Finitely-generated modules over a principal ideal domain

November 6, 2014

Let R be a commutative ring throughout. Usually R will be an integral domain and even a principal ideal domain, but these assumptions will be made explicitly. Since R is commutative, there is no distinction between left, right and 2-sided ideals. In particular, for every ideal I we have a quotient ring R/I. F always denotes a field.

Our goal is to prove the classification theorem for finitely-generated modules over a principal ideal domain, which comes in two versions: *elementary divisors* and *invariant factors*. For $R = \mathbb{Z}$ this gives a classification of finitely-generated abelian groups, while for $R = F[x]$ it can be use to classify $n \times n$-matrices up to similarity.

1 Overview of principal ideal domains

1.1 Definition and examples

A principal ideal domain is an integral domain in which every ideal is *principal*, i.e. generated by a single element. The two most well-known examples are $\mathbb{Z}$ and $F[x]$ (remember that F is a field throughout this installment of the notes). Other examples include the formal power series ring $F[[x]]$, the Laurent polynomials $F[x, x^{-1}]$, and the following:

Example: Let p be a prime and let $\mathbb{Z}_{(p)}$ denote the subring of $\mathbb{Q}$ consisting of fractions whose denominator (in lowest terms) is not divisible by p. One can check that $\mathbb{Z}_{(p)}$ is a principal ideal domain by checking that its only ideals are the zero ideal and the ideals (p^n), $n \geq 0$. This ring is called "the integers localized at p". One property of note is that it has a unique maximal ideal, namely (p).

Example: the p-adic integers. This example is included for your cultural edification, and is not required reading. Let $\mathbb{Z}_p$ denote the subring of $\prod_{n=1}^{\infty} \mathbb{Z}/p^n$ consisting of sequences $(a_1, a_2, ...)$ such that $a_{n+1} = a_n \bmod p^n$ for all ≥ 1. This ring is called the *p-adic integers*. (If you are accustomed to using the notation $\mathbb{Z}_p$ for the integers mod p, please cease doing so immediately. The integers mod p should be denoted $\mathbb{Z}/p$.) Note there is an inclusion $\mathbb{Z} \subset \mathbb{Z}_p$ given by sending n to its sequence of residue classes mod p^n. One can show that as in the previous example, the only ideals are the zero ideal and the ideals (p^n), $n \geq 0$. In particular every ideal is principal. Moreover $\mathbb{Z}_p$ is an integral domain (check this!), hence a principal ideal domain.

The p-adic integers were invented for number theoretic purposes. The idea is that when trying to analyze a Diophantine equation such as $x^m + y^m = z^m$, one might want to first look for solutions mod p, mod p^2 etc. The ring $\mathbb{Z}_p$ provides a mechanism for encoding solutions mod *all* powers of p. This ring is widely applied in other contexts as well, including representation theory and algebraic topology.

1.2 Principal ideal domains and the Big Picture

Algebraic geometry and algebraic number theory are two of the oldest branches of mathematics (especially if you delete the word "algebraic"), and much of abstract algebra grew out of these two sources. Let's take a look at how principal ideal domains arise in each case.

In algebraic geometry the prototypical ring is a polynomial ring $F[x_1, ..., x_n]$. So we see right away that principal ideal domains have limited application: $F[x_1, ..., x_n]$ is a principal ideal domain if and only if $n \leq 1$. Here one can easily check that for $n > 1$ the ideal $(x_1, ..., x_n)$ is not principal (how could $x_1, ..., x_n$ all be divisible by a fixed polynomial f?). In fact $F[x_1, ..., x_n]$ is in some sense an "n-dimensional" ring, a concept to be made precise in spring but that certainly makes intuitive sense if you think of $F = \mathbb{R}$. Principal ideal domains are in this same sense "1-dimensional" and hence in geometry apply mainly to curves.

To illustrate, consider $F = \mathbb{R}$ and the curve $y = x^3$. The *coordinate ring* R of this curve is the ring of polynomial functions on it, or to be precise $\mathbb{R}[x, y]/I$ where I is the ideal of all functions vanishing on our curve. In fact I is the ideal generated by $(y - x^3)$, so $R = \mathbb{R}[x, y]/(y - x^3)$. I claim that R is a principal ideal domain, and in fact this example was set up to make it easy: $R \cong \mathbb{R}[x]$ as $\mathbb{R}$-algebras (check this!—it works for any curve that is a graph $y = f(x)$), so is a principal ideal domain.

Now consider the curve $y^2 = x^3$, which has a cusp at the origin. Now the coordinate ring $S = \mathbb{R}[x, y]/(y^2 - x^3)$ is again an integral domain, and is again "1-dimensional" in the sense that I haven't defined, but *it is not a principal ideal domain.* One easy way to see this is to first show that S is isomorphic to the subalgebra T of $\mathbb{R}[x]$ with $\mathbb{R}$-basis $\{x^i : i \neq 1\}$. In T the span of x^i for $i > 0$ is an ideal generated by x^2, x^3 but clearly cannot be generated by one element. The amazing and beautiful fact is that this is not a coincidence; for curves the failure of the principal ideal domain condition is essentially equivalent to the existence of singularities such as the cusp!

In algebraic number theory the prototypical ring is $\mathbb{Z}$. More generally one considers subrings of $\mathbb{C}$ that are finitely-generated (and hence free) as abelian groups. Familiar examples include the Gaussian integers $\mathbb{Z}[i]$ and the cyclotomic integers $\mathbb{Z}[\xi_n]$, where $\xi_n = e^{2\pi i/n}$. All of these rings are "1-dimensional" integral domains, but they need not be principal ideal domains. A standard textbook example that is not a principal ideal domain is $R = \mathbb{Z}\sqrt{-5}$. I won't give the proof here; we'll come back to this in Winter. Upon seeing this example, any reasonable person's first question is: "Why $\sqrt{-5}$?" What does $\sqrt{-5}$ have to do with the price of bananas?[1] The fact is that it is very difficult to say whether a "number ring" of this type is a principal ideal domain or not, and purely algebraic methods are not adequate for the task—even though the problem itself *is* purely algebraic. Analytic number theory—zeta

[1] Very little, unless your grocery store sells bananas at irrational, imaginary prices.

functions, L-functions and the like—must be brought to bear, and such methods lie entirely outside our scope.

