Notes on Primality Testing And Public Key Cryptography Part 1: Randomized Algorithms Miller–Rabin and Solovay–Strassen Tests

Jean Gallier and Jocelyn Quaintance
Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104, USA
e-mail: jean@cis.upenn.edu

February 27, 2019

Contents

Chapter 1

Introduction

1.1 Prime Numbers and Composite Numbers

Prime numbers have fascinated mathematicians and more generally curious minds for thousands of years. What is a prime number? Well, $2, 3, 5, 7, 11, 13, \ldots, 9973$ are prime numbers. The defining property of a prime number p is that *it is a positive integer $p \geq 2$ that is only divisible by 1 and p.* Equivalently, p is prime if and only if p is a positive integer $p \geq 2$ that is not divisible by any integer m such that $2 \leq m < p$. A positive integer $n \geq 2$ which is not prime is called *composite*. Observe that the number 1 is considered neither a prime nor a composite. For example, $6 = 2 \cdot 3$ is composite. Is $3\,215\,031\,751$ composite? Yes, because

$$3\,215\,031\,751 = 151 \cdot 751 \cdot 28351.$$

The above number has the remarkable property of being the only composite integer less than $25 \cdot 10^9$ which a strong pseudoprime simultaneously to the bases $2, 3, 5, 7$; see Definition 5.5, and Ribenboim [18] (Chapter 2, Section XI).

Even though the definition of primality is very simple, the structure of the set of prime numbers is highly nontrivial. The prime numbers are the basic building blocks of the natural numbers because of the following theorem bearing the impressive name of *fundamental theorem of arithmetic*.

Theorem 1.1. *Every natural number $n \geq 2$ has a unique factorization*

$$n = p_1^{i_1} p_2^{i_2} \cdots p_k^{i_k},$$

where the exponents $i_1, \ldots, i_k$ are positive integers and $p_1 < p_2 < \cdots < p_k$ are primes.

Every book on number theory has a proof of Theorem 1.1. The proof is not difficult and uses induction. It has two parts. The first part shows the existence of a factorization. The second part shows its uniqueness. For example, see Apostol [1] (Chapter 1, Theorem 1.10).

How many prime numbers are there? Many! In fact, infinitely many.

Theorem 1.2. *The set of prime numbers is infinite.*

Proof. We give three proofs. These proofs only use the fact that every integer greater than 1 has some prime divisor.

(1) (Euclid) Suppose that $p_1 = 2 < p_2 = 3 < \cdots < p_m$ are all the primes. Consider $N = p_1 p_2 \cdots p_m + 1$. The number N must be divisible by some prime p ($p = N$ is possible). Then p must be one of the p_i, so $p = p_i$ divides $N - p_1 p_2 \cdots p_m = 1$, contradicting the fact that $p_i \geq 2$.

(2) (Kummer) Suppose that $p_1 = 2 < p_2 = 3 < \cdots < p_m$ are all the primes, as in (1), but this time let $N = p_1 p_2 \cdots p_m$. Observe that $N > 2$. The number $N - 1$ must be divisible by one of the primes p_i ($p_i = N - 1$ is possible). But if p_i divides $N - 1$, then p_i divides $N - (N - 1) = 1$, a contradiction.

(3) (Hermite) We prove that for every natural number $n \geq 2$, there is some prime $p > n$. Consider $N = n! + 1$. The number N must be divisible by some prime p ($p = N$ is possible). Any prime p dividing N is distinct from $2, 3, \ldots, n$, since otherwise p would divide $N - n! = 1$, a contradiction.

There are many more proofs; see Ribenboim [18]. □

The problem of determining whether a given integer is prime is one of the better known and most easily understood problems of pure mathematics. This problem has caught the interest of mathematicians again and again for centuries. However, it was not until the 20th century that questions about primality testing and factoring were recognized as problems of practical importance, and a central part of applied mathematics. The advent of cryptographic systems that use large primes, such as RSA, was the main driving force for the development of fast and reliable methods for primality testing. Indeed, as we see in Chapter 2, in order to create RSA keys, one needs to produce large prime numbers. How do we do that?

1.2 Methods for Primality Testing

The general strategy to test whether an integer $n > 2$ is prime or composite is to choose some property, say A, implied by primality, and to search for a counterexample a to this property for the number n, namely some a for which property A fails. We look for properties for which checking that a candidate a is indeed a countexample can be done quickly.

Typically, together with the number n being tested for primality, some candidate counterexample a is supplied to an algorithm which runs a test to determine whether a is really a counterexample to property A for n. If the test says that a is a counterexample, also called a *witness*, then we know for sure that n is composite. If the algorithm reports that a is not a witness to the fact that n is composite, does this imply that n is prime? Unfortunately, no.

This is because, there may be some composite number n and some candidate counterexample a for which the test says that a is not a countexample. Such a number a is called a *liar*. The other reason is that we haven't tested all the candidate counterexamples a for n.

The remedy is to make sure that we pick a property A such that if n is composite, then at least some candidate a is not a liar, and to test all potential countexamples a. The difficulty is that trying all candidate countexamples can be too expensive to be practical.

The following analogy may be helpful to understand the nature of such a method. Suppose we have a population and we are interested in determining whether some individual is rich or not (we will say that someone who is not rich is poor). Every individual n has several bank accounts a, and there is a test to check whether a bank account a has a negative balance. The test has the property that if it is applied to an individual n and to one of its bank accounts a, and if it is positive (it says that account a has a negative balance), then the individual n is definitely poor. Note that we are assuming that a rich person is honest, namely that all bank accounts of a rich person have a nonnegative balance. This may be an unrealistic assumption. But if the test is negative (which means that account a has a nonnegative balance), this does not imply that n is rich.

The problem is that the test may not be 100% reliable. It is possible that an individual n is poor, yet the test is negative for account a (account a has a nonnegative balance). We may also not have tested all the accounts of n.

One way to deal with this problem is to use probabilities. If we know that the conditional probability that the test is positive for some account a given that n is poor is greater than $p \geq 1/2$, then we can apply the test to ℓ accounts chosen independently at random. It is easy to show that the conditional probability that the test is negative ℓ times given that an individual n is poor is less than $(1-p)^\ell$. For p close to 1 and ℓ large enough, this probability is very small. Thus, if we have high confidence in the test (p is close to 1) and if an individual n is poor, it is very unlikely that the test will be negative ℓ times.

Actually, what we would really like to know is the conditional probability that the individual n is rich given that the test is negative ℓ times. If the probability that an individual n is rich is known, then the above conditional probability can be computed using Bayes's rule. We will show how to do this later. A Monte Carlo algorithm does not give a definite answer. However, if ℓ is large enough (say $\ell = 100$), then the conditional probability that the property of interest holds (here, n is rich), given that the test is negative ℓ times, is very close to 1. In other words, if ℓ is large enough and if the test is negative ℓ times, then we have high confidence that n is rich.

There are two classes of primality testing algorithms:

(1) Algorithms that try all possible countexamples, and for which the test does not lie. These algorithms give a definite answer: n is prime or n is composite. Until 2002, no algorithms running in polynomial time, were known. The situation changed in 2002 when a paper with the title "PRIMES is in **P**," by Agrawal, Kayal and Saxena,

appeared on the website of the Indian Institute of Technology at Kanpur, India. In this paper, it was shown that testing for primality has a deterministic (nonrandomized) algorithm that runs in polynomial time.

We will not discuss algorithms of this type here, and instead refer the reader to Crandall and Pomerance [3] and Ribenboim [18].

(2) Randomized algorithms. To avoid having problems with infinite events, we assume that we are testing numbers in some large finite interval $\mathcal{I}$. Given any positive integer $m \in \mathcal{I}$, some candidate witness a is chosen at random. We have a test which, given m and a potential witness a, determines whether or not a is indeed a witness to the fact that m is composite. Such an algorithm is a *Monte Carlo* algorithm, which means the following:

(1) If the test is positive, then $m \in \mathcal{I}$ is composite. In terms of probabilities, this is expressed by saying that the conditional probability that $m \in \mathcal{I}$ is composite given that the test is positive is equal to 1. If we denote the event that some positive integer $m \in \mathcal{I}$ is composite by C, then we can express the above as

$$\mathsf{Pr}(C \mid \text{test is positive}) = 1.$$

(2) If $m \in \mathcal{I}$ is composite, then the test is positive for at least 50% of the choices for a. We can express the above as

$$\mathsf{Pr}(\text{test is positive} \mid C) \geq \frac{1}{2}.$$

This gives us a degree of confidence in the test.

The contrapositive of (1) says that if $m \in \mathcal{I}$ is prime, then the test is negative. If we denote by P the event that some positive integer $m \in \mathcal{I}$ is prime, then this is expressed as

$$\mathsf{Pr}(\text{test is negative} \mid P) = 1.$$

If we repeat the test ℓ times by picking independent potential witnesses, then the conditional probability that the test is negative ℓ times given that n is composite, written $\mathsf{Pr}(\text{test is negative } \ell \text{ times} \mid C)$, is given by

$$\begin{aligned}
\mathsf{Pr}(\text{test is negative } \ell \text{ times} \mid C) &= \mathsf{Pr}(\text{test is negative} \mid C)^{\ell} \\
&= (1 - \mathsf{Pr}(\text{test is positive} \mid C))^{\ell} \\
&\leq \left(1 - \frac{1}{2}\right)^{\ell} \\
&= \left(\frac{1}{2}\right)^{\ell},
\end{aligned}$$

where we used Property (2) of a Monte Carlo algorithm that

$$\mathsf{Pr}(\text{test is positive} \mid C) \geq \frac{1}{2}$$

and the independence of the trials. This confirms that if we run the algorithm ℓ times, then $\mathsf{Pr}(\text{test is negative } \ell \text{ times} \mid C)$ is very small. In other words, it is very unlikely that the test will lie ℓ times (is negative) given that the number $m \in \mathcal{I}$ is composite.

If the probabilty $\mathsf{Pr}(P)$ of the event P is known, which requires knowledge of the distribution of the primes in the interval $\mathcal{I}$, then the conditional probability

$$\mathsf{Pr}(P \mid \text{test is negative } \ell \text{ times})$$

can be determined using Bayes's rule. We do this in Section 5.4.

Our Monte Carlo algorithm does not give a definite answer. However, if ℓ is large enough (say $\ell = 100$), then the conditional probability that the number n being tested is prime given that the test is negative ℓ times, is very close to 1.

1.3 Some Tests for Compositeness

The algorithms that we will discuss test three kinds of properties:

(1) The *Fermat test*. For any odd integer $n \geq 5$, pick randomly some $a \in \{2, \ldots, n-2\}$, and test whether

$$a^{n-1} \not\equiv 1 \pmod{n}.$$

If the test is positive, then return n is composite, else n is a "probable prime."

(2) The *Miller–Rabin test*. For any odd integer $n \geq 5$, write $n - 1 = 2^k t$ with t odd and $k \geq 1$, pick randomly some $a \in \{2, \ldots, n-2\}$, and test whether

(a) $a^t \not\equiv \pm 1 \pmod{n}$, and

(b) $a^{2^i t} \not\equiv n - 1 \pmod{n}$, for all i with $1 \leq i \leq k-1$.

If the test is positive, then return n is composite, else n is a "probable prime."

(3) The *Euler test*. For any odd integer $n \geq 5$, pick randomly some $a \in \{2, \ldots, n-2\}$, and test whether

$$\left(\frac{a}{n}\right) a^{(n-1)/2} \not\equiv 1 \pmod{n}.$$

If the test is positive, then return n is composite, else n is a "probable prime." The expression $\left(\frac{a}{n}\right)$ is the *Jacobi symbol*. It is a generalization of the *Legendre symbol*. These symbols have to do with quadratic residues. Given any integer $n \geq 2$, an integer

m such that $\gcd(m, n) = 1$ is said to be a *quadratic residue* mod n (or a *square* mod n) if the congruence

$$x^2 \equiv m \pmod{n}$$

has a solution. Let p be an odd prime. For any integer m, the *Legendre symbol* $\left(\frac{m}{p}\right)$ is defined as follows:

$$\left(\frac{m}{p}\right) = \begin{cases} +1 & \text{if } m \text{ is a quadratic residue modulo } p \\ -1 & \text{if } m \text{ is a quadratic nonresidue modulo } p \\ 0 & \text{if } p \text{ divides } m. \end{cases}$$

The Jacobi symbol $\left(\frac{m}{P}\right)$ is defined for a positive odd integer $P \geq 3$ in terms of the prime factorization of P; see Definition 6.3.

The remarkable fact about the Legendre symbol is that it gives us an efficient method for testing whether a number m is a quadratic residue mod n without actually solving the congruence $x^2 \equiv m \pmod{n}$. The Jacobi symbol gives us an even more efficient method which avoids factoring. The reason is that there is an unexpected and deep relationship between the symbols $\left(\frac{p}{q}\right)$ and $\left(\frac{q}{p}\right)$, known as the *law of quadratic reciprocity*.

The law of quadratic reciprocity was conjectured by Legendre and proved by Gauss, who gave no less than seven proofs. It is one of the gems of number theory, and we will prove it in Section 6.7.

Property (1) of a Monte Carlo algorithm holds for all three tests. Next we need to show that Property (2) holds. For this, it is helpful to define the following sets of liars: for every odd composite $n \geq 3$, write $n - 1 = 2^k t$ with t odd and $k \geq 1$,

$$\begin{aligned} L_n^F &= \{a \in \{1 \leq a \leq n-1\} \mid a^{n-1} \equiv 1 \pmod{n}\}, \\ L_n^{MR} &= \{a \in \{1, \ldots, n-1\}, \text{either } a^t \equiv 1 \pmod{n}, \\ &\qquad \text{or } a^{2^i t} \equiv n-1 \pmod{n}, \text{ for some } i \text{ with } 0 \leq i \leq k-1\} \\ L_n^E &= \{a \in \{1, \ldots, n\} \mid \left(\frac{a}{n}\right) a^{(n-1)/2} \equiv 1 \pmod{n}\}. \end{aligned}$$

The set L_n^F is called the set of F-liars (Fermat liars), the set L_n^{MR} is called the set of MR-liars (Miller–Rabin liars) and the set L_n^E is called the set of E-liars (Euler liars).

It is easy to see that all three sets of liars are subsets of the multiplicative group $(\mathbb{Z}/n\mathbb{Z})^*$ of invertible elements of the ring $\mathbb{Z}/n\mathbb{Z}$. The order of this group is $\varphi(n)$, a famous function due Euler, where $\varphi(n)$ is the number of integers a with $1 \leq a \leq n$ such that $\gcd(a, n) = 1$. Obviously, $\varphi(n) < n$ if $n > 1$.

Now if we could prove that our sets of liars are proper subsets of $(\mathbb{Z}/n\mathbb{Z})^*$ of size at most $\varphi(n)/2$, then we woud be done, because the conditional probability that a is a liar given that n is composite would be at most $\varphi(n)/(2n) < 1/2$.

It turns out that both L_n^F and L_n^E are subgroups of $(\mathbb{Z}/n\mathbb{Z})^*$, but unfortunately L_n^{MR} is not closed under multiplication. If we can prove that L_n^F and L_n^E are proper subgroups of $(\mathbb{Z}/n\mathbb{Z})^*$, then by Lagrange's theorem we are done, because the order of a subgroup divides the order of the group, so a proper subgroup has order at most $\varphi(n)/2$.

Solovay and Strassen proved that L_n^E is a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$; see Theorem 6.12. This is a nice an nontrivial proof.

As to L_n^F, unfortunately there are composites n for which $L_n^F = (\mathbb{Z}/n\mathbb{Z})^*$; all numbers $a \in \{1\ldots, n-1\}$ are liars! Such trouble makers are called *Carmichael numbers*. The smallest one is $561 = 3 \times 11 \times 17$. More bad news: there are infinitely many Carmichael numbers; see Section 5.3.

The Miller–Rabin test, which is a stronger version of the Fermat test, is immune to Carmichael numbers. Indeed, even though L_n^{MR} is not a group, it is contained is a subgroup $\overline{\mathcal{S}}(n)$ of $(\mathbb{Z}/n\mathbb{Z})^*$ of the form

$$\overline{\mathcal{S}}(n) = \{a \in (\mathbb{Z}/n\mathbb{Z})^* \mid a^m \equiv \pm 1 \pmod{n}\},$$

for some suitable m (depending on n), such that m divides $n-1$. Monier and Rabin proved that the subgroup $\overline{\mathcal{S}}(n)$ is a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$, and that if $n > 9$, then the order of $\overline{\mathcal{S}}(n)$ is at most $\varphi(n)/4 \leq (n - 1/4)$; see Theorem 5.13. This is a beautiful proof that mixes combinatorial and number theoretic ideas. We also show that $L_n^{MR} \subseteq L_n^E$; see Section 6.9.

Having some powerful methods for testing for primality, we show in Chapter 2 how prime numbers can be used for public key cryptography, and in particular we present the RSA system.

The investigation of primality testing algorithms and cryptographic methods provides wonderful and strong motivations for delving more deeply into number theory.

One will quickly realize that in order to get more than a superficial understanding of randomized algorithms for primality testing, one needs to know some basic properties of groups, rings, and fields, and in particular properties of cyclic groups. In particular, the multiplicative group $(\mathbb{Z}/n\mathbb{Z})^*$ of invertible elements of the ring $\mathbb{Z}/n\mathbb{Z}$ plays an important role. It is crucial to know when the group $(\mathbb{Z}/n\mathbb{Z})^*$ is cyclic, which means that it is generated by a single element called a *primitive root*. A famous theorem of Gauss tells us that the group $(\mathbb{Z}/n\mathbb{Z})^*$ has a primitive root iff $n = 2, 4, p^m$, or $2p^m$ where p is an odd prime. We give a complete proof of this result in Sections 4.4 and 4.5. In Sections 4.1, 4.2, and 4.3, we provide a review of groups, rings, and fields.

Quadratic residues, the Legendre symbol, and the Jacobi symbol, play a crucial role in the Solovay–Strassen test. We also present a randomized algorithm for finding the square root of a number which is a quadratic residue modulo an odd prime. One of the jewels of number theory is the *law of quadratic reciprocity*, which was also proved by Gauss (in fact,

he gave seven proofs). The law of quadratic reciprocity relates the symbols $\left(\frac{p}{q}\right)$ and $\left(\frac{q}{p}\right)$, and yields a fast method to evaluate the Legendre (and the Jacobi) symbol.

Even though it is not absolutely necessary to know how to prove the law of quadratic reciprocity to understand the Solovay–Strassen test, we feel that it would be a shame not to include a proof, so we do. In fact, we give two proofs. The second one, due to Eisenstein (1845), is particularly original because is uses a trigonometric identity.

Our philosophy is that primality testing and cryptographic methods give us a great excuse to present some deep and beautiful mathematics, with an emphasis on number theory. We also believe that it is important to prove everything we state, and we (mostly) do!

Two excellent references on cryptography and its mathematical underpinnings are Hoffstein, Pipher and Silverman [8], and Shoup [21]. A more advanced treatment is given in Crandall and Pomerance's remarkable book [3] and in Ribenboim's delightful book [18]; Dietzfelbinger [4] is also very good but less encyclopedic. An easy going and delightful introduction to number theory is found in Silverman [22]. More advanced presentations are given in Apostol [1], Niven, Zuckerman, and Montgomery [16], and Ireland and Rosen [9]. Serre's book [20] is another great source for those intetested in advanced topics in number theory. For those interested in original sources, Dirichlet–Dedekind [12] is a real jewel. This book is based on a manuscript of Dirichlet but was actually written by Richard Dedekind and published in 1863 after Dirichlet's death in 1859. The English translation is by John Stillwell. The reader will be pleasantly surprised to see how clear and lively the style is, and will find a masterly exposition of many of the results from Gauss's famous *Disquisitiones Arithmeticae* [7]. Incidently, if you can get your hands on a translation of Gauss's masterpiece, you will experience what it is to be exposed to pure genius.

Sorry, we will not discuss applications of elliptic curves and lattice methods to primality testing, factoring, and cryptographic methods, in these notes. Perhaps in another set of notes ...

Chapter 2

Public Key Cryptography

2.1 Public Key Cryptography; The RSA System

Ever since written communication was used, people have been interested in trying to conceal the content of their messages from their adversaries. This has led to the development of techniques of secret communication, a science known as *cryptography*.

The basic situation is that one party, A, say Albert, wants to send a message to another party, J, say Julia. However, there is a danger that some ill-intentioned third party, Machiavelli, may intercept the message and learn things that he is not supposed to know about and as a result, do evil things. The original message, understandable to all parties, is known as the *plain text*. To protect the content of the message, Albert *encrypts* his message. When Julia receives the encrypted message, she must *decrypt* it in order to be able to read it. Both Albert and Julia share some information that Machiavelli does not have, a *key*. Without a key, Machiavelli, is incapable of decrypting the message and thus, to do harm.

There are many schemes for generating keys to encrypt and decrypt messages. We are going to describe a method involving *public and private keys* known as the *RSA Cryptosystem*, named after its inventors, Ronald Rivest, Adi Shamir, and Leonard Adleman (1978), based on ideas by Diffie and Hellman (1976). We highly recommend reading the orginal paper by Rivest, Shamir, and Adleman [19]. It is beautifully written and easy to follow. A very clear, but concise exposition can also be found in Koblitz [10]. An encyclopedic coverage of cryptography can be found in Menezes, van Oorschot, and Vanstone's *Handbook* [14].

The RSA system is widely used in practice, for example in SSL (Secure Socket Layer), which in turn is used in https (secure http). Any time you visit a "secure site" on the Internet (to read e-mail or to order merchandise), your computer generates a public key and a private key for you and uses them to make sure that your credit card number and other personal data remain secret. Interestingly, although one might think that the mathematics behind such a scheme is very advanced and complicated, this is not so. Therefore, in this section, we are going to explain the basics of RSA.

The first step is to convert the plain text of characters into an integer. This can be done

easily by assigning distinct integers to the distinct characters, for example, by converting each character to its ASCII code. From now on, we assume that this conversion has been performed.

The next and more subtle step is to use modular arithmetic. We assume that the reader has some familiarity with basic facts of arithmetic (greatest common divisors, *etc.*). A "gentle" exposition is given in Gallier [6], Chapter 5. We pick a (large) positive integer m and perform arithmetic modulo m. Let us explain this step in more detail.

Recall that for all $a, b \in \mathbb{Z}$, we write $a \equiv b \pmod{m}$ iff $a - b = km$, for some $k \in \mathbb{Z}$, and we say that *a and b are congruent modulo m*. We already know that congruence is an equivalence relation but it also satisfies the following properties.

Proposition 2.1. *For any positive integer m, for all $a_1, a_2, b_1, b_2 \in \mathbb{Z}$, the following properties hold. If $a_1 \equiv b_1 \pmod{m}$ and $a_2 \equiv b_2 \pmod{m}$, then*

(1) $a_1 + a_2 \equiv b_1 + b_2 \pmod{m}$.

(2) $a_1 - a_2 \equiv b_1 - b_2 \pmod{m}$.

(3) $a_1 a_2 \equiv b_1 b_2 \pmod{m}$.

Proof. We only check (3), leaving (1) and (2) as easy exercises. Because $a_1 \equiv b_1 \pmod{m}$ and $a_2 \equiv b_2 \pmod{m}$, we have $a_1 = b_1 + k_1 m$ and $a_2 = b_2 + k_2 m$, for some $k_1, k_2 \in \mathbb{Z}$, and so

$$a_1 a_2 = (b_1 + k_1 m)(b_2 + k_2 m) = b_1 b_2 + (b_1 k_2 + k_1 b_2 + k_1 m k_2)m,$$

which means that $a_1 a_2 \equiv b_1 b_2 \pmod{m}$. A more elegant proof consists in observing that

$$\begin{aligned} a_1 a_2 - b_1 b_2 &= a_1(a_2 - b_2) + (a_1 - b_1)b_2 \\ &= (a_1 k_2 + k_1 b_2)m, \end{aligned}$$

as claimed. □

Proposition 2.1 allows us to define addition, subtraction, and multiplication on equivalence classes modulo m.

Definition 2.1. Given any positive integer m, we denote by $\mathbb{Z}/m\mathbb{Z}$ the set of equivalence classes modulo m. If we write $\overline{a}$ for the equivalence class of $a \in \mathbb{Z}$, then we define addition, subtraction, and multiplication on residue classes as follows:

$$\begin{aligned} \overline{a} + \overline{b} &= \overline{a + b} \\ \overline{a} - \overline{b} &= \overline{a - b} \\ \overline{a}\,\overline{b} &= \overline{ab}. \end{aligned}$$

The above operations make sense because $\overline{a+b}$ does not depend on the representatives chosen in the equivalence classes $\overline{a}$ and $\overline{b}$, and similarly for $\overline{a-b}$ and $\overline{ab}$. Each equivalence class $\overline{a}$ contains a unique representative from the set of remainders $\{0, 1, \ldots, m-1\}$, modulo m, so the above operations are completely determined by $m \times m$ tables. Using the arithmetic operations of $\mathbb{Z}/m\mathbb{Z}$ is called *modular arithmetic*.

For an arbitrary m, the set $\mathbb{Z}/m\mathbb{Z}$ is an algebraic structure known as a *ring*. Addition and subtraction behave as in $\mathbb{Z}$ but multiplication is stranger. For example, when $m = 6$,

$$\begin{aligned} 2 \cdot 3 &= 0 \\ 3 \cdot 4 &= 0, \end{aligned}$$

inasmuch as $2 \cdot 3 = 6 \equiv 0 \pmod 6$, and $3 \cdot 4 = 12 \equiv 0 \pmod 6$. Therefore, it is not true that every nonzero element has a multiplicative inverse. However, it is known (see Gallier [6], Chapter 5) that a nonzero integer a has a multiplicative inverse iff $\gcd(a, m) = 1$ (use the Bézout identity). For example,

$$5 \cdot 5 = 1,$$

because $5 \cdot 5 = 25 \equiv 1 \pmod 6$.

As a consequence, when m is a prime number, every nonzero element not divisible by m has a multiplicative inverse. In this case, $\mathbb{Z}/m\mathbb{Z}$ is more like $\mathbb{Q}$; it is a finite *field*. However, note that in $\mathbb{Z}/m\mathbb{Z}$ we have

$$\underbrace{1 + 1 + \cdots + 1}_{m \text{ times}} = 0$$

(because $m \equiv 0 \pmod m$), a phenomenom that does not happen in $\mathbb{Q}$ (or $\mathbb{R}$).

The RSA method uses modular arithmetic. One of the main ingredients of public key cryptography is that one should use an encryption function, $f \colon \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$, which is easy to compute (i.e., can be computed efficiently) but such that its inverse f^{-1} is practically impossible to compute unless one has *special additional information*. Such functions are usually referred to as *trapdoor one-way functions*. Remarkably, *exponentiation modulo* m, that is, the function, $x \mapsto x^e \bmod m$, is a trapdoor one-way function for suitably chosen m and e.

Thus, we claim the following.

(1) Computing $x^e \bmod m$ can be done efficiently .

(2) Finding x such that

$$x^e \equiv y \pmod m$$

with $0 \leq x, y \leq m-1$, is hard, unless one has extra information about m. The function that finds an eth root modulo m is sometimes called a *discrete logarithm*.

We explain shortly how to compute $x^e \bmod m$ efficiently using the *square and multiply* method also known as *repeated squaring*.

As to the second claim, actually, no proof has been given yet that this function is a one-way function but, so far, this has not been refuted either.

Now, what's the trick to make it a trapdoor function?

What we do is to pick two distinct large prime numbers, p and q (say over 200 decimal digits), which are "sufficiently random" and we let

$$m = pq.$$

Next, we pick a random e, with $1 < e < (p-1)(q-1)$, relatively prime to $(p-1)(q-1)$.

Because $\gcd(e, (p-1)(q-1)) = 1$, there is some d with $1 < d < (p-1)(q-1)$, such that $ed \equiv 1 \pmod{(p-1)(q-1)}$.

Then, we claim that to find x such that

$$x^e \equiv y \pmod{m},$$

we simply compute $y^d \bmod m$, and this can be done easily, as we claimed earlier. The reason why the above "works" is that

$$x^{ed} \equiv x \pmod{m}, \qquad (*)$$

for all $x \in \mathbb{Z}$, which we prove later.

Setting up RSA

In summary to set up RSA for Albert (A) to receive encrypted messages, perform the following steps.

1. Albert generates two distinct large and sufficiently random primes, p_A and q_A. They are kept secret.

2. Albert computes $m_A = p_A q_A$. This number called the *modulus* will be made public.

3. Albert picks at random some e_A, with $1 < e_A < (p_A - 1)(q_A - 1)$, so that $\gcd(e_A, (p_A - 1)(q_A - 1)) = 1$. The number e_A is called the *encryption key* and it will also be public.

4. Albert computes the inverse, $d_A = e_A^{-1}$ modulo $(p_A - 1)(q_A - 1)$, of e_A. This number is kept secret. The pair (d_A, m_A) is Albert's *private key* and d_A is called the *decryption key*.

5. Albert publishes the pair (e_A, m_A) as his *public key*.

Encrypting a Message

Now, if Julia wants to send a message, x, to Albert, she proceeds as follows. First, she splits x into chunks, $x_1, \ldots, x_k$, each of length at most $m_A - 1$, if necessary (again, I assume that x has been converted to an integer in a preliminary step). Then she looks up Albert's public key (e_A, m_A) and she computes

$$y_i = E_A(x_i) = x_i^{e_A} \bmod m_A,$$

for $i = 1, \ldots, k$. Finally, she sends the sequence $y_1, \ldots, y_k$ to Albert. This encrypted message is known as the *cyphertext*. The function E_A is Albert's *encryption function*.

