1 The Integers

The fundamental object in elementary number theory is the integers. Our goal in this chapter is to define the integers axiomatically and to develop some basic properties of primes and divisibility. We then introduce some other number systems, in order to contrast their properties with those of the integers.

1.1 The Integers, Axiomatically

• We are all quite familiar with the integers \( \mathbb{Z} \), consisting of the natural numbers \( \mathbb{N} \) (1, 2, 3, 4, \ldots), along with their negatives (\(-1, -2, -3, -4, \ldots\)) and zero (0).

• But it is not quite so simple to prove things about the integers without a solid set of properties to work from.

• In order to put everything on rigorous ground, we define the integers using the following axioms:

  (R1) The operations \(+\) and \(\cdot\) are associative: thus, \(a + (b + c) = (a + b) + c\) and \(a \cdot (b \cdot c) = (a \cdot b) \cdot c\).

  (R2) The operations \(+\) and \(\cdot\) are commutative: thus, \(a + b = b + a\) and \(a \cdot b = b \cdot a\).

  (R3) There is an additive identity 0 satisfying \(a + 0 = a\) for all \(a \in \mathbb{Z}\).

  (R4) Every \(a \in \mathbb{Z}\) has an additive inverse \(-a\) satisfying \(a + (-a) = 0\).

  (R5) The operation \(\cdot\) distributes over \(+\): thus, \(a \cdot (b + c) = a \cdot b + a \cdot c\).

  (R6) There is a multiplicative identity 1 \(\neq 0\), satisfying \(1 \cdot a = a\) for all \(a \in \mathbb{Z}\).

  Furthermore, there is a subset of \(\mathbb{Z}\), called \(\mathbb{N}\), such that

  (N1) For every \(a \in \mathbb{Z}\), precisely one of the following holds: \(a \in \mathbb{N}\), \(a = 0\), or \((-a) \in \mathbb{N}\).

  (N2) \(\mathbb{N}\) is closed under \(+\) and \(\cdot\).

  (N3) Every nonempty subset \(S\) of \(\mathbb{N}\) contains a smallest element: that is, an element \(x \in S\) such that if \(y \in S\), then either \(y = x\) or \(y - x \in \mathbb{N}\).

• Remark: As is natural, we can define the binary operation of subtraction by setting \(a - b = a + (-b)\).

• Definition: Using the definition of \(\mathbb{N}\), we can define a relation “<” by saying \(a < b\) if and only if \(b - a \in \mathbb{N}\). (We define \(b > a\) to be the same thing.) The axioms (N1) and (N2) ensure that this symbol behaves in the way we would expect an inequality symbol to behave: for any \(a\) and \(b\), exactly one of \(a < b\), \(a = b\), or \(b < a\) holds, \(a < b\) and \(b < c\) imply \(a < c\), and \(a < b\) with \(0 < c\) implies \(ac < bc\).

• Definition: The axiom (N3) is often called the well-ordering principle. It is the axiom that differentiates the integers from other number systems (such as the rational numbers or the real numbers, which obey all of the other axioms).
1.2 Basic Arithmetic and Induction

- Using the axioms for \( \mathbb{Z} \), we can establish all of the properties of basic arithmetic. Doing this is not especially difficult: it merely requires applying a few of the axioms and some case analysis. Here are some examples:
  
  - The additive and multiplicative identities are unique. Additive inverses are unique.
  - For all \( a \in \mathbb{Z} \), \( 0 \cdot a = 0 \), \( (-1) \cdot a = -a \), and \( -(a) = a \).
  - For any \( a \) and \( b \), \( b \cdot a = a + (b-1) \cdot a \). Thus, \( 2 \cdot a = a + a \), and so forth.
  - The multiplicative identity \( 1 \in \mathbb{N} \).
  - If \( ab = 0 \), then \( a = 0 \) or \( b = 0 \). If \( ab = 1 \), either \( a = b = 1 \) or \( a = b = -1 \).

- A rather obvious yet bizarrely important property of the integers is the following result:

- **Proposition:** There are no integers between 0 and 1.
  
  - **Proof:** Let \( S = \{ r \in \mathbb{N} : 0 < r < 1 \} \). If \( S \) is empty, we are done, so assume \( S \neq \emptyset \).
  - By the well-ordering principle, \( S \) has a minimal element \( r \).
  - Now observe that since \( 0 < r < 1 \), \( 0 < r^2 < r < 1 \). But this is a contradiction, because \( r^2 \) is then a positive integer less than \( r \), but \( r \) was assumed to be minimal.

- We can now establish the validity of “proof by induction”:

- **Proposition:** If \( S \) is a set of positive integers such that \( 1 \in S \), and \( n \in S \) implies \( (n+1) \in S \), then \( S = \mathbb{N} \).
  
