Jordan normal form

1 Principal ideal domains

A commutative ring R is an *integral domain* if it has no zero divisors, i.e., $a \cdot b = 0$ for some $a, b \in R$ implies $a = 0$ or $b = 0$. An integral domain R is a principal ideal domain (PID) if every ideal in R is generated by a single element. Examples of PID-s include $\mathbb{Z}$ and polynomial rings $F[x]$ of a single variable over a field F. Both examples have a *norm function* $N : R \setminus \{0\} \to \mathbb{N}$ that satisfies the *Euclidean property*: for all $a \in R$ and $b \in R \setminus \{0\}$ there is a $q \in R$ and and $r \in R$ such that $a = qb + r$ and $r = 0$ or $N(r) < N(b)$. (For $R = \mathbb{Z}$, the norm function may be chosen to be the absolute value, for $R = F[x]$ the norm function may be chosen as the degree function.) Such domains are called *Euclidean domains* and they have a *Euclidean algorithm*.

Lemma 1 *Every Euclidean domain R is a PID.*

Proof: Let I be a nonzero ideal of R and let $b \in I \setminus \{0\}$ be an element whose norm is smallest among the norms of elements of $I \setminus \{0\}$. We claim that $I = (b)$. In fact, for any $a \in I$, we may write $a = qb + r$ such that $r = 0$ or $N(r) < N(b)$. If $r = 0$ then $a = qb \in (b)$. If $r \neq 0$ then $r = a - qb \in I \setminus \{0\}$ has smaller norm than b, in contradiction with the choice of b. Thus we must have $a \in (b)$ for all $a \in I$, and so $I \subseteq (b)$. The other inclusion $(b) \subseteq I$ is obvious. ◇

An element $p \in R$ is *prime* if $p \mid ab$ implies $p \mid a$ or $p \mid b$. An element $i \in R$ is *irreducible* if it can not be written in the form $i = i_1 \cdot i_2$ without at least one of i_1 and i_2 being a unit (=a divisor of 1). If $a = ub$ where u is a unit, then we call a and b *associates* and use the notation $a \sim b$. It is easy to see that $\sim$ is an equivalence relation, and that $a \sim b$ if and only if $(a) = (b)$.

Lemma 2 *In every integral domain R all primes are irreducible.*

Proof: Assume p is a prime and $p = a \cdot b$. Then either $p \mid a$ or $p \mid b$, w.l.o.g. $p \mid a$. Thus $a = r \cdot p$ for some $r \in R$, and we have $p = rpb$. Since R is an integral domain, we may simplify by p and get $1 = rb$. Thus b is a unit, $a \sim p$. ◇

In a PID the converse holds as well.

Lemma 3 *In a PID every irreducible is prime.*

Proof: Assume p is irreducible and that $p \mid a \cdot b$. If $p \mid a$, we are done. Otherwise, consider the ideal (p, a). This is generated by a single element d. The element p is irreducible and a multiple of d, thus either $d \sim p$ or d is a unit. If $d \sim p$ then $d \mid a$ implies $p \mid a$, a contradiction. Thus d must be a unit, w.l.o.g. $d = 1$. Hence $(p, a) = (1)$ and so $(pb, ab) = (b)$. Now p divides pb and ab, so p divides b. ◇

Theorem 1 *Every PID R is a unique factorization domain (UFD): every $r \in R$ may be written as a product of powers of primes. This decomposition is unique up to taking associates.*

2 Finitely generated modules over principal ideal domains

An R-module M is *finitely generated* if there is a finite set $\{m_1, \ldots, m_k\} \subseteq M$ such that each $m \in M$ may be written as $m = r_1 m_1 + \cdots + r_k m_k$ for some $r_1, \ldots, r_k \in R$. An R-module is *cyclic* if it is generated by a single element of M.

The main theorem is the following.