Decrypting a Message

In order to decrypt the message $y_1, \ldots, y_k$ that Julia sent him, Albert uses his private key (d_A, m_A) to compute each

$$x_i = D_A(y_i) = y_i^{d_A} \bmod m_A,$$

and this yields the sequence $x_1, \ldots, x_k$. The function D_A is Albert's *decryption function*.

Similarly, in order for Julia to receive encrypted messages, she must set her own public key (e_J, m_J) and private key (d_J, m_J) by picking two distinct primes p_J and q_J and e_J, as explained earlier.

The beauty of the scheme is that the sender only needs to know the public key of the recipient to send a message but an eavesdropper is unable to decrypt the encoded message unless he somehow gets his hands on the secret key of the receiver.

Let us give a concrete illustration of the RSA scheme using an example borrowed from Silverman [22] (Chapter 18). We write messages using only the 26 upper-case letters A, B, ..., Z, encoded as the integers A = 11, B = 12, ..., Z = 36. It would be more convenient to have assigned a number to represent a blank space but to keep things as simple as possible we do not do that.

Say Albert picks the two primes $p_A = 12553$ and $q_A = 13007$, so that $m_A = p_A q_A = 163,276,871$ and $(p_A - 1)(q_A - 1) = 163,251,312$. Albert also picks $e_A = 79921$, relatively prime to $(p_A - 1)(q_A - 1)$ and then finds the inverse d_A, of e_A modulo $(p_A - 1)(q_A - 1)$ using the extended Euclidean algorithm (more details are given in Section 2.3) which turns out to be $d_A = 145,604,785$. One can check that

$$145,604,785 \cdot 79921 - 71282 \cdot 163,251,312 = 1,$$

which confirms that d_A is indeed the inverse of e_A modulo $163,251,312$.

Now, assume that Albert receives the following message, broken in chunks of at most nine digits, because $m_A = 163,276,871$ has nine digits.

$$145387828 \qquad 47164891 \qquad 152020614 \qquad 27279275 \qquad 35356191.$$

Albert decrypts the above messages using his private key (d_A, m_A), where $d_A = 145,604,785$, using the repeated squaring method (described in Section 2.3) and finds that

$$\begin{aligned} 145387828^{145,604,785} &\equiv 30182523 \pmod{163,276,871} \\ 47164891^{145,604,785} &\equiv 26292524 \pmod{163,276,871} \\ 152020614^{145,604,785} &\equiv 19291924 \pmod{163,276,871} \\ 27279275^{145,604,785} &\equiv 30282531 \pmod{163,276,871} \end{aligned}$$

$$35356191^{145,604,785} \equiv 122215 \pmod{163,276,871}$$

which yields the message

$$30182523\ \ 26292524\ \ 19291924\ \ 30282531\ \ 122215,$$

and finally, translating each two-digit numeric code to its corresponding character, to the message

$$\text{T H O M P S O N I S I N T R O U B L E}$$

or, in more readable format

$$\text{Thompson is in trouble}$$

It would be instructive to encrypt the decoded message

$$30182523\ \ 26292524\ \ 19291924\ \ 30282531\ \ 122215$$

using the public key $e_A = 79921$. If everything goes well, we should get our original message

$$145387828 \qquad 47164891 \qquad 152020614 \qquad 27279275 \qquad 35356191$$

back.

Let us now explain in more detail how the RSA system works and why it is correct.

2.2 Correctness of The RSA System

We begin by proving the correctness of the inversion formula $(*)$. For this, we need a classical result known as *Fermat's little theorem*.

This result was first stated by Fermat in 1640 but apparently no proof was published at the time and the first known proof was given by Leibnitz (1646–1716). A different proof was given by Ivory in 1806 and this is the proof that we give here. It has the advantage that it can be easily generalized to Euler's version (1760) of Fermat's little theorem.

Theorem 2.2. *(Fermat's Little Theorem) If p is any prime number, then the following two equivalent properties hold.*

Figure 2.1: Pierre de Fermat, 1601–1665

(1) For every integer $a \in \mathbb{Z}$, if a is not divisible by p, then we have

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

(2) For every integer $a \in \mathbb{Z}$, we have

$$a^p \equiv a \pmod{p}.$$

Furthermore, (2) implies (1).

Proof. (1) Consider the integers

$$a,\ 2a,\ 3a,\ \ldots,\ (p-1)a$$

and let

$$r_1,\ r_2,\ r_3,\ \ldots,\ r_{p-1}$$

be the sequence of remainders of the division of the numbers in the first sequence by p. Because $\gcd(a, p) = 1$, none of the numbers in the first sequence is divisible by p, so $1 \leq r_i \leq p-1$, for $i = 1, \ldots, p-1$. We claim that these remainders are all distinct. If not, then say $r_i = r_j$, with $1 \leq i < j \leq p-1$. But then, because

$$ai \equiv r_i \pmod{p}$$

and

$$aj \equiv r_j \pmod{p},$$

we deduce that

$$aj - ai \equiv r_j - r_i \pmod{p},$$

and because $r_i = r_j$, we get,

$$a(j-i) \equiv 0 \pmod{p}.$$

This means that p divides $a(j-i)$, but $\gcd(a,p)=1$ so, by Euclid's proposition, p must divide $j-i$. However $1 \leq j-i < p-1$, so we get a contradiction and the remainders are indeed all distinct.

There are $p-1$ distinct remainders and they are all nonzero, therefore we must have

$$\{r_1, r_2, \ldots, r_{p-1}\} = \{1, 2, \ldots, p-1\}.$$

Using Property (3) of congruences (see Proposition 2.1), we get

$$a \cdot 2a \cdot 3a \cdots (p-1)a \equiv 1 \cdot 2 \cdot 3 \cdots (p-1) \,(\mathrm{mod}\, p);$$

that is,

$$(a^{p-1}-1) \cdot (p-1)! \equiv 0 \,(\mathrm{mod}\, p).$$

Again, p divides $(a^{p-1}-1) \cdot (p-1)!$, but because p is relatively prime to $(p-1)!$, it must divide $a^{p-1}-1$, as claimed.

(2) If $\gcd(a,p)=1$, we proved in (1) that

$$a^{p-1} \equiv 1 \,(\mathrm{mod}\, p),$$

from which we get

$$a^p \equiv a \,(\mathrm{mod}\, p),$$

because $a \equiv a (\mathrm{mod} p)$. If a is divisible by p, then $a \equiv 0 (\mathrm{mod} p)$, which implies $a^p \equiv 0 (\mathrm{mod} p)$, and thus, that

$$a^p \equiv a \,(\mathrm{mod}\, p).$$

Therefore, (2) holds for all $a \in \mathbb{Z}$ and we just proved that (1) implies (2). Finally, if (2) holds and if $\gcd(a,p)=1$, as p divides $a^p - a = a(a^{p-1}-1)$, it must divide $a^{p-1}-1$, which shows that (1) holds and so, (2) implies (1). □

It is now easy to establish the correctness of RSA.

Proposition 2.3. *For any two distinct prime numbers p and q, if e and d are any two positive integers such that*

1. $1 < e, d < (p-1)(q-1)$,
2. $ed \equiv 1 \,(\mathrm{mod}\, (p-1)(q-1))$,

then for every $x \in \mathbb{Z}$ we have

$$x^{ed} \equiv x \,(\mathrm{mod}\, pq).$$

Proof. Because p and q are two distinct prime numbers, by Euclid's proposition it is enough to prove that both p and q divide $x^{ed} - x$. We show that $x^{ed} - x$ is divisible by p, the proof of divisibility by q being similar.

By Condition (2), we have

$$ed = 1 + (p-1)(q-1)k,$$

with $k \geq 1$, inasmuch as $1 < e, d < (p-1)(q-1)$. Thus, if we write $h = (q-1)k$, we have $h \geq 1$ and

$$\begin{aligned} x^{ed} - x &\equiv x^{1+(p-1)h} - x \,(\mathrm{mod}\, p) \\ &\equiv x((x^{p-1})^h - 1) \,(\mathrm{mod}\, p) \\ &\equiv x(x^{p-1} - 1)((x^{p-1})^{h-1} + (x^{p-1})^{h-2} + \cdots + 1) \,(\mathrm{mod}\, p) \\ &\equiv (x^p - x)((x^{p-1})^{h-1} + (x^{p-1})^{h-2} + \cdots + 1) \,(\mathrm{mod}\, p) \\ &\equiv 0 \,(\mathrm{mod}\, p), \end{aligned}$$

because $x^p - x \equiv 0 \,(\mathrm{mod}\, p)$, by Fermat's little theorem. □

Remark: Of course, Proposition 2.3 holds if we allow $e = d = 1$, but this not interesting for encryption. The number $(p-1)(q-1)$ turns out to be the number of positive integers less than pq that are relatively prime to pq. For any arbitrary positive integer, m, the number of positive integers less than m that are relatively prime to m is given by the *Euler φ function* (or *Euler totient*), denoted φ (see Niven, Zuckerman, and Montgomery [16], Section 2.1, for basic properties of φ).

Fermat's little theorem can be generalized to what is known as *Euler's formula*: For every integer a, if $\gcd(a, m) = 1$, then

$$a^{\varphi(m)} \equiv 1 \,(\mathrm{mod}\, m).$$

Because $\varphi(pq) = (p-1)(q-1)$, when $\gcd(x, \varphi(pq)) = 1$, Proposition 2.3 follows from Euler's formula. However, that argument does not show that Proposition 2.3 holds when $\gcd(x, \varphi(pq)) > 1$ and a special argument is required in this case.

It can be shown that if we replace pq by a positive integer m that is square-free (does not contain a square factor) and if we assume that e and d are chosen so that $1 < e, d < \varphi(m)$ and $ed \equiv 1 \,(\mathrm{mod}\, \varphi(m))$, then

$$x^{ed} \equiv x \,(\mathrm{mod}\, m)$$

for all $x \in \mathbb{Z}$ (see Proposition 4.33).

We see no great advantage in using this fancier argument and this is why we used the more elementary proof based on Fermat's little theorem.

Proposition 2.3 immediately implies that the decrypting and encrypting RSA functions D_A and E_A are mutual inverses for any A. Furthermore, E_A is easy to compute but, without extra information, namely, the trapdoor d_A, it is practically impossible to compute $D_A = E_A^{-1}$. That D_A is hard to compute without a trapdoor is related to the fact that factoring a large number, such as m_A, into its factors p_A and q_A is hard. Today, it is practically impossible to factor numbers over 300 decimal digits long. Although no proof has been given so far, it is believed that factoring will remain a hard problem. So, even if in the next few years it becomes possible to factor 300-digit numbers, it will still be impossible to factor 400-digit numbers. RSA has the peculiar property that it depends both on the fact that primality testing is easy but that factoring is hard. What a stroke of genius!

2.3 Algorithms for Computing Powers and Inverses Modulo m

First, we explain how to compute $x^n \bmod m$ efficiently, where $n \geq 1$. Let us first consider computing the nth power x^n of some positive integer. The idea is to look at the parity of n and to proceed recursively. If n is even, say $n = 2k$, then

$$x^n = x^{2k} = (x^k)^2,$$

so, compute x^k recursively and then square the result. If n is odd, say $n = 2k + 1$, then

$$x^n = x^{2k+1} = (x^k)^2 \cdot x,$$

so, compute x^k recursively, square it, and multiply the result by x.

What this suggests is to write $n \geq 1$ in binary, say

$$n = b_\ell \cdot 2^\ell + b_{\ell-1} \cdot 2^{\ell-1} + \cdots + b_1 \cdot 2^1 + b_0,$$

where $b_i \in \{0, 1\}$ with $b_\ell = 1$ or, if we let $J = \{j \mid b_j = 1\}$, as

$$n = \sum_{j \in J} 2^j.$$

Then we have

$$x^n \equiv x^{\sum_{j \in J} 2^j} = \prod_{j \in J} x^{2^j} \bmod m.$$

This suggests computing the residues r_j such that

$$x^{2^j} \equiv r_j \,(\bmod\, m),$$

because then,

$$x^n \equiv \prod_{j \in J} r_j \,(\bmod\, m),$$

where we can compute this latter product modulo m two terms at a time.

For example, say we want to compute $999^{179} \bmod 1763$. First, we observe that

$$179 = 2^7 + 2^5 + 2^4 + 2^1 + 1,$$

and we compute the powers modulo 1763:

$$\begin{aligned}
999^{2^1} &\equiv 143 \pmod{1763}\\
999^{2^2} &\equiv 143^2 \equiv 1056 \pmod{1763}\\
999^{2^3} &\equiv 1056^2 \equiv 920 \pmod{1763}\\
999^{2^4} &\equiv 920^2 \equiv 160 \pmod{1763}\\
999^{2^5} &\equiv 160^2 \equiv 918 \pmod{1763}\\
999^{2^6} &\equiv 918^2 \equiv 10 \pmod{1763}\\
999^{2^7} &\equiv 10^2 \equiv 100 \pmod{1763}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
999^{179} &\equiv 999 \cdot 143 \cdot 160 \cdot 918 \cdot 100 \pmod{1763}\\
&\equiv 54 \cdot 160 \cdot 918 \cdot 100 \pmod{1763}\\
&\equiv 1588 \cdot 918 \cdot 100 \pmod{1763}\\
&\equiv 1546 \cdot 100 \pmod{1763}\\
&\equiv 1219 \pmod{1763},
\end{aligned}$$

and we find that

$$999^{179} \equiv 1219 \pmod{1763}.$$

Of course, it would be impossible to exponentiate 999^{179} first and then reduce modulo 1763. As we can see, the number of multiplications needed is bounded by $2\log_2 n$, which is quite good.

The above method can be implemented without actually converting n to base 2. If n is even, say $n = 2k$, then $n/2 = k$ and if n is odd, say $n = 2k + 1$, then $(n - 1)/2 = k$, so we have a way of dropping the unit digit in the binary expansion of n and shifting the remaining digits one place to the right without explicitly computing this binary expansion. Here is an algorithm for computing $x^n \bmod m$, with $n \geq 1$, using the *repeated squaring* method.

An Algorithm to Compute $x^n \bmod m$ Using Repeated Squaring

begin
$u := 1$; $a := x$;

```
  while n > 1 do
    if even(n) then e := 0 else e := 1;
    if e = 1 then u := a · u mod m;
    a := a^2 mod m; n := (n − e)/2
  endwhile;
  u := a · u mod m
end
```

The final value of u is the result. The reason why the algorithm is correct is that after j rounds through the while loop, $a = x^{2^j} \bmod m$ and

$$u = \prod_{i \in J \mid i<j} x^{2^i} \bmod m,$$

with this product interpreted as 1 when $j = 0$.

Observe that the while loop is only executed $n - 1$ times to avoid squaring once more unnecessarily and the last multiplication $a \cdot u$ is performed outside of the while loop. Also, if we delete the reductions modulo m, the above algorithm is a fast method for computing the nth power of an integer x and the time speed-up of not performing the last squaring step is more significant. We leave the details of the proof that the above algorithm is correct as an exercise.

Let us now consider the problem of computing efficiently the inverse of an integer a, modulo m, provided that $\gcd(a, m) = 1$. Full details are given in Gallier [6], Chapter 5.

The extended Euclidean algorithm can be used to find some integers x, y, such that

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

where a and b are any two positive integers. In our situation, $a = m$ and $b = a$ and we only need to find y (we would like a positive integer).

When using the Euclidean algorithm for computing $\gcd(m, a)$, with $2 \leq a < m$, we compute the following sequence of quotients and remainders.

$$\begin{aligned}
m &= aq_1 + r_1 \\
a &= r_1 q_2 + r_2 \\
r_1 &= r_2 q_3 + r_3 \\
&\vdots \\
r_{k-1} &= r_k q_{k+1} + r_{k+1} \\
&\vdots \\
r_{n-3} &= r_{n-2} q_{n-1} + r_{n-1} \\
r_{n-2} &= r_{n-1} q_n + 0,
\end{aligned}$$

with $n \geq 3$, $0 < r_1 < b$, $q_k \geq 1$, for $k = 1, \ldots, n$, and $0 < r_{k+1} < r_k$, for $k = 1, \ldots, n-2$. Observe that $r_n = 0$. If $n = 2$, we have just two divisions,

$$\begin{aligned} m &= aq_1 + r_1 \\ a &= r_1 q_2 + 0, \end{aligned}$$

with $0 < r_1 < b$, $q_1, q_2 \geq 1$, and $r_2 = 0$. Thus, it is convenient to set $r_{-1} = m$ and $r_0 = a$.

It can be shown (Gallier [6], Chapter 5) that if we set

$$\begin{aligned} x_{-1} &= 1 \\ y_{-1} &= 0 \\ x_0 &= 0 \\ y_0 &= 1 \\ x_{i+1} &= x_{i-1} - x_i q_{i+1} \\ y_{i+1} &= y_{i-1} - y_i q_{i+1}, \end{aligned}$$

for $i = 0, \ldots, n-2$, then

$$mx_{n-1} + ay_{n-1} = \gcd(m, a) = r_{n-1},$$

and so, if $\gcd(m, a) = 1$, then $r_{n-1} = 1$ and we have

$$ay_{n-1} \equiv 1 \,(\mathrm{mod}\, m).$$

Now, y_{n-1} may be greater than m or negative but we already know how to deal with that. This suggests reducing modulo m during the recurrence and we are led to the following recurrence.

$$\begin{aligned} y_{-1} &= 0 \\ y_0 &= 1 \\ z_{i+1} &= y_{i-1} - y_i q_{i+1} \\ y_{i+1} &= z_{i+1} \bmod m \quad \text{if} \quad z_{i+1} \geq 0 \\ y_{i+1} &= m - ((-z_{i+1}) \bmod m) \quad \text{if} \quad z_{i+1} < 0, \end{aligned}$$

for $i = 0, \ldots, n-2$.

It is easy to prove by induction that

$$ay_i \equiv r_i \,(\mathrm{mod}\, m)$$

for $i = 0, \ldots, n-1$ and thus, if $\gcd(a, m) > 1$, then a does not have an inverse modulo m, else

$$ay_{n-1} \equiv 1 \,(\mathrm{mod}\, m)$$

and y_{n-1} is the inverse of a modulo m such that $1 \leq y_{n-1} < m$, as desired. Note that we also get $y_0 = 1$ when $a = 1$.

We leave this proof as an exercise. Here is an algorithm.

An Algorithm for Computing the Inverse of a Modulo m

Given any natural number a with $1 \leq a < m$ and $\gcd(a, m) = 1$, the following algorithm returns the inverse of a modulo m as y.

begin
  $y := 0$; $v := 1$; $g := m$; $r := a$;
  $pr := r$; $q := \lfloor g/pr \rfloor$; $r := g - pr\,q$; (divide g by pr, to get $g = pr\,q + r$)
  if $r = 0$ **then**
    $y := 1$; $g := pr$
  else
    $r = pr$;
    while $r \neq 0$ **do**
      $pr := r$; $pv := v$;
      $q := \lfloor g/pr \rfloor$; $r := g - pr\,q$; (divide g by pr, to get $g = pr\,q + r$)
      $v := y - pv\,q$;
      if $v < 0$ **then**
        $v := m - ((-v) \bmod m)$
      else
        $v = v \bmod m$
      endif
      $g := pr$; $y := pv$
    endwhile;
  endif;
  inverse$(a) := y$
end

For example, we used the above algorithm to find that $d_A = 145,604,785$ is the inverse of $e_A = 79921$ modulo $(p_A - 1)(q_A - 1) = 163,251,312$.

The remaining issues are how to choose large random prime numbers p, q, and how to find a random number e, which is relatively prime to $(p-1)(q-1)$. For this, we rely on a deep result of number theory known as the *prime number theorem*.

2.4 Finding Large Primes; Signatures; Safety of RSA

Roughly speaking, the prime number theorem ensures that the density of primes is high enough to guarantee that there are many primes with a large specified number of digits.

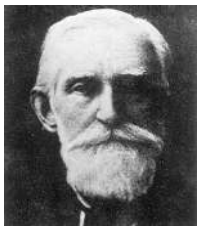


Figure 2.2: Pafnuty Lvovich Chebyshev, 1821–1894 (left), Jacques Salomon Hadamard, 1865–1963 (middle), and Charles Jean de la Vallée Poussin, 1866–1962 (right)

The relevant function is the *prime counting function* $\pi(n)$.

Definition 2.2. The *prime counting function* π is the function defined so that

$$\pi(n) = \text{number of prime numbers } p, \text{ such that } p \leq n,$$

for every natural number $n \in \mathbb{N}$.

Obviously, $\pi(0) = \pi(1) = 0$. We have $\pi(10) = 4$ because the primes no greater than 10 are $2, 3, 5, 7$ and $\pi(20) = 8$ because the primes no greater than 20 are $2, 3, 5, 7, 11, 13, 17, 19$. The growth of the function π was studied by Legendre, Gauss, Chebyshev, and Riemann between 1808 and 1859. By then, it was conjectured that

$$\pi(n) \sim \frac{n}{\ln(n)},^{1}$$

for n large, which means that

$$\lim_{n \mapsto \infty} \pi(n) \Big/ \frac{n}{\ln(n)} = 1.$$

However, a rigorous proof was not found until 1896. Indeed, in 1896, Jacques Hadamard and Charles de la Vallée-Poussin independendly gave a proof of this "most wanted theorem," using methods from complex analysis. These proofs are difficult and although more elementary proofs were given later, in particular by Erdös and Selberg (1949), those proofs are still quite hard. Thus, we content ourselves with a statement of the theorem.

Theorem 2.4. *(Prime Number Theorem) For n large, the number of primes $\pi(n)$ no larger than n is approximately equal to $n/\ln(n)$, which means that*

$$\lim_{n \mapsto \infty} \pi(n) \Big/ \frac{n}{\ln(n)} = 1.$$

[1] We use $\ln(n)$ to denote the logarithm of n to the base e, known as the *natural logarithm* of n.


Figure 2.3: Paul Erdös, 1913–1996 (left), Atle Selberg, 1917–2007 (right)

For a rather detailed account of the history of the prime number theorem (for short, *PNT*), we refer the reader to Ribenboim [18] (Chapter 4).

As an illustration of the use of the PNT, we can estimate the number of primes with 200 decimal digits. Indeed this is the difference of the number of primes up to 10^{200} minus the number of primes up to 10^{199}, which is approximately

$$\frac{10^{200}}{200 \ln 10} - \frac{10^{199}}{199 \ln 10} \approx 1.95 \cdot 10^{197}.$$

Thus, we see that there is a huge number of primes with 200 decimal digits. The number of natural numbers with 200 digits is $10^{200} - 10^{199} = 9 \cdot 10^{199}$, thus the proportion of 200-digit numbers that are prime is

$$\frac{1.95 \cdot 10^{197}}{9 \cdot 10^{199}} \approx \frac{1}{460}.$$

Consequently, among the natural numbers with 200 digits, roughly one in every 460 is a prime.

Beware that the above argument is not entirely rigorous because the prime number theorem only yields an approximation of $\pi(n)$ but sharper estimates can be used to say how large n should be to guarantee a prescribed error on the probability, say 1%.

The implication of the above fact is that if we wish to find a random prime with 200 digits, we pick at random some natural number with 200 digits and test whether it is prime. If this number is not prime, then we discard it and try again, and so on. On the average, after 460 trials, a prime should pop up,

This leads us the question: How do we test for primality?

Primality testing has also been studied for a long time. Remarkably, Fermat's little theorem yields a test for nonprimality. Indeed, if $p > 1$ fails to divide $a^{p-1} - 1$ for some natural number a, where $2 \leq a \leq p - 1$, then p cannot be a prime. The simplest a to try is $a = 2$. From a practical point of view, we can compute $a^{p-1} \bmod p$ using the method of repeated squaring and check whether the remainder is 1.

Figure 2.4: Robert Daniel Carmichael, 1879–1967

But what if p fails the Fermat test? Unfortunately, there are natural numbers p, such that p divides $2^{p-1} - 1$ and yet, p is composite. For example $p = 341 = 11 \cdot 31$ is such a number.

Actually, 2^{340} being quite big, how do we check that $2^{340} - 1$ is divisible by 341?

We just have to show that $2^{340} - 1$ is divisible by 11 and by 31. We can use Fermat's little theorem. Because 11 is prime, we know that 11 divides $2^{10} - 1$. But,

$$2^{340} - 1 = (2^{10})^{34} - 1 = (2^{10} - 1)((2^{10})^{33} + (2^{10})^{32} + \cdots + 1),$$

so $2^{340} - 1$ is also divisible by 11.

As to divisibility by 31, observe that $31 = 2^5 - 1$, and

$$2^{340} - 1 = (2^5)^{68} - 1 = (2^5 - 1)((2^5)^{67} + (2^5)^{66} + \cdots + 1),$$

so $2^{340} - 1$ is also divisible by 31.

A number p that is not a prime but behaves like a prime in the sense that p divides $2^{p-1} - 1$, is called a *pseudo-prime*. Unfortunately, the Fermat test gives a "false positive" for pseudo-primes.

Rather than simply testing whether $2^{p-1} - 1$ is divisible by p, we can also try whether $3^{p-1} - 1$ is divisible by p and whether $5^{p-1} - 1$ is divisible by p, and so on.

Unfortunately, there are composite natural numbers p, such that p divides $a^{p-1} - 1$, for all positive natural numbers a with $\gcd(a, p) = 1$. Such numbers are known as *Carmichael numbers*. The smallest Carmichael number is $p = 561 = 3 \cdot 11 \cdot 17$. The reader should try proving that, in fact, $a^{560} - 1$ is divisible by 561 for every positive natural number a, such that $\gcd(a, 561) = 1$, using the technique that we used to prove that 341 divides $2^{340} - 1$.

It turns out that there are infinitely many Carmichael numbers. Again, for a thorough introduction to primality testing, pseudo-primes, Carmichael numbers, and more, we highly recommend Ribenboim [18] (Chapter 2). An excellent (but more terse) account is also given in Koblitz [10] (Chapter V).

Still, what do we do about the problem of false positives? The key is to switch to *probabilistic methods*. Indeed, if we can design a method that is guaranteed to give a false positive with probablity less than 0.5, then we can repeat this test for randomly chosen as and reduce the probability of false positive considerably. For example, if we repeat the experiment 100 times, the probability of false positive is less than $2^{-100} < 10^{-30}$. This is probably less than the probability of hardware failure.

Various probabilistic methods for primality testing have been designed. One of them is the Miller–Rabin test, another the APR test, and yet another the Solovay–Strassen test. Since 2002, it has been known that primality testing can be done in polynomial time. This result is due to Agrawal, Kayal, and Saxena and known as the AKS test solved a long-standing problem; see Dietzfelbinger [4] and Crandall and Pomerance [3] (Chapter 4). Remarkably, Agrawal and Kayal worked on this problem for their senior project in order to complete their bachelor's degree. It remains to be seen whether this test is really practical for very large numbers.

A very important point to make is that these primality testing methods *do not* provide a factorization of m when m is composite. This is actually a crucial ingredient for the security of the RSA scheme. So far, it appears (and it is hoped) that *factoring* an integer is a much harder problem than testing for primality and all known methods are incapable of factoring natural numbers with over 300 decimal digits (it would take centuries).

For a comprehensive exposition of the subject of primality-testing, we refer the reader to Crandall and Pomerance [3] (Chapters 3 and 4) and again, to Ribenboim [18] (Chapter 2) and Koblitz [10] (Chapter V). We give a thorough presentation of the Miller–Rabin and the Solovay–Strassen tests in Chapters 5 and 6 (with complete proofs).

Going back to the RSA method, we now have ways of finding the large random primes p and q by picking at random some 200-digit numbers and testing for primality. Rivest, Shamir, and Adleman also recommend to pick p and q so that they differ by a few decimal digits, that both $p-1$ and $q-1$ should contain large prime factors and that $\gcd(p-1, q-1)$ should be small. The public key, e, relatively prime to $(p-1)(q-1)$ can also be found by a similar method: Pick at random a number, $e < (p-1)(q-1)$, which is large enough (say, greater than $\max\{p, q\}$) and test whether $\gcd(e, (p-1)(q-1)) = 1$, which can be done quickly using the extended Euclidean algorithm. If not, discard e and try another number, and so on. It is easy to see that such an e will be found in no more trials than it takes to find a prime; see Lovász, Pelikán, and Vesztergombi [13] (Chapter 15), which contains one of the simplest and clearest presentations of RSA that we know of. Koblitz [10] (Chapter IV) also provides some details on this topic as well as Menezes, van Oorschot, and Vanstone's *Handbook* [14].

If Albert receives a message coming from Julia, how can he be sure that this message does not come from an imposter? Just because the message is signed "Julia" does not mean that it comes from Julia; it could have been sent by someone else pretending to be Julia, inasmuch as all that is needed to send a message to Albert is Albert's public key, which is known to everybody. This leads us to the issue of *signatures*.

There are various schemes for adding a signature to an encrypted message to ensure that the sender of a message is really who he or she claims to be (with a high degree of confidence). The trick is to make use of the sender's keys. We propose two scenarios.

1. The sender, Julia, encrypts the message x to be sent with *her own private key*, (d_J, m_J), creating the message $D_J(x) = y_1$. Then, Julia adds her signature, "Julia", at the end of the message y_1, encrypts the message "y_1 Julia" using *Albert's public key*, (e_A, m_A), creating the message $y_2 = E_A(y_1 \text{ Julia})$, and finally sends the message y_2 to Albert.

