Number Theory

Muchang Bahng

Fall 2021

Contents

An introductory course in number theory. Much of the material introduced in this chapter can be found in other sections, especially those about Euclidean and Integral domains which are generalizations of the integers.

We begin by stating the well ordering principle of the natural numbers $\mathbb{N}$.

Theorem 0.1 (Well-Ordering Principle)

Every nonempty set S of nonnegative integers contains a least element. That is, there exists some integer $a \in S$ such that $a \leq b$ for all $b \in S$.

This leads to the following.

Theorem 0.2 (Archimedean Property)

If a, b are any positive integers, there exists a positive integer n such that $na \geq b$.

0.0.1 Induction

We provide three methods of proof.

Theorem 0.3 (Induction Principle)

Given $P(n)$, a property depending on a positive integer n,
1. if $P(n_0)$ is true for some positive integer n_0, and
2. if for every $k \geq n_0$, $P(k)$ true implies $P(k+1)$ true,

then $P(n)$ is true for all $n \geq n_0$.

Theorem 0.4 (Strong Induction Principle)

Given $P(n)$, a property depending on a positive integer n,
1. if $P(n_0), P(n_0+1), ..., P(n_0+m)$ are true for some positive integer n_0 and nonnegative integer m, and
2. if for every $k > n_0 + m$, $P(j)$ true for all $n_0 \leq j \leq k$ implies $P(k)$ true,

then $P(n)$ is true for all $n \geq n_0$.

Theorem 0.5 (Infinite Descent)

Given $P(n)$, a property depending on a positive integer n, assume that $P(n)$ is false for a set of integers $\mathcal{S}$. Let the smallest element of $\mathcal{S}$ be n_0. If $P(n_0)$ false implies $P(k)$ false, where $k < n_0$, then by contradiction, $P(n)$ is true for all n.

Note that the method of infinite descent is based off of the well ordering principle.

In some cases (especially in the Putnam exam), sometimes a creative use of induction will be requied. For example, you can first induct on a subset $\mathcal{S}$ of $\mathbb{N}$, then induct backwards (proving $P(n)$ true given $P(n+1)$ true), or use a double induction argument where you induct on two variables instead of one.

1 Divisibility Theory and Primes

A huge portion of number theory rests on the following theorem/algorithm.

Theorem 1.1 (Division Algorithm)

Given integers a, b with $b > 0$, there exist unique integers q, r satisfying

$$a = qb + r, \;\; 0 \leq r < b$$

The integers q and r are called the *quotient* and *remainder* in the division of a by b, respectively.

Proof. This statement can be quite obvious, but a rigorous proof requires the use of the well-ordering principle and proof by contradiction.

Definition 1.1

Let a and b be given integers, with at least one of them different from zero. The *greatest common divisor of a and b*, denoted by $\gcd(a, b)$, is the positive integer d satisfying

1. $d|a$ and $d|b$
2. If $c|a$ and $c|b$, then $c \leq d$

Note that 0 is divisible by every number.

Theorem 1.2

Given integers a, b not both of which are 0, there exist integers x and y such that

$$\gcd(a, b) = ax + by$$

Proof. Consider the set S of all positive linear combinations of a and b. Note that S is nonempty and is a subset of $\mathbb{N}$.

$$S \equiv \{au + bv \mid au + bv > 0, \; u, v \in \mathbb{Z}\}$$

From the well-ordering principle, S must contain a smallest element d. Thus, from the definition of S, there exist integers x and y for which $d = ax + by$. We claim that $d = \gcd(a, b)$.
Using the division algorithm, we can obtain integers q, r such that $a = qd + r$, where $0 \leq r < d$. Then, r can be written in the form

$$\begin{aligned} r = a - qd &= a - q(ax + by) \\ &= a(1 - qx) + b(-qy) \end{aligned}$$

If $r > 0$, then this representation of r above would simply mean that $r \in S$, contradicting the fact that d is the smallest element in S. So, $r = 0 \implies a = qd$, which implies $d|a$. By similar reasoning, $d|b$, which makes d a common divisor of a and b.
Now, if c is an arbitrary positive common divisor of the integers a and b, then $c|(ax + by)$; that is, $c|d$. Since $d \geq c$ for all c, $d = \gcd(a, b)$.

Corollary 1.3

If a and b are given nonzero integers, then the set

$$T \equiv \{ax + by \mid x, y \in \mathbb{Z}\}$$

is precisely the set of all multiples of $d = \gcd(a, b)$.

Definition 1.2

Two integers a and b, not both of which are zero, are said to be *relatively prime* whenever $\gcd(a, b) = 1$.

Theorem 1.4

Let a, b be nonzero integers. Then a and b are relatively prime if and only if there exist $x, y \in \mathbb{Z}$ such that

$$1 = ax + by$$

Proof. This is a direct result of the previous corollary.

This result directly leads to an observation that may be useful in some situations.

Corollary 1.5

If $\gcd(a, b) = d$, then $\gcd(a/d, b/d) = 1$.

Proof. Since it is possible to find integers x, y such that

$$d = ax + by$$

Upon dividing the Diophantine equation by d, we obtain

$$1 = \left(\frac{a}{d}\right)x + \left(\frac{b}{d}\right)y$$

where a/d and b/d are integers. Using the previous theorem, the two are relatively prime.

1.1 The Euclidean Algorithm

Here we introduce an algorithm that finds the greatest common divisors of two arbitrary integers. Without loss of generality, we can assume that $a, b > 0$ when finding

$$\gcd(a, b)$$

We will need to following lemma.

Lemma 1.6

If $a = qb + r$, then $\gcd(a, b) = \gcd(b, r)$.

Proof. If $d = \gcd(a, b)$, then the relations $d|a$ and $d|b$ together imply that $d|(a - qb)$, or $d|r$. Thus, d is a common divisor of both b and r. On the other hand, if c is an arbitrary common divisor of both b and r, then $c|(qb+r)$, whence $c|a \implies c$ is a common divisor of both a and b, so that $c \leq d$. So, $c \leq d$.

Using the result of this lemma, we can calculate

$$\begin{aligned} a &= q_1 b + r_1, & 0 < r_1 < b \\ b &= q_2 r_1 + r_2, & 0 < r_2 < r_1 \\ r_1 &= q_3 r_2 + r_3, & 0 < r_3 < r_2 \\ &..., & ... \\ r_{n-2} &= q_n r_{n-1} + r_n, & 0 < r_n < r_{n-1} \\ r_{n-1} &= q_{n+1} r_n + 0, \end{aligned}$$

and find that

$$\gcd(a,b) = \gcd(b, r_1) = ... = \gcd(r_{n-1}, r_n) = \gcd(r_n, 0) = r_n$$

Example 1.1

Let us calculate $\gcd(12378, 3054)$ using the Euclidean algorithm. The appropriate calculations produces the following:

$$\begin{aligned} 12378 &= 4 \cdot 3054 + 162 \\ 3054 &= 18 \cdot 162 + 138 \\ 162 &= 1 \cdot 138 + 24 \\ 24 &= 1 \cdot 18 + 6 \\ 18 &= 3 \cdot 6 + 0 \end{aligned}$$

Therefore, $\gcd(12378, 3054) = 6$. To represent 6 as a linear combination of the integers 12378 and 3054, we start with the second to last equation and substitute in remainders.

$$\begin{aligned} 6 &= 24 - 18 \\ &= 24 - (138 - 5 \cdot 24) \\ &= 6 \cdot 24 - 138 \\ &= 6(162 - 138) - 138 \\ &= 6 \cdot 162 - 7 \cdot 138 \\ &= 6 \cdot 162 - 7(3054 - 18 \cdot 162) \\ &= 132 \cdot 162 - 7 \cdot 3054 \\ &= 132(12378 - 4 \cdot 3054) - 7 \cdot 3054 \\ &= 132 \cdot 12378 + (-535) \cdot 3054 \end{aligned}$$

Thus, we have

$$\gcd(12378, 3054) = 6 = 12378x + 3054y$$

where $x = 132, y = -535$.

Theorem 1.7 (Lame)

The number of steps required in the Euclidean Algorithm is at most 5 times the number of digits in the smaller integer.

Theorem 1.8

For positive integers a, b,

$$\gcd(a,b)\,\text{lcm}(a,b) = ab$$

Proof. Let $d = \gcd(a, b)$. This allows us to express $a = dr$ and $b = ds$ for some $r, s \in \mathbb{N}$. If

$$m = \frac{ab}{d}$$

then $m = as = rb$, which makes m a positive common multiple of both a and b. Now, let c be a positive integer that is a common multiple of a and b, say, $c = au + bv$. Since there exist integers x, y satisfying $d = ax + by$, we get

$$\frac{c}{m} = \frac{cd}{ab} = \frac{c(ax + by)}{ab} = \frac{c}{b}x + \frac{c}{a}y = vx + uy$$

This means that $m|c$ and so $m \leq c$.

The significance of the previous theorem is that it makes the calculation of the least common multiple dependent on the greatest common divisor, which can be calculated using the Euclidean algorithm. For example,

$$\text{lcm}(3054, 12378) = \frac{3054 \cdot 12378}{6} = 6300402$$

Corollary 1.9

For any choice of positive integers a, b, $\text{lcm}(a, b) = ab$ if and only if $\gcd(a, b) = 1$.

1.2 The Diophantine Equation ax+by=c

It is customary to call a Diophantine equation any equation in one or more variables that is to be solved in the integers. The simplest type of Diophantine equation is

$$ax + by = c$$

Theorem 1.10

The linear Diophantine equation $ax + by = c$ has a solution if and only if $d|c$, where $d = \gcd(a, b)$. If x_0, y_0 is a particular solution to this equation, then the general solution can be paramaterized as

$$x = x_0 + \left(\frac{b}{d}\right)t, \; y = y_0 - \left(\frac{a}{d}\right)t, \quad t \in \mathbb{Z}$$

To find a particular solution, we apply Euclidean's algorithm to the coefficients a, b and work backwards to find a linear combination of a and b to get $\gcd(a, b)$. Then we multiply it according to the proper scalar to find the values of x, y.

Example 1.2

Consider the linear Diophantine equation.

$$172x + 20y = 1000$$

We apply Euclidean's algorithm to calculate $\gcd(172, 20)$.

$$\begin{aligned} 172 &= 8 \cdot 20 + 12 \\ 20 &= 1 \cdot 12 + 8 \\ 12 &= 1 \cdot 8 + 4 \\ 8 &= 2 \cdot 4 \end{aligned}$$

So, $\gcd(172, 20) = 4$. Since $4|1000$, a solution to this equation exists. Moreover, by working backwards, we have

$$\begin{aligned} 4 &= 12 - 8 \\ &= 12 - (20 - 12) \\ &= 2 \cdot 12 - 20 \\ &= 2(172 - 8 \cdot 20) - 20 \\ &= 2 \cdot 172 + (-17) \cdot 20 \end{aligned}$$

By multiplying both sides of $4 = 2 \cdot 172 + (-17) \cdot 20$ by 250, we get

$$1000 = 500 \cdot 172 + (-4250) \cdot 20$$

So, $x = 500, y = -4250$ is one solution to the equation. All other solutions are expressed by

$$\begin{aligned} x &= 500 + \frac{20}{4}t = 500 + 5t \\ y &= -4250 - \frac{172}{4}t = -4250 - 43t \end{aligned}$$

Corollary 1.11

If $\gcd(a, b) = 1$, and if x_0, y_0 is a particular solution of the linear Diophantine equation $ax + by = c$, then all solutions are given by

$$x = x_0 + bt, \;\; y = y_0 - at$$

Systems of linear equations can also be solved accordingly with a bit of modification.

Example 1.3

To solve the system

$$5x + 3y + \frac{1}{3}z = 100, \; x + y + z = 100$$

by eliminating one of the unknowns by substituting $z = 100 - x - y$, we are left with the equation

$$5x + 3y + \frac{1}{3}(100 - x - y) = 100 \implies 7x + 4y = 100$$

1.3 The Fundamental Theorem of Arithmetic

Definition 1.3

An integer $p > 1$ is called a *prime number* if its only positive divisors are 1 and p.

Theorem 1.12 (Fundamental Theorem of Arithmetic)

Every positive integer $n > 1$ can be expressed as a product of primes. This representation is unique up to the order in which the factors occur.

The process of putting a number into this form is called *prime factorization*.

Corollary 1.13

Any positive integer $n > 1$ can be written uniquely in a *canonical form*

$$n = \prod_{i=1}^{r} p_i^{k_i} = p_1^{k_1} p_2^{k_2} ... p_r^{k_r}$$

where, for $i = 1, 2, ..., r$, each k_i is a positive integer and each p_i is a prime, with $p_1 < p_2 < ... < p_r$.

We now present a method of identifying whether a certain number is prime or not.

Theorem 1.14 (Sieve of Eratosthenes)

If an integer $a > 1$ is not divisible by any prime $p \leq \sqrt{a}$, then a is prime.

Theorem 1.15 (Euclid)

There is an infinite number of primes.

Proof. Assume that there are a fininte number of primes $p_1, ..., p_n$. Consider the number

$$P = p_1 p_2 ... p_n + 1$$

Clearly, P is not divisible by any of the p_i's

We can actually put an upper bound on the nth (smallest) prime.

Theorem 1.16

If p_n is the nth prime number, then

$$p_n \leq 2^{2^{n-1}}$$

However, by 1854, a much better bound was formed.

Theorem 1.17

$$p_n \leq 2^n$$

Definition 1.4

A *repunit* is an integer written (in decimal notation) as a string of 1's, such as $11, 111, 1111, ...$ Let R_n denote the repunit with n digits. Every repunit is in the form

$$R_n = \frac{10^n - 1}{9}$$

The first seven repunit primes are

$$R_2, R_{19}, R_{23}, R_{317}, R_{1031}, R_{49081}, R_{86453}$$

1.4 The Goldbach Conjecture

We now introduce some progress on identifying some pattern in primes. We have already established our first claim: that there are an infinite number of primes. We can claim even further.

Theorem 1.18

The sum of the reciprocals of the primes diverges to infinity. That is, given the set of all primes $\mathbb{P} \subset \mathbb{N}$,

$$\sum_{p \in \mathbb{P}} p = \infty$$

Definition 1.5

A *twin prime* is a pair of primes (p, q) such that $q - p = 2$.

Theorem 1.19 (Twin Prime Conjecture)

There are an infinite number of twin primes.

Twin primes get much more scarce as numbers get bigger. The largest known twin prime in 2002 is

$$33219825 \cdot 2^{169690} \pm 1$$

with 51090 digits long.

Theorem 1.20 (Brun)

The sum of the reciprocals of the twin primes converges to a sum, known as *Brun's constant*. Brun's constant is approximately

$$1.902160583209 \pm 0.000000000781$$

based on all twin primes less than 2×10^{16}.

Theorem 1.21 (Zhang, 2014)

There are an infinite number of prime pairs differing by 246.

We now state one of the oldest and most well-known conjectures in number theory.

Theorem 1.22 (Goldbach Conjecture, 1742)

Every even positive integer greater than 2 is the sum of two prime numbers.

The numerical data supporting the Goldbach conjecture is overwhelming, and many mathematicians believe that it is true. We provide more claims about primes.

Theorem 1.23

There are an infinite number of primes in the form $4n + 3$.

Theorem 1.24 (Dirichlet)

If a and b are relatively prime positive integers, then the arithmetic progression

$$a, a+b, a+2b, a+3b$$

contains infinite many primes.

For example, this theorem tells us that there are an infinite number of primes ending in 999 since they appear in the arithmetic progression $1000n + 999$, where

$$\gcd(1000, 999) = 1$$

Theorem 1.25

There exists arbitrarily long but finite arithmetic progressions consisting only of prime numbers. The longest progression found to date is the 22 primes

$$11410337850553 + 4609098694200n, \quad 0 \leq n \leq 21$$

The prime factorization of the common difference between the terms is

$$2^3 \cdot 3 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19 \cdot 23 \cdot 1033$$

which is divisible by 9699690, the product of the primes less than 22. This leads to the following theorem.

Theorem 1.26

If all the $n > 2$ terms of the arithmetic progression

$$p, p+d, p+2d, ..., p+(n-1)d$$

2 The Theory of Congruences

Definition 2.1

Let n be a fixed positive integer. Two integers a and b are said to be *congruent modulo* n, denoted

$$a \equiv b \pmod{n}$$

if $n|(a-b)$; that is, if there exists an integer k such that $a - b = kn$.

The following is clearly a relation within the set of integers. That is,

1. $a \equiv a \pmod{n}$
2. $a \equiv b \pmod{n} \implies b \equiv a \pmod{n}$
3. $a \equiv b \pmod{n}, b \equiv c \pmod{n} \implies a \equiv c \pmod{n}$

This furthermore partitions the integers into *congruence classes*.

Theorem 2.1

For arbitrary integers a and b, $a \equiv b \pmod{n}$ if and only if a and b have the same nonnegative remainder when divided by n.

Proof. Trivial.

Since the integers are naturally endowed with the operations of addition and multiplication, we can conclude even further results about congruences.

Theorem 2.2

Let $n > 1$ be fixed and a, b, c, d be arbitrary integers. Then

1. $a \equiv b \pmod{n}, c \equiv d \pmod{n} \implies a + c \equiv b + d \pmod{n}$
2. $a \equiv b \pmod{n}, c \equiv d \pmod{n} \implies ac \equiv bd \pmod{n}$
3. $a \equiv b \pmod{n} \implies a^k \equiv b^k \pmod{n}$ for any positive integer k

All three can be combined to get the following. Let

$$P(x) = \sum_{k=0}^{m} c_k x^k$$

be a polynomial function of x with integral coefficients c_k. If $a \equiv b \pmod{n}$, then $P(a) \equiv P(b) \pmod{n}$.

However, note that congruences do not hold when integers are divided! Note the example

$$2 \equiv 8 \pmod{6} \not\implies 1 \equiv 4 \pmod{6}$$

The following theorem must be used.