  - **Proof:** Let \( T = \mathbb{N} \setminus S \), the set of elements of \( \mathbb{N} \) not in \( S \). If \( T \) is empty, we are done, so assume \( T \neq \emptyset \).
  - By the well-ordering principle, \( T \) has a minimal element \( r \).
  - Since \( r \) is positive, there are three possibilities: \( 0 < r < 1 \), \( r = 1 \), or \( 1 < r \).
  - Since there are no positive integers between 0 and 1, and \( 1 \in S \), the only remaining possibility is that \( 1 < r \). But then \( 0 < r - 1 \), so \( r - 1 \) is a positive integer. Since \( r - 1 < r \) and \( r \) is minimal, we see that \( r - 1 \in S \). But the hypotheses on \( S \) then imply \( r \in S \), which is a contradiction since we assumed \( r \in T \).
  - Hence \( T = \emptyset \), so \( S = \mathbb{N} \).

- One typically phrases a proof by induction in the following manner: if \( P(n) \) is a proposition such that the “base case” \( P(1) \) holds, and we can also show the “inductive step” that \( P(n) \) implies \( P(n+1) \) for all \( n \geq 1 \), then \( P(k) \) is true for every positive integer \( k \).

- **Example:** Show that \( 1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6} \).
  
  - **Proof:** We show this by induction on \( n \).
  - **Base Case:** \( n = 1 \): both sides are clearly equal to 1 if \( n = 1 \).
  - **Inductive Step:** Assuming that \( 1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6} \),

  we want to show that

  \[
  1^2 + 2^2 + \cdots + n^2 + (n+1)^2 = \frac{(n+1)(n+2)(2n+3)}{6}.
  \]

  - We add \((n+1)^2\) to the left-hand side of the first equality: this gives

  \[
  1^2 + 2^2 + \cdots + n^2 + (n+1)^2 = \frac{n(n+1)(2n+1)}{6} + (n+1)^2
  = (n+1) \cdot \frac{2n^2 + n + 6n + 6}{6}
  = (n+1)(n+2)(2n+3)
  \]

  which is the desired result.
We will also note that induction is equivalent to “strong induction”: in strong induction, to show that \( P(n+1) \) holds in the inductive step, one assumes that \( P(k) \) holds for all \( k \leq n \), rather than merely the fact that \( P(n) \) alone holds.

1.3 Divisibility and the Euclidean Algorithm

- We have constructed three of the operations of standard arithmetic: +, −, and ·. We now turn to division.

- **Definition:** If \( a \neq 0 \), we say that \( a \) **divides** \( b \), written \( a \mid b \), if there is an integer \( k \) with \( b = ka \).
  - **Examples:** \( 2 \mid 4 \), \( (−7) \mid 7 \), and \( 6 \mid 0 \).

- There are a number of basic properties of divisibility that follow immediately from the definition and properties of arithmetic:
  - If \( a \mid b \), then \( a \mid bc \) for any \( c \).
  - If \( a \mid b \) and \( b \mid c \), then \( a \mid c \).
  - If \( a \mid b \) and \( a \mid c \), then \( a \mid (xb + yc) \) for any \( x \) and \( y \).
  - If \( a \mid b \) and \( b \mid a \), then \( a = \pm b \).
  - If \( a \mid b \) and \( a, b > 0 \), then \( a \leq b \).
  - For any \( m \neq 0 \), \( a \mid b \) is equivalent to \( (ma) \mid (mb) \).

- If \( 0 < b < a \) and \( b \) does not divide \( a \), we can still attempt to divide \( a \) by \( b \) to obtain a quotient and remainder: this is a less-explicit version of the long-division algorithm familiar from elementary school. Formally:

- **Theorem:** If \( a \) and \( b \) are positive integers, then there exist unique integers \( q \) and \( r \) such that \( a = qb + r \) with \( 0 \leq r < b \). Furthermore, \( r = 0 \) if and only if \( b \mid a \).
  - **Proof:** The last statement follows immediately from the first part.
  - To show existence, let \( T \) be the intersection of the set \( S = \{a + kb, k \in \mathbb{Z}\} \) with the positive integers. Observe that since \( a \in S \), \( T \) is nonempty.
  - Let \( r \) be the minimal element of \( T \): then \( 0 \leq r \), and since \( r - a \) is not in \( T \) by minimality, we also have \( r < a \). But since \( r \) is in the set \( S \), we must have \( r = a - qb \) for some integer \( q \).
  - For uniqueness, suppose \( qb + r = a = q'b + r' \) with \( 0 \leq r, r' < b \). Then \( −b < r − r' < b \), but we can write \( r − r' = b(q' − q) \), so dividing through by \( b \) yields \( −1 < q' − q < 1 \). But since \( q' − q \) is an integer and there are no integers between 0 and 1 (or -1 and 0), it must be the case that \( q' = q \) and \( r' = r \).
  - **Example:** If \( a = 25 \) and \( b = 4 \), then the set \( S = \{-7,−3,1,5,…,21,25,29,33,…\} \), and \( T = \{1,5,9,…\} \). The minimal element of \( T \) is \( r = 1 \), and then we obtain \( q = \frac{a − r}{b} = 6 \). And indeed, we have \( 25 = 6 \cdot 4 + 1 \).
  - In practice, of course, we would not actually construct the sets \( S \) and \( T \) to determine \( q \) and \( r \): we would just numerically compute \( 25/4 \) and round down to the nearest integer to find \( q \).

