Chapter 5

Rings

Our overarching goal, laid out in Chapter 2, is to show that there are quintic polynomials whose roots are *not* expressible in terms of its coefficients using just the operations of addition, subtraction, multiplication, division, and the extraction of roots (thus implying that there is no "quintic formula" that is analogous to the quadratic formula). We decided that we would say that such polynomials are *not* solvable by radicals, and in Chapter 4, we finally were able to write down a formal definition of this term. We also proved there are many polynomials that are solvable by radicals. But, how do we show that a polynomial is *not* solvable by radicals? We start by taking a closer look at polynomials.

5.1 Abstract rings

As we investigate polynomials, it will be useful to harness (and abstract) the algebraic properties that they possesses. For example, if we add two polynomials in $\mathbb{Q}[x]$, we obtain a polynomial that is again in $\mathbb{Q}[x]$, and similarly for multiplication. Let's explore the structure of $F[x]$ in general (where F is any field).

Definition 5.1. Let F be a field. The structure $(F[x], +, \cdot)$ consists of the set $F[x]$ together with the operations $+$ and $\cdot$ defined as follows. Let $p(x) = a_0 + a_1 x + \cdots + a_m x^m$ and $q(x) = b_0 + b_1 x + \cdots + b_n x^n$ with $m \leq n$.

- **Addition:** $p(x)+q(x) = (a_0+b_0)+(a_1+b_1)x+\cdots+(a_n+b_n)x^n$ (where $a_k = 0$ when $k > m$).
- **Multiplication:** $p(x) \cdot q(x) = \displaystyle\sum_{k=0}^{m+n} c_k x^k$ where $c_k = a_0 b_k + a_1 b_{k-1} + \cdots + a_{k-1} b_1 + a_k b_0$.

We often refer to the entire structure $(F[x], +, \cdot)$ as simply $F[x]$.

In the definition of polynomial multiplication above, $p(x) \cdot q(x)$ is just the result of applying the distributive law repeatedly to $(a_0 + a_1 x + \cdots + a_m x^m) \cdot (b_0 + b_1 x + \cdots + b_n x^n)$ and then grouping according to the powers of x. We will see that the operations of polynomial addition and multiplication have many familiar properties; let's prove a couple.

Problem 5.2. Using the definitions of polynomial addition and multiplication together with properties of fields, prove that for all fields F, both addition and multiplication in $F[x]$ are commutative.

Problem 5.3. Let F be any field. Find an additive identity for $F[x]$, and prove that it works. Also, if $p(x) = a_0 + a_1x + \cdots + a_mx^m$ is an arbitrary polynomial in $F[x]$, find its additive inverse, and prove that it works.

Problem 5.4. Which elements of $\mathbb{Q}[x]$ have a multiplicative inverse? Which do not? Justify your answer.

As should be becoming clear, many of the properties of F transfer to $F[x]$ (but not all!). Let's record some of those properties in a fact, which we will not prove. The existence of multiplicative inverses is notably absent.

Fact 5.5. Let F be any field. The following are true for $F[x]$.

- **Addition Laws:** Addition is associative and commutative. There is a unique additive identity, namely the constant zero polynomial, and every polynomial $p(x)$ has a unique additive inverse, denoted $-p(x)$.
- **Multiplication Laws:** Multiplication is associative and commutative. There is a unique multiplicative identity, namely the constant polynomial 1.
- **Distributivity Laws:** For all $p(x), q(x), r(x) \in F[x]$, $p(x)(q(x)+r(x)) = p(x)q(x)+p(x)r(x)$ and $(q(x) + r(x))p(x) = q(x)p(x) + r(x)p(x)$.

5.1.1 Definition and first examples

Since $F[x]$ lacks multiplicative inverses for many of its elements, it does not form a field. Nevertheless, motivated by our desire to study polynomials, we will abstract the structure that is present so that we can prove theorems about polynomials over any field, instead of working one field at a time. However, before we do, it's worth noting that there are many other structures that are not fields but do satisfy the laws in Fact 5.5—perhaps the most prominent one is the integers $\mathbb{Z}$. We arrive at the definition of a ring.

Definition 5.6. A **ring** is a structure $(R, +, \cdot)$ consisting of a set R together with two binary operations $+$ and $\cdot$ (which we call *addition* and *multiplication*) such that for some element $0 \in R$ the following axioms hold.

- **Addition Axioms:** Addition is associative and commutative; the element 0 is an additive identity; every $x \in R$ has an additive inverse with respect to 0, denoted $-x$.
- **Multiplication Axioms:** Multiplication is associative.
- **Distributivity Axioms:** For all $x, y, z \in R$, $x(y + z) = xy + xz$ and $(y + z)x = yx + zx$.

In the case that multiplication is commutative, R is called a **commutative ring**, and in the case that there is a multiplicative identity, R is called a **ring with unity** (or ring with 1).

The notion of a ring is quite general, and the terminology "commutative ring" and "ring with unity" highlight some of the additional properties that $F[x]$ has, but arbitrary rings may not. But notice that fields have all of these properties *and more*. The next definition is meant to highlight this.

Definition 5.7. A **division ring** is a ring with unity such that every nonzero element has a multiplicative inverse.

Problem 5.8. Fill in each box of the table below with **Y**es or **N**o. Assume that + and · are defined "as usual" for each set.*

	ring	commutative ring	ring with unity	division ring	field
$\mathbb{Z}$					
$2\mathbb{Z}$					
$\mathbb{N}$					
$\mathbb{Q}$					
$\mathbb{H}$					
$\mathbb{Z}_6$					
$\mathbb{R}[x]$					
$\{a+bi \mid a,b \in \mathbb{Q}\}$					
$\{a+bi \mid a,b \in \mathbb{Z}\}$					

5.1.2 Basic properties

Many of the basic properties of fields hold also for rings, with essentially the same proofs, so we will just take them as fact.

Fact 5.9 (Compare with Fact 3.49)**.** Let R be a ring.

(1) The additive identity is unique. If there exists a multiplicative identity, it is unique.

(2) Additive inverses are unique. If an element has a multiplicative inverse, it is unique.

Fact 5.10 (Compare with Fact 3.50)**.** Let R be a ring.

(1) For all $x \in R$, $x \cdot 0 = 0 = 0 \cdot x$.

(2) For all $x, y \in R$, $(-x)y = -(xy)$ and $x(-y) = -(xy)$.

(3) If R contains at least two elements and has a multiplicative identity, then the additive and multiplicative identities are different, i.e. $0 \neq 1$.

*$2\mathbb{Z}$ denotes the even integers. The operations are usual integer addition and multiplication.

Let's explore one further property that fields possess but is not listed above: for all x and y in a field, if $xy = 0$, then $x = 0$ or $y = 0$.

Definition 5.11. Let R be a ring. An element $a \in R$ is called a **zero divisor** if a is nonzero and there exists a nonzero $b \in R$ such that $ab = 0$. A ring is called an **integral domain** if it is a commutative ring with unity containing at least two elements but *no zero divisors*.

As remarked above, fields do not have zero divisors, so every field is indeed an integral domain. However, the prototypical integral domain (which explains the choice of name) is $\mathbb{Z}$. Let's look for others.

Problem 5.12. For each of the following rings, determine if there are zero divisors, and if so, find them all. Is the ring an integral domain?

(1) $\mathbb{Z}_5$

(2) $\mathbb{Z}_{10}$

(3) $\mathbb{H}$

(4) $\mathbb{R}[x]$

When working with integral domains, the following property is key.

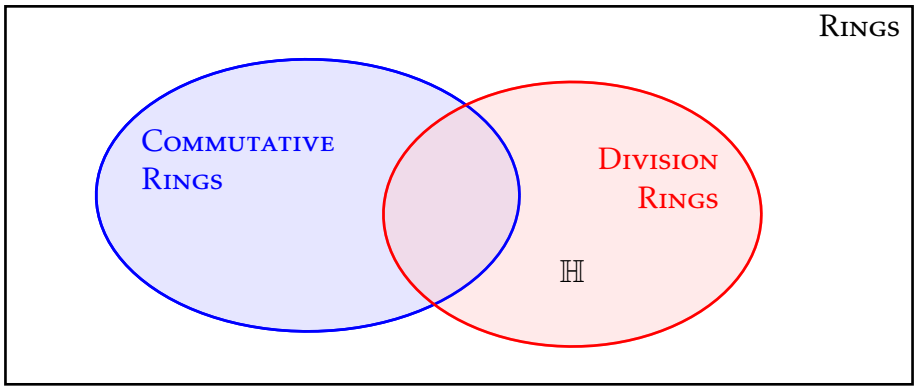
Theorem 5.13 (Cancellation Property)**.** Let R be an integral domain. For all $a, b, c \in R$, if $ab = ac$, then either $a = 0$ or $b = c$.