Theorem 2.3

If $ca \equiv cb \pmod{n}$, then $a \equiv b \pmod{n/d}$, where $d = \gcd(c, n)$.

This states that if $\gcd(c, n) = 1$, then we can divide both sides by c without a change in modulus.

Corollary 2.4

If $ca \equiv cb \pmod{n}$ and $\gcd(c, n) = 1$, then $a \equiv b \pmod{n}$.

Corollary 2.5

If $ca \equiv cb \pmod{p}$ where p is a prime number, then $a \equiv b \pmod{p}$.

Definition 2.2

A number in the digit form

$$\overline{a_n a_{n-1} ... a_0}$$

in *base* m is calculated to be in the form

$$\overline{a_n a_{n-1} ... a_-} = \sum_{i=0}^{n} a_i m^i = a_0 + a_1 m^1 + a_2 m^2 + ... + a_n m^n$$

With this, we can prove requirements of divisibility of numbers by $3, 9$, and 11.

Theorem 2.6

Let $N = \overline{a_n a_{n-1} ... a_0}$ be the decimal (base 10) representation of a the positive integer N. Then
1. $3|N$ if and only if $3 \Big| \sum_{i=0}^n a_i$
2. $9|N$ if and only if $9 \Big| \sum_{i=0}^n a_i$
3. $11|N$ if and only if $11 \Big| \sum_{i=0}^n (-1)^i a_i$

Proof. We can see that

$$\begin{aligned} \overline{a_n a_{n-1} ... a_0} = \sum_{i=0}^n a_i 10^i &\equiv \sum_{i=0}^n a_i (1)^i \pmod 3 \\ &\equiv \sum_{i=0}^n a_i (1)^i \pmod 9 \\ &\equiv \sum_{i=0}^n a_i (-1)^i \pmod{11} \end{aligned}$$

2.1 Linear Congruences

Definition 2.3

An equation of linear congruence is of form

$$ax \equiv b \pmod n$$

where the solutions are equivalence classes of integers $[x]$. Two integers in the same equivalence class are counted as the same solution.

Theorem 2.7

The linear congruence $ax \equiv b \pmod n$ has a solution if and only if $d|b$, where $d = \gcd(d, n)$. If $d|b$, then it has d distinct solutions of equivalence classes.
Furthermore, if x_0 is a particular solution, then the $d = \gcd(a, n)$ incongruent solutions are

$$x_0, x_0 + \frac{n}{d}, x_0 + 2\Big(\frac{n}{d}\Big), ..., x_0 + (d-1)\Big(\frac{n}{d}\Big)$$

Corollary 2.8

If $\gcd(a, n) = 1$, then the linear congruence $ax \equiv b \pmod n$ has a unique solution modulo n.

Example 2.1

Consider the equation $18x \equiv 30 \pmod{42}$. Since $\gcd(18, 42) = 6$ and $6|30$, there are exactly 6 solutions that are incongruent modulo 42. One solution is $x = 4$, so the rest of them are

$$x \equiv 4 + \frac{42}{6} t \equiv 4 + 7t \pmod{42}, \quad t = 0, 1, 2, 3, 4, 5$$

which is the equivalence classes

$$x \equiv 4, 11, 18, 25, 32, 39 \pmod{42}$$

Theorem 2.9 (Chinese Remainder Theorem)

Let $n_1, n_2, ..., n_r$ be positive integers such that $\gcd(n_i, n_j) = 1$ for $i \neq j$. Then, the system of linear congruences

$$\begin{aligned} x &\equiv a_1 \pmod{n_1} \\ x &\equiv a_2 \pmod{n_2} \\ &\cdots \\ x &\equiv a_r \pmod{n_r} \end{aligned}$$

has a simultaneous solution, which is unique modulo the integer $n_1 n_2 ... n_r$.

Proof. Define the product $n = n_1 n_2 ... n_r$. For each $k = 1, 2, ..., r$, let

$$N_k = \frac{n}{n_k} = n_1 n_2 ... n_{k-1} n_{k+1} ... n_r$$

By hypothesis, all n_i are relatively prime, so, $\gcd(N_k, n_k) = 1$. According to the theory of a single linear congruence, it is therefore possibly to solve the congruence $N_k x \equiv 1 \pmod{n_k}$; denote the unique solution as x_k. We claim that the integer

$$\bar{x} = \sum_{i=1}^{r} a_i N_i x_i$$

is a simultaneous solution of the given system. Since $N_i \equiv 0 \pmod{n_k}$ for $i \neq k$, we have

$$\bar{x} = \sum_{i=1}^{r} a_i N_i x_i \equiv a_k N_k x_k \pmod{n_k}$$

But since the integer x_k was chosen to satisfy the congruence $N_k x \equiv 1 \pmod{n_k}$, this forces

$$\bar{x} \equiv a_k \pmod{n_k}$$

which shows that a solution exists. As for uniqueness, suppose that x' is any other integer satisfying the congruences. Then,

$$\bar{x} \equiv a_k \equiv x' \pmod{n_k}, \quad k = 1, 2, ..., r$$

and so $n_k | \bar{x} - x'$ for each k. Since $\gcd(n_i, n_j) = 1$, this implies that

$$\Big(\prod_{i=1}^{r} n_i \Big) \Big| (\bar{x} - x')$$

which implies that $\bar{x} \equiv x' \pmod{n}$.

Example 2.2

Let us solve the system

$$\begin{aligned} x &\equiv 2 \pmod 3 \\ x &\equiv 3 \pmod 5 \\ x &\equiv 2 \pmod 7 \end{aligned}$$

We have $n = 3 \cdot 5 \cdot 7$ and so

$$N_1 = 35, N_2 = 21, N_3 = 15$$

leading to the linear congruences

$$\begin{aligned} 35x &\equiv 1 \pmod 3 \\ 21x &\equiv 1 \pmod 5 \\ 15x &\equiv 1 \pmod 7 \end{aligned}$$

The solutions to these equations are $x_1 = 2, x_2 = 1, x_3 = 1$, respectively. Thus, a solution of the original system is given by

$$x = 2 \cdot 35 \cdot 2 + 3 \cdot 21 \cdot 1 + 2 \cdot 15 \cdot 1 = 233$$

Taking modulo 105, we get the unique solution $x = 233 \equiv 23 \pmod{105}$.

Definition 2.4

A linear congruence equation in two variables is of the form

$$ax + by \equiv c \pmod n$$

This congruence has a solution if and only if $\gcd(a, b, n) | c$.

We briefly describe the process of solving the equation when either one of a or b is relatively prime to n. Without loss of generality, let $\gcd(a, n) = 1$. Then, we can express the congruence as

$$ax \equiv c - by \pmod n$$

and for each of the n incongruent values of y, we are guaranteed a unique solution for x.

Example 2.3

Given the equation

$$7x + 4y \equiv 5 \pmod{12}$$

since $\gcd(7, 12) = 1$, we change the equation to

$$7x \equiv 5 - 4y \pmod{12}$$

Using casework by substituting each of the 12 possible incongruent values of y, we can reduce the above to a linear equation in one variable. For instance, leting $y \equiv 5 \pmod{12}$ produces the equation

$$7x \equiv -15 \pmod{12} \implies -5x \equiv -15 \implies x \equiv 3 \pmod{12}$$

Therefore, $(x, y) \equiv (3, 5)$ is one out of the 12 solutions.

We now shift towards solving systems of these equations.

Theorem 2.10

The system of linear congruences

$$\begin{aligned} ax + by &\equiv r \pmod{n} \\ cx + dy &\equiv s \pmod{n} \end{aligned}$$

has a unique solution modulo n whenever $\gcd(ad - bc, n) = 1$.

Proof. Let us multiply the first congruence of the system by d, the second congruence by b, and subtract the lower result form the upper. We then get

$$(ad - bc)x \equiv dr - bs \pmod{n}$$

Since by hypothesis, $\gcd(ad - bc, n) = 1$, this ensures that the congruence

$$(ad - bc)z \equiv 1 \pmod{n}$$

has a unique solution; call it t. When we multiply this to the first equation, we get

$$x \equiv t(dr - bs) \pmod{n}$$

Similarly, we can get a value for y:

$$y \equiv t(as - cr) \pmod{n}$$

Since we have described an explicit formula for the solutions x, y, we are done.

Notice that we can interpret this system as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \equiv \begin{pmatrix} r \\ s \end{pmatrix} \pmod{n}$$

For those with a bit of background in algebra, we can interpret the matrix of cofficients as a linear endomorphism of the quotient space of lattices $\mathbb{Z}^2/\sim$, where $\sim$ is the congruence relation.

Example 2.4

We use the formulas gotten in the previous proof to find the solutions of the system:

$$\begin{aligned} 7x + 3y &\equiv 10 \pmod{16} \\ 2x + 5y &\equiv 9 \pmod{16} \end{aligned}$$

Since $\gcd(7 \cdot 5 - 2 \cdot 3, 16) = \gcd(29, 16) = 1$, a solution exists. Multiplying the first congruence by 5, the second by 3, and subtracting the second from the first gives the equation

$$29x \equiv 23 \pmod{16} \implies 13x \equiv 7 \pmod{6}$$

producing the solution $x \equiv 3 \pmod{16}$. When we eliminate the x variable, we get the equation

$$29y \equiv 43 \pmod{16} \implies y \equiv 7 \pmod{16}$$

So, the unique solution to the system is

$$x \equiv 3 \pmod{16}, \quad y \equiv 7 \pmod{16}$$

2.2 Fermant's Little Theorem and Pseudoprimes

Theorem 2.11 (Fermant's Little Theorem)

Let p be a prime and suppose that $p \nmid a$. Then,

$$a^{p-1} \equiv 1 \pmod{p}$$

Proof. We consider the first $p-1$ positive multiples of a.

$$a, 2a, 3a, ..., (p-1)a$$

None of these numbers is congruent modulo p to any other, nor is any congruent to zero, since if any were, then

$$ra \equiv sa \pmod{p}, \;\; 1 \leq r < s \leq p-1$$

then a could be canceled to give $r \equiv s \pmod{p}$, which is impossible. Therefore, the previous set of integers must be congruent modulo p to $1, 2, 3, ..., p-1$, taken in some orer. Multiplying these congruences together gives

$$a^{p-1}(p-1)! \equiv (p-1)! \pmod{p}$$

Since $p \nmid (p-1)!$, we can divide both sides by $(p-1)!$ without changing the modulo to get

$$a^{p-1} \equiv 1 \pmod{p}$$

We can state this theorem in a slightly more general way by not requiring that p does not divide a.

Corollary 2.12

If p is prime, then $a^p \equiv a \pmod{p}$ for any integer a.

Ancient Chinese mathematicians conjectured that n is prime if and only if $n|(2^n - 2)$, which held true up to 340. However, $n = 341$ provides a counterexample to this claim, but numbers n that satisfy $n|(2^n - 2)$ are prime often enough to merit a name.

Definition 2.5

A composite integer n is called a *pseudoprime* if $n|(2^n - 2)$.

Theorem 2.13

If n is an odd pseudoprime, then

$$M_n = 2^n - 1$$

is a larger one.

Corollary 2.14

There are an infinite number of pseudoprimes.

Proof. The previous theorem allows us to construct an infinite sequence of increasing odd pseudoprimes.

The first four are $341, 561, 645$, and 1105.

Definition 2.6

More generally, a composite integer n for which

$$a^n \equiv a \pmod{n}$$

is called a *pseudoprime to the base a*. When $a = 2$, n is simply said to be a pseudoprime.

Theorem 2.15

There are infinitely many pseudoprimes to any given base.

Even though there are an infinite number of pseudoprimes, they are much rarer than regular primes. Indeed, there are only 245 pseudoprimes and $78,498$ primes smaller than $1,000,000$.

Definition 2.7

Composite numbers n that are pseudoprimes to every base a are called *absolute pseudoprimes*.

Lemma 2.16

561 is an absolute pseudoprime.

Proof. Note that $561 = 3 \cdot 11 \cdot 17$, and notice that $\gcd(a, 561) = 1$ gives

$$\gcd(a, 3) = \gcd(a, 11) = \gcd(a, 17) = 1$$

Using Fermant's little theorem, we get the congruences

$$a^2 \equiv 1 \pmod{3},\ a^{10} \equiv 1 \pmod{11},\ a^{16} \equiv 1 \pmod{17}$$

which implies

$$\begin{aligned} a^{560} \equiv (a^2)^{280} \equiv 1 \pmod{3} \\ a^{560} \equiv (a^{10})^{56} \equiv 1 \pmod{11} \\ a^{560} \equiv (a^{16})^{35} \equiv 1 \pmod{17} \end{aligned}$$

So, we have $a^{560} \equiv 1 \pmod{561}$, where $\gcd(a, 561) = 1$. So, $a^{561} \equiv a \pmod{561}$ for all a.

The next absolute pseudoprimes are

$$\begin{aligned} 1105 &= 5 \cdot 13 \cdot 17 \\ 2821 &= 7 \cdot 13 \cdot 31 \\ 15841 &= 7 \cdot 31 \cdot 73 \\ ... &= ... \\ 16046641 &= 13 \cdot 37 \cdot 73 \cdot 457 \end{aligned}$$

Now, we present a theorem that provides a means for producing absolute pseudoprimes.

Theorem 2.17

Let n be a composite square-free integer, say $p_1...p_n$, where the p_i are distinct primes. If

$$(p_i - 1)\big|(n-1) \text{ for } i = 1, 2, ..., r$$

then n is an absolute pseudoprime.

Proof. Suppose that a is an integer such that $\gcd(a, n) = 1$, so that $\gcd(a, p_i) = 1$ for all i. Then, Fermant's theorem yields

$$p_i\big|a^{p_i-1} - 1 \implies p_i\big|(a^n - a)$$

for all a and for all $i = 1, 2, ..., r$. So, we end up with $n\big|(a^n - a)$, making n an absolute pseudoprime.

There are 43 absolute pseudoprimes less than $1,000,000$ and $105,212$ less than 10^{15}.

Theorem 2.18 (Wilson's Theorem)

p is a prime number if and only if

$$(p-1)! \equiv -1 \pmod{p}$$

Proof. ($\rightarrow$) We can check by hand that the cases $p = 2$ and $p = 3$ are evident. Take $p > 3$. Suppose that a is any one of the $p - 1$ positive integers

$$1, 2, 3, ..., p-1$$

and consider the linear congruence $ax \equiv 1 \pmod{p}$. Since $\gcd(a, p) = 1$, there is a unique solution modulo p, call it a'. So, there is a unique integer a', with $1 \leq a' \leq p-1$ satisfying $aa' \equiv 1 \pmod{p}$. Now, note that because p is prime, $a = a'$ if and only if $a = 1$ or $a = p-1$, since this would lead to the congruence $a^2 \equiv 1 \pmod{p}$. If we omit the numbers 1 and $p-1$, we claim that the remaining $(p-3)/2$ numbers can be multiplied together to be congruent to 1. That is, we can group the remaining integers $2, 3, ..., p-2$ into pairs a, a' where $a \neq a'$, such that their product $aa' \equiv 1 \pmod{p}$. It is a fact that

$$2 \cdot 3 \cdot ... \cdot (p-2) \equiv 1 \pmod{p} \iff (p-2)! \equiv 1 \pmod{p}$$

We multiply by $p - 1$ to obtain the congruence

$$(p-1)! \equiv p - 1 \equiv -1 \pmod{p}$$

($\leftarrow$) The converse will not be proven here.

Example 2.5

Let us take $p = 13$. Then, we get

$$11! = (2 \cdot 7)(3 \cdot 9)(4 \cdot 10)(5 \cdot 8)(6 \cdot 11) \equiv 1 \cdot 1 \cdot 1 \cdot 1 \cdot 1 \equiv 1 \pmod{13}$$

which implies that

$$12! \equiv 12 \equiv -1 \pmod{13}$$

Definition 2.8

A *quadratic congruence* is a congruence of the form

$$ax^2 + bx + c \equiv 0 \pmod{n}, \quad a \neq 0 \pmod{n}$$

An application of Wilson's theorem goes into the following claim.

Theorem 2.19

The quadratic congruence $x^2 + 1 \equiv 0 \pmod{p}$, where p is an odd prime, has a solution if and only if $p \equiv 1 \pmod{4}$.

We finally end with a generalization of Fermant's theorem by stating Euler's theorem.

Theorem 2.20 (Euler's Theorem)

If $n \geq 1$ and $\gcd(a, n) = 1$, then

$$a^{\varphi(n)} \equiv 1 \pmod{n}$$

Fermant's theorem is then a corollary of Euler's theorem.

Corollary 2.21 (Fermant's Little Theorem)

If p is prime and p does not divide a, then $a^{p-1} \equiv 1 \pmod{p}$.

Proof. If p is prime, then $\varphi(p) = p - 1$. So,

$$a^{p-1} \equiv a^{\varphi(p)} \equiv 1 \pmod{p}$$

2.3 Fermant-Kraitchik Factorization Method

3 Number Theoretic Functions

3.1 Sum and Number of Divisors

Definition 3.1

A *number-theoretic* (or *arithmetic*) function is a function whose domain is the set of positive integers. That is, it is a function

$$F : \mathbb{Z} \longrightarrow X$$

for arbitrary X (not necessarily $\mathbb{Z}$).

Two of the most common arithmetic functions are defined below.

Definition 3.2

Given a positive integer n, let $\tau(n)$ denote the number of positive divisors of n and let $\sigma(n)$ denote the sum of these divisors.

We can also interpret τ and σ as

$$\sum_{d|n} f(d)$$

where the subscript on the summation denotes all divisors d of n and f is some function. For instance,

$$\sum_{d|20} f(d) = f(1) + f(2) + f(4) + f(5) + f(10) + f(20)$$

With this, τ and σ can be expressed in the form

$$\tau(n) = \sum_{d|n} 1, \quad \sigma(n) = \sum_{d|n} d$$

The following theorem provides a well known method to compute τ.

