \text{diag}_{k,n}(d_1, \ldots, d_m)$, $P \in \text{GL}_k(R)$, and $Q \in \text{GL}_n(R)$. Write

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = Q \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$
$$\begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix} = P^{-1} \begin{pmatrix} u_1 \\ \vdots \\ v_k \end{pmatrix}.$$

Then, $\{y_1, \ldots, y_n\}$ form a basis for M, and $\{v_1, \ldots, v_k\}$ generate N; since we may write

$$\begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix} = D \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

it follows that $k = m$, $v_i = d_i y_i$, and we are done. Linear independence follows from the fact that R is an integral domain and that $\{y_1, \ldots, y_n\}$ is a basis. □

Theorem (Classification: Invariant Factors Form)**:** Let R be a PID, and M an arbitrary finitely generated R-module. Then,

$$M \cong \left(\bigoplus_{i=1}^{\ell} R/(a_i) \right) \oplus R^s,$$

where $\ell, s \in \mathbb{Z}_{\geq 0}$, $a_1, \ldots, a_\ell \in R$ are nonzero non-units, and we have $(a_1) \supseteq (a_2) \supseteq \cdots \supseteq (a_\ell)$.

Moreover, ℓ, s are unique, and $a_1, \ldots, a_\ell$ are unique up to associates.

Proof. We show existence; uniqueness will follow from the discussion after the proof of the classification in elementary divisors form.

Let $\{x_1, \dots, x_n\}$ be generators for M; then, there exists a surjective R-module homomorphism $\pi\colon F = R^n \to M$ given by $(r_1, \dots, r_n) \mapsto \sum_{i=1}^n r_i x_i$. Let $N = \ker(\pi)$, so that $M \cong F/N$.

By the compatible basis theorem, there are $y_1, \dots, y_n \in F$ and $a_1, \dots, a_m \in R$ with $m \leq n$, $(a_1) \supseteq (a_2) \supseteq \cdots \supseteq (a_m) \neq 0$ such that $\{y_1, \dots, y_n\}$ are a basis for F and $\{a_1 y_1, \dots, a_m y_m\}$ are a basis for N.

By definition, we then have

$$F = \bigoplus_{i=1}^n Ry_i$$
$$N = \bigoplus_{i=1}^m (a_i) y_i.$$

Now, since the $(a_i)y_i$ are submodules of Ry_i, we have

$$\begin{aligned} M &\cong F/N \\ &= \bigoplus_{i=1}^n Ry_i \Big/ \bigoplus_{i=1}^m (a_i)y_i \\ &= \bigoplus_{i=1}^m R/(a_i) \oplus R^{n-m-1}. \end{aligned}$$

Moreover, $(a_i) = R$ if a_i is a unit, so we may remove such terms. □

Theorem (Classification, Invariant Factors)**:** Let M be a finitely generated module over a PID. Then, there are primes $p_1, \dots, p_t \in R$ not necessarily distinct, and $d_1, \dots, d_t \in \mathbb{N}$ such that

$$M \cong \left(\bigoplus_{i=1}^t R/\left(p_i^{d_i}\right) \right) \oplus R^s.$$

Moreover, s, t are unique, and $p_1^{d_1}, \dots, p_t^{d_t}$ are unique up to associates and permutations. We call the collection $\left\{p_1^{d_1}, \dots, p_t^{d_t}\right\}$ the elementary divisors for M.

Proof. It suffices to consider the case of $M = R/(a)$, where a is a nonzero non-unit in R. Since PIDs are UFDs, we may then write

$$a = u \prod_{i=1}^t p_i^{d_i},$$

where u is a unit and the p_i are pairwise non-associate primes. Then, we have that $\left(p_i^{d_i}\right)$ and $\left(p_j^{d_j}\right)$ for $i \neq j$ are comaximal ideals, meaning that by the Chinese Remainder Theorem, we may write

$$R/(a) \cong \bigoplus_{i=1}^t R/\left(p_i^{d_i}\right).$$

□

As for uniqueness, we start by showing uniqueness of the free part, which is simpler. Observe that if we let $F = \text{frac}(R)$ be the field of fractions, then

$$F \otimes_R M = F \otimes \left(\left(\bigoplus_{i=1}^t R/\left(p_i^{d_i}\right) \right) \oplus R^s \right)$$

$$\cong F^s,$$

meaning that we may define $s \coloneqq \dim(F \otimes_R M)$. Since dimensions of vector spaces are invariant, it follows that s is well-defined.