Problem 5.14. What properties of integral domains did you use in your proof of Theorem 5.13? Can you rewrite the theorem to be more general? Try.

Let's pause to collect and organize all of our new definitions.

Problem 5.15. Complete the following Venn Diagram by adding in shapes for each of the following terms. Try to provide examples that live in each of the gaps, but we have not encountered enough examples (in these notes) to cover all gaps yet.

- Fields
- Rings
- Commutative Rings
- Rings with unity
- Division Rings
- Integral domains



5.1.3 Units

Unless R is actually a division ring, not all elements of R will have a multiplicative inverse. Let's explore those elements that *do* have an inverse.

Definition 5.16. Let R be a ring with unity containing at least two elements. Then, $u \in R$ is called a **unit** if u has a multiplicative inverse. The set of all units in R is denoted $U(R)$.

Problem 5.17. For each of the following rings, find all of the units, i.e. determine $U(R)$.

(1) $\mathbb{Z}$

(2) $\mathbb{Z}_5$

(3) $\mathbb{R}$

(4) $\mathbb{R}[x]$

Problem 5.18. Consider the ring $\mathbb{Z}_{20}$. Find all units of $\mathbb{Z}_{20}$ and also find all zero divisors. What do you notice?

Problem 5.19. Let n be a positive integer. Make a conjecture about $U(\mathbb{Z}_n)$ by filling in the blank: $U(\mathbb{Z}_n) = \{a \in \mathbb{Z}_n \mid$ (fill in the blank) $\}$. What evidence do you have?

Theorem 5.20. Let R be a ring with unity containing at least two elements. If $u \in R$ is a unit, then u is *not* a zero divisor.

Problem 5.21. Either prove or disprove the *converse* of Theorem 5.20.

Theorem 5.22. Let R be a ring with unity containing at least two elements. Then $(U(R), \cdot)$ is a group.

5.2 An aside: matrix rings

Matrix rings are really the prototypical ring with unity. Although you may have only seen matrices with real entries, it turns out that we can do matrix arithmetic with other types of entries, e.g. entries from $\mathbb{C}$ or $\mathbb{Z}$. In fact, the usual matrix addition and multiplication makes sense when the entries come from any ring.

Definition 5.23. Let R be a ring and n a positive integer. Then $M_n(R)$ denotes the set of all $n \times n$ matrices whose entries come from R. The structure $(M_n(R), +, \cdot)$ consists of the set $M_n(R)$ of all $n \times n$ matrices whose entries come from R, together with the operations of usual matrix addition and matrix multiplication.

Problem 5.24. Provide examples of matrices satisfying each of the following conditions.

(1) $A \in M_3(\mathbb{C})$ but $A \notin M_3(\mathbb{R})$

(2) $B \in M_2(\mathbb{H})$ but $B \notin M_2(\mathbb{C})$

(3) $C \in M_2\left(\mathbb{Q}\left(\sqrt{5}\right)\right)$ but $C \notin M_2(\mathbb{Q})$

(4) $D \in M_2(\mathbb{R}[x])$ but $D \notin M_2(\mathbb{R})$

Problem 5.25. Verify that $M_2(\mathbb{Z})$ is closed under matrix multiplication.

The next fact shows that $M_n(R)$ is a ring with unity (for each positive n). Afterward, we will explore some of the other ring properties we discussed above.

Fact 5.26. Let R be any ring. The following are true for $M_n(R)$.

- **Addition Laws:** Addition is associative and commutative. There is a unique additive identity, namely the matrix with all entries equal to 0, and every matrix A has a unique additive inverse, denoted $-A$.
- **Multiplication Laws:** Multiplication is associative. There is a unique multiplicative identity, namely the matrix with 1's on the main diagonal and 0's everywhere else.
- **Distributivity Laws:** For all $A, B, C \in M_n(R)$, $A(B+C) = AB + AC$ and $(B+C)A = BA + CA$.

Problem 5.27. Is $M_2(\mathbb{R})$ commutative? Prove your answer.

Problem 5.28. Does $M_2(\mathbb{R})$ have zero divisors? Prove your answer.

The collection of units in a matrix ring forms a group with respect to matrix multiplication by Theorem 5.22. It is a very important object and even has a special name.

Definition 5.29. Let R be a ring and n a positive integer. The **general linear group** over the ring R, denoted $\mathrm{GL}_n(R)$, is the group of units in the ring $M_n(R)$.

Problem 5.30. Show that $\begin{bmatrix} i & 3 \\ 0 & i \end{bmatrix} \in \mathrm{GL}_2(\mathbb{C})$ by finding a multiplicative inverse for it. Also, find two different matrices in $M_2(\mathbb{C})$ that are *not* in $\mathrm{GL}_2(\mathbb{C})$.

5.3 Polynomial rings

Our study of rings was motivated by our desire to learn more about polynomials, and we now dive a little deeper into the theory of polynomial rings. Ultimately, we will focus on polynomials rings $F[x]$ where F is a field. In this section, we will see that $F[x]$ behaves in many ways like the integers $\mathbb{Z}$: $F[x]$ is an integral domain, there is a division algorithm for $F[x]$, there exists a greatest common divisor for polynomials, and there is a notion of primes and prime factorizations. Let's start with some important terminology.

Definition 5.31. Let R be a ring, and let $p(x) \in R[x]$ be a *nonzero* polynomial. If $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ with $a_n \neq 0$, then n is called the **degree** of $p(x)$, denoted $\deg p(x)$. In words, $\deg p(x)$ is the highest power of x in $p(x)$ with a nonzero coefficient. The degree of the zero polynomial is undefined.

Problem 5.32. Determine the degree of each of the following polynomials.

(1) $q(x) = 4x^5 + 2x^2 + 5 - 8x^2 + 2x^5$ in the ring $\mathbb{Z}[x]$

(2) $r(x) = 4x^5 + 2x^2 + 5 - 14x^2 + 2x^5$ in the ring $\mathbb{Z}_6[x]$

(3) $p(t) = (3t^2 - \sqrt{2})(-1 + 2t - t^3)$ in the ring $\mathbb{R}[t]$

(4) $s(x) = (5-i)^8 - (5-s)^8$ in the ring $\mathbb{C}[s]$

The degree function is incredibly useful when working with polynomials—let's prove a couple of properties about it.

Theorem 5.33. Let R be a ring. If $p(x)$ and $q(x)$ are nonzero polynomials $R[x]$, then $\deg(p(x)+q(x)) \leq \max(\deg p(x), \deg q(x))$ or $\deg(p(x)+q(x))$ is undefined.

Problem 5.34. Give an example of polynomials $p(x), q(x) \in \mathbb{Q}[x]$ such that $\deg(p(x)+q(x)) < \max(\deg p(x), \deg q(x))$.

Theorem 5.35. Let D be an integral domain. If $p(x)$ and $q(x)$ are nonzero polynomials $D[x]$, then $\deg(p(x)q(x)) = \deg p(x) + \deg q(x)$, and in particular, $\deg(p(x)q(x))$ is defined.

Problem 5.36. Give an example of nonzero polynomials $p(x), q(x) \in \mathbb{Z}_{10}[x]$ such that $\deg(p(x)q(x)) \neq \deg p(x) + \deg q(x)$. Why does this not contradict Theorem 5.35?

Corollary 5.37. If D is an integral domain, then $D[x]$ is an integral domain.

5.3.1 Division algorithm

Here we explore what it means for one polynomial to divide another as well as the idea of a quotient and remainder for division. These should be familiar from previous classes for $\mathbb{R}[x]$, but here we see that they generalize to arbitrary $F[x]$ for F a field.

