Math 533 Winter 2021, Lecture 13: Root adjunction and finite fields

website: `https://www.cip.ifi.lmu.de/~grinberg/t/21w/`

1. Monoid algebra and polynomials ([DF, Chapter 9]) (cont'd)

1.1. Adjoining roots (cont'd)

In Lecture 12, we have seen a few examples of the construction in which we start with a commutative ring R and a polynomial $b \in R[x]$, and construct the quotient ring $R[x]/b$. To recall, the bottom line of this construction is "throw a new root of b into the ring R and see what happens". Often, this produces a ring extension of R – i.e., a larger ring that contains R as a subring. (For example, this happens if $R = \mathbb{R}$ and $b = x^2 + 1$; this is how Cardano defined the complex numbers.) However, this doesn't always go well. Sometimes, what happens instead is that the ring R collapses to a trivial ring (e.g., if $b = 1$) or at least becomes smaller (e.g., we have $(\mathbb{Z}/6)[x]/(2x-1) \cong \mathbb{Z}/3$). Sometimes, the ring loses some of its properties: e.g., if we throw a new root of $x^2 - 1$ into the field $\mathbb{Q}$, then the resulting ring $\mathbb{Q}[x]/\left(x^2-1\right)$ not only fails to be a field, but even fails to be an integral domain (indeed, we have seen that this ring is isomorphic to $\mathbb{Q} \times \mathbb{Q}$).

Let us put these things in order. First, let us show that the residue class $\overline{x}$ in $R[x]/b$ is a root of b, so that our construction really creates a root of b:

Proposition 1.1.1. Let $b \in R[x]$ be a polynomial. (Recall that R is still a fixed commutative ring.)

(a) The projection map

$$\begin{aligned} \pi_b : R[x] &\to R[x]/b, \\ p &\mapsto \overline{p} \end{aligned}$$

is an $R[x]$-algebra morphism, and thus an R-algebra morphism.

(b) The map[1]

$$\begin{aligned} R &\to R[x]/b, \\ r &\mapsto \overline{r} \end{aligned}$$

is an R-algebra morphism.

(c) We have $p[\overline{x}] = \overline{p}$ for any $p \in R[x]$.

(d) The element $\overline{x} \in R[x]/b$ is a root of b.

None of this is difficult to prove, but the following proposition will make the proof (even) more comfortable:

Proposition 1.1.2 ("Polynomials commute with algebra morphisms")**.** Let A and B be two R-algebras. Let $f : A \to B$ be an R-algebra morphism. Let $a \in A$. Let $p \in R[x]$ be a polynomial. Then,

$$f(p[a]) = p[f(a)].$$

Proof of Proposition 1.1.2. Let us give a proof by example: Set $p = 5x^4 + x^3 + 7x^1$. Then, $p[a] = 5a^4 + a^3 + 7a^1$ and $p[f(a)] = 5f(a)^4 + f(a)^3 + 7f(a)^1$. Thus, the claim we have to prove rewrites as

$$f\left(5a^4 + a^3 + 7a^1\right) = 5f(a)^4 + f(a)^3 + 7f(a)^1.$$

But this follows easily from the fact that f is an R-algebra morphism: Indeed,

$$\begin{aligned} f\left(5a^4 + a^3 + 7a^1\right) &= f\left(5a^4\right) + f\left(a^3\right) + f\left(7a^1\right) && \text{(since } f \text{ respects addition)} \\ &= 5f\left(a^4\right) + f\left(a^3\right) + 7f\left(a^1\right) && \text{(since } f \text{ respects scaling)} \\ &= 5f(a)^4 + f(a)^3 + 7f(a)^1 && \text{(since } f \text{ respects powers)}. \end{aligned}$$

The rigorous proof in the general case is LTTR. □

Proof of Proposition 1.1.1. **(a)** This follows from the general fact (proved back in Lecture 12) that the canonical projection from an R-algebra to its quotient is an R-algebra morphism. Note that we need to apply this fact to $R[x]$ instead of R here, in order to conclude that the map in question is an $R[x]$-algebra morphism.