On the other hand, I suppose a reasonable person's first question might instead be: "Why do we care whether a number ring is a principal ideal domain or not?" The interesting fact is that for a number ring R (properly defined), being a principal ideal domain is equivalent to having unique factorization into irreducible elements. Erroneous 19th-century "proofs" of Fermat's Last Theorem, some of which even made it into the literature, were based on the naive assumption that for a prime p the cyclotomic integer ring $\mathbb{Z}[\xi_p]$ has unique factorization. Alas, this is true only for $p \leq 19$ (again, the proof of this fact requires analytic methods).

Thus the principal ideal domain property is related to both singularities of algebraic curves and to unique factorization in number rings. When you add to this that principal ideal domains are the simplest kind of commutative ring after fields, and that large parts of linear algebra appear as corollaries of their module theory, no reasonable person can deny that they are worthy objects of study. (The proof of that final claim is by contradiction.)

1.3 Basic properties

Recall that two elements a, b of an integral domain R are *associates* if $a = ub$ for some unit u, or equivalently $(a) = (b)$ i.e. the principal ideals they generate are the same. Note this is an equivalence relation. Recall also that a nonzero nonunit $x \in R$ is *irreducible* if whenever $x = ab$ either a or b is a unit, and is *prime* if the ideal (x) is a prime ideal, i.e. whenever $x|ab$ either $x|a$ or $x|b$. Every prime element is irreducible but the converse is false, in general (it isn't obvious that there are counterexamples; we will return to this point later).

Theorem 1.1 *Let R be a principal ideal domain. Then*

a) Every irreducible element of R is prime.

b) R is a unique factorization domain; i.e. every nonzero nonunit x can be factored as a product of irreducible elements: $x = p_1...p_n$, and this factorization is unique up to ordering and associates.

c) Every nonzero prime ideal is maximal.

d) Every ascending chain of ideals $I_1 \subset I_2 \subset ...$ stabilizes, i.e. there is an N such that $I_n = I_{n+1}$ for $n \geq N$.

Remarks: 1. The p_i's in (b) need not be distinct. Often we prefer to rewrite the factorization in the form $x = q_1^{i_1}...q_r^{i_r}$ where the q_i's are pairwise non-associate irreducible elements.

2. In (b), the "unique up to associates" clause can be eliminated by expressing the factorization in terms of principal ideals: $(x) = (p_1)...(p_n)$, uniquely up to ordering.

3. A commutative ring satisfying property (d) is called a *noetherian* ring.[2] Such rings will be studied in depth in Spring. The proof of (d) is easy, and we may as well give the more general statement:

Proposition 1.2 *Let R be any commutative ring. Then R is noetherian if and only if every ideal in R is finitely-generated.*

[2]Whether to capitalize the "n" in noetherian or not is a matter of taste. Such rings are named after Emmy Noether, and as I see it, it is a great honor to have your name turned into an adjective and even greater honor to lose the capital. Similarly I usually write "artinian" instead of "Artinian".

Proof: Suppose R is noetherian, and I is an ideal. If I is not finitely-generated then it contains a countably infinite set of elements x_i no finite subset of which generates I. So the ascending chain $(x_1) \subset (x_1, x_2) \subset ...$ fails to stabilize, contradiction.

Conversely suppose every ideal is finitely-generated and $I_1 \subset I_2 \subset ...$ is an ascending chain of ideals. Then $\cup_k I_k$ is an ideal, and hence is generated by some finite subset S. Then $S \subset I_N$ for some N, hence $I_n = I_{n+1}$ for all $n \geq N$.

Principal ideal domains have many further properties in common with $\mathbb{Z}$, for example concerning greatest common divisors, least common multiples and so on (but keep track of associates!). Consult your undergraduate algebra text or the references on reserve for details.

2 Finitely-generated modules over noetherian rings

Let R be a commutative ring. Although our immediate interest is in principal ideal domains, we may as well work more generally here, as the general case is no harder.

Proposition 2.1 *The property "finitely-generated R-module" is preserved by quotient modules, extensions and finite direct sums, but not in general by submodules.*

Proof: The quotient module case is clear. "Preserved by extensions" means that if $0 \longrightarrow L \longrightarrow M \longrightarrow N \longrightarrow 0$ is a short exact sequence and L, N are finitely-generated, then M is finitely-generated. This case too is easy, and left to the reader. The assertion for finite direct sum follows from the extension case in the usual way (a direct sum of two modules can be written as an extension in the evident way, then induct on the number of summands).

For a counterexample to preservation by submodules, take $R = F[x_1, x_2, ...]$ to be a polynomial ring on a countably infinite set of variables x_i. Let I denote the ideal generated by all the x_i's. Then R is a finitely-generated R-module (it is generated by one element, namely 1) and I is a submodule, but I is not finitely-generated. Any finite subset $f_1, ..., f_m$ of I consists of polynomials in only finitely many of the variables x_i, say $x_1, ..., x_n$. If this set generates I then $x_{n+1} = \sum_{i=1}^{m} r_i f_i$ for some $r_i \in R$, which is absurd (to check carefully this alleged absurdity, note that each r_i also involves only finitely many x_i's, so you've reduced to a calculation in a polynomial ring in finitely many variables).

An R-module M is *noetherian* if every ascending chain of submodules $M_1 \subset M_2 \subset ...$ stabilizes.

Proposition 2.2 *M is noetherian if and only if every submodule of M is finitely-generated.*

The proof is identical to the proof just given for ideals. Now, here is one very useful property of noetherian rings:

Proposition 2.3 *Suppose R is noetherian. Then every submodule of a finitely-generated R-module is finitely-generated. In particular this is true for principal ideal domains.*

Proof: The assertion is true for R as a module over itself, by definition of a noetherian ring and the preceeding proposition. Hence it is true for any finite direct sum of copies of R, and hence is true for any finitely-generated free module. So it is true for any quotient of a finitely-generated free module. But any finitely-generated module is such a quotient, QED.