Definition 5.38. Let R be a ring, and let $a, b \in R$. We say that b **divides** a (or b is a **divisor** of a) if there exists some $q \in R$ such that $a = bq$.

Problem 5.39. Consider the polynomial $p(x) = x^2 - 1$ in $\mathbb{Q}[x]$.

(1) Does $x+1$ divide $p(x)$ in $\mathbb{Q}[x]$? Why or why not?

(2) Does $p(x)$ divide $x+1$ in $\mathbb{Q}[x]$? Why or why not?

(3) Does 3 divide $p(x)$ in $\mathbb{Q}[x]$? Why or why not?

Theorem 5.40. Let $p(x) \in R[x]$ with R a ring. If $c \in R$ and $(x-c)$ divides $p(x)$, then $p(c) = 0$.

Even if $b(x)$ does not divide $a(x)$, it can still be useful to perform the division to obtain a quotient and remainder.

Problem 5.41. Consider the polynomials $a(x) = x^4 + x^3 - 8x + 5$ and $b(x) = x^2 - 3$ in $\mathbb{Q}[x]$. Use polynomial long division to show that $b(x)$ does not divide $a(x)$. What is the quotient and what is the remainder? Write $a(x)$ as $a(x) = b(x)q(x) + r(x)$ for some $q(x), r(x) \in \mathbb{Q}[x]$ with $\deg r(x) < \deg b(x)$.

The "division algorithm" (Theorem 5.43) formalizes what results from long division. And, it turns out that it is true for polynomials over any field (not just $\mathbb{Q}$). The next lemma prepares for the proof of the division algorithm.

Lemma 5.42. Let F be a field, and let $a(x), b(x) \in F[x]$ with $\deg a(x) \geq \deg b(x)$. Assume that $a(x) = a_0 + a_1 x + \cdots + a_n x^n$ with $a_n \neq 0$ and $b(x) = b_0 + b_1 x + \cdots + b_m x^m$ with $b_m \neq 0$. Set $a_1(x) = a(x) - b(x) a_n b_m^{-1} x^{n-m}$. Then $\deg a_1(x) < \deg a(x)$, and $b(x)$ divides $a(x) - a_1(x)$.

Suppose we are trying to divide $b(x)$ into $a(x)$. How do we find the quotient and the remainder? Well, if the degree of $a(x)$ is smaller than the degree of $b(x)$ there is nothing to do (and if $a(x) = 0$, there is also nothing to do). Otherwise, we can use the previous lemma to produce a polynomial $a_1(x)$ such that $\deg a_1(x) < \deg a(x)$ and $b(x)$ divides $a(x) - a_1(x)$, or in other words, $a(x) - a_1(x) = b(x)q_1(x)$ for some $q_1(x)$. Now, suppose we repeat the process and divide $b(x)$ into the resulting $a_1(x)$ to produce $a_2(x)$ and $q_2(x)$. Continuing in this fashion, we produce $a_2, q_2, a_3, q_3 \ldots, a_k, q_k$, stopping once the degree of $a_k(x)$ becomes smaller than the degree of $b(x)$ (or $a_k(x) = 0$). In total, we get something like the following.

$$\begin{aligned} a(x) - a_1(x) &= b(x)q_1(x) \\ a_1(x) - a_2(x) &= b(x)q_2(x) \\ a_2(x) - a_3(x) &= b(x)q_3(x) \\ &\vdots \\ a_{k-1}(x) - a_k(x) &= b(x)q_k(x) \end{aligned}$$

Adding the above equations together and moving things around, we arrive at

$$a(x) = b(x)\left(q_1(x) + q_2(x) + q_3(x) + \cdots + q_k(x)\right) + a_k(x)$$

with the degree of $a_k(x)$ being less than the degree of $b(x)$. Thus, $a_k(x)$ is the remainder and $q_1(x) + \cdots + q_k(x)$ the quotient. This is the rough idea behind the division algorithm.

Theorem 5.43 (Division algorithm for $F[x]$)**.** Let F be a field, and let $a(x), b(x) \in F[x]$ with $b(x) \neq 0$. Then there exist $q(x), r(x) \in F[x]$ such that

$$a(x) = b(x)q(x) + r(x)$$

with $\deg r(x) < \deg b(x)$ or $r(x) = 0$.

The division algorithm is the theoretical analogue of long division. If you want to divide concrete polynomials, use long division, but if you want to prove something about divisibility for arbitrary polynomials, use the division algorithm. It is often used to prove a polynomial $b(x)$ actually divides another polynomial $a(x)$. The strategy is to apply the division algorithm to produce the equation $a(x) = b(x)q(x) + r(x)$ (with $\deg r(x) < \deg b(x)$ or $r(x) = 0$) and then use this to show that, in fact, $r(x) = 0$, implying that $a(x) = b(x)q(x)$ as desired. Let's try using this approach to prove the converse of Theorem 5.40.

Theorem 5.44. Let $p(x) \in F[x]$ for F a field. If $c \in F$ and $p(c) = 0$, then $(x - c)$ divides $p(x)$.

Problem 5.45. Consider the polynomial $p(x) = x^2 + x + 3$ in $\mathbb{Z}_5[x]$. Compute $p(c)$ for each $c \in \mathbb{Z}_5$, and use the results to determine which polynomials of the form $(x - c)$ divide $p(x)$. Factor $p(x)$ into a product of degree 1 polynomials in $\mathbb{Z}_5[x]$, if possible.

Problem 5.46. Consider the polynomial $p(x) = x^2 + x + 1$ in $\mathbb{Z}_5[x]$. Explain why $p(x)$ cannot be factored into a product of degree 1 polynomials in $\mathbb{Z}_5[x]$.

5.3.2 Greatest common divisors

The fact that there is a division algorithm for $F[x]$ (Theorem 5.43) is a rather special property for a ring to possess, and it has several important consequences. The first one we'll explore is the existence of a "greatest common divisor" for two polynomials, and our first order of business is to try to decide on a reasonable definition of this.

Problem 5.47. What are the common divisors of 6 and -9 in $\mathbb{Z}$? Which one is the greatest common divisor?

Problem 5.48. Consider the polynomials $a(x) = 2x^2 - 2$ and $b(x) = 2x^2 + 2x - 4$ in $\mathbb{Q}[x]$.

(1) Show that $x-1$ is a common divisor of $a(x)$ and $b(x)$ by finding $q(x), s(x) \in \mathbb{Q}[x]$ such that $a(x) = (x-1)q(x)$ and $b(x) = (x-1)s(x)$.

(2) Show that $-2(x-1)$ is a common divisor of $a(x)$ and $b(x)$ by finding $q(x), s(x) \in \mathbb{Q}[x]$ such that $a(x) = -2(x-1)q(x)$ and $b(x) = -2(x-1)s(x)$.

(3) Show that $100(x-1)$ is a common divisor of $a(x)$ and $b(x)$ by finding $q(x), s(x) \in \mathbb{Q}[x]$ such that $a(x) = 100(x-1)q(x)$ and $b(x) = 100(x-1)s(x)$.

Which one, if any, would be a good choice as the "greatest common divisor"?

The previous problem highlights that there are several (actually, infinitely many) choices for the "greatest common divisor" of two polynomials. Our choice for which one we call the greatest common divisor is, in some sense, the simplest one.

Definition 5.49. A polynomial $p(x)$ of degree n is called **monic** if the coefficient of x^n (i.e. the leading coefficient) is 1.

For example, $7 - 2x + x^2$ is monic, since the coefficient of x^2 is 1. However, neither $7 - 2x + 3x^2$ nor $7 - 2x - x^2$ are monic.

Definition 5.50. Let F be a field, and let $a(x), b(x) \in F[x]$ be nonzero polynomials. A polynomial $d(x) \in F[x]$ is called a **greatest common divisor** of $a(x)$ and $b(x)$ if

(1) $d(x)$ is monic,

(2) $d(x)$ divides both $a(x)$ and $b(x)$,

(3) if $h(x)$ divides both $a(x)$ and $b(x)$, then $h(x)$ divides $d(x)$.

