Unit NT

Number Theory and Cryptography

Section 1: Basic Facts About Numbers

In this section, we shall take a look at some of the most basic properties of $\mathbb{Z}$, the set of integers. We look at properties related to parity (even, odd), prime factorization, irrationality of square roots, and modular arithmetic.

First we recall some standard notation for sets of various basic types of numbers.

- $\mathbb{R}$ denotes the real numbers,
- $\mathbb{Z}$ denotes the integers,
- $\mathbb{Q}$ denotes the rational numbers (ratios of integers),
- $\mathbb{N}$ denotes the nonnegative integers (the "natural numbers"),
- $\mathbb{N}^+$ denotes the nonzero natural numbers (the positive integers),
- $\mathbb{N}_2^+$ denotes the set of natural numbers greater than or equal to 2.

Note that $\mathbb{R} - \mathbb{Q}$ is the set of irrational numbers.

Example 1 (Odd and even integers) A basic subdivision of $\mathbb{Z}$ is into the odd integers and the even integers. An element of $\mathbb{Z}$ is even if it is "of the form $2t$," where $t \in \mathbb{Z}$. An element of $\mathbb{Z}$ is odd if it is not even. The odd integers are all of the form $2t + 1$, where $t \in \mathbb{Z}$. (This should be proved, but we will not do so.) The phrase "of the form $2t$" can be written precisely as

$$\forall\, n \in \mathbb{Z},\ (n \text{ is even}) \text{ if and only if } (\exists\, t \in \mathbb{Z} \text{ such that } n = 2t).$$

The most elementary mathematical facts about odd and even integers concern the *closure properties*.[1] Here is the closure property for multiplication:

The integers m and n are both odd if and only if mn is odd.

(Equivalently, by negating both sides of "if and only if," at least one the integers m or n is even if and only if mn is even.) To show the "only if" part, suppose that if m and n are both odd, say $m = 2j+1$ and $m = 2k+1$. Then $mn = 4jk+2j+2k+1 = 2(2jk+j+k)+1$ is of the form $2t + 1$ where $t = 2jk + j + k$. Thus, mn is odd. To show the "if" part, we use the inverse. Suppose that at least one of m or n is even. Without loss of generality, we may suppose that m is even, say $m = 2j$. Then $mn = 2jn$ is of the form $2t$ where $t = jn$. Thus, mn is even. A similar statement for addition is that, for integers m and n, $m + n$ is odd if and only if one of them is odd and the other is even.

[1] A function on $S \times S$ has the closure property on S if its image is contained in S. Here S is the odd integers and the function is multiplication.

From the closure property for multiplication of odd integers, you can prove by induction that for any $k \geq 1$, and any integer m, m^k is odd if and only if m is odd. Logically equivalent is that m^k is even if and only if m is even. The fact that m^k is odd if m is odd can also be proved using the binomial theorem, which you should have seen in high school:

$$(x+y)^k = \sum_{i=1}^{k} \binom{k}{i} x^i y^{k-i}.$$

Since m is odd, $m = 2j+1$ for some integer j. Let $x = 2j$ and $y = 1$. Written another way,

$$m^k = (2j+1)^k = 1 + (2j)^1 \binom{k}{1} + (2j)^2 \binom{k}{2} + \cdots + (2j)^k \binom{k}{k}.$$

In this form m^k is obviously 1 plus an even integer and hence odd. $\square$

Prime Numbers and Factorization

Most mathematicians would agree that the most important concept in number theory is the notion of a prime.

Definition 1 (Prime and composite numbers) *A natural number n is prime if $n \geq 2$ and the only divisors of n are n and 1. We denote the set of prime numbers by $\mathbb{P}$. An integer $n \geq 2$ that is not prime is composite.*

The number 2 is the smallest prime and the only even prime. The other primes less than 20 are 3, 5, 7, 11, 13, 17, 19.

Example 2 (Prime factorization of any integer $n \geq 2$) Consider the integer 226512. It ends in 2 so it is divisible by 2. (We say that "n is divisible by m," indicated by the notation $m \mid n$, if $n = qm$ for some integer q.) In fact, $226512/2 = 113256$. We can divide by 2 again, $113256/2 = 56628$; and again, $56628/2 = 28314$; and again, $38314/2 = 14157$. That's it. We can't divide by 2 anymore, so we have $226512 = 2^4 \times 14157$. But, it is easy to check that 14157 is divisible by 3 to get 4719 which is again divisible by 3 to get 1573. That's it for dividing by 3, so we have $226512 = 2^4 \times 3^2 \times 1573$. Continuing in this manner, we end up with $226512 = 2^4 \times 3^2 \times 11^2 \times 13$. We have written 226512 as a product of primes. Also, the notation $m \nmid n$ means that n is not divisible by m.

Can every integer greater than 1 be written as a product of primes? What about a single prime p? It is convenient to adopt the terminology that a single prime p is a product of one prime, itself.[2]

[2] We could go even further and say that 1 is also can be written as an empty product. In fact, mathematicians do this: They say that an empty sum is 0 and an empty product is 1. You may think this strange, but you've already seen it with exponents: The notation a^n stands for the product of n copies of a. Thus a^0 is the product of no copies of a, and you learned that we define $a^0 = 1$ when you studied exponents. This is done so that the rule $a^{n+m} = a^n a^m$ will work when $n = 0$.

In Unit IS (Induction, Sequences and Series) we use induction to prove the assertion $\mathcal{A}(n)$ for every integer $n \geq 2$ where

$$\mathcal{A}(n) = \text{“}n \text{ is a product of primes.”}$$

You might find it helpful to read the first two pages of Unit IS at this time. We start (base case) with $n = 2$, which is a prime and hence a product of primes. The induction hypothesis is the following:

“Suppose that for some $n > 2$, the assertion $\mathcal{A}(k)$ is true for all k such that $2 \leq k < n$.”

Assume the induction hypothesis and consider n. If n is a prime, then it is a product of primes (itself). Otherwise, $n = st$ where $1 < s < n$ and $1 < t < n$. By the induction hypothesis, s and t are each a product of primes. Hence $n = st$ is a product of primes. Thus $\mathcal{A}(n)$ is true and the assertion is proved by induction.

If $n \geq 2$ is an integer, the notation $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$ is commonly used to designate its prime factorization, where $p_1, p_2, \ldots p_k$ are distinct primes and all $e_i > 0$. In other words, each prime factor is raised to its highest power that divides n. Thus, $226512 = 2^4 \times 3^2 \times 1573^1$. Of course, exponents with value 1 are usually omitted, thus 1573^1 would be written 1573.

It is important to note (We won't give a proof.) that prime factorization is *unique* in the following sense. Suppose one student correctly computes a prime factorization of n and gets $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$ where she has ordered the prime factors so that $p_1 < p_2 < \cdots < p_k$. Suppose that another student also correctly computes a prime factorization of n and gets $n = q_1^{f_1} q_2^{f_2} \cdots q_j^{f_j}$ with $q_1 < q_2 < \cdots < q_j$, then $k = j$, $q_i = p_i$, and $e_i = f_i$, for $i = 1, \ldots, k$. Let's call this a theorem:

Theorem 1 (Unique prime factorization) *Every integer $n \geq 2$ can be factored into a product of primes. This factorization is unique in the sense that any two such factorizations differ only in the order in which the primes are written.*

Sometimes people think it is “obvious” that prime factorization is unique. That's not true. There are sets other than the integers where prime factorization can be defined, but it may not be unique.[3] The assumption that it is unique was used in a “proof” of Fermat's Last Theorem about a century ago. Of course, the proof was false because factorization was not unique in the set being studied. Understanding the problem led to what is known as “algebraic number theory,” which eventually led to a correct proof of Fermat's Last Theorem. □

[3] When $a, b \in \mathbb{Z}$, complex numbers of the form $a + b\sqrt{-5}$ are a type of “algebraic integer.” The set of these “integers” is denoted by $\mathbb{Z}\left[\sqrt{-5}\right] = \left\{ a + b\sqrt{-5} \;\middle|\; a, b \in \mathbb{Z} \right\}$. We have

$$6 = 2 \times 3 = \left(1 + \sqrt{-5}\right)\left(1 - \sqrt{-5}\right).$$

Since 2, 3, $1 + \sqrt{-5}$ and $1 - \sqrt{-5}$ cannot be factored further in $\mathbb{Z}\left[\sqrt{-5}\right]$, they are “primes.” Hence prime factorization is not unique for $\mathbb{Z}\left[\sqrt{-5}\right]$. The desire for uniqueness led to the concept of “ideals” in $\mathbb{Z}\left[\sqrt{-5}\right]$ and the development of “algebraic number theory.”

Now that we know that every integer $n \geq 2$ is a product of powers of primes, we can show

Theorem 2 (Infinitely many primes) *There are infinitely many primes.*

Proof: Suppose that there were only finitely many primes, say the k primes

$$\mathbb{P} = \{p_1, p_2, \ldots, p_k\}.$$

Consider the integer $n = (p_1 p_2 \cdots p_k) + 1$ gotten by taking the product of all of the primes in $\mathbb{P}$ and adding one. Clearly, $n \notin \mathbb{P}$ (it's too big). That means n is a product of primes. Let p be one of the prime factors of n. Hence n/p is an integer. For $p_i \in \mathbb{P}$, dividing n by p_i leaves a remainder of 1 and so n/p_i is not an integer. Since n/p is an integer and n/p_i is not, we cannot have $p = p_i$. Hence $p \notin \mathbb{P}$. Contradiction! Thus there cannot be finitely many primes. $\square$

Prime factorization can be used to prove things that apparently do not depend on primes. Our next example illustrates this.

