Math 533 Winter 2021, Lecture 12: Polynomial rings

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

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

1.1. Univariate polynomials (cont'd)

1.1.1. Degrees and coefficients (cont'd)

Reminder: R denotes a fixed commutative ring.[1]

Let me recall that the zero polynomial $0 = 0x^0 + 0x^1 + 0x^2 + \cdots$ has degree $-\infty$ by definition. This $-\infty$ is not a number, but we agree that $-\infty$ is smaller than any integer and does not change if you add an integer to it (i.e., we have $(-\infty) + m = -\infty$ for any $m \in \mathbb{Z}$).

Remark 1.1.1. Let $n \in \mathbb{N}$. Then,

$$\begin{aligned}
&\{f \in R[x] \mid \deg f \leq n\} \\
&= \left\{f \in R[x] \mid f = a_0x^0 + a_1x^1 + \cdots + a_nx^n \text{ for some } a_0, a_1, \ldots, a_n \in R\right\} \\
&= \operatorname{span}\left(x^0, x^1, \ldots, x^n\right).
\end{aligned}$$

This is clearly an R-submodule of $R[x]$.

Corollary 1.1.2. Let $p, q \in R[x]$. Then,

$$\deg(p+q) \leq \max\{\deg p, \deg q\} \qquad \text{and} \tag{1}$$
$$\deg(p-q) \leq \max\{\deg p, \deg q\} \tag{2}$$

Proof. Let $n = \max\{\deg p, \deg q\}$. Let N denote the subset $\{f \in R[x] \mid \deg f \leq n\}$ of $R[x]$. Then, we know from Remark 1.1.1 that N is an R-submodule of $R[x]$. Moreover, the definition of n shows that $\deg p \leq n$, so that $p \in N$. Similarly,

[1] By the way, it absolutely is possible to define polynomials over a noncommutative ring, if you are sufficiently careful. (In particular, this includes polynomials over matrix rings; these are rather useful in linear algebra. The indeterminates in such polynomials commute with all elements of R.) We have defined the notion of an R-algebra only for commutative rings R, but there are ways to adapt it to the general setup; alternatively, it is possible to redo the construction of the polynomial ring by hand without using R-algebras. See [ChaLoi21, §1.3] for the latter approach.

$q \in N$. Hence, $p+q \in N$ (since N is an R-submodule of $R[x]$); in other words, $\deg(p+q) \leq n$. In other words, $\deg(p+q) \leq \max\{\deg p, \deg q\}$ (since $n = \max\{\deg p, \deg q\}$). Similarly, we can find $\deg(p-q) \leq \max\{\deg p, \deg q\}$. This proves Corollary 1.1.2. □

Remark 1.1.3. The polynomials of degree ≤ 0 are precisely the constant polynomials – i.e., the elements of R (embedded into $R[x]$ as explained in Lecture 11).

The following proposition collects some properties of products of univariate polynomials:

Proposition 1.1.4. Let $p, q \in R[x]$.

(a) We have $\deg(pq) \leq \deg p + \deg q$.

(b) We have $\deg(pq) = \deg p + \deg q$ if $p \neq 0$ and the leading coefficient of p is a unit.

(c) We have $\deg(pq) = \deg p + \deg q$ if R is an integral domain.

(d) If $n, m \in \mathbb{N}$ satisfy $n \geq \deg p$ and $m \geq \deg q$, then

$$\left[x^{n+m}\right](pq) = \left[x^n\right](p) \cdot \left[x^m\right](q).$$

(e) If $pq = 0$ and $p \neq 0$ and if the leading coefficient of p is a unit, then $q = 0$.

Corollary 1.1.5. If R is an integral domain, then the polynomial ring $R[x]$ is an integral domain.

Proof of Proposition 1.1.4. We will give an informal "proof by example". Rigorous arguments can be found in various places[2].

Let p and q be two polynomials of degrees $\deg p = 2$ and $\deg q = 3$. Write p and q as $p = ax^2 + bx + c$ and $q = dx^3 + ex^2 + fx + g$ (with $a, b, c, \ldots, g \in R$). Then,

$$\begin{aligned} pq &= \left(ax^2 + bx + c\right)\left(dx^3 + ex^2 + fx + g\right) \\ &= adx^5 + (\text{lower powers of } x). \end{aligned} \tag{3}$$

Thus, $\deg(pq) \leq 5 = 2 + 3 = \deg p + \deg q$. This proves Proposition 1.1.4 **(a)**.

Moreover, $a \neq 0$ (since $\deg p = 2$) and $d \neq 0$ (since $\deg q = 3$). If R is an integral domain, then this entails $ad \neq 0$ and therefore $\deg(pq) = 5$ (by (3)).

[2]Can they? I'm not so sure any more; apparently everyone just handwaves them away or leaves them to the reader (e.g., Bourbaki writes about part **(a)** that "the proof is immediate"). A while ago I have written up proofs for parts **(a)** and **(d)** in [Grinbe19] (where they appear as parts **(a)** and **(b)** of Lemma 3.12), albeit only in the particular case when p is monic (but the proofs can easily be generalized). A generalization of parts **(b)** and **(c)** also appears in [ChaLoi21, Proposition (1.3.12)].

This proves Proposition 1.1.4 **(c)**. On the other hand, if a is a unit, then we also have $ad \neq 0$ (because otherwise, we would have $ad = 0$ and thus $a^{-1} \underbrace{ad}_{=0} = 0$, which would contradict $a^{-1}ad = d \neq 0$) and therefore $\deg(pq) = 5$ (by (3)). This proves Proposition 1.1.4 **(b)** (since a is the leading coefficient of p).

The equality (3) shows that the coefficient of x^5 in pq is ad, and no higher powers of x than x^5 appear in pq. That is, we have $[x^5](pq) = \underbrace{a}_{=[x^2](p)} \underbrace{d}_{=[x^3](q)} = [x^2](p) \cdot [x^3](q)$, and we have $[x^i](pq) = 0$ for all $i > 5$. This quickly yields Proposition 1.1.4 **(d)**.

To prove Proposition 1.1.4 **(e)**, we assume the contrary. Thus, $pq = 0$ and $p \neq 0$ and the leading coefficient of p is a unit, but $q \neq 0$. Then, Proposition 1.1.4 **(b)** yields $\deg(pq) = \underbrace{\deg p}_{\geq 0} + \underbrace{\deg q}_{\geq 0} \geq 0$. However, $pq = 0$, so $\deg(pq) = \deg 0 = -\infty < 0$. These two inequalities clearly contradict each other, and our proof of Proposition 1.1.4 **(e)** is complete. □

Proof of Corollary 1.1.5. Assume that R is an integral domain. Let $p, q \in R[x]$ be nonzero. Then, Proposition 1.1.4 **(c)** yields $\deg(pq) = \underbrace{\deg p}_{\geq 0} + \underbrace{\deg q}_{\geq 0} \geq 0$, and thus $pq \neq 0$ (since $pq = 0$ would yield $\deg(pq) = \deg 0 = -\infty < 0$). Thus, we have shown that $pq \neq 0$ for any nonzero $p, q \in R[x]$. In other words, $R[x]$ is an integral domain. □

If R is not an integral domain, then polynomials over R can behave rather strangely. For example, over $\mathbb{Z}/4$, we have

$$\left(\overline{1} + \overline{2}x\right)^2 = \overline{1} + \overline{4}x + \overline{4}x^2 = \overline{1} \qquad \left(\text{since } \overline{4} = \overline{0}\right).$$

So the degree of a polynomial can decrease when it is squared!

1.1.2. Division with remainder

The most important feature of univariate polynomials is division with remainder:

Theorem 1.1.6 (Division-with-remainder theorem for polynomials)**.** Let $b \in R[x]$ be a nonzero polynomial whose leading coefficient is a unit. Let $a \in R[x]$ be any polynomial.

(a) Then, there is a **unique** pair (q, r) of polynomials in $R[x]$ such that

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

(b) Moreover, this pair satisfies $\deg q \leq \deg a - \deg b$.

The polynomials q and r in Theorem 1.1.6 are called the **quotient** and the **remainder** obtained when dividing a by b. Note that if $\deg a < \deg b$, then the quotient q is 0 whereas the remainder r is a. The quotient and the remainder become interesting when $\deg a \geq \deg b$.