Theorem 3.1

Given a positive integer n, let its prime factorization be

$$n = \prod_i p_i^{k_i}$$

Then, the divisors of n are precisely those integers d of the form

$$d = \prod_i p_i^{a_i}, \;\; 0 \leq a_i \leq k_i \text{ for } i = 1, 2, ..., r$$

Corollary 3.2

If the prime factorization of n is $n = p_1^{k_1} p_2^{k_2} ... p_r^{k_r}$, then

$$\tau(n) = \prod_i (k_i + 1)$$

$$\sigma(n) = \prod_i \frac{p_i^{k_i+1} - 1}{p_i - 1}$$

Proof. The evaluation for $\tau(n)$ is trivial, since each divisor can be made by "choosing" from the $k_i + 1$ choices for the exponent a_i. To evaluate $\sigma(n)$, consider the product

$$\prod_i \Big(\sum_{j=0}^{k_i} p_i^j \Big) = \prod_i \left(1 + p_i + p_i^2 + ... + p_i^{k_i}\right)$$

and notice that each divisor of n appears once and only once as a term in the expansion of this product.

Theorem 3.3

The product of the positive divisors of a positive integer n is equal to $n^{\tau(n)/2}$. That is,

$$n^{\tau(n)} = \left(\prod_{d|n} d \right)^2$$

Note that given positive integer m, n,

$$\tau(mn) \neq \tau(m) \cdot \tau(n) \text{ and } \sigma(mn) = \sigma(m) \cdot \sigma(n)$$

That is, τ and σ are not multiplicative in general! However, there is a certain circumstance when they are multiplicative.

Definition 3.3

Within the context of number theory, a number theoretic function f is said to be *multiplicative* if

$$f(mn) = f(m)f(n)$$

whenever $\gcd(m, n) = 1$.

Theorem 3.4

τ and σ are multiplicative functions.

Proof. Since m and n are coprime, the prime factorization of m does not "overlap" that of n in such a way that none of the exponents are the same between m and n.

We can prove a more general results on multiplicative functions. T

Lemma 3.5

If $\gcd(m, n) = 1$, then the set of positive divisors mn consists of all products $d_1 d_2$, where $d_1 | m$ and $d_2 | n$, and $\gcd(d_1, d_2) = 1$. Furthermore, these products are all distinct.

Theorem 3.6

If f is a multiplicative function and F is defined by

$$F(n) = \sum_{d|n} f(d)$$

then F is also multiplicative.

Proof. Let m, n be coprime. By the previous lemma, every divisor of mn can be written as $d_1 d_2$. By definition of a multiplicative function, $f(d_1 d_2) = f(d_1)f(d_2)$, which implies

$$\begin{aligned} F(mn) &= \sum_{d_1|m, d_2|n} f(d_1)f(d_2) \\ &= \left(\sum_{d_1|m} f(d_1) \right) \left(\sum_{d_2|n} f(d_2) \right) \\ &= F(m)F(n) \end{aligned}$$

From this result, we can see that since the corresponding f's in the summation representation of τ and σ are multiplicative, the functions themselves are multiplicative.

3.2 The Mobius Inversion Formula

Definition 3.4

For a positive integer n, we define the *Mobius μ-function* as

$$\mu(n) = \begin{cases} 1 & n = 1 \\ 0 & p^2 | n \text{ for some prime } p \\ (-1)^r & n = p_1 p_2 ... p_r, \text{ where } p_i \text{ are distinct primes} \end{cases}$$

In words, this definition states that $\mu(n) = 0$ if n is not a square-free integer, whereas $\mu(n) = (-1)^r$ if n is square-free with r prime factors.

Example 3.1

Say $n = 30$. Then $\mu(30) = \mu(2 \cdot 3 \cdot 5) = (-1)^3 = -1$. The first few values of μ are

$$\mu(1) = 1,\ \mu(2) = -1,\ \mu(3) = -1,\ \mu(4) = 0,\ \mu(5) = -1,\ \mu(6) = (-1)^2 = 1$$

Lemma 3.7

μ is a multiplicative function. (Note that multiplicative only applies to arguments that are relatively prime)

What happens if we sum all of the divisors of n with μ applied to it?

Theorem 3.8

For each positive integer $n \geq 1$,

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & n = 1 \\ 0 & n > 1 \end{cases}$$

Example 3.2

$$\begin{aligned} \sum_{d|10} \mu(d) &= \mu(1) + \mu(2) + \mu(5) + \mu(10) \\ &= 1 + (-1) + (-1) + 1 = 0 \end{aligned}$$

The significance of the Mobius function is shown in the following theorem.

Theorem 3.9 (Mobius Inversion Formula)

Let F and f be two number theoretic functions related by the formula

$$F(n) = \sum_{d|n} f(d)$$

Then,

$$f(n) = \sum_{d|n} \mu(d)\, F\Big(\frac{n}{d}\Big) = \sum_{d|n} \mu\Big(\frac{n}{d}\Big)\, F(d)$$

Example 3.3

Let us use $n = 10$. We see that

$$\begin{aligned}\sum_{d|10}\Big(\sum_{c|(10/d)} \mu(d)\, f(c)\Big) &= \mu(1)\big(f(1)+f(2)+f(5)+f(10)\big)\\ &\quad + \mu(2)\big(f(1)+f(5)\big) + \mu(5)\big(f(1)+f(2)\big) + \mu(10)f(1)\\ &= f(1)\big(\mu(1)+\mu(2)+\mu(5)+\mu(10)\big)\\ &\quad + f(2)\big(\mu(1)+\mu(5)\big) + f(5)\big(\mu(1)+\mu(2)\big) + f(10)\mu(1)\\ &= \sum_{c|10}\Big(\sum_{d|10/c} f(c)\mu(d)\Big)\end{aligned}$$

Lemma 3.10

If F is a multiplicative function and

$$F(n) = \sum_{d|n} f(d)$$

then f is also multiplicative.

3.3 The Greatest Integer Function

Definition 3.5

For an arbitrary real number x, we denote as $[x]$, called the *floor function*, the largest integer less than or equal to x. That is, $[x]$ is the unique integer satisfying

$$x - 1 < [x] \leq x$$

Clearly, every real number x can be written as

$$x = [x] + \theta, \;\; 0 \leq \theta < 1$$

Given an integer n, we now introduce a method in finding the highest power k of p prime such that p^k divides $n!$.

Theorem 3.11

If n is a positive integer and p a prime, then the highest power k of p that divides $n!$ is

$$\sum_{k=1}^{\infty} \left[\frac{n}{p^k}\right]$$

where the series is infinite, because $[n/p^k] = 0$ for $p^k > n$.

Example 3.4

The greatest power of 2 that can divide 50! is

$$\begin{aligned}&[50/2] + [50/2^2] + [50/2^3] + [50/2^4] + [50/2^5] \\ &\quad = 25 + 12 + 6 + 3 + 1 \\ &\quad = 47\end{aligned}$$

So, 2^{47} divides 50!, but 2^{48} does not.

Lemma 3.12

If n and r are positive integers with $1 \leq r < n$, then the binomial coefficient

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!}$$

is also an integer.

Proof. We prove this using the floor function. Note that for any real numbers a, b, we have $[a+b] \geq [a] + [b]$. In particular, for each prime factor p of $r!(n-r)!$,

$$\left[\frac{n}{p^k}\right] \geq \left[\frac{r}{p^k}\right] + \left[\frac{n-r}{p^k}\right], \quad k = 1, 2, ...$$

Summing them over k, we get

$$\sum_{k\geq 1}\left[\frac{n}{p^k}\right] \geq \sum_{k\geq 1}\left[\frac{r}{p^k}\right] + \sum_{k\geq 1}\left[\frac{n-r}{p^k}\right]$$

The left hand side gives the exponent of the highest power of the prime p that divides $n!$, while the right hand side equals the highest power of this prime contained in $r!(n-r)!$. Hence, p appears in the numerator at least as many times in the denominator. Since this holds true for all p, $r!(n-r)!$ must divide $n!$, making the binomial coefficient an integer.

Corollary 3.13

For a positive integer r, the product of any r consecutive positive integers is divisible by $r!$.

Proof. The product of r consecutive integers, the largest of which is n, is

$$n(n-1)(n-2)...(n-r+1)$$

Now, we have

$$n(n-1)...(n-r+1) = \left(\frac{n!}{r!(n-r)!}\right)r!$$

Since $n!/r!(n-r)!$ is an integer, $r!$ must divide the product $n(n-1)...(n-r+1)$.

We incorporate the floor function into the topic of number theoretic functions.

Theorem 3.14

Let f and F be number theoretic functions such that

$$F(n) = \sum_{d|n} f(d)$$

Then, for any positive integer N,

$$\sum_{n=1}^{N} F(n) = \sum_{k=1}^{N} f(k) \left[\frac{N}{k} \right]$$

This allows us to compute τ and σ with the following corollaries.

Corollary 3.15

If N is a positive integer, then

$$\sum_{n=1}^{N} \tau(n) = \sum_{n=1}^{N} \left[\frac{N}{n} \right]$$

Corollary 3.16

If N is a positive integer, then

$$\sum_{n=1}^{N} \sigma(n) = \sum_{n=1}^{N} n \left[\frac{N}{n} \right]$$

Example 3.5

Consider the case when $N = 6$. Then,

$$\sum_{n=1}^{6} \tau(n) = \sum_{n=1}^{6} \left[\frac{6}{n} \right] = 6 + 3 + 2 + 1 + 1 + 1 = 14$$

We also have

$$\sum_{n=1}^{6} \sigma(n) = \sum_{n=1}^{6} n \left[\frac{6}{n} \right] = 1 \cdot 6 + 2 \cdot 3 + 3 \cdot 2 + 4 \cdot 1 + 5 \cdot 1 + 6 \cdot 1 = 33$$

3.4 Euler's Totient (Phi) Function

Definition 3.6

For $n \geq 1$, let $\varphi(n)$ denote the number of positive integers not exceeding n that are relatively prime to n.

For example, $\varphi(30) = 8$, since there are a total of 8 integers. Explicitly listing them out gives

$$1, 7, 11, 13, 15, 19, 23, 29$$

Clearly, there is an upper bound for φ. That is,

$$\varphi(n) \leq n - 1$$

with equality reaching if n is prime. That is, if we graph $\big(n, \varphi(n)\big)$, all points will be bounded in the lower triangular region of the first quadrant.

Theorem 3.17

Algebraically, $\varphi(n)$ gives the order for the multiplicative group of integer modulo n, which is isomorphic to the multiplicative group $\mathbb{Z}/n\mathbb{Z}$. That is,

$$\varphi(n) = \text{card}\left(\frac{\mathbb{Z}}{n\mathbb{Z}}\right)$$

Lemma 3.18

If p is prime and $k > 0$, then

$$\varphi(p^k) = p^k - p^{k-1} = p^k\left(1 - \frac{1}{p}\right)$$

Lemma 3.19

φ is a multiplicative function.

These two leads to the following theorem that describes a method to compute $\varphi(n)$.

Theorem 3.20

If the integer $n > 1$ has the prime factorization

$$n = p_1^{k_1} p_2^{k_2} ... p_r^{k_r}$$

then

$$\begin{aligned} \varphi(n) &= \big(p_1^{k_1} - p_1^{k_1-1}\big)\big(p_2^{k_2} - p_2^{k_2-1}\big)...\big(p_r^{k_r} - p_r^{k_r-1}\big) \\ &= n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)...\left(1 - \frac{1}{p_r}\right) \end{aligned}$$

Example 3.6

To calculate $\varphi(360)$, note that $360 = 2^3 \cdot 3^2 \cdot 5$, so

$$\varphi(360) = 360\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{5}\right) = 96$$

Notice that except for $\varphi(1)$ and $\varphi(2)$, the values of $\varphi(n)$ are always even.

Theorem 3.21

For $n > 2$, $\varphi(n)$ is an even integer.

Proof. In the case when n is a power of 2; that is, $n = 2^k$, then

$$\varphi(n) = \varphi(2^k) = 2^k\left(1 - \frac{1}{2}\right) = 2^{k-1}$$

If n is not a power of 2, then it is divisible by an odd prime p. So, we can write $n = p^k m$ for some $k \geq 1$ and m, where $\gcd(p^k, m) = 1$. Using the multiplicative property of φ, we get

$$\varphi(n) = \varphi(p^k)\varphi(m) = p^{k-1}(p-1)\varphi(m)$$

where $p - 1$ is even, so $\varphi(n)$ is also even.

One interesting property of the totient function is that the sum of the values of $\varphi(d)$ as d ranges over the positive divisors of n is equal to n itself.

Theorem 3.22 (Gauss)

For each positive integer $n \geq 1$,

$$n = \sum_{d|n} \varphi(d)$$

which is the sum being added over all positive divisors of n.

Proof. The integers between 1 and n can be separated into classes as follows. If d is a positive divisor of n, we put the integer m in the class S_d provided that $\gcd(m, n) = d$. That is,

$$S_d = \{m \mid \gcd(m, n) = d, 1 \leq m \leq n\}$$

Now, $\gcd(m, n) = d$ if and only if $\gcd(m/d, n/d) = 1$. Thus, the number of integers in the class S_d is equal to the number of positive integers not exceeding n/d that are relatively prime to n/d, which is just equal to $\varphi(n/d)$. Since each of the integers $1, 2, ..., n$ lies in exactly one class S_d, we get the formula

$$n = \sum_{d|n} \text{card}(S_d) = \sum_{d|n} \varphi\left(\frac{n}{d}\right)$$

But as d runs through all positive divisors of n, so does n/d, implying that

$$\sum_{d|n} \varphi\left(\frac{n}{d}\right) = \sum_{d|n} \varphi(d)$$

Example 3.7

Let $n = 10$. Then the classes S_d are

$$\begin{aligned} S_1 &= \{1, 3, 7, 9\} \\ S_2 &= \{2, 4, 6, 8\} \\ S_5 &= \{5\} \\ S_{10} &= \{10\} \end{aligned}$$

These contain $\varphi(10) = 4, \varphi(5) = 4, \varphi(2) = 1, \varphi(1) = 1$ integers, respectively. Therefore,

$$\begin{aligned}\sum_{d|10} \varphi(d) &= \varphi(10) + \varphi(5) + \varphi(2) + \varphi(1) \\ &= 4 + 4 + 1 + 1 = 10\end{aligned}$$

Theorem 3.23

For $n > 1$, the sum of the positive integers less than n and relatively prime to n is

$$\frac{1}{2} n\, \varphi(n)$$

Proof. Let $a_1, a_2, ..., a_{\varphi(n)}$ be the positive integers less than n and relatively prime to n. Because $\gcd(a, n) = 1$ if and only if $\gcd(n - a, n) = 1$, the numbers

$$n - a_1, n - a_2, ..., n - a_{\varphi(n)}$$

are equal in some order to $a_1, a_2, ..., a_{\varphi(n)}$. Thus,

$$\begin{aligned}a_1 + a_2 + ... + a_{\varphi(n)} &= (n - a_1) + (n - a_2) + ... + (n - a_{\varphi(n)}) \\ &= n\, \varphi(n) - (a_1 + a_2 + ... + a_{\varphi(n)}\end{aligned}$$

This implies that

$$2(\sum_{i=1}^{\varphi(n)} a_i = n\, \varphi(n)$$

Example 3.8

When $n = 30$, the $\varphi(30) = 8$ integers that are less than 30 and relatively prime to it are

$$1, 7, 11, 13, 17, 19, 23, 29$$

This is consistent with the theorem, since

$$1 + 7 + 11 + 13 + 17 + 19 + 23 + 29 = \frac{1}{2} \cdot 30 \cdot 8$$

Also, note the pairings:

$$1 + 29 = 30,\ 7 + 23 = 3-,\ 11 + 19 = 30,\ 13 + 17 = 30$$

This final theorem provides an application of the Mobius inversion formula.

Theorem 3.24

For any positive integer n,

$$\varphi(n) = n \sum_{d|n} \frac{\mu(d)}{d}$$

Proof. We apply the inversion formula to

$$F(n) = n = \sum_{d|n} \varphi(d)$$

to get

$$\begin{aligned} \varphi(n) &= \sum_{d|n} \mu(d)\, F\left(\frac{n}{d}\right) \\ &= \sum_{d|n} \mu(d)\, \frac{n}{d} \end{aligned}$$

4 Primitive Roots and Indices

With Euler's theorem, we know that $a^{\varphi(n)} \equiv 1 \pmod{n}$, whenever $\gcd(a, n) = 1$. However, there are often powers smaller than $a^{\varphi(n)}$ that are congruent to 1 modulo n.

Definition 4.1

Let $n > 1$ and $\gcd(a, n) = 1$. The *order of a modulo n* is the smallest positive integer k such that $a^k \equiv 1$.

Example 4.1

Consider the successive powers of 2 modulo 7.

$$2^1 \equiv 2, \;\; 2^2 \equiv 4, \; 2^3 \equiv 1, \; 2^4 \equiv 2, ...$$

So, the integer 2 has order 3 modulo 7.

Lemma 4.1

If two integers are congruent modulo n, then they have the same order modulo n. For if $a \equiv b \pmod{n}$ and $a^k \equiv 1 \pmod{n}$, then $a^k \equiv b^k \pmod{n}$, implying that $b^k \equiv 1 \pmod{n}$.

Also note that our definition of order modulo n concerns only integers a for which $\gcd(a, n) = 1$. Indeed, if $\gcd(a, n) > 1$, then we see that the linear congruence $ax \equiv 1 \pmod{n}$ has no solution, meaning that the relation $a^k \equiv 1 \pmod{n}$ cannot hold. With this in mind, one can deduce the following theorem.

Theorem 4.2

Let the integer a have order k modulo n. Then $a^h \equiv 1 \pmod{n}$ if and only if $k|h$; in particular, $k|\varphi(n)$.

Another basic result.

Theorem 4.3

If the integer a has order k modulo n, then $a^i \equiv a^j \pmod{n}$ if and only if $i \equiv j \pmod{k}$.