Example 3 (For all $n \in \mathbb{N}$, $\sqrt{n}$ is either an integer or irrational) The integer 36 is nice because $\sqrt{36} = 6$ and 6 is an integer. Thus 36 is called a *perfect square*. A perfect square is an integer whose square root is also an integer. Suppose $\sqrt{n}$ is not an integer. How "bad" is it? For example, maybe, though not an integer, $\sqrt{n}$ is rational; that is, $\sqrt{n} = a/b$ for some integers a and b. Sadly, that can't happen. We prove this by contradiction

Suppose $\sqrt{n} = a/b$ where $b \geq 2$ and we have cancelled common factors from the numerator and denominator. Since $\sqrt{n} = a/b$, we have $nb^2 = a^2$. Let p be a prime factor of b (p exists since $b \geq 2$). Since prime factorization is unique, p is a prime factor of $nb^2 = a^2$. On the other hand, since p is a prime factor of b, it is not a prime factor of a since we have cancelled common factors to get a and b. So far, we have shown that p is a prime factor of a^2 but not a prime factor of a. In the next paragraph, we show that this is a contradiction.

For any integer x, if the prime factorization of x is $x = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$ then the prime factorization of x^2 is $x^2 = p_1^{2e_1} p_2^{2e_2} \cdots p_k^{2e_k}$. In other words, any integer x has exactly the same prime divisors as its square, x^2. Apply this with $x = a$. We have proved

Theorem 3 (Irrational square roots) *For all $n \in \mathbb{N}$, $\sqrt{n}$ is either an integer or irrational.*

We can use this to get a lot of irrational numbers. Suppose that $k^2 < n < (k+1)^2$ for some $k \in \mathbb{N}$. Taking square roots, we have $k < \sqrt{n} < k+1$. Thus $\sqrt{n}$ cannot be an integer and so it must be irrational. In particular $\sqrt{2}$, $\sqrt{3}$, $\sqrt{5}$, $\sqrt{6}$, $\sqrt{7}$, $\sqrt{8}$ are all irrational.[4]

[4] Some classical Greeks were bothered by this. They thought there should be a basic unit of length such that all the lines in a geometrical construction were integer multiples of that length, but they could prove that this was impossible: By the Pythagorean Theorem, the diagonal of a unit square has length $\sqrt{2}$, which they knew was irrational. If the side of the square were b basic units long and the diagonal were a, then $\sqrt{2} = a/b$.

There are some basic properties of irrational and rational numbers lurking beneath the surface here. *If the product xy of two numbers is irrational, one of the numbers must be irrational.* Equivalently (the contrapositive), if x and y are both rational, say $x = a/b$ and $y = c/d$, then $xy = ac/bd$ is rational. Likewise, *if the sum $x + y$ of two numbers is irrational, one of the numbers must be irrational* (prove this).

Some students think these statements mean that the product of two nonzero irrational numbers is irrational and the sum of two irrational numbers is irrational, both statements are false: $\sqrt{2} \times \sqrt{2} = 2$ and $(-\sqrt{2}) + \sqrt{2} = 0$. It is true, however, that *if $x \neq 0$ is rational and y is irrational, then the product xy is irrational.* To prove this statement, use the contrapositive. If $xy = a/b$ then $y = a/bx$. Since $x \neq 0$ is rational, say $x = c/d$, this implies that $y = ad/cb$ is rational. $\square$

Example 4 (The rational numbers are countable) We want to show that we can create a list $a_1, a_2, a_3, \ldots$ such that every rational number appears on the list. We do this as follows:

Step 1. Start the list with $0, 1/1, -1/1$ and set $k = 3$.

Step 2. Append to the list all rational numbers in reduced form where the sum of the numerator and denominator (ignoring signs) is k. Begin with the largest numerators and proceed to the smallest, listing positive numbers and then negative ones. (Thus, for $k = 3$ we append $2/1, 1/2, -1/2, -2/1$ and for $k = 4$ we append $3/1, 1/3, -1/3, -3/1$.)

Step 3. Increase k by one and go to Step 2.

The list begins

$$
\begin{array}{llllll}
 & a_1 = 0, & a_2 = 1/1, & a_3 = -1/1, & \\
k = 3: & a_4 = 2/1, & a_5 = 1/2, & a_6 = -1/2, & a_7 = -2/1, \\
k = 4: & a_8 = 3/1, & a_9 = 1/3, & a_{10} = -1/3, & a_{11} = -3/1, \\
k = 5: & a_{12} = 4/1, & a_{13} = 3/2, & a_{14} = 2/3, & a_{15} = 1/4, \\
 & a_{15} = -1/4, & a_{16} = -2/3, & a_{17} = -3/2, & a_{18} - 4/1,
\end{array}
$$

Note that each rational number occurs exactly once in the list. In some sense, the number of rational numbers is the same as the number of positive integers since we have one rational number for each positive integer (the subscript of a)!

Because we can form such a list, we say that the set of rational numbers is *countable*. More simply, people say that the rationals are countable. $\square$

Example 5 (The real numbers are *not* countable) We must show that it is impossible to form a list of the real numbers. How can we do this? We must show that, *no matter what list of real numbers we have*, there is some real number that is not on the list.

Suppose we have a list $a_1, a_2, \ldots$ of real numbers. Let d_k be the kth digit after the decimal point in a_k. For example, if $a_4 = 2.718281828\ldots$ (the number e), then $d_4 = 2$. If $d_k = 1$, let $b_k = 2$ and, if $d_k \neq 1$, let $b_k = 1$. Look at the number $r = 0.b_1b_2b_3\ldots$. We claim it is not in the list. Why is this? Suppose someone claims, for example that $a_{99} = r$. By definition, d_{99} is the ninety-ninth digit of a_{99} after the decimal point. Since $b_{99} \neq d_{99}$, the numbers r and a_{99} differ in their ninety-ninth digits. Thus $r \neq a_{99}$.

Arguments of this type are called *diagonal arguments*. Why is this? A picture can help. Here $*$ stands for a digit we are not interested in and we have dropped all the digits before the decimal points.

$$\begin{aligned} a_1 &= .d_1 * * * * * \ldots \\ a_2 &= .* \ d_2 * * * * \ldots \\ a_3 &= .* * \ d_3 * * * \ldots \\ a_4 &= .* * * \ d_4 * * \ldots \\ a_5 &= .* * * * \ d_5 * \ldots \end{aligned}$$

The digits $d_1, d_2, \ldots$ that we are changing appear in a diagonal pattern. The diagonal is not always so straightforward in a diagonal argument. $\square$

Remainders and Modular Arithmetic

We all know from elementary school that if we divide one integer x by another $d > 0$, we get a quotient q and a remainder r, where $0 \le r < d$. In other words, $x = qd + r$, $0 \le r < d$. For example, if $x = 234$ and $d = 21$, then $q = 11$ and $r = 3$. Thus, $234 = 11 \times 21 + 3$. There are 21 possible remainders that can be gotten by dividing some randomly chosen integer by 21. These remainders belong to the set $\{0, 1, 2, \ldots, 20\}$. The set $\mathbb{Z}$ of all integers can be partitioned (divided up) into 21 subsets

$$21\mathbb{Z}, \quad 21\mathbb{Z} + 1, \quad 21\mathbb{Z} + 2, \ldots, 21\mathbb{Z} + 20$$

according to these remainders. Note that, for a set S of numbers $aS + b = \{as + b \mid s \in S\}$ so that $21\mathbb{Z} + 4 = \{\ldots, -17, 4, 25, \ldots\}$. We have just seen that 234 belongs to the subset $21\mathbb{Z} + 3$. (The set $21\mathbb{Z} + 3$ equals $\{3 + 21k \mid k = 0, \pm 1, \pm 2, \ldots\}$.) For general $d > 0$, instead of $d = 21$, we get

$$d\mathbb{Z}, \quad d\mathbb{Z} + 1, \quad d\mathbb{Z} + 2, \ldots, d\mathbb{Z} + (d-1)$$

The sets $d\mathbb{Z} + j$ are called *residue classes modulo d*.

If $x = qd + r$, $0 \le r < d$, then we denote this fact by x modulo $d = r$ or by $x \bmod d = r$. In this usage, "mod" is called a *binary operation*. Given any pair of integers x and $d > 0$, computing $x \bmod d$ always results in some integer r, $0 \le r < d$.

The word "mod" is also used to convey the information that "x and x' belong to the same residue class mod d." The notation is $x = x' \pmod{d}$ or $x \ne x' \pmod{d}$ to express the facts (respectively) that "x and x' belong to the same residue class mod d," or, "x and x' do not belong to the same residue class mod d." Often you will see $\equiv$ used instead of $=$ in these expressions.

Because of the possible confusion between these two uses, we will use the C programming language notation for the binary operation. Let's summarize all this in a definition.

Definition 2 (Residue classes and "mod") *Let $d \ge 2$ be an integer For $0 \le j < d$ the set $d\mathbb{Z} + j = \{nd + j \mid n \in \mathbb{Z}\}$ is called a residue class modulo d. The notation "mod" is used in two ways:*

- $x = x' \pmod{d}$ *This means that x and x' belong to the same reside class modulo d. In other words, when x and x' are divided by d they have the same remainder. We say that x and y are equal modulo d (or* mod *d). For reasons we will learn later, this is referred to as "using mod as an equivalence relation." The notation $x \equiv x' \pmod{d}$ is also used to indicate that x and y are equal modulo d. If the value of d is clear, people often write $x \equiv x'$, omitting* $(\bmod\ d)$.
- $x \bmod d = r$ *or* $x \,\%\, d = r$ *This means that when x is divided by d the remainder is r where $0 \le r < d$. Used this way, "mod" is a binary operator. To avoid confusion, we will use the C programming language notation $r = x \,\%\, d$.*

Since the two uses of "mod" involve different placement of "mod," you should not be confused as to which use is intended.