Thus, in Problem 5.48, the greatest common divisor of the polynomials $a(x)$ and $b(x)$ is $x-1$. That said, we don't yet know that a greatest common divisor always exists, but let's start by showing that if one exists, there is only one.

Lemma 5.51. Let F be a field, and let $a(x), b(x) \in F[x]$ be nonzero polynomials. If $d_1(x)$ and $d_2(x)$ are greatest common divisors of $a(x)$ and $b(x)$, then $d_1(x) = d_2(x)$.

We now work towards the existence of a greatest common divisor for arbitrary polynomials in $F[x]$ (for arbitrary fields). The proof of this result is tightly tied to analyzing certain combinations of the polynomials $a(x)$ and $b(x)$. Let's explore this a bit.

Problem 5.52. Consider the polynomials $a(x) = 2x^2 - 2$ and $b(x) = 2x^2 + 2x - 4$ in $\mathbb{Q}[x]$. Define

$$I = \{f(x)a(x) + g(x)b(x) \mid f(x), g(x) \in \mathbb{Q}[x]\}.$$

(1) Write down 5 different polynomials that are in the set I.

(2) Show that $x - 1$ divides an arbitrary polynomial in I.

The idea behind the second part of Problem 5.52 can be used to prove the following general result about sets like I.

Theorem 5.53. Let F be a field, and let $a(x), b(x) \in F[x]$ be nonzero polynomials. Define

$$I = \{f(x)a(x) + g(x)b(x) \mid f(x), g(x) \in F[x]\}.$$

If $h(x)$ divides both $a(x)$ and $b(x)$, then $h(x)$ divides every $c(x) \in I$.

The existence and uniqueness of greatest common divisors in $F[x]$ is presented in the following fact.

Fact 5.54. Let F be a field, and let $a(x), b(x) \in F[x]$ be nonzero polynomials. There exists a unique greatest common divisor of $a(x)$ and $b(x)$, and if $d(x)$ is the greatest common divisor, then

$$d(x) = f(x)a(x) + g(x)b(x),$$

for some $f(x), g(x) \in F[x]$.

The proof of this fact is interesting, but let's content ourselves to just outline it. The approach is fairly straight forward. Define

$$I = \{f(x)a(x) + g(x)b(x) \mid f(x), g(x) \in F[x]\}.$$

Theorem 5.53 tells us that every common divisor of $a(x)$ and $b(x)$ is a divisor of every polynomial in I. Thus, if I contains a monic, common divisor of $a(x)$ and $b(x)$, it must be the greatest common divisor. So, we look for a common divisor of $a(x)$ and $b(x)$ in I. And to do this, the key idea is to choose a polynomial of smallest degree in I.

Let m be the smallest degree of all nonzero polynomials in I (which exists by the Well-Ordering Property of the natural numbers). Choosing any polynomial of degree m in I, we can divide out the leading coefficient to get a *monic* polynomial $d(x)$, which we can show is still in I. The polynomial $d(x)$ will be the greatest common divisor.

To see that $d(x)$ divides $a(x)$, we use Theorem 5.43 (the division algorithm) to write $a(x) = d(x)q(x) + r(x)$ for $q(x), r(x) \in F[x]$ with $\deg r(x) < \deg d(x)$ or $r(x) = 0$. Towards a contradiction, assume $r(x) \neq 0$. Now, since $d(x) \in I$, there exist $f_d(x), g_d(x) \in F[x]$ such that

$$\begin{aligned} r(x) &= a(x) - d(x)q(x) \\ &= a(x) - [f_d(x)a(x) + g_d(x)b(x)]q(x) \\ &= [1 - f_d(x)q(x)]a(x) + [-g_d(x)q(x)]b(x) \\ &\in I. \end{aligned}$$

Since $\deg r(x) < \deg d(x)$, this contradicts the fact that $d(x)$ had the smallest possible degree of all polynomials in I. Thus, $r(x) = 0$, and $d(x)$ divides $a(x)$. A similar argument shows that $d(x)$ also divides $b(x)$, so $d(x)$ is a monic, common divisor of $a(x)$ and $b(x)$. And, Theorem 5.53 shows that $d(x)$ is a greatest common divisor. But then, it is the unique greatest common divisor by Lemma 5.51.

Using Fact 5.54, we can rewrite the set I is a very nice way.

Corollary 5.55. Let F be a field, and let $a(x), b(x) \in F[x]$ be nonzero polynomials. Define

$$I = \{f(x)a(x) + g(x)b(x) \mid f(x), g(x) \in F[x]\}.$$

If $d(x)$ is the greatest common divisor of $a(x)$ and $b(x)$, then $I = \{p(x)d(x) \mid p(x) \in F[x]\}$.

With similar ideas as in the proof of Fact 5.54, one can prove the following fact that characterizes the greatest common divisor in several different ways.

Fact 5.56. Let F be a field, and let $a(x), b(x) \in F[x]$ be nonzero polynomials. Define

$$I = \{f(x)a(x) + g(x)b(x) \mid f(x), g(x) \in F[x]\}.$$

For any polynomial $d(x) \in F[x]$, the following are equivalent:

(1) $d(x)$ is the greatest common divisor of $a(x)$ and $b(x)$;

(2) $d(x)$ is a monic common divisor of $a(x)$ and $b(x)$, and $d(x) \in I$;

(3) $d(x)$ is a monic, and $I = \{p(x)d(x) \mid p(x) \in F[x]\}$.

So, you may be wondering: how do we compute the greatest common divisor of two polynomials? First think about how you would compute the greatest common divisor of 168 and 180. Really, think about it... Many people will factor both 168 and 180 into primes and then multiply the prime factors they have in common. This works for integers, and in fact, it will also work for polynomials once we develop the notion of a "prime polynomial." However, there is another approach, which in general is way more efficient: the Euclidean Algorithm. We will not develop it here, but you are encouraged to look it up (perhaps starting on Wikipedia).

5.3.3 Irreducible polynomials

We now develop the analogous notion of a prime number for polynomials, which will be an *irreducible* polynomial. The concept of irreducibility makes sense quite generally, so we start by defining it for any integral domain. Recall that, by Corollary 5.37, $F[x]$ is always an integral domain when F is a field.

To motivate the definition, think about how prime integers are defined: $p \in \mathbb{Z}$ is prime if (1) $p > 1$ and (2) $p = ab$ implies that $a = \pm 1$ or $b = \pm 1$. Since the units of $\mathbb{Z}$ are precisely ± 1, the second condition could be rewritten as "$p = ab$ implies that a or b is a unit." Also, since we don't want 1 (or -1) to be considered prime, the first condition is mostly captured by ensuring that "p is not zero and not a unit."

Definition 5.57. Let D be an integral domain. An element $p \in D$ is **irreducible** if

- $p \neq 0$ and $p \notin U(D)$, and
- for all $a, b \in D$, $p = ab$ implies $a \in U(D)$ or $b \in U(D)$.

The element p is **reducible** if it is not irreducible; that is if $p = 0$, $p \in U(D)$, or there exist $a, b \in D$ such that $p = ab$ and $a, b \notin U(D)$.

Problem 5.58. Use Definition 5.57 to show that a field has no irreducible elements.

Problem 5.59. What are the irreducible elements in $\mathbb{Z}$?

In order to investigate irreducibility in an integral domain D, we need to know its units. Our overarching goal is to better understand polynomials, so let's start there.

Theorem 5.60. Let F be a field. Then $p(x)$ is a unit in $F[x]$ if and only if $\deg p(x) = 0$.

Let's rewrite our definition of reducibility in a more useable form for polynomials.

Theorem 5.61. Let F be a field, and let $p(x)$ be a nonconstant polynomial in $F[x]$. Then $p(x)$ is reducible if and only if $\deg p(x) > 0$ and there exist polynomials $a(x), b(x) \in F[x]$ such that $p(x) = a(x)b(x)$ with $\deg a(x) < \deg p(x)$ and $\deg b(x) < \deg p(x)$.

Problem 5.62. Determine if $p(x)$ is reducible or irreducible in the given ring. If it's reducible, write down a factorization.