   When Albert receives the encrypted message y_2 claiming to come from *Julia*, first he decrypts the message using *his private key* (d_A, m_A). He will see an encrypted message, $D_A(y_2) = y_1$ Julia, with the legible signature, *Julia*. He will then delete the signature from this message and decrypt the message y_1 using *Julia's public key* (e_J, m_J), getting $x = E_J(y_1)$. Albert will know whether someone else faked this message if the result is garbage. Indeed, only Julia could have encrypted the original message x with her private key, which is only known to her. An eavesdropper who is pretending to be Julia would not know Julia's private key and so, would not have encrypted the original message to be sent using Julia's secret key.

2. The sender, Julia, first adds her signature, "Julia", to the message x to be sent and then, she encrypts the message "x Julia" with *Albert's public key* (e_A, m_A), creating the message $y_1 = E_A(x \text{ Julia})$. Julia also encrypts the original message x using *her private key* (d_J, m_J) creating the message $y_2 = D_J(x)$, and finally she sends the pair of messages (y_1, y_2).

   When Albert receives a pair of messages (y_1, y_2), claiming to have been sent by Julia, first Albert decrypts y_1 using *his private key* (d_A, m_A), getting the message $D_A(y_1) = x$ Julia. Albert finds the signature, Julia, and then decrypts y_2 using *Julia's public key* (e_J, m_J), getting the message $x' = E_J(y_2)$. If $x = x'$, then Albert has serious assurance that the sender is indeed Julia and not an imposter.

The last topic that we would like to discuss is the *security* of the RSA scheme. This is a difficult issue and many researchers have worked on it. As we remarked earlier, the security of RSA hinges on the fact that factoring is hard. It has been shown that if one has a method for breaking the RSA scheme (namely, to find the secret key d), then there is a probabilistic method for finding the factors p and q, of $m = pq$ (see Koblitz [10], Chapter IV, Section 2, or Menezes, van Oorschot, and Vanstone [14], Section 8.2.2). If p and q are chosen to be large enough, factoring $m = pq$ will be practically impossible and so it is unlikely that RSA can be cracked. However, there may be other attacks and, at present, there is no proof that RSA is fully secure.

Observe that because $m = pq$ is known to everybody, if somehow one can learn $N = (p-1)(q-1)$, then p and q can be recovered. Indeed $N = (p-1)(q-1) = pq - (p+q) + 1 =$

$m - (p + q) + 1$ and so,

$$\begin{aligned} pq &= m \\ p + q &= m - N + 1, \end{aligned}$$

and p and q are the roots of the quadratic equation

$$X^2 - (m - N + 1)X + m = 0.$$

Thus, a line of attack is to try to find the value of $(p - 1)(q - 1)$. For more on the security of RSA, see Menezes, van Oorschot, and Vanstone's *Handbook* [14].

Chapter 3

Primality Testing Using Randomized Algorithms; Introduction

In article 329 of his famous *Disquisitiones Arithmeticae* [7] (published in 1801, when he was 24 years old), C.F. Gauss writes (in Latin!):

> "The problem of distinguishing prime numbers from composite numbers and resolving the latter into their prime factors is known to be one of the most important and useful in arithmetic. It has engaged the industry and wisdom of ancient and moderm geometers to such an extent that it would be superfluous to discuss the problem at length. Nevertherless we must confess that all methods that have been proposed thus far are either restricted to very special cases or are so laborious and difficult that even for numbers that do not exceed the limits of tables constructed by estimable men, they try the patience of even the practiced calculator. And these methods do not apply at all to larger numbers ... The techniques that were previously known would require intolerable labor even for the most indefatigable calculator."

The problem of determining whether a given integer is prime is one of the better known and most easily understood problems of pure mathematics. This problem has caught the interest of mathematicians again and again for centuries. However, it was not until the 20th century that questions about primality testing and factoring were recognized as problems of practical importance, and a central part of applied mathematics. The advent of cryptographic systems that use large primes, such as RSA, was the main driving force for the development of fast and reliable methods for primality testing. Indeed, as we saw in earlier sections of these notes, in order to create RSA keys, one needs to produce large prime numbers. How do we do that?

One method is to produce a random string of digits (say of 200 digits), and then to test whether this number is prime or not. As we explained earlier, by the Prime Number Theorem, among the natural numbers with 200 digits, roughly one in every 460 is a prime. Thus, it should take at most 460 trials (picking at random some natural number with 200

digits) before a prime shows up. Note that we need a mechanism to generate random numbers, an interesting and tricky problem, but for now, we postpone discussing random number generation.

It remains to find methods for testing an integer for primality, and perhaps for factoring composite numbers.

In 1903, at the meeting of the American Mathematical Society, F.N. Cole came to the blackboard and, without saying a word, wrote down

$$2^{67} - 1 = 147573952589676412927 = 193707721 \times 761838257287,$$

and then used long multiplication to multiply the two numbers on the right-hand side to prove that he was indeed correct. Afterwards, he said that figuring this out had taken him "three years of Sundays." Too bad laptops did not exist in 1903.

The moral of this tale is that *checking* that a number is composite can be done quickly (that is, in polynomial time), but *finding* a factorization is hard. In general, it requires an exhaustive search. Another important observation is that most efficient tests for compositeness *do not* produce a factorization. For example, Lucas had already shown that $2^{67} - 1$ is composite, but without finding a factor.

In fact, although this has not been proved, factoring appears to be a much harder problem than primality testing, which is a good thing since the safety of many cryptographic systems depends on the assumption that factoring is hard!

As we explained in the introduction, most algorithms for testing whether an integer n is prime actually test for compositeness. This is because tests for compositeness usually try to find a counterexample to some property, say A, implied by primality. If such a counterexample can be guessed, then it is cheap to check that property A fails, and then we know for sure that n is composite. We also have a *witness* (or certificate) that n is composite. If the algorithm fails to show that n is composite, does this imply that n is prime? Unfortunately, no. This is because, in general, the algorithm has not tested all potential countexamples. So, how do we fix the algorithm?

One possibility is to try systematically all potential countexamples. If the algorithm fails on all counterexamples, then the number n has to be prime. The problem with this approach is that the number of counterexamples is generally too big, and this method is not practical. Methods of this kind are presented in Crandall and Pomerance [3] and Ribenboim [18].

Another approach is to use a randomized algorithm. Typically, a counterexample is some number a randomly chosen from the set $\{2, \ldots, n-2\}$, and the algorithm performs a test on a and n to determine whether a is a counterexample. If the test is positive, then for sure n is composite, and a is a witness to the fact that n is composite. If the test is negative, then the algorithm does not find n to be composite, and we can call it again several times, each time picking (independently from previous trials) another random number a. If the algorithm ever reports a positive test, then for sure n is a composite. But what if we call the

algorithm say 20 times, and every time the test is negative (which means that the algorithm does not find n to be composite 20 times). Can we be sure that n is a prime?

Not necessarily, because the test performed by the algorithm may not be 100% reliable. If n is prime, the test performed by the algorithm on every a is negative (as it should), but there may be some composite n and some a for which the test is negative. Such a number a is called a *liar*, because it fools the test. Even though n is composite, a does not trigger the test to be positive, to indicate that n is indeed composite. But if the conditional probability that the test performed by the algorithm is positive given that n is composite is large enough, say at least $1/2$, then it can be shown that the conditional probability that n is composite, given that the test performed by the algorithm is negative 20 times, is less than $\ln(n) \cdot (1/2)^{20}$ (see Section 5.4).[1] In summary, if we run the algorithm ℓ times (for ℓ large enough, say $\ell = 100$) on some number n, and if each time the test performed by the algorithm is negative, then we can be very confident that n is prime. Such kind of randomized algorithm is called a *Monte Carlo algorithm*.

Several randomized algorithms for primality testing have been designed, including the Miller–Rabin and the Solovay–Strassen tests, to be discussed in Chapters 5 and 6. Then, in the summer of 2002, a paper with the title "PRIMES is in **P**," by Agrawal, Kayal and Saxena, appeared on the website of the Indian Institute of Technology at Kanpur, India. In this paper, it was shown that testing for primality has a deterministic (nonrandomized) algorithm that runs in polynomial time. Finally, the long-standing open problem of "deciding whether primality testing is in **P**" was settled in this amazing paper, by an algorithm usually referred to as the *AKS algorithm*. We will not discuss this algorithm in these notes (but, perhaps in another set of notes ...).

[1] Recall that we use $\ln(n)$ to denote the logarithm of n to the base e, known as the *natural logarithm* of n.

Chapter 4

Basic Facts About Groups, Rings, Fields, and Number Theory

This chapter provides a review of the mathematical background needed to thoroughly understand the randomized algorithms for primality testing presented in the following chapters, especially the proofs. This includes some basics on groups, the structure of cyclic groups, rings, fields, and finite fields. The multiplicative groups $(\mathbb{Z}/n\mathbb{Z})^*$ of invertible elements of the rings $\mathbb{Z}/n\mathbb{Z}$ play a particularly important role. It is crucial to know when the group $(\mathbb{Z}/n\mathbb{Z})^*$ is cyclic, which means that it is generated by a single element called a *primitive root*. A famous result due to Gauss says that the group $(\mathbb{Z}/n\mathbb{Z})^*$ has a primitive root iff $n = 2, 4, p^m$, or $2p^m$ where p is an odd prime. We give a complete proof of this result in Sections 4.4 and 4.5.

Readers familiar with groups, rings and fields should probably skip Sections 4.1, 4.2, and 4.3. However, the reader may want to read Sections 4.4 and 4.5, skipping proofs the first time, before reading Chapter 5. The material in these two sections is classical and very beautiful. Similarly, the reader may want to read Section 4.7, omitting proofs the first time, before reading Chapter 6.

4.1 Groups, Subgroups, Cosets

Definition 4.1. A *group* is a set G equipped with a binary operation $\cdot\colon G \times G \to G$ that associates an element $a \cdot b \in G$ to every pair of elements $a, b \in G$, and having the following properties: $\cdot$ is associative, has an identity element $e \in G$, and every element in G is invertible (w.r.t. $\cdot$). More explicitly, this means that the following equations hold for all $a, b, c \in G$:

(G1) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$. (associativity);

(G2) $a \cdot e = e \cdot a = a$. (identity);

(G3) For every $a \in G$, there is some $a^{-1} \in G$ such that $a \cdot a^{-1} = a^{-1} \cdot a = e$. (inverse).

A group G is *abelian* (or *commutative*) if

$$a \cdot b = b \cdot a \quad \text{for all } a, b \in G.$$

A set M together with an operation $\cdot\colon M \times M \to M$ and an element e satisfying only Conditions (G1) and (G2) is called a *monoid*. For example, the set $\mathbb{N} = \{0, 1, \ldots, n, \ldots\}$ of natural numbers is a (commutative) monoid under addition. However, it is not a group.

Some examples of groups are given below.

Example 4.1.

1. The set $\mathbb{Z} = \{\ldots, -n, \ldots, -1, 0, 1, \ldots, n, \ldots\}$ of integers is an abelian group under addition, with identity element 0. However, $\mathbb{Z}^* = \mathbb{Z} - \{0\}$ is not a group under multiplication.

2. The set $\mathbb{Q}$ of rational numbers (fractions p/q with $p, q \in \mathbb{Z}$ and $q \neq 0$) is an abelian group under addition, with identity element 0. The set $\mathbb{Q}^* = \mathbb{Q} - \{0\}$ is also an abelian group under multiplication, with identity element 1.

3. Given any nonempty set S, the set of bijections $f\colon S \to S$, also called *permutations of* S, is a group under function composition (i.e., the multiplication of f and g is the composition $g \circ f$), with identity element the identity function id_S. This group is not abelian as soon as S has more than two elements. The permutation group of the set $S = \{1, \ldots, n\}$ is often denoted $\mathfrak{S}_n$ and called the *symmetric group* on n elements.

4. For any natural number $n \geq 1$, the set $\mathbb{Z}/n\mathbb{Z}$ of residues modulo n as in Definition 2.1 is an abelian group under addition modulo n.

5. The set of $n \times n$ invertible matrices with real (or complex) coefficients is a group under matrix multiplication, with identity element the identity matrix I_n. This group is called the *general linear group* and is usually denoted by $\mathbf{GL}(n, \mathbb{R})$ (or $\mathbf{GL}(n, \mathbb{C})$).

6. The set of $n \times n$ invertible matrices A with real (or complex) coefficients such that $\det(A) = 1$ is a group under matrix multiplication, with identity element the identity matrix I_n. This group is called the *special linear group* and is usually denoted by $\mathbf{SL}(n, \mathbb{R})$ (or $\mathbf{SL}(n, \mathbb{C})$).

7. The set of $n \times n$ matrices Q with real coefficients such that

$$QQ^\top = Q^\top Q = I_n$$

is a group under matrix multiplication, with identity element the identity matrix I_n; we have $Q^{-1} = Q^\top$. This group is called the *orthogonal group* and is usually denoted by $\mathbf{O}(n)$.

8. The set of $n \times n$ invertible matrices Q with real coefficients such that

$$QQ^\top = Q^\top Q = I_n \quad \text{and} \quad \det(Q) = 1$$

is a group under matrix multiplication, with identity element the identity matrix I_n; as in (6), we have $Q^{-1} = Q^\top$. This group is called the *special orthogonal group* or *rotation group* and is usually denoted by $\mathbf{SO}(n)$.

The groups in (5)–(8) are nonabelian for $n \geq 2$, except for $\mathbf{SO}(2)$ which is abelian (but $\mathbf{O}(2)$ is not abelian).

It is customary to denote the operation of an abelian group G by $+$, in which case the inverse a^{-1} of an element $a \in G$ is denoted by $-a$.

The identity element of a group is *unique*. In fact, we can prove a more general fact:

Proposition 4.1. *If a binary operation $\cdot : M \times M \to M$ is associative and if $e' \in M$ is a left identity and $e'' \in M$ is a right identity, which means that*

$$e' \cdot a = a \quad \text{for all} \quad a \in M \tag{G2l}$$

and

$$a \cdot e'' = a \quad \text{for all} \quad a \in M, \tag{G2r}$$

then $e' = e''$.

Proof. If we let $a = e''$ in equation (G2l), we get

$$e' \cdot e'' = e'',$$

and if we let $a = e'$ in equation (G2r), we get

$$e' \cdot e'' = e',$$

and thus

$$e' = e' \cdot e'' = e'',$$

as claimed. □

Proposition 4.1 implies that the identity element of a monoid is unique, and since every group is a monoid, the identity element of a group is unique. Furthermore, every element in a group has a *unique inverse*. This is a consequence of a slightly more general fact:

Proposition 4.2. *In a monoid M with identity element e, if some element $a \in M$ has some left inverse $a' \in M$ and some right inverse $a'' \in M$, which means that*

$$a' \cdot a = e \tag{G3l}$$

and

$$a \cdot a'' = e, \tag{G3r}$$

then $a' = a''$.

Proof. Using (G3l) and the fact that e is an identity element, we have

$$(a' \cdot a) \cdot a'' = e \cdot a'' = a''.$$

Similarly, Using (G3r) and the fact that e is an identity element, we have

$$a' \cdot (a \cdot a'') = a' \cdot e = a'.$$

However, since M is monoid, the operation $\cdot$ is associative, so

$$a' = a' \cdot (a \cdot a'') = (a' \cdot a) \cdot a'' = a'',$$

as claimed. □

Remark: Axioms (G2) and (G3) can be weakened a bit by requiring only (G2r) (the existence of a right identity) and (G3r) (the existence of a right inverse for every element) (or (G2l) and (G3l)). It is a good exercise to prove that the group axioms (G2) and (G3) follow from (G2r) and (G3r).

Definition 4.2. If a group G has a finite number n of elements, we say that G is a group of *order* n. If G is infinite, we say that G has *infinite order*. The order of a group is usually denoted by $|G|$ (if G is finite).

Given a group G, for any two subsets $R, S \subseteq G$, we let

$$RS = \{r \cdot s \mid r \in R,\ s \in S\}.$$

In particular, for any $g \in G$, if $R = \{g\}$, we write

$$gS = \{g \cdot s \mid s \in S\},$$

and similarly, if $S = \{g\}$, we write

$$Rg = \{r \cdot g \mid r \in R\}.$$

From now on, we will drop the multiplication sign and write $g_1 g_2$ for $g_1 \cdot g_2$.

Definition 4.3. Let G be a group. For any $g \in G$, define L_g, the *left translation by* g, by $L_g(a) = ga$, for all $a \in G$, and R_g, the *right translation by* g, by $R_g(a) = ag$, for all $a \in G$.

The following simple fact is often used.

Proposition 4.3. *Given a group G, the translations L_g and R_g are bijections.*

Proof. We show this for L_g, the proof for R_g being similar.

If $L_g(a) = L_g(b)$, then $ga = gb$, and multiplying on the left by g^{-1}, we get $a = b$, so L_g injective. For any $b \in G$, we have $L_g(g^{-1}b) = gg^{-1}b = b$, so L_g is surjective. Therefore, L_g is bijective. □

Definition 4.4. Given a group G, a subset H of G is a *subgroup of* G iff

(1) The identity element e of G also belongs to H ($e \in H$);

(2) For all $h_1, h_2 \in H$, we have $h_1 h_2 \in H$;

(3) For all $h \in H$, we have $h^{-1} \in H$.

The proof of the following proposition is left as an exercise.

Proposition 4.4. *Given a group G, a subset $H \subseteq G$ is a subgroup of G iff H is nonempty and whenever $h_1, h_2 \in H$, then $h_1 h_2^{-1} \in H$.*

If the group G is finite, then the following criterion can be used.

Proposition 4.5. *Given a finite group G, a subset $H \subseteq G$ is a subgroup of G iff*

(1) $e \in H$;

(2) H is closed under multiplication.

Proof. We just have to prove that Condition (3) of Definition 4.4 holds. For any $a \in H$, since the left translation L_a is bijective, its restriction to H is injective, and since H is finite, it is also bijective. Since $e \in H$, there is a unique $b \in H$ such that $L_a(b) = ab = e$. However, if a^{-1} is the inverse of a in G, we also have $L_a(a^{-1}) = aa^{-1} = e$, and by injectivity of L_a, we have $a^{-1} = b \in H$. □

Example 4.2.

1. For any integer $n \in \mathbb{Z}$, the set
$$n\mathbb{Z} = \{nk \mid k \in \mathbb{Z}\}$$
is a subgroup of the group $\mathbb{Z}$.

2. The set of matrices
$$\mathbf{GL}^+(n, \mathbb{R}) = \{A \in \mathbf{GL}(n, \mathbb{R}) \mid \det(A) > 0\}$$
is a subgroup of the group $\mathbf{GL}(n, \mathbb{R})$.

3. The group $\mathbf{SL}(n, \mathbb{R})$ is a subgroup of the group $\mathbf{GL}(n, \mathbb{R})$.

4. The group $\mathbf{O}(n)$ is a subgroup of the group $\mathbf{GL}(n, \mathbb{R})$.

5. The group $\mathbf{SO}(n)$ is a subgroup of the group $\mathbf{O}(n)$, and a subgroup of the group $\mathbf{SL}(n, \mathbb{R})$.

6. It is not hard to show that every 2×2 rotation matrix $R \in \mathbf{SO}(2)$ can be written as

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad \text{with} \quad 0 \leq \theta < 2\pi.$$

   Then $\mathbf{SO}(2)$ can be considered as a subgroup of $\mathbf{SO}(3)$ by viewing the matrix

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

   as the matrix

$$Q = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

7. The set of 2×2 upper-triangular matrices of the form

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \quad a, b, c \in \mathbb{R},\ a, c \neq 0$$

   is a subgroup of the group $\mathbf{GL}(2, \mathbb{R})$.

8. The set V consisting of the four matrices

$$\begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$$

   is a subgroup of the group $\mathbf{GL}(2, \mathbb{R})$ called the *Klein four-group*.

Definition 4.5. If H is a subgroup of G and $g \in G$ is any element, the sets of the form gH are called *left cosets of H in G* and the sets of the form Hg are called *right cosets of H in G*. The left cosets (resp. right cosets) of H induce an equivalence relation $\sim$ defined as follows: For all $g_1, g_2 \in G$,

$$g_1 \sim g_2 \quad \text{iff} \quad g_1 H = g_2 H$$

(resp. $g_1 \sim g_2$ iff $Hg_1 = Hg_2$). Obviously, $\sim$ is an equivalence relation.

Now, we claim the following fact:

Proposition 4.6. *Given a group G and any subgroup H of G, we have $g_1 H = g_2 H$ iff $g_2^{-1} g_1 H = H$ iff $g_2^{-1} g_1 \in H$, for all $g_1, g_2 \in G$.*

Proof. If we apply the bijection $L_{g_2^{-1}}$ to both $g_1 H$ and $g_2 H$ we get $L_{g_2^{-1}}(g_1 H) = g_2^{-1} g_1 H$ and $L_{g_2^{-1}}(g_2 H) = H$, so $g_1 H = g_2 H$ iff $g_2^{-1} g_1 H = H$. If $g_2^{-1} g_1 H = H$, since $1 \in H$, we get $g_2^{-1} g_1 \in H$. Conversely, if $g_2^{-1} g_1 \in H$, since H is a group, the left translation $L_{g_2^{-1} g_1}$ is a bijection of H, so $g_2^{-1} g_1 H = H$. Thus, $g_2^{-1} g_1 H = H$ iff $g_2^{-1} g_1 \in H$. □

It follows that the equivalence class of an element $g \in G$ is the coset gH (resp. Hg). Since L_g is a bijection between H and gH, the cosets gH all have the same cardinality. The map $L_{g^{-1}} \circ R_g$ is a bijection between the left coset gH and the right coset Hg, so they also have the same cardinality. Since the distinct cosets gH form a partition of G, we obtain the following fact:

Proposition 4.7. *(Lagrange) For any finite group G and any subgroup H of G, the order h of H divides the order n of G.*

Definition 4.6. Given a finite group G and a subgroup H of G, if $n = |G|$ and $h = |H|$, then the ratio n/h is denoted by $(G : H)$ and is called the *index of H in G*.

The index $(G : H)$ is the number of left (and right) cosets of H in G. Proposition 4.7 can be stated as

$$|G| = (G : H)|H|.$$

The set of left cosets of H in G (which, in general, is **not** a group) is denoted G/H. The "points" of G/H are obtained by "collapsing" all the elements in a coset into a single element.

Example 4.3.

1. Let n be any positive integer, and consider the subgroup $n\mathbb{Z}$ of $\mathbb{Z}$ (under addition). The coset of 0 is the set $\{0\}$, and the coset of any nonzero integer $m \in \mathbb{Z}$ is

   $$m + n\mathbb{Z} = \{m + nk \mid k \in \mathbb{Z}\}.$$

   By dividing m by n, we have $m = nq + r$ for some unique r such that $0 \leq r \leq n - 1$. But then we see that r is the smallest positive element of the coset $m + n\mathbb{Z}$. This implies that there is a bijection betwen the cosets of the subgroup $n\mathbb{Z}$ of $\mathbb{Z}$ and the set of residues $\{0, 1, \ldots, n - 1\}$ modulo n, or equivalently a bijection with $\mathbb{Z}/n\mathbb{Z}$.

2. The cosets of $\mathbf{SL}(n, \mathbb{R})$ in $\mathbf{GL}(n, \mathbb{R})$ are the sets of matrices

   $$A\,\mathbf{SL}(n, \mathbb{R}) = \{AB \mid B \in \mathbf{SL}(n, \mathbb{R})\},\ A \in \mathbf{GL}(n, \mathbb{R}).$$

   Since A is invertible, $\det(A) \neq 0$, and we can write $A = (\det(A))^{1/n}((\det(A))^{-1/n}A)$ if $\det(A) > 0$ and $A = (-\det(A))^{1/n}((-\det(A))^{-1/n}A)$ if $\det(A) < 0$. But we have $(\det(A))^{-1/n}A \in \mathbf{SL}(n, \mathbb{R})$ if $\det(A) > 0$ and $-(-\det(A))^{-1/n}A \in \mathbf{SL}(n, \mathbb{R})$ if $\det(A) < 0$, so the coset $A\,\mathbf{SL}(n, \mathbb{R})$ contains the matrix

   $$(\det(A))^{1/n} I_n \quad \text{if} \quad \det(A) > 0, \quad -(-\det(A))^{1/n} I_n \quad \text{if} \quad \det(A) < 0.$$

   It follows that there is a bijection between the cosets of $\mathbf{SL}(n, \mathbb{R})$ in $\mathbf{GL}(n, \mathbb{R})$ and $\mathbb{R}$.

3. The cosets of $\mathbf{SO}(n)$ in $\mathbf{GL}^+(n,\mathbb{R})$ are the sets of matrices

$$A\,\mathbf{SO}(n) = \{AQ \mid Q \in \mathbf{SO}(n)\},\ A \in \mathbf{GL}^+(n,\mathbb{R}).$$

It can be shown (using the polar form for matrices) that there is a bijection between the cosets of $\mathbf{SO}(n)$ in $\mathbf{GL}^+(n,\mathbb{R})$ and the set of $n \times n$ symmetric, positive, definite matrices; these are the symmetric matrices whose eigenvalues are strictly positive.

4. The cosets of $\mathbf{SO}(2)$ in $\mathbf{SO}(3)$ are the sets of matrices

$$Q\,\mathbf{SO}(2) = \{QR \mid R \in \mathbf{SO}(2)\},\ Q \in \mathbf{SO}(3).$$

The group $\mathbf{SO}(3)$ moves the points on the sphere S^2 in $\mathbb{R}^3$, namely for any $x \in S^2$,

$$x \mapsto Qx \quad \text{for any rotation } Q \in \mathbf{SO}(3).$$

Here,

$$S^2 = \{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}.$$

Let $N = (0,0,1)$ be the north pole on the sphere S^2. Then it is not hard to show that $\mathbf{SO}(2)$ is precisely the subgroup of $\mathbf{SO}(3)$ that leaves N fixed. As a consequence, all rotations QR in the coset $Q\,\mathbf{SO}(2)$ map N to the same point $QN \in S^2$, and it can be shown that there is a bijection between the cosets of $\mathbf{SO}(2)$ in $\mathbf{SO}(3)$ and the points on S^2. The surjectivity of this map has to do with the fact that the action of $\mathbf{SO}(3)$ on S^2 is transitive, which means that for any point $x \in S^2$, there is some rotation $Q \in \mathbf{SO}(3)$ such that $QN = x$.

It is tempting to define a multiplication operation on left cosets (or right cosets) by setting

$$(g_1H)(g_2H) = (g_1g_2)H,$$

but this operation is not well defined in general, unless the subgroup H possesses a special property. In Example 4.3, it is possible to define multiplication of cosets in (1), but it is not possible in (2) and (3).

The property of the subgroup H that allows defining a multiplication operation on left cosets is typical of the kernels of group homomorphisms, so we are led to the following definition.

Definition 4.7. Given any two groups G and G', a function $\varphi\colon G \to G'$ is a *homomorphism* iff

$$\varphi(g_1g_2) = \varphi(g_1)\varphi(g_2), \quad \text{for all } g_1, g_2 \in G.$$

Taking $g_1 = g_2 = e$ (in G), we see that

$$\varphi(e) = e',$$

and taking $g_1 = g$ and $g_2 = g^{-1}$, we see that

$$\varphi(g^{-1}) = (\varphi(g))^{-1}.$$

Example 4.4.

1. The map $\varphi\colon \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ given by $\varphi(m) = m \bmod n$ for all $m \in \mathbb{Z}$ is a homomorphism.
2. The map $\det\colon \mathbf{GL}(n,\mathbb{R}) \to \mathbb{R}$ is a homomorphism because $\det(AB) = \det(A)\det(B)$ for any two matrices A, B. Similarly, the map $\det\colon \mathbf{O}(n) \to \mathbb{R}$ is a homomorphism.

If $\varphi\colon G \to G'$ and $\psi\colon G' \to G''$ are group homomorphisms, then $\psi \circ \varphi\colon G \to G''$ is also a homomorphism. If $\varphi\colon G \to G'$ is a homomorphism of groups, and if $H \subseteq G$, $H' \subseteq G'$ are two subgroups, then it is easily checked that

$$\text{Im } H = \varphi(H) = \{\varphi(g) \mid g \in H\}$$

is a subgroup of G' and

$$\varphi^{-1}(H') = \{g \in G \mid \varphi(g) \in H'\}$$

is a subgroup of G. In particular, when $H' = \{e'\}$, we obtain the *kernel*, Ker φ, of φ.

Definition 4.8. If $\varphi\colon G \to G'$ is a homomorphism of groups, and if $H \subseteq G$ is a subgroup of G, then the subgroup of G',

$$\text{Im } H = \varphi(H) = \{\varphi(g) \mid g \in H\},$$

is called the *image of H by φ*, and the subgroup of G,

$$\text{Ker } \varphi = \{g \in G \mid \varphi(g) = e'\},$$

is called the *kernel* of φ.

Example 4.5.

1. The kernel of the homomorphism $\varphi\colon \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ is $n\mathbb{Z}$.
2. The kernel of the homomorphism $\det\colon \mathbf{GL}(n,\mathbb{R}) \to \mathbb{R}$ is $\mathbf{SL}(n,\mathbb{R})$. Similarly, the kernel of the homomorphism $\det\colon \mathbf{O}(n) \to \mathbb{R}$ is $\mathbf{SO}(n)$.