- **Definition:** If \( d \mid a \) and \( d \mid b \), then \( d \) is a **common divisor** of \( a \) and \( b \). If \( a \) and \( b \) are not both zero, then there are only a finite number of common divisors: the largest one is called the greatest common divisor, or \( \text{gcd} \), and denoted by \( \text{gcd}(a, b) \).
  - **Warning:** Many authors use the notation \( (a, b) \) to denote the gcd of \( a \) and \( b \) – this stems from the notation used for ideals in ring theory. The author of these notes generally dislikes using this notation and will write \( \text{gcd} \) explicitly, since otherwise it is easy to confuse the \( \text{gcd} \) with an ordered pair \( (a, b) \).
  - **Example:** The positive divisors of 30 are 1, 2, 3, 5, 6, 10, 15, 30. The positive divisors of 42 are 1, 2, 3, 6, 7, 14, 21, 42. The common (positive) divisors are 1, 2, 3, and 6, and the gcd is therefore 6.

- **Theorem:** If \( d = \text{gcd}(a, b) \), then there exist integers \( x \) and \( y \) with \( d = ax + by \): in fact, the gcd is the smallest positive such linear combination.
This theorem says that the greatest common divisor of two integers is an integral linear combination of those integers.

**Proof:** Without loss of generality assume \( a \neq 0 \), and let \( S = \{as + bt : s, t \in \mathbb{Z}\} \cap \mathbb{N} \).

Clearly \( S \neq \emptyset \) since one of \( a \) and \( -a \) is in \( S \), so now let \( l = ax + by \) be the minimal element of \( S \).

We claim that \( l \mid b \).

- Apply the division algorithm to write \( b = ql + r \) for some \( 0 \leq r < l \).
- Observe that \( r = b - ql = b - q(ax + by) = a(-qx) + b(1 - qy) \) is a linear combination of \( a \) and \( b \). It is not negative, but it also cannot be positive because otherwise it would necessarily be less than \( l \), and \( l \) is minimal.
- Hence \( r = 0 \), so \( l \mid b \).

By a symmetric argument, \( l \mid c \), and so \( l \) is a common divisor of \( a \) and \( b \), whence \( l \leq d \).

But now since \( d \mid a \) and \( d \mid b \) we can wrote \( a = dk_a \) and \( b = dk_b \) for some integers \( k_a \) and \( k_b \), and then \( l = ax + by = dk_a x + dk_b y = d(k_a x + k_b y) \).

Therefore \( d \mid l \), so in particular \( d \leq l \) since both are positive. Since \( l \leq d \) as well, we conclude \( l = d \).

**Corollary:** If \( l \mid a \) and \( l \mid b \), then \( l \) divides \( \gcd(a, b) \).

**Proof:** Since \( l \mid a \) and \( l \mid b \), \( l \) divides any linear combination of \( a \) and \( b \); in particular, it divides the gcd.

- As an example: we saw above that the gcd of 30 and 42 is 6, and indeed we can see that \( 3 \cdot 30 - 2 \cdot 42 = 6 \). The other common divisors are 1, 2, and 3, and indeed they all divide 6.
- As another example: because \( 6 \cdot 24 - 11 \cdot 13 = 1 \), we see that 24 and 13 have greatest common divisor 1, since their gcd must divide any linear combination. Having a gcd of 1 occurs often enough that we give this situation a name:

**Definition:** If \( \gcd(a, b) = 1 \), we say \( a \) and \( b \) are relatively prime.

Using these results we can quickly prove a number of useful facts about greatest common divisors:

- If \( m > 0 \), then \( m \cdot \gcd(a, b) = \gcd(ma, mb) \): we can write \( \gcd(ma, mb) = \min_{x, y \in \mathbb{Z}} [(max + mby) \cap \mathbb{N}] = m \cdot \min_{x, y \in \mathbb{Z}} [(ax + by) \cap \mathbb{N}] = m \cdot \gcd(a, b) \).

- If \( d > 0 \) divides both \( a \) and \( b \), then \( \gcd(a/d, b/d) = \gcd(a, b)/d \): simply apply the above result to \( a/d \) and \( b/d \) with \( m = d \), and rearrange.