We conclude with one more way to think about the noetherian condition.

Proposition 2.4 *M is a noetherian R-module if and only if every nonempty set of submodules of M has a maximal element, with respect to ordering by inclusion.*

Proof: Suppose M is noetherian. Let S be a nonempty sset of submodules of M, and suppose S has no maximal element. Then we can inductively construct an increasing chain $M_1 \subset M_2 \subset ...$ with each $M_i \in S$ and the inclusions proper, contradiction.

Conversely if the stated maximal condition holds and $M_1 \subset M_2...$ is an ascending chain, the set of M_i's has a maximal element M_k and hence the chain stabilizes starting at M_k.

3 Modules over an integral domain

Recall that an *integral domain* is a commutative ring with no (non-zero) zero-divisors, i.e. if $xy = 0$ then either $x = 0$ or $y = 0$. Here again our main interest is in principal ideal domains, but the general case is no harder for the results considered here.

3.1 Torsion modules and torsion-free modules

Let R be an integral domain, and let M be a left R-module. An element $x \in M$ is a *torsion element* if there is a nonzero $r \in R$ such that $rx = 0$. Recall that the *annihilator ideal* of x is defined by $ann\, x = \{r \in R : rx = 0\}$; note that it is a left ideal. Thus x is a torsion element if and only if $ann\, x \neq 0$.

Note that the map $R \longrightarrow M$ given by $r \mapsto rx$ is an R-module homomorphism, and factors uniquely through an injective R-module homomorphism $R/ann\, x \longrightarrow M$.

M is a *torsion module* if every $x \in M$ is a torsion element, and is *torsion-free* if no nonzero element is a torsion element.

Examples. 1. If $R = \mathbb{Z}$, so a module is just an abelian group, x is a torsion element for the module if and only if it is an element of finite order in the usual group-theoretic sense.

2. A free module over an integral domain is torsion-free. In particular, every F-vector space is a torsion-free F-module.

3. If $I \subset R$ is an ideal, then I is a torsion-free R-module (since R is an integral domain). If I is non-zero, then R/I is a torsion module.

Remark: The definitions "torsion element", "torsion module" etc. make sense over any ring. But if the ring has zero-divisors, then even R as a module over itself has torsion, since any zero-divisor is a torsion element. As a result the concept has limited utility, and we prefer to confine it to integral domains.

The full subcategories of torsion modules and torsion-free modules will be denoted respectively **tR-mod** and **tfR-mod**.

3.2 The fraction module of a torsion-free module

As we have mentioned several times, F-modules—i.e., vector spaces—are among the easiest modules to work with, and therefore it is a common strategy to extract information about general R-modules M from various vector spaces associated to M. We have already seen one way to do this: Since our ring is commutative, for any maximal ideal I we know that R/I is a field, and M/IM is an (R/I)-module. For example, we showed that if M is a free R-module then its rank is the dimension of M/IM, thereby proving the rank is well-defined. But if we think about $R = \mathbb{Z}$, for instance, another possibility presents itself. Suppose we have a free abelian group $\oplus_{i=1}^n \mathbb{Z}$. Rather than reducing it mod a prime, we might want to include it in the $\mathbb{Q}$-vector space $\oplus_{i=1}^n \mathbb{Q}$. This leads us to the general concept of a "fraction module", analogous to the fraction field of an integral domain. It is a special case of a much more general procedure known as "localization", which we'll study extensively in spring.

Let R be an integral domain, with fraction field F, and let M be a torsion-free R-module. Then one can construct a "fraction module" M_F from M in exactly the same way that F is constructed from R: We define M_F to be the set of equivalence classes of pairs (written as fractions) $\frac{x}{a}$ with $x \in M$ and $a \in R - 0$, where the equivalence relation is given by $\frac{x}{a} = \frac{y}{b}$ if $bx = ay$. Then the usual rule for addition of fractions makes M_F an abelian group, and moreover it is a vector space over F with multiplication $\frac{a}{b} \cdot \frac{x}{c} = \frac{ax}{bc}$. Note that there is an inclusion $M \longrightarrow M_F$ given by $x \mapsto \frac{x}{1}$, and it is a homomorphism of R-modules.

The upshot of this discussion is that we have a functor $tf R\text{-mod} \longrightarrow F\text{-mod}$ given by $M \mapsto M_F$ (the definition of induced maps, and verification that they satisfy the two conditions for a functor, is left to the reader).

Proposition 3.1 *a) $M \mapsto M_F$ preserves direct sums. More precisely, for any collection of torsion-free R-modules M_α, the natural map $\oplus_\alpha (M_\alpha)_F \longrightarrow (\oplus_\alpha M_\alpha)_F$ is an isomorphism of F-modules.*

b) If M is a free R-module, then $rank_R M = dim_F M_F$.

Proof: Exercise. *Caution*: In the case of an infinite collection of M_α's, part (a) is false if direct sum is replaced by direct product (can you think of a counterexample?). So your proof should make it clear where you are using the finiteness condition in the definition of direct sum.

In view of part (b), for any torsion-free M we *define* $rank_R M = dim_F M_F$.

The functor M_F is better behaved than reduction mod an ideal. It has the especially nice property that it preserves short exact sequences:

Proposition 3.2 *Let $0 \longrightarrow L \longrightarrow M \longrightarrow N \longrightarrow 0$ be a short exact sequence of R-modules. Then $0 \longrightarrow L_F \longrightarrow M_F \longrightarrow N_F \longrightarrow 0$ is short exact.*

Proof: See the exercises, where you will also investigate why the analogous statement for reduction mod an ideal fails.

Proposition 3.3 *Let M be a finitely-generated torsion-free R-module. Then M has finite rank.*

Proof: There exists a free R-module P of finite rank n and a surjective homomorphism $\phi : P \longrightarrow M$. Applying Proposition 3.2 to the short exact sequence $0 \longrightarrow Ker\,\phi \longrightarrow P \longrightarrow M \longrightarrow 0$, we conclude that $rank_R M = dim_F M_F \leq n$.

