Lecture 32: PIDs–or, similarities between $\mathbb{Z}$ and $F[t]$.

Exercise 32.1. Let R be a commutative ring, and let $x_1, \ldots, x_n$ be a finite collection of elements. Define for yourself the *ideal generated by* $x_1, \ldots, x_n$. Prove that it's an ideal.

Answer: Since we have n elements in R, they uniquely defined a module homomorphism

$$R^{\oplus n} \to R.$$

We let the *ideal generated by* $x_1, \ldots, x_n$ be the image of this homomorphism. By homework, the image of R is a submodule of R, and by definition, a submodule of R is an ideal.

Definition 32.2. We let $(x_1, \ldots, x_n) \subset R$ denote the ideal generated by the elements $x_1, \ldots, x_n$. Explicitly, it is the set of all elements in R that can be expressed as

$$a_1 x_1 + \ldots + a_n x_n$$

for $a_i \in R$.

Chit-chat 32.3. Let F be a field. The point of this lecture is to show that $\mathbb{Z}$ and $F[t]$ are very similar rings. This may be a surprising statement at first glance, but we'll see what we mean. Let me say one important thing: There is an analogy between

(1) The *size* of an integer (or the log of the size of an integer), and
(2) The *degree* of a polynomial.

For instance, for any two integers $x, y \in \mathbb{Z}$, we have that

$$\log(|xy|) = \log|x| + \log|y|.$$

And for any two polynomials in $F[t]$, we have that

$$\deg(fg) = \deg f + \deg g.$$

The *size* of integers allows us to use induction when we want to prove statements about all integers. Though we've taken log above, log preserves order, so the multiplicative property above is still useful for inductive proofs. Likewise, the *degree* of a polynomial will allow us to prove statements about all polynomials by induction.

1. Review

Let me make some preliminary definitions, one of which is new.

Definition 32.4. Let R be a commutative ring. A *zero divisor* is an element $x \in R$ such that

$$xy = 0$$

for some $y \neq 0$.

Example 32.5. Here are some simple examples:

(1) If R has more than one element, then 0 is always a zero divisor, since $0y = 0$ for any $y \in R$. (R needs to have more than one element to guarantee that y can be chosen to be non-zero.)
(2) If $R = \mathbb{Z}/n\mathbb{Z}$ where n is not prime, then choose two integers x, y so that $xy = n$, where neither x nor y is ± 1. Then $\overline{x}$ and $\overline{y}$ are zero divisors in R, for $\overline{x} \neq 0, \overline{y} \neq 0$, but $\overline{xy} = \overline{n} = 0$.

And from last time:

Definition 32.6. A commutative ring R is called a *principal ideal domain*, or *PID*, if

(1) If $I \subset R$ is any ideal, then $I = (x)$ for some $x \in R$.
(2) The only zero divisor in R is 0.

Remark 32.7. The word *domain* means there are no non-zero zero divisors. Sometimes you'll hear the term *integral domain*, which means a commutative ring with no non-zero zero divisors.

The "principal ideal" part of the term means that every ideal is "principal"—i.e., generated by one element.

Example 32.8. By far these are the two most important examples of principal ideal domains:

(1) $R = \mathbb{Z}$. We know any subgroup of $\mathbb{Z}$ is of the form $n\mathbb{Z}$ for some n. Moreover, $n\mathbb{Z} = (n)$, since by definition, any element of $n\mathbb{Z}$ is of the form an for some integer a. Since any ideal is in particular a subgroup of R, we have that every ideal in $\mathbb{Z}$ is principal.
(2) $R = F[t]$ for F a field. Then $F[t]$ is a PID by a theorem I stated last time, which I now recall:

Theorem 32.9. Let F be a field. Then any ideal $I \subset F[t]$ is generated by a single element.

This is a problem in your take-home midterm.

2. The Euclidean algorithm

One major reason that $\mathbb{Z}$ and $F[t]$ are such similar rings is that they both have a division-remainder algorithm, or the *Euclidean algorithm*. Recall the following statement, which we have known since the cradle:

Theorem 32.10 (Divisions and remainders for integers)**.** Let x be an integer, and n any other integer. Then there exists integers q, r such that

$$x = nq + r$$

where $0 \leq r < n$.