Example 1.1.7. Let us first give a **non-example:** Let $R = \mathbb{Z}$ and $b = 2$ (a constant polynomial) and $a = x$. The leading coefficient of b is not a unit (since 2 is not a unit in $\mathbb{Z}$), so we don't expect the theorem to hold. And indeed: we cannot write $a = qb + r$ with $\deg r < \deg b$. Indeed, this would mean $x = q \cdot 2 + r$ with $\deg r < 0$ (since the constant polynomial 2 has degree $\deg 2 = 0$); but this is impossible, since this would entail $x = q \cdot 2$, which would contradict the fact that x is not divisible by 2.

Proof of Theorem 1.1.6. **(a)** Again, we shall give a proof by example. (For a rigorous proof, see [Grinbe19, Theorem 3.16 and Lemma 3.19] or [Ford21, Theorem 3.6.4] or [ChaLoi21, Theorem (1.3.15)] or [Knapp16, Proposition 1.12] or [DF, §9.2, Thm 3]. Note that some of these sources assume that b is monic, and others assume that R is a field; however, the proofs easily adapt to our general case.)

Let u be the leading coefficient of b. Then, u is a unit (by assumption), so it has an inverse u^{-1}. Scaling the polynomial b by u^{-1} results in a monic polynomial $u^{-1}b$ (since its leading coefficient u gets multiplied by u^{-1}). Thus, we can replace b by the monic polynomial $u^{-1}b$ without changing much (the degree of b remains the same, and the required equality $a = qb + r$ becomes $a = qu \cdot \left(u^{-1}b\right) + r$, so that we have to multiply the q in the desired pair (q, r) by u). Hence, we WLOG assume that b is monic.

We are doing a proof by example, so let us assume that $\deg a = 3$ and $\deg b = 2$. Thus, we can write a and b as $a = cx^3 + dx^2 + ex + f$ and $b = x^2 + gx + h$ for some $c, d, e, f, g, h \in R$ (since b is monic).

We now repeatedly subtract appropriate multiples of b from a in order to decrease its degree:

$$\begin{aligned}
& a = cx^3 + dx^2 + ex + f \\
\Longrightarrow\ & a - (cx)\, b = (d - cg)\, x^2 + (e - ch)\, x + f \\
& \qquad \Big(\text{here, we have subtracted } (cx)\, b \text{ to kill off the } cx^3 \text{ term}\Big) \\
\Longrightarrow\ & a - (cx)\, b - (d - cg)\, b = (e - ch - (d - cg)\, g)\, x + (f - (d - cg)\, h) \\
& \qquad \Big(\text{here, we have subtracted } (d - cg)\, b \text{ to kill off the } (d - cg)\, x^2 \text{ term}\Big).
\end{aligned}$$

Thus,

$$\begin{aligned}
a &= (cx)\, b + (d - cg)\, b + (e - ch - (d - cg)\, g)\, x + (f - (d - cg)\, h) \\
&= (cx + (d - cg))\, b + (e - ch - (d - cg)\, g)\, x + (f - (d - cg)\, h).
\end{aligned}$$

Setting $q := cx + (d - cg)$ and $r := (e - ch - (d - cg)\, g)\, x + (f - (d - cg)\, h)$, we can rewrite this as

$$a = qb + r.$$

Note that $\deg r < \deg b$ (since any polynomial of degree $\geq \deg b$ could still be reduced further by subtracting a multiple of b from it).

Thus we have found a pair (q, r) of polynomials satisfying

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

It remains to prove its uniqueness. In other words, we have to prove that if (q_1, r_1) and (q_2, r_2) are two pairs of polynomials satisfying

$$\begin{aligned} &a = q_1 b + r_1 \quad \text{and} \quad \deg r_1 < \deg b \quad \text{and} \\ &a = q_2 b + r_2 \quad \text{and} \quad \deg r_2 < \deg b, \end{aligned}$$

then $(q_1, r_1) = (q_2, r_2)$. To prove this, we fix two such pairs (q_1, r_1) and (q_2, r_2). Then, we have

$$q_1 b + r_1 = a = q_2 b + r_2,$$

so that $q_1 b - q_2 b = r_2 - r_1$. In other words, $(q_1 - q_2)\, b = r_2 - r_1$. Hence,

$$\begin{aligned} \deg\left((q_1 - q_2)\, b\right) &= \deg\left(r_2 - r_1\right) \leq \max\left\{\deg r_2, \deg r_1\right\} \qquad \text{(by (2))} \\ &< \deg b \qquad \left(\text{since } \deg r_2 < \deg b \text{ and } \deg r_1 < \deg b\right). \end{aligned}$$

However, recall that the leading coefficient of b is a unit. Hence, if the polynomial $q_1 - q_2$ was nonzero, then Proposition 1.1.4 **(b)** would entail

$$\deg\left((q_1 - q_2)\, b\right) = \underbrace{\deg\left(q_1 - q_2\right)}_{\geq 0} + \deg b \geq \deg b,$$

which would contradict $\deg\left((q_1 - q_2)\, b\right) < \deg b$. So $q_1 - q_2$ must be zero; i.e., we have $q_1 = q_2$. Using $(q_1 - q_2)\, b = r_2 - r_1$, we furthermore obtain $r_2 - r_1 = \underbrace{(q_1 - q_2)}_{=0} b = 0$, so that $r_1 = r_2$. Hence, $(q_1, r_1) = (q_2, r_2)$. This completes the proof of the uniqueness of (q, r). Thus, Theorem 1.1.6 **(a)** is proved.

(b) You can obtain Theorem 1.1.6 **(b)** by a careful analysis of the construction of the pair (q, r) in our proof of part **(a)**. Indeed, each of the terms of q was originally a factor that we multiplied to b in order to reduce a; however, the highest power of x in a was $x^{\deg a}$, so the factors we used did not contain any powers of x higher than $x^{\deg a - \deg b}$.

Alternatively, you can prove Theorem 1.1.6 **(b)** independently of part **(a)**: Let (q, r) be a pair of polynomials in $R\left[x\right]$ such that

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

We must prove that $\deg q \leq \deg a - \deg b$. Assume the contrary. Thus, $\deg q > \deg a - \deg b$. Therefore, in particular, $q \neq 0$ (since $q = 0$ would entail $\deg q =$

$\deg 0 = -\infty \leq \deg a - \deg b$), so that $\deg q \geq 0$. However, the leading coefficient of b is a unit; thus, Proposition 1.1.4 **(b)** yields that

$$\deg(bq) = \deg b + \deg q > \deg a \qquad (\text{since } \deg q > \deg a - \deg b).$$

Also,

$$\deg(bq) = \deg b + \underbrace{\deg q}_{\geq 0} \geq \deg b > \deg r \qquad (\text{since } \deg r < \deg b).$$

Combining these two inequalities, we obtain

$$\deg(bq) > \max\{\deg a, \deg r\}.$$

But from $a = qb + r$, we obtain $a - r = qb = bq$, so that $bq = a - r$. Hence,

$$\deg(bq) = \deg(a - r) \leq \max\{\deg a, \deg r\} \qquad (\text{by (2)}),$$

which contradicts $\deg(bq) > \max\{\deg a, \deg r\}$. This contradiction shows that our assumption was wrong; thus, Theorem 1.1.6 **(b)** is proven. □

We record two automatic corollaries of Theorem 1.1.6:

Corollary 1.1.8. Let $b \in R[x]$ be a monic polynomial. Let $a \in R[x]$ be any polynomial.

(a) Then, there is a **unique** pair (q, r) of polynomials in $R[x]$ such that

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

(b) Moreover, this pair satisfies $\deg q \leq \deg a - \deg b$.

Proof. The polynomial b is monic; thus, its leading coefficient is a unit (since 1 is a unit). Hence, Theorem 1.1.6 applies. □

Corollary 1.1.9. Let F be a field. Let $b \in F[x]$ be any nonzero polynomial. Let $a \in F[x]$ be any polynomial.

(a) Then, there is a **unique** pair (q, r) of polynomials in $F[x]$ such that

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

(b) Moreover, this pair satisfies $\deg q \leq \deg a - \deg b$.