Let's use these ideas to show that even over integral domains, a submodule of a free module need not be free. Take $R = K(x, y)$, where K is a field, and let I denote the ideal (x, y). Then I is a submodule of the free module R, but I claim it is not free. Now, one could prove this by elementary *ad hoc* fiddling around, but I want to sketch a slicker approach. Let F denote the fraction field of R. Then $I_F = R_F = F$, so if I is free then it has rank one. On the other hand I itself is a maximal ideal, with residue field $R/I = K$, so if I is free its then rank is also $dim_K(I/I^2)$. But I^2 is just the K-vector space spanned by the homogeneous polynomials of degree > 1, so the images of x, y in I/I^2 form a K-basis. Hence I has rank 2, contradiction.

4 Finitely-generated torsion-free modules over a principal ideal domain

Lemma 4.1 *Let R be a principal ideal domain. Then every finitely-generated R-submodule Q of the fraction field F is free of rank one.*

Proof: Suppose Q is generated by $\frac{a_i}{b_i}$, $1 \leq i \leq m$, and let $x = \frac{1}{b_1 b_2 \ldots b_m}$. Then $Q \subset Rx$, and Rx is a free R-module of rank one. Since R is a principal ideal domain, it follows that Q is also free of rank one.

We are now in a position to prove:

Theorem 4.2 *Let R be a principal ideal domain, and let M be a finitely-generated torsion-free R-module. Then M is a finitely-generated free module.*

Proof: Let $n = rank_R M$. If $n = 1$, the theorem is equivalent to the lemma.

If $n > 1$, choose an F-linear map $\lambda : M_F \longrightarrow F$. Then its restriction $\lambda_M : M \longrightarrow F$ is a nonzero R-linear map. So there is a short exact sequence of R-modules $0 \longrightarrow K \longrightarrow M \longrightarrow L \longrightarrow 0$, where K, L are respectively the kernel and image of λ_M. By the lemma L is free of rank one, so by the freeness $M \cong K \oplus L$. Then $rank_R K = n - 1$, so by induction we can assume K is R-free of rank $\leq n - 1$. Hence M is free of rank $\leq n$, as desired.

Theorem 4.3 *Let R be a principal ideal domain. Then every submodule of a free module is free.*

Proof: Let M be a free module, $N \subset M$ a submodule. If M is finitely-generated, then since R is noetherian, N is also finitely-generated. Since N is torsion-free it is therefore free by the previous theorem. The case of an arbitrary free module is harder; for a proof see e.g. [Hungerford].

5 Torsion modules

Proposition 5.1 *Let R be a principal ideal domain, let x be a nonzero nonunit and let $x = q_1^{i_1}...q_r^{i_r}$ be a factorization into pairwise nonassociate irreducibles q_i. Then*

$$R/(x) \cong \oplus_{k=1}^{r} R/(q_k^{i_k}).$$

Proof: Since the finite direct sum is the same as the product, we get a module homomorphism $\phi : R/(x) \longrightarrow \oplus_{k=1}^{r} R/(q_k^{i_k})$ whose k-th component is reduction mod $q_k^{i_k}$. If $\phi(\overline{r}) = 0$ (where $\overline{r}$ denotes reduction mod (x)) then $q_k^{i_k} | r$ for all k, and since R is a unique factorization domain this forces $x|r$ and $\overline{r} = 0$. Hence ϕ is injective. Since source and target have the same finite length $\sum_{k=1}^{r} i_k$, ϕ is an isomorphism.

Let R be a principal ideal domain and let M be a torsion R-module. Thus for every $x \in M$ there is a nonzero $r \in R$ such that $rx = 0$. If $r = p^k$ for some irreducible p and $k \geq 0$, we call r a *p-torsion element*. It is easy to see that the subset of all p-torsion elements forms a submodule, which we denote by $t_p M$. Note that this is an abuse of notation, that would be better written $t_{(p)} M$ to emphasize that it only depends on the ideal (p) and not on the particular choice of generator p of the ideal (equivalently, choice of p within its associate class). But we'll indulge in it nonetheless, to avoid notational clutter.

Proposition 5.2 *Let R be a principal ideal domain and let M be a torsion R-module. Then $M = \oplus_p t_p M$, where (p) ranges over the nonzero prime ideals of R (or more efficiently, over those (p) such that $t_p M \neq 0$).*

Proof: Suppose $(p) \neq (q)$ and $x \in t_p M \cap t_q M$. Then $p^j x = 0 = q^k x$ for some j, k. Since p^j, q^k are relatively prime, $(p^j, q^k) = (1)$. Hence $1 \cdot x = 0$ and $x = 0$. This shows $\oplus_p t_p \subset M$. To show the inclusion is an equality, let $x \in M$ and $rx = 0$. Factor r into prime powers: $r = p_1^{i_1}...p_m^{i_m}$ and let $r_j = r/p^{i_j}$. Then $gcd\,(r_1, ..., r_m) = 1$, so there are elements $c_i \in R$ with $\sum c_i r_i = 1$. Moreover, r_i is a p_i-torsion element and $x = \sum c_i r_i x \in \oplus_{i=1}^{m} M_{p_i}$, completing the proof.

Note that if $\phi : M \longrightarrow N$ is a homomorphism of R-modules, then ϕ preserves torsion submodules and indeed preserves p-torsion submodules: $\phi(t_p M) \subset t_p N$.

If a torsion module M is a finite direct sum of cyclic modules of the form $R/(p^k)$ for p irreducible (equivalently prime, since R is a principal ideal domain), we call the set of primes p that occurs the *torsion primes*, the natural numbers k that occur for a given p the *p-exponents*, and define the *multiplicity* $m_{p,k}$ to be the number of times $R/(p^k)$ occurs as a summand in the given decomposition. Note carefully that so far we only know that the set of torsion primes is independent of the choice of such a decomposition; we will see shortly that the exponents and multiplicities also depend only on M, not the particular decomposition.