(b) The map

$$\begin{aligned} R &\to R[x]/b, \\ r &\mapsto \overline{r} \end{aligned}$$

is the composition of the projection map π_b from part **(a)** with the inclusion map

$$\begin{aligned} R &\to R[x], \\ r &\mapsto r = rx^0. \end{aligned}$$

[1] Note the difference between the maps in part **(a)** and in part **(b)**: The map in part **(a)** takes as input a polynomial $p \in R[x]$, whereas the map in part **(b)** takes as input a scalar $r \in R$ (and treats it as a constant polynomial, i.e., as $rx^0 \in R[x]$). If you regard R as a subring of $R[x]$, you can thus view the map in part **(b)** as a restriction of the map in part **(a)**.

Thus, it is a composition of two R-algebra morphisms (since both π_b and the inclusion map are R-algebra morphisms). Hence, it is an R-algebra morphism itself[2]. This proves Proposition 1.1.1 **(b)**.

(c) Here is an abstract argument: Let $p \in R[x]$. The projection map π_b from Proposition 1.1.1 **(a)** is an R-algebra morphism (by Proposition 1.1.1 **(a)**). Hence, Proposition 1.1.2 (applied to $A = R[x]$ and $B = R[x]/b$ and $a = x$ and $f = \pi_b$) yields

$$\pi_b(p[x]) = p[\pi_b(x)]. \tag{1}$$

However, the definition of π_b yields $\pi_b(p[x]) = \overline{p[x]} = \overline{p}$ (since $p[x] = p$) and $\pi_b(x) = \overline{x}$. Hence, (1) rewrites as $\overline{p} = p[\overline{x}]$. This proves Proposition 1.1.1 **(c)**.

Alternatively, you can prove it directly by writing p as $p = \sum_{i=0}^{n} p_i x^i$ with $p_i \in R$. (Indeed, if you do this, then the claim rewrites as $\sum_{i=0}^{n} p_i \overline{x}^i = \overline{\sum_{i=0}^{n} p_i x^i}$; but this is an easy consequence of how the quotient $R[x]/b$ was defined.)

(d) Proposition 1.1.1 **(c)** (applied to $p = b$) yields $b[\overline{x}] = \overline{b} = \overline{0}$ (since $b \in bR[x]$). In other words, $\overline{x}$ is a root of b. This proves Proposition 1.1.1 **(d)**. □

Next, for a large class of polynomials $b \in R[x]$ (including the monic ones, and all the nonzero polynomials over a field), we are going to show how $R[x]/b$ looks like as an R-module:

Theorem 1.1.3. Let $m \in \mathbb{N}$. Let $b \in R[x]$ be a polynomial of degree m such that its leading coefficient $[x^m]\, b$ is a unit. Then:

(a) Each element of $R[x]/b$ can be uniquely written in the form

$$a_0\overline{x^0} + a_1\overline{x^1} + \cdots + a_{m-1}\overline{x^{m-1}} \qquad \text{with } a_0, a_1, \ldots, a_{m-1} \in R.$$

(b) The m vectors $\overline{x^0}, \overline{x^1}, \ldots, \overline{x^{m-1}}$ form a basis of the R-module $R[x]/b$. Thus, this R-module $R[x]/b$ is free of rank $m = \deg b$.

(c) Assume that $m > 0$. Then, the R-algebra morphism[3]

$$\begin{aligned} R &\to R[x]/b, \\ r &\mapsto \overline{r} \end{aligned}$$

is injective. Therefore, R can be viewed as an R-subalgebra (thus a subring) of $R[x]/b$ if we identify each $r \in R$ with its image $\overline{r} \in R[x]/b$.

(d) In particular, under the assumption that $m > 0$, there exists a commutative ring that contains R as a subring and that contains a root of b.

[2]Indeed, there is an easy fact (which we never stated, but which is completely straightforward to prove after what we have seen) that any composition of two R-algebra morphisms is itself an R-algebra morphism.

[3]This is the map from Proposition 1.1.1 **(b)**.

Proof. **(a)** Let $\alpha \in R[x]/b$. Then, $\alpha = \overline{a}$ for some polynomial $a \in R[x]$. Consider this a. The division-with-remainder theorem for polynomials (see Lecture 12) tells us that there is a unique pair (q, r) of polynomials in $R[x]$ such that

$$a = qb + r \qquad \text{and} \qquad \deg r < \deg b.$$

Consider this pair (q, r). Then, in $R[x]/b$, we have $\overline{a} = \overline{r}$ (since $a = qb + r$ entails $a - r = qb = bq \in bR[x]$).