Remark 32.11. We used this heavily when we proved that the only subgroups of $\mathbb{Z}$ are of the form $n\mathbb{Z}$.

The following is the analogous statement for polynomials, where (log of) the size of an integer is replaced by the degree of a polynomial.

Theorem 32.12. Let F be a field, and $g \in F[t]$ a polynomial with coefficients in F. Then for any polynomial $f \in F[t]$, there exists polynomials $q, r \in F[t]$ so that

$$g = fq + r$$

where $0 \leq \deg r < \deg f$.

Remark 32.13. That is, we can always divide a polynomial g by another polynomial f, and look at the remainder.

Chit-chat 32.14. I'll go over the proof of this. I think this isn't a standard part of high school math, and a lot of people only learn this through math competitions. I, for instance, never learned this until I was in college!

Proof. If $\deg g < \deg f$, we are finished, simply by setting

$$g = f0 + g.$$

That is, we can't divide a smaller-degree polynomial by a bigger-degree polynomial, so we just end up dividing trivially, and the remainder is g itself. So we need to prove the case when $\deg g \geq \deg f$.

We proceed by induction on the degree of the polynomial g. That is, having fixed f, we have seen that the statement is true for all g with $\deg g < \deg f$ (the base cases). We will assume it true for all g with $\deg g \leq e - 1$, and prove it true for those g with $\deg g = e$.

Let

$$f = a_d t^d + \ldots + a_1 t + a_0, \qquad a_d \neq 0$$

and

$$g = b_e t^e + \ldots + b_1 t + b_0, \qquad b_e \neq 0$$

so that f and g are degree d and e polynomials, respectively. Since both $a_d, b_e \in F$ are non-zero, and since F is a field, there exists a unique number q_{e-d} so that

$$q_{e-d} a_d = b_e.$$

So consider the polynomial

$$q_{e-d} f = b_e t^d + q_{e-d} a_{d-1} t^{d-1} + \ldots + q_{e-d} a_1 t + q_{e-d} a_0.$$

Multiply this polynomial by t^{e-d} to obtain

$$q_{e-d} t^{e-d} f = b_e t^e + q_{e-d} a_{d-1} t^{e-1} + \ldots q_{e-d} a_1 t^{e-d+1} + q_{e-d} a_0 t^{e-d}.$$

Note that this polynomial has the same degree as g, and the same highest-degree coefficient b_e that g has. So we can subtract it from g to obtain a lower-degree polynomial, $g' := g - q_{e-d} t^{e-d} f$. By induction on degree, there is a polynomial Q and a polynomial r so that

$$g' = Qf + r$$

where r has degree less than f. Then we can write

$$\begin{aligned} g &= (q_{e-d} t^{d-e}) f + g - (q_{e-d} t^{d-e}) f \\ &= (q_{e-d} t^{d-e}) f + g' \\ &= (q_{e-d} t^{d-e}) f + Qf + r \\ &= (q_{e-d} t^{d-e} + Q) f + r \end{aligned}$$

so set

$$q = q_{e-d} t^{d-e} + Q,$$

and we have

$$g = qf + r$$

where $\deg r < \deg f$. We are finished. □

3. Primes and factorization in PIDs

We've used the word "unit" in passing, but I want to write it down for the record.

Definition 32.15. An element $x \in R$ is called a *unit* if there is some $y \in R$ for which

$$xy = yx = 1_R.$$

Example 32.16. The units of $\mathbb{Z}$ are the elements ± 1. Likewise, the units of a field F are the non-zero elements of F.

Proposition 32.17. Let $R = F[t]$. Then the units of R are the constant, non-zero polynomials.

PROOF. If $fg = 1$, we must have that $\deg f + \deg g = \deg 1 = 0$. Hence both $\deg f$ and $\deg g$ must be zero—i.e., f and g must be constant. But constant polynomials form a subring $F \subset F[t]$, so two constant polynomials can multiply to one if and only if they are non-zero (since any non-zero element in F has a multiplicative inverse). □

Now I want to generalize the notion of being a prime in $\mathbb{Z}$ to arbitrary rings.

