PRINCIPAL IDEAL DOMAINS

RICHARD CUDNEY

1. Unique factorization

1.1. **Introduction.** If you like, you can consider these notes as an exposition of algebra over Z, masked in obfuscatory terminology. They merely recast all of the usual arguments from elementary arithmetic in the more abstract language of rings, ideals, and modules. However, I hope it will become clear how the language of ideals clarifies these kinds of problems.

1.1.1. *Preliminary notions.* In these notes, all rings are commutative. k always stands for a field.

Definition 1.2. *An element u of a ring R is called a unit if there exists an element v of R such that $uv = 1$. v is called the inverse of u.*

Exercise 1.3. *The inverse of a unit is unique.*

Exercise 1.4. *The only units in Z are ± 1. The only units in k[x] are the non-zero elements of k.*

Exercise 1.5. *A ring is a field iff all non-zero elements are units.*

Definition 1.6. *An element a of a ring is called irreducible if it is not a unit and $a = bc$ implies that b or c is a unit.*

This definition is our generalization of the definition of a prime integer to arbitrary rings. Unfortunately this is distinct from the following definition:

Definition 1.7. *An element a of a ring is called prime if the ideal it generates is prime. In other words, if $a|bc$ implies that $a|b$ or $a|c$.*

Definition 1.8. *A ring is called a domain if (0) is a prime ideal. In other words, if $ab = 0$ implies that $a = 0$ or $b = 0$.*

Exercise 1.9. *In a domain, primes are irreducible.*

1.10. **The unique factorization theorem for PIDs.** We give the following definition of rings for which our ideal unique factorization theorem is true:

Definition 1.11. *A domain R is called a unique factorization domain(abbreviated UFD) if every element has a factorization into irreducibles, and this factorization is unique up to multiplying the irreducibles by units.*

It is proved in elementary number theory that Z is a UFD. To prove this fact, one generally proceeds by using the division algorithm to show using the Euclidean algorithm that linear diophantine equations can be solved, from which the fact that prime numbers are prime in the above sense follows. One uses the same method to

Date: September 14, 2004.

show that k[x] is a UFD. We can define the most general ring to which this sequence of reasoning applies:

Definition 1.12. *A domain R is called a Euclidean domain if there exists a function from R to the natural numbers denoted by absolute value, which takes non-zero elements of R to positive numbers, such that for any a and b in R, there exist q and r such that $a = bq + r$ and $|r| < |b|$.*

This definition is not satisfactory because it involves an auxiliary function that is not well-defined. We get a more intrinsic definition by abstracting the solvability of linear diophantine equations:

Definition 1.13. *A domain R is called a principal ideal domain(abbreviated PID) if all ideals are generated by one element.*

How does this generalize solvability of linear diophantine equations? For example, suppose a and b have no common non-unit factor, or in other words, are relatively prime. (a, b), the ideal generated by a and b is equal to (d) for some d. $d|a$ and $d|b$ so d is a unit, so $(a, b) = (1)$. Thus since $1 \in (a, b)$, we can find x and y such that $ax + by = 1$.

Proposition 1.14. *Euclidean domains are PIDs.*

Proof: Let A be an ideal in a Euclidean domain R. Let a be an element of A of minimal non-zero Euclidean norm. Such an a exists, because otherwise $A = (0)$, which is a principal ideal. Now let b be any other element of A. Since R is Euclidean, we can find q and r such that $b = aq + r$, with $|r| < |a|$. Since a and b are in A, so is r, so $r = 0$ since it has smaller norm than a. Thus $A = (a)$.

From now on, except for the very end of the notes, we will forget about the notion of a Euclidean ring and generalize properties of Z to arbitrary PIDs.

1.14.1. *Existence of factorizations in PIDs.* It is usual for a Euclidean norm to be multiplicative, and such is the case is nearly every example. This property is used to prove that factorizations into irreducibles exist-for if an element a is not irreducible, it has a non-trivial factor a_1, and by multiplicativity, each of these factors has smaller norm, so by using induction, we can show that any element has a factorization into irreducibles. But we want to show the existence of irreducible factorizations for arbitrary PIDs, so a different argument is necessary.