We have $\deg r < \deg b = m$; thus, we can write r in the form $r = r_0 x^0 + r_1 x^1 + \cdots + r_{m-1} x^{m-1}$ for some $r_0, r_1, \ldots, r_{m-1} \in R$. Consider these $r_0, r_1, \ldots, r_{m-1}$. We have

$$\begin{aligned}\alpha = \overline{a} = \overline{r} &= \overline{r_0 x^0 + r_1 x^1 + \cdots + r_{m-1} x^{m-1}} \\ &\qquad \left(\text{since } r = r_0 x^0 + r_1 x^1 + \cdots + r_{m-1} x^{m-1}\right) \\ &= r_0 \overline{x^0} + r_1 \overline{x^1} + \cdots + r_{m-1} \overline{x^{m-1}}\end{aligned}$$

(since the scaling and the addition of the quotient algebra $R[x]/b$ were inherited from $R[x]$).

Thus, we have represented our $\alpha \in R[x]/b$ in the form

$$a_0 \overline{x^0} + a_1 \overline{x^1} + \cdots + a_{m-1} \overline{x^{m-1}} \qquad \text{with } a_0, a_1, \ldots, a_{m-1} \in R$$

(namely, for $a_i = r_i$). It remains to show that this representation is unique.

This can be shown by walking the above proof backwards and using the uniqueness part of the division-with-remainder theorem. Here are the details: Assume that

$$\alpha = b_0 \overline{x^0} + b_1 \overline{x^1} + \cdots + b_{m-1} \overline{x^{m-1}} \qquad \text{with } b_0, b_1, \ldots, b_{m-1} \in R$$

is some representation of α in the above form. We must then show that this representation is actually the representation that we constructed above – i.e., that we have $b_i = r_i$ for each $i \in \{0, 1, \ldots, m-1\}$.

Indeed, define a polynomial $s \in R[x]$ by $s = b_0 x^0 + b_1 x^1 + \cdots + b_{m-1} x^{m-1}$. Then, $\deg s \leq m - 1 < m = \deg b$. Also,

$$\overline{a} = \alpha = b_0 \overline{x^0} + b_1 \overline{x^1} + \cdots + b_{m-1} \overline{x^{m-1}} = \overline{b_0 x^0 + b_1 x^1 + \cdots + b_{m-1} x^{m-1}} = \overline{s}$$

(since $b_0 x^0 + b_1 x^1 + \cdots + b_{m-1} x^{m-1} = s$). In other words, $a - s \in bR[x]$. In other words,

$$a - s = bd \qquad \text{for some } d \in R[x].$$

Consider this d. Thus, $a = bd + s = db + s$. Now, the pair (d, s) is a pair of polynomials in $R[x]$ satisfying $a = db + s$ and $\deg s < \deg b$. This means that it satisfies the exact conditions that the pair (q, r) was asked to satisfy. However, the division-with-remainder theorem for polynomials said that the pair (q, r)

satisfying those conditions was unique. Hence, we must have $(d, s) = (q, r)$ (since (d, s) satisfies the same conditions as (q, r)). Thus, $d = q$ and $s = r$.

Now,

$$b_0 x^0 + b_1 x^1 + \cdots + b_{m-1} x^{m-1} = s = r = r_0 x^0 + r_1 x^1 + \cdots + r_{m-1} x^{m-1}.$$

Comparing coefficients in these polynomials, we conclude that $b_i = r_i$ for each $i \in \{0, 1, \ldots, m-1\}$ (since $\left(x^0, x^1, x^2, \ldots\right)$ is a basis of the R-module $R[x]$). This is what we needed to show. Theorem 1.1.3 **(a)** is thus proved.