Corollary 4.4

If a has order k modulo n, then the integers $a, a^2, a^3, ..., a^k$ are incongruent modulo n.

Proof. If $a^i \equiv a^j \pmod{n}$ for $1 \leq i \leq j \leq k$, then the theorem ensures that $i \equiv j \pmod{k}$. But this is impossible unless $i = j$.

Theorem 4.5

If the integer a has order k modulo n and $h > 0$, then a^h has order $k/\gcd(h, k)$ modulo n.

Corollary 4.6

Let a have order k modulo n. Then a^h also has order k if and only if $\gcd(h, k) = 1$.

Example 4.2

2 has order 12 modulo 13. Calculations show that the orders of 2^2 and 2^3 are 6 and 4, respectively, which is consistent with the result that

$$6 = \frac{12}{\gcd(2, 12)}, \quad 4 = \frac{12}{\gcd(3, 12)}$$

Moreoever, the integers that also have order 12 modulo 13 are

$$2^1 \equiv 2, \ 2^5 \equiv 6, \ 2^7 \equiv 11, \ 2^{11} \equiv 7 \pmod{13}$$

Definition 4.2

If an integer a has the largest order possible, then we call it a *primitive root of* n. That is, if $\gcd(a, n) = 1$ and a is of order $\varphi(n)$ modulo n, then a is a *primitive root* of n.

Example 4.3

Listing out all the positive multiplies of 3, we can see that 3 is a primitive root of 7 since it has an order of $\varphi(7) = 6$.

$$3^1 \equiv 3, \ 3^2 \equiv 2, \ 3^3 \equiv 6, \ 3^4 \equiv 4, \ 3^5 \equiv 5, \ 3^6 \equiv 1$$

Another primitive root of 7 is 5, since it also has an order of $\varphi(7) = 6$

$$5^1 \equiv 5, \ 5^2 \equiv 4, \ 5^3 \equiv 6, \ 5^4 \equiv 2, \ 5^5 \equiv 3, \ 5^6 \equiv 1 \pmod{7}$$

However, no other primitive roots exist for 7. Try 4,

$$4^1 \equiv 4, \ 4^2 \equiv 2, \ 4^3 \equiv 1 \pmod{7}$$

which has an order of $3 \neq \varphi(7)$.

In fact, primitive roots exist for any prime modulus, since Euler's theorem combined with the fact that any number less than a prime is coprime with the prime itself. There are plenty of primitive roots for composite numbers, though.

Example 4.4

2 is a primitive root of 9. Note that $\varphi(9) = 6$

$$2^1 \equiv 2,\ 2^2 \equiv 4,\ 2^3 \equiv 8,\ 2^4 \equiv 7,\ 2^5 \equiv 5,\ 2^6 \equiv 1$$

However, it is more often the case that a number is not a primitive root.

Theorem 4.7

If the Fermant number $F_n = 2^{2^n} + 1$ with $n \geq 2$ is a prime, then 2 is a not a primitive root of F_n.

Proof. We factorize $F_{n+1} = 2^{2^{n+1}} + 1 = (2^{2^n} + 1)(2^{2^n} - 1)$, which implies that

$$2^{2^{n+1}} \equiv 1 \pmod{F_n}$$

This means that the order of 2 modulo F_n does not exceed 2^{n+1}. But if F_n is assumed to be prime, then

$$\varphi(F_n) = F_n - 1 = 2^{2^n}$$

but we can prove (by induction) that $2^{2^n} > 2^{n+1}$ whenever $n > 1$. Thus, the order of 2 modulo F_n is smaller than $\varphi(F_n)$ and by definition 2 cannot be a primitive root of F_n.

The following theorem is immensely useful.

Theorem 4.8

Let $\gcd(a, n) = 1$ and let $a_1, a_2, ..., a_{\varphi(n)}$ be the positive integers less than n and relatively prime to n. If a is a primitive root of n, then

$$a, a^2, ..., a^{\varphi(n)}$$

are congruent modulo n to $a_1, a_2, ..., a_{\varphi(n)}$ in some order.

Proof. Since a is relatively prime to n the same holds for all the powers of a, meaning that each a^k is congruent modulo n to some one of the a_i. But since the $\varphi(n)$ numbers in the set $\{a, a^2, ..., a^{\varphi(n)}\}$ are incongruent, these powers must represent some permutation of the integers $a_1, a_2, ..., a_{\varphi(n)}$.

Corollary 4.9

If n has a primitive root, then it has exactly $\varphi\big(\varphi(n)\big)$ of them.

Proof. Suppose that a is a primitive root of n. By the theorem, any other primitive root of n is found among the members of the set $\{a, a^2, ..., a^{\varphi(n)}\}$. But the number of powers $a^k, 1 \leq k \leq \varphi(n)$, that have order $\varphi(n)$ is equal to the number of integers k for which $\gcd\big(k, \varphi(n)\big) = 1$. There are $\varphi\big(\varphi(n)\big)$ such integers.

4.1 Primitive Roots for Primes

Theorem 4.10 (Lagrange)

If p is prime and

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0, \quad a_n \not\equiv 0 \pmod{p}$$

is a polynomial of degree $n \geq 1$ with integral coefficients, then the congruence

$$f(x) \equiv 0 \pmod{p}$$

has at most n incongruent solutions modulo p.

Corollary 4.11

If p is a prime number and $d|(p-1)$, then the congruence

$$x^d - 1 \equiv 0 \pmod{p}$$

has exactly d solutions.

Theorem 4.12

If p is a prime number and $d|(p-1)$, then there are exactly $\varphi(d)$ incongruent integers having order d modulo p.

Corollary 4.13

If p is prime, then there are exactly $\varphi(p-1)$ incongruent primitive roots of p.

Definition 4.3

Let $\chi(p)$ denote the smallest positive primitive root of the prime p.

The first few values of χ is

$$\begin{array}{cccccc} \chi(2)=1 & \chi(3)=2 & \chi(5)=2 & \chi(7)=3 & \chi(11)=2 & \chi(13)=2 \\ \chi(17)=3 & \chi(19)=2 & \chi(23)=5 & \chi(29)=2 & \chi(31)=3 & \chi(37)=2 \\ \chi(41)=6 & \chi(43)=3 & \chi(47)=5 & \chi(53)=2 & \chi(59)=2 & \chi(61)=2 \\ \chi(67)=2 & \chi(71)=7 & \chi(73)=5 & \chi(79)=3 & \chi(83)=2 & \chi(89)=3 \end{array}$$

The table suggests, although not proven, that there exist an infinite number of primes p for which $\chi(p) = 2$. Looking at the distribution of values more statistically, we can see that $\chi(p) \leq 19$ for all $p < 200$. Additionally, among the first 19862 odd primes up to 223051, $\chi(p) \leq 6$ holds for about 80% of these primes; $\chi(p) = 2$ about 37% of the time and $\chi(p) = 3$ about 23% of the time.

4.2 Primitive Roots for Composite Numbers

We state a few results.

Theorem 4.14

For $k \geq 3$, the integer 2^k has no primitive roots.

Proof. We start by showing that if a is an odd integer, then for $k \geq 3$

$$a^{2^{k-2}} \equiv 1 \pmod{2^k}$$

If $k = 3$, this congruence becomes $a^2 \equiv 1 \pmod 8$, which is true. For $k > 3$ we proceed by induction on k. Assume that the congruence holds for some integer k. Then

$$a^{2^{k-2}} \equiv 1 \pmod{2^k} \implies a^{2^{k-2}} = 1 + b2^k$$

where $b \in \mathbb{Z}$. Squaring both sides, we get

$$\begin{aligned} a^{2^{k-1}} = \left(a^{2k-2}\right)^2 &= 1 + 2(b2^k) + (b2^k)^2 \\ &= 1 + 2^{k+1}(b + b^2 2^{k-1}) \\ &\equiv 1 \pmod{2^{k+1}} \end{aligned}$$

meaning that the congruence holds for $n + 1$ and so for all $n > 3$. Now, the integers that are relatively prime to 2^k are precisely the odd integers, so $\varphi(2^k) = 2^{k-1}$, which is also equivalent to $2 \cdot 2^{k-2}$. So, if a is an odd integer and $k \geq 3$, then by the congruence just proved,

$$a^{\varphi(2^k)/2} \equiv 1 \pmod{2^k}$$

and consequently, there are no primitive roots of 2^k.

Theorem 4.15

If $\gcd(m, n) = 1$, where $m > 2, n > 2$, then the integer mn has no primitive roots.

Corollary 4.16

The integer n fails to have a primitive if either
1. n is divisible by two odd primes, or
2. n is of the form $2^m p_k$, where p is an odd prime and $m \geq 2$.

This allows us to reduce our search for primitive roots to the integers $2, 4, p^k$, and $2p^k$, where p is an odd prime. The following theorem says the rest.

Theorem 4.17

An integer $n > 1$ has a primitive root if and only if

$$n = 2, 4, p^k, \text{ or } 2p^k$$

where p is an odd prime.

4.3 The Theory of Indices

Definition 4.4

Let r be a primitive root of n. If $\gcd(a, n) = 1$, then the smallest positive integer k such that $a \equiv r^k \pmod{n}$ is called the *index of a relative to* r, denoted by $\text{ind}_r a$.

Clearly, $1 \leq \text{ind}_r a \leq \varphi(n)$, and

$$r^{\text{ind}_r a} \equiv a \pmod{n}$$

The notation $\text{ind}_r a$ is meaningless unless $\gcd(a, n) = 1$.

Example 4.5

The integer 2 is a primitive root of 5, and

$$2^1 \equiv 2 \quad 2^2 \equiv 4 \quad 2^3 \equiv 3 \quad 2^4 \equiv 1 \pmod{5}$$

If follows that

$$\text{ind}_2 1 = 4 \quad \text{ind}_2 2 = 1 \quad \text{ind}_2 3 = 3 \quad \text{ind}_2 4 = 2$$

Note that the way the index operation behaves is very similar to the logarithmic function.

Theorem 4.18

If n has a primitive root r and $\text{ind}_r a$ denote the index of a relative to r, then the following properties hold.

1. $\text{ind}_r(ab) \equiv \text{ind}_r a + \text{ind}_r b \pmod{\varphi(n)}$
2. $\text{ind}_r a^k \equiv k \, \text{ind}_r a \pmod{\varphi(n)}$ for $k > 0$
3. $\text{ind}_r 1 \equiv 0 \pmod{\varphi(n)}$, $\text{ind}_r \equiv 1 \pmod{\varphi(n)}$

The theory of indices can be used to solve certain types of congruences. For example, the binomial congruence

$$x^k \equiv a \pmod{n}, \quad k \geq 2$$

where n is a positive integer having a primitive root and $\gcd(a, n) = 1$ is entirely equivalent to the linear congruence

$$k \operatorname{ind} x \equiv \operatorname{ind} a \pmod{\varphi(n)}$$

Theorem 4.19

Let n be an integer possessing a primitive root and let $\gcd(a, n) = 1$. Then the congruence $x^k \equiv a \pmod{n}$ has a solution if and only if

$$a^{\varphi(n)/d} \equiv 1 \pmod{n}$$

where $d = \gcd\big(k, \varphi(n)\big)$. If it has a solution, then there are exactly d solutions modulo n.

Corollary 4.20

Let p be a prime and let $\gcd(a, p) = 1$. Then the congruence $x^k \equiv a \pmod{p}$ has a solution if and only if

$$a^{(p-1)/d} \equiv 1 \pmod{p}$$

where $d = \gcd(k, p-1)$.

5 Introduction to Cryptography

The practice of encrypting and decrypting messages is called cryptography. Codes are called *ciphers*, the information to be concealed is called *plaintext*, and after transformation to a secret form, a message is called *ciphertext*.

5.1 Common Cipher Methods

We now describe one of the most ancient and simplest of all encryption techniques, named after the Roman emperor Julius Caesar.

5.1.1 Caesar Cipher

Let us assign the English alphabet into digits from 00 to 25.

A	B	C	D	E	F	G	H	I	J	K	L	M
00	01	02	03	04	05	06	07	08	09	10	11	12
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
13	14	15	16	17	18	19	20	21	22	23	24	25

Then, if P is the digital equivalent of a plaintext letter and C is the digital equivalent of the corresponding ciphertext letter, then

$$C \equiv P + d \pmod{26}$$

where d is how much the alphabet "shifts."

The plaintext message CAESAR WAS GREAT can be digitized to

$$02 \quad 00 \quad 04 \quad 18 \quad 00 \quad 17 \quad 22 \quad 00 \quad 18 \quad 06 \quad 17 \quad 04 \quad 00 \quad 19$$

and using the congruence $C \equiv P + 3 \pmod{26}$, this becomes the ciphertext

$$05 \quad 03 \quad 07 \quad 21 \quad 03 \quad 20 \quad 25 \quad 03 \quad 21 \quad 09 \quad 20 \quad 07 \quad 03 \quad 22$$

which translates to FDHVDU ZDV JUHDW.

To recover the plaintext, the procedure is to simply reverse the means of the congruence

$$P \equiv C - 3 \equiv C + 23 \pmod{26}$$

This cipher is extermely simple and therefore, insecure. This is an example of a *monoalphabetic cipher*, an encryption scheme in which each letter of the original message is replaced by the same cipher substitute. Such cipher systems are extremely vulnerable to statistical methods of attack because they preserve the frequency (i.e. relative commonness) of individual letters.

5.1.2 Vigenere Cipher

One of the simplest and most famous example of a *polyalphabetic cipher* (a cipher that transformed a plaintext letter into more than one ciphertext equivalent) is the *Vigenere cipher*. In this case, the standard alphabet is digitized with number 00 to 25, and the communicating parties agree on an easily remembered word or phrase, called the keyword. The digitized version of the keyword is arranged below the numerical plaintext of the message and added together to produce the ciphertext.

Let the plaintext be ATTACK AT ONCE, with the keyword READY. The numerical version of READY is 17 04 00 03 24. We write the numerical plaintext on the top row and repeating sequences of the numerical version of READY below.

00	19	19	00	02	10	00	19	14	13	02	04
17	04	00	03	24	17	04	00	03	24	17	04

When the columns are added modulo 26, we get

17	23	19	03	00	01	04	19	17	11	19	08

or, converted to letters, RXTDAB ET RLTI.

Note that a given letter of plaintext is represented by different letters in ciphertext. The double T in the word ATTACK no longer appears as a double letter when ciphered.

In general, any sequence of n letters with numerical equivalents $b_1, b_2, ..., b_n$ $(00 \leq b_i \leq 25)$ can serve as the keyword. The plaintext message can be expressed as successive blocks $P_1P_2P_3...P_n$ of n two-digit integers P_i, and then converted to ciphertext blocks $C_1C_2...C_n$ by means of the congruences

$$C_i \equiv P_i + b_i \pmod{26}, \quad 1 \leq i \leq n$$

Decryption is carried out by simply reversing it.

$$P_i \equiv C_i - b_i \pmod{26}, \quad 1 \leq i \leq n$$

A weakness in the Vigenere algorithm is that once the length of the keyword has been determined, a coded message can be regarded as a number of separate monoalphabetic ciphers, each subject to straightforward frequency analysis. Then rather than using a single word that is to be repeated, people have used what is called a *running key*, which is a random assignment of ciphertext letters to plaintext letters. A popular procedure for generating such keys is to use the text of a book, and the system was thought to be secure until algorithms were generated that broke those codes.

However, a modification of using what is now called the *autokey* has made it more secure. This approach makes use of the plaintext message itself in constructing the encryption key. The idea is to start the keyword with a short *seed* or *prime* (generally a single letter) followed by the plaintext, whose ending is truncated by the length of the seed. Conveniently, this only requires the two communicating groups to remember the one letter key.

Assume that the message

ONE IF BY DAWN

is to be encrypted. Taking the letter K as the seed, the keyword becomes

KONEIFBYDAW

Now we can convert both to numerical form, obtaining the array

14	13	04	08	05	01	24	03	00	22	13
10	14	13	04	08	05	01	24	03	00	22

and adding them up modulo 26 gives

24	01	17	12	13	06	25	01	03	22	09

or changing back to letters,

YBR MN GZ BDWJ

We can decipher the message by first converting it to its numerical form. Suppose that the plaintext is $P_1P_2...P_n$ and the ciphertext is $C_1C_2...C_n$. If S indicates the seed, then the first letter of the plaintext is gotten with

$$P_1 = C_1 - S \pmod{26}$$

For the following letters, we use

$$P_k \equiv C_k - P_{k-1} \pmod{26}, \quad 2 \leq k \leq n$$

Doing this recovers

$$\begin{aligned} P_1 &\equiv 24 - 10 \equiv 14 \pmod{26} &&\implies P_1 = O \\ P_2 &\equiv 01 - 14 \equiv 13 \pmod{26} &&\implies P_2 = N \\ P_3 &\equiv 17 - 13 \equiv 04 \pmod{26} &&\implies P_3 = E \\ &... \end{aligned}$$

5.1.3 Hill's Cipher

An even better security system is to divide the plaintext message into blocks of n letters (possibly filling out hte last block by adding dummy letters such as Xs), and then encrypt block by block by using a system of n linear congruences in n variables. In its simplest form, when $n = 2$, the procedure takes two successive letters and transforms their numerical equivalents P_1P_2 into a block C_1C_2 of ciphertext numbers via the pair of congruences.

$$\begin{aligned} C_1 &\equiv aP_1 + bP_2 \pmod{26} \\ C_2 &\equiv cP_1 + dP_2 \pmod{26} \end{aligned}$$

In order to permit decipherment (that is, for the system to be solvable), the four coefficients a, b, c, d must be selected so that $\gcd(ad - bc, 26) = 1$.