Example 6 (A fact about remainders) There is something important about remainders that they may not have discussed in elementary school. Suppose $x = qd+r$ and $x' = q'd+r'$. Then, subtracting and dividing by d gives

$$\frac{x-x'}{d} = \frac{(q-q')d+(r-r')}{d} = q-q' + \frac{r-r'}{d}\,.$$

Note that since $0 \le r < d$ and $0 \le r' < d$ we must have $0 \le |r-r'| < d$. This means that the only way that $\frac{r-r'}{d}$ can be an integer is that $|r-r'| = 0$ or $r = r'$. This seems like a trivial point, but it is very important. It means that x and x' have the same remainder when divided by d (i.e., belong to the same residue class mod d) if and only if d divides $x-x'$. For example 7666 and 7652 belong to the same residue class modulo 7 since $7666 - 7652 = 14$, which is 0 modulo 7. $\square$

The notation $x = x' \pmod{d}$ behaves like equality in many ways. The following theorem lists three of them.

Theorem 4 (Arithmetic with mod) *The notation $x = x'$ (mod d) behaves like equality for addition, subtraction and multiplication. In other words, if $x = x'$ (mod d) and $y = y'$ (mod d) then*

$$x+y = x'+y' \pmod{d}, \quad x-y = x'-y' \pmod{d} \quad \text{and} \quad xy = x'y' \pmod{d}.$$

We talk about *addition modulo d* or simply *modular addition*, and similarly for subtraction and multiplication. Notice that we did not say $x/y = x'/y' \bmod d$. It is not true in general. For example, $2 = 8 \pmod{6}$ and $2 = 2 \pmod{6}$ but $2/2 \ne 8/2 \pmod{6}$.

Proof: We prove addition. By definition $x+y = x'+y' \pmod{d}$ means that $(x+y) - (x'+y')$ is divisible by d. But

$$\frac{(x+y)-(x'+y')}{d} = \frac{(x-x')+(y-y')}{d} = \frac{x-x'}{d} + \frac{y-y'}{d}\,.$$

Since $x = x' \pmod d$ and $y = y' \pmod d$, both $x - x'$ and $y - y'$ are divisible by d. Thus, $(x + y) - (x' + y')$ is divisible by d.

The proof for subtraction is nearly the same as for addition, so we omit it.

We now prove multiplication. Again, we show that $xy - x'y'$ is divisible by d:

$$\frac{xy - x'y'}{d} = \frac{x(y - y') + y'(x - x')}{d} = x\frac{y - y'}{d} + y'\frac{x - x'}{d}\,.$$

Since, $x = x' \pmod d$ and $y = y' \pmod d$, both $x - x'$ and $y - y'$ are divisible by d. Thus, $xy - x'y'$ is divisible by d. $\square$

Example 7 (Powers of $d\mathbb{Z} + 1$) Suppose $x \in d\mathbb{Z} + 1$. We could equally well write this as $x \bmod d = 1$ or $x = 1 \pmod d$ or even just $x \equiv 1$ provided we know we are doing arithmetic modulo d. We claim that $x^n \equiv 1$ for all $n \in \mathbb{N}$. The proof is by induction on n.

For $n = 0$, $x^0 = 1$ and so $x^0 \equiv 1$. For $n = 1$, $x^1 = x$ and so $x^1 \equiv 1$ since we are given that $x \equiv 1$. For $n > 1$, $x^n = (x^{n-1})x$. By induction $x^{n-1} \equiv 1$. By the theorem, $x^{n-1}x \equiv 1 \times 1 = 1$. We are done.

When $d = 2$, you should be able to see that this simply states that powers of odd numbers are odd, a fact we proved in Example 1. $\square$

Example 8 (Using modular arithmetic cleverly) There are smart ways and dumb ways to use Theorem 4. It is interesting to look first at a dumb way, just to see the power of these statements. Suppose you want to find the remainder when the number $N = 113 \times (167 + 484) + 192 \times 145$ is divided by 21. That is, we wish to know $N \pmod{21}$. A friend says he is going to help. He tells you that $113 = 95180 \pmod{21}$, $167 = 5159244761 \pmod{21}$, $484 = 9073 \pmod{21}$, $192 = 207441 \pmod{21}$ and $145 = 19857871 \pmod{21}$. He suggests you substitute those larger numbers for the original numbers in the expression $N = 113 \times (167 + 484) + 192 \times 145$ to get

$$M = 95180 \times (5159244761 + 9073) + 207441 \times 19857871\,.$$

He assures you that, if you compute M and divide by 21 you will get the desired remainder r. He says he would like to borrow your car while you do the computations. After several hours work, you get $M = 495177116538231$. Dividing by 21 gives 15 as a remainder. Thus, $r = 15$, so $N \pmod{21} = 15$. That is the right answer but it is a dumb way to do it!

Another way is to just compute

$$N = 113 \times (167 + 484) + 192 \times 145 = 101403$$

and divide that by 21 to get the remainder 15. That is not too dumb.

Another way is to note that $113 = 8 \pmod{21}$, $167 = 20 \pmod{21}$, $484 = 1 \pmod{21}$, $192 = 3 \pmod{21}$, $145 = 19 \pmod{21}$. Substitute those for the corresponding numbers to get $L = 8(20 + 1) + 3 * 19 = 225$. Now divide 225 by 21 to get 15 as the remainder.

A modification on the above is to note that $20 = -1 \pmod{21}$ and $19 = -2 \pmod{21}$ to get $L' = 8(-1 + 1) + 3(-2) = -6$. Dividing -6 by 21 gives a remainder of 15. Did you learn that in elementary school? The remainder r must always be positive, $0 \le r < 21$. Thus, writing $-6 = q \times 21 + r$ gives $-6 = (-1) \times 21 + 15$. Do you see the power of these techniques? Don't be afraid to use them (wisely). Note that they apply to multiplying and adding, not dividing. For example, $484 = 1 \pmod{21}$, $22 = 1 \pmod{21}$, but $484/21 \pmod{21} \neq 1/1 \pmod{21}$. The number $484/21$ is not even an integer. $\square$

The Floor and Ceiling Functions

In computer science, many basic concepts are naturally expressed in terms of integer values (e.g., running time, input size, memory blocks) but are analyzed by functions that return real numbers. The conversion of the real numbers to integers that have direct meaning in terms of original problems sometimes involves the special functions "floor" and "ceiling."

Let $x \in \mathbb{R}$ be a real number. The *floor function* of x, denoted by $\lfloor x \rfloor$, is the largest integer less than or equal to x. It is also called the *greatest integer* function. The *ceiling function* of x, denoted by $\lceil x \rceil$, is the least integer greater than or equal to x. It is also called the *least integer* function.

Here are some examples:

$$\begin{gathered}
\lceil 2.8 \rceil = 3, \qquad \lceil 5 \rceil = 5, \qquad \lceil -2.8 \rceil = -2, \\
\lfloor 2.8 \rfloor = 2, \qquad \lfloor 5 \rfloor = 5, \qquad \lfloor -2.8 \rfloor = -3, \\
\lceil 55 + 2.8 \rceil = 55 + \lceil 2.8 \rceil = 55 + 3 = 58, \\
\lfloor -5.6 \rfloor = -6 = -\lceil -(-5.6) \rceil,
\end{gathered}$$

Geometrically, the idea is simple. The floor of x moves you to the next integer less than or equal to x on the number line. The ceiling moves you to the next integer greater than or equal to x. For computation, notice that

$$\begin{gathered}
\forall\, n \in \mathbb{Z},\ \forall\, x \in \mathbb{R},\ \lfloor n + x \rfloor = n + \lfloor x \rfloor. \\
\forall\, n \in \mathbb{Z},\ \forall\, x \in \mathbb{R},\ \lceil n + x \rceil = n + \lceil x \rceil.
\end{gathered}$$

This is easily shown and we omit the proof. Note also that

$$\lfloor x \rfloor = -\lceil -x \rceil \quad \text{and} \quad \lceil x \rceil = -\lfloor -x \rfloor.$$

For example, $\lfloor 2.1 \rfloor = -\lceil -2.1 \rceil$.

For proofs and exercises, it is often helpful to know that any real number can be written as the sum of an integer n and a fraction f, $-1 < f < +1$. Thus, $4.9 = 4 + 0.9$, $-3.6 = -3 - 0.6 = -4 + 0.4$. If $x = n + f$, then, since $\lfloor x \rfloor = n + \lfloor f \rfloor$ and $\lceil x \rceil = n + \lceil f \rceil$. you only have to think about the fractional part in your computations. For example,

$$\begin{gathered}
\lfloor 4.9 \rfloor = 4 + \lfloor 0.9 \rfloor = 4 + 0 = 4, \\
\lceil -3.6 \rceil = -4 + \lceil 0.4 \rceil = -4 + 1 = -3.
\end{gathered}$$

If you prefer, $\lceil -3.6 \rceil = -3 + \lceil -.6 \rceil = -3 + 0 = -3$.

Exercises for Section 1

1.1. Prove the statement if true, otherwise find a counterexample.

(a) For all natural numbers x and y, $x + y$ is odd if one of x and y even and the other is odd.

(b) For all natural numbers x and y, if $x + y$ is odd then one of x and y even and the other is odd.

1.2. Prove the statement if true, otherwise find a counterexample.

(a) The difference of any two odd integers is odd.

(b) If the sum of two integers is even, one of them must be even.

1.3. Prove the statement if true, otherwise find a counterexample.

(a) The product of two integers is even if and only if at least one of them is even.

(b) The product of two integers is odd if and only if at least one of them is odd.

1.4. Prove the statement if true, otherwise find a counterexample.

(a) For any integers m and n, $m^3 - n^3$ is even if and only if $m - n$ is even.

(b) For any integers m and n, $m^3 - n^3$ is odd if and only if $m - n$ is odd.

1.5. Prove the statement if true, otherwise find a counterexample.

(a) For all integers $n > 2$, $n^3 - 8$ is composite.

(b) For all integers n, $(-1)^n = -1$ if and only if n is odd.

1.6. Prove the statement if true, otherwise find a counterexample.

(a) $\forall\, n \in \mathbb{Z}$, $n^2 + n + 5$ is odd.