(b) This is just Theorem 1.1.3 **(a)**, rewritten in terms of modules and bases. In some more detail:

- Each element of $R[x]/b$ can be written in the form

  $$a_0 \overline{x^0} + a_1 \overline{x^1} + \cdots + a_{m-1} \overline{x^{m-1}} \qquad \text{with } a_0, a_1, \ldots, a_{m-1} \in R$$

  (according to Theorem 1.1.3 **(a)**). In other words, each element of $R[x]/b$ is an R-linear combination of $\overline{x^0}, \overline{x^1}, \ldots, \overline{x^{m-1}}$. Thus, the list $\left(\overline{x^0}, \overline{x^1}, \ldots, \overline{x^{m-1}}\right)$ spans the R-module $R[x]/b$.

- Each element of $R[x]/b$ can be **uniquely** represented in the form

  $$a_0 \overline{x^0} + a_1 \overline{x^1} + \cdots + a_{m-1} \overline{x^{m-1}} \qquad \text{with } a_0, a_1, \ldots, a_{m-1} \in R$$

  (according to Theorem 1.1.3 **(a)**). Hence, in particular, the zero vector $\overline{0} \in R[x]/b$ can be **uniquely** represented in this form. But it is clear how to represent $\overline{0}$ in this form: We just write

  $$\overline{0} = 0\overline{x^0} + 0\overline{x^1} + \cdots + 0\overline{x^{m-1}}.$$

  Since we have just said that $\overline{0}$ can be **uniquely** represented in this form, we thus conclude that this is the **only** way to represent $\overline{0}$ in this form. In other words, if $\overline{0}$ has been represented in the form $a_0 \overline{x^0} + a_1 \overline{x^1} + \cdots + a_{m-1} \overline{x^{m-1}}$ with $a_0, a_1, \ldots, a_{m-1} \in R$, then we must have $a_0 = a_1 = \cdots = a_{m-1} = 0$. In other words, if $a_0, a_1, \ldots, a_{m-1} \in R$ satisfy $a_0 \overline{x^0} + a_1 \overline{x^1} + \cdots + a_{m-1} \overline{x^{m-1}} = \overline{0}$, then $a_0 = a_1 = \cdots = a_{m-1} = 0$. But this is saying precisely that the list $\left(\overline{x^0}, \overline{x^1}, \ldots, \overline{x^{m-1}}\right)$ is R-linearly independent.

Thus, we have shown that the list $\left(\overline{x^0}, \overline{x^1}, \ldots, \overline{x^{m-1}}\right)$ is R-linearly independent and spans $R[x]/b$. In other words, this list is a basis of $R[x]/b$. This proves Theorem 1.1.3 **(b)**.

(c) We know (from Proposition 1.1.1 **(b)**) that the map

$$\begin{aligned} R &\to R[x]/b, \\ r &\mapsto \overline{r} \end{aligned}$$

is an R-algebra morphism. We only need to show that it is injective. It clearly suffices to show that its kernel is $\{0\}$ (because we know that an R-module morphism is injective if and only if its kernel is $\{0\}$).

So let r be in the kernel of this morphism. We must prove that $r = 0$.

Since r is in the kernel of the above morphism, we have $\overline{r} = 0$ in $R[x]/b$. In other words, r is a multiple of b. In other words, $r = bc$ for some polynomial $c \in R[x]$. Consider this c. From $r = bc$, we obtain $\deg r = \deg(bc) = \deg b + \deg c$ (by one of the propositions of Lecture 12, since the leading coefficient of b is a unit). Thus, $\deg b + \deg c = \deg r \leq 0$ (since r is constant). However, $\deg b = m > 0$ by assumption. Hence, $\deg b > 0 \geq \deg b + \deg c$. This entails $\deg c < 0$. This means that $c = 0$, whence $r = b\underbrace{c}_{=0} = 0$.

Forget that we fixed r. We thus have proved that if r is in the kernel of our morphism, then $r = 0$. Hence, the kernel of our morphism is $\{0\}$ (since 0 is clearly in its kernel). Thus, the morphism is injective, and Theorem 1.1.3 **(c)** is proven.

