MODULES OVER PIDS

In this note, we record for my Math 830 class the structure theorem for finitely generated modules over principal ideal domains. The most basic form of the theorem can be stated as follows.

Structure Theorem for finitely generated modules over a PID. *Let R be a PID and M a finitely generated R-module. Then M is a direct sum of cyclic modules.*

The next definition plays a key role in the structure theorem.

Definition A. Let R be an integral domain, M an R-module.

(i) Set $T(M) := \{m \in M \mid rm = 0, \text{ for some non-zero } r \in R\}$. It is easy to check that $T(M)$ is a submodule of M. $T(M)$ is called the *torsion submodule* of M.
(ii) M is *torsion-free* if $T(M) = 0$. M is a *torsion module* if $M = T(M)$.

If R is an integral domain, then a finitely generated free R-module F is torsion free. Indeed, if the rank of F is n, then $F \cong R^n$, which is clearly torsion-free. The following crucial ingredient to the structure theorem we seek shows that, over a PID, torsion-free modules are free. It follows that any $M \subseteq F$ is also torsion-free. The next theorem shows that the converse to this holds over a PID.

Theorem B. *Let R be a* PID and M a finitely generated R-module.

(i) If F is a finitely generated free R-module of rank r and $N \subseteq F$ is a submodule, then N is a finitely generated free R-module of rank less than or equal to r.
(ii) If M is torsion-free, then M is a free R-module.
(iii) If M can be generated by n elements, then any submodule of M can be generated by n elements or less. In particular, a submodule of a cyclic module is cyclic.
(iv) There exists a submodule $F \subseteq M$ such that F is a free R-module and $M = F \oplus T(M)$.

Proof. In order to make the notation a bit less confusing, henceforth (throughout these notes), if x is an element of an R-module, we will write $\langle x \rangle$ for the cyclic submodule generated by x and if $a \in R$, we will write aR for the principal ideal generated by a. For (i), we induct on r. If $r = 1$, then $F = \langle x \rangle$ for some $x \in R$. If $N = 0$, there is nothing to prove. Otherwise, since $N \subseteq Rx$, we consider the ideal $J := \{t \in R \mid tx \in N\}$. We then have $J = aR$, for some $a \in R$. We claim $N = \langle ax \rangle$, which is a free R-module, since F is torsion-free. Clearly $\langle ax \rangle \subseteq N$, by definition of a. On the other hand, if $n \in N$, then $n = rx$, for some $r \in J$, so $r = r'a$, for some $r' \in R$. Thus, $n = tx = (r'a)x = r'(ax)$, which shows that $N \subseteq \langle ax \rangle$, which gives what we want.

Suppose $r > 1$. Let $x_1, \ldots, x_r$ be a basis for F. Set $G := \langle x_1, \ldots, x_{r-1} \rangle$, a free module of rank $r - 1$. By induction on r, $N \cap G$ is either (0) or a free R-module of rank $r - 1$ or less. Now every element n in N can be written in the form $n = a_1x_1 + \cdots + a_{r-1}x_{r-1} + a_rx_r$, with each $a_j \in R$. If, for every element in N, $a_r = 0$, then $N \subseteq G$, and we are done by induction on r. Otherwise, let J denote the ideal of R generated by all of the coefficients of x_r as n varies over the elements of N. J is clearly an ideal of R. Thus, $J = aR$, for sone $a \in R$. By definition of J, there exists $n_0 \in N$ of the form $n_0 = s_1x_1 + \cdots + s_{r-1}x_{r-1} + ax_r$.

We claim $N = (N \cap G) \oplus \langle n_0 \rangle$. If so, then, on the one hand, $\langle n_0 \rangle$ is a free R-module of rank one. On the other hand, $N \cap G$ has a basis consisting of $r - 1$ or fewer elements. Putting these bases together gives a basis for N having no more than n elements (see Homework 5), which is what we want. For the claim, take $n = c_1x_1 + \cdots c_{r-1}x_{r-1} + c_rx_r$ in N. If $c_r = 0$, then $n \in N \cap G$. Otherwise, by definition of J, $c_r = da$, for some $d \in R$. it follows that $n - dn_0 \in N \cap G$. Therefore, $n \in (N \cap G) + \langle n_0 \rangle$, and therefore $N = (N \cap G) + \langle n_0 \rangle$. On the other hand, suppose $t_1x_1 + \cdots t_{r-1}x_{r-1} = sn_0$ belongs to $(N \cap G) \cap \langle n_0 \rangle$. Then

$$t_1x_1 + \cdots + t_{r-1}x_{r-1} = ss_1x_1 + \cdots + ss_{r-1}x_{r-1} + sax_r.$$

Since the x_j are linearly independent, this gives $sax_r = 0$, forcing $s = 0$. Thus, $sn_0 = 0$, showing $(N \cap G) \cap \langle n_0 \rangle = 0$, and thus, $N = (N \cap G) \oplus \langle n_0 \rangle$, as required.

For part (ii) Suppose $M = \langle x_1, \ldots, x_n \rangle$ and k is the largest integer such that k elements in the set $\{x_1, \ldots, x_n\}$ are linearly independent over R. Without loss of generality, we may assume these elements are $x_1, \ldots, x_k$. Set $F := \langle x_1, \ldots, x_k \rangle$, a free R-module. Then for each x_j with $k < j \leq n$, there exists $0 \neq a_j \in R$ such that $a_j x_j \in F$. Set $a := a_{k+1} \cdots a_n$. It follows that $aM \subseteq F$. By part (i), aM is a free R-module. However, since M is torsion-free, then map $\phi : M \to aM$ given by $\phi(m) = am$ is an isomorphism of R-modules. Thus, M is a free R-module. Note that the rank of M is k, since on the one hand, $\langle x_1, \ldots, x_k \rangle$ is a free module contained in M, so $k \leq \text{rank}(M)$, while on the other hand, $aM \subseteq \langle x_1, \ldots, x_k \rangle$, so $\text{rank}(aM) \leq k$. Since $M \cong aM$, we must have $\text{rank}(M) = k$.

For part (iii), Suppose $M = \langle x_1, \ldots, x_n \rangle$. Then we have a surjective homomorphism $\phi : R^n \to M$, which takes $(r_1, \ldots, r_n)$ to $r_1 x_1 + \cdots + r_n x_n$. Let N be a submodule of M. Then $\phi^{-1}(N)$ is a submodule of R^n. By part (i), $\phi^{-1}(N)$ is a finitely generated free R-module, whose rank is less than or equal to n. Thus, N is a homomorphic image of an R-module generated by n or fewer elements, and hence N is generated by n or fewer elements.