(b) $\forall\, n \in \mathbb{Z}$, $(6(n^2 + n + 1) - (5n^2 - 3)$ is a perfect square).

(c) $\exists\, M > 0, \forall\, n > M$, $(n^2 - n + 11$ is prime).

(d) There is a unique prime p of the form $n^2 + 2n - 3$.

1.7. Prove the statement if true, otherwise find a counterexample.

(a) For all integers $n > 0$, either n is a perfect square or, $n = x + y$ where x and y are perfect squares or, $n = x + y + z$ where x, y, and z perfect squares.

(b) The product of four consecutive positive integers is never a perfect square.

1.8. Prove the statement if true, otherwise find a counterexample.

(a) For all distinct positive integers m and n, either $m^{1/2} + n^{1/2}$ and $m^{1/2} - n^{1/2}$ are both rational or both irrational.
Hint: Consider $\left(m^{1/2} + n^{1/2}\right)\left(m^{1/2} - n^{1/2}\right)$.

(b) For all distinct positive integers, if either $m^{1/2}+n^{1/2}$ or $m^{1/2}-n^{1/2}$ are rational then both m and n are perfect squares.

(c) For all distinct positive integers m and n, both m and n are perfect squares if and only if $m + 2m^{1/2}n^{1/2} + n$ is a perfect square.

(d) Which of (a), (b) and (c) are true if $m \neq n$ is changed to $m = n$?

1.9. Prove that an integer $n > 1$ is composite if and only if p divides n for some prime $p \leq n^{1/2}$.

1.10. Write the following rational numbers as the ratio a/b of two integers a and $b > 0$.

(a) 3.1415

(b) 0.30303030...

(c) 6.32152152152152...

1.11. Let $x \in \mathbb{R}$ satisfy the equation $\frac{ax+b}{cx+d} = 1$ where a, b, c, and d are rational. Is x rational? Explain.

1.12. In each case, if the statement is true, prove it, if false, give a counterexample.

(a) The sum of three consecutive integers is zero (mod 3).

(b) The product of two even integers is zero (mod 4).

(c) An integer is divisible by 16 only if it is divisible by 8.

(d) For all odd integers n, $3n + 3$ is divisible by 6.

1.13. In each case, if the statement is true, prove it, if false, give a counterexample.

(a) $\forall\, a, b, c \in \mathbb{Z}$, if $a \mid b$ then $a \mid bc$.

(b) $\forall\, a, b, c \in \mathbb{Z}$, if $a \mid b$ and $b \mid c$, then $a \mid c$

(c) $\forall\, a, b, c \in \mathbb{Z}$, if $a \mid c$ then $ab \mid c$.

1.14. In each case, if the statement is true, prove it, if false, give a counterexample.

(a) $\forall\, a, b, c \in \mathbb{Z}$, if $a \mid (b + c)$ then $a \mid b$ and $a \mid c$.

(b) $\forall\, a, b, c \in \mathbb{Z}$, if $a \mid bc$ then $a \mid b$ or $a \mid c$.

(c) $\forall\, a, b \in \mathbb{Z}$, if $a \mid b$ then $a^2 \mid b^2$.

(d) $\forall\, a, b \in \mathbb{Z}$, if $a \mid 6\,b$ then $a \mid 6$ or $a \mid b$.

1.15. In each case, factor the given number into a product of powers of distinct primes.

(a) 1404. (b) 9702. (c) 89250.

1.16. Let $n = p_1^{e_1} \cdots p_k^{e_k}$ be the factorization of n into powers of distinct primes. Let $m \geq 1$ be an integer.

(a) What is the factorization of n^m into powers of distinct primes?

(b) If $s > 0$ is an integer but $s^{1/m}$ is not, must $s^{1/m}$ be irrational? Explain your answer.

1.17. In each case, factor the given number into a product of powers of distinct primes. Recall that $n! = n(n-1)(n-2)\cdots 1$ is the product of the first n integers.

(a) 20!. How many terminal zeros in this number?

(b) $(20!)^2$. How many terminal zeros in this number?

(c) $(20!)^3$. How many terminal zeros in this number?

1.18. Prove that if x is a nonzero natural number then $3 \mid x$ if and only if 3 divides the sum of the decimal digits of x.

1.19. Prove or give a counterexample: The product of any four consecutive integers is equal to 0 (mod 8).

1.20. Prove that, for all integers $n > 1$, $n^2 - 3 \neq 0 \pmod 4$.

1.21. Prove that, for all odd integers n, $n^4 = 1 \pmod{16}$.

1.22. If $m - n$ has remainder 0 when divided by d does that mean the m and n each have the same remainder when divided by d? Support your answer by giving a counterexample or a proof.

1.23. For all integers m, n, a, b, if $m \bmod d = a$ and $n \bmod d = b$ does that mean that $(m + n) \bmod d = a + b$?

1.24. (a) Prove: If $j = k \pmod d$, then $d\mathbb{Z} + j = d\mathbb{Z} + k$.

(b) Prove: If $j \neq k \pmod d$, then $(d\mathbb{Z} + j) \cap (d\mathbb{Z} + k)$ is the empty set.

1.25. If $a > 0$, $\log_a(x)$ is the unique number such that $a^{\log_a(x)} = x$.

(a) Suppose that p and q are two different primes. Prove that $\log_p(q)$ is irrational.

(b) Is the result in (a) true if p and q are allowed to be composite numbers? Justify your answer.

(c) For integers k and m, prove that $\log_a(b) = k/m$ if and only if $a^k = b^m$.

1.26. In each case, if the statement is true, prove it, if false, give a counterexample.

(a) $\forall\, x, y \in \mathbb{R}$, $(\lfloor x - y \rfloor = \lfloor x \rfloor - \lfloor y \rfloor)$.

(b) $\forall\, x \in \mathbb{R}$, $\forall\, k \in \mathbb{Z}$, $(\lfloor x - k \rfloor = \lfloor x \rfloor - k)$.

(c) $\forall\, x \in \mathbb{R}$, $k \in \mathbb{N}$, $(\lfloor x^k \rfloor = \lfloor x \rfloor^k)$.

1.27. In each case, if the statement is true, prove it, if false, give a counterexample.

(a) $\forall\, n \in \mathbb{Z}$, $k \in \mathbb{N}^+$, $(\lfloor \frac{n}{k} \rfloor = \frac{n-r}{k})$ where $r = n \,\%\, k)$.

(b) $\forall\, x \in \mathbb{R}$, $\forall\, a, b \in \mathbb{N}^+$, $(\lfloor ax + b \rfloor = a\lfloor x \rfloor + b)$.

1.28. Prove each of the following statements or give a counterexample.

(a) $\forall\, x \in \mathbb{R} - \mathbb{Z}$, $(\lfloor x \rfloor + \lfloor -x \rfloor = -1)$.

(b) $\forall\, x \in \mathbb{R} - \mathbb{Z}$, $(\lceil x \rceil + \lceil -x \rceil = +1)$.

Section 2: Cryptography and Secrecy

Cryptography is concerned with secret messages. Cryptanalysis is the name for the general area of breaking secret codes so the messages can be read. This general topic represents a vast body of knowledge. We begin by introducing the basic ideas and problems. Then we take time out to study some number theory functions that are useful for cryptography on the internet. Finally, we look at two protocols that are currently used — Diffie-Hellman and RSA.

Basic Ideas

Suppose that Alice wishes to send a message to Bob in such a way that anyone else receiving her message will not be able to understand it. She can communicate in code. There are three pieces of data involved:

- The *plaintext*, which is what Alice wants to tell Bob.
- The *ciphertext*, which is the message Alice actually sends Bob.
- The *key*, which tells how to convert plaintext to ciphertext and vice versa. Since the key is known to Alice and Bob, it is sometimes called the *shared key*.

The rules for converting can be thought of as functions. If $\mathcal{P}$ is the set of all possible plaintext messages and $\mathcal{C}$ is the set of all possible ciphertext messages, then the key K determines a function $f_K : \mathcal{P} \to \mathcal{C}$ that Alice uses to *encipher* the message. Bob uses the inverse function f_K^{-1} to *decipher* the message. Notice that, in order to decipher, f_K^{-1} must exist. Thus f_K must be an injection. The next example illustrates a simple scheme for doing this.

Example 9 (A simple code) Instead of Alice and Bob, we have two factories A and B that are going to exchange goods. There are 64 different items (coded $0, 1, 2, \ldots 63$) to be shipped and four methods of shipping (regular mail represented by the code 00; priority mail, code 01; air mail, code 10; and next day air, code 11). A shipment request looks something like 10101001. The two least significant bits, 01 in this case specify the method of shipping and the other six bits the item in base 2 (101010 or item 42 in this case).

The factories want to keep the orders they are requesting from each other secret from their competitors. To keep things secret, the factories agree on a simple encipherment procedure. They agree on a fixed eight bit binary string that they share as a secret. Here is the secret string that they happen to choose: $K = 11000111$. This is the *shared key*, also called the *secret key* or, simply, the key.

Factory A wants to place order $r = 10101001$ with factory B. To do this, the folks at A add r to K bit-by-bit using addition mod 2. That is, $0 + 0 = 0$, $0 + 1 = 1 + 0 = 1$, $1 + 1 = 0$. Here is what happens:

$$\begin{array}{ll} 10101001 & \text{plaintext} \\ 11000111 & \text{key } K \\ \hline 01101110 & \text{ciphertext} \end{array}$$

The first line is the message, the second line is the key, and the third line is the mod 2 bit-by-bit sum of the message and the key. We have just computed $f_K(10101001)$. Actually, this is done in the computer. When someone wants to place an order, they type in 10101001. The computer does the addition and sends the result to factory B.