The following characterization of the injectivity of a group homomorphism is used all the time.

Proposition 4.8. *If $\varphi\colon G \to G'$ is a homomorphism of groups, then $\varphi\colon G \to G'$ is injective iff* Ker $\varphi = \{e\}$. *(We also write* Ker $\varphi = (0)$.*)*

Proof. Assume φ is injective. Since $\varphi(e) = e'$, if $\varphi(g) = e'$, then $\varphi(g) = \varphi(e)$, and by injectivity of φ we must have $g = e$, so Ker $\varphi = \{e\}$.

Conversely, assume that Ker $\varphi = \{e\}$. If $\varphi(g_1) = \varphi(g_2)$, then by multiplication on the left by $(\varphi(g_1))^{-1}$ we get

$$e' = (\varphi(g_1))^{-1}\varphi(g_1) = (\varphi(g_1))^{-1}\varphi(g_2),$$

and since φ is a homomorphism $(\varphi(g_1))^{-1} = \varphi(g_1^{-1})$, so

$$e' = (\varphi(g_1))^{-1}\varphi(g_2) = \varphi(g_1^{-1})\varphi(g_2) = \varphi(g_1^{-1}g_2).$$

This shows that $g_1^{-1}g_2 \in \text{Ker } \varphi$, but since Ker $\varphi = \{e\}$ we have $g_1^{-1}g_2 = e$, and thus $g_2 = g_1$, proving that φ is injective. □

Definition 4.9. We say that a group homomorphism $\varphi\colon G \to G'$ is an *isomorphism* if there is a homomorphism $\psi\colon G' \to G$, so that

$$\psi \circ \varphi = \text{id}_G \quad \text{and} \quad \varphi \circ \psi = \text{id}_{G'}. \tag{$\dagger$}$$

If φ is an isomorphism we say that the groups G and G' are *isomorphic*. When $G' = G$, a group isomorphism is called an *automorphism*.

The reasoning used in the proof of Proposition 4.2 shows that if a a group homomorphism $\varphi\colon G \to G'$ is an isomorphism, then the homomorphism $\psi\colon G' \to G$ satisfying Condition ($\dagger$) is unique. This homomorphism is denoted φ^{-1}.

The left translations L_g and the right translations R_g are automorphisms of G.

Suppose $\varphi\colon G \to G'$ is a bijective homomorphism, and let φ^{-1} be the inverse of φ (as a function). Then for all $a, b \in G$, we have

$$\varphi(\varphi^{-1}(a)\varphi^{-1}(b)) = \varphi(\varphi^{-1}(a))\varphi(\varphi^{-1}(b)) = ab,$$

and so

$$\varphi^{-1}(ab) = \varphi^{-1}(a)\varphi^{-1}(b),$$

which proves that φ^{-1} is a homomorphism. Therefore, we proved the following fact.

Proposition 4.9. *A bijective group homomorphism $\varphi\colon G \to G'$ is an isomorphism.*

Observe that the property

$$gH = Hg, \quad \text{for all } g \in G. \tag{$*$}$$

is equivalent by multiplication on the right by g^{-1} to

$$gHg^{-1} = H, \quad \text{for all } g \in G,$$

and the above is equivalent to

$$gHg^{-1} \subseteq H, \quad \text{for all } g \in G. \tag{$**$}$$

This is because $gHg^{-1} \subseteq H$ implies $H \subseteq g^{-1}Hg$, and this for all $g \in G$.

Proposition 4.10. *Let $\varphi\colon G \to G'$ be a group homomorphism. Then $H = \text{Ker } \varphi$ satisfies Property ($**$), and thus Property ($*$).*

Proof. We have

$$\varphi(ghg^{-1}) = \varphi(g)\varphi(h)\varphi(g^{-1}) = \varphi(g)e'\varphi(g)^{-1} = \varphi(g)\varphi(g)^{-1} = e',$$

for all $h \in H = \text{Ker } \varphi$ and all $g \in G$. Thus, by definition of $H = \text{Ker } \varphi$, we have $gHg^{-1} \subseteq H$. □

Definition 4.10. For any group G, a subgroup N of G is a *normal subgroup* of G iff

$$gNg^{-1} = N, \quad \text{for all } g \in G.$$

This is denoted by $N \lhd G$.

Proposition 4.10 shows that the kernel Ker φ of a homomorphism $\varphi\colon G \to G'$ is a normal subgroup of G.

Observe that if G is abelian, then *every* subgroup of G is normal.

Consider Example 4.2. Let $R \in \mathbf{SO}(2)$ and $A \in \mathbf{SL}(2, \mathbb{R})$ be the matrices

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Then

$$A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

and we have

$$ARA^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix},$$

and clearly $ARA^{-1} \notin \mathbf{SO}(2)$. Therefore $\mathbf{SO}(2)$ is not a normal subgroup of $\mathbf{SL}(2, \mathbb{R})$. The same counter-example shows that $\mathbf{O}(2)$ is not a normal subgroup of $\mathbf{GL}(2, \mathbb{R})$.

Let $R \in \mathbf{SO}(2)$ and $Q \in \mathbf{SO}(3)$ be the matrices

$$R = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Then

$$Q^{-1} = Q^{\top} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

and we have

$$QRQ^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Observe that $QRQ^{-1} \notin \mathbf{SO}(2)$, so $\mathbf{SO}(2)$ is not a normal subgroup of $\mathbf{SO}(3)$.

Let T and $A \in \mathbf{GL}(2, \mathbb{R})$ be the following matrices

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

We have

$$A^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = A,$$

and

$$ATA^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

The matrix T is upper triangular, but ATA^{-1} is not, so the group of 2×2 upper triangular matrices is not a normal subgroup of $\mathbf{GL}(2, \mathbb{R})$.

Let $Q \in V$ and $A \in \mathbb{GL}(2, \mathbb{R})$ be the following matrices

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

We have

$$A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

and

$$AQA^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix}.$$

Clearly $AQA^{-1} \notin V$, which shows that the Klein four group is not a normal subgroup of $\mathbf{GL}(2, \mathbb{R})$.

The reader should check that the subgroups $n\mathbb{Z}$, $\mathbf{GL}^+(n, \mathbb{R})$, $\mathbf{SL}(n, \mathbb{R})$, and $\mathbf{SO}(n, \mathbb{R})$ as a subgroup of $\mathbf{O}(n, \mathbb{R})$, are normal subgroups.

If N is a normal subgroup of G, the equivalence relation $\sim$ induced by left cosets (see Definition 4.5) is the same as the equivalence induced by right cosets. Furthermore, this equivalence relation is a *congruence*, which means that: For all $g_1, g_2, g_1', g_2' \in G$,

(1) If $g_1N = g_1'N$ and $g_2N = g_2'N$, then $g_1g_2N = g_1'g_2'N$, and

(2) If $g_1N = g_2N$, then $g_1^{-1}N = g_2^{-1}N$.

As a consequence, we can define a group structure on the set $G/\sim$ of equivalence classes modulo $\sim$, by setting

$$(g_1N)(g_2N) = (g_1g_2)N.$$

Definition 4.11. Let G be a group and N be a normal subgroup of G. The group obtained by defining the multiplication of (left) cosets by

$$(g_1N)(g_2N) = (g_1g_2)N, \quad g_1, g_2 \in G$$

is denoted G/N, and called the *quotient of G by N*. The equivalence class gN of an element $g \in G$ is also denoted $\overline{g}$ (or $[g]$). The map $\pi\colon G \to G/N$ given by

$$\pi(g) = \overline{g} = gN$$

is a group homomorphism called the *canonical projection*.

Since the kernel of a homomorphism is a normal subgroup, we obtain the following very useful result.

Proposition 4.11. *Given a homomorphism of groups $\varphi\colon G \to G'$, the groups $G/\mathrm{Ker}\ \varphi$ and $\mathrm{Im}\ \varphi = \varphi(G)$ are isomorphic.*

Proof. Since φ is surjective onto its image, we may assume that φ is surjective, so that $G' = \mathrm{Im}\ \varphi$. We define a map $\overline{\varphi}\colon G/\mathrm{Ker}\ \varphi \to G'$ as follows:

$$\overline{\varphi}(\overline{g}) = \varphi(g), \quad g \in G.$$

We need to check that the definition of this map does not depend on the representative chosen in the coset $\overline{g} = g\,\mathrm{Ker}\ \varphi$, and that it is a homomorphism. If g' is another element in the coset $g\,\mathrm{Ker}\ \varphi$, which means that $g' = gh$ for some $h \in \mathrm{Ker}\,\varphi$, then

$$\varphi(g') = \varphi(gh) = \varphi(g)\varphi(h) = \varphi(g)e' = \varphi(g),$$

since $\varphi(h) = e'$ as $h \in \mathrm{Ker}\ \varphi$. This shows that

$$\overline{\varphi}(\overline{g'}) = \varphi(g') = \varphi(g) = \overline{\varphi}(\overline{g}),$$

so the map $\overline{\varphi}$ is well defined. It is a homomorphism because

$$\begin{aligned}\overline{\varphi}(\overline{g}\,\overline{g'}) &= \overline{\varphi}(\overline{gg'})\\ &= \varphi(gg')\\ &= \varphi(g)\varphi(g')\\ &= \overline{\varphi}(\overline{g})\overline{\varphi}(\overline{g'}).\end{aligned}$$

The map $\overline{\varphi}$ is injective because $\overline{\varphi}(\overline{g}) = e'$ iff $\varphi(g) = e'$ iff $g \in \mathrm{Ker}\ \varphi$, iff $\overline{g} = \overline{e}$. The map $\overline{\varphi}$ is surjective because φ is surjective. Therefore $\overline{\varphi}$ is a bijective homomorphism, and thus an isomorphism, as claimed. □

Proposition 4.11 is called the *first isomorphism theorem*.

A useful way to construct groups is the *direct product* construction.

Definition 4.12. Given two groups G an H, we let $G \times H$ be the Cartestian product of the sets G and H with the multiplication operation $\cdot$ given by

$$(g_1, h_1) \cdot (g_2, h_2) = (g_1 g_2, h_1 h_2).$$

It is immediately verified that $G \times H$ is a group called the *direct product* of G and H.

Similarly, given any n groups $G_1, \ldots, G_n$, we can define the direct product $G_1 \times \cdots \times G_n$ is a similar way.

If G is an abelian group and $H_1, \ldots, H_n$ are subgroups of G, the situation is simpler. Consider the map

$$a \colon H_1 \times \cdots \times H_n \to G$$

given by

$$a(h_1, \ldots, h_n) = h_1 + \cdots + h_n,$$

using $+$ for the operation of the group G. It is easy to verify that a is a group homomorphism, so its image is a subgroup of G denoted by $H_1 + \cdots + H_n$, and called the *sum* of the groups H_i. The following proposition will be needed.

Proposition 4.12. *Given an abelian group G, if H_1 and H_2 are any subgroups of G such that $H_1 \cap H_2 = \{0\}$, then the map a is an isomorphism*

$$a \colon H_1 \times H_2 \to H_1 + H_2.$$

Proof. The map is surjective by definition, so we just have to check that it is injective. For this, we show that $\operatorname{Ker} a = \{(0, 0)\}$. We have $a(a_1, a_2) = 0$ iff $a_1 + a_2 = 0$ iff $a_1 = -a_2$. Since $a_1 \in H_1$ and $a_2 \in H_2$, we see that $a_1, a_2 \in H_1 \cap H_2 = \{0\}$, so $a_1 = a_2 = 0$, which proves that $\operatorname{Ker} a = \{(0, 0)\}$. □

Under the conditions of Proposition 4.12, namely $H_1 \cap H_2 = \{0\}$, the group $H_1 + H_2$ is called the *direct sum* of H_1 and H_2; it is denoted by $H_1 \oplus H_2$, and we have an isomorphism $H_1 \times H_2 \cong H_1 \oplus H_2$.

4.2 Cyclic Groups

Given a group G with unit element 1, for any element $g \in G$ and for any natural number $n \in \mathbb{N}$, define g^n as follows:

$$g^0 = 1$$
$$g^{n+1} = g \cdot g^n.$$

For any integer $n \in \mathbb{Z}$, we define g^n by

$$g^n = \begin{cases} g^n & \text{if } n \geq 0 \\ (g^{-1})^{(-n)} & \text{if } n < 0. \end{cases}$$

The following properties are easily verified:

$$\begin{aligned} g^i \cdot g^j &= g^{i+j} \\ (g^i)^{-1} &= g^{-i} \\ g^i \cdot g^j &= g^j \cdot g^i, \end{aligned}$$

for all $i, j \in \mathbb{Z}$.

Define the subset $\langle g \rangle$ of G by

$$\langle g \rangle = \{g^n \mid n \in \mathbb{Z}\}.$$

The following proposition is left as an exercise.

Proposition 4.13. *Given a group G, for any element $g \in G$, the set $\langle g \rangle$ is the smallest abelian subgroup of G containing g.*

Definition 4.13. A group G is *cyclic* iff there is some element $g \in G$ such that $G = \langle g \rangle$. An element $g \in G$ with this property is called a *generator* of G.

The Klein four group V of Example 4.2 is abelian, but not cyclic. This is because V has four elements, but all the elements different from the identity have order 2.

Cyclic groups are quotients of $\mathbb{Z}$. For this, we use a basic property of $\mathbb{Z}$. Recall that for any $n \in \mathbb{Z}$, we let $n\mathbb{Z}$ denote the set of multiples of n,

$$n\mathbb{Z} = \{nk \mid k \in \mathbb{Z}\}.$$

Proposition 4.14. *Every subgroup H of $\mathbb{Z}$ is of the form $H = n\mathbb{Z}$ for some $n \in \mathbb{N}$.*

Proof. If H is the trivial group $\{0\}$, then let $n = 0$. If H is nontrivial, for any nonzero element $m \in H$, we also have $-m \in H$ and either m or $-m$ is positive, so let n be the smallest positive integer in H. By Proposition 4.13, $n\mathbb{Z}$ is the smallest subgroup of H containing n. For any $m \in H$ with $m \neq 0$, we can write

$$m = nq + r, \quad \text{with} \quad 0 \leq r < n.$$

Now, since $n\mathbb{Z} \subseteq H$, we have $nq \in H$, and since $m \in H$, we get $r = m - nq \in H$. However, $0 \leq r < n$, contradicting the minimality of n, so $r = 0$, and $H = n\mathbb{Z}$. □

Given any cyclic group G, for any generator g of G, we can define a mapping $\varphi\colon \mathbb{Z} \to G$ by $\varphi(m) = g^m$. Since g generates G, this mapping is surjective. The mapping φ is clearly a group homomorphism, so let $H = \operatorname{Ker}\varphi$ be its kernel. By a previous observation, $H = n\mathbb{Z}$ for some $n \in \mathbb{Z}$, so by the first homomorphism theorem, we obtain an isomorphism

$$\overline{\varphi}\colon \mathbb{Z}/n\mathbb{Z} \longrightarrow G$$

from the quotient group $\mathbb{Z}/n\mathbb{Z}$ onto G. Obviously, if G has finite order, then $|G| = n$. In summary, we have the following result.

Proposition 4.15. *Every cyclic group G is either isomorphic to $\mathbb{Z}$, or to $\mathbb{Z}/n\mathbb{Z}$, for some natural number $n > 0$. In the first case, we say that G is an infinite cyclic group, and in the second case, we say that G is a cyclic group of order n.*

The quotient group $\mathbb{Z}/n\mathbb{Z}$ consists of the cosets $m + n\mathbb{Z} = \{m + nk \mid k \in \mathbb{Z}\}$, with $m \in \mathbb{Z}$, that is, of the equivalence classes of $\mathbb{Z}$ under the equivalence relation $\equiv$ defined such that

$$x \equiv y \quad \text{iff} \quad x - y \in nZ \quad \text{iff} \quad x \equiv y \pmod{n}.$$

We also denote the equivalence class $x + n\mathbb{Z}$ of x by $\overline{x}$, or if we want to be more precise by $[x]_n$. The group operation is given by $\overline{x} + \overline{y} = \overline{x + y}$. For every $x \in \mathbb{Z}$, there is a unique representative, $x \bmod n$ (the nonnegative remainder of the division of x by n) in the class $\overline{x}$ of x, such that $0 \leq x \bmod n \leq n - 1$. For this reason, we often identify $\mathbb{Z}/n\mathbb{Z}$ with the set $\{0, \ldots, n-1\}$. To be more rigorous, we can give $\{0, \ldots, n-1\}$ a group structure by defining $+_n$ such that

$$x +_n y = (x + y) \bmod n.$$

Then, it is easy to see that $\{0, \ldots, n-1\}$ with the operation $+_n$ is a group with identity element 0 isomorphic to $\mathbb{Z}/n\mathbb{Z}$. The additions tables of $\mathbb{Z}/n\mathbb{Z}$ for $n = 2, 3, 4, 5, 6, 7$ are shown below.

$n = 2$

$+$	0	1
0	0	1
1	1	0

$n = 3$

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$n = 4$

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$n = 5$

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$n = 6$

$+$	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

$n = 7$

$+$	0	1	2	3	4	5	6
0	0	1	2	3	4	5	6
1	1	2	3	4	5	6	0
2	2	3	4	5	6	0	1
3	3	4	5	6	0	1	2
4	4	5	6	0	1	2	3
5	5	6	0	1	2	3	4
6	6	0	1	2	3	4	5

We can also define a multiplication operation $\cdot$ on $\mathbb{Z}/n\mathbb{Z}$ as follows:

$$\overline{a} \cdot \overline{b} = \overline{ab} = \overline{ab \bmod n}.$$

Then, it is easy to check that $\cdot$ is abelian, associative, that 1 is an identity element for $\cdot$, and that $\cdot$ is distributive on the left and on the right with respect to addition. This makes $\mathbb{Z}/n\mathbb{Z}$ into a *commutative ring*. We usually suppress the dot and write $\overline{a}\,\overline{b}$ instead of $\overline{a} \cdot \overline{b}$. The multiplication tables of $\mathbb{Z}/n\mathbb{Z}$ for $n = 2, 3, \ldots, 9$ are shown below. Since $0 \cdot m = m \cdot 0 = 0$ for all m, these tables are only given for nonzero arguments.

$n = 2$

$\cdot$	1
1	1

$n = 3$

$\cdot$	1	2
1	1	2
2	2	1

$n = 4$

$\cdot$	**1**	2	**3**
1	**1**	2	**3**
2	2	0	2
3	**3**	2	**1**

$n = 5$

$\cdot$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

$n = 6$

$\cdot$	**1**	2	3	4	**5**
1	**1**	2	3	4	**5**
2	2	4	0	2	4
3	3	0	3	0	3
4	4	2	0	4	2
5	**5**	4	3	2	**1**

$n = 7$

$\cdot$	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	1	3	5
3	3	6	2	5	1	4
4	4	1	5	2	6	3
5	5	3	1	6	4	2
6	6	5	4	3	2	1

$n = 8$

$\cdot$	**1**	2	**3**	4	**5**	6	**7**
1	**1**	2	**3**	4	**5**	6	**7**
2	2	4	6	0	2	4	6
3	**3**	6	**1**	4	**7**	2	**5**
4	4	0	4	0	4	0	4
5	**5**	2	**7**	4	**1**	6	**3**
6	6	4	2	0	6	4	2
7	**7**	6	**5**	4	**3**	2	**1**

$n = 9$

$\cdot$	**1**	**2**	3	**4**	**5**	6	**7**	**8**
1	**1**	**2**	3	**4**	**5**	6	**7**	**8**
2	**2**	**4**	6	**8**	**1**	3	**5**	**7**
3	3	6	0	3	6	0	3	6
4	**4**	**8**	3	**7**	**2**	6	**1**	**5**
5	**5**	**1**	6	**2**	**7**	3	**8**	**4**
6	6	3	0	6	3	0	6	3
7	**7**	**5**	3	**1**	**8**	6	**4**	**2**
8	**8**	**7**	6	**5**	**4**	3	**2**	**1**

Examining the above tables, we observe that for $n = 2, 3, 5, 7$, which are primes, every element has an inverse, and thus $\mathbb{Z}/n\mathbb{Z} - \{0\}$ is an abelian group under multiplication. For

$n = 4, 6, 8, 9$, the elements a that have an inverse are precisely those that are relatively prime to the modulus n (that is, $\gcd(a, n) = 1$). The subset of these elements, shown in boldface, forms an abelian group under multiplication.

These observations hold in general.

Proposition 4.16. *Given any integer $n \geq 1$, for any $a \in \mathbb{Z}$, the residue class $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$ is invertible with respect to multiplication iff* $\gcd(a, n) = 1$.

Proof. If $\overline{a}$ has inverse $\overline{b}$ in $\mathbb{Z}/n\mathbb{Z}$, then $\overline{a}\,\overline{b} = 1$, which means that

$$ab \equiv 1 \pmod{n},$$

that is $ab = 1 + nk$ for some $k \in \mathbb{Z}$, which is the Bezout identity

$$ab - nk = 1$$

and implies that $\gcd(a, n) = 1$. Conversely, if $\gcd(a, n) = 1$, then by Bezout's identity there exist $u, v \in \mathbb{Z}$ such that

$$au + nv = 1,$$

so $au = 1 - nv$, that is,

$$au \equiv 1 \pmod{n},$$

which means that $\overline{a}\,\overline{u} = 1$, so $\overline{a}$ is invertible in $\mathbb{Z}/n\mathbb{Z}$. □

Definition 4.14. The group (under multiplication) of invertible elements of the ring $\mathbb{Z}/n\mathbb{Z}$ is denoted by $(\mathbb{Z}/n\mathbb{Z})^*$. Note that this group is abelian and only defined if $n \geq 2$.

The *Euler φ-function* plays an important role in the theory of the groups $(\mathbb{Z}/n\mathbb{Z})^*$.

Definition 4.15. Given any positive integer $n \geq 1$, the *Euler φ-function* (or Euler *totient function*) is defined such that $\varphi(n)$ is the number of integers a, with $1 \leq a \leq n$, which are relatively prime to n; that is, with $\gcd(a, n) = 1$.[1]

Then, by Proposition 4.16, we see that the group $(\mathbb{Z}/n\mathbb{Z})^*$ has order $\varphi(n)$.

For $n = 2$, $(\mathbb{Z}/2\mathbb{Z})^* = \{1\}$, the trivial group. For $n = 3$, $(\mathbb{Z}/3\mathbb{Z})^* = \{1, 2\}$, and for $n = 4$, we have $(\mathbb{Z}/4\mathbb{Z})^* = \{1, 3\}$. Both groups are isomorphic to the group $\{-1, 1\}$. For $n = 6 = 2 \cdot 3$, we have $\varphi(6) = 2$, which is confirmed since $(\mathbb{Z}/6\mathbb{Z})^* = \{1, 5\}$, for $n = 8$, we have $\varphi(8) = 4$, which is confirmed since $(\mathbb{Z}/8\mathbb{Z})^* = \{1, 3, 5, 7\}$, and for $n = 9 = 3 \cdot 3$, we have $\varphi(9) = 6$, which is confirmed since $(\mathbb{Z}/9\mathbb{Z})^* = \{1, 2, 4, 5, 7, 8\}$.

Since $\gcd(a, n) = 1$ for every $a \in \{1, \ldots, n-1\}$ iff n is prime, by Proposition 4.16 we see that $(\mathbb{Z}/n\mathbb{Z})^* = \mathbb{Z}/n\mathbb{Z} - \{0\}$ iff n is prime.

[1] We allow $a = n$ to accomodate the special case $n = 1$.

Even though in principle a finite cyclic group has a very simple structure, finding a generator for a finite cyclic group is generally hard. For example, it turns out that the multiplicative group $(\mathbb{Z}/p\mathbb{Z})^*$ is a cyclic group when p is prime, but no efficient method for finding a generator for $(\mathbb{Z}/p\mathbb{Z})^*$ is known (besides a brute-force search).

Examining the multiplication tables for $(\mathbb{Z}/n\mathbb{Z})^*$ for $n = 3, 4, \ldots, 9$, we can check the following facts:

1. 2 is a generator for $(\mathbb{Z}/3\mathbb{Z})^*$.
2. 3 is a generator for $(\mathbb{Z}/4\mathbb{Z})^*$.
3. 2 is a generator for $(\mathbb{Z}/5\mathbb{Z})^*$.
4. 5 is a generator for $(\mathbb{Z}/6\mathbb{Z})^*$.
5. 3 is a generator for $(\mathbb{Z}/7\mathbb{Z})^*$.
6. Every element of $(\mathbb{Z}/8\mathbb{Z})^*$ satisfies the equation $a^2 = 1 \pmod 8$, thus $(\mathbb{Z}/8\mathbb{Z})^*$ has no generators.
7. 2 is a generator for $(\mathbb{Z}/9\mathbb{Z})^*$.

More generally, the multiplicative groups $(\mathbb{Z}/p^k\mathbb{Z})^*$ and $(\mathbb{Z}/2p^k\mathbb{Z})^*$ are cyclic groups when p is an odd prime and $k \geq 1$. A generator of the group $(\mathbb{Z}/n\mathbb{Z})^*$ (when there is one), is called a *primitive root modulo* n. As an exercise, the reader should check that the next value of n for which $(\mathbb{Z}/n\mathbb{Z})^*$ has no generator is $n = 12$. The existence of primitive roots is thoroughly investigated in Section 4.4.

The notion of order an element in a group plays an important role.

Definition 4.16. Given a group G, for any $g \in G$, the *order of g in G*, denoted by $\mathrm{ord}_G(g)$, is either infinite if the cyclic group $\langle g \rangle$ is infinite, or defined so that $\mathrm{ord}_G(g) = |\langle g \rangle|$ if $\langle g \rangle$ has finite order.

The following characterization of the order of an element will be needed.

Proposition 4.17. *Given a group G and an element $g \in G$, if g has finite order, then $\mathrm{ord}_G(g) = s$ is characterized as follows: s is the smallest positive integer such that $g^s = 1$. Furthermore, $g, g^2, \ldots, g^s = 1$ are all distinct, and for any positive integer m such that $g^m = 1$, then s divides m.*

Proof. Assume $\langle g \rangle$ has order s. By proposition 4.15, we have an isomorphism $\varphi\colon \mathbb{Z}/s\mathbb{Z} \to \langle g \rangle$ with $\varphi(1) = g$. Consequently, $\langle g \rangle = \{1 = g^s, g, g^2, \ldots, g^{s-1}\}$, where these elements are all distinct, so s is indeed the smallest positive integer such that $g^s = 1$.

Conversely, if s is the least positive integer such that $g^s = 1$, then $g, g^2, \ldots, g^s = 1$ are all distinct, since otherwise we would have $g^i = g^j$ for some i, j with $1 \leq i < j \leq s$, and then we would have

$$g^{j-i} = 1$$

with $0 < j - i < s$, contradicting the minimality of s.

For any $n \in \mathbb{N}$, we can write $n = sq + r$, with $0 \leq r < s$, and we get

$$g^n = g^{sq+r} = (g^s)^q \cdot g^r = g^r.$$

Consequently, $\langle g \rangle = \{1, g, \ldots, g^{s-1}\}$, and $\langle g \rangle$ has order s.

If $g^m = 1$, then writing $m = sq + r$, with $0 \leq r < s$, we get

$$1 = g^m = g^{sq+r} = (g^s)^q \cdot g^r = g^r,$$

so $g^r = 1$ with $0 \leq r < s$, contradicting the minimality of s, so $r = 0$ and s divides m. □

The next proposition deals with subgroups of cyclic groups.

Proposition 4.18. *Let $G = \langle g \rangle$ be a finite cyclic group of order n and let H be any subgroup of G.*

(a) The group H is cyclic and generated by some element g^k, where $k \geq 1$ is the least integer such that $g^k \in H$.

(b) The order $d = |H|$ of H divides n and $n = dk$.

(c) We have $H = \{a \in G \mid a^d = 1\}$, with d from (b).

(d) For every $d \geq 1$, the set

$$H_d = \{a \in G \mid a^d = 1\}$$

is a cyclic subgroup of G of order $\gcd(n, d)$.

(e) For every divisor d of n, there is a unique cyclic subgroup H of order d given by

$$H = \{a \in G \mid a^d = 1\}.$$

Proof. If $H = \{1\}$, then all claims are true with $k = n$ and $d = 1$. From now on, assume that $|H| > 1$, and pick $g^k \in H$ with $k \geq 1$ minimal. Since $|H| > 1$, we must have $k < n$.

(a) For any element $g^m \in H$, we can write $m = kq + r$, with $0 \leq r < k$. Then, we have

$$g^m = g^{kq+r} = (g^k)^q \cdot g^r,$$

and since $g^m, g^k \in H$, we have $g^r = (g^k)^{-q} \cdot g^m \in H$. However, $0 \leq r < k$, contradicting the minimality of k, so $r = 0$. It follows that $H = \langle g^k \rangle$ is cyclic.

(b) Let us prove that k divides n. Let $s = \gcd(k, n)$. By Bezout's theorem, we can write

$$s = ku + nv$$

for some $u, v \in \mathbb{Z}$. Then, since $g^n = 1$, we have

$$g^s = g^{ku+nv} = (g^k)^u \cdot (g^n)^v = (g^k)^u,$$

which shows that $g^s \in H$. Since k is the least positive integer such that $g^k \in H$, we must have $s = k$; that is, k divides n. But then, g^k must have order $d = n/k$, since the order of g^k is the smallest natural number h such that $g^{kh} = 1$, and since $n = dk$ is the order of g, it must divide hk, which means that d must divide h, and so $h = d$.