Theorem 2 *Let R be a PID and M a finitely generated R-module. Then M may be written as a finite direct sum of cyclic modules $M_1 \oplus \cdots \oplus M_k$ such that each M_i is either isomorphic to R (torsion free), or to $R/(p^\alpha)$ for some prime power p^α. This decomposition is unique up to the numbers of M_i's that are isomorphic to R or a given $R/(p^\alpha)$.*

Example 1 Every Abelian group is a $\mathbb{Z}$-module. Thus we obtain that every finitely generated Abelian group is a direct sum of copies of $\mathbb{Z}$ and copies of $\mathbb{Z}_{p^\alpha}$'s, and that this decomposition is essentially unique.

Our main application of Theorem 2 is to show that every linear operator $T : V \to V$ on finite dimensional complex vector space V has a *Jordan normal form.* For that purpose consider first a finite dimensional vector space V over an arbitrary field and a linear map $T : V \to V$.

Lemma 4 *The scalar multiplication defined as $p(x) \cdot v := p(T)(v)$ turns V into a unitary $F[x]$-module.*

Since $F[x]$ is a PID, Theorem 2 is applicable, and we may write V as a direct sum of cyclic $F[x]$-modules.

Lemma 5 *A subspace $U \leq V$ is an $F[x]$-submodule if and only if it is a T-invariant subspace.*

Lemma 6 *A subspace $U \leq V$ is a cyclic $F[x]$-module if and only if it is a T-cyclic submodule of V.*

Note that, no cyclic $F[x]$-submodule of V can be isomorphic to $F[x]$, since V is a finite dimensional, and $F[x]$ is an infinite dimensional vector space. Thus every cyclic $F[x]$-submodule U of V is appearing in this instance of Theorem 2 is isomorphic to $F[x]/(p(x)^\alpha)$ for some irreducible polynomial $p(x) \in F[x]$. Note that the characteristic polynomial of x acting on $F[x]/(p(x)^\alpha)$ is $p(x)^\alpha$.

Assume from now on that $F = \mathbb{C}$. Every irreducible polynomial of $\mathbb{C}[x]$ is linear. Thus every cyclic $\mathbb{C}[x]$-submodule appearing in Theorem 2 applied to $R = \mathbb{C}[x]$ is of the form $\mathbb{C}[x]/((x-\lambda)^\alpha$ for some $\lambda \in \mathbb{C}$ and $\alpha \in \mathbb{N}$. Let v be the generator of such a submodule U. Then v, $(T-\lambda I)v$, $\ldots$, $(T-\lambda I)^{\alpha-1}v$ are linearly independent, whereas $(T-\lambda I)^\alpha v = 0$. The submodule U is the linear span of v, $(T-\lambda I)v$, $\ldots$, $(T-\lambda I)^{\alpha-1}v$, and these vectors form a basis of U. Since

$$T(T-\lambda I)^k(v) = (T-\lambda I)^{k+1}(v) + \lambda(T-\lambda I)^k(v) \quad \text{for } k = 1, 2, \ldots, \alpha - 1.$$

the matrix of T_U in this basis is

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

Such a matrix is called a *Jordan block.* Theorem 2 implies that the matrix of every linear map $T : V \to V$ on a finite dimensional complex vector space V may be written as a block-diagonal matrix of Jordan blocks.

3 The proof of the main theorem

In this section R is a PID.

Lemma 7 *Let P be an $n \times n$ matrix with entries from R. Then P has an inverse if and only if* $\det(P)$ *is a unit.*

Proof: If $P \cdot Q = I$ then $\det(P)\det(Q) = 1$. Conversely, if $\det(P)$ is a unit, then $\det(P)^{-1} \in R$ and P^{-1} may be obtained "using Cramer's rule". ◇

Lemma 8 *Assume that the ideal $(r_1, \ldots, r_n)$ is generated by the single element δ. ("The GCD of $r_1, \ldots, r_n$ is δ"). Then there is an $n \times n$ matrix with entries from R whose first row is $r_1, \ldots, r_n$ and whose determinant is δ.*