The next theorem completes the classification of finitely-generated torsion modules.

Theorem 5.3 *Let R be a principal ideal domain and let M be a finitely-generated torsion R-module. Then M is isomorphic to a direct sum of cyclic modules of the form $R/(p^k)$, p irreducible. Moreover the set of prime ideals (p), exponents k and multiplicities $m_{p,k}$ that occur are uniquely determined by M, up to ordering.*

Proof: We first prove the existence. By Proposition 5.2, we reduce at once to the case M a p-torsion module for some fixed prime ideal (p). Since M is a finitely-generated torsion module, it has finite length ℓ, and we proceed by induction on ℓ. Let $x \in M$ be an element of maximal p-torsion order n. In other words, $p^n w = 0$ for all $w \in M$ and $p^{n-1}x \neq 0$.

Lemma 5.4 *The short exact sequence $0 \longrightarrow Rx \longrightarrow M \longrightarrow M/Rx \longrightarrow 0$ splits, so $M \cong Rx \oplus M/Rx$.*

Note that the existence part of the theorem follows immediately, since M/Rx has lower length and hence is a direct sum of cyclic modules by induction.

To prove the lemma, recall that the sequence splits provided there is a homomorphism $\phi : M \longrightarrow Rx$ such that $\phi|_{Rx} = Id$. Let $L \subset M$ be a submodule containing Rx, and suppose we have a splitting defined on L; i.e. a homomorphism $\psi : L \longrightarrow Rx$ such that $\psi|_{Rx} = Id$. If $L \neq M$, I claim that ψ can be extended to a splitting defined on a submodule $\tilde{L}$ properly containing L. Since $\ell(\tilde{L}) > \ell(L)$, and M has finite length, after a finite number of steps we obtain the desired splitting ϕ.

To prove the claim, suppose $L \neq M$. Then there is a $y \notin L$ such that $py \in L$. Note that $p^{n-1}\psi(py) = 0$, and therefore $\psi(py) = pz$ for some $z \in Rx$ (since $Rx \cong R/(p^n)$). Set $\tilde{L} = L + Ry$. Let $\eta : L \oplus Ry \longrightarrow Rx$ denote the homomorphism given by $\eta = \psi$ on L and by $\eta(y) = z$ on Ry. Since the latter two homomorphisms agree on $L \cap Ry$, η factors uniquely through a homomorphism $\tilde{\psi} : \tilde{L} \longrightarrow Rx$. By construction $\tilde{\psi}$ is a splitting extending ψ. This proves the claim and completes the proof of existence in the theorem.

We now turn to the proof of uniqueness. Let $a := (a_1, a_2, ...)$ be a sequence of natural numbers all but finitely many of which are zero. Let

$$M_a = \oplus_i (R/p^i)^{a_i}.$$

Note there are only finitely many nonzero summands. What we need to show is that if $M_{\underline{a}} \cong M_{\underline{b}}$, then $\underline{a} = \underline{b}$; in other words, $\underline{a}$ is an isomorphism invariant. To do this we'll attach to $M_{\underline{a}}$ another sequence $L = (\ell_1, ...)$ that is clearly isomorphism invariant, and then show L determines $\underline{a}$.

For any R-module M and $r \in R$, let $M[r] = \{x \in M : rx = 0\}$. For M finitely-generated we then set $\ell_k = \ell(M[p^k])$. It is clear that an isomorphism $M \cong N$ induces an isomorphism $M[r] \cong N[r]$, so in particular ℓ_k is an isomorphism invariant for all k. Now let n be maximal such that $a_n \neq 0$ and observe:

$\ell_1 = a_1 + a_2 ... + a_n$
$\ell_2 = \ell_1 + a_2 + ... + a_n$
.

.
. $\ell_n = \ell_{n-1} + a_n$.

Starting with the last equation we can solve inductively for the a_i's in terms of the ℓ_i's. Hence $\underline{a}$ is an isomorphism invariant, as claimed.

This completes the proof of the classification theorem for finitely-generated torsion modules.

Well, not quite, because the theorem has a variant form in terms of the so-called "invariant factors". To motivate the invariant factor form, note that the elementary divisor form of the theorem given above has one feature that could be regarded as a "flaw": In general the number of summands (or number of generators) obtained is much larger than necessary. For example, even a cyclic module $R/(x)$ appears as the direct sum of k cyclic modules, where k is the number of distinct irreducibles dividing x. (Think of e.g. $\mathbb{Z}/60 = \mathbb{Z}/4 \oplus \mathbb{Z}/3 \oplus \mathbb{Z}/5$.) The invariant factor form has the advantage that the number of cyclic summands in the decomposition is as small as possible. Having said this, it is only fair to point that for many consumers of the theory, the elementary divisor form is by far the more useful. I for one have used the elementary divisor form thousands of times, almost every day, whereas I very rarely use invariant factors. This is because I generally prefer whenever possible to work "one prime at a time". If only one torsion prime (=irreducible for principal ideal domains) occurs in the given module, the two forms coincide.

Now, here is the Invariant Factor Decomposition:

Theorem 5.5 *Let M be a finitely-generated torsion module over the principal ideal domain R. Then*

$$M \cong R/(x_1) \oplus R/(x_2) \oplus ... \oplus R/(x_m),$$

where $x_1|x_2|...|x_m$ (i.e. x_1 divides x_2 and so on). Moreover the number m and the ideals (x_i) (called the invariant factors) are uniquely determined by M.

