FUNDAMENTAL CONCEPTS OF ALGEBRA

Donald L. White
Department of Mathematical Sciences
Kent State University

Release 3.0
January 12, 2009

Contents

1 Number Systems **1**
1.1 The Basic Number Systems 1
1.2 Complex Numbers 7
1.3 Algebraic Properties of Number Systems 17
1.4 Sets and Equivalence Relations 25
1.5 Formal Constructions of Number Systems 28

2 Basic Number Theory **35**
2.1 Principle of Mathematical Induction 35
2.2 Divisibility of Integers 41
2.3 Division Algorithm and Greatest Common Divisor 44
2.4 Properties of the Greatest Common Divisor 50
2.5 Prime Numbers 56
2.6 Prime Factorizations and Divisibility 62
2.7 Congruence 67
2.8 Congruence and Divisibility Tests 75

3 Polynomials **84**
3.1 Algebraic Properties of Polynomials 84
3.2 Binomial Coefficients and Binomial Theorem 92
3.3 Divisibility and Polynomials 98
3.4 Synthetic Division 109
3.5 Factors and Roots of Polynomials 112
3.6 Irreducible Polynomials 124
3.7 Irreducible Polynomials as Primes 132

A Trigonometry Review **139**

B Answers to Selected Problems **141**

Chapter 1

Number Systems

In this chapter we study the basic arithmetic and algebraic properties of the familiar number systems the integers, rational numbers, real numbers, and the possibly less familiar complex numbers. We will consider which algebraic properties these number systems have in common as well as the ways in which they differ.

We will use the following notation to denote sets of numbers.

$$\begin{aligned}
\mathbb{N} &= \{1, 2, 3, \ldots\} = \text{Natural Numbers} \\
\mathbb{Z} &= \{0, \pm 1, \pm 2, \pm 3, \ldots\} = \text{Integers} \\
\mathbb{Q} &= \left\{ \frac{a}{b} \;\middle|\; a, b \in \mathbb{Z},\, b \neq 0 \right\} = \text{Rational Numbers} \\
\mathbb{R} &= \text{Real Numbers} \\
\mathbb{C} &= \text{Complex Numbers}
\end{aligned}$$

1.1 The Basic Number Systems

The first numbers anyone learns about are the "counting numbers" or **natural numbers** $1, 2, 3, \ldots$, which we will denote by $\mathbb{N}$. We eventually learn about the basic operations of addition and multiplication of natural numbers. These operations are examples of **binary operations**, that is, operations that combine any two natural numbers to obtain another natural number. Addition and multiplication of natural numbers satisfy some very nice properties, such as commutativity, associativity, and the distributive law, which we will study more formally in a later section.

The other familiar arithmetic operations of subtraction and division are really just the "inverse operations" of addition and multiplication, and will not be considered as basic operations. (Although multiplication of natural numbers is really just repeated addition, this is a much less obvious interpretation in other number systems.) If we only wish to consider the natural numbers, we quickly encounter problems with subtraction and division. These operations can be performed on pairs of natural numbers only in some cases. For example, $3 - 5$ and $3 \div 5$ are not natural numbers.

In order to be able to subtract, we introduce the number 0 and the "negatives" of the natural numbers to obtain the set of **integers** $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$. The number 0 acts as a

neutral element or identity element for addition, because for any integer a, $a+0=a$. The negative of any integer a acts as an inverse for a relative to addition, because $a+(-a)=0$.

Of course, thinking of 0 and $-a$ in terms of addition leads to some interesting questions. Why is 0 times any number equal to 0? Why is the product of two negative numbers a positive number? More generally, what does multiplication by a negative number really mean? Is it still repeated addition? We will return to these questions later.

The operations of addition and multiplication in $\mathbb{Z}$ still satisfy the same properties as in the set of natural numbers $\mathbb{N}$, but $\mathbb{Z}$ has an identity element and inverses for addition. This allows us to subtract any integer from any other and obtain another integer.

Division can still only be performed on certain pairs of integers, however. Although the number 1 acts as a neutral element or identity element for multiplication in $\mathbb{Z}$, since $1 \cdot a = a$ for every integer a, the set $\mathbb{Z}$ does not have inverses relative to multiplication for most of its elements. In order to be able to divide, we must introduce fractions and obtain the set $\mathbb{Q}$ of rational numbers, with operations defined as follows.

Definition 1.1.1 *The set $\mathbb{Q}$ of* **rational numbers** *is the set of all quotients of integers (i.e., fractions),*

$$\mathbb{Q} = \left\{ \frac{a}{b} \;\middle|\; a, b \in \mathbb{Z},\, b \neq 0 \right\},$$

and we define

i. $\dfrac{a}{b} = \dfrac{a'}{b'}$ *if and only if* $ab' = ba'$,

ii. $\dfrac{a}{b} + \dfrac{c}{d} = \dfrac{ad+bc}{bd}$,

iii. $\dfrac{a}{b} \cdot \dfrac{c}{d} = \dfrac{ac}{bd}$.

Note that since $a = \frac{a}{1}$ for an integer a, every integer is also a rational number and we have $\mathbb{Z} \subseteq \mathbb{Q}$. The operations of addition and multiplication in $\mathbb{Q}$ still satisfy all of the properties as in the set of integers $\mathbb{Z}$.

If $q \neq 0$ is a rational number, say $q = \frac{a}{b}$, then $a \neq 0$, and $r = \frac{b}{a}$ is also a rational number. Since

$$q \cdot r = \frac{a}{b} \cdot \frac{b}{a} = \frac{ab}{ba} = 1,$$

the rational number $\frac{b}{a}$ is an inverse for $\frac{a}{b}$ relative to multiplication. That is, if $q = \frac{a}{b} \neq 0$ is a rational number, the multiplicative inverse of q is $q^{-1} = \frac{b}{a}$, the **reciprocal** of q.

A proper construction of the set $\mathbb{R}$ of real numbers requires tools from analysis beyond the scope of this text. A less rigorous description of $\mathbb{R}$ in terms of decimal expansions will suffice for our purposes. We will first recall the basics of decimal expansions and discuss decimal expansions of rational numbers.

A positive integer m can always be written in its decimal form and expressed as a sum of multiples of non-negative powers of 10:

$$m = n_k n_{k-1} \ldots n_2 n_1 n_0 = n_k \cdot 10^k + n_{k-1} \cdot 10^{k-1} + \cdots + n_2 \cdot 10^2 + n_1 \cdot 10^1 + n_0 \cdot 10^0.$$

Similarly, a positive number $r < 1$ with a terminating decimal expansion can be written as a sum of multiples of negative powers of 10:

$$r = 0.d_1d_2\ldots d_{k-1}d_k = d_1 \cdot 10^{-1} + d_2 \cdot 10^{-2} + \cdots + d_{k-1} \cdot 10^{-(k-1)} + d_k \cdot 10^{-k}.$$

If the decimal expansion does not terminate, then r is an "infinite sum" of multiples of negative powers of 10:

$$r = 0.d_1d_2d_3\ldots = d_1 \cdot 10^{-1} + d_2 \cdot 10^{-2} + d_3 \cdot 10^{-3} + \cdots.$$

In general, a (real) number can be written as a finite sum of multiples of non-negative powers of 10 plus a (possibly) infinite sum of multiples of negative powers of 10, and the coefficients in this sum are the digits in the decimal expansion of the number.

You may recall that the set of rational numbers defined in Definition 1.1.1 can also be described in terms of decimal expansions. The rational numbers are precisely those (real) numbers with terminating or repeating decimal expansions. For example, $3/8 = 0.375$ or $9/7 = 1.285714285714\ldots = 1.\overline{285714}$. In order to verify this characterization of rational numbers, we must show that if $\frac{a}{b}$ is any rational number, then the decimal expansion of $\frac{a}{b}$ either terminates or is a repeating decimal, and that every terminating or repeating decimal is the decimal expansion of a rational number $\frac{a}{b}$.

We will first verify that every rational number $\frac{a}{b}$ has either a terminating or repeating decimal expansion. First note that we may assume $\frac{a}{b}$ is positive and $a < b$. (Why?) The decimal expansion of $\frac{a}{b}$ is obtained by performing the long division $a \div b$.

In the algorithm for long division, we place a decimal point to the right of a and append just enough zeros to obtain a number larger than a (without the decimal). We then divide, placing the quotient above the division sign and obtaining a remainder r with $0 \leqslant r < b$. (This is possible by the Division Algorithm, which we will study formally in §2.3.) The algorithm is then repeated to divide r by b and so on, as demonstrated for $3 \div 17$ in the following example:

$$\begin{array}{rcccccc}
 & 0. & 1 & 7 & 6 & 4 & \ldots \\
\cline{2-7}
17\,) & 3. & 0 & & & & \\
 & 1 & 7 & & & & \\
\cline{2-3}
 & \mathbf{1} & \mathbf{3} & 0 & & & \\
 & 1 & 1 & 9 & & & \\
\cline{3-4}
 & & \mathbf{1} & \mathbf{1} & 0 & & \\
 & & 1 & 0 & 2 & & \\
\cline{4-5}
 & & & & \mathbf{8} & 0 & \\
 & & & & . & . & .
\end{array}$$

The boldface numbers in the example are the remainders.

If a remainder of 0 is obtained at some stage, then the decimal expansion terminates, as for $\frac{3}{8}$:

$$\begin{array}{rcccc}
 & 0. & 3 & 7 & 5 \\
\cline{2-5}
8\,) & 3. & 0 & & \\
 & 2 & 4 & & \\
\cline{3-4}
 & & \mathbf{6} & 0 & \\
 & & 5 & 6 & \\
\cline{4-5}
 & & & \mathbf{4} & 0 \\
 & & & 4 & 0 \\
\cline{5-5}
 & & & & 0
\end{array}$$

If the remainder is never zero, then each remainder r satisfies $1 \leqslant r \leqslant b - 1$. Thus there are only finitely many different possible remainders, and at some point a remainder must repeat. Once a

remainder repeats, the same sequence of quotients and remainders must repeat forever, and the decimal expansion is repeating, as for $\frac{7}{54}$:

$$
\begin{array}{rl}
 & 0.\ 1\ 2\ 9\ 6\ 2\ 9\ 6\ \ldots \\
\hline
54\,) & 7.\ 0 \\
 & \ 5\ 4 \\
 & \ \mathbf{1}\ \mathbf{6}\ 0 \\
 & \ 1\ 0\ 8 \\
 & \ \ \mathbf{5}\ \mathbf{2}\ 0 \\
 & \ \ 4\ 8\ 6 \\
 & \ \ \ \mathbf{3}\ \mathbf{4}\ 0 \\
 & \ \ \ 3\ 2\ 4 \\
 & \ \ \ \ \mathbf{1}\ \mathbf{6}\ 0 \\
 & \ \ \ \ 1\ 0\ 8 \\
 & \ \ \ \ \ \mathbf{5}\ \mathbf{2}\ 0 \\
 & \ \ \ \ \ 4\ 8\ 6 \\
 & \ \ \ \ \ \ \mathbf{3}\ \mathbf{4}\ 0 \\
 & \ \ \ \ \ \ 3\ 2\ 4 \\
 & \ \ \ \ \ \ \ \mathbf{1}\ \mathbf{6} \\
 & \ \ \ \ \ \ \ .\ .\ .
\end{array}
$$

Notice that 16 is the first remainder to repeat in this example and the corresponding quotient, 2, is the start of the repeating decimal.

Exploration: Under what conditions on a fraction a/b in lowest terms will the decimal expansion be terminating? Investigate this question by searching in number theory texts or Internet sources.

Conversely, we must verify that every terminating or repeating decimal expansion is the decimal expansion of a rational number. Again, we may assume the decimal number is positive and less than one. (Why?) A terminating decimal with k decimal places, say $.d_1d_2\ldots d_k$, can be written as the fraction

$$\frac{d_1d_2\ldots d_k}{10^k}.$$

We say that the **period** of a repeating decimal is k if the length of the shortest repeating sequence of digits is k. For example, the period of $0.454545\ldots = 0.\overline{45}$ is 2 and the period of $0.1234563456\ldots = 0.12\overline{3456}$ is 4.

A repeating decimal of period k, say

$$R = 0.d_1d_2\ldots d_j\overline{r_1r_2\ldots r_k},$$

can be expressed as a fraction, that is, a rational number, as follows. First, multiply R by 10^k to obtain 10^kR. This has the effect of moving the decimal point k places to the right, or equivalently, shifting the digits of R k places to the left. Because the period of R is k, the digits of R and 10^kR will be the same after some decimal place. Thus the number

$$10^kR - R = (10^k - 1)R$$

will be a terminating decimal, hence equal to some fraction T. Therefore $(10^k - 1)R = T$ and

$$R = \frac{T}{10^k - 1}$$

is a rational number.

Example: Write the repeating decimal $R = 0.12345345\ldots = 0.12\overline{345}$ as a fraction.

The period of R is 3, so we calculate $10^3R = 1000R$:

$$\begin{aligned} R &= 0\,.\,12\,345\,345 \\ 1000R &= 123\,.\,45\,345\,345, \end{aligned}$$

thus

$$1000R - R = 999R = 123.33 = \frac{12333}{100}$$

and so

$$R = \frac{12333}{999 \cdot 100} = \frac{12333}{99900} = \frac{4111}{33300}.$$

□

The set $\mathbb{R}$ of **real numbers** consists of all possible decimal expansions. We have shown that the rational numbers are precisely those real numbers with either terminating or repeating decimal expansions. As there are clearly decimal expansions that are not repeating (for example, $0.01011011101111\ldots$), not all real numbers are rational. Those real numbers that do not have terminating or repeating decimal expansions, and therefore are not rational, are called **irrational numbers**. Thus $\mathbb{R}$ consists of the rational numbers along with the irrational numbers.

The irrational numbers are real numbers that cannot be expressed as a quotient of two integers. Some well-known examples of irrational numbers are $\sqrt{2}$, e, and π, but there are many others. In fact, in a sense that can be made precise, most real numbers are irrational.

For computational purposes, we usually approximate irrational numbers by rational numbers. For example, your calculator probably uses the approximation 3.141592654 (a rational number) for π in calculations.

This approximation is sufficiently accurate for most purposes, but π, or any other irrational number, can be approximated to within any desired degree of accuracy by a rational number. Probably the easiest way to see this is to note that truncating the decimal expansion of the irrational number results in a terminating decimal, hence a rational number, and the greater the number of decimal places used, the closer the approximation will be.

In particular, if I is an irrational number and R is the rational number obtained by using the digits of I to the left of the decimal and the first k digits of I to the right of the decimal, then $0 < I - R < 10^{-k}$. Thus R approximates I to within 10^{-k}. For example, if $R = 3.141592653589793$, then

$$0 < \pi - R < 10^{-15} = 0.000000000000001.$$

Finally, we note that it is not possible to determine whether a number is rational or irrational from any terminating decimal approximation. For example, a calculator with a 10-digit display will show $e \approx 2.718281828$, which certainly appears to be a repeating decimal, although e is in fact irrational. (The next digit in the decimal expansion of e is 4.) On the other hand, the calculator shows $1/17 \approx .0588235294$, which shows no evidence of repetition, although $1/17$ is clearly rational. To verify that a given number I is irrational, it is necessary to prove that there cannot be integers a and b such that $I = a/b$.

§1.1 Exercises

1. Use long division to find the repeating decimal expansion of the following rational numbers. **Show your work on the long division.**

   (a) $\dfrac{5}{101}$ (c) $\dfrac{17}{135}$

   (b) $\dfrac{47}{110}$ (d) $\dfrac{5}{14}$

2. Convert the following repeating decimal expansions to fractions in lowest terms.

   (a) $0.393939\ldots = 0.\overline{39}$

   (b) $4.302302302\ldots = 4.\overline{302}$

   (c) $57.13478478478\ldots = 57.13\overline{478}$

   (d) $102.102537253725372\ldots = 102.102\overline{5372}$

3. Show that $1 = 0.999999\ldots = 0.\overline{9}$.

4. Explain why the period of a rational number a/b with a repeating decimal is at most $b - 1$.

5. By Definition 1.1.1, two fractions $\frac{a}{b}$ and $\frac{a'}{b'}$ are equal if and only if $ab' = ba'$. Show that if $\frac{a}{b} = \frac{a'}{b'}$ and $\frac{c}{d} = \frac{c'}{d'}$, then

   (a) $\dfrac{a}{b} + \dfrac{c}{d} = \dfrac{a'}{b'} + \dfrac{c'}{d'}$, that is, $\dfrac{ad + bc}{bd} = \dfrac{a'd' + b'c'}{b'd'}$ and

   (b) $\dfrac{a}{b} \cdot \dfrac{c}{d} = \dfrac{a'}{b'} \cdot \dfrac{c'}{d'}$, that is, $\dfrac{ac}{bd} = \dfrac{a'c'}{b'd'}$.

   (This exercise shows that addition and multiplication of rational numbers are "well-defined," so that the sum or product of two rational numbers does not depend on the particular representation of the numbers as fractions.)

6. Write a paragraph with your explanation to a middle school or high school student as to why 0 times any number is 0.

7. Write a paragraph with your explanation to a middle school or high school student as to why the product of two negative numbers is a positive number.

1.2 Complex Numbers

In this section we introduce the arithmetic and geometry of complex numbers.

Definition 1.2.1 *The set* $\mathbb{C}$ *of* **complex numbers** *is the set of symbols* $a+bi$*, where* a *and* b *are real numbers, and we define:*

i. $a+bi = c+di$ *if and only if* $a=c$ *and* $b=d$*,*

ii. $(a+bi)+(c+di) = (a+c)+(b+d)i$*,*

iii. $(a+bi)\cdot(c+di) = (ac-bd)+(ad+bc)i$*.*

Note that the definition of multiplication implies (with $a=c=0$, $b=d=1$) that $i^2=-1$. Therefore we think of i as the square root of -1. Of course $(-i)^2=-1$ as well, so we define $i=\sqrt{-1}$, the principal square root of -1. If c is a positive real number, we also define $\sqrt{-c}=\sqrt{c}\cdot i$.

With the convention that $i^2=-1$, multiplication of complex numbers follows the usual rules for multiplying binomials:

$$(a+bi)\cdot(c+di) = ac+adi+bci+bdi^2 = (ac-bd)+(ad+bc)i.$$

Example: If $w=3+5i$ and $z=2-7i$, we have

$$w+z = (3+5i)+(2-7i) = (3+2)+(5-7)i = 5-2i$$

and

$$w\cdot z = (3+5i)\cdot(2-7i) = [3(2)-5(-7)]+[3(-7)+5(2)]i = 41-11i.$$

If a is a real number, we can write $a=a+0i$ and consider a to be a complex number as well. Hence $\mathbb{R}$ is a subset of $\mathbb{C}$. In fact, we have the following containments among the sets we have considered:

$$\mathbb{N}\subseteq\mathbb{Z}\subseteq\mathbb{Q}\subseteq\mathbb{R}\subseteq\mathbb{C}.$$

In particular, the numbers $0=0+0i$ and $1=1+0i$ are elements of $\mathbb{C}$ and are the identity elements for addition and multiplication, respectively, in $\mathbb{C}$. If $z=a+bi$ then $-z=-a-bi$ is its additive inverse. It is an easy exercise to verify these statements by direct calculations.

In order to discuss multiplicative inverses and division in $\mathbb{C}$, we require more definitions.

Definition 1.2.2 *If* $z=a+bi$ *is a complex number, we define*

i. *the* **real part** *of* z *is* $\mathrm{Re}(z)=a$*,*

ii. *the* **imaginary part** *of* z *is* $\mathrm{Im}(z)=b$*.*

(Note that the imaginary part of z is b and NOT bi.)

Examples:

1. $\mathrm{Re}(3+5i) = 3$ and $\mathrm{Im}(3+5i) = 5$.
2. $\mathrm{Re}(2-7i) = 2$ and $\mathrm{Im}(2-7i) = -7$.
3. $\mathrm{Re}(3i) = 0$ and $\mathrm{Im}(3i) = 3$.
4. $\mathrm{Re}(7) = 7$ and $\mathrm{Im}(7) = 0$.

Definition 1.2.3 *If $z = a+bi$ is a complex number, the* **(complex) conjugate** *of z is $\overline{z} = a-bi$.*

Examples:

1. $\overline{3+5i} = 3-5i$.
2. $\overline{2-7i} = 2+7i$.
3. $\overline{3i} = -3i$.
4. $\overline{7} = 7$.

The following properties of conjugates are easy to verify using the definitions.

Proposition 1.2.4 *If z and w are complex numbers, then*

i. $\overline{z+w} = \overline{z} + \overline{w}$

ii. $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$.

Proof. (i) Let $z = a+bi$ and $w = c+di$, so that $\overline{z} = a-bi$ and $\overline{w} = c-di$. We have

$$\begin{aligned}
\overline{z+w} &= \overline{(a+bi)+(c+di)} \\
&= \overline{(a+c)+(b+d)i} \text{ by Definition 1.2.1 (ii),} \\
&= (a+c)-(b+d)i \text{ by Definition 1.2.3,} \\
&= (a-bi)+(c-di) \text{ by Definition 1.2.1 (ii),} \\
&= \overline{z}+\overline{w},
\end{aligned}$$

and so $\overline{z+w} = \overline{z}+\overline{w}$ as claimed.

(ii) The proof of (ii) is similar and is left as an exercise. (See Exercise 1.2.11.) □

Proposition 1.2.5 *If $z = a+bi$ is a complex number, then*

i. $z + \overline{z} = 2a = 2\mathrm{Re}(z)$

ii. $z - \overline{z} = 2bi = 2\mathrm{Im}(z) \cdot i$.

Proof. (i) Let $z = a+bi$ so that $\overline{z} = a-bi$ and $\mathrm{Re}(z) = a$. We have

$$\begin{aligned}
z+\overline{z} &= (a+bi)+(a-bi) \\
&= (a+a)+(b-b)i \text{ by Definition 1.2.1 (ii),} \\
&= 2a+0i \\
&= 2a,
\end{aligned}$$

and so $z+\overline{z} = 2a = 2\mathrm{Re}(z)$ as claimed.

(ii) The proof of (ii) is similar and is left as an exercise. (See Exercise 1.2.12.) □

The proof of the next result is a computation similar to the previous proofs and is left as an exercise. (See Exercise 1.2.13.)

Proposition 1.2.6 *If $z = a + bi$ is a complex number, then $z \cdot \overline{z} = a^2 + b^2$, a non-negative real number.*

We now consider multiplicative inverses and division of complex numbers. If $z = a + bi$ is a non-zero complex number, then a or b is non-zero, and $a^2 + b^2$ is a non-zero real number. By the proposition above, we have

$$\frac{z\overline{z}}{a^2 + b^2} = 1,$$

hence

$$z \cdot \frac{\overline{z}}{a^2 + b^2} = z \cdot \left(\frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}\, i \right) = 1.$$

We therefore have:

Proposition 1.2.7 *If $z = a+bi$ is a non-zero complex number, then there is a complex number z^{-1} such that $zz^{-1} = 1$. In particular,*

$$z^{-1} = \frac{\overline{z}}{z\overline{z}} = \frac{\overline{z}}{a^2 + b^2} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}\, i.$$

A quotient $\frac{a+bi}{c+di}$ of two complex numbers can be written in the standard form by multiplying the numerator and denominator by the conjugate of the denominator, which leaves a real number in the denominator. This procedure is similar to "rationalizing" a denominator containing a root. That is,

$$\frac{a + bi}{c + di} = \frac{(a + bi)(c - di)}{(c + di)(c - di)} = \frac{(ac + bd) + (bc - ad)i}{c^2 + d^2} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}\, i.$$

It is best to learn the procedure demonstrated here, and not to memorize this final formula.

Examples:

1. If $z = 2 + 3i$, then the multiplicative inverse or reciprocal of z is

$$z^{-1} = \frac{1}{2 + 3i} = \frac{1}{2 + 3i} \cdot \frac{2 - 3i}{2 - 3i} = \frac{2 - 3i}{2^2 + 3^2} = \frac{2 - 3i}{13} = \frac{2}{13} - \frac{3}{13}\, i.$$

2. If $w = -1 + 3i$ and $z = 2 - 5i$, the the quotient w/z is

$$\frac{w}{z} = \frac{-1 + 3i}{2 - 5i} = \frac{-1 + 3i}{2 - 5i} \cdot \frac{2 + 5i}{2 + 5i} = \frac{(-2 - 15) + (-5 + 6)i}{2^2 + 5^2} = \frac{-17 + i}{29} = -\frac{17}{29} + \frac{1}{29}\, i.$$

Geometry of Complex Numbers

We often represent real numbers geometrically as points on a number line. Similarly, complex numbers are represented as points in the **complex plane**. We use the usual coordinate plane (xy-plane) with the x-axis as the **real axis** and the y-axis as the **imaginary axis**. The number $z = a+bi$ is represented by the point with coordinates (a, b), as in Figure 1.1. Thus the real number $a = a + 0i$ is represented by the point $(a, 0)$ on the real axis, and the real number line coincides with the real axis.

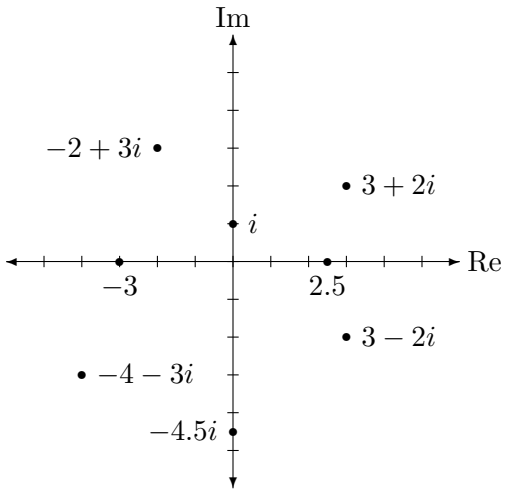


Figure 1.1: The Complex Plane

With these conventions, $\overline{z}$ is the reflection of z in the real axis (x-axis). Also, the distance of $z = a + bi$ from the origin is $\sqrt{a^2 + b^2}$. Recall that the absolute value of a real number is the distance from the number to the origin on the number line. Accordingly, we make the following definition for complex numbers.

Definition 1.2.8 *If $z = a + bi$ is a complex number, the* **absolute value** *or* **modulus** *of z, denoted $|z|$, is the distance from z to the origin in the complex plane. Thus*

$$|z| = \sqrt{a^2 + b^2} = \sqrt{z\overline{z}}.$$

Note that this also says $z\overline{z} = |z|^2$.

Examples:

1. $|3 + 2i| = \sqrt{3^2 + 2^2} = \sqrt{13}$.
2. $|4 - 5i| = \sqrt{4^2 + (-5)^2} = \sqrt{16 + 25} = \sqrt{41}$.
3. $|-7i| = \sqrt{0^2 + (-7)^2} = \sqrt{49} = 7$.
4. $|-6| = \sqrt{(-6)^2 + 0^2} = \sqrt{36} = 6$.

We now consider the geometric interpretations of addition and multiplication of complex numbers. Let $z = a + bi$ and $w = c + di$ be complex numbers represented by the points $P = (a, b)$ and $Q = (c, d)$ in the complex plane, and let $O = (0, 0)$ be the origin. Assume that the points P, Q, and O are not all on the same line. Then the point $S = (a + c, b + d)$ representing the sum $z+w = (a+c)+(b+d)i$ is the endpoint of the diagonal of the parallelogram with the line segments $\overline{OP}$ and $\overline{OQ}$ as sides. This is illustrated in the example in Figure 1.2. (The reader who is familiar with linear algebra should notice that this is the geometric description of vector addition in the vector space $\mathbb{R}^2$.)

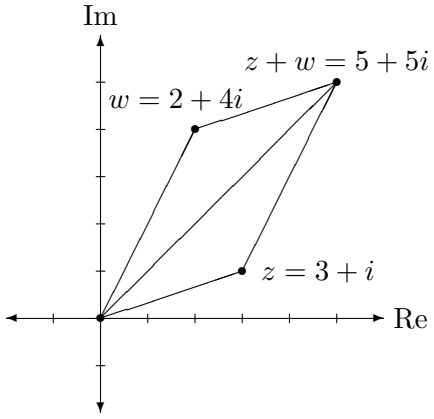


Figure 1.2: Addition of Complex Numbers

Exercise: Describe geometrically the sum $z + w$ in the case where the points P, Q, and O lie on the same line. Experiment with some specific examples and consider separately the cases where P and Q lie on the same side of the origin and where they lie on opposite sides of the origin.

The additive inverse $-z = -a - bi$ of a complex number $z = a + bi$ is the reflection of z in the origin. Using the geometric descriptions of addition and additive inverses, along with the fact that $w - z = w + (-z)$, we obtain a geometric interpretation of subtraction of complex numbers, as illustrated by the example in Figure 1.3.

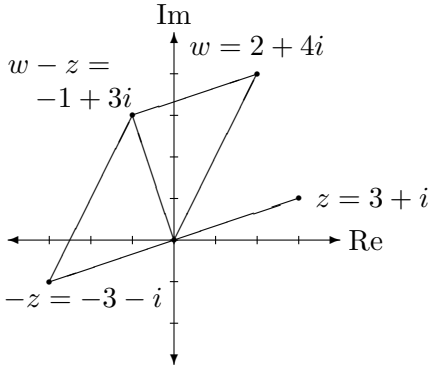


Figure 1.3: Subtraction of Complex Numbers

In order to describe multiplication of complex numbers geometrically, it will be convenient to use trigonometry and the "polar form" of complex numbers. A point $P = (a, b)$ in the plane can be uniquely determined by its distance r from the origin O and the angle θ between the positive x-axis and the segment $\overline{OP}$. We call (r, θ) the **polar coordinates** of the point P.

We can similarly obtain a polar representation of a complex number $z = a + bi$. As noted previously, the distance of z from the origin is $r = |z| = \sqrt{a^2 + b^2}$, the modulus of z. We will also need the following definition.

Definition 1.2.9 *The angle θ between the positive real axis and the line segment from the origin to the complex number z is the* **argument** *of z, denoted* arg z.

Note that the angle arg z is not unique, adding any integer multiple of 2π will yield another argument. The polar form of $z = a + bi$ is illustrated in Figure 1.4.

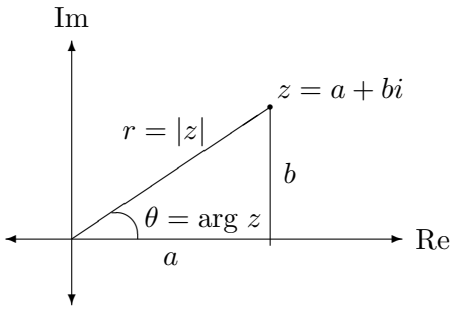


Figure 1.4: Polar Form of Complex Numbers

The polar and standard representations of z are related by the equations

$$a = r\cos\theta, \quad b = r\sin\theta$$

and

$$r = |z| = \sqrt{a^2 + b^2}, \quad \tan\theta = \frac{b}{a}.$$

Therefore, the complex number z can be written in the **polar form**

$$z = r(\cos\theta + i\sin\theta).$$

Note that $r = |z|$ is a non-negative real number and $\cos\theta + i\sin\theta$ is a complex number of modulus $\sqrt{\cos^2\theta + \sin^2\theta} = 1$, hence lies on the unit circle centered at the origin.

Examples:

1. Find the polar form of the complex number $z = 3 + 3\sqrt{3}\,i$.

   We have $r = \sqrt{3^2 + (3\sqrt{3})^2} = \sqrt{9 + 27} = \sqrt{36} = 6$ and $\tan\theta = 3\sqrt{3}/3 = \sqrt{3}$. Since z is in the first quadrant and $\tan\theta = \sqrt{3}$, we have $\theta = \pi/3$. Therefore, the polar form of z is

   $$z = 6\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right).$$

2. Find the polar form of the complex number $w = -2\sqrt{3} + 2i$.

   We have $r = \sqrt{(-2\sqrt{3})^2 + 2^2} = \sqrt{12+4} = \sqrt{16} = 4$ and $\tan\theta = 2/(-2\sqrt{3}) = -1/\sqrt{3}$. The reference angle is $\theta' = \arctan(1/\sqrt{3}) = \pi/6$, and since w is in the second quadrant, we have $\theta = \pi - \pi/6 = 5\pi/6$. Therefore, the polar form of w is

$$w = 4\left(\cos\frac{5\pi}{6} + i\sin\frac{5\pi}{6}\right).$$

The next result says that in order to multiply two complex numbers, we multiply their moduli and *add* their arguments.

Theorem 1.2.10 *If* $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$ *and* $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$, *then*

$$z_1 z_2 = r_1 r_2[\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)].$$

Proof. Suppose we have two complex numbers $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$ and $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$. Their product is

$$\begin{aligned} z_1 z_2 &= r_1(\cos\theta_1 + i\sin\theta_1)\cdot r_2(\cos\theta_2 + i\sin\theta_2) \\ &= r_1 r_2[(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i(\cos\theta_1\sin\theta_2 + \sin\theta_1\cos\theta_2)]. \end{aligned}$$

By the angle sum formulas for sine and cosine (see Appendix A), we have

$$\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 = \cos(\theta_1+\theta_2)$$

and

$$\cos\theta_1\sin\theta_2 + \sin\theta_1\cos\theta_2 = \sin(\theta_1+\theta_2).$$

Substituting yields $z_1 z_2 = r_1 r_2[\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)]$, as claimed. □

Using this theorem and mathematical induction (see §2.1), we obtain the following corollary.

Corollary 1.2.11 *If* $z = r(\cos\theta + i\sin\theta)$ *and* n *is a positive integer, then*

$$z^n = r^n(\cos n\theta + i\sin n\theta).$$

In the case where $r = 1$, Corollary 1.2.11 becomes

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$$

and is known as **de Moivre's Theorem**. This says that the nth power of a complex number z of modulus 1 (i.e., a number on the unit circle) is the number on the unit circle whose argument is n times the argument of z. See §2.1, Example 3, for the formal proof.

Examples:

1. Use polar form to calculate $z \cdot w$ for $z = 3 + 3\sqrt{3}\,i$ and $w = -2\sqrt{3} + 2i$.

   We saw in the examples above that

   $$z = 6\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right)$$

   and

   $$w = 4\left(\cos\frac{5\pi}{6} + i\sin\frac{5\pi}{6}\right).$$

   Therefore, by Theorem 1.2.10,

   $$\begin{aligned} z \cdot w &= 6 \cdot 4\left[\cos\left(\frac{\pi}{3} + \frac{5\pi}{6}\right) + i\sin\left(\frac{\pi}{3} + \frac{5\pi}{6}\right)\right] \\ &= 24\left(\cos\frac{7\pi}{6} + i\sin\frac{7\pi}{6}\right) \\ &= 24\left(-\frac{\sqrt{3}}{2} - \frac{1}{2}\,i\right) \\ &= -12\sqrt{3} - 12i. \end{aligned}$$

2. Use polar form to calculate $(1 - i)^{27}$.

   First convert $1 - i$ to polar form. We have $r = \sqrt{1^2 + (-1)^2} = \sqrt{2}$ and $\tan\theta = -1/1 = -1$. The reference angle is then $\pi/4$ and and since $1 - i$ is in the fourth quadrant, we have $\theta = 2\pi - \pi/4 = 7\pi/4$. Hence the polar form is

   $$1 - i = \sqrt{2}\left(\cos\frac{7\pi}{4} + i\sin\frac{7\pi}{4}\right).$$

   Therefore, by Corollary 1.2.11,

   $$\begin{aligned} (1-i)^{27} &= \left[\sqrt{2}\left(\cos\frac{7\pi}{4} + i\sin\frac{7\pi}{4}\right)\right]^{27} \\ &= (\sqrt{2})^{27}\left[\cos\left(27 \cdot \frac{7\pi}{4}\right) + i\sin\left(27 \cdot \frac{7\pi}{4}\right)\right] \\ &= 2^{27/2}\left(\cos\frac{189\pi}{4} + i\sin\frac{189\pi}{4}\right) \\ &= 2^{13} \cdot 2^{1/2}\left[\cos\left(\frac{5\pi}{4} + 23 \cdot 2\pi\right) + i\sin\left(\frac{5\pi}{4} + 23 \cdot 2\pi\right)\right] \\ &= 8192\sqrt{2}\left(\cos\frac{5\pi}{4} + i\sin\frac{5\pi}{4}\right) \\ &= 8192\sqrt{2}\left(-\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\,i\right) \\ &= -8192(1 + i). \end{aligned}$$

If z is a complex number, then $-z = (-1) \cdot z$, and the additive inverse can be interpreted geometrically in terms of multiplication. We have $|-1| = 1$ and $\arg(-1) = \pi$, hence by Theorem 1.2.10, $|-z| = |z|$ and $\arg(-z) = \pi + \arg z$. It follows that $-z$ is the reflection of z in the origin, as noted previously.

Both Theorem 1.2.10 and de Moivre's Theorem may be more easily understood if we consider the complex exponential function $f(z) = e^z$. Using infinite series (and trigonometric functions), it is possible to extend the definition of the natural exponential function from the real numbers to the complex numbers and to obtain **Euler's Formula**

$$e^{i\theta} = \cos\theta + i\sin\theta,$$

where θ is a real number. Thus a complex number of modulus 1 and argument θ can be expressed as $e^{i\theta}$, and a complex number z of modulus r and argument θ can be expressed as

$$z = r(\cos\theta + i\sin\theta) = re^{i\theta}.$$

Theorem 1.2.10 then follows from the usual laws of exponents. If $z_1 = r_1e^{i\theta_1}$ and $z_2 = r_2e^{i\theta_2}$, then

$$z_1z_2 = r_1e^{i\theta_1}r_2e^{i\theta_2} = r_1r_2e^{i(\theta_1+\theta_2)}$$

and if $z = re^{i\theta}$, then

$$z^n = (re^{i\theta})^n = r^ne^{in\theta}.$$

Euler's Formula also implies a nice relation among the important numbers 0, 1, e, π, and i. Letting $\theta = \pi$, Euler's formula becomes

$$e^{i\pi} = \cos\pi + i\sin\pi = -1$$

or

$$e^{i\pi} + 1 = 0.$$

---

§1.2 Exercises

1. Determine the real part, the imaginary part, the complex conjugate, and the modulus of each of the following.

   (a) $3 + 5i$ (c) $-4 + i$

   (b) $7 - 2i$ (d) 5

2. Perform the indicated operations. Write answers in the form $a + bi$, where a and b are real numbers.

   (a) $(2 + 3i) + (-3 + 4i)$ (c) $(4+i)+(3+2i)+(4-5i)$

   (b) $(5 - 2i) - (3 + 7i)$ (d) $(2+4i)-(1+2i)+(3-2i)$

3. Compute the following products and write in the form $a+bi$, where a and b are real numbers.

   (a) $(2 + 3i)(-3 + 4i)$ (c) $(2 + 3i)(2 - 3i)$

   (b) $(7 + 2i)(3 - 2i)$ (d) $i(5 - 7i)$

4. Find the multiplicative inverse of each of the following complex numbers. Write answers in the form $a + bi$, where a and b are real numbers.

   (a) $3 + 4i$ (c) $2 + 3i$

   (b) $7 - 2i$ (d) $7i$

5. Compute the following quotients and write in the form $a+bi$, where a and b are real numbers.

   (a) $\dfrac{5 + 4i}{3 + 2i}$

   (b) $(3 + 2i)/(5 + 4i)$

   (c) $(-3 + 4i) \div (2 - i)$

6. Find the modulus and argument of each of the following complex numbers and write the numbers in polar form.

   (a) $1 - i$ (c) $3 - 3\sqrt{3}\,i$ (e) $5i$

   (b) $-\sqrt{3} - i$ (d) $-\sqrt{2} + \sqrt{2}\,i$ (f) -7

7. Let $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$ and $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$. Using Theorem 1.2.10, deduce a formula for the polar form of the quotient z_1/z_2.

8. Let $z = r(\cos\theta + i\sin\theta)$. Using Theorem 1.2.10, deduce a formula for the polar form of the reciprocal $1/z$.

9. Let $z = 1 + \sqrt{3}\,i$, so $|z| = 2$ and $\arg z = \frac{\pi}{3}$, and $w = -2\sqrt{2} + 2\sqrt{2}\,i$, so $|w| = 4$ and $\arg w = \frac{3\pi}{4}$. Find the modulus and argument of each of the following complex numbers and write the numbers in polar form.

   (a) $z \cdot w$ (c) $z^2 \cdot w$ (e) z/w

   (b) w^2 (d) z^5 (f) $1/z$

10. Use Theorem 1.2.10 to evaluate the following powers. Write your answers in the standard form $a + bi$, with a and b *exact* real numbers and without trigonometric functions.

    (a) $(1 + i)^{10}$ (c) $\left(-\dfrac{1}{2} + \dfrac{\sqrt{3}}{2}\,i\right)^{17}$

    (b) $\left(\dfrac{\sqrt{3}}{2} + \dfrac{1}{2}\,i\right)^{14}$ (d) $\left(\dfrac{\sqrt{2}}{2} + \dfrac{\sqrt{2}}{2}\,i\right)^{8}$

11. Verify Proposition 1.2.4 (ii).

12. Verify Proposition 1.2.5 (ii).

13. Verify Proposition 1.2.6.

14. Show that if z is a complex number then $\overline{\overline{z}} = z$.

15. Let z be a complex number. Show that $z \in \mathbb{R}$ if and only if $\overline{z} = z$.

1.3 Algebraic Properties of Number Systems

In this section, we discuss various algebraic properties that our number systems share. Before reading further, consider the following questions.

Class Preparation Problems:

1. What properties of addition or multiplication of natural numbers are used in the following equations or calculations? Write out the properties carefully. *Why* are they true?

   (a) $5 + (7 + 9) = 5 + (9 + 7)$

   (b) $6 + (2 + 3) = (2 + 3) + 6$

   (c) $5 + (7 + 9) = (5 + 7) + 9$

   (d) Compute $3 + 4 + 8$.

   (e) $2 \cdot (3 \cdot 5) = 2 \cdot (5 \cdot 3)$

   (f) Compute $3 \cdot 4 \cdot 5$.

   (g) $2 \cdot (3 \cdot 5) = (3 \cdot 2) \cdot 5$

   (h) $3 \cdot (5 + 7) = 3 \cdot 5 + 3 \cdot 7$

2. Besides those demonstrated above, do you know any other "basic" properties of addition or multiplication of natural numbers?

3. Does subtraction or division make sense in $\mathbb{N}$?

4. Now consider the set of integers $\mathbb{Z}$. Are the properties discussed for $\mathbb{N}$ also true for $\mathbb{Z}$? How could we define 0 and negative numbers in terms of addition? How could we define subtraction?

5. With 0 and $-n$ defined via addition, use the known properties of addition and multiplication to show:

   (a) $0 \cdot n = 0$ for all $n \in \mathbb{Z}$.

   (b) $(-1) \cdot a = -a$ for all $a \in \mathbb{Z}$.

   (c) $(-1) \cdot (-1) = 1$, or more generally, $(-a) \cdot (-b) = a \cdot b$ for all $a, b \in \mathbb{Z}$.

6. With the definitions discussed above, we can define subtraction in terms of addition in $\mathbb{Z}$. Does subtraction satisfy all of the same properties as addition? If not, which ones fail?

7. In order to be able to subtract, we extended from $\mathbb{N}$ to $\mathbb{Z}$. What set is needed in order to allow us to divide?

8. Are all of the properties of addition and multiplication for $\mathbb{Z}$ also valid for $\mathbb{Q}$? What additional properties does $\mathbb{Q}$ have that $\mathbb{Z}$ does not?

9. How can we define division in $\mathbb{Q}$ in terms of multiplication? Does division satisfy all of the same properties as multiplication? If not, which ones fail?

10. Do addition and multiplication in $\mathbb{R}$ and $\mathbb{C}$ satisfy the same properties as in $\mathbb{Q}$?

The questions above concern basic properties of addition and multiplication in the number systems we have discussed. The following list summarizes the properties that may (or may not) be satisfied by a set S on which addition and multiplication are defined.

Definition 1.3.1 (Algebraic Properties) *Let S be a set on which addition $(+)$ and multiplication $(\cdot)$ are defined. We define the following (potential) properties.*

Properties of Addition:

i. **Closure under Addition:** *$a + b$ is in S for all a and b in S.*

ii. **Associative Law of Addition:** *$a + (b + c) = (a + b) + c$ for all a, b, and c in S.*

iii. **Commutative Law of Addition:** *$a + b = b + a$ for all a and b in S.*

iv. **Additive Identity:** *There is an element 0 in S such that $a + 0 = 0 + a = a$ for all a in S.*

v. **Additive Inverses:** *For each element a in S, there is an element $-a$ in S such that $a + (-a) = (-a) + a = 0$.*

Properties of Multiplication:

vi. **Closure under Multiplication:** *$a \cdot b$ is in S for all a and b in S.*

vii. **Associative Law of Multiplication:** *$a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all a, b, and c in S.*

viii. **Commutative Law of Multiplication:** *$a \cdot b = b \cdot a$ for all a and b in S.*

ix. **Multiplicative Identity:** *There is an element 1 in S such that $a \cdot 1 = 1 \cdot a = a$ for all a in S.*

x. **Multiplicative Inverses:** *For each element $a \neq 0$ in S, there is an element a^{-1} in S such that $a \cdot (a^{-1}) = (a^{-1}) \cdot a = 1$.*

Property Relating Addition and Multiplication:

xi. **Distributive Laws:**
$c \cdot (a + b) = c \cdot a + c \cdot b$ for all a, b, and c in S,
$(a + b) \cdot c = a \cdot c + b \cdot c$ for all a, b, and c in S.

The two distributive laws are the **left** and **right** distributive laws, respectively. If property (viii) holds in S, so that multiplication is commutative, only one of the distributive laws is necessary, as the other follows from commutativity of multiplication.

The set $\mathbb{N}$ of natural numbers satisfies all of these properties except (iv), (v), and (x). The set $\mathbb{W} = \mathbb{N} \cup \{0\}$ does have an additive identity, since the number 0 satisfies property (iv). However, $\mathbb{W}$ does not contain an additive inverse for any of its elements except 0, hence property (v) is not satisfied by $\mathbb{W}$.

The set $\mathbb{Z}$ of integers contains the natural numbers as well as 0 and the negative integers. The integers satisfy all properties satisfied by the natural numbers. The number 0 is the additive identity of $\mathbb{Z}$ since $a + 0 = 0 + a = a$ for every integer a. If a is any integer and $-a$ is its negative, then $a + (-a) = (-a) + a = 0$, so $-a$ is the additive inverse of a. Thus $\mathbb{Z}$ satisfies properties (iv) and (v). The number 1 is the multiplicative identity of $\mathbb{Z}$ since $1 \cdot a = a \cdot 1 = a$ for every integer a. Unless $a = 1$ or $a = -1$, however, there is no *integer* b such that $a \cdot b = b \cdot a = 1$, and therefore property (x) is not satisfied by $\mathbb{Z}$.

The set $\mathbb{Q}$ of rational numbers contains the integers as well as all quotients of integers, or fractions. Using Definition 1.1.1 and assuming the known properties of integers, it can be shown that $\mathbb{Q}$ satisfies properties (i)–(ix) and (xi) of Definition 1.3.1 (see examples below). If $\frac{a}{b}$ is a non-zero rational number, then $a \neq 0$ so the reciprocal $\frac{b}{a}$ is also a rational number, and $\frac{a}{b} \cdot \frac{b}{a} = \frac{b}{a} \cdot \frac{a}{b} = 1$. Hence $\frac{b}{a}$ is the multiplicative inverse of $\frac{a}{b}$, and $\mathbb{Q}$ also satisfies property (x).

Examples:

1. Assuming the known properties of $\mathbb{Z}$, show that multiplication in $\mathbb{Q}$ is commutative.

   Proof. Let $\frac{a}{b}$ and $\frac{c}{d}$ be any rational numbers, so $a, b, c, d \in \mathbb{Z}$. We then use properties of $\mathbb{Z}$ in Definition 1.3.1 to show that $\frac{a}{b} \cdot \frac{c}{d} = \frac{c}{d} \cdot \frac{a}{b}$. We have

   $$\begin{aligned} \frac{a}{b} \cdot \frac{c}{d} &= \frac{ac}{bd} \text{ by Definition 1.1.1 (iii),} \\ &= \frac{ca}{db} \text{ by 1.3.1 (viii), commutativity of multiplication in } \mathbb{Z}, \\ &= \frac{c}{d} \cdot \frac{a}{b} \text{ by Definition 1.1.1 (iii),} \end{aligned}$$

   and so multiplication in $\mathbb{Q}$ is commutative. □

2. Assuming the known properties of $\mathbb{Z}$, show that addition in $\mathbb{Q}$ is associative.

   Proof. Let $\frac{a}{b}$, $\frac{c}{d}$, and $\frac{e}{f}$ be any rational numbers, so $a, b, c, d, e, f \in \mathbb{Z}$. By definition of addition in $\mathbb{Q}$ (Definition 1.1.1 (ii)), we have

   $$\left(\frac{a}{b} + \frac{c}{d}\right) + \frac{e}{f} = \frac{ad + bc}{bd} + \frac{e}{f} = \frac{(ad + bc)f + (bd)e}{(bd)f}$$

   and

   $$\frac{a}{b} + \left(\frac{c}{d} + \frac{e}{f}\right) = \frac{a}{b} + \frac{cf + de}{df} = \frac{a(df) + b(cf + de)}{b(df)}.$$

   We use properties of $\mathbb{Z}$ in Definition 1.3.1 to show that these expressions are equal. We have

   $$\begin{aligned} \left(\frac{a}{b} + \frac{c}{d}\right) + \frac{e}{f} &= \frac{(ad + bc)f + (bd)e}{(bd)f} \text{ as above,} \\ &= \frac{[(ad)f + (bc)f] + b(de)}{b(df)} \text{ by 1.3.1 (xi), (vii) in } \mathbb{Z}, \\ &= \frac{a(df) + [b(cf) + b(de)]}{b(df)} \text{ by 1.3.1 (ii), (vii) in } \mathbb{Z}, \\ &= \frac{a(df) + b(cf + de)}{b(df)} \text{ by 1.3.1 (xi) in } \mathbb{Z}, \\ &= \frac{a}{b} + \left(\frac{c}{d} + \frac{e}{f}\right) \text{ as above.} \end{aligned}$$

   Therefore $(\frac{a}{b} + \frac{c}{d}) + \frac{e}{f} = \frac{a}{b} + (\frac{c}{d} + \frac{e}{f})$ and addition in $\mathbb{Q}$ is associative. □

It can be shown that the set $\mathbb{R}$ of real numbers also satisfies properties (i)–(xi) of Definition 1.3.1 with the usual addition and multiplication. Using the definitions of addition and multiplication in $\mathbb{C}$ from Definition 1.2.1 and the known properties of the real numbers, it can be shown that $\mathbb{C}$ satisfies properties (i)–(ix) and (xi) of Definition 1.3.1 (see examples below). Proposition 1.2.7 says that $\mathbb{C}$ also satisfies property (x) of Definition 1.3.1. Hence $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ satisfy all of the properties in Definition 1.3.1.

Example: Assuming the known properties of $\mathbb{R}$, show that multiplication in $\mathbb{C}$ is commutative.

Proof. Let $a+bi$ and $c+di$ be any complex numbers, so $a, b, c, d \in \mathbb{R}$. We use properties of $\mathbb{R}$ in Definition 1.3.1 to show that $(a+bi)\cdot(c+di) = (c+di)\cdot(a+bi)$. We have

$$\begin{aligned}
(a+bi)\cdot(c+di) &= (ac-bd)+(ad+bc)i \text{ by Definition 1.2.1 (iii)},\\
&= (ca-db)+(da+cb)i \text{ by 1.3.1 (viii) in } \mathbb{R},\\
&= (ca-db)+(cb+da)i \text{ by 1.3.1 (iii) in } \mathbb{R},\\
&= (c+di)\cdot(a+bi) \text{ by Definition 1.2.1 (iii)},
\end{aligned}$$

and so multiplication in $\mathbb{Q}$ is commutative. □

Definition 1.3.2 *Let S be a set on which addition $(+)$ and multiplication $(\cdot)$ are defined.*

i. *If properties* (i)–(vii) *and* (xi) *of Definition 1.3.1 are satisfied in S, we say that S is a* **ring**.

ii. *If S is a ring and* (viii) *is also satisfied, we say that S is a* **commutative ring**.

iii. *If S is a ring and* (ix) *is also satisfied, we say that S is a* **ring with identity** *or a* **ring with 1**.

iv. *A set S satisfying all of the properties* (i)–(xi) *of Definition 1.3.1 is called a* **field**.

Thus $\mathbb{N}$ is NOT a ring because properties (iv) and (v) are not satisfied. By the discussion above, we have the following result.

Theorem 1.3.3 *The set of integers $\mathbb{Z}$ is a commutative ring with 1, and $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are fields.*

Exercise: Consider other sets you have studied in mathematics on which addition and multiplication are defined (for example, various sets of functions, polynomials, matrices, etc.). Are any of these sets rings or fields?

We are already familiar with the rings $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ and recognize that 0, $-a$, 1, satisfy properties (iv), (v), and (ix), respectively, and in $\mathbb{Q}$ or $\mathbb{R}$, $a^{-1} = 1/a$ satisfies property (x). In a general ring S, these elements are *defined* by the corresponding properties.

The additive identity element "0" is any element of S that satisfies property (iv), whether it looks like something we would call zero or not, and the additive inverse "$-a$" is just the element we add to a in order to get 0. Thus 0 and $-a$ are defined in terms of addition. In most rings, including $\mathbb{C}$, there is no concept of a "positive" or "negative" element.

We also define **subtraction** in a ring S in terms of addition. If a and b are elements of S, we define $a-b$ to be the element $a+(-b)$ of S, where of course $-b$ is the additive inverse of b.