When factory B's computer receives the ciphertext, it adds the shared key to the ciphertext as follows:

$$\begin{array}{ll} 01101110 & \text{ciphertext} \\ 11000111 & \text{key } K \\ \hline 10101001 & \text{plaintext} \end{array}$$

This reverses the process and reveals the correct order from factory A. Pretty nifty — the function and its inverse are the same, i.e. $f_K^{-1} = f_K$. □

In the previous example, $f_K^{-1} = f_K$. This makes programming easier since the software for deciphering is the same as the software for enciphering. As a result, many systems are designed to have $f_K^{-1} = f_K$.

There is a problem with our simple system (other than the fact that it's too simple): We can only send an 8-bit message.

- What if we want to send English instead of bits? This is no problem since computers store *everything* as bits. For example, text is stored using ASCII.
- What if we want to send longer messages? Well, we could break it into pieces that are 8-bits long and add the key to each 8-bit piece. For reasons we won't go into, using the same key K for each 8-bit piece is bad. Therefore there should be some rule for changing K. A simple rule is to replace the K for the current piece with $3K \bmod 2^8$ for the next piece.

Example 10 (Industrial espionage) Let's return to our factories that have been happily communicating secretly with each other.

Suppose Joe, who does industrial espionage for a competitor is able to intercept the ciphertext as it passes over the internet. He wants to know what orders are being placed; that is, he wants to find the plaintext. (He knows how to interpret the plaintext since lots of people at factories A and B know what it means.)

Joe manages to get an employee to place a fake order, say 11110000.

$$\begin{array}{ll} 11110000 & \text{plaintext} \\ \underline{11000111} & \underline{\text{key } K} \\ 00110111 & \text{ciphertext} \end{array}$$

Bob intercepts the ciphertext and adds it to the plaintext as follows:

$$\begin{array}{ll} 00110111 & \text{ciphertext} \\ \underline{11110000} & \underline{\text{plaintext}} \\ 11000111 & \text{key } K \end{array}$$

Now Joe has the key. Clever guy!

Except that the key and messages are much longer and the function f_K is not so simple, this sort of stuff goes on in the real world all of the time. For example, K might be anywhere from 64 to 128 bits, so there are anywhere from 2^{64} to 2^{128} possibilities for K.

You might ask why Joe didn't just get an employee to tell him key. The key is in the computer program. Only a few people, if any, know what it is. Well then, how did Joe know that f_K was plaintext plus key? In the real world, people use standard encryption algorithms (i.e., standard functions) that are public knowledge. When your computer browser is in secure mode, it is using a standard algorithm that Joe knows about. □

How can a company prevent Joe from getting their secrets this way? When we're thinking about this, we should imagine that the key is longer (64 to 128 bits) and that the plaintext is much longer. Here are some possibilities.

- Make it harder for Joe to get K.
  - We could improve employee loyalty. This may be difficult. A more reliable solution would be preferred.
  - We could invent an encryption system so that, even with plaintext and ciphertext, it is hard for Joe to compute K. Later, we'll discuss a way to do this.
- Change K frequently.
  - Sending out a new K may be feasible with two factories. It's much harder if there are a hundred — there are logistic and security problems. Why can't we simply encrypt the new K and send it out? Because, if Joe has the old K, he can read the message and get the new one.
  - When two computers want to communicate, have them decide on a K for that communication. This sounds impossible since Joe could eavesdrop. Later, we'll discuss a way to do this.
- Make Joe's knowledge of K useless.
  - We could invent an encryption system so that, even with K and ciphertext it is hard for Joe to compute plaintext without some additional (secret) information. Later, we'll discuss a way to do this.

The gcd, lcm and ϕ Functions

We now discuss some number theory functions that are important in cryptography. After we understand them, we'll use them in the Diffie-Hellman and RSA protocols.

Definition 3 (Greatest common divisor and least common multiple) *If k, n and n/k are integers, we write $k \mid n$ (read "k divides n") and we call k a* divisor *of n and we call n a* multiple *of k. The* greatest common divisor *of m and n is the largest (positive) k such that k is a divisor of m and k is a divisor of n. It is denoted by $\gcd(m,n)$. The* least common multiple *of m and n is the smallest positive integer k such that k is a multiple of m and k is a multiple of n. It is denoted by $\operatorname{lcm}(m,n)$.*

For example, if $m = 6$, its positive divisors are 1, 2, 3 and 6. Its positive multiples are 6, 12, 18, ... The greatest common divisor of 6 and 9 is 3, written $\gcd(6,9) = 3$. Similarly, $\operatorname{lcm}(6,9) = 18$.

The $\gcd(120,26) = 2$. It is also the case that $5 \times 120 - 23 \times 26 = 2$. In other words, there are integers $a = 5$ and $b = -23$ such that $am + bn = \gcd(m,n)$ where $m = 120$ and $n = 26$. This is a fact that is true for any m and n. That is, we claim

Theorem 5 (The gcd as a linear combination) *The greatest common divisor of m and n is a linear combination, with integral coefficients, of m and n.*

Corollary (All common divisors) *An integer k divides m and n if and only if it divides* $\gcd(m,n)$.

Proof: We can see why this must be true without knowing how to compute the coefficients a and b. The set $S = \{Am + Bn \mid A, B \in \mathbb{Z}, Am + Bn > 0\}$ is a nonempty set of positive integers (since $|m| \in S$) and therefore has a least element (by common sense at this point). Let $am + bn = L$ be this least element. Note that $L \mid m$. If not, we would have $m = qL + r$, $0 < r < L$. Thus,

$$r = m - qL = m - q(am + bn) = (1 - qa)m - (qb)n \in S.$$

This would contradict the minimality of L since $0 < r < L$. Similarly, $L \mid n$. Thus, L is a common divisor of m and n. Any integer x that is a common divisor of m and n divides any element $Am + Bn$ of S and thus $x \mid L$. Thus, $L = \gcd(m,n)$ is the greatest common divisor of m and n. This proves that $am + bn = \gcd(m,n)$.

In the last couple of sentences of the previous paragraph, we concluded that, if x divides both m n, then $x \mid \gcd(m,n)$. Conversely, suppose $x \mid \gcd(m,n)$. This means that x divides both m and n. This proves the corollary. $\square$

Example 11 (Some properties of gcd **and** lcm**)** Let $n > 0$ and $m > 0$ be positive integers and let $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$ and $m = p_1^{f_1} p_2^{f_2} \cdots p_k^{f_k}$ be factorizations of m and n into primes where some of the exponents f_i or e_i may be zero (in order to make k and the list of p_i the same for both factorizations). For example, $n = 6500 = 2^2 \times 5^3 \times 13$ and $m = 24696 = 2^3 \times 3^2 \times 7^3$ would, using this convention, be written as $n = 2^2 \times 3^0 \times 5^3 \times 7^0 \times 13^1$ and $m = 2^3 \times 3^2 \times 5^0 \times 7^3 \times 13^0$. The following theorem is the general result of which this example is a special case. We will not prove it. You should think carefully about the example and make up some of your own until you see why the theorem is true.

Theorem 6 (Computing gcd **and** lcm**)** *If* $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$ *and* $m = p_1^{f_1} p_2^{f_2} \cdots p_k^{f_k}$, *then*

$$\gcd(m,n) = p_1^{\min(e_1,f_1)} p_2^{\min(e_2,f_2)} \cdots p_k^{\min(e_k,f_k)}$$

and

$$\operatorname{lcm}(m,n) = p_1^{\max(e_1,f_1)} p_2^{\max(e_2,f_2)} \cdots p_k^{\max(e_k,f_k)}.$$

Applying this to

$$6500 = 2^2 \times 3^0 \times 5^3 \times 7^0 \times 13^1 \quad \text{and} \quad 24696 = 2^3 \times 3^2 \times 5^0 \times 7^3 \times 13^0$$

gives

$$\gcd(6500, 24696) = 2^2 \times 3^0 \times 5^0 \times 7^0 \times 13^0 = 4$$

and

$$\operatorname{lcm}(6500, 24696) = 2^3 \times 3^2 \times 5^3 \times 7^3 \times 13^1 = 40131000.$$

This is really pretty easy!

The theorem has various consequences.

- Every divisor $d = p_1^{d_1} p_2^{d_2} \cdots p_k^{d_k}$ of m and n has $d_i \le e_i$ and $d_i \le f_i$. Thus $d_i \le \min(e_i, f_i)$ and so d is also a divisor of $\gcd(m,n)$. That is, every common divisor of m and n is a divisor of $\gcd(m,n)$. (We also proved this in the process of proving Theorem 5.) Conversely, every divisor of $\gcd(m,n)$ is a common divisor of m and n.
- Similarly, every common multiple of m and n is a multiple of $\operatorname{lcm}(m,n)$. Conversely, every multiple of $\operatorname{lcm}(m,n)$ is a common multiple of m and n.
- $\gcd(m,n)\operatorname{lcm}(m,n) = mn$ because $\min(e_i,f_i) + \max(e_i,f_i) = e_i + f_i$ and so the p_i term in $\gcd(m,n)\operatorname{lcm}(m,n)$ is
$$p_i^{\min(e_i,f_i)} p_i^{\max(e_i,f_i)} = p_i^{\min(e_i,f_i)+\max(e_i,f_i)} = p_i^{e_i+f_i} = p_i^{e_i} p_i^{f_i}.$$
- If d is a common divisor of m and n, then $\gcd(m/d, n/d) = \gcd(m,n)/d$. In particular, when $d = \gcd(m,n)$, we have $\gcd(m/d, n/d) = 1$. We omit the proof. $\square$

The one thing you have to do to use the previous method for computing greatest common divisors and least common multiples is to factor n and m into primes. That can be difficult for big numbers. This method for computing gcd and lcm is more of theoretical or conceptual interest than of practical interest. Commonly available software for your computer will compute the gcd and the lcm quickly and efficiently for most integers that you may be interested in, without having to factor the integers. In the next example, we discuss the method that the software uses.

Example 12 (The Euclidean algorithm) Suppose we want to compute $\gcd(330, 156)$. Here's a "magical" procedure for doing it.