(c) From (b), $H = \{g^k, g^{2k}, \ldots, g^{dk} = 1\}$, and we have $(g^{jk})^d = (g^{dk})^j = 1$, which shows that every $a \in H$ satisfies the equation $a^d = 1$. Conversely, if $a \in H$ satisfies $a^d = 1$, since $a = g^i$ for some i, we have $g^{id} = 1$, and since g has order n, the number $n = kd$ must divide id, which means that k must divide i. Consequently, $a = (g^k)^{i/k} \in H$.

(d) It is immediately verified that H_d is a subgroup of G. We have $a = g^i \in H_d$ iff $(g^i)^d = g^{id} = 1$. Write $r = \gcd(d, n)$, $n = n_1 r$ and $d = d_1 r$. Then $\gcd(n_1, d_1) = 1$. Since g has order n, the number $n = n_1 r$ divides $id = id_1 r$, so n_1 divides id_1. Since $\gcd(n_1, d_1) = 1$, the number n_1 divides i, and since $1 \leq i \leq n$, we conclude that $i = n_1, 2n_1, \ldots, rn_1 = n$. Therefore, H_d has order $r = \gcd(d, n)$.

(e) This follows immediately from (d). □

Proposition 4.19. *Let $G = \langle g \rangle$ be a finite cyclic group of order n. Then we have:*

(a) For any $a \in G$, the order $\mathrm{ord}_G(a)$ of a divides n.

(b) For any i, with $1 \leq i \leq n$, the order of g^i is $n/\gcd(i, n)$.

(c) For every divisor d of n, the group G contains $\varphi(d)$ elements of order d. In particular, a cyclic group of order n has $\varphi(n)$ generators.

Proof. (a) The order $\mathrm{ord}_G(a)$ of a is the order of the cyclic group $\langle a \rangle$, and by Lagrange's theorem (Proposition 4.7), $\mathrm{ord}_G(a)$ divides n.

(b) Write $k = \gcd(i, n)$, $i = i_1 k$, and $n = n_1 k$. The order d of g^i is the smallest positive integer such that $(g^i)^d = g^{id} = 1$. Since g has order n, the number $n = n_1 k$ must divide $id = i_1 kd$, so that n_1 divides $i_1 d$. Since $\gcd(i_1, n_1) = 1$, the number n_1 must divide d, and so $d = n_1 = n/k$, as claimed.

(c) By (b), we need to know how many $i \in \{1, \ldots, n\}$ have the property $n/\gcd(i, n) = d$, or equivalently

$$\gcd(i, n) = n/d = k.$$

Obviously, i must be of the form $i = jk$, with $1 \leq j \leq d$. Now,

$$k = \gcd(i, n) = \gcd(jk, dk) = k \gcd(j, d),$$

so $\gcd(j, d) = 1$. But, there are $\varphi(d)$ integers $i \in \{1, \ldots, d\}$ such that $\gcd(j, d) = 1$, which yields (c). □

Here is another useful proposition.

Proposition 4.20. *For any abelian group G, if a is an element of finite order n_1, b is an element of finite order n_2, and $\gcd(n_1, n_2) = 1$, then $a + b$ has order $n_1 n_2$.*

Proof. The first step is to prove that $\langle a \rangle \cap \langle b \rangle = \{0\}$. This is because $\langle a \rangle \cap \langle b \rangle$ is a subgroup of both $\langle a \rangle$ and $\langle b \rangle$, so by Lagrange's theorem, the order m of $\langle a \rangle \cap \langle b \rangle$ divide both n_1 and n_2. Since $\gcd(n_1, n_2) = 1$, we must have $m = 1$. Next, we claim that if $k(a + b) = 0$, then $ka = kb = 0$. This is because if $k(a + b) = 0$, then $ka = -kb$, so $ka, kb \in \langle a \rangle \cap \langle b \rangle = \{0\}$, which means that $ka = 0$ and $kb = 0$. Now, the order of $a + b$ is the smallest positive integer s such that $s(a + b) = 0$. From what we just proved, $sa = 0$ and $sb = 0$, and since n_1 and n_2 are the orders of a and b respectively, n_1 and n_2 must divide s. Since $\gcd(n_1, n_2) = 1$, we conclude that $n_1 n_2$ divides s. On the other hand, since n_1 and n_2 are the orders of a and b respectively, $n_1 a = 0$ and $n_2 b = 0$, so $n_1 n_2 (a + b) = n_2 n_1 a + n_1 n_2 b = 0$, and since s is the least positive integer such that $s(a + b) = 0$, we see that s divides $n_1 n_2$, so we must have $s = n_1 n_2$. □

We can now prove the following important fact.

Proposition 4.21. *For every integer $n \geq 1$, we have*

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

Proof. By Proposition 4.18 (e), for every divisor d of n, there is a unique cyclic subgroup C_d of $\mathbb{Z}/n\mathbb{Z}$ of order d, and let Φ_d be the set of generators of C_d. Clearly, the sets Φ_d are pairwise disjoint. Every $g \in \mathbb{Z}/n\mathbb{Z}$ generates a cyclic group $\langle g \rangle$ of some order d, which by Proposition 4.18 (b) and (e) must be the cyclic subgroup C_d for some divisor d of n, so g is a generator for C_d, which means that $g \in \Phi_d$. It follows that the subsets Φ_d form a partition of $\mathbb{Z}/n\mathbb{Z}$, and since by Proposition 4.19, each group C_d has $\varphi(d)$ generators, we conclude that

$$n = |\mathbb{Z}/n\mathbb{Z}| = \sum_{d|n} |\Phi_d| = \sum_{d|n} \varphi(d),$$

as claimed. □

Proposition 4.21 yields a very useful characterization of cyclic groups. The proof is due to J.P. Serre; see Serre [20].

Theorem 4.22. *Let G be a finite group of order n. Then, G is cyclic iff for every divisor d of n, there are at most d elements $a \in G$ such that $a^d = 1$.*

Proof. If G is cyclic, we proved in Proposition 4.18 that for every divisor d of n there is a unique subgroup of order d given by $H_d = \{a \in G \mid a^d = 1\}$.

Let us now prove the converse. If there is some $x \in G$ of order d, then the subgroup $\langle x \rangle = \{x, x^2, \ldots, x^d = 1\}$ is cyclic of order d, and the d elements in $\langle x \rangle$ satisfy the equation $a^d = 1$. If some $y \in G$ satisfies the equation $y^d = 1$, then we already have d solutions in $\langle x \rangle$, so $y \in \langle x \rangle$. In particular, all elements of G of order d are generators of $\langle x \rangle$, and there are $\varphi(d)$ such elements. Hence, the number of elements of G of order d is either 0 or $\varphi(d)$. If it were 0 for some divisor d of n, then the formula

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

from Proposition 4.21 would say that G has strictly less than n elements, a contradiction. Therefore, for every divisor d of n, there are $\varphi(d)$ elements of order d. In particular, for $n = d$, we have an element x of order n, which shows that $G = \langle x \rangle$ is cyclic. □

We also have the following simple result which yields a short proof of a result of Euler.

Proposition 4.23. *If G is any finite group of order n, then the order of any element $g \in G$ divides n. Thus,*

$$g^n = 1, \quad \textit{for all } g \in G.$$

Proof. The cyclic subgroup $\langle g \rangle$ is a subgroup of G, so by Lagrange's theorem, its order k divides the order of G. By Proposition 4.17, we have $g^k = 1$, and since k divides n we get $g^n = 1$. □

For any integer $n \geq 2$, let $(\mathbb{Z}/n\mathbb{Z})^*$ be the group of invertible elements of the ring $\mathbb{Z}/n\mathbb{Z}$. This is a group of order $\varphi(n)$. Then, Proposition 4.23 yields the following result.

Theorem 4.24. *(Euler) For any integer $n \geq 2$ and any $a \in \{1, \ldots, n-1\}$ such that $\gcd(a, n) = 1$, we have*

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

In particular, if n is a prime, then $\varphi(n) = n - 1$, and we get Fermat's little theorem.

Theorem 4.25. *(Fermat's little theorem) For any prime p and any $a \in \{1, \ldots, p-1\}$, we have*

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

4.3 Rings and Fields

The groups $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}/n\mathbb{Z}$, and $\mathrm{M}_n(\mathbb{R})$ are more than abelian groups, they are also commutative rings. Furthermore, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are fields. We now introduce rings and fields.

Definition 4.17. A *ring* is a set A equipped with two operations $+\colon A \times A \to A$ (called *addition*) and $*\colon A \times A \to A$ (called *multiplication*) having the following properties:

(R1) A is an abelian group w.r.t. $+$;

(R2) $*$ is associative and has an identity element $1 \in A$;

(R3) $*$ is distributive w.r.t. $+$.

The identity element for addition is denoted 0, and the additive inverse of $a \in A$ is denoted by $-a$. More explicitly, the axioms of a ring are the following equations which hold for all $a, b, c \in A$:

$$a + (b + c) = (a + b) + c \qquad \text{(associativity of } +) \tag{4.1}$$
$$a + b = b + a \qquad \text{(commutativity of } +) \tag{4.2}$$
$$a + 0 = 0 + a = a \qquad \text{(zero)} \tag{4.3}$$
$$a + (-a) = (-a) + a = 0 \qquad \text{(additive inverse)} \tag{4.4}$$
$$a * (b * c) = (a * b) * c \qquad \text{(associativity of } *) \tag{4.5}$$
$$a * 1 = 1 * a = a \qquad \text{(identity for } *) \tag{4.6}$$
$$(a + b) * c = (a * c) + (b * c) \qquad \text{(distributivity)} \tag{4.7}$$
$$a * (b + c) = (a * b) + (a * c) \qquad \text{(distributivity)} \tag{4.8}$$

The ring A is *commutative* if

$$a * b = b * a \quad \text{for all } a, b \in A.$$

From (4.7) and (4.8), we easily obtain

$$a * 0 = 0 * a = 0 \tag{4.9}$$
$$a * (-b) = (-a) * b = -(a * b). \tag{4.10}$$

Note that (4.9) implies that if $1 = 0$, then $a = 0$ for all $a \in A$, and thus, $A = \{0\}$. The ring $A = \{0\}$ is called the *trivial ring*. A ring for which $1 \neq 0$ is called *nontrivial*. The multiplication $a * b$ of two elements $a, b \in A$ is often denoted by ab.

Example 4.6.

1. The additive groups $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, are commutative rings.

2. For any positive integer $n \in \mathbb{N}$, the group $\mathbb{Z}/n\mathbb{Z}$ is a group under addition. We can also define a multiplication operation by

$$\overline{a} \cdot \overline{b} = \overline{ab} = \overline{ab \bmod n},$$

for all $a, b \in \mathbb{Z}$. The reader will easily check that the ring axioms are satisfied, with $\overline{0}$ as zero and $\overline{1}$ as multiplicative unit. The resulting ring is denoted by $\mathbb{Z}/n\mathbb{Z}$.[2]

3. The group $\mathbb{R}[X]$ of polynomials in one variable with real coefficients is a ring under multiplication of polynomials. It is a commutative ring.

4. Let d be any positive integer. If d is not divisible by any integer of the form m^2, with $m \in \mathbb{N}$ and $m \geq 2$, then we say that d is *square-free*. For example, $d = 1, 2, 3, 5, 6, 7, 10$ are square-free, but $4, 8, 9, 12$ are not square-free. If d is any square-free integer and if $d \geq 2$, then the set of real numbers

$$\mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d} \in \mathbb{R} \mid a, b \in \mathbb{Z}\}$$

is a commutative a ring. If $z = a + b\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$, we write $\overline{z} = a - b\sqrt{d}$. Note that $z\overline{z} = a^2 - db^2$.

5. Similarly, if $d \geq 1$ is a positive square-free integer, then the set of complex numbers

$$\mathbb{Z}[\sqrt{-d}] = \{a + ib\sqrt{d} \in \mathbb{C} \mid a, b \in \mathbb{Z}\}$$

is a commutative ring. If $z = a + ib\sqrt{d} \in \mathbb{Z}[\sqrt{-d}]$, we write $\overline{z} = a - ib\sqrt{d}$. Note that $z\overline{z} = a^2 + db^2$. The case where $d = 1$ is a famous example that was investigated by Gauss, and $\mathbb{Z}[\sqrt{-1}]$, also denoted $\mathbb{Z}[i]$, is called the ring of *Gaussian integers*.

6. The group of $n \times n$ matrices $\mathrm{M}_n(\mathbb{R})$ is a ring under matrix multiplication. However, it is not a commutative ring.

7. The group $\mathcal{C}(a, b)$ of continuous functions $f\colon (a, b) \to \mathbb{R}$ is a ring under the operation $f \cdot g$ defined such that

$$(f \cdot g)(x) = f(x)g(x)$$

for all $x \in (a, b)$.

Definition 4.18. Given a ring A, for any element $a \in A$, if there is some element $b \in A$ such that $b \neq 0$ and $ab = 0$, then we say that a is a *zero divisor*. A ring A is an *integral domain* (or an *entire ring*) if $0 \neq 1$, A is commutative, and $ab = 0$ implies that $a = 0$ or $b = 0$, for all $a, b \in A$. In other words, an integral domain is a nontrivial commutative ring with no zero divisors besides 0.

[2]The notation $\mathbb{Z}_n$ is sometimes used instead of $\mathbb{Z}/n\mathbb{Z}$ but it clashes with the notation for the *n-adic integers* so we prefer not to use it.

Example 4.7.

1. The rings $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, are integral domains.
2. The ring $\mathbb{R}[X]$ of polynomials in one variable with real coefficients is an integral domain.
3. For any positive integer, $n \in \mathbb{N}$, we have the ring $\mathbb{Z}/n\mathbb{Z}$. Observe that if n is composite, then this ring has zero-divisors. For example, if $n = 4$, then we have
$$2 \cdot 2 \equiv 0 \pmod 4.$$
The reader should prove that $\mathbb{Z}/n\mathbb{Z}$ is an integral domain iff n is prime (use Proposition 4.16).
4. If d is a square-free positive integer and if $d \geq 2$, the ring $\mathbb{Z}[\sqrt{d}]$ is an integral domain. Similarly, if $d \geq 1$ is a square-free positive integer, the ring $\mathbb{Z}[\sqrt{-d}]$ is an integral domain. Finding the invertible elements of these rings is a very interesting problem.
5. The ring of $n \times n$ matrices $\mathrm{M}_n(\mathbb{R})$ has zero divisors.

A homomorphism between rings is a mapping preserving addition and multiplication (and 0 and 1).

Definition 4.19. Given two rings A and B, a *homomorphism between A and B* is a function $h\colon A \to B$ satisfying the following conditions for all $x, y \in A$:
$$\begin{aligned} h(x+y) &= h(x) + h(y) \\ h(xy) &= h(x)h(y) \\ h(0) &= 0 \\ h(1) &= 1. \end{aligned}$$

Actually, because B is a group under addition, $h(0) = 0$ follows from
$$h(x+y) = h(x) + h(y).$$

Example 4.8.

1. If A is a ring, for any integer $n \in \mathbb{Z}$, for any $a \in A$, we define $n \cdot a$ by
$$n \cdot a = \underbrace{a + \cdots + a}_{n}$$
if $n \geq 0$ (with $0 \cdot a = 0$) and
$$n \cdot a = -(-n) \cdot a$$
if $n < 0$. Then, the map $h\colon \mathbb{Z} \to A$ given by
$$h(n) = n \cdot 1_A$$
is a ring homomorphism (where 1_A is the multiplicative identity of A).

2. Given any real $\lambda \in \mathbb{R}$, the evaluation map $\eta_\lambda \colon \mathbb{R}[X] \to \mathbb{R}$ defined by

$$\eta_\lambda(f(X)) = f(\lambda)$$

for every polynomial $f(X) \in \mathbb{R}[X]$ is a ring homomorphism.

Definition 4.20. A ring homomorphism $h \colon A \to B$ is an *isomorphism* iff there is a ring homomorphism $g \colon B \to A$ such that $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$. An isomorphism from a ring to itself is called an *automorphism*.

As in the case of a group isomorphism, the homomorphism g is unique and denoted by h^{-1}, and it is easy to show that a bijective ring homomorphism $h \colon A \to B$ is an isomorphism.

Definition 4.21. Given a ring A, a subset A' of A is a *subring* of A if A' is a subgroup of A (under addition), is closed under multiplication, and contains 1.

For example, we have the following sequence in which every ring on the left of an inlcusion sign is a subring of the ring on the right of the inclusion sign:

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

The ring $\mathbb{Z}$ is a subring of both $\mathbb{Z}[\sqrt{d}]$ and $\mathbb{Z}[\sqrt{-d}]$, the ring $\mathbb{Z}[\sqrt{d}]$ is a subring of $\mathbb{R}$ and the ring $\mathbb{Z}[\sqrt{-d}]$ is a subring of $\mathbb{C}$.

If $h \colon A \to B$ is a homomorphism of rings, then it is easy to show for any subring A', the image $h(A')$ is a subring of B, and for any subring B' of B, the inverse image $h^{-1}(B')$ is a subring of A.

As for groups, the *kernel* of a ring homomorphism $h \colon A \to B$ is defined by

$$\mathrm{Ker}\ h = \{a \in A \mid h(a) = 0\}.$$

Just as in the case of groups, we have the following criterion for the injectivity of a ring homomorphism. The proof is identical to the proof for groups.

Proposition 4.26. *If $h \colon A \to B$ is a homomorphism of rings, then $h \colon A \to B$ is injective iff* $\mathrm{Ker}\ h = \{0\}$. *(We also write* $\mathrm{Ker}\ h = (0)$.*)*

The kernel of a ring homomorphism is an abelian subgroup of the additive group A, but in general it is not a subring of A, because it may not contain the multiplicative identity element 1. However, it satisfies the following closure property under multiplication:

$$ab \in \mathrm{Ker}\ h \quad \text{and} \quad ba \in \mathrm{Ker}\ h \quad \text{for all } a \in \mathrm{Ker}\ h \text{ and all } b \in A.$$

This is because if $h(a) = 0$, then for all $b \in A$ we have

$$h(ab) = h(a)h(b) = 0h(b) = 0 \quad \text{and} \quad h(ba) = h(b)h(a) = h(b)0 = 0.$$

Definition 4.22. Given a ring A, an additive subgroup $\mathfrak{I}$ of A satisfying the property below

$$ab \in \mathfrak{I} \quad \text{and} \quad ba \in \mathfrak{I} \quad \text{for all } a \in \mathfrak{I} \text{ and all } b \in A \tag{$*_{\text{ideal}}$}$$

is called a *two-sided ideal*. If A is a commutative ring, we simply say an *ideal*.

It turns out that for any ring A and any two-sided ideal $\mathfrak{I}$, the set $A/\mathfrak{I}$ of additive cosets $a + \mathfrak{I}$ (with $a \in A$) is a ring called a *quotient ring*. Then we have the following analog of Proposition 4.11, also called the *first isomorphism theorem*.

Proposition 4.27. *Given a homomorphism of rings $h\colon A \to B$, the rings $A/\mathrm{Ker}\, h$ and $\mathrm{Im}\, h = h(A)$ are isomorphic.*

A field is a commutative ring K for which $A - \{0\}$ is a group under multiplication.

Definition 4.23. A set K is a *field* if it is a ring and the following properties hold:

(F1) $0 \neq 1$;

(F2) $K^* = K - \{0\}$ is a group w.r.t. $*$ (i.e., every $a \neq 0$ has an inverse w.r.t. $*$);

(F3) $*$ is commutative.

If $*$ is not commutative but (F1) and (F2) hold, we say that we have a *skew field* (or *noncommutative field*).

Note that we are assuming that the operation $*$ of a field is commutative. This convention is not universally adopted, but since $*$ will be commutative for most fields we will encounter, we may as well include this condition in the definition.

Example 4.9.

1. The rings $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are fields.
2. The set of (formal) fractions $f(X)/g(X)$ of polynomials $f(X), g(X) \in \mathbb{R}[X]$, where $g(X)$ is not the null polynomial, is a field.
3. The ring $\mathcal{C}(a, b)$ of continuous functions $f\colon (a, b) \to \mathbb{R}$ such that $f(x) \neq 0$ for all $x \in (a, b)$ is a field.
4. Using Proposition 4.16, it is easy to see that the ring $\mathbb{Z}/p\mathbb{Z}$ is a field iff p is prime.
5. If d is a square-free positive integer and if $d \geq 2$, the set

$$\mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} \in \mathbb{R} \mid a, b \in \mathbb{Q}\}$$

is a field. If $z = a + b\sqrt{d} \in \mathbb{Q}(\sqrt{d})$ and $\overline{z} = a - b\sqrt{d}$, then it is easy to check that if $z \neq 0$, then $z^{-1} = \overline{z}/(z\overline{z})$.

6. Similarly, If $d \geq 1$ is a square-free positive integer, the set of complex numbers

$$\mathbb{Q}(\sqrt{-d}) = \{a + ib\sqrt{d} \in \mathbb{C} \mid a, b \in \mathbb{Q}\}$$

is a field. If $z = a + ib\sqrt{d} \in \mathbb{Q}(\sqrt{-d})$ and $\overline{z} = a - ib\sqrt{d}$, then it is easy to check that if $z \neq 0$, then $z^{-1} = \overline{z}/(z\overline{z})$.

Definition 4.24. A homomorphism $h\colon K_1 \to K_2$ between two fields K_1 and K_2 is just a homomorphism between the rings K_1 and K_2.

However, because K_1^* and K_2^* are groups under multiplication, a homomorphism of fields must be injective.

Proof. First, observe that for any $x \neq 0$,

$$1 = h(1) = h(xx^{-1}) = h(x)h(x^{-1})$$

and

$$1 = h(1) = h(x^{-1}x) = h(x^{-1})h(x),$$

so $h(x) \neq 0$ and

$$h(x^{-1}) = h(x)^{-1}.$$

But then, if $h(x) = 0$, we must have $x = 0$. Consequently, h is injective. □

Definition 4.25. A field homomorphism $h\colon K_1 \to K_2$ is an *isomorphism* iff there is a homomorphism $g\colon K_2 \to K_1$ such that $g \circ f = \mathrm{id}_{K_1}$ and $f \circ g = \mathrm{id}_{K_2}$. An isomorphism from a field to itself is called an *automorphism*.

Then, just as in the case of rings, g is unique and denoted by h^{-1}, and a bijective field homomorphism $h\colon K_1 \to K_2$ is an isomorphism.

Definition 4.26. Since every homomorphism $h\colon K_1 \to K_2$ between two fields is injective, the image $f(K_1)$ of K_1 is a subfield of K_2. We say that K_2 is an *extension* of K_1.

For example, $\mathbb{R}$ is an extension of $\mathbb{Q}$ and $\mathbb{C}$ is an extension of $\mathbb{R}$. The fields $\mathbb{Q}(\sqrt{d})$ and $\mathbb{Q}(\sqrt{-d})$ are extensions of $\mathbb{Q}$, the field $\mathbb{R}$ is an extension of $\mathbb{Q}(\sqrt{d})$ and the field $\mathbb{C}$ is an extension of $\mathbb{Q}(\sqrt{-d})$.

Definition 4.27. A field K is said to be *algebraically closed* if every polynomial $p(X)$ with coefficients in K has some root in K; that is, there is some $a \in K$ such that $p(a) = 0$.

It can be shown that every field K has some minimal extension Ω which is algebraically closed, called an *algebraic closure* of K. For example, $\mathbb{C}$ is the algebraic closure of $\mathbb{R}$. The algebraic closure of $\mathbb{Q}$ is called the *field of algebraic numbers*. This field consists of all complex numbers that are zeros of a polynomial with coefficients in $\mathbb{Q}$.

Definition 4.28. Given a field K and an automorphism $h\colon K \to K$ of K, it is easy to check that the set

$$\mathrm{Fix}(h) = \{a \in K \mid h(a) = a\}$$

of elements of K fixed by h is a subfield of K called the *field fixed by* h.

For example, if $d \geq 2$ is square-free, then the map $c\colon \mathbb{Q}(\sqrt{d}) \to \mathbb{Q}(\sqrt{d})$ given by

$$c(a + b\sqrt{d}) = a - b\sqrt{d}$$

is an automorphism of $\mathbb{Q}(\sqrt{d})$, and $\mathrm{Fix}(c) = \mathbb{Q}$.

If K is a field, we have the ring homomorphism $h\colon \mathbb{Z} \to K$ given by $h(n) = n \cdot 1$. If h is injective, then K contains a copy of $\mathbb{Z}$, and since it is a field, it contains a copy of $\mathbb{Q}$. In this case, we say that K has *characteristic* 0. If h is not injective, then $h(\mathbb{Z})$ is a subring of K, and thus an integral domain, the kernel of h is a subgroup of $\mathbb{Z}$, which by Proposition 4.14 must be of the form $p\mathbb{Z}$ for some $p \geq 1$. By the first isomorphism theorem, $h(\mathbb{Z})$ is isomorphic to $\mathbb{Z}/p\mathbb{Z}$ for some $p \geq 1$. But then, p must be prime since $\mathbb{Z}/p\mathbb{Z}$ is an integral domain iff it is a field iff p is prime. The prime p is called the *characteristic* of K, and we also says that K is of *finite characteristic*.

Definition 4.29. If K is a field, then either

(1) $n \cdot 1 \neq 0$ for all integer $n \geq 1$, in which case we say that K has *characteristic* 0, or

(2) There is some smallest prime number p such that $p \cdot 1 = 0$ called the *characteristic* of K, and we say K is of *finite characteristic*.

A field K of characteristic 0 contains a copy of $\mathbb{Q}$, thus is infinite. As we will see in Section 4.7, a finite field has nonzero characteristic p. However, there are infinite fields of nonzero characteristic.

If K_2 is a field extension of K_1, then K_2 is a vector space over K_1.

Definition 4.30. If K_2 is a field extension of K_1 and if the K_1-vector space K_2 has finite dimension m, we say that K_2 *is an extension of degree* m *over* K_1. The degree of K_2 over K_1 is denoted by $[K_2 : K_1]$.

For example, the fields $\mathbb{Q}(\sqrt{d})$ and $\mathbb{Q}(\sqrt{-d})$ have degree 2 over $\mathbb{Q}$.

Finite fields can be completely classified, which is the object Section 4.7.

4.4 Primitive Roots

In this section, we prove that certain multiplicative groups of the form $(\mathbb{Z}/n\mathbb{Z})^*$ are cyclic. It turns out that the group $(\mathbb{Z}/n\mathbb{Z})^*$ is cyclic if $n = 2, 4, p^m$, and $2p^m$, where p is an odd prime and $m \geq 1$. A generator for $(\mathbb{Z}/n\mathbb{Z})^*$ is called a *primitive root modulo* n. This terminology goes back to Euler, and is also used by Gauss in his *Disquisitiones Arithmeticae* [7]; see Article 57. In fact, it is remarkable that most of the results of this section are due to Gauss. Translations of the *Disquisitiones Arithmeticae* are available, for example, in French, and we highly recommend reading Articles 52 through 93. Gauss' style is strikingly lively and clear. Basically all the results of this section are also proved in another famous book, namely the *Vorlesungen über Zahlentheorie*, by Lejeune–Dirichlet [12]. This book was actually written by Richard Dedekind and published in 1863 after Dirichlet's death in 1859. The English translation is by John Stillwell. We were amazed to see that most contemporary books on number theory, including Apostol excellent's book [1], give proofs of the existence of primitive roots, and proofs of the quadratic reciprocity theorem, which are basically Dirichlet's proofs.

First, we review a basic structure theorem for the rings of the form $\mathbb{Z}/n\mathbb{Z}$. For this, we need the following form of the Chinese remainder theorem.

Theorem 4.28. *(Chinese remainder theorem, abstract version) For any integer $n \geq 1$, if $n = n_1 \cdots n_r$ where the n_i are relatively prime in pair, which means that $\gcd(n_i, n_j) = 1$ for all $i \neq j$, then we have an isomorphism*

$$\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_r\mathbb{Z}.$$

Proof. Consider the map $\varphi\colon \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_r\mathbb{Z}$ given by

$$\varphi(a) = (a \bmod n_1, \ldots, a \bmod n_r).$$

The map φ is a homomorphism, so let's determine its kernel $\operatorname{Ker}\varphi$. We have $\varphi(a) = (0, \ldots, 0)$ iff

$$a \equiv 0 \pmod{n_i}, \quad i = 1, \ldots, n_r,$$

and since the n_i are pairwise relatively prime, this is equivalent to

$$a \equiv 0 \pmod{n_1 \cdots n_r}.$$

Thus, $\operatorname{Ker}\varphi = n\mathbb{Z}$, and we get an injection

$$\overline{\varphi}\colon \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_r\mathbb{Z}.$$

However, $|\mathbb{Z}/n\mathbb{Z}| = n = n_1 \cdots n_r$ and $|\mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_r\mathbb{Z}| = n_1 \cdots n_r$, which shows that $\overline{\varphi}$ is a bijection, and thus an isomorphism. □

Theorem 4.28 does not explicitly tells us how to solve a system of congruences

$$\begin{aligned} x &\equiv b_1 \pmod{n_1} \\ &\vdots \\ x &\equiv b_r \pmod{n_r}, \end{aligned}$$

but the following version of the Chinese remainder theorem tells us how to do so.