For part (iv) let's first observe that $M/T(M)$ is torsion-free. Indeed, suppose $r \cdot \overline{m} = \overline{0}$, for $\overline{m} \in M/T(M)$, and $0 \neq r \in R$. Then $rm \in T(M)$, so there exists $0 \neq r' \in R$ such that $r'(rm) = 0$. Thus, $(r'r)m = 0$. Since $r'r \neq 0$, $m \in T(m)$, so that $\overline{m} = \overline{0}$ in $M/T(M)$. Thus, $M/T(M)$ is torsion-free, and finitely generated. By part (ii), $M/T(M)$ is a finitely generated free R-module. Thus, the canonical map $M \to M/T(M)$ is a surjective map onto a finitely generated free R-module. Since the kernel of this map is $T(M)$, by problem 4 on Homework 1, there exists a free R-submodule $F \subseteq M$ such that $M = F \oplus T(M)$. □

Regarding part (iv) of Theorem B, $F \subseteq M$ may not be unique, as it depends upon the map $j : M/T(M) \to M$, but its rank is unique. To see this just note that if

$$F' \oplus T(M) = M = F \oplus T(M),$$

with F' a free R-module, then modding out $T(M)$, we have $F' \cong M/T(M) \cong F$, so that F' and F are isomorphic free modules and therefore have the same rank. We can then define the *rank of* M to be the rank of F. Note also, that when we write $M = F \oplus T(M)$, if $\{x_1, \ldots, x_r\}$ is a basis for F, then $F = Rx_1 \oplus \cdots \oplus Rx_r$ is a direct sum of cyclic R-modules. Thus, to finish the structure theorem for finitely generated modules over a PID, we only have to show that a finitely generated torsion module is a direct sum of cyclic modules.

We begin with a definition.

Definition C. Let R be a commutative ring and M an R-module. For $x \in M$ we define the *annihilator* of x to be the set $\text{ann}(x) := \{r \in R \mid rx = 0\}$. We set $\text{ann}(M) := \{r \in R \mid rx = 0, \text{ for all } x \in M\}$, the *annihilator of* M. Both annihilators are ideals of R. Note that M is torsion-free if and only if $\text{ann}(x) = 0$, for all x and M is a torsion module if and only if $\text{ann}(x) \neq 0$, for all $x \in M$.

Proposition D. *Let R be a* PID *and M a finitely generated, torsion R-module.*

(i) $\text{ann}(M) = aR \neq 0$, for some $a \in R$. In particular, a annihilates M and divides any element of R annihilating M.

(ii) If $\text{ann}(M) = aR$ and p is a prime dividing a, then $M(p) := \{x \in M \mid p^j x = 0 \text{ for some } j \geq 1\}$ is a non-zero submodule with $\text{ann}(M(p)) = p^e R$, for some $e \geq 1$.

(iii) Suppose $x, z \in M$ satisfy $\text{ann}(x) = p^e R = \text{ann}(z)$, for $p \in R$ a prime and $e \geq 1$. If $\langle x \rangle \subseteq \langle z \rangle$, then $\langle x \rangle = \langle z \rangle$.

(iv) Suppose $M = \langle x \rangle$ is a cyclic module with $\text{ann}(M) = p^e R$, with $p \in R$ prime and $e \geq 1$. Then for any non-zero submodule $N \subseteq M$, we have $N = \langle p^d x \rangle$ for some $0 \leq d < e$.

(v) Suppose M is annihilated by $p \in R$, p a prime element. Then M is a direct sum of cyclic submodules.

Proof. For (i), suppose M is generated over R by $x_1, \ldots, x_n$. For each i there exists $0 \neq a_i \in R$ such $a_i x_i = 0$. If we set $a_0 := a_1 \cdots a_n$, then we have that $a_0 x_i = 0$ for all i and hence $a_0 x = 0$, for all $x \in M$. Note that a_0 is a no-zero element in $\text{ann}(M)$. Since R is a PID, $\text{ann}(M)$ is principal, so we write $\text{ann}(M) = aR$, for $0 \neq a \in R$. Moreover, if $bx = 0$, for all $x \in M$, then $a \mid b$ (since $b \in aR$), which gives what we want.

(ii) We first note that $M(p) \neq 0$. Write $a = pa'$. Note that $a' \notin Ra = \text{ann}(M)$, since a' is not a multiple of a, so there exists $0 \neq z \in M$ such that $a'z \neq 0$. Thus, $a'z \in M(p)$. Suppose $x, y \in M(p)$ and $p^j x = 0 = p^k y$, then $p^{j+k}(x + y) = 0$. Moreover $p^j(rx) = 0$, for all $r \in R$. Therefore $M(p)$ is a submodule of M. By

Proposition 33.2, $M(p)$ is a finitely generated module. If $v_1, \ldots, v_c$ generate $M(p)$, let $e \geq 1$ be the least exponent such that $p^e v_j = 0$, for all j. Then, clearly, $p^e \in \text{ann}(M(p))$. Suppose $\text{ann}(M(p)) = cR$, for $c \in R$. Then, by Part (i), $c \mid p^e$. By the unique factorization property, $c = p^j$, for some $j \geq 1$. By definition, $j \geq e$, and thus $c \in p^e R$. Therefore, $\text{ann}(M(p)) = p^e R$.

For (iii), we have $x = rz$, for some $r \in R$. If $p \mid r$, then $p^{e-1}x = 0$, a contradiction (since p^{e-1} is not a multiple of p^e). Thus, $p \nmid r$, so we can write $1 = ur + vp^e$, with $u, v \in R$. We then have

$$ux = urz = z - vp^e z = z,$$

showing that $z \in \langle x \rangle$, which implies $\langle z \rangle = \langle x \rangle$.

For (iv), by Theorem B (iii), N is a cyclic module, say $N = \langle n \rangle$. Since $N \subseteq \langle x \rangle$, we have $n = rx$, for some $r \in R$. Since R has the unique factorization property, we may write $r = r_0 p^c$, for $r_0 \in R$ not divisible by p. Thus, $n = r_0 p^c x$. Since $n \neq 0$, we must have $0 \leq c < e$. On the one hand, we have $\langle n \rangle \subseteq \langle p^c x \rangle$. On the other hand, we may write $1 = ur_0 + vp^{e-c}$, since p does not divide r_0. Multiplying this equation by $p^c x$, and using the fact that $p^e x = 0$, we have, $p^c x = ur_0 p^c x$. Thus, $p^c x = un$, showing that $p^c x \in \langle n \rangle$, and thus $\langle p^c x \rangle \subseteq \langle n \rangle$, which gives $\langle p^c x \rangle = \langle n \rangle = N$, which is what we want.

For (v), let $n \geq 1$ be such that M can be generated by n elements and M cannot be generated by fewer than n elements. Suppose $M = \langle x_1, \ldots, x_n \rangle$. We show by induction on n that $M = \langle x_1 \rangle \oplus \cdots \oplus \langle x_n \rangle$. If $n = 1$, there is nothing to prove. Suppose $n > 1$. Set $M' := \langle x_1, \ldots, x_{n-1} \rangle$. Clearly M' cannot be generated by fewer than $n-1$ elements, otherwise these elements together with x_n would generated M, contradicting the choice of n. Thus, $M' = \langle x_1 \rangle \oplus \cdots \oplus \langle x_{n-1} \rangle$. It suffices to show $M = M' \oplus \langle x_n \rangle$. Clearly $M = M' + \langle x_n \rangle$. Suppose $z \in M' \cap \langle x_n \rangle$. If $z \neq 0$, $\langle z \rangle \subseteq \langle x_n \rangle$, so by part (iii), $\langle z \rangle = \langle x_n \rangle$, so $x_n \in \langle z \rangle \subseteq M'$, which is a contradiction - since this would imply $M = M'$. Therefore, $z = 0$ and thus, $M' \cap \langle x_n \rangle = 0$, which gives $M = M' \oplus \langle x_n \rangle$, as required. □

Remark E. Let M be an R-module over the PID R. Suppose $x \in T(M)$ and let $\text{ann}(x) = aR$. Then we have a surjective R-module map $\phi : R \to \langle x \rangle$ whose kernel is aR. Thus, $\langle x \rangle \cong R/aR.$, which looks exactly like the isomorphism we obtain when we have a cyclic group. When $a = p$ is prime, then $\langle x \rangle$ is the module analogue of a cyclic groups of order p: p kills its generator, and the group has no proper subgroups.

