Fundamental theorem of modules over a PID and applications

Travis Schedler, WOMP 2007

September 11, 2007

0.1 The fundamental theorem of modules over PIDs

- A PID (Principal Ideal Domain) is an integral domain (=ring without zero-divisors) such that every ideal is principal (=generated by a single element).
- Examples: any field $\mathbf{k}$; $\mathbf{k}[x]$, $\mathbb{Z}$, $\mathbb{Z}[i]$ (the latter three use the **Euclidean algorithm** and are more generally examples of Euclidean domains); $\mathbb{Z}[\frac{1+\sqrt{-19}}{2}]$ (a PID that is **not** a Euclidean domain).
- Any PID is a Noetherian UFD (Unique Factorization Domain). Here:
  - Noetherian means that every ideal is finitely generated;
  - Equivalently, every finitely-generated module is a Noetherian module (all submodules are finitely-generated).

The main goal of this talk is to prove and apply the theorem

Theorem 0.1.1. *If R is a PID, and M any finitely-generated R-module, then*

$$M \cong \bigoplus_{i=1}^{m} R/s_i, \tag{0.1.2}$$

where $s_1 \mid s_2 \mid \cdots \mid s_m$.

Corollary 0.1.3. *We may alternatively write M as*

$$M \cong R^{\oplus r} \oplus R/p_1^{r_1} \oplus \cdots \oplus R/p_n^{r_n}, \tag{0.1.4}$$

for some primes p_i and integers $r_i \geq 1, r \geq 0$.

Proof of Corollary 0.1.3. Simply apply unique factorization, and observe that

$$R/(p_1^{r_1} \cdots p_j^{r_j}) \cong R/p_1^{r_1} \oplus \cdots \oplus R/p_j^{r_j}. \tag{0.1.5}$$

□

0.2 Applications of the fundamental theorem The following is immediate:

Theorem 0.2.1. *Any finitely-generated abelian group A may be represented as a direct sum*

$$A \cong \bigoplus_{i=1}^{m} \mathbb{Z}/s_m, \tag{0.2.2}$$

with $s_1 \mid s_2 \mid \cdots \mid s_m$. Alternatively, we may write

$$A \cong \mathbb{Z}^{\oplus r} \oplus \mathbb{Z}/p_1^{r_1} \oplus \cdots \oplus \mathbb{Z}/p_n^{r_n}. \tag{0.2.3}$$

The following consequences of Theorem 0.1.1 (and Corollary 0.1.3) are somewhat less obvious.

Theorem 0.2.4 (Rational canonical form)**.** *Let M be any square matrix over a field $\mathbf{k}$. Then, after changing basis, we may rewrite M as a block-diagonal matrix, with blocks M_i of the form*

$$\begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & -a_{i,0} \\ 1 & 0 & 0 & 0 & \cdots & -a_{i,1} \\ 0 & 1 & 0 & 0 & \cdots & -a_{i,2} \\ \vdots & 0 & \ddots & \ddots & & \vdots \\ 0 & \cdots & & & 1 & -a_{i,d_i-1} \end{pmatrix}, \tag{0.2.5}$$

for polynomials

$$s_i = x^{d_i} + a_{i,d_i-1}x^{d_i-1} + \cdots + a_{i,1}x + a_{i,0} \tag{0.2.6}$$

such that

$$s_1 \mid s_2 \mid \cdots s_m, \quad \chi(M) = s_1 \cdot s_2 \cdot \cdots \cdot s_m, \tag{0.2.7}$$

$\chi(M) =$ the characteristic polynomial of M.

Proof. The main idea is as follows: View a matrix M as a linear transformation $T : V \to V$. Now, for any polynomial $P(x)$, we have the linear operator

$$P(T) : V \to V, \tag{0.2.8}$$

viewing $T^m = T \circ T \circ \cdots \circ T$, m times (i.e., multiplication is composition of operators).

Thus, we may view V as an $\mathbf{k}[x]$-module, and thus may apply Corollary 0.1.3 to describe the structure of V. We obtain

$$V \cong \mathbf{k}[x]/s_1 \oplus \cdots \oplus \mathbf{k}[x]/s_m, \tag{0.2.9}$$

for some polynomials $s_1 \mid \cdots \mid s_m$. Let us write

$$V = V_1 \oplus \cdots \oplus V_m, \tag{0.2.10}$$

for some vector subspaces such that

$$V_i \cong \mathbf{k}[x]/s_i. \tag{0.2.11}$$

Picking bases for each V_i yields a basis of V in which M is block-diagonal:

$$\begin{pmatrix} M_1 & 0 & 0 & \cdots & 0 \\ 0 & M_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & & & M_n \end{pmatrix}, \tag{0.2.12}$$

where each M_i is the matrix for $T|_{V_i}$.