(1) $p(x) = x^2 + 1$ in $\mathbb{C}[x]$

(2) $p(x) = x^2 + 1$ in $\mathbb{Q}[x]$

(3) $p(x) = x^2 + 1$ in $\mathbb{Z}_2[x]$

(4) $p(x) = x^2 + 1$ in $\mathbb{Z}_3[x]$

Let's catalog a couple of general irreducibility/reducibility results for polynomials of small degree.

Theorem 5.63. Let F be a field. If $\deg p(x) = 1$, then $p(x)$ is irreducible.

Theorem 5.64. Let F be a field. If $\deg p(x) = 2, 3$, then $p(x)$ is reducible if and only if $p(x)$ has a root in F.

Problem 5.65. Determine if $p(x)$ is reducible or irreducible in the given ring. If it's reducible, write down a factorization.

(1) $p(x) = x^3 - 2$ in $\mathbb{Q}[x]$

(2) $p(x) = x^3 - 2$ in $\mathbb{Z}_5[x]$

Problem 5.66. Determine if each of the following polynomials are reducible or irreducible in the given ring.

(1) $p(x) = x^3 - 8$ in $\mathbb{Q}[x]$

(2) $p(x) = x^3 - 8$ in $\mathbb{Z}_5[x]$

(3) $r(x) = x^4 - 8x^2 + 15$ in $\mathbb{Q}[x]$

(4) $r(x) = x^4 - 8x^2 + 15$ in $\mathbb{Z}_5[x]$

To solidify the analogy between irreducible elements and primes, let's prove a factorization theorem.

Theorem 5.67. If F is a field, then any polynomial of positive degree in $F[x]$ can be written as a product of polynomials that are irreducible in $F[x]$.

As you know, in the integers every number greater than or equal to 2 can be factored into a product of primes in a way that is *unique up to reordering the factors*. There is a similar uniqueness result for polynomials: any polynomial of positive degree in $F[x]$ can be written as a product of irreducible polynomials in a way that is unique up to reordering the factors and multiplying each factor by a unit.

Problem 5.68. Let $p(x) = 6x^4 - 7x^3 + 15x^2 - 21x - 9$. Then the following are two different factorizations of $p(x)$ into irreducibles in $\mathbb{Q}[x]$:

- $p(x) = (2x-3)(3x+1)(x^2+3)$, and
- $p(x) = (x+\frac{1}{3})(2x^2+6)(3x-\frac{9}{2})$.

Explain why the factorizations are the same after possibly reordering the factors and multiplying each factor by a unit.

5.4 Subrings

We now return to general ring theory. As with groups and fields, the notion of a subring is fundamental.

Definition 5.69. Let $(R, +, \cdot)$ be a ring, and let S be a subset of R. Then S is a **subring** of R if S is a ring in its own right with respect to operations $+$ and $\cdot$ *inherited from* R.

As with fields, many of the properties of the ring R automatically pass to a subset S (e.g. associativity), leaving only a handful of the ring axioms to actually be verified.

Theorem 5.70. Let R be a ring, and let $S \subseteq R$. Then S is a subring of R if and only if

(1) S is nonempty;

(2) for all $x, y \in S$, $x + y \in S$ and $xy \in S$; and

(3) for all $x \in S$, $-x \in S$.

Problem 5.71. Determine if each of the following subsets of $\mathbb{Q}[x]$ are actually subrings.

(1) $A = \{p(x) \mid p(x) = c \text{ for some } c \in \mathbb{Q}\}$ (i.e. the set of constant polynomials)

(2) $B = \{p(x) \mid p(x) = 0 \text{ or } \deg p(x) \leq 1\}$ (i.e. the set of linear polynomials)

(3) $\mathbb{Z}[x]$

(4) $I = \{f(x)x^2 + g(x)(1+x^5) \mid f(x), g(x) \in \mathbb{Q}[x]\}$

Problem 5.72. Explain why $\mathbb{Z}_5$ is *not* a subring of $\mathbb{Z}$.

Examples of subrings of $\mathbb{C}$ include $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}$, and $\mathbb{Q}(i)$. These examples can, in turn, be used to create subrings of polynomial rings and matrix rings .

Theorem 5.73. If S is a subring of R, then

(1) $S[x]$ is a subring of $R[x]$, and

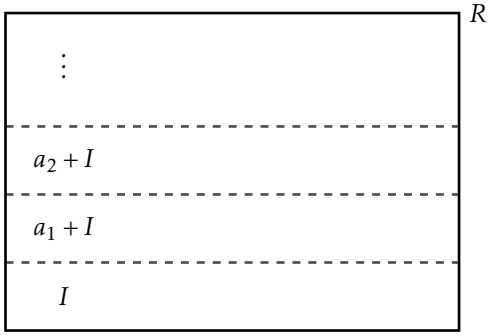
(2) $M_n(S)$ is a subring of $M_n(R)$.

5.5 Ideals and quotients

We now turn our attention to a special class of subrings known as *ideals*. The motivation for studying ideals of a ring is the same as for studying normal subgroups of a group: they give rise to quotients.

Let's explore the extra properties that a subring might need to ensure that the set of cosets can be given the structure of a ring. To start out, let's just assume that I is an *additive subgroup* of the ring R. Since $(R,+)$ is abelian, I is automatically a normal subgroup of $(R,+)$.

Now, let's consider the set of cosets of I in R, which we write as R/I. Recall that $R/I = \{a+I \mid a \in R\}$ where $a+I = \{a+y \mid y \in I\}$. Remember, that the set of cosets will partition R. One way to picture this is given below—it's followed by a couple of important properties about R/I from group theory.



Fact 5.74. Let I be an additive subgroup of a ring R. Then for all $a,b \in R$

(1) $a+I = b+I$ if and only if $a-b \in I$ if and only if $a \in b+I$; and

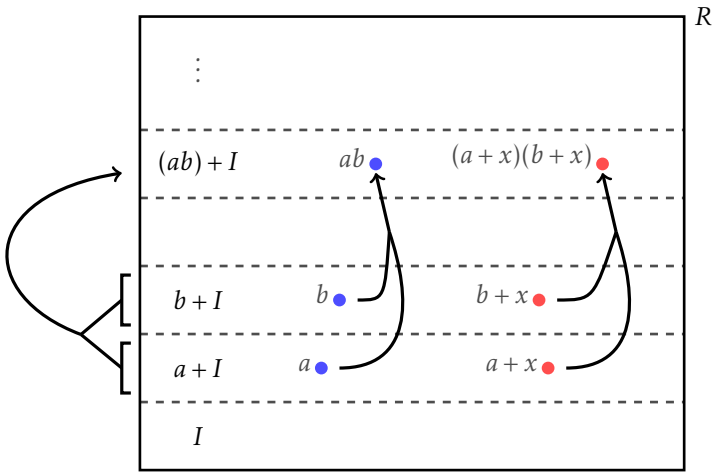
(2) either $(a+I)\cap(b+I) = \emptyset$ or $a+I = b+I$.

The goal is to understand when R/I can be given the structure of a ring. To do this, we need to decide how to add and multiply cosets. We would like to define $(a+I)+(b+I) = (a+b)+I$ and $(a+I)(b+I) = ab+I$, but the worry is that these operations are not well-defined. That is, the coset $a+I$ goes by many names (since $a+I = a'+I$ for every $a' \in a+I$), so we have to make sure that our definitions for the operations do not depend on which names we use for the cosets.

Now, as mentioned before, I is a *normal* subgroup of $(R,+)$, so we know that coset addition is well-defined. Let's see what we need for multiplication. Fix two arbitrary cosets $a+I$ and $b+I$. Then, for all $x,y \in I$, $a+I = (a+x)+I$ and $b+I = (b+y)+I$. Thus, in order for coset multiplication to be well-defined, we need to ensure that

$$ab+I = (a+x)(b+y)+I \text{ for all } a,b \in R \text{ and all } x,y \in I.$$

The desired picture is as follows.