- If \( a \) and \( b \) are both relatively prime to \( m \), then so is \( ab \): there exist \( x_1, y_1, x_2, y_2 \) with \( ax_1 + my_1 = 1 \) and \( bx_2 + my_2 = 1 \). Multiplying yields \( ab(x_1x_2) + m(y_1bx_2 + y_2ax_1 + my_1y_2) = 1 \), meaning that \( ab \) and \( m \) are relatively prime.

- For any integer \( x \), \( \gcd(a, b) = \gcd(a, b + ax) \): the set of linear combinations of \( a \) and \( b \) is the same as the set of integral linear combinations of \( a \) and \( b + ax \).

- If \( c \mid ab \) and \( b, c \) are relatively prime, then \( c \mid a \): by the first property listed, \( \gcd(ab, ac) = a \cdot \gcd(b, c) = a \). Since \( c \mid ab \) and \( c \mid ac \), we conclude \( c \mid a \).

- One question raised by the previous theorems is: how can we actually compute the gcd, except by actually writing down lists of common divisors? And how can we compute the gcd as a linear combination of the original integers? Both questions have a nice answer:

**Theorem (Euclidean Algorithm):** Given integers \( 0 < b < a \), repeatedly apply the division algorithm as follows, until a remainder of zero is obtained:

\[
\begin{align*}
a &= q_1b + r_1 \\
b &= q_2r_1 + r_2 \\
r_1 &= q_3r_2 + r_3 \\
&\vdots \\
r_{k-1} &= q_kr_{k-1} + r_{k+1} \\
r_k &= q_{k+1}r_{k+1}.
\end{align*}
\]

This algorithm can be used to compute the gcd of two integers in a systematic way.
Then \( \gcd(a, b) \) is equal to the last nonzero remainder, \( r_{k+1} \). Furthermore, by successively solving for the remainders and plugging in the previous equations, \( r_{k+1} \) can be explicitly written as a linear combination of \( a \) and \( b \).

- **Proof:** First observe that the algorithm will eventually terminate, because \( b > r_1 > r_2 > \cdots \geq 0 \), and the well-ordering principle dictates that there cannot exist an infinite decreasing sequence of nonnegative integers.
  - We now claim that \( \gcd(a, b) = \gcd(b, r_1) \): this follows because \( \gcd(b, r_1) = \gcd(b, a - q_1b) = \gcd(b, a) \) from the gcd properties we proved earlier.
  - Now we can repeatedly apply this fact to see that \( \gcd(a, b) = \gcd(b, r_1) = \gcd(r_1, r_2) = \cdots = \gcd(r_k, r_{k+1}) = r_{k+1} \) since \( r_{k+1} \) divides \( r_k \).
  - The correctness of the algorithm for computing the gcd as a linear combination follows by an easy induction.

- **Example:** Find the gcd of 30 and 42 using the Euclidean Algorithm, and write the gcd explicitly as a linear combination of 30 and 42.
  - First, we divide:
    
    \[
    \begin{align*}
    42 &= 1 \cdot 30 + 12 \\
    30 &= 2 \cdot 12 + 6 \\
    12 &= 2 \cdot 6.
    \end{align*}
    \]
    Since 6 is the last nonzero remainder, it is the gcd.
  - For the linear combination, we solve for the remainders:
    
    \[
    \begin{align*}
    12 &= 42 - 1 \cdot 30 \\
    6 &= 30 - 2 \cdot 12 = 30 - 2 \cdot (42 - 1 \cdot 30) = 3 \cdot 30 - 2 \cdot 42
    \end{align*}
    \]
    so we obtain \( \boxed{6 = 3 \cdot 30 - 2 \cdot 42} \).

- **Definition:** If \( a \mid l \) and \( b \mid l \), \( l \) is a common multiple of \( a \) and \( b \). Among all (nonnegative) common multiples of \( a \) and \( b \), the smallest such \( l \) is called the least common multiple of \( a \) and \( b \).
  - **Example:** The least common multiple of 30 and 42 is 210.
  - The least common multiple is often mentioned in elementary school in the context of adding fractions (for finding the “least common denominator”).

- The least common multiple has fewer nice properties than the gcd. It does obey the relation \( m \cdot \text{lcm}(a, b) = \text{lcm}(ma, mb) \):
  - Since \( ma \) divides \( \text{lcm}(ma, mb) \), we can write \( \text{lcm}(ma, mb) = mk \) for some integer \( k \). Then \( ma \mid mk \) and \( mb \mid mk \), so \( a \) and \( b \) both divide \( k \), whence \( k \geq l \), where \( l = \text{lcm}(ma, mb) \).
  - On the other hand, certainly \( ma \) and \( mb \) divide \( ml \), so \( ml \geq mk \). We must therefore have \( l = k \).