Proof. The polynomial b is nonzero; thus, its leading coefficient is a unit (since any nonzero element of the field F is a unit). Hence, Theorem 1.1.6 applies. □

The following simple proposition is the polynomial analogue of the classical fact that a positive integer b divides an integer a if and only if the remainder a leaves when divided by b is 0:

Proposition 1.1.10. Let $b \in R[x]$ be a nonzero polynomial whose leading coefficient is a unit. Let $a \in R[x]$ be any polynomial. Let q and r be the quotient and the remainder obtained when dividing a by b. Then, we have the logical equivalence $(r = 0) \iff (b \mid a \text{ in } R[x])$.

Proof. The definition of quotient and remainder yields $a = qb + r$. Hence, if $r = 0$, then $a = qb + \underbrace{r}_{=0} = qb$ and thus $b \mid a$ in $R[x]$. This proves the "$\Longrightarrow$" direction of the required equivalence. It thus remains to prove the "$\Longleftarrow$" direction.

So we assume that $b \mid a$ in $R[x]$. We need to show that $r = 0$.

We have assumed $b \mid a$ in $R[x]$. In other words, there exists a $c \in R[x]$ such that $a = cb$. Consider this c. We have $a = cb = bc = bc + 0$ and $\deg 0 = -\infty < \deg b$. Thus, $(c, 0)$ is a pair $(\widetilde{q}, \widetilde{r})$ of polynomials in $F[x]$ such that $a = \widetilde{q}b + \widetilde{r}$ and $\deg \widetilde{r} < \deg b$. But (q, r) is also such a pair (by the definition of quotient and remainder). However, Corollary 1.1.8 shows that there is a **unique** such pair. In particular, any two such pairs must be identical. Thus, the two pairs (q, r) and $(c, 0)$ must be identical. That is, we have $q = c$ and $r = 0$. In particular, $r = 0$; this completes the proof of the "$\Longleftarrow$" direction. Proposition 1.1.10 is thus proven. □

1.1.3. Roots

Let's now talk about roots of polynomials.

Definition 1.1.11. Let A be an R-algebra. Let $p \in R[x]$. An element $a \in A$ is said to be a **root** of p if $p(a) = 0$.

This is a rather wide notion of roots. For example, the matrix $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathbb{Q}^{2\times 2}$ is a root of the polynomial $x^2 \in \mathbb{Q}[x]$, since the square of this matrix is 0.

Proposition 1.1.12. Let p be a polynomial in $R[x]$. Let $a \in R$. Then, we have the following logical equivalence:

$$(a \text{ is a root of } p) \iff (x - a \mid p \text{ in } R[x]).$$

Before we prove this proposition, let us repeat a theorem from Lecture 11 that will be used in the proof:

Theorem 1.1.13. Let A be an R-algebra. Let $a \in A$. Then, the map

$$\begin{aligned} R[x] &\to A, \\ p &\mapsto p[a] \end{aligned}$$

is an R-algebra morphism. In particular, for any two polynomials $p, q \in R[x]$, we have

$$(pq)[a] = p[a] \cdot q[a]; \tag{4}$$
$$(p+q)[a] = p[a] + q[a]. \tag{5}$$

Proof of Proposition 1.1.12. The polynomial $x - a$ is monic. Hence, Corollary 1.1.8 **(a)** (applied to p and $x - a$ instead of a and b) shows that there is a **unique** pair (q, r) of polynomials in $R[x]$ such that

$$p = q \cdot (x - a) + r \qquad \text{and} \qquad \deg r < \deg (x - a).$$

Consider this pair (q, r). From $\deg r < \deg (x - a) = 1$, we see that $\deg r \leq 0$, which means that r is a constant. In other words, $r \in R$.

Now, let us substitute a for x on both sides of the equality $p = q \cdot (x - a) + r$. Thus we get

$$p[a] = q[a] \cdot (a - a) + r[a]. \tag{6}$$

It is worth going through this equality in some more detail. Namely, we have $p = q \cdot (x - a) + r$, so that

$$\begin{aligned} p[a] &= (q \cdot (x - a) + r)[a] = (q \cdot (x - a))[a] + r[a] \\ &\qquad \left(\begin{array}{c} \text{by (5), applied to } R, q \cdot (x-a) \text{ and } r \\ \text{instead of } A, p \text{ and } q \end{array} \right) \\ &= q[a] \cdot \underbrace{(x - a)[a]}_{\substack{= a - a \\ \text{(by the definition} \\ \text{of an evaluation)}}} + r[a] \\ &\qquad \left(\begin{array}{c} \text{by (4), applied to } R, q \text{ and } x - a \\ \text{instead of } A, p \text{ and } q \end{array} \right) \\ &= q[a] \cdot (a - a) + r[a]. \end{aligned}$$

Thus, (6) has been proven in detail.

Now, (6) becomes

$$p[a] = q[a] \cdot \underbrace{(a - a)}_{=0} + r[a] = r[a] = r \qquad (\text{since } r \text{ is a constant}).$$

Now, we have the following chain of equivalences:

$$\begin{aligned} (a \text{ is a root of } p) &\iff (p[a] = 0) \qquad (\text{by the definition of a root}) \\ &\iff (r = 0) \qquad (\text{since } p[a] = r) \\ &\iff (x - a \mid p \text{ in } R[x]) \end{aligned}$$

(by Proposition 1.1.10, applied to $x - a$ and p instead of b and a). This proves Proposition 1.1.12. $\square$

The following theorem is often known as the **easy half of the FTA (Fundamental Theorem of Algebra)**:

Theorem 1.1.14. Let R be an integral domain. Let $n \in \mathbb{N}$. Then, any nonzero polynomial $p \in R[x]$ of degree $\leq n$ has at most n roots in R. (We are not counting the roots with multiplicity here.)

Proof. We induct on n. The *base case* ($n = 0$) is obvious (indeed, a nonzero polynomial of degree ≤ 0 must be constant, and thus cannot have any roots to begin with).

Induction step: Let m be a positive integer. Assume (as the induction hypothesis) that Theorem 1.1.14 holds for $n = m - 1$. We must prove that Theorem 1.1.14 holds for $n = m$.

So let $p \in R[x]$ be a nonzero polynomial of degree $\leq m$. We must prove that p has at most m roots in R.

Indeed, assume the contrary. Thus, p has $m+1$ distinct roots $a_1, a_2, \ldots, a_{m+1}$ in R (and possibly more, but we will only need these $m+1$).

In particular, a_{m+1} is a root of p, so that we have $x - a_{m+1} \mid p$ in $R[x]$ (by Proposition 1.1.12, applied to $a = a_{m+1}$). That is, there exists a polynomial $q \in R[x]$ such that $p = (x - a_{m+1}) \cdot q$. Consider this q. Now, it is easy to see that $a_1, a_2, \ldots, a_m$ are roots of q (indeed, this uses the fact that $a_1, a_2, \ldots, a_{m+1}$ are **distinct** roots of p and that R is an integral domain[3]). Hence, the polynomial q has at least m roots in R (since these m roots $a_1, a_2, \ldots, a_m$ are distinct). Also, the polynomial q is nonzero (since otherwise, we would have $q = 0$ and thus $p = (x - a_{m+1}) \cdot \underbrace{q}_{=0} = 0$, contradicting the fact that p is nonzero).