Similarly, the multiplicative identity "1" and multiplicative inverse "a^{-1}" are defined in terms of multiplication by properties (ix) and (x), respectively. If S satisfies property (viii), so that multiplication in S is commutative, and S satisfies properties (ix) and (x), we can define **division** in terms of multiplication as well. For elements a and b of S with $b \neq 0$, we define $a \div b$ or a/b to be the element $a \cdot b^{-1}$ of S, where b^{-1} is the multiplicative inverse of b.

In the rational numbers $\mathbb{Q}$, for example, if $q = \frac{a}{b}$ and $r = \frac{c}{d} \neq 0$, then we have $r^{-1} = \frac{d}{c}$, the reciprocal of r, and $q \div r = q \cdot r^{-1}$, hence

$$\frac{a}{b} \div \frac{c}{d} = \frac{a}{b} \cdot \left(\frac{c}{d}\right)^{-1} = \frac{a}{b} \cdot \frac{d}{c}.$$

This explains the rule for dividing fractions we all learned in school.

Division is not usually defined in a ring that is not commutative. Since the elements $b^{-1} \cdot a$ and $a \cdot b^{-1}$ may not be equal, the expression $a \div b$ would be ambiguous.

Note that any results we can derive from properties (i)–(vii) and (xi) of Definition 1.3.1 will hold in *any* ring. This is the main advantage of such "abstraction" in algebra. Any results derived from a common list of properties will hold in any system satisfying those properties, and it is not necessary to reprove the same results repeatedly for different systems. This is essentially the same philosophy behind letting x stand for any number, one of the earliest cases of abstraction we encounter in learning algebra.

For example, the following basic results are familiar properties of the real numbers, but can be proved in a much more general context.

Proposition 1.3.4 *If S is a ring, then the following hold.*

i. *The additive identity element 0 of S is unique.*

ii. *If $a \in S$, then the additive inverse of a is unique.*

iii. *If $a \in S$, then $-(-a) = a$.*

Proof. (i) Suppose 0_a and 0_b are both identity elements for S, that is, both satisfy property (iv) of Definition 1.3.1. We have $0_a + 0_b = 0_b$ since 0_a is an additive identity, and $0_a + 0_b = 0_a$ since 0_b is an additive identity. Hence $0_a = 0_a + 0_b = 0_b$, and so $0_a = 0_b$. Thus there is only one additive identity element.

(ii) Let a be an element of S and suppose both b and b' are additive inverses for a; that is, both satisfy property (v) of Definition 1.3.1. Thus we have $b + a = 0$ and $a + b' = 0$, and so

$$\begin{aligned} b &= b + 0 \text{ by Definition 1.3.1 (iv),} \\ &= b + (a + b') \text{ by Definition 1.3.1 (v),} \\ &= (b + a) + b' \text{ by Definition 1.3.1 (ii),} \\ &= 0 + b' \text{ by Definition 1.3.1 (v),} \\ &= b' \text{ by Definition 1.3.1 (iv).} \end{aligned}$$

Hence $b = b'$ and so there is only one additive inverse for a.

(iii) Let a be an element of S and $-a$ the additive inverse of a. By part (ii) of this proposition and Definition 1.3.1 (v), $-(-a)$ is the unique element b of S satisfying $b+(-a) = (-a)+b = 0$. But also by Definition 1.3.1 (v), we have that $a + (-a) = (-a) + a = 0$ and by uniqueness of additive inverses, it follows that $-(-a) = a$. □

Proposition 1.3.5 *If S is a ring with 1, then the following hold.*

i. *The multiplicative identity element 1 of S is unique.*

ii. *If a is a non-zero element of S, then the multiplicative inverse of a is unique.*

iii. *If a is a non-zero element of S, then $(a^{-1})^{-1} = a$.*

Proof. The proof is similar to the proof of Proposition 1.3.4 above and is left as an exercise. □

The next results are restatements of properties of $\mathbb{Z}$ mentioned in the class preparation problems. They can be proved for more general rings, and then the same properties hold in $\mathbb{Z}$ as a special case.

Proposition 1.3.6 *If S is any ring, then $0 \cdot s = 0$ for all $s \in S$.*

Proof. The proof is left as an exercise. □

Proposition 1.3.7 *If S is any ring with 1, then the following hold:*

i. *$(-1) \cdot s = -s$ for all $s \in S$,*

ii. *$(-1) \cdot (-1) = 1$.*

Proof. (i) Let s be an element of S. By Proposition 1.3.4, $-s$ is the unique element $b \in S$ satisfying $s + b = b + s = 0$. It therefore suffices to prove that $s + (-1) \cdot s = (-1) \cdot s + s = 0$. Since addition in S is commutative (see Definition 1.3.2), we only need to show that $s + (-1) \cdot s = 0$. We have

$$\begin{aligned} s + (-1) \cdot s &= 1 \cdot s + (-1) \cdot s \text{ by Definition 1.3.1 (ix)}, \\ &= (1 + (-1)) \cdot s \text{ by Definition 1.3.1 (xi)}, \\ &= 0 \cdot s \text{ by Definition 1.3.1 (v)}, \\ &= 0 \text{ by Proposition 1.3.6}, \end{aligned}$$

Hence $s + (-1) \cdot s = 0$ and so $(-1) \cdot s = -s$.

(ii) By part (i) of this proposition, $(-1) \cdot (-1) = -(-1)$. By Proposition 1.3.4 (iii), $-(-1) = 1$. Therefore $(-1) \cdot (-1) = 1$, as claimed. □

Note that these propositions involve multiplication by 0 or an additive inverse. Both 0 and additive inverses are defined completely in terms of addition, so the proof of any result concerning multiplication by these elements must use the distributive laws, the only property relating addition and multiplication.

§1.3 Exercises

1. State which properties of addition and/or multiplication of natural numbers are used in the following equations. (More than one may be necessary.) All variables stand for natural numbers.

   (a) $3x + (4y + z) = 3x + (z + 4y)$

   (b) $5(x + y) = 5x + 5y$

   (c) $3(xy) = (3x)y$

   (d) $2x + (2y + z) = 2(x + y) + z$

   (e) $5(xy) = (5y)x$

2. Using Definition 1.1.1 and the fact that properties (i)–(ix) and (xi) of Definition 1.3.1 hold in the ring $\mathbb{Z}$ of integers, prove that addition in $\mathbb{Q}$ is commutative; that is,

$$\frac{a}{b} + \frac{c}{d} = \frac{c}{d} + \frac{a}{b}.$$

3. Using Definition 1.1.1 and the fact that properties (i)–(ix) and (xi) of Definition 1.3.1 hold in the ring $\mathbb{Z}$ of integers, prove that multiplication in $\mathbb{Q}$ is associative; that is,

$$\frac{a}{b} \cdot \left(\frac{c}{d} \cdot \frac{e}{f}\right) = \left(\frac{a}{b} \cdot \frac{c}{d}\right) \cdot \frac{e}{f}.$$

4. Using Definition 1.1.1 and the fact that properties (i)–(ix) and (xi) of Definition 1.3.1 hold in the ring $\mathbb{Z}$ of integers, prove that

$$\frac{a}{c} + \frac{b}{c} = \frac{a + b}{c}.$$

5. Using Definition 1.1.1 and the fact that properties (i)–(ix) and (xi) of Definition 1.3.1 hold in the ring $\mathbb{Z}$ of integers, prove that the distributive law holds in $\mathbb{Q}$; that is,

$$\frac{a}{b} \cdot \left(\frac{c}{d} + \frac{e}{f}\right) = \frac{a}{b} \cdot \frac{c}{d} + \frac{a}{b} \cdot \frac{e}{f}.$$

6. Using Definition 1.2.1 and the fact that properties (i)–(xi) of Definition 1.3.1 hold in the field $\mathbb{R}$ of real numbers, prove

   (a) addition in $\mathbb{C}$ is commutative; that is, $(a + bi) + (c + di) = (c + di) + (a + bi)$, and

   (b) addition in $\mathbb{C}$ is associative; that is,

$$(a + bi) + ((c + di) + (e + fi)) = ((a + bi) + (c + di)) + (e + fi).$$

7. Using Definition 1.2.1 and the fact that properties (i)–(xi) of Definition 1.3.1 hold in the field $\mathbb{R}$ of real numbers, prove that the distributive law holds in $\mathbb{C}$; that is,

$$(a + bi) \cdot ((c + di) + (e + fi)) = (a + bi) \cdot (c + di) + (a + bi) \cdot (e + fi).$$

8. Which properties of Definition 1.3.1 are satisfied by the set $S = 2\mathbb{Z}$ of *even* integers? Is $2\mathbb{Z}$ a ring? If so, is it a ring with 1? Justify your answers.

9. The subset $\mathbb{Z}[i] = \{a + bi \mid a, b \text{ are integers}\}$ of $\mathbb{C}$ is called the ring of **Gaussian integers**. Show that $\mathbb{Z}[i]$ is a commutative ring with 1. Is $\mathbb{Z}[i]$ a field? Explain.

10. Let F be the set of all functions f from $\mathbb{R}$ to $\mathbb{R}$, and define addition and multiplication "pointwise" as in algebra or calculus. That is, for functions f and g, define $f + g$ and $f \cdot g$ by $(f + g)(x) = f(x) + g(x)$ and $(f \cdot g)(x) = f(x) \cdot g(x)$ for all real numbers x.

    Which properties of Definition 1.3.1 are satisfied by F? Which properties are satisfied by the subset C of F consisting of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$? Justify your answers.

11. Prove Proposition 1.3.5.
    [Hint: Imitate the proof of Proposition 1.3.4.]

12. Prove Proposition 1.3.6.
    [Hint: Start with the fact that $0 + 0 = 0$ (why?) and use the distributive law.]

1.4 Sets and Equivalence Relations

In order to describe more rigorous constructions of our number systems, we need to introduce the concept of an equivalence relation on a set. We will first recall some terminology and notation from set theory. Our approach to set theory will be somewhat informal.

Definition 1.4.1 *A* **set** *A is a collection of objects. The objects in A are called the* **elements** *of A. If a is an element of A, we write $a \in A$ and if b is not an element of A we write $b \notin A$.*

A set of particular importance is the set containing nothing.

Definition 1.4.2 *The* **empty set** *is the (unique) set with no elements, denoted $\varnothing$.*

Definition 1.4.3 *Let A and B be sets.*

i. *We say A and B are* **equal***, denoted $A = B$, if the elements of A and B are the same.*

ii. *If each element of A is also an element of B, we say A is a* **subset** *of B and write $A \subseteq B$.*

iii. *If $A \subseteq B$ but $A \neq B$, we say A is a* **proper subset** *of B, denoted $A \subsetneq B$.*

Thus for sets A and B, $A \subseteq B$ means that if $a \in A$ then $a \in B$, and $A \subsetneq B$ means that $a \in B$ for every $a \in A$ and there is some $b \in B$ such that $b \notin A$. If $A \subseteq B$, we also say A is **contained in** B, and if $A \subsetneq B$, we say A is **properly contained in** B. Note that $\varnothing \subseteq A$ for any set A.

Be careful to distinguish between the symbols $\in$ and $\subseteq$. The symbol $\in$ is used only with elements and $\subseteq$ is used only with sets. For example, if $A = \{1, 2, 3\}$, then 1 is an *element* of A, but $\{1\}$ is the *subset* of A consisting of a single element 1. We write $1 \in A$ and $\{1\} \subseteq A$, but both of $1 \subseteq A$ and $\{1\} \in A$ are incorrect.

The following result provides the most common method for showing that two sets are equal.

Proposition 1.4.4 *Two sets A and B are equal if and only if $A \subseteq B$ and $B \subseteq A$.*

The proposition says that we can show $A = B$ by showing that if $a \in A$ then $a \in B$ and if $b \in B$ then $b \in A$.

We next consider some methods for constructing new sets from old.

Definition 1.4.5 *Let A, B, $A_1, A_2, \ldots, A_n$ be sets.*

i. *The* **union** *of A and B is the set $A \cup B = \{x \mid x \in A \text{ or } x \in B\}$. More generally, the union of $A_1, A_2, \ldots, A_n$ is the set*

$$A_1 \cup A_2 \cup \cdots \cup A_n = \bigcup_{i=1}^{n} A_i = \{x \mid x \in A_i \text{ for some } i\}.$$

ii. *The* **intersection** *of A and B is the set $A \cap B = \{x \mid x \in A \text{ and } x \in B\}$. More generally, the intersection of $A_1, A_2, \ldots, A_n$ is the set*

$$A_1 \cap A_2 \cap \cdots \cap A_n = \bigcap_{i=1}^{n} A_i = \{x \mid x \in A_i \text{ for all } i\}.$$

The word "or" is used in the inclusive sense. That is, $x \in A$ or $x \in B$ means that x is an element of *at least* one of the sets A, B, and could be an element of both.

Definition 1.4.6 *Two sets A, B are* **disjoint** *if $A \cap B = \varnothing$. A set A is the* **disjoint union** *of sets $A_1, A_2, \ldots, A_n$ if $A = A_1 \cup A_2 \cup \cdots \cup A_n$ and $A_i \cap A_j = \varnothing$ for all $i \neq j$. In this case, we also say the sets $A_1, A_2, \ldots, A_n$ form a* **partition** *of A.*

Definition 1.4.7 *Let A and B be sets. The* **Cartesian product** *of A and B is the set*

$$A \times B = \{(a, b) \mid a \in A,\, b \in B\}$$

consisting of all ordered pairs (a, b), where $a \in A$ and $b \in B$.

The elements of $A \times B$ are *ordered* pairs, and two elements (a_1, b_1) and (a_2, b_2) are equal if and only if $a_1 = a_2$ and $b_1 = b_2$. Thus if $A = \{1, 2\} = B$ for example, then $(1, 2)$ and $(2, 1)$ are different elements of $A \times B$.

We are now prepared to define equivalence relations.

Definition 1.4.8 *A* **relation** *R on a set A is a subset of $A \times A$.*

Although it is best to formally define a relation as a set, we often avoid the formality by thinking of elements of A being "related" to each other. In practice, we usually write aRb and say a is **related** to b if (a, b) is an element of the relation R.

Care must be taken, however, not to read too much into the word "related." There is nothing in the definition, for example, that says that if a is related to b, then b must be related to a. It is possible that (a, b) is an element of R but that (b, a) is not.

We will be primarily interested in a particular type of relation called an equivalence relation.

Definition 1.4.9 *An* **equivalence relation** *on a set A is a relation $\sim$ satisfying, for all a, b, and c in A:*

i. $a \sim a$ **(reflexive property)**.

ii. *If $a \sim b$ then $b \sim a$* **(symmetric property)**.

iii. *If $a \sim b$ and $b \sim c$, then $a \sim c$* **(transitive property)**.

If $\sim$ is an equivalence relation on A and $a \sim b$, we say a and b are **equivalent**. Equivalence is a generalization of equality. If a and b are equivalent, they need not be equal, but they are "the same" relative to some property. Note that equality is an equivalence relation.

Example: We can define a relation $\sim$ on the set $\mathbb{R}$ of real numbers by $a \sim b$ if $|a| = |b|$. It is easily verified that $\sim$ is an equivalence relation on $\mathbb{R}$ (do it!). Under this relation, 3 and -3 are equivalent although they are not equal. They are "the same" in that they are the same distance from the origin. □

It is often useful to consider all elements of a set that are equivalent to a given element under an equivalence relation to be the same. We give the set of such elements a name.

Definition 1.4.10 *If $\sim$ is an equivalence relation on a set A and a is an element of A, the* **equivalence class** *of a is the set $[a] = \{b \in A \mid a \sim b\}$.*

The equivalence class of a is just the set of all elements of A that are equivalent to a. In the example above, $[a] = \{a, -a\}$. Thus $[3] = \{3, -3\} = [-3]$. The fact that $3 \sim -3$ and $[3] = [-3]$ is not a coincidence.

Proposition 1.4.11 *Let $\sim$ be an equivalence relation on a set A. For a and b in A, $[a] = [b]$ if and only if $a \sim b$.*

The proposition says that if two elements of A are equivalent, then their equivalence classes are the same. By the defining properties of an equivalence relation, all elements in a given equivalence class must be equivalent (verify!).

Notice that every element a of A is in some equivalence class, namely $[a]$. Hence A is the union of the equivalence classes. The next result says that this is a disjoint union, so the equivalence classes partition the set A into disjoint subsets.

Proposition 1.4.12 *Let $\sim$ be an equivalence relation on a set A. If a and b are elements of A, then either $[a] = [b]$ or $[a] \cap [b] = \varnothing$.*

---

§1.4 Exercises

In the following exercises, A, B, and C are sets.

1. Show that if $A \subseteq C$ and $B \subseteq C$, then $A \cup B \subseteq C$.
2. Show that if $C \subseteq A$ and $C \subseteq B$, then $C \subseteq A \cap B$.
3. Show that if $A \subseteq B$, then $A \cup C \subseteq B \cup C$.
4. Show that if $A \subseteq B$, then $A \cap C \subseteq B \cap C$.
5. Show that $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
6. Show that $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
7. Verify that the relation $\sim$ on $\mathbb{R}$, defined by $a \sim b$ if and only if $|a| = |b|$, is an equivalence relation.
8. Verify that the relation $\sim$ on $\mathbb{R}$, defined by $a \sim b$ if and only if $a - b$ is rational, is an equivalence relation.
9. Verify that the relation $\sim$ on the set of fractions $\{\frac{a}{b} \mid a, b \in \mathbb{Z},\ b \neq 0\}$, defined by $\frac{a}{b} \sim \frac{c}{d}$ if and only if $ad = bc$, is an equivalence relation.
10. Show that if $\sim$ is an equivalence relation on a set A and $a \sim b$, then $[a] = [b]$.
11. Show that if $\sim$ is an equivalence relation on a set A and a and b are elements of A such that $[a] \cap [b]$ is not empty, then $a \sim b$. (This proves Proposition 1.4.12.)

1.5 Formal Constructions of Number Systems

Today, most of us are comfortable with the number systems we have discussed thus far, with the possible exception of the complex numbers. This has not always been the case, however. Ideas such as 0, negative numbers, irrational numbers, and complex numbers have historically been much more difficult to grasp than the natural numbers. In order to believe that these objects exist, it is useful to have concrete constructions that behave in the expected ways.

In this section we present a more formal and rigorous construction of some of our number systems. We will begin by giving a construction of the field $\mathbb{Q}$ of rational numbers assuming we have the ring $\mathbb{Z}$ of integers. We will then construct the integers $\mathbb{Z}$ from the natural numbers $\mathbb{N}$ and then the complex numbers $\mathbb{C}$ from the real numbers $\mathbb{R}$.

A model for the natural numbers can be constructed using a set of axioms (the Peano axioms), but this construction is much more abstract and will not be included here. We also will not include a construction of the real numbers because such a construction would require tools from analysis that are beyond the scope of this text.

Construction of the Rational Numbers

Although we should logically construct the integers first and then construct the rational numbers from them, the ideas behind the construction of the rational numbers are more familiar to us. The differences between this construction and the usual understanding of the rational numbers as fractions mainly concern notation and the formalization of the idea of equivalent fractions. We will therefore first consider the construction of the rational numbers. We will assume all properties previously discussed for the ring of integers. The rational numbers will be constructed as a set of equivalence classes of ordered pairs of integers.

The construction given here is a special case of the construction of the "field of fractions of an integral domain" studied in ring theory. It will also allow us to obtain other fields starting with rings similar to the integers, such as rings of polynomials.

Definition 1.5.1 *Let $\mathcal{Q} = \{(a, b) \mid a, b \in \mathbb{Z},\ b \neq 0\}$ and define a relation $\sim$ on $\mathcal{Q}$ by $(a, b) \sim (a', b')$ if and only if $ab' = ba'$.*

Proposition 1.5.2 *The relation $\sim$ defined on $\mathcal{Q}$ is an equivalence relation.*

Definition 1.5.3 *For (a, b) in $\mathcal{Q}$, let $[(a, b)]$ be the equivalence class of (a, b) under the relation $\sim$. Define $\boldsymbol{Q}$ to be the set of equivalence classes of elements of $\mathcal{Q}$, that is,*

$$\boldsymbol{Q} = \{\, [(a, b)] \mid (a, b) \in \mathcal{Q} \,\}.$$

The best way to understand this construction and what follows is to think of the element (a, b) of $\mathcal{Q}$ as the specific fraction $\frac{a}{b}$. For (a, b) and (a', b') in $\mathcal{Q}$, we have $(a, b) \sim (a', b')$ if and only if $ab' = ba'$, hence if and only if $\frac{a}{b} = \frac{a'}{b'}$. The equivalence relation $\sim$ can therefore be viewed as the usual equivalence of fractions.

The element $[(a, b)]$ of $\boldsymbol{Q}$ then can be viewed as the rational number represented by $\frac{a}{b}$, or by any other equivalent fraction. We can consider $\mathbb{Z}$ to be a subset of $\boldsymbol{Q}$ by identifying an integer z with the element $[(z, 1)]$ of $\boldsymbol{Q}$.

The issue of equality of fractions is more formally addressed in this construction by taking the elements of $\boldsymbol{Q}$ to be equivalence classes of elements of $\mathcal{Q}$, or of fractions. Thus the rational number $0.666\ldots = 0.\overline{6}$ is represented by the single equivalence class $[(2,3)]$ instead of by the various equivalent fractions $\frac{2}{3}$, $\frac{4}{6}$, $\frac{6}{9}$, ...

Keep in mind that the elements of $\boldsymbol{Q}$ are equivalence classes, hence are sets. For example,

$$[(2,3)] = \{(2,3), (-2,-3), (4,6), (-4,-6), (6,9), \ldots\}.$$

For (a,b) in $\mathcal{Q}$, $[(a,b)] = [(ac,bc)]$ for every non-zero integer c, just as $\frac{a}{b} = \frac{ac}{bc}$.

Next, we must define addition and multiplication on $\boldsymbol{Q}$ so that $\boldsymbol{Q}$ behaves like the rational numbers $\mathbb{Q}$. Compare the following definition to Definition 1.1.1.

Definition 1.5.4 *For $\boldsymbol{x} = [(a,b)]$ and $\boldsymbol{y} = [(c,d)]$ in $\boldsymbol{Q}$, we define addition $\boldsymbol{x}+\boldsymbol{y}$ and multiplication $\boldsymbol{x} \cdot \boldsymbol{y}$ as follows:*

i. $\boldsymbol{x} + \boldsymbol{y} = [(a,b)] + [(c,d)] = [(ad+bc, bd)]$

ii. $\boldsymbol{x} \cdot \boldsymbol{y} = [(a,b)] \cdot [(c,d)] = [(ac,bd)]$.

Note that the addition and multiplication of $\boldsymbol{x}$ and $\boldsymbol{y}$ appear to depend on the particular representatives (a,b) and (c,d) chosen for the equivalence classes. In order for these definitions to be useful, it is necessary to know that this is not the case.

Proposition 1.5.5 *If $[(a,b)] = [(a',b')]$ and $[(c,d)] = [(c',d')]$ in $\boldsymbol{Q}$, then*

$$[(a,b)] + [(c,d)] = [(a',b')] + [(c',d')]$$

and

$$[(a,b)] \cdot [(c,d)] = [(a',b')] \cdot [(c',d')].$$

That is, addition and multiplication in $\boldsymbol{Q}$ are **well-defined**.

By the definitions of addition and multiplication in $\boldsymbol{Q}$ and the closure of $\mathbb{Z}$ under addition and multiplication, it is clear that $\boldsymbol{Q}$ is closed under both operations. The next results can be verified by direct calculations.

Proposition 1.5.6 *The following hold in $\boldsymbol{Q}$.*

i. *The element $\mathbf{0} = [(0,1)]$ is the additive identity element of $\boldsymbol{Q}$.*

ii. *If $\boldsymbol{x} = [(a,b)]$ is an element of $\boldsymbol{Q}$, then the additive inverse of $\boldsymbol{x}$ is $-\boldsymbol{x} = [(-a,b)]$.*

We have that $[(a,b)] = \mathbf{0}$ if and only if $(a,b) \sim (0,1)$, hence if and only if $a \cdot 1 = b \cdot 0$, that is, $a = 0$. Hence if $[(a,b)] \neq \mathbf{0}$, then $a \neq 0$ and $[(b,a)]$ is also an element of $\boldsymbol{Q}$. This is used in the next result.

Proposition 1.5.7 *The following hold in* $\boldsymbol{Q}$*.*

i. *The element* $\mathbf{1} = [(1,1)]$ *is the multiplicative identity element of* $\boldsymbol{Q}$*.*

ii. *If* $\boldsymbol{x} = [(a,b)] \neq \mathbf{0}$ *is an element of* $\boldsymbol{Q}$*, then the multiplicative inverse of* $\boldsymbol{x}$ *is* $\boldsymbol{x}^{-1} = [(b,a)]$*.*

Using the definitions of $\boldsymbol{Q}$, addition, and multiplication, and the known properties of $\mathbb{Z}$, it can be shown that both addition and multiplication in $\boldsymbol{Q}$ are commutative and associative, and that the distributive law holds in $\boldsymbol{Q}$. The proofs are very similar to those in Exercises 1.3.2–1.3.5. Along with the results and remarks above, this yields the following theorem.

Theorem 1.5.8 *With* $\boldsymbol{Q}$*, addition, and multiplication defined as above,* $\boldsymbol{Q}$ *satisfies properties* (i)–(xi) *of Definition 1.3.1. That is,* $\boldsymbol{Q}$ *is a field.*

If we identify the element $[(a,b)]$ of $\boldsymbol{Q}$ with the rational number represented by the fraction $\frac{a}{b}$ (or any equivalent fraction), $\boldsymbol{Q}$ is "algebraically the same" as $\mathbb{Q}$. In abstract algebra, we would say that the two fields are **isomorphic**.

Construction of the Integers

We next consider the construction of the integers from the natural numbers. We will assume all properties previously discussed for the natural numbers $\mathbb{N}$. The integers will be constructed as a set of equivalence classes of ordered pairs of natural numbers.

Definition 1.5.9 *Let* $\mathcal{Z} = \{(a,b) \mid a,b \in \mathbb{N}\}$ *and define a relation* $\sim$ *on* $\mathcal{Z}$ *by* $(a,b) \sim (a',b')$ *if and only if* $a + b' = b + a'$*.*

Proposition 1.5.10 *The relation* $\sim$ *defined on* $\mathcal{Z}$ *is an equivalence relation.*

Definition 1.5.11 *For* (a,b) *in* $\mathcal{Z}$*, let* $[(a,b)]$ *be the equivalence class of* (a,b) *under the relation* $\sim$*. Define* $\boldsymbol{Z}$ *to be the set of equivalence classes of elements of* $\mathcal{Z}$*, that is,*

$$\boldsymbol{Z} = \{\, [(a,b)] \mid (a,b) \in \mathcal{Z} \,\}.$$

Note the similarity in the definitions of $\mathcal{Z}$, $\sim$, and $\boldsymbol{Z}$ to the definitions of $\mathcal{Q}$, $\sim$, and $\boldsymbol{Q}$. The only difference is that multiplication in the definition of $\sim$ for $\mathcal{Q}$ (i.e., $ab' = ba'$) is replaced by addition in the definition of $\sim$ for $\mathcal{Z}$ (i.e., $a + b' = b + a'$). This should not be a surprise. We constructed $\boldsymbol{Q}$ from $\mathbb{Z}$ in order to obtain multiplicative inverses and we are constructing $\boldsymbol{Z}$ from $\mathbb{N}$ in order to obtain additive inverses.

In order to better understand this construction, think of the element $[(a,b)]$ of $\boldsymbol{Z}$ as the integer $a - b$. For (a,b) and (a',b') in $\mathcal{Q}$, we have $(a,b) \sim (a',b')$ if and only if $a + b' = b + a'$, hence if and only if $a - b = a' - b'$. Hence an element of $\boldsymbol{Z}$ corresponds to exactly one integer.

We can consider $\mathbb{N}$ to be a subset of $\boldsymbol{Z}$ by identifying a natural number n with the element $[(n+1,1)]$ of $\boldsymbol{Z}$. We thereby identify $\mathbb{N}$ with the subset $\{\, [(a,b)] \mid a > b \,\}$ of $\boldsymbol{Z}$, identify the integer 0 with the element $[(1,1)]$ $(= [(n,n)]$ for all $n \in \mathbb{N})$, and identify the negative integers with the subset $\{\, [(a,b)] \mid a < b \,\}$ of $\boldsymbol{Z}$.

Again, keep in mind that the elements of $\boldsymbol{Z}$ are equivalence classes, hence are sets. For example,

$$[(3,5)] = \{\, [(n, n+2)] \mid n \in \mathbb{N} \,\} = \{(1,3),(2,4),(3,5),(4,6),\ldots\}.$$

For (a,b) in $\mathcal{Z}$, $[(a,b)] = [(a+c,b+c)]$ for every natural number c.

Next, we define addition and multiplication on $\boldsymbol{Z}$ so that $\boldsymbol{Z}$ behaves like the integers $\mathbb{Z}$.

Definition 1.5.12 *For $\boldsymbol{x} = [(a,b)]$ and $\boldsymbol{y} = [(c,d)]$ in $\boldsymbol{Z}$, we define addition $\boldsymbol{x}+\boldsymbol{y}$ and multiplication $\boldsymbol{x}\cdot\boldsymbol{y}$ as follows:*

i. $\boldsymbol{x}+\boldsymbol{y} = [(a,b)] + [(c,d)] = [(a+c, b+d)]$

ii. $\boldsymbol{x}\cdot\boldsymbol{y} = [(a,b)]\cdot[(c,d)] = [(ac+bd, ad+bc)]$.

Recalling our interpretation of an element $[(m,n)]$ of $\boldsymbol{Z}$ as the integer $m-n$, the definitions of addition and multiplication in $\boldsymbol{Z}$ are inspired by the facts that $(a-b)+(c-d) = (a+c)-(b+d)$ and $(a-b)(c-d) = (ac+bd)-(ad+bc)$ in $\mathbb{Z}$.

The addition and multiplication of $\boldsymbol{x}$ and $\boldsymbol{y}$ are defined in terms of the particular representatives (a,b) and (c,d) chosen for the equivalence classes. In order for these definitions to be useful, it is necessary to know that the definitions do not actually depend on these choices.

Proposition 1.5.13 *If $[(a,b)] = [(a',b')]$ and $[(c,d)] = [(c',d')]$ in $\boldsymbol{Z}$, then*

$$[(a,b)] + [(c,d)] = [(a',b')] + [(c',d')]$$

and

$$[(a,b)]\cdot[(c,d)] = [(a',b')]\cdot[(c',d')].$$

That is, addition and multiplication in $\boldsymbol{Z}$ are well-defined.

By the definitions of addition and multiplication in $\boldsymbol{Z}$ and the closure of $\mathbb{N}$ under addition and multiplication, it is clear that $\boldsymbol{Z}$ is closed under both operations. The next results can be verified by direct calculations.

Proposition 1.5.14 *The following hold in $\boldsymbol{Z}$.*

i. *The element $\mathbf{0} = [(1,1)]$ is the additive identity element of $\boldsymbol{Z}$.*

ii. *If $\boldsymbol{x} = [(a,b)]$ is an element of $\boldsymbol{Z}$, then the additive inverse of $\boldsymbol{x}$ is $-\boldsymbol{x} = [(b,a)]$.*

Proposition 1.5.15 *The element $\mathbf{1} = [(2,1)]$ is the multiplicative identity element of $\boldsymbol{Z}$.*

Most elements of $\boldsymbol{Z}$ do not have multiplicative inverses. For example, let $\boldsymbol{x} = [(3,1)]$ (so $\boldsymbol{x}$ corresponds to the natural number 2). If there were a multiplicative inverse for $\boldsymbol{x}$, then there would be natural numbers c and d such that $[(3,1)]\cdot[(c,d)] = \mathbf{1} = [(2,1)]$. By definition of multiplication in $\boldsymbol{Z}$ and properties of equivalence classes (Proposition 1.4.11), this implies $(3c+d, 3d+c) \sim (2,1)$. Note that $[(2,1)] = \{(n+1,n) \mid n \in \mathbb{N}\}$ (verify!), hence this implies $(3c+d, 3d+c) = (n+1, n)$ for some natural number n. Therefore $3c+d = n+1$ and $3d+c = n$. Substituting the second equation in the first, we obtain $3c+d = 3d+c+1$, or $2c = 2d+1$. But c and d are natural numbers, so $2c$

is even and $2d+1$ is odd, hence $2c$ cannot equal $2d+1$. Therefore no such natural numbers exist and $\boldsymbol{x} = [(3,1)]$ has no multiplicative inverse in $\boldsymbol{Z}$. In fact, it can be shown that $[(2,1)]$ and $[(1,2)]$ are the *only* elements of $\boldsymbol{Z}$ that have multiplicative inverses. (What are their inverses?)

Using the definitions of $\boldsymbol{Z}$, addition, and multiplication, and the known properties of $\mathbb{N}$, it can be shown that both addition and multiplication in $\boldsymbol{Z}$ are commutative and associative, and that the distributive law holds in $\boldsymbol{Z}$. Along with the results and remarks above, this yields the following theorem.

Theorem 1.5.16 *With $\boldsymbol{Z}$, addition, and multiplication defined as above, $\boldsymbol{Z}$ satisfies properties* (i)–(ix) *and* (xi) *of Definition 1.3.1. That is, $\boldsymbol{Z}$ is a commutative ring with* 1.

Identifying the element $[(a,b)]$ of $\boldsymbol{Z}$ with the integer $a-b$, we see that $\boldsymbol{Z}$ is algebraically the same as $\mathbb{Z}$. That is, the two rings are isomorphic.

Construction of the Complex Numbers

Our definition of $\mathbb{C}$ as the "set of symbols $a+bi$" (Definition 1.2.1) was somewhat unsatisfactory. We give here a more rigorous construction of the complex numbers as ordered pairs of real numbers. Admittedly, the difference is subtle, but this construction avoids the use of formal symbols as elements of our set. It will not be necessary to use equivalence classes as in the two previous constructions. We will assume all of the field properties of the real numbers $\mathbb{R}$.

Definition 1.5.17 *Let $\boldsymbol{C} = \mathbb{R} \times \mathbb{R}$ and for $\boldsymbol{x} = (a,b)$ and $\boldsymbol{y} = (c,d)$, define addition $\boldsymbol{x}+\boldsymbol{y}$ and multiplication $\boldsymbol{x}\cdot\boldsymbol{y}$ as follows:*

i. $\boldsymbol{x}+\boldsymbol{y} = (a,b)+(c,d) = (a+c, b+d)$

ii. $\boldsymbol{x}\cdot\boldsymbol{y} = (a,b)\cdot(c,d) = (ac-bd, ad+bc)$.

Compare this to Definition 1.2.1. We will view the element (a,b) of $\boldsymbol{C}$ as the complex number $a+bi$. A real number a corresponds to the element $(a,0)$ of $\boldsymbol{C}$, and $(0,1)$ corresponds to the number i, so $(0,1)^2 = (-1,0)$ corresponds to -1 as desired.

This algebraic construction of $\boldsymbol{C}$ corresponds precisely to our earlier geometric interpretation of $\mathbb{C}$ as the complex plane. In both cases, the complex number $a+bi$ is interpreted as the ordered pair (a,b) of real numbers. The definitions of addition and multiplication in the two interpretations are the same.

Definition 1.5.17 and the closure of $\mathbb{R}$ under addition and multiplication imply that $\boldsymbol{C}$ is closed under both operations. The following results give the additive and multiplicative identities and inverses, and are easy to verify.

Proposition 1.5.18 *The following hold in $\boldsymbol{C}$.*

i. *The element $\mathbf{0} = (0,0)$ is the additive identity element of $\boldsymbol{C}$.*

ii. *If $\boldsymbol{x} = (a,b)$ is an element of $\boldsymbol{C}$, then the additive inverse of $\boldsymbol{x}$ is $-\boldsymbol{x} = (-a,-b)$.*

We have that $(a,b) = \mathbf{0}$ if and only if $a = b = 0$. Hence if $(a,b) \neq \mathbf{0}$, then at least one of a or b is non-zero, and so a^2+b^2 is a non-zero real number. This is used in the next result.

Proposition 1.5.19 *The following hold in* $\boldsymbol{C}$*.*

i. *The element* $\mathbf{1} = (1, 0)$ *is the multiplicative identity element of* $\boldsymbol{C}$*.*

ii. *If* $\boldsymbol{x} = (a, b) \neq \mathbf{0}$ *is an element of* $\boldsymbol{C}$*, then the multiplicative inverse of* $\boldsymbol{x}$ *is*

$$\boldsymbol{x}^{-1} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right).$$

Using Definition 1.5.17, and the known properties of $\mathbb{R}$, it can be shown that both addition and multiplication in $\boldsymbol{C}$ are commutative and associative, and that the distributive law holds in $\boldsymbol{C}$. The proofs are very similar to those in Exercises 1.3.6–1.3.7. Along with the results and remarks above, this yields the following theorem.

Theorem 1.5.20 *With* $\boldsymbol{C}$*, addition, and multiplication defined as above,* $\boldsymbol{C}$ *satisfies properties* (i)–(xi) *of Definition 1.3.1. That is,* $\boldsymbol{C}$ *is a field.*

The fields $\boldsymbol{C}$ and $\mathbb{C}$ are seen to be isomorphic by identifying the element (a, b) of $\boldsymbol{C}$ with the complex number $a + bi$.

---

§1.5 Exercises

1. Show that the relation $\sim$ defined on $\mathcal{Q}$ by $(a, b) \sim (a', b')$ if and only if $ab' = ba'$ is an equivalence relation.

2. Prove Proposition 1.5.5, that addition and multiplication in $\boldsymbol{Q}$ are well-defined. This is equivalent to proving that if $(a, b) \sim (a', b')$ and $(c, d) \sim (c', d')$ in $\mathcal{Q}$, then

   (a) $(ad + bc, bd) \sim (a'd' + b'c', b'd')$ and

   (b) $(ac, bd) \sim (a'c', b'd')$.

3. Prove Proposition 1.5.6. That is, show that for all $[(a, b)] \in \boldsymbol{Q}$,

   (a) $[(0, 1)] + [(a, b)] = [(a, b)] + [(0, 1)] = [(a, b)]$ and

   (b) $[(a, b)] + [(-a, b)] = [(-a, b)] + [(a, b)] = [(0, 1)]$.

4. Prove Proposition 1.5.7. That is, show that

   (a) $[(1, 1)] \cdot [(a, b)] = [(a, b)] \cdot [(1, 1)] = [(a, b)]$ for all $[(a, b)] \in \boldsymbol{Q}$ and

   (b) $[(a, b)] \cdot [(b, a)] = [(b, a)] \cdot [(a, b)] = [(1, 1)]$ for all $[(a, b)] \in \boldsymbol{Q}$ with $[(a, b)] \neq \mathbf{0}$.

5. Show that the relation $\sim$ defined on $\mathcal{Z}$ by $(a, b) \sim (a', b')$ if and only if $a + b' = b + a'$ is an equivalence relation.

6. Prove Proposition 1.5.13, that addition and multiplication in $\boldsymbol{Z}$ are well-defined. This is equivalent to proving that if $(a, b) \sim (a', b')$ and $(c, d) \sim (c', d')$ in $\mathcal{Z}$, then

   (a) $(a + c, b + d) \sim (a' + c', b' + d')$ and

   (b) $(ac + bd, ad + bc) \sim (a'c' + b'd', a'd' + b'c')$.

7. Prove Proposition 1.5.14. That is, show that for all $[(a, b)] \in \boldsymbol{Z}$,

   (a) $[(1, 1)] + [(a, b)] = [(a, b)] + [(1, 1)] = [(a, b)]$ and

   (b) $[(a, b)] + [(b, a)] = [(b, a)] + [(a, b)] = [(1, 1)]$.

8. Prove Proposition 1.5.15. That is, show that $[(2, 1)] \cdot [(a, b)] = [(a, b)] \cdot [(2, 1)] = [(a, b)]$ for all $[(a, b)] \in \boldsymbol{Q}$.

9. Show that no element of $\boldsymbol{Z}$ other than $[(2, 1)]$ or $[(1, 2)]$ has a multiplicative inverse in $\boldsymbol{Z}$. Find the inverses of $[(2, 1)]$ and $[(1, 2)]$.

10. Let $\mathbf{0} = [(1, 1)]$ and $\boldsymbol{x} = [(a, b)]$ in $\boldsymbol{Z}$. Show that $\mathbf{0} \cdot \boldsymbol{x} = \mathbf{0}$. (This shows that zero times any "integer" is zero.)

11. Let $\mathbf{1} = [(2, 1)]$, $\boldsymbol{x} = [(a, b)]$, and $\boldsymbol{y} = [(c, d)]$ in $\boldsymbol{Z}$. Show that $(-\boldsymbol{x}) \cdot (-\boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y}$. Deduce that in particular, $(-\mathbf{1}) \cdot (-\mathbf{1}) = \mathbf{1}$ and $(-\mathbf{1}) \cdot \boldsymbol{x} = -\boldsymbol{x}$. (This shows that the product of two "negative integers" is a "positive integer.")

Chapter 2

Basic Number Theory

In this chapter, we introduce some basic ideas from Number Theory, the study of properties of the natural numbers and integers. We will concentrate on properties related to divisibility of integers, including greatest common divisors, prime numbers, and prime factorizations.

2.1 Principle of Mathematical Induction

Consider the following questions before reading further.

Class Preparation Problems:

1. What is "inductive" reasoning? Can it be used to prove mathematical statements?
2. Verify the following statements for as many cases as you can:
   (a) $2^{2^n} + 1$ is prime for every integer $n \geqslant 0$.
   (b) If n is any even integer greater than 4, then n is the sum of two odd primes.
3. Are the statements above true? How many cases would we need to check in order to *prove* they are true? How many cases to prove false?

A mathematical proof involves "deductive" reasoning. New statements are deduced logically from known definitions or theorems. If the hypotheses are true and the logical argument is sound, the conclusion must be true.

Inductive reasoning, on the other hand, attempts to make general conclusions based on specific observations. While this is a good method for constructing conjectures of what *might* be true, it cannot be used to *prove* mathematical statements. After making some specific observations and making a reasonable conjecture, it is still necessary to use a deductive argument based on known results in order to prove the conjecture. Many statements are true in a few specific cases but turn out not to be true in general. It is not possible to prove a general statement with examples.

The numbers $F_n = 2^{2^n} + 1$, for $n \geqslant 0$ an integer, are called **Fermat numbers**. In the 1600s, Fermat observed that the first five of these numbers,

$$F_0 = 3,\ F_1 = 5,\ F_2 = 17,\ F_3 = 257,\ F_4 = 65537,$$

are prime, and conjectured that F_n is prime for all $n \geqslant 0$, but was unable to prove this conjecture. The conjecture was not true, as

$$F_5 = 2^{2^5} + 1 = 4,294,967,297 = (641)(6,700,417)$$

is not prime. In fact, no other prime Fermat number has yet been found. The complete prime factorizations of the Fermat numbers F_n for $5 \leqslant n \leqslant 11$ are known and the numbers F_n for $5 \leqslant n \leqslant 32$ are known to be composite. Thus the observation that F_n is prime for $0 \leqslant n \leqslant 4$ is rather misleading.

One version of the **Goldbach Conjecture** states that every even integer greater than 4 is the sum of two odd primes. The conjecture has been verified for all even integers up to $4 \cdot 10^{14}$ (J. Richstein, *Math. Comput.* **70** (2001), 1745–1750), and there are as yet unpublished claims that it has been verified up to $3 \cdot 10^{17}$. Although the conjecture is believed to be true, it remains a conjecture because despite this considerable evidence, it has not been proved. It remains possible that there is some even integer greater than 4 that is *not* the sum of two odd primes.

The Principle of Mathematical Induction is a method for proving statements about natural numbers (or integers). It is not inductive reasoning as discussed above. It is based on the following property of natural numbers that we will assume.

Theorem 2.1.1 (Well-ordering Principle) *If S is a non-empty set of natural numbers, then S has a smallest element.*

Note that this is a special property of the natural numbers. It is not true, for example, for the set of positive real numbers or the set of positive rational numbers. It can be generalized to the set of integers: any non-empty set of integers that is bounded below has a smallest element.

Theorem 2.1.2 (Principle of Mathematical Induction) *Let S be a set of natural numbers. If*

i. *1 is in S, and*

ii. *whenever k is in S, $k+1$ is also in S,*

then S is the set of all natural numbers (that is, every natural number n is in S).

Proof. We prove the theorem by contradiction, using the Well-ordering Principle. That is, we assume the hypotheses of the theorem and suppose the conclusion is false. We then proceed to reach a contradiction.

Assume (i) and (ii) hold and suppose S is not the entire set $\mathbb{N}$ of natural numbers. The set T of natural numbers *not* in S is therefore a non-empty set of natural numbers. Hence T has a smallest element m by the Well-ordering Principle.

Now m is the smallest natural number *not* in S. Since 1 is in S by (i), it follows that $m > 1$, so $k = m - 1$ is a natural number. Since $k < m$, we have that k is in S. But then (ii) implies $k + 1 = m$ is in S, which is a contradiction. Hence our assumption that S is not all of $\mathbb{N}$ must be false, and the theorem is proved. □

We usually do not formally consider a set S of natural numbers when using mathematical induction. The following is an alternative version.

Theorem 2.1.3 (Principle of Mathematical Induction (alternative version)) *Let $P(n)$ be a statement about natural numbers. If*

i. *$P(1)$ is true, and*

ii. *whenever $P(k)$ is true, $P(k+1)$ is also true,*

then $P(n)$ is true for all $n \geqslant 1$.

Notes:

1. It is easy to see that the two versions of the Principle of Mathematical Induction are equivalent by letting S be the set of all natural numbers n for which $P(n)$ is true.

2. Statement (i) is called the **base step** of the induction. The assumption in (ii) that $k \in S$, or $P(k)$ is true, is called the **inductive hypothesis**, and the proof of (ii) is called the **inductive step**.

3. It is not necessary to start the induction at 1. In (i), if we replace $1 \in S$ with $n_0 \in S$, then the conclusion is that $n \in S$ for all $n \geqslant n_0$. (Or, if we replace "$P(1)$ is true" with "$P(n_0)$ is true," the conclusion is that $P(n)$ is true for all $n \geqslant n_0$.) In fact, S can be a set of integers or $P(n)$ a statement about integers, and n_0 can be any integer. That is, n_0 can even be 0 or negative.

Examples:

1. Use induction to show that $1^3 + 2^3 + 3^3 + \cdots + n^3 = \dfrac{n^2(n+1)^2}{4}$ for $n \geqslant 1$.

   Proof. (i) We first show that the statement is true for $n = 1$. When $n = 1$, the formula becomes $1^3 = \frac{1^2(1+1)^2}{4} = \frac{4}{4} = 1$, and so the statement is true.

   (ii) Next, we assume the statement is true for $n = k$; that is,

   $$1^3 + 2^3 + 3^3 + \cdots + k^3 = \frac{k^2(k+1)^2}{4} \quad (*)$$

   and show that this implies the statement is true for $n = k + 1$; that is,

   $$\begin{aligned} 1^3 + 2^3 + 3^3 + \cdots + k^3 + (k+1)^3 &= \frac{(k+1)^2((k+1)+1)^2}{4} \\ &= \frac{(k+1)^2(k+2)^2}{4}. \end{aligned}$$

We have

$$\begin{aligned} 1^3+2^3+3^3+\cdots+k^3+(k+1)^3 &= \frac{k^2(k+1)^2}{4}+(k+1)^3 \text{ by } (*), \\ &= \frac{k^2(k+1)^2+4(k+1)^3}{4} \\ &= \frac{(k+1)^2[k^2+4(k+1)]}{4} \\ &= \frac{(k+1)^2[k^2+4k+4]}{4} \\ &= \frac{(k+1)^2(k+2)^2}{4}. \end{aligned}$$

Hence if the statement is true for $n=k$, then the statement is true for $n=k+1$.

Since (i) and (ii) hold, the statement is true for all $n \geqslant 1$ by the Principle of Mathematical Induction. □

2. Use induction to show that $\dfrac{1}{1\cdot 2}+\dfrac{1}{2\cdot 3}+\dfrac{1}{3\cdot 4}+\cdots+\dfrac{1}{n(n+1)}=\dfrac{n}{n+1}$ for $n \geqslant 1$.

Proof. (i) We first show that the statement is true for $n=1$. When $n=1$, the formula becomes $\frac{1}{1\cdot 2}=\frac{1}{1+1}$, and so the statement is true.

(ii) Next, we assume the statement is true for $n=k$; that is,

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\frac{1}{3\cdot 4}+\cdots+\frac{1}{k(k+1)}=\frac{k}{k+1} \quad (*)$$

and show that this implies the statement is true for $n=k+1$; that is,

$$\begin{aligned} \frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\cdots+\frac{1}{k(k+1)}+\frac{1}{(k+1)(k+2)} &= \frac{k+1}{((k+1)+1)} \\ &= \frac{k+1}{k+2}. \end{aligned}$$

We have

$$\begin{aligned} \frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\cdots+\frac{1}{k(k+1)}+\frac{1}{(k+1)(k+2)} &= \frac{k}{k+1}+\frac{1}{(k+1)(k+2)} \text{ by } (*), \\ &= \frac{k(k+2)+1}{(k+1)(k+2)} \\ &= \frac{k^2+2k+1}{(k+1)(k+2)} \\ &= \frac{(k+1)^2}{(k+1)(k+2)} \\ &= \frac{k+1}{k+2}. \end{aligned}$$

Hence if the statement is true for $n=k$, then the statement is true for $n=k+1$.

Since (i) and (ii) hold, the statement is true for all $n \geqslant 1$ by the Principle of Mathematical Induction. □

3. Prove de Moivre's Theorem (Corollary 1.2.11 with $r = 1$); i.e., show that

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$$

for all $n \geqslant 1$.

Proof. (i) We first show that the statement is true for $n = 1$. When $n = 1$, the statement becomes $(\cos\theta + i\sin\theta)^1 = \cos 1\theta + i\sin 1\theta$ or equivalently, $\cos\theta + i\sin\theta = \cos\theta + i\sin\theta$, which is clearly true.

(ii) Next, we assume the statement is true for $n = k$; that is,

$$(\cos\theta + i\sin\theta)^k = \cos k\theta + i\sin k\theta \quad (*)$$

and show that this implies the statement is true for $n = k + 1$; that is,

$$(\cos\theta + i\sin\theta)^{k+1} = \cos[(k+1)\theta] + i\sin[(k+1)\theta].$$

We have

$$\begin{aligned}
(\cos\theta + i\sin\theta)^{k+1} &= (\cos\theta + i\sin\theta)^k \cdot (\cos\theta + i\sin\theta) \\
&= (\cos k\theta + i\sin k\theta) \cdot (\cos\theta + i\sin\theta) \text{ by } (*), \\
&= \cos(k\theta + \theta) + i\sin(k\theta + \theta) \text{ by Theorem 1.2.10}, \\
&= \cos[(k+1)\theta] + i\sin[(k+1)\theta],
\end{aligned}$$

and therefore $(\cos\theta + i\sin\theta)^{k+1} = \cos[(k+1)\theta] + i\sin[(k+1)\theta]$. Hence if the statement is true for $n = k$, then the statement is true for $n = k + 1$.

Since (i) and (ii) hold, the statement is true for all $n \geqslant 1$ by the Principle of Mathematical Induction. □

4. Use induction to show that $n! > 2^n$ for $n \geqslant 4$. (Recall that $n! = n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1$.)

Proof. (i) We first show that the statement is true for $n = 4$. When $n = 4$, the inequality becomes $4! > 2^4$, and since $4! = 24$ and $2^4 = 16$, the statement is true. (Note that the statement is actually false for $n = 1, 2, 3$.)

(ii) Next, we assume the statement is true for $n = k \geqslant 4$; that is,

$$k! > 2^k \quad (*)$$

and show that this implies the statement is true for $n = k + 1$; that is,

$$(k+1)! > 2^{k+1}.$$

We have

$$\begin{aligned}
(k+1)! &= (k+1) \cdot k! \\
&> (k+1) \cdot 2^k \text{ by } (*), \\
&> 2 \cdot 2^k \text{ since } k \geqslant 4, \\
&= 2^{k+1},
\end{aligned}$$

and therefore $(k+1)! > 2^{k+1}$. Hence if the statement is true for $n = k$, then the statement is true for $n = k + 1$.

Since (i) and (ii) hold, the statement is true for all $n \geqslant 4$ by the Principle of Mathematical Induction. □

---

§2.1 Exercises

Prove the following statements using the Principle of Mathematical Induction.

1. $1 + 2 + 3 + \cdots + n = \dfrac{n(n+1)}{2}$ for $n \geqslant 1$.
2. $1^2 + 2^2 + 3^2 + \cdots + n^2 = \dfrac{n(n+1)(2n+1)}{6}$ for $n \geqslant 1$.
3. $1 + 2 + 2^2 + 2^3 + \cdots + 2^{n-1} = 2^n - 1$ for $n \geqslant 1$.
4. $1 + 3 + 5 + \cdots + (2n-1) = n^2$ for $n \geqslant 1$.
5. $1 \cdot 2 + 2 \cdot 3 + 3 \cdot 4 + \cdots + n(n+1) = \dfrac{n(n+1)(n+2)}{3}$ for $n \geqslant 1$.
6. $\dfrac{1}{1 \cdot 3} + \dfrac{1}{3 \cdot 5} + \dfrac{1}{5 \cdot 7} + \cdots + \dfrac{1}{(2n-1)(2n+1)} = \dfrac{n}{2n+1}$ for $n \geqslant 1$.
7. $\left(\dfrac{1}{2} + 1\right) \cdot \left(\dfrac{1}{2} + \dfrac{1}{2}\right) \cdot \left(\dfrac{1}{2} + \dfrac{1}{3}\right) \cdot \left(\dfrac{1}{2} + \dfrac{1}{4}\right) \cdots \left(\dfrac{1}{2} + \dfrac{1}{n}\right) = \dfrac{(n+1)(n+2)}{2^{n+1}}$ for $n \geqslant 1$.
8. $\dfrac{d}{dx}\, x^n = nx^{n-1}$ for $n \geqslant 1$, assuming only $\dfrac{d}{dx}\, x = 1$ and the product rule for differentiation.
9. If x is a positive real number, then $(1 + x)^n \geqslant 1 + nx$ for all natural numbers $n \geqslant 1$.
10. If x and y are real numbers, then $(xy)^n = x^n y^n$ for all natural numbers $n \geqslant 1$.
11. $4^n - 1$ is divisible by 3 for all $n \geqslant 0$.
12. A set with n elements has exactly 2^n subsets for $n \geqslant 0$.
13. $2^n > n$ for all $n \geqslant 1$.
14. $n! > n^2$ for $n \geqslant 4$.