- **Proposition:** If \( a, b > 0 \), the gcd and lcm satisfy \( \gcd(a, b) \cdot \text{lcm}(a, b) = ab \).
  - Thus, if we want to calculate the lcm of two arbitrary integers, we can just compute the gcd using the Euclidean Algorithm, and then apply the result of this proposition to get the lcm.
  - **Proof:** First suppose \( a \) and \( b \) are relatively prime, and let \( l \) be a common multiple. Since \( a \mid l \) we can write \( l = ak \) for some integer \( k \): then since \( b \mid ak \) and \( \gcd(a, b) = 1 \), we conclude by properties of divisibility that \( b \mid k \), meaning that \( k \geq b \) and thus \( l \geq ab \). But clearly \( ab \) is a common multiple of \( a \) and \( b \), so it is the least common multiple.
  - In the general case, let \( d = \gcd(a, b) \). Then \( \gcd(a/d, b/d) = 1 \), so by the above we see that \( \text{lcm}(a/d, b/d) = ab/d^2 \). Then \( \gcd(a, b) \cdot \text{lcm}(a, b) = d \cdot d \text{lcm}(a/d, b/d) = ab \), as desired.
1.4 Primes and Unique Factorization

- **Definition:** If \( p > 1 \) is an integer, we say it is prime if there is no \( d \) with \( 1 < d < p \) such that \( d \mid p \). (In other words, if \( p \) has no proper divisors.) If \( n > 1 \) is not prime, we say it is composite.
  - The first few primes are 2, 3, 5, 7, 11, 13, 17, 19, and so forth. 1 is neither prime nor composite.

- **Proposition:** Every positive integer \( n > 1 \) can be written as a product of primes (where a “product” is allowed to have only one term).
  - The representation of \( n \) as a product of primes is called the **prime factorization** of \( n \). (For example, the prime factorization of 6 is \( 6 = 2 \cdot 3 \).) We will show in a moment that it is unique up to reordering the terms.
  - **Proof:** We use strong induction on \( n \). The result clearly holds if \( n = 2 \), since 2 is prime.
  - Now suppose \( n > 2 \). If \( n \) is prime, we are done, so assume that \( n \) is not prime, hence composite. By definition, there exists a \( d \) with \( 1 < d < n \) such that \( d \mid n \); then \( n/d \) is an integer satisfying \( 1 < n/d < n \). By the strong induction hypothesis, both \( d \) and \( n/d \) can be written as a product of primes; multiplying these two products then yields \( n \) as a product of primes.

- **Proposition:** If \( p \) is prime and \( p \mid ab \), then \( p \mid a \) or \( p \mid b \).
  - **Proof:** If \( p \mid a \) we are done, so assume that \( p \nmid a \). Consider \( \gcd(a, p) \): it divides \( p \), hence is either 1 or \( p \), but it is not \( p \) because \( p \) does not divide \( a \). Therefore, \( \gcd(a, p) = 1 \), so \( a \) and \( p \) are relatively prime. Then since \( p \mid ab \) and \( a \), \( p \) are relatively prime, we see that \( p \mid b \).

- **Theorem (Fundamental Theorem of Arithmetic):** Every integer \( n > 1 \) can be factored into a product of primes, and this factorization is unique up to reordering of the factors.
  - **Proof:** Suppose \( n \) is minimal and has two different factorizations: \( n = p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_l \). If any of the primes \( p_i \) and \( q_j \) were equal, we could cancel the corresponding terms and obtain a smaller \( n \), so \( p_i \neq q_j \) for any \( j \leq j \leq l \).
  - But since \( p_i \) is prime and divides \( q_1 q_2 \cdots q_l \), by repeated application of the previous proposition we see that \( p_i \) must divide one of \( q_1, q_2, \ldots, q_l \); say, \( q_i \). But the only divisors of \( q_i \) are 1 and \( q_i \), and \( p_i \) cannot be either of them. This is a contradiction, and we are done.

- To save space, we group equal primes together when actually writing out the canonical prime factorization: thus, \( 12 = 2^2 \cdot 3 \), \( 720 = 2^4 \cdot 3^2 \cdot 5 \), and so forth. More generally, we often write the prime factorization in the form \( n = \prod p_i^{n_i} \), where the \( p_i \) are some (finite) set of primes and the \( n_i \) are their corresponding exponents.