Next, we show the uniqueness of the elementary divisor decomposition. We let

$$\begin{aligned} M &\cong M_0 \\ &\coloneqq \left(\bigoplus_{i=1}^{t} R/\left(p_i^{d_i}\right) \right) \oplus R^s \end{aligned}$$

as in the elementary divisor decomposition. We let $\varphi \colon M \to M_0$ be a fixed isomorphism. Given $x, y \in R$, we will write $x \sim y$ whenever x and y are associates. For all primes p and all $d \in \mathbb{N}$, define

$$\begin{aligned} \mu_{p,d}(M_0) &= \left| \left\{ i \;\middle|\; p_i^{d_i} \sim p^d \right\} \right| \\ &= |\{ i \mid d_i = d, p_i \sim p \}|. \end{aligned}$$

We will show that $\mu_{p,d}$ is entirely determined by p, d, M. We will let $W = \{([p], d) \mid p \text{ is a prime}, d \in \mathbb{N}\}$.

Notice that

$$t = \sum_{([p],d) \in W} \mu_{p,d}(M_0).$$

Furthermore, we have $\varphi\left(p^d M\right) = p^d M_0$. We will examine the quotient $p^d M / p^{d+1} M$. Using the isomorphism φ and trundling through much grunt-work, we find that

$$p^d M_0 / p^{d+1} M_0 \cong \left(\bigoplus_{i=1}^{t} p^d \left(R/\left(p_i^{d_i}\right) \right) / p^{d+1} \left(R/\left(p_i^{d_i}\right) \right) \right) \oplus (R/(p))^s.$$

Examining the quotients in the torsion part, we observe that, as a general rule, we have

$$\begin{aligned} z(R/I) &= \{ z(r + I) \mid r \in R \} \\ &= \{ zr + I \mid r \in R \} \\ &= ((z) + I)/I \\ &\coloneqq \{ u + I \mid u \in (z) \}, \end{aligned}$$

so in our case, this yields

$$z(R/(y)) = (\gcd(z, y))/(y).$$

Therefore, returning to the quotient $p^d M_0 / p^{d+1} M_0$, we observe that if p, p_i are not associate, then the corresponding quotient is zero, and if p, p_i are associates, then

$$p^d\left(R/\left(p^{d_i}\right)\right) = \left(p^{\min(d, d_i)}\right)/\left(p^{d_i}\right),$$

meaning that the corresponding quotient is isomorphic to

$$\left(p^{\min(d, d_i)}\right)/\left(p^{\min(d+1, d_i)}\right).$$

If $d \geq d_i$, then the quotient is zero, and if $d < d_i$, then the quotient is $\left(p^d\right)/\left(p^{d+1}\right) \cong R/(p)$. Since prime ideals are maximal in PIDs, we have that

$$p^d M / p^{d+1} M \cong (R/(p))^{g_{p,d}}$$

where

$$\begin{aligned} g_{p,d} &= \dim_{R/(p)}\left(p^d M/p^{d+1}M\right) \\ &= s + |\{i \mid p_i \sim p, d_i > d\}| \\ &= s + \sum_{k \geq d+1} \mu_{p,k}(M_0). \end{aligned}$$

Observe then that for any $d \in \mathbb{N}$, we have

$$\begin{aligned} \mu_{p,d}(M_0) &= g_{p,d-1} - g_{p,d} \\ &= \dim_{R/(p)}\left(p^{d-1}M/p^dM\right) - \dim_{R/(p)}\left(p^dM/p^{d+1}M\right) \end{aligned}$$

Since these quantities are uniquely determined by M, it follows that the elementary divisors decomposition of M is unique.

As for the uniqueness of the invariant factors decomposition, we write

$$M \cong \left(\bigoplus_{i=1}^{\ell} R/(a_i)\right) \oplus R^s,$$

and observe that M has the multiset of elementary divisors given by

$$E = \{\text{prime powers in the prime factorization of the } a_i\}.$$

Replacing primes by associates, we may write E as a collection of $\left\{p_i^{b_{i,j}} \;\middle|\; 1 \leq i \leq h, 1 \leq j \leq n_i\right\}$, where $p_1, \ldots, p_h$ are pairwise non-associates, and $b_{i,1} \geq b_{i,2} \geq \cdots \geq b_{i,n_i}$.

Since E is unique by the uniqueness of the elementary divisors decomposition, and $(a_1) \supseteq (a_2) \supseteq \cdots \supseteq (a_\ell)$, with the a_i non-associate, we have then that

$$\begin{aligned} a_\ell &= \prod_{j=1}^{h} p_j^{b_j,1} \\ a_{\ell-1} &= \prod_{j=1}^{h} p_j^{b_{j,2}} \\ &\vdots \end{aligned}$$

so that the invariant factors are themselves unique.