2.2 Divisibility of Integers

Definition 2.2.1 *Let a and b be integers, with $a \neq 0$. We say a* **divides** *b, and write $a \mid b$, if $b = na$ for some integer n. If a does* not *divide b, we write $a \nmid b$.*

If a divides b, we also say that a is a **divisor** or **factor** of b, that b is a **multiple** of a, or that b is **divisible** by a. It is important to note that the notation $a \mid b$ stands for the statement that a divides b; it is NOT a number. In particular $a \mid b$ is not the fraction a/b.

Note that if $a \mid b$, then the rational number $\frac{b}{a}$ is an integer, since if $b = na$ for some integer n, then $\frac{b}{a} = n$. When working with integers, however, it is usually best to use the operation of multiplication and not division. It is therefore preferable to use the definition above in order to prove results about divisibility.

We will need to be able to use several important basic properties of divisibility. These are listed among the (potential) properties below. Consider the following problems before reading further.

Class Preparation Problems: Determine if each of the following statements is TRUE or FALSE. For each *true* statement, give a proof. For each *false* statement, find an example of integers for which the statement is false. (Answers appear on the next page.)

1. $0 \mid 0$.
2. If $a \neq 0$, then $a \mid 0$.
3. If $a \neq 0$, then $0 \mid a$.
4. If $a \neq 0$, then $a \mid a$.
5. If a and b are positive and $a \mid b$, then $a \leqslant b$.
   More generally, if a and b are any non-zero integers and $a \mid b$, then $|a| \leqslant |b|$.
6. If a is any integer, then $1 \mid a$.
7. If $a \mid 1$, then $a = \pm 1$.
8. If $a \mid b$, then $b \mid a$.
9. If $a \mid b$ and $b \mid a$, then $b = \pm a$.
10. If $a \mid b$ and $b \mid c$, then $a \mid c$.
11. If $a \mid b$ and $a \mid c$, then $a^2 \mid bc$.
12. If $a \mid b$ and $d \mid c$, then $ad \mid bc$.
13. If $a \mid bc$, then $a \mid b$ or $a \mid c$.
14. If $a \mid c$ and $b \mid c$, then $ab \mid c$.
15. If $a \mid b$ and $a \mid c$, then $a \mid b + c$.
16. If $a \mid b + c$, then $a \mid b$ or $a \mid c$.

Statements 1, 3, 8, 13, 14, and 16 above are FALSE, the other statements are true. Statement 11 is just a special case of statement 12 (let $d = a$). We restate the remaining true properties in the following theorem.

Theorem 2.2.2 (Divisibility Properties) *If a, b, and c are integers, then the following properties hold.*

i. *If $a \neq 0$, then $a \mid 0$.*

ii. *If $a \neq 0$, then $a \mid a$.*

iii. *If a is any integer, then $1 \mid a$.*

iv. *If a and b are non-zero and $a \mid b$, then $|a| \leqslant |b|$.*

v. *If $a \mid 1$, then $a = \pm 1$.*

vi. *If $a \mid b$ and $b \mid a$, then $b = \pm a$.*

vii. *If $a \mid b$ and $b \mid c$, then $a \mid c$.*

viii. *If $a \mid b$ and $d \mid c$, then $ad \mid bc$.*

ix. *If $a \mid b$ and $a \mid c$, then $a \mid b + c$.*

Proof. Property (i) follows from the definition and the fact that $0 = 0 \cdot a$. Properties (ii) and (iii) both follow from $a = 1 \cdot a$.

For (iv), note that if $a \mid b$, then $b = na$ for some integer n, and so $|b| = |n| \cdot |a|$. Since a and b are non-zero, n is also non-zero and $|n| \geqslant 1$ because n is an integer. Thus $|b| = |n| \cdot |a| \geqslant 1 \cdot |a| = |a|$ and $|b| \geqslant |a|$ as claimed.

For (v), we have by (iv) that if $a \mid 1$, then $|a| \leqslant 1$ and $a \neq 0$. Since a is an integer, this leaves only the possibility that $|a| = 1$ and so $a = \pm 1$. Similarly, if $a \mid b$ and $b \mid a$, then $|a| \leqslant |b|$ and $|b| \leqslant |a|$, hence $|a| = |b|$. Again, since a and b are integers, this implies $b = \pm a$ and (vi) holds.

To prove (vii), we observe that if $a \mid b$ and $b \mid c$, then by definition, $b = na$ and $c = mb$ for some integers m and n. Substituting, we have

$$c = m(na) = (mn)a$$

by associativity of multiplication in $\mathbb{Z}$. Since $\mathbb{Z}$ is closed under multiplication, mn is an integer, and so $c = (mn)a$ implies $a \mid c$ by definition.

Similarly, if $a \mid b$ and $d \mid c$, then $b = na$ and $c = md$ for integers m and n, and so

$$bc = (na)(md) = (nm)(ad)$$

by the associativity and commutativity of multiplication in $\mathbb{Z}$. Again, since $\mathbb{Z}$ is closed under multiplication, nm is an integer and $bc = (nm)(ad)$ implies $ad \mid bc$, and so (viii) holds.

Finally, if $a \mid b$ and $a \mid c$, then $b = ma$ and $c = na$ for some integers m and n. Thus

$$b + c = ma + na = (m + n)a$$

by the distributive law in $\mathbb{Z}$. Since $\mathbb{Z}$ is closed under addition, $m+n$ is an integer and $b+c = (m+n)a$ implies $a \mid b + c$ by definition and (ix) holds. □

The following theorem is very useful and important. Its proof is nearly identical to the proof of Theorem 2.2.2 (ix) above and is left as an exercise (see Exercise 2.2.7).

Theorem 2.2.3 (Combination Theorem) *If $a \mid b$ and $a \mid c$, then $a \mid bx + cy$ for all integers x and y.*

Theorem 2.2.2 (ix) is really just a special case of the Combination Theorem, with $x = y = 1$. By setting $x = 1$ and $y = -1$ in the Combination Theorem, it follows that if $a \mid b$ and $a \mid c$, then $a \mid b - c$ as well. This observation leads to the following useful corollary, whose proof is left as an exercise (see Exercise 2.2.9).

Corollary 2.2.4 *If $a \mid r + s$ and $a \mid r$, then $a \mid s$.*

In other words, if a divides a sum of two integers and divides one of the two integers, then a must divide the other.

Remark: It is an easy exercise to show that if $a \mid b$, then $(-a) \mid b$. Also, $a \mid (-b)$, and $(-a) \mid (-b)$. Thus the sign of an integer has no effect on divisibility, and the divisors of an integer come in pairs (a positive divisor and its negative). We will therefore usually consider only positive divisors of non-negative integers. Keep in mind, however, that nearly all of the results we prove are valid for negative integers as well.

---

§2.2 Exercises

1. Show that if $a \mid b$, then
   (a) $(-a) \mid b$,
   (b) $a \mid (-b)$,
   (c) $(-a) \mid (-b)$.
2. Use Definition 2.2.1 to prove that if $a \mid b$ and c is an integer, then $a \mid bc$.
3. Use Definition 2.2.1 to prove that if $a \mid c$ and $b \mid c$, then $ab \mid c^2$.
4. Find integers a, b, and c such that $a \mid bc$, but $a \nmid b$ and $a \nmid c$.
5. Find integers a, b, and c such that $a \mid c$ and $b \mid c$, but $ab \nmid c$.
6. Show that if $c \neq 0$ and $ac \mid bc$, then $a \mid b$.
7. Prove the Combination Theorem (Theorem 2.2.3):
   If $a \mid b$ and $a \mid c$, then $a \mid bx + cy$ for all integers x and y.
8. Find integers a, b, and c such that $a \mid b + c$, but $a \nmid b$ and $a \nmid c$.
9. Prove that if $a \mid r + s$ and $a \mid r$, then $a \mid s$.
10. Use induction to prove that $4 \mid 5^n - 1$ for all $n \geqslant 1$.
11. Use induction to prove that $5 \mid 6^n - 1$ for all $n \geqslant 1$.
12. Show that if a is a fixed positive integer, then $a \mid (a + 1)^n - 1$ for all $n \geqslant 1$.

2.3 Division Algorithm and Greatest Common Divisor

The following questions are intended primarily to prompt review of long division and lead to ideas on the Division Algorithm, greatest common divisors, and the Euclidean Algorithm. Try to answer them *before* reading further.

Class Preparation Problems:

1. Use long division to find the quotient and remainder for the following division problems:

$$1027 \div 8$$

$$1737 \div 9$$

2. How do you know when the long division process is finished?

3. Write an equation relating 1027, 8, and the quotient and remainder from the division $1027 \div 8$. Do the same for 1737 and 9.

4. Characterize the property $a \mid b$ in terms of the quotient and remainder in the division $b \div a$.

5. Use a calculator to find the quotient and remainder for the following division problems:

$$5746 \div 90$$

$$97635682 \div 8923$$

6. Explain your method for finding the quotient and remainder in the problems above. Why does your method work?

7. Devise a method for using *subtraction* to divide a positive integer b by a positive integer a to find the quotient and remainder.

8. How would you define the greatest common divisor (GCD) of two integers a and b? Are any restrictions on a and b necessary?

9. What methods do you know for finding the greatest common divisor of two integers? Make up some examples using your methods.

10. Methods taught in school for finding the greatest common divisor of integers a and b usually involve knowing all divisors or the prime factorizations of a and b. Do these methods work well for finding the GCD of 2145 and 546? What about the GCD of 100980 and 124488?

Long Division and the Division Algorithm

When we use long division to divide an integer b (the **dividend**) by an integer a (the **divisor**), the process continues until the remainder is less than $|a|$. (Note that the word "divisor" has a different meaning here than when we called a a divisor of b if $a \mid b$.) This guarantees the uniqueness of the quotient and remainder — there cannot be two different correct answers to a long division problem. The following theorem precisely states this property of division of integers. The Division Algorithm is usually proved by using the long division procedure along with induction. The proof is omitted here.

Theorem 2.3.1 (Division Algorithm) *If a and b are integers with $a > 0$, then there exist unique integers q and r satisfying $b = qa + r$ and $0 \leqslant r < a$.*

Remarks:

1. The integer q is called the **quotient** and r is the **remainder** on division of b by a.
2. The uniqueness of the remainder r is guaranteed by the condition that $0 \leqslant r < a$. There are many integers q and r such that $b = qa + r$, but only one pair q, r also satisfying $0 \leqslant r < a$. This uniqueness will be of great importance to us in our further study of properties of the integers.
3. In terms of rational numbers, the equation $b = qa + r$ can be rewritten as $\frac{b}{a} = q + \frac{r}{a}$. The condition that $0 \leqslant r < a$ guarantees that $0 \leqslant \frac{r}{a} < 1$. Thus q is the integer part of $\frac{b}{a}$ and $\frac{r}{a}$ is the fractional part.
4. The condition that $a \mid b$ is equivalent to the condition that the remainder r is 0.
5. It is not necessary to assume that a is positive in the Division Algorithm, if we assume that $a \neq 0$ and replace the condition $0 \leqslant r < a$ with $0 \leqslant r < |a|$. We will usually only consider division of positive integers, however, so assuming $a > 0$ will cause no serious problems.

The Division Algorithm is an extremely important theorem that will be the basis for much of what follows.

Greatest Common Divisor

A **common divisor** of two integers a and b is an integer that divides both a and b (that is, it is a divisor of both a and b). Naturally, the greatest common divisor of a and b is the *largest* among all common divisors. The precise definition is as follows.

Definition 2.3.2 *Let a and b be integers, at least one of which is not 0. The* **greatest common divisor** *of a and b is the (positive) integer d satisfying*

i. *$d \mid a$ and $d \mid b$,*

ii. *if $c \mid a$ and $c \mid b$, then $c \leqslant d$.*

Remarks:

1. We use the notation (a, b) for the greatest common divisor of a and b. We also frequently use the abbreviation GCD.
2. Since $1 \mid a$ and $1 \mid b$, condition (ii) of the definition implies $d \geqslant 1$. It is therefore not really necessary to specify that the GCD is a *positive* integer in the definition.
3. Condition (i) of the definition says that d is a *common divisor* of a and b. Condition (ii) says that d is larger than any other common divisor.
4. The GCD of two integers a and b is unique. (Exercise: Prove this.)

Finding all divisors of a and b and searching among them for the largest integer dividing both is not practical for large numbers. It is also not computationally feasible to use methods that require the prime factorizations of a and b, unless they are known ahead of time. A more computationally efficient method is the Euclidean Algorithm, demonstrated in the following example.

Example: Find the GCD of 657 and 306.

Start by using the Division Algorithm to find the quotient and remainder when dividing 657 by 306. For each successive step, divide the divisor of the current step by the remainder of the current step. Continue until the remainder is 0.

$$657 = 2 \times \underline{306} + \underline{45} \tag{2.1}$$
$$306 = 6 \times \underline{45} + \underline{36} \tag{2.2}$$
$$45 = 1 \times \underline{36} + \underline{9} \tag{2.3}$$
$$36 = 4 \times 9 + 0 \tag{2.4}$$

Here 9 is the last non-zero remainder, and we claim that $(657, 306) = 9$.

We first show condition (i) of the definition holds, that $9 \mid 657$ and $9 \mid 306$. Equation 2.4 above implies $9 \mid 36$ and by basic properties of divisibility we know $36 \mid 36$. Thus Equation 2.3 and the Combination Theorem together imply that $9 \mid 45$. We now have $9 \mid 36$ and $9 \mid 45$, so Equation 2.2 and the Combination Theorem imply $9 \mid 306$. Finally, since $9 \mid 45$ and $9 \mid 306$, Equation 2.1 and the Combination Theorem imply $9 \mid 657$. Hence (i) holds. (Note: We obviously could have shown that 9 divides both 657 and 306 by computation, but the method demonstrated here works to prove the result in general.)

We next show that condition (ii) of the definition holds, that if $c \mid 657$ and $c \mid 306$, then $c \leqslant 9$. By the equations above, we have:

$$45 = 657 - 2 \times 306 \tag{2.5}$$
$$36 = 306 - 6 \times 45 \tag{2.6}$$
$$9 = 45 - 1 \times 36 \tag{2.7}$$

If $c \mid 657$ and $c \mid 306$, then Equation 2.5 and the Combination Theorem imply that $c \mid 45$. Equation 2.6 and the Combination Theorem then imply that $c \mid 36$, and finally Equation 2.7 and the Combination Theorem imply $c \mid 9$. Hence by Theorem 2.2.2 (iv), we have $c \leqslant 9$, and condition (ii) holds.

We have shown that 9 satisfies the definition of the GCD of 657 and 306. Hence $(657, 306) = 9$ as claimed. □

The general method for finding the greatest common divisor of two integers is the Euclidean Algorithm, described as follows.

Theorem 2.3.3 (Euclidean Algorithm) *Let A and B be positive integers, with $B \geqslant A$. Use the Division Algorithm to obtain the following system of equations:*

$$\begin{array}{lcll} B & = & Q_1 A + R_1, & 0 < R_1 < A \\ A & = & Q_2 R_1 + R_2, & 0 < R_2 < R_1 \\ R_1 & = & Q_3 R_2 + R_3, & 0 < R_3 < R_2 \\ \vdots & \vdots & \vdots & \vdots \\ R_{n-2} & = & Q_n R_{n-1} + R_n, & 0 < R_n < R_{n-1} \\ R_{n-1} & = & Q_{n+1} R_n. & \end{array}$$

The last non-zero remainder, R_n, is the greatest common divisor of A and B.

Remarks:

1. The procedure used in the example above can be used to show that R_n divides both A and B, and if $c \mid A$ and $c \mid B$ then $c \mid R_n$, hence in particular $c \leqslant R_n$. It then follows from the definition that $R_n = (A, B)$.

2. If $d = (A, B)$, then by definition d is *greater than* any other common divisor. In the proof of the Euclidean Algorithm, we see that in fact d is *divisible* by any other common divisor. This stronger result will be stated more formally later.

We will give a more direct proof of the Euclidean Algorithm, using the following lemma and induction.

Lemma 2.3.4 *If a and b are non-zero integers, and q and r are integers such that $b = qa + r$, then $(b, a) = (a, r)$.*

Proof. If $c \mid a$ and $c \mid r$, then by the Combination Theorem, $c \mid qa + r$; that is, $c \mid b$. Conversely, if $c \mid b$ and $c \mid a$, then $c \mid b - qa$; that is, $c \mid r$. Therefore, the set of common divisors of a and r is precisely the same as the set of common divisors of b and a. The largest number in this set of common divisors is then equal to (a, r) and to (b, a), hence $(a, r) = (b, a)$. □

Proof of Theorem 2.3.3. We will first show that $(B, A) = (R_i, R_{i+1})$ for all $i = 1, \ldots, n-1$. The proof is by induction on i.

By Lemma 2.3.4, $B = Q_1 A + R_1$ implies that $(B, A) = (A, R_1)$ and $A = Q_2 R_1 + R_2$ implies that $(A, R_1) = (R_1, R_2)$. Hence $(B, A) = (R_1, R_2)$ and the claim holds for $i = 1$.

Now assume the claim is true for $i = k$; that is, $(B, A) = (R_k, R_{k+1})$. Since

$$R_k = Q_{k+2} R_{k+1} + R_{k+2},$$

Lemma 2.3.4 implies $(R_k, R_{k+1}) = (R_{k+1}, R_{k+2})$. Thus $(B, A) = (R_{k+1}, R_{k+2})$, and the claim is true for $i = k + 1$.

By the Principle of Mathematical Induction, we therefore have that $(B, A) = (R_i, R_{i+1})$ for all $i = 1, \ldots, n-1$. In particular, $(B, A) = (R_{n-1}, R_n)$.

Finally, $R_n \mid R_n$ and $R_n \mid R_{n-1}$ (as $R_{n-1} = Q_{n+1}R_n$), and if c is any other common divisor, then $c \mid R_n$ and so $c \leqslant R_n$. Hence $(B, A) = (R_{n-1}, R_n) = R_n$ and the theorem is proved. □

Another important property of the GCD of two integers is that it can be written as a combination of the integers in a particular way. In general, the remainder in each equation in the Euclidean Algorithm can be written as a combination of the dividend and divisor of the equation. Working backwards through the algorithm, we can then always write the GCD of a and b as an integer combination of a and b, that is, $(a, b) = ax + by$ for some integers x and y. This result will be stated formally and proved in the next section (see Theorem 2.4.5). The procedure is demonstrated in the following example.

Example: We used the Euclidean Algorithm to find $(657, 306)$, and had the following equations:

$$\begin{aligned} 657 &= 2 \times 306 + 45 && (2.8)\\ 306 &= 6 \times 45 + 36 && (2.9)\\ 45 &= 1 \times 36 + 9 && (2.10)\\ 36 &= 4 \times 9 + 0 && (2.11) \end{aligned}$$

We begin with Equation 2.10, writing the remainder 9 as a combination of the previous dividend 45 and previous remainder 36. Continue the procedure with each previous equation:

$$\begin{aligned} 9 &= 45 - 1(36) \text{ by Equation 2.10,}\\ &= 45 - 1[306 - 6(45)] \text{ by Equation 2.9,}\\ &= (-1)306 + 7(45)\\ &= (-1)306 + 7[657 - 2(306)] \text{ by Equation 2.8,}\\ &= 7(657) - 15(306). \end{aligned}$$

Thus we can write $9 = (657, 306) = 657(7) + 306(-15)$. □

§2.3 Exercises

1. Use *long division* to find the quotient and remainder for the following divisions, and write the equation of the form $b = qa + r$ for each.
   (a) $4752 \div 35$
   (b) $9976 \div 43$
2. Use a calculator to find the quotient and remainder for the following divisions, and write the equation of the form $b = qa + r$ for each.
   (a) $2351487 \div 5726$
   (b) $84637851 \div 7498$
3. Show that at least one of any three consecutive integers must be divisible by 3.
   [Hint: Let b, $b+1$, and $b+2$ be the three consecutive integers. By the Division Algorithm (Theorem 2.3.1), there are three cases to consider: $b = 3q$, $b = 3q+1$, or $b = 3q+2$, for some integer q.]
4. Use the Euclidean Algorithm to find $(2145, 546)$.
5. Use the Euclidean Algorithm to find $(3054, 162)$.
6. Use the Euclidean Algorithm to find $(5967, 1540)$.
7. Use the Euclidean Algorithm to find $(272, 119)$ and then find integers x and y such that $(272, 119) = 272x + 119y$.
8. Use the Euclidean Algorithm to find $(495, 210)$ and then find integers x and y such that $(495, 210) = 495x + 210y$.
9. Use the Euclidean Algorithm to find $(264, 189)$ and then find integers x and y such that $(264, 189) = 264x + 189y$.
10. Use the Euclidean Algorithm to find $(510, 414)$ and then find integers x and y such that $(510, 414) = 510x + 414y$.
11. Show that if $(a, b) = 1$ and $c \mid a$, then $(c, b) = 1$.
12. Show that if a and b are integers, at least one of which is non-zero, then $(-a, b) = (a, b)$.
    [Hint: Let $d = (-a, b)$ and show that d satisfies the conditions of Definition 2.3.2 for (a, b).]
13. Show that if a and b are integers with $a > 0$ and $a \mid b$, then $(a, b) = a$.
    [Hint: Show that a satisfies the conditions of Definition 2.3.2.]
14. Use Exercise 13 to show that if a and b are integers with $a > 0$, then the following hold:
    (a) $(a, 0) = a$,
    (b) $(1, b) = 1$,
    (c) $(a, a) = a$.

2.4 Properties of the Greatest Common Divisor

The following basic property of the GCD is easily verified using the definition of GCD and properties of divisibility.

Proposition 2.4.1 *Let a and b be integers with $a > 0$. If $a \mid b$, then $(a, b) = a$.*

Proof. We show that parts (i) and (ii) of the definition of GCD (Definition 2.3.2) are satisfied by a. Since $a \neq 0$, we know that $a \mid a$ (Theorem 2.2.2 (ii)), and we are given that $a \mid b$. Hence part (i) of the definition holds with $d = a$.

Suppose now that c is an integer such that $c \mid a$ and $c \mid b$. Since $a > 0$ and $c \mid a$, we have $c \leqslant a$ by Theorem 2.2.2 (iv). Hence part (ii) of the definition holds with $d = a$, and so $(a, b) = a$ as claimed. □

Applying the proposition to certain special cases, we get an immediate corollary.

Corollary 2.4.2 *If a and b are integers with $a > 0$, then*

i. $(a, 0) = a$,

ii. $(1, b) = 1$,

iii. $(a, a) = a$.

Proof. Since $a > 0$, we have by Theorem 2.2.2 that $a \mid 0$, $1 \mid b$, and $a \mid a$. Thus (i), (ii), and (iii), respectively, follow from Proposition 2.4.1. □

We noted previously that the divisors of an integer and its negative are the same. Therefore the signs of the integers a and b do not affect the GCD of a and b, and so the following proposition holds.

Proposition 2.4.3 *If a and b are integers, at least one of which is not 0, then*

$$(a, b) = (a, -b) = (-a, b) = (-a, -b).$$

In an earlier example, we used the Euclidean Algorithm to show $(657, 306) = 9$. In the process, we actually showed that if $c \mid 657$ and $c \mid 306$, then not only is $c \leqslant 9$, but in fact $c \mid 9$. Moreover, we showed that $(657, 306)$ could be written as an integer combination of 657 and 306. In particular, we found that $(657, 306) = 9 = 657(7) + 306(-15)$. Both of these observations are true in general and can be proved by working through the equations arising in the Euclidean Algorithm (see Theorem 2.3.3), as demonstrated in the examples.

We will prove both of these properties of the GCD in another way, using the following important characterization of the GCD.

Lemma 2.4.4 *Let a and b be integers, at least one of which is not 0. The GCD of a and b is the* smallest *positive integer that can be written in the form $ar + bs$ for integers r and s.*

Proof. Let a and b be integers, at least one of which is not 0, and define the set

$$\mathcal{S} = \{ar + bs \mid r, s \in \mathbb{Z} \text{ and } ar + bs > 0\}.$$

Since at least one of a or b is non-zero, at least one of a, $-a$, b, or $-b$ is *positive* and is therefore in $\mathcal{S}$. In particular, $\mathcal{S}$ is a non-empty set of natural numbers, hence has a smallest element d by the Well-ordering Principle (Theorem 2.1.1). Since $d \in \mathcal{S}$, we have $d = ax + by$ for some integers x, y.

The conclusion of the theorem is that this positive integer d is the GCD of a and b. We will show that d satisfies parts (i) and (ii) of Definition 2.3.2, hence $d = (a, b)$.

We first show (i), that $d \mid a$ and $d \mid b$. By the Division Algorithm (Theorem 2.3.1), we can write $a = qd + r$ for integers q and r with $0 \leqslant r < d$. Moreover, we can express r as

$$r = a - qd = a - q(ax + by) = a(1 - qx) + b(-qy).$$

Since $1 - qx$ and $-qy$ are integers, if $r > 0$, then $r \in \mathcal{S}$. However, d was chosen as the *smallest* element of $\mathcal{S}$ and $r < d$. Hence r is not an element of $\mathcal{S}$, and so $r = 0$ and $a = qd$. Therefore $d \mid a$. The proof that $d \mid b$ is nearly identical and is left as an exercise.

Finally, we show (ii), that if c is an integer such that $c \mid a$ and $c \mid b$, then $c \leqslant d$. Suppose $c \mid a$ and $c \mid b$. Since $d = ax + by$ with $x, y \in \mathbb{Z}$, the Combination Theorem (Theorem 2.2.3) says that $c \mid d$. Since $c \mid d$ and $d > 0$, it follows from Theorem 2.2.2 (iv) that $c \leqslant d$ as claimed. □

The two properties of the GCD mentioned above now follow easily from this lemma.

Theorem 2.4.5 *If a and b are integers, at least one of which is not 0, then there are integers x and y such that $(a, b) = ax + by$.*

Proof. This follows immediately from Lemma 2.4.4. □

This theorem states that the GCD of a and b can be written as an integer combination of a and b. Note, however, that the GCD is *not* the only integer that can be written as such a combination. For example, since $9 = 657(7) + 306(-15)$, we also have $45 = 657(35) + 306(-75)$. In fact, *any* multiple of 9 can be written as a combination of 657 and 306, as $n \cdot 9 = 657(7n) + 306(-15n)$.

More generally, we have the following characterization of integers that can be written in the form $ar + bs$, with r and s integers.

Corollary 2.4.6 *Let a and b be integers, at least one of which is not 0. Let $d = (a, b)$ and let m be an integer. The integer m can be written in the form $m = ar + bs$, with r and s integers, if and only if m is a multiple of d (that is, $d \mid m$).*

Proof. If $m = ar + bs$ for some integers r and s, then since $d \mid a$ and $d \mid b$, the Combination Theorem implies $d \mid m$. Thus m is a multiple of d.

Conversely, suppose m is a multiple of d, so that $m = nd$ for some integer n. Since $d = ax + by$ for some $x, y \in \mathbb{Z}$ by Theorem 2.4.5, we have $m = nd = a(nx) + b(ny)$, and nx, ny are integers. Hence d can be expressed in the form $ar + bs$, with $r = nx$ and $s = ny$ integers. □

In particular, the fact that $m = ar + bs$ for some integers r and s does *not* imply that $m = (a, b)$. It only implies m is a *multiple* of (a, b).

Theorem 2.4.7 *Let a and b be integers, at least one of which is not 0. If $c \mid a$ and $c \mid b$, then $c \mid (a, b)$.*

Proof. By Theorem 2.4.5, we can write $(a, b) = ax + by$ with $x, y \in \mathbb{Z}$. If $c \mid a$ and $c \mid b$, then by the Combination Theorem, we have $c \mid ax + by$, and so $c \mid (a, b)$ as claimed. □

Thus, not only is (a, b) greater than any other common divisor of a and b, (a, b) is also divisible by any other common divisor. By basic properties of divisibility, we know that if $c \mid (a, b)$, then $c \leqslant (a, b)$ as well (since (a, b) is positive). We could therefore obtain an alternate definition of the GCD by replacing $c \leqslant d$ in condition (ii) by $c \mid d$. This does not guarantee that d is positive, however, so $d > 0$ would also have to be assumed. We state this equivalent characterization of the GCD as a theorem.

Theorem 2.4.8 *Let a and b be integers, at least one of which is not 0, and let d be a positive integer. Then $d = (a, b)$ if and only if*

i. *$d \mid a$ and $d \mid b$, and*

ii. *if $c \mid a$ and $c \mid b$, then $c \mid d$.*

Relatively Prime Pairs of Integers

To study further properties of divisibility and the GCD, we will consider a special case.

Definition 2.4.9 *We say two integers a and b are* **relatively prime** *(or* **coprime***) if $(a, b) = 1$.*

Work through the following problems prior to reading further.

Class Preparation Problems:

1. Derive the consequences of Lemma 2.4.4, Theorem 2.4.5, and Corollary 2.4.6 in the case where a and b are relatively prime.

2. Show that $(5, 8) = 1$ and find integers x and y so that $5x + 8y = 1$.

3. Express 2, -13, and 37 in the form $5r + 8s$, with r and s integers. Are there any integers m that *cannot* be written in this form? Explain.

4. Suppose you want to put a certain number of gallons of water into a pool, and you only have an 8 gallon bucket and a 5 gallon bucket (with no markings) to measure the water. How could you measure 7 gallons of water into the pool? What about 11 gallons? For any positive integer m, how could you measure m gallons into the pool?

5. Suppose now that you have only a 9 gallon bucket and a 12 gallon bucket. Could you measure 15 gallons of water into the pool? What about 20 gallons? What integer numbers of gallons of water can be measured?

6. What relationship between the bucket sizes is necessary in order to guarantee that every positive integer number of gallons of water can be measured with two buckets? (Given sufficient water supply, of course.)

7. Note that $6 = 132(8) + 105(-10)$. Can we conclude that $(132, 105) = 6$? Why or why not?

8. Now note that $1 = 130(16) + 231(-9)$. Can we conclude that $(130, 231) = 1$? Why or why not?

9. Verify that $(6, 10) = 2$. Also find the following GCDs:

$$\begin{aligned} (12, 20) &= (2 \cdot 6, 2 \cdot 10) \\ (30, 50) &= (5 \cdot 6, 5 \cdot 10) \\ (42, 70) &= (7 \cdot 6, 7 \cdot 10) \\ (60, 100) &= (10 \cdot 6, 10 \cdot 10). \end{aligned}$$

   Do these calculations suggest a general result?

10. Note again that $(6, 10) = 2$. What is $(\frac{6}{2}, \frac{10}{2}) = (3, 5)$? Observe also that $(24, 54) = 6$. What is $(\frac{24}{6}, \frac{54}{6}) = (4, 9)$? Do these examples suggest a general result?

We now study further properties of the GCD and derive some consequences of previous theorems in the case that a and b are relatively prime.

Theorem 2.4.10 *If $(a, b) = 1$, then there are integers x and y such that $ax + by = 1$.*

Proof. This is Theorem 2.4.5 with $(a, b) = 1$. □

Corollary 2.4.11 *If $(a, b) = 1$ and m is* any *integer, then there are integers r and s such that $m = ar + bs$.*

Proof. Since $(a, b) = 1$ and $1 \mid m$ for every integer m, this follows from Corollary 2.4.6. □

We saw before that in general $m = ax+by$, with x and y integers, does not imply that $m = (a, b)$. We were able to describe all integers that can be expressed in this form. If $1 = ax + by$ for some integers x and y, however, we can conclude that $(a, b) = 1$. We get the converse of Theorem 2.4.10 in this special case.

Theorem 2.4.12 *If there are integers x and y such that $ax + by = 1$, then $(a, b) = 1$.*

Proof. Since $1 = ax + by$ with $x, y \in \mathbb{Z}$, it follows from Corollary 2.4.6 that $(a, b) \mid 1$. Hence by Theorem 2.2.2 (v), $(a, b) = \pm 1$, and since $(a, b) > 0$, we have $(a, b) = 1$. □

The next results are the general theorems suggested by Problems 9 and 10 above.

Theorem 2.4.13 *If $(a,b) = d$ and m is any positive integer, then $(ma, mb) = md$; i.e. $(ma, mb) = m(a,b)$.*

Proof. We are given that $(a,b) = d$ and need to show that $(ma, mb) = md$. Therefore, by Theorem 2.4.8, we need to show

i. $md \mid ma$ and $md \mid mb$, and

ii. if $c \mid ma$ and $c \mid mb$, then $c \mid md$.

(i) Since $(a,b) = d$, we have $d \mid a$ and $d \mid b$. Hence $a = dr$ and $b = ds$ for some $r, s \in \mathbb{Z}$. Thus $ma = (md)r$ and $mb = (md)s$, and so $md \mid ma$ and $md \mid mb$.

(ii) Since $(a,b) = d$, we have $d = ax + by$ for some $x, y \in \mathbb{Z}$, by Theorem 2.4.5. Therefore $md = (ma)x + (mb)y$, and so if $c \mid ma$ ad $c \mid mb$, then $c \mid md$ by the Combination Theorem. □

Theorem 2.4.14 *If $(a,b) = d$, then $(\frac{a}{d}, \frac{b}{d}) = 1$.*

Proof. Since $(a,b) = d$, both $\frac{a}{d}$ and $\frac{b}{d}$ are integers. Therefore, by Theorem 2.4.13,

$$d = (a,b) = \left(d \cdot \frac{a}{d}, d \cdot \frac{b}{d}\right) = d \cdot \left(\frac{a}{d}, \frac{b}{d}\right).$$

Since $d \neq 0$, we can divide both sides of the equation by d to obtain $1 = (\frac{a}{d}, \frac{b}{d})$. □

Note that in general, *both* a and b must be divided by (a,b) in order to obtain a pair of relatively prime integers. For example, $(12, 10) = 2$ and, consistent with the theorem,

$$\left(\frac{12}{2}, \frac{10}{2}\right) = (6,5) = 1.$$

However,

$$\left(\frac{12}{2}, 10\right) = (6, 10) = 2 \neq 1.$$

On previous exercises, we showed that, in general, (1) $a \mid bc$ does *not* imply $a \mid b$ or $a \mid c$, and that (2) $a \mid c$ and $b \mid c$ does *not* imply $ab \mid c$. We do get these implications if we add an appropriate condition, however.

Theorem 2.4.15 (Euclid's Lemma) *If $a \mid bc$ and $(a,b) = 1$, then $a \mid c$.*

Proof. Since $(a,b) = 1$, we have $1 = ax + by$ for some $x, y \in \mathbb{Z}$, by Theorem 2.4.10. Thus

$$\begin{aligned} c &= (ax + by)c \\ &= (ax)c + (by)c \\ &= a(xc) + (bc)y. \end{aligned}$$

Since $a \mid a$ and $a \mid bc$, we have $a \mid a(xc) + (bc)y$ by the Combination Theorem, and hence $a \mid c$. □

Theorem 2.4.16 *If $a \mid c$ and $b \mid c$, and $(a, b) = 1$, then $ab \mid c$.*

Proof. Since $(a, b) = 1$, we have $1 = ax + by$ for some $x, y \in \mathbb{Z}$, by Theorem 2.4.10. Since $a \mid c$ and $b \mid c$, we have $c = am$ and $c = bn$ for some $m, n \in \mathbb{Z}$. Thus

$$\begin{aligned} c &= c(ax + by) \\ &= c(ax) + c(by) \\ &= (ca)x + (cb)y \\ &= (bna)x + (amb)y \\ &= ab(nx + my). \end{aligned}$$

All of m, n, x, and y are integers and $\mathbb{Z}$ is closed under addition and multiplication, hence $nx + my$ is an integer. Thus $c = ab(nx + my)$ implies $ab \mid c$. □

These results can be generalized as in the following theorems. The proofs are nearly identical to those of Theorems 2.4.15 and 2.4.16 above, and are left as exercises (see Exercises 2.4.2 and 2.4.3).

Theorem 2.4.17 *If $a \mid bc$ and $(a, b) = d$, then $a \mid cd$.*

Theorem 2.4.18 *If $a \mid c$ and $b \mid c$, and $(a, b) = d$, then $ab \mid cd$.*

---

§2.4 Exercises

1. Assume that $8 = (56, 72) = 56(4) + 72(-3)$. Determine which of the following integers can be expressed in the form $56r + 72s$, with r and s integers. For those that can, find r and s, and for those that cannot, explain why.

   (a) -16 (a) 42 (c) 70

   (b) -28 (b) 64 (d) -88

2. Show that if $a \mid bc$ and $(a, b) = d$, then $a \mid cd$; that is, prove Theorem 2.4.17.
   [Hint: Adapt the proof of Euclid's Lemma (Theorem 2.4.15), using the fact that $d = ax + by$ for some integers x and y.]

3. Show that if $a \mid c$ and $b \mid c$, and $(a, b) = d$, then $ab \mid cd$; that is, prove Theorem 2.4.18.
   [Hint: Adapt the proof of Theorem 2.4.16, using the fact that $d = ax + by$ for some integers x and y.]

4. Explain how you could measure out exactly 13 ounces of water given a 7 ounce cup and a 9 ounce cup. Is there any positive integer m for which it would be impossible to measure m ounces of water? (Assume an unlimited supply of water.) **Explain.**

5. Which of the amounts of water below could be measured given a 6 ounce cup and a 9 ounce cup, and which could not? Fully **explain** your answer.

   (a) 3 ounces (c) 11 ounces

   (b) 8 ounces (d) 21 ounces

2.5 Prime Numbers

Work through the following questions prior to reading further.

Class Preparation Problems:

1. What is a prime number?
2. What is a composite number?
3. Is 1 a prime?
4. Are there any even primes? If so, how many?
5. How can we determine if a number is prime or not? That is, given an integer n, how would you go about deciding whether or not n is prime? Is it necessary to check all positive integers less than or equal to n as potential divisors?
6. How many primes are there? Finitely many? Infinitely many?
7. Is there a largest prime number?
8. We have seen that in general, if $a \mid bc$ then a may or may not divide b or c. If p is prime and $p \mid bc$, does p necessarily divide a or b? Why or why not? Can you find an example where p is prime and $p \mid bc$, but $p \nmid b$ and $p \nmid c$?
9. If p is prime, can $\sqrt{p}$ be a rational number?
10. If n is a positive integer, under what conditions is $\sqrt{n}$ a rational number? Can $\sqrt{n}$ be rational and *not* be an integer?
11. What is meant by the *prime factorization* of a positive integer?
12. Can every positive integer be written as a product of prime numbers?
13. Can a positive integer be written as a product of primes in more than one way (other than the obvious variation of changing the order of the primes)?

Definition 2.5.1 *An integer $p > 1$ is a* **prime number** *(or a* **prime***) if the only positive divisors of p are 1 and p. If an integer $n > 1$ is not prime, we say n is a* **composite number**.

Remarks:

1. The first few primes are 2, 3, 5, 7, 11, 13, 17, 19, 23. The only even prime is 2, since every even number has 2 as a factor.
2. The largest *known* prime is $2^{43,112,609} - 1$, with 12,978,189 digits, discovered in August 2008.
3. The websites `www.mersenne.org` and `www.utm.edu/research/primes` are excellent sources for information on prime numbers.
4. An integer $n > 1$ is composite if and only if $n = a \cdot b$, for some integers a and b with $1 < a < n$ and $1 < b < n$.

Theorem 2.5.2 (Prime Divisor Principle) *If n is an integer and $n > 1$, then there is a prime p that divides n.*

Proof. Let $\mathcal{S}$ be the set of all divisors d of n with $d > 1$. Since $n \mid n$ and $n > 1$, we have $n \in \mathcal{S}$ and so $\mathcal{S}$ is non-empty. By the Well-ordering Principle (Theorem 2.1.1), $\mathcal{S}$ has a smallest element, say p. Thus p is the smallest divisor of n that is greater than 1. We will show that p is a prime.

Suppose p is *not* prime. Then, since $p > 1$, we have $p = ab$, where a and b are integers satisfying $1 < a < p$ and $1 < b < p$. Now $p = ab$, so $a \mid p$, and we chose p to be a divisor of n, so $p \mid n$. Thus by transitivity of divisibility (Theorem 2.2.2 (vii)), we have $a \mid n$. But $1 < a < p$ and p is the smallest divisor of n greater than 1, a contradiction. Therefore, our assumption that p is *not* prime must be false, and so p is prime. □

In order to prove that a given positive integer n is prime using the definition, it is necessary to verify that no integer a with $1 < a < n$ is a divisor of n. By the Prime Divisor Principle, it is only necessary to determine whether each *prime* less than n is a divisor of n. The next results further reduce the amount of work required.

Lemma 2.5.3 *If $n > 1$ is a composite number, then there is a prime p with $p \mid n$ and $p \leqslant \sqrt{n}$.*

Proof. Since n is composite, we know that $n = ab$ for some $a, b \in \mathbb{Z}$ with $1 < a < n$ and $1 < b < n$. We may assume without loss of generality that $a \leqslant b$. If $a > \sqrt{n}$, then we have $b > \sqrt{n}$ and so

$$n = ab > \sqrt{n}\sqrt{n} = n,$$

a contradiction. Hence $a \leqslant \sqrt{n}$.

By the Prime Divisor Principle, there is a prime p such that $p \mid a$. Since $p \mid a$ and $a \mid n$, Theorem 2.2.2 (vii) implies $p \mid n$. Since $p \mid a$ and $a > 0$, Theorem 2.2.2 (iv) implies $p \leqslant a$. We showed that $a \leqslant \sqrt{n}$, hence $p \leqslant \sqrt{n}$. □

The following theorem is an equivalent restatement of the lemma.

Theorem 2.5.4 (Prime Test) *Let $n > 1$ be an integer. If no prime p with $p \leqslant \sqrt{n}$ divides n, then n is prime.*

Proof. Since $n > 1$, either n is prime or n is composite. If n were composite, then Lemma 2.5.3 implies there would be a prime divisor p of n with $p \leqslant \sqrt{n}$. By hypothesis, this is not the case, and so n is prime. □

An ancient method for listing all primes up to a given number n is the **Sieve of Eratosthenes**. First write down all the numbers from 2 up to n. Start by crossing out all multiples of 2. The first number not crossed out is a prime (in this step, 3). Next cross out all multiples of this next prime. Again, the next number not crossed out is prime. Continue this process until the next number remaining is greater than $\sqrt{n}$. All of the numbers not crossed out are then primes.

Unfortunately, this method has several drawbacks. What possible drawbacks to this method do you see?

We mentioned above that the largest currently *known* prime is $2^{43,112,609} - 1$. However, there can be no absolute largest prime, as the following result of Euclid implies.

Theorem 2.5.5 *There are infinitely many primes.*

Proof. Suppose that there are only finitely many primes, say p_1, p_2, $p_3, \ldots, p_k$ are all of them. By Theorem 2.5.2, the integer

$$n = p_1 \cdot p_2 \cdot p_3 \cdots p_k + 1$$

has a prime divisor p, which must be in the given list of primes, say $p = p_i$. Now $p \mid n$ and $p_i \mid p_1 \cdot p_2 \cdot p_3 \cdots p_k$, hence $p \mid p_1 \cdot p_2 \cdot p_3 \cdots p_k$. The Combination Theorem then implies

$$p \mid n - p_1 \cdot p_2 \cdot p_3 \cdots p_k,$$

that is, $p \mid 1$. But this implies $p = \pm 1$, contradicting the fact that p is prime. Hence our assumption that there are only finitely many primes must be false. □

The next theorem is a special case of Euclid's Lemma where the divisor is prime.

Theorem 2.5.6 (Euclid's Lemma for Primes) *If p is a prime and $p \mid ab$, then $p \mid a$ or $p \mid b$.*

Proof. Since p is prime and (p, a) is a positive divisor of p, we have that either $(p, a) = 1$ or $(p, a) = p$. If $(p, a) = 1$, then $p \mid b$ by Euclid's Lemma (Theorem 2.4.15). If $(p, a) = p$, then $p \mid a$ by definition of GCD. Hence either $p \mid a$ or $p \mid b$. □

Corollary 2.5.7 *If p is prime and $p \mid a_1 a_2 \cdots a_n$, then $p \mid a_i$ for some i.*

Proof. We proceed by induction on the number n of factors. If $n = 1$, the hypothesis becomes $p \mid a_1$ and the conclusion is obvious. Hence the result is true if $n = 1$.

Assume now that the result holds for $n = k$; that is,

if $p \mid a_1 a_2 \cdots a_k$ then $p \mid a_i$ for some $i = 1, 2, \ldots, k$ (*)

and show that the result holds for $n = k + 1$; that is,

if $p \mid a_1 a_2 \cdots a_k a_{k+1}$ then $p \mid a_i$ for some $i = 1, 2, \ldots, k + 1$.

Let $a = a_1 a_2 \cdots a_k$ and $b = a_{k+1}$. If $p \mid a_1 a_2 \cdots a_k a_{k+1}$, then $p \mid ab$. By Theorem 2.5.6, we have that $p \mid a$ or $p \mid b$. If $p \mid a$, then $p \mid a_1 a_2 \cdots a_k$, and by the inductive hypothesis (*), $p \mid a_i$ for some $i = 1, 2, \ldots, k$. If $p \mid b$, then $p \mid a_{k+1}$. Hence, in any case, $p \mid a_i$ for some $i = 1, 2, \ldots, k + 1$. Therefore, if the statement is true for $n = k$, then it is true for $n = k + 1$, hence the result holds for all $n \geqslant 1$ by the Principle of Mathematical Induction. □

Corollary 2.5.8 *If p is prime and $p \mid q_1 q_2 \cdots q_n$, where $q_1, q_2, \ldots, q_n$ are primes, then $p = q_i$ for some i.*

Proof. By the previous corollary, if $p \mid q_1 q_2 \cdots q_n$, then $p \mid q_i$ for some $i = 1, 2, \ldots, n$. Since q_i is prime, its only positive divisors are 1 and q_i, and since $p > 1$, this implies $p = q_i$. □

A nice application of Euclid's Lemma is one of the proofs of the irrationality of $\sqrt{2}$, or more generally of the square root of a prime.

Theorem 2.5.9 *If p is a prime, then $\sqrt{p}$ is irrational.*

Proof. Let p be prime and suppose $\sqrt{p}$ is rational. We can then write $\sqrt{p} = \frac{a}{b}$, where a and b are integers and the fraction is in lowest terms, that is, $(a, b) = 1$. It follows that $a^2 = pb^2$.

We have $p \mid a^2$, or $p \mid a \cdot a$, and so $p \mid a$ by Theorem 2.5.6. Thus $a = pm$ for some integer m, and substituting in the equation above yields $p^2m^2 = pb^2$. Since $p \neq 0$ we can divide both sides of this equation by p to obtain $pm^2 = b^2$. As before, this implies $p \mid b^2$, and so by Theorem 2.5.6, we have $p \mid b$.

We have shown that the prime p divides both a and b, contradicting the fact that $(a, b) = 1$. Hence our assumption that $\sqrt{p}$ is rational must be false. □

Using properties of the GCD, we can generalize this result to characterize the integers whose square roots are rational.

Theorem 2.5.10 *If n is a positive integer, then either n is a perfect square (that is, $\sqrt{n}$ is an integer) or $\sqrt{n}$ is irrational.*

Proof. Let $n \in \mathbb{Z}$ and suppose $\sqrt{n}$ is rational. We can then write $\sqrt{n} = \frac{a}{b}$ with $(a, b) = 1$, thus $a = b\sqrt{n}$. By Theorem 2.4.10, $1 = ax + by$ for some integers x and y. Multiplying both sides of this equation by $\sqrt{n}$ and substituting using $a = b\sqrt{n}$ yields

$$\begin{aligned} \sqrt{n} &= a\sqrt{n}x + b\sqrt{n}y \\ &= (b\sqrt{n})\sqrt{n}x + ay \\ &= bnx + ay. \end{aligned}$$

Since a, b, n, x, and y are all integers and $\mathbb{Z}$ is closed under multiplication and addition, we have that $bnx + ay = \sqrt{n}$ is an integer. We have shown that if $\sqrt{n}$ is rational, then $\sqrt{n}$ is an integer, and the theorem follows. □

The next result is an extremely important property of integers (as the name suggests) that we usually take for granted.

Theorem 2.5.11 (Fundamental Theorem of Arithmetic) *Every integer $n > 1$ is either a prime or is a product of primes. The expression of n as a product of primes is unique except for the order in which the factors are written.*

Proof. We first show such a factorization exists. By the Well-ordering Principle (Theorem 2.1.1), if there is an integer greater than 1 that is *not* a prime or a product of primes, then there is a smallest such integer, say m. Since $m > 1$ and m is not prime, $m = ab$ for some integers a and b with $1 < a < m$ and $1 < b < m$. By the minimality of m, each of a, b is either prime or a product of primes. Hence $ab = m$ is a product of primes, contradicting the choice of m. Therefore, our assumption that there is an integer greater than 1 that is neither prime nor a product of primes must be false.

To prove uniqueness, suppose n can be written as a product of primes in two ways. Let $p_1, p_2, \ldots, p_r$ be all of the distinct primes that appear in at least one of the two factorizations. Writing products of repeated prime factors as powers, and recalling that $p_i^0 = 1$, we can write the two factorizations in the form

$$n = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r} = p_1^{b_1} p_2^{b_2} \cdots p_r^{b_r},$$

where $a_i, b_i \in \mathbb{Z}$ and $a_i \geqslant 0$, $b_i \geqslant 0$ for all i (and, for each i, at least one of a_i, b_i is non-zero). We must show that $a_i = b_i$ for all i.

Suppose that $a_i \neq b_i$ for some i. We may assume, without loss of generality, that $i = 1$ and $b_1 < a_1$. Dividing both factorizations by $p_1^{b_1}$ yields

$$p_1^{a_1 - b_1} p_2^{a_2} \cdots p_r^{a_r} = p_2^{b_2} \cdots p_r^{b_r}.$$

Now $b_1 < a_1$, hence $a_1 - b_1 > 0$ and so p_1 divides the factorization on the left. By equality, we must also have

$$p_1 \mid p_2^{b_2} \cdots p_r^{b_r}.$$

By Corollary 2.5.8, this implies $p_1 = p_j$ for some $2 \leqslant j \leqslant r$, contradicting the fact that $p_1, p_2, \ldots, p_r$ are distinct primes. Hence $a_i = b_i$ for all i and the expression of n as a product of primes is unique, except for the order of the factors. □

By prescribing that the primes in a prime factorization be written in increasing order and writing repeated products as powers, we obtain an absolutely unique prime factorization for each integer greater than 1.

Corollary 2.5.12 *Every integer $n > 1$ can be expressed in exactly one way in the form*

$$n = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r},$$

where $p_1 < p_2 < \cdots < p_r$ are primes and $a_i \geqslant 1$ for all i.

Definition 2.5.13 *The expression of n as a product of prime powers satisfying the conditions in Corollary 2.5.12 is called the* **canonical prime factorization** *of n.*

Example: If $n = 246,960$, then n can be written as a product of primes in various ways:

$$\begin{aligned} 246,960 &= 2 \cdot 3 \cdot 5 \cdot 7 \cdot 2 \cdot 3 \cdot 7 \cdot 2 \cdot 7 \cdot 2 \\ &= 7 \cdot 7 \cdot 7 \cdot 5 \cdot 3 \cdot 3 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \\ &= 7 \cdot 5 \cdot 3 \cdot 2 \cdot 7 \cdot 3 \cdot 2 \cdot 7 \cdot 2 \cdot 2 \\ &= 2 \cdot 2 \cdot 2 \cdot 2 \cdot 3 \cdot 3 \cdot 5 \cdot 7 \cdot 7 \cdot 7. \end{aligned}$$

The theorem says that although we can rearrange the *order* of the product any way we like, every expression of $246,960$ as a product of primes will involve the primes 2, 3, 5, 7, and no others, and there will be four factors of 2, two factors of 3, one factor of 5, and three factors of 7. Ordering the primes and writing in terms of powers as in the corollary, we obtain the *canonical* prime factorization

$$246,960 = 2^4 \cdot 3^2 \cdot 5 \cdot 7^3,$$

which is unique. □

§2.5 Exercises

1. Use the Prime Test (Theorem 2.5.4) to determine which of the following integers are prime.
   (a) 157
   (b) 239
   (c) 513
   (d) 667
   (e) 2003
2. Show that a 2-digit number n is prime if and only if n is *not* divisible by 2, 3, 5, or 7.
3. Show that if n is a 3-digit composite number then n has a prime divisor p with $p \leq 31$.
4. Use the Sieve of Eratosthenes to find all primes less than 100.
5. Find the canonical prime factorizations of the following integers.
   (a) 338
   (b) 1547
   (c) 2700
6. Prove the following.
   (a) If n is a perfect square, then every exponent in the canonical prime factorization of n is *even*.
   (b) If every exponent in the canonical prime factorization of n is *even*, then n is a perfect square.
7. Without doing any calculations, explain why a right triangle cannot have sides with lengths 2, 3, and 3.6. Could there be a right triangle with sides of lengths 2, 3, and 3.605551275? Explain.
8. Show that if a right triangle has two sides of integer length, then the length of the third side is either an integer or is irrational.