However, Proposition 1.1.4 **(c)** (or Proposition 1.1.4 **(b)**, if you wish) yields

$$\deg\left((x - a_{m+1}) \cdot q\right) = \underbrace{\deg\left(x - a_{m+1}\right)}_{=1} + \deg q = 1 + \deg q,$$

[3]Here is the proof in detail: Let $i \in \{1, 2, \ldots, m\}$. We must show that a_i is a root of q. Note that $i \neq m+1$ (since $i \in \{1, 2, \ldots, m\}$) and thus $a_i \neq a_{m+1}$ (since $a_1, a_2, \ldots, a_{m+1}$ are distinct). Substituting a_i for x in the equality $p = (x - a_{m+1}) \cdot q$, we find

$$p[a_i] = (a_i - a_{m+1}) \cdot q[a_i]$$

(formally speaking, this relies on a similar argument as we used to prove (6)). Hence,

$$(a_i - a_{m+1}) \cdot q[a_i] = p[a_i] = 0 \qquad \text{(since } a_i \text{ is a root of } p\text{)}.$$

Since R is an integral domain, this entails that we have $a_i - a_{m+1} = 0$ or $q[a_i] = 0$. Since $a_i - a_{m+1} = 0$ is impossible (because $a_i \neq a_{m+1}$), we thus conclude that $q[a_i] = 0$. In other words, a_i is a root of q. Qed.

so that

$$\deg q = \deg\left(\underbrace{(x-a_{m+1})\cdot q}_{=p}\right) - 1 = \underbrace{\deg p}_{\substack{\leq m \\ \text{(since } p \text{ has degree } \leq m\text{)}}} - 1 \leq m-1.$$

In other words, the polynomial q has degree $\leq m-1$. Hence, by the induction hypothesis, we can apply Theorem 1.1.14 to q and $m-1$ instead of p and n. We thus conclude that q has at most $m-1$ roots in R. This contradicts the fact that q has at least m roots in R (which we have shown above). This contradiction completes the induction step, and so we are done proving Theorem 1.1.14. □

1.1.4. Application to $\mathbb{Z}/p$

Let me show an application of this theorem to finite fields. We will need the following fact:

Theorem 1.1.15 (Fermat's little theorem, short FℓT)**.** Let p be a prime number. Let $a \in \mathbb{Z}$. Then, $a^p \equiv a \bmod p$.

Proof. If $a \equiv 0 \bmod p$, then this is clear (since we have $a^p \equiv 0^p \equiv 0 \equiv a \bmod p$ in this case). So let us WLOG assume that $a \not\equiv 0 \bmod p$. Hence, the residue class $\overline{a} \in \mathbb{Z}/p$ is nonzero. Therefore, $\overline{a}$ is a unit of the ring $\mathbb{Z}/p$ (since $\mathbb{Z}/p$ is a field, so that every nonzero element of $\mathbb{Z}/p$ is a unit). In other words, $\overline{a} \in (\mathbb{Z}/p)^\times$.

However, the units of the ring $\mathbb{Z}/p$ are $\overline{1}, \overline{2}, \ldots, \overline{p-1}$ (again because every nonzero element of $\mathbb{Z}/p$ is a unit). Thus, in particular, there are $p-1$ of them. This shows that the group $(\mathbb{Z}/p)^\times$ has order $p-1$. Hence, Lagrange's theorem (from group theory)[4] shows that $u^{p-1} = 1$ for each $u \in (\mathbb{Z}/p)^\times$. Applying this to $u = \overline{a}$, we obtain $\overline{a}^{p-1} = 1$. Hence, $\overline{a^p} = \overline{a}^p = \underbrace{\overline{a}^{p-1}}_{=1}\cdot\overline{a} = \overline{a}$. In other words, $a^p \equiv a \bmod p$. Theorem 1.1.15 is thus proven. □

Now, let us reword this theorem in the language of polynomials. First, we consider the polynomial

$$x^p - x \in (\mathbb{Z}/p)[x].$$

Theorem 1.1.15 yields that all evaluations of this polynomial at elements of $\mathbb{Z}/p$ are 0 (in fact, for each $a \in \mathbb{Z}$, we have $(x^p - x)[\overline{a}] = \overline{a}^p - \overline{a} = \overline{a^p - a} = \overline{0}$, since Theorem 1.1.15 yields $a^p \equiv a \bmod p$). The polynomial itself is not zero, and this is no surprise: It is a degree-p polynomial, so it can afford to have p roots in $\mathbb{Z}/p$ without being forced by Theorem 1.1.14 to be the zero polynomial. However, it is "dangerously close"; if its degree was even a little bit smaller

[4] Recall that this theorem says the following: If G is a finite group of order m (for some $m \in \mathbb{N}$), then $u^m = 1$ for each $u \in G$ (where we are writing G multiplicatively, so that 1 denotes the neutral element of G).

than p, then we would obtain a contradiction. We can exploit this to extract a nice corollary. To this end, we define the more sophisticated polynomial

$$f := (x^p - x) - \underbrace{\prod_{u \in \mathbb{Z}/p} (x-u)}_{= (x-\overline{0})(x-\overline{1})\cdots(x-\overline{(p-1)})} \qquad \in (\mathbb{Z}/p)[x].$$

This polynomial f has degree $\leq p-1$ (check this![5]). But it still has (at least) p roots in $\mathbb{Z}/p$; indeed, the p elements $\overline{0}, \overline{1}, \ldots, \overline{p-1}$ are roots of f, since each $a \in \{0, 1, \ldots, p-1\}$ satisfies

$$f[\overline{a}] = \underbrace{(\overline{a}^p - \overline{a})}_{\substack{=0 \\ \text{(by Theorem 1.1.15)}}} - \underbrace{\prod_{u \in \mathbb{Z}/p} (\overline{a} - u)}_{\substack{=0 \\ \text{(since one of the factors} \\ \text{in this product is } \overline{a}-\overline{a}=0)}} = 0 - 0 = 0.$$

If the polynomial f was nonzero, then this would contradict Theorem 1.1.14 (since $\mathbb{Z}/p$ is a field and thus an integral domain). Hence, f must be zero. Since we defined f to be the difference $x^p - x - \prod_{u \in \mathbb{Z}/p} (x-u)$, we thus conclude that $x^p - x = \prod_{u \in \mathbb{Z}/p} (x-u)$. Let us state this as a proposition:

Proposition 1.1.16. Let p be a prime number. Then,

$$x^p - x = \prod_{u \in \mathbb{Z}/p} (x-u) \qquad \text{in the polynomial ring } (\mathbb{Z}/p)[x].$$

Now, let us milk this for consequences. We have

$$\begin{aligned}
\prod_{u \in \mathbb{Z}/p} (x-u) &= (x-\overline{0})(x-\overline{1})\cdots\left(x-\overline{(p-1)}\right) \\
&= x \underbrace{(x-\overline{1})(x-\overline{2})\cdots\left(x-\overline{(p-1)}\right)}_{\substack{= (-\overline{1})(-\overline{2})\cdots(-\overline{(p-1)})\cdot x^0 + (\text{higher powers of } x) \\ \text{(here, ``higher powers of } x\text{'' means ``any powers of } x \text{ higher than } x^0\text{'')}} \\
&= x\left((-\overline{1})(-\overline{2})\cdots\left(-\overline{(p-1)}\right)\cdot x^0 + (\text{higher powers of } x)\right) \\
&= (-\overline{1})(-\overline{2})\cdots\left(-\overline{(p-1)}\right)\cdot x^1 + (\text{higher powers of } x).
\end{aligned}$$

[5] *Proof.* Both polynomials $x^p - x$ and $\prod_{u \in \mathbb{Z}/p} (x-u)$ have degree p and leading coefficient 1. Thus, when you subtract the polynomial $\prod_{u \in \mathbb{Z}/p} (x-u)$ from $x^p - x$, the x^p terms of both polynomials cancel, and what remains is a linear combination of $x^0, x^1, \ldots, x^{p-1}$ – that is, a polynomial of degree $\leq p-1$.

Thus, the coefficient of x^1 in the polynomial $\prod\limits_{u \in \mathbb{Z}/p} (x-u)$ is

$$\begin{aligned}\left(-\overline{1}\right)\left(-\overline{2}\right)\cdots\left(-\overline{(p-1)}\right) &= \overline{(-1)^{p-1}\cdot(1\cdot 2\cdot\cdots\cdot(p-1))} \\ &= \overline{(-1)^{p-1}\cdot(p-1)!}.\end{aligned}$$

On the other hand, the coefficient of x^1 in the polynomial $x^p - x$ is $\overline{-1}$ (since $p > 1$). But these two coefficients must be equal (since Proposition 1.1.16 says that the polynomials $\prod\limits_{u \in \mathbb{Z}/p} (x-u)$ and $x^p - x$ are equal). Hence, $\overline{(-1)^{p-1}\cdot(p-1)!} = \overline{-1}$. In other words,

$$(-1)^{p-1}\cdot(p-1)! \equiv -1 \bmod p.$$

If we multiply this congruence by $(-1)^{p-1}$, then the left hand side becomes $(p-1)!$ (since $(-1)^{p-1}\cdot(-1)^{p-1} = 1$), and thus we get

$$(p-1)! \equiv (-1)^{p-1}\cdot(-1) = (-1)^p \equiv -1 \bmod p$$

(by Theorem 1.1.15, applied to $a = -1$). Thus, we have proved Wilson's theorem (from Lecture 7) again!