(d) Assume that $m > 0$. The ring $R[x]/b$ contains a root of b (namely, $\overline{x}$, according to Proposition 1.1.1 **(d)**), and also contains "a copy of R", in the sense that there is an injective ring morphism from R to $R[x]/b$ (namely, the one we constructed in Theorem 1.1.3 **(c)**). If we replace this copy of R by the original R (by replacing each $\overline{r} \in R[x]/b$ with the corresponding $r \in R$), then we obtain a ring that contains R as a subring but also contains a root of b. This proves Theorem 1.1.3 **(d)**. □

Let us summarize: We have generalized the construction of $\mathbb{C}$. Namely, we have found a way to "adjoin" a root of a polynomial $b \in R[x]$ to a commutative ring R by forming the quotient ring $R[x]/b$. This latter ring is always a commutative ring and an R-algebra. Moreover, if b is "nice" (that is, we have $\deg b > 0$, and the leading coefficient of b is a unit), then this latter ring $R[x]/b$ will contain R as a subring (by Theorem 1.1.3 **(c)**) and also will be a free R-module of rank $\deg b$ (by Theorem 1.1.3 **(b)**). If b is not as "nice", then the ring $R[x]/b$ may fail to contain R as a subring (even though it still is an R-algebra), and may be smaller than R or even trivial.

1.2. Field extensions from adjoining roots

Let F be a field. Then, any non-constant univariate polynomial $b \in F[x]$ is "nice" in the sense of the preceding paragraph, so that $F[x]/b$ is a commutative ring that contains F as a subring and that contains a root of b. When will this ring $F[x]/b$ be a field?

We first state a simple fact about the units of $F[x]$:

Proposition 1.2.1. The units of the polynomial ring $F[x]$ are precisely the nonzero constant polynomials.

Proof. Any nonzero constant polynomial is a unit of $F[x]$ (since it is a unit of F). Conversely, any unit of $F[x]$ must be a nonzero constant polynomial[4]. $\square$

Recall (from Lecture 12) that $F[x]$ is a Euclidean domain, hence a PID (by Lecture 6), hence a UFD (by Lecture 7). Furthermore, an element $p \in F[x]$ is prime[5] if and only if it is irreducible (by one of the results in Lecture 6, since $F[x]$ is a PID). The notion of "irreducible" in $F[x]$ is precisely the classical concept of an irreducible polynomial:

Proposition 1.2.2. Let $p \in F[x]$. Then, p is irreducible if and only if p is non-constant and cannot be written as a product of two non-constant polynomials.

Proof. The definition of "irreducible" says that p is irreducible if and only if p is nonzero and not a unit and has the property that whenever $a, b \in F[x]$ satisfy $ab = p$, at least one of a and b must be a unit.

In view of Proposition 1.2.1, this can be rewritten as follows: p is irreducible if and only if p is nonzero and not a nonzero constant polynomial and has the property that whenever $a, b \in F[x]$ satisfy $ab = p$, at least one of a and b must be a nonzero constant polynomial.

We can declutter this statement (e.g., "nonzero and not a nonzero constant polynomial" can be shortened to "non-constant"), and thus obtain the following: p is irreducible if and only if p is non-constant and has the property that whenever $a, b \in F[x]$ satisfy $ab = p$, at least one of a and b must be constant. In other words, p is irreducible if and only if p is non-constant and cannot be written as a product of two non-constant polynomials. $\square$

Now, we can characterize when a quotient ring of the form $F[x]/p$ is a field:

Theorem 1.2.3. Let $p \in F[x]$. Then, the ring $F[x]/p$ is a field if and only if p is irreducible.

For example, the irreducible polynomial x^2+1 over $\mathbb{R}$ yields the field $\mathbb{R}[x]/\left(x^2+1\right)$ (which is $\cong \mathbb{C}$), but the non-irreducible polynomial x^2-1 over $\mathbb{R}$ yields the non-field $\mathbb{R}[x]/\left(x^2-1\right) \cong \mathbb{R} \times \mathbb{R}$.

Theorem 1.2.3 is analogous to the fact that $\mathbb{Z}/n$ is a field (for a positive integer n) if and only if n is prime. Just like the latter fact, it is a particular case of the following general property of PIDs:

[4] *Proof.* Let u be a unit of $F[x]$. We must show that u is a nonzero constant polynomial.