Proposition F. *Let R be a* PID *and M a finitely generated, torsion R-module. Suppose $\text{ann}(M) = aR$, and $a = p_1^{e_1} \cdots p_r^{e_r}$, for primes $p_i \in R$ and $e_i \geq 1$. Then*

(i) $M = M(p_1) \oplus \cdots \oplus M(p_r)$.
(ii) $\text{ann}(M(p_i)) = p_i^{e_i} R$.

Proof. For (i), set $s_i := \Pi_{j \neq i} p_j^{e_j}$, for $1 \leq i \leq r$. Since the GCD of the s_i equals 1, the ideal generated by the s_i is R. Thus, we may write $1 = t_1 s_1 + \cdots + t_r s_r$. For any $x \in M$, we have $x = (t_1 s_1 x) + \cdots + (t_r s_r x)$. Since $p_i^{e_i} \cdot (t_i s_i x) = 0$, each $t_i s_i x \in M(p_i)$. This shows that $M = M(p_1) + \cdots + M(p_r)$. On the other hand, by the previous proposition, each $M(p_i)$ is annihilated by a power of p_i, so we take α_i to be the least power of p_i annihilating $M(p_i)$. Then there exist $c, d \in R$ such $1 = cp_i^{\alpha_i} + du_i$, where $u_i = \Pi_{j \neq i} p_j^{\alpha_j}$. Now suppose $y \in M(p_i) \cap (\sum_{j \neq i} M(p_j))$. Then $y = (cp_i^{\alpha_i} y) + (du_i y)$. Since $y \in M(p_i)$, $cp_i^{\alpha_i} y = 0$, while on the other hand, $du_i y = 0$, since $y \in \sum_{j \neq i} M(p_j)$. Thus, $y = 0$, showing $M = M(p_1) \oplus \cdots \oplus M(p_r)$, as required. □

For (ii), by the previous paragraph, $\text{ann}(M(p_i)) = p_i^{\alpha_i} R$, where α_i to be the least power of p_i annihilating $M(p_i)$. Set $a' := p_i^{\alpha_1} \cdots p_r^{\alpha_r}$. Then a' annihilates M, since every element in M is a sum of elements from the $M(p_i)$. Thus, $a \mid a'$. It follows that each $\alpha_i \geq e_i$. Now suppose, for example, $\alpha_1 > e_1$. Then there exists a non-zero $x \in M(p_1)$ such that $p_1^{e_1} x \neq 0$. Since $p_1^{e_1} x$ is annihilated by $p_2^{e_2} \cdots p_r^{e_r}$, the first part of the proof shows that $p_1^{e_1} x$ belongs to $M(p_2) + \cdots + M(p_r)$, contradicting the directness of the sum in part (i). Thus, $p_1^{e_1}$ annihilates $M(p_1)$ and we have $\alpha_1 = e_1$. Similarly, $\alpha_j = e_j$ for $2 \leq j \leq r$. □

In Proposition D (v) we showed that if M is a finitely generated module over a PID having the property that M is annihilated by a prime element, then M is a direct sum of cyclic modules. Theorem G below extends this to modules annihilated by a power of a prime element, and is the most difficult part of the structure theorem for modules over a PID.

Theorem G. *Let R be a PID and M a finitely generated R-module with $ann(M) = p^eR$, where $p \in R$ is prime and $e \geq 1$. Then M is a direct sum of cyclic modules. In fact, there exist $x_1, \ldots, x_n \in R$ and $e = e_1 \geq \cdots \geq e_n$ such that $M = \langle x_1 \rangle \oplus \cdots \oplus \langle x_n \rangle$, with $ann(x_i) = p^{e_i}R$, for all i.*

Proof. We begin by noting that there exists $0 \neq x \in M$ such that $p^{e-1}x \neq 0$. Otherwise, p^{e-1} annihilates every x in M, and thus is divisible by p^e, which cannot happen. So we start with $0 \neq x$ such that $p^{e-1}x \neq 0$, i.e., $\text{ann}(x) = \text{ann}(M)$. If $M = \langle x \rangle$, we are done. If $M \neq \langle x \rangle$, we set $x_1 := x$ and claim there is a submodule M_1 such that $M = \langle x_1 \rangle \oplus M_1$. Suppose we could always find such an M_1 whenever a cyclic submodule has the same annihilator as the module. Then, taking $x_2 \in M_1$ so that $\text{ann}(x_2) = \text{ann}(M_1)$, either $M_1 = \langle x_2 \rangle$, and thus, $M = \langle x_1 \rangle \oplus \langle x_2 \rangle$ or there exists $M_2 \subseteq M_1$ such that $M_1 = \langle x_2 \rangle \oplus M_2$, so that $M = \langle x_1 \rangle \oplus \langle x_2 \rangle \oplus M_2$. If we apply the construction inductively, then we have a chain of submodules,

$$\langle x_1 \rangle \subseteq \langle x_1 \rangle \oplus \langle x_2 \rangle \subseteq \langle x_1 \rangle \oplus \langle x_2 \rangle \oplus \langle x_3 \rangle \subseteq \cdots .$$

Since M satisfies the ascending chain condition, this process must stop when M is a direct sum of cyclic submodules. For the statement about annihilators, first note that since p^eR is in the annihilator of every element and submodule of M, the annihilator of every element and submodule divides p^e and is thus generated by a power of p. Moreover, since $M_{i+1} \subseteq M_i$, $\text{ann}(M_i) \subseteq \text{ann}(M_{i+1})$, and therefore, if $\text{ann}(M_i) = p^{e_i}R$ and $\text{ann}(M_{i+1}) = p^{e_{i+1}}R$, $e_i \geq e_{i+1}$. Since $\text{ann}(x_i) = \text{ann}(M_i)$, the statement concerning annihilators follows.

Thus, we must prove the following statement: If M is a finitely generated R-module with $\text{ann}(M) = p^eR$ and $x \in M$ satisfies $\text{ann}(x) = p^eR$, then there exists a submodule $K \subseteq M$ such that $M = \langle x \rangle \oplus K$. To see this, we will show that there exists an R-module homomorphism $\alpha : M \to \langle x \rangle$ that is the identity on $\langle x \rangle$. Suppose α exists. Set K to be the kernel of α. Let $m \in M$. Then $\alpha(m) \in Rx$, and hence $\alpha(\alpha(m)) = \alpha(m)$. Thus, $\alpha(m - \alpha(m)) = \alpha(m) - \alpha(\alpha(m)) = 0$, so that $m - \alpha(m) \in K$. Thus, $m \subseteq \langle x \rangle + K$, since $\alpha(m) \in \langle x \rangle$, by definition. Therefore, $M = \langle x \rangle + K$. Suppose $rx \in K$ belongs to $\langle x \rangle \cap K$. Then $rx = \alpha(rx) = 0$. Thus, $\langle x \rangle \cap K = 0$, showing $M = \langle x \rangle \oplus K$.