For example, let us Hill encrypt the messages BUY NOW with blocks of 2 letters through the system

$$\begin{aligned} C_1 &\equiv 2P_1 + 3P_2 \pmod{26} \\ C_2 &\equiv 5P_1 + 8P_2 \pmod{26} \end{aligned}$$

The first block BU is numerically equivalent to 01 20, which is encrpyted by

$$\begin{aligned} 2(01) + 3(20) &\equiv 62 \equiv 10 \pmod{26} \\ 5(01) + 8(20) &\equiv 165 \equiv 09 \pmod{26} \end{aligned}$$

Doing this for the additional blocks YN and OW, we get the completed ciphertext

$$10 \quad 09 \quad 09 \quad 16 \quad 16 \quad 12$$

which can be expressed as KJJQQM. Deciphering the message requires solving the original system of congruences for P_1 and P_2 in terms of C_1 and C_2. After calculation, we get

$$\begin{aligned} P_1 &\equiv 8C_1 - 3C_2 \pmod{26} \\ P_2 &\equiv -5C_1 + 2C_2 \pmod{26} \end{aligned}$$

For the block 10 09 of ciphertext, we calculate

$$\begin{aligned} P_1 &\equiv 8(10) - 3(09) \equiv 53 \equiv 01 \pmod{26} \\ P_2 &\equiv -5(10) + 2(09) \equiv -32 \equiv 20 \pmod{26} \end{aligned}$$

Indeed, the block 01 20 represents BU. Doing this for the rest of the numbers returns the plaintext.

5.1.4 Verman Cipher

Another way of representing the letters of the alphabet is with binary numbers.

$$\begin{array}{lll} A = 11000 & J = 11010 & S = 10100 \\ B = 10011 & K = 11110 & T = 00001 \\ C = 01110 & L = 11110 & U = 11100 \\ D = 10010 & M = 00111 & V = 01111 \\ E = 10000 & N = 00110 & W = 11001 \\ F = 10110 & O = 00011 & X = 10111 \\ G = 01011 & P = 01101 & Y = 10101 \\ H = 00101 & Q = 11101 & Z = 10001 \\ I = 01100 & R = 01010 & \end{array}$$

For example, a plaintext message ACT NOW would be translated into a sequence of binary digits

$$11000011100000100110000111100 1$$

Then, both parties would have some type of encryption key of an arbitrary sequence of 0s and 1s with the same length as that of the numerical plaintext. For example, a random key can be generated as

$$10100101110010001000111100 1011$$

Then, by adding the key onto the numerical unencrypted message modulo 2, we get the encrypted message

$$01100110010010101110111111 0010$$

The security of this cipher is extremely high, especially if a new key is generated after every use (called a *one-time system*).

5.1.5 RSA Encryption

In conventional cryptographic systems, the sender and receiver jointly have a secret *key*. The sender uses the key to encrypt the plaintext to be sent, and the receiver uses the same key to decrypt the ciphertext obtained.

Public-key cryptography differs from conventional cryptography in that it uses two keys: encryption key and a decryption key. Although the two keys effect inverse operations and are therefore related, there is no easily computed method of deriving the decryption key from the encryption key. Thus, the encryption key can be made public without compromising the decryption key. That is, each user can encrypt messages, but only the intended recipient (whose decryption key is kept secret) can decipher them. A major advantage of a public-key cryptosystem is that it is unnecessary for senders and receivers to exchange a key in advance of their decision to communicate with each other.

In 1977, R. Rivest, A. Shamir, and L. Adleman proposed a public key system called *RSA*, named after their initials. Its security depends on the assumption that in the current state of computer technology, the factorization of composite numbers with large prime factors is prohibitively time-consuming.

Each user of the RSA system chooses a pair of distinct primes p and q, large enough that the factorization of their product $n = pq$, *called the enciphering modulus*, is beyond all current computational capabilities. For instance, picking p and q with 200 digits each would produce a number n with approximately 400 digits. Having selected n, the user then chooses a random positive integer k, called the *enciphering exponent*, satisfying

$$\gcd\big(k, \varphi(n)\big) = 1$$

The pair (n, k) (but not the factors p, q of n) is placed in a public file as the user's personal encryption key. This allows anyone else in the communication network to encrypt and send a message to that individual.

The encrpytion process begins with digitizing an alphabet. An example would be

$$\begin{array}{cccc} A=00 & K=10 & U=20 & 1=30 \\ B=01 & L=11 & V=21 & 2=31 \\ C=02 & M=12 & W=22 & 3=32 \\ D=03 & N=13 & X=23 & 4=33 \\ E=04 & O=14 & Y=24 & 5=34 \\ F=05 & P=15 & Z=25 & 6=35 \\ G=06 & Q=16 & ,=26 & 7=36 \\ H=07 & R=17 & .=27 & 8=37 \\ I=08 & S=18 & ?=28 & 9=38 \\ J=09 & T=19 & 0=29 & !=39 \end{array}$$

and 99 indicating a space between words. For example, the message

$$\text{The brown fox is quick}$$

is transformed into the numerical string

$$M = 1907049901171422139905142399081899162008021027$$

It is assumed that the plaintext number $M < n$, where n is, again, the enciphering modulus. Otherwise, it would be impossible the distinguish M from any larger integer congruent to it modulo n. When the message is too long to be handled as a single number $M < n$, then M is broken up into blocks of digits $M_1, M_2, ..., M_s$ of appropriate size, and each block is encrypted separately.

Looking up the intended recipient's encryption key (n, k) in the public directory, the sender disguises the plaintext number M as a ciphertext number r by raising M to the kth power and then reducing the result modulo n. That is,

$$M^k \equiv r \pmod{n}$$

From this step, it is obvious why $M < n$; it is wasn't, then it would be impossible to deduce M from r. This encrpytion method is very fast on high speed computers. Since k can be any integer such that $\gcd(k, \varphi(n)) = 1$, a obvious recommended choice of k is to be any prime larger than both p and q.

At the other end, the authorized recipient deciphers the transmitted information by first determining the integer j, the secret *recovery exponent*, for which

$$kj \equiv 1 \pmod{\varphi(n)}$$

Because $\gcd(k, \varphi(n)) = 1$, this linear congruence has a unique solution modulo $\varphi(n)$. In fact, the Euclidean algorithm produces j as a solution x to the equation

$$kx + \varphi(n)y = 1$$

The recovery exponent can only be calculated by someone who knows both k and $\varphi(n) = (p-1)(q-1)$ and hence, knows the prime factors p and q. So, j is secure from a third party. Now, by calculating r^j modulo n and assuming that $\gcd(n, M) = 1$ to use Euler's theorem, the recipient can see that

$$\begin{aligned} r^j \equiv (M^k)^j &\equiv M^{1+\varphi(n)t} \\ &\equiv M\big(M^{\varphi(n)}\big)^t \equiv M \cdot 1^t \equiv M \pmod{n} \end{aligned}$$

In other words, raising the ciphertext number to the jth power and reducing it modulo n recovers the original plaintext number M.

In the unlikely even that M and n are not coprime, we can actually prove that

$$r^j \equiv M \pmod{p} \text{ and } r^j \equiv M \pmod{q}$$

which yields the desired congruence $r^j \equiv M \pmod{n}$. Again, the major advantage to this encryption system is that it does not require the knowledge of the two primes p and q; it only requires the product n.

We work through an example with the RSA public-key algorithm. We first select two primes

$$p = 29, \quad q = 53$$

of an unrealistically small size for example purposes. In reality, p and q would be large enough to fill up a considerable portion of this page. Our enciphering modulus of $n = 29 \cdot 53 = 1537$, and $\varphi(n) = 28 \cdot 52 = 1456$. Since $\gcd(47, 1456) = 1$, we may choose $k = 47$ to be the enciphering exponent. Then, the recovery exponent, the unique integer j satisfying the congruence $kj \equiv 1 \pmod{\varphi(n)}$, is $j = 31$. The encrypt the message

$$\text{NO WAY} \implies M = 131499220024$$

Now, since $n = 1537$, we want each block to be an integer less than 1537. Given this restriction, it seems reasonable to split M into blocks of three digits each. The first block, 131 encrypts as the ciphertext number

$$131^{47} \equiv 0570 \pmod{1537}$$

At the other end, the authorized recipient, knowing that the recovery exponent is $j = 31$, begins to recover the plaintext number by computing

$$570^{31} \equiv 131 \pmod{1537}$$

The total ciphertext of our message is

$$0570\ 1222\ 0708\ 1341$$

The security of the RSA system rests on what is known as the *work factor*, the expected amount of computer time needed to factor the product of two large primes. Factoring is computationally more difficult than distinguishing between primes and composites, so at least up to current times, this system is secure. Even if computers get better, we can just choose larger primes.

In 1977, the three inventors of the system submitted a ciphertext message to *Scientific American* which depended on a 129-digit enciphering modulus that was the product of two primes of approximately the same length. The large number acquired the name RSA-129. Taking into account the most powerful factoring methods and fastest computers available at that time, it was estimate that at least 40 quadrillion years would be required to break down RSA-129, but with increasing computing power, it was broken after 17 years in 1994.

5.2 The Merkle-Hellman Knapsack Cryptosystem

The *Knapsack problem*, or the *subset sum problem*, in combinatorics is as follows: Given a knapsack of volume V and n items of various volumes $a_1, a_2, ..., a_n$, can a subset of these items be found that will completely fill the knapsack? Slightly modified, for positive integers $a_1, a_2, ..., a_n$ and a sum V, solve the equation

$$V = \sum_i a_i x_i$$

where $x_i \in \{0, 1\}$ for $i = 1, 2, ..., n$.

There may be no, one, or multiple solutions, but finding a solution to a randomly chosen knapsack problem is notoriously difficult. None of the known methods for attacking the problem are substantially less time-consuming than bashing through all 2^n possibilities for $x_1, x_2, ..., x_n$.

Example 5.1

The knapsack problem

$$22 = 3x_1 + 7x_2 + 9x_3 + 11x_4 + 20x_5$$

has no solution, but the problem

$$27 = 3x_1 + 7x_2 + 9x_3 + 11x_4 + 20x_5$$

has two distinct solutions

$$x_2 = x_3 = x_4 = 1, \;\; x_1 = x_5 = 0$$

and

$$x_2 = x_5 = 1, \;\; x_1 = x_3 = x_4 = 0$$

However, if the sequence of integers $a_1, a_2, ..., a_n$ happens to have some special properties, then the knapsack problem becomes much easier to solve.

Definition 5.1

A sequence $a_1, a_2, ..., a_n$ is *superincreasing* when each a_i is larger than the sum of all the preceding ones; that is,

$$a_i > \sum_{j=1}^{i} a_j, \;\; i = 2, 3, ..., n$$

A simple example of a knapsack problem with a superincreasing sequence is

$$V = x_1 + 2x_2 + 4x_3 + ... + 2^n x_n, \;\; V < 2^{n+1}$$

Knapsack problems with superincreasing sequences are uniquely solvable if they are solvable at all. The general algorithm goes as such: Suppose that we wish to solve the Knapsack problem

$$V = a_1x_1 + a_2x_2 + ... + a_nx_n$$

where $a_1, ..., a_n$ is superincreasing. Assume that V can be obtained by using some subset of the sequence so that V is not larger than the sum $a_1 + ...a_n$. Working from right to left in our sequence, we begin by letting $x_n = 1$. If $V \geq a_n$ and $x_n = 0$ if $V < a_n$. Then, obtain $x_{n-1}, x_{n-2}, ..., x_1$ in turn by choosing

$$x_i = \begin{cases} 1 & \text{if } V - (a_{i+1}x_{i+1} + ... + a_nx_n) \leq a_i \\ 0 & \text{if } V - (a_{i+1}x_{i+1} + ... + a_nx_n) < a_i \end{cases}$$

Example 5.2

We have the superincreasing knapsack problem

$$28 = 3x_1 + 5x_2 + 11x_3 + 20x_4 + 41x_5$$

We start with the largest coefficient 41. Since $41 > 28$, $x_5 = 0$. The next largest coefficient is 20, with $20 < 28$. The sum of the preceeding coefficients is $3 + 5 + 11 < 28$, so that these cannot fill the knapsack. Therefore 20 must be included in the sum and $x_4 = 1$. Knowing the values of x_4 and x_5, the problem is reduced to

$$8 = 3x_1 + 5x_2 + 11x_3$$

Since $11 > 8$, $x_3 = 0$, meaning that $x_1 = x_2 = 1$ to sum up to 8. Therefore, the solution is

$$x_1 = x_2 = x_4 = 1, \;\; x_3 = x_5 = 0$$

A public-key encryption system is based off of this knapsack problem. A typical user of the system starts by choosing a superincreasing sequence $a_1, a_2, ..., a_n$. He or she also selects a modulus $m > 2a_n$ and a multiplier

a, with $0 < a < m$ and $\gcd(a, m) = 1$. This ensures that the congruence

$$ax \equiv 1 \pmod{m}$$

has a unique solution, say $x \equiv c \pmod{m}$. Finally, we form the sequence of integers $b_1, b_2, ..., b_n$, defined by

$$b_i \equiv aa_i \pmod{m}, \quad i = 1, 2, ..., n$$

where $0 < b_i < m$. Carrying out this last transformation generally destroys the superincreasing property of the a_i's. The user keeps the original sequence $a_1, a_2, ..., a_n$ and the numbers m and a, but publishes $b_1, b_2, ..., b_n$ in a public directory. As the reader would expect, this sequence of b_i's serves as the encryption key.

We will use the following binary representation of the alphabet.

$A = 00000$	$J = 01001$	$S = 10010$
$B = 00001$	$K = 01010$	$T = 10011$
$C = 00010$	$L = 01011$	$U = 10100$
$D = 00011$	$M = 01100$	$V = 10101$
$E = 00100$	$N = 01101$	$W = 10110$
$F = 00101$	$O = 01110$	$X = 10111$
$G = 00110$	$P = 01111$	$Y = 11000$
$H = 00111$	$Q = 10000$	$Z = 11001$
$I = 01000$	$R = 10001$	

For example, the message First Place would be converted into the numerical representation

$$M = 00101\ 0100\ 10001\ 10010\ 10011\ 01111\ 01011\ 00000\ 00010\ 00100$$

The sender then splits this string into an arbitrary number of blocks of n binary digits (remember that n is the length of the sequences a_i and b_i), with the last block being filled out with 1s at the end if necessary. The public encrypting sequence $b_1, b_2, ..., b_n$ is used to transform the given plaintext block, say

$$x_1x_2x_3...x_n$$

into the sum

$$S = b_1x_1 + b_2x_2 + ... + b_nx_n$$

and the encryption is complete for that block. We do this for the rest of the blocks to encrypt the rest of the message. Now, since because each x_i is either 0 or 1, the problem of recreating the plaintext block from S is equivalent to solving the apparently difficult knapsack problem (remember that the new sequence $b_1, b_2, ..., b_n$ is not superincreasing anymore).

Once the authorized receiver receives this knapsack problem, he/she can change it into an easy one using the private key. Knowing c and m, the recipient can compute

$$S' \equiv cS \pmod{m}, \quad 0 \leq S' < m$$

and by expanding, we get

$$\begin{aligned} S' &\equiv cb_1x_1 + cb_2x_2 + ... + cb_nx_n \pmod{m} \\ &\equiv caa_1x_1 + caa_2x_2 + ... + caa_nx_n \pmod{m} \end{aligned}$$

Now, $ca \equiv 1 \pmod{m}$, so the previous congruence becomes

$$S' \equiv a_1x_1 + a_2x_2... + a_nx_n \pmod{m}$$

But due to the conditions that $m > 2a_n > a_1 + ... + a_n$ and that $0 \leq S' < m$, the congruence can be simplified to the equality

$$S' = a_1x_1 + a_2x_2 + ... + a_nx_n$$

Since S' and the superincreasing a_i's are given, the solution to this superincreasing knapsack problem can be easily computed, allowing us to recover the plaintext block $x_1x_2...x_n$ of n of the binary digits. Doing this for all the blocks entirely decrypts the message.