2.6 Prime Factorizations and Divisibility

Work through the following problems prior to reading further.

Class Preparation Problems:

1. Find the prime factorizations of all positive divisors of $72 = 2^3 \cdot 3^2$ (namely, 1, 2, 3, 4, 6, 8, 9, 12, 18, 24, 36, and 72).
2. Find the prime factorizations of some non-divisors of 72. For example, find the prime factorizations of 16, 27, 32, and 54, which are all divisible by the same primes as 72, and the prime factorizations of 15 and 20.
3. Compare the exponents in the prime factorizations of divisors and non-divisors of 72 with those in the factorization of 72.
4. Suppose $a = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ and $b = p_1^{b_1} p_2^{b_2} \cdots p_r^{b_r}$, where $p_1, p_2, \ldots, p_r$ are distinct primes. Use the examples above to make a conjecture about the conditions on the exponents that are required in order that $a \mid b$.
5. How many positive divisors does 72 have? Write down all divisors of 180 with their prime factorizations. How many are there? Can you use these examples to guess a formula for the number of divisors of an integer in terms of its prime factorization?
6. What is meant by the least common multiple of two integers a and b?
7. Do you know or can you derive a formula for the greatest common divisor and least common multiple of a and b in terms of the prime factorizations of a and b? Why does the formula work?
8. How are the GCD and LCM of a and b related? Why?

In the canonical prime factorization of an integer, we require that all of the exponents be positive, otherwise the factorization would not be unique. For example, $12 = 2^2 \cdot 3 = 2^2 \cdot 3 \cdot 5^0 = 2^2 \cdot 3 \cdot 7^0 \cdot 11^0$. When comparing two integers, however, it is often useful to allow the exponents to be zero so both integers can be written as products of powers of the same primes.

Proposition 2.6.1 (Prime Factorizations for Comparison) *Let a and b be positive integers and let $p_1 < p_2 < \cdots < p_r$ be all of the primes dividing a, b, or both. We can then write $a = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ and $b = p_1^{b_1} p_2^{b_2} \cdots p_r^{b_r}$, with $a_i \geqslant 0$ and $b_i \geqslant 0$ for all i.*

Example: The canonical prime factorizations of $a = 4116$ and $b = 94864$ are

$$a = 4116 = 2^2 \cdot 3 \cdot 7^3 \text{ and } b = 94864 = 2^4 \cdot 7^2 \cdot 11^2.$$

Thus the distinct primes dividing a or b are 2, 3, 7, and 11. Notice that a does not have a factor of 11 and b does not have a factor of 3. The Prime Factorizations for Comparison are then

$$a = 4116 = 2^2 \cdot 3 \cdot 7^3 \cdot 11^0 \text{ and } b = 94864 = 2^4 \cdot 3^0 \cdot 7^2 \cdot 11^2.$$

□

We can use this to characterize divisibility in terms of prime factorizations.

Theorem 2.6.2 *Let $a = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ and $b = p_1^{b_1} p_2^{b_2} \cdots p_r^{b_r}$, with the p_i distinct primes and $a_i \geqslant 0$, $b_i \geqslant 0$ for all i. Then $a \mid b$ if and only if $a_i \leqslant b_i$ for all i.*

Proof. Suppose first that $a \mid b$, so that $b = na$ for some integer n. If p is any prime divisor of n, then $p \mid n$ and $n \mid b$, hence $p \mid b$ by Theorem 2.2.2 (vii). Therefore, every prime divisor of n is one of the p_i, and we can write $n = p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$, where $n_i \geqslant 0$ for each i.

Writing $b = na$ in terms of the prime factorizations, we have

$$\begin{aligned} p_1^{b_1} p_2^{b_2} \cdots p_r^{b_r} &= (p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r})(p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}) \\ &= p_1^{n_1+a_1} p_2^{n_2+a_2} \cdots p_r^{n_r+a_r}, \end{aligned}$$

and so $b_i = n_i + a_i$ for each i, by uniqueness of factorization (see Theorem 2.5.11). Since $n_i \geqslant 0$ for each i, we have $a_i \leqslant n_i + a_i = b_i$ for each i.

Conversely, suppose $a_i \leqslant b_i$, so that $b_i - a_i \geqslant 0$, for each i. Thus $m = p_1^{b_1-a_1} p_2^{b_2-a_2} \cdots p_r^{b_r-a_r}$ is an integer. We then have

$$\begin{aligned} b &= p_1^{b_1} p_2^{b_2} \cdots p_r^{b_r} \\ &= (p_1^{b_1-a_1} p_2^{b_2-a_2} \cdots p_r^{b_r-a_r})(p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}) \\ &= ma. \end{aligned}$$

Hence $b = ma$ and $m \in \mathbb{Z}$, and so $a \mid b$. □

By the theorem, the positive divisors of $b = p_1^{b_1} p_2^{b_2} \cdots p_r^{b_r}$ are all possible integers of the form $a = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$, where for each i, $0 \leqslant a_i \leqslant b_i$.

Example: The positive divisors of $540 = 2^2 \cdot 3^3 \cdot 5$ are the integers of the form $2^i \cdot 3^j \cdot 5^k$ with $0 \leqslant i \leqslant 2$, $0 \leqslant j \leqslant 3$, and $0 \leqslant k \leqslant 1$. These are as follows:

$$\begin{array}{rrrr} 1 = 2^0 \cdot 3^0 \cdot 5^0 & 3 = 2^0 \cdot 3^1 \cdot 5^0 & 9 = 2^0 \cdot 3^2 \cdot 5^0 & 27 = 2^0 \cdot 3^3 \cdot 5^0 \\ 2 = 2^1 \cdot 3^0 \cdot 5^0 & 6 = 2^1 \cdot 3^1 \cdot 5^0 & 18 = 2^1 \cdot 3^2 \cdot 5^0 & 54 = 2^1 \cdot 3^3 \cdot 5^0 \\ 4 = 2^2 \cdot 3^0 \cdot 5^0 & 12 = 2^2 \cdot 3^1 \cdot 5^0 & 36 = 2^2 \cdot 3^2 \cdot 5^0 & 108 = 2^2 \cdot 3^3 \cdot 5^0 \\ & & & \\ 5 = 2^0 \cdot 3^0 \cdot 5^1 & 15 = 2^0 \cdot 3^1 \cdot 5^1 & 45 = 2^0 \cdot 3^2 \cdot 5^1 & 135 = 2^0 \cdot 3^3 \cdot 5^1 \\ 10 = 2^1 \cdot 3^0 \cdot 5^1 & 30 = 2^1 \cdot 3^1 \cdot 5^1 & 90 = 2^1 \cdot 3^2 \cdot 5^1 & 270 = 2^1 \cdot 3^3 \cdot 5^1 \\ 20 = 2^2 \cdot 3^0 \cdot 5^1 & 60 = 2^2 \cdot 3^1 \cdot 5^1 & 180 = 2^2 \cdot 3^2 \cdot 5^1 & 540 = 2^2 \cdot 3^3 \cdot 5^1 \end{array}$$

The total number of positive divisors is obtained by multiplying the number of choices for each of i, j, and k, hence is $3 \cdot 4 \cdot 2 = 24$. □

The number of divisors of b depends only on the exponents in the prime factorization of b. Counting the number of possible combinations of exponents yields the following result.

Theorem 2.6.3 *If $b = p_1^{b_1} p_2^{b_2} \cdots p_r^{b_r}$ is the canonical prime factorization of b, then the number of positive divisors of b is $(b_1 + 1)(b_2 + 1) \cdots (b_r + 1)$.*

Proof. By Theorem 2.6.2, the positive divisors of b are all possible integers of the form

$$a = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r},$$

where for each i, $0 \leqslant a_i \leqslant b_i$. There are then $b_i + 1$ choices for the exponent a_i. The number of positive divisors is the the total number of possible combinations of exponents, which is the product of the number of choices for each exponent, i.e., $(b_1 + 1)(b_2 + 1) \cdots (b_r + 1)$. □

Note that this works even with the more general prime factorization above, since if $b_i = 0$ for some i, then $b_i + 1 = 1$ and multiplying the number of divisors by $b_i + 1$ has no effect.

A concept closely related to the greatest common divisor is the least common multiple.

Definition 2.6.4 *Let a and b be non-zero integers. The* **least common multiple** *of a and b is the positive integer m, denoted $m = [a, b]$, satisfying*

i. *$a \mid m$ and $b \mid m$,*

ii. *if $a \mid c$ and $b \mid c$ with $c > 0$, then $m \leqslant c$.*

Because any negative common multiple of a and b will be less than any positive one, we need to include the condition that m is positive and to include $c > 0$ in part (ii) of the definition.

Using the characterization of divisibility above, we can express both the GCD and LCM of a and b in terms of prime factorizations.

Theorem 2.6.5 *Let $a = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ and $b = p_1^{b_1} p_2^{b_2} \cdots p_r^{b_r}$, with the p_i distinct primes and $a_i \geqslant 0$, $b_i \geqslant 0$ for all i. Then*

a. *$(a, b) = p_1^{d_1} p_2^{d_2} \cdots p_r^{d_r}$, where $d_i = \min\{a_i, b_i\}$ for all i, and*

b. *$[a, b] = p_1^{m_1} p_2^{m_2} \cdots p_r^{m_r}$, where $m_i = \max\{a_i, b_i\}$ for all i.*

Proof. GCD: Let $d = p_1^{d_1} p_2^{d_2} \cdots p_r^{d_r}$, where $d_i = \min\{a_i, b_i\}$ for each i. We show that $d = (a, b)$ by showing that d satisfies parts (i) and (ii) of the definition of GCD (Definition 2.3.2).

(i) Since $d_i = \min\{a_i, b_i\}$, we have $d_i \leqslant a_i$ and $d_i \leqslant b_i$ for all i. Therefore, it follows from Theorem 2.6.2 that $d \mid a$ and $d \mid b$, and so (i) holds.

(ii) Let c be a positive integer satisfying $c \mid a$ and $c \mid b$. If p is any prime divisor of c, then by Theorem 2.2.2 (vii), p must be one of the p_i. Thus we can write $c = p_1^{c_1} p_2^{c_2} \cdots p_r^{c_r}$, where $c_i \geqslant 0$ for each i. Since $c \mid a$ and $c \mid b$, Theorem 2.6.2 implies that $c_i \leqslant a_i$ and $c_i \leqslant b_i$ for all i. Hence

$$c_i \leqslant \min\{a_i, b_i\} = d_i$$

for each i, and so $c \mid d$, again by Theorem 2.6.2. Thus $c \leqslant d$ and (ii) holds.

LCM: Let $m = p_1^{m_1} p_2^{m_2} \cdots p_r^{m_r}$, where $m_i = \max\{a_i, b_i\}$ for each i. We show that $m = [a, b]$ by showing that m satisfies parts (i) and (ii) of the definition of LCM (Definition 2.6.4).

(i) Since $m_i = \max\{a_i, b_i\}$, we have $m_i \geqslant a_i$ and $m_i \geqslant b_i$ for all i. Therefore, it follows from Theorem 2.6.2 that $a \mid m$ and $b \mid m$, and so (i) holds.

(ii) Let c be a positive integer satisfying $a \mid c$ and $b \mid c$. Each prime divisor of a or b will also divide c, but c may have other prime divisors as well. Thus we can write $c = p_1^{c_1} p_2^{c_2} \cdots p_r^{c_r} \cdot k$, where each $c_i \geqslant 0$ and k is an integer such that $p_i \nmid k$ for each i.

Since $a \mid c$ and $b \mid c$, Theorem 2.6.2 implies that $a_i \leqslant c_i$ and $b_i \leqslant c_i$ for all i. Hence

$$m_i = \max\{a_i, b_i\} \leqslant c_i$$

for each i. It follows that $m \mid p_1^{c_1} p_2^{c_2} \cdots p_r^{c_r}$, again by Theorem 2.6.2, and therefore that $m \mid c$, since $p_1^{c_1} p_2^{c_2} \cdots p_r^{c_r} \mid c$. Thus $m \leqslant c$, as $c > 0$, and so (ii) holds. □

Example: Let $a = 1815156 = 2^2 \cdot 3^3 \cdot 7^5$ and $b = 3600 = 2^4 \cdot 3^2 \cdot 5^2$. Writing a and b in terms of the same primes, we have

$$a = 2^2 \cdot 3^3 \cdot 5^0 \cdot 7^5$$
$$b = 2^4 \cdot 3^2 \cdot 5^2 \cdot 7^0$$

Hence, by the theorem,

$$(a, b) = 2^2 \cdot 3^2 \cdot 5^0 \cdot 7^0$$
$$[a, b] = 2^4 \cdot 3^3 \cdot 5^2 \cdot 7^5$$

and of course the GCD can be rewritten as $(a, b) = 2^2 \cdot 3^2$. Notice also that

$$\begin{aligned}
(a, b)[a, b] &= (2^2 \cdot 3^2 \cdot 5^0 \cdot 7^0)(2^4 \cdot 3^3 \cdot 5^2 \cdot 7^5) \\
&= 2^{2+4} \cdot 3^{2+3} \cdot 5^{0+2} \cdot 7^{0+5} \\
&= 2^{2+4} \cdot 3^{3+2} \cdot 5^{0+2} \cdot 7^{5+0} \\
&= (2^2 \cdot 3^3 \cdot 5^0 \cdot 7^5)(2^4 \cdot 3^2 \cdot 5^2 \cdot 7^0) \\
&= a \cdot b,
\end{aligned}$$

and so $(a, b)[a, b] = a \cdot b$. □

We showed before that the GCD of two integers is divisible by any common divisor. Similarly, any common multiple of two integers is a multiple of the LCM.

Corollary 2.6.6 *Let a and b be non-zero integers and let $m = [a, b]$. If c is any integer satisfying $a \mid c$ and $b \mid c$, then $m \mid c$.*

Proof. This is shown in part (ii) for the LCM in the proof of Theorem 2.6.5. □

The characterization of the GCD and LCM in Theorem 2.6.5 implies a very important relationship between the two numbers, as suggested in the example above.

Corollary 2.6.7 *If a and b are positive integers, then $(a, b) \cdot [a, b] = a \cdot b$.*

Proof. We use the notation of Theorem 2.6.5. For each i, $d_i = \min\{a_i, b_i\}$ and $m_i = \max\{a_i, b_i\}$. Hence d_i is one of a_i or b_i and m_i is the other. In any case, we have $a_i + b_i = d_i + m_i$. Therefore,

$$\begin{aligned}
(a, b) \cdot [a, b] &= (p_1^{d_1} p_2^{d_2} \cdots p_r^{d_r})(p_1^{m_1} p_2^{m_2} \cdots p_r^{m_r}) \\
&= p_1^{d_1+m_1} p_2^{d_2+m_2} \cdots p_r^{d_r+m_r} \\
&= p_1^{a_1+b_1} p_2^{a_2+b_2} \cdots p_r^{a_r+b_r} \\
&= (p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r})(p_1^{b_1} p_2^{b_2} \cdots p_r^{b_r}) \\
&= a \cdot b,
\end{aligned}$$

and so $(a, b) \cdot [a, b] = a \cdot b$. □

Note that the corollary is equivalent to the statement that

$$[a, b] = \frac{a \cdot b}{(a, b)}.$$

Therefore, it is not necessary to know the prime factorizations of a and b in order to compute the LCM. The Euclidean algorithm can be used to find (a, b), and then this formula allows us to calculate $[a, b]$ easily.

---

§2.6 Exercises

1. Let $n = 2^{14} \cdot 3^{23} \cdot 5^{35}$. Determine whether each of the following integers divides n. Explain your answers.
   (a) $a = 2^8 \cdot 3^{25} \cdot 5^2$
   (b) $b = 2^{13} \cdot 3^{22} \cdot 5^{34}$
   (c) $c = 3^{19} \cdot 5^2 \cdot 7$
2. Write out all positive divisors of $600 = 2^3 \cdot 3 \cdot 5^2$.
3. Determine the number of positive divisors of the following integers.
   (a) $145546856 = 2^3 \cdot 7^2 \cdot 13^5$
   (b) $5384464553 = 13^2 \cdot 17 \cdot 37^4$
4. Determine the number of positive divisors of the following integers.
   (a) 15125
   (b) 33750
5. For the following pairs of integers a, b, find (a, b) and $[a, b]$, and verify that $(a, b) \cdot [a, b] = a \cdot b$.
   (a) $a = 2^5 \cdot 5^3 \cdot 7^4 \cdot 13^8$, $b = 2^7 \cdot 3^4 \cdot 7^3 \cdot 11^{17}$
   (b) $a = 2^3 \cdot 3^2 \cdot 5^6 \cdot 7^4 \cdot 11$, $b = 2^2 \cdot 3^3 \cdot 5^5 \cdot 7^7 \cdot 17$
6. Given that $(16191, 8481) = 771$, find $[16191, 8481]$. *Do not* factor the integers. Explain your answer.
7. Given that $(25193, 46787) = 3599$, find $[25193, 46787]$. *Do not* factor the integers. Explain your answer.
8. Use the Euclidean algorithm to find the GCD and LCM of 963 and 657. (Do not factor.)
9. Use the Euclidean algorithm to find the GCD and LCM of 510 and 414. (Do not factor.)

2.7 Congruence

Recall that the Division Algorithm (Theorem 2.3.1) says that if a and b are integers with $a > 0$, then there are *unique* integers q and r with $b = qa + r$ and $0 \leqslant r < a$. This says that if we fix an integer $n > 0$, then every integer m can be written in exactly one way in the form $m = qn + r$, with $r = 0, 1, 2, \ldots, n-1$ (and q an integer).

Thus, for a fixed integer n, the remainder on division of m by n is uniquely determined by m. We can therefore classify all integers according to their remainder on division by n, putting two integers in the same category if they leave the same remainder on division by n.

For example, let $n = 3$. Every integer leaves a remainder of 0, 1, or 2 on division by 3. Hence every integer can be written in exactly one of the forms $3k$, $3k+1$, or $3k+2$, with k an integer.

Example: Show that every perfect square is of the form $4k$ or $4k+1$.
Proof. If n is a perfect square, then $n = m^2$ for some integer m. By the Division Algorithm, m is of one of the forms $4q$, $4q+1$, $4q+2$, or $4q+3$.

If m is of the form $4q$, then

$$n = m^2 = (4q)^2 = 16q^2 = 4(4q^2).$$

Since $k = 4q^2$ is an integer, n is of the form $4k$.

If m is of the form $4q+1$, then

$$n = m^2 = (4q+1)^2 = 16q^2 + 8q + 1 = 4(4q^2 + 2q) + 1.$$

Since $k = 4q^2 + 2q$ is an integer, n is of the form $4k+1$.

If m is of the form $4q+2$, then

$$n = m^2 = (4q+2)^2 = 16q^2 + 16q + 4 = 4(4q^2 + 4q + 1).$$

Since $k = 4q^2 + 4q + 1$ is an integer, n is of the form $4k+1$.

If m is of the form $4q+3$, then

$$n = m^2 = (4q+3)^2 = 16q^2 + 24q + 9 = (16q^2 + 24q + 8) + 1 = 4(4q^2 + 6q + 2) + 1.$$

Since $k = 4q^2 + 6q + 2$ is an integer, n is of the form $4k+1$.

Therefore, in all possible cases, n is of the form $4k$ or $4k+1$. □

Discussing integers "of the form $nk + r$" or "with remainder r on division by n" becomes cumbersome very quickly. The following result gives an easier characterization of this idea.

Proposition 2.7.1 *Let n be a positive integer. Two integers a and b leave the same remainder on division by n if and only if $n \mid a - b$.*

Proof. If a and b leave the same remainder on division by n, then $a = q_1 n + r$ and $b = q_2 n + r$, for some $q_1, q_2 \in \mathbb{Z}$. Thus

$$a - b = (q_1 n + r) - (q_2 n + r) = q_1 n + r - q_2 n - r = (q_1 - q_2)n,$$

and since $q_1 - q_2$ is an integer, $a - b = (q_1 - q_2)n$ implies $n \mid a - b$.

Conversely, suppose $n \mid a-b$. By the Division Algorithm, we have $a = q_1 n + r_1$ and $b = q_2 n + r_2$, for some $q_1, r_1, q_2, r_2 \in \mathbb{Z}$ with $0 \leqslant r_1 < n$ and $0 \leqslant r_2 < n$. Hence

$$a - b = (q_1 n + r_1) - (q_2 n + r_2) = q_1 n + r_1 - q_2 n - r_2 = (q_1 - q_2)n + (r_1 - r_2),$$

and so

$$r_1 - r_2 = (a - b) - (q_1 - q_2)n.$$

Since $n \mid a - b$ and $n \mid n$, the Combination Theorem implies $n \mid r_1 - r_2$, and so either $r_1 - r_2 = 0$ or $n \leqslant |r_1 - r_2|$ by Theorem 2.2.2 (iv). However, because $0 \leqslant r_1 < n$ and $0 \leqslant r_2 < n$, we have $0 \leqslant |r_1 - r_2| < n$, hence $r_1 - r_2 = 0$ and $r_1 = r_2$ as claimed. □

Definition 2.7.2 *Let n be a positive integer. We say integers a and b are* **congruent modulo** *n, and write $a \equiv b \pmod{n}$, if and only if $n \mid a - b$.*

The integer n in the definition is called the **modulus**, and will always be a positive integer. The concept of two numbers being congruent is only of interest if the numbers are integers. Thus whenever we use the notation $a \equiv b \pmod{n}$, it will always be assumed that n is a positive integer and that a and b are integers, even if not explicitly mentioned.

The proposition above says that $a \equiv b \pmod{n}$ if and only if a and b leave the same remainder on division by n. Moreover, we have the following equivalent conditions (the first being simply the definition):

$$\begin{aligned} a \equiv b \pmod{n} &\iff n \mid a - b \\ &\iff a \text{ and } b \text{ have the same remainder on division by } n \\ &\iff a \text{ and } b \text{ have the same "form" } qn + r \\ &\iff a = b + cn \text{ for some } c \in \mathbb{Z} \\ &\iff b = a + dn \text{ for some } d \in \mathbb{Z}. \end{aligned}$$

Example: $55 \equiv 29 \pmod{13}$ because $55 - 29 = 26 = 2 \cdot 13$. Observe also that

$$\begin{aligned} 55 &= 4(13) + 3 \\ 29 &= 2(13) + 3, \end{aligned}$$

so 55 and 29 have the same remainder, 3, on division by 13. We also have

$$\begin{aligned} 55 &= 29 + 2(13) \\ 29 &= 55 + (-2)(13). \end{aligned}$$

Notice that $55 \equiv 3 \pmod{13}$ and $29 \equiv 3 \pmod{13}$; that is, each integer is congruent modulo 13 to its remainder on division by 13. □

The definition of congruence and the Division Algorithm imply the following basic properties of congruence suggested in the example above.

Theorem 2.7.3 *If n is a positive integer and a is any integer, then the following hold:*

i. *There is a unique integer r, with $0 \leqslant r \leqslant n-1$, such that $a \equiv r \pmod{n}$.*

ii. *$a \equiv 0 \pmod{n}$ if and only if $n \mid a$.*

Proof. (i) By the Division Algorithm, we can write $a = qn + r$ for unique integers q and r with $0 \leqslant r \leqslant n-1$. Since $a - r = qn$, we have $n \mid a - r$ and $a \equiv r \pmod{n}$.

(ii) By definition, $a \equiv 0 \pmod{n}$ if and only if $n \mid a - 0$, that is, $n \mid a$. □

Definition 2.7.4 *The integer r with $0 \leqslant r \leqslant n-1$ and $a \equiv r \pmod{n}$ is called the* **least (non-negative) residue** *of a modulo n.*

Note that the least non-negative residue of a modulo n is simply the remainder on division of a by n. For example, 3 is the least non-negative residue of 55 (and of 29) modulo 13, as seen in the example above.

The next theorem says that congruence is an equivalence relation on the set of integers (see Definition 1.4.9).

Theorem 2.7.5 *If n is a positive integer and a, b, and c are any integers, then the following hold:*

i. *$a \equiv a \pmod{n}$.*

ii. *If $a \equiv b \pmod{n}$, then $b \equiv a \pmod{n}$.*

iii. *If $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $a \equiv c \pmod{n}$.*

Proof. (i) We have $a - a = 0$, and since $n \neq 0$, n divides 0. Hence $n \mid a - a$ and $a \equiv a \pmod{n}$.

(ii) If $a \equiv b \pmod{n}$, then $n \mid a - b$. Hence $n \mid -(a - b)$, that is, $n \mid b - a$, and so $b \equiv a \pmod{n}$.

(iii) If $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $n \mid a - b$ and $n \mid b - c$. Therefore, by the Combination Theorem, $n \mid (a - b) + (b - c)$, and since $(a - b) + (b - c) = a - c$, we have $n \mid a - c$. Hence $a \equiv c \pmod{n}$. □

Algebraic Properties of Congruence

Most of the manipulations that can be done with (integer) equations can also be done with congruences. Our next result says that we can add or subtract the same integer on both sides of a congruence, or multiply both sides of a congruence by the same integer.

Theorem 2.7.6 *If $a \equiv b \pmod{n}$, then for any integer c,*

i. *$a + c \equiv b + c \pmod{n}$*

ii. *$ac \equiv bc \pmod{n}$.*

Proof. This is an easy exercise using the definitions and also follows from Theorem 2.7.7 below. □

More generally, we can add, subtract, or multiply not just the same integer but also *congruent* integers on both sides of a congruence.

Theorem 2.7.7 *Let $a \equiv b \pmod{n}$. If $c \equiv d \pmod{n}$, then*

i. $a + c \equiv b + d \pmod{n}$

ii. $ac \equiv bd \pmod{n}$.

Proof. (i) Since $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, we have $n \mid a - b$ and $n \mid c - d$. By the Combination Theorem, this implies $n \mid (a-b)+(c-d)$, and since $(a-b)+(c-d) = (a+c)-(b+d)$, we also have $n \mid (a + c) - (b + d)$. Hence $a + c \equiv b + d \pmod{n}$ by definition of congruence.

(ii) Again, we know $n \mid a - b$ and $n \mid c - d$, and we need to show that $n \mid ac - bd$. Observe that

$$ac - bd = ac - bc + bc - bd = c(a - b) + b(c - d).$$

By the combination Theorem, $n \mid c(a - b) + b(c - d)$, hence $n \mid ac - bd$, and so $ac \equiv bd \pmod{n}$ by definition of congruence. □

This theorem, along with induction, also implies that we can raise both sides of a congruence to the same positive integer power.

Corollary 2.7.8 *If $a \equiv b \pmod{n}$, then $a^k \equiv b^k \pmod{n}$ for every positive integer k.*

Proof. The statement is obviously true if $k = 1$. Assume the statement is true for $k = \ell$; that is, if $a \equiv b \pmod{n}$, then $a^\ell \equiv b^\ell \pmod{n}$. By Theorem 2.7.7 (ii), these two congruences imply

$$a \cdot a^\ell \equiv b \cdot b^\ell \pmod{n},$$

or equvalently $a^{\ell+1} \equiv b^{\ell+1} \pmod{n}$. Hence if the statement is true for $k = \ell$, then it is true for $k = \ell + 1$, and therefore it is true for all $k \geqslant 1$ by the Principle of Mathematical Induction. □

Combining the results above, we obtain the following.

Corollary 2.7.9 *If $P(x) = c_n x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0$ is a polynomial in x with integer coefficients and $a \equiv b \pmod{n}$, then $P(a) \equiv P(b) \pmod{n}$, that is,*

$$c_n a^n + c_{n-1}a^{n-1} + \cdots + c_1 a + c_0 \equiv c_n b^n + c_{n-1}b^{n-1} + \cdots + c_1 b + c_0 \pmod{n}.$$

Remarks:

1. We cannot *divide* both sides of a congruence by the same integer, or cancel an integer from both sides, in general. For example, $2 \cdot 7 \equiv 2 \cdot 3 \pmod{8}$, but $7 \not\equiv 3 \pmod{8}$.

2. Similarly, we know that if $a \cdot b = 0$, then $a = 0$ or $b = 0$, but the analogous statement for congruences is false. For example, $3 \cdot 4 \equiv 0 \pmod{6}$, but $3 \not\equiv 0 \pmod{6}$ and $4 \not\equiv 0 \pmod{6}$.

The reason we cannot divide both sides of a congruence by an integer in general is that not every integer has a multiplicative inverse modulo n. As we noted when discussing the properties of our various number systems, division is actually multiplication by a multiplicative inverse.

Definition 2.7.10 *Let n be a positive integer. We say that an integer a has a* **multiplicative inverse modulo n** *if there exists an integer r such that $ra \equiv 1 \pmod{n}$.*

Suppose a is an integer with a multiplicative inverse r modulo n. In this case, we could multiply both sides of a congruence by r in order to "divide" by a. Using the fact that there are integers r and s such that $ar + ns = 1$ if and only if $(a, n) = 1$, we can characterize the integers that have multiplicative inverses modulo n.

Theorem 2.7.11 *Let n be a positive integer. The integer a has a multiplicative inverse modulo n if and only if $(a, n) = 1$.*

Proof. Suppose first that a has a multiplicative inverse r modulo n. We have $ra \equiv 1 \pmod{n}$, hence $n \mid ra - 1$. It follows that $ra - 1 = dn$ and $a(r) + n(-d) = 1$ for some integer d. Theorem 2.4.12 now implies $(a, n) = 1$.

Conversely, if $(a, n) = 1$, then Theorem 2.4.10 says there exist integers x and y such that $ax + ny = 1$. Hence $ax - 1 = (-y)n$, and so $n \mid ax - 1$. Therefore, $ax \equiv 1 \pmod{n}$ and x is a multiplicative inverse modulo n. □

Note that the multiplicative inverse of a modulo n, if it exists, is *not* unique. If r is an inverse for a and $r \equiv r' \pmod{n}$, then r' is also a multiplicative inverse for a modulo n because, by Theorem 2.7.7 (ii), $r'a \equiv ra \equiv 1 \pmod{n}$.

For small values of a and n, an inverse of a can usually be found easily by trial and error. For larger values, the Euclidean Algorithm can be used to find integers x and y so that $ax + ny = 1$, and then x is an inverse of a as in the proof of the theorem.

Example: Since $(9, 14) = 1$, there is a multiplicative inverse for 9 modulo 14.

We need to find an integer r such that $9r \equiv 1 \pmod{14}$, that is, such that $9r = m \cdot 14 + 1$ for some integer m. The first few integers of the form $m \cdot 14 + 1$ are 15, 29, 43, 57, 71, 85, and 99, and we observe that 99 is the first of these divisible by 9. We have $9 \cdot 11 = 99 = 7 \cdot 14 + 1$, hence $9 \cdot 11 \equiv 1 \pmod{14}$, and so 11 is a multiplicative inverse for 9 modulo 14. □

As noted above, it is the integers with multiplicative inverses modulo n that can be "cancelled" from both sides of a congruence. This is stated more formally in the next theorem. This theorem also implies that if r and r' are both multiplicative inverses for a modulo n, then $r \equiv r' \pmod{n}$. Thus, even though the multiplicative inverse is not unique, all of the inverses must be congruent modulo n.

Theorem 2.7.12 *Let n be a positive integer and let a, b, and c be integers. If $ab \equiv ac \pmod{n}$ and $(a, n) = 1$, then $b \equiv c \pmod{n}$.*

Proof. Let $ab \equiv ac \pmod{n}$ and let $(a, n) = 1$. By Theorem 2.7.11, there is a multiplicative inverse r for a. Hence $r(ab) \equiv r(ac) \pmod{n}$ by Theorem 2.7.7 (ii), so $(ra)b \equiv (ra)c \pmod{n}$, and so $1 \cdot b \equiv 1 \cdot c \pmod{n}$. Finally, this says $b \equiv c \pmod{n}$, as claimed. □

The theorem above can also be proved directly using Euclid's Lemma (Theorem 2.4.15). If $ab \equiv ac \pmod{n}$, then $n \mid ab - ac$, hence $n \mid a(b - c)$. By Eulcid's Lemma, if $(a, n) = 1$, then this implies $n \mid b - c$, and so $b \equiv c \pmod{n}$.

Similarly, the following corollary is simply Euclid's Lemma stated in the language of congruences, as $ab \equiv 0 \pmod{n}$ means $n \mid ab$ and $b \equiv 0 \pmod{n}$ means $n \mid b$.

Corollary 2.7.13 *Let n be a positive integer and let a and b be integers. If $ab \equiv 0 \pmod{n}$ and $(a, n) = 1$, then $b \equiv 0 \pmod{n}$.*

The next corollary is the restatement of Euclid's Lemma for Primes (Theorem 2.5.6) in the language of congruences. It says that if the modulus is a prime, then the situation in Remark 2 above cannot happen.

Corollary 2.7.14 *Let p be a prime number and let a and b be integers. If $ab \equiv 0 \pmod{p}$, then $a \equiv 0 \pmod{p}$ or $b \equiv 0 \pmod{p}$.*

Modular Arithmetic and the Ring $\mathbb{Z}_n$

You may be familiar with "modular arithmetic" or "clock arithmetic" modulo n using the set of integers $N = \{0, 1, \ldots, n-1\}$. This set is a complete set of residues mod n by Theorem 2.7.3; that is, every integer is congruent modulo n to exactly one integer in the set.

In modular arithmetic, we define addition and multiplication on N as follows. For a and b in N, $a+b$ is the unique element s of N such that $a+b \equiv s \pmod{n}$, and ab is the unique element m of N such that $ab \equiv m \pmod{n}$. This is sometimes called "clock arithmetic" because addition of times on a clock is simply addition modulo 12 (but with 12 replacing 0 in N, since $12 \equiv 0 \pmod{12}$).

It can be shown that that N is a commutative ring with 1 under the operations of addition and multiplication modulo n. The proof is rather tedious and somewhat tricky, however, due in part to the requirement that we reduce sums and products modulo n to get an element of N. This is similar to problems encountered in proving that the set $\mathbb{Q}$ of rational numbers is a field, arising from the existence of many equivalent expressions of a given fraction. As with our formal construction of the rational numbers in §1.5, we can use equivalence classes to construct a ring algebraically equivalent (i.e., isomorphic) to N. An advantage of this alternate construction is that the proof that it yields a ring is much more straightforward.

Recall that Theorem 2.7.5 implies that congruence modulo n is an equivalence relation. The set of equivalence classes will be the underlying set for our ring construction.

Definition 2.7.15 *Let n be a positive integer. The* **congruence class** $\overline{a}$ *of an integer a modulo n is the equivalence class of a under the equivalence relation $\equiv$, thus*

$$\overline{a} = \{b \in \mathbb{Z} \mid a \equiv b \pmod{n}\}.$$

The set $\mathbb{Z}_n = \{\overline{a} \mid a \in \mathbb{Z}\}$ of congruence classes modulo n is called the set of **integers mod** n.

We must define the operations of addition and multiplication on the set $\mathbb{Z}_n$.

Definition 2.7.16 *For $\overline{a}$ and $\overline{b}$ in $\mathbb{Z}_n$ we define addition $(+)$ and multiplication $(\cdot)$ by*

$$\overline{a} + \overline{b} = \overline{a + b}$$

and

$$\overline{a} \cdot \overline{b} = \overline{a \cdot b}.$$

For example, if $n = 7$, then $\overline{4}+\overline{5} = \overline{9} = \overline{2}$ and $\overline{4}\cdot\overline{5} = \overline{20} = \overline{6}$. Note, however, that there are many different representatives for a given congruence class (see Proposition 1.4.11). The definitions of addition and multiplication seem to depend on the class representative. For $n = 7$, we have $\overline{4} = \overline{11}$ and $\overline{5} = \overline{19}$. If we use 11 and 19 as class representatives in place of 4 and 5, respectively, we get $\overline{11}+\overline{19} = \overline{30} = \overline{2}$ and $\overline{11}\cdot\overline{19} = \overline{209} = \overline{6}$. In this case, we get the same sum and product using either set of representatives.

In fact, this is true in any case. It follows from Theorem 2.7.7 that addition and multiplication are well-defined in $\mathbb{Z}_n$, as stated in the following result.

Proposition 2.7.17 *The operations of addition and multiplication in $\mathbb{Z}_n$ are well-defined. That is, if $\overline{a} = \overline{a'}$ and $\overline{b} = \overline{b'}$, then $\overline{a}+\overline{b} = \overline{a'}+\overline{b'}$ and $\overline{a}\cdot\overline{b} = \overline{a'}\cdot\overline{b'}$.*

Proof. If $\overline{a} = \overline{a'}$ and $\overline{b} = \overline{b'}$, then $a \equiv a' \pmod{n}$ and $b \equiv b' \pmod{n}$. By Theorem 2.7.7, we have $a + b \equiv a' + b' \pmod{n}$ and $a \cdot b \equiv a' \cdot b' \pmod{n}$ and therefore

$$\overline{a}+\overline{b} = \overline{a+b} = \overline{a'+b'} = \overline{a'}+\overline{b'}$$

and

$$\overline{a}\cdot\overline{b} = \overline{a\cdot b} = \overline{a'\cdot b'} = \overline{a'}\cdot\overline{b'},$$

as claimed. □.

Using the definitions of addition and multiplication in $\mathbb{Z}_n$ and the fact that $\mathbb{Z}$ is a commutative ring with 1, it is straightforward to verify the next theorem.

Theorem 2.7.18 *The set $\mathbb{Z}_n$ under the operations defined in Definition 2.7.16 satisfies properties* (i)–(ix) *and* (xi) *of Definition 1.3.1 and is therefore a commutative ring with* 1.

There is an interesting difference between the rings $\mathbb{Z}_n$ and $\mathbb{Z}$. As you are aware, for a and b in $\mathbb{Z}$, we have $ab = 0$ if and only if $a = 0$ or $b = 0$. This is not the case in $\mathbb{Z}_n$. For example, in $\mathbb{Z}_6$, we have $\overline{3} \neq \overline{0}$ and $\overline{4} \neq \overline{0}$, but $\overline{3}\cdot\overline{4} = \overline{12} = \overline{0}$.

In particular, this implies $\mathbb{Z}_6$ cannot be a field. If $\overline{3}$ had a multiplicative inverse $\overline{r}$, then we would have

$$\begin{aligned}
\overline{0} &= \overline{r}\cdot\overline{0} \\
&= \overline{r}\cdot(\overline{3}\cdot\overline{4}) \\
&= (\overline{r}\cdot\overline{3})\cdot\overline{4} \\
&= \overline{1}\cdot\overline{4} \\
&= \overline{4}
\end{aligned}$$

so that $\overline{4} = \overline{0}$, a contradiction.

More generally, if n is not a prime, we can write $n = ab$, where a and b are integers with $1 < a < n$ and $1 < b < n$. Thus $\overline{a} \neq \overline{0}$ and $\overline{b} \neq \overline{0}$, but $\overline{a}\cdot\overline{b} = \overline{n} = \overline{0}$. Hence, if n is not a prime, then $\mathbb{Z}_n$ is not a field.

This also follows from Theorem 2.7.11, which essentially describes the elements of $\mathbb{Z}_n$ that have multiplicative inverses.

Theorem 2.7.19 *Let n be a positive integer. An element $\overline{a}$ of $\mathbb{Z}_n$ has a multiplicative inverse (i.e., an element $\overline{r}$ with $\overline{r} \cdot \overline{a} = \overline{1}$) if and only if $(a, n) = 1$.*

For example, in $\mathbb{Z}_6 = \{\overline{0}, \overline{1}, \overline{2}, \overline{3}, \overline{4}, \overline{5}\}$, the elements with multiplicative inverses are $\overline{1}$ and $\overline{5}$. (What are their inverses?)

Now $\mathbb{Z}_n$ is a field if and only if every non-zero element has a multiplicative inverse, so if and only if every integer not divisible by n is relatively prime to n. This holds if and only if n is prime. We therefore have the following result.

Theorem 2.7.20 *Let n be a positive integer. The ring $\mathbb{Z}_n$ is a field if and only if n is a prime.*

This theorem gives us an infinite family of *finite* fields $\mathbb{Z}_p$, p a prime, in addition to the infinite fields $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$.

---

§2.7 Exercises

1. Find the least non-negative residue of the given integer (that is, find r with $0 \leqslant r \leqslant n-1$, so that $a \equiv r \pmod{n}$).

   (a) 46735, mod 7 (c) 56485, mod 4 (e) 65386, mod 19

   (b) 458, mod 37 (d) 11466, mod 21

2. Show that every perfect square is congruent to 0, 1, or 4 modulo 8.

3. Show that if a, b, and c are integers such that $a^2 + b^2 = c^2$, then at least one of a or b must be even. [Hint: Use the Example on page 67.]

   (Note: This also says that if a right triangle has sides of integer lengths, then at least one of the legs must be of *even* length.)

4. Show that if $a \equiv b \pmod{n}$ and $m \mid n$, then $a \equiv b \pmod{m}$.
   [Hint: Use the definition of congruence and basic properties of divisibility.]

5. Show that if $a \equiv b \pmod{n}$ and c is a positive integer, then $ca \equiv cb \pmod{cn}$.
   [Hint: Use the definitions of congruence and divisibility.]

6. By Corollary 2.7.8, if $a \equiv b \pmod{n}$, then $a^k \equiv b^k \pmod{n}$ for any positive integer k. Show that the converse is false by finding integers a, b, and n such that $a^2 \equiv b^2 \pmod{n}$ but $a \not\equiv b \pmod{n}$.

7. Use induction to show that $4^n \equiv 3n + 1 \pmod{9}$ for all $n \geqslant 0$.

8. Use induction to show that $2^{2^n} + 1 \equiv 5 \pmod{12}$ for all $n \geqslant 1$.

9. Let n be a positive integer, a an integer, and let $\overline{a}$ be the congruence class of a modulo n. Show that $\overline{a} = \{a + kn \mid k \in \mathbb{Z}\}$.

10. Show that $\mathbb{Z}_n$ has precisely n elements.

11. Determine which elements of $\mathbb{Z}_{10}$ have multiplicative inverses, and find the inverse of each.

2.8 Congruence and Divisibility Tests

Any non-negative integer m can be written in the form

$$m = d_k d_{k-1} \dots d_2 d_1 d_0,$$

where d_0 is the ones or units digit of m, d_1 is the tens digit, d_2 is the hundreds digit, and so forth. In this notation,

$$m = d_k d_{k-1} \dots d_2 d_1 d_0 = d_k \cdot 10^k + d_{k-1} \cdot 10^{k-1} + \cdots + d_2 \cdot 10^2 + d_1 \cdot 10^1 + d_0 \cdot 10^0.$$

We will refer to d_k as the first digit of m and d_0 as the last digit of m.

Example: The integer $m = 725986$ has $k = 6$ digits,

$$m = 7 \cdot 10^5 + 2 \cdot 10^4 + 5 \cdot 10^3 + 9 \cdot 10^2 + 8 \cdot 10^1 + 6 \cdot 10^0,$$

and we have

$$d_5 = 7,\ d_4 = 2,\ d_3 = 5,\ d_2 = 9,\ d_1 = 8, \text{ and } d_0 = 6,$$

in the notation above. □

Notation: In all of the results below, $m = d_k d_{k-1} \dots d_2 d_1 d_0$ denotes a positive integer and $d_k, d_{k-1}, \dots, d_1, d_0$ are the digits of m.

There are nice shortcuts for finding the least residue of an integer modulo some small integers, and therefore for determining when an integer is divisible by these small integers. The corollaries below give tests for divisibility by 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, and 13.

Congruence and Divisibility by Powers of 2, 5, and 10

We first consider congruence and divisibility by powers of 2, 5, and 10. The tests for all of these are based on the fact that 2, 5, and 10 all divide 10, so that $2^n \mid 10^\ell$, $5^n \mid 10^\ell$, and $10^n \mid 10^\ell$ for all $\ell \geqslant n$. The idea behind the proofs of the tests is demonstrated in the following examples.

Examples: Let $m = 578551$ and find the least non-negative residue of m modulo 4, modulo 125, and modulo 10000.

1. Residue modulo $4 = 2^2$: Observe that $2^2 \mid 10^2$, i.e., $4 \mid 100$, so that $100 \equiv 0 \pmod 4$. Hence

$$\begin{aligned} m &= (5785)(100) + 51 \\ &\equiv 5785 \cdot 0 + 51 \pmod 4 \\ &\equiv 51 \pmod 4 \\ &\equiv 3 \pmod 4, \end{aligned}$$

since $51 = 12 \cdot 4 + 3$. Therefore, the least non-negative residue of m modulo 4 is 3.

2. Residue modulo $125 = 5^3$: Observe that $5^3 \mid 10^3$, i.e., $125 \mid 1000$, so that $1000 \equiv 0 \pmod{125}$. Hence

$$\begin{aligned} m &= (578)(1000) + 551 \\ &\equiv 578 \cdot 0 + 551 \pmod{125} \\ &\equiv 551 \pmod{125} \\ &\equiv 51 \pmod{125}, \end{aligned}$$

since $551 = 4 \cdot 125 + 51$. Therefore, the least non-negative residue of m modulo 125 is 51.

3. Residue modulo $10000 = 10^4$: Obviously, $10^4 \mid 10^4$, so that $10000 \equiv 0 \pmod{10000}$. Hence

$$\begin{aligned} m &= (57)(10000) + 8551 \\ &\equiv 57 \cdot 0 + 8551 \pmod{10000} \\ &\equiv 8551 \pmod{10000}. \end{aligned}$$

Therefore, the least non-negative residue of m modulo 10000 is 8551. □

The first theorem says that m is congruent modulo 2^n, 5^n, and 10^n to the number made up of the last n digits of m.

Theorem 2.8.1 *If $m = d_k d_{k-1} \ldots d_2 d_1 d_0$ and n is a positive integer, then the following hold:*

i. $m \equiv d_{n-1} d_{n-2} \ldots d_2 d_1 d_0 \pmod{2^n}$,

ii. $m \equiv d_{n-1} d_{n-2} \ldots d_2 d_1 d_0 \pmod{5^n}$,

iii. $m \equiv d_{n-1} d_{n-2} \ldots d_2 d_1 d_0 \pmod{10^n}$.

Proof. Observe that since $2 \cdot 5 = 10$, we have $2^n \cdot 5^n = 10^n$, and moreover for $\ell \geqslant n$,

$$10^\ell = 2^\ell \cdot 5^\ell = (2^n \cdot 5^n)(2^{\ell-n} \cdot 5^{\ell-n}).$$

Since $2^{\ell-n} \cdot 5^{\ell-n}$ is an integer, this implies 2^n, 5^n, and 10^n all divide 10^ℓ, and so 10^ℓ is congruent to 0 modulo each of 2^n, 5^n, and 10^n, for all $\ell \geqslant n$. Therefore, if Z is any one of 2^n, 5^n, or 10^n, then

$$\begin{aligned} m &= d_k d_{k-1} \ldots d_2 d_1 d_0 \\ &= d_k \cdot 10^k + d_{k-1} \cdot 10^{k-1} + \cdots + d_n \cdot 10^n + d_{n-1} \cdot 10^{n-1} + \cdots + d_1 \cdot 10^1 + d_0 \cdot 10^0 \\ &\equiv d_k \cdot 0 + d_{k-1} \cdot 0 + \cdots + d_n \cdot 0 + d_{n-1} \cdot 10^{n-1} + \cdots + d_1 \cdot 10^1 + d_0 \cdot 10^0 \pmod{Z} \\ &\equiv 0 + 0 + \cdots + 0 + d_{n-1} \cdot 10^{n-1} + \cdots + d_1 \cdot 10^1 + d_0 \cdot 10^0 \pmod{Z} \\ &\equiv d_{n-1} \cdot 10^{n-1} + \cdots + d_1 \cdot 10^1 + d_0 \cdot 10^0 \pmod{Z} \\ &\equiv d_{n-1} d_{n-2} \ldots d_2 d_1 d_0 \pmod{Z}. \end{aligned}$$

Hence $m \equiv d_{n-1} d_{n-2} \ldots d_2 d_1 d_0 \pmod{Z}$ for $Z = 2^n$, $Z = 5^n$, or $Z = 10^n$, as claimed. □

In particular, notice that since $0 \leqslant d_{n-1} d_{n-2} \ldots d_2 d_1 d_0 < 10^n$, this number made up of the last n digits of m is precisely the least non-negative residue of m modulo 10^n.

Examples: The Theorem says that for $m = 578551$ we know immediately that

$$\begin{aligned} m &\equiv 51 \,(\text{mod } 2^2), \\ m &\equiv 551 \,(\text{mod } 5^3), \\ m &\equiv 8551 \,(\text{mod } 10^4), \end{aligned}$$

without having to do the initial computations we did in the examples above. □

Recalling that $Z \mid m$ if and only if $m \equiv 0 \,(\text{mod } Z)$, we have the following corollary.

Corollary 2.8.2 *Let m and n be positive integers and let Z be one of 2^n, 5^n, or 10^n. Then $Z \mid m$ if and only if Z divides the number made up of the last n digits of m.*

We obtain divisibility tests for 2, 5, and 10 by taking $n = 1$ in the corollary above.

Corollary 2.8.3 (Divisibility Tests for 2, 5, and 10)

If $m = d_k d_{k-1} \ldots d_2 d_1 d_0$, then the following hold:

i. $2 \mid m$ *if and only if* d_0 *is even,*

ii. $5 \mid m$ *if and only if* $d_0 = 0$ *or* $d_0 = 5$,

iii. $10 \mid m$ *if and only if* $d_0 = 0$.

Proof. By Corollary 2.8.2, one of 2, 5, or 10 will divide m if and only if it divides the last digit, d_0, of m. Thus $2 \mid m$ if and only if d_0 is divisible by 2, i.e., is even. Hence (i) holds. Since 0 and 5 are the only one-digit numbers divisible by 5, and 0 is the only one-digit number divisible by 10, statements (ii) and (iii) hold. □

More generally, observe that Corollary 2.8.2 implies 10^n divides m if and only if the last n digits of m are all 0.

Divisibility tests for 4 and 8 follow from Corollary 2.8.2 by taking $n = 2$ and $n = 3$, respectively.

Corollary 2.8.4 (Divisibility Tests for 4 and 8)

If $m = d_k d_{k-1} \ldots d_2 d_1 d_0$, then the following hold:

i. $4 \mid m$ *if and only if* $4 \mid d_1 d_0$,

ii. $8 \mid m$ *if and only if* $8 \mid d_2 d_1 d_0$.

Examples:

1. Since $4 \nmid 54$, we have that $4 \nmid 1875625654$.
2. Since $8 \mid 744$ (check!), we have that $8 \mid 917863265744$.
3. Since $625 \mid 4375$ (check!) and $625 = 5^4$, we have that $625 \mid 17423591\underline{4375}$. □

Congruence and Divisibility by 3 and 9

The proofs of the tests for congruence modulo 3 and 9 use the fact that $10 \equiv 1 \pmod 3$ and $10 \equiv 1 \pmod 9$, hence $10^\ell \equiv 1 \pmod 3$ and $10^\ell \equiv 1 \pmod 9$ for all $\ell \geqslant 1$.

Theorem 2.8.5 *If $m = d_k d_{k-1} \dots d_2 d_1 d_0$, then*

$$m \equiv d_k + d_{k-1} + \cdots + d_2 + d_1 + d_0 \pmod 3$$

and

$$m \equiv d_k + d_{k-1} + \cdots + d_2 + d_1 + d_0 \pmod 9;$$

that is, m is congruent to the sum of its digits modulo 3 and modulo 9.

Proof. We have $10 \equiv 1 \pmod 3$ and $10 \equiv 1 \pmod 9$, hence by Corollary 2.7.8,

$$10^\ell \equiv 1^\ell \equiv 1 \pmod 3 \text{ and } 10^\ell \equiv 1^\ell \equiv 1 \pmod 9$$

for every positive integer ℓ. Hence if $Z = 3$ or $Z = 9$, then

$$\begin{aligned} m &= d_k d_{k-1} \dots d_2 d_1 d_0 \\ &= d_k \cdot 10^k + d_{k-1} \cdot 10^{k-1} + \cdots + d_2 \cdot 10^2 + d_1 \cdot 10^1 + d_0 \cdot 10^0 \\ &\equiv d_k \cdot 1 + d_{k-1} \cdot 1 + \cdots + d_2 \cdot 1 + d_1 \cdot 1 + d_0 \cdot 1 \pmod Z \\ &\equiv d_k + d_{k-1} + \cdots + d_2 + d_1 + d_0 \pmod Z, \end{aligned}$$

as claimed. □

Again recalling that $Z \mid m$ if and only if $m \equiv 0 \pmod Z$, we have the following corollary.

Corollary 2.8.6 (Divisibility Tests for 3 and 9)

If $m = d_k d_{k-1} \dots d_2 d_1 d_0$, then the following hold:

i. $3 \mid m$ *if and only if* $3 \mid d_k + d_{k-1} + \cdots + d_2 + d_1 + d_0$,

ii. $9 \mid m$ *if and only if* $9 \mid d_k + d_{k-1} + \cdots + d_2 + d_1 + d_0$.

In other words, 3 divides m if and only if 3 divides the sum of the digits of m, and 9 divides m if and only if 9 divides the sum of the digits of m.

Examples:

1. If $m = 7854623$, then

$$m \equiv 7 + 8 + 5 + 4 + 6 + 2 + 3 \equiv 35 \equiv 2 \pmod 3$$

and

$$m \equiv 7 + 8 + 5 + 4 + 6 + 2 + 3 \equiv 35 \equiv 8 \pmod 9.$$

In particular, $3 \nmid 7854623$ and $9 \nmid 7854623$.

2. If $m = 1748235$, then

$$m \equiv 1 + 7 + 4 + 8 + 2 + 3 + 5 \equiv 30 \equiv 0 \pmod 3$$

and

$$m \equiv 1 + 7 + 4 + 8 + 2 + 3 + 5 \equiv 30 \equiv 3 \pmod 9.$$

In particular, $3 \mid 30$, so $3 \mid 1748235$, but $9 \nmid 1748235$. □

Since $(2, 3) = 1$ and $(4, 3) = 1$, we can use Theorem 2.4.16 and combine the divisibility tests for 2 and 3 and for 4 and 3 to obtain divisibility tests for 6 and 12, respectively.