Proof: Induction on n. Assume the Lemma is true for some n, and consider $(r_1, \ldots, r_{n+1}) = (\delta_{n+1})$. Introducing δ_n for a generator of $(r_1, \ldots, r_n)$ we may write $\delta_{n+1} = \delta_n \cdot \xi + r_{n+1} \cdot \eta$ for some $\xi, \eta \in R$. By our induction hypothesis there is an $n \times n$ matrix whose first row is $r_1, \ldots, r_n$ and whose determinant is δ_n. Add now $(-\eta r_1/\delta_n, \ldots, -\eta r_n/\delta_n, \xi)$ as the last row to the matrix and $(r_{n+1}, 0, \ldots, 0, \xi)^T$ as the last column, to obtain the following matrix:

$$\begin{pmatrix} r_1 & \cdots & r_n & r_{n+1} \\ & & & 0 \\ & & & \vdots \\ & & & 0 \\ -\eta r_1/\delta_n & \cdots & -\eta r_n/\delta_n & \xi \end{pmatrix}$$

Expanding by the last column yields that the determinant of this matrix is $\xi \cdot \delta_n + r_{n+1} \cdot \eta = \delta_{n+1}$. ◇

Lemma 9 *Assume A is an $n \times n$ matrix with entries from R. Then there are invertible matrices P and Q with entries from R such that PAQ is a diagonal matrix diag$(\alpha_1, \ldots, \alpha_n)$ such that $\alpha_i \mid \alpha_{i+1}$ holds for $i = 1, 2, \ldots, n-1$.*

Proof: We proceed by induction on n. Let $(r_1, \ldots, r_n)^T$ be the first column of A and assume that a generator of the ideal $(r_1, \ldots, r_n)$ is $\alpha = \xi_1 r_1 + \cdots \xi_n r_n$. Then we have

$$1 = \xi_1 \frac{r_1}{\alpha} + \cdots \xi_n \frac{r_n}{\alpha}$$

and so $(\xi_1, \ldots, \xi_n) = 1$. By Lemma 8, there is a matrix whose first row is $(\xi_1, \ldots, \xi_n)$ and whose determinant is 1. By Lemma 7 such a matrix is invertible. Multiplying A by such a matrix from the right we obtain a matrix whose first column contains the GCD of the column entries in its first row. Repeat the same procedure for the first row, and we obtain a new matrix whose first row contains the GCD of the row entries in the first column. Now repeat the same procedure for the first column again and so on. At the end of each phase the first row first column entry becomes a divisor of the first row first column entry of the previous phase. Since R has unique factorization, this procedure can not keep giving proper divisors for ever, since the number of prime factors (counted with multiplicities) of the first row first entry can not decrease indefinitely. Hence after finitely many steps we obtain a first row first column entry δ that divides all entries in the first row and all entries in the first column. If δ does not divide all entries in the matrix, say it does not divide an entry in the k-th row, exchange the first and the k-th row and restart the procedure. (Note that δ remains in the first column, so after two steps we get a proper divisor of δ in the first row first column.) Again, we can not have an infinite descending chain of proper divisors, thus after finitely many steps we arrive at a matrix whose only nonzero entry in the first row and in the first column is the first row, first column entry α_1 that divides all other entries in the matrix. By our induction hypothesis, we may transform the matrix formed by rows $2, 3, \ldots, n$ and columns $2, 3, \ldots, n$ into a diagonal matrix $\operatorname{diag}(\alpha_2, \ldots, \alpha_n)$ such that $\alpha_i \mid \alpha_{i+1}$ holds for $i = 2, \ldots, n-1$. $\diamond$

Lemma 10 *The entries $\alpha_1, \ldots, \alpha_n$ in the previous lemma are unique, up to taking associates.*

Proof: Let $\Delta_i(A)$ be the GCD of all $i \times i$ minors of A. Using Cauchy-Binet, it is easy to show that $\Delta_i(AB)$ is a multiple of $\Delta_i(A)$ for all $n \times n$ matrix A with entries in R. Assume now $B = PAQ$ where P and Q are invertible. From $B = PAQ$ we have $\Delta_i(A) \mid \Delta_i(B)$ for all i, from $A = P^{-1}BQ^{-1}$ we have $\Delta_i(B) \mid \Delta_i(A)$ for all i. Thus $\Delta_i(B) \sim \Delta_i(A)$ for all i. Assume now that B is $\operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ such that $\alpha_i \mid \alpha_{i+1}$ holds for $i = 1, 2, \ldots, n-1$. Then $\Delta_i(B) = \alpha_1 \cdots \alpha_i$ and so we must have $\alpha_1 \sim \Delta_1(A)$ and $\alpha_i \sim \Delta_i(A)/\Delta_{i-1}(A)$ for $i = 2, \ldots, n$. $\diamond$

An R-module is *free* if it is the direct sum of modules isomorphic to R.