We provide an example with low-level sequences. Suppose that a typical user of this cryptosystem selects as a secret key the superincreasing sequences $3, 5, 11, 20, 41$, the modulus 85, and the multiplier $a = 44$. Each member of the superincreasing sequence is multiplied by 44 and reduced modulo 85 to yield

$$\begin{aligned} 44 \cdot 3 &\equiv 47 \pmod{85} \\ 44 \cdot 5 &\equiv 50 \pmod{85} \\ 44 \cdot 11 &\equiv 59 \pmod{85} \\ 44 \cdot 20 &\equiv 30 \pmod{85} \\ 44 \cdot 41 &\equiv 19 \pmod{85} \end{aligned}$$

These five numbers $47, 50, 59, 30, 19$ is submitted to the public directory. Someone who wants to send a plaintext message to the user, such as

HELP US

first converts it into the following binary digits.

$$M = 00111\ 00100\ 01011\ 01111\ 10100\ 10010$$

Then, since the length of the sequence is 5, the entire string is broken up into blocks of digits of length 5. Using the listed public key to encrypt, the sender transforms the successive blocks into

$$\begin{aligned} 108 &= 47 \cdot 0 + 50 \cdot 0 + 59 \cdot 1 + 30 \cdot 1 + 19 \cdot 1 \\ 59 &= 47 \cdot 0 + 50 \cdot 0 + 59 \cdot 1 + 30 \cdot 0 + 19 \cdot 0 \\ 99 &= 47 \cdot 0 + 50 \cdot 1 + 59 \cdot 0 + 30 \cdot 1 + 19 \cdot 1 \\ 158 &= 47 \cdot 0 + 50 \cdot 1 + 59 \cdot 1 + 30 \cdot 1 + 19 \cdot 1 \\ 106 &= 47 \cdot 1 + 50 \cdot 0 + 59 \cdot 1 + 30 \cdot 0 + 19 \cdot 0 \\ 77 &= 47 \cdot 1 + 50 \cdot 0 + 59 \cdot 0 + 30 \cdot 1 + 19 \cdot 0 \end{aligned}$$

Therefore, the transmitted ciphertext consists of the sequence of positive integers.

$$108 \quad 59 \quad 99 \quad 158 \quad 106 \quad 77$$

To read the message, the legitimate receiver first solves the congruence $44x \equiv 1 \pmod{85}$ to get the value of c, which is $x \equiv 29 \pmod{85}$. Then, each ciphertext number is multiplied by 29 and reduced modulo 85 to produce a superincreasing knapsack problem.

$$\begin{aligned} 29 \cdot 108 &\equiv 72 \pmod{85} \\ 29 \cdot 59 &\equiv 11 \pmod{85} \\ 29 \cdot 99 &\equiv 66 \pmod{85} \\ 29 \cdot 158 &\equiv 77 \pmod{85} \\ 29 \cdot 106 &\equiv 14 \pmod{85} \\ 29 \cdot 77 &\equiv 23 \pmod{85} \end{aligned}$$

which produces six corresponding knapsack problems with superincreasing sequences for each calculation. Each problem can be easily computed to get the corresponding solutions

$$\begin{aligned} 72 &= 3x_1 + 5x_2 + 11x_3 + 20x_4 + 41x_5 &&\implies (x_1, x_2, x_3, x_4, x_5) = (0, 0, 1, 1, 1) \\ 11 &= 3x_1 + 5x_2 + 11x_3 + 20x_4 + 41x_5 &&\implies (x_1, x_2, x_3, x_4, x_5) = (0, 0, 1, 0, 0) \\ 66 &= 3x_1 + 5x_2 + 11x_3 + 20x_4 + 41x_5 &&\implies (x_1, x_2, x_3, x_4, x_5) = (0, 1, 0, 1, 1) \\ 77 &= 3x_1 + 5x_2 + 11x_3 + 20x_4 + 41x_5 &&\implies (x_1, x_2, x_3, x_4, x_5) = (0, 1, 1, 1, 1) \\ 14 &= 3x_1 + 5x_2 + 11x_3 + 20x_4 + 41x_5 &&\implies (x_1, x_2, x_3, x_4, x_5) = (1, 0, 1, 0, 0) \\ 23 &= 3x_1 + 5x_2 + 11x_3 + 20x_4 + 41x_5 &&\implies (x_1, x_2, x_3, x_4, x_5) = (1, 0, 0, 1, 0) \end{aligned}$$

This cryptosystem aroused a great deal of interest because it was based on a provably difficult problem. However in 1982, Shamir invented a reasonably fast algorithm for solving a knapsack problem. The weakness of the system is that the public encryption key $b_1, b_2, ..., b_n$ is too special; that is, multiplying by a and reducing modulo m does not completely disguise the sequence $a_1, a_2, ..., a_n$. The system can be modified by iterating the modular multiplication method with different values of a and m so that the public and private sequences differ by several transformations, but even this was successfully broken by 1985. Although most variations of the Merkle-Hellman scheme have been shown to be insecure, there are a few that have resisted.

5.3 An Application of Primitive Roots to Cryptography

Most modern cryptography systems rely on the presumed difficulty of solving some particular number theoretic problem within a reasonable length of time.

5.3.1 ElGamal Encryption

In 1985, Taher ElGamal introduced a method of encrypting messages based on a version of the discrete logarithm problem, which is stated as follows: Find the integer $0 < x < \varphi(n)$, if it exists, that is the solution to the congruence

$$r^x \equiv y \pmod{n}$$

for given r, y, n. The exponent x is said to be the *discrete logarithm of y to the base r, modulo n*. By requiring that the base r be a primitive root of prime number n, it is guaranteed that y will always have a well-defined logarithm; that is, a solution x will always exist (by definition of the primitive root, and $x = \varphi(n) - 1$ when n is prime, at the very least). Note that merely requiring n to be prime guarantees that $x = \varphi(n)$ to be a solution by Euler's theorem, but there may exist no solutions that are less than $\varphi(n)$. The logarithm could be found by exhaustive search; that is, by calculating the successive powers of r until $y \equiv r^x \pmod{n}$ is reached However, this would not be practical for large n.

A typical user begins by selecting a prime number p along with one of its primitive roots r. Then an integer k with $2 \leq k \leq p-2$ is randomly chosen to serve as the secret key. Then, a is calculated as such.

$$a \equiv r^k \pmod{p}, \quad 0 \leq a \leq p-1$$

The triple of integers (p, r, a) becomes the person's public key, but the value of the exponent k is not revealed. It is also impractical for an unauthorized third party to calculate k since it would require them to solve a discrete logarithm problem that would be nearly impossible for large values of a and p.

Example 5.3

An individual begins by picking the prime $p = 113$ and its smallest primitive root $r = 3$. The choice $k = 37$ is then made for the integer satisfying $2 \leq 37 \leq 111$. Then $a \equiv 3^{37} \pmod{113}$ is calculated

$$a \equiv 3^{37} \equiv 3^1 \cdot 3^4 \cdot 3^{37} \equiv 3 \cdot 81 \cdot 28 \equiv 24 \pmod{113}$$

The triple $(113, 2, 24)$ serves as the public key, while the integer 37 becomes the secret deciphering key.

Now, assume that a message is to be sent to someone who has a public key (p, r, a) and also the corresponding private key k. We first convert the original message into a numerical equivalent with, say the standard convention that

$$A = 00 \quad B = 01 \quad ... \quad Z = 25$$

It is assumed that $M < p$. If $M \geq p$, then M is split into successive blocks, each block containing the same number of digits (which must be even since the numerical representation all have an even number of digits). Depending on how big the prime p is (which determines how big the blocks can get), it may be necessary to add extra digits (sometimes $25 = z$) to fill out the final block. Let B denote the first block. Then, the

sender, who is aware of the recipient's public key, arbitrarily selects an integer $2 \leq j \leq p-2$ and computes two values:

$$C_1 \equiv r^j \pmod{p}, \;\; C_2 \equiv Ba^j \pmod{p}, \; 0 \leq C_1, C_2 \leq p-1$$

The encrypted ciphertext of the block B is the pair of integers (C_1, C_2). For greater security, it is possible for the choice of j to be changed from block to block. The recipient of the ciphertext can then recover the block B by using the secret key k using the following identity. The recipient first evaluates $C_1^{p-1-k} \pmod{p}$ and then $P \equiv C_2 C_1^{p-1-k}$. Then the two values are multiplied together.

$$\begin{aligned} P \equiv C_2 C_1^{p-1-k} &\equiv (Ba^j)(r^j)^{p-1-k} \\ &\equiv B(r^k)^j (r^{j(p-1)-jk}) \\ &\equiv B(r^{p-1})^j \\ &\equiv B \pmod{p} \end{aligned}$$

where the final congruence results from the Fermant identity $r^{p-1} \equiv 1 \pmod{p}$. Therefore, the decryption can be carried out by someone who knows the value of k.

We work though an example with a reasonably small prime number for simplicity. Assume that the user wishes the deliver the message

SELL NOW

to a receiver who has the secret key $k = 15$ and public encryption key $(p, r, a) = (43, 3, 22)$, where $22 \equiv 3^{15} \pmod{43}$. The plaintext is first converted to the string of digits

$$M = 18\ 01\ 11\ 11\ 13\ 14\ 22$$

To create the ciphertext, the sender selects an integer j satisfying $2 \leq j \leq 41$, say $j = 23$, and then calculates

$$C_1 = r^j \equiv 3^{23} \equiv 34 \pmod{43} \text{ and } a^j \equiv 22^{23} \equiv 32 \pmod{43}$$

So, the product $C_1 B \equiv 32B \pmod{43}$ is computed for each two-digit block B of M. Doing this for all 7 blocks modulo 43.

$$\begin{array}{llll} 32 \cdot 18 \equiv 17 & 32 \cdot 04 \equiv 42 & 32 \cdot 11 \equiv 08 & 32 \cdot 11 \equiv 08 \\ 32 \cdot 13 \equiv 29 & 32 \cdot 14 \equiv 18 & 32 \cdot 22 \equiv 16 & \end{array}$$

We get the ciphertext

$$(34, 17)\ (34, 42)\ (34, 08)\ (34, 08)\ (34, 29)\ (34, 18)\ (34, 16)$$

The receiver, who knows that $k = 15$, decrypts it by first calculating

$$C_1^{p-1-k} \equiv 34^{27} \equiv 39 \pmod{43}$$

Then, this is multiplied modulo 43 to the second entry in the ciphertext pair.

$$\begin{array}{llll} 39 \cdot 17 \equiv 18 & 39 \cdot 42 \equiv 04 & 39 \cdot 08 \equiv 11 & 39 \cdot 08 \equiv 11 \\ 39 \cdot 29 \equiv 13 & 39 \cdot 18 \equiv 14 & 39 \cdot 16 \equiv 22 & \end{array}$$

which produces the plaintext in numerical form.

5.3.2 Digital Signatures

To confirm the integrity of a message, that is to confirm that the incoming message was sent by an authorized person, the sender must provide a *digital signature*. Fortunately, the ElGamal cryptosystem allows for an efficient procedure for authenticating messages.

Consider a user (sender) of the system who has a public key (p, r, a), private key k, and encrypted message M. The first step toward supplying a signature is to choose an integer $1 \leq j \leq p-1$ where $\gcd(j, p-1) = 1$. Let B be the first block (and later blocks) of the ciphertext message. The user computes

$$c \equiv r^j \pmod{p}, \; 0 \leq j \leq p-1$$

and then obtains a solution of the linear congruence

$$jd + kc \equiv B \pmod{p-1} \implies jd \equiv B - kc,\ 0 \leq d \leq p-2$$

The solution d can be found using the Euclidean algorithm. The pair of integers (c, d) is the required digital signature appended to the message. Note that while c can be made by anyone, the integer d can be created only by someone who knows the private key k, the random integer j, and the encoded message M. What really matters is that the sender knows k.

The recipient uses the sender's public key (p, r, a) to confirm the purported signature. By calculating the two values

$$V_1 \equiv a^c c^d \pmod{p} \text{ and } V_2 \equiv r^B \pmod{p},\ \ 0 \leq V_1, V_2 \leq p-1$$

the signature is accepted as legitimate if $V_1 = V_2$, since (if the actual value of d is the solution to the linear congruence $jd + kc \equiv B \pmod{p-1}$),

$$\begin{aligned} V_1 \equiv a^c c^d &\equiv (r^k)^c (r^j)^d \\ &\equiv r^{kc+jd} \\ &\equiv r^B \equiv V_2 \pmod{p} \end{aligned}$$

In other words, this signature verifies that the sender actually has the key k (which must be needed to get the proper value of d). Note that this does not require the receiver to know the key.

For example, a sender having public key $(43, 3, 22)$ and private key $k = 15$ wants to sign and reply to the message SELL NOW. This is carried out by first choosing an integer $0 \leq j \leq 42$ with $\gcd(j, 42) = 1$; say $j = 25$. If the first block of the encoded reply is $B = 13$, then the person calculates

$$c \equiv 3^{25} \equiv 5 \pmod{43}$$

and solves the congruence

$$25d \equiv 13 - 5 \cdot 15 \pmod{42}$$

to get $d \equiv 16 \pmod{42}$. The digital signature is therefore $(5, 16)$. On its arrival, the signature is confirmed by checking the equality of integers V_1 and V_2.

$$\begin{aligned} V_1 &\equiv 22^5 \cdot 5^{16} \equiv 39 \cdot 40 \equiv 12 \pmod{43} \\ V_2 &\equiv 3^{13} \equiv 12 \pmod{43} \end{aligned}$$

6 Perfect Numbers and Mersenne Primes

Definition 6.1

A *proper divisor* of an integer n are all of its divisors except n itself.

Definition 6.2

A positive integer n is said to be *perfect* if n is equal to the sum of its proper divisors.

We can also express it in the following way. Let $\sigma(n)$ be the sum of all of its divisors. Then, a perfect number is an integer n such that

$$\sigma(n) = 2n$$

Example 6.1

Some examples of proper divisors are:

$$\sigma(6) = 1 + 2 + 3 + 6 = 2 \cdot 6$$
$$\sigma(28) = 1 + 2 + 4 + 7 + 14 + 28 = 2 \cdot 28$$

Let P_k be the kth proper divisor, then

$$P_3 = 496$$
$$P_4 = 8128$$
$$P_5 = 33550336$$
$$P_6 = 8589869056$$
$$P_7 = 137438691328$$
$$P_8 = 2305843008139952128$$
$$P_9 = 2658455991569831744654692615953842176$$

It is not known whether there are a finite number or an infinite number of perfect numbers. We proceed to find some patterns in the form of perfect numbers.

Theorem 6.1

If $2^k - 1$ is prime $(k > 1)$, then

$$n = 2^{k-1}(2^k - 1)$$

is perfect and every even perfect number is of this form.

Therefore, the problem of finding even perfect numbers is reduced to the search of all primes of the form $2^k - 1$. That is, upon finding a Mersenne prime, we just multiply it by the corresponding multiple of 2 to get a perfect number.

Definition 6.3

Numbers of the form

$$M_n = 2^n - 1, \quad n \geq 1$$

are called *Mersenne numbers*. Mersenne numbers that are also prime are called *Mersenne primes*.

Lemma 6.2

If $a^k - 1$ is prime $(a > 0, k \geq 2)$, then $a = 2$ and k is prime.

Proof. Since

$$a^k - 1 = (a - 1)(a^{k-1} + a^{k-2} + ... + a + 1)$$

where

$$a^{k-1} + a^{k-2} + ... + a + 1 \geq a + 1 > 1$$

the other factor of $a^k - 1$ (which is assumed to be prime) must be 1. So, $a - 1 = 1 \implies a = 2$. To prove k prime, assume that it is composite. Then, we can write $k = rs$, where $r, s > 1$. Then,

$$\begin{aligned} a^k - 1 &= (a^r)^s - 1 \\ &= (a^r - 1)(a^{r(s-1)} + a^{r(s-2)} + ... + a^r + 1) \end{aligned}$$

where both factors are clearly greater than 1. This violates that $a^k - 1$ must be prime, so our assumption that k is composite is false.

We can write the first six Mersenne primes (also perfect numbers) as

$$\begin{aligned} P_1 &= 2(2^2 - 1) \\ P_2 &= 2^2(2^3 - 1) \\ P_3 &= 2^4(2^5 - 1) \\ P_4 &= 2^6(2^7 - 1) \\ P_5 &= 2^{12}(2^{13} - 1) \\ P_6 &= 2^{16}(2^{17} - 1) \\ P_7 &= 2^{18}(2^{19} - 1) \\ P_8 &= 2^{30}(2^{31} - 1) \\ P_9 &= 2^{66}(2^{67} - 1) \end{aligned}$$

This leads to the question of whether there are an infinite number primes of the type $2^p - 1$, where p is a prime.

Theorem 6.3 (Conjecture)

There exists an infinite number of Mersenne primes of form

$$2^p - 1, \quad p \text{ prime}$$

If this conjecture is true, then this would imply that there exists an infinite number of (even) perfect numbers. We can also prove results on the digits of even perfect numbers. So far, there are a total of 51 Mersenne primes found, with the largest being

$$2^{82589933} - 1$$

with $24,862,048$ digits when written in base-10. It is also the largest known prime as of November 2020.

Theorem 6.4

An even perfect number n ends in the digit 6 or 8. That is,

$$n \equiv 6 \pmod{10} \text{ or } n \equiv 8 \pmod{10}$$

Even better, every even perfect number ends in 6 or 28.

One property that was noticed was that substituting some Mersenne primes for n in the formula $2^n - 1$ produces a higher Mersenne prime. This works for the first four Mersenne primes $3, 7, 31$, and 127.

$$2^2 - 1 = 3 \implies 2^3 - 1 = 7$$

It was conjectured that if the number M_n is prime, then M_{M_n} is also prime, but this was shown to false when

$$M_{M_{13}} = 2^{M_{13}} - 1 = 2^{8191} - 1$$

was shown to be composite.

The final type of numbers is a *Fermant number*.

Definition 6.4

A *Fermant number* is an integer of the form

$$F_n = 2^{2^n} + 1, \quad n \geq 0$$

If F_n is prime, then it is said to be a *Fermant prime*.

The first five Fermant numbers are indeed prime, but F_5 was shown to be composite.

$$\begin{aligned} F_0 &= 2^{2^0} + 1 = 3 \\ F_1 &= 2^{2^1} + 1 = 5 \\ F_2 &= 2^{2^2} + 1 = 17 \\ F_3 &= 2^{2^3} + 1 = 257 \\ F_4 &= 2^{2^4} + 1 = 65537 \\ F_5 &= 2^{2^5} + 1 = 4294967297 \end{aligned}$$

Theorem 6.5

The Fermant number F_5 is divisible by 641.

Proof. By letting $a = 2^7$ and $b = 5$, we have

$$1 + ab = 641$$

We can see that

$$1 + ab - b^4 = 1 + (a - b^3)b = 1 + 3b = 2^4$$

This implies that

$$\begin{aligned} F_5 = 2^{2^5} + 1 &= 2^{32} + 1 \\ &= 2^4 a^4 + 1 \\ &= (1 + ab - b^4)a^4 + 1 \\ &= (1 + ab)a^4 + (1 - a^4 b^4) \\ &= (1 + ab)\big(a^4 + (1 - ab)(1 + a^2 b^2)\big) \end{aligned}$$

which gives $641 | F_5$.

It is not known whether there are an infinite number of Fermant primes, or even if there is at least one Fermant prime beyond F_4. But there is a useful property about Fermant numbers in that they are relatively prime to each other.

Lemma 6.6

For distinct Fermant numbers F_n, F_m, where $n, m \geq 0$,

$$\gcd(F_m, F_n) = 1$$

One final result we have is about the divisors of Fermant numbers.

Theorem 6.7

Any prime divisor p of the Fermant number $F_n = 2^{2^n} + 1$, where $n \geq 2$, is of the form

$$p = k \cdot 2^{n+2} + 1$$

7 Certain Nonlinear Diophantine Equations

Definition 7.1

A *Pythagorean triple* is a set of three integers x, y, z such that

$$x^2 + y^2 = z^2$$

Theorem 7.1

All the solutions of the Pythagorean equation

$$x^2 + y^2 = z^2$$

satisfying the conditions

$$\gcd(x, y, z) = 1, \;\; 2 \big| x, \;\; x, y, z > 0$$

are given by the formulas

$$x = 2st, \;\; y = s^2 - t^2, \;\; z = s^2 + t^2$$

for integers $s > t > 0$ such that $\gcd(s, t) = 1$ and $s \not\equiv t \pmod 2$.