1.1.5. $F[x]$ is a Euclidean domain

Let us go back to the case of polynomials over a general field. I next record an abstract consequence of Corollary 1.1.9 **(a)**:

Theorem 1.1.17. Let F be a field. Then, the polynomial ring $F[x]$ is a Euclidean domain (with the Euclidean norm $N : F[x] \to \mathbb{N}$ being "almost" the degree function, in the sense that $N(p) = \max\{\deg p, 0\}$ for any $p \in F[x]$), thus a PID, thus a UFD.

Proof. Define a map $N : F[x] \to \mathbb{N}$ by

$$N(p) = \max\{\deg p, 0\} \qquad \text{for any } p \in F[x].$$

Then, Corollary 1.1.9 **(a)** shows that N is a Euclidean norm on the ring $F[x]$. Hence, $F[x]$ is a Euclidean domain (since Corollary 1.1.5 shows that $F[x]$ is an integral domain). Thus, $F[x]$ is a PID (since we know from Lecture 6 that every Euclidean domain is a PID) and a UFD (since we know from Lecture 7 that every PID is a UFD). □

Note that the "UFD" part of Theorem 1.1.17 is not a very constructive result; there is no general algorithm for actually finding a prime factorization of a polynomial (i.e., for factoring a polynomial into irreducible polynomials)

that works over any field. There are reasonably good algorithms for prime factorization in $\mathbb{Q}[x]$, however.

Theorem 1.1.17 entails, in particular, that univariate polynomials over a field have gcds and lcms. Moreover, the analogue of Bezout's theorem holds:

Theorem 1.1.18 (Bezout's theorem for polynomials)**.** Let F be a field. Let $a, b \in F[x]$ be two polynomials. Then, for any choice of $\gcd(a,b)$, there exist two polynomials $u, v \in F[x]$ such that $ua + vb = \gcd(a,b)$.

Proof. This is a general fact that holds in every PID (but not in every UFD). To wit, let us set $R = F[x]$, and recall that R is a PID (by Theorem 1.1.17). Recall how we proved the existence of a gcd (in Lecture 6): Namely, we argued that there exists a $c \in R$ satisfying $aR + bR = cR$ (since R is a PID, so that the ideal $aR + bR$ of R must be principal), and then we proved that this c is a gcd of a and b. Now, assume that we have chosen some gcd of a and b, and denoted it by $\gcd(a,b)$. This $\gcd(a,b)$ is not necessarily identical to c, but it is clearly associate to c, since we have shown (in Lecture 6) that any two gcds of a and b are associate. Thus, $\gcd(a,b) = cu$ for some unit u of R. Consider this u. Now,

$$\gcd(a,b) = c\underbrace{u}_{\in R} \in cR = aR + bR.$$

In other words, there exist some $u, v \in R$ such that $\gcd(a,b) = au + bv$. In other words, there exist some $u, v \in R$ such that $\gcd(a,b) = ua + vb$. This proves Theorem 1.1.18. □

Warning: Multivariate polynomial rings (like $\mathbb{Q}[x,y]$) are not PIDs (and thus not Euclidean domains either). For example, the ideal $x\mathbb{Q}[x,y] + y\mathbb{Q}[x,y]$ is not principal. (Check this! This ideal is easily seen to consist of all polynomials whose constant term (= coefficient of x^0y^0) is 0, but these polynomials are not the multiples of a single polynomial.) However, multivariate polynomial rings over fields (and, more generally, over UFDs) are still UFDs. This is a deeper result than the ones we have proved above (see, e.g., [DF, §9.3, Corollary 8] or [Ford21, Theorem 3.7.4] or [ChaLoi21, Corollary (2.6.7)] or [Knapp16, Corollary 8.21 and Remark after it] for proofs). As a consequence, polynomials over a field (or a UFD) have gcds; however, they don't generally satisfy Bezout's theorem unless the polynomials are univariate polynomials over a field.

Univariate polynomial rings over non-fields (like $\mathbb{Z}[x]$) behave similarly: They are not PIDs, but they are UFDs when the base ring is a UFD. (That is, if R is a UFD, then so is $R[x]$.)

1.2. Intermezzo: quotients of R-algebras

In preparation for the next section, let me quickly introduce quotients of R-algebras. I have previously defined quotients of rings modulo ideals, and quotients of R-modules modulo submodules. These two concepts can be combined to obtain quotients of R-algebras modulo ideals:

Theorem 1.2.1. Let A be an R-algebra. Let I be an ideal of A. Then:

(a) The ideal I is also an R-submodule of A.

(b) The quotient ring A/I and the quotient R-module A/I fit together to form an R-algebra.

(c) The canonical projection $A \to A/I$ (which sends each $a \in A$ to its residue class $\overline{a} = a + I$) is an R-algebra morphism (from the original R-algebra A to the R-algebra A/I that we just constructed in part **(b)**).

Proof. **(a)** We already know that I is closed under addition and contains zero (since I is an ideal). So we must only show that I is closed under scaling. In other words, we must show that $ri \in I$ for each $r \in R$ and $i \in I$. But this is easy: If $r \in R$ and $i \in I$, then

$$r \underbrace{i}_{=1_A \cdot i} = r \cdot 1_A \cdot i = \underbrace{(r \cdot 1_A)}_{\in A} \cdot \underbrace{i}_{\in I} \in I$$

(since I is an ideal of A).

(b) LTTR. (You just need to verify the "scale-invariance of multiplication" axiom, but this is straightforward.)

(c) We already know that this canonical projection is a ring morphism and an R-module morphism; thus, it is an R-algebra morphism. □

Let us next recall the universal property of quotient rings (Lecture 3), which is the tool of choice from constructing ring morphisms out of a quotient ring:

Theorem 1.2.2 (Universal property of quotient rings)**.** Let R be a ring. Let I be an ideal of R.

Let S be a ring. Let $f : R \to S$ be a ring morphism. Assume that $f(I) = 0$ (this is shorthand for saying that $f(a) = 0$ for all $a \in I$). Consider the canonical projection $\pi : R \to R/I$. Then, there is a unique ring morphism $f' : R/I \to S$ satisfying $f = f' \circ \pi$.

We can adapt this theorem to R-algebras with just trivial modifications (alas, we have to rename R and S as A and B, since R already means something different):

Theorem 1.2.3 (Universal property of quotient algebras)**.** Let A be an R-algebra. Let I be an ideal of A.

Let B be an R-algebra. Let $f : A \to B$ be an R-algebra morphism. Assume that $f(I) = 0$ (this is shorthand for saying that $f(a) = 0$ for all $a \in I$). Consider the canonical projection $\pi : A \to A/I$. Then, there is a unique R-algebra morphism $f' : A/I \to B$ satisfying $f = f' \circ \pi$.

Proof. Adapt the argument that we used to prove Theorem 1.2.2. The only new thing we need to check is that the map f' constructed in the proof is R-linear; but this is just as straightforward as showing that this map is a ring morphism. □

1.3. Adjoining roots

1.3.1. Examples

Remember how the complex numbers were first introduced by Cardano (back in the 16th century). Nowadays we define them as pairs of real numbers; this is a straightforward process (first you define addition and multiplication and zero and unity; then you show that the ring axioms hold). But this is the modern definition; the original vision was different: Cardano essentially proposed to imagine that there is a new number called i that satisfies $i^2 = -1$ but otherwise behaves like the numbers we know. So you're allowed to form arbitrary polynomials in i, but you have to equate i^2 to -1, so you never end up getting anything more complicated than numbers of the form $a + bi$ with $a, b \in \mathbb{R}$ (since any higher power of i can be reduced to ± 1 or $\pm i$ using the $i^2 = -1$ rule). Thus, it makes sense to encode complex numbers as pairs, but this is merely one way of encoding them.