Proposition 1 *Assume M is a free R-module, isomorphic to the direct sum of n-copies of R. Then every submodule N of M is free.*

Proof: First we show by induction on n that N may be generated by n elements. Consider the set

$$I := \{r \in R \ : \ \exists r_2, \ldots, r_n \in R \ (r, r_2, \cdots, r_n) \in N\}.$$

Obviously, I is an ideal of R. Since R is a PID, I is generated by an element ε. By the definition of I, there are entries $\varepsilon_2, \ldots, \varepsilon_n$ such that $(\varepsilon, \varepsilon_2, \ldots, \varepsilon_n) \in N$. Furthermore, for any $(r_1, \ldots, r_n) \in N$ the first coordinate r_1 is a multiple of ε. Subtracting the appropriate multiple of $(\varepsilon, \varepsilon_2, \ldots, \varepsilon_n)$ from $(r_1, \ldots, r_n)$ we get an element of N whose first coordinate is zero. We obtained that N is generated by $(\varepsilon, \varepsilon_2, \ldots, \varepsilon_n)$ and the submodule

$$N_0 := \{(r_1, \ldots, r_n) \in N \ : \ r_1 = 0\}.$$

By our induction hypothesis, N_0 may be generated by $n-1$ elements, and so N may be generated by n elements.

Assume now that N is generated by $f_1, \ldots, f_n$ and that M is freely generated by $e_1, \ldots, e_n$. Then each f_i may be uniquely written as $f_i = \sum_{j=1}^n a_{i,j} e_j$. Let us denote the matrix with entries $a_{i,j}$ by A. By Lemma 9 there are invertible matrices P and Q such that PAQ is a diagonal matrix $\text{diag}(\alpha_1, \ldots, \alpha_n)$ such that $\alpha_i \mid \alpha_{i+1}$ holds for $i = 1, 2, \ldots, n-1$. Replacing A with PAQ corresponds to replacing the basis $e_1, \ldots, e_n$ of M by some basis $e'_1, \ldots, e'_n$ and and the generating system $f_1, \ldots, f_n$ by some generating system $f'_1, \ldots, f'_n$. (Note that $e'_1, \ldots, e'_n$ is still a basis since any nontrivial linear relation $r_1 \cdot e'_1 + \cdots + r_n e'_n = 0$ would induce a nontrivial linear relation among the e_j-s.) Assume that $\alpha_1, \ldots, \alpha_k \neq 0$ but $\alpha_{k+1} = \cdots = \alpha_n = 0$. This implies $f'_{k+1} = \cdots = f'_n = 0$ and that for $j \leq k$, f'_j is a nonzero multiple of e'_j. As a consequence N is freely generated by $f'_1, \ldots, f'_k$. ◇

Theorem 3 *Every finitely generated R-module M has a finite basis, i.e., it is the direct sum of finitely many cyclic R-modules.*

Proof: Assume M is generated by $u_1, \ldots, u_n$. Consider the homomorphism $\phi : R^n \to M$, sending $(r_1, \ldots, r_n)$ into $r_1 u_1 + \cdots + r_n u_n$. By the proof of the previous Proposition, the kernel of ϕ is a free submodule of R^n, and R^n has a basis $e'_1, \ldots, e'_n$ such that $\alpha_1 e'_1, \ldots, \alpha_k e'_k$ form a basis of the kernel of ϕ for some integer k and some $\alpha_1, \ldots, \alpha_k \in R$. Consider now the generating system $\phi(e'_1), \ldots, \phi(e'_n)$ of M. We have

$$r_1 \phi(e'_1) + \cdots + r_n \phi(e'_n) = 0$$

if and only if $r_1 e'_1 + \cdots + r_n e'_n$ belongs to the kernel of ϕ which is equivalent to