Corollary 7.2

The radius of the inscribed circle of a Pythagorean triangel is always an integer.

7.1 Fermant's Last Theorem

Theorem 7.3

The Diophantine equation

$$x^4 + y^4 = z^2$$

has no solution in the positive integers x, y, z.

Proof. Assume that there exists a positive solution x_0, y_0, z_0. Without loss of generality, suppose also that $\gcd(x_0, y_0) = 1$. Then, we express the equation as

$$(x_0^2)^2 + (y_0^2)^2 = z_0^2$$

meaning that x_0^2, y_0^2, z_0 must be a Pythagorean triple and must be (without loss of generality of the order of x_0 and y_0)

$$x_0^2 = 2st$$
$$y_0^2 = s^2 - t^2$$
$$z_0 = s^2 + t^2$$

where $s > t > 0$ are relatively prime integers and exactly one of s and t is even. Note that since y_0 is

odd, $y_0^2 \equiv 1 \pmod 4$. If s is even, then

$$1 \equiv y_0^2 = s^2 - t^2 \equiv 0 - 1 \equiv 3 \pmod 4$$

which is an impossibility. Therefore, s must be odd and so t is even; denote $t = 2r$. Then, the equation $x_0^2 = 2st$ becomes $x_0^2 = 4sr$, which says that

$$\left(\frac{x_0}{2}\right)^2 = sr$$

But note that since $\gcd(s, r) = 1$ (due to $\gcd(s, t) = 1$) and sr is a perfect square, this must imply that each of the integers s and r are both perfect squares. Denote them by $s = z_1^2, r = w_1^2$. Now, since

$$t^2 + y_0^2 = s^2$$

and $\gcd(s, t) = 1$, it follows that $\gcd(t, y_0, s) = 1$, making them a Pythagorean triple. With t even, we get

$$\begin{aligned} t &= 2uv \\ y_0 &= u^2 - v^2 \\ s &= u^2 + v^2 \end{aligned}$$

for relatively prime integers $u > v > 0$. Now, the relation

$$uv = \frac{t}{2} = r = w_1^2$$

implies that u and v are both squares, so denote $u = x_1^2, v = y_1^2$. When these values are substituted in the equation $s = u^2 + v^2$, we get

$$z_1^2 = s = u^2 + v^2 = x_1^4 + y_1^4$$

and we are back at the same equation again. But now, consider the inequality

$$0 < z_1 \leq z_1^2 = s \leq s^2 < s^2 + t^2 = z_0$$

Therefore, starting with one solution x_0, y_0, z_0, we have proved the existence of another solution x_1, y_1, z_1 such that $0 < z_1 < z_0$. Repeating the argument, this would lead to a third solution x_2, y_2, z_2 and so forth, which provides an infinite decreasing sequence of positive integers

$$z_0 > z_1 > z_2 > ...$$

But since there is only a finite supply of positive integers less than z_0, a contradiction occurs, so no solution does exist.

An immediate result is the following corollary.

Corollary 7.4

The equation $x^4 + y^4 = z^4$ has no solution in the positive integers.

Proof. (x_0, y_0, z_0) being a positive solution implies that (x_0, y_0, z_0^2) is a solution of $x^4 + y^4 = z^2$, which contradicts the previous theorem.

If $n > 2$, then n is either a power of 2 or divisible by an odd prime p. In the first case, $n = 4k$ and the

Fermant equation $x^n + y^n = z^n$ can be written as

$$(x^k)^4 + (y^k)^4 = (z^k)^4$$

which does not have a solution by the previous corollary. When $n = pk$, the Fermant equation is the same as

$$(x^k)^p + (y^k)^p = (z^k)^p$$

So, if it could be shown that the equation $x^p + y^p = z^p$ has no solution, then, there would exist no solutions for $x^n + y^n = z^n$. After more than 300 years of effort, Fermant's conjecture turned out to be true (proved in 1995).

Theorem 7.5 (Fermant's Last Theorem)

There exist no solution to the Diophantine equation

$$x^n + y^n = z^n$$

for all integers $n > 2$. For $n = 1, 2$, there is clearly an infinite number of solutions.

Theorem 7.6 (Fermant)

The Diophantine equation

$$x^4 - y^4 = z^2$$

has no solution in the positive integers x, y, z.

Theorem 7.7

The area of a Pythagorean triangle can never be equal to a perfect square.

Proof. Assume that a solution exists with side lengths x, y and hypotenuse length z such that $x^2 + y^2 = z^2$. Then, the area of the triangle is $\frac{1}{2}xy$ and let it be equal to u^2 for some $u \in \mathbb{N}$. Then, $2xy = 4u^2$, and adding/subtracting the equation into $x^2 + y^2 = z^2$, we get

$$(x + y)^2 = z^2 + 4u^2, \;\; (x - y)^2 = z^2 - 4u^2$$

When these last two equations are multiplied together, we get

$$(x^2 - y^2)^2 = z^4 - 16u^4 = z^4 - (2u)^4$$

But this contradicts the fact that there exists solutions to the equation $x^4 - y^4 = z^2$, so no such u can exist.

8 Representation of Integers as Sums of Squares

8.1 Sums of Two Squares

A common question is to find whether every integer can be expressed as a sum of squares, and if so, what is the minimum number of squares (including 0^2) that one needs to express an integer? It turns out to be 4 (e.g. $7 = 2^2 + 1^2 + 1^2 + 1^2$), but we will first explore the necessary and sufficient conditions that a positive integer be representable as the sum of two squares.

Lemma 8.1

If m and n are each the sum of two squares, then so it their product mn.

Proof. If $m = a^2 + b^2$ and $n = c^2 + d^2$, then

$$mn = (a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (ad - bc)^2$$

Clearly, not every prime can be written as the sum of two squares, since if this were indeed true, then by the previous lemma, every number can be written as a sum of squares (which contradicts our counterexample that $7 = 2^2 + 1^2 + 1^2 + 1^2$).

Theorem 8.2

No prime p of the form $4k + 3$ is a sum of two squares.

Proof. $a \equiv 0, 1, 2, 3 \pmod 4$ for all $a \in \mathbb{N} \implies a^2 \equiv 0, 1 \pmod 4$. This means that

$$a^2 + b^2 \equiv 0, 1, 2 \pmod 4$$

Lemma 8.3 (Thue's Lemma)

Let p be a prime and let $\gcd(a, p) = 1$. Then, the congruence

$$ax \equiv y \pmod p$$

admits a solution x_0, y_0, where

$$0 < |x_0| < \sqrt{p} \text{ and } 0 < |y_0| < \sqrt{p}$$

Theorem 8.4 (Fermant)

An odd prime p is expressible as a sum of two squares if and only if $p \equiv 1 \pmod 4$.

Corollary 8.5

Any prime p of the form $4k + 1$ can be represented uniquely, up to order of the summands, as a sum of two squares.

The following is a statement about representing integers as the *difference* of two squares.

Theorem 8.6

A positive integer n can be represented as the difference of two squares if and only if n is not of the form $4k + 2$.

Proof. Because $a^2 \equiv 0, 1 \pmod 2$ for integers a, it follows that

$$a^2 - b^2 \equiv 0, 1, 3 \pmod 4$$

Corollary 8.7

An odd prime is the difference of two successive squares.

Proof. We can put p in the form

$$p = \left(\frac{p+1}{2}\right)^2 - \left(\frac{n-1}{2}\right)^2$$

For example,

$$11 = 6^2 - 5^2, \;\; 17 = 9^2 - 8^2, \;\; 29 = 15^2 - 14^2$$

8.2 Sums of More Than Two Squares

Expanding the allowed number of summands to three squares allows to broaden the amount of integers expressible as sums of squares. For example,

$$14 = 3^2 + 2^2 + 1^2, \;\; 33 = 5^2 + 2^2 + 2^2, \;\; 67 = 7^2 + 3^2 + 3^2$$

But we can guarantee that there exists integers that are still not expressible as the sum of two squares.

Theorem 8.8

No positive integer of the form $4^n(8m+7)$ can be represented as the sum of three squares.

Proof. For any integer a, $a^2 \equiv 0, 1, 4 \pmod 8$, which implies that

$$a^2 + b^2 + c^2 \equiv 0, 1, 2, 3, 4, 5, 6 \pmod 8$$

for any integers a, b, c. So, there exist no solutions for $a^2 + b^2 + c^2 = 8m + 7$. Now, suppose that $n \geq 1$ and solutions exist to the equation

$$a^2 + b^2 + c^2 = 4^n(8m+7)$$

Then, all three integers a, b, c must be even (choosing exactly one to be even leads to an inconsistency when doing $\pmod 4$). So, substituting, $a = 2a_1, b = 2b_1, c = 2c_1$, we get

$$a_1^2 + b_1^2 + c_1^2 = 4^{n-1}(8m+7)$$

We can do this until one of the a_i, b_i, or c_i are odd or $n = 1$. In either case, this leads to a contradiction.

To prove that every number p can be written as the sum of four squares, we need the following two lemmas.

Lemma 8.9 (Euler)

If the integers m and n are each the sum of four squares, then mn is likewise representable as sums of four squares.

Proof. A straightforward, yet tedious calculation shows this.

$$\begin{aligned} mn &= (a_1^2 + a_2^2 + a_3^2 + a_4^2)(b_1^2 + b_2^2 + b_3^2 + b_4^2) \\ &= (a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4)^2 \\ &+ (a_1b_2 - a_2b_1 + a_3b_4 - a_4b_3)^2 \\ &+ (a_1b_3 - a_2b_4 - a_3b_1 + a_4b_2)^2 \\ &+ (a_1b_4 + a_2b_3 - a_3b_2 - a_4b_1)^2 \end{aligned}$$

Lemma 8.10

If p is an odd prime, then the congruence

$$x^2 + y^2 + 1 \equiv 0 \pmod{p}$$

has a solution x_0, y_0 where

$$0 \leq x_0 \leq (p-1)/2, \;\; 0 \leq y_0 \leq (p-1)/2$$

This leads to the theorem we've been waiting for.

Theorem 8.11

Any prime p can be written as the sum of four squares.

By prime factorizing every number $n > 1$ and using Euler's lemma, we get.

Corollary 8.12 (Lagrange)

Any positive integer n can be written as the sum of four squares, some of which may be 0.

These results have a natural extension to sums of higher powers. In fact, the minimum number of kth powers needed to produce a representation of every natural number is denoted $g(k)$.

Theorem 8.13

Every positive integer can be expressed as the sum of 9 cubes. That is, $g(3) = 9$.

However, only the numbers

$$\begin{aligned} 23 &= 2^3 + 2^3 + 1^3 + 1^3 + 1^3 + 1^3 + 1^3 + 1^3 + 1^3 \\ 239 &= 4^3 + 4^3 + 3^3 + 3^3 + 3^3 + 3^3 + 1^3 + 1^3 + 1^3 \end{aligned}$$

are the only integers that actually require as many as 9 cubes in their representation. We can claim something even stronger.

Theorem 8.14 (Linnik)

There are only a finite number of integers that require at least 8 cubes in their represenations.

Theorem 8.15

Every positive integer can be expressed as the sum of 53 fourth powers. That is, $g(4) = 19$. Furthermore, $g(5) = 37$.

For higher numbers n, the following result was proved.

Theorem 8.16

For all but a finite number of integers $n \geq 6$, the following formula holds.

$$g(k) = \left\lfloor \left(\frac{3}{2}\right)^k \right\rfloor + 2^k - 2$$

However, there is strong evidence that this theorem holds for all p.

9 Fibonacci Numbers

Definition 9.1

The *Fibonacci sequence* is defined recursively as

$$u_n = \begin{cases} 1 & n = 1, 2 \\ u_{n-1} + u_{n-2} & n \geq 3 \end{cases}$$

The first few Fibonacci numbers are

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, ...$$

Theorem 9.1

In the Fibonnaci sequence, $\gcd(u_n, u_{n+1}) = 1$ for every $n \geq 1$.

Proof. Suppose that $d > 1$ and that the integer d divides both u_n and u_{n-1}. Then, it divides $u_{n-2} = u_n - u_{n-1}$, and doing this recursively, this implies that $d|u_1$, which is false since $u_1 = 1$.

Theorem 9.2

Except u_1, u_2, u_6, and u_{12}, each Fibonacci number has a "new" prime factor; that is, a prime factor that does not occur in any Fibonacci number with a smaller subscript.

Theorem 9.3

For $m, n \geq 1$, u_{mn} is divisible by u_m.

Theorem 9.4

The greatest common divisor of two Fibonacci numbers is also a Fibonacci number. In fact,

$$\gcd(u_m, u_n) = u_d, \text{ where } d = \gcd(n, m)$$

Corollary 9.5

In the Fibonacci sequence, $u_m \mid u_n$ if and only if $m|n$ for $n \geq m \geq 3$.

The following theorem shows a result in expressing integers as sums of Fibonacci numbers.

Theorem 9.6 (Zeckendorf Representation)

Any positive integer N can be expressed as a sum of distinct Fibonacci numbers, no two of which are consecutive. That is,

$$N = u_{k_1} + u_{k_2} + ... + u_{k_r}$$

where $k_1 \geq 2$ and $k_{j+1} \geq k_j + 2$ for $j = 1, 2, ..., r-1$.

Using linear algebra, the explicit representation of Fibonacci numbers is evident.

Theorem 9.7 (Binet's Formula)

For every Fibonacci number u_n,

$$u_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right) \implies u_n = \frac{\alpha^n - \beta^n}{\alpha - \beta}$$

where

$$\alpha = \frac{1+\sqrt{5}}{2},\ \beta = \frac{1-\sqrt{5}}{2}$$

Proof. The first formula can be found using linear algebra.

One useful application of Binet's formula is to produce new Fibonacci numbers from old ones.

Corollary 9.8

We claim that

$$u_{n+2}^2 - u_n^2 = u_{2n+2}$$

Proof. Since $\alpha\beta = 1$, we have $(\alpha\beta)^{2k} = 1$,

$$\begin{aligned} u_{n+2}^2 - u_n^2 &= \left(\frac{\alpha^{n+2} - \beta^{n+2}}{\alpha - \beta}\right)^2 - \left(\frac{\alpha^n - \beta^n}{\alpha - \beta}\right)^2 \\ &= \frac{\alpha^{2(n+2)} - 2 + \beta^{2(n+2)}}{(\alpha-\beta)^2} - \frac{\alpha^{2n} - 2 + \beta^{2n}}{(\alpha-\beta)^2} \\ &= \frac{\alpha^{2(n+2)} + \beta^{2(n+2)} - \alpha^{2n} - \beta^{2n}}{(\alpha-\beta)^2} \\ &= \frac{(\alpha^2 - \beta^2)(\alpha^{2n+2} - \beta^{2n+2})}{(\alpha-\beta)^2} \\ &= (\alpha+\beta)\left(\frac{\alpha^{2n+2} - \beta^{2n+2}}{\alpha - \beta}\right) \\ &= 1 \cdot u_{2n+2} = u_{2n+2} \end{aligned}$$

Another one.

Corollary 9.9

We claim that

$$u_{2n+1}u_{2n-1} - 1 = u_{2n}^2$$

Proof. We calculate

$$\begin{aligned} u_{2n+1}u_{2n-1} &= \left(\frac{\alpha^{2n+1} - \beta^{2n+1}}{\sqrt{5}}\right)\left(\frac{\alpha^{2n-1} - \beta^{2n-1}}{\sqrt{5}}\right) - 1 \\ &= \frac{1}{5}\big(\alpha^{4n} + \beta^{4n} - (\alpha\beta)^{2n-1}\alpha^2 - (\alpha\beta)^{2n-1}\beta^2 - 5\big) \\ &= \frac{1}{5}\big(\alpha^{4n} + \beta^{4n} + (\alpha^2 + \beta^2) - 5\big) \end{aligned}$$

Since $\alpha^2 + \beta^2 = 3$, we have

$$\begin{aligned} \frac{1}{5}\big(\alpha^{4n} + \beta^{4n} - 2\big) &= \frac{1}{5}\big(\alpha^{4n} + \beta^{4n} - 2(\alpha\beta)^{2n}\big) \\ &= \left(\frac{\alpha^{2n} - \beta^{2n}}{\sqrt{5}}\right)^2 = u_{2n}^2 \end{aligned}$$

Corollary 9.10

Binet's formula can be modified to

$$u_n = \left[\frac{\alpha^n}{\sqrt{5}} + \frac{1}{2}\right]$$

Proof. Since $0 < |\beta| < 1$, we see that

$$|\beta^n| = |\beta|^n < 1 \text{ for } n \geq 1$$

Therefore, we have

$$\begin{aligned} \left|u_n - \frac{\alpha^n}{\sqrt{5}}\right| &= \left|\frac{\alpha^n - \beta^n}{\sqrt{5}} - \frac{\alpha^n}{\sqrt{5}}\right| \\ &= \frac{|\beta^n|}{\sqrt{5}} < \frac{1}{\sqrt{5}} < \frac{1}{2} \end{aligned}$$

Therefore, we can view u_n as the largest integer not exceeding $\frac{\alpha^n}{\sqrt{5}} + \frac{1}{2}$, leading to the formula

$$u_n = \left[\frac{\alpha^n}{\sqrt{5}} + \frac{1}{2}\right]$$

We also introduce two final theorems concerning prime factors of Fibonacci numbers.

Theorem 9.11

For any prime $p > 5$, either

$$p \,|\, u_{p-1} \text{ or } p \,|\, u_{p+1}$$

but not both.