Of course, Cardano's original vision is not a rigorous definition; just as easily you could introduce a number j satisfying $0j = 1$, and thus collapse the entire number system (since this new number would let you argue that $1 = 0j = (0+0)\,j = 0j + 0j = 1 + 1 = 2$). So, if we want to make Cardano's definition rigorous, we have to rewrite it algebraically. One way to do this is to define $\mathbb{C}$ as the quotient ring

$$\mathbb{R}[x] \,/\, \left(x^2+1\right)\mathbb{R}[x].$$

In fact, we start with $\mathbb{R}[x]$ because our complex numbers should be polynomials in a single symbol i (which will be represented by the indeterminate x in $\mathbb{R}[x]$); but then we quotient out the ideal $\left(x^2+1\right)\mathbb{R}[x]$ since we want i^2+1 (and thus also each multiple of i^2+1) to be 0 in our complex numbers.

To be on the safe side, let us show that this quotient ring $\mathbb{R}[x] \,/\, \left(x^2+1\right)\mathbb{R}[x]$ is isomorphic to the complex numbers $\mathbb{C}$ as we know them (i.e., defined in the modern way, as pairs of real numbers).

First of all, we introduce a shorthand:

Convention 1.3.1. If R is a commutative ring, and if $a \in R$, then the quotient ring R/aR will be abbreviated as R/a. We are already using a particular case of this notation, as we are writing $\mathbb{Z}/n$ for $\mathbb{Z}/n\mathbb{Z}$ when n is an integer.

So we want to prove that $\mathbb{R}[x] \,/\, \left(x^2+1\right) \cong \mathbb{C}$ as rings – and even better, as $\mathbb{R}$-algebras. Let's be a little bit more precise:

Proposition 1.3.2. We have $\mathbb{R}[x] \,/\, \left(x^2+1\right) \cong \mathbb{C}$ as $\mathbb{R}$-algebras. More concretely: There is an $\mathbb{R}$-algebra isomorphism

$$\begin{aligned} \mathbb{R}[x] \,/\, \left(x^2+1\right) &\to \mathbb{C}, \\ \overline{p} &\mapsto p[i]. \end{aligned}$$

Proof. We already know that $\mathbb{C}$ is an $\mathbb{R}$-algebra. Thus, Theorem 1.1.13 (applied to $R = \mathbb{R}$ and $A = \mathbb{C}$ and $a = i$) yields that the map

$$\begin{aligned} f : \mathbb{R}[x] &\to \mathbb{C}, \\ p &\mapsto p[i] \end{aligned}$$

is an $\mathbb{R}$-algebra morphism. This map f sends the principal ideal $\left(x^2+1\right)\mathbb{R}[x]$ to 0, because for each $q \in \mathbb{R}[x]$, we have

$$f\left(\left(x^2+1\right)\cdot q\right) = \left(\left(x^2+1\right)\cdot q\right)[i] = \underbrace{\left(i^2+1\right)}_{=0}\cdot q[i] = 0.$$

Hence, Theorem 1.2.3 (applied to $R = \mathbb{R}$, $A = \mathbb{R}[x]$, $I = \left(x^2+1\right)\mathbb{R}[x]$ and $B = \mathbb{C}$) shows there is a unique $\mathbb{R}$-algebra morphism

$$f' : \mathbb{R}[x] / \left(x^2+1\right) \to \mathbb{C}$$

satisfying $f = f' \circ \pi$, where $\pi : \mathbb{R}[x] \to \mathbb{R}[x] / \left(x^2+1\right)$ is the canonical projection. Consider this f'. The equality $f = f' \circ \pi$ means that each $p \in \mathbb{R}[x]$ satisfies

$$f(p) = \left(f' \circ \pi\right)(p) = f'\left(\underbrace{\pi(p)}_{=\overline{p}}\right) = f'\left(\overline{p}\right),$$

so that

$$f'\left(\overline{p}\right) = f(p) = p[i] \qquad \text{(by the definition of } f\text{)}. \tag{7}$$

Now, why is f' an isomorphism?

It's not hard to see that f' is surjective: Indeed, any $z \in \mathbb{C}$ can be written as $z = a + bi$ for some $a, b \in \mathbb{R}$, and then we have $z = a + bi = f'\left(\overline{a+bx}\right)$ (since (7) yields $f'\left(\overline{a+bx}\right) = (a+bx)[i] = a + bi$).

Now, how can we prove that f' is injective? Since f' is $\mathbb{R}$-linear, it suffices to show that $\operatorname{Ker}\left(f'\right) = \{0\}$ (by a lemma in Lecture 9).

Let $u \in \operatorname{Ker}\left(f'\right)$. Thus, $u \in \mathbb{R}[x] / \left(x^2+1\right)$, so that $u = \overline{p}$ for some $p \in \mathbb{R}[x]$. Consider this p.

However, Theorem 1.1.6 **(a)** (applied to $R = \mathbb{R}$, $b = x^2+1$ and $a = p$) yields that there is a unique pair (q, r) of polynomials in $\mathbb{R}[x]$ such that

$$p = q \cdot \left(x^2+1\right) + r \qquad \text{and} \qquad \deg r < \deg\left(x^2+1\right).$$

Consider this pair (q, r). From $\deg r < \deg\left(x^2+1\right) = 2$, we see that the polynomial r can be written as $a + bx$ for some $a, b \in \mathbb{R}$. Consider these a, b. From

$p = q \cdot \left(x^2+1\right) + r$, we obtain $p - r = q \cdot \left(x^2+1\right) \in \left(x^2+1\right)\mathbb{R}[x]$; thus, $\overline{p} = \overline{r}$ in the quotient ring $\mathbb{R}[x] / \left(x^2+1\right)$. Now,

$$\begin{aligned} u = \overline{p} = \overline{r} &= \overline{a+bx} \qquad (\text{since } r = a+bx), \qquad \text{so that} \\ f'(u) &= f'\left(\overline{a+bx}\right) = (a+bx)[i] \qquad (\text{by (7)}) \\ &= a + bi. \end{aligned}$$

Hence, $a + bi = f'(u) = 0$ (since $u \in \operatorname{Ker}(f')$). Since $a, b \in \mathbb{R}$, this entails $a = b = 0$ (since the complex numbers 1 and i are $\mathbb{R}$-linearly independent). Thus, $u = \overline{a+bx}$ rewrites as $u = \overline{0+0x} = 0 \in \{0\}$.

Forget that we fixed u. We thus have shown that $u \in \{0\}$ for each $u \in \operatorname{Ker}(f')$. In other words, $\operatorname{Ker}(f') \subseteq \{0\}$. Since the reverse inclusion $\{0\} \subseteq \operatorname{Ker}(f')$ is obvious, we thus conclude that $\operatorname{Ker}(f') = \{0\}$. As we have said, this entails that f' is injective.

Now we know that the map f' is injective and surjective. Hence, f' is bijective, i.e., invertible. Since every invertible $\mathbb{R}$-algebra morphism is an $\mathbb{R}$-algebra isomorphism[6], we thus conclude that f' is an $\mathbb{R}$-algebra isomorphism. This proves Proposition 1.3.2. □

Note the use of polynomial division (with remainder) in our above proof of Proposition 1.3.2. It has a natural usefulness in the study of quotient rings of $\mathbb{R}[x]$, just as integer division (with remainder) is crucial to the study of quotient rings of $\mathbb{Z}$.

Similarly to Proposition 1.3.2, we can reveal further quotient rings of polynomial rings as certain rings we know:

Proposition 1.3.3. **(a)** Recall the ring $\mathbb{Z}[i]$ of Gaussian integers. We have $\mathbb{Z}[x] / \left(x^2+1\right) \cong \mathbb{Z}[i]$ as $\mathbb{Z}$-algebras. More concretely: There is a $\mathbb{Z}$-algebra isomorphism

$$\begin{aligned} \mathbb{Z}[x] / \left(x^2+1\right) &\to \mathbb{Z}[i], \\ \overline{p} &\mapsto p[i]. \end{aligned}$$

(b) Recall the ring $\mathbb{S} = \mathbb{Q}\left[\sqrt{5}\right] = \left\{a + b\sqrt{5} \mid a, b \in \mathbb{Q}\right\}$ (a subring of $\mathbb{R}$). We have $\mathbb{Q}[x] / \left(x^2-5\right) \cong \mathbb{S}$ as $\mathbb{Q}$-algebras. More concretely: There is a $\mathbb{Q}$-algebra isomorphism

$$\begin{aligned} \mathbb{Q}[x] / \left(x^2-5\right) &\to \mathbb{S}, \\ \overline{p} &\mapsto p\left[\sqrt{5}\right]. \end{aligned}$$

[6]This is proved in the same way as we showed that every invertible ring morphism is a ring isomorphism.