Theorem 4.29. *(Chinese remainder theorem) For any integer $n \geq 1$, if $n = n_1 \cdots n_r$ where the n_i are relatively prime in pair, which means that $\gcd(n_i, n_j) = 1$ for all $i \neq j$, for any $b_1, \ldots, b_r \in \mathbb{Z}$, there exists a unique x with $0 \leq x \leq n-1$ such that*

$$\begin{aligned} x &\equiv b_1 \pmod{n_1} \\ &\vdots \\ x &\equiv b_r \pmod{n_r}. \end{aligned}$$

Proof. Let $m_i = n/n_i$, for $i = 1, \ldots, r$. Since the n_i are pairwise relatively prime, we have $\gcd(m_i, n_i) = 1$, so m_i has a unique inverse m_i' modulo n_i; that is,

$$m_i m_i' \equiv 1 \pmod{n_i}.$$

Let

$$x = b_1 m_1 m_1' + \cdots + b_r m_r m_r'.$$

We claim that x is a solution of our congruences. Indeed, since each m_j contains the factor n_i if $i \neq j$, we have

$$b_1 m_1 m_1' + \cdots + b_r m_r m_r' \equiv b_i m_i m_i' \pmod{n_i},$$

and since $m_i m_i' \equiv 1 \pmod{n_i}$, we get

$$b_1 m_1 m_1' + \cdots + b_r m_r m_r' \equiv b_i \pmod{n_i},$$

as required. The uniqueness of x follows from Theorem 4.28. We can also observe that if x, y are two solutions such that $0 \leq x, y \leq n-1$, then $x \equiv y \pmod{n_i}$ for $i = 1, \ldots, r$, which implies $x \equiv y \pmod{n}$, and thus $x = y$. □

Interestingly, Theorem 4.28 also applies to the group $(\mathbb{Z}/n\mathbb{Z})^*$ of units (invertible elements) of the ring $\mathbb{Z}/n\mathbb{Z}$. Note that we must have $n \geq 2$.

Theorem 4.30. *For any integer $n > 1$, if $n = n_1 \cdots n_r$ where the n_i are relatively prime in pair, which means that $\gcd(n_i, n_j) = 1$ for all $i \neq j$, then we have an isomorphism*

$$(\mathbb{Z}/n\mathbb{Z})^* \cong (\mathbb{Z}/n_1\mathbb{Z})^* \times \cdots \times (\mathbb{Z}/n_r\mathbb{Z})^*.$$

Proof. By Theorem 4.28, we have an isomorphism

$$\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_r\mathbb{Z}.$$

However, an element $(a_1, \ldots, a_r)$ of the product ring $\mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_r\mathbb{Z}$ is invertible iff each a_i is invertible in $\mathbb{Z}/n_i\mathbb{Z}$, which shows that the above isomorphism induces a group isomorphism

$$(\mathbb{Z}/n\mathbb{Z})^* \cong (\mathbb{Z}/n_1\mathbb{Z})^* \times \cdots \times (\mathbb{Z}/n_r\mathbb{Z})^*,$$

as claimed. □

As a corollary of Theorem 4.30, since the group $(\mathbb{Z}/n_i\mathbb{Z})^*$ has order $\varphi(n_i)$, we obtain the multiplicative property of the Euler φ-function.

Proposition 4.31. *For any two positive integers m, n, if $\gcd(m, n) = 1$, then*

$$\varphi(mn) = \varphi(m)\varphi(n).$$

If p is prime then $\varphi(p) = p - 1$. Also, if $k \geq 2$ and if p is prime, then the numbers between 1 and p^k not relatively prime to p^k are of the form ps with $1 \leq s \leq p^{k-1}$ so $\varphi(p^k) = p^k - p^{k-1} = p^{k-1}(p-1)$. Since these equations are used a lot, we record them below.

For every prime number p, for all $m \geq 2$, we have

$$\varphi(p) = p - 1, \quad \varphi(p^m) = p^{m-1}(p-1). \tag{$*_\varphi$}$$

Using Proposition 4.31 and the above equations, we can compute $\varphi(n)$ for every n (we start with $\varphi(1) = 1$). Since every positive integer $n > 1$ has a unique prime factorization

$$n = p_1^{k_1} \cdots p_r^{k_r},$$

we get

$$\varphi(n) = p_1^{k_1-1} \cdots p_r^{k_r-1}(p_1 - 1) \cdots (p_r - 1) = n\left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_r}\right).$$

As an application of Proposition 4.31 and Theorem 4.24, we prove the property mentioned in the Remark after Proposition 2.3. We begin with the following result.

Proposition 4.32. *For any positive integer $m \geq 2$ and any integer a such that $\gcd(a, m) = 1$, for any two positive integers d, e, if $ed \equiv 1 \pmod{\varphi(m)}$, then $a^{ed} \equiv a \pmod{m}$.*

Proof. Since $ed \equiv 1 \pmod{\varphi(m)}$ and $e, d, m > 0$, we can write $ed = 1 + k\varphi(m)$ for some integer $k \geq 0$. Since $\gcd(a, m) = 1$, by Euler's theorem (Theorem 4.24)

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

so we have

$$a^{ed} \equiv a^{1+k\varphi(m)} \equiv aa^{k\varphi(m)} \equiv a(a^{\varphi(m)})^k \equiv a\,1^k \equiv a \pmod{m},$$

as claimed. □

Recall that an integer m is *square-free* if it is not divisible by any integer d^2 with $d \geq 2$. Using Proposition 4.32, we can prove the following result from Niven, Zuckerman, and Montgomery [16] (Section 2.5, Problem 4).

Proposition 4.33. *Let $m \geq 2$ be any positive square-free integer. For any two positive integers d, e, if $ed \equiv 1 \pmod{\varphi(m)}$, then $a^{ed} \equiv a \pmod{m}$ for all integers $a \in \mathbb{Z}$.*

Proof. First observe that if g and m are positive integers such that g divides m, then $\varphi(g)$ divides $\varphi(m)$. This is because we can write $g = p_1^{i_1} \cdots p_h^{i_h}$ and $m = p_1^{j_1} \cdots p_h^{j_h} p_{h+1}^{j_{h+1}} \cdots p_k^{j_k}$, for some distinct primes $p_1, \ldots, p_h, p_{h+1}, \ldots, p_k$, with $h \leq k$, $1 \leq i_1 \leq j_1, \ldots, 1 \leq i_h \leq j_h$, and $j_{h+1}, \ldots, j_k \geq 1$, and we have

$$\begin{aligned}\varphi(g) &= p_1^{i_1-1} \cdots p_h^{i_h-1}(p_1 - 1) \cdots (p_h - 1)\\ \varphi(m) &= p_1^{j_1-1} \cdots p_h^{j_h-1} p_{h+1}^{j_{h+1}-1} \cdots p_k^{j_k-1}(p_1 - 1) \cdots (p_h - 1)(p_{h+1} - 1) \cdots (p_k - 1),\end{aligned}$$

so $\varphi(g)$ divides $\varphi(m)$.

If $a = 0$, then result is trivial, so assume that $a \neq 0$. Let $g = \gcd(a, m)$ and write $m = gm_1$. Since m is square-free, m can be expressed as $m = p_1 \cdots p_k$ for *distinct* primes p_i, and since $m = gm_1$, the numbers g and m_1 have no prime factor in common so $\gcd(m_1, g) = 1$. Since $g = \gcd(a, m)$, we also have $\gcd(a, m_1) = 1$.

By hypothesis $ed \equiv 1 \pmod{\varphi(m)}$, and since both g and m_1 divide m, by the previous observation $\varphi(g)$ divides $\varphi(m)$ and $\varphi(m_1)$ divides $\varphi(m)$. Therefore

$$ed \equiv 1 \pmod{\varphi(g)} \quad \text{and} \quad ed \equiv 1 \pmod{\varphi(m_1)}.$$

Since $\gcd(a, m_1) = 1$, by proposition 4.32 applied to a and m_1, we get

$$a^{ed} \equiv a \pmod{m_1}. \tag{$*_1$}$$

Since g divides a, trivially we have

$$a^{ed} \equiv a \pmod{g}. \tag{$*_2$}$$

Since $\gcd(m_1, g) = 1$, by the Chinese reminder theorem applied to $(*_1)$ and $(*_2)$, we deduce that

$$a^{ed} \equiv a \pmod{m},$$

as claimed. □

Observe that if $m = pq$ where p and q are two distinct primes, then Proposition 2.3 is a consequence of Proposition 4.33. We now return to the existence of primitive roots.

Theorem 4.30 reduces the study of the group $(\mathbb{Z}/n\mathbb{Z})^*$ to the stucture of the groups $(\mathbb{Z}/p^k\mathbb{Z})^*$, where p is a prime and $k \geq 1$. The case $p = 2$ is exceptional, but the case where p is an odd prime is nice; namely, $(\mathbb{Z}/p^k\mathbb{Z})^*$ is a cyclic group. We begin with the case $k = 1$.

Theorem 4.34. *(Gauss) For every odd prime p, the group $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic of order $p-1$. It has $\varphi(p-1)$ generators.*

Proof. We use Theorem 4.22 applied to $G = (\mathbb{Z}/p\mathbb{Z})^*$ and $n = \varphi(p) = p-1$. Since p is prime, $\mathbb{Z}/p\mathbb{Z}$ is a field, for every divisor d of $p-1$, the equation $x^d - 1 = 0$ has at most d roots in $\mathbb{Z}/p\mathbb{Z}$, and a fortiori in $(\mathbb{Z}/p\mathbb{Z})^*$. Here, we used the fact known from algebra that every nonzero polynomial of degree d with coefficients in a field has a most d roots. Therefore, $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic and has $\varphi(p-1)$ generators. □

Definition 4.31. For any positive integer n, the integers $a \in \mathbb{Z}$ such that $a \bmod n$ is a generator of $(\mathbb{Z}/n\mathbb{Z})^*$ are called *primitive roots* mod n.

Remark: Gauss' proof is not all that different from the one we gave. For every divisor d of $p-1$, Gauss defines $\psi(d)$ as the number of integers a, with $1 \leq a \leq p-1$, that have order d, and then proves that $\psi(d) = \varphi(d)$. For this, he proves Proposition 4.21; see Articles 52–56 of the *Disquisitiones Arithmeticae* [7]. Gauss also warns about the danger of regarding as established, statements which are not proved. He goes on to say that "nobody has attempted to prove Theorem 4.34, except Euler," and that Euler has talked extensively about the necessity of proving it, but that his proof is flawed in two respects! A version of the same proof is also given in Dirichlet–Dedekind [12] (Chapter 2, Section 30).

Ribenboim reports that Gauss proposes an algorithm for finding a primitive root modulo p in Articles 73 and 74 in the *Disquisitiones Arithmeticae* [7]; see Ribenboim[18] (Chapter 2, Section II). The algorithm is as follows:

Step 1. Pick any integer a with $2 \leq a \leq p-1$, and find the order t of a, that is, the least positive integer such that $a^t \equiv 1 \pmod{p}$. If a has order $p-1$, then it is a primitive root modulo p. Otherwise, go to the next step.

Step 2. Find any any number b, with $2 \leq b \leq p-1$, such that $b \not\equiv a^i \pmod{p}$, for $i = 1, \ldots, t$. Let u be the order of b, the least positive integer such that $b^u \equiv 1 \pmod{p}$. I claim that u does not divide t.

This is because if u divides t, since $b^u \equiv 1 \pmod{p}$, we would get $b^t \equiv 1 \pmod{p}$, but since the congruence $X^t \equiv 1 \pmod{p}$ has t solutions $(a, a^2, \ldots, a^t)$, then we would have $b \equiv a^i \pmod{p}$ for some i with $1 \leq i \leq t$, a contradiction. If $u = p-1$, then b is a primitive root. Otherwise, let y be the least common multiple of t and u. Then, we can split y as $y = mn$, where $\gcd(m, n) = 1$, m divides t, and n divides u. As explained by Gauss in a footnote, m and n can be obtained from prime factorizations of t and u. All prime powers only in t are included in m, all prime powers only in u are included in n, and prime powers both in t and u are included in m or n, it doesn't matter. Then, $a' \equiv a^{t/m} \pmod{p}$ has order m, $b' \equiv b^{u/n} \pmod{p}$ has order n, and because $\gcd(m, n) = 1$, the element $c = a'b'$ has order $y = mn > t$ modulo p. If $mn = p-1$, then c is a primitive root modulo p. Otherwise, go back to Step 2 with $a = c$ and $t = y$.

Since $y > t$ in Step 2, the order of t keeps increasing while dividing $p-1$, so eventually $t = p-1$, and a primitive root is found. Gauss illutrates this process for $p = 73$, and finds the primitive root 5. Gauss' algorithm requires factoring y as mn with $\gcd(m, n) = 1$, and this step requires prime factorizations of t and u. For large p, this is not a practical method. Still, it is impressive that Gauss gave an algorithm for finding a primitive root over 200 years ago.

The above algorithm does not necessarily yield the smallest primitive root g_p modulo p. It is known that $g_p > C \log p$ for infinitely many primes (for some constant C), and that $g_p < p^{0.499}$ for all $p > e^{e^{24}}$ (see Ribenboim [18], Chapter 2, Section II).

We now consider the case where $n = p^m$, with p prime and $m \geq 2$. We follow the beautiful exposition given in Apostol [1]. As we mentioned earlier, this exposition is extremely close to Dirichlet's presentation (as written up by Dedekind) [12]. The following technical proposition is needed.

Proposition 4.35. *For any odd prime p, let g be a primitive root modulo p such that*

$$g^{p-1} \not\equiv 1 \pmod{p^2}.$$

Then, for all $i \geq 2$, we have

$$g^{\varphi(p^{i-1})} \not\equiv 1 \pmod{p^i}.$$

Proof. We proceed by induction on i. The base case $i = 2$ is the hypothesis since $\varphi(p) = p-1$ as p is prime. For the induction step, assume that

$$g^{\varphi(p^{i-1})} \not\equiv 1 \pmod{p^i}. \tag{$*$}$$

By Euler's theorem,

$$g^{\varphi(p^{i-1})} \equiv 1 \pmod{p^{i-1}},$$

so we have

$$g^{\varphi(p^{i-1})} = 1 + kp^{i-1}$$

for some $k \in \mathbb{Z}$, and p does not divide k because of $(*)$. Raising the above equation to the pth power, since $\varphi(p^{i-1}) = p^{i-1} - p^{i-2}$ by $(*_\varphi)$, we get $p\varphi(p^{i-1}) = p^i - p^{i-1} = \varphi(p^i)$, and using the binomial formula

$$\begin{aligned} g^{\varphi(p^i)} &= (1 + kp^{i-1})^p = 1 + kp^i + k^2 \frac{p(p-1)}{2} p^{2(i-1)} + rp^{3(i-1)} \\ &= 1 + kp^i + k^2 \frac{p-1}{2} p^{2i-1} + rp^{3(i-1)}, \end{aligned}$$

for some $r \in \mathbb{Z}$. Now, $2i - 1 \geq i + 1$ and $3i - 3 \geq i + 1$ since $i \geq 2$, so we get the congruence

$$g^{\varphi(p^i)} = 1 + kp^i \pmod{p^{i+1}},$$

where p does not divide k, and therefore

$$g^{\varphi(p^i)} \not\equiv 1 \pmod{p^{i+1}},$$

establishing the induction step. □

The next step if to "promote" a primitive root modulo p to a primitive root modulo p^m. For this, we use the following proposition.

Proposition 4.36. *For any odd prime p, there is a primitive root g modulo p such that*

$$g^{p-1} \not\equiv 1 \pmod{p^2}. \tag{$*$}$$

Proof. Let g be any primitive root modulo p. If $(*)$ holds, we are done. Otherwise, $g^{p-1} \equiv 1 \pmod{p^2}$, in which case we consider $g_1 = g + p$. Since g is a primitive root modulo p and since $g + p \equiv g \pmod{p}$, the integer g_1 is also a primitive root modulo p, and we claim that it satisfies $(*)$. By the binomial formula we have

$$\begin{aligned} g_1^{p-1} &= (g+p)^{p-1} \\ &= g^{p-1} + (p-1)g^{p-2}p + tp^2, \\ &= g^{p-1} - g^{p-2}p + (t + g^{p-2})p^2, \end{aligned}$$

for some $t \in \mathbb{Z}$, and because $g^{p-1} \equiv 1 \pmod{p^2}$, we get

$$\begin{aligned} g_1^{p-1} &\equiv g^{p-1} - pg^{p-2} \pmod{p^2} \\ &\equiv 1 - pg^{p-2} \pmod{p^2}. \end{aligned}$$

But, we cannot have $pg^{p-2} \equiv 0 \pmod{p^2}$, for this would imply that $g^{p-2} \equiv 0 \pmod{p}$, contradicting the fact that g is a primitive root modulo p. Therefore, $g_1^{p-1} \not\equiv 1 \pmod{p^2}$, as claimed. □

Finally, we can prove that primitive roots modulo p^m exist.

Proposition 4.37. *For any odd prime p, a primitive root g modulo p is a primitive root modulo p^m for all $m \geq 2$ iff*

$$g^{p-1} \not\equiv 1 \pmod{p^2}. \tag{$*$}$$

Proof. Suppose that g is a primitive root modulo p^m for all $m \geq 1$. In particular, g is a primitive root modulo p^2. We have (by Fermat's little theorem)

$$g^{p-1} \equiv 1 \pmod{p},$$

and since $\varphi(p^2) = p(p-1) > p - 1$, if

$$g^{p-1} \equiv 1 \pmod{p^2},$$

then g can't be a primitive root modulo p^2, so $g^{p-1} \not\equiv 1 \pmod{p^2}$ must hold.

Conversely, assume that the primitive root g modulo p satisfies $(*)$. We prove that g is a primitive root modulo p^m for all $m \geq 2$. Let t be the order of g in $(\mathbb{Z}/p^m\mathbb{Z})^*$. We need to prove that

$$t = \varphi(p^m).$$

Since $g^t \equiv 1 \pmod{p^m}$, we also have $g^t \equiv 1 \pmod{p}$, and since g has order $p-1$ modulo p, we conclude that $p-1$ divides t, so we can write

$$t = q(p-1)$$

for some $q \in \mathbb{Z}$. Since g is a primitive root modulo p, we have $\gcd(g,p) = 1$, which implies $\gcd(g,p^m) = 1$, and by Euler's Theorem we have $g^{\varphi(p^m)} \equiv 1 \pmod{p^m}$, and since t is the order of g modulo p^m, the number t must divide $\varphi(p^m) = p^{m-1}(p-1)$; that is, $q(p-1)$ divides $p^{m-1}(p-1)$, so q divides p^{m-1}. Therefore, we can write

$$t = p^b(p-1), \quad \text{with} \quad b \leq m-1.$$

If we can prove that $b = m-1$, then we are done.

Assume by contradiction that $b < m-1$. If so, $b \leq m-2$ and $t = p^b(p-1)$ divides $p^{m-2}(p-1) = \varphi(p^{m-1})$. As a consequence, from $g^t \equiv 1 \pmod{p^m}$, we get

$$g^{\varphi(p^{m-1})} \equiv 1 \pmod{p^m}.$$

However, since by assumption

$$g^{p-1} \not\equiv 1 \pmod{p^2},$$

Proposition 4.35 implies that

$$g^{\varphi(p^{i-1})} \not\equiv 1 \pmod{p^i} \quad \text{for all} \quad i \geq 2,$$

a contradiction. Therefore, $b = m-1$ and the proof is complete. □

Putting Propositions 4.36 and 4.37 together, and using the fact that

$$\varphi(\varphi(p^m)) = \varphi(p^{m-1}(p-1)) = \varphi(p^{m-1})\varphi(p-1) = p^{m-2}(p-1)\varphi(p-1),$$

we obtain our theorem.

Theorem 4.38. *(Gauss) For every odd prime p and every integer $m \geq 2$, the group $(\mathbb{Z}/p^m\mathbb{Z})^*$ is cyclic of order $p^{m-1}(p-1)$. Furthermore, it has $\varphi(\varphi(p^m)) = p^{m-2}(p-1)\varphi(p-1)$ primitive roots.*

Remark: Gauss proves Theorem 4.38 in Articles 82–89 of the *Disquisitiones Arithmeticae* [7]. The above proof is basically Dedekind's proof [12] (Supplement V).

The case $n = 2p^m$ is easily handled.

Theorem 4.39. *For every odd prime p and every integer $m \geq 1$, the group $(\mathbb{Z}/2p^m\mathbb{Z})^*$ is cyclic. In fact, $(\mathbb{Z}/2p^m\mathbb{Z})^* \cong (\mathbb{Z}/p^m\mathbb{Z})^*$. Furthermore, there exist odd primitive roots g modulo p^m, and each such g is also a primitive root modulo $2p^m$.*

Proof. Since p is an odd prime, $\gcd(2, p) = 1$, so Theorem 4.30 yields an isomorphism

$$(\mathbb{Z}/2p^m\mathbb{Z})^* \cong (\mathbb{Z}/2\mathbb{Z})^* \times (\mathbb{Z}/p^m\mathbb{Z})^* \cong (\mathbb{Z}/p^m\mathbb{Z})^*,$$

since $(\mathbb{Z}/2\mathbb{Z})^*$ is the trivial group $\{1\}$.

If g is a primitive root modulo p^m, since $g + p^m \equiv g \pmod{p^m}$, the integer $g + p^m$ is also a primitive root modulo p^m, and since p is odd, either g or $g + p^m$ is odd (p^m is odd). Let g be an odd primitive root modulo p^m and let t be its order modulo $2p^m$. We need to prove that $t = \varphi(2p^m) = \varphi(2)\varphi(p^m) = \varphi(p^m)$. Now, t must divide $\varphi(2p^m) = \varphi(p^m)$, since $\gcd(g, p^m) = 1$ and g odd implies that $\gcd(g, 2p^m) = 1$ so by Euler's theorem $g^{\varphi(2p^m)} \equiv 1 \pmod{2p^m}$. On the other hand, $g^t \equiv 1 \pmod{2p^m}$, which implies $g^t \equiv 1 \pmod{p^m}$, so $\varphi(p^m)$ divides t since g is a primitive root modulo p^m (it has order $\varphi(p^m)$ modulo p^m). Therefore, $t = \varphi(p^m) = \varphi(2p^m)$, as claimed. □

In summary, we proved that primitive roots exist if $n = 2, 4, p^m$, or $2p^m$ where p is an odd prime. In the next section, we show that primitive roots do not exist in all the other cases.

4.5 Which Groups $(\mathbb{Z}/n\mathbb{Z})^*$ Have Primitive Roots

We begin with the case $p = 2^m$ with $m \geq 3$. Observe that by $(*_\varphi)$, we have $\varphi(2^m) = 2^{m-1}(2-1) = 2^{m-1}$, so if $m \geq 2$, then $\varphi(2^m)/2 = 2^{m-2}$.

Proposition 4.40. *(Gauss) If a is an odd integer, then for all $m \geq 3$, we have*

$$a^{\varphi(2^m)/2} \equiv a^{2^{m-2}} \equiv 1 \pmod{2^m}.$$

Therefore, there are no primitive roots modulo 2^m.

Proof. We proceed by induction on m. When $m = 3$, we need to show that $a^2 \equiv 1 \pmod 8$, if a is odd. This is because a is of the form $a = 2k + 1$,

$$(2k+1)^2 = 4k^2 + 4k + 1 = 4k(k+1) + 1,$$

and $k(k + 1)$ is always even.

For the induction step, the induction hypothesis says that

$$a^{2^{m-2}} = 1 + 2^m t,$$

for some $t \in \mathbb{Z}$. Squaring both sides, we get

$$a^{2^{m-1}} = 1 + 2^{m+1}t + 2^{2m}t^2,$$

so

$$a^{2^{m-1}} \equiv 1 \pmod{2^{m+1}},$$

establishing the induction step. □

Remark: Gauss proves Proposition 4.40 in Article 90 of the *Disquisitiones Arithmeticae* [7]. It also appears in Dirichlet–Dedekind [12] (Supplement V).

In fact, primitive roots do not exist in all the following cases.

Proposition 4.41. *Given any integer $n \geq 2$, if n is not of the form $n = 2, 4, p^m$, or $2p^m$, where p is an odd prime, then for any integer a with $\gcd(a, n) = 1$, we have*

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

Therefore, there are no primitive roots modulo n.

Proof. We already proved that primitive roots do not exist if $n = 2^m$ with $m \geq 3$. Therefore, we may assume that n has a factorization of the form

$$n = 2^k p_1^{k_1} \cdots p_s^{k_s},$$

where the p_i are odd primes, $s \geq 1$, $k_i \geq 1$, and $k \geq 0$. Furthermore, since n is not of the form $n = 2, 4, p^m$, or $2p^m$, we have $k \geq 2$ if $s = 1$, and $s \geq 2$ if $k = 0, 1$. We have

$$\varphi(n) = \varphi(2^k)\varphi(p_1^{k_1}) \cdots \varphi(p_s^{k_s}).$$

Pick $a \in \mathbb{Z}$ such that $\gcd(a, n) = 1$. We need to prove that

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

Let g be a primitive root modulo $p_1^{k_1}$, and write

$$a \equiv g^i \pmod{p_1^{k_1}}.$$

Then, we have

$$a^{\varphi(n)/2} \equiv g^{i\varphi(n)/2} \equiv g^{t\varphi(p_1^{k_1})} \pmod{p_1^{k_1}},$$

with

$$t = i\varphi(2^k)\varphi(p_2^{k_2}) \cdots \varphi(p_s^{k_s})/2.$$

We claim that t is an integer.

If $k \geq 2$, then $\varphi(2^k) = 2^{k-1}$ is even, so t is an integer. If $k = 0$ or $k = 1$, then $s \geq 2$ and the factor $\varphi(p_2^{k_2}) = p_2^{k_2-1}(p_2 - 1)$ is even, so t is also an integer.

Since

$$g^{\varphi(p_1^{k_1})} \equiv 1 \pmod{p_1^{k_1}},$$

from

$$a^{\varphi(n)/2} \equiv g^{t\varphi(p_1^{k_1})} \pmod{p_1^{k_1}},$$

we obtain

$$a^{\varphi(n)/2} \equiv 1 \pmod{p_1^{k_1}}.$$

A similar proof shows that

$$a^{\varphi(n)/2} \equiv 1 \pmod{p_i^{k_i}}$$

for $i = 1, \ldots s$. We still need to prove that a similar congruence holds modulo 2^k.

If $k \geq 3$, since $\gcd(a, n) = 1$, the number a must be odd, and by Proposition 4.40, we have

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

Since $\varphi(2^k)$ divides $\varphi(n)$, we get

$$a^{\varphi(n)/2} \equiv 1 \pmod{2^k}, \quad k \geq 3.$$

If $k \leq 2$, then we check directly that

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

If $k = 1$ or $k = 2$, since $\gcd(a, n) = 1$, the number a must be odd. If $k = 1$, since $\varphi(2^1) = \varphi(2) = 1$ and a is odd, we have $a \equiv 1 \pmod 2$ as desired. If $k = 2$, since a is odd, either $a = 4s + 1$ or $a = 4s + 3$ for some integer s, but then since $\varphi(2^2) = 2$,

$$a^2 \equiv (4s + 1)^2 = 16s^2 + 8s + 1 \equiv 1 \pmod 4$$

and

$$a^2 \equiv (4s + 3)^2 = 16s^2 + 24s + 9 \equiv 1 \pmod 4,$$

as desired.

But if $k \leq 2$, then $s \geq 2$, so

$$\varphi(n) = \varphi(2^k)\varphi(p_1^{k_1}) \cdots \varphi(p_s^{k_s}) = \varphi(2^k)p_1^{k_1-1}(p_1 - 1)\varphi(p_2^{k_2}) \cdots \varphi(p_s^{k_s}) = 2r\varphi(2^k),$$

for some integer r. Thus, $\varphi(2^k)$ divides $\varphi(n)/2$, and

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

holds for $k \leq 2$. In summary, the congruences

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

hold for $i = 1, \ldots, s$ and $k \geq 0$. Since the moduli are pairwise relatively prime, we obtain

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

as claimed. □

Putting everything together, we have the following remarkable result, most of which is due to Gauss.

Theorem 4.42. *The group $(\mathbb{Z}/n\mathbb{Z})^*$ is cyclic of order $\varphi(n)$ iff $n = 2, 4, p^m$, or $2p^m$, where p is an odd prime and $m \geq 1$. There are $\varphi(\varphi(n))$ primitive roots modulo n.*

Surprisingly, even in the case where $n = p$ is an odd prime, there is no known criterion to determine whether an integer a is a primitive root modulo p. For example, we don't know how to determine if 2 is a primitive root modulo p, other than by computing all powers 2^i modulo p. In fact, we have the following conjecture made by Emil Artin around 1920 (see Silverman [22], Chapter 21):

Artin's Conjecture. The number 2 is a primitive root for infinitely many primes.

Also, it is easy to see that a perfect square (a number of the form a^2) and -1 are not primitive roots. Artin also made the following conjecture.

The Generalized Artin Conjecture. Every integer which is not a perfect square and is different from -1 is a primitive root for infinitely many primes.

It has been shown by Christopher Hooley (1967) that if the Extended Riemann Hypothesis (ERH) holds, then the generalized Artin conjecture also holds. For a brief decription of the ERH, see Section 5.6.

More can be said in the "bad" case $n = 2^m$ with $m \geq 3$, Amazingly, 5 plays a special role.