- **Proposition:** If \( a = \prod p_i^{a_i} \) and \( b = \prod p_i^{b_i} \), then \( a \mid b \) if and only if \( a_i \leq b_i \) for each \( i \). In particular, \( \gcd(a, b) = \prod p_i^{\min(a_i, b_i)} \) and \( \operatorname{lcm}(a, b) = \prod p_i^{\max(a_i, b_i)} \).
  - **Proof:** We observe that if \( b = ak \) and \( k = \prod p_i^{k_i} \), then \( a_i + k_i = b_i \). Since all exponents are nonnegative, saying that such an integer \( k \) exists is equivalent to saying that \( a_i \leq b_i \) for all \( i \).
  - The statements about the \( \gcd \) and \( \operatorname{lcm} \) then follow immediately, since (for example) the exponent of \( p_i \) in the \( \gcd \) is the largest integer that is \( \leq a_i \) and \( \leq b_i \), which is a more convoluted way of saying the minimum of \( a_i \) and \( b_i \).

- One question we might have is: how many primes are there? The most basic answer to this question is that there are infinitely many primes:

- **Theorem (Euclid):** There are infinitely many prime numbers.
  - **Proof:** Suppose there are only finitely many prime numbers \( p_1, p_2, \ldots, p_k \), and consider \( n = p_1 p_2 \cdots p_k + 1 \).
  - Since \( n \) is bigger than each \( p_i \), \( n \) cannot be prime (since it would necessarily have to be on the list).
  - Therefore \( n \) is composite. Consider the prime factorization of \( n \); necessarily at least one prime on the list must appear in it: say \( p_i \).
o. Since $p_i$ also divides $p_1 p_2 \cdots p_k$, we see that $p_i$ therefore divides $n - p_1 p_2 \cdots p_k = 1$. But this is a contradiction. Hence there are infinitely many primes.

- At this stage, we will briefly mention a few of the most famous results and open problems relating to prime numbers:
  
  o (Prime Number Theorem) Euclid’s result, while extremely elegant, does not tell us much about the actual primes themselves: for example, it does not say anything about how common the primes are. Are most numbers prime? Or are most numbers composite? A more rigorous way to frame this question is: let $\pi(n)$ be the number of primes in the interval $[1, n]$. How fast does $\pi(n)$ increase as $n$ increases: does it grow like $n$, or $\sqrt{n}$, or something else? The answer is given by the so-called “Prime Number Theorem”: $\pi(n) \sim \frac{n}{\log(n)}$, where log denotes the natural logarithm. (The notation $f(n) \sim g(n)$ means that as $n \to \infty$, the limit $\lim_{n \to \infty} f(n)/g(n) = 1$.)
  
  o (Twin Primes) Another question is: how close do primes get? It is obvious that 2 is the only even prime, so aside from 2 and 3, any pair of primes has to differ by at least 2: such pairs are called “twin primes”. One can write down a long list of twin primes: $(3, 5)$, $(5, 7)$, $(11, 13)$, $(17, 19)$, $(29, 31)$, $(41, 43)$, $(59, 61)$, and so forth. Are there infinitely many? The answer is not known, although twin primes are expected to be quite rare. However, it has been proven (as of August 2014) that there exist infinitely many pairs of primes $(p_1, p_2)$ such that $|p_2 - p_1| \leq 246$.
  
  o (Goldbach’s Conjecture) One can observe that $2 + 2 = 4$, $3 + 3 = 6$, $3 + 5 = 8$, $3 + 7 = 10$, $5 + 7 = 12$, $3 + 11 = 14$, $3 + 13 = 16$, $5 + 13 = 18$, $7 + 13 = 20$, and so forth. It appears that every even number (bigger than 4) can be written as the sum of two primes. It is not known whether this pattern continues, although it has been numerically verified for every even integer less than $10^{18}$. In 2013, a proof that every odd integer greater than 7 can be written as a sum of three primes was announced. (This result is weaker than Goldbach’s conjecture, but it is of the same type.)

1.5 The Binomial Theorem

- Definition: Let $x$ be a real number and $k \geq 0$ be an integer. The binomial coefficient $\binom{x}{k}$ is defined to be the product $\binom{x}{k} = \frac{x(x - 1)(x - 2) \cdots (x - k + 1)}{k(k - 1)(k - 2) \cdots 1} = \frac{x(x - 1)(x - 2) \cdots (x - k + 1)}{k!}$.

- Proposition: If $n$ is an integer and $S$ is a set containing $n$ elements, then $\binom{n}{k} = \frac{n!}{k!(n - k)!}$ is the number of possible ways to choose a subset of $k$ elements of $S$. In particular, $\binom{n}{k}$ is an integer for all $n$ and $k$.

  - Proof: Imagine choosing the $k$ elements from $S$, in sequential order.
  - There are $n$ choices for the first element, $n - 1$ choices for the second element, and in general there are $n - j + 1$ choices for the $j$th element, for a total of $n(n - 1) \cdots (n - k + 1)$ ways of choosing the $k$ elements.
  - Since each of the $k$! ways of choosing these specific $k$ elements in a different order yield the same subset of $k$ elements, we must divide by $k!$ to remove duplicate subsets from the count.
  - Thus, the total number of subsets is $\frac{n(n - 1) \cdots (n - k + 1)}{k!} = \binom{n}{k}$.