After distributing, we have $ab+I = ab+ay+xb+xy+I$, which simplifies to $I = ay+xb+xy+I$. By Fact 5.74, we see that what we really need to ensure is that

$$ay+xb+xy \in I \text{ for all } a,b \in R \text{ and all } x,y \in I.$$

In particular, this has to be true when $a = b = 0$, which implies that $xy \in I$ for all $x,y \in I$, so I needs to be closed under multiplication, hence a subring. So, let's assume that I is a subring. Then, since $xy \in I$, $ay+xb+xy \in I$ reduces to $ay+xb \in I$. So, assuming that I is a subring, our previous condition becomes

$$ay+xb \in I \text{ for all } a,b \in R \text{ and all } x,y \in I.$$

Now, if a is arbitrary and $b = 0$, then we see that $ay \in I$ for all $a \in R$ and $y \in I$. Similarly, we find that $xb \in I$ for all $b \in R$ and $x \in I$. These are new properties. In words, I must be closed under (left and right) multiplication by elements from R.

In conclusion, if I is a subring that is closed under multiplication by elements from R then the above addition and multiplication for R/I is well-defined, making R/I a ring. The converse is also true. As such, we give these special subrings a special name.

Definition 5.75. Let R be a ring, and let $I \subseteq R$. Then I is an **ideal** of R if

(1) I is a subring; and

(2) for all $r \in R$ and all $a \in I$, both $ra \in I$ and $ar \in I$.

Let's also summarize our work above about defining operations on the quotient R/I.

Fact 5.76. Let R be a ring and let I be an ideal of R. Then R/I is a ring under the binary operations defined as follows. For all $a, b \in R$,

- $(a+I)+(b+I) = (a+b)+I$;
- $(a+I)(b+I) = (ab)+I$.

Problem 5.77. For each subset of the given ring, determine if the subset is an ideal, a subring, or neither.

	ideal	subring	neither
$\mathbb{Z} \subset \mathbb{Q}$			
$2\mathbb{Z} \subset \mathbb{Z}$			
$\{\text{odd integers}\} \subset \mathbb{Z}$			
$\{0,3,6,9\} \subset \mathbb{Z}_{12}$			
$\{p(x) \mid p(0) = 0\} \subset \mathbb{Q}[x]$			
$\{\text{constant polynomials}\} \subset \mathbb{Q}[x]$			

Problem 5.78. Let $I = \{(x^2+1)p(x) \mid p(x) \in \mathbb{Q}[x]\}$.

(1) Show that I is an ideal of $\mathbb{Q}[x]$.

(2) Write out 5 elements of I, each with a different degree.

(3) Explain why I contains polynomials of every degree larger than or equal to 2.

(4) Let $a(x) = x^4 + 3x + 5$. Write out 5 elements of the coset $a(x) + I$.

(5) Find some $b(x)$ in the coset $a(x) + I$ such that $\deg b(x) = 1$.

(6) Explain why $(x+I)^2 = -1 + I$ in the ring $\mathbb{Q}[x]/I$.

Problem 5.79. Let $I = \{(x^2+1)p(x) \mid p(x) \in \mathbb{Q}[x]\}$. Show that every coset $a(x) + I \in \mathbb{Q}[x]/I$ can be represented as $a(x) + I = r(x) + I$ for some $r(x) \in \mathbb{Q}[x]$ where $\deg r(x) < 2$ or $r(x) = 0$.

Problem 5.80. Recall that $6\mathbb{Z} = \{6z \mid z \in \mathbb{Z}\}$.

(1) Show that $6\mathbb{Z}$ is an ideal of $\mathbb{Z}$.

(2) Show that $a + I = b + I$ if and only if $a \equiv_6 b$.

(3) Show that for all $a + I \in \mathbb{Z}/6\mathbb{Z}$, $a + I = r + I$ for some $r \in \mathbb{Z}$ with $0 \leq r < 6$.

Problem 5.81. Let $I = \{3q \mid q \in \mathbb{Q}\}$.

(1) Show that I is an ideal of $\mathbb{Q}$.

(2) Show that $1 \in I$, and use this to explain why $I = \mathbb{Q}$.

Let's record some observations from the previous problems.

Theorem 5.82. Let R be a commutative ring, and let $a \in R$. The set $I = \{ar \mid r \in R\}$ is an ideal of R.

In the previous theorem, the set $\{ar \mid r \in R\}$ should be thought of as the set of all multiples of a, and it is often denoted aR (as in $2\mathbb{Z}$).

Theorem 5.83. Assume R is a ring with unity. Let I be an ideal of R. If I contains a unit of R, then $I = R$.

Theorem 5.84. Let R be a ring. Then $\{0\}$ and R are ideals of R.

Theorem 5.85. Assume R is a commutative ring with $1 \neq 0$. Then R is a field if and only if the only ideals of R are $\{0\}$ and R.

Theorem 5.86. Let I be an ideal of a ring R.

(1) If R is a commutative ring, then R/I is commutative ring.

(2) If R is a ring with unity, then R/I is a ring with unity.

5.5.1 Generating ideals

As with groups and fields, we will want to generate subobjects from subsets. Generating ideals will be more useful for us that subrings, so we will only focus on ideals. Regarding notation, it is common to use (A) for the ideal generated by A instead of $\langle A \rangle$, and we will follow that convention. We begin with intersections.

Theorem 5.87. If I and J are ideals of a ring R, then $I \cap J$ is an ideal of R.

This can be generalized to arbitrary intersections.

Theorem 5.88. If $\mathcal{C}$ is any collection of ideals of a ring R, then the intersection of all ideals from $\mathcal{C}$ is again an ideal of R.

Now, if A is any subset of R, we can let $\mathcal{C}$ be the collection of all ideals containing A, to see that the intersection of all ideals containing A is an ideal, and it must be the smallest ideal containing A. This leads to the following definition.

Definition 5.89. Suppose A is a subset of a ring R. The **ideal of R generated by A**, denoted (A), is defined to be the intersection of all ideals containing A. An ideal generated by one element is a **principal ideal**. If $A = \{a_1, \ldots a_k\}$, we often write $(a_1, \ldots a_k)$ in place of (A).

Problem 5.90. Consider the ring $\mathbb{Z}$.

(1) Recall that $2\mathbb{Z}$ is an ideal of $\mathbb{Z}$. Now use the definition of (2) to explain why $(2) \subseteq 2\mathbb{Z}$.

(2) Use that (2) is an ideal containing 2 to explain why $2\mathbb{Z} \subseteq (2)$. Conclude that $(2) = 2\mathbb{Z}$.

(3) Use that $(6, 10)$ is an ideal containing 6 and 10, to write down 5 elements of $(6, 10)$.

(4) Use the definition of $(6,10)$ to explain why $(6,10) \subseteq (2)$.

(5) Use Bézout's lemma (Fact 3.52) to show that $(2) \subseteq (6,10)$. Conclude that $(6,10) = (2)$.

(6) For arbitrary $m, n \in \mathbb{Z}$, do you think $(m,n) = (a)$ for some $a \in \mathbb{Z}$? Why or why not?

Let's work to abstract some of what we discovered in this problem.

Theorem 5.91. If R is a commutative ring with unity, then $(a) = \{ar \mid r \in R\}$.

Theorem 5.91 says that (a) is precisely the set of all multiples of a. Or, in other words, (a) is the set of all elements that are divisible by a. In particular, $(n) = n\mathbb{Z}$ in the ring $\mathbb{Z}$. But what about (m,n)?

Theorem 5.92. If $m, n \in \mathbb{Z}$ are nonzero, then $(m,n) = (d)$ where $d = \gcd(m,n)$.

In words, an ideal of $\mathbb{Z}$ that is generated by two elements can actually be generated by a single element. But more is true. The method for constructing the greatest common divisor for two elements can be easily adapted to show that *any* ideal of $\mathbb{Z}$ can be generated by a single element, which yields the following fact.

Fact 5.93. If I is any ideal of $\mathbb{Z}$, then I is a principle ideal. Moreover, if I is not the zero ideal, then $I = (d)$ if and only if d has the smallest possible absolute value among all nonzero elements of I.

In fact, a similar result holds in any ring with a division algorithm. Importantly for us, this applies to polynomial rings over fields. Let's prove the result for ideals generated by two elements and leave the general case as a fact.

Theorem 5.94. Let F be a field. If $a(x), b(x) \in F[x]$ are nonzero, then $(a(x), b(x)) = (d(x))$ where $d(x) = \gcd(a(x), b(x))$.