Proposition 4.43. *For any integers x, y, if $x \equiv 1 + 4y \pmod 8$, then*

$$x^{2^k} \equiv 1 + 2^{k+2}y \pmod{2^{k+3}},$$

for all $k \geq 0$.

Proof. We proceed by induction on k. The case $k = 0$ is the hypothesis. For the induction step, it is enough to prove that if $a \equiv 1+2^{k+1}b \pmod{2^{k+2}}$ for some $k \geq 1$, then $a^2 \equiv 1+2^{k+2}b \pmod{2^{k+3}}$.

If $a \equiv 1 + 2^{k+1}b \pmod{2^{k+2}}$, then $a = 1 + 2^{k+1}b + c2^{k+2}$, for some c, so we get

$$\begin{aligned} a^2 &= (1 + 2^{k+1}b + c2^{k+2})^2 \\ &= (1 + 2^{k+1}(b + 2c))^2 \\ &= 1 + 2^{k+2}(b + 2c) + 2^{2k+2}(b + 2c)^2 \\ &= 1 + 2^{k+2}b + 2^{k+3}c + 2^{2k+2}(b + 2c)^2, \end{aligned}$$

and because $k \geq 1$, we have $2k + 2 \geq k + 3$, so we get

$$a^2 \equiv 1 + 2^{k+2}b \pmod{2^{k+3}},$$

establishing the induction step. □

Observe that if we set $x = 5$ and $y = 1$, then $5 \equiv 5 \pmod 8$, so by Proposition 4.43, we have

$$5^{2^k} \equiv 1 + 2^{k+2} \pmod{2^{k+3}}, \quad \text{for all } k \geq 0.$$

On the other hand, since 5 is odd, by Proposition 4.40, we have

$$5^{2^{m-2}} \equiv 1 \pmod{2^m}.$$

Therefore, 5 has order 2^{m-2} modulo 2^m. We can use this fact to prove the following result (following Bourbaki [2], Chapter VII, §2.4). This result is more or less implicit in Article 91 of the *Disquisitiones Arithmeticae* [7]. It is explicitly proved in Dirichlet–Dedekind [12] (Supplement V).

Theorem 4.44. *For any $m \geq 3$, the group $(\mathbb{Z}/2^m\mathbb{Z})^*$ is isomorphic to the direct product $\{-1, 1\} \times \langle 5\rangle$ of the cyclic subgroup $\{-1, 1\}$ generated by -1 and the cyclic subgroup $\langle 5\rangle$ of order 2^{m-2} generated by 5.*

Proof. For $m \geq 3$, we have the homomorphism $\pi\colon (\mathbb{Z}/2^m\mathbb{Z})^* \to (\mathbb{Z}/4\mathbb{Z})^*$ given by

$$\pi(a \bmod 2^m) = a \bmod 4.$$

The kernel of this homomorphism is the subgroup $U(2^m)$ of $(\mathbb{Z}/2^m\mathbb{Z})^*$ given by

$$U(2^m) = \{a \bmod 2^m \mid a \equiv 1 \pmod 4\},$$

and so $U(2^m)$ has order 2^{m-2}. By the first isomorphism theorem, $(\mathbb{Z}/4\mathbb{Z})^*$ is isomorphic to $(\mathbb{Z}/2^m\mathbb{Z})^*/U(2^m)$, and by Lagrange's theorem, since $(\mathbb{Z}/4\mathbb{Z})^*$ has order 2 and $U(2^m)$ has order 2^{m-2}, we conclude that $(\mathbb{Z}/2^m\mathbb{Z})^*$ has order 2^{m-1}.

We already know that the cyclic subgroup group $\langle 5\rangle$ generated by 5 has order 2^{m-2} modulo 2^m, and since $5 \equiv 1 \pmod 4$, we see that $5 \in U(2^m)$, and thus $U(2^m) = \langle 5\rangle$. We claim that $-1 \notin \langle 5\rangle$. This follows because $\pi(-1) = -1 \bmod 4$, and $-1 \not\equiv 1 \pmod 4$, so -1 does not belong to the kernel $U(2^m)$ of π, which is equal to $\langle 5\rangle$.

I

Consequently, if $H = \{-1, 1\}$ is the subgroup generated by -1, we have $H \cap \langle 5\rangle = \{0\}$. By Proposition 4.12, we have an isomorphism

$$\{-1, 1\} \times \langle 5\rangle \cong \{-1, 1\} \oplus \langle 5\rangle.$$

Now, $(\mathbb{Z}/2^m\mathbb{Z})^*$ has order 2^{m-1}, the subgroup $\langle 5\rangle$ has order 2^{m-2}, and $\{-1, 1\}$ has order 2, so

$$(\mathbb{Z}/2^m\mathbb{Z})^* = \{-1, 1\} \oplus \langle 5\rangle$$

and we have an isomorphism $(\mathbb{Z}/2^m\mathbb{Z})^* \cong \{-1, 1\} \times \langle 5\rangle$. □

Remarks: Another way to prove Theorem 4.38 is to proceed as follows (following Bourbaki [2], Chapter VII, §2.4). We have the homomorphism $\pi\colon (\mathbb{Z}/p^m\mathbb{Z})^* \to (\mathbb{Z}/p\mathbb{Z})^*$ given by

$$\pi(a \bmod p^m) = a \bmod p.$$

The kernel of this homomorphism is the subgroup $U(p^m)$ of $(\mathbb{Z}/p^m\mathbb{Z})^*$ given by

$$U(p^m) = \{a \bmod p^m \mid a \equiv 1 \pmod p\},$$

and so $U(p^m)$ has order p^{m-1}. By the first isomorphism theorem, $(\mathbb{Z}/p\mathbb{Z})^*$ is isomorphic to $(\mathbb{Z}/p^m\mathbb{Z})^*/U(p^m)$, and by Lagrange's theorem, since $(\mathbb{Z}/p\mathbb{Z})^*$ has order $p-1$ and $U(p^m)$ has order p^{m-1}, we conclude that $(\mathbb{Z}/p^m\mathbb{Z})^*$ has order $p^{m-1}(p-1)$.

Next, we show that $p+1$ has order p^{m-1} in $(\mathbb{Z}/p^m\mathbb{Z})^*$, and since $p+1 \equiv 1 \pmod p$, $p+1$ is a generator for $U(p^m)$. For this we prove that if p is an odd prime and $x \equiv 1 + py \pmod{p^2}$, then $x^{p^k} \equiv 1 + p^{k+1}y \pmod{p^{k+2}}$, for all $k \geq 0$.

Then, using a primitive root of $(\mathbb{Z}/p\mathbb{Z})^*$, we can find an element y of order $p-1$ in $(\mathbb{Z}/p^m\mathbb{Z})^*$. By Proposition 4.20, since $\gcd(p^{m-1}, p-1) = 1$, we conclude that $(p+1)y$ has order $p^{m-1}(p-1) = \varphi(p^m)$, so $(p+1)y$ is a primitive root modulo p^m.

4.6 The Lucas Theorem, PRIMES is in NP

In this section we discuss an application of the existence of primitive roots in $(\mathbb{Z}/p\mathbb{Z})^*$ where p is an odd prime, known an the $n-1$ *test*. This test due to E. Lucas determines whether a positive odd integer n is prime or not by examining the prime factors of $n-1$ and checking some congruences.

The $n-1$ test can be described as the construction of a certain kind of tree rooted with n, and it turns out that the number of nodes in this tree is bounded by $2\log_2 n$, and that the number of modular multiplications involved in checking the congruences is bounded by $2\log_2^2 n$.

Recall that when we talk about the complexity of algorithms dealing with numbers, we assume that all inputs (to a Turing machine) are strings representing these numbers, typically in base 2. Since the length of the binary representation of a natural number $n \geq 1$ is $\lfloor \log_2 n \rfloor + 1$ (or $\lceil \log_2(n+1) \rceil$, which allows $n = 0$), the complexity of algorithms dealing with (nonzero) numbers m, n, *etc.* is expressed in terms of $\log_2 m, \log_2 n$, *etc.* Recall that for any real number $x \in \mathbb{R}$, the *floor of* x is the greatest integer $\lfloor x \rfloor$ that is less that or equal to x, and the *ceiling of* x is the least integer $\lceil x \rceil$ that is greater that or equal to x. If we choose to represent numbers in base 10, since for any base b we have $\log_b x = \ln x / \ln b$, we have

$$\log_2 x = \frac{\ln 10}{\ln 2} \log_{10} x.$$

Since $(\ln 10)/(\ln 2) \approx 3.322 \approx 10/3$, we see that the number of decimal digits needed to represent the integer n in base 10 is approximately 30% of the number of bits needed to represent n in base 2.

Since the Lucas test yields a tree such that the number of modular multiplications involved in checking the congruences is bounded by $2\log_2^2 n$, it is not hard to show that testing whether or not a positive integer n is prime, a problem denoted PRIMES, belongs to the complexity class **NP**. This result was shown by V. Pratt [17] (1975), but Peter Freyd told me that it was "folklore." Of course, since 2002, thanks to the AKS algorithm, we know that PRIMES actually belongs to the class **P**, but this is a much harder result.

Here is Lehmer's version of the Lucas result, from 1876.

Theorem 4.45. *(Lucas theorem) Let n be a positive integer with $n \geq 2$. Then n is prime iff there is some integer $a \in \{1, 2, \ldots, n-1\}$ such that the following two conditions hold:*

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

(2) If $n > 2$, then $a^{(n-1)/q} \not\equiv 1 \pmod{n}$ for all prime divisors q of $n-1$.

Proof. First, assume that Conditions (1) and (2) hold. If $n = 2$, since 2 is prime, we are done. Thus assume that $n \geq 3$, and let r be the order of a. We claim that $r = n-1$. The condition $a^{n-1} \equiv 1 \pmod{n}$ implies that r divides $n-1$. Suppose that $r < n-1$, and let q be a prime divisor of $(n-1)/r$ (so q divides $n-1$). Since r is the order or a we have $a^r \equiv 1 \pmod{n}$, so we get

$$a^{(n-1)/q} \equiv a^{r(n-1)/(rq)} \equiv (a^r)^{(n-1)/(rq)} \equiv 1^{(n-1)/(rq)} \equiv 1 \pmod{n},$$

contradicting Condition (2). Therefore, $r = n-1$, as claimed.

We now show that n must be prime. Now $a^{n-1} \equiv 1 \pmod{n}$ implies that a and n are relatively prime so by Euler's Theorem (Theorem 4.24),

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

Since the order of a is $n-1$, we have $n-1 \leq \varphi(n)$. If $n \geq 3$ is not prime, then n has some prime divisor p, but n and p are integers in $\{1, 2, \ldots, n\}$ that are not relatively prime to n, so by definition of $\varphi(n)$, we have $\varphi(n) \leq n-2$, contradicting the fact that $n-1 \leq \varphi(n)$. Therefore, n must be prime.

Conversely, assume that n is prime. If $n = 2$, then we set $a = 1$. Otherwise, pick a to be any primitive root modulo p. □

Clearly, if $n > 2$ then we may assume that $a \geq 2$. The main difficulty with the $n-1$ test is not so much guessing the primitive root a, but finding a *complete prime factorization* of $n-1$. However, as a nondeterministic algorithm, the $n-1$ test yields a "proof" that a number n is indeed prime which can be represented as a tree, and the number of operations

needed to check the required conditions (the congruences) is bounded by $c \log_2^2 n$ for some positive constant c, and this implies that testing primality is in **NP**.

Before explaining the details of this method, we sharpen slightly Lucas' theorem to deal only with odd prime divisors.

Theorem 4.46. *Let n be a positive odd integer with $n \geq 3$. Then n is prime iff there is some integer $a \in \{2, \ldots, n-1\}$ (a guess for a primitive root modulo n) such that the following two conditions hold:*

(1b) $a^{(n-1)/2} \equiv -1 \pmod{n}$.

(2b) If $n-1$ is not a power of 2, then $a^{(n-1)/2q} \not\equiv -1 \pmod{n}$ for all odd prime divisors q of $n-1$.

Proof. Assume that Conditions (1b) and (2b) of Theorem 4.46 hold. Then we claim that Conditions (1) and (2) of Theorem 4.45 hold. By squaring the congruence $a^{(n-1)/2} \equiv -1 \pmod{n}$, we get $a^{n-1} \equiv 1 \pmod{n}$, which is Condition (1) of Theorem 4.45. Since $a^{(n-1)/2} \equiv -1 \pmod{n}$, Condition (2) of Theorem 4.45 holds for $q = 2$. Next, if q is an odd prime divisor of $n-1$, let $m = a^{(n-1)/2q}$. Condition (1b) means that

$$m^q \equiv a^{(n-1)/2} \equiv -1 \pmod{n}.$$

Now if $m^2 \equiv a^{(n-1)/q} \equiv 1 \pmod{n}$, since q is an odd prime, we can write $q = 2k+1$ for some $k \geq 1$, and then

$$m^q \equiv m^{2k+1} \equiv (m^2)^k m \equiv 1^k m \equiv m \pmod{n},$$

and since $m^q \equiv -1 \pmod{n}$, we get

$$m \equiv -1 \pmod{n}$$

(regardless of whether n is prime or not). Thus we proved that if $m^q \equiv -1 \pmod{n}$ and $m^2 \equiv 1 \pmod{n}$, then $m \equiv -1 \pmod{n}$. By contrapositive, we see that if $m \not\equiv -1 \pmod{n}$, then $m^2 \not\equiv 1 \pmod{n}$ or $m^q \not\equiv -1 \pmod{n}$, but since $m^q \equiv a^{(n-1)/2} \equiv -1 \pmod{n}$ by Condition (1a), we conclude that $m^2 \equiv a^{(n-1)/q} \not\equiv 1 \pmod{n}$, which is Condition (2) of Theorem 4.45. But then, Theorem 4.45 implies that n is prime.

Conversely, assume that n is an odd prime, and let a be any primitive root modulo n. Then by little Fermat we know that

$$a^{n-1} \equiv 1 \pmod{n},$$

so

$$(a^{(n-1)/2} - 1)(a^{(n-1)/2} + 1) \equiv 0 \pmod{n}.$$

Since n is prime, either $a^{(n-1)/2} \equiv 1 \pmod{n}$ or $a^{(n-1)/2} \equiv -1 \pmod{n}$, but since a generates $(\mathbb{Z}/n\mathbb{Z})^*$, it has order $n-1$, so the congruence $a^{(n-1)/2} \equiv 1 \pmod{n}$ is impossible, and

Condition (1b) must hold. Similarly, if we had $a^{(n-1)/2q} \equiv -1 \pmod{n}$ for some odd prime divisor q of $n-1$, then by squaring we would have

$$a^{(n-1)/q} \equiv 1 \pmod{n},$$

and a would have order at most $(n-1)/q < n-1$, which is absurd. □

If n is an odd prime, we can use Theorem 4.46 to build recursively a tree which is a proof, or certificate, of the fact that n is indeed prime. We first illustrate this process with the prime $n = 1279$.

Example 4.10. If $n = 1279$, then we easily check that $n - 1 = 1278 = 2 \cdot 3^2 \cdot 71$. We build a tree whose root node contains the triple $(1279, ((2,1),(3,2),(71,1)),3)$, where $a = 3$ is the guess for a primitive root modulo 1279. In this simple example, it is clear that 3 and 71 are prime, but we must supply proofs that these number are prime, so we recursively apply the process to the odd divisors 3 and 71.

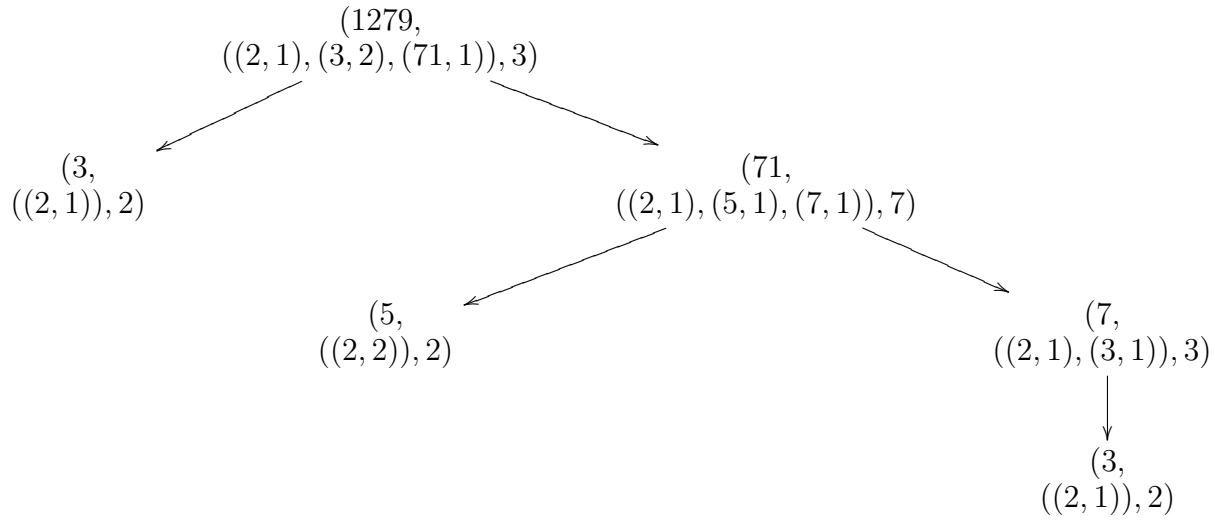
Since $3 - 1 = 2^1$ is a power of 2, we create a one-node tree $(3, ((2,1)), 2)$, where $a = 2$ is a guess for a primitive root modulo 3. This is a leaf node.

Since $71-1 = 70 = 2\cdot 5\cdot 7$, we create a tree whose root node is $(71, ((2,1),(5,1),(7,1)),7)$, where $a = 7$ is the guess for a primitive root modulo 71. Since $5 - 1 = 4 = 2^2$, and $7 - 1 = 6 = 2 \cdot 3$, this node has two successors $(5, ((2,2)), 2)$ and $(7, ((2,1),(3,1)),3)$, where 2 is the guess for a primitive root modulo 5, and 3 is the guess for a primitive root modulo 7.

Since $4 = 2^2$ is a power of 2, the node $(5, ((2,2)), 2)$ is a leaf node.

Since $3 - 1 = 2^1$, the node $(7, ((2,1),(3,1)),3)$ has a single successor, $(3, ((2,1)), 2)$, where $a = 2$ is a guess for a primitive root modulo 3. Since $2 = 2^1$ is a power of 2, the node $(3, ((2,1)), 2)$ is a leaf node.

To recap, we obtain the following tree:



We still have to check that the relevant congruences hold at every node. For the root node $(1279, ((2,1),(3,2),(71,1)),3)$, we check that

$$3^{1278/2} \equiv 3^{864} \equiv -1 \pmod{1279} \tag{1b}$$
$$3^{1278/(2\cdot 3)} \equiv 3^{213} \equiv 775 \pmod{1279} \tag{2b}$$
$$3^{1278/(2\cdot 71)} \equiv 3^{9} \equiv 498 \pmod{1279}. \tag{2b}$$

Assuming that 3 and 71 are prime, the above congruences check that Conditions (1a) and (2b) are satisfied, and by Theorem 4.46 this proves that 1279 is prime. We still have to certify that 3 and 71 are prime, and we do this recursively.

For the leaf node $(3, ((2,1)), 2)$, we check that

$$2^{2/2} \equiv -1 \pmod{3}. \tag{1b}$$

For the node $(71, ((2,1),(5,1),(7,1)),7)$, we check that

$$7^{70/2} \equiv 7^{35} \equiv -1 \pmod{71} \tag{1b}$$
$$7^{70/(2\cdot 5)} \equiv 7^{7} \equiv 14 \pmod{71} \tag{2b}$$
$$7^{70/(2\cdot 7)} \equiv 7^{5} \equiv 51 \pmod{71}. \tag{2b}$$

Now, we certified that 3 and 71 are prime, assuming that 5 and 7 are prime, which we now establish.

For the leaf node $(5, ((2,2)), 2)$, we check that

$$2^{4/2} \equiv 2^{2} \equiv -1 \pmod{5}. \tag{1b}$$

For the node $(7, ((2,1),(3,1)),3)$, we check that

$$3^{6/2} \equiv 3^{3} \equiv -1 \pmod{7} \tag{1b}$$
$$3^{6/(2\cdot 3)} \equiv 3^{1} \equiv 3 \pmod{7}. \tag{2b}$$

We have certified that 5 and 7 are prime, given that 3 is prime, which we finally verify.

At last, for the leaf node $(3, ((2,1)), 2)$, we check that

$$2^{2/2} \equiv -1 \pmod{3}. \tag{1b}$$

The above example suggests the following definition.

Definition 4.32. Given any odd integer $n \geq 3$, a *pre-Lucas tree for* n is defined inductively as follows:

(1) It is a one-node tree labeled with $(n, ((2, i_0)), a)$, such that $n - 1 = 2^{i_0}$, for some $i_0 \geq 1$ and some $a \in \{2, \ldots, n-1\}$.

(2) If $L_1, \ldots, L_k$ are k pre-Lucas (with $k \geq 1$), where the tree L_j is a pre-Lucas tree for some odd integer $q_j \geq 3$, then the tree L whose root is labeled with $(n, ((2, i_0), ((q_1, i_1), \ldots, (q_k, i_k)), a)$ and whose jth subtree is L_j is a *pre-Lucas tree* for n if

$$n - 1 = 2^{i_0} q_1^{i_1} \cdots q_k^{i_k},$$

for some $i_0, i_1, \ldots, i_k \geq 1$, and some $a \in \{2, \ldots, n-1\}$.

Both in (1) and (2), the number a is a guess for a primitive root modulo n.

A pre-Lucas tree for n is a *Lucas tree for* n if the following conditions are satisfied:

(3) If L is a one-node tree labeled with $(n, ((2, i_0)), a)$, then

$$a^{(n-1)/2} \equiv -1 \pmod{n}.$$

(4) If L is a pre-Lucas tree whose root is labeled with $(n, ((2, i_0), ((q_1, i_1), \ldots, (q_k, i_k)), a)$, and whose jth subtree L_j is a pre-Lucas tree for q_j, then L_j is a Lucas tree for q_j for $j = 1, \ldots, k$, and

(a) $a^{(n-1)/2} \equiv -1 \pmod{n}$.

(b) $a^{(n-1)/2q_j} \not\equiv -1 \pmod{n}$ for $j = 1, \ldots, k$.

Since Conditions (3) and (4) of Definition 4.32 are Conditions (1b) and (2b) of Theorem, 4.46, we see that Definition 4.32 has been designed in such a way that Theorem 4.46 yields the following result.

Theorem 4.47. *An odd integer $n \geq 3$ is prime iff it has some Lucas tree.*

The issue is now to see how long it takes to check that a pre-Lucas tree is a Lucas tree. Of course, exponentiation modulo n is performed by repeated squaring, as explained in Section 2.3. In that section, we observed that computing $x^m \bmod n$ requires at most $2 \log_2 m$ modular multiplications. Using this fact we obtain the following result.

Proposition 4.48. *If p is any odd prime, then any pre-Lucas tree L for p has at most $\log_2 p$ nodes, and the number $M(p)$ of modular multiplications required to check that the pre-Lucas tree L is a Lucas tree is less than $2 \log_2^2 p$.*

Proof. Let $N(p)$ be the number of nodes in a pre-Lucas tree for p. We proceed by complete induction. If $p = 3$, then $p - 1 = 2^1$, any pre-Lucas tree has a single node, and $1 < \log_2 3$.

Suppose the results holds for any odd prime less than p. If $p - 1 = 2^{i_0}$, then any Lucas tree has a single node, and $1 < \log_2 3 < \log_2 p$. If $p - 1$ has the prime factorization

$$p - 1 = 2^{i_0} q_1^{i_1} \cdots q_k^{i_k},$$

then by the induction hypothesis, each pre-Lucas tree L_j for q_j has less than $\log_2 q_j$ nodes, so

$$N(p) = 1 + \sum_{j=1}^{k} N(q_j) < 1 + \sum_{j=1}^{k} \log_2 q_j = 1 + \log_2(q_1 \cdots q_k) \leq 1 + \log_2\left(\frac{p-1}{2}\right) < \log_2 p,$$

establishing the induction hypothesis.

If r is one of the odd primes in the pre-Lucas tree for p, and $r < p$, then there is some other odd prime q in this pre-Lucas tree such that r divides $q-1$ and $q \leq p$. We also have to show that at some point, $a^{(q-1)/2r} \not\equiv -1 \pmod{q}$ for some a, and at another point, that $b^{(r-1)/2} \equiv -1 \pmod{r}$ for some b. Using the fact that the number of modular multiplications required to exponentiate to the power m is at most $2\log_2 m$, we see that the number of multiplications required by the above two exponentiations does not exceed

$$2\log_2\left(\frac{q-1}{2r}\right) + 2\log_2\left(\frac{r-1}{2}\right) < 2\log_2 q - 4 < 2\log_2 p.$$

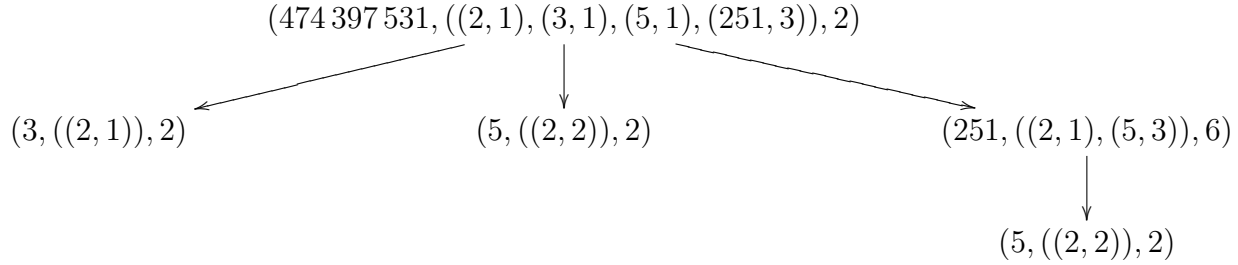
As a consequence, we have

$$M(p) < 2\log_2\left(\frac{p-1}{2}\right) + (N(p)-1)2\log_2 p < 2\log_2 p + (\log_2 p - 1)2\log_2 p = 2\log_2^2 p,$$

as claimed. □

The following impressive example is from Pratt [17].

Example 4.11. Let $n = 474\,397\,531$. It is easy to check that $n - 1 = 474\,397\,531 - 1 = 474\,397\,530 = 2 \cdot 3 \cdot 5 \cdot 251^3$. We claim that the following is a Lucas tree for $n = 474\,397\,531$:



To verify that the above pre-Lucas tree is a Lucas tree, we check that 2 is indeed a primitive root modulo 474 397 531 by computing (using `Mathematica`) that

$$2^{474\,397\,530/2} \equiv 2^{237\,198\,765} \equiv -1 \pmod{474\,397\,531} \tag{1}$$

$$2^{474\,397\,530/(2\cdot 3)} \equiv 2^{79\,066\,255} \equiv 9\,583\,569 \pmod{474\,397\,531} \tag{2}$$

$$2^{474\,397\,530/(2\cdot 5)} \equiv 2^{47\,439\,753} \equiv 91\,151\,207 \pmod{474\,397\,531} \tag{3}$$

$$2^{474\,397\,530/(2\cdot 251)} \equiv 2^{945\,015} \equiv 282\,211\,150 \pmod{474\,397\,531}. \tag{4}$$

The number of modular multiplications is: 27 in (1), 26 in (2), 25 in (3) and 19 in (4).

We have $251 - 1 = 250 = 2 \cdot 5^3$, and we verify that 6 is a primitive root modulo 251 by computing:

$$6^{250/2} \equiv 6^{125} \equiv -1 \pmod{251} \tag{5}$$

$$6^{250/(2\cdot 5)} \equiv 6^{10} \equiv 175 \pmod{251}. \tag{6}$$

The number of modular multiplications is: 6 in (5), and 3 in (6).

We have $5 - 1 = 4 = 2^2$, and 2 is a primitive root modulo 5, since

$$2^{4/2} \equiv 2^2 \equiv -1 \pmod{5}. \tag{7}$$

This takes one multiplication.

We have $3 - 1 = 2 = 2^1$, and 2 is a primitive root modulo 3, since

$$2^{2/2} \equiv 2^1 \equiv -1 \pmod{3}. \tag{8}$$

This takes 0 multiplications.

Therefore, 474 397 531 is prime.

As nice as it is, Proposition 4.48 is deceiving, because *finding* a Lucas tree is hard.

Remark: Pratt [17] presents his method for finding a certificate of primality in terms of a proof system. Although quite elegant, we feel that this method is not as transparent as the method using Lucas trees, which we adapted from Crandall and Pomerance [3]. Pratt's proofs can be represented as trees, as Pratt sketches in Section 3 of his paper. However, Pratt uses the basic version of Lucas' theorem, Theorem 4.45, instead of the improved version, Theorem 4.46, so his proof trees have at least twice as many nodes as ours.

The following nice result was first shown by V. Pratt [17] in 1975.

Theorem 4.49. *The problem* PRIMES *(testing whether an integer is prime) is in* **NP**.