$$r_1 e'_1 + \cdots + r_n e'_n = s_1 \alpha_1 e'_1 + \cdots + s_k \alpha_k e'_k$$

for some $s_1, \ldots, s_k \in R$. Since $e'_1, \ldots, e'_n$ is a free basis of R^n, this is only possible if $r_{k+1} = \cdots = r_n = 0$ and $r_j = s_j \alpha_j$ for $j \leq k$. But then we have

$$r_1 \phi(e'_1) = \cdots = r_n \phi(e'_n) = 0$$

component-wise. ◇

Proposition 2 *Every cyclic R module is either isomorphic to R or isomorphic to a direct sum of R-modules isomorphic to cyclic modules of the form $R/(p^\alpha)$ where p is a prime in R.*

Proof: If $M = (m)$ is a cyclic R-module then the kernel of the surjective homomorphism $R \to M$ sending 1 into m is an ideal I of R and M is isomorphic to R/I. If $I = 0$ then M is isomorphic to R, otherwise M is isomorphic to $R/(r)$ for some $r \in R \setminus \{0\}$. This element r has a unique prime factorization $r = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ and it is easy to see that $R/(r)$ is isomorphic to $R/(p_1^{\alpha_1}) \times \cdots \times R/(p_k^{\alpha_k})$ ("Chinese remainder theorem"). ◇

As a consequence of Theorem 3 and Proposition 2 we have the existence part of Theorem 2. For the uniqueness part let us introduce the *submodule of torsion elements*

$$\mathcal{T}(M) := \{m \in M \ : \exists r \in R \setminus \{0\} \ \ rm = 0\}.$$

It is easy to see that $\mathcal{T}(M)$ is a submodule and that $M/\mathcal{T}(M)$ is *torsion free*, i.e., $\mathcal{T}(M/\mathcal{T}(M)) = 0$. Thus $M/\mathcal{T}(M)$ is the direct sum of finitely many copies of R, and the number of these copies must be the same irrespective of the basis, otherwise the transition matrix from one basis to the other can not be invertible. Thus we may assume that $\mathcal{T}(M) = M$, i.e., each element of M is a torsion element. Given a prime $p \in R$, the direct sum of all summands of isomorphic to $R/(p^\alpha)$ for some α forms the submodule

$$M_p := \{m \in M \ : \exists k \geq 0 \ \ p^k \cdot m = 0\}.$$

Here M_p is a submodule of M, defined in a "basis independent" way. Thus we may assume $M = M_p$ for some prime $p \in R$. Let us introduce now the submodules

$$N_k := \{m \in M \ : \ p^k \cdots m = 0\}.$$

Since $M = N_k$, for some appropriately high value of k, it is sufficient to show the uniqueness by induction on k. If $M = N_1$, then M is a vector space over the field $R/(p)$, and so all of its bases have the same number of elements. Assume the statement is true for $M = N_{k-1}$ and consider $M = N_k$. Assume a basis of M is of the form $a_1, \ldots a_r; b_1, \ldots, b_s$ where $pa_i = 0$ for all i and $pb_j \neq 0$ for all j. Then $pb_1, \ldots pb_s$ is a basis of pM, a module satisfying $p^{k-1}m = 0$ for all $m \in pM$. By our induction hypothesis the uniqueness holds for the orders of $pb_1, \ldots pb_s$. The submodule

$$M[p] := \{m \in M \ : \ pm = 0\}$$

has basis

$$a_1, \ldots, a_r, \frac{|b_1|}{p} b_1, \ldots, \frac{|b_s|}{p} b_s.$$

Here $|b|$ is the generator of the ideal $\{r \in R \ : \ rb = 0\}$. The elements $\frac{|b_1|}{p} b_1, \ldots, \frac{|b_s|}{p} b_s$ form a basis of $M[p] \cap pM$. Thus $a_1, \ldots, a_r$ is a basis of $M[p]/(M[p] \cap pM)$, a vector space over $R/(p)$. Thus uniqueness holds also for $a_1, \ldots, a_r$.