Invariant Factors and Rational Canonical Form

Now that we've proven the classification of finitely generated modules over PIDs, we will use the special case of $F[x]$-modules to determine the rational canonical form.

Given a field F and a finite-dimensional vector space V over F, with $T\colon V \to V$ an F-linear transformation, we want to find a basis β of V such that $[T]_\beta$ is a simple form. The first one we will construct is the Rational Canonical Form, which exists over any field.

Note that if we let $x \cdot v = Tv$, then we may turn V into an $F[x]$-module. This is the only $F[x]$-module structure on V that extends the original vector space structure with x acting as T.

Since V is finite-dimensional over F, it is finitely generated as an F-module, so it is finitely-generated as an $F[x]$-module. Therefore, we may write

$$V \cong \widetilde{V}$$

$$\coloneqq \bigoplus_{i=1}^{n} \frac{F[x]}{(a_i(x))},$$

with $a_1(x)|a_2(x)|\cdots|a_m(x)$ nonzero non-unit elements. Note that there is no free summand since $F[x]$ itself is infinite-dimensional as a vector space over F, so since V is assumed to be finite-dimensional, it follows that there is no free summand.

We may safely assume that the a_i are monic, so that the a_i are unique. We call them the invariant factors of this linear transformation.

We let $\varphi\colon V \to \widetilde{V}$ be an isomorphism of $F[x]$-modules, which gives rise to the following commutative diagram.

$$\begin{array}{ccc} V & \xrightarrow{T} & V \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\varphi} \\ \widetilde{V} & \xrightarrow[\varphi\circ T\circ\varphi^{-1}]{} & \widetilde{V} \end{array}$$

Note that if β is a basis of V, then $\varphi(\beta)$ is a basis for $\varphi(V)$, and we have $[T]_\beta = \left[\varphi\circ T\circ\varphi^{-1}\right]_{\varphi(\beta)}$, by the definition of the basis.

Note that if $w\in\widetilde{V}$, then the action of x on w is given by

$$\begin{aligned} x\cdot w &= x\cdot\varphi\big(\varphi^{-1}(w)\big) \\ &= \varphi\big(x\cdot\varphi^{-1}(w)\big) \\ &= \varphi\big(T\circ\varphi^{-1}(w)\big) \\ &= \big(\varphi\circ T\circ\varphi^{-1}\big)(w). \end{aligned}$$

Therefore, we may assume that in fact $V=\widetilde{V}$. Writing

$$V_i = F[x]/(a_i(x)),$$

then we have

$$V=\bigoplus_{i=1}^{m} V_i,$$

and if β_i is a basis for V_i, then

$$[T]_\beta = \operatorname{diag}\Big([T|_{V_1}]_{\beta_1},\ldots,[T|_{V_m}]_{\beta_m}\Big).$$

Therefore, it suffices to consider the case of one invariant factor,

$$V=F[x]/(a(x)),$$

where $a(x)$ is a nonconstant monic polynomial of degree $d>0$. Write

$$a(x)=x^d+\sum_{i=0}^{d-1}a_ix^i.$$

Then, for any $p(x)\in F[x]$, we define $\overline{p(x)}=p(x)+(a(x))$, and observe that $\left\{\overline{1},\overline{x},\ldots,\overline{x}^{d-1}\right\}=\beta$ is a basis for V, and we have $T\big(\overline{x}^i\big)=\overline{x}^{i+1}$, and we have the matrix form

$$[T]_\beta = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & \ddots & \cdots & \vdots & -a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & \vdots & -a_{d-2} \\ 0 & 0 & \cdots & 1 & -a_{d-1} \end{pmatrix}$$

This matrix is known as the *companion matrix* of the polynomial $a(x)$, and is denoted $C_{a(x)}$. In the general case, we have the block-diagonal matrix

$$[T]_\beta = \operatorname{diag}\big(C_{a_1(x)}, \dots, C_{a_m(x)}\big).$$

The matrix of this form is known as the *rational canonical form* of the linear transformation T.

Theorem: Let V be a finite-dimensional vector space over F, and let $T\colon V \to V$. Then, there exists a basis β of V and non-constant monic polynomials $a_1, \dots, a_m(x) \in F[x]$ with $a_1(x) | \cdots | a_m(x)$ such that

$$[T]_\beta = \operatorname{diag}\big(C_{a_1(x)}, \dots, C_{a_m(x)}\big)$$

is in block-diagonal form, where $C_{a_i(x)}$ is the companion matrix of $p(x)$.

Moreover, the number m and the polynomials $a_i(x)$ are unique. The polynomials $a_i(x)$ are called the invariant factors of T.