We know that u is a unit of $F[x]$; hence, there exists some $v \in F[x]$ satisfying $uv = 1$. Consider this v. From $uv = 1 \neq 0$, we obtain $u \neq 0$, so that u is nonzero. Hence, $\deg(uv) = \deg u + \deg v$ (by a proposition from Lecture 12, since F is an integral domain). Moreover, from $uv = 1$, we obtain $\deg(uv) = \deg 1 = 0$, so that $0 = \deg(uv) = \deg u + \underbrace{\deg v}_{\geq 0} \geq \deg u$,

which entails that u is constant. Thus, u is a nonzero constant polynomial, qed.

[5] See Lecture 6 for the definitions of prime and irreducible elements of an integral domain.

Theorem 1.2.4. Let R be a PID. Let $p \in R$. Then, the ring R/p is a field if and only if p is irreducible.

Proof. $\Longrightarrow$: LTTR.

$\Longleftarrow$: Assume that p is irreducible. We must show that R/p is a field.

First of all, p is not a unit (since p is irreducible), so that 1 is not a multiple of p. Hence, $\overline{1} \neq \overline{0}$ in R/p. In other words, the ring R/p is not trivial. This ring is furthermore commutative (since R is commutative).

Now, let $\alpha \in R/p$ be a nonzero. We shall prove that α is a unit.

Write α as $\overline{a}$ for some $a \in R$. Then, $\overline{a} = \alpha \neq 0$ in R/p (since α is nonzero), so that $p \nmid a$.

Now, recall that R is a PID, so that any ideal of R is principal. In particular, this entails that the ideal $aR + pR$ is principal. In other words, there exists some $g \in R$ such that $aR + pR = gR$. Consider this g. According to Lecture 6, we can conclude from $aR + pR = gR$ that g is a gcd of a and p. Thus, $g \mid a$ and $g \mid p$.

However, p is irreducible; hence, every divisor of p is either a unit or associate to p (indeed, this is easily seen to be a consequence of the definition of "irreducible"[6]). Thus, g is either a unit or associate to p (since $g \mid p$). However, g cannot be associate to p (because if g was associate to p, then we would have $p \mid g \mid a$, which would contradict $p \nmid a$). Hence, g must be a unit. So it has an inverse g^{-1}.

But $g = g \cdot 1 \in gR = aR + pR$. In other words, there exist two elements $u, v \in R$ such that $g = au + pv$. Consider these u, v. Then, $g = au + pv = ua + pv$, so that

$$\overline{g} = \overline{ua + pv} = \overline{ua} \qquad \text{(since } pv \in pR\text{)}$$

in R/p. Therefore,

$$\overline{g^{-1}u} \cdot \overline{a} = \overline{g^{-1}} \cdot \overline{u} \cdot \overline{a} = \overline{g^{-1}} \cdot \underbrace{\overline{ua}}_{=\overline{g}} = \overline{g^{-1}} \cdot \overline{g} = \overline{g^{-1}g} = \overline{1}.$$

But this equality shows that $\overline{g^{-1}u}$ is an inverse of $\overline{a}$ in the ring R/p (because we know that R/p is commutative, so that we don't need to check $\overline{a} \cdot \overline{g^{-1}u} = \overline{1}$ as well). Thus, $\overline{a}$ is a unit. In other words, α is a unit (since $\alpha = \overline{a}$).

Forget that we fixed α. We thus have shown that any nonzero $\alpha \in R/p$ is a unit. In other words, R/p is a field (since R/p is a nontrivial commutative ring). $\square$

As a consequence of Theorem 1.2.4, we can now "adjoin" a root of an irreducible polynomial to a field without destroying its field-ness: Namely, if we have a field F and some irreducible polynomial $b \in F[x]$, then the quotient ring $F[x]/b$ will be a field that contains F as a subring and that contains a root of

[6] Indeed: If d is a divisor of p, then there exists an $e \in R$ such that $p = de$. Consider this e. From $p = de$, we conclude that d or e is a unit (since p is irreducible). In the first case, d is a unit; in the second case, d is associate to p.

b. This generalizes Cardano's definition of $\mathbb{C}$, but can also be applied to adjoin roots to fields other than $\mathbb{R}$.

Example 1.2.5. The polynomial $x^2+1 \in (\mathbb{Z}/3)[x]$ is irreducible. (Indeed, $\mathbb{Z}/3$ being a finite field, we could verify this by going through all nonconstant polynomials of degree < 2 and checking that none of them divides x^2+1.)