Theorem 9.12

Let $p \geq 7$ be a prime for which $p \equiv 2 \pmod 5$ of $p \equiv 4 \pmod 5$. If $2p-1$ is also prime, then

$$(2p-1) \mid u_p$$

Example 9.1

$u_1 9 = 37 \cdot 113$, where $19 \equiv 4 \pmod 5$. $u_{37} = 73 \cdot 330929$, where $37 \equiv 2 \pmod 5$.

10 Continued Fractions

Definition 10.1

A *partition* of a positive integer n is a way of writing n as a sum of positive integers, with order being irrelevant.. Let $p(n)$ denote the total number of partitions of n.

Example 10.1

The

Theorem 10.1 (Hardy-Ramanujan)

For large n, the partition function satisfies the relation

$$p(n) \approx \frac{e^{c\sqrt{n}}}{4n\sqrt{3}}, \quad c = \pi\sqrt{\frac{2}{3}}$$

Theorem 10.2 (Ramanujan)

With the partition function p and any integer n, we have

$$p(5k+4) \equiv 0 \pmod 5 \tag{1}$$
$$p(7k+5) \equiv 0 \pmod 7 \tag{2}$$
$$p(11k+6) \equiv 0 \pmod{11} \tag{3}$$

Theorem 10.3 (Ramanujan)

The constant π can be calculated with the infinite series.

$$\frac{1}{\pi} = \frac{\sqrt{8}}{9801} \sum_{n=0}^{\infty} \frac{(4n)!}{(n!)^4} \frac{[1103 + 26390n]}{396^{4n}}$$

Each successive term in the series adds roughly 8 more correct digits! The efficiency of this series has made it possible to calculate millions of digits of π. Another series is

$$\frac{1}{\pi} = \sum_{n=0}^{\infty} \binom{2n}{n}^3 \frac{42n+5}{2^{12n+4}}$$

10.1 Finite Continued Fractions

Definition 10.2

A *finite continued fraction* is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{... + \cfrac{...}{a_{n-1} + \cfrac{1}{a_n}}}}$$

where $a_0, a_1, ..., a_n$ are all real numbers, all of which except possible a_0 are positive. The a_k's are called the *partial denominators* of this fraction. If the a_k's are all integers, then the fraction is called a *simple finite continued fraction.*

Theorem 10.4

Any rational number can be written as a finite simple continued fraction with the algorithm presented in the proof.

Proof. Let a/b, where $b > 0$ be an arbitrary rational number. Euclid's algorithm for finding the greatest common divisor of a and b gives us the equations

$$\begin{aligned}
a &= ba_0 + r_1 && 0 < r_1 < b \\
b &= r_1 a_1 + r_2 && 0 < r_2 < r_1 \\
r_1 &= r_2 a_2 + r_3 && 0 < r_3 < r_2 \\
&... && ... \\
r_{n-1} &= r_{n-1} a_{n-1} + r_n && 0 < r_n < r_{n-1} \\
r_{n-1} &= r_n a_n + 0
\end{aligned}$$

We can rewrite it in the following way.

$$\begin{aligned}
\frac{a}{b} &= a_0 + \frac{r_1}{b} = a_0 + \frac{1}{\frac{b}{r_1}} \\
\frac{b}{r_1} &= a_1 + \frac{r_2}{r_1} = a_1 + \frac{1}{\frac{r_1}{r_2}} \\
\frac{r_1}{r_2} &= a_2 + \frac{r_3}{r_2} = a_2 + \frac{1}{\frac{r_2}{r_3}} \\
... &= ... \\
\frac{r_{n-1}}{r_n} &= a_n
\end{aligned}$$

Then by substituting the equations below to the one above it starting from the third equation, we can get

$$\frac{a}{b} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cfrac{r_2}{r_3}}}}$$

Continuing in from the bottom equation, we get

$$\frac{a}{b} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{... + \cfrac{...}{a_{n-1} + \cfrac{1}{a_n}}}}$$

Example 10.2

We find the representation of 19/51 as a continued fraction. We use Euclid's algorithm to get

$$\begin{aligned} 51 = 2 \cdot 19 + 13 &\implies \frac{51}{19} = 2 + \frac{13}{19} \\ 19 = 1 \cdot 13 + 6 &\implies \frac{19}{13} = 1 + \frac{6}{13} \\ 13 = 2 \cdot 6 + 1 &\implies \frac{13}{6} = 2 + \frac{1}{6} \\ 6 = 6 \cdot 1 + 0 &\implies \frac{6}{6} = 1 \end{aligned}$$

After the substitutions, we get

$$\begin{aligned} \frac{19}{51} = \frac{1}{\frac{51}{9}} &= \frac{1}{2 + \frac{13}{19}} \\ &= ... \\ &= \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{2 + \frac{1}{6}}}} \end{aligned}$$

Definition 10.3

The continued fraction made from $[a_0; a_1, a_2, ...a_n]$ by cutting off the expansion after the kth partial denominator a_k is called the *kth convergent* of the given continued fraction and denoted by C_k. That is,

$$C_k = [a_0; a_1, a_2, ..., a_k], \;\; 1 \leq k \leq n$$

Much of the labor of computing convergents of a finite continued fraction can be avoided by establishing certain formulas for their numerators and denominators.

Theorem 10.5

Given a finite continued fraction $[a_0; a_1, a_2, ..., a_n]$, let

$$\begin{aligned} p_0 &= p_0 & q_0 &= 1 \\ p_1 &= a_1 a_0 + 1 & q_1 &= a_1 \\ p_k &= a_k p_{k-1} + p_{k-2} & q_k &= a_k q_{k-1} + q_{k-2} \end{aligned}$$

for $k = 2, 3, ..., n$. Then, the kth convergent of the fraction has the value

$$C_k = \frac{p_k}{q_k},\ k = 0, 1, ..., n$$

Proof. We can manually check that this is true for $k = 0, 1, 2$. Assume that it is true for $k = m$, where $2 \leq m$. Then,

$$C_m = \frac{p_m}{q_m} = \frac{a_m p_{m-1} + p_{m-2}}{a_m q_{m-1} + q_{m-2}}$$

Note that the integers $p_{m-1}, q_{m-1}, p_{m-2}, q_{m-2}$ depend on the first $m-1$ partial denominators $a_1, \dots, a_{m-1}$ and there are independent of the value of a_m. The equation above therefore remains true if we replace a_m with $a_m + \frac{1}{a_{m+1}}$.

$$\left[a_0; a_1, a_2, \dots, a_m + \frac{1}{a_{m+1}}\right] = \frac{\left(a_m + \frac{1}{a_{m+1}}\right)p_{m-1} + p_{m-2}}{\left(a_m + \frac{1}{a_{m+1}}\right)q_{m-1} + q_{m-2}}$$

But this above mth convergent is really just equal to the $(m+1)$th convergent C_{m+1} since the final term on the bottom of the continued fraction is replaced with one more continuation. This means that

$$\begin{aligned} C_{m+1} &= \frac{\left(a_m + \frac{1}{a_{m+1}}\right)p_{m-1} + p_{m-2}}{\left(a_m + \frac{1}{a_{m+1}}\right)q_{m-1} + q_{m-2}} \\ &= \frac{a_{m+1}(a_m p_{m-1} + p_{m-2}) + p_{m-1}}{a_{m+1}(a_m q_m + q_{m-2}) + q_{m-1}} \\ &= \frac{a_{m+1}p_m + p_{m-1}}{a_{m+1}q_m + q_{m-1}} \end{aligned}$$

Which is the desired formula for C_{m+1}. So, the equation is satisfied at $k = m+1$.

Theorem 10.6

If $C_k = p_k/q_k$ is the kth convergent of the finite simple continued fraction $[a_0; a_1, \dots, a_n]$, then

$$p_k q_{k-1} - q_k p_{k-1} = (-1)^{k-1},\ \ 1 \leq k \leq n$$

Proof. We use induction on k. The base case for $k = 1$ holds true since

$$p_1 q_0 - q_1 p_0 = (a_1 a_0 + 1) \cdot 1 - a_1 \cdot a_0 = 1 = (-1)^{1-1}$$

Now, assuming that the formula is true for some $k = m$, then

$$\begin{aligned} p_{m+1}q_m - q_{m+1}p_m &= (a_{m+1}p_m + p_{m-1})q_m - (a_{m+1}q_m + q_{m-1})p_m \\ &= -(p_m q_{m-1} - q_m p_{m-1}) \\ &= -(-1)^{m-1} = (-1)^m \end{aligned}$$

Corollary 10.7

For $1 \leq k \leq n$, p_k and q_k are relatively prime.

Proof. If $d = \gcd(p_k, q_k) \neq 1$, then this implies that the left hand side has factor d, which must mean that the right hand side also has factor d. But the right hand side is ± 1, leading to a contradiction.

Example 10.3

Consider the continued fraction $[0; 1, 1, ..., 1]$. The first few convergents are

$$C_0 = 0/1,\ C_1 = 1/1,\ C_2 = 1/2,\ C_3 = 2/3,\ C_4 = 3/5,\ ...$$

Because the numerator p_k and denominator q_k of the kth convergent is expressed

$$p_k = 1 \cdot p_{k-1} + p_{k-2} = p_{k-1} + p_{k-2}$$
$$q_k = 1 \cdot q_{k-1} + q_{k-2} = q_{k-1} + q_{k-2}$$

we can see that the numerator and denominator forms a Fibonacci sequence. That is,

$$C_k = \frac{u_k}{u_{k+1}},\ \ k \geq 2$$

where u_k denotes the kth Fibonacci number.

Here is another useful property of convergents.

Lemma 10.8

If q_k is the denominator of the kth convergent C_k of the simple continued fraction $[a_0; a_1, ..., a_n]$, then $q_{k-1} \leq q_k$ for $1 \leq k \leq n$, with strict inequality satisfied when $k > 1$.

Proof. We prove by induction. When $k = 1$,

$$q_0 = 1 \leq a_1 = q_1$$

Assume that it is true for $k = m$. Then,

$$q_{m+1} = a_{m+1}q_m + q_{m-1} > a_{m+1}q_m \geq 1 \cdot q_m = q_m$$

which implies that the inequality is true for $k = m + 1$.

Theorem 10.9

The convergents with even subscripts form a strictly increasing sequence.

$$C_0 < C_2 < C_4 < ...$$

The convergents with odd subscripts form a strictly decreasing sequence.

$$C_1 > C_3 > C_5 > ...$$

Every convergent with an odd subscript is greater than every convergent with an even subscript.

Proof. Using the previous theorems, we calculate that

$$\begin{aligned} C_{k+2} - C_k &= (C_{k+2} - C_{k+1}) + (C_{k+1} - C_k) \\ &= \left(\frac{p_{k+2}}{q_{k+2}} - \frac{p_{k+1}}{q_{k+1}} \right) + \left(\frac{p_{k+1}}{q_{k+1}} - \frac{p_k}{q_k} \right) \\ &= \frac{(-1)^{k+1}}{q_{k+2}q_{k+1}} + \frac{(-1)^k}{q_{k+1}q_k} \\ &= \frac{(-1)^k (q_{k+2} - q_k)}{q_k q_{k+1} q_{k+2}} \end{aligned}$$

Since $q_k > 0$ for all k and by using the previous lemma that $q_{k+2} - q_k > 0$, $C_{k+2} - C_k$ has the same algebraic sign as $(-1)^k$. So,

1. If k is even, then $C_{k+2} - C_k$ has the same sign as 1 and is thus positive, which means that

$$C_0 < C_2 < C_4 < ...$$

2. If k is odd, then $C_{k+2} - C_k$ has the same sign as -1 and is thus negative, which means that

$$C_1 > C_3 > C_5 > ...$$

To show that any odd numbered convergent C_{2r-1} is greater than any even numbered convergent C_{2s}, we divide the equation $p_k q_{k-1} - q_k p_{k-1} = (-1)^{k-1}$ by $q_k q_{k-1}$ to get

$$C_k - C_{k-1} = \frac{p_k}{q_k} - \frac{p_{k-1}}{q_{k-1}} = \frac{(-1)^{k-1}}{q_k q_{k-1}}$$

This means that $C_{2j} < C_{2j-1}$. Therefore, we can put together various inequalities and combine our results so far to get

$$C_{2s} < C_{2s+2r} < C_{2s+2r-1} < C_{2r-1}$$

From this, we can see that subsequent convergents alternatingly underestimate and overestimate the true value of the rational number n.

10.2 Infinite Continued Fractions

Definition 10.4

An *infinite continued fraction* is an expression of the form

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{a_3 + ...}}}$$

where $a_0, a_1, a_2, ...$ and $b_1, b_2, b_3, ...$ are real numbers. An *infinite simply continued fraction* has form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + ...}}}$$

which, for compactness, is denoted $[a_0; a_1, a_2, a_3, ...]$. If $a_0, a_1, ...$ is an infinite sequence of integers, all

positive except possibly a_0, then the infinite simple continued fraction $[a_0; a_1, a_2, ...]$ has the value

$$\lim_{n\to\infty}[a_0; a_1, a_2, ..., a_n]$$

Theorem 10.10 (Brouncker)

The infinite product

$$\frac{4}{\pi} = \frac{3\cdot 3\cdot 5\cdot 5\cdot 7\cdot 7\cdot ...}{2\cdot 4\cdot 4\cdot 6\cdot 6\cdot 8\cdot ...}$$

can be converted into the identity

$$\frac{4}{\pi} = 1 + \cfrac{1^2}{2 + \cfrac{3^2}{2 + \cfrac{5^2}{2 + \cfrac{7^2}{2 + ...}}}}$$

However, this calculation is not computationally efficient.

Theorem 10.11 (Ramanujan)

$$e^{2\pi/5}\left(\sqrt{\frac{5+\sqrt{5}}{2}} - \frac{1+\sqrt{5}}{2}\right) = \cfrac{1}{1 + \cfrac{e^{-2\pi}}{1 + \cfrac{e^{-4\pi}}{1 + \cfrac{e^{-6\pi}}{1 + ...}}}}$$

Definition 10.5

If an infinite simple continued fraction contains a block of partial denominators $a_1, a_2, ..., a_r$, then we can write it as

$$[a_0; \overline{a_1, a_2, ..., a_n}]$$

Theorem 10.12

The value of any infinite continued fraction is an irrational number.

Theorem 10.13

Two infinite continued fractions $[a_0; a_1, a_2, ...]$ and $[b_0; b_1, b_2, ...]$ are equal if and only if $a_i = b_i$ for $i = 0, 1, 2, ...$

Corollary 10.14

Two distinct infinite continued fractions represent two distinct irrational numbers.

Theorem 10.15

Every irrational number has a unique representation as an infinite continued fraction, the representation being obtained from the continued fraction algorithm described in the following proof.

Proof. Given an arbitrary irrational number x_0, we would want to identify it with a certain sequence $[a_0; a_1, a_2, ...]$ such that the continued fraction determined by the sequence x_0. We first define

$$x_1 = \frac{1}{x_0 - [x_0]}, \;\; x_2 = \frac{1}{x_1 - [x_1]}, \;\; x_3 = \frac{1}{x_2 - [x_2]}, ...$$

and then take

$$a_0 = [x_0], \;\; a_1 = [x_1], \;\; a_2 = [x_2], \;\; a_3 = [x_3], ...$$

In general, the a_k are given inductively by

$$a_k = [x_k], \;\; x_{k+1} = \frac{1}{x_k - a_k}$$

Clearly, x_{k+1} is irrational if x_k is irrational. Since x_0 is irrational, every x_k is irrational. Thus,

$$0 < x_k - a_k = x_k - [x_k] < 1 \implies x_{k+1} = \frac{1}{x_k - a_k} > 1$$

with $a_{k+1} = [x_{k+1}] \geq 1$ for all $k \geq 0$. This leads to an infinite sequence of integers $a_0, a_1, ...$, all positive except possibly for a_0. Now, by defining x_k in the form

$$x_k = a_k + \frac{1}{x_{k+1}}$$

through successive substitutions, we get

$$\begin{aligned} x_0 &= a_0 + \frac{1}{x_1} \\ &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{x_2}} \\ &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{x_3}}} \\ &= ... \\ &= [a_0; a_1, a_2, ..., a_n, x_{n+1}] \end{aligned}$$

Corollary 10.16

If p_n/q_n is the n convergent to the irrational number x, then

$$\left| x - \frac{p_n}{q_n} \right| < \frac{1}{q_{n+1} q_n} < \frac{1}{q_n^2}$$

Combining these results, we can see that the following map

$$\rho : \mathbb{Z}^\omega \longrightarrow \mathbb{R} \setminus \mathbb{Z}$$

that maps sequences to infinite continued fractions is bijective.

Example 10.4

To calculate the infinite fraction form of $\pi = 3.141592...$, we use the algorithm to get

$$\begin{aligned}
x_0 &= \pi = 3 + (\pi - 3) && a_0 = 3 \\
x_1 &= \frac{1}{x_0 - [x_0]} = \frac{1}{0.14159265...} = 7.06251330... && a_1 = 7 \\
x_2 &= \frac{1}{x_1 - [x_1]} = \frac{1}{0.06251330...} = 15.99659440... && a_2 = 15 \\
x_3 &= \frac{1}{x_2 - [x_2]} = \frac{1}{0.99659440...} = 1.00341723... && a_3 = 1 \\
x_4 &= \frac{1}{x_3 - [x_3]} = \frac{1}{0.00341723...} = 292.63467... && a_4 = 292 \\
&... && ...
\end{aligned}$$

Thus, the infinite continued fraction for π starts with

$$\pi = [3; 7, 15, 1, 292, ...]$$

But unlike most irrational numbers, there is no explicit pattern that gives a complete sequence of a_n.

Theorem 10.17 (Euler)

Here are nice representations of $e = 2.71828...$, which does have a pattern of even integers occurring in order and separated by two 1's.

$$e = [2; 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, ...]$$

Moreover, the following representations have partial denominators that form an arithmetic progression.

$$\frac{e-1}{e+1} = [0; 2, 6, 10, 14, 18, ...] \qquad \frac{e^2-1}{e^2+1} = [0; 1, 3, 5, 7, 9, ...]$$