Corollary 2.8.7 (Divisibility Tests for 6 and 12)

If $m = d_k d_{k-1} \dots d_2 d_1 d_0$, then the following hold:

i. $6 \mid m$ *if and only if* d_0 *is even and* $3 \mid d_k + d_{k-1} + \cdots + d_2 + d_1 + d_0$,

ii. $12 \mid m$ *if and only if* $4 \mid d_1 d_0$ *and* $3 \mid d_k + d_{k-1} + \cdots + d_2 + d_1 + d_0$.

Proof. Since $(2, 3) = 1$, Theorem 2.4.16 says that $2 \cdot 3 \mid m$ if and only if $2 \mid m$ and $3 \mid m$. Similarly, since $(4, 3) = 1$, we have $4 \cdot 3 \mid m$ if and only if $4 \mid m$ and $3 \mid m$. Hence $6 \mid m$ if and only if m passes the divisibility tests for both 2 and 3, and $12 \mid m$ if and only if m passes the divisibility tests for both 4 and 3. □

Congruence and Divisibility by 11

The tests for congruence modulo 11 and divisibility by 11 are based on the fact that 10 is congruent to -1 modulo 11. We will require the following notation. For $m = d_k d_{k-1} \dots d_2 d_1 d_0$ a positive integer, denote

$$A = d_0 - d_1 + d_2 - d_3 + \cdots + (-1)^{k-1} d_{k-1} + (-1)^k d_k = \sum_{i \text{ even}} d_i - \sum_{j \text{ odd}} d_j.$$

Thus A is the alternating sum of the digits of m. Note that the signs alternate beginning with a *plus* for d_0 and a *minus* for d_1.

Theorem 2.8.8 *If $m = d_k d_{k-1} \dots d_2 d_1 d_0$, then $m \equiv A \pmod{11}$.*

Proof. We have $10 \equiv -1 \pmod{11}$, hence by Corollary 2.7.8, $10^\ell \equiv (-1)^\ell \pmod{11}$ for all $\ell \geqslant 0$. Thus $10^\ell \equiv 1 \pmod{11}$ if ℓ is even and $10^\ell \equiv -1 \pmod{11}$ if ℓ is odd. Therefore, we have

$$\begin{aligned} m &= d_k d_{k-1} \dots d_2 d_1 d_0 \\ &= d_k \cdot 10^k + d_{k-1} \cdot 10^{k-1} + \cdots + d_3 \cdot 10^3 + d_2 \cdot 10^2 + d_1 \cdot 10^1 + d_0 \cdot 10^0 \\ &\equiv d_k(-1)^k + d_{k-1}(-1)^{k-1} + \cdots + d_3(-1)^3 + d_2(-1)^2 + d_1(-1)^1 + d_0(-1)^0 \pmod Z \\ &\equiv (-1)^k d_k + (-1)^{k-1} d_{k-1} + \cdots - d_3 + d_2 - d_1 + d_0 \pmod Z \\ &\equiv A \pmod Z, \end{aligned}$$

as claimed. □

Recalling as before that $Z \mid m$ if and only if $m \equiv 0 \pmod{Z}$, we obtain a divisibility test for 11.

Corollary 2.8.9 (Divisibility Test for 11)
If $m = d_k d_{k-1} \ldots d_2 d_1 d_0$, then $11 \mid m$ if and only if $11 \mid A$.

When computing A to test for congruence modulo 11, it is essential to *add* the digits in the places corresponding to *even* powers of 10 and *subtract* those in the places corresponding to *odd* powers of 10. Adding and subtracting in the other order yields $-A$. While this is not relevant for the divisibility test (as $11 \mid -A$ if and only if $11 \mid A$), it is important in order to correctly determine the congruence of m modulo 11.

Examples:

1. If $m = 2\underline{3}8\underline{5}7\underline{2}1\underline{6}3\underline{4}$, (with digits corresponding to *even* powers of 10 underlined), then

$$\begin{aligned} A &= 4-3+6-1+2-7+5-8+3-2 \\ &= (3+5+2+6+4)-(2+8+7+1+3) \\ &= 20-21 \\ &= -1. \end{aligned}$$

   Hence $m \equiv -1 \equiv 10 \pmod{11}$, and so the least non-negative residue of 2385721634 mod 11 is 10. In particular, $11 \nmid 2385721634$.

2. If $m = \underline{2}9\underline{3}8\underline{2}7\underline{6}4\underline{4}$, (again with digits corresponding to *even* powers of 10 underlined), then

$$\begin{aligned} A &= 4-4+6-7+2-8+3-9+2 \\ &= (2+3+2+6+4)-(9+8+7+4) \\ &= 17-28 \\ &= -11. \end{aligned}$$

   Hence $A = -11$ and $11 \mid A$, and so $11 \mid 293827644$. □

Congruence and Divisibility by 7 and 13

The essential fact behind the tests for congruence modulo 7 and 13 is that $1001 = 7 \cdot 11 \cdot 13$, hence $10^3 \equiv -1 \pmod{7}$ and $10^3 \equiv -1 \pmod{13}$. We will require the following notation. For $m = d_k d_{k-1} \ldots d_2 d_1 d_0$ a positive integer, denote

$$\begin{aligned} T &= (100d_2 + 10d_1 + d_0) - (100d_5 + 10d_4 + d_3) + (100d_8 + 10d_7 + d_6) - \cdots \\ &= \sum_{i \text{ even}} (100d_{3i+2} + 10d_{3i+1} + d_{3i}) - \sum_{j \text{ odd}} (100d_{3j+2} + 10d_{3j+1} + d_{3j}) \end{aligned}$$

or, writing $100a + 10b + c$ as the three-digit number abc,

$$\begin{aligned} T &= (d_2 d_1 d_0) - (d_5 d_4 d_3) + (d_8 d_7 d_6) - \cdots \\ &= \sum_{i \text{ even}} d_{3i+2} d_{3i+1} d_{3i} - \sum_{j \text{ odd}} d_{3j+2} d_{3j+1} d_{3j} \end{aligned}$$

Thus T is the alternating sum of triples of digits of m. Note that $d_2 d_1 d_0$, for example, is a three-digit number and is *not* equal to $d_2 + d_1 + d_0$ or $d_2 \cdot d_1 \cdot d_0$.

Theorem 2.8.10 *If $m = d_k d_{k-1} \dots d_2 d_1 d_0$, then $m \equiv T \pmod 7$ and $m \equiv T \pmod{13}$.*

Proof. Since $10^3 = 7 \cdot 11 \cdot 13 - 1$, we have $10^3 \equiv -1 \pmod 7$ and $10^3 \equiv -1 \pmod{13}$. Hence, by Corollary 2.7.8, it follows that

$$10^{3\ell} \equiv (-1)^\ell \pmod 7 \text{ and } 10^{3\ell} \equiv (-1)^\ell \pmod{13}.$$

Let $Z = 7$ or $Z = 13$. we then have $10^{3\ell} \equiv 1 \pmod Z$ if ℓ is even and $10^{3\ell} \equiv -1 \pmod Z$ if ℓ is odd. Notice also that

$$d_{3\ell+2} \cdot 10^{3\ell+2} + d_{3\ell+1} \cdot 10^{3\ell+1} + d_{3\ell} \cdot 10^{3\ell} = (100 \cdot d_{3\ell+2} + 10 \cdot d_{3\ell+1} + d_{3\ell}) \cdot 10^{3\ell}.$$

By appending zeros to the left of the number m if necessary, we may assume that $k = 3r + 2$ for some non-negative integer r. We then have

$$\begin{aligned}
m &= d_k d_{k-1} \dots d_8 d_7 d_6 d_5 d_4 d_3 d_2 d_1 d_0 \\
&= d_{3r+2} d_{3r+1} d_{3r} \dots d_8 d_7 d_6 d_5 d_4 d_3 d_2 d_1 d_0 \\
&= d_{3r+2} \cdot 10^{3r+2} + d_{3r+1} \cdot 10^{3r+1} + d_{3r} \cdot 10^{3r} + \dots + d_8 \cdot 10^8 + d_7 \cdot 10^7 + d_6 \cdot 10^6 \\
&\qquad + d_5 \cdot 10^5 + d_4 \cdot 10^4 + d_3 \cdot 10^3 + d_2 \cdot 10^2 + d_1 \cdot 10^1 + d_0 \cdot 10^0 \\
&= (100 d_{3r+2} + 10 d_{3r+1} + d_{3r}) 10^{3r} + \dots + (100 d_8 + 10 d_7 + d_6) 10^6 \\
&\qquad + (100 d_5 + 10 d_4 + d_3) 10^3 + (100 d_2 + 10 d_1 + d_0) 10^0 \\
&\equiv (100 d_{3r+2} + 10 d_{3r+1} + d_{3r})(-1)^r + \dots + (100 d_8 + 10 d_7 + d_6)(-1)^2 \\
&\qquad + (100 d_5 + 10 d_4 + d_3)(-1)^1 + (100 d_2 + 10 d_1 + d_0)(-1)^0 \pmod Z \\
&\equiv (-1)^r (d_{3r+2} d_{3r+1} d_{3r}) + \dots + (d_8 d_7 d_6) - (d_5 d_4 d_3) + (d_2 d_1 d_0) \pmod Z \\
&\equiv T \pmod Z,
\end{aligned}$$

as claimed. □

Recalling as before that $Z \mid m$ if and only if $m \equiv 0 \pmod Z$, we obtain divisibility tests for 7 and 13.

Corollary 2.8.11 (Divisibility Tests for 7 and 13)

If $m = d_k d_{k-1} \dots d_2 d_1 d_0$, then the following hold:

i. $7 \mid m$ *if and only if* $7 \mid T$,

ii. $13 \mid m$ *if and only if* $13 \mid T$.

As with the computation of the alternating sum A, it is essential to compute T using the correct signs in order to apply the tests for congruence modulo 7 and 13. Note also that since $11 \mid 1001$, we also have $m \equiv T \pmod{11}$ and $11 \mid m$ if and only if $11 \mid T$. This test is much less convenient than the tests for 11 given in Theorem 2.8.8 and Corollary 2.8.9, however.

Example: Let $m = 23,587,614,934$. Since the number of digits is not divisible by 3, we can append a 0 on the left to obtain $m = 023,\underline{587},614,\underline{934}$, where the triples of digits that are *added* in the alternating sum are underlined. We have

$$\begin{aligned} T &= 934 - 614 + 587 - 23 \\ &= (587 + 934) - (23 + 614) \\ &= 1521 - 637 \\ &= 884. \end{aligned}$$

Hence $m \equiv 884 \equiv 2 \pmod 7$, since $884 = 126 \cdot 7 + 2$, and therefore the least non-negative residue of $23,587,614,934$ modulo 7 is 2. In particular, $7 \nmid 23,587,614,934$.

Also, $m \equiv 884 \pmod{13}$. Since $884 = 68 \cdot 13$, we have $13 \mid 884$, hence $m \equiv 0 \pmod{13}$ and $13 \mid 23,587,614,934$. □

It is also possible to combine tests we have derived to obtain divisibility tests for 14 and 15 (see the exercises), giving us divisibility tests for all positive integers up to 16.

---

§2.8 Exercises

On Exercises 1–4, use the theorems on congruence modulo 2^n, 3, 7, 9, 11, and 13 to find the least non-negative residues of the given integers. **Justify your answers.**

1. (a) $a = 476532189318$, mod 4
   (b) $b = 23765981235$, mod 8
   (c) $c = 351487629538$, mod 16
2. (a) $a = 472356734512$, mod 3
   (b) $b = 472356734512$, mod 9
   (c) $c = 324562783713$, mod 3
   (d) $d = 324562783713$, mod 9
3. (a) $a = 34781526247$, mod 11
   (b) $b = 123456789012$, mod 11
   (c) $c = 632475268196$, mod 11
4. (a) $a = 347815623107$, mod 7
   (b) $b = 347815623107$, mod 13
   (c) $c = 28473265918$, mod 7
   (d) $d = 28473265918$, mod 13

5. Use the divisibility tests for 2 and 7 to derive a test for divisibility by 14. Prove that the test is valid.

6. Use the divisibility tests for 3 and 5 to derive a test for divisibility by 15. Prove that the test is valid.

7. Use the divisibility tests for 2 and 9 to derive a test for divisibility by 18. Prove that the test is valid.

8. Show that a positive integer $m = d_k d_{k-1} \dots d_2 d_1 d_0$ is divisible by 20 if and only if the number $d_1 d_0$ made up of the last two digits of m is divisible by 20, or, equivalently, that $20 \mid m$ if and only if d_1 is even and $d_0 = 0$.

On Exercises 9–13, determine which, if any, of the given integers a, b, c are divisible by the indicated integer n. **Show your work and justify your answers.**

9. Determine if divisible by $n = 4$.

   (a) $a = 478563289358$
   (b) $b = 12354456724$
   (c) $c = 352148763376$

10. Determine if divisible by $n = 3$ and if divisible by $n = 9$.

    (a) $a = 21437856252$
    (b) $b = 54637281274$
    (c) $c = 42315768543$

11. Determine if divisible by $n = 6$.

    (a) $a = 47835624312$
    (b) $b = 65348127214$
    (c) $c = 27135248145$

12. Determine if divisible by $n = 11$.

    (a) $a = 41783526413$
    (b) $b = 615837429152$
    (c) $c = 724356712859$

13. Determine if divisible by $n = 7$ and if divisible by $n = 13$.

    (a) $a = 98239072918$
    (b) $b = 199885455861$
    (c) $c = 182443992562$

Chapter 3

Polynomials

In this chapter, we study the algebra of polynomials in a variable x with coefficients in a commutative ring S. We can define addition and multiplication of polynomials, and we will see that the set of polynomials has algebraic properties very similar to those of the integers. We will consider algebraic properties such as those in Definition 1.3.1, divisibility of polynomials and greatest common divisors as we did for integers in Chapter 2, as well as roots and factors of polynomials and irreducible polynomials, which are analogous to prime numbers.

3.1 Algebraic Properties of Polynomials

In this section, we will discuss polynomials with coefficients in some commutative ring S. For our purposes, the ring S will be one of $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{Z}_p$, where p is a prime. Unless otherwise stated, all definitions and results are valid in all five cases.

Definition 3.1.1 *A* **polynomial** *in the variable x with coefficients in the commutative ring S is an algebraic expression of the form*

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

where $n \geqslant 0$ is an integer and a_i is an element of S for all i.

The set of all polynomials in x with coefficients in S is denoted $S[x]$. We will need the following basic terminology.

Definition 3.1.2 *Let $p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial in $S[x]$.*

i. *The elements a_i are called the* **coefficients** *of $p(x)$.*

ii. *The monomials $a_i x^i$ are the* **terms** *of $p(x)$. The term with the highest power of x and non-zero coefficient is the* **leading term** *of $p(x)$, and a_0 is the* **constant term**.

iii. *The* **leading coefficient** *of $p(x)$ is the coefficient of the leading term of $p(x)$.*

iv. *We call $p(x)$ a* **monic polynomial** *if the leading coefficient of $p(x)$ is 1.*

v. *If $a_i = 0$ for all $i \geqslant 1$, so that $p(x) = a_0$, we call $p(x)$ a* **constant polynomial**. *By identifying an element a_0 of S with the constant polynomial $p(x) = a_0$, we will consider S to be a subset of $S[x]$.*

vi. *If $a_i = 0$ for all i, so that $p(x) = 0$, we call $p(x)$ the* **zero polynomial**.

vii. *If a_i is non-zero, we call i the* **degree** *of the monomial $a_i x^i$. If $p(x)$ is a non-zero polynomial, the* **degree** *of the polynomial $p(x)$ is defined to be the degree of the leading term of $p(x)$. We define the degree of the zero polynomial to be $-\infty$. We denote the degree of $p(x)$ by* $\deg p(x)$.

viii. *A polynomial of degree 1 is called a* **linear** *polynomial, a polynomial of degree 2 is a* **quadratic** *polynomial, and a polynomial of degree 3 is a* **cubic** *polynomial.*

Example: Let $p(x) = 3x^5 - 6x^4 + 7x^2 - 9x + 8$.

- The coefficients of $p(x)$ are $a_5 = 3$, $a_4 = -6$, $a_3 = 0$, $a_2 = 7$, $a_1 = -9$, and $a_0 = 8$.
- The leading term of $p(x)$ is $3x^5$ and the leading coefficient is 3.
  (Note that this is the case regardless of the order in which the terms are written. The leading term is the term with the highest power of x.)
- The degree of $p(x)$ is 5. □

Remark: By Definition 3.1.2 (vii), the degree of a *non-zero* constant polynomial is 0. We have defined the degree of the zero polynomial to be $-\infty$. Another approach used by some is to simply say that the zero polynomial has no degree. In any case, the degree of the zero polynomial *cannot* be defined to be 0 or any other integer. Otherwise, certain desirable properties of the degree will not be valid. Also, according to our definition, the zero polynomial has no leading term or leading coefficient.

In many ways, the set $S[x]$ is algebraically very similar to the set $\mathbb{Z}$ of integers. In particular, we can define addition and multiplication of polynomials and show that these operations satisfy the same basic algebraic properties in $S[x]$ as they do in $\mathbb{Z}$.

To compare or add two polynomials of different degrees, we can write both using all powers of x up to the higher degree, using 0 as a coefficient when necessary.

Definition 3.1.3 (Equality) *Two polynomials $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ and $q(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0$ in $S[x]$ are* **equal** *if and only if $a_i = b_i$ for all i.*

Definition 3.1.4 (Addition) *Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ and $q(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0$ be polynomials in $S[x]$. We define the* **sum** *$p(x) + q(x)$ by*

$$p(x) + q(x) = (a_n + b_n)x^n + (a_{n-1} + b_{n-1})x^{n-1} + \cdots + (a_1 + b_1)x + (a_0 + b_0).$$

Definition 3.1.5 (Multiplication) *Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ and $q(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$ be polynomials in $S[x]$. We define the* **product** *$p(x) \cdot q(x)$ by*

$$p(x) \cdot q(x) = c_r x^r + c_{r-1}x^{r-1} + \cdots + c_1 x + c_0,$$

where $r = n + m$ and, for each i,

$$c_i = a_i b_0 + a_{i-1}b_1 + \cdots + a_1 b_{i-1} + a_0 b_i.$$

Examples:

1. If $p(x) = 5x^3 + 4x^2 - 7x + 3$ and $q(x) = 6x^4 + 2x^3 + x - 9$, then we can write

$$p(x) = 0x^4 + 5x^3 + 4x^2 - 7x + 3,$$

$$q(x) = 6x^4 + 2x^3 + 0x^2 + 1x - 9,$$

and so

$$\begin{aligned} p(x) + q(x) &= (0+6)x^4 + (5+2)x^3 + (4+0)x^2 + (-7+1)x + (3-9) \\ &= 6x^4 + 7x^3 + 4x^2 - 6x - 6. \end{aligned}$$

Thus we add polynomials by simply adding corresponding coefficients.

2. For $a(x) = 2x^3 - 5x^2 + 7$ and $b(x) = 3x^2 + 2x - 4$, we have deg $a(x) = 3$ and deg $b(x) = 2$, hence deg $(a(x)b(x)) = 3 + 2 = 5$. The coefficients of $a(x)$ and $b(x)$ are

$$\begin{array}{ll} a_0 = 7 & b_0 = -4 \\ a_1 = 0 & b_1 = 2 \\ a_2 = -5 & b_2 = 3 \\ a_3 = 2 & b_3 = 0 \\ a_4 = 0 & b_4 = 0 \\ a_5 = 0 & b_5 = 0. \end{array}$$

The coefficients of the product $a(x)b(x)$ are therefore

$$\begin{aligned} c_0 &= a_0b_0 = 7 \cdot (-4) = -28 \\ c_1 &= a_1b_0 + a_0b_1 = 0 \cdot (-4) + 7 \cdot 2 = 14 \\ c_2 &= a_2b_0 + a_1b_1 + a_0b_2 = (-5) \cdot (-4) + 0 \cdot 2 + 7 \cdot 3 = 41 \\ c_3 &= a_3b_0 + a_2b_1 + a_1b_2 + a_0b_3 = 2 \cdot (-4) + (-5) \cdot 2 + 0 \cdot 3 + 7 \cdot 0 = -18 \\ c_4 &= a_4b_0 + a_3b_1 + a_2b_2 + a_1b_3 + a_0b_4 \\ &= 0 \cdot (-4) + 2 \cdot 2 + (-5) \cdot 3 + 0 \cdot 0 + 7 \cdot 0 = -11 \\ c_5 &= a_5b_0 + a_4b_1 + a_3b_2 + a_2b_3 + a_1b_4 + a_0b_5 \\ &= 0 \cdot (-4) + 0 \cdot 2 + 2 \cdot 3 + (-5) \cdot 0 + 0 \cdot 0 + 7 \cdot 0 = 6, \end{aligned}$$

and so

$$a(x)b(x) = c_5x^5 + c_4x^4 + c_3x^3 + c_2x^2 + c_1x + c_0 = 6x^5 - 11x^4 - 18x^3 + 41x^2 + 14x - 28.$$

This example is intended to illustrate the notation in the *formal* definition of multiplication. In practice, polynomials are multiplied by multiplying every term of the first by every term of the second (using laws of exponents and multiplying coefficients), and then collecting like terms. This procedure essentially uses a "distributive law" several times. □

The other algebraic properties of the integers are valid in $S[x]$ as well.

Theorem 3.1.6 (Polynomial Properties) *If $S = \mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{Z}_p$, p a prime, then $S[x]$ satisfies the following:*

Properties of Addition:

i. Closure under Addition: *$p(x) + q(x)$ is in $S[x]$ for all $p(x)$ and $q(x)$ in $S[x]$.*

ii. Associative Law of Addition:

$$p(x) + [q(x) + r(x)] = [p(x) + q(x)] + r(x)$$

for all $p(x)$, $q(x)$, and $r(x)$ in $S[x]$.

iii. Commutative Law of Addition: *$p(x) + q(x) = q(x) + p(x)$ for all $p(x)$ and $q(x)$ in $S[x]$.*

iv. Additive Identity: *There is a polynomial $Z(x)$ in $S[x]$ such that*

$$p(x) + Z(x) = Z(x) + p(x) = p(x)$$

for all $p(x)$ in $S[x]$.

v. Additive Inverses: *For each polynomial $p(x)$ in $S[x]$, there is a polynomial $-p(x)$ in $S[x]$ such that*

$$p(x) + (-p(x)) = (-p(x)) + p(x) = Z(x).$$

Properties of Multiplication:

vi. Closure under Multiplication: *$p(x) \cdot q(x)$ is in $S[x]$ for all $p(x)$ and $q(x)$ in $S[x]$.*

vii. Associative Law of Multiplication:

$$p(x) \cdot [q(x) \cdot r(x)] = [p(x) \cdot q(x)] \cdot r(x)$$

for all $p(x)$, $q(x)$, and $r(x)$ in $S[x]$.

viii. Commutative Law of Multiplication: *$p(x) \cdot q(x) = q(x) \cdot p(x)$ for all $p(x)$ and $q(x)$ in $S[x]$.*

ix. Multiplicative Identity: *There is a polynomial $I(x)$ in $S[x]$ such that*

$$p(x) \cdot I(x) = I(x) \cdot p(x) = p(x)$$

for all $p(x)$ in $S[x]$.

Property Relating Addition and Multiplication:

x. Distributive Law:

$$p(x) \cdot [q(x) + r(x)] = p(x) \cdot q(x) + p(x) \cdot r(x)$$

for all $p(x)$, $q(x)$, and $r(x)$ in $S[x]$.

Compare these properties to the properties satisfied by $\mathbb{Z}$ in the list in Definition 1.3.1. Note that $Z(x)$ is the zero polynomial, $Z(x) = 0$, and $I(x)$ is the constant polynomial 1, $I(x) = 1$. If $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, then the additive inverse of $p(x)$ is

$$-p(x) = (-1)p(x) = -a_n x^n - a_{n-1}x^{n-1} - \cdots - a_1 x - a_0.$$

Theorem 3.1.6 can be proved using the algebraic properties of S and the remarks above.

Example: Show that $p(x) + q(x) = q(x) + p(x)$ for all $p(x), q(x) \in S[x]$; that is, addition of polynomials is commutative.

Proof. Let $p(x) = a_n x^n + \cdots + a_1 x + a_0$ and $q(x) = b_n x^n + \cdots + b_1 x + b_0$. We have

$$\begin{aligned} p(x) + q(x) &= (a_n x^n + \cdots + a_1 x + a_0) + (b_n x^n + \cdots + b_1 x + b_0) \\ &= (a_n + b_n)x^n + \cdots + (a_1 + b_1)x + (a_0 + b_0) \text{ by Definition 3.1.4,} \\ &= (b_n + a_n)x^n + \cdots + (b_1 + a_1)x + (b_0 + a_0) \text{ by commutativity of addition in } S, \\ &= (b_n x^n + \cdots + b_1 x + b_0) + (a_n x^n + \cdots + a_1 x + a_0) \text{ by Definition 3.1.4,} \\ &= q(x) + p(x), \end{aligned}$$

and so $p(x) + q(x) = q(x) + p(x)$ as claimed. □

Corollary 3.1.7 *If $S = \mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{Z}_p$, p a prime, then $S[x]$ is a commutative ring with* 1.

Before considering the existence of multiplicative inverses, we prove some useful properties of the degrees of polynomials.

Theorem 3.1.8 *If $p(x)$, $q(x)$ are polynomials in $S[x]$, then* $\deg(p(x)q(x)) = \deg p(x) + \deg q(x)$.

Proof. If $p(x) = 0$ or $q(x) = 0$, then $p(x)q(x) = 0$. Thus either $\deg p(x) = -\infty$ or $\deg q(x) = -\infty$, and $\deg(p(x)q(x)) = -\infty$. We then have

$$\deg(p(x)q(x)) = -\infty = \deg p(x) + \deg q(x)$$

and the equation holds.

We may now assume $p(x)$ and $q(x)$ are both non-zero. Let $p(x) = a_n x^n + \cdots + a_1 x + a_0$, with $a_n \neq 0$, and let $q(x) = b_m x^m + \cdots + b_1 x + b_0$, with $b_m \neq 0$, so that $\deg p(x) = n$ and $\deg q(x) = m$. Notice that since S is one of $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{Z}_p$, where p is a prime, and $a_n \neq 0$, $b_m \neq 0$, we have $a_n b_m \neq 0$. Thus, by Definition 3.1.5, the leading term of $p(x)q(x)$ is $a_n b_m x^{n+m}$, and so

$$\deg(p(x)q(x)) = n + m = \deg p(x) + \deg q(x)$$

as claimed. □

Notice that this theorem is valid even if one of $p(x)$ or $q(x)$ is the zero polynomial. This would not be true had we attempted to define the degree of the zero polynomial to be 0 or any other integer.

The next results follow easily from the theorem, but can also be proved directly using the definition of multiplication as in the proof of the theorem.

Corollary 3.1.9 *If $p(x)$ and $q(x)$ are non-zero polynomials in $S[x]$, then $p(x) \cdot q(x)$ is non-zero.*

Proof. If $p(x)$ and $q(x)$ are non-zero, then $\deg\, p(x) = n$ and $\deg\, q(x) = m$ for some non-negative integers n and m. Thus $\deg\,(p(x)q(x)) = n + m$ is also a non-negative integer, and so $p(x)q(x)$ is non-zero. □

Note that the corollary also implies that if $p(x) \cdot q(x) = 0$, then $p(x) = 0$ or $q(x) = 0$.

Corollary 3.1.10 *Let S be one of $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{Z}_p$, where p is a prime. A polynomial $p(x)$ in $S[x]$ has a multiplicative inverse in $S[x]$ if and only if $p(x)$ is a non-zero constant polynomial.*

Proof. If $p(x) = a_0$ is a non-zero constant polynomial, then a_0 is in S. By hypothesis, S is a field, and so a_0 has a multiplicative inverse a_0^{-1} in S. Thus the constant polynomial $q(x) = a_0^{-1}$ is in $S[x]$, and $p(x)q(x) = a_0 a_0^{-1} = 1$. Since $I(x) = 1$ is the multiplicative identity element of $S[x]$, this means $q(x)$ is the multiplicative inverse of $p(x)$.

Conversely suppose $p(x) \in S[x]$ has a multiplicative inverse. This means $p(x)q(x) = 1$ for some $q(x) \in S[x]$, and so

$$0 = \deg\, 1 = \deg\,(p(x)q(x)) = \deg\, p(x) + \deg\, q(x).$$

In particular, neither $\deg\, p(x)$ nor $\deg\, q(x)$ can be $-\infty$, so both $p(x)$ and $q(x)$ are non-zero polynomials. Hence $\deg\, p(x)$ and $\deg\, q(x)$ are both non-negative integers and their sum is 0. It follows that $\deg\, p(x) = \deg\, q(x) = 0$, and so $p(x)$ is a constant polynomial. □

The degree of the sum of two polynomials is not as easy to specify as the degree of a product, in general. We can say the following, however.

Theorem 3.1.11 *If $p(x)$, $q(x)$ are polynomials in $S[x]$, then*

$$\deg\,(p(x) + q(x)) \leqslant \max\{\deg\, p(x), \deg\, q(x)\}.$$

Proof. Let $p(x) = a_n x^n + \cdots + a_1 x + a_0$ and $q(x) = b_m x^m + \cdots + b_1 x + b_0$, where $a_n \neq 0$ and $b_m \neq 0$, so that $\deg\, p(x) = n$ and $\deg\, q(x) = m$. We may assume without loss of generality that $n \geqslant m$, so that $\max\{\deg\, p(x), \deg\, q(x)\} = n$ (and b_n may be 0). By Definition 3.1.4, the leading term of $p(x) + q(x)$ will be $(a_n + b_n)x^n$, provided $a_n + b_n \neq 0$, and will be of degree less than n if $a_n + b_n = 0$. In any case, we have

$$\deg\,(p(x) + q(x)) \leqslant n = \max\{\deg\, p(x), \deg\, q(x)\}$$

as claimed. □

In fact, the degree of the sum is *equal* to the maximum of the two degrees, unless $p(x)$ and $q(x)$ have the same degree and the leading coefficients are the negatives of each other.

By Corollary 3.1.10, even if $S = \mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{Z}_p$, non-constant polynomials in $S[x]$ will not have multiplicative inverses. Hence $S[x]$ is never a field. However, it is possible to construct a field containing $S[x]$ in much the same way we constructed the field $\mathbb{Q}$ containing the integers $\mathbb{Z}$.

The formal construction is analogous to the construction of $\boldsymbol{Q}$ in §1.5. We start with the set

$$\mathcal{R} = \{(a(x), b(x)) \mid a(x), b(x) \in S[x],\ b(x) \neq 0\}$$

of ordered pairs of polynomials with the second polynomial non-zero. The relation $\sim$ defined on $\mathcal{R}$ by $(a(x), b(x)) \sim (a'(x), b'(x))$ if and only if $a(x)b'(x) = b(x)a'(x)$ is an equivalence relation.

We define

$$\boldsymbol{R} = \{\, [(a(x), b(x))] \mid (a(x), b(x)) \in \mathcal{R} \,\}$$

to be the set of equivalence classes. For $X = [(a(x), b(x))]$ and $Y = [(c(x), d(x))]$ in $\boldsymbol{R}$, we define addition $X + Y$ and multiplication $X \cdot Y$ by

$$X + Y = [(a(x), b(x))] + [(c(x), d(x))] = [(a(x)d(x) + b(x)c(x), b(x)d(x))]$$

and

$$X \cdot Y = [(a(x), b(x))] \cdot [(c(x), d(x))] = [(a(x)c(x), b(x)d(x))].$$

It is then straightforward, but tedious, to verify that $\boldsymbol{R}$ is a field. Identifying a polynomial $p(x)$ in $S[x]$ with the equivalence class $[(p(x), 1)]$ in $\boldsymbol{R}$, we consider $S[x]$ to be be contained in the field $\boldsymbol{R}$.

Less formally, we can view an ordered pair $(a(x), b(x))$ in $\mathcal{R}$ as the rational expression $\dfrac{a(x)}{b(x)}$, and its equivalence class $[(a(x), b(x))]$ as the rational function represented by all equivalent quotients of polynomials. We then obtain the more familiar field of rational functions, denoted $S(x)$.

Theorem 3.1.12 *Let S be one of $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{Z}_p$, p a prime, and let*

$$S(x) = \left\{ \frac{a(x)}{b(x)} \;\middle|\; a(x), b(x) \in S[x],\ b(x) \neq 0 \right\}$$

be the set of rational functions with coefficients in S. Then $S(x)$ is a field with addition and multiplication defined by

$$\frac{a(x)}{b(x)} + \frac{c(x)}{d(x)} = \frac{a(x)d(x) + b(x)c(x)}{b(x)d(x)}$$

and

$$\frac{a(x)}{b(x)} \cdot \frac{c(x)}{d(x)} = \frac{a(x)c(x)}{b(x)d(x)}.$$

The fields $S(x)$ and $\boldsymbol{R}$ are algebraically equivalent (i.e., isomorphic). By identifying a polynomial $p(x)$ with the rational function $\dfrac{p(x)}{1}$, we consider $S[x]$ to be contained in $S(x)$, just as $\mathbb{Z}$ is contained in $\mathbb{Q}$.

§3.1 Exercises

1. Determine the degree and leading coefficient of each of the following polynomials:
   (a) $p_1(x) = 5x^3 - 3x^2 + 2$
   (b) $p_2(x) = 7x^2 - 9x^5 + 4x^3 - 6$
   (c) $p_3(x) = 6$
   (d) $p_4(x) = 0$
   (e) $p_5(x) = x^{473} - 5$

2. Use only Definition 3.1.5 to find the product of the polynomials $a(x) = 2x^3 + 5x^2 + 4x + 7$ and $b(x) = 3x^2 + 6x + 8$. That is, identify a_i and b_i for each i and write the coefficients c_i of the product in the form $c_i = a_i b_0 + \cdots + a_0 b_i$ as in the definition. Include any 0 coefficients as well.

3. (a) For $p(x) = 5x^4 - 3x^3 + 2x^2 + 7x - 4$ and $q(x) = x^5 - 3x^4 + 4x^2 - 8x$, find $p(x) + q(x)$ and $p(x) - q(x)$.
   (b) For $p(x) = 2x^3 + 3x$ and $q(x) = x^2 - 5$, find $p(x)q(x)$.
   (c) For $p(x) = 4x^3 - 3x^2 + 1$ and $q(x) = x^2 + 2x + 5$, find $p(x)q(x)$.
   (d) For $p(x) = x^2 + 3$, find $p(x)^2$.

4. Use Definition 3.1.4 and properties of real numbers to show that if $a(x) = a_n x^n + \cdots + a_1 x + a_0$ and $b(x) = b_n x^n + \cdots + b_1 x + b_0$ are polynomials in $\mathbb{R}[x]$, then $a(x) + b(x) = b(x) + a(x)$. That is, verify the commutative law of polynomial addition. [Note: It will be necessary to use the commutative law of addition of real numbers. Be sure to indicate where this is used.]

5. Use Definition 3.1.4 and properties of real numbers to show that if $a(x) = a_n x^n + \cdots + a_1 x + a_0$, $b(x) = b_n x^n + \cdots + b_1 x + b_0$, and $c(x) = c_n x^n + \cdots + c_1 x + c_0$ are polynomials in $\mathbb{R}[x]$, then $a(x) + [b(x) + c(x)] = [a(x) + b(x)] + c(x)$. That is, verify the associative law of polynomial addition. [Note: It will be necessary to use the associative law of addition of real numbers. Be sure to indicate where this is used.]

6. Use Definition 3.1.5 and properties of real numbers to show that if $a(x) = a_m x^m + \cdots + a_1 x + a_0$ and $b(x) = b_n x^n + \cdots + b_1 x + b_0$ are polynomials in $\mathbb{R}[x]$, then $a(x) \cdot b(x)$ is a polynomial in $\mathbb{R}[x]$. That is, verify that $\mathbb{R}[x]$ is closed under multiplication. [Note: It will be necessary to use the closure of $\mathbb{R}$ under addition and multiplication. Be sure to indicate where this is used.]

7. Show that if $a(x)$, $b(x)$, and $q(x)$ are non-zero polynomials in $S[x]$ such that $b(x) = q(x)a(x)$, then $\deg a(x) \leqslant \deg b(x)$. [Hint: Use Theorem 3.1.8.]

3.2 Binomial Coefficients and Binomial Theorem

We will now consider the special polynomial product $(1+x)^n$ in $\mathbb{Z}[x]$. If we were to expand this product, we would obtain a polynomial in x of degree n with integer coefficients. The coefficients are the binomial coefficients.

Definition 3.2.1 *Let $n \geqslant 0$ be an integer and let r be an integer with $0 \leqslant r \leqslant n$. The* **binomial coefficient** $\binom{n}{r}$ *is defined to be the coefficient of x^r in the polynomial $(1+x)^n$.*

That is, if $(1+x)^n = c_0 + c_1 x + \cdots + c_{n-1}x^{n-1} + c_n x^n$, then $\binom{n}{r} = c_r$.

Example: By direct computation, we have

$$(1+x)^4 = 1 + 4x + 6x^2 + 4x^3 + x^4,$$

and so

$$\binom{4}{0} = 1,\ \binom{4}{1} = 4,\ \binom{4}{2} = 6,\ \binom{4}{3} = 4,\ \binom{4}{4} = 1,$$

by Definition 3.2.1. □

Remarks:

1. The binomial coefficient $\binom{n}{r}$ is pronounced "n choose r," due to the fact that it is also equal to the number of ways to choose a set of r objects from an n element set. We can relate this interpretation to the coefficients of $(1+x)^n$ as follows. The polynomial $(1+x)^n$ is a sum of terms, each of which is a product of either an x or a 1 chosen from each of the n factors of $(1+x)(1+x)\cdots(1+x)$. An x^r term arises by choosing an x from r factors and 1 from the rest, and the coefficient of x^r is the number of different ways to do this. Hence the coefficient $\binom{n}{r}$ is the number of ways to choose r factors from the set of n factors.

2. The polynomial $1+x$ is in the set $\mathbb{Z}[x]$, and by the Polynomial Properties Theorem (Theorem 3.1.6 (vi)), $\mathbb{Z}[x]$ is closed under multiplication. It follows that $(1+x)^n$ is in $\mathbb{Z}[x]$, and the binomial coefficient $\binom{n}{r}$ is always an integer.

It is easy to check by direct calculation that the leading coefficient and the constant term of $(1+x)^n$ are both 1, which implies the following result.

Proposition 3.2.2 *For every integer $n \geqslant 0$,*

$$\binom{n}{0} = \binom{n}{n} = 1.$$

The proposition above allows us to compute the first and last binomial coefficients for a given value of n. Those in between can then be calculated inductively using the following very important theorem.

Theorem 3.2.3 (Pascal's Rule) *For integers n, r, with $0 < r < n+1$,*

$$\binom{n+1}{r} = \binom{n}{r-1} + \binom{n}{r}.$$

Proof. By definition, we must show that if $(1+x)^n = c_0 + c_1x + \cdots + c_{n-1}x^{n-1} + c_nx^n$, so that $\binom{n}{r-1} = c_{r-1}$ and $\binom{n}{r} = c_r$, then the coefficient $\binom{n+1}{r}$ of x^r in $(1+x)^{n+1}$ is $c_{r-1} + c_r$.

We can calculate $(x+1)^{n+1}$ as follows:

$$\begin{aligned}
(1+x)^{n+1} &= (1+x)(1+x)^n \\
&= (1+x)^n + x(1+x)^n \\
&= (c_0 + c_1x + \cdots + c_{n-1}x^{n-1} + c_nx^n) + x(c_0 + c_1x + \cdots + c_{n-1}x^{n-1} + c_nx^n) \\
&= c_0 + c_1x + c_2x^2 + \cdots + c_{n-1}x^{n-1} + c_nx^n + \\
&\qquad c_0x + c_1x^2 + c_2x^3 + \cdots + c_{n-1}x^n + c_nx^{n+1} \\
&= c_0 + (c_0+c_1)x + (c_1+c_2)x^2 + \cdots + (c_{n-2}+c_{n-1})x^{n-1} + (c_{n-1}+c_n)x^n + c_nx^{n+1}.
\end{aligned}$$

Therefore, if $0 < r < n+1$, the coefficient of x^r in $(1+x)^{n+1}$ is indeed $c_{r-1} + c_r$ as desired, and so Pascal's Rule holds. □

Example: We saw in the previous example that $\binom{4}{1} = 4$ and $\binom{4}{2} = 6$, hence Pascal's Rule, with $n = 4$ and $r = 2$, says that

$$\binom{5}{2} = \binom{4}{1} + \binom{4}{2} = 4 + 6 = 10.$$

This can be verified directly by expanding $(1+x)^5$. □

Using Pascal's Rule, the binomial coefficients can be calculated and put in a triangular array, called **Pascal's Triangle:**

$$\begin{array}{ccccccccccccccccccc}
 & & & & & & & & 1 & & & & & & & & & & \\
 & & & & & & & 1 & & 1 & & & & & & & & & \\
 & & & & & & 1 & & 2 & & 1 & & & & & & & & \\
 & & & & & 1 & & 3 & & 3 & & 1 & & & & & & & \\
 & & & & 1 & & 4 & & 6 & & 4 & & 1 & & & & & & \\
 & & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & & & & \\
 & & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & & & & \\
 & \iddots & & \iddots & & \iddots & & \iddots & \vdots & \ddots & & \ddots & & \ddots & & \ddots & & &
\end{array}$$

The binomial coefficient $\binom{n}{r}$ is placed in row n, position r (where the rows and positions are numbered starting with 0). Thus the "border" entries are 1 by Proposition 3.2.2, and Pascal's Rule says that each "interior" entry is the sum of the two entries above it.

Using Pascal's Rule and induction, we can derive a formula to compute the binomial coefficients directly. Recall that for a positive integer m, we define m-**factorial**, denoted $m!$, as

$$m! = m(m-1)(m-2)\cdots(3)(2)(1),$$

and we define $0! = 1$.

Theorem 3.2.4 *If n and r are integers with $0 \leqslant r \leqslant n$, then*

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

Proof. We proceed by induction on n. If $n = 0$, then $r = 0$ as well, and we have

$$\binom{0}{0} = 1 = \frac{0!}{0!(0-0)!}.$$

Therefore, the formula holds with $n = 0$.

Now assume the formula holds for $n = k$; that is,

$$\binom{k}{r} = \frac{k!}{r!(k-r)!} \quad (*)$$

for all $0 \leqslant r \leqslant k$. We must show that, given $(*)$, the formula holds for $n = k+1$; that is,

$$\binom{k+1}{r} = \frac{(k+1)!}{r!((k+1)-r)!}$$

for all $0 \leqslant r \leqslant k+1$.

Using Proposition 3.2.2 and the fact that $0! = 1$, we have

$$\frac{(k+1)!}{0!((k+1)-0)!} = \frac{(k+1)!}{0!(k+1)!} = \frac{(k+1)!}{(k+1)!} = 1 = \binom{k+1}{0}$$

and

$$\frac{(k+1)!}{(k+1)!((k+1)-(k+1))!} = \frac{(k+1)!}{(k+1)!\,0!} = \frac{(k+1)!}{(k+1)!} = 1 = \binom{k+1}{k+1},$$

and so the formula holds for $n = k+1$ with $r = 0$ and $r = k+1$. We may now assume $0 < r < k+1$, so that Pascal's Rule (Theorem 3.2.3) applies. We have

$$\begin{aligned}
\binom{k+1}{r} &= \binom{k}{r-1} + \binom{k}{r} \text{ by Pascal's Rule,} \\
&= \frac{k!}{(r-1)!(k-(r-1))!} + \frac{k!}{r!(k-r)!} \text{ by } (*), \\
&= \frac{k!}{(r-1)!(k-r+1)(k-r)!} + \frac{k!}{r(r-1)!(k-r)!} \\
&= \frac{k!}{(r-1)!(k-r)!}\left[\frac{1}{(k-r+1)} + \frac{1}{r}\right] \\
&= \frac{k!}{(r-1)!(k-r)!} \cdot \frac{r+(k-r+1)}{r(k-r+1)} \\
&= \frac{k!}{(r-1)!(k-r)!} \cdot \frac{k+1}{r(k-r+1)} \\
&= \frac{(k+1)k!}{r(r-1)!(k-r+1)(k-r)!} \\
&= \frac{(k+1)!}{r!(k-r+1)!} = \frac{(k+1)!}{r!((k+1)-r)!},
\end{aligned}$$

and so the formula holds for $n = k+1$. Therefore, the formula holds for all $n \geqslant 0$ by the Principle of Mathematical Induction. □

We noted in Proposition 3.2.2 that for $n \geqslant 0$,

$$\binom{n}{0} = \binom{n}{n} = 1.$$

Using Theorem 3.2.4, it is easy to see that for $n \geqslant 1$,

$$\binom{n}{1} = \binom{n}{n-1} = n.$$

More generally, we have the following corollary, whose proof is left as an exercise (see Exercise 3.2.6).

Corollary 3.2.5 *If n and r are integers with $0 \leqslant r \leqslant n$, then*

$$\binom{n}{r} = \binom{n}{n-r}.$$

Using the formula above and recalling the definition of binomial coefficients, we have:

Corollary 3.2.6 *If n is an integer with $n \geqslant 0$, then*

$$(1+x)^n = \sum_{r=0}^{n} \binom{n}{r} x^r = \sum_{r=0}^{n} \frac{n!}{r!(n-r)!} x^r.$$

More generally, we have the Binomial Theorem, a formula for expanding the product $(a+b)^n$.

Theorem 3.2.7 (Binomial Theorem) *If n is an integer with $n \geqslant 0$, then*

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^{n-r} b^r = \sum_{r=0}^{n} \frac{n!}{r!(n-r)!} a^{n-r} b^r.$$

Proof. If $a = 0$, the left side of the equation becomes b^n and the right side reduces to $\binom{n}{n} b^n$, and since $\binom{n}{n} = 1$, the formula holds. We may therefore assume $a \neq 0$.

Writing $(a+b) = a(1+\frac{b}{a})$ and letting $x = \frac{b}{a}$ in Corollary 3.2.6, we have

$$(a+b)^n = a^n \left(1 + \frac{b}{a}\right)^n = a^n \sum_{r=0}^{n} \binom{n}{r} \left(\frac{b}{a}\right)^r = \sum_{r=0}^{n} \binom{n}{r} a^n \cdot \frac{b^r}{a^r} = \sum_{r=0}^{n} \binom{n}{r} a^{n-r} b^r,$$

and the first equality in the formulas holds. The second equality in the formula follows by replacing the binomial coefficient by the expression from Theorem 3.2.4. □

Example: We expand $(3x-2)^5$ using the Binomial Theorem with $n=5$, $a=3x$, and $b=-2$:

$$\begin{aligned}
(3x-2)^5 &= \sum_{r=0}^{5} \binom{5}{r} (3x)^{5-r}(-2)^r \\
&= \binom{5}{0}(3x)^5 + \binom{5}{1}(3x)^4(-2) + \binom{5}{2}(3x)^3(-2)^2 + \\
&\qquad \binom{5}{3}(3x)^2(-2)^3 + \binom{5}{4}(3x)(-2)^4 + \binom{5}{5}(-2)^5 \\
&= (3x)^5 + 5(3x)^4(-2) + 10(3x)^3(-2)^2 + 10(3x)^2(-2)^3 + 5(3x)(-2)^4 + (-2)^5 \\
&= 243x^5 + 5\cdot 81x^4\cdot(-2) + 10\cdot 27x^3\cdot 4 + 10\cdot 9x^2\cdot(-8) + 5\cdot 3x\cdot 16 + (-32) \\
&= 243x^5 - 810x^4 + 1080x^3 - 720x^2 + 240x - 32.
\end{aligned}$$

Hence $(3x-2)^5 = 243x^5 - 810x^4 + 1080x^3 - 720x^2 + 240x - 32$. □

---

§3.2 Exercises

1. Evaluate the following binomial coefficients. Show your work.

   (a) $\binom{7}{3}$ (c) $\binom{8}{2}$ (e) $\binom{9}{5}$

   (b) $\binom{6}{4}$ (d) $\binom{8}{6}$

2. Find the coefficient

   (a) of x^{16} in $(x+1)^{20}$,
   (b) of x^9 in $(x+1)^{15}$,
   (c) of x^7 in $(x+1)^{18}$.

3. Use the Binomial Theorem to expand the following powers:

   (a) $(2x+3)^4$
   (b) $(2x+3)^5$
   (c) $(2x+3)^6$.

4. Show that if n and r are integers with $0 \leqslant r \leqslant n$, then $r!(n-r)! \mid n!$.

5. Show that if $n \geqslant 1$, then $\binom{n}{1} = \binom{n}{n-1} = n$.

6. Show that if n and r are integers with $0 \leqslant r \leqslant n$, then $\binom{n}{r} = \binom{n}{n-r}$.

On Exercises 7 and 8, prove the statements by using the Binomial Theorem to evaluate $(a+b)^n$ for an appropriate choice of a and b.

7. Show that if n is an integer and $n \geqslant 1$, then

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n-1} + \binom{n}{n} = 2^n.$$

8. Show that if n is an integer and $n \geqslant 1$, then

$$\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \cdots + (-1)^n \binom{n}{n} = 0.$$

3.3 Divisibility and Polynomials

In this section we consider the divisibility properties of polynomials, including the Division Algorithm and greatest common divisors. Most of the results on divisibility of integers can be translated almost word for word into analogous results for polynomials. Compare the definitions and results below to those in §2.2.

Note: So that we can always divide coefficients of polynomials, we will now assume that the ring S of coefficients is one of $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{Z}_p$, p a prime, but *not* $\mathbb{Z}$. Thus we assume that S is a field.

In all of the results below, $a(x)$, $b(x)$, $c(x)$, etc., will denote polynomials in $S[x]$. Recall that constants (elements of S) are considered to be elements of $S[x]$ by identifying them with constant polynomials.

Definition 3.3.1 *Let $a(x)$ and $b(x)$ be polynomials in $S[x]$, with $a(x) \neq 0$. We say that $a(x)$* **divides** *$b(x)$, and write $a(x) \mid b(x)$, if $b(x) = q(x)a(x)$ for some polynomial $q(x)$ in $S[x]$.*

As with integers, if $a(x) \mid b(x)$, we also say $a(x)$ is a **divisor** or **factor** of $b(x)$, that $b(x)$ is a **multiple** of $a(x)$, or that $b(x)$ is **divisible** by $a(x)$.

The following basic properties are proved in exactly the same way as the analogous results for integers.

Theorem 3.3.2 *The following properties hold for polynomials $a(x)$, $b(x)$, $c(x)$, and $d(x)$ in $S[x]$.*

i. *If $a(x) \neq 0$, then $a(x) \mid 0$.*

ii. *If $a(x) \neq 0$, then $a(x) \mid a(x)$.*

iii. *If $a(x)$ is any polynomial, then $1 \mid a(x)$.*

iv. *If $a(x) \mid b(x)$ and $b(x) \mid c(x)$, then $a(x) \mid c(x)$.*

v. *If $a(x) \mid b(x)$ and $d(x) \mid c(x)$, then $a(x)d(x) \mid b(x)c(x)$.*

vi. **(Combination Theorem)** *If $a(x) \mid b(x)$ and $a(x) \mid c(x)$, then $a(x) \mid b(x)f(x) + c(x)g(x)$ for all polynomials $f(x)$, $g(x)$ in $S[x]$.*

Proof. Property (i) follows from the definition and the fact that $0 = 0 \cdot a(x)$. Properties (ii) and (iii) both follow from $a(x) = 1 \cdot a(x)$ and the fact that $1 \in S[x]$. The proofs of (iv) and (v) are left as exercises (see Exercises 3.3.3 and 3.3.4).

For (vi), if $a(x) \mid b(x)$ and $a(x) \mid c(x)$, then $b(x) = a(x)m(x)$ and $c(x) = a(x)n(x)$ for some $m(x), n(x) \in S[x]$. Hence we have

$$b(x)f(x) + c(x)g(x) = (a(x)m(x))f(x) + (a(x)n(x))g(x) = a(x)[m(x)f(x) + n(x)g(x)].$$

Since $S[x]$ is closed under multiplication and addition, $m(x)f(x) + n(x)g(x)$ is in $S[x]$, and so this implies $a(x) \mid b(x)f(x) + c(x)g(x)$ by definition. □

We noted that for integers, the signs of the integers do not affect divisibility. The analogous result for polynomials involves constant multiples.

Proposition 3.3.3 *If $a(x) \mid b(x)$ and $k \neq 0$ is an element of S, then $k \cdot a(x) \mid b(x)$.*

Proof. Let $a(x) \mid b(x)$ and $0 \neq k \in S$. Then $b(x) = a(x)m(x)$ for some $m(x) \in S[x]$, and since S is a field, k has a multiplicative inverse k^{-1} in S, hence also in $S[x]$. We have

$$b(x) = a(x)m(x) = (k \cdot k^{-1})a(x)m(x) = k \cdot a(x)(k^{-1} \cdot m(x)),$$

and since $S[x]$ is closed under multiplication, $k^{-1} \cdot m(x) \in S[x]$ and this implies $k \cdot a(x) \mid b(x)$ by definition. □

Note that the proof of the proposition depends on the fact that S is a *field* and thus every non-zero element k in S has a multiplicative inverse in S (and hence in $S[x]$). This is also necessary for the following corollary.

Corollary 3.3.4 *If $k \neq 0$ is an element of S and $b(x)$ is in $S[x]$, then $k \mid b(x)$.*

Proof. Since $1 \mid b(x)$ and $k \cdot 1 = k$, substituting $a(x) = 1$ in Proposition 3.3.3 implies the result. □

We showed that if a and b are positive integers and $a \mid b$, then $a \leqslant b$. The degree induces a partial order on the set of polynomials, and we obtain the following analogous result.