Definition 32.18. An element $x \in R$ is called *prime*, or *irreducible*, if

(1) x is not a unit, and
(2) the only elements dividing x are units, or unit multiples of x. Explicitly, if
$$x = ab$$
for some $a, b \in R$, then either a or b must be a unit.

Example 32.19. Here are some examples of primes in rings:

(1) Let $R = \mathbb{Z}$. If x is a prime number, or the negative of a prime number, then the only numbers dividing x are ± 1 and $\pm x$. Necessarily, if $x = ab$, then either a or b must equal ± 1, which are the units of $\mathbb{Z}$. Hence the prime elements of $\mathbb{Z}$ (under this definition) are prime numbers or their negatives. Note that zero is not a unit.
(2) Let $R = F[t]$. Then the only units of $F[t]$ are constant, non-zero polynomials. So f is a prime, or irreducible, if and only if the only polynomials dividing f have equal degree to f, or are constant polynomials.
(3) As a subexample, if $\deg f = 1$, then f is irreducible. For if $gh = f$, then $\deg g + \deg h = \deg f = 1$. But this means that one of g or h must have degree 0. That is, *any linear polynomial is irreducible.*

Theorem 32.20 (Unique factorization for PIDs)**.** Let R be a PID. Then for any non-zero element $x \in R$, there exists a finite collection of distinct prime elements $p_1, \ldots, p_k \in R$ so that
$$x = p_1^{n_1} p_2^{n_2} \ldots p_k^{n_k}, \qquad n_i \geq 1$$
and so that no p_i is a unit multiple of p_j for $i \neq j$. The n_i are unique, and the elements p_i are unique up to multiplication by units and reordering.

Example 32.21. As a consequence:

(1) If $R = \mathbb{Z}$, recall that a prime element of R is simply a prime number, or a negative of a prime number. Thus the theorem is saying that any integer $x \in \mathbb{Z}$ can be written as a product of powers of primes:

$$x = p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}.$$

If each p_i is taken to be a positive prime number, this is often called the prime factorization of x. In context of the theorem, however, note we could replace p_1 and p_2 by $-p_1$ and $-p_2$, and we would still be able to express x as a product of powers of primes. In this sense, the choice of the p_i is only unique up to multiplying by units. Of course, for the integers, we can choose to order each p_i so that the $p_i < p_{i+1}$ and we have a preferred ordering, but this is not true in general PIDs.

(2) If $R = F[t]$, this is saying that every polynomial can be written as a product of irreducible polynomials p_i:

$$f = p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}.$$

(3) As an example, if $F = \mathbb{C}$, then any polynomial can be written as a product of linear polynomials:

$$f = (t - \alpha_1)^{n_1} \dots (t - \alpha_k)^{n_k}.$$

I caution you that for other fields, we may not be able to decompose f into *linear* polynomials. See your midterm.

Proof of the Theorem. Let $f \in R$. If f is irreducible, we are finished, just by writing $p_1 = f$ and

$$f = p_1.$$

Otherwise, f factors into a product

$$f = a_1 b_1$$

where a_1, b_1 are not unit multiples of f. This means that f is contained in the ideal (a_1, b_1) generated by a_1 and b_1. If each of a_1 and b_1 is prime, we are finished; otherwise we can factor them further so $a_1 = c_2 d_2$ and $b = e_2 f_2$. This means that f is in the ideal generated by (c_2, d_2, e_2, f_2). We see this process might go on indefinitely, and we might keep dividing up elements further without ever reaching a prime element. However, since R is a PID, each of these ideals—$(f), (a_1, b_1), (c_2, d_2, e_2, f_2)$, etc.—are generated by a single element. So let's let g_n denote the element that generates the ideal we obtain at the nth step of this subdivision process. We have that

$$(f) = (g_0) \subset (g_1) \subset (g_2) \subset \dots.$$

Well, note that

$$J := \bigcup_{i=0}^{\infty} (g_i)$$

is an ideal again! (Try proving this yourself.) Since R is a PID, this must also be generated by a single element:

$$J = (g)$$

but $g \in J$ if and only if $g \in (g_n)$ for some finite n, by definition of union. This means that

$$(g_n) \subset J = (g) \subset (g_n).$$

That is,

$$(g_n) = J$$

so the subdivision process stops at some finite n. This means that f can be written as a product of primes

$$f = \prod p_i.$$

□