Recall that containment of ideals is the opposite of divisibility. So if we have $a_1|a$, $a_2|a_1$, we have an ascending chain of ideals

$$(a) \subset (a_1) \subset (a_1) \ldots$$

If none of these factors a_i are ever irreducible, then this chain keeps increasing. However, if we knew that it stopped at some point, at the point where it stopped a_k would be irreducible. Thus every element in R would be divisible by some irreducible, and then by dividing a by irreducibles successively, we could construct another ascending chain, the termination of which would imply that a has a factorization in terms of irreducibles.

Definition 1.15. *A ring is called Noetherian if all ascending chains of ideals terminate.*

What the discussion above shows is the following:

Proposition 1.16. *In Noetherian rings any element can be factored into irreducibles.*

Proposition 1.17. *PIDs are Noetherian.*

Proof: Let $(a_1) \subset (a_s) \subset \ldots$ be an ascending chain of ideals. $\cup(a_i)$ is an ideal, so it is (a) for some a. a is in one of the a_i, so at that point the chain terminates.

Corollary 1.18. *In PIDs every element can be factored into irreducibles.*

Exercise 1.19. *Generalize the above proof to show that a ring is Noetherian iff all of its ideals are finitely generated.*

1.19.1. *Uniqueness of factorizations in PIDs.*

Proposition 1.20. *For any domain, factorizations into primes are unique up to unit multiples.*

Proof: Suppose $r = p_1 \ldots p_n = q_1 \ldots q_m$. Since p_1 is prime, it divides some q_j. Thus $q_j = ap_1$. Since q_j is irreducible and p_1 is not a unit, a is a unit. Now divide both factorizations of r by p_1. An induction argument on n will complete the proof.

Proposition 1.21. *In a PID, irreducibles are prime.*

Proof: Let p be an irreducible element in a PID R. Suppose $p|ab$, but p does not divide a. Then p and a are relatively prime, so we can solve $ax + py = 1$. Then $bax + bpy = b$. So $p|b$.

Corollary 1.22. *PIDs are UFDs.*

1.23. **Some UFDs that are not PIDs.**
In this section we will show that if R is a UFD then R[x] is as well. By induction, this well tell us that all of polynomial rings in a finite number of variables over a PID are UFDs, and in particular, all polynomial rings in a finite number of variables over a field.

First, we want to prove that factorizations into irreducible elements exist in R[x]. It is true that if R is Noetherian then R[x] is as well-this statement is known as the Hilbert basis theorem-but this is not useful for us. Not all UFDs are Noetherian, and there is a more direct way to show that irreducible factorizations exist.

Proposition 1.24. *If R is a UFD, then every element in R[x] can be factored into irreducibles.*

Proof: We will induct on the degree of a polynomial. If $\deg(f) > 0$, it is an element of R, and so can be factored into irreducibles of R, which are necessarily irreducibles of R[x]. Now, if f has degree n ¿ 0, factor out the greatest common factor of all the coefficients of f, which makes sense sinec R is a UFD, and then $f = ag$, where g is divisible by no non-units in R. Thus if g is not irreducible, every proper factor is a polynomial of lower degree.

To continue the proof, we need to introduce an auxiliary construction-the fraction field of a domain. This is the ring generated by symbols (a, b) subject to the relation $(a, b) = (c, d)$ if $ad = bc$ and with addition and subtraction defined as though (a, b) were a fraction a/b. Call K the fraction field of R. By the last section, K[x] is a PID. So we are very close to having what we want.

Lemma 1.25. *Primes in R are prime in R[x]*

Proof: It is equivalent to showing that R[x]/pR[x] is a domain, if p is a prime in R. However, R[x]/pR[x] is easily seen to be isomorphic to (R/pR)[x], and R/pR is a domain, so (R/pR)[x] is one as well.

Lemma 1.26. *Let $f \in R$. If $f = gh$, g and h in K[x], then there exist e and d in K such that $eg \in R[x]$, $dh \in R[x]$ and $ed = 1$.*

Proof: Clearly we can find a and b such $agbh = cf$ for some c in R. If c is a unit we are done. Otherwise some prime p divides c, and by the previous lemma, it divides ag or bh. So dividing a or b by c, we can reduce the length of prime factors of c. By induction we are done.