Theorem 3.3.5 *If $a(x) \mid b(x)$ and $b(x) \neq 0$, then* $\deg\, a(x) \leqslant \deg\, b(x)$.

Proof. Since $a(x) \mid b(x)$ and $b(x) \neq 0$, there exists a non-zero polynomial $m(x) \in S[x]$ such that $b(x) = m(x)a(x)$. By Theorem 3.1.8, we have

$$\deg\, b(x) = \deg\,(m(x)a(x)) = \deg\, m(x) + \deg\, a(x).$$

Since $m(x) \neq 0$, we know $\deg\, m(x) \geqslant 0$, and so

$$\deg\, b(x) = \deg\, m(x) + \deg\, a(x) \geqslant \deg\, a(x).$$

Therefore, $\deg\, b(x) \geqslant \deg\, a(x)$. □

We showed that if two integers a, b divide each other, then $b = \pm a$. The analogue for polynomials is the following result.

Theorem 3.3.6 *If $a(x) \mid b(x)$ and $b(x) \mid a(x)$, then $b(x) = k \cdot a(x)$ for some constant k in S.*

Proof. Since $a(x) \mid b(x)$ and $b(x) \mid a(x)$, we have $b(x) = m(x)a(x)$ and $a(x) = n(x)b(x)$ for some $m(x), n(x) \in S[x]$. Hence

$$a(x) = n(x)b(x) = n(x)m(x)a(x),$$

and so

$$0 = n(x)m(x)a(x) - a(x) = [n(x)m(x) - 1]a(x).$$

By hypothesis, we know $a(x) \neq 0$, hence $n(x)m(x) - 1 = 0$ by Corollary 3.1.9. Thus $n(x)m(x) = 1$, and so $n(x)$ and $m(x)$ are non-zero constants by Corollary 3.1.10. In particular, $m(x) = k \in S$ and $b(x) = k \cdot a(x)$. □

Observe that the following corollary is equivalent to Corollary 3.1.10 because $b(x) \mid 1$ is equivalent to the existence of a multiplicative inverse for $b(x)$ in $S[x]$.

Corollary 3.3.7 *If $b(x) \mid 1$, then $b(x)$ is a non-zero constant polynomial.*

Proof. Since $1 \mid b(x)$, we may substitute $a(x) = 1$ in Theorem 3.3.6 to obtain $b(x) = k \cdot 1 = k$ for some non-zero constant $k \in S$. □

Analogous to the Division Algorithm for integers, a polynomial in $S[x]$ can be divided by another polynomial using long division to obtain a quotient and a remainder. In this case, the process continues until the remainder has lower degree than the divisor. As before, the quotient and remainder are unique.

Example: We use long division to divide $b(x) = 2x^3 - 5x^2 + 6x - 3$ by $a(x) = x^2 + 4x$ and find the quotient $q(x)$ and remainder $r(x)$. The procedure is similar to long division of integers and is as follows:

1. Divide the leading term of the dividend by the leading term of the divisor. In this case, this is $2x^3/x^2 = 2x$.
2. Multiply the divisor by the result of Step (1). In this case, we obtain $2x(x^2+4x) = 2x^3+8x^2$.
3. Subtract the result of Step (2) from the dividend. Here, this yields
$$(2x^3 - 5x^2 + 6x - 3) - (2x^3 + 8x^2) = -13x^2 + 6x - 3.$$
4. Using the result of Step (3) as the new dividend (of lower degree), repeat the steps until the difference obtained in Step (3) is of strictly lower degree than the degree of the divisor.

The sum of the terms obtained from Step (1) will be the quotient $q(x)$ and the final remainder in Step (4) will be the remainder $r(x)$. The complete procedure for our example is given below.

$$\begin{array}{rrrrrrrr}
 & 2x & - & 13 & & & & \\
\cline{2-8}
x^2+4x\,\big) & 2x^3 & - & 5x^2 & + & 6x & - & 3 \\
 & 2x^3 & + & 8x^2 & & & & \\
\cline{2-8}
 & & - & 13x^2 & + & 6x & - & 3 \\
 & & - & 13x^2 & - & 52x & & \\
\cline{2-8}
 & & & & & 58x & - & 3
\end{array}$$

We have $q(x) = 2x - 13$ and $r(x) = 58x - 3$. Observe that

$$\deg(58x - 3) = 1 < 2 = \deg(x^2 + 4x),$$

so that $\deg r(x) < \deg a(x)$. It is easy to check by direct calculation that

$$2x^3 - 5x^2 + 6x - 3 = (2x - 13)(x^2 + 4x) + (58x - 3),$$

that is, $b(x) = q(x)a(x) + r(x)$. □

The Division Algorithm for polynomials gives a precise statement of the existence and uniqueness of a quotient and remainder. Again, the proof involves the long division procedure along with induction, and is omitted here.

Theorem 3.3.8 (Division Algorithm) *If $a(x)$ and $b(x)$ are polynomials in $S[x]$ with $a(x) \neq 0$, then there exist unique polynomials $q(x)$ and $r(x)$ in $S[x]$ satisfying $b(x) = q(x)a(x) + r(x)$ and either $r(x) = 0$ or $\deg\, r(x) < \deg\, a(x)$.*

As with integers, $b(x)$ is the **dividend**, $a(x)$ is the **divisor**, $q(x)$ is the **quotient**, and $r(x)$ is the **remainder**. The condition that $a(x) \mid b(x)$ is equivalent to the condition that $r(x) = 0$. Since we defined $\deg\, 0 = -\infty$, the case $r(x) = 0$ is included in the case $\deg\, r(x) < \deg\, a(x)$, but we include the possibility $r(x) = 0$ separately in the theorem for emphasis.

Greatest Common Divisor

Definition 3.3.9 *Let $a(x)$ and $b(x)$ be polynomials in $S[x]$, at least one of which is not zero. The* **greatest common divisor** *of $a(x)$ and $b(x)$ is the monic polynomial $d(x)$ satisfying*

i. *$d(x) \mid a(x)$ and $d(x) \mid b(x)$,*

ii. *if $c(x) \mid a(x)$ and $c(x) \mid b(x)$, then $\deg\, c(x) \leqslant \deg\, d(x)$.*

We denote the GCD of $a(x)$ and $b(x)$ as $d(x) = (a(x), b(x))$. The GCD of $a(x)$ and $b(x)$ is the *monic* polynomial of highest degree that divides both $a(x)$ and $b(x)$. It is not difficult to see that a GCD must exist, but it is not obvious from the definition that it must be unique. The GCD is found using the Euclidean Algorithm. The procedure is nearly identical to that used for integers.

Theorem 3.3.10 (Euclidean Algorithm) *Let $A(x)$ and $B(x)$ be non-zero polynomials in $S[x]$. Use the Division Algorithm to obtain the following system of equations:*

$$\begin{array}{rcll}
B(x) & = & Q_1(x)A(x) + R_1(x), & 0 \leqslant \deg\, R_1(x) < \deg\, A(x) \\
A(x) & = & Q_2(x)R_1(x) + R_2(x), & 0 \leqslant \deg\, R_2(x) < \deg\, R_1(x) \\
R_1(x) & = & Q_3(x)R_2(x) + R_3(x), & 0 \leqslant \deg\, R_3(x) < \deg\, R_2(x) \\
\vdots & \vdots & \vdots & \vdots \\
R_{n-2}(x) & = & Q_n(x)R_{n-1}(x) + R_n(x), & 0 \leqslant \deg\, R_n(x) < \deg\, R_{n-1}(x) \\
R_{n-1}(x) & = & Q_{n+1}(x)R_n(x). &
\end{array}$$

Let $R_n(x)$ be the last non-zero remainder. If r is the leading coefficient of $R_n(x)$, then $\dfrac{1}{r}R_n(x)$ is the greatest common divisor of $A(x)$ and $B(x)$.

Remarks: Using the Combination Theorem, it can be shown that

1. $R_n(x) \mid A(x)$ and $R_n(x) \mid B(x)$, and
2. if $c(x) \mid A(x)$ and $c(x) \mid B(x)$, then $c(x) \mid R_n(x)$.

If r is the leading coefficient of $R_n(x)$, then $\frac{1}{r}R_n(x)$ is monic and, by Proposition 3.3.3 and basic properties of divisibility, also satisfies conditions (1) and (2). If $c(x) \mid \frac{1}{r}R_n(x)$, then by Theorem 3.3.5 we know $\deg c(x) \leqslant \deg \frac{1}{r}R_n(x)$. It follows that $\frac{1}{r}R_n(x)$ is the GCD of $A(x)$ and $B(x)$.

The procedure used for the verification of the Euclidean Algorithm suggested in the remarks above is the same as that used in the example in §2.3 on page 46. It is notationally rather cumbersome to do this in general, and we will instead give a proof along the lines of the proof of Theorem 2.3.3. Before proceeding with the proof, we demonstrate the procedure with an example.

Example: We find the GCD of $A(x) = 6x^3 - 10x^2 - 30x + 18$ and $B(x) = 3x^4 + x^3 - 23x^2 - 23x + 6$. Using long division to find quotients and remainders, we obtain the following equations:

$$\begin{aligned} 3x^4 + x^3 - 23x^2 - 23x + 6 &= ((1/2)x + 1)(6x^3 - 10x^2 - 30x + 18) + (2x^2 - 2x - 12) \\ 6x^3 - 10x^2 - 30x + 18 &= (3x - 2)(2x^2 - 2x - 12) + (2x - 6) \\ 2x^2 - 2x - 12 &= (x + 2)(2x - 6). \end{aligned}$$

The last non-zero remainder is $R(x) = 2x - 6$. Since the leading coefficient of $R(x)$ is 2, we have

$$(A(x), B(x)) = \frac{1}{2}R(x) = \frac{1}{2}(2x - 6),$$

and so $(A(x), B(x)) = x - 3$. □

The main result we need for the proof of the Euclidean Algorithm is the following lemma, analogous to Lemma 2.3.4 for integers.

Lemma 3.3.11 *If $a(x)$ and $b(x)$ are non-zero polynomials in $S[x]$, and $q(x)$ and $r(x)$ are polynomials in $S[x]$ such that $b(x) = q(x)a(x) + r(x)$, then $(b(x), a(x)) = (a(x), r(x))$.*

Proof. If $c(x) \mid a(x)$ and $c(x) \mid r(x)$, then by the Combination Theorem, $c(x) \mid q(x)a(x) + r(x)$; that is, $c(x) \mid b(x)$. Conversely, if $c(x) \mid b(x)$ and $c(x) \mid a(x)$, then $c(x) \mid b(x) - q(x)a(x)$; that is, $c(x) \mid r(x)$. Therefore, the set of common divisors of $a(x)$ and $r(x)$ is precisely the same as the set of common divisors of $b(x)$ and $a(x)$. Thus the monic polynomial of highest degree in this set of common divisors is equal to $(a(x), r(x))$ and to $(b(x), a(x))$, hence $(a(x), r(x)) = (b(x), a(x))$. □

Proof of Theorem 3.3.10. We will first show that

$$(B(x), A(x)) = (R_i(x), R_{i+1}(x))$$

for all $i = 1, \ldots, n - 1$. The proof is by induction on i.

By Lemma 3.3.11, $B(x) = Q_1(x)A(x) + R_1(x)$ implies

$$(B(x), A(x)) = (A(x), R_1(x))$$

and $A(x) = Q_2(x)R_1(x) + R_2(x)$ implies

$$(A(x), R_1(x)) = (R_1(x), R_2(x)).$$

Hence

$$(B(x), A(x)) = (R_1(x), R_2(x))$$

and the claim holds for $i = 1$.

Now assume the claim is true for $i = k$; that is,

$$(B(x), A(x)) = (R_k(x), R_{k+1}(x)).$$

Since $R_k(x) = Q_{k+2}(x)R_{k+1}(x) + R_{k+2}(x)$, Lemma 3.3.11 implies

$$(R_k(x), R_{k+1}(x)) = (R_{k+1}(x), R_{k+2}(x)).$$

Thus

$$(B(x), A(x)) = (R_{k+1}(x), R_{k+2}(x)),$$

and the claim is true for $i = k + 1$.

By the Principle of Mathematical Induction, we therefore have that

$$(B(x), A(x)) = (R_i(x), R_{i+1}(x))$$

for all $i = 1, \ldots, n - 1$. In particular,

$$(B(x), A(x)) = (R_{n-1}(x), R_n(x)).$$

Finally, $\frac{1}{r}R_n(x)$ is monic and, by Proposition 3.3.3, $\frac{1}{r}R_n(x) \mid R_n(x)$ and $\frac{1}{r}R_n(x) \mid R_{n-1}(x)$ (as $R_{n-1}(x) = Q_{n+1}(x)R_n(x)$). If $c(x)$ is any other common divisor, then $c(x) \mid R_n(x)$ and so

$$\deg\, c(x) \leqslant \deg\, R_n(x) = \deg\, \frac{1}{r}R_n(x).$$

Hence

$$(B(x), A(x)) = (R_{n-1}(x), R_n(x)) = \frac{1}{r}R_n(x)$$

and the theorem is proved. □

The GCD of two polynomials satisfies the basic properties satisfies by the GCD of integers. In particular, the GCD can be written as a combination of the two polynomials, and any common divisor must also divide the GCD. We first need the following characterization of the GCD, analogous to Lemma 2.4.4.

Lemma 3.3.12 *Let $a(x)$ and $b(x)$ be polynomials in $S[x]$, at least one of which is not zero. The GCD of $a(x)$ and $b(x)$ is the monic polynomial in $S[x]$ of* smallest degree *that can be written in the form $a(x)s(x) + b(x)t(x)$ for $s(x), t(x) \in S[x]$.*

Proof. Let $a(x)$ and $b(x)$ be polynomials in $S[x]$, at least one of which is not 0, and define the set

$$\mathcal{P} = \{a(x)s(x) + b(x)t(x) \mid s(x), t(x) \in S[x] \text{ and } a(x)s(x) + b(x)t(x) \neq 0\}.$$

Since at least one of $a(x)$ or $b(x)$ is non-zero, at least one of $a(x)$ or $b(x)$ is therefore in $\mathcal{P}$, and so $\mathcal{P}$ is non-empty. In particular, the set of *degrees* of polynomials in $\mathcal{P}$ is a non-empty set of non-negative integers, hence has a smallest element by the Well-ordering Principle (Theorem 2.1.1).

Thus there is a polynomial $d_1(x) \in \mathcal{P}$ with smallest possible degree. If a is the leading coefficient of $d_1(x)$, then $d(x) = \frac{1}{a}d_1(x)$ is monic, is an element of $\mathcal{P}$, and is of the same degree as $d_1(x)$. Thus $d(x)$ is a monic polynomial of smallest degree in $\mathcal{P}$, and so $d(x) = a(x)m(x) + b(x)n(x)$ for some $m(x), n(x) \in S[x]$.

The conclusion of the theorem is that this polynomial $d(x)$ is the GCD of $a(x)$ and $b(x)$. We will show that $d(x)$ satisfies parts (i) and (ii) of Definition 3.3.9, hence $d(x) = (a(x), b(x))$.

We first show (i), that $d(x) \mid a(x)$ and $d(x) \mid b(x)$. By the Division Algorithm (Theorem 3.3.8), we can write $a(x) = q(x)d(x) + r(x)$ for $q(x), r(x) \in S[x]$ with $r(x) = 0$ or $0 \leqslant \deg r(x) < \deg d(x)$. Moreover, we can express $r(x)$ as

$$\begin{aligned} r(x) &= a(x) - q(x)d(x) \\ &= a(x) - q(x)(a(x)m(x) + b(x)n(x)) \\ &= a(x)(1 - q(x)m(x)) + b(x)(-q(x)n(x)). \end{aligned}$$

Since $1 - q(x)m(x)$ and $-q(x)n(x)$ are in $S[x]$, if $r(x) \neq 0$, then $r(x) \in \mathcal{P}$. However, $d(x)$ was chosen to have *smallest degree* among the elements of $\mathcal{P}$ and $\deg r(x) < \deg d(x)$. Hence $r(x)$ is not an element of $\mathcal{P}$, and so $r(x) = 0$ and $a(x) = q(x)d(x)$. Therefore $d(x) \mid a(x)$. The proof that $d(x) \mid b(x)$ is nearly identical and is left as an exercise.

Finally, we show (ii), that if $c(x)$ is a polynomial in $S[x]$ such that $c(x) \mid a(x)$ and $c(x) \mid b(x)$, then $\deg c(x) \leqslant \deg d(x)$. Suppose $c(x) \mid a(x)$ and $c(x) \mid b(x)$. Since $d(x) = a(x)m(x) + b(x)n(x)$ with $m(x), n(x) \in S[x]$, the Combination Theorem (Theorem 3.3.2 (vi)) says that $c(x) \mid d(x)$. Since $c(x) \mid d(x)$ and $d(x) \neq 0$, it follows from Theorem 3.3.5 that $\deg c(x) \leqslant \deg d(x)$ as claimed. □

The lemma implies the first of the basic properties of the GCD mentioned above.

Theorem 3.3.13 *If $a(x)$ and $b(x)$ are polynomials in $S[x]$, at least one of which is not zero, then there are polynomials $f(x)$, $g(x)$ in $S[x]$ such that $(a(x), b(x)) = a(x)f(x) + b(x)g(x)$.*

Proof. This follows immediately from Lemma 3.3.12. □

As was the case for integers, the only polynomials that can be written as a combination of $a(x)$ and $b(x)$ are the multiples of $(a(x), b(x))$.

Theorem 3.3.14 *Let $d(x) = (a(x), b(x))$. A polynomial $m(x)$ can be expressed in the form*

$$m(x) = a(x)s(x) + b(x)t(x),$$

with $s(x)$, $t(x)$ in $S[x]$, if and only if $d(x) \mid m(x)$.

Proof. If

$$m(x) = a(x)s(x) + b(x)t(x)$$

for some polynomials $s(x)$ and $t(x)$ in $S[x]$, then since $d(x) \mid a(x)$ and $d(x) \mid b(x)$, the Combination Theorem implies $d(x) \mid m(x)$. Thus $m(x)$ is a multiple of $d(x)$.

Conversely, suppose $m(x)$ is a multiple of $d(x)$, so that $m(x) = n(x)d(x)$ for some $n(x) \in S[x]$. Since

$$d(x) = a(x)u(x) + b(x)v(x)$$

for some $u(x), v(x) \in S[x]$ by Theorem 3.3.13, we have

$$m(x) = n(x)d(x) = a(x)(n(x)u(x)) + b(x)(n(x)v(x)),$$

and $n(x)u(x)$, $n(x)v(x)$ are in $S[x]$. Hence $d(x)$ can be expressed in the form $a(x)s(x) + b(x)t(x)$, with $s(x) = n(x)u(x)$ and $t(x) = n(x)v(x)$ in $S[x]$. □

In particular, note that if $m(x) = a(x)s(x) + b(x)t(x)$ for some $s(x)$, $t(x)$ in $S[x]$, it does *not* generally follow that $m(x) = (a(x), b(x))$.

In order to write $(a(x), b(x))$ as a combination of $a(x)$ and $b(x)$, we work backwards through the Euclidean Algorithm, writing each remainder as a combination of the two previous remainders. This is the same procedure that was used for integers.

Example: For $A(x) = 6x^3 - 10x^2 - 30x + 18$ and $B(x) = 3x^4 + x^3 - 23x^2 - 23x + 6$, we find polynomials $f(x)$, $g(x)$ such that $(A(x), B(x)) = A(x)f(x) + B(x)g(x)$. Recall from the example on page 102 that $(A(x), B(x)) = x - 3$, and the Euclidean Algorithm yields the following equations:

$$3x^4 + x^3 - 23x^2 - 23x + 6 = ((1/2)x + 1)(6x^3 - 10x^2 - 30x + 18) + (2x^2 - 2x - 12) \quad (3.1)$$

$$6x^3 - 10x^2 - 30x + 18 = (3x - 2)(2x^2 - 2x - 12) + (2x - 6) \quad (3.2)$$

$$2x^2 - 2x - 12 = (x + 2)(2x - 6). \quad (3.3)$$

Beginning with Equation 3.2 and solving for the remainder in terms of the dividend and divisor, we have

$$\begin{aligned} 2x - 6 &= (6x^3 - 10x^2 - 30x + 18) - (3x - 2)(2x^2 - 2x - 12) \text{ by Equation 3.2,} \\ &= (6x^3 - 10x^2 - 30x + 18) - \\ &\qquad (3x - 2)[(3x^4 + x^3 - 23x^2 - 23x + 6) - ((1/2)x + 1)(6x^3 - 10x^2 - 30x + 18)] \\ &\qquad \text{by Equation 3.1,} \\ &= (6x^3 - 10x^2 - 30x + 18)[1 + (3x - 2)((1/2)x + 1)] - \\ &\qquad (3x - 2)(3x^4 + x^3 - 23x^2 - 23x + 6) \\ &= (6x^3 - 10x^2 - 30x + 18)[(3/2)x^2 + 2x - 1)] + (3x^4 + x^3 - 23x^2 - 23x + 6)(-3x + 2). \end{aligned}$$

Recalling that the GCD is $x - 3 = (1/2)(2x - 6)$, we divide both sides of the equation by 2 to obtain

$$x - 3 = (6x^3 - 10x^2 - 30x + 18)\left[\frac{3}{4}x^2 + x - \frac{1}{2}\right] + (3x^4 + x^3 - 23x^2 - 23x + 6)\left[\frac{-3}{2}x + 1\right].$$

Hence $f(x) = (3/4)x^2 + x - (1/2)$ and $g(x) = (-3/2)x + 1$. □

Theorem 3.3.15 *Let $a(x)$ and $b(x)$ be polynomials, at least one of which is not zero. If $c(x) \mid a(x)$ and $c(x) \mid b(x)$, then $c(x) \mid (a(x), b(x))$.*

Proof. By Theorem 3.3.13, we can write $(a(x), b(x)) = a(x)s(x) + b(x)t(x)$ with $s(x), t(x) \in S[x]$. If $c(x) \mid a(x)$ and $c(x) \mid b(x)$, then by the Combination Theorem, we have $c(x) \mid a(x)s(x) + b(x)t(x)$, and so $c(x) \mid (a(x), b(x))$ as claimed. □

The previous theorem implies an alternate definition of the GCD. We state this equivalent characterization of the GCD as a theorem.

Theorem 3.3.16 *Let $a(x)$ and $b(x)$ be polynomials, at least one of which is not zero, and let $d(x)$ be a monic polynomial. Then $d(x) = (a(x), b(x))$ if and only if $d(x)$ satisfies*

i. *$d(x) \mid a(x)$ and $d(x) \mid b(x)$, and*

ii. *if $c(x) \mid a(x)$ and $c(x) \mid b(x)$, then $c(x) \mid d(x)$.*

To justify the use of the phrase "the GCD" in the Euclidean Algorithm and in the remarks above, we need to know that there is only one polynomial satisfying the definition. This is the reason the definition requires a monic polynomial. Note that any non-zero constant multiple of $(a(x), b(x))$ will satisfy conditions (i) and (ii) of the definition.

Theorem 3.3.17 *The greatest common divisor of two polynomials is unique.*

Proof. Let $a(x)$, $b(x)$ be polynomials in $S[x]$. Suppose both $d_1(x)$ and $d_2(x)$ in $S[x]$ are monic and satisfy conditions (i) and (ii) of Theorem 3.3.16. In particular, both are common divisors of $a(x)$ and $b(x)$. We need to show that $d_1(x) = d_2(x)$.

Since $d(x) = d_1(x)$ satisfies condition (ii) and $d_2(x)$ is a common divisor of $a(x)$ and $b(x)$, we have $d_2(x) \mid d_1(x)$. Similarly, since $d(x) = d_2(x)$ satisfies condition (ii) and $d_1(x)$ is a common divisor of $a(x)$ and $b(x)$, we have $d_1(x) \mid d_2(x)$.

Now by Theorem 3.3.6, we have $d_2(x) = k \cdot d_1(x)$ for some $k \in S$. Since $d_1(x)$ is monic, k is the leading coefficient of $k \cdot d_1(x)$, hence also of $d_2(x)$. But $d_2(x)$ is also monic, and so $k = 1$ and we have $d_2(x) = d_1(x)$. □

The other basic properties of the GCD of integers in §2.4 can also be stated and proved for polynomials, with the appropriate changes in wording. Some of these will appear as exercises.

Relatively Prime Pairs of Polynomials

Definition 3.3.18 *Two polynomials $a(x)$, $b(x)$ in $S[x]$ are* **relatively prime** *if $(a(x), b(x)) = 1$.*

Remark: Polynomials $a(x)$ and $b(x)$ are relatively prime if they have no *non-constant* common factors. For example, the polynomials $a(x) = 2x + 2 = 2(x + 1)$ and $b(x) = 4x - 2 = 2(2x - 1)$ have the constant 2 as a common factor, but no non-constant common factors. Although 2 satisfies conditions (i) and (ii) of the definition of GCD, it is not monic, and $\frac{1}{2} \cdot 2 = 1$ is the GCD. Hence $2x + 2$ and $4x - 2$ are relatively prime.

By Theorem 3.3.13, we know that if $a(x)$ and $b(x)$ are relatively prime, then 1 can be written as a combination of $a(x)$ and $b(x)$. The converse also holds in the case of relatively prime polynomials.

Theorem 3.3.19 *Let $a(x)$ and $b(x)$ be polynomials in $S[x]$. There exist polynomials $f(x)$, $g(x)$ in $S[x]$ such that $a(x)f(x) + b(x)g(x) = 1$ if and only if $(a(x), b(x)) = 1$.*

Proof. If $(a(x), b(x)) = 1$, then $a(x)f(x) + b(x)g(x) = 1$ for some $f(x), g(x) \in S[x]$ by Theorem 3.3.13. Conversely, if $a(x)f(x) + b(x)g(x) = 1$ for some $f(x), g(x) \in S[x]$, then by Theorem 3.3.14, $(a(x), b(x)) \mid 1$. Hence, by Theorem 3.3.7, $(a(x), b(x)) = k$ for some $k \in S$, and since $(a(x), b(x))$ is monic, $k = 1$. □

Finally, we have some related divisibility properties. The proofs of these properties are almost identical to those of the analogous results for integers.

Theorem 3.3.20 (Euclid's Lemma for Polynomials) *If $a(x) \mid b(x)c(x)$ and $(a(x), b(x)) = 1$, then $a(x) \mid c(x)$.*

Proof. Since $(a(x), b(x)) = 1$, we have $1 = a(x)s(x) + b(x)t(x)$ for some $s(x), t(x) \in S[x]$, by Theorem 3.3.13. Thus

$$\begin{aligned} c(x) &= [a(x)s(x) + b(x)t(x)]c(x) \\ &= [a(x)s(x)]c(x) + [b(x)t(x)]c(x) \\ &= a(x)[s(x)c(x)] + [b(x)c(x)]t(x). \end{aligned}$$

Since $a(x) \mid a(x)$ and $a(x) \mid b(x)c(x)$, we have

$$a(x) \mid a(x)[s(x)c(x)] + [b(x)c(x)]t(x)$$

by the Combination Theorem, and hence $a(x) \mid c(x)$. □

Theorem 3.3.21 *If $a(x) \mid c(x)$ and $b(x) \mid c(x)$, and $(a(x), b(x)) = 1$, then $a(x)b(x) \mid c(x)$.*

Proof. Since $(a(x), b(x)) = 1$, we have $1 = a(x)s(x) + b(x)t(x)$ for some $s(x), t(x) \in S[x]$, by Theorem 3.3.13. Since $a(x) \mid c(x)$ and $b(x) \mid c(x)$, we have $c(x) = a(x)m(x)$ and $c(x) = b(x)n(x)$ for some $m(x), n(x) \in S[x]$. Thus

$$\begin{aligned} c(x) &= c(x)[a(x)s(x) + b(x)t(x)] \\ &= c(x)[a(x)s(x)] + c(x)[b(x)t(x)] \\ &= [c(x)a(x)]s(x) + [c(x)b(x)]t(x) \\ &= [b(x)n(x)a(x)]s(x) + [a(x)m(x)b(x)]t(x) \\ &= a(x)b(x)[n(x)s(x) + m(x)t(x)]. \end{aligned}$$

All of $m(x)$, $n(x)$, $s(x)$, and $t(x)$ are in $S[x]$ and $S[x]$ is closed under addition and multiplication, hence $n(x)s(x) + m(x)t(x)$ is in $S[x]$. Thus

$$c(x) = a(x)b(x)[n(x)s(x) + m(x)t(x)]$$

implies $a(x)b(x) \mid c(x)$. □

More generally, we have the following results, whose proofs are similar to the proofs of Theorem 3.3.20 and Theorem 3.3.21 and are left to the exercises (see Exercises 3.3.11 and 3.3.12).

Theorem 3.3.22 *If $a(x) \mid b(x)c(x)$ and $(a(x), b(x)) = d(x)$, then $a(x) \mid c(x)d(x)$.*

Theorem 3.3.23 *If $a(x) \mid c(x)$ and $b(x) \mid c(x)$, and $(a(x), b(x)) = d(x)$, then $a(x)b(x) \mid c(x)d(x)$.*

§3.3 Exercises

For all exercises, S denotes $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$, and $a(x)$, $b(x)$, $c(x)$, and $d(x)$ denote polynomials in $S[x]$.

1. Use Definition 3.3.1 to show that if $a(x) \mid c(x)$ and $b(x) \mid c(x)$, then $a(x)b(x) \mid (c(x))^2$.
2. Show that if $a(x) \mid b(x)$ and $c(x)$ is any polynomial in $S[x]$, then $a(x) \mid b(x)c(x)$.
3. Show that if $a(x) \mid b(x)$ and $d(x) \mid c(x)$, then $a(x)d(x) \mid b(x)c(x)$.
4. Show that if $a(x) \mid b(x)$ and $b(x) \mid c(x)$, then $a(x) \mid c(x)$.
5. Use long division to find the quotient and remainder when $b(x)$ is divided by $a(x)$. Write $b(x) = q(x)a(x) + r(x)$, and expand the right hand side to check your answers.
   (a) $b(x) = 8x^4 + 6x^2 - 3x + 1$
   $a(x) = 2x^2 - x + 2$
   (b) $b(x) = x^3 + 3x^2 + 4x + 3$
   $a(x) = 3x + 6$

For Exercises 6–8, find the greatest common divisor of the polynomials $a(x)$ and $b(x)$, and find polynomials $f(x)$ and $g(x)$ so that $(a(x), b(x)) = a(x)f(x) + b(x)g(x)$.

6. $a(x) = x^2 - 1$
   $b(x) = 2x^7 - 4x^5 + 2$
7. $a(x) = x + 3$
   $b(x) = x^3 - 2x + 4$
8. $a(x) = x^4 - 4x^2 - 3x + 6$
   $b(x) = x^3 + x^2 - x - 10$
9. Show that if $a(x)$ is monic and $a(x) \mid b(x)$, then $(a(x), b(x)) = a(x)$.
   [Hint: Show that $a(x)$ satisfies the definition of the GCD of $a(x)$ and $b(x)$.]
10. Show that if k is a non-zero element of S, then $(k \cdot a(x), b(x)) = (a(x), b(x))$.
    [Hint: Let $d(x) = (a(x), b(x))$ and show that $d(x)$ satisfies the definition of the GCD of $k \cdot a(x)$ and $b(x)$. You may need to use Exercise 2 above.]
11. Show that if $a(x) \mid b(x)c(x)$ and $(a(x), b(x)) = d(x)$, then $a(x) \mid c(x)d(x)$; that is, prove Theorem 3.3.22.
12. Show that if $a(x) \mid c(x)$ and $b(x) \mid c(x)$, and $(a(x), b(x)) = d(x)$, then $a(x)b(x) \mid c(x)d(x)$; that is, prove Theorem 3.3.23.

3.4 Synthetic Division

Long division of a polynomial $p(x)$ by a monic polynomial $x-a$ of degree 1 can be greatly simplified by a process called **synthetic division**. This is essentially just long division with all unnecessary information deleted. The procedure is described below with an example.

We will use synthetic division to find the quotient and remainder when $p(x) = 2x^4 - 10x^3 + 13x^2 - 4$ is divided by $x - 3$.

1. Write the *zero* of the divisor (in this case, 3) next to an "upside down division sign:"

$$\begin{array}{r|l} 3 & \\ \hline \end{array}$$

2. Write the coefficients of $p(x)$ (*including all* 0 *coefficients*) in descending degree order:

$$\begin{array}{r|lllll} 3 & 2 & -10 & 13 & 0 & -4 \\ \hline \end{array}$$

3. Under the division sign, draw a line and write the leading coefficient below the line:

$$\begin{array}{r|lllll} 3 & 2 & -10 & 13 & 0 & -4 \\ \hline & 2 & & & & \end{array}$$

4. Multiply the zero of the divisor (3) by the next coefficient below the line and write the product above the line in the next column:

$$\begin{array}{r|lllll} 3 & 2 & -10 & 13 & 0 & -4 \\ & & 6 & & & \\ \hline & 2 & & & & \end{array}$$

5. *Add* the numbers in the next column and write the sum below the line:

$$\begin{array}{r|lllll} 3 & 2 & -10 & 13 & 0 & -4 \\ & & 6 & & & \\ \hline & 2 & -4 & & & \end{array}$$

6. Repeat steps 4 and 5 until the last column is reached:

$$\begin{array}{r|rrrrr} 3 & 2 & -10 & 13 & 0 & -4 \\ & & 6 & -12 & 3 & 9 \\ \hline & 2 & -4 & 1 & 3 & 5 \end{array}$$

7. Interpret the result. The number in the last column (5) is the remainder. The other numbers (2, -4, 1, and 3) are the coefficients of the quotient, in descending degree order. In this example, the quotient is $q(x) = 2x^3 - 4x^2 + x + 3$ and the remainder is $r(x) = 5$. Hence

$$2x^4 - 10x^3 + 13x^2 - 4 = (2x^3 - 4x^2 + x + 3)(x - 3) + 5.$$

Remarks:

1. This procedure works *only* if the divisor is of the form $x-a$, where a is an element of S (that is, monic of degree 1). It does *not* apply, for example, to division of $p(x)$ by x^2-2 or $3x-5$.

2. The number to the left of the "division sign" is the *zero* of the divisor, that is, the value of x that would make the divisor zero. Hence for $x-3$ we write 3 and for $x+2$ we would write -2. In general, for $x-a$ we write a. (Note that $x+2=x-(-2)$.)

3. A coefficient of 0 *must* be included for any "missing" terms in $p(x)$. If $\deg p(x)=n$, there will be $n+1$ coefficients. In the example above, $\deg p(x)=4$ and there are 5 coefficients, for x^4, x^3, x^2, x, and the constant term.

4. The coefficients must be written in descending degree order. Hence for dividing

$$p(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0 \quad \text{by} \quad x-a$$

   we write:

$$\begin{array}{r|cccccc} a & a_n & a_{n-1} & \cdots & a_1 & a_0 \\ \hline \end{array}$$

5. The numbers in each column above the line are *added* to obtain the number below the line, *not* subtracted.

6. If $\deg p(x)=n$ and $p(x)$ is divided by $x-a$, of degree 1, then the quotient $q(x)$ will be of degree $n-1$ and the remainder $r(x)$ must have degree less than 1. Hence $q(x)$ has exactly one less coefficient than $p(x)$, and $r(x)$ will always be a constant.

Example: Divide $p(x)=3x^6+3x^5+2x^4+6x^3-x+2$ by $a(x)=x+1$. Since there is no degree 2 term in $p(x)$, the coefficient of x^2 in $p(x)$ is 0 and

$$p(x)=3x^6+3x^5+2x^4+6x^3+0x^2-x+2.$$

Since $a(x)=x+1=x-(-1)$, the zero of $a(x)$ is -1. We therefore start with:

$$\begin{array}{r|ccccccc} -1 & 3 & 3 & 2 & 6 & 0 & -1 & 2 \\ \hline \end{array}$$

Following the steps described above, we obtain:

$$\begin{array}{r|ccccccc} -1 & 3 & 3 & 2 & 6 & 0 & -1 & 2 \\ \hline & & -3 & 0 & -2 & -4 & 4 & -3 \\ \hline & 3 & 0 & 2 & 4 & -4 & 3 & -1 \end{array}$$

The quotient $q(x)$ is of degree $6-1=5$, with coefficients 3, 0, 2, 4, -4, and 3, hence

$$\begin{aligned} q(x) &= 3x^5+0x^4+2x^3+4x^2-4x+3 \\ &= 3x^5+2x^3+4x^2-4x+3. \end{aligned}$$

The remainder is $r(x)=-1$. Hence, we have

$$3x^6+3x^5+2x^4+6x^3-x+2=(3x^5+2x^3+4x^2-4x+3)(x+1)-1.$$

§3.4 Exercises

Use synthetic division to divide $p(x)$ by $a(x)$ and find the quotient $q(x)$ and remainder $r(x)$.

1. $p(x) = 3x^3 - 4x^2 - 2x + 7$,
   $a(x) = x - 2$
2. $p(x) = x^5 - 2x^4 - 5x^3 + 3x^2 + 4x + 1$,
   $a(x) = x + 3$
3. $p(x) = 2x^4 + 3x^3 - 2x^2 + 4x - 3$,
   $a(x) = x - \frac{1}{2}$
4. $p(x) = 3x^3 - 4x^2 + 3x - 7$,
   $a(x) = x + \frac{1}{3}$
5. $p(x) = 3x^3 - 4x - 5$,
   $a(x) = x - 3$
6. $p(x) = 2x^5 - 4x^3 + 2x + 3$,
   $a(x) = x - 1$
7. $p(x) = x^7 - x^6 + 2x^5 - x^3 + x^2 + 2x + 4$,
   $a(x) = x + 1$

3.5 Factors and Roots of Polynomials

Up to this point, we have viewed polynomials essentially as formal algebraic objects and have studied the algebraic properties of addition, multiplication, and division of polynomials. We now will also view polynomials as functions from the ring S to itself. Unless stated otherwise, S will denote one of $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{Z}_p$, p a prime, hence S is a field.

Definition 3.5.1 *If $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ is a polynomial in $S[x]$ and b is an element of S, then $p(x)$* **evaluated** *at b is*

$$p(b) = a_n b^n + a_{n-1}b^{n-1} + \cdots + a_1 b + a_0.$$

If $p(b) = 0$, we say b is a **root** *of $p(x)$ or a* **zero** *of $p(x)$.*

Since S is closed under addition and multiplication, $p(b)$ is an element of S. Thus $p(x)$ represents the function from S to S sending an element b in S to the element $p(b)$ in S. Evaluation of polynomials is compatible with addition and multiplication of polynomials in the sense of the next theorem, whose proof is straightforward but tedious.

Theorem 3.5.2 *Let $f(x)$ and $g(x)$ be polynomials in $S[x]$, and let $s(x) = f(x) + g(x)$ and $p(x) = f(x)g(x)$ be their sum and product, respectively. If b is an element of S, then $s(b) = f(b) + g(b)$ and $p(b) = f(b)g(b)$.*

This result is not as vacuous, or as obvious, as it may appear. For the product, it says that if we multiply the *polynomials* $f(x)$ and $g(x)$ to obtain the polynomial $p(x)$, and then evaluate $p(b)$, we get the same result as we would by evaluating $f(b)$ and $g(b)$, and then multiplying the *elements* $f(b)$ and $g(b)$ of S. For example, let $f(x) = 3x^2 + 5x + 2$ and $g(x) = x^2 - 4$, and let $b = 3$. Then

$$p(x) = f(x)g(x) = 3x^4 + 5x^3 - 10x^2 - 20x - 8.$$

The theorem says that

$$p(3) = 3 \cdot 3^4 + 5 \cdot 3^3 - 10 \cdot 3^2 - 20 \cdot 3 - 8$$

is equal to

$$f(3)g(3) = (3 \cdot 3^2 + 5 \cdot 3 + 2)(3^2 - 4).$$

This very much depends on the fact that multiplication in the ring S of coefficients is commutative. If multiplication in S were non-commutative, the result would be false. (Can you see why?)

We noted previously that if a polynomial $p(x)$ is divided by $x - a$, then the remainder, being of degree less than 1 by the Division Algorithm, must be a constant, that is, an element of S. The next result specifies what this element of S must be.

Theorem 3.5.3 (Remainder Theorem) *Let $p(x)$ be a polynomial in $S[x]$ and let a be an element of S. The remainder on division of $p(x)$ by $x - a$ is $p(a)$. Hence, $p(x) = q(x)(x - a) + p(a)$.*

Proof. By the Division Algorithm, we have $p(x) = q(x)(x - a) + r(x)$ with $q(x), r(x) \in S[x]$ and $\deg r(x) < \deg (x - a) = 1$. Thus $r(x)$ is a constant, say $r(x) = r \in S$. By Theorem 3.5.2, we then have

$$p(a) = q(a)(a - a) + r = q(a) \cdot 0 + r = r,$$

and so $r = p(a)$ as claimed. □

This result gives us an often very efficient method for evaluating a polynomial $p(x)$ at an element a of S. If we use synthetic division to divide $p(x)$ by $x - a$, then $p(a)$ is the remainder.

Example: Evaluate $p(x) = 3x^6 + 3x^5 + 2x^4 + 6x^3 - x + 2$ at $x = 1$.

By the Remainder Theorem, $p(1)$ is the remainder on dividing $p(x)$ by $x - 1$, which we can find by synthetic division:

$$\begin{array}{c|ccccccc} 1 & 3 & 3 & 2 & 6 & 0 & -1 & 2 \\ \hline & & 3 & 6 & 8 & 14 & 14 & 13 \\ \hline & 3 & 6 & 8 & 14 & 14 & 13 & 15 \end{array}$$

The remainder is 15 and therefore $p(1) = 15$. We can verify that this method worked in this case by evaluating $p(1)$ directly:

$$p(1) = 3 \cdot 1^6 + 3 \cdot 1^5 + 2 \cdot 1^4 + 6 \cdot 1^3 - 1 + 2 = 3 + 3 + 2 + 6 - 1 + 2 = 15,$$

and so $p(1) = 15$ as claimed. □

Since a polynomial $f(x)$ divides $p(x)$ if and only if the remainder on division of $p(x)$ by $f(x)$ is 0, the Remainder Theorem implies the following.

Theorem 3.5.4 (Factor Theorem) *Let $p(x)$ be a polynomial in $S[x]$ and let a be an element of S. Then $x - a \mid p(x)$ if and only if $p(a) = 0$. That is, $x - a$ is a* factor *of $p(x)$ if and only if a is a* root *of $p(x)$.*

Proof. By the Remainder Theorem, $p(a)$ is the remainder on dividing $p(x)$ by $x-a$. Since $x-a \mid p(x)$ if and only if the remainder is 0, this implies $x - a \mid p(x)$ if and only if $p(a) = 0$. □

Example: Let $p(x) = x^4 - 3x^2 - 5x - 3$. We have

$$p(1) = 1^4 - 3 \cdot 1^2 - 5 \cdot 1 - 3 = 1 - 3 - 5 - 3 = -10,$$

hence 1 is not a root and $x - 1$ is not a factor of $p(x)$. However,

$$p(-1) = (-1)^4 - 3 \cdot (-1)^2 - 5 \cdot (-1) - 3 = 1 - 3 + 5 - 3 = 0,$$

and so -1 is a root and $x - (-1) = x + 1$ is a factor of $p(x)$. Note that although these values were easy to compute directly, we would obtain more information by using synthetic division to evaluate $p(-1)$. We have

$$\begin{array}{c|ccccc} -1 & 1 & 0 & -3 & -5 & -3 \\ \hline & & -1 & 1 & 2 & 3 \\ \hline & 1 & -1 & -2 & -3 & 0 \end{array}$$

which shows that $x + 1$ is a factor of $p(x)$, but also shows that the other factor is $x^3 - x^2 - 2x - 3$. Hence $p(x) = (x + 1)(x^3 - x^2 - 2x - 3)$. □

Recall that if $S = \mathbb{R}$, and $p(x)$ is a polynomial in $\mathbb{R}[x]$, the **graph** of $p(x)$ is the set of points $(a, p(a))$ in the xy-plane. The x-**intercepts** of $p(x)$ are the x coordinates of the points where the graph of $p(x)$ intersects the x-axis. Thus the x-intercepts are the numbers $x = a$ such that $p(a) = 0$. We therefore have the following corollary.

Corollary 3.5.5 *Let $p(x)$ be a polynomial in $\mathbb{R}[x]$ and let a be a real number. The following are equivalent.*

i. *$x - a$ is a factor of $p(x)$.*

ii. *a is a root of $p(x)$.*

iii. *$x = a$ is an x-intercept of $p(x)$.*

If a is a root of $p(x)$, it is possible that $(x-a)^m$ is a factor of $p(x)$ for some $m > 1$. In this case, we say a is a **repeated root** of $p(x)$. More specifically, we make the following definition.

Definition 3.5.6 *Let $p(x)$ be a polynomial in $S[x]$ and let a be a root of $p(x)$ in S. If $(x-a)^m$ is a factor of $p(x)$ but $(x-a)^{m+1}$ is not a factor of $p(x)$, we say a is a root of $p(x)$ of* **multiplicity** *m.*

The correspondence between factors and roots of $p(x)$ allows us to factor $p(x)$ if its roots are known, or to find the roots of $p(x)$ if the linear factors are known.

Theorem 3.5.7 *Let $p(x)$ be a polynomial in $S[x]$ of degree n, with leading coefficient a. If $p(x)$ has n roots $r_1, r_2, \ldots, r_n$ in S (including possibly repeated roots), then*

$$p(x) = a(x - r_1)(x - r_2) \cdots (x - r_n).$$

Sketch of Proof. First write $p(x) = a \cdot p_1(x)$, where $p_1(x) = \frac{1}{a}p(x)$ is monic. (Note that $p(x)$ and $p_1(x)$ have the same roots, as $p(r) = a \cdot p_1(r) = 0$ if and only if $p_1(r) = 0$.) By the Factor Theorem, since r_1 is a root of $p_1(x)$, we have

$$p(x) = a(x - r_1)q_1(x)$$

for some $q_1(x) \in S[x]$. Then r_2 is a root of $q_1(x)$ (see Exercise 3.5.9), and so

$$p(x) = a(x - r_1)(x - r_2)q_2(x)$$

for some $q_2(x) \in S[x]$. Continuing in this way yields the desired factorization. □

Corollary 3.5.8 *If $p(x)$ is a polynomial in $S[x]$ of degree n, then $p(x)$ has* at most *n roots in S.*

Proof. If $r_1, r_2, \ldots, r_m$ are roots of $p(x)$ in S, then

$$q(x) = (x - r_1)(x - r_2) \cdots (x - r_m)$$

is a factor of $p(x)$ of degree m. Hence by Theorem 3.3.5, we have

$$m = \deg\, q(x) \leqslant \deg\, p(x) = n,$$

and so $m \leqslant n$ and the number, m, of roots of $p(x)$ is at most n. □

For the remainder of this section, we will assume S is one of $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$. If $p(x)$ is a quadratic polynomial (i.e., of degree 2), the roots of $p(x)$ can be found with the Quadratic Formula.

Theorem 3.5.9 (Quadratic Formula) *Let $p(x) = ax^2 + bx + c$ be a polynomial in $S[x]$, with $a \neq 0$. The roots of $p(x)$ are then*

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Proof. We complete the square to solve the equation $ax^2 + bx + c = 0$. Dividing both sides by the leading coefficient a and then subtracting $\frac{c}{a}$ from both sides yields

$$x^2 + \frac{b}{a}\,x = -\frac{c}{a}.$$

Complete the square on the left side by adding $\left(\frac{b}{2a}\right)^2$ to both sides to obtain

$$x^2 + \frac{b}{a}\,x + \left(\frac{b}{2a}\right)^2 = \left(\frac{b}{2a}\right)^2 - \frac{c}{a}.$$

The left side is now a perfect square and we have

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2}{4a^2} - \frac{c}{a} = \frac{b^2 - 4ac}{4a^2}.$$

Taking square roots of both sides yields

$$x + \frac{b}{2a} = \pm\sqrt{\frac{b^2 - 4ac}{4a^2}} = \frac{\pm\sqrt{b^2 - 4ac}}{2a}.$$

Finally, subtracting $\frac{b}{2a}$ from both sides, we obtain

$$x = -\frac{b}{2a} + \frac{\pm\sqrt{b^2 - 4ac}}{2a} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

as claimed. □

There are similar, but much more complicated, formulas for the roots of cubic (degree 3) and quartic (degree 4) polynomials in terms of radicals involving the coefficients. It can be proved that no such general formula can exist for polynomials of degree 5 or higher.

The number $b^2 - 4ac$ is called the **discriminant** of $p(x)$. Note that the roots given by the formula are in S if and only if $b^2 - 4ac$ is the square of an element of S. Hence there are three possibilities:

1. If $b^2 - 4ac$ is not the square of an element of S, then $p(x)$ has no roots in S.
2. If $b^2 - 4ac \neq 0$ and is the square of an element of S, then $p(x)$ has two distinct roots in S.
3. If $b^2 - 4ac = 0$, then $p(x)$ has one root in S of multiplicity 2.

In particular, suppose $p(x) = ax^2 + bx + c$ is a quadratic polynomial with integer coefficients. The roots of $p(x)$ are rational if and only if $b^2 - 4ac$ is a perfect square, that is, the square of an integer. The roots of $p(x)$ are real if and only if $b^2 - 4ac \geqslant 0$. In any case, the roots of $p(x)$ are complex.

The Quadratic Formula, along with Theorem 3.5.7, allows us to factor any quadratic polynomial with roots in S. If $p(x) = ax^2 + bx + c$ has roots r_1 and r_2 in S, then

$$p(x) = a(x - r_1)(x - r_2).$$

Thus the problem of factoring a quadratic polynomial or finding its roots is solved by the formula.

Examples:

1. Factor $p(x) = 2x^2 + x + 5$.

   Note that the discriminant is $1^2 - 4(2)(5) = -39 < 0$, so there are no real roots. The roots are

   $$x = \frac{-1 \pm \sqrt{-39}}{4},$$

   that is,

   $$x = -\frac{1}{4} + \frac{\sqrt{39}}{4}\, i \ \text{ and } \ x = -\frac{1}{4} - \frac{\sqrt{39}}{4}\, i.$$

   Therefore, since the leading coefficient of $p(x)$ is 2, we have

   $$p(x) = 2x^2 + x + 5 = 2\left[x - \left(-\frac{1}{4} + \frac{\sqrt{39}}{4}\, i\right)\right]\left[x - \left(-\frac{1}{4} - \frac{\sqrt{39}}{4}\, i\right)\right].$$

2. Factor $p(x) = 6x^2 - 19x - 36$.

   In this case, the discriminant is $(-19)^2 - 4(6)(-36) = 1225 = 35^2$, hence $p(x)$ has two distinct *rational* roots. The roots are

   $$x = \frac{19 \pm \sqrt{1225}}{12} = \frac{19 \pm 35}{12},$$

   that is,

   $$x = \frac{54}{12} = \frac{9}{2} \ \text{ and } \ x = -\frac{16}{12} = -\frac{4}{3}.$$

   Therefore, since the leading coefficient of $p(x)$ is 6, we have

   $$p(x) = 6x^2 - 19x - 36 = 6\left[x - \frac{9}{2}\right]\left[x - \left(-\frac{4}{3}\right)\right] = 6\left(x - \frac{9}{2}\right)\left(x + \frac{4}{3}\right).$$

   We will further refine this factorization below. □

Polynomials with Integer Coefficients

In general, finding exact values of irrational or complex roots of a polynomial is a difficult problem. The problem of finding rational roots of a polynomial with integer coefficients, if they exist, is relatively easy. In this section, we will state some results that are very useful for finding roots and factors of polynomials with integer coefficients.

Note that if $p(x)$ has rational coefficients and m is the LCM of the denominators of the coefficients of $p(x)$, then $m \cdot p(x)$ has integer coefficients. The polynomials $p(x)$ and $m \cdot p(x)$ have the same roots (why?), hence these results can also be adapted to find roots of polynomials with rational coefficients.

We saw above how to factor a quadratic polynomial $p(x) = ax^2 + bx + c$ with roots r_1 and r_2. If the coefficients of $p(x)$ are integers and the roots r_1 and r_2 are rational (not necessarily integers), then it is actually possible to find *integers* a_1, b_1, a_2, and b_2 such that

$$p(x) = a(x - r_1)(x - r_2) = (a_1x - b_1)(a_2x - b_2).$$

Thus $a = a_1a_2$, $b_1 = a_1r_1$, and $b_2 = a_2r_2$.

Example: In the example above, we obtained

$$p(x) = 6x^2 - 19x - 36 = 6\left(x - \frac{9}{2}\right)\left(x + \frac{4}{3}\right).$$

In this case, $a = 6$, $r_1 = 9/2$ and $r_2 = -4/3$. If we let $a_1 = 2$ and $a_2 = 3$, so that

$$a = 6 = 2 \cdot 3 = a_1a_2,$$

then

$$b_1 = a_1r_1 = 2(9/2) = 9$$

and

$$b_2 = a_2r_2 = 3(-4/3) = -4,$$

and we have

$$p(x) = 6\left(x - \frac{9}{2}\right)\left(x + \frac{4}{3}\right) = 2\left(x - \frac{9}{2}\right) \cdot 3\left(x + \frac{4}{3}\right) = (2x - 9)(3x + 4).$$

Thus the polynomial $p(x)$ can be written as a product of factors with integer coefficients. □

This says that if a quadratic polynomial with integer coefficients can be factored as a product of polynomials with rational coefficients, then it can be factored as a product of polynomials with integer coefficients. This is, in fact, true for polynomials of any degree by the following result, known as Gauss's Lemma.