Thus, Theorem 1.2.4 yields that $(\mathbb{Z}/3)[x] / \left(x^2+1\right)$ is a field. This field is a free $\mathbb{Z}/3$-module of rank 2 (by Theorem 1.1.3 **(b)**), and thus is isomorphic to $(\mathbb{Z}/3)^2 = (\mathbb{Z}/3) \times (\mathbb{Z}/3)$ as a $\mathbb{Z}/3$-module (but not as a ring, of course). Hence, the size of this field is $\left|(\mathbb{Z}/3)^2\right| = |\mathbb{Z}/3|^2 = 3^2 = 9$.

Thus, we have found a finite field of size 9. We have obtained it from $\mathbb{Z}/3$ in the same way as $\mathbb{C}$ was obtained from $\mathbb{R}$: by adjoining a square root of -1.

Incidentally, this field can also be constructed as $\mathbb{Z}[i]/3$.

2. Finite fields

2.1. Basics

Example 1.2.5 may make you wonder: what finite fields can we find? We know that for each prime p, the quotient ring $\mathbb{Z}/p$ is a field of size p; thus, we know a finite field of any prime size. Now we have found a finite field of size 9, too. What other finite fields exist?

Let's first grab the low-hanging fruit:

Proposition 2.1.1. Let p be a prime number. Then:

(a) There exists an irreducible polynomial $b \in (\mathbb{Z}/p)[x]$ of degree 2 over $\mathbb{Z}/p$.

(b) There exists a finite field of size p^2.

Proof. We write F for $\mathbb{Z}/p$. Thus, F is a field and satisfies $|F| = |\mathbb{Z}/p| = p$.

(a) If $p = 2$, then we can take $b = x^2+x+1$; it is easy to check that this b is irreducible.

Thus, WLOG assume that $p \neq 2$. Hence, $p > 2$. Thus, $\overline{1} \neq \overline{-1}$ in $\mathbb{Z}/p$. In other words, $\overline{1} \neq \overline{-1}$ in F (since $\mathbb{Z}/p = F$). The map

$$\begin{aligned} F &\to F, \\ a &\mapsto a^2 \end{aligned}$$

is not injective (since $\overline{1}^2 = \overline{-1}^2$ but $\overline{1} \neq \overline{-1}$), and thus cannot be surjective (by the pigeonhole principle). Thus, there exists some $u \in F$ that is not in the image of this map. In other words, there exists some $u \in F$ that is not a square. Consider such a u. Then, the polynomial $x^2 - u$ has no roots in F.

Now it is not hard to prove that the polynomial $x^2 - u$ is irreducible.[7] This proves Proposition 2.1.1 **(a)** (since $x^2 - u \in (\mathbb{Z}/p)[x]$ is an irreducible polynomial of degree 2).

(b) Proposition 2.1.1 **(a)** yields that there exists an irreducible polynomial $b \in F[x]$ of degree 2 over F (since $\mathbb{Z}/p = F$). Consider this b. Theorem 1.2.3 (applied to b instead of p) then yields that the ring $F[x]/b$ is a field. Moreover, $F[x]/b$ is a free F-module of rank 2 (by Theorem 1.1.3 **(b)**), and thus is isomorphic to F^2 as a F-module, and therefore has size $\left|F^2\right| = |F|^2 = p^2$ (since $|F| = p$). Hence, $F[x]/b$ is a finite field of size p^2. This proves Proposition 2.1.1 **(b)**. □

By more complicated but somewhat similar arguments[8], we can also see that there exists a finite field of size p^3 for any prime p. This suggests generalizing to p^m; but this is much harder. Indeed, a nonconstant polynomial over F of degree ≤ 3 will always be irreducible if it has no roots in F (check this!); however, for polynomials of degree ≥ 4, this is no longer the case (fun exercise: prove that the polynomial $x^4 + 4 \in \mathbb{Q}[x]$ is not irreducible, despite of course not having any roots over $\mathbb{Q}$). Thus, our trick for finding irreducible polynomials will no longer work for degrees > 3. We can still find a field of size p^4 by applying our trick twice (first get a finite field of size p^2, then proceed to find an irreducible

[7] *Proof.* Assume that we have written $x^2 - u$ as a product fg of two non-constant polynomials $f, g \in F[x]$. We shall derive a contradiction.