To find $\alpha : M \to \langle x \rangle$ which is the identity on $\langle x \rangle$, let $\mathcal{C}$ denote the collection of submodules $N \subseteq M$ containing $\langle x \rangle$ for which there exists a homomorphism $\gamma : N \to \langle x \rangle$ which is the identity on $\langle x \rangle$. Note that $\langle x \rangle$ belongs to $\mathcal{C}$ by just taking the identity map on $\langle x \rangle$, so $\mathcal{C}$ is not empty. Then $\mathcal{C}$ has a maximal element, say N, together with a homomorphism $\alpha : N \to \langle x \rangle$, which is the identity on $\langle x \rangle$. We claim $N = M$. If so, then we are done. Suppose not. Take $m \in M \backslash N$ such that $pm \in N$. Then $p^em = 0$, so that $p^{e-1}\alpha(pm) = 0$. Now, since $\alpha(pm) \in \langle x \rangle$, we may write $\alpha(pm) = rx$, for some $r \in R$. Thus, $0 = p^{e-1}(rx) = (p^{e-1}r)x$, so $p^{e-1}r \in \text{ann}(x) = p^eR$. Thus, $p^{e-1}r$ is divisible by p^e, so r is divisible by p. Thus, we may write $r = r_0p$ and therefore $\alpha(pm) = p(r_0x)$ [1]. Set $z := r_0x$, so that

$$\alpha(pm) = pz \in Rx.$$

We now define $\gamma : N + \langle m \rangle \to \langle x \rangle$ as follows: $\gamma(n + rm) = \alpha(n) + rz$, for all $n \in N$ and $r \in R$. If γ is well defined, then the fact that γ extends α and $N + \langle m \rangle$ is strictly larger that N contradicts the maximality of N. Thus, we must have $N = M$, which gives what we want.

Finally, to see that γ is well defined, suppose that $n + rm = n' + r'm$, for $r, r' \in R$ and $n, n' \in N$. Then $(r' - r)m \in N$. Set $J := \{s \in R \mid sm \in N\}$. This is a principal ideal, so $J = aR$, say. Since $p \in J$, $p \in aR$. Thus, $a \mid p$. This can only happen if a is a unit multiple of p, and hence $aR = pR$. Therefore, $r' - r = tp$,

[1] Note: Here is where we are using the fact that $\text{ann}(x) = \text{ann}(M)$. If $\text{ann}(x) = p^cR$, with $c < e$, and the only power of p annihilating m is e, then the equation $(p^{e-1}r)x = 0$ does not tell us anything, since $p^{e-1}x$ is already 0.

so that $r' = r + tp$. Thus, $n + rm = n' + (r + tp)m$ so that $n = n' + tpm$. Therefore

$$\begin{aligned}\gamma(n + rm) &= \gamma((n' + tpm) + rm)\\ &= \alpha(n' + tpm) + rz\\ &= \alpha(n') + t\alpha(pm) + rz\\ &= \alpha(n') + tpz + rz\\ &= \alpha(n') + (r + tp)z\\ &= \alpha(n') + r'z\\ &= \gamma(n' + rm).\end{aligned}$$

This shows that γ is well defined and provides the contradiction we sought. This completes the proof of the theorem. □

We now have all of the ingredients for the structure theorem for finitely generated torsion modules over a PID. Note, then, that this yields a structure theorem for finite abelian groups.

Structure Theorem for Torsion modules over a PID. *Let R be a* PID *and M a finitely generated torsion module over R, i.e., $M = T(M)$. Then M is a direct sum of cyclic modules. In fact, if we write $ann(M) = aR$ and $a = p_1^{e_1} \cdots p_r^{e_r}$, with each $p_i \in R$ prime and $e_i \geq 1$, then there exist integers $n_1, \ldots, n_r \geq 1$ and for each $1 \leq i \leq r$, integers $e_i = e_{i,1} \geq \cdots \geq e_{i,n_i}$ and elements $x_{i,1}, \ldots x_{i,n_i} \in M$ such that $ann(x_{i,j}) = p_i^{e_{i,j}} R$ and*

$$M = \langle x_{1,1}\rangle \oplus \cdots \oplus \langle x_{1,n_1}\rangle \oplus \cdots \oplus \langle x_{r,1}\rangle \oplus \cdots \oplus \langle x_{r,n_r}\rangle.$$

Moreover, we have

$$M \cong (R/p_1^{e_{1,1}}R) \oplus \cdots \oplus (R/p_1^{e_{1,n_1}}R) \oplus \cdots \oplus (R/p_r^{e_{r,1}}R) \oplus \cdots \oplus (R/p_r^{e_{r,n_r}}R).$$

This second decomposition may be thought of as an external direct sum.

Proof. The first statement is immediate from Proposition F and Theorem G. The second statement follows from the first and the fact that the homomorphism $\phi_{i,j} : R \to \langle x_{i,j}\rangle$ defined by $\phi(r) = rx_{i,j}$ is a surjective R-module homomorphism with kernel $p_i^{e_{i,j}}R$. □

For a finite abelian group G, one says that G has *index* $a \geq 2$ if G has elements of order a and the order of every element of G is divisible by a. This is the same thing as saying that, when we view G as a $\mathbb{Z}$-module, $\text{ann}(G) = a\mathbb{Z}$. Thus, the following corollary is immediate from the structure theorem above.

Corollary H. Let G be a finite abelian group of index $a = p_1^{e_1} \cdots p_r^{e_r}$, where $p_1, \ldots, p_r$ are prime. Then, for each $1 \leq i \leq r$, there exists $e_i = e_{i1} \geq \cdots \geq e_{is_i}$ such that

$$G \cong (\mathbb{Z}_{p_1^{e_{1,1}}} \oplus \cdots \oplus \mathbb{Z}_{p_1^{e_{1,n_1}}}) \oplus \cdots \oplus (\mathbb{Z}_{p_r^{e_{r,1}}} \oplus \cdots \oplus \mathbb{Z}_{p_r^{e_{r,n_r}}}).$$

Putting together Theorem B with the structure theorem for torsion modules, we obtain the full structure theorem for finitely generated modules over a PID.