- We form a sequence that starts 330, 156.
- To get the next term in the sequence, divide 156 into 330 and keep the remainder: 330, 156, 18.
- To get the next term in the sequence, divide 18 into 156 and keep the remainder: 330, 156, 18, 12.
- To get the next term in the sequence, divide 12 into 18 and keep the remainder: 330, 156, 18, 12, 6.
- To get the next term in the sequence, divide 6 into 12 and keep the remainder: 330, 156, 18, 12, 6, 0.

Since we've reached zero, we stop and the term just before it (namely six) is the greatest common divisor. We could have started with 156, 330. Then we would have 156, 330, 156, 18, 12, 6, 0.

We need to formulate this in general and we need to prove that it works; that is, it isn't magic.

Here's the general procedure. Given two numbers $m > 0$ and $n > 0$, let $X_1 = m$ and $X_2 = n$. Define X_{k+1} to be the remainder when X_{k-1} is divided by X_k. Since X_{k+1} is a remainder, $X_{k+1} < X_k$. Thus we have $X_2 > X_3 > \cdots$. This eventually must reach zero, say $X_{t+1} = 0$. Then $\gcd(m, n) = X_t$. This is known as the *Euclidean algorithm*.

Why does it work? We claim that $\gcd(X_{k+1}, X_k) = \gcd(X_k, X_{k-1})$ for $k = 2, 3, \ldots, t$. Before proving this, let's see why it tells us that the algorithm works. We have

$$\gcd(m, n) = \gcd(X_1, X_2) = \gcd(X_2, X_3) = \cdots = \gcd(X_t, X_{t+1}) = \gcd(X_t, 0) = X_t,$$

where $\gcd(X_t, 0) = X_t$ since all numbers divide zero.

Now for the proof of the claim. Since X_{k+1} is the remainder after dividing X_{k-1} by X_k, it follows that $X_{k+1} = X_{k-1} - qX_k$ where q is the quotient when we divide X_{k-1} by X_k. Our claim states that

$$\gcd(X_{k-1} - qX_k,\ X_k) = \gcd(X_k, X_{k-1}).$$

More generally, we claim that $\gcd(a,\ b - ca) = \gcd(a, b)$ for any integers a, b, c. Suppose $d \mid a$ and $d \mid b$, then $a = Ad$ and $b = Bd$ for some integers A and B. Then

$$b - ca = Bd - cAd = (B - cA)d \quad \text{and so} \quad d \mid (b - ca).$$

Suppose $d \mid a$ and $d \mid (b - ca)$, then $a = Ad$ and $b - ca = Cd$ for some integers A and C. Then

$$b = (b - ca) + ca = Cd + cAd = (C + cA)d \quad \text{and so} \quad d \mid b.$$

We've now shown that d is a common divisor of a and b if and only if it is a common divisor of a and $b - ca$. This completes the proof. □

Example 13 (The Euclidean algorithm and Theorem 5) In Theorem 5 we showed that there are $a, b \in \mathbb{Z}$ so that $\gcd(m,n) = am + bn$, but we had no idea how to compute a and b. The Euclidean algorithm, with a slight modification, allows us to compute the a and b. Suppose we start with $m = X_1$ and $n = X_2$ and apply the Euclidean algorithm to get $X_t = d = \gcd(m,n)$:

$$X_1 > X_2 > X_3 > \cdots > X_t > X_{t+1} = 0.$$

Let $Q_2, Q_3, \ldots, Q_{t-1}$ be the list of quotients associated with the nonzero remainders in this list. Thus, $X_{i-1} = Q_i X_i + X_{i+1}$ for $i = 2, \ldots, t-1$. Note that $X_{t-2} = Q_{t-1}X_{t-1} + X_t$ so $\gcd(m,n) = X_t = X_{t-2} - Q_{t-1}X_{t-1}$. If $t = 3$ we would have $am + bn = \gcd(m,n)$ with $a = 1$, $b = -Q_{t-1}$, and our work would be done!

If $t > 3$, we can continue in the same way. We still have $X_t = X_{t-2} - Q_{t-1}X_{t-1}$. We also have $X_{t-1} = X_{t-3} - Q_{t-2}X_{t-2}$. If we substitute the second equation into the first, we get $X_t = \gcd(m,n)$ as a linear combination with integral coefficients of X_{t-3} and X_{t-2}. If $t = 4$, we are done. Otherwise, using $X_{t-2} = X_{t-4} - Q_{t-3}X_{t-3}$, we get X_t as a linear combination of X_{t-3} and X_{t-4}. Note that we are working our way towards getting $X_t = \gcd(m,n)$ as a linear combination with integral coefficients of X_1 and X_2. At this point we abandon the general discussion and move to an example.

Consider $X_1 = 60$ and $X_2 = 13$. Here is the list of nonzero remainders produced by the Euclidean algorithm:

$$60 > 13 > 8 > 5 > 3 > 2 > 1.$$

Thus, $t = 7$ and $\gcd(60, 13) = 1$. We kept track of the quotients: 4, 1, 1, 1, 1. To make it easier to see the connection between quotients and remainders we can write them in this way

$$\begin{array}{ccccccccccccc} 60 & > & 13 & > & 8 & > & 5 & > & 3 & > & 2 & > & 1 \\ & & 4 & & 1 & & 1 & & 1 & & 1 & & \end{array}$$

where we see that $60 = 4 \times 13 + 8$, $13 = 1 \times 8 + 5$, …, $3 = 1 \times 2 + 1$. Now we start working backwards. $1 = 3 - 1 \times 2$, $2 = 5 - 1 \times 3$, so $1 = 2 \times 3 - 1 \times 5$. Next we have $3 = 8 - 1 \times 5$, so

$$1 = 2(8 - 5) - 1 \times 5 = 2 \times 8 - 3 \times 5.$$

Next, $5 = 13 - 1 \times 8$, so $1 = 2 \times 8 - 3(13 - 8) = 5 \times 8 - 3 \times 13$. Finally, $8 = 60 - 4 \times 13$, so $1 = 5 \times 60 - 23 \times 13$. This is the final answer: $1 = \gcd(m,n) = am + bn$ where $m = 60$, $n = 13$, $a = 5$, and $b = -13$. You should make up some examples on your own and carry out this computation. $\square$

The positive integers $k = 1, 5, 7, 11$ are less than 12 and have no common factors with 12 (i.e., are *relatively prime to* 12). Another way to say this is $\gcd(k, 12) = 1$. The four numbers, 1, 5, 7, and 11 are the only numbers k with $\gcd(k, 12) = 1$ and $1 \le k \le 12$. For this reason, we say $\phi(12) = 4$. More generally:

Definition 4 (The Euler ϕ function) *We define a function $\phi(n)$, with domain the positive integers, to be the number of integers k, $1 \le k \le n$, such that $\gcd(k,n) = 1$. This function is called the Euler ϕ function.*

Example 14 (Properties of the Euler ϕ function) We have noted that $\phi(12) = 4$. Since $\gcd(1,1) = 1$, we have $\phi(1) = 1$. For any prime p, we have $\phi(p) = p - 1$ because $\gcd(k,p) = 1$ for $k = 1, 2, \ldots, p-1$.

Suppose $n = pq$ is the prime factorization of n and $p \neq q$. We can list the positive integers less than n that are *not* relatively prime to n. There are two classes of such numbers. The q multiples of p: $p, 2p, 3p, \ldots, qp$ and the p multiples of q: $q, 2q, 3q, \ldots, pq$. Except for $qp = pq$, these two lists have no numbers in common (why?). Thus, the total number of positive integers less than or equal to n that *are not* relatively prime to n is $q + p - 1$. Thus, the number of number less than or equal to $n = pq$ that *are* relatively prime to n is $pq - (p + q - 1) = (p-1)(q-1)$.

The set of numbers less than n that *are* relatively prime to n has a name. It is called the *group of units of* n and the numbers in that set are called *units*. The reason for this name is beyond the scope of our course, but does not involve difficult concepts. The Euler ϕ function and the group of units come into computer science in connection with computer security. It is the basis for a certain type of encryption known as *RSA* (discussed below) and is used in a common encryption protocol called PGP (Pretty Good Privacy). The key property that makes the group of units useful in this context is that $a^{\phi(n)} = 1 \pmod{n}$ whenever a is a unit. We won't prove this fact, but let's look at an example. Suppose $n = 12$. We know that $\phi(12) = 4$ and that the units are $\{1, 5, 7, 11\}$. Clearly $1^{\phi(12)} = 1 \pmod{12}$. What about the other units? We have $5^2 = 25 = 1 \pmod{12}$. Thus $5^4 = 1^2 = 1 \pmod{12}$. We could do the same calculations for 7 and 11. Here's another way. Since $7 = -5 \pmod{12}$, $7^4 = (-1)^4 5^4 = 5^4 = 1 \pmod{12}$. Likewise, $11 = -1 \pmod{12}$ and so $11^4 = (-1)^4 = 1 \pmod{12}$. You may have noticed that $a^2 = 1 \pmod{12}$ for all units a. There's no guarantee that $\phi(n)$ is the least power for which $a^{\phi(n)} = 1 \pmod{n}$ for all units a.

If $n = pq$ then, since $\phi(n) = (p-1)(q-1)$, this property becomes

$$m^{(p-1)(q-1)} = 1 \pmod{pq} \quad \text{when} \quad \gcd(m, pq) = 1.$$

This fact will be important in our discussion of the RSA protocol. $\square$

Cryptography on the Internet

Suppose two people — Alice and Bob — wish to communicate secretly, but anyone can eavesdrop on there conversation. How can they do this? We already saw in Example 9 how they could do this, and we saw how some problems could arise because of espionage. There's another problem we haven't mentioned. What if Alice and Bob don't have a secret key K that they both know?

Cryptography on the internet addresses this. It uses "public-information algorithms": No prior secret communication between Alice and Bob is needed — it's all done publicly. There are two approaches in use.