It remains to describe each M_i. Let us suppose that $d_i = \deg s_i$. Then, $V_i \cong \mathbf{k}[x]/s_i$ is a vector space of dimension d_i, with basis $1, x, x^2, \dots, x^{d_i-1}$ (using the isomorphism). Using this basis, we obtain the matrix (0.2.5), whose characteristic polynomial is s_i.

It remains to verify the second part of (0.2.7). This follows because the characteristic polynomial of a block-diagonal matrix is the product of the characteristic polynomials of the blocks. □

Theorem 0.2.13 (Jordan form)**.** *Let M be any square matrix over a field $\mathbf{k}$. Then, after changing basis, we may rewrite M as a block-diagonal matrix, with blocks of the form*

$$\begin{pmatrix} D & 0 & 0 & 0 & \cdots & 0 \\ Y & D & 0 & 0 & \cdots & 0 \\ 0 & Y & D & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & & \vdots \\ 0 & \cdots & & & Y & D \end{pmatrix}, \tag{0.2.14}$$

where D is a matrix whose characteristic polynomial is irreducible, and Y is the elementary matrix

$$\begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & 0 \end{pmatrix}, \tag{0.2.15}$$

*of the same size as D. (We may choose D to be in **rational canonical form**,* (0.2.5)*.)*[1]

Corollary 0.2.16. *If $\mathbf{k}$ is algebraically closed, then the Jordan blocks have the form*

$$\begin{pmatrix} \lambda & 0 & 0 & 0 & \cdots & 0 \\ 1 & \lambda & 0 & 0 & \cdots & 0 \\ 0 & 1 & \lambda & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & & \vdots \\ 0 & \cdots & & & 1 & \lambda \end{pmatrix}, \tag{0.2.17}$$

for some $\lambda \in \mathbf{k}$.

Proof of Corollary 0.2.16. If $\mathbf{k}$ is algebraically closed, any irreducible characteristic polynomial must be linear, and thus the matrix D must be a 1×1-matrix. □

Remark 0.2.18. We don't actually need $\mathbf{k}$ to be algebraically closed for the corollary, as long as the characteristic polynomial splits into linear factors (i.e., all eigenvalues are in $\mathbf{k}$).

Proof of Theorem 0.2.13. Using Corollary 0.1.3 (instead of Theorem 0.1.1), we obtain

$$V \cong \mathbf{k}[x]^{\oplus r} \oplus \mathbf{k}[x]/p_1^{r_1} \oplus \cdots \oplus \mathbf{k}[x]/p_n^{r_n}, \tag{0.2.19}$$

[1] We can think of D as a single number: a root α of its characteristic polynomial. This makes sense if we view the vector space on which D acts as isomorphic to $\mathbf{k}[x]/\chi(D) = \mathbf{k}[\alpha]$, with the matrix acting as multiplication by α. If $\mathbf{k}$ is perfect, then we can also replace Y by I and view the whole matrix as acting on a $\mathbf{k}[x]/\chi(D)$-vector space (with a Jordan form like (0.2.17)); but not in general if $\chi(D)$ is not separable over $\mathbf{k}$.

where p_i are irreducible polynomials, and $r_i \geq 1$. We must have $r = 0$ since V is finite-dimensional.

Let us again choose subspaces $V_i \subset V$ so that $V_i \cong \mathbf{k}[x]/p_i^{r_i}$ as a $\mathbf{k}[x]$-module, and $V = V_1 \oplus \cdots \oplus V_n$. It remains only to show that (0.2.14) is a matrix for $T|_{V_i}$ with the appropriate choice of basis. We use the basis

$$\{x^j \cdot p_i^\ell\}_{0 \leq j < d_i; 0 \leq \ell < r_i}. \tag{0.2.20}$$

It is obvious that this is a basis because it consists of one polynomial for every degree between 0 and $d_i \cdot r_i - 1 = \deg(p_i^{r_i}) - 1$.