Structure theorem for finitely generated modules over a PID. *Let R be a* PID *and M be a finitely generated module over R. Then their exists unique integers $s, r \geq 0, \geq 0$, unique primes $p_1, \ldots, p_r \in R$ and unique positive integers $e_{i,1} \geq \cdots \geq e_{i,n_i}$, for $1 \leq i \leq r$, such that*

$$M \cong R^s \oplus (R/p_1^{e_{1,1}}R) \oplus \cdots \oplus (R/p_1^{e_{1,n_1}}R) \oplus \cdots \oplus (R/p_r^{e_{r,1}}R) \oplus \cdots \oplus (R/p_r^{e_{r,n_r}}R).$$

Proof. By Theorem B, and the remark following it, $M = F \oplus T(M)$, where F is a free R-module whose rank depends only on M. Thus, if the rank of F is s, $F \cong R^s$. By the structure theorem for torsion modules, if we write $\text{ann}(T(M)) = aR$, and $a = p_1^{e_1} \cdots p_r^{e_r}$ with each $p_i \in R$ prime and $e_i \geq 1$, then

$$T(M) \cong (R/p_1^{e_{1,1}}R) \oplus \cdots \oplus (R/p_1^{e_{1,n_1}}R) \oplus \cdots \oplus (R/p_r^{e_{r,1}}R) \oplus \cdots \oplus (R/p_r^{e_{r,n_r}}R),$$

for unique $\{e_{i,j}\}$, as prescribed. □

Remark I. In the statement of the full structure theorem, we have chosen to express M isomorphic to the external direct sum of cyclic modules. Of course, given Theorem B and the structure theorem for finitely generated torsion modules, it is clear that M is also the internal direct sum of a free module and cyclic

submodules as described in the structure theorem for finitely generated torsion modules. Note that it can be the case that $s = 0$, in which case M is a torsion module, or $r = 0$, in which case M is a free R-module. The ideals $p_i^{e_{ij}}$ appearing in the structure theorems are called the *elementary divisors* of M, and are uniquely determined by M, as the following proposition shows.

Proposition J. *Let R be a PID and A, B finitely generated R-modules. Then $A \cong B$ if and only if A and B have the same rank and same elementary divisors.*

Here, by the rank of A (or B), we mean the rank of any free submodule $F \subseteq A$ such that $A = F \oplus T(A)$. By the comment following the proof of Theorem B, the rank of A is well defined.

Proof of Proposition J. It is easy to see that if A and B have the same rank and elementary divisors, then they are isomorphic. Conversely, suppose $\phi : A \to B$ is an R-module homomorphism. Write $A = F \oplus T(A)$ and $B = G \oplus T(B)$, where $F \subseteq A$ and $G \subseteq G$ are free submodules. We first note that $\phi(T(A)) = T(B)$. Suppose $a \in T(A)$ and $0 \neq c$ satisfies $ca = 0$. Then $c\phi(a) = 0$, showing $\phi(a) \in T(B)$, whence ϕ maps $T(A)$ to $T(B)$. Given $b \in T(B)$ there exists $0 \neq d \in R$ such that $db = 0$. Since ϕ is surjective, $b = \phi(a')$, for some $a' \in A$. Thus, $\phi(da') = 0$. Since ϕ is injective $da' = 0$, showing $a' \in T(A)$. Thus, $b \in \phi(T(A))$. Moreover, we clearly have that ϕ restricted to $T(A)$ is an isomorphism between $T(A)$ and $T(B)$. In addition, ϕ induces an isomorphism $\overline{\phi} : A/T(A) \to B/T(B)$. This shows that F and G are isomorphic and hence A and B have the same rank.

Thus we may begin again, and assume that A and B are torsion R-modules. Since $A \cong B$, they have the same annihilator. Set $\mathrm{ann}(A) := aR =: \mathrm{ann}(B)$ and write $a = p_1^{e_1} \cdots p_r^{e_r}$. Then $A = A(p_1) \oplus \cdots \oplus A(p_r)$ and $B = B(p_1) \oplus \cdots \oplus B(p_r)$, and the same proof as in the previous paragraph shows that for each $1 \leq i \leq r$, ϕ restricted to $A(p_i)$ gives an isomorphism from $A(p_i)$ to $B(p_i)$. Thus we may begin once again assuming $\mathrm{ann}(A) = p^e R = \mathrm{ann}(B)$, for some prime $p \in R$ and $e \geq 1$.

Let us write $A = \langle x_1 \rangle \oplus \cdots \oplus \langle x_n \rangle$, with $\mathrm{ann}(x_i) = p^{e_i} R$, for $1 \leq i \leq n$ and $e = e_1 \geq \cdots \geq e_n$. Similarly, we can write $B = \langle y_1 \rangle \oplus \cdots \oplus \langle y_s \rangle$, with $\mathrm{ann}(y_i) = p^{f_i} R$, for $1 \leq i \leq s$ and $e = f_1 \geq \cdots \geq f_s$. We first show that $n = s$.

Consider the vector spaces A/pA and B/pB over $\overline{R} := R/pR$. Since $\phi(pA) = pB$, we have an induced isomorphism of vector spaces

$$\hat{\phi} : \langle x_1 \rangle / p\langle x_1 \rangle \oplus \cdots \oplus \langle x_n \rangle / p\langle x_n \rangle = A/pA \to B/pB = \langle y_1 \rangle / p\langle y_1 \rangle \oplus \cdots \oplus \langle y_s \rangle / p\langle y_s \rangle.$$

However, each $\langle x_i \rangle / p\langle x_i \rangle$ is isomorphic to R/pR and similarly for each $\langle y_j \rangle / p\langle y_j \rangle$. Thus, A/pA is an n-dimensional vector space over $\overline{R}$ and B/pB is s-dimensional. Therefore, $n = s$.

We now proceed by induction on e to see that $e_i = f_i$, for all $1 \leq i \leq n$. The case $e = 1$ is clear, since $e = e_1 \geq \cdots \geq e_n$, and similarly for the f_j. Suppose $e > 1$, but $e_{i+1} = \cdots e_n = 1$ and $f_{j+1} = \cdots = f_n = 1$. Now, as before, we have pA and pB are isomorphic, since $\phi(pA) = pB$. Moreover,

$$pA = p\langle x_1 \rangle \oplus \cdots \oplus p\langle x_i \rangle = \langle px_1 \rangle \oplus \cdots \oplus \langle px_i \rangle$$

and

$$pB = p\langle y_1 \rangle \oplus \cdots \oplus p\langle y_\rangle = \langle py_1 \rangle \oplus \cdots \oplus \langle py_j \rangle.$$

Now, $\mathrm{ann}(pA) = p^{e-1} R = \mathrm{ann}(pB)$, so by induction, $i = j$, which then gives $e_c = f_c = 1$, for $i + 1 \leq c \leq n$. In addition, $\mathrm{ann}(px_s) = p^{e_s} R$ and $\mathrm{ann}(py_s) = p^{f_s - 1} R$, for $1 \leq s \leq i$. The induction hypothesis then gives $e_s - 1 = f_s - 1$, and hence $e_s = f_s$, for $1 \leq s \leq i$, which completes the proof. □

There is another standard way to present the structure theorem for finitely generated modules over a principal ideal domain, in which the decomposition of $T(M)$ takes a different form. Suppose we write

$$T(M) \cong (R/p_1^{e_{1,1}} R) \oplus \cdots \oplus (R/p_1^{e_{1,n_1}} R) \oplus \cdots \oplus (R/p_r^{e_{r,1}} R) \oplus \cdots \oplus (R/p_r^{e_{r,n_r}} R),$$

as in the structure theorem. One can use the *Chinese Remainder Theorem*, CRT for short, for PIDs, to rearrange the terms in the decomposition above. Recall, that for the integers, the CRT states that if $n, m \in \mathbb{Z}$

are relatively prime, then $\mathbb{Z}_{nm} \cong \mathbb{Z}_n \oplus \mathbb{Z}_m$.[2] Since GCDs and Bezout's Principle hold in R, given $r, s \in R$ with GCD equal to 1, we have $R/rsR \cong R/rR \oplus R/sR$. This isomorphism easily extends to finitely many elements that are pairwise relatively prime. Thus, if we set $a_1 := p_1^{e_{1,1}} \cdots p_r^{e_{r,1}}$, we have that

$$R/p_1^{e_{1,2}}R \oplus \cdots \oplus R/p_r^{e_{r,1}}R \cong R/a_1R.$$

Now set $a_2 := p_1^{e_{1,2}} \cdots p_2^{e_{r,2}}$. We then have $a_2|a_1$ and

$$R/p_1^{e_{1,2}}R \oplus \cdots \oplus R/p_r^{e_{r,2}}R \cong R/a_2R,$$

and hence

$$R/p_1^{e_{1,2}}R \oplus \cdots \oplus R/p_r^{e_{r,1}}R \oplus R/p_1^{e_{1,2}}R \oplus \cdots \oplus R/p_r^{e_{r,2}}R \cong R/a_1R \oplus R/a_2R.$$