Theorem 3.5.10 (Gauss's Lemma) *Let $p(x)$ be a polynomial with integer coefficients. If $p(x) = A(x)B(x)$, where $A(x)$ and $B(x)$ are non-constant polynomials with rational coefficients, then there exist rational numbers r and s such that $a(x) = rA(x)$ and $b(x) = sB(x)$ have integer coefficients, and $p(x) = a(x)b(x)$.*

The proof of Gauss's Lemma requires some concepts we have not discussed and will be omitted here. A proof can be found in almost any book on abstract algebra. The following example illustrates the notation in the theorem.

Example: Suppose we factor the polynomial $p(x) = 6x^5 + 11x^3 + 12x^2 - 10x - 8$ as

$$p(x) = 6x^5 + 11x^3 + 12x^2 - 10x - 8 = \left(\frac{15}{2}\,x^2 - 5\right)\left(\frac{4}{5}\,x^3 + 2x + \frac{8}{5}\right),$$

so that $A(x) = \left(\frac{15}{2}\,x^2 - 5\right)$ and $B(x) = \left(\frac{4}{5}\,x^3 + 2x + \frac{8}{5}\right)$ in the notation of Gauss's Lemma. Thus $p(x) \in \mathbb{Z}[x]$ is factored into a product of polynomials with *rational* coefficients. If we let $r = 2/5$ and $s = 5/2$, then

$$a(x) = rA(x) = \frac{2}{5}\left(\frac{15}{2}\,x^2 - 5\right) = 3x^2 - 2$$

and

$$b(x) = sB(x) = \frac{5}{2}\left(\frac{4}{5}\,x^3 + 2x + \frac{8}{5}\right) = 2x^3 + 5x + 4.$$

Thus, since $rs = (2/5)(5/2) = 1$, we have

$$\begin{aligned} p(x) &= A(x)B(x) \\ &= \left(\frac{15}{2}\,x^2 - 5\right)\left(\frac{4}{5}\,x^3 + 2x + \frac{8}{5}\right) \\ &= \frac{2}{5}\left(\frac{15}{2}\,x^2 - 5\right)\frac{5}{2}\left(\frac{4}{5}\,x^3 + 2x + \frac{8}{5}\right) \\ &= (3x^2 - 2)(2x^3 + 5x + 4) \\ &= a(x)b(x). \end{aligned}$$

Hence

$$p(x) = 6x^5 + 11x^3 + 12x^2 - 10x - 8 = (3x^2 - 2)(2x^3 + 5x + 4)$$

and we have written $p(x)$ as a product of polynomials with *integer* coefficients. □

According to the Factor Theorem, in order to factor a polynomial completely, we need to find all of its roots. For a polynomial $p(x)$ with integer coefficients, it is best to try to find all rational roots first. The next result restricts this search to a relatively small number of potential roots.

Theorem 3.5.11 (Rational Root Test) *Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial with integer coefficients. If r/s is a rational root of $p(x)$ with $(r, s) = 1$, then $r \mid a_0$ and $s \mid a_n$.*

Proof. Let r/s be a rational root of $p(x)$, where $r, s \in \mathbb{Z}$ and $(r, s) = 1$. Since r/s is a root of $p(x)$, we have

$$\begin{aligned} 0 &= p(r/s) \\ &= a_n\left(\frac{r}{s}\right)^n + a_{n-1}\left(\frac{r}{s}\right)^{n-1} + \cdots + a_1\left(\frac{r}{s}\right) + a_0 \\ &= a_n\frac{r^n}{s^n} + a_{n-1}\frac{r^{n-1}}{s^{n-1}} + \cdots + a_1\frac{r}{s} + a_0. \end{aligned}$$

Multiplying both sides of the equation by s^n yields

$$a_n r^n + a_{n-1} r^{n-1} s + a_{n-2} r^{n-2} s^2 + \cdots + a_2 r^2 s^{n-2} + a_1 r s^{n-1} + a_0 s^n = 0. \quad (*)$$

Thus we have

$$\begin{aligned} a_n r^n &= -(a_{n-1} r^{n-1} s + a_{n-2} r^{n-2} s^2 + \cdots + a_2 r^2 s^{n-2} + a_1 r s^{n-1} + a_0 s^n) \\ &= -s(a_{n-1} r^{n-1} + a_{n-2} r^{n-2} s + \cdots + a_2 r^2 s^{n-3} + a_1 r s^{n-2} + a_0 s^{n-1}), \end{aligned}$$

which implies $s \mid a_n r^n$. If any prime divides both s and r^n, then it also divides r by Theorem 2.5.6. Since $(s, r) = 1$, there is no such prime and it follows that $(s, r^n) = 1$. Now by Euclid's Lemma (Theorem 2.4.15), since $s \mid a_n r^n$ and $(s, r^n) = 1$, we have $s \mid a_n$.

Similarly, Equation $(*)$ implies

$$a_0 s^n = -r(a_n r^{n-1} + a_{n-1} r^{n-2} s + a_{n-2} r^{n-3} s^2 + \cdots + a_2 r s^{n-2} + a_1 s^{n-1}),$$

and so $r \mid a_0 s^n$. As before, $(r, s) = 1$ implies $(r, s^n) = 1$, and Euclid's Lemma implies $r \mid a_0$. □

Letting $a_n = 1$ in the theorem, we have the following important special case of the Rational Root Test.

Corollary 3.5.12 *If $p(x) = x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ is a* monic *polynomial with integer coefficients, then any rational root d of $p(x)$ is an integer, and $d \mid a_0$.*

Proof. If $d = r/s$ is a rational root of $p(x)$ with $(r, s) = 1$, then by Theorem 3.5.11, $s \mid 1$ and $r \mid a_0$. Thus $s = \pm 1$ and $d = \pm r$ is an integer with $d \mid a_0$. □

The Rational Root Test gives us a procedure to find all rational roots of $p(x)$. Find all positive divisors r of the constant term a_0 and all positive divisors s of the leading coefficient a_n. The only potential rational roots are then the numbers $\pm r/s$. Evaluate $p(x)$ at each of these values of x to determine which are actually roots. (In general, many of these numbers will *not* be roots.)

Examples:

1. Find all potential rational roots of $p(x) = x^5 - 5x^3 + 2x^2 - 7x + 6$.

   By Corollary 3.5.12, all rational roots are integer divisors of the constant term, 6. Therefore, the potential rational roots are ± 1, ± 2, ± 3, and ± 6.

2. Find all potential rational roots of $p(x) = 2x^8 + 3x^4 + 12$.

   By the Rational Root Test, the potential rational roots are the numbers $\pm r/s$, where r is a positive divisor of the constant term, 12, and s is a divisor of the leading coefficient, 2. Thus

   $$r = 1,\ 2,\ 3,\ 4,\ 6,\ \text{or } 12$$

   and

   $$s = 1 \text{ or } 2.$$

   Therefore, the potential rational roots are

   $$\pm\frac{r}{s} = \pm 1,\ \pm 2,\ \pm 3,\ \pm 4,\ \pm 6,\ \pm 12,\ \pm\frac{1}{2},\ \pm\frac{3}{2}.$$

   Of course, not all of the potential rational roots will be actual roots. □

By Corollary 3.5.5, the graph of $p(x)$ may be used to help locate the roots. After finding the potential rational roots, graph the function $y = p(x)$ on a graphing calculator or computer. The actual roots are the x-intercepts, so any potential rational roots that are clearly not x-intercepts can be eliminated from the list of potential roots.

Note that for those numbers a that *appear* to be x-intercepts on the graph, it is still necessary to verify computationally that $p(a) = 0$ or algebraically that $x - a$ is a factor of $p(x)$. The graph on a calculator is merely an approximation of the graph of the function. It suggests numbers that *may* be roots, but it is impossible to determine with certainty which numbers *are* roots from such a graph without computational or algebraic verification. Moreover, $p(x)$ could have irrational roots, and it is generally not possible to guess the exact values of irrational roots from even a very accurate graph.

The procedure can also be simplified somewhat by factoring after each root is found. If r/s is a rational root of $p(x)$ with $(r, s) = 1$, then Gauss's Lemma implies $p(x)$ can be factored as $p(x) = (sx - r)q(x)$, where $q(x)$ also has integer coefficients. The roots of $p(x)$ and $q(x)$, other than possibly r/s, are the same. Since $q(x)$ has lower degree than $p(x)$, it may be easier to factor. Moreover, the leading coefficient and constant term of $q(x)$ will be divisors of those of $p(x)$, so $q(x)$ may have fewer potential rational roots than $p(x)$.

Examples:

1. Factor $p(x) = 9x^3 - 3x^2 - 5x + 2$ completely.

   By the Rational Root Test, the potential rational roots are the numbers $\pm r/s$, where r is a positive divisor of the constant term, 2, and s is a positive divisor of the leading coefficient, 9. Thus $r = 1$ or 2 and $s = 1$, 3, or 9. Therefore, the potential rational roots are

   $$\pm\frac{r}{s} = \pm 1,\ \pm 2,\ \pm\frac{1}{3},\ \pm\frac{2}{3},\ \pm\frac{1}{9},\ \pm\frac{2}{9}.$$

   Graphing the function $p(x) = 9x^3 - 3x^2 - 5x + 2$ on a calculator, it appears that possibly $1/3$ and $2/3$ may be roots. We use synthetic division to check whether either of these is a root. First, we check $1/3$:

   $$\begin{array}{r|rrrr} 1/3 & 9 & -3 & -5 & 2 \\ & & 3 & 0 & -5/3 \\ \hline & 9 & 0 & -5 & 1/3 \end{array}$$

   showing that $1/3$ is *not* a root (as the remainder is not 0).

   Next, we check $2/3$:

   $$\begin{array}{r|rrrr} 2/3 & 9 & -3 & -5 & 2 \\ & & 6 & 2 & -2 \\ \hline & 9 & 3 & -3 & 0 \end{array}$$

   showing that $2/3$ is a root and that

   $$\begin{aligned} p(x) &= \left(x - \frac{2}{3}\right)(9x^2 + 3x - 3) \\ &= \left(x - \frac{2}{3}\right) \cdot 3(3x^2 + x - 1) \\ &= (3x - 2)(3x^2 + x - 1). \end{aligned}$$

Finally, we factor the quadratic polynomial $q(x) = 3x^2 + x - 1$. Observe first that the discriminant of $q(x)$ is $1^2 - 4(3)(-1) = 13$, which is positive but is not the square of an integer. Therefore, the remaining roots are real but irrational. By the Quadratic Formula, the roots are

$$x = \frac{-1 + \sqrt{13}}{6} \text{ and } x = \frac{-1 - \sqrt{13}}{6}.$$

Therefore, since the leading coefficient of $q(x)$ is 3, we have

$$q(x) = 3x^2 + x - 1 = 3\left[x - \left(\frac{-1 + \sqrt{13}}{6}\right)\right]\left[x - \left(\frac{-1 - \sqrt{13}}{6}\right)\right],$$

and we have

$$p(x) = (3x - 2)(3x^2 + x - 1) = 3(3x - 2)\left[x - \left(\frac{-1 + \sqrt{13}}{6}\right)\right]\left[x - \left(\frac{-1 - \sqrt{13}}{6}\right)\right]$$

over the real numbers.

2. Factor $p(x) = x^4 + 2x^3 - 10x^2 - 41x - 60$ completely.

By Corollary 3.5.12, the potential rational roots of $p(x)$ are the integer divisors of 60, hence are

$$\pm 1,\ \pm 2,\ \pm 3,\ \pm 4,\ \pm 5,\ \pm 6,\ \pm 10,\ \pm 12,\ \pm 15,\ \pm 20,\ \pm 30,\ \pm 60.$$

Graphing the function $p(x) = x^4 + 2x^3 - 10x^2 - 41x - 60$ on a calculator, it appears that possibly -3 and 4 may be roots. We use synthetic division to check whether either of these is a root. First, we check -3:

$$\begin{array}{r|rrrrr} -3 & 1 & 2 & -10 & -41 & -60 \\ & & -3 & 3 & 21 & 60 \\ \hline & 1 & -1 & -7 & -20 & 0 \end{array}$$

showing that -3 is a root and $p(x) = (x + 3)(x^3 - x^2 - 7x - 20)$.

If 4 is a root of $p(x)$, then it is also a root of $q(x) = x^3 - x^2 - 7x - 20$. We check whether 4 is a root:

$$\begin{array}{r|rrrr} 4 & 1 & -1 & -7 & -20 \\ & & 4 & 12 & 20 \\ \hline & 1 & 3 & 5 & 0 \end{array}$$

showing that 4 is a root of $q(x)$ and $q(x) = (x - 4)(x^2 + 3x + 5)$. Thus

$$p(x) = (x + 3)(x^3 - x^2 - 7x - 20) = (x + 3)(x - 4)(x^2 + 3x + 5).$$

Finally, we factor the quadratic polynomial $r(x) = x^2 + 3x + 5$. Observe first that the discriminant of $r(x)$ is $3^2 - 4(1)(5) = -11 < 0$, and so the remaining roots are complex and not real. By the Quadratic Formula, the roots are

$$x = \frac{-3 \pm \sqrt{-11}}{2},$$

that is,

$$x = -\frac{3}{2} + \frac{\sqrt{11}}{2}\,i \;\text{ and }\; x = -\frac{3}{2} - \frac{\sqrt{11}}{2}\,i.$$

Therefore, since $r(x)$ is monic, we have

$$r(x) = x^2 + 3x + 5 = \left[x - \left(-\frac{3}{2} + \frac{\sqrt{11}}{2}\,i\right)\right]\left[x - \left(-\frac{3}{2} - \frac{\sqrt{11}}{2}\,i\right)\right],$$

and so

$$p(x) = (x+3)(x-4)(x^2+3x+5) = (x+3)(x-4)\left[x - \left(-\frac{3}{2} + \frac{\sqrt{11}}{2}\,i\right)\right]\left[x - \left(-\frac{3}{2} - \frac{\sqrt{11}}{2}\,i\right)\right]$$

over the complex numbers. □

---

§3.5 Exercises

1. Use synthetic division to find the quotient and remainder when $p(x)$ is divided by $f(x)$.
   (a) $p(x) = 3x^3 - 12x^2 - 9x + 1,\ f(x) = x - 5.$
   (b) $p(x) = x^5 - 9x^3 + 2x^2 + x - 11,\ f(x) = x + 3.$

2. Use synthetic division and the Remainder Theorem to evaluate $p(a)$ and determine if $x - a$ is a factor of $p(x)$.
   (a) $p(x) = x^3 - x^2 + x + 5,\ a = -1.$
   (b) $p(x) = x^4 + 3x^3 - 16x^2 - 27x + 63,\ a = 3.$

3. Use the Quadratic Formula to help in factoring $p(x)$. If the roots are rational, write the factors with *integer* coefficients. If the roots are real but not rational, write the factors with real number coefficients. If the roots are not real, factor over $\mathbb{C}$.
   (a) $p(x) = 6x^2 - 11x - 72$
   (b) $p(x) = 3x^2 + 5x - 7$
   (c) $p(x) = 2x^2 + 5x + 4$

4. Find all *potential* rational roots of $p(x)$. (You need not determine which are actual roots.)
   (a) $p(x) = x^5 - 4x^3 + 18$
   (b) $p(x) = 3x^8 + 6x^2 - 5x - 10$
   (c) $p(x) = 4x^4 + 3x^3 + 2x^2 + 6$

5. Factor the following polynomials completely (over $\mathbb{C}$, if necessary).
   (a) $p(x) = 2x^3 + 3x^2 - 32x + 15$
   (b) $p(x) = x^3 + 5x^2 + x - 10$
6. Factor $p(x) = 3x^4 + 5x^3 - 2x^2 - 3x + 1$ completely (over $\mathbb{C}$, if necessary).
7. Factor $p(x) = 2x^4 - 11x^3 + 20x^2 - 14x + 3$ completely (over $\mathbb{C}$, if necessary).
8. Factor $p(x) = 2x^4 - 17x^3 + 43x^2 - 37x + 6$ completely (over $\mathbb{C}$, if necessary).
9. Let $p(x) = (x - r)q(x)$, where $p(x)$ and $q(x)$ are polynomials in $S[x]$ and r is an element of S. Show the following.
   (a) If a is a root of $p(x)$ with $a \neq r$, then a is a root of $q(x)$.
   (b) If b is any root of $q(x)$, then b is a root of $p(x)$.

   [Hint: Use the definition of root, *not* the Factor Theorem.]
10. Let $f(x)$ and $g(x)$ be polynomials in $S[x]$. Prove that if $(f(x), g(x)) = 1$, then $f(x)$ and $g(x)$ have no root in common.
11. Let $p(x)$ be a polynomial in $S[x]$. Show that if $p(x)$ has a repeated root r, then r is also a root of the derivative $p'(x)$ of $p(x)$.
    [Hint: Write $p(x) = (x - r)^2 q(x)$, where $q(x)$ is in $S[x]$. Use the product rule to find $p'(x)$, and show r is a root of $p'(x)$.]

3.6 Irreducible Polynomials

We will now study polynomials that cannot be factored, called irreducible polynomials. The irreducible polynomials will play essentially the same role in our study of polynomials that prime numbers did in our study of the integers. In this section, we will primarily be concerned with deriving criteria which will help us to recognize irreducible polynomials.

Definition 3.6.1 *Let $p(x)$ be a non-constant polynomial in $S[x]$. We say $p(x)$ is* **irreducible over** *S (or* **irreducible in** *$S[x]$) if $p(x)$ cannot be factored as $p(x) = A(x)B(x)$ with $A(x)$, $B(x)$ non-constant polynomials in $S[x]$. If $p(x)$ has such a factorization, we say $p(x)$ is* **reducible over** *S.*

In other words, $p(x)$ is irreducible if the only divisors of $p(x)$ are constants and constant multiples of $p(x)$. Note that $p(x)$ is reducible over S if and only if there are polynomials $A(x)$ and $B(x)$ in $S[x]$ such that $1 \leqslant \deg A(x) < \deg p(x)$, $1 \leqslant \deg B(x) < \deg p(x)$, and $p(x) = A(x)B(x)$.

We should also point out that for polynomials in $\mathbb{Z}[x]$, the definition of irreducible polynomial given here is slightly different than that usually given for an "irreducible element" in a ring. Our definition would say that the polynomial $2x - 2$ is irreducible over $\mathbb{Z}$. The ring theoretic definition would say that this polynomial is *not* an irreducible element of $\mathbb{Z}[x]$ because $2x - 2 = 2(x - 2)$, and the factor 2 does not have a multiplicative inverse in $\mathbb{Z}[x]$, hence is a "legitimate" factor. Our definition does agree with the standard definition in case S is a field, however, and the slight difference in case $S = \mathbb{Z}$ will cause no difficulties.

Note: Whether or not $p(x)$ is irreducible depends very much on the ring S of coefficients. For example, $x^2 - 2$ is irreducible over $\mathbb{Q}$, but $x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$ over $\mathbb{R}$ or $\mathbb{C}$. Similarly, $x^2 + 1$ is irreducible over $\mathbb{Q}$ and over $\mathbb{R}$, but is reducible over $\mathbb{C}$, as $x^2 + 1 = (x - i)(x + i)$.

For now, the ring S of coefficients will be one of $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{Z}_p$, p a prime, hence S is always a *field*. Although we will consider polynomials with integer coefficients, there will be no loss in considering such polynomials to be in $\mathbb{Q}[x]$. This is due to the fact that irreducibility over $\mathbb{Z}$ and over $\mathbb{Q}$ are equivalent by the next result, which follows from Gauss's Lemma.

Theorem 3.6.2 *A polynomial $p(x)$ with integer coefficients is irreducible over $\mathbb{Z}$ if and only if $p(x)$ is irreducible over $\mathbb{Q}$.*

Proof. Let $p(x)$ be a polynomial with integer coefficients. We prove the equivalent statement that $p(x)$ is reducible over $\mathbb{Z}$ if and only if $p(x)$ is reducible over $\mathbb{Q}$.

If $p(x)$ is reducible over $\mathbb{Z}$, then $p(x) = a(x)b(x)$, where $a(x)$ and $b(x)$ are non-constant polynomials in $\mathbb{Z}[x]$. Since $\mathbb{Z}[x]$ is contained in $\mathbb{Q}[x]$, we have $a(x), b(x) \in \mathbb{Q}[x]$ as well, and so $p(x)$ is reducible over $\mathbb{Q}$.

Conversely, if $p(x)$ is reducible over $\mathbb{Q}$, then $p(x) = A(x)B(x)$, where $A(x)$ and $B(x)$ are non-constant polynomials in $\mathbb{Q}[x]$. By Gauss's Lemma (Theorem 3.5.10), there exist polynomials $a(x)$ and $b(x)$ in $\mathbb{Z}[x]$ that are constant multiples of $A(x)$ and $B(x)$, respectively, hence are non-constant, such that $p(x) = a(x)b(x)$. Therefore, $p(x)$ is reducible over $\mathbb{Z}$. □

Recall that since we are assuming S is a field, multiplication by a non-zero constant has no effect on divisibility of polynomials in $S[x]$. It follows that multiplication by a non-zero constant has no effect on irreducibility.

Theorem 3.6.3 *Let $p(x)$ be a polynomial in $S[x]$ and let k be a non-zero element of S. Then $k \cdot p(x)$ is irreducible over S if and only if $p(x)$ is irreducible over S.*

We next state some results that are useful in determining whether a given polynomial is irreducible. The first follows immediately from the definition and the fact that the degree of a product is the sum of the degrees of the factors.

Theorem 3.6.4 *Every polynomial in $S[x]$ of degree 1 is irreducible over S.*

The Factor Theorem implies the following result.

Theorem 3.6.5 *Let $p(x)$ be a polynomial in $S[x]$ with $\deg p(x) > 1$. If $p(x)$ has a root in S then $p(x)$ is reducible over S.*

Proof. By the Factor Theorem (Theorem 3.5.4), if $a \in S$ is a root of $p(x)$, then $p(x) = (x-a)b(x)$ for some $b(x) \in S[x]$. Hence, by Theorem 3.1.8,

$$1 < \deg\, p(x) = \deg\,(x-a)b(x) = \deg\,(x-a) + \deg\, b(x) = 1 + \deg\, b(x),$$

and so $\deg\, b(x) > 0$. Thus both $x - a$ and $b(x)$ are non-constant polynomials in $S[x]$, and hence $p(x)$ is reducible over S. □

The theorem implies that if $p(x)$ is irreducible over S, then $p(x)$ has no root in S. The converse is false in general, however. If $p(x)$ has no root in S, it is not necessarily true that $p(x)$ is irreducible. For example, $p(x) = x^4 - 4 = (x^2-2)(x^2+2)$ has no roots in $\mathbb{Q}$, but $p(x)$ is clearly not irreducible over $\mathbb{Q}$. Thus showing that a polynomial does not have a root in S is not sufficient to prove that the polynomial is irreducible, in general. If the degree of the polynomial is small enough, however, this will be sufficient to prove irreducibility.

Theorem 3.6.6 *A polynomial $p(x)$ in $S[x]$ of degree 2 or 3 is irreducible over S if and only if $p(x)$ has no root in S.*

Proof. Let $p(x)$ be a polynomial in $S[x]$ of degree 2 or 3. We prove the equivalent statement that $p(x)$ has a root in S if and only if $p(x)$ is reducible over S. If $p(x)$ has a root in S, then, since $\deg\, p(x) > 1$, Theorem 3.6.5 implies that $p(x)$ is reducible over S.

Conversely, if $p(x)$ is reducible over S, then $p(x) = f(x)g(x)$ for some $f(x), g(x) \in S[x]$ with $\deg\, f(x) \geqslant 1$ and $\deg\, g(x) \geqslant 1$. Now if both $\deg\, f(x)$ and $\deg\, g(x)$ are *greater* than 1, hence are at least 2, then

$$\deg\, p(x) = \deg\, f(x)g(x) = \deg\, f(x) + \deg\, g(x) \geqslant 4.$$

Therefore, since the degree of $p(x)$ is 2 or 3, at least one of $f(x)$, $g(x)$ is of degree 1.

We may assume without loss of generality that $\deg\, f(x) = 1$, and so $f(x) = ax + b$ for some $a, b \in S$ with $a \neq 0$. Since S is a field, $-b/a$ is in S and

$$p(-b/a) = f(-b/a)g(-b/a) = (a(-b/a) + b)g(-b/a) = (-b+b)g(-b/a) = 0 \cdot g(-b/a) = 0.$$

Hence $-b/a \in S$ is a root of $p(x)$. □

Using this theorem and previous results on the quadratic formula and on rational roots of polynomials, it is easy to determine if a polynomial of degree 2 or 3 with integer coefficients is irreducible over $\mathbb{Q}$.

Examples:

1. The polynomial $p(x) = 2x^3 - 5x^2 + 1$ is irreducible over $\mathbb{Q}$.

   Proof. Since $\deg p(x) = 3$, Theorem 3.6.6 says we need only show that $p(x)$ has no roots in $\mathbb{Q}$. By the Rational Root Test (Theorem 3.5.11), the only possible rational roots are $\pm r/s$, where $r \mid 1$ and $s \mid 2$. Thus $r = 1$ and $s = 1$ or 2, and so the possible rational roots are ± 1 and $\pm 1/2$. It is easy to verify that

   $$p(1) = -2,\ p(-1) = 4,\ p(1/2) = -5/4,\ \text{and } p(-1/2) = 13/4.$$

   Hence $p(x)$ has no roots in $\mathbb{Q}$ and is therefore irreducible over $\mathbb{Q}$. □

2. A quadratic polynomial $ax^2 + bx + c$ with integer coefficients and $a \neq 0$ is irreducible over $\mathbb{Q}$ if and only if the discriminant $b^2 - 4ac$ is *not* a perfect square.

   Proof. Since $\deg p(x) = 2$, Theorem 3.6.6 says that $p(x)$ is irreducible over $\mathbb{Q}$ if and only if $p(x)$ has no roots in $\mathbb{Q}$. By the Quadratic Formula, the roots of $p(x)$ are in $\mathbb{Q}$ if and only if $\sqrt{b^2 - 4ac}$ is rational. By Theorem 2.5.10, $\sqrt{b^2 - 4ac}$ is rational if and only if it is an integer. Thus $p(x)$ has a root in $\mathbb{Q}$ if and only if $b^2 - 4ac$ is a perfect square, and so $p(x)$ has *no* roots in $\mathbb{Q}$ (hence is irreducible) if and only if $b^2 - 4ac$ is *not* a perfect square. □

Determining whether a polynomial of degree greater than 3 is irreducible is a more difficult problem. In particular, it is *not* sufficient to simply show that the polynomial has no rational roots. This is necessary, as it shows that there are no factors of degree 1, but it does not eliminate the possibility that there may be factors of higher degree.

Examples:

1. Let $p(x) = x^4 - 2x^2 + 3$. By the Rational Root Test, the only possible rational roots are ± 1 and ± 3. It is easily verified that none of these is a root, and so $p(x)$ has no rational roots. However, since $p(x) = (x^2 - 3)(x^2 + 1)$, it is clear that $p(x)$ is *not* irreducible over $\mathbb{Q}$.

2. The polynomial $p(x) = x^4 + 3x^3 + x^2 + 4x + 2$ is irreducible over $\mathbb{Q}$.

   Proof. By the Rational Root Test (Theorem 3.5.11), the only possible rational roots of $p(x)$ are integer divisors of 2, hence are ± 1 and ± 2. We have

   $$p(1) = 11,\ p(-1) = -3,\ p(2) = 54,\ \text{and } p(-2) = -10,$$

   hence $p(x)$ has no rational roots and therefore no linear factors.

   Since $\deg p(x) = 4$, if $p(x)$ is reducible over $\mathbb{Q}$, it must factor as a product of two quadratic polynomials. By Gauss's Lemma (Theorem 3.5.10), since the coefficients of $p(x)$ are integers and $p(x)$ is monic, if $p(x)$ is reducible then

   $$p(x) = (x^2 + ax + b)(x^2 + cx + d),$$

where a, b, c, and d are integers. Hence we have

$$p(x) = x^4 + 3x^3 + x^2 + 4x + 2 = x^4 + (a+c)x^3 + (ac+b+d)x^2 + (ad+bc)x + bd.$$

Equating corresponding coefficients yields the system of equations

$$\begin{aligned} a + c &= 3 && (3.4)\\ ac + b + d &= 1 && (3.5)\\ ad + bc &= 4 && (3.6)\\ bd &= 2. && (3.7) \end{aligned}$$

Since $bd = 2$ by Equation 3.7 and $b, d \in \mathbb{Z}$, we have either that one of b or d is 1 and the other is 2, or one is -1 and the other is -2. Hence, in any case, $b + d = \pm 3$. By Equation 3.4, $c = 3 - a$, and so

$$ac = a(3-a) = 3a - a^2.$$

If $b + d = 3$, then substituting $ac = 3a - a^2$ and $b + d = 3$ in Equation 3.5, we obtain

$$(3a - a^2) + 3 = 1,$$

or

$$a^2 - 3a - 2 = 0.$$

The discriminant of this quadratic equation is $(-3)^2 - 4(1)(-2) = 17$, which is not a perfect square. Hence the solutions are not rational.

If $b + d = -3$, then substituting $ac = 3a - a^2$ and $b + d = -3$ in Equation 3.5, we obtain

$$(3a - a^2) - 3 = 1,$$

or

$$a^2 - 3a + 4 = 0.$$

The discriminant of this quadratic equation is $(-3)^2 - 4(1)(4) = -7$, which is negative. Hence the solutions are not real numbers.

In either case, a is not an integer. Therefore, no integers a, b, c, d can satisfy all of Equations 3.4–3.7, and hence $p(x)$ is irreducible over $\mathbb{Q}$. □

The following result can be very helpful in proving that a polynomial with integer coefficients is irreducible over $\mathbb{Q}$.

Theorem 3.6.7 (Eisenstein's Criterion) *Let $a(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial with integer coefficients. If there is a prime number p such that*

$$p \mid a_0,\ p \mid a_1, \ldots,\ p \mid a_{n-1}$$

but

$$p \nmid a_n \text{ and } p^2 \nmid a_0,$$

then $a(x)$ is irreducible over $\mathbb{Q}$.

Proof. If $\deg a(x) = n = 1$, then $a(x)$ is irreducible over $\mathbb{Q}$, so we may assume $n \geqslant 2$. We suppose $a(x)$ is *not* irreducible and reach a contradiction.

By Gauss's Lemma (Theorem 3.5.10), if $a(x)$ is not irreducible over $\mathbb{Q}$, then $a(x)$ factors over $\mathbb{Z}$. Hence there is a factorization

$$a(x) = (b_r x^r + b_{r-1}x^{r-1} + \cdots + b_1 x + b_0)(c_s x^s + c_{s-1}x^{s-1} + \cdots + c_1 x + c_0) \quad (*)$$

where $1 \leqslant r \leqslant n-1$ and $1 \leqslant s \leqslant n-1$; $r + s = n$; $b_i, c_j \in \mathbb{Z}$ for all i, j; $b_r \neq 0$ and $c_s \neq 0$.

We have $p \mid a_0$ and $a_0 = b_0 c_0$, and so $p \mid b_0 c_0$. By Euclid's Lemma for Primes (Theorem 2.5.6), this implies $p \mid b_0$ or $p \mid c_0$. But, since $p^2 \nmid a_0$, p cannot divide both b_0 and c_0. Hence p divides exactly one of b_0, c_0. We may therefore assume without loss of generality that $p \mid b_0$ and $p \nmid c_0$.

Since $p \nmid a_n$ and $a_n = b_r c_s$, we have that $p \nmid b_r$. Thus $p \mid b_0$ and $p \nmid b_r$, and so there is a smallest integer k such that $p \nmid b_k$; that is $p \nmid b_k$ for some k with $1 \leqslant k \leqslant r \leqslant n-1$, but $p \mid b_i$ for $0 \leqslant i \leqslant k-1$.

By definition of polynomial multiplication (Definition 3.1.5), we have

$$a_k = b_k c_0 + b_{k-1}c_1 + \cdots + b_1 c_{k-1} + b_0 c_k.$$

Since $k \leqslant n-1$, we know by hypothesis that $p \mid a_k$. Also, $p \mid b_i$ for $i = 0, 1, \ldots, k-1$. Hence by the Combination Theorem (Theorem 2.2.3), we have

$$p \mid a_k - (b_{k-1}c_1 + \cdots + b_0 c_k);$$

that is, $p \mid b_k c_0$. By Euclid's Lemma for Primes, this implies $p \mid b_k$ or $p \mid c_0$, a contradiction. Hence there is no factorization of the form $(*)$, and so $a(x)$ is irreducible over $\mathbb{Q}$. □

We showed previously that the square root of a prime number must be irrational. The next result, which is an easy consequence of Eisenstein's Criterion, implies that every root of a prime number is irrational.

Corollary 3.6.8 *If p is a prime number, then $x^n - p$ is irreducible over $\mathbb{Q}$ for every $n \geqslant 1$.*

Proof. The prime p satisfies the hypotheses of Eisenstein's Criterion. □

Corollary 3.6.9 *If n is any positive integer, then there is an irreducible polynomial in $\mathbb{Q}[x]$ of degree n.*

Proof. By the previous corollary, $x^n - 2$ is irreducible over $\mathbb{Q}$ of degree n. □

The second corollary says that there are irreducible polynomials over $\mathbb{Q}$ of every possible degree. This differs considerably from the situation for polynomials over $\mathbb{R}$ or $\mathbb{C}$. The following very deep theorem, first proved by Gauss, says that the only irreducible polynomials with complex coefficients are those of degree 1.

Theorem 3.6.10 (Fundamental Theorem of Algebra) *If $p(x)$ is a non-constant polynomial in $\mathbb{C}[x]$, then $p(x)$ has at least one root in $\mathbb{C}$.*

Corollary 3.6.11 *No polynomial in $\mathbb{C}[x]$ of degree greater than 1 is irreducible over $\mathbb{C}$. Every non-constant polynomial in $\mathbb{C}[x]$ can be factored as a product of linear polynomials in $\mathbb{C}[x]$.*

The Fundamental Theorem of Algebra says that the field $\mathbb{C}$ of complex numbers is **algebraically closed**. This means roughly that no larger field can be obtained from $\mathbb{C}$ by algebraic operations such as taking roots. For example, we obtained $\mathbb{C}$ from $\mathbb{R}$ by "adjoining" $\sqrt{-1}$, but any such operation on complex numbers will just yield complex numbers. There are fields that properly contain $\mathbb{C}$, such as the field $\mathbb{C}(x)$ of rational functions in x with complex coefficients (see Theorem 3.1.12). Such fields cannot be obtained by adjoining roots of complex polynomials to $\mathbb{C}$, however.

The situation for polynomials in $\mathbb{R}[x]$ is only slightly less restrictive. It is easy to show, using the Intermediate Value Theorem (as studied in Calculus), that every polynomial in $\mathbb{R}[x]$ of odd degree has a root in $\mathbb{R}$. Thus, in particular, there are no irreducible polynomials of odd degree greater than 1 in $\mathbb{R}[x]$. It is possible, however, for a polynomial in $\mathbb{R}[x]$ of even degree to have no real roots. For example, the polynomial $f(x) = x^4 + 3x^2 + 2$ has no real roots. As noted above, this does *not* imply $f(x)$ is irreducible, and in fact $f(x) = (x^2 + 1)(x^2 + 2)$.

A polynomial of degree 2 in $\mathbb{R}[x]$ with a negative discriminant has no real roots and is therefore irreducible over $\mathbb{R}$. Hence there are irreducible polynomials in $\mathbb{R}[x]$ of degree 1 and of degree 2. By the Quadratic Formula, if a quadratic polynomial with real coefficients has complex roots, they are conjugates of each other. It is true more generally that the complex roots of a polynomial with real coefficients of any degree must occur in complex conjugate pairs. We can show, using this result and the Fundamental Theorem of Algebra, that no polynomial in $\mathbb{R}[x]$ of degree greater than 2 is irreducible over $\mathbb{R}$.

Theorem 3.6.12 *Let $p(x)$ be a polynomial with real coefficients. If z is a complex root of $p(x)$, then the complex conjugate $\overline{z}$ of z is also a root of $p(x)$.*

Proof. Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, with $a_i \in \mathbb{R}$ for all i, and let $z \in \mathbb{C}$ be a root. Thus we have

$$p(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 = 0. \quad (*)$$

By the definition of the complex conjugate (Definition 1.2.3), it is clear that the conjugate of any real number is the number itself. Hence $\overline{a_i} = a_i$ for each i. Also, by Theorem 1.2.4, the conjugate of a sum is the sum of the conjugates and the conjugate of a product is the product of the conjugates. From this, it also follows that $\overline{z^m} = (\overline{z})^m$ for any positive integer m. Therefore, we have

$$\begin{aligned}
p(\overline{z}) &= a_n(\overline{z})^n + a_{n-1}(\overline{z})^{n-1} + \cdots + a_1\overline{z} + a_0 \\
&= \overline{a_n}(\overline{z})^n + \overline{a_{n-1}}(\overline{z})^{n-1} + \cdots + \overline{a_1}\,\overline{z} + \overline{a_0} \text{ since } a_i \in \mathbb{R} \text{ for all } i, \\
&= \overline{a_n}\,\overline{z^n} + \overline{a_{n-1}}\,\overline{z^{n-1}} + \cdots + \overline{a_1}\,\overline{z} + \overline{a_0} \text{ by Theorem 1.2.4 (ii)}, \\
&= \overline{a_n z^n} + \overline{a_{n-1}z^{n-1}} + \cdots + \overline{a_1 z} + \overline{a_0} \text{ by Theorem 1.2.4 (ii)}, \\
&= \overline{a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0} \text{ by Theorem 1.2.4 (i)}, \\
&= \overline{0} \text{ by } (*), \\
&= 0.
\end{aligned}$$

Hence $p(\overline{z}) = 0$ and $\overline{z}$ is a root of $p(x)$. □

Theorem 3.6.13 (Fundamental Theorem of Algebra – Real Number Version) *Every non-constant polynomial in $\mathbb{R}[x]$ can be factored as a product of linear and irreducible quadratic polynomials in $\mathbb{R}[x]$.*

Proof. Let $p(x)$ be a polynomial with coefficients in $\mathbb{R}$. By the Fundamental Theorem of Algebra, $p(x)$ factors as a product of linear factors over the complex numbers $\mathbb{C}$. By Theorem 3.6.12, the non-real roots come in complex conjugate pairs. Therefore, if $a \in \mathbb{R}$ is the leading coefficient of $p(x)$, then over $\mathbb{C}$, $p(x)$ can be factored as

$$p(x) = a(x - r_1)(x - r_2)\cdots(x - r_m)(x - z_1)(x - \overline{z_1})(x - z_2)(x - \overline{z_2})\cdots(x - z_n)(x - \overline{z_n}),$$

where $r_1, r_2, \ldots, r_m$ are the real roots of $p(x)$ and $z_1, \overline{z_1}, z_2, \overline{z_2}, \ldots, z_n, \overline{z_n}$ are the complex, non-real, roots of $p(x)$.

For each $i = 1, \ldots, n$, we have

$$(x - z_i)(x - \overline{z_i}) = x^2 - (z_i + \overline{z_i})x + z_i\overline{z_i}.$$

By Proposition 1.2.5 (i), $z_i + \overline{z_i} = 2\mathrm{Re}(z_i) \in \mathbb{R}$ and by Proposition 1.2.6 and Definition 1.2.8, $z_i\overline{z_i} = |z_i|^2 \in \mathbb{R}$. Therefore, we have

$$p(x) = a(x - r_1)\cdots(x - r_m)(x^2 - 2\mathrm{Re}(z_1)x + |z_1|^2)\cdots(x^2 - 2\mathrm{Re}(z_n)x + |z_n|^2),$$

and the coefficients of each factor are real numbers. Hence this factorization into a product of linear and quadratic factors is over the real numbers $\mathbb{R}$. □

In particular, the theorem implies the following corollary.

Corollary 3.6.14 *No polynomial of degree greater than 2 with real coefficients is irreducible over $\mathbb{R}$.*

§3.6 Exercises

1. Find all irreducible polynomials of degree 2 or 3 in $\mathbb{Z}_2[x]$.
2. Find all monic irreducible polynomials of degree 2 or 3 in $\mathbb{Z}_3[x]$.
3. Show that the condition that p be a prime in Eisenstein's Criterion is necessary by finding a polynomial $a(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ and an integer m such that

$$m \mid a_0,\ m \mid a_1,\ \ldots,\ m \mid a_{n-1},\ m \nmid a_n,\ \text{and } m^2 \nmid a_0,$$

   but $a(x)$ is NOT irreducible.
4. Show that $f(x) = x^4 + x^3 - 2x - 4$ does not have a rational root and show that $f(x)$ is NOT irreducible over $\mathbb{Q}$.

Show that the following polynomials are irreducible over $\mathbb{Q}$.

5. $f(x) = 3x^2 + 5x - 4$
6. $f(x) = x^3 + 3x^2 + 2$
7. $f(x) = 2x^3 + 4x + 5$
8. $f(x) = x^{11} - 3x^4 + 12x^3 + 36x - 6$
9. $f(x) = 3x^7 - 10x^5 + 50x^4 - 40x^2 + 20$
10. $f(x) = 5x^4 - 7x + 7$
11. $f(x) = x^4 + 2x^3 + x + 1$
12. $f(x) = x^4 + 5x^2 + 3x + 2$
13. $f(x) = x^4 + 3x^3 + x^2 + 3x + 5$

3.7 Irreducible Polynomials as Primes

In this section, S will always denote one of $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$.

As noted previously, the irreducible polynomials are the analogues in $S[x]$ of the prime numbers in $\mathbb{Z}$. Many of the results we proved about prime numbers can be restated in terms of irreducible polynomials. Compare the results in this section to those on primes in §2.5 and §2.6. The first result is analogous to the Prime Divisor Principle (Theorem 2.5.2).

Theorem 3.7.1 (Irreducible Factor Principle) *If $a(x)$ is a non-constant polynomial in $S[x]$, then there is an irreducible polynomial in $S[x]$ that divides $a(x)$.*

Proof. Let $\mathcal{S}$ be the set of all degrees of non-constant divisors of $a(x)$ in $S[x]$. Since $a(x) \mid a(x)$ and $a(x)$ is non-constant, we have $\deg a(x) \in \mathcal{S}$ and so $\mathcal{S}$ is non-empty. By the Well-ordering Principle (Theorem 2.1.1), $\mathcal{S}$ has a smallest element, say m, and so there is a non-constant divisor $d(x)$ of $a(x)$ in $S[x]$ of degree $m \geqslant 1$. Thus $d(x)$ is a non-constant divisor of $a(x)$ of smallest degree. We will show that $d(x)$ is irreducible.

Suppose $d(x)$ is *not* irreducible. Then, since $d(x)$ is non-constant, we have $d(x) = f(x)g(x)$, where $f(x)$ and $g(x)$ are non-constant polynomials in $S[x]$ satisfying $\deg f(x) < \deg d(x)$ and $\deg g(x) < \deg d(x)$. Now $d(x) = f(x)g(x)$, so $f(x) \mid d(x)$, and we chose $d(x)$ to be a divisor of $a(x)$, so $d(x) \mid a(x)$. Thus by transitivity of divisibility (Theorem 3.3.2 (iv)), we have $f(x) \mid a(x)$. But $\deg f(x) < \deg d(x)$ and $d(x)$ is a non-constant divisor of $a(x)$ of smallest degree, a contradiction. Therefore, our assumption that $d(x)$ is *not* irreducible must be false, and so $d(x)$ is an irreducible divisor of $a(x)$. □

The next theorem is a special case of Euclid's Lemma for Polynomials where the divisor is irreducible. It is analogous to Euclid's Lemma for Primes (Theorem 2.5.6).

Theorem 3.7.2 (Euclid's Lemma for Irreducible Polynomials) *If $p(x)$ is an irreducible polynomial in $S[x]$ and $p(x) \mid a(x)b(x)$, where $a(x)$, $b(x)$ are in $S[x]$, then $p(x) \mid a(x)$ or $p(x) \mid b(x)$.*

Proof. Since $p(x)$ is irreducible and $(p(x), a(x))$ is a divisor of $p(x)$, we have either $(p(x), a(x)) = 1$ or $(p(x), a(x)) = k \cdot p(x)$ for some non-zero constant $k \in S$. If $(p(x), a(x)) = 1$, then $p(x) \mid b(x)$ by Euclid's Lemma for Polynomials (Theorem 3.3.20). If $(p(x), a(x)) = k \cdot p(x)$, then $k \cdot p(x) \mid a(x)$ by definition of GCD. Since $p(x) = \frac{1}{k}(k \cdot p(x))$, it follows that $p(x) \mid a(x)$ by Proposition 3.3.3. Hence either $p(x) \mid a(x)$ or $p(x) \mid b(x)$. □

Corollary 3.7.3 *If $p(x)$ is an irreducible polynomial in $S[x]$ and $p(x) \mid a_1(x)a_2(x)\cdots a_n(x)$, where the $a_i(x)$ are in $S[x]$, then $p(x) \mid a_i(x)$ for some i.*

Proof. We proceed by induction on the number n of factors. If $n = 1$, the hypothesis becomes $p(x) \mid a_1(x)$ and the conclusion is obvious. Hence the result is true if $n = 1$.

Assume now that the result holds for $n = k$; that is,

$$\text{if } p(x) \mid a_1(x)a_2(x)\cdots a_k(x) \text{ then } p(x) \mid a_i(x) \text{ for some } i = 1, 2, \ldots, k \ (*)$$

and show that the result holds for $n = k + 1$; that is,

$$\text{if } p(x) \mid a_1(x)a_2(x)\cdots a_k(x)a_{k+1}(x) \text{ then } p(x) \mid a_i(x) \text{ for some } i = 1, 2, \ldots, k + 1.$$

Let $a(x) = a_1(x)a_2(x)\cdots a_k(x)$ and $b(x) = a_{k+1}(x)$. If $p(x) \mid a_1(x)a_2(x)\cdots a_k(x)a_{k+1}(x)$, then $p(x) \mid a(x)b(x)$. By Theorem 3.7.2, we have that $p(x) \mid a(x)$ or $p(x) \mid b(x)$. If $p(x) \mid a(x)$, then $p(x) \mid a_1(x)a_2(x)\cdots a_k(x)$, and by the inductive hypothesis (*), $p(x) \mid a_i(x)$ for some $i = 1, 2, \ldots, k$. If $p(x) \mid b(x)$, then $p(x) \mid a_{k+1}(x)$. Hence, in any case, $p(x) \mid a_i(x)$ for some $i = 1, 2, \ldots, k+1$. Therefore, if the statement is true for $n = k$, then it is true for $n = k+1$, hence the result holds for all $n \geqslant 1$ by the Principle of Mathematical Induction. □

Corollary 3.7.4 *If $p(x)$ is an irreducible polynomial in $S[x]$ and $p(x) \mid q_1(x)q_2(x)\cdots q_n(x)$, where $q_1(x), q_2(x), \ldots, q_n(x)$ are irreducible polynomials in $S[x]$, then $p(x) = k \cdot q_i(x)$ for some i and some number k in S.*

Proof. By Corollary 3.7.3, if $p(x) \mid q_1(x)q_2(x)\cdots q_n(x)$, then $p(x) \mid q_i(x)$ for some $i = 1, 2, \ldots, n$. Since $q_i(x)$ is irreducible, its only non-constant divisors are multiples of $q_i(x)$ itself, and since $p(x)$ is irreducible, hence non-constant, this implies $p(x) = k \cdot q_i(x)$ for some $k \in S$. □

We also get a polynomial version of the Fundamental Theorem of Arithmetic (Theorem 2.5.11).

Theorem 3.7.5 (Unique Factorization) *Let $a(x)$ be a non-constant polynomial in $S[x]$. Then either $a(x)$ is irreducible or $a(x)$ is a product of irreducible polynomials in $S[x]$. In case $a(x)$ is not irreducible,*

$$a(x) = \alpha \cdot p_1(x)^{a_1} p_2(x)^{a_2} \cdots p_r(x)^{a_r},$$

where α is the leading coefficient of $a(x)$, the $p_i(x)$ are distinct, irreducible, monic polynomials in $S[x]$, and $a_i \geqslant 1$ for all i. This factorization is unique except for the order in which the factors are written.

Proof. We first show such a factorization exists. By the Well-ordering Principle (Theorem 2.1.1), if there is a non-constant polynomial that is *not* irreducible or a product of irreducible polynomials, then there is one of smallest degree, say $m(x)$. Since $m(x)$ is non-constant and is not irreducible, $m(x) = f(x)g(x)$ for some non-constant polynomials $f(x)$ and $g(x)$ with $\deg f(x) < \deg m(x)$ and $\deg g(x) < \deg m(x)$. By the minimality of $\deg m(x)$, each of $f(x)$, $g(x)$ is either irreducible or a product of irreducible polynomials. Hence $f(x)g(x) = m(x)$ is a product of irreducible polynomials, contradicting the choice of $m(x)$. Therefore, our assumption that there is a non-constant polynomial that is neither irreducible nor a product of irreducible polynomials must be false.

To prove uniqueness, suppose $a(x)$ can be written as a product of irreducible polynomials in two ways. Let $p_1(x), p_2(x), \ldots, p_r(x)$ be all of the distinct monic irreducible polynomials that appear in at least one of the two factorizations. Writing products of repeated irreducible factors as powers (recalling that $p_i(x)^0 = 1$) and denoting by α the leading coefficient of $a(x)$, we can write the two factorizations in the form

$$a(x) = \alpha \cdot p_1(x)^{a_1} p_2(x)^{a_2} \cdots p_r(x)^{a_r} = \alpha \cdot p_1(x)^{b_1} p_2(x)^{b_2} \cdots p_r(x)^{b_r},$$

where $a_i, b_i \in \mathbb{Z}$ and $a_i \geqslant 0$, $b_i \geqslant 0$ for all i (and, for each i, at least one of a_i, b_i is non-zero). We must show that $a_i = b_i$ for all i.

Suppose $a_i \neq b_i$ for some i. We may assume, without loss of generality, that $i = 1$ and $b_1 < a_1$. Dividing both factorizations by $p_1(x)^{b_1}$ yields

$$p_1(x)^{a_1 - b_1} p_2(x)^{a_2} \cdots p_r(x)^{a_r} = p_2(x)^{b_2} \cdots p_r(x)^{b_r}.$$

Now $b_1 < a_1$, hence $a_1 - b_1 > 0$ and so $p_1(x)$ divides the factorization on the left. By equality, we must also have

$$p_1(x) \mid p_2(x)^{b_2} \cdots p_r(x)^{b_r}.$$

By Corollary 3.7.4, this implies $p_1(x) = k \cdot p_j(x)$ for some $2 \leqslant j \leqslant r$ and some constant $k \in S$. Since both $p_1(x)$ and $p_j(x)$ are monic, $k = 1$ and $p_1(x) = p_j(x)$, contradicting the fact that the polynomials $p_1(x), p_2(x), \ldots, p_r(x)$ are distinct. Hence $a_i = b_i$ for all i and the expression of $a(x)$ as a product of irreducible polynomials is unique, except for the order of the factors. □

This theorem states that every non-constant polynomial in $S[x]$ has a "canonical" factorization as a product of irreducible polynomials. As we did for integers (Proposition 2.6.1), we can compare two polynomials by allowing the exponents a_i of the irreducible factors to be 0, and then writing both polynomials as products of powers of the same irreducible polynomials. We can use this to characterize divisibility in terms of factorizations into products of irreducible polynomials (compare to Theorem 2.6.2).

Theorem 3.7.6 *Let $a(x) = \alpha p_1(x)^{a_1} p_2(x)^{a_2} \cdots p_r(x)^{a_r}$ and $b(x) = \beta p_1(x)^{b_1} p_2(x)^{b_2} \cdots p_r(x)^{b_r}$, with the $p_i(x)$ distinct, monic, irreducible polynomials in $S[x]$, α and β non-zero numbers in S, and $a_i \geqslant 0$, $b_i \geqslant 0$ for all i. Then $a(x) \mid b(x)$ if and only if $a_i \leqslant b_i$ for all i.*

Proof. Suppose first that $a(x) \mid b(x)$, so that $b(x) = f(x)a(x)$ for some polynomial $f(x) \in S[x]$. If $p(x)$ is any irreducible factor of $f(x)$, then $p(x) \mid f(x)$ and $f(x) \mid b(x)$, hence $p(x) \mid b(x)$ by Theorem 3.3.2 (iv). Therefore, every monic irreducible factor of $f(x)$ is one of the $p_i(x)$, and we can write

$$f(x) = \gamma p_1(x)^{n_1} p_2(x)^{n_2} \cdots p_r(x)^{n_r},$$

where $n_i \geqslant 0$ for each i and $\gamma \in S$ is a non-zero constant.

Writing $b(x) = f(x)a(x)$ in terms of the irreducible factorizations, we have

$$\begin{aligned} \beta p_1(x)^{b_1} p_2(x)^{b_2} \cdots p_r(x)^{b_r} &= \gamma(p_1(x)^{n_1} p_2(x)^{n_2} \cdots p_r(x)^{n_r}) \cdot \alpha(p_1(x)^{a_1} p_2(x)^{a_2} \cdots p_r(x)^{a_r}) \\ &= \gamma\alpha p_1(x)^{n_1+a_1} p_2(x)^{n_2+a_2} \cdots p_r(x)^{n_r+a_r}, \end{aligned}$$

and so $b_i = n_i + a_i$ for each i, by uniqueness of factorization (see Theorem 3.7.5). Since $n_i \geqslant 0$ for each i, we have $a_i \leqslant n_i + a_i = b_i$ for each i.