- The binomial coefficients properly belong more to combinatorics, since they appear quite frequently in counting problems. However, we often require them in algebra as well:

- Theorem (The Binomial Theorem): For any integer $n \geq 1$, $(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.

  - Proof: If we multiply out the product $\prod_{i=1}^{n} (x + y)$, we will obtain a total of $2^n$ terms, each of which is a string consisting of $n$ terms, each of which is either an $x$ or a $y$.
  - The number of such strings which are equivalent to $x^k y^{n-k}$ is the same as the number of strings that have exactly $k$ terms that are an $x$. 

7
• Specifying such a string is the same as specifying the \( k \) positions in the string that are an \( x \), which is in turn the same as the number of subsets of \( \{1, 2, \ldots, n\} \) having exactly \( k \) elements: but this is simply \( \binom{n}{k} \).

• There are a number of identities satisfied by the binomial coefficients. Two particularly important ones that can be shown directly using algebra are Pascal’s identity \( \binom{x}{k} + \binom{x}{k-1} = \binom{x+1}{k} \), which holds for any real \( x \), and the symmetry identity \( \binom{n}{k} = \binom{n}{n-k} \), which holds for any nonnegative integer \( n \).

  ○ Remark: Using Pascal’s identity, one can give an inductive proof of the Binomial Theorem.

### 1.6 Rings and Other Number Systems

• One of our goals in number theory is to examine properties of the integers that generalize to other number systems. To do this in a standardized way, we need to discuss some basic properties of rings.

  ○ A fuller discussion of the theory of rings belongs to abstract algebra, but the language of algebra provides the best setting in which to compare different number systems.

  ○ We will only give a brief overview here, leaving a more extensive discussion of generalizations to a later chapter.

### Definition:

A **commutative ring** is any set \( R \) that has two (closed) binary operations \( + \) and \( \cdot \) that satisfy the five axioms (R1)-(R5):

- **(R1)** The operations \( + \) and \( \cdot \) are associative: thus, \( a + (b + c) = (a + b) + c \) and \( a \cdot (b \cdot c) = (a \cdot b) \cdot c \).

- **(R2)** The operations \( + \) and \( \cdot \) are commutative: thus, \( a + b = b + a \) and \( a \cdot b = b \cdot a \).

- **(R3)** There is an additive identity \( 0 \) satisfying \( a + 0 = a \) for all \( a \in R \).

- **(R4)** Every \( a \in R \) has an additive inverse \( -a \) satisfying \( a + (-a) = 0 \).

- **(R5)** The operation \( \cdot \) distributes over \( + \): thus, \( a \cdot (b + c) = a \cdot b + a \cdot c \).

• **Definition:** If a ring also satisfies axiom (R6), we say it is a **commutative ring with identity**.

  **(R6)** There is a multiplicative identity \( 1 \neq 0 \), satisfying \( 1 \cdot a = a \) for all \( a \in R \).

• **Definition:** If a commutative ring with identity further satisfies the axiom (F), it is called a **field**.

  **(F)** Every nonzero \( a \in R \) has a multiplicative inverse \( a^{-1} \), satisfying \( a \cdot a^{-1} = 1 \).

• **Remark:** Fields may be a familiar object from linear algebra, as fields are a central underlying component of a vector space.

• Here are some prototypical examples of rings:

  ○ The integers \( \mathbb{Z} \) are a commutative ring with identity.

  ○ The set of even integers is a commutative ring that does not have an identity.

  ○ The rational numbers \( \mathbb{Q} \), the real numbers \( \mathbb{R} \), and the complex numbers \( \mathbb{C} \) are all fields.

  ○ The set of complex numbers of the form \( a + bi \) where \( a, b \in \mathbb{Z} \) are a commutative ring with identity. This ring is denoted \( \mathbb{Z}[i] \) (read as: “\( \mathbb{Z} \) adjoin \( i \)” and is also often called the Gaussian integers. Explicitly, the addition and multiplication behave as follows: \((a+bi)+(c+di) = (a+c)+(b+d)i\), and \((a+bi) \cdot (c+di) = (ac-bd) + (ad+bc)i\).

  ○ The set of real numbers of the form \( a + b\sqrt{2} \) where \( a, b \in \mathbb{Z} \) are a commutative ring with identity, and denoted \( \mathbb{Z}[\sqrt{2}] \). The addition and multiplication are defined in the same way as the Gaussian integers.