Proof: We first prove the existence, proceeding "from the top down", meaning that we first find x_m, then x_{m-1} etc. Let $p_1, ...p_r$ be the torsion primes of M, and write the elementary divisor decomposition as follows: First, $M = \oplus_{i=1}^r M_i$, where M_i is the p_i-torsion submodule. For each i we can further decompose M_i as a direct sum of cyclic modules of the form $R/(p_i^k)$. Let $d(i)$ denote the maximal exponent k that occurs for p_i, and set

$$y_1 = \prod_{i=1}^{r} p_i^{d(i)}.$$

Using Proposition 5.1 we then have a decomposition

$$M = R/(y_1) \oplus N,$$

where N is obtained by deleting one summand $R/(p_i^{d(i)}$ from each M_i. Repeat the process on N, and continue. The result is a decomposition

$$M = \oplus_{i=1}^{m} R/(y_i)$$

for some m. Moreover it is clear that $y_i | y_{i-1}$, since the maximal exponents $d(i)$ can only decrease at each stage of the process. Let $x_1, ..., x_m$ denote the y_i's in reverse order yields the existence.

We now turn to the proof of uniqueness. Suppose we are given a decomposition as in the theorem: $M = M/(x_1) \oplus ... \oplus M/(x_m)$ with $x_1 | x_2 | ... | x_m$. This data is uniquely determined by the $m \times r$ matrix a_{ij} defined by

$$(x_i) = \prod_j (p_j)^{a_{ij}}.$$

The matrix $A := a_{ij}$ has the properties:

(1) $a_{1j} \neq 0$ for at least one j (i.e. the first row is not identically zero).
(2) $a_{ij} \leq a_{i,j+1}$ (the a_{ij}'s are weakly increasing down the columns).
(3) The list of nonzero entries in each column depends only on M, not the choice of invariant factor decomposition (this follows from the uniqueness clause in the elementary divisors decomposition).

Uniqueness now follows formally from these properties. For suppose $M = M/(y_1) \oplus ... \oplus M/(y_n)$ is another such decomposition, with matrix $B := b_{ij}$. Choose j so that $a_{1j} \neq 0$. Then the j-th column of B has the same list of nonzero entries as the j-th column of A by (3), although *a priori* there could be some zero entries at the start. Hence $m \leq n$. Reversing the roles of A, B, we conclude that $m = n$. Then conditions (2), (3) clearly force $a_{ij} = b_{ij}$ for all i, j, and hence $(x_i) = (y_i)$ for all i, as desired.

That m is the minimal number of summands in a decomposition into cyclic modules is left as an exercise.

6 The classification theorem for finitely-generated modules over a principal ideal domain

We've now done all the hard work for the main theorem.

Theorem 6.1 *Let R be a principal ideal domain and let M be a finitely-generated R-module. Then $M \cong N \oplus tM$, where N is a free R-module of finite rank and tM is the torsion submodule and hence classified by either its elementary divisors or its invariant factors as above. Moreover rank N depends only on M.*

Proof: Recall that tM denotes the torsion submodule of M. We have a short exact sequence $0 \longrightarrow tM \longrightarrow M \longrightarrow M/tM \longrightarrow 0$. Since M/tM is a finitely-generated torsion-free module and R is a principal ideal domain, M/tM is a free module of finite rank. By an exercise in the module notes, the sequence therefore splits, so there is an isomorphism $M \longrightarrow M/tM \oplus tM$ and

we can take $N = M/tM$. We have already classified tM. Finally $rank\, N = rank\,(M/tM) = dim_F(M/tM)_F$ (where F is the fraction field of R), so the rank depends only on M. QED!

We define $rank\, M = rank\,(M/tM)$. The finite list of prime ideals occuring in the decomposition of tM are the *torsion primes*. The exponents i such that $R/(p^i)$ occurs in the elementary divisor decomposition are the *exponents*, or p-exponents if we want to specify p. Finally the *multipicity* of a given R/p^i is the number of times it occurs in the elementary divisor decomposition. Thus M is uniquely determined by its rank, torsion primes, exponents and multiplicities.

Here are two important special cases. We focus on the torsion submodule, since that's the more complicated part.

Example. Take $R = \mathbb{Z}$. Then the theorem classifies finitely-generated abelian groups: If A is such a group, then it is isomorphic to a direct sum of $\mathbb{Z}^n$ and cyclic modules $\mathbb{Z}/p^i$ for some finite list of primes p, exponents i and multiplicites p. This data uniquely determines A up to isomorphism. For example, suppose we have a finite abelian group of order 72. Then we know it is one of six possible groups, up to isomorphism: The 2-torsion subgroup is $(\mathbb{Z}/2)^3$, $(\mathbb{Z}/2) \oplus \mathbb{Z}/4$ or $\mathbb{Z}/8$, while the 3-torsion subgroup is $(\mathbb{Z}/3)^2$ or $\mathbb{Z}/9$. Note also that a finitely-generated abelian group is a torsion group if and only if it is finite.

Example. Take $R = F[x]$. Then a finitely-generated $F[x]$-module is a torsion module if and only if it is finite-dimensional (when the ring is an F-algebra, dimension always refers to F-dimension unless otherwise specified). The proof of this claim is one of the exercises below. The nonzero prime ideals are the ideals generated by irreducible polynomials, so any finitely-generated torsion module is isomorphic to a direct sum of modules of the form $F[x]/f^i$ with f an irreducible polynomial. The polynomial ring case will be studied in detail in the next installment of the notes.

Finally, we discuss an important special class of principal ideal domains: the *local* principal ideal domains.

A commutative ring is *local* if it has a unique maximal ideal $\mathfrak{m}$. Let's pause to observe a simple fact:

Proposition 6.2 *Let R be a commutative ring, $I \subset R$ a proper ideal. Then R is local with I as its unique maximal ideal if and only if every $r \notin I$ is a unit.*

Proof: Suppose R is local with unique maximal ideal I, and $r \notin I$. If r is not a unit then it is contained in some maximal ideal J, a contradiction since $J \neq I$.

Conversely suppose every $r \notin I$ is a unit. Then I is maximal because if $I \subset J$ is a proper inclusion of ideals, then J contains a unit and hence $J = R$. Moreover if K is any maximal ideal, then it can't contain any units, hence $K \subset J$ and by maximality $K = J$.