- Somehow Alice and Bob can develop a secret key even though someone is eavesdropping on their conversation. In this process, Alice and Bob usually play similar roles and so this is known as *symmetric encryption.*

- Alice can make known to the world data that allows people to encrypt messages to send to her but makes it hard for people other than Alice to decrypt them. Bob can do the same. Since this information (the key) is publicly known, this approach is called *public key cryptography.*

These approaches depend on what are called *trapdoor functions*. A trapdoor function is an invertible function g such that, given $g(x)$ it is hard to compute x. Such functions are also called *one-way functions*, but this is a bit misleading since it suggests that g is not invertible. We will discuss protocols that use two different trapdoor functions.

Example 15 (Discrete logs and better encryption) There are many ways to design a system such that, knowing the plaintext and ciphertext, it is still hard to recover the key. The method we describe here is not actually used, but it lays some of the groundwork for our next example.

If you use your calculator, you can easily compute $11^7 = 19487171$. If you know that 19487171 is of the form 11^x, for some x, you can equally well use your calculator to get x. From high school, you should remember that $x = \log_{11}(19487171)$. Probably, you would do that calculation using the LOG or LN button on your calculator as follows: LOG(19487171)/LOG(11) = 7. In any case, it is pretty easy. But, a seemingly innocent modification makes this sort of calculation very difficult in many cases.

If we compute $11^t \,\%\, 163$ for $t = 0, 2, \ldots, 161$, we get each of the numbers $1, 2, \ldots, 162$ exactly once — but they are in a mixed up order. Instead of 11^7, let's compute $11^7 \,\%\, 163$. The answer is 32. Now its not so easy to solve the equation $32 = 11^x \,\%\, 163$ for x even though we know there is a unique x between 0 and 161. For small numbers like these, it can be done by trying all $0 \le x < 162$. But, for big numbers with hundreds of digits, it seems to be all but impossible by any presently available methods. This problem of recovering an exponent from an exponentiated expression after it has been reduced modulo some number is called the *discrete logarithm problem* and the exponent is called the *discrete logarithm.*

Here is how we might use discrete logarithms to make it very hard for Joe's espionage when Alice and Bob have a secret key K. We choose a large modulus p that never changes. When someone wants to send a message P, the computer chooses a "base" b at random and computes $b^K \,\%\, p$. Call the result of this computation L.

The computer uses L to encrypt P by whatever method is being used for encryption. Thus, the computer obtains $f_L(P) = C$. It sends b and C. The computer at the other end computes $b^K \,\%\, p$ to obtain L and uses it to decrypt the message.

What can the spy Joe do? Suppose the encryption method is the one used in Example 10: We simply write L as a binary number and add it bitwise to the message P. Since the modulus p is fixed, we'll assume Joe knows what it is. As before, Joe gets his friend to send a message, so he has P, C and b for this particular message — call them P_1, C_1 and b_1. From P_1 and C_1, Joe recovers L_1. Later, someone else sends a message P_2. The computer chooses a random b_2, computes $b_2^K \,\%\, p = L_2$ and C_2. By eavesdropping Joe gets b_2 and C_2.

- To decrypt the message, Joe needs to find L_2 so that he can add it bitwise to C_2.
- To get L_2 he needs K because $L_2 = b_2^K \pmod{p}$ and he knows b_2.

- To get K he needs to solve the discrete log problem because he has b_1 and L_1 and $b_1^K = L_1 \pmod{p}$.

This is too hard, so Joe gives up.

There was nothing special about adding L bitwise to P. Whatever method was used, Joe would still want to recover K and so would need to carry out the steps in the previous paragraph. $\square$

Suppose the values of b and p are known and fixed. The function g, defined by $g(n) = b^n \,\%\, p$, is thought to be a trapdoor function. Finding n from $g(n)$ is referred to as computing the discrete log of b^n. As remarked in the previous example, computing the discrete log is believed to be very difficult. Thus g is believed to be a trapdoor function.

Suppose Alice and Bob want to communicate over the internet in secrecy, but have no shared key K. They must somehow construct K even though Joe can read their communications.

Example 16 (Diffie-Hellman: a symmetric key-exchange protocol) Here is how two computers can use the difficulty of the discrete log problem to generate a key K that they will share. Everyone agrees on a modulus p that is built into a program all computers can use. They also agree on a base b. Thus everyone, including the spy Joe, knows p and b. For purposes of illustration, we take $p = 163$ and $b = 11$. The values actually used on the internet are *much* bigger. We call the two computers that want to communicate A and B.

Computer A chooses, in secret, a random number s with $1 < s < p - 1$. Let us say 13 is chosen by A. Then A computes $b^s \,\%\, p = S$ and sends S to computer B. In our example, $S = 19$ since $11^{13} \,\%\, 163 = 19$. Meanwhile, B carries out the same process, choosing t and computing T, which it sends to A. Let us say B chooses $t = 23$. Thus B computes[5] $T = 11^{23} \,\%\, 163 = 116$.

Where are we now? Both computers and the spy Joe know that $S = 19$ and $T = 116$. Only computer A knows that $s = 13$ and only computer B knows that $t = 23$. In general, the public information is b, p, S and T; however, s and t are *not* public information since they were never sent over the internet.

What do the computers do now? Computer A uses its secret number s and computes $T^s \,\%\, p = K$. In our case, $116^{13} \,\%\, 163 = 154$, so $K = 154$. Likewise, B computes $S^t \,\%\, p = K$, which is $19^{23} \,\%\, 163 = 154$ in our case. That's amazing — A and B have the same number! Why is this? With all calculations modulo p, we have

$$T^s = (b^t)^s = b^{ts} = (b^s)^t = S^t \pmod{p}.$$

Where does this leave Joe? The obvious way for him to get key is to find either s or t since he already knows S and T. To find s, he needs to solve the discrete log problem $b^s = S \pmod{p}$. Likewise for T. Maybe there is a clever way for Joe to get K easily from b, p, S and T. At the present time, nobody knows of any such method, so Joe is stuck.

[5] The following computations and others like it can be done by using software packages such as GNU-MP, Maple® and Mathematica®. If you have to do it on a pocket calculator, it's best to do it in steps taking advantage of the properties of modular arithmetic.

The method of key exchange just discussed is called the Diffie-Hellman algorithm. It was discovered in 1976 and was the first public-information algorithm invented — invented *in public* that is! Apparently, the same algorithm, as well as other, later-to-be-discovered algorithms (such as RSA — Rivest, Shamir, Adleman, published by them in 1978), were discovered by British cryptanalysts working in secret in the Communications-Electronics Security Group in Britain during the early 1970's. Working in that group, Malcolm Williamson discovered the "Diffie-Hellman" algorithm in 1974. $\square$

Our next example is based on the difficulty of factoring. In this case, g is the function from pairs of primes $p < q$ to their product; that is, $g(p,q) = pq$. This is believed to be a trapdoor function when both p and q are large. To put this another way, all known methods of factoring take a long time. The protocol in this example is due to Rivest, Shamir and Adleman and so is called the RSA protocol.

***Example 17 (The RSA protocol)** This encryption system is based on the choice of some integer N that is a product of two primes. Suppose we take $N = 77$. We see easily that $N = pq$ where $p = 7$, $q = 11$. In real applications of this protocol p and q are primes with hundreds of digits, so given $N = pq$, it is very hard (or so it seems with present techniques) to factor N to get p and q. This is where the security of this method resides. Let's pretend that Alice makes known to the public her integer 77, and that Bob wants to send her a message. Suppose the spy Joe can't figure out how to factor 77. (In RSA this is true because much larger primes are used and multiplication is believed to be a trapdoor function.)

Alice is going to make known some more information. She picks two numbers e and d such that $ed = 1 \pmod{60}$. Why 60? Because $60 = (p-1)(q-1) = \phi(77)$. Suppose Alice picks $e = 13$ and $d = 37$. In this case $e\,d = 13 \times 37 = 481$. Check it out: $481 = 1 \pmod{60}$. She makes known to the public $e = 13$ and keeps $d = 37$ secret. Since Joe can't factor 77, he can't get the values $p = 7$ and $q = 11$. Hence Joe can't get the number $(p-1)(q-1) = 60$, and so he can't figure out that $d = 37$, given the publicly displayed number $e = 13$.

By the way, we didn't say how Alice chose the pair $e = 13$ and $d = 37$. Well, she just picked the e because it "seemed like a nice number." So that's her choice, as long as $\gcd(e, 60) = 1$. Clearly $\gcd(13, 60) = 1$, so she did all right there. To pick the $d = 37$ she used the method in Example 13 applied to $m = 13$, $n = 60$. You should reconstruct her calculations.

So now we have all that Alice is willing to tell the world: $N = 77$ and $e = 13$. In other words N and e are Alice's public information. The factorization $N = pq$ and the value of d are *not* public information because they were not sent over the internet.

In this encryption scheme, the only messages that Bob can send are numbers M, $1 \le M < 77 = N$, where $\gcd(M, N) = 1$. There are $\phi(77) = 60$ such messages. For example, Bob may decide to send the message $M = 5$. To send his message, he looks at Alice's public information (77 and 13) and sends $M^e \,\%\, 77 = 5^{13} \,\%\, 77$. You can easily check on your calculator that $5^{13} = 26 \pmod{77}$. In general, $M^e \,\%\, N$ is sent by Bob. Call it C.

So now Alice receives the message 26. Here is what she does to decrypt the message. She computes $26^{37} \,\%\, 77$ and gets 5. Recall that 37 was her secret number paired with 13.

This is the RSA protocol.

Suppose Joe intercepts C by eavesdropping. (In this case, the value was 26.) What can he do? If he knew $d = 37$, his life would be simple since he could do what Alice has done to decrypt the message. As far as is known, he'd have to be able to factor N in order to compute d — too hard! Could he do something else? Nobody knows of anything Joe could do that would not be hard.