Continuing in this fashion, we have the existence of $a_1, \ldots, a_d \in R$ such that $a_d \mid a_{d-1} \mid \cdots \mid a_1$ and

$$M \cong R^s \oplus R/a_1R \oplus R/a_2R \oplus \cdots \oplus R/a_dR.$$

The ideals $a_1R, \ldots a_dR$ are called the *invariant factors of* R and are uniquely determined by M.

Application to Linear Algebra. Let V be a finite dimensional vector space over the field F and $T : V \to V$ a linear transformation. Then V becomes an $F[x]$-module under the action of T: For $p(x) \in F[x]$ and $v \in V$, one defines $p(x) \cdot v$ to be $p(T)(v)$. It is easy to check that this makes V into a finitely generated $F[x]$-module. Since $F[x]$ is a PID, we may apply the structure theorem for finitely generated modules over a PID. Let us first note that V is a torsion $F[x]$-module. For this, set $d := \dim(V)$ and take $0 \neq v \in V$. Then $v, T(v), \cdots, T^d(v)$ are linearly dependent vectors, and thus, there exist $\alpha_0, \ldots, \alpha_d \in F$ such that $\alpha_0 v + \alpha_1 T(v) + \cdots + \alpha_d T^d(v) = 0$. Setting $p(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_d x^d$, this yields $p(T)(v) = 0$, or in module notation, $p(x)v = 0$. Since V is a finitely generated torsion module over $F[x]$, its annihilator is non-zero. If we set $q(x)R := \text{ann}(V)$, then $q(T) = 0$, since $q(T)(v) = 0$, for all $v \in V$. Note that if $p(x) \in F[x]$ and $p(T) = 0$, then $p(x)$ annihilates V and thus $q(x) \mid p(x)$. For this reason, $q(x)$ is called the *minimal polynomial of* T.

Now take $v \in V$. Then $\text{ann}(v)$ is easily seen to be generated by the polynomial $q_v(x)$ of least degree such that $q_v(T)(v) = 0$, and moreover $q_v(x)$ divides any polynomial $p(x)$ such that $p(T)(v) = 0$. What is $\langle v \rangle$, the cyclic submodule of V generated by v ? By definition it is $\{f(x)v \mid f(x) \in F[x]\} = \{f(T)(v) \mid f(x) \in F[x]\}$. Let $c := \deg(q_v(x))$. Then, for any $p(x) \in F[x]$ with degree greater than or equal to c, we can write $p(x) = h(x)q_v(x) + r(x)$, with $\deg r(x) < c$, so that $p(T)(v) = h(T)q_v(T)(v) + r(T)(v) = r(T)(v)$. It follows easily from this, that $B := \{v, T(v), \ldots, T^{c-1}(v)\}$ forms a basis for $\langle v \rangle$, when we think of $\langle v \rangle$ as a vector space over F. [3] Note that by definition, $\langle v \rangle$ is invariant under T, i.e., $T|_{\langle v \rangle}$ is a linear transformation from $\langle v \rangle$ to itself. What is the matrix of $T|_{\langle v \rangle}$ with respect to the basis B? To find this, write $q_v(x) = x^c + \alpha_1 x^{c-1} + \cdots + \alpha_c$. Then we have:

$$\begin{aligned}
T(v) &= 0 \cdot v + 1 \cdot T(v) + 0 \cdot T^2(c) + 0 \cdot T^3(v) + \cdots + 0 \cdot T^{c-1}(v) \\
T(T(v)) &= 0 \cdot v + 0 \cdot T(v) + 1 \cdot T^2(v) + 0 \cdot T^3(v) + \cdots + 0 \cdot T^{c-1}(v) \\
\vdots &= \vdots \\
T(T^{c-1}(v)) &= -\alpha_c \cdot v + \alpha_{c-1} \cdot T(v) + \alpha_{c-2} \cdot T^2(v) - \cdots - \alpha_1 \cdot T^{c-1}(v),
\end{aligned}$$

the last inequality following from $q_v(T)(v) = 0$. Thus, the matrix of $T|_{\langle v \rangle}$ with respect to B is

$$\begin{pmatrix}
0 & 0 & 0 & \cdots & -a_0 \\
1 & 0 & 0 & \cdots & -\alpha_1 \\
0 & 1 & 0 & \cdots & -\alpha_2 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & -\alpha_{c-1}
\end{pmatrix},$$

[2] To so this, let $\phi : \mathbb{Z} \to \mathbb{Z}_n \oplus \mathbb{Z}_m$ be the map that sends $a \in \mathbb{Z}$ to $(\overline{a}, \overline{a}) \in \mathbb{Z}_n \oplus \mathbb{Z}_m$. Since n, m are relatively prime, there exist $r \in n\mathbb{Z}$ and $s \in m\mathbb{Z}$ such that $1 = r + s$. Thus, $\overline{1} = \overline{s}$ in $\mathbb{Z}_n$ and $\overline{1} = \overline{r}$ in $\mathbb{Z}_m$. Given $(\overline{a}, \overline{b}) \in \mathbb{Z}_n \oplus \mathbb{Z}_m$, we have $\phi(sa + rb) = (\overline{sa + rb}, \overline{sa + rb}) = (\overline{sa}, \overline{rb}) = (\overline{a}, \overline{b})$, showing ϕ is surjective. It is easy to see that the kernel of ϕ is $nm\mathbb{Z}$, so the required isomorphism holds.

[3] These vectors clearly span $\langle v \rangle$ as a vector space over F, and any non-trivial dependence relation among them would yield a polynomial $g(x)$ of degree less than c such that $g(T)(v) = 0$.

the *Companion Matrix of* $q_v(x)$, denoted $C(q_v(x))$. Note that this $c \times c$ matrix depends only on $q_v(x)$, so that given any $f(x) \in F[x]$ we may define its companion matrix $C(f(x))$ analogously.

We now state the *Rational Canonical Form Theorem* for T and note that it is really just a restatement of the structure theorem for finitely generated torsion modules over a PID in the case that $F[x]$ is the ring, V is the module and the ring action is defined by $T : V \to V$.