Now, acting by T on the element $x^j \cdot p_i^\ell$ just is multiplication by x. It remains only to note that

$$x \cdot (x^{d_i - 1} \cdot p_i^\ell) = p_i^{\ell+1} - (a_0 + a_1 \cdot x + \cdots + a_{d_i - 1} \cdot x^{d_i - 1}) \cdot p_i^\ell. \tag{0.2.21}$$

□

0.3 Relationship to semisimple and nilpotent decomposition When $\mathbf{k}$ is algebraically closed, Jordan form is a decomposition of our square matrix M into

$$M = S + N, \tag{0.3.1}$$

where S is semisimple (given by the diagonal in the basis (0.2.17), with all other entries zero), and N is nilpotent (given by the subdiagonal in the basis (0.2.17), with all other entries zero). Here, a matrix N is nilpotent iff $N^j = 0$ for some $j \geq 1$, and S is semisimple iff it is diagonalizable over the algebraic closure of $\mathbf{k}$.

More generally, we only need $\mathbf{k}$ to be perfect:

Theorem 0.3.2. *Let $\mathbf{k}$ be a perfect field. The matrix M may be uniquely written as*

$$M = S + N, \tag{0.3.3}$$

where S is semisimple and N is nilpotent, in such a way that S and N are $\mathbf{k}$-polynomials of M. The matrices M, S, and N all commute with each other.

We may interpret this theorem when M is written as (0.2.14): the part with D's is the matrix S, and the part with Y's is the matrix N. It is not difficult to check that this S is semisimple. It is clear that S and N commute and hence that they commute with M.

To show that S is a polynomial in M is not really simplified using (0.2.14). Thus, the following actually proves the theorem without needing Jordan form (but using (0.2.14) perhaps gives a visual picture of what is going on):

Proof. First, suppose

$$\chi(M) = p_1^{s_1} \cdots p_n^{s_n} \tag{0.3.4}$$

for distinct irreducibles p_i. For each i, let q_i be a polynomial such that

$$p_1^{s_1} \cdots p_{i-1}^{s_{i-1}} q_i p_{i+1}^{s_{i+1}} \cdots p_n^{s_n} \equiv 1 \pmod{p_i^{s_i}}. \tag{0.3.5}$$

Let $P_i(x)$ be the above polynomial. Then, $P_i(M)$ must be the diagonal matrix (as in (0.2.12)) which has 0's all along the diagonal except at the M_i-block, where it is the identity. So, we may reduce to the case when $n = 1$, and $\chi(M) = p_1^{s_1}$.

Now, if $\mathbf{k}$ is algebraically closed, then it is clear that the semisimple part S is a scalar, which is a polynomial of M.

If $\mathbf{k}$ is only assumed to be perfect, we use Galois theory (covered in the next lecture): Let K be the splitting field over $\mathbf{k}$ of the polynomial p_q. Then, $p_1 = q_1 q_2 \cdots q_\ell$ for some linear polynomials q_i. We may assume that

$$q_i = x - \lambda_i, \tag{0.3.6}$$

with $\{\lambda_i\}$ an orbit of the Galois group of K over $\mathbf{k}$ (this orbit has size $[K : \mathbf{k}]$). Then, using the above procedure, we have $S = P(M)$ for

$$P(x) = \sum_{\sigma \in \mathrm{Gal}(K/\mathbf{k})} \sigma(\lambda_1 Q_1)(x) \in \mathbf{k}[x]. \tag{0.3.7}$$

For uniqueness, we note that if $S + N = S' + N'$ for S, N, S', N' all polynomials of M, then $S - S' = N' - N$. Since N', N and S, S' commute, the LHS is semisimple and the RHS is nilpotent, which shows they are both zero. $\square$

Remark 0.3.8. We don't actually need $\mathbf{k}$ to be perfect, as long as all the irreducible factors of $\chi(M)$ are separable (i.e., have distinct roots over $\bar{\mathbf{k}}$). Otherwise, we can canonically decompose $M = S + N$ over the separating field of $\chi(M)$ over $\mathbf{k}$ (the minimal extension such that all irreducible factors of $\chi(M)$ are separable). Without passing to such an extension, we can get a Jordan/rational canonical form, but it won't necessarily yield an $S + N$ one: for example, $\begin{pmatrix} 0 & y \\ 1 & 0 \end{pmatrix}$ has irreducible characteristic polynomial $x^2 + y$ over $\mathbb{F}_2(y)$, but it is equivalent to $\begin{pmatrix} \sqrt{y} & 0 \\ 1 & \sqrt{y} \end{pmatrix}$ over $\mathbb{F}_2(\sqrt{y})$.