Modules over local principal ideal domains are especially simple, because in a principal ideal domain every nonzero prime ideal is maximal, and therefore if R is a local principal ideal domain that is not a field, it has exactly two prime ideals: 0 and $\mathfrak{m}$. Consequently every torsion module is isomorphic to a direct sum of cyclic modules of the form $R/\mathfrak{m}^i$. Following are some examples, with details left to the reader:

Example. $\mathbb{Z}_{(p)}$ is a local principal ideal domain. Moreover $\mathbb{Z}_{(p)}/(p^i) \cong \mathbb{Z}/p$ as abelian groups. So every finitely-generated torsion module over $\mathbb{Z}_{(p)}$ is a direct sum of $\mathbb{Z}/p^i$'s.

Example. The formal power series ring $F[[x]]$ is a local principal ideal domain. The proof is an exercise below.

7 Classification of submodules and homomorphisms for finitely-generated free modules

Throughout this section, R is a principal ideal domain, and M is a finitely-generated free R-module of rank m.

The key result, from which the others will be derived, is a "canonical form" for submodules of M.

Theorem 7.1 *Let N be a nonzero submodule of M, of rank n. Then there is a basis $e_1, ..., e_m$ for M and elements $a_1|a_2...|a_n$ of R such that $a_1e_1, ..., a_ne_n$ is a basis for N.*

Moreover, the ideals (a_i) are uniquely determined by N.

Before giving the proof, some comments are in order.

- We know that N is a finitely-generated free module.
- If R is a field then the theorem is elementary linear algebra, with all $a_i = 1$ (any basis of N extends to a basis of M in that case).
- We anticipate some kind of induction on rank, perhaps making a judicious choice of surjective $\lambda : M \longrightarrow R$ (many such exist, by the freeness of M). For this purpose it is useful to consider *all* homomorphisms $M \longrightarrow R$, i.e. the R-dual of M, denoted $M^* := Hom_R(M, R)$. Note that M^* is itself an R-module, with $(r \cdot \lambda)(x) = r\lambda(x)$ for $r \in R$, $x \in M$.
- For many purposes the divisibility condition $a_1|a_2...$ is not important. The key point is having the simple "diagonal" form.
- When $m = n$, be sure not to confuse this with diagonalizability. A square matrix A over R is *diagonalizable* if it is *similar* to a diagonal matrix, meaning that there is a single basis such that the matrix of A in this basis is diagonal. The theorem says you make A diagonal if you allow different bases in the source and target, which is a far weaker statement.

Proof of Theorem: We prove the existence (uniqueness is left as an exercise). For each $\mu \in M^*$, $\mu(N) \subset R$ is an ideal. Since R is noetherian, the collection of all such ideals has a maximal element $\lambda(N)$, and since R is a principal ideal domain we have $\lambda(N) = (a_1)$ for some a_1. So there is an $x_1 \in N$ with $\lambda(x_1) = a_1$.

Claim 1. For any $\mu \in M^*$, $a_1|\mu(x_1)$.

Let $b = \mu(x_1)$, and let $d = gcd(a_1, b)$. Then by a general fact about principal ideal domains there are elements $r, s \in R$ such that $ra_1 + sb = d$. Set $\xi = r\lambda + s\mu \in M^*$; then $\xi(x_1) = d$. But $d|a_1$ so $(a_1) \subset (d)$ forcing $(a_1) = (d)$ by maximality. Hence $a_1|b$, proving Claim 1.

Claim 2. There is an $e_1 \in M$ with $x_1 = a_1e_1$ and $\lambda(e_1) = 1$.

To see this, choose any basis $f_1, ..., f_m$ of M (thereby identifying M with R^m) and write $x_1 = \sum c_i f_i$. Since the projections on each factor of R^m are elements of M^*, by Claim 1 we conclude that $a_1|c_i$ for all i. This proves the first statement of Claim 2, and applying λ to the equation $x_1 = a_1e_1$ prove the second.

So not only is there a short exact sequence $0 \longrightarrow Ker\,\lambda \longrightarrow M \longrightarrow R \longrightarrow 0$, and not only does it split; even better, we have an explicit splitting $R \xrightarrow{\cong} Re_1$. Hence $M = Ker\,\lambda \oplus Re_1$.

Warning! Dangerous curve! If a module L decomposes as $L = L_1 \oplus L_2$, and $K \subset L$ is a submodule, in no way, shape or form does it follow that $K = (K \cap L_1) \oplus (K \cap L_2)$. Indeed this fails miserably even for $R = \mathbb{R}$ and the good old xy-plane of yore, which we think of as the direct sum of its two axes. Now think of the submodule defined by $y = x$.

The moral of our story is that *Proof-by-wishful-thinking* is not adequate to conclude

Claim 4: $N = (N \cap Ker\,\lambda) \oplus (N \cap Re_1)$.

To prove it we need to show that if $y \in N$ and $y = y_1 + y_2$ with $y_1 \in Ker\,\lambda$ and $y_2 \in Re_1$, then $y_1, y_2 \in N$. A moment's reflection reveals that it suffices to show one of the two is in N. So consider $y_2 = re_1$ for some $r \in R$. Then $r = \lambda(y_2) = \lambda(y) \in \lambda(N) = (a_1)$, hence $y_2 \in N$ and claim 4 is proved.

By induction on rank we can assume that $Ker\,\lambda$ has a basis $e_2, ..., e_m$ such that $a_2e_2, ..., a_ne_n$ is a basis for $N \cap Ker\,\lambda$, and $a_2|a_3|....$ It only remains to show:

Claim 5: $a_1|a_2$.

This is proved similarly to Claim 1. Let $d = gcd(a_1, a_2)$. Then there exist $r_1, r_2 \in R$ such that $r_1a_1 + r_2a_2 = d$. Let $\pi_2 \in M^*$ denote projection on the second coordinate and set $\mu = r\lambda + s\pi_2$. Then $\mu(a_1e_1 + a_2e_2) = d$, and since $(a_1) \subset (d)$ we must have $(a_1) = (d)$ by maximality. Then $a_1|a_2$ as desired.