Rational Canonical Form via elementary divisors. *Let V be a finite dimensional vector space over the field F and $T : V \to V$ a linear transformation. Factor the minimal polynomial of T as $q(x) = p_1(x)^{e_1} \cdots p_r(x)^{e_r}$, with each $p_i(x)$ irreducible over F. Then V is a direct sum of cyclic subspaces. In particular, for each $1 \leq i \leq r$ there exist positive integers $e_i = e_{i,1} \geq \cdots \geq e_{i,n_i}$, and a basis $\mathcal{B}$ for V such that A, the matrix of T with respect to $\mathcal{B}$, has the the block diagonal form*

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix},$$

and for each $1 \leq i \leq r$,

$$A_i = \begin{pmatrix} C(p_i(x)^{e_{i,1}}) & 0 & \cdots & 0 \\ 0 & C(p_i(x)^{e_{i,2}}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C(p_i(x)^{e_{i,n_i}}) \end{pmatrix}. \quad \square$$

We note that from the module perspective, $V = V(p_1(x)) \oplus \cdots \oplus V(p_r(x))$. Each $V(p_i(x))$ is a direct sum of cyclic subspaces that yield a companion matrix of the form $C(p_i(x)^{e_{i,j}})$, and each A_i is the matrix of $T|_{V(p_i(x))}$ with respect to the union of the cyclic bases from each of those cyclic subspaces.

An alternate approach to Theorem G. The proof of Theorem G above conceptually follows the same approach to the Rational Canonical Form Theorem taken in the Fall Math 790 class, in the following sense: If M is a finitely generated module over the PID R with $p \in R$ prime, and $\text{ann}(M) = p^e R$, then this corresponds to the case that the linear operator T on the finite dimensional vector space V over the field F has minimal polynomial $p(x)^e \in F[x]$, with $p(x)$ irreducible over F. Taking $x \in M$ with $\text{ann}(x) = \text{ann}(M)$ corresponds to taking a maximal vector $v \in V$. In the module case, we seek a summand K such that $M = \langle x \rangle \oplus K$ and in the linear algebra case we seek a T-invariant complement of $\langle v, T \rangle$, the cyclic subspace of V determined by v and T. In each case, induction finishes the proof. The approach below to Theorem G relies more on a commutative algebra perspective, so we begin with a standard rseult.

Nakayama's Lemma. *Let R be a commutative ring, M a finitely generated R-module, and $J \subseteq R$ the Jacobson radical of R. Suppose $N \subseteq M$ is a submodule and $M = N + JM$. Then $N = M$.*

Proof. We first note that $M/JM = (JM + N)/N = J(M/N)$, so we have a finitely generated R-module $A := M/N$ satisfying $A = JA$. If we show $A = 0$, then $M = N$. Starting again, suppose n is the least number of non-zero elements required to generate A and $A = \langle x_1, \ldots, x_n \rangle$. Then $x_1 \in M = JM$, so we can write $x_1 = j_1 x_1 + \cdots + j_n x_n$, for some $j_i \in J$. Thus, $(1 - j_1)x_1 = j_2 x_2 + \cdots + j_n x_n$. Since $1 - j_1$ is a unit, we have $x_1 \in \langle x_2, \ldots, x_n \rangle$, which implies $A = \langle x_2, \ldots, x_n \rangle$, contradicting the minimality of n. Thus, $A = 0$, and $M = N$. $\square$

Corollary J. *Suppose R has a unique maximal ideal P and M is a finitely generated R-module. Set $\tilde{R} := R/P$ and $\tilde{M} := M/PM$, so that $\tilde{M}$ is a finite dimensional vector space over $\tilde{R}$. Take $x_1, \ldots, x_n$ in M. Then $x_1, \ldots, x_n$ is a minimal generating set for M, ie., $x_1, \ldots, x_n$ generate M, but no subset of the x_j generates M, if and only if $\tilde{x_1}, \ldots, \tilde{x_n}$ forms a basis for $\tilde{M}$.*

Here, we are writing $\tilde{x_j}$ for the class of x_j in $\tilde{M}$. It follows that every minimal generating set for M has the same number of elements, namely the dimension of the vector space $\tilde{M}$ over $\tilde{R}$.

Proof. First assume that $x_1, \ldots, x_n$ is a minimal generating set for M. We clearly have that $\tilde{x_1}, \ldots, \tilde{x_s}$ span $\tilde{M}$. Suppose we have $\tilde{r_1}\tilde{x_1} + \cdots + \tilde{r_n}\tilde{x_n} \equiv 0$ in $\tilde{M}$. We want each $\tilde{r_j} \equiv 0$ in $\tilde{R}$, in other words, in R, we

should have $r_j \in P$, for all j. In M, we have $r_1x_1 + \cdots + r_nx_n = z$, for some $z \in PM$. Writing $z = \sum_j t_jx_j$, with each $t_i \in P$, we have $\sum_{i=1}^{s}(r_i - t_i)x_i = 0$. Suppose, for example, $\tilde{r_1} \not\equiv 0$ in $\tilde{R}$. Then $r_1 \notin P$, and thus, $r_1 - t_1 \notin P$, so that $r_1 - t_1$ is a unit in R. From the equation $\sum_i(r_i - t_i)x_i = 0$, it follows that x_1 is in the submodule of M generated by $x_2, \ldots, x_n$. This gives $M = \langle x_2, \ldots, x_n\rangle$, contradicting the minimality assumption. It follows, that $\tilde{r_1} \equiv 0$ in $\tilde{R}$, and similarly, $\tilde{r_i} \equiv 0$, for all i, showing that that $\{\tilde{x_1}, \ldots, \tilde{x_n}\}$ is a basis for $\tilde{M}$.

Conversely, suppose $\tilde{x_1}, \ldots, \tilde{x_s}$ is a basis for $\tilde{M}$. Set $N := \langle x_1, \ldots, x_n\rangle$. Since the $\tilde{x_j}$ span $\tilde{M}$, we have $M/PM = (N + PM)/PM$, so $M = N + PM$. Thus, by Nakayama's lemma, $N = M$, i.e., $x_1, \ldots, x_n$ generate M. Suppose this generating set is not minimal, say $x_1 = r_2x_2 + \cdots + r_nx_n$, for some $r_i \in R$. In $\tilde{M}$ we have $\tilde{x_1} \equiv \tilde{r_2}\tilde{x_2} + \cdots + \tilde{r_n}x_n$, contradicting the linear independence of the $\tilde{x_j}$ Thus, $x_1, \ldots, x_n$ is a minimal generating set for M. □

For the remainder of this note, we assume that R is a PID, and M is a finitely generated R-module with $\mathrm{ann}(M) = p^eR$, for $p \in R$ prime. We begin with a couple of observation regarding M.

Remarks K. (i) From Homework 1, we have that M is also an $\overline{R} := R/p^eR$-module, and moreover, for any residue class $\overline{r} \in \overline{R}$, $\overline{r}x = rx$, for all $x \in M$.

(ii) By the correspondence theorem between ideals of R and $\overline{R}$, it is easily seen that $\overline{R}$ has just one maximal ideal, namely $p\overline{R}$. Since the action of R on M is the same as the action of $\overline{R}$ on M, it follows from Corollary J that $x_1 \ldots, x_n \in M$ is a minimal generating set for M if and only if their images in M/pM form a basis for M/pM over the field R/pR. Thus, the number of elements in a minimal generating set for M as an R-module is well defined.