Remark 0.3.9. There is a multiplicative analogue: any invertible matrix M may be uniquely written as $M = SU$, where S is semisimple and U is unipotent, and S, U are $\mathbf{k}$-polynomials of M (assuming $\chi(M)$ is separable over $\mathbf{k}$). Indeed, simply set $U = S^{-1}(S + N)$ in the above theorem.

0.4 Proof of the fundamental theorem The fundamental theorem may be proved by using a generalization of Gaussian elimination of matrices.

First, let us reduce the problem to one that involves matrices. Given any finitely-generated module M over a PID R, being finitely generated means that

$$R^{\oplus d} \twoheadrightarrow M, \tag{0.4.1}$$

for some $d \geq 0$. Now, let the kernel of this map be K. Because any PID is Noetherian, K must be finitely generated as well, so we have

$$R^{\oplus f} \twoheadrightarrow K \hookrightarrow R^{\oplus d} \twoheadrightarrow M, \tag{0.4.2}$$

so that M is the cokernel of the composition

$$R^{\oplus f} \to R^{\oplus d}. \tag{0.4.3}$$

Claim 0.4.4. *We can choose new R-bases of $R^{\oplus f}, R^{\oplus d}$ such that the map $R^{\oplus f} \to R^{\oplus d}$ has the form (if $f = d$):*

$$\begin{pmatrix} s_1 & 0 & 0 & \cdots & 0 \\ 0 & s_2 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & & \vdots \\ 0 & \cdots & & & s_n \end{pmatrix}, \tag{0.4.5}$$

for $s_1 \mid s_2 \mid \cdots \mid s_n$ (if $f \neq d$, then add zeros to the right or bottom of the above matrix).

From the claim we deduce the theorem, since the cokernel of the above matrix is $\bigoplus_{i=1}^{n} R/s_i$.

Proof of Claim 0.4.4. Given any $d \times f$-matrix M, changing bases on the domain and range is the same as replacing M by a new matrix

$$M' = AMB, \tag{0.4.6}$$

where A, B are invertible matrices of the appropriate size.

Now, given

$$\begin{pmatrix} a_{00} & \cdots & a_{0f} \\ \vdots & \ddots & \vdots \\ a_{d0} & \cdots & a_{df} \end{pmatrix}, \tag{0.4.7}$$

we would like to apply the Gaussian elimination (along rows and columns) together with permutations to obtain a matrix of the desired form.

To show this, let s_1 be any generator of the ideal $(a_{ij})_{0 \le i \le d, 0 \le j \le f}$.

Claim 0.4.8. *Using row and column operations and permutations, we can ensure that $a_{00} \mid s_1$.*

Using this claim, we may kill off a_{i0}, a_{0i} for $i \neq 0$. Now, we are left with a matrix

$$\begin{pmatrix} a_{00} & 0 & \cdots & 0 \\ 0 & a_{11} & \cdots & a_{1f} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{f1} & \cdots & a_{fd} \end{pmatrix}, \tag{0.4.9}$$

where $a_{00} \mid a_{ij}$ for all i, j. Then, by applying induction, we obtain a matrix of the desired form. □

Proof of Claim 0.4.8. It is enough to show that, for any entries $a_{ij}, a_{k\ell}$, that we can change bases so that $\gcd(a_{ij}, a_{k\ell})$[2] appears somewhere in the matrix (using permutations).

It is enough to show that

$$\begin{pmatrix} a \\ b \end{pmatrix} \tag{0.4.10}$$

can be multiplied by invertible matrices so that $\gcd(a, b)$ appears. Suppose that $\gcd(a, b) = x \cdot a + y \cdot b$. Clearly, $(x, y) = (1)$ (otherwise we could divide this equation and get a gcd with fewer prime factors). Then, if $x \cdot u + y \cdot v = 1$, the matrix

$$\begin{pmatrix} x & y \\ -v & u \end{pmatrix} \tag{0.4.11}$$

is invertible, and

$$\begin{pmatrix} x & y \\ -v & u \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \gcd(a, b) \\ -v \cdot a + u \cdot b \end{pmatrix}, \tag{0.4.12}$$

as desired. □

[2] We use the abusive notation $\gcd(x, y)$ for any element such that $(x, y) = (\gcd(x, y))$. This is unique up to multiplication by a unit.