The theorem can be reformulated in the following way:

Theorem 7.2 *Let M, N be as in the previous theorem, and let $e_1, ..., e_m$ be a fixed basis for M. Then there is an R-module automorphism $\phi : M \xrightarrow{\cong} M$ such that $\phi(N)$ is the span of $a_1e_1, ..., a_ne_n$ and $a_1|a_2...|a_n$.*

Moreover the ideals (a_i) are uniquely determined by N.

This follows from Theorem 7.1 in the usual way as for vector spaces, interpreting a change of basis for M as an automorphism.

Next we classify R-module homomorphisms $N \longrightarrow M$, where N is free of rank n and M is free of rank m (we no longer assume N is a submodule, and either n or m could be the larger of the two). We will state our results in three equivalent forms, starting with the most

"concrete" and proceeding to the categorical version. Let $k = min\,(m, n)$ and let $(a_1, ..., a_k)$ be elements of R (we allow $a_i = 0$). If $k = m$ we define the matrix $A(a_1, ..., a_k) \in M_{mn}R$ to be the matrix whose lefthand $k \times k$ block is diagonal with diagonal entries $a_1, ..., a_k$, and zero elsewhere, and similarly for $k = n$ using the upper $k \times k$ block. For example, if $m = 3$, $n = 4$ we mean the matrix

$$\begin{pmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 \\ 0 & 0 & a_3 & 0 \end{pmatrix}$$

Theorem 7.3 *Let $\phi : N \longrightarrow M$ be a homomorphism. Then there are bases $e_1, ..., e_n$ for N and $f_1, ..., f_m$ for M such that the matrix of ϕ in these bases has the form $A(a_1, ..., a_k)$, and $a_1|a_2...|a_k$.*

Moreover the ideals (a_i) are uniquely determined by ϕ.

Note that the divisibility condition forces any 0's to come at the end. We sketch the proof and leave the details to the reader. Applying Theorem 7.1 to $Im\,\phi$, we can find a basis f_i for M such that $a_1 f_1, ..., a_q f_q$ is a basis for $Im\,\phi$, where $q = rank\,(Im\,\phi)$. Here $q \leq k$ and we set $a_i = 0$ for $i > q$. Since $Im\,\phi$ is free, the short exact sequence

$$0 \longrightarrow Ker\,\phi \longrightarrow N \longrightarrow Im\,\phi \longrightarrow 0$$

splits, so there is a submodule $L \subset N$ such that $N = L \oplus Ker\,\phi$ and $\phi|_L$ is an isomorphism onto $Im\,\phi$. We then take a basis e_i for M such that $\phi(e_i) = a_i f_i$ for $i \leq q$, and $\{e_i : i < q\}$ is a basis for $Ker\,\phi$. Uniqueness is left to the reader.

The second version follows at once from the first, interpreting changes of basis in terms of automorphisms.

Theorem 7.4 *Let $\phi : R^n \longrightarrow R^m$ be a homomorphism. Then there are automorphisms B of R^n and C of R^m such that $B\phi C = A(a_1, ..., a_k)$ (with divisibility and uniqueness as before).*

Finally, we recognize that what we're really doing is classifying certain morphisms up to isomorphism in the category of morphisms (see the category theory notes).

Theorem 7.5 *Let $\phi : N \longrightarrow M$ be a homomorphism of finitely-generated free modules. Then ϕ is isomorphic in the category of morphisms to some $A(a_1, ..., a_k) : R^n \longrightarrow R^m$, with divisibility and uniqueness as before.*

In other words, there is a commutative diagram

$$\begin{array}{ccc} N & \xrightarrow{\phi} & M \\ {\scriptstyle\alpha}\downarrow & & \downarrow{\scriptstyle\beta} \\ R^n & \xrightarrow[A]{} & R^m \end{array}$$

with α, β isomorphisms and $A = A(a_1, ..., a_k)$.

8 Exercises

1. Use principal ideal domain theory to classify up to isomorphism the irreducible representations of C_p (p prime) over: a) $\mathbb{C}$; b) $\mathbb{R}$; c) $\mathbb{Q}$. For part (c) you may assume the cyclotomic polynomial $1 + x + x^2 + ... + x^{p-1}$ is irreducible in $\mathbb{Q}[x]$.

Remark: The same method can be used for any cyclic group C_n; you might think about this as an optional exercise.

2. Show that the number m occuring in the invariant factor decomposition is the minimal number of summands in a decomposition of M into cyclic modules.

3. Show that a finitely-generated $F[x]$-module is a torsion module if and only if it is finite dimensional. Do not use the classification theorem; do it directly.

4. Show that the formal power series ring $F[[x]]$ is a local principal ideal domain.

5. Let $A : \mathbb{Z}^n \longrightarrow \mathbb{Z}^n$ be a group homomorphism. Show that *coker* A is finite if and only if *det* $A \neq 0$, in which case $|coker\, A| = |det\, A|$ (|| denotes order on the left, absolute value on the right.)

Notes: 1. The *cokernel coker* ϕ of any module homomorphism $\phi : M \longrightarrow N$ is $N/(Im\, \phi)$.

2. The determinant is defined by regarding A as an $n \times n$ matrix over $\mathbb{Z}$ (hence over $\mathbb{Q}$).

3. *Optional Challenge Problem.* Formulate and prove an analogous statement over an arbitrary principal ideal domain R . Here the biggest challenge is the "formulate" step, i.e. determining what the right analogue would be. The cokernel and determinant are defined in the same way, but over a general principal ideal domain neither the "order" of a module A nor the "absolute value" of an element of R make any sense at all. So, and this is often the case in the real world of mathematics, the puzzle is to figure out what the right definitions are.

6. a) How many isomorphism classes are there of abelian groups of order 900? Describe them.

b) How many isomorphism classes of $\mathbb{F}_2[x]$-modules of dimension 2 are there? Describe them.

7. Explain briefly why the existence part of the "elementary divisor" form of the classification of finitely-generated torsion modules follows from the "invariant factor" form.

8. Use the canonical form for submodules Theorem 7.1 to prove the existence part of Theorem 5.5 (the "invariant factor" form of the classification).

9. Prove the uniqueness statement in the canonical form Theorem 7.1.