Some of you might think that Joe had to solve the discrete log problem rather than the factoring problem since he saw $M^e \,\%\, N$. In the discrete log problem for $M^e \bmod N$, we know M and want to find e. Joe's problem is just the reverse — he knows e and wants to find M. This is believed to be a hard problem and is believed to be equivalent to factoring.

Why does Alice's decryption method work? In general, she is sent C, which is $M^e \,\%\, N$, and computes $C^d = (M^e)^d = M^{ed} \pmod{N}$. Recall, that $ed = 1 \pmod{\phi(N)}$. Thus $ed = 1 + k\phi(N)$ for some integer k. Hence

$$M^{ed} = M^{1+k\phi(N)} = M \times \left(M^{\phi(N)}\right)^k.$$

Since $\gcd(M, N) = 1$, M is a unit (see Example 14) and so, by the property at the end of Example 14, $M^{\phi(N)} = 1 \bmod N$. Thus

$$M^{ed} = M \times (1)^k = M \pmod{N}.$$

Since $1 \le M < N$, we have recovered M exactly, not just "mod N."

The RSA protocol is not a good scheme for amateurs to set up for themselves. The primes p and q must be hundreds of digits. The numbers e and d must be chosen with care. We must be sure that $\gcd(M, N) = 1$. Since $\phi(N)/N = (1 - \frac{1}{p})(1 - \frac{1}{q})$, it is close to 1 for large p and q. So the probability of selecting a bad message M (one that is not relatively prime to N) is quite small. $\square$

Exercises for Section 2

2.1. Use the Euclidean algorithm to find all common divisors of

(a) 1001 and 544 (b) 3510 and 652

2.2. Find all common divisors of 252 and 180 using the Euclidean algorithm.

2.3. How many common divisors are there of 59400 and 16200?

2.4. Using the Euclidean algorithm, find A and B such that $Am + Bn = \gcd(m.n)$ where $m = 252$ and $n = 180$.

2.5. Using the Euclidean algorithm, find A and B such that $Am+Bn = \gcd(m.n)$ where $m = 59400$ and $n = 16200$.

2.6. Using the Euclidean algorithm, find A and B such that $Am+Bn = \gcd(m.n)$ where $m = 163$ and $n = 86$.

2.7. Prove that $\gcd(a, b)$ divides $\mathrm{lcm}(a, b)$.

2.8. In each case find $\mathrm{lcm}(120, 108)$ (a) by prime factorization and (b) by the Euclidean algorithm.

2.9. Suppose a and b are positive integers. Prove directly from the definition of the least common multiple that $a \mid b$ if and only if $\mathrm{lcm}(a, b) = b$.

2.10. Following Example 16, suppose $p = 163$, $b = 11$. Computer A still chooses 13, but B chooses 15 instead of 23. What is the common key that results?

2.11. Suppose that, in Example 16, one of the computers chooses 1. Explain how the spy Joe can detect that and get their shared key.

***2.12.** Suppose that N is a prime in the RSA protocol of Example 17. How can the spy Joe find the message M if he has e, N and the encrypted message $M^e \,\%\, N = C$?

***2.13.** Using the same numbers as in Example 17, decrypt the message 2.

***2.14.** Consider the RSA protocol (Example 17). Suppose that $N = 5 \times 13$ and $e = 7$. What is d?

***2.15.** Consider the RSA protocol (Example 17). Explain why d and e must both be chosen to be odd.

Number Theory and Cryptography

Multiple Choice Questions for Review

In each case there is one correct answer (given at the end of the problem set). Try to work the problem first without looking at the answer. Understand both why the correct answer is correct and why the other answers are wrong.

1. "If $k > 1$ then $2^k - 1$ is not a perfect square." Which of the following is a correct proof?

(a) If $2^k - 1 = n^2$ then $2^{k-1} - 1 = (n-1)^2$ and $\frac{n^2+1}{(n-1)^2+1} = \frac{2^k}{2^{k-1}} = 2$. But this latter ratio is 2 if and only if $n = 1$ or $n = 3$. Thus, $2^k - 1 = n^2$ leads to a contradiction.

(b) If $2^k - 1 = n^2$ then $2^k = n^2 + 1$. Since 2 divides n^2, 2 does not divide $n^2 + 1$. This is a contradiction since obviously 2 divides 2^k.

(c) $2^k - 1$ is odd and an odd number which is a perfect square can't differ from a power of two by one.

(d) $2^k - 1$ is odd and an odd number can never be a perfect square.

(e) If $2^k - 1 = n^2$ then n is odd. If $n = 2j + 1$ then $2^k - 1 = (2j + 1)^2 = 4j^2 + 4j + 1$ which implies that 2^k, $k > 1$ is divisible by 2 but not by 4. This is a contradiction.

2. The repeating decimal number $3.14159265265265\ldots$ written as a ratio of two integers a/b is

(a) 313845111/99990000

(b) 313844841/99900000

(c) 313845006/99990000

(d) 313845106/99900000

(e) 313845123/99000000

3. Which of the following statements is true:

(a) A number is rational if and only if its square is rational.

(b) An integer n is odd if and only if $n^2 + 2n$ is odd.

(c) A number is irrational if and only if its square is irrational.

(d) A number n is odd if and only if $n(n + 1)$ is even

(e) At least one of two numbers x and y is irrational if and only if the product xy is irrational.

4. Which of the following statements is true:

(a) A number k divides the sum of three consecutive integers n, $n + 1$, and $n + 2$ if and only if it divides the middle integer $n + 1$.

(b) An integer n is divisible by 6 if and only if it is divisible by 3.

(c) For all integers a, b, and c, $a \mid bc$ if and only if $a \mid b$ and $a \mid c$.

(d) For all integers a, b, and c, $a \mid (b + c)$ if and only if $a \mid b$ and $a \mid c$.

(e) If r and s are integers, then $r \mid s$ if and only if $r^2 \mid s^2$.

5. For all $N \geq 0$, if $N = k(k+1)(k+2)$ is the product of three consecutive non-negative integers then for some integer $s > k$, N is divisible by a number of the form

(a) $s^2 - 1$

(b) $s^2 - 2$

(c) s^2

(d) $s^2 + 1$

(e) $s^2 + 2$

6. To one percent accuracy, the number of integers n in the list $0^4, 1^4, 2^4, \ldots, 1000^4$ such that $n \,\%\, 16 = 1$ is

(a) 20 percent

(b) 50 percent

(c) 30 percent

(d) 35 percent

(e) 25 percent

7. Which of the following statements is TRUE:

(a) For all odd integers n, $\lceil n/2 \rceil = \frac{n+1}{2}$.

(b) For all real numbers x and y, $\lceil x + y \rceil = \lceil x \rceil + \lceil y \rceil$.

(c) For all real numbers x, $\lceil x^2 \rceil = (\lceil x \rceil)^2$.

(d) For all real numbers x and y, $\lfloor x + y \rfloor = \lfloor x \rfloor + \lfloor y \rfloor$.

(e) For all real numbers x and y, $\lfloor xy \rfloor = \lfloor x \rfloor \lfloor y \rfloor$.

8. Which of the following statements is logically equivalent to the statement, "If a and $b \neq 0$ are rational numbers and $r \neq 0$ is an irrational number, then $a + br$ is irrational."

(a) If a and $b \neq 0$ are rational and $r \neq 0$ is real, then $a + br$ is rational only if r is irrational.

(b) If a and $b \neq 0$ are rational and $r \neq 0$ is real, then $a + br$ is irrational only if r is irrational.

(c) If a and $b \neq 0$ are rational and $r \neq 0$ is real, then r is rational only if $a + br$ is rational.

(d) If a and $b \neq 0$ are rational and $r \neq 0$ is real, then $a + br$ is rational only if r is rational.

(e) If a and $b \neq 0$ are rational and $r \neq 0$ is real, then $a + br$ is irrational only if r is rational.

9. The number of primes of the form $|n^2 - 6n + 5|$ where n is an integer is

(a) 0 (b) 1 (c) 2 (d) 3 (e) 4

Number Theory and Cryptography

10. The Euclidean Algorithm is used to produce a sequence $X_1 > X_2 > \cdots > X_{k-1} > X_k = 0$ of positive integers where each X_t, $2 < t \leq k$, is the remainder gotten by dividing X_{t-2} by X_{t-1}. If $X_{k-1} = 45$ then the set of all (positive) common divisors of X_1 and X_2 is

(a) $\{1, 3, 5\}$

(b) $\{1, 3, 5, 9, 15, \}$

(c) $\{1, 9, 15, 45\}$

(d) $\{1, 3, 5, 15\}$

(e) $\{1, 3, 5, 9, 15, 45\}$

11. Let L be the least common multiple of 175 and 105. Among all of the common divisors $x > 1$ of 175 and 105, let D be the smallest. Which is correct of the following:

(a) $D = 5$ and $L = 1050$

(b) $D = 5$ and $L = 35$

(c) $D = 7$ and $L = 525$

(d) $D = 5$ and $L = 525$

(e) $D = 7$ and $L = 1050$

12. The Euclidean Algorithm is used to produce a sequence $X_1 > X_2 > X_3 > X_4 > X_5 = 0$ of positive integers where $X_t = q_{t+1}X_{t+1} + X_{t+2}, t = 1, 2, 3$. The quotients are $q_2 = 3$, $q_3 = 2$, and $q_4 = 2$. Which of the following is correct?

(a) $\gcd(X_1, X_2) = -2X_1 + 6X_2$

(b) $\gcd(X_1, X_2) = -2X_1 - 6X_2$

(c) $\gcd(X_1, X_2) = -2X_1 - 7X_2$

(d) $\gcd(X_1, X_2) = 2X_1 + 7X_2$

(e) $\gcd(X_1, X_2) = -2X_1 + 7X_2$

Answers: **1** (e), **2** (d), **3** (b), **4** (e), **5** (a), **6** (b), **7** (a), **8** (d), **9** (c), **10** (e), **11** (d), **12** (e).