Exercise 1.27. *Irreducible elements of R[X] are irreducible in K[x].*

Proposition 1.28. *Irreducible elements of R[x] are prime.*

Proof: Let f be an irreducible element of R[x]. If f is of degree 0, the above lemma tells us that it is prime. Suppose f ¿ 0. Suppose $f|gh$. Then by the exercise, $f|g$ or $f|h$ in K[x]. Suppose without loss of generality that $g = kf$, k in K[x]. Then $g = laf$, for some $a \in$ K such that af is in R[x], and some $l \in$ R[x]. Let $a = b/c$, where b and c have no common factor. Then c divides f, so c is a unit, because f is an irreducible polynomial of positive degree. So $a \in R$. So $f|g$.

2. Modules over PIDs

2.1. **Introduction.** Recall the classification for finitely generated abelian groups. It states that any finitely abelian group is the direct sum of cyclic groups of prime power order and infinite order, with the number of factors of each kind well-defined up to isomorphism. This section will generalize this statement to a theorem about finitely generated modules over an arbitrary PID. Unlike what a simpleminded analogy with the above section may suggest, here the existence of a decomposition is much more difficult and useful than the uniqueness result. The reason is that there is a very general uniqueness theorem for decompositions, which you will see in first quarter algebra, which I think is much more illuminating than the easier ad hoc argument used to show uniqueness in this simple case. Thus the bulk of this text will focus on the existence result.

Definition 2.2. *A module M is free if it is isomorphic to a direct sum of some(possibly infinite) number of copies of R.*

Definition 2.3. *An element m of M is called torsion if there exists a non-zero element r of R such that $rm = 0$.*

Definition 2.4. *A module is called torsion if all of its elements are torsion. A module is torsion free if none of its non-zero elements are torsion.*

Definition 2.5. *If M is a module, let M^{tor} be the submodule of torsion elements.*

Definition 2.6. *A module is called cyclic if it is generated by a single element.*

2.6.1. *Noetherian modules.* We will be construct our decomposition in several steps. We will want to know that all the submodules and quotient modules we construct are finitely generated, so that we can reduce the problem to simpler cases. The best way to see this is by the material contained in this sequence of exercises. An exercise further below suggests a different method of proving this material that is more hands-on.

Definition 2.7. *An R module is called Noetherian if all ascending chains of submodules terminate.*

Exercise 2.8. *A module M is Noetherian if and only if all submodules are finitely generated. In particular, a Noetherian module is finitely generated.*

Exercise 2.9. *If M is Noetherian, so are all of its submodules and quotients.*

Exercise 2.10. *Let N be a submodule of M. Then M is Noetherian iff both N and M/N are.*

Exercise 2.11. *Over a Noetherian ring, all finitely generated modules are Noetherian.*

Theorem 2.12. *Over a PID R , any finitely generated module is a direct sum of a free module and cyclic modules, of prime-power torsion.*

The proof of this theoremwill occupy the remainder of these notes.

2.13. **Free modules.** It follows from the theorem above that a torsion-free finitely generated module over a PID will be free. Proving this statement will be our first task. For knowing this fact will give us more-that we can split off torsion as a direct summand, from the general fact that whenever we have a free quotient of a module M, M can be written as a direct sum of the kernel and the quotient. This result applies if we take the quotient of M by M^{tor}, which will be torsion-free and hence free.

Lemma 2.14. *Let N be a submodule of M, and let $\frac{M}{N} = F$ be free. Then M is isomorphic to $N \oplus F$.*

Proof: Let π be the projection from M to F. By mapping each generator of f to an inverse image, we can construct a map $s : F \to M$ such that $\pi \circ s = Id_F$. Thus s is an isomorphism onto its image $s(F)$. Now clearly $s(F)$ and N generate M, and have 0 intersection, so we have our splitting.