Proof. **(a)** Analogous to the proof of Proposition 1.3.2.
(b) Analogous to the proof of Proposition 1.3.2. □

Proposition 1.3.2 and Proposition 1.3.3 suggest that when we start with a ring R and a polynomial $b \in R[x]$, then the quotient ring $R[x]/b$ is (in some way) an "extension" of R by a root of b, in the sense that it contains R as a subring (at least up to isomorphism) but also contains a root of b (namely, $\overline{x}$). Thus, we can hope that by taking the quotient ring $R[x]/b$, we can "adjoin" a root of b to the ring R even if b has no root over R (just as Cardano defined the complex numbers by "adjoining" a root of x^2+1 to $\mathbb{R}$).

The following example (in which we take a quotient of $\mathbb{Z}[x]$ by a constant polynomial) dashes some cold water on this hope, at least in its general form:

Proposition 1.3.4. **(a)** We have $(\mathbb{Z}[x])/m \cong (\mathbb{Z}/m)[x]$ as $\mathbb{Z}$-algebras (i.e., as rings) for any integer m.
(b) The ring $(\mathbb{Z}[x])/1$ is trivial.

Proof sketch. **(a)** Let m be an integer. Then, the principal ideal $m\mathbb{Z}[x]$ of $\mathbb{Z}[x]$ consists of all polynomials whose all coefficients are multiples of m. Thus, it is easy to see that the map

$$f : (\mathbb{Z}[x])/m \to (\mathbb{Z}/m)[x],$$
$$\overline{a_0x^0 + a_1x^1 + a_2x^2 + \cdots} \mapsto \overline{a_0}x^0 + \overline{a_1}x^1 + \overline{a_2}x^2 + \cdots$$

is well-defined and is a $\mathbb{Z}$-algebra isomorphism. This proves Proposition 1.3.4 **(a)**.

(b) More generally: If R is any ring, then the ring $R/1$ is trivial. This is because the principal ideal $1R$ of R is the whole ring R, so there is only one coset modulo this ideal. □

Proposition 1.3.4 **(a)** (applied to $m = 2$) shows that if we take the quotient ring of $\mathbb{Z}[x]$ modulo (the principal ideal generated by) the constant polynomial 2, then we don't get an "extension" of $\mathbb{Z}$; what we instead get is the polynomial ring $(\mathbb{Z}/2)[x]$, in which (unlike in $\mathbb{Z}$) we have $1+1=0$ (so it certainly cannot contain a copy of $\mathbb{Z}$ as a subring). But if you think about this carefully, you will realize that this perfectly agrees with the idea of "adjoining a root". Indeed, to "adjoin" a root of the constant polynomial 2 to $\mathbb{Z}$ means to introduce a new "number" x satisfying $2 = 0$. The equation $2 = 0$ tells us nothing about the number x (so it remains completely unconstrained), but collapses all even integers to 0, thus leaving us with the ring $(\mathbb{Z}/2)[x]$. This is precisely what Proposition 1.3.4 **(a)** told us. Likewise, "adjoining" a root of 1 to $\mathbb{Z}$ causes $1 = 0$, which renders the ring trivial (since any element of a ring is a multiple of 1); this agrees with Proposition 1.3.4 **(b)**.

The examples so far have taught us that – yes – we can "adjoin" a root of a polynomial to a commutative ring R, but we don't always get an extension of

R (although we do always get an R-algebra). In the next lecture we will see a (sufficient) criterion for when we do.

Here is another natural question: What happens if we "adjoin" a root of a polynomial b that already has a root in R ? For example, let us take the polynomial $x^2 - 1$ over $\mathbb{Q}$ (which has 1 and -1 as roots). It turns out that the resulting quotient ring $\mathbb{Q}[x] / \left(x^2 - 1\right)$ is a good friend of ours by now:

Proposition 1.3.5. Recall the group algebra $\mathbb{Q}[C_2]$ of the cyclic group C_2 from Lecture 11. Then,

$$\mathbb{Q}[x] / \left(x^2 - 1\right) \cong \mathbb{Q}[C_2] \cong \mathbb{Q} \times \mathbb{Q} \qquad \text{as } \mathbb{Q}\text{-algebras.}$$

Proof. In Lecture 11, we have seen that the group algebra $\mathbb{Q}[C_2]$ has a basis (e_1, e_u) (as a $\mathbb{Q}$-module). By one of our conventions, we can write 1 and u for e_1 and e_u, so that this basis becomes $(1, u)$. We also know (from Lecture 11) that $\mathbb{Q}[C_2] \cong \mathbb{Q} \times \mathbb{Q}$ as $\mathbb{Q}$-algebras. It thus remains to prove that $\mathbb{Q}[x] / \left(x^2 - 1\right) \cong \mathbb{Q}[C_2]$.

Note the similarity between $\mathbb{Q}[C_2]$ and $\mathbb{C}$:

- The $\mathbb{Q}$-module $\mathbb{Q}[C_2]$ has basis $(1, u)$, with $u^2 = 1$.
- The $\mathbb{R}$-module $\mathbb{C}$ has basis $(1, i)$, with $i^2 = -1$.

This suggests that we just copypaste our above proof of Proposition 1.3.2, replacing $\mathbb{R}$, $\mathbb{C}$ and i by $\mathbb{Q}$, $\mathbb{Q}[C_2]$ and u and occasionally flipping signs. This is precisely what we are now going to do (but in a smaller font, to avoid wasting paper).

Theorem 1.1.13 (applied to $R = \mathbb{Q}$ and $A = \mathbb{Q}[C_2]$ and $a = u$) yields that the map

$$\begin{aligned} f : \mathbb{Q}[x] &\to \mathbb{Q}[C_2], \\ p &\mapsto p[u] \end{aligned}$$

is a $\mathbb{Q}$-algebra morphism. This map f sends the principal ideal $\left(x^2 - 1\right)\mathbb{Q}[x]$ to 0, because for each $q \in \mathbb{Q}[x]$, we have

$$f\left(\left(x^2 - 1\right) \cdot q\right) = \left(\left(x^2 - 1\right) \cdot q\right)[u] = \underbrace{\left(u^2 - 1\right)}_{\substack{=0 \\ (\text{since } u^2 = 1)}} \cdot q[u] = 0.$$

Hence, Theorem 1.2.3 (applied to $R = \mathbb{Q}$, $A = \mathbb{Q}[x]$, $I = \left(x^2 - 1\right)\mathbb{Q}[x]$ and $B = \mathbb{Q}[C_2]$) shows there is a unique $\mathbb{Q}$-algebra morphism

$$f' : \mathbb{Q}[x] / \left(x^2 - 1\right) \to \mathbb{Q}[C_2]$$

satisfying $f = f' \circ \pi$, where $\pi : \mathbb{Q}[x] \to \mathbb{Q}[x] / \left(x^2 - 1\right)$ is the canonical projection. Consider this f'. The equality $f = f' \circ \pi$ means that each $p \in \mathbb{Q}[x]$ satisfies

$$f(p) = \left(f' \circ \pi\right)(p) = f'\left(\underbrace{\pi(p)}_{=\overline{p}}\right) = f'\left(\overline{p}\right),$$

so that

$$f'(\overline{p}) = f(p) = p[u] \qquad \text{(by the definition of } f\text{)}. \tag{8}$$

Now, why is f' an isomorphism?

It's not hard to see that f' is surjective: Indeed, any $z \in \mathbb{Q}[C_2]$ can be written as $z = a + bu$ for some $a, b \in \mathbb{Q}$, and then we have $z = a + bu = f'\left(\overline{a + bx}\right)$ (since (8) yields $f'\left(\overline{a + bx}\right) = (a + bx)[u] = a + bu$).

Now, how can we prove that f' is injective? Since f' is $\mathbb{Q}$-linear, it suffices to show that $\operatorname{Ker}(f') = \{0\}$ (by a lemma in Lecture 9).

Let $u \in \operatorname{Ker}(f')$. Thus, $u \in \mathbb{Q}[x] / (x^2 - 1)$, so that $u = \overline{p}$ for some $p \in \mathbb{Q}[x]$. Consider this p.