Fact 5.95. Let F be a field. If I is any ideal of $F[x]$, then I is a principle ideal. Moreover, if I is not the zero ideal, then $I = (d(x))$ if and only if $d(x)$ has the smallest possible degree among all nonzero elements of I.

Problem 5.96. Consider $a(x) = -x^2 - 3x + 10$, $b(x) = 2x^2 + 8x - 10$, and $c(x) = x^3 - 2$ in $\mathbb{Q}[x]$.

(1) Find a $d(x) \in \mathbb{Q}[x]$ such that $(a(x), b(x)) = (d(x))$. Is $(a(x), b(x)) = \mathbb{Q}[x]$? Explain.

(2) Find a $d'(x) \in \mathbb{Q}[x]$ such that $(a(x), c(x)) = (d'(x))$. Is $(a(x), c(x)) = \mathbb{Q}[x]$? Explain.

We saw that $\mathbb{Z}$ and $F[x]$ have a special property: every ideal is a principal ideal. This does not happen in every ring, as we'll see, so rings with this property get a special name.

Definition 5.97. An integral domain is called a **principal ideal domain (PID)** if every ideal is a principal ideal.

Thus, $\mathbb{Z}$ and $F[x]$ (for F a field) are examples of PIDs. Let's show that $\mathbb{Z}[x]$ is not.

Problem 5.98. Consider the set $I := \{2f(x) + xg(x) \mid f(x), g(x) \in \mathbb{Z}[x]\}$ in the ring $\mathbb{Z}[x]$. The set I is an ideal of $\mathbb{Z}[x]$; you do *not* need to prove this. Let's show that I is not principal. Towards a contradiction, assume that $I = (d(x))$ for some $d(x) \in \mathbb{Z}[x]$.

(1) Use the fact that $2 \in I = (d(x))$ to show that $d(x)$ is a constant polynomial and, moreover, that $d(x) = \pm 1, \pm 2$.

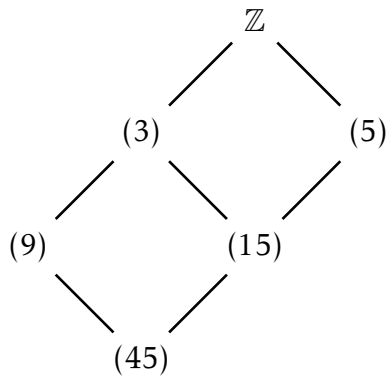
(2) Explain why every polynomial in I has a constant term that is even, and use this to show that in fact $d(x) = \pm 2$.

(3) Now, we also know that $x \in I = (d(x))$. Why is this a contradiction?

Let's return to $\mathbb{Z}$, and put together what we have learned about its ideals. First, by Fact 5.93, every ideal of $\mathbb{Z}$ is a principal ideal, so every ideal of $\mathbb{Z}$ is of the form (n) for some $n \in \mathbb{Z}$. Moreover, Theorem 5.91 tell us that (n) is just the set of all multiple of n, so $(n) = n\mathbb{Z}$. Thus, we know all of the ideals of $\mathbb{Z}$, and we know that they have a nice form. But how do they fit together? Let's explore this.

Theorem 5.99. Let R be a commutative ring with unity, and let $a, b \in R$. Then $(a) \subseteq (b)$ if and only if b divides a.

This theorem can be used to quickly draw portions of the lattice of ideals of $\mathbb{Z}$. For example, suppose we want to draw all of the ideals that contain the ideal (45). Every ideal is principal; suppose (n) is an ideal containing (45). By Theorem 5.99, n must divide 45. So, looking at all divisors of 45 (and noticing that $(-n) = (n)$), we get the following lattice.



Problem 5.100. Draw the the lattice of ideals of $\mathbb{Z}$ that contain the ideal (60).

Let's focus on the ideals of $\mathbb{Z}$ that are at the top of the lattice but below $\mathbb{Z}$.

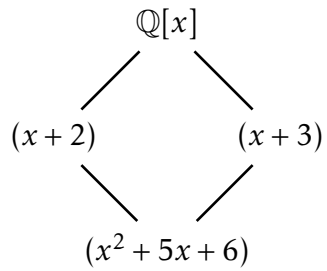
Definition 5.101. An ideal M of a ring R is called a **maximal ideal** if $M \neq R$ and the only ideals containing M are M and R.

Theorem 5.102. An ideal I of $\mathbb{Z}$ is maximal if and only if $I = (p)$ for some prime $p \in \mathbb{Z}$.

Since $F[x]$ (for F a field) is also PID, it's relatively easy to study the ideals of $F[x]$ too. As with $\mathbb{Z}$, every ideal is a principal ideal, and we can use Theorem 5.99 to see how they fit together. There is one preliminary result that will help us avoid redundancies.

Theorem 5.103. Let R be a commutative ring with unity. If $a \in R$ and $u \in U(R)$, then $(a) = (ua)$.

Now, suppose we want to find all of the ideals of $\mathbb{Q}[x]$ that contain the ideal (x^2+5x+6). As before, Theorem 5.99 tells us that we should look at divisors of $x^2 + 5x + 6$ in $\mathbb{Q}[x]$. Noting that $x^2 + 5x + 6 = (x + 2)(x + 3)$, we get the following.



Problem 5.104. Draw the the lattice of ideals of $\mathbb{Q}[x]$ that contain the ideal $(x^4 + x^2)$.

Theorem 5.105. Let F be a field. An ideal I of $F[x]$ is maximal if and only if $I = (p(x))$ for some irreducible polynomial $p(x) \in F[x]$.

5.6 Homomorphisms

As with groups, we use homomorphisms (and isomorphisms) to compare rings and fields.

Definition 5.106. Let R and S be rings. A map $\phi : R \to S$ is called a **ring homomorphism** if the following are true for all $a, b \in R$:

(1) $\phi(a + b) = \phi(a) + \phi(b)$;

(2) $\phi(ab) = \phi(a)\phi(b)$.

If ϕ is a bijection, then ϕ is called an **isomorphism**, in which case, we say that R and S are **isomorphic rings** and write $R \cong S$.

Problem 5.107. Determine which of the following are ring homomorphisms. Explain.

(1) $\phi : \mathbb{Z} \to 3\mathbb{Z}$ defined by $\phi(n) = 3n$

(2) $\alpha : \mathbb{C} \to \mathbb{C}$ defined by $\alpha(a + bi) = a - bi$

(3) $\beta : \mathbb{Z} \to \mathbb{Z}$ defined by $\beta(a) = a^3$

(4) $f : \mathbb{Z}_3 \to \mathbb{Z}_3$ defined by $f(a) = a^3$

(5) $g : \mathbb{C} \to D_2(\mathbb{R})$ defined by $g(a + bi) = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, where $D_2(\mathbb{R}) = \left\{ \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix} \mid x, y \in \mathbb{R} \right\}$

(6) $h : \mathbb{Q}[x] \to \mathbb{C}$ defined by $h(p(x)) = p(0)$

Problem 5.108. Which of the homomorphisms in Problem 5.107 were isomorphisms?

In Problem 5.107(6), we saw that "evaluating at zero" was a homomorphism from $\mathbb{Q}[x]$ to $\mathbb{C}$. At its core, this fact simply rests on how we add and multiply polynomials and does not require us to evaluate specifically at zero. For example, if $f(x), g(x) \in R[x]$ (for a ring R) and $c \in R$, then writing out $f(x) = a_0 + \cdots + a_m x^m$ and $g(x) = b_0 + \cdots + b_m x^m$ (with some a_i and b_j possible zero), we can show that $(f+g)(c) = f(c) + g(c)$. The analogous statement holds for multiplication too. We'll add in the details in the next theorem to see that "evaluating at c" is a homomorphism (whether or not $c = 0$).

Theorem 5.109 (Evaluation homomorphism)**.** Let R be a ring, and let $c \in R$. Define a function $\mathcal{E}_c : R[x] \to R$ by the rule $\mathcal{E}_c(p(x)) = p(c)$. Then $\mathcal{E}_c$ is a homomorphism.

Let's prove some general properties about homomorphisms.