Proposition 2.15. *A finitely generated torsion-free module is isomorphic to a submodule of a free module.*

Proof: Let e_i be a set of generators of M, and let M_i be the submodule generated by x_j with $j \leq i$. Because M is torsion free M_1 is a free module. We want to successively show that M_{i+1} is isomorphic to a submodule of a free module, given that M_i is. So we need to prove the following: Given N, x, such that $N + (x) = M$, and N is isomorphic to a submodule of a free module, then so is M, to prove the induction step. If $N \cap (x) = 0$, then clearly we can extend by adding another generator to the free module that N is a submodule of and mapping x to it. This map is well defined because there are no relations between x and M.

Now, note that (x) is isomorphic to R, because x is not torsion, so $M \cap (x)$ is isomorphic to some ideal in R, so it is generated by some ax because R is a PID. Then all relations between x where x has a non-zero coefficient are multiples of a single one, $ax = n$, up to relations between elements of N, so to extend our map to M, we just need to extend it in a way compatible with this single relation. We can do this in an obvious way if we divide the generators of the free module N is a submodule of by a, which will still leave us with a free module.

Proposition 2.16. *A finitely generated submodule of a free module is free.*

Proof: Let M be a submodule of R^n. Let e_i be the generators of R^n. Let M_i be the intersection of M with the submodule E_i of R^n generated by e_j, $j \leq i$.

We will prove successively that M_{i+1} is free, supposing that M_i is. $\frac{M \cap E_{i+1}}{M \cap E_i}$ is isomorphic to $\frac{(M \cap E_{i+1}) + E_i}{E_i}$, which is isomorphic to a submodule of $\frac{E_{i+1}}{E_i} = R$, so it is free, being isomorphic to an ideal of R, which is necessarily princiapl. Thus by the splitting lemma above, M_{i+1} is isomorphic to $R \oplus M_i$, or M_i, depending on whether the quotient was free of rank 0 or 1.

Corollary 2.17. *A finitely generated torsion-free R module is free.*

Corollary 2.18. *M^{tor} splits off as a direct summand of M.*

Proposition 2.19. *If R^k is isomorphic to R^l then $l = k$.*

Proof: Let m be a maximal ideal of R. Then $(R/m)^k$ is isomorphic to $R^k/(mR^k)$, which is isomorphic to $R^l/(mR^l)$, which is isomorphic to $(R/mR)^l$. Thus $(R/m)^k$ and $(R/m)^l$ are isomorphic as R/m modules. R/m is a field, so $k = l$ by the fundamental theorem of linear algebra.

2.20. **Decomposing torsion modules.**

2.20.1. *The canonical decomposition.* Now that we have decomposed our module into the direct sum of a free part and a torsion part, we want to decompose the torsion part. The first step is to decompose it into pieces of prime power torsion.

Let M be a finitely generated torsion module. Let e_i be a set of generators, and let a_i be a non-zero element of R that kills e_i. Then $\prod a_i = m$ is a non-zero element of R that kills M. Let $m = p^n q$ where q is not divisible by p.

Proposition 2.21. *Let M be a module killed $n = ab$, such that $(a, b) = 1$. Then if M_a is the submodule killed by a, and M_b the submodule killed by b, then the natural map $M_a \oplus M_b \to M$ is an isomorphism.*

Proof: To show that a module is the direct sum of two submodules, we need to show two things: that the modules have zero intersection, and that together they generate M. Now, since $(a, b) = 1$, can solve $ax + by = 1$. So for any $m \in$ M, $axm + bym = m$. $axm \in M_b$ and $bym \in M_a$, so $M_a + M_b = M$. Now suppose $m \in M_a \cap M_b$. Then $m = xam + bym = 0 + 0 = 0$.