Conversely, suppose $a_i \leqslant b_i$, so that $b_i - a_i \geqslant 0$, for each i. Also, since S is a field and $\alpha \neq 0$, we have $\beta/\alpha \in S$. Thus

$$g(x) = (\beta/\alpha) p_1(x)^{b_1-a_1} p_2(x)^{b_2-a_2} \cdots p_r(x)^{b_r-a_r}$$

is a polynomial in $S[x]$. We then have

$$\begin{aligned} b(x) &= \beta p_1(x)^{b_1} p_2(x)^{b_2} \cdots p_r(x)^{b_r} \\ &= (\beta/\alpha)(p_1(x)^{b_1-a_1} p_2(x)^{b_2-a_2} \cdots p_r(x)^{b_r-a_r}) \cdot \alpha(p_1(x)^{a_1} p_2(x)^{a_2} \cdots p_r(x)^{a_r}) \\ &= g(x)a(x). \end{aligned}$$

Hence $b(x) = g(x)a(x)$ and $g(x) \in S[x]$, and so $a(x) \mid b(x)$. □

As we did for integers, we can also define the least common multiple of two polynomials and compute the LCM in terms of the factorizations into products of irreducible polynomials.

Definition 3.7.7 *Let $a(x)$ and $b(x)$ be non-zero polynomials in $S[x]$. The* **least common multiple** *of $a(x)$ and $b(x)$ is the monic polynomial $m(x)$ in $S[x]$, denoted $m(x) = [a(x), b(x)]$, satisfying*

i. *$a(x) \mid m(x)$ and $b(x) \mid m(x)$, and*

ii. *if $a(x) \mid c(x)$ and $b(x) \mid c(x)$ for a non-zero polynomial $c(x)$ in $S[x]$, then* $\deg\, m(x) \leqslant \deg\, c(x)$.

Note that any constant multiple of $m(x)$ (including 0) will satisfy conditions (i) and (ii) of the definition. The LCM is defined to be the *monic* polynomial satisfying these conditions so that the LCM will be unique and non-zero.

The next result, analogous to Theorem 2.6.5 for integers, allows us to compute the GCD and LCM of two polynomials using the factorizations into products of irreducible polynomials.

Theorem 3.7.8 *Let $a(x) = \alpha p_1(x)^{a_1} p_2(x)^{a_2} \cdots p_r(x)^{a_r}$ and $b(x) = \beta p_1(x)^{b_1} p_2(x)^{b_2} \cdots p_r(x)^{b_r}$, with the $p_i(x)$ distinct, monic, irreducible polynomials in $S[x]$, α and β non-zero numbers in S, and $a_i \geqslant 0$, $b_i \geqslant 0$ for all i. Then*

a. *$(a(x), b(x)) = p_1(x)^{d_1} p_2(x)^{d_2} \cdots p_r(x)^{d_r}$, where $d_i = \min\{a_i, b_i\}$ for all i, and*

b. *$[a(x), b(x)] = p_1(x)^{m_1} p_2(x)^{m_2} \cdots p_r(x)^{m_r}$, where $m_i = \max\{a_i, b_i\}$ for all i.*

Proof. GCD: Let

$$d(x) = p_1(x)^{d_1} p_2(x)^{d_2} \cdots p_r(x)^{d_r},$$

where $d_i = \min\{a_i, b_i\}$ for each i, so that $d(x) \in S[x]$ is monic. We show that $d(x) = (a(x), b(x))$ by showing that $d(x)$ satisfies parts (i) and (ii) of Theorem 3.3.16 (the "alternate" definition of the GCD).

(i) Since $d_i = \min\{a_i, b_i\}$, we have $d_i \leqslant a_i$ and $d_i \leqslant b_i$ for all i. Therefore, it follows from Theorem 3.7.6 that $d(x) \mid a(x)$ and $d(x) \mid b(x)$, and so (i) holds.

(ii) Let $c(x)$ be a polynomial in $S[x]$ satisfying $c(x) \mid a(x)$ and $c(x) \mid b(x)$. If $p(x)$ is any monic irreducible factor of $c(x)$, then by Theorem 3.3.2 (iv), $p(x)$ must be one of the $p_i(x)$. Thus we can write

$$c(x) = \gamma p_1(x)^{c_1} p_2(x)^{c_2} \cdots p_r(x)^{c_r},$$

where $\gamma \in S$ and $c_i \geqslant 0$ for each i. Since $c(x) \mid a(x)$ and $c(x) \mid b(x)$, Theorem 3.7.6 implies that $c_i \leqslant a_i$ and $c_i \leqslant b_i$ for all i. Hence

$$c_i \leqslant \min\{a_i, b_i\} = d_i$$

for each i, and so $c(x) \mid d(x)$, again by Theorem 3.7.6, and (ii) holds.

LCM: Let

$$m(x) = p_1(x)^{m_1} p_2(x)^{m_2} \cdots p_r(x)^{m_r},$$

where $m_i = \max\{a_i, b_i\}$ for each i, so that $m(x) \in S[x]$ is monic. We show that $m(x) = [a(x), b(x)]$ by showing that m satisfies parts (i) and (ii) of the definition of LCM (Definition 3.7.7).

(i) Since $m_i = \max\{a_i, b_i\}$, we have $m_i \geqslant a_i$ and $m_i \geqslant b_i$ for all i. Therefore, it follows from Theorem 3.7.6 that $a(x) \mid m(x)$ and $b(x) \mid m(x)$, and so (i) holds.

(ii) Let $c(x)$ be a non-zero polynomial in $S[x]$ satisfying $a(x) \mid c(x)$ and $b(x) \mid c(x)$. Each irreducible factor of $a(x)$ or $b(x)$ will also divide $c(x)$, but $c(x)$ may have other irreducible factors as well. Thus we can write

$$c(x) = p_1(x)^{c_1} p_2(x)^{c_2} \cdots p_r(x)^{c_r} \cdot k(x),$$

where each $c_i \geqslant 0$ and $k(x) \in S[x]$ is a polynomial such that $p_i(x) \nmid k(x)$ for each i.

Since $a(x) \mid c(x)$ and $b(x) \mid c(x)$, Theorem 3.7.6 implies that $a_i \leqslant c_i$ and $b_i \leqslant c_i$ for all i. Hence

$$m_i = \max\{a_i, b_i\} \leqslant c_i$$

for each i. It follows that

$$m(x) \mid p_1(x)^{c_1} p_2(x)^{c_2} \cdots p_r(x)^{c_r},$$

again by Theorem 3.7.6, and therefore that $m(x) \mid c(x)$, since $p_1(x)^{c_1} p_2(x)^{c_2} \cdots p_r(x)^{c_r} \mid c(x)$. Thus $\deg m(x) \leqslant \deg c(x)$ by Theorem 3.3.5, as $c(x) \neq 0$, and so (ii) holds. □

Example: Let $S = \mathbb{Q}$ and let $a(x) = (x-2)^2(x^5+3)^3(x-7)^5$ and $b(x) = (x-2)^4(x^5+3)^2(x^{17}-5)^2$. (Note that the non-linear factors are irreducible over $\mathbb{Q}$ by Eisenstein's Criterion, Theorem 3.6.7.) Writing $a(x)$ and $b(x)$ in terms of the same irreducible factors, we have

$$a(x) = (x-2)^2(x^5+3)^3(x^{17}-5)^0(x-7)^5$$

$$b(x) = (x-2)^4(x^5+3)^2(x^{17}-5)^2(x-7)^0$$

Hence, by the theorem,

$$\begin{aligned} (a(x), b(x)) &= (x-2)^2(x^5+3)^2(x^{17}-5)^0(x-7)^0 = (x-2)^2(x^5+3)^2 \text{ and} \\ [a(x), b(x)] &= (x-2)^4(x^5+3)^3(x^{17}-5)^2(x-7)^5. \end{aligned}$$

Notice also that

$$\begin{aligned} (a(x), b(x))[a(x), b(x)] &= [(x-2)^2(x^5+3)^2] \cdot [(x-2)^4(x^5+3)^3(x^{17}-5)^2(x-7)^5] \\ &= (x-2)^{2+4}(x^5+3)^{2+3}(x^{17}-5)^{0+2}(x-7)^{0+5} \\ &= (x-2)^{2+4}(x^5+3)^{3+2}(x^{17}-5)^{0+2}(x-7)^{5+0} \\ &= [(x-2)^2(x^5+3)^3(x-7)^5] \cdot [(x-2)^4(x^5+3)^2(x^{17}-5)^2] \\ &= a(x)b(x), \end{aligned}$$

and so $(a(x), b(x))[a(x), b(x)] = a(x)b(x)$. □

We showed that the GCD of two polynomials is divisible by any common divisor. Similarly, any common multiple of two polynomials is a multiple of the LCM.

Corollary 3.7.9 *Let $a(x)$ and $b(x)$ be non-zero polynomials in $S[x]$ and let $m(x) = [a(x), b(x)]$. If $c(x)$ is a polynomial in $S[x]$ satisfying $a(x) \mid c(x)$ and $b(x) \mid c(x)$, then $m(x) \mid c(x)$.*

Proof. This is shown in part (ii) for the LCM in the proof of Theorem 3.7.8. □

The characterization of the GCD and LCM in Theorem 3.7.8 implies a very important relationship between the two polynomials, as suggested in the example above.

Corollary 3.7.10 *If $a(x)$ and $b(x)$ are monic polynomials in $S[x]$, then*

$$(a(x), b(x)) \cdot [a(x), b(x)] = a(x) \cdot b(x).$$

Proof. We use the notation of Theorem 3.7.8. For each i, $d_i = \min\{a_i, b_i\}$ and $m_i = \max\{a_i, b_i\}$. Hence d_i is one of a_i or b_i and m_i is the other. In any case, we have $a_i + b_i = d_i + m_i$. Therefore,

$$\begin{aligned}
(a(x), b(x)) \cdot [a(x), b(x)] &= (p_1(x)^{d_1} p_2(x)^{d_2} \cdots p_r(x)^{d_r})(p_1(x)^{m_1} p_2(x)^{m_2} \cdots p_r(x)^{m_r}) \\
&= p_1(x)^{d_1+m_1} p_2(x)^{d_2+m_2} \cdots p_r(x)^{d_r+m_r} \\
&= p_1(x)^{a_1+b_1} p_2(x)^{a_2+b_2} \cdots p_r(x)^{a_r+b_r} \\
&= (p_1(x)^{a_1} p_2(x)^{a_2} \cdots p_r(x)^{a_r})(p_1(x)^{b_1} p_2(x)^{b_2} \cdots p_r(x)^{b_r}) \\
&= a(x) \cdot b(x),
\end{aligned}$$

and so $(a(x), b(x)) \cdot [a(x), b(x)] = a(x) \cdot b(x)$. □

The only reason $a(x)$ and $b(x)$ must be monic in the corollary is that by definition the left side of the equation is monic. More generally, if $a(x)$ and $b(x)$ are any non-zero polynomials, with leading coefficients α and β, respectively, then we have

$$(a(x), b(x)) \cdot [a(x), b(x)] = \frac{1}{\alpha\beta} \cdot a(x) \cdot b(x).$$

§3.7 Exercises

1. Let $f(x) = 5(x-8)^7(x^2+6)^9(x^3-7)^4 \in \mathbb{Q}[x]$. Determine whether each of the following polynomials divides $f(x)$. **Explain your answers.**

   (a) $a(x) = (x-8)^6(x^2+6)^5(x^3-7)^4$

   (b) $b(x) = 5(x-8)(x^2+6)^2(x^3-7)^5$

   (c) $c(x) = 3(x-8)^2(x^2+6)^2$

   (d) $d(x) = (x-6)^3(x^2+8)^4(x^3-7)$

2. Find a formula for the number of monic divisors of a polynomial $a(x)$ in $S[x]$ whose canonical factorization is $a(x) = \alpha \cdot p_1(x)^{a_1}p_2(x)^{a_2}\cdots p_r(x)^{a_r}$. **Justify your answer.**

3. Find the greatest common divisor and least common multiple of each of the following pairs of polynomials $a(x)$, $b(x)$ in $\mathbb{Q}[x]$.

   (a) $a(x) = 6(x-2)^5(x-4)^7(x^2+5)^2$
   $b(x) = 9(x-2)^3(x^2+5)^8(x^3+5)^4$

   (b) $a(x) = 2(x-1)^2(x-2)^3(x-3)^4(x-4)^5$
   $b(x) = 4(x-2)^6(x-3)^5(x-4)^4(x-5)^3$

For Exercises 4–6, use the Euclidean Algorithm to find the greatest common divisor and the least common multiple of the polynomials $a(x)$ and $b(x)$.

4. $a(x) = x^2 - 1$
   $b(x) = 2x^7 - 4x^5 + 2$

5. $a(x) = x + 3$
   $b(x) = x^3 - 2x + 4$

6. $a(x) = x^4 - 4x^2 - 3x + 6$
   $b(x) = x^3 + x^2 - x - 10$

Appendix A

Trigonometry Review

The following information is needed for working with the polar form of complex numbers in §1.2.

Definitions:

- π radians $= 180^\circ$
- Sine, cosine, and tangent of an acute angle θ in terms of ratios of sides of a right triangle:
$$\sin\theta = \frac{\text{opp}}{\text{hyp}}, \qquad \cos\theta = \frac{\text{adj}}{\text{hyp}}, \qquad \tan\theta = \frac{\text{opp}}{\text{adj}}.$$
- Sine, cosine, and tangent of an angle θ with initial side on the positive x-axis and the point (x, y) on its terminal side, with $r = \sqrt{x^2 + y^2}$:
$$\sin\theta = \frac{y}{r}, \qquad \cos\theta = \frac{x}{r}, \qquad \tan\theta = \frac{y}{x}.$$

Basic Identities:

- $\tan\theta = \dfrac{\sin\theta}{\cos\theta}$
- $\sin^2\theta + \cos^2\theta = 1$
- $\sin(-\theta) = -\sin\theta$
  $\cos(-\theta) = \cos\theta$
- $\sin(\theta + 2\pi) = \sin\theta$
  $\cos(\theta + 2\pi) = \cos\theta$

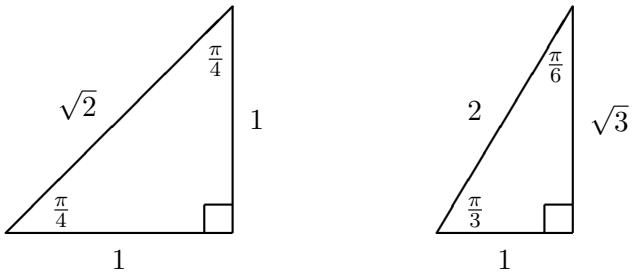
Angle Sum Formulas:

- $\sin(x + y) = \sin x \cos y + \cos x \sin y$
- $\cos(x + y) = \cos x \cos y - \sin x \sin y$

Values: Sine, cosine, and tangent at all special angles; i.e., at all multiples of:

$$\frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{\pi}{2}, \pi.$$

These can be memorized or calculated as needed using the right triangles below:



Any answers involving these values should be given as **exact** values. For example,

$$\sin\left(\frac{2\pi}{3}\right) = \frac{\sqrt{3}}{2},$$

and *not* 0.8660254038.

Appendix B

Answers to Selected Problems

§1.1

1. (a) $0.\overline{0495}$
   (b) $0.4\overline{27}$
   (c) $0.1\overline{259}$
   (d) $0.3\overline{571428}$
2. (a) 13/33
   (b) 4298/999
   (c) 1141553/19980
   (d) 102092327/999900

§1.2

1.

	Re(z)	Im(z)	$\overline{z}$	$\lvert z\rvert$
(a)	3	5	$3-5i$	$\sqrt{34}$
(b)	7	-2	$7+2i$	$\sqrt{53}$
(c)	-4	1	$-4-i$	$\sqrt{17}$
(d)	5	0	5	5

2. (a) $-1+7i$
   (b) $2-9i$
   (c) $11-2i$
   (d) 4
3. (a) $-18-i$
   (b) $25-8i$
   (c) 13
   (d) $7+5i$

§1.2, continued.

4. (a) $\frac{3}{25}-\frac{4}{25}\,i$
   (b) $\frac{7}{53}+\frac{2}{53}\,i$
   (c) $\frac{2}{13}-\frac{3}{13}\,i$
   (d) $-\frac{1}{7}\,i$
5. (a) $\frac{23}{13}+\frac{2}{13}\,i$
   (b) $\frac{23}{41}-\frac{2}{41}\,i$
   (c) $-2+i$
6.

	$\lvert z\rvert$	arg z	Polar Form
(a)	$\sqrt{2}$	$\frac{7\pi}{4}$	$\sqrt{2}\,(\cos\frac{7\pi}{4}+i\sin\frac{7\pi}{4})$
(b)	2	$\frac{7\pi}{6}$	$2\,(\cos\frac{7\pi}{6}+i\sin\frac{7\pi}{6})$
(c)	6	$\frac{5\pi}{3}$	$6\,(\cos\frac{5\pi}{3}+i\sin\frac{5\pi}{3})$
(d)	2	$\frac{3\pi}{4}$	$2\,(\cos\frac{3\pi}{4}+i\sin\frac{3\pi}{4})$
(e)	5	$\frac{\pi}{2}$	$5\,(\cos\frac{\pi}{2}+i\sin\frac{\pi}{2})$
(f)	7	π	$7\,(\cos\pi+i\sin\pi)$

7. $\frac{z_1}{z_2}=\frac{r_1}{r_2}(\cos(\theta_1-\theta_2)+i\sin(\theta_1-\theta_2))$
8. $\begin{aligned}\frac{1}{z} &= \frac{1}{r}(\cos(-\theta)+i\sin(-\theta))\\ &= \frac{1}{r}(\cos(\theta)-i\sin(\theta))\end{aligned}$

§1.2, continued.

9.

	$\lvert c\rvert$	Arg	Polar
(a)	8	$\frac{13\pi}{12}$	$8\left(\cos\frac{13\pi}{12}+i\sin\frac{13\pi}{12}\right)$
(b)	16	$\frac{3\pi}{2}$	$16\left(\cos\frac{3\pi}{2}+i\sin\frac{3\pi}{2}\right)$
(c)	16	$\frac{17\pi}{12}$	$16\left(\cos\frac{17\pi}{12}+i\sin\frac{17\pi}{12}\right)$
(d)	32	$\frac{5\pi}{3}$	$32\left(\cos\frac{5\pi}{3}+i\sin\frac{5\pi}{3}\right)$
(e)	$\frac{1}{2}$	$-\frac{5\pi}{12}$	$\frac{1}{2}(\cos(-\frac{5\pi}{12})+i\sin(-\frac{5\pi}{12}))$
(f)	$\frac{1}{2}$	$-\frac{\pi}{3}$	$\frac{1}{2}(\cos(-\frac{\pi}{3})+i\sin(-\frac{\pi}{3}))$

10. (a) $32i$
    (b) $\frac{1}{2}+\frac{\sqrt{3}}{2}\,i$
    (c) $-\frac{1}{2}-\frac{\sqrt{3}}{2}\,i$
    (d) 1

§1.3

1. (a) Commutative Law of Addition
   (b) Distributive Law
   (c) Associative Law of Multiplication
   (d) Associative Law of Addition & Distributive Law
   (e) Commutative Law of Multiplication & Associative Law of Multiplication

8. $2\mathbb{Z}$ satisfies all except (ix) and (x).
   $2\mathbb{Z}$ is a ring but not a ring with 1.

9. $\mathbb{Z}[i]$ is not a field.

10. Both F and C satisfy all except (x).

§2.2

4. If $a=12$, $b=4$, $c=6$, then $a=12$ divides $bc=24$, but $12\nmid 4$ and $12\nmid 6$.

5. Both $a=6$ and $b=10$ divide $c=30$, but $ab=6\cdot 10=60$ and $60\nmid 30$.

8. If $a=2$, $b=5$, $c=7$, then $a=2$ divides $b+c=12$, but $2\nmid 5$ and $2\nmid 7$.

§2.3

1. (a) $q=135$, $r=27$, $4752=135(35)+27$.
   (b) $q=232$, $r=0$, $9976=232(43)+0$ or $9976=232(43)$.

2. (a) $q=410$, $r=3827$,
   $2351487=410(5726)+3827$.
   (b) $q=11288$, $r=427$,
   $84637851=11288(7498)+427$.

4. 39

5. 6

6. 1

7. $(272,119)=17$
   $17=272(-3)+119(7)$

8. $(495,210)=15$
   $15=495(3)+210(-7)$

9. $(264,189)=3$
   $3=264(-5)+189(7)$

10. $(510,414)=6$
    $6=510(13)+414(-16)$

§2.4

1. (a) $56(-8)+72(6)=-16$
   (b) -28 cannot be expressed in this form.
   (c) 42 cannot be expressed in this form.
   (d) $56(32)+72(-24)=64$
   (e) 70 cannot be expressed in this form.
   (f) $56(-44)+72(33)=-88$

5. (a) could
   (b) could not
   (c) could not
   (d) could

§2.5

1. (a) prime
   (b) prime
   (c) not prime
   (d) not prime
   (e) is prime

4. The primes less than 100 are:

2	3	5	7	11
13	17	19	23	29
31	37	41	43	47
53	59	61	67	71
73	79	83	89	97

5. (a) $2 \cdot 13^2$
   (b) $7 \cdot 13 \cdot 17$
   (c) $2^2 \cdot 3^3 \cdot 5^2$

§2.6

1. (a) a does not divide n.
   (b) b divides n.
   (c) c does not divide n.

2. The positive divisors of $600 = 2^3 \cdot 3 \cdot 5^2$ are:

$$1 = 2^0 \cdot 3^0 \cdot 5^0 \qquad 2 = 2^1 \cdot 3^0 \cdot 5^0$$
$$5 = 2^0 \cdot 3^0 \cdot 5^1 \qquad 10 = 2^1 \cdot 3^0 \cdot 5^1$$
$$25 = 2^0 \cdot 3^0 \cdot 5^2 \qquad 50 = 2^1 \cdot 3^0 \cdot 5^2$$

$$4 = 2^2 \cdot 3^0 \cdot 5^0 \qquad 8 = 2^3 \cdot 3^0 \cdot 5^0$$
$$20 = 2^2 \cdot 3^0 \cdot 5^1 \qquad 40 = 2^3 \cdot 3^0 \cdot 5^1$$
$$100 = 2^2 \cdot 3^0 \cdot 5^2 \qquad 200 = 2^3 \cdot 3^0 \cdot 5^2$$

$$3 = 2^0 \cdot 3^1 \cdot 5^0 \qquad 6 = 2^1 \cdot 3^1 \cdot 5^0$$
$$15 = 2^0 \cdot 3^1 \cdot 5^1 \qquad 30 = 2^1 \cdot 3^1 \cdot 5^1$$
$$75 = 2^0 \cdot 3^1 \cdot 5^2 \qquad 150 = 2^1 \cdot 3^1 \cdot 5^2$$

$$12 = 2^2 \cdot 3^1 \cdot 5^0 \qquad 24 = 2^3 \cdot 3^1 \cdot 5^0$$
$$60 = 2^2 \cdot 3^1 \cdot 5^1 \qquad 120 = 2^3 \cdot 3^1 \cdot 5^1$$
$$300 = 2^2 \cdot 3^1 \cdot 5^2 \qquad 600 = 2^3 \cdot 3^1 \cdot 5^2$$

§2.6, continued.

3. (a) 72
   (b) 30

4. (a) 12
   (b) 40

5. (a) $(a, b) = 2^5 \cdot 7^3$
   $[a, b] = 2^7 \cdot 3^4 \cdot 5^3 \cdot 7^4 \cdot 11^{17} \cdot 13^8$
   (b) $(a, b) = 2^2 \cdot 3^2 \cdot 5^5 \cdot 7^4$
   $[a, b] = 2^3 \cdot 3^3 \cdot 5^6 \cdot 7^7 \cdot 11 \cdot 17$

6. 178101

7. 327509

8. $(963, 657) = 9$
   $[963, 657] = 70299$

9. $(510, 414) = 6$
   $[510, 414] = 35190$

§2.7

1. (a) 3
   (b) 14
   (c) 1
   (d) 0
   (e) 7

6. If $a = 1$, $b = -1$, $n = 3$, then $(-1)^2 = 1$; $1^2 \equiv (-1)^2 (\text{mod } 3)$ but $1 \not\equiv -1 (\text{mod } 3)$.

11. $(\overline{1})^{-1} = \overline{1}$
    $(\overline{3})^{-1} = \overline{7}$
    $(\overline{7})^{-1} = \overline{3}$
    $(\overline{9})^{-1} = \overline{9}$

§2.8

1. (a) 2
   (b) 3
   (c) 2
2. (a) 1
   (b) 4
   (c) 0
   (d) 6
3. (a) 6
   (b) 7
   (c) 2
4. (a) 1
   (b) 4
   (c) 6
   (d) 6
9. (a) $4 \nmid 478563289358$
   (b) $4 \mid 12354456724$
   (c) $4 \mid 352148763376$
10. (a) $3 \mid 21437856252$, $9 \mid 21437856252$
    (b) $3 \nmid 54637281274$, $9 \nmid 54637281274$
    (c) $3 \mid 42315768543$, $9 \nmid 42315768543$
11. (a) $6 \mid 47835624312$
    (b) $6 \nmid 65348127214$
    (c) $6 \nmid 27135248145$
12. (a) $11 \nmid 41783526413$
    (b) $11 \mid 615837429152$
    (c) $11 \nmid 724356712859$
13. (a) $7 \mid 98239072918$, $13 \nmid 98239072918$
    (b) $7 \mid 199885455861$, $13 \mid 199885455861$
    (c) $7 \nmid 182443992562$, $13 \mid 182443992562$

§3.1

1.

	Deg.	Coeff.
(a)	3	5
(b)	5	-9
(c)	0	6
(d)	$-\infty$	none
(e)	473	1

2. Coefficients of $a(x)$, $b(x)$:
   $a_0 = 7, \quad b_0 = 8$
   $a_1 = 4, \quad b_1 = 6$
   $a_2 = 5, \quad b_2 = 3$
   $a_3 = 2, \quad b_3 = 0$
   $a_4 = 0, \quad b_4 = 0$
   $a_5 = 0, \quad b_5 = 0$

   Coefficients of $a(x)b(x)$:
   $c_0 = 56$
   $c_1 = 74$
   $c_2 = 85$
   $c_3 = 58$
   $c_4 = 27$
   $c_5 = 6$

   Product $a(x)b(x)$:
   $6x^5 + 27x^4 + 58x^3 + 85x^2 + 74x + 56$
3. (a) $p(x)+q(x) = x^5+2x^4-3x^3+6x^2-x-4$
   $p(x) - q(x) = -x^5 + 8x^4 - 3x^3 - 2x^2 + 15x - 4$
   (b) $2x^5 - 7x^3 - 15x$
   (c) $4x^5 + 5x^4 + 14x^3 - 14x^2 + 2x + 5$
   (d) $x^4 + 6x^2 + 9$

§3.2

1. (a) 35
   (b) 15
   (c) 28
   (d) 28
   (e) 126

2. (a) $\binom{20}{16} = 4845$
   (b) $\binom{15}{9} = 5005$
   (c) $\binom{18}{7} = 31824$

3. (a) $16x^4 + 96x^3 + 216x^2 + 216x + 81$
   (b) $32x^5 + 240x^4 + 720x^3 + 1080x^2 + 810x + 243$
   (c) $64x^6 + 576x^5 + 2160x^4 + 4320x^3 + 4860x^2 + 2916x + 729$

§3.3

5. (a) $q(x) = 4x^2 + 2x$
   $r(x) = -7x + 1$
   $b(x) = (4x^2 + 2x) \cdot a(x) + (-7x + 1)$
   (b) $q(x) = \frac{1}{3}x^2 + \frac{1}{3}x + \frac{2}{3}$
   $r(x) = -1$
   $b(x) = \left(\frac{1}{3}x^2 + \frac{1}{3}x + \frac{2}{3}\right) \cdot a(x) + (-1)$

6. $(a(x), b(x)) = x - 1$
   $x - 1 = a(x)(x^5 - x^3 - x) + b(x)(-\frac{1}{2})$

7. $(a(x), b(x)) = 1$
   $1 = a(x)(\frac{1}{17}x^2 - \frac{3}{17}x + \frac{7}{17}) + b(x)(-\frac{1}{17})$

8. $(a(x), b(x)) = x - 2$
   $x - 2 = a(x)(\frac{1}{18}x + \frac{2}{9}) + b(x)(-\frac{1}{18}x^2 - \frac{3}{18}x + \frac{1}{3})$

§3.4

1. $q(x) = 3x^2 + 2x + 2$
   $r(x) = 11$

2	3	−4	−2	7
		6	4	4
	3	2	2	11

2. $q(x) = x^4 - 5x^3 + 10x^2 - 27x + 85$
   $r(x) = -254$

−3	1	−2	−5	3	4	1
		−3	15	−30	81	−255
	1	−5	10	−27	85	−254

3. $q(x) = 2x^3 + 4x^2 + 4$
   $r(x) = -1$

$\frac{1}{2}$	2	3	−2	4	−3
		1	2	0	2
	2	4	0	4	−1

4. $q(x) = 3x^2 - 5x + (14/3)$
   $r(x) = -77/9$

$-\frac{1}{3}$	3	−4	3	−7
		−1	5/3	−14/9
	3	−5	14/3	−77/9

5. $q(x) = 3x^2 + 9x + 23$
   $r(x) = 64$

3	3	0	−4	−5
		9	27	69
	3	9	23	64

6. $q(x) = 2x^4 + 2x^3 - 2x^2 - 2x$
   $r(x) = 3$

1	2	0	−4	0	2	3
		2	2	−2	−2	0
	2	2	−2	−2	0	3

7. $q(x) = x^6 - 2x^5 + 4x^4 - 4x^3 + 3x^2 - 2x + 4$
   $r(x) = 0$

−1	1	−1	2	0	−1	1	2	4
		−1	2	−4	4	−3	2	−4
	1	−2	4	−4	3	−2	4	0

§3.5

1. (a) $q(x) = 3x^2 + 3x + 6$
   $r(x) = 31$

$$\begin{array}{r|rrrr} 5 & 3 & -12 & -9 & 1 \\ & & 15 & 15 & 30 \\ \hline & 3 & 3 & 6 & 31 \end{array}$$

   (b) $q(x) = x^4 - 3x^3 + 2x - 5$
   $r(x) = 4$

$$\begin{array}{r|rrrrrr} -3 & 1 & 0 & -9 & 2 & 1 & -11 \\ & & -3 & 9 & 0 & -6 & 15 \\ \hline & 1 & -3 & 0 & 2 & -5 & 4 \end{array}$$

2. (a) $p(-1) = 2$, $x + 1$ is not a factor.

$$\begin{array}{r|rrrr} -1 & 1 & -1 & 1 & 5 \\ & & -1 & 2 & -3 \\ \hline & 1 & -2 & 3 & 2 \end{array}$$

   (b) $p(3) = 0$, $x - 3$ is a factor.

$$\begin{array}{r|rrrrr} 3 & 1 & 3 & -16 & -27 & 63 \\ & & 3 & 18 & 6 & -63 \\ \hline & 1 & 6 & 2 & -21 & 0 \end{array}$$

3. (a) $(2x - 9)(3x + 8)$
   (b) $3\left[x - \left(\frac{-5+\sqrt{109}}{6}\right)\right]\left[x - \left(\frac{-5-\sqrt{109}}{6}\right)\right]$
   (c) $2\left[x - \left(-\frac{5}{4} + \frac{\sqrt{7}}{4}\,i\right)\right]\left[x - \left(-\frac{5}{4} - \frac{\sqrt{7}}{4}\,i\right)\right]$

4. (a) ±1, ±2, ±3, ±6, ±9, ±18
   (b) ±1, ±2, ±5, ±10, $\pm1/3$, $\pm2/3$, $\pm5/3$, $\pm10/3$
   (c) ±1, ±2, ±3, ±6, $\pm1/2$, $\pm3/2$, $\pm1/4$, $\pm3/4$

5. (a) $(x - 3)(2x - 1)(x + 5)$
   (b) $(x+2)\left[x - \left(\frac{-3+\sqrt{29}}{2}\right)\right]\left[x - \left(\frac{-3-\sqrt{29}}{2}\right)\right]$

6. $(x+1)(3x-1)\left[x - \left(\frac{-1+\sqrt{5}}{2}\right)\right]\left[x - \left(\frac{-1-\sqrt{5}}{2}\right)\right]$

7. $(x-1)(2x-3)\left[x - \left(\frac{3+\sqrt{5}}{2}\right)\right]\left[x - \left(\frac{3-\sqrt{5}}{2}\right)\right]$

8. $(x - 2)(2x - 3)\left[x - \left(\frac{5+\sqrt{21}}{2}\right)\right]\left[x - \left(\frac{5-\sqrt{21}}{2}\right)\right]$

§3.6

1. $x^2 + x + 1$, $x^3 + x + 1$, $x^3 + x^2 + 1$

2. $x^2 + 1$, $x^2 + x + 2$, $x^2 + 2x + 2$,
   $x^3 + 2x + 1$, $x^3 + 2x^2 + 1$,
   $x^3 + x^2 + 2x + 1$, $x^3 + 2x^2 + x + 1$,
   $x^3 + 2x + 2$, $x^3 + x^2 + 2$,
   $x^3 + x^2 + x + 2$, $x^3 + 2x^2 + 2x + 2$

3. $x^2 + 12x + 36$ with $m = 12$ is one example.

§3.7

1. (a) $a(x) \mid f(x)$
   (b) $b(x) \nmid f(x)$
   (c) $c(x) \mid f(x)$
   (d) $d(x) \nmid f(x)$

2. $(a_1 + 1)(a_2 + 1) \cdots (a_r + 1)$

3. (a) $((a(x), b(x)) = (x - 2)^3(x^2 + 5)^2$
   $[(a(x), b(x)] =$
   $(x - 2)^5(x - 4)^7(x^2 + 5)^8(x^3 + 5)^4$
   (b) $((a(x), b(x)) = (x - 2)^3(x - 3)^4(x - 4)^4$
   $[(a(x), b(x)] =$
   $(x-1)^2(x-2)^6(x-3)^5(x-4)^5(x-5)^3$

4. $(a(x), b(x)) = x - 1$
   $[a(x), b(x)] = x^8 + x^7 - 2x^6 - 2x^5 + x + 1$

5. $(a(x), b(x)) = 1$
   $[a(x), b(x)] = x^4 + 3x^3 - 2x^2 - 2x + 12$

6. $(a(x), b(x)) = x - 2$
   $[a(x), b(x)] =$
   $x^6 + 3x^5 + x^4 - 15x^3 - 23x^2 + 3x + 30$

Index

$[(a(x), b(x))]$, 90
$[(a, b)]$, 28
$[a(x), b(x)]$, 135
$[a, b]$, 64
$[a]$, 27
$(a(x), b(x))$, 101
(a, b), 46
$\mathbf{0}$, 29
$\mathbf{1}$, 30
$-a$, 18
$A \cap B$, 25
$A \cup B$, 25
$A \subseteq B$, 25
$A \times B$, 26
$A \subsetneq B$, 25
$\frac{a(x)}{b(x)}$, 90
$a(x) \mid b(x)$, 98
$a + bi$, 7
aRb, 26
$a \equiv b \pmod{n}$, 68
$a \sim b$, 26
a^{-1}, 18
$a \mid b$, 41
$a \nmid b$, 41
$\arg z$, 12
$\mathbb{C}$, 1
$\mathbb{C}(x)$, 129
$\boldsymbol{C}$, 32
$\frac{d}{dx}$, 40
$\deg p(x)$, 85
$\in$, 25
$\notin$, 25
e, 5
e^z, 15
F_n, 35
$I(x)$, 87
i, 7
$\mathrm{Im}(z)$, 7
$\mathbb{N}$, 1
$n!$, 39
$\varnothing$, 25
$P(n)$, 37
π, 5
$\mathbb{Q}$, 1
$\mathbb{Q}[x]$, 124
$\boldsymbol{Q}$, 28
$\mathcal{Q}$, 28
$\mathbb{R}$, 1
$\mathbb{R}[x]$, 91
$\mathbb{R}^2$, 11
$\boldsymbol{R}$, 90
$\mathcal{R}$, 90
$\mathrm{Re}(z)$, 7
$S(x)$, 90
$S[x]$, 84
$\mathbb{W}$, 18
$2\mathbb{Z}$, 24
$Z(x)$, 87
$\mathbb{Z}$, 1
$\mathbb{Z}[x]$, 92
$\mathbb{Z}_2[x]$, 131
$\mathbb{Z}_3[x]$, 131
$\mathbb{Z}_n$, 72
$\boldsymbol{Z}$, 30
$\mathcal{Z}$, 30
$\overline{z}$, 8
$|z|$, 10
$\sqrt{-1}$, 7
$\sqrt{-c}$, 7

absolute value
    complex number, 10
    real number, 10
abstraction, 21

addition, 1, 2
  associative law, 18, 30, 32, 33, 87, 91
  closure, 18, 29, 31, 32, 87, 91
  commutative law, 18, 30, 32, 33, 87, 91
  identity for, 2, 7, 18, 20, 21, 29, 31, 32, 87
    uniqueness of, 21
  in $\boldsymbol{C}$, 32
  in $\boldsymbol{Q}$, 29, 30, 33
  in $\boldsymbol{R}$, 90
  in $\boldsymbol{Z}$, 31, 32, 34
  in $\mathbb{Z}_n$, 72, 73
  in $S(x)$, 90
  inverse for, 2, 7, 18, 20, 21, 29, 31, 32, 87
    uniqueness of, 21
  modular, 72
  of complex numbers, 7, 15
  of functions, 24
  of polynomials, 85
  of rational numbers, 2
  of real numbers, 20
additive identity, *see* identity
additive inverse, *see* inverse
algebraic properties, 18, 87
  of addition, 18, 87
  of congruence, 69–72
  of multiplication, 18, 87
  of number systems, 17–24
  of polynomials, 84–92
algebraically closed, 129
angle, 12
angle sum formulas, 13
argument of complex number, 13, 15, 16
  definition, 12
associative law, 1
  addition, 18, 30, 32, 33, 87, 91
  multiplication, 18, 30, 32, 33, 87

base step, *see* induction
binary operation, 1
binomial coefficient, 92, 93, 95, 96
  definition, 92
Binomial Theorem, 95–97

$\mathbb{C}$, 1
$\boldsymbol{C}$, 32
calculus, 24, 129
Cartesian product, 26
circular reasoning, *see* reasoning, circular
clock arithmetic, 72
closure
  addition, 18, 29, 31, 32, 87, 91
  multiplication, 18, 29, 31, 32, 87, 91
coefficients
  binomial, 92, 93, 95, 96
    definition, 92
  complex, 128
  integer, 115–117, 120, 122, 124
  leading, 84, 85, 91, 92, 101, 102, 109, 114, 119, 120, 133, 137
  of polynomials, 84–86, 92, 96, 98, 109, 110, 112, 115, 117, 124
  real, 129
  zero, 109, 110
Combination Theorem
  for integers, 43, 46
  for polynomials, 98, 101
common divisor
  of integers, 45–47, 52, 65
  of polynomials, 136
common multiple
  of integers, 64, 65
  of polynomials, 136
commutative law, 1
  addition, 18, 30, 32, 33, 87, 91
  multiplication, 18, 21, 30, 32, 33, 87
complex numbers, 1, 7–16, 28, 32, 33, 129
  addition, 7, 15
    associative, 23
    commutative, 23
    geometric interpretation, 11
  additive inverse
    geometric interpretation, 11, 15
  definition, 7
  distributive law, 23
  division, 9, 16
  equality, 7
  geometric representation, 10–15, 32
  multiplication, 7, 15
    geometric interpretation, 11–13, 15
    powers, 13
  multiplicative inverse, 9, 16

polar form, 12, 16
reciprocal, 16
subtraction
geometric interpretation, 11
complex plane, 10, 32
complex root, *see* root of a polynomial
composite number, 36, 56, 61
definition, 56
congruence, 67–74
algebraic properties, 69–72
and cancellation, 70
as equivalence relation, 69
congruence class, 73, 74
definition, 72
representative, 73
congruence tests, 75–83
mod 10^n, 76
mod 11, 79, 82
mod 13, 81, 82
mod 2^n, 76, 82
mod 3, 78, 82
mod 5^n, 76
mod 7, 81, 82
mod 9, 78, 82
congruent modulo n, 70, 72
definition, 68
conjugate, 9, 15, 129
definition, 8
geometric interpretation, 10
coprime, 52
counting numbers, 1

de Moivre's Theorem, 13, 39
decimal expansion, 2, 3
non-repeating, 5
non-terminating, 3
of irrational numbers, 5
of rational numbers, 3–6
repeating, 3–6
terminating, 3–6
deductive reasoning, 35
degree, 85, 91, 92, 100, 109, 110, 112, 114, 115, 120
definition, 85
odd, 129
of a monomial, 85
of constant, 85
of product, 88, 89, 125
of sum, 89
of zero, 85, 88, 101
properties, 88
derivative, 123
differentiation, 40
digits, 3–5, 75
alternating sum of, 79
hundreds, 75
ones, 75
tens, 75
discriminant, 115, 129
distributive law, 1, 18, 22, 30, 32, 33, 86, 87
left, 18
right, 18
dividend
integer, 45, 48
polynomial, 101
divides
integer, 43
definition, 41
polynomial, 113, 132
definition, 98
divisibility
of integers, 41–43, 62, 74
and prime factorizations, 62–66
and signs, 43, 99
properties, 42, 46, 50, 52, 74, 98
of polynomials, 98–108
and constant multiples, 99, 125
and irreducible factorization, 134
properties, 98, 102, 107, 108
divisibility tests, 75–83
for 10, 77
for 10^n, 77
for 11, 80, 83
for 12, 79
for 13, 81, 83
for 14, 83
for 15, 83
for 18, 83
for 2, 77
for 20, 83

for 2^n, 77, 83
for 3, 78, 83
for 4, 77
for 5, 77
for 5^n, 77
for 6, 79, 83
for 7, 81, 83
for 8, 77
for 9, 78, 83
divisible
integer, 41, 52
polynomial, 98, 136
division, 1, 2, 7, 21, 70
of complex numbers, 9, 16
of fractions, 21
of integers, 3
long, 3, 6, 44, 45, 49
of polynomials, 100
long, 100, 102, 108, 109
of rational numbers, 21
synthetic, 109–111, 113, 122
Division Algorithm
for integers, 3, 45–47, 49, 67, 69, 100
for polynomials, 101, 112
divisor
common, *see* common divisor
greatest common, *see* greatest common divisor
integer, 41, 43, 45, 46, 48, 50, 56, 62, 66, 119, 120
number of, 63, 66
polynomial, 98, 101, 109, 110, 124
number of, 138
prime, 61

e, 5
Eisenstein's Criterion, 127, 128
element, 25, 26
empty set, 25, 27
definition, 25
equality, 26
of complex numbers, 7
of fractions, 29
of polynomials, 85
of rational numbers, 2
of sets, 25
equivalence class, 27–31, 72, 90
definition, 27
equivalence relation, 26–28, 30, 33, 69, 72, 90
definition, 26
equivalent, 26, 27
fractions, 28, 30
Eratosthenes, Sieve of, *see* Sieve of Eratosthenes
Euclid, 58
Euclid's Lemma
for integers, 54, 55, 58
for irreducible polynomials, 132
for polynomials, 107, 132
for primes, 58, 72, 132
Euclidean Algorithm
for integers, 46–50, 66
for polynomials, 101, 102, 105, 106, 138
Euler's Formula, 15
evaluation of polynomials, 112, 113, 119, 122
exponential function, 15

factor
integer, 41
polynomial, 98, 113, 114, 120, 122–124, 133
linear, 114, 129, 130
quadratic, 130
Factor Theorem, 113, 118, 125
factorial, 39, 93
factorization
canonical, 134, 138
irreducible, 133
and divisibility, 134
Fermat, 35
Fermat numbers, 35, 36
field, 20, 23, 24, 28, 30, 32, 33, 72–74, 89, 90, 98, 124, 125, 129
definition, 20
finite, 74
of fractions, 28
fractions, 2, 4, 6, 19, 28, 29
division, 21
equality, 29
equivalent, 28, 30
lowest terms, 6, 59
functions, 20, 24, 112

continuous, 24
Fundamental Theorem
of Algebra, 128, 129
Real Number Version, 130
of Arithmetic, 59

Gauss, 128
Gauss's Lemma, 117, 120, 124
Gaussian integers, 24
GCD, *see* greatest common divisor
Goldbach Conjecture, 36
graph, 113, 120
greatest common divisor, 46
of integers, 44–50, 59, 62, 64
alternate characterization, 52
and prime factorizations, 64, 66
and signs, 50
as combination, 51, 55
definition, 45
properties, 50–55
relation to LCM, 65, 66
uniqueness of, 46
of polynomials, 101, 102, 106, 108, 138
alternate characterization, 106
and irreducible factorization, 135
as combination, 103, 104, 106, 108
definition, 101
properties, 106, 108
relation to LCM, 137
uniqueness of, 101, 106

i, 7
identity
additive, 2, 7, 18, 20, 21, 29, 31, 32, 87
definition, 18
uniqueness of, 21
multiplicative, 2, 7, 18, 21, 22, 30, 31, 33, 87
definition, 18
uniqueness of, 22
imaginary axis, 10
imaginary part, 15
definition, 7
induction, 13, 43, 47, 74, 93
base step, 37
inductive step, 37
Principle of Mathematical, 35–40
inductive hypothesis, 37
inductive reasoning, 35, 36
infinite series, 15
infinite sum, 3
integers, 1, 2, 5, 18, 19, 23, 28, 30, 32, 35–37, 41, 45, 73, 74, 85, 87, 89, 124
divisibility, 41–43
properties, 42
even, 24, 36
mod n, 72
integral domain, 28
intercepts, 113, 114, 120
Intermediate Value Theorem, 129
intersection, 25
inverse
additive, 2, 7, 18, 20–22, 29, 31, 32, 87
definition, 18
uniqueness of, 21
multiplicative, 2, 7, 9, 18, 19, 21, 22, 30–34, 70, 73, 74, 124
definition, 18
modulo n, 70, 71
uniqueness of, 22
inverse operation, 1
irrational numbers, 5, 28, 59, 128
decimal expansion, 5
rational approximations, 5
irrational root, *see* root of a polynomial
irreducible element, 124
Irreducible Factor Principle, 132
isomorphic, 30, 32, 33, 72, 90

laws of exponents, 15, 86
LCM, *see* least common multiple
least common multiple
of integers, 62, 64, 117
and prime factorizations, 64, 66
definition, 64
relation to GCD, 65, 66
of polynomials, 138
and irreducible factorization, 135
definition, 135
relation to GCD, 137
uniqueness of, 135
least residue, *see* residue

linear algebra, 11

matrices, 20
modular arithmetic, 72
modulus
  of complex number, 12, 13, 15, 16
    definition, 10
  of congruence, 68, 72
monomial, 84
multiple
  common, *see* common multiple
  integer, 41, 51
  least common, *see* least common multiple
  polynomial, 98, 104
    constant, 99, 106
multiplication, 1, 2
  associative law, 18, 30, 32, 33, 87
  closure, 18, 29, 31, 32, 87, 91
  commutative law, 18, 21, 30, 32, 33, 87
  identity for, 2, 7, 18, 21, 22, 30, 31, 33, 87
    uniqueness of, 22
  in $\boldsymbol{C}$, 32
  in $\boldsymbol{Q}$, 29, 30, 33
  in $\boldsymbol{R}$, 90
  in $\boldsymbol{Z}$, 31, 32, 34
  in $\mathbb{Z}_n$, 72, 73
  in $S(x)$, 90
  inverse for, 2, 7, 9, 18, 19, 21, 22, 30–34, 70, 73, 74, 124
    uniqueness of, 22
  modular, 72
  of complex numbers, 7, 15
  of functions, 24
  of polynomials, 85, 89, 91, 92
  of rational numbers, 2
  of real numbers, 20
multiplicative identity, *see* identity
multiplicative inverse, *see* inverse
multiplicity, 114

$\mathbb{N}$, 1
n choose r, 92
natural numbers, 1, 2, 18, 23, 28, 30–32, 35–37, 40
negative numbers, 1, 2, 6, 18, 20, 28, 30, 34
number line, 10
Number Theory, 35
numbers
  complex, *see* complex numbers
  integer, *see* integers
  irrational, *see* irrational numbers
  natural, *see* natural numbers
  prime, *see* prime numbers
  rational, *see* rational numbers
  real, *see* real numbers

or, 26
ordered pair, 26, 28, 30, 32
origin, 10

partition, 26, 27
Pascal's Rule, 93
Pascal's Triangle, 93
Peano axioms, 28
perfect square, 59, 61, 67, 74, 115
period, 4, 6
π, 5
polar coordinates, 12
polar form, *see* complex numbers
polynomials, 20, 28, 70, 84, 90, 112, 113
  addition, 85, 112
    associative law, 87, 91
    closure, 87, 112
    commutative law, 87, 91
    identity for, 87
    inverse for, 87, 88
  algebraic properties, 84–92
  constant, 85, 88, 89, 98–100, 110, 112, 124
  cubic, 85
    roots of, 115
  definition, 84
  distributive law, 87
  equality, 85
  fifth degree
    roots of, 115
  irreducible, 124–138
    definition, 124
    over $\mathbb{C}$, 129
    over $\mathbb{Q}$, 124–128, 131
    over $\mathbb{R}$, 124, 129, 130
    over $\mathbb{Z}$, 124
    over $\mathbb{Z}_2$, 131

over $\mathbb{Z}_3$, 131
linear, 85
monic, 84, 101, 102, 106, 108–110, 119, 133, 135, 137, 138
multiplication, 85, 89, 91, 92, 112
associative law, 87
closure, 87, 91, 92, 112
commutative law, 87
identity for, 87
inverse for, 88, 89
non-constant, 89, 117, 124
non-zero, 89, 91, 100
quadratic, 85, 115–117, 129
factors of, 116, 117
roots of, 115–117
quartic
roots of, 115
reducible, 124, 125
definition, 124
with integer coefficients, 116–122, 126, 127
factors of, 116
roots of, 116
with rational coefficients, 117
roots of, 117
zero, 85, 88
prime, *see* prime numbers
Prime Divisor Principle, 57, 132
prime factorization, 36
and divisibility, 62–66
and greatest common divisor, 64, 66
and least common multiple, 64, 66
canonical, 61–63
definition, 60
exponents in, 62, 63
for comparison, 62
uniqueness of, 59, 60
prime numbers, 35, 36, 56–61, 72–74, 124, 127, 128, 132
definition, 56
even, 56
infinitely many, 58
largest known, 56, 58
websites, 56
Prime Test, 57, 61
product rule, 40, 123

$\mathbb{Q}$, 1
$\boldsymbol{Q}$, 28
$\mathcal{Q}$, 28
Quadratic Formula, 115, 116, 122, 126
quotient
of complex numbers, 9, 16
of integers, 2, 3, 5, 19, 44–46, 49
of polynomials, 90, 101, 108–110, 122

$\mathbb{R}$, 1
$\boldsymbol{R}$, 90
$\mathcal{R}$, 90
radicals, 115
rational expression, 90
rational functions, 90
field of, 90, 129
rational numbers, 1–6, 19, 21, 28–30, 36, 41, 45, 59, 72, 89, 117
addition, 2
commutative, 23
well-defined, 6
decimal expansion, 3–6
definition, 2
distributive law, 23
division, 21
equality, 2
multiplication, 2
associative, 23
well-defined, 6
multiplicative inverse, 2
rational root, *see* root of a polynomial
Rational Root Test, 118
for monic polynomials, 119
rationalizing, 9
real axis, 10, 12
real numbers, 1, 2, 5, 7, 9, 15, 20, 21, 23, 24, 28, 32, 36, 40, 114
addition, 20
multiplication, 20
real part, 15
definition, 7
real root, *see* root of a polynomial
reasoning, circular, *see* circular reasoning
reciprocal, 2, 16, 19, 21
reflexive, 26

related, 26
relation, 26–28, 30
  definition, 26
relatively prime
  integers, 53, 74
    definition, 52
  polynomials, 106, 123
    definition, 106
remainder
  integer, 3, 44–49, 67, 68
    uniqueness of, 45
  polynomial, 101, 105, 108–110, 112, 113, 122
Remainder Theorem, 112, 113, 122
residue
  complete set, 72
  least (non-negative), 74, 75, 82
    definition, 69
ring, 20–24, 28, 32, 72, 74, 98, 112, 124
  commutative, 20, 24, 32, 72, 73, 84, 88
    definition, 20
  definition, 20
  non-commutative, 21
  with identity, 20, 22, 24, 32, 72, 73, 88
    definition, 20
root
  of a number, 129
  of a polynomial, 113–116, 123, 125, 129
    complex, 115, 116, 122, 128, 129
    definition, 112
    distinct, 115
    integer, 119
    irrational, 116, 120
    multiplicity, 114, 115
    potential rational, 119, 120, 122
    rational, 115–120, 122, 126, 131
    real, 115, 129
    repeated, 114, 123
  of a prime, 128
  square
    of a prime, 59, 128
    of an integer, 59

set, 25–27, 29, 31, 36, 40, 72
  containment, 25
    proper, 25
  definition, 25
  disjoint, 26
  empty, 25, 27
  equality, 25
  non-empty, 36
Sieve of Eratosthenes, 57, 61
square, 115
subset, 25, 26, 28, 30, 40, 85
  definition, 25
  disjoint, 27
  proper, 25
subtraction, 1, 2, 20
symmetric, 26
synthetic division, 109–111, 113, 122

terms, 84, 92, 110
  constant, 84, 92, 110, 119, 120
  leading, 84, 85
transitive, 26
trigonometry, 12

union, 25, 27
  disjoint, 26, 27
Unique Factorization Theorem, 133
unit circle, 12, 13

variable, 84
vector addition, 11
vector space, 11

well-defined, 6, 29, 31, 33, 34, 73
Well-ordering Principle, 36, 57, 132

$\mathbb{Z}$, 1
$\boldsymbol{Z}$, 30
$\mathcal{Z}$, 30
zero of a polynomial, 109, 110
  definition, 112