However, Theorem 1.1.6 **(a)** (applied to $R = \mathbb{Q}$, $b = x^2 - 1$ and $a = p$) yields that there is a unique pair (q, r) of polynomials in $\mathbb{Q}[x]$ such that

$$p = q \cdot (x^2 - 1) + r \qquad \text{and} \qquad \deg r < \deg\left(x^2 - 1\right).$$

Consider this pair (q, r). From $\deg r < \deg\left(x^2 - 1\right) = 2$, we see that the polynomial r can be written as $a + bx$ for some $a, b \in \mathbb{Q}$. Consider these a, b. From $p = q \cdot (x^2 - 1) + r$, we obtain $p - r = q \cdot (x^2 - 1) \in (x^2 - 1)\mathbb{Q}[x]$; thus, $\overline{p} = \overline{r}$ in the quotient ring $\mathbb{Q}[x] / (x^2 - 1)$. Now,

$$\begin{aligned} u = \overline{p} = \overline{r} = \overline{a + bx} \qquad & \text{(since } r = a + bx\text{)}, \qquad \text{so that} \\ f'(u) = f'\left(\overline{a + bx}\right) = (a + bx)[u] \qquad & \text{(by (8))} \\ = a + bu. & \end{aligned}$$

Hence, $a + bu = f'(u) = 0$ (since $u \in \operatorname{Ker}(f')$). Since $a, b \in \mathbb{Q}$, this entails $a = b = 0$ (since the vectors 1 and u in $\mathbb{Q}[C_2]$ are $\mathbb{Q}$-linearly independent). Thus, $u = \overline{a + bx}$ rewrites as $u = \overline{0 + 0x} = 0 \in \{0\}$.

Forget that we fixed u. We thus have shown that $u \in \{0\}$ for each $u \in \operatorname{Ker}(f')$. In other words, $\operatorname{Ker}(f') \subseteq \{0\}$. Since the reverse inclusion $\{0\} \subseteq \operatorname{Ker}(f')$ is obvious, we thus conclude that $\operatorname{Ker}(f') = \{0\}$. As we have said, this entails that f' is injective.

Now we know that the map f' is injective and surjective. Hence, f' is bijective, i.e., invertible. Since every invertible $\mathbb{Q}$-algebra morphism is a $\mathbb{Q}$-algebra isomorphism[7], we thus conclude that f' is an $\mathbb{Q}$-algebra isomorphism. Hence, $\mathbb{Q}[x] / (x^2 - 1) \cong \mathbb{Q}[C_2]$. As we said, this proves Proposition 1.3.5. □

In our proofs of Proposition 1.3.2, 1.3.5 and 1.3.3 (even though I left the latter to the reader), we used that the leading coefficients of the polynomials we were quotienting out were units. Indeed, this is what allowed us to apply Theorem 1.1.6 **(a)**, which was a crucial step in proving that f' is injective. Describing quotient rings becomes much more complicated when the leading coefficient of the polynomial is not a unit. Sometimes it is nevertheless possible. Here is a particularly well-behaved example:

[7] This is proved in the same way as we showed that every invertible ring morphism is a ring isomorphism.

Proposition 1.3.6. Fix a nonzero integer m. Define the ring R_m as in exercise 1 on homework set #1; that is, R_m is the subring

$$\left\{ r \in \mathbb{Z} \mid \text{there exists an } m \in \mathbb{N} \text{ satisfying } m^k r \in \mathbb{Z} \right\}$$

of $\mathbb{Q}$. Then,

$$\mathbb{Z}[x] / (mx - 1) \cong R_m \qquad \text{as } \mathbb{Z}\text{-algebras (i.e., as rings).}$$

More concretely: There is a $\mathbb{Z}$-algebra isomorphism

$$\begin{aligned} \mathbb{Z}[x] / (mx - 1) &\to R_m, \\ \overline{p} &\mapsto p\left[\frac{1}{m}\right]. \end{aligned}$$

Proof sketch. Intuitively, this should be exactly what you expect: According to our "adjoining roots" philosophy, the ring $\mathbb{Z}[x] / (mx - 1)$ is what you get if you "adjoin" a root of the polynomial $mx - 1$ to $\mathbb{Z}$. But such a root would behave like the rational number $\frac{1}{m}$; so it is no surprise that the resulting ring would be isomorphic to R_m (since R_m is really just "the numbers you can get if you start with the integers and also allow multiplying by $\frac{1}{m}$"). This, of course, is not a proof.

An actual proof can be done along the following lines:

1. Show that a $\mathbb{Z}$-algebra morphism

$$\begin{aligned} \alpha : \mathbb{Z}[x] / (mx - 1) &\to R_m, \\ \overline{p} &\mapsto p\left[\frac{1}{m}\right] \end{aligned}$$

   exists. This is similar to the corresponding part of the proof of Proposition 1.3.2 (where we called the corresponding morphism f' rather than α); the main roles are played by Theorem 1.1.13 and Theorem 1.2.3.

2. (Optional:) Show that this morphism α is surjective. (In fact, each element of R_m has the form $\frac{a}{m^k}$ for some $a \in \mathbb{Z}$ and some $k \in \mathbb{N}$, and thus equals $\alpha\left(\overline{ax^k}\right)$.)

3. Don't waste your time trying to show that α is injective; there is no quick way to prove this directly.

4. Show that there is a map

$$\begin{aligned} \beta : R_m &\to \mathbb{Z}\left[x\right] / \left(mx-1\right), \\ \frac{a}{m^k} &\mapsto \overline{ax^k} \qquad \left(\text{where } a \in \mathbb{Z} \text{ and } k \in \mathbb{N}\right). \end{aligned}$$

(You need to show that this is well-defined – i.e., that if an element of R_m has been written in the form $\frac{a}{m^k}$ in two different ways, then the resulting residue classes $\overline{ax^k}$ will be equal.)

5. Show that β is a $\mathbb{Z}$-algebra morphism. (This is an exercise in bringing fractions to a common denominator.)

6. Show that $\beta \circ \alpha = \mathrm{id}$. (Indeed, $\beta \circ \alpha$ is a $\mathbb{Z}$-algebra morphism, since β and α are $\mathbb{Z}$-algebra morphisms. Moreover, it is easy to show that $\left(\beta \circ \alpha\right)\left(\overline{x}\right) = \overline{x}$. Hence, $\left(\beta \circ \alpha\right)\left(\sum_{i=0}^{n} c_i \overline{x}^i\right) = \sum_{i=0}^{n} c_i \overline{x}^i$ for each $n \in \mathbb{N}$ and any coefficients $c_0, c_1, \ldots, c_n \in \mathbb{Z}$ (since $\beta \circ \alpha$ is a $\mathbb{Z}$-algebra morphism). But this is saying that $\beta \circ \alpha = \mathrm{id}$, since every element of $\mathbb{Z}\left[x\right] / \left(mx-1\right)$ can be written as $\sum_{i=0}^{n} c_i \overline{x}^i$ for some $n \in \mathbb{N}$ and some coefficients $c_0, c_1, \ldots, c_n \in \mathbb{Z}$.)

7. Show that $\alpha \circ \beta = \mathrm{id}$. (Indeed, if you have done Step 2, then this follows from $\beta \circ \alpha = \mathrm{id}$. Otherwise, show it directly.)

8. Conclude from Steps 6 and 7 that the maps α and β are mutually inverse, and thus α is invertible. Since α is a $\mathbb{Z}$-algebra morphism, this entails that α is a $\mathbb{Z}$-algebra isomorphism, and you are done. □

References

[ChaLoi21] Antoine Chambert-Loir, *(Mostly) Commutative Algebra*, 27 January 2021.
`https://webusers.imj-prg.fr/~antoine.chambert-loir/publications/teach/sv-commalg.pdf`

[Ford21] Timothy J. Ford, *Abstract Algebra*, 2 February 2021.
`http://math.fau.edu/ford/preprints/Algebra_Book/Algebra_Book.pdf`

[Grinbe19] Darij Grinberg, *Regular elements of a ring, monic polynomials and "lcm-coprimality"*, 5 August 2019.
`https://www.cip.ifi.lmu.de/~grinberg/algebra/regpol.pdf`

[Knapp16] Anthony W. Knapp, *Basic Algebra*, Digital 2nd edition 2016.
`http://www.math.stonybrook.edu/~aknapp/download.html`