  ○ More generally, if \( D \) is any integer, the set of complex numbers of the form \( a + b\sqrt{D} \) for \( a, b \in \mathbb{Z} \) forms a ring, denoted \( \mathbb{Z}[\sqrt{D}] \). Associated to this ring is a particularly important map called the norm map, which is defined as follows: \( N(a+b\sqrt{D}) = (a+b\sqrt{D}) \cdot (a-b\sqrt{D}) = a^2 - Db^2 \). Observe that this function takes values in the integers, and that it is also multiplicative: \( N(rs) = N(r)N(s) \) for any \( r, s \in R \).
○ The set $\mathbb{R}[x]$ of polynomials with real coefficients in the variable $x$ is a commutative ring with identity, under the standard polynomial addition and multiplication. More generally, if $F$ is any field, the set $F[x]$ of polynomials with coefficients in $F$ forms a commutative ring with identity.

○ **Definition:** If $R$ is a commutative ring with identity, then an element $a \in R$ is a **unit** if there exists an element $b \in R$ such that $ab = 1$.

○ **Example:** In the integers, there are only two units, 1 and $-1$.

○ **Example:** In any field, every nonzero element is a unit (because, by definition, every nonzero element has a multiplicative inverse).

○ **Example:** In the Gaussian integers, the units are 1, $-1$, $i$, and $-i$. This can be seen as follows: suppose that $a + bi$ is a unit, with $(a + bi)(c + di) = 1$. Applying the norm map $N(a + bi) = a^2 + b^2$ to both sides yields $(a^2 + b^2)(c^2 + d^2) = 1$. But both of these quantities are nonnegative integers, so it must be the case that $a^2 + b^2 = c^2 + d^2 = 1$. But the only integral solutions to this equation are $(a, b) = (\pm 1, 0)$ and $(0, \pm 1)$, meaning that the only units are $\pm 1$ and $\pm i$.

○ **Example:** In the ring $\mathbb{Z}[\sqrt{2}]$, the integers 1 and $-1$ are units, but the element $\sqrt{2} + 1$ is also a unit, because $(\sqrt{2} + 1) \cdot (\sqrt{2} - 1) = 1$. This in fact implies that any power of $(\sqrt{2} + 1)$ is a unit in this ring, since its inverse is the corresponding power of $(\sqrt{2} - 1)$: so we see that this ring has infinitely many units, unlike the Gaussian integers. (We will see later that these are essentially the only units in this ring.) Note that $\mathbb{Z}[\sqrt{2}]$ is not a field, however, because $\sqrt{2}$ is not a unit.

○ **Example:** In the ring of polynomials $\mathbb{R}[x]$, the units are the nonzero constant polynomials. To see this we observe that if $p(x) \cdot q(x) = 1$, then $p(x)$ cannot have positive degree.

○ We can generalize the observations about units in $\mathbb{Z}[i]$ and $\mathbb{Z}[\sqrt{2}]$.

○ **Proposition:** For a fixed $D$, an element $r$ in the ring $\mathbb{Z}[\sqrt{D}]$ is a unit if and only if $N(r) = \pm 1$.

○ **Proof:** Suppose $r = a + b\sqrt{D}$.

○ If $N(r) = \pm 1$, then we see that $r \cdot (a + b\sqrt{D}) = \pm 1$, so (by multiplying by $-1$ if necessary) we obtain a multiplicative inverse for $r$.

○ Conversely, suppose $r$ is a unit and $rs = 1$. Taking norms yields $N(r)N(s) = 1$. Since $N(r)$ and $N(s)$ are both integers, we see that $N(r)$ must either be 1 or $-1$.

○ **Definition:** We can adapt the concept of divisibility directly into the setting of a general ring $R$: if $a, b \in R$, we say that $a | b$ if there exists some $k \in R$ such that $b = ak$.

○ **Example:** In $\mathbb{Z}[i]$, the element $2 + i$ divides 5, because $5 = (2 + i)(2 - i)$.

○ **Example:** In $\mathbb{R}[x]$, the polynomial $3x + 6$ divides the polynomial $x^2 - x - 6 = \frac{1}{3}(3x + 6) \cdot (x - 1)$.

○ **Warning:** If $R$ does not have a 1, bizarre things can occur with divisibility. For example, let $R$ be the ring consisting of the even integers. Then in this ring, it is not the case that $2|6$, because there is no even integer $k$ such that $6 = 2k$. Indeed, it is not even the case that $2|2$ in this ring!

○ In the case that $R$ does have a 1, we retain some of the properties of divisibility that we had over the integers:

○ If $a | b$, then $a | bc$ for any $c$ in $R$.

○ If $a | b$ and $b | c$, then $a | c$.

○ If $a | b$ and $a | c$, then $a | (xb + yc)$ for any $x$ and $y$ in $R$.

○ We will continue our examination of these properties in other rings in a later chapter.

Well, you're at the end of my handout. Hope it was helpful.

Copyright notice: This material is copyright Evan Dummit, 2014. You may not reproduce or distribute this material without my express permission.