Theorem 5.110. Let $\phi : R \to S$ be a ring homomorphism. Then

(1) $\phi(0) = 0$; and

(2) for all $a \in R$, $\phi(-a) = -\phi(a)$.

Theorem 5.111. Suppose that R is a ring with unity. Let $\phi : R \to S$ be a ring homomorphism. Then

(1) $\phi(1) = 1$; and

(2) for all $a \in U(R)$, $\phi(a^{-1}) = (\phi(a))^{-1}$.

Theorem 5.112 (Composition of homomorphisms)**.** Let $\phi : R_1 \to R_2$ and $\psi : R_2 \to R_3$ be ring homomorphisms. Then $\psi \circ \phi : R_1 \to R_3$ is also a ring homomorphism.

There are two important sets attached to homomorphisms: the kernel and the image.

Definition 5.113. Let $\phi : R \to S$ be a ring homomorphism.

- The **kernel** of ϕ, denote $\ker \phi$, is defined to be $\ker \phi = \{a \in R \mid \phi(a) = 0\}$.
- The **image** of ϕ, denoted $\phi(R)$, is defined to be $\phi(R) = \{b \in S \mid b = \phi(a) \text{ for some } a \in R\}$.

Problem 5.114. Determine the kernel and image of each homomorphism.

(1) $\alpha : \mathbb{Q}[x] \to \mathbb{C}$ defined by $\alpha(p(x)) = p(0)$

(2) $\beta : \mathbb{Z} \to \mathbb{Z}_5$ defined by $\beta(n) = n \pmod 5$

Problem 5.115. Let I be an ideal of a ring R. Define a function $\pi : R \to R/I$ by $\pi(r) = r + I$.

(1) Show that π is a homomorphism.

(2) What is the kernel of π?

(3) What is the image of π?

As with groups, the kernel and image of a homomorphism have special properties and can be used to detect if a function is injective (one-to-one) or surjective (onto).

Theorem 5.116. Let $\phi : R \to S$ be a ring homomorphism. Then $\ker(\phi)$ is an ideal of R, and $\phi(R)$ is a subring of S.

Theorem 5.117. Let $\phi : R \to S$ be a ring homomorphism. Then

(1) ϕ is injective if and only if $\ker\phi = \{0\}$, and

(2) ϕ is surjective if and only if $\phi(R) = S$.

Problem 5.118. By Theorem 5.109, the "evaluation at i" function from $\mathbb{C}[x]$ to $\mathbb{C}$ is a homomorphism. We can restrict the domain to $\mathbb{Q}[x]$ to obtain a homomorphism from $\mathcal{E}_i : \mathbb{Q}[x] \to \mathbb{C}$ defined by $\mathcal{E}_i(p(x)) = p(i)$, where i denotes the usual complex number.

(1) Show that $x^2 + 1$ is in $\ker\mathcal{E}_i$.

(2) Since $\ker\mathcal{E}_i$ is an ideal and $\mathbb{Q}[x]$ is a PID, $\ker\mathcal{E}_i = (m(x))$ for some $m(x) \in \mathbb{Q}[x]$. Use the fact that $x^2 + 1$ is irreducible in $\mathbb{Q}[x]$, to show that $m(x) = a(x^2 + 1)$ for some nonzero constant $a \in \mathbb{Q}$.

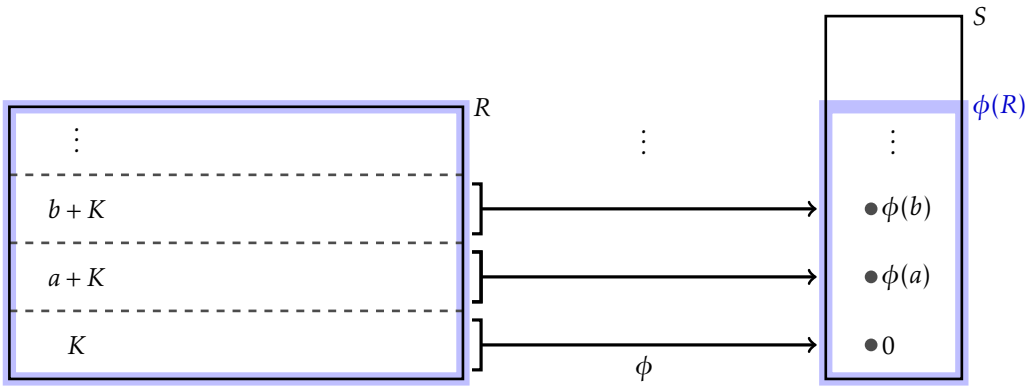
(3) Explain why $\ker\mathcal{E}_i = (x^2 + 1)$ (where $(x^2 + 1)$ is the ideal generated by $x^2 + 1$).

(4) Explain why the image of $\mathcal{E}_i$ is contained in $\mathbb{Q}(i)$.

5.6.1 First Isomorphism Theorem

The proof of the First Isomorphism Theorem for rings is essentially the same as for groups. Suppose we have a ring homomorphism $\phi : R \to S$. In order for ϕ to be an isomorphism, it must also be one-to-one and onto, but it may very well not be. However, with some slight adjustments, we can modify it to be both one-to-one and onto.

First, to make ϕ surjective, we change the codomain from S to $\phi(R)$. Technically, we have a different function now, but we will still call it ϕ. Anyway, $\phi : R \to \phi(R)$ is now a surjective homomorphism (by the definition of the image). But how do we make it one-to-one? By Theorem 5.117, we need the kernel to be trivial. To accomplish, we work with the quotient ring $R/\ker\phi$. In this way, we collapse all elements in the kernel to a single element (the coset $\ker\phi$) that maps to 0. Setting $K := \ker\phi$, the picture is like this.



Theorem 5.119 (First Isomorphism Theorem for Rings)**.** If $\phi : R \to S$ is a ring homomorphism, then $R/\ker\phi \cong \phi(R)$.

Problem 5.120. Use the First Isomorphism Theorem together Problem 5.114 to prove each of the following.

(1) $\mathbb{Q}[x]/(x) \cong \mathbb{Q}$ (where (x) is the ideal generated by x)

(2) $\mathbb{Z}/5\mathbb{Z} \cong \mathbb{Z}_5$

We now work towards a very nice and useable criterion to determine when a quotient ring is actually a field. Our approach will rely on Theorem 5.85, so we need to be able determine the relationship between ideals of R and ideals of the quotient R/I. We will investigate this via homomorphisms, which we can then apply to R/I using Problem 5.115.

Theorem 5.121. Let $\phi : R \to S$ be a surjective ring homomorphism. If I is an ideal of R, then $\phi(I) = \{b \in S \mid b = \phi(a) \text{ for some } a \in I\}$ (called the **image** of I) is an ideal of S.

Theorem 5.122. Let $\phi : R \to S$ be a ring homomorphism. If J is an ideal of S, then $\phi^{-1}(J) := \{a \in R \mid \phi(a) \in J\}$ (called the **inverse image** of J) is an ideal of R and, moreover, $\ker\phi \subseteq \phi^{-1}(J)$.

Combining Theorems 5.121 and 5.122 with Theorem 5.85 yields the following result.

Theorem 5.123. Let R be a commutative ring with $1 \neq 0$. Suppose $\phi : R \to S$ is a surjective ring homomorphism. Then $\ker\phi$ is maximal ideal of R if and only if S is field.

In light of Problem 5.115, our desired criterion for when a quotient ring is a field is a relatively quick consequence of the previous result.

Theorem 5.124. Let R be a commutative ring with $1 \neq 0$. Then I is maximal ideal of R if and only if R/I is field.

Problem 5.125. Let's revisit Problem 5.118.

(1) Explain why $\mathbb{Q}[x]/(x^2+1)$ is a field.

(2) Show that $\mathbb{Q}[x]/(x^2+1) \cong \mathbb{Q}(i)$.

Let's end this section (as well as this decidedly long chapter!) by mentioning that there is an analog of Theorem 5.124 characterizing when R/I is an integral domain. The condition is that I is a so-called *prime* ideal. The result is tangential to our story, so we'll leave it for another time (see Wikipedia or most any other book on abstract algebra).