Thus $M = M_q + M_{p^n}$. So by induction on the number of prime factors of m, we can express M as the direct sum of the submodules of elements of prime power torsion, for the different primes. Note that these submodules are well-defined, in the sense that if we have an isomorphism from N to M the p-power torsion part of N maps isomorphically to the p-power torsion part of M. One might also say that this splitting of a torsion module into the parts with torsion at different primes is canonical. More basically, the torsion part of M is well-defined in this sense, and thus the quotient $\frac{M}{M^{\text{tor}}}$ is also well-defined, in that if M is isomorphic to N then this induces an isomorphism $\frac{M}{M^{\text{tor}}} \to \frac{N}{N^{\text{tor}}}$. Thus, the free part of a module M over a PID is well-defined as a module-though not as a submodule of M, since to construct the map from $\frac{M}{M^{\text{tor}}}$ to M we needed to make arbitrary choices.

2.21.1. *Elementary row and column operations.* Finally, we want to decompose modules of prime power torsion into cyclic modules. One way to find this decomposition is to use elementary row and column and operations, reminiscent of linear algebra techniques. Let us pick a finite set of generators of M, x_i. Also, pick n such that M is killed by p^n. Then we can think of M as an $S = \frac{R}{p^n R}$ module. We can consider these generators as giving a surjective map from S^n to M, where n is the number of generators. The kernel K is finitely generated, by the results above, or the exercise at the end of this section, so we can write down a finite set of elements r_i that generate K. These are relations between the different generators, and since these generate the kernel, all relations between the generators are linear combinations of these r_i. These relations, or in other words, this submodule of R^n clearly determines M up to isomorphism. We can thus encode the data of M in a finite array, where the entries in a given column are the entries of a fixed generator x_i in the different relations, and each row corresponds to a relation r_i.

The three following operations clearly do not change the submodule generated by the r_i: permuting the r_i, multiplying a relation by a unit, and adding a multiple of a relation to a different relation. In terms of the array, these correspond to permuting the rows, multiplying a row by a unit, and a multiple of one row to another.

We can also apply similar operations to the set of generators: we can multiply a generator by a unit, permute the generators, or add a multiple of one generator to another. Thus we can permute the columns, multiply a column by a unit, or add a multiple of one column to another.

Now we will show that one can split off a cyclic direct summand of M. Given the array of relations, let a_{jk} be the entry divisible by the least power of p. Now any other a_{jl} can be written as ap^m, where $m \geq n$, and $(a, p) = 1$. Thus a is a unit in S. Thus we can kill a_{jl} by subtracting a multiple of a_{jk}. Thus, by repeating, we can make all entries in jth row except a_{jk} vanish, and by the same process, can kill everything else in the kth column. Then a_{jk} will sit alone in its row and column, so that x_k is not subject to any relations involving the other generators, and is only subject to multiples of the relation $a_{jk}x_k = 0$. Thus the submodule generated by x_k and the submodule generated by the other generators are disjoint, and so we can split off the cyclic submodule generated by x_k as a direct summand. Now by using induction on the number of generators, we can complete our construction of the desired splitting of a finitely generated module over a PID.

Exercise 2.22. *Use elementary row operations to prove that a submodule of a finitely generated module over a Noetherian ring is finitely generated directly, without the results of the subsection above. It may be useful to include at least one zero relation in your array.*

2.22.1. *Decomposition of modules over a Euclidean domain.* It turns out that one can use the above elementary operations to accomplish the entire theorem over Euclidean rings, given the Chinese remainder theorem. Here we give a brief sketch of this method, which can be considered as an algorithm for putting an array of generators and relations into diagonal form.

Let us be given the array representating the relations on a given set of generators for a finitely generated module M. Let a a_{jk} be the element of smallest non-zero Euclidean norm. Then by combining the division algorithm with elementary row and column operations, we can make all of the elements in the same row and column as a_{jk} have smaller norm than a_{jk}. Now repeat this, and the minimum norm keeps shrinking, so that process must eventually stop. At that point, the element of minimum non-zero norm must divide all of the entries in the same row and column, so we can make all of those entries 0. Thus we have a generator x subject only to a relation $dx = 0$, and not having any non-trivial relation with the other generators. Thus clearly we can split off the submodule generated by x as a direct summand, and this submodule will be isomorphic to A/dA. Now repeat. By induction on the number of generators, we are done.

It turns out that by slightly expanding the collection of allowed elementary operations slightly, one can give a similar proof for arbitrary PIDs. For this proof, see Jacobson's Basic Algebra I.