Corollary: Let K/F be a field extension. Suppose $A, B \in \operatorname{Mat}_n(F)$. Then, A and B are F-similar if and only if A and B are K-similar.

Proof. The forward direction is clear.

In the reverse direction, if A and B are not F-similar, then the RCF of A is not equal to the RCF of B over F, so they are not equal over K, so A and B are not similar over K. □

Theorem: If A is a matrix, then the following Smith Normal Form computes the RCF of A:

$$\operatorname{SNF}(xI - A) = \operatorname{diag}(1, \dots, 1, a_1(x), \dots, a_m(x)),$$

where there are $n - m$ instances of 1.

Proof. First, it can be shown that if A is the companion matrix of $a(x) \in F[x]$, then

$$\operatorname{SNF}(xI - A) = \operatorname{diag}(1, \dots, 1, a(x)).$$

Then, if A is an arbitrary matrix in rational canonical form, then we reduce to the first case by restricting the row and column operations to each block.

Finally, if A is a general matrix, then if we set $B = \operatorname{RCF}(A)$, then A and B have the same invariant factors, so $\operatorname{SNF}(xI - B)$ is in the desired form. We observe then that

$$\operatorname{SNF}\big(P(xI - B)P^{-1}\big) = \operatorname{SNF}(xI - A),$$

so we are done. □

Now, we can relate the RCF to the minimal and characteristic polynomials of linear transformations.

Definition: Let F be a field, $n \in \mathbb{N}$, $T\colon V \to V$ a linear map between n-dimensional vector spaces. The polynomial $\chi_T(x) \coloneqq \det(xI - T)$ is called the *characteristic polynomial* of T.

Definition: Let F, V, T be as above. The *minimal polynomial* of T, denoted $\mu_T(x)$, is the unique monic polynomial of minimal degree such that $\mu_T(T) = 0$.

For existence, observe that $\left\{1, T, T^2, \dots, T^{n^2}\right\}$ are linearly dependent in the space $\operatorname{end}_F(V)$, so there is some dependence relation among them. We may then take the dependence relation with smallest degree.

For uniqueness, observe that if there are two such minimal polynomials, then their difference is a polynomial of strictly smaller degree, so it must be equal to 0.

Definition: Let $V = V_T$ be the vector space V considered as an $F[x]$-module, where we let $x \cdot v = Tv$. Define

$$\begin{aligned} I_T &\coloneqq \operatorname{ann}_{F[x]}(V_T) \\ &= \{p(x) \in F[x] \mid p(x) \cdot v = 0 \text{ for all } v \in V\} \\ &= \{p(x)\mathbb{F}[x] \mid p(T) = 0\}. \end{aligned}$$

Then, I_T is a nonzero ideal in $F[x]$, so I_T is generated by a single element (since $F[x]$ is a PID); we may also define μ_T to be the unique monic polynomial that generates I_T.

Theorem: Let $T\colon V \to V$ be a linear map, and let $a_1(x), \dots, a_m(x)$ be its invariant factors. Then,

$$\begin{aligned}\chi_T(x) &= \prod_{k=1}^{m} a_k(x) \\ \mu_T(x) &= a_m(x).\end{aligned}$$

Proof. Choose a basis β of V, and let $A = [T]_\beta$. Then, there is some unit $u \in (F[x])^\times$ such that

$$\begin{aligned}\chi_T(x) &= \chi_A(x) \\ &= \det(xI - A) \\ &= u \det(\text{SNF}(xI - A)) \\ &= u \prod_{k=1}^{m} a_k(x),\end{aligned}$$

but since $\chi_T(x)$ is monic, $u = 1$.

By the invariant factors decomposition of V, we have

$$V \cong \bigoplus_{i=1}^{m} F[x]/(a_i(x)),$$

with $(a_1(x)) \supseteq \cdots \supseteq (a_m(x))$. If $p(x) \in F[x]$, then $p(x) \in \text{ann}(V_T)$ if and only if $p(x) \in \text{ann}(F[x]/(a_i(x)))$ for each $i = 1, \dots, m$, which holds if and only if $p(x) \in (a_i(x))$ for each i. Thus, $a_m(x)$ is a generator for $(\mu_T(x))$, so since $a_m(x)$ is also monic, it follows that $a_m(x) = \mu_T(x)$. □

Corollary (Cayley–Hamilton Theorem)**:** If $T\colon V \to V$ is a linear transformation between finite-dimensional vector spaces, then $\mu_T(x) | \chi_T(x)$.

Furthermore, there is some exponent k such that $\chi_T(x) | (\mu_T(x))^k$.