Indeed, we have assumed that $x^2 - u = fg$; hence, $\deg\left(x^2 - u\right) = \deg(fg) = \deg f + \deg g$ (since F is an integral domain). Thus, $\deg f + \deg g = \deg\left(x^2 - u\right) = 2$. Since $\deg f$ and $\deg g$ are positive integers (because f and g are non-constant), this entails that $\deg f$ and $\deg g$ must equal 1 (since the only pair of positive integers that add up to 2 is $(1, 1)$). Thus, in particular, $\deg f = 1$. Hence, $f = ax + b$ for some $a, b \in F$ with $a \neq 0$. Consider these a, b. From $f = ax + b$, we obtain $f\left[\frac{-b}{a}\right] = a \cdot \frac{-b}{a} + b = 0$. Thus, the polynomial f has a root in F (namely, $\frac{-b}{a}$). Hence, the polynomial $x^2 - u$ has a root in F as well (indeed, $f \mid fg = x^2 - u$, so that every root of f is also a root of $x^2 - u$). This contradicts the fact that the polynomial $x^2 - u$ has no roots in F.

Thus, we have found a contradiction stemming from our assumption that $x^2 - u$ is a product fg of two non-constant polynomials $f, g \in F[x]$. Hence, $x^2 - u$ cannot be written as such a product. In other words, $x^2 - u$ is irreducible (since $x^2 - u$ is a non-constant polynomial). Qed.

[8] Not too similar! It is not true that the map

$$F \to F,$$
$$a \mapsto a^3$$

is always non-surjective when $F = \mathbb{Z}/p$ for $p > 3$. Instead, you have to argue the existence of an irreducible polynomial $b \in (\mathbb{Z}/p)[x]$ of degree 3 over $\mathbb{Z}/p$ by a counting argument: Show that the total number of monic degree-3 polynomials in $(\mathbb{Z}/p)[x]$ is p^3, whereas the total number of monic degree-3 polynomials in $(\mathbb{Z}/p)[x]$ that can be written as a product of a degree-1 and a degree-2 polynomial is smaller than p^3; thus, at least one monic degree-3 polynomial cannot be written as such a product.

polynomial of degree 2 over that field), and by induction we can find fields of sizes $p^8, p^{16}, p^{32}, \ldots$. But we don't get a field of size p^5 this way.

So do such fields exist?

2.2. The characteristic of a field

Leaving prime powers aside for a moment, what about fields of size 6 ? It turns out that such fields don't exist, for a fairly simple reason. Fields have an important invariant, the so-called **characteristic**:

Definition 2.2.1. Let F be a field. The **characteristic** of F is an integer called $\operatorname{char} F$, which is defined as follows:

- If there exists a positive integer n such that $n \cdot 1_F = 0$, then $\operatorname{char} F$ is defined to be the **smallest** such n.
- If such an n does not exist, then $\operatorname{char} F$ is defined to be 0.

Roughly speaking, $\operatorname{char} F$ is "how often you have to add 1_F to itself to obtain 0" (with the caveat that we define it to be 0 if you never obtain 0 by adding 1_F to itself). Here are some examples:

- We have $\operatorname{char} \mathbb{Q} = 0$, since there exists no positive integer n such that $n \cdot 1_{\mathbb{Q}} = 0$. For the same reason, $\operatorname{char} \mathbb{R} = 0$ and $\operatorname{char} \mathbb{C} = 0$.
- For any prime p, we have $\operatorname{char}(\mathbb{Z}/p) = p$. Indeed, $p \cdot 1_{\mathbb{Z}/p} = p \cdot \overline{1} = \overline{p \cdot 1} = \overline{p} = \overline{0}$ in $\mathbb{Z}/p$, but every positive integer $n < p$ satisfies $n \cdot 1_{\mathbb{Z}/p} = n \cdot \overline{1} = \overline{n \cdot 1} = \overline{n} \neq \overline{0}$ in $\mathbb{Z}/p$.
- For our fields F of size p^2 or p^3, we also have $\operatorname{char} F = p$, since they contain $\mathbb{Z}/p$ as subrings.