(iii) If S is an arbitrary commutative ring, and A is an S-module with submodules $B_1, \ldots, B_r$ satisfying $A = B_1 + \cdots + B_r$, it is straight forward to check that $A = B_1 \oplus \cdots \oplus B_r$ if and only if whenever $b_1 + \cdots + b_r = 0$, for $b_i \in B_i$, then $b_i = 0$, for all i. In particular, if each $B_i = \langle x_i\rangle$, then $A = B_1 \oplus \cdots \oplus B_r$ if and only if whenever $s_1x_1 + \cdots + s_rx_r = 0$, each $s_ix_i = 0$.

The next lemma is reminiscent of the proof of Cauchy's theorem for abelian groups.

Lemma L. *Let S be an integral domain, L an S-module, $x \in L$ such that $ann(L) = aS = ann(x)$. Set $\overline{L} := L/\langle x\rangle$. Suppose $z \in L$ satisfies $ann(\overline{z}) = bS$. Then there exists $t \in L$ such that $\overline{t} = \overline{z}$ in $\overline{L}$ and $ann(t) = bS$.*

Proof. It is enough to find $t \in L$ such that $bt = 0$ and $\overline{t} = \overline{z}$, for then $bS \subseteq \mathrm{ann}(t)$. On the other hand, for $r \in S$, $rt = 0$ implies $r\overline{t} = r\overline{b} \equiv 0$ in $\overline{L}$, so $r \in bR$. Thus, $\mathrm{ann}(t) = bR$.

Now, $bz = fx$, for some $f \in S$, and since $az = 0$, $a\overline{z} \equiv 0$, so $a = \gamma b$, for some $\gamma \in S$. Thus, we have, $0 = az = \gamma bz = \gamma fx$, which implies $\gamma f \in \mathrm{ann}(x) = aS$. Thus, $\gamma f = \tau a$, for some $\tau \in S$, and thus, $\gamma f = \tau\gamma b$, showing that $f = \tau b$. Set $t := z - \tau x$. Then $\overline{t} = \overline{z}$ in $\overline{L}$. Moreover,

$$bt = b(z - \tau x) = bz - b\tau x = bz - fx = 0,$$

so $b \in \mathrm{ann}(t)$, as required. □

Theorem G Revisited. *Let R be a PID and M a finitely generated R-module with $ann(M) = p^eR$, where $p \in R$ is prime and $e \geq 1$. Then M is a direct sum of cyclic modules. In fact, there exist $x_1, \ldots, x_n \in R$ and $e = e_1 \geq \cdots \geq e_n$ such that $M = \langle x_1\rangle \oplus \cdots \oplus \langle x_n\rangle$, with $ann(x_i) = p^{e_i}R$, for all i.*

Proof. Let n denote the minimal number of generators of M, which is well defined by Remark K(ii) above. We induct on n to show that M is the direct sum of n cyclic submodules with the required annihilators, the case $n = 1$ being trivial. Now suppose $n > 1$. Take $0 \neq x \in M$ such that $p^{e-1}x \neq 0$. Such an x exists, otherwise $p^{e-1} \in \mathrm{ann}(M) = p^eR$, which cannot happen. Thus, $p^ex = 0$, and it follows that $p^e \in \mathrm{ann}(x)$. On the other hand, $\mathrm{ann}(x) = cR$, for some $c \in R$, so we can write $p^e = rc$, for $r \in R$. Unique factorization implies that c must be a unit times p^i, for some i, and this forces c to be a unit multiple of p^e. Thus, $\mathrm{ann}(x) = p^eR = \mathrm{ann}(M)$.

We now note that x can be extended to a minimal generating set for M. Since $\mathrm{ann}(x) = p^eR$, we cannot have $x \in pM$, otherwise $p^{e-1}x = 0$. Thus, the image of x in the R/pR vector space M/pM is non-zero.

It can therefore be extended to a basis of M/pM. The pre-images of these basis elements in M form a minimal generating set for M as a module over R/p^eR, by Corollary J, and hence they form a minimal generating set of M as a module over R. Let us write $x, y_2, \ldots, y_n$ for this minimal generating set. Set $\overline{M} := M/\langle x \rangle$. Then $\overline{M}$ is a finitely generated R-module over R and since $p^eM = 0$, $p^e\overline{M} = 0$, and this forces $\operatorname{ann}(\overline{M}) = p^f$, for some $1 \leq f \leq e$. Now, $\overline{M}$ is minimally generated by $n-1$ elements, namely, the residue classes of $y_2, \ldots, y_n$. By induction on n, there exist $z_2, \ldots, z_n \in M$ such that $\overline{M} = \langle \overline{z_2} \rangle \oplus \cdots \oplus \langle \overline{z_n} \rangle$, and moreover, there exist $f = e_2 \geq \cdots \geq e_n$ such that $\operatorname{ann}(\overline{z_i}) = p_i^{e_i}R$, for all $2 \leq i \leq n$. By Lemma L, there exist $x_2, \ldots, x_n \in M$ such that $\overline{x_i} = \overline{z_i}$ and $\operatorname{ann}(x_i) = p_i^{e_i}R$, for all $2 \leq i \leq n$. If we set $x_1 := x$, we are done if we show $M = \langle x_1 \rangle \oplus \langle x_2 \rangle \oplus \cdots \oplus \langle x_n \rangle$. For this, we must show that $x_1, \ldots, x_n$ generate M and if $r_1x_1 + \cdots + r_nx_n = 0$, then each $r_ix_i = 0$.

Take $h \in M$. Then, $\overline{h} = \sum_{i=2}^n r_i\overline{x_i}$ in $\overline{M}$, for some $r_i \in R$, since $\overline{x_2}, \ldots, \overline{x_n}$ generate $\overline{M}$. Therefore, $h - \sum_{i=2}^n r_ix_i = r_1x_1$, for some $r_1 \in R$. It follows that $h = \sum_{i=1} r_ix_i$, showing $M = \langle x_1, \ldots, x_n \rangle$. In other words, $M = \langle x_1 \rangle + \cdots + \langle x_n \rangle$. Now, suppose $r_1x_1 + \cdots + r_nx_n = 0$, for $r_i \in R$. Then, $r_2\overline{x_2} + \cdots + r_n\overline{x_n} \equiv 0$ in $\overline{M}$. By the direct sum property for $\overline{M}$, each $r_i\overline{x_i} \equiv 0$ in $\overline{x_i}$. Thus, for each $2 \leq i \leq n$, $r_i \in \operatorname{ann}(\overline{x_i}) = \operatorname{ann}(x_i)$, and hence $r_ix_i = 0$, for $2 \leq i \leq n$. But then $r_1x_1 = 0$. Therefore, by Remark K (iii), $M = \langle x_1 \rangle \oplus \cdots \oplus \langle x_n \rangle$.

Finally, we have that for $2 \leq i \leq n$, $\operatorname{ann}(x_i) = p_i^{e_i}R$, with $e_2 \geq \cdots \geq e_n$. Moreover, we have $p^eM = 0$, so $p^e\langle x_i \rangle = 0$, for $2 \leq i \leq n$. Thus, $p^e \in p_i^{e_i}R$, and hence $e \geq e_i$, for all such i. Setting $p_1 := p$ and $e_1 := e$ finishes the proof. $\square$