Proof. Since all even integers besides 2 are composite, we can restrict out attention to odd integers $n \geq 3$. By Theorem 4.47, an odd integer $n \geq 3$ is prime iff it has a Lucas tree. Given any odd integer $n \geq 3$, since all the numbers involved in the definition of a pre-Lucas tree are less than n, there is a finite (very large) number of pre-Lucas trees for n. Given a guess of a Lucas tree for n, checking that this tree is a pre-Lucas tree can be performed in $O(\log_2 n)$, and by Proposition 4.48, checking that it is a Lucas tree can be done in $O(\log_2^2 n)$. Therefore PRIMES is in **NP**. □

Of course, checking whether a number n is composite is in **NP**, since it suffices to guess to factors n_1, n_2 and to check that $n = n_1 n_2$, which can be done in polynomial time in $\log_2 n$. Therefore, PRIMES $\in$ **NP** $\cap$ **coNP**. As we said earlier, this was the situation until the discovery of the AKS algorithm, which places PRIMES in **P**.

Remark: Altough finding a primitive root modulo p is hard, we know that the number of primitive roots modulo p is $\varphi(\varphi(p))$. If p is large enough, this number is actually quite large. According to Crandal and Pomerance [3] (Chapter 4, Section 4.1.1), if p is a prime and if $p > 200560490131$, then p has more than $p/(2 \ln \ln p)$ primitive roots.

The Lucas test yields a method to check whether certain numbers known as Fermat numbers are prime. These are numbers for which the prime factorization of $n-1$ is a power of 2.

Example 4.12. First, let us find a tree for $n = 17$. Since $17 - 1 = 16 = 2^4$, we get the one-node tree $(17, ((2, 4)), 3)$, where 3 is a guess for a primitive root of 17. Since

$$3^{16/2} \equiv 3^8 \equiv -1 \pmod{17},$$

we have a proof that $17 = 2^4 + 1 = 2^{2^2} + 1$ is prime.

Let us now find a tree for $n = 257$. Since $257 - 1 = 256 = 2^8$, we have the one-node tree $(257, ((2, 8)), 3)$. We can check that

$$3^{256/2} = 3^{128} \equiv -1 \pmod{257},$$

which proves that $257 = 2^8 + 1 = 2^{2^3} + 1$ is prime.

Finally, let us now find a tree for $n = 65537$. Since $65537 - 1 = 65536 = 2^{16}$, we have the one-node tree $(65537, ((2, 16)), 3)$. We can check that

$$3^{65536/2} = 3^{32768} \equiv -1 \pmod{65537},$$

which proves that $65537 = 2^{16} + 1 = 2^{2^4} + 1$ is prime.

The numbers $F_k = 2^{2^k} + 1$ are known as *Fermat numbers*. We just verified that $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$, and $F_4 = 65537$, are prime.

Is F_k prime for any $k \geq 5$? Since 2^{2^k} grows very fast with k, this is a very hard problem.

Euler found that $F_5 = 2^{32} + 1 = 4294967297 = 641 \cdot 6700417$, so F_5 is not prime. We determined that F_2, F_3, F_4 are prime by applying Theorem 4.46 with $a = 3$. Actually, it can be shown that the converse is true, and there is a criterion due to Pepin (1877) to determine whether F_k is prime.

Theorem 4.50. *(Pepin test) For all $k \geq 1$, the Fermat number $F_k = 2^{2^k} + 1$ is prime iff $3^{(F_k - 1)/2} \equiv -1 \pmod{F_k}$.*

Proof. One direction of the Pepin test follows from Theorem 4.46 with $a = 3$. The converse is shown using the Legendre symbol and the law of quadratic reciprocity discussed in Chapter 6. The reader can skip this proof until she/he has read Chapter 6, or at least Sections 6.2 and 6.5.

Assume that F_k is prime, the goal is to prove that

$$\left(\frac{3}{F_k}\right) = -1,$$

where $\left(\frac{3}{F_k}\right)$ is the Legendre symbol. First, observe that if $k \geq 1$, we have

$$2^{2^k} \equiv 2^{2 2^{k-1}} \equiv 4^{2^{k-1}} \equiv 1 \pmod 3.$$

Consequently $F_k = 2^{2^k} + 1 \equiv 2 \pmod 3$. Since $k \geq 1$, we also have $F_k \equiv 1 \pmod 4$. It follows that

$$\frac{3-1}{2}\frac{F_k - 1}{2} \equiv 0 \pmod 2,$$

so by quadratic reciprocity (Theorem 6.13) and since $F_k \equiv 2 \pmod 3$, by Proposition 6.6, since $3 \equiv 3 \pmod 8$, we have

$$\left(\frac{3}{F_k}\right) = \left(\frac{F_k}{3}\right) = \left(\frac{2}{3}\right) = -1.$$

By Euler's criterion (Theorem 6.3), we have

$$3^{\frac{F_k - 1}{2}} \equiv \left(\frac{3}{F_k}\right) \equiv -1 \pmod{F_k},$$

as claimed. □

There is also a nice result restricting the kinds of prime factors that a Fermat number may have.

Theorem 4.51. *(Euler, Lucas) For any integer $n \geq 2$, every prime factor p of $F_n = 2^{2^n} + 1$ must satisfy $p \equiv 1 \pmod{2^{n+2}}$.*

Proof. Let p be a prime factor of F_n, and let h be the least positive integer such that $2^h \equiv 1 \pmod p$. Since p divides $2^{2^n} + 1$, we have $2^{2^n} \equiv -1 \pmod p$, which implies that $h = (2^n)^2 = 2^{n+1}$. By Fermat's little theorem, $2^{p-1} \equiv 1 \pmod p$, so h being the least positive integer such that $2^h \equiv 1 \pmod p$, we deduce that 2^{n+1} divides $p - 1$. Since $n \geq 2$, this implies that $p \equiv 1 \pmod 8$. By Proposition 6.6, we have

$$\left(\frac{2}{p}\right) = 1,$$

so 2 is a quadratic residue modulo p, and by Euler's criterion (Theorem 6.3), we have

$$2^{(p-1)/2} \equiv \left(\frac{2}{p}\right) \equiv 1 \pmod{p},$$

so $h = 2^{n+1}$ divides $(p-1)/2$, which means that $p \equiv 1 \pmod{2^{n+2}}$. □

For $n = 5$, it happens that the prime $p = 641 = 1 + 128 \cdot 5 = 1 + 2^7 \cdot 5$, so by Theorem 4.51, the prime 641 is a factor of the Fermat number $F_5 = 2^{2^5} + 1 = 2^{32} + 1$. This is how Euler found that F_5 was composite.

The Pepin test and Theorem 4.51 have been used to show that $F_5, \ldots, F_{24}$ are composite. It has also been established by other methods that $F_{25}, \ldots, F_{32}$ are composite, but whether any number F_k is prime for $k \geq 33$ is a famous open problem.

It is interesting to observe that any prime factor of F_n is not a prime factor of any of its predecessors $F_0, \ldots, F_{n-1}$. This is because for $n \geq 1$, we have

$$F_0 F_1 \cdots F_{n-1} = F_n - 2,$$

which is easily shown by induction using the fact that

$$\begin{aligned}(F_n - 2)F_n &= F_n^2 - 2F_n \\ &= (2^{2^n} + 1)^2 - 2(2^{2^n} + 1) \\ &= 2^{2^{n+1}} + 22^{2^n} + 1 - 22^{2^n} - 2 \\ &= 2^{2^{n+1}} - 1 = F_{n+1} - 2.\end{aligned}$$

Any prime factor of F_n and F_j for $j < n$ would divide 2, but the F_k are odd, so this is impossible. As a corollary we obtain an amusing proof of the fact that there are infinitely many primes.

For more on Fermat numbers, see Crandal and Pomerance [3] (Chapter 1, Section 1.3.2).

4.7 The Structure of Finite Fields

Suppose K is a field of characteristic p. For every i, with $0 \leq i \leq p$, the binomial coefficient $\binom{p}{i}$ is given by

$$\binom{p}{i} = \frac{p!}{i!(p-i)!},$$

so if $1 \leq i \leq p-1$, we have

$$i\binom{p}{i} = p\binom{p-1}{i-1}.$$

Since $1 \leq i \leq p-1$ and p is prime, we have $\gcd(p, i) = 1$, and so p divides $\binom{p}{i}$.

Proposition 4.52. *If K is a field of characteristic p, the map (Frobenius map) $\sigma\colon K \to K$ given by*

$$\sigma(a) = a^p$$

is an isomorphism of K onto a subfield of K denoted K^p.

Proof. Since K is commutative, it is clear that $\sigma(ab) = \sigma(a)\sigma(b)$. Obviously $\sigma(0) = 0$ and $\sigma(1) = 1$. By the binomial formula and using the fact that p divides $\binom{p}{i}$ for $i = 1, \ldots, p-1$, since K has characteristic p, we have $\binom{p}{i} = 0$ for $i = 1, \ldots, p-1$, so we have

$$\begin{aligned}\sigma(a+b) &= (a+b)^p \\ &= a^p + \sum_{i=1}^{p-1} \binom{p}{i} a^{p-i}b^i + b^p \\ &= a^p + b^p = \sigma(a) + \sigma(b).\end{aligned}$$

Therefore, σ is a homomorphism, and as we remarked earlier, it is injective. □

The field $\mathbb{Z}/p\mathbb{Z}$ with p prime is also denoted by $\mathbb{F}_p$. Here is the structure theorem for finite fields (after J.P. Serre; see Serre [20]).

Theorem 4.53. *Let K be a finite field.*

(i) The field K is of characteristic $p \geq 2$ (p prime). If K is of degree m over $\mathbb{F}_p$, then K has $q = p^m$ elements.

(ii) Let p be any prime, let m be any natural number $m \geq 1$, and write $q = p^m$. For any algebraicaly closed field Ω of characteristic p, there exists a unique subfield $\mathbb{F}_q$ of Ω with q elements. The map $\sigma_q\colon \Omega \to \Omega$ given by $\sigma_q(x) = x^q$ is an automorphism of Ω, and the field $\mathbb{F}_q$ is the set of roots of the polynomial $X^q - X$; that is, $\mathbb{F}_q = \mathrm{Fix}(\sigma_q)$.

(iii) Every finite field with $q = p^m$ elements is isomorphic to $\mathbb{F}_q$.

Proof. (i) Since K is finite, the map $\mathbb{Z} \longrightarrow K$ given by $n \mapsto n \cdot 1$ cannot be injective, so K must have characteristic $p \geq 2$, and it contains $\mathbb{F}_p$ as a subfield. If K has dimension m as a vector space over $\mathbb{F}_p$, then it is obvious that K has p^m elements.

(ii) We know from Proposition 4.52 that the map $\sigma\colon \Omega \to \Omega$ given by $\sigma(x) = x^p$ is an injective homomorphism. Since $\sigma_q = \sigma^m$, the map σ_q is also an injective homomorphism. Since Ω is algebraically closed, for any $a \in K$, the polynomial $X^q - a$ has a root in Ω, which shows that σ_q is also surjective, thus an automorphism of Ω. Then, the field $\mathbb{F}_q$ fixed by σ_q is a subfield of Ω. Since $\mathbb{F}_q$ is also the set of roots of the polynomial $X^q - X$, it has at most q roots. We claim that $F(X) = X^q - X$ has simple roots. From a result of algebra, this is the case if the derivative $F'(X)$ of $F(X)$ is not the zero polynomial. But, since we are in characteristic p and $m \geq 1$, we have

$$F'(X) = qX^{q-1} - 1 = pp^{m-1}X^{q-1} - 1 = -1$$

so $F'(X)$ is not zero. Therefore, $F(X)$ has exactly q roots, and $\mathbb{F}_q$ has $q = p^m$ elements.

If K is any other subfield of Ω with q elements, since the multiplicative group K^* of K is a finite group of order $q-1$, we have

$$x^{q-1} = 1, \quad \text{for all} \quad x \in K^*,$$

and so

$$x^q - x = 0 \quad \text{for all} \quad x \in K,$$

which shows that K is fixed by σ_q, and so $K \subseteq \mathbb{F}_q$. Since $|K| = |\mathbb{F}_q| = q$, we must have $K = \mathbb{F}_q$.

(iii) If K is a finite field with $q = p^m$ elements, then the reasoning in (ii) shows that K is the set of roots of the polynomial $F(X) = X^q - X$. This means that K is the splitting field of $\mathbb{F}_p$ (the smallest field extension of $\mathbb{F}_p$ in which $F(X)$ has all its roots). But, as Ω is algebraically closed and contains a copy of $\mathbb{F}_p$, it contains a splitting field K' of $\mathbb{F}_p$. Since any two splitting fields are isomorphic (see Lang [11], Chapter 5), the field K can be embedded in Ω (as K'), so by (ii) K is isomorphic to $\mathbb{F}_q$. □

Using Theorem 4.22, we obtain the following important result.

Theorem 4.54. *For every prime p and every integer $m \geq 1$, the multiplicative group $\mathbb{F}_{p^m}^*$ of the finite field $\mathbb{F}_{p^m}$ is a cyclic group with $p^m - 1$ elements.*

Proof. For any divisor d of $p^m - 1$, the polynomial $X^d - 1$ has at most d roots in $\mathbb{F}_{p^m}^*$, therefore by Theorem 4.22, the group $\mathbb{F}_{p^m}^*$ is cyclic. □

Any generator of $\mathbb{F}_{p^m}^*$ is called a *primitive root of unity* (to be more precise, a primitive $(p^m - 1)$th root of unity). Observe that the proof of Theorem 4.54 actually shows that *every* finite subgroup of the multiplicative subgroup K^* of any field K is cyclic.

Chapter 5

The Miller–Rabin Test

This chapter is heavily inspired by Dietzfelbinger [4] and Crandall and Pomerance [3]. The Miller–Rabin test makes use of two basic properties of the prime numbers:

(1) *Fermat's little theorem*, which says that if p is a prime and if a is any integer which is not a multiple of p, then

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

Usually, we assume that $1 \leq a \leq p-1$.

(2) If p is a prime, then 1 has only trivial square roots, which means that the only solutions a with $1 \leq a \leq p-1$ of the congruence

$$a^2 \equiv 1 \pmod{p}$$

are $a = 1$ and $a = p-1 \equiv -1 \pmod{p}$.

Property (2) is proved as follows.

Proof. Observe that if $a^2 \equiv 1 \pmod{p}$, then $a^2 - 1 = (a+1)(a-1)$ is divisible by p, and since p is prime, either p divides $a-1$ or p divides $a+1$. Because $1 \leq a \leq p-1$, we conclude that $a = 1$ or $a = p-1$. On the other hand, 1 and $p-1$ are always square roots of unity modulo p (even if p is not prime), since $1^2 \equiv 1 \pmod{p}$ and $(p-1)^2 \equiv (-1)^2 \equiv 1 \pmod{p}$. □

Since computing $a \bmod n$ is cheap, a first approach to test whether a number n is prime is to see if (2) fails. To understand how practical this method is, we need to know how many square roots modulo n the equation $x^2 \equiv 1 \pmod{n}$ has. This is the purpose of the next section.

Recall that if $n \geq 2$, the group $(\mathbb{Z}/n\mathbb{Z})^*$ is the multiplicative group of units of the ring $\mathbb{Z}/n\mathbb{Z}$; that is,

$$(\mathbb{Z}/n\mathbb{Z})^* = \{a \in \mathbb{N} \mid 1 \leq a \leq n-1,\ \gcd(a,n) = 1\}.$$

Here and in several places later, with a slight abuse of notation, we are identifying the equivalence class $\overline{a}$ of a with its representative $a \bmod n$. The order (number of elements) of $(\mathbb{Z}/n\mathbb{Z})^*$ is $\varphi(n)$, where $\varphi(n)$ is the number of integers a, with $1 \leq a \leq n$, which are relatively prime to n ($\gcd(a, n) = 1$).

5.1 Square Roots of Unity

It turns out that 1 and -1 are the only square roots of unity modulo n iff n is of the form 4, p^m, or $2p^m$, where p is an odd prime.[1] To prove this fact, we use the following proposition.

Proposition 5.1. *If p is an odd prime, then there are exactly two square roots of unity modulo p^m and $2p^m$ ($m \geq 1$), namely 1 and -1. There is a unique square root of unity modulo 2 (i.e. 1), two square roots of unity modulo 4 (i.e. ± 1), and four square root of unity modulo 2^m if $m \geq 3$, namely ± 1 and $2^{m-1} \pm 1$.*

Proof. If $x^2 \equiv 1 \pmod{n}$, then x is its own invere modulo n, so $x \bmod n \in (\mathbb{Z}/n\mathbb{Z})^*$, and we may assume that $x \in (\mathbb{Z}/n\mathbb{Z})^*$. First, assume that $n = p^m$ with p an odd prime. In this case, we know that primitive roots modulo p^m exist, so pick one, say g. Then, every $x \in (\mathbb{Z}/p^m\mathbb{Z})^*$ can be written as $x = g^i$, with $1 \leq i \leq \varphi(p^m) = p^{m-1}(p-1)$, and $x^2 \equiv 1 \pmod{p^m}$ is equivalent to $g^{2i} \equiv 1 \pmod{p^m}$. Since g has order $\varphi(p^m)$, the congruence $g^{2i} \equiv 1 \pmod{p^m}$ holds iff $\varphi(p^m) = p^{m-1}(p-1)$ divides $2i$, that is, iff $p^{m-1}(p-1)/2$ divides i (since p is odd). Since $1 \leq i \leq p^{m-1}(p-1)$, there are only two possibilities: $i = p^{m-1}(p-1)/2$ and $i = p^{m-1}(p-1)$, which correspond to $x = -1$ and $x = 1$.

The case $n = 2p^m$ is analogous, since primitive roots also exist and since $\varphi(2p^m) = \varphi(p^m)$.

The cases $n = 2$ and $n = 4$ are clear.

Assume that $n = 2^m$ with $m \geq 3$. We are seeking solutions of the congruence $x^2 \equiv 1 \pmod{2^m}$, with $1 \leq x \leq 2^m - 1$. Note that

$$(2^{m-1} + x)^2 \equiv 2^{2m-2} + 2^m x + x^2 \equiv x^2 \pmod{2^m}$$

since $m \geq 3$. Therefore, it is sufficient to find solutions x such that $1 \leq x \leq 2^{m-1} - 1$. We have $x^2 \equiv 1 \pmod{2^m}$ iff $(x-1)(x+1) \equiv 0 \pmod{2^m}$, so there are three mutally exclusive possibilities:

1. $x \equiv 1 \pmod{2^m}$. Since $1 \leq x \leq 2^m - 1$, we must have $x = 1$.

2. $x \equiv -1 \equiv 2^m - 1 \pmod{2^m}$. Since $1 \leq x \leq 2^{m-1} - 1$, this case is impossible.

3. $x - 1 = h2^i$ and $x + 1 = k2^{m-i}$, with $1 \leq i \leq m - 2$ and $h, k > 0$.

[1] I thank Peter Freyd for communicating this result to me.

In the third case, we deduce that

$$x = h2^{i-1} + k2^{m-i-1}$$
$$1 = k2^{m-i-1} - h2^{i-1}.$$

If $2 \leq i \leq m-2$, then 1 is divisible by 2, which is absurd. Therefore $i = 1$.

If $i = 1$, since $x+1 = k2^{m-i}$, we have $x = k2^{m-1}-1$, and since $k > 0$ and $1 \leq x \leq 2^{m-1}-1$, we must have $k = 1$, so

$$x = 2^{m-1} - 1.$$

Since $m \geq 3$, we have $2^{m-1} - 1 \not\equiv 1 \pmod{2^m}$, and $2^{m-1} - 1$ is a square root of unity distinct from 1.

In summary, we proved that there are exactly two square roots of unity $x = 1$ and $x = 2^{m-1} - 1$ such that $1 \leq x \leq 2^{m-1} - 1$. Since $x + 2^{m-1}$ is also a square root of unity, $2^{m-1} + 1$ and $2^{m-1} + 2^{m-1} - 1 = 2^m - 1 \equiv -1 \pmod{m}$, are also squares roots of unity, so there are exactly four square roots of unity modulo 2^m; namely ± 1 and $2^{m-1} \pm 1$. □

Remark: The fact that there are precisely four square roots of unity modulo 2^m when $m \geq 3$ follows immediately from the fact that $(\mathbb{Z}/2^m\mathbb{Z})^*$ is isomorphic to the direct product of the two cyclic subgroups $\{-1, 1\}$ and $\langle 5 \rangle$, both of even order (see Theorem 4.44).

Now, we can determine the exact number of square roots of unity modulo n.

Theorem 5.2. *For any natural number $n > 1$, if the prime factorization of n is*

$$n = 2^m p_1^{j_1} \cdots p_k^{j_k},$$

where $p_1, \ldots, p_k$ are distinct odd primes and $m + k \geq 1$, then the number s of disinct square roots of unity modulo n is given by

$$s = \begin{cases} 2^k & \text{if } m = 0 \text{ and } k \geq 1 \text{ or } m = 1 \text{ and } k \geq 0 \\ 2^{k+1} & \text{if } m = 2 \text{ and } k \geq 0 \\ 2^{k+2} & \text{if } m \geq 3 \text{ and } k \geq 0. \end{cases}$$

Proof. First, consider the case where $m = 0$. Since $p_1, \ldots, p_k$ are pairwise relatively prime, the congruence $x^2 \equiv 1 \pmod{n}$ is equivalent to the k congruences

$$x^2 \equiv 1 \pmod{p_1^{j_1}}$$
$$\vdots$$
$$x^2 \equiv 1 \pmod{p_1^{j_k}}.$$

From Proposition 5.1, each congruence $x^2 \equiv 1 \pmod{p_1^{j_i}}$ has the two solutions $x = 1$ and $x = -1$ modulo p^{j_i}. By the Chinese remainder theorem, there is a bijection between the

set of solutions x modulo n and the set of k tuples of solutions $(x_1, \ldots, x_k)$ where x_i is a solution modulo p^{j_i}, and since there are 2^k solutions $(x_1, \ldots, x_k)$ with $x_i = \pm 1$, there are 2^s square roots modulo n.

If $k = 0$, then Proposition 5.1 says that the congruence $x^2 \equiv 1 \pmod{2^m}$ has one solution if $m = 1$, two solutions if $m = 2$, and 4 solutions if $m \geq 3$.

If $m \geq 1$ and $k \geq 1$, since $2, p_1, \ldots, p_k$ are pairwise relatively prime, the congruence $x^2 \equiv 1 \pmod{n}$ is equivalent to the $k + 1$ congruences

$$\begin{aligned} x^2 &\equiv 1 \pmod{2^m} \\ x^2 &\equiv 1 \pmod{p_1^{j_1}} \\ &\vdots \\ x^2 &\equiv 1 \pmod{p_1^{j_k}}. \end{aligned}$$

Again, we use the Chinese remainder theorem. Each congruence $x^2 \equiv 1 \pmod{p_1^{j_i}}$ has the two solutions 1 and -1, and the congruence $x^2 \equiv 1 \pmod{2^m}$ has one solution if $m = 1$, two solutions if $m = 2$, and 4 solutions if $m \geq 3$. Therefore, there are 2^k square roots if $m = 1$, $2 \times 2^k = 2^{k+1}$ square roots if $m = 2$, and $4 \times 2^k = 2^{k+2}$ square roots if $m = 3$. □

For example, if $n = 91 = 7 \times 13$, then $27^2 = 729 = 8 \times 91 + 1$, so 27 is a square root of 1 (mod 91). The other nontrivial square root of 1 (mod 91) is 64.

If we find some nontrivial square root of unity a modulo n, then we know that n is composite (and a is a witness to the fact that n is composite). Unfortunately, if n is composite, unless the number k of distinct primes dividing n is large, the number of nontrivial square roots of unity modulo n (at most $2^{k+2} - 2$) is a lot smaller than n, so it is not practical to test a randomly chosen $a \in \{2, \ldots, n-2\}$. Therefore we consider making use of (1) for a more practical test.

5.2 The Fermat Test; F-Witnesses and F-Liars

Going back to (1), observe that if $n \geq 2$, then by the binomial formula, we have

$$(n-1)^{n-1} \equiv (-1)^{n-1} \pmod{n}.$$

Consequently, if $n \geq 3$ is odd, then

$$(n-1)^{n-1} \equiv 1 \pmod{n},$$

and if $n \geq 4$ is even, then

$$(n-1)^{n-1} \equiv -1 \pmod{n}.$$

(In the special case where $n = 2$, we have $1 \equiv -1 \pmod 2$.) Now, for any natural number $n \geq 4$, if $2 \leq a \leq n-2$ and if

$$a^{n-1} \not\equiv 1 \pmod n,$$

then n is not prime, since Fermat's little theorem does not hold. Since all primes except 2 are odd integers, we only need to test odd integers for compositeness and this suggests the following test:

Fermat test: For any odd integer $n \geq 5$, pick randomly some $a \in \{2, \ldots, n-2\}$.

If $a^{n-1} \not\equiv 1 \pmod n$, then return "$n$ is composite," else return "n is a probable prime."

Of course, we compute exponentiation modulo n using fast algorithms based on repeated squaring.

Definition 5.1. Let $n \in \mathbb{N}$ be any integer such that $n \geq 3$.

(1) An integer a such that $2 \leq a \leq n-1$ is called a *Fermat witness*, for short an *F-witness* for n, if $a^{n-1} \not\equiv 1 \pmod n$.

(2) If n is an odd composite, then an integer a with $1 \leq a \leq n-1$ is a *Fermat liar*, for short an *F-liar* for n, if $a^{n-1} \equiv 1 \pmod n$. The set of F-liars for n is denoted by L_n^F.

Every even number $n \geq 4$ has $n-1$ has an F-witness. This is a bit of an overkill, since every positive even number, except 2, is a composite. The number 1 is a trivial F-liar, and by a previous observation, when n is an odd composite, $n-1$ is always an F-liar.

Definition 5.2. A composite number $n \geq 4$ such $a \geq 2$ is an F-liar for n is called a *Fermat pseudoprime base a* (for short, a *pseudoprime base a*).

It can be checked that 2 is an F-witness for all integers $n \geq 3$ up to $n = 340$. However, for $n = 341 = 11 \times 31$, we get

$$2^{340} \equiv 1 \pmod{341},$$

so 2 is an F-liar for 341, and 341 is a pseudoprime base 2. If we try $a = 3$, we find that

$$3^{340} \equiv 56 \pmod{341},$$

so 3 is an F-witness for 341, and 341 is not a pseudoprime base 3. On the other hand, it is easy to check that $91 = 7 \times 13$ is not a pseudoprime base 2 because $2^{90} \equiv 64 \pmod{91}$, but it is a pseudoprime base 3 because $3^{90} \equiv 1 \pmod{91}$.

The above considerations suggest the following question: if $n \geq 3$ is a (odd) composite, does it necessarily have some F-witness? The answer is yes, but this is not of practical use.

Proposition 5.3. *For any integer $n \geq 2$, the following properties hold:*

(a) For any integer a such that $1 \leq a \leq n-1$, if $a^r \equiv 1 \pmod n$ for some $r \geq 1$, then $a \in (\mathbb{Z}/n\mathbb{Z})^*$.

(b) If $a^{n-1} \equiv 1 \pmod{n}$ *for all* a *with* $1 \leq a \leq n-1$*, then* n *is prime.*

Proof. (a) if $r = 1$, then we must have $a = 1$, and if $r \geq 2$, then $a^{r-1}a \equiv 1 \pmod{n}$ shows that a is a unit, so in both cases $a \in (\mathbb{Z}/n\mathbb{Z})^*$.

(b) If $a^{n-1} \equiv 1 \pmod{n}$ for all a with $1 \leq a \leq n-1$, then by (a), we have $(\mathbb{Z}/n\mathbb{Z})^* = \{1, \ldots, n-1\}$, so $\gcd(a, n) = 1$ for all $a = 2, \ldots, n-1$, which implies that n is prime. □

By Proposition 5.3 (b), if $n \geq 4$ is a composite, then it must have some F-witness. Furthermore, by (a), the $n - 1 - \varphi(n)$ elements of the set

$$\{1, \ldots, n-1\} - (\mathbb{Z}/n\mathbb{Z})^* = \{a \in \mathbb{N} \mid 2 \leq a \leq n-2 \mid \gcd(a, n) > 1\}$$

must be all F-witnesses ($a^{n-1} \not\equiv 1 \pmod{n}$).

Unfortunately, this set is very slim for many composite numbers. For example, if $n = pq$ is the product of two distinct primes p and q, then this sets contains $pq - 1 - (p-1)(q-1) = p + q - 2$ elements. If p and q are roughly equal, then $p + q - 2$ is very small in comparison to $n = pq$.

The case $n = 91 = 7 \times 13$ gives us a concrete idea of what is going on. There are 18 F-witnesses not in $(\mathbb{Z}/91\mathbb{Z})^*$ (multiples of 7 and 13):

$$7, 14, 21, 28, 35, 42, 49, 56, 63, 70, 77, 84$$
$$13, 26, 39, 52, 65, 78.$$

There are 36 F-witnesses in $(\mathbb{Z}/91\mathbb{Z})^*$:

$$2, 5, 6, 8, 11, 15, 18, 19, 20, 24, 31, 32,$$
$$33, 34, 37, 41, 44, 45, 46, 47, 50, 54, 57, 58,$$
$$59, 60, 67, 71, 72, 73, 76, 80, 83, 85, 86, 89.$$

Finally, there and 36 F-liars (necessarily in $(\mathbb{Z}/91\mathbb{Z})^*$);

$$1, 3, 4, 9, 10, 12, 16, 17, 22, 23, 25, 27,$$
$$29, 30, 36, 38, 40, 43, 48, 51, 53, 55, 61, 62,$$
$$64, 66, 68, 69, 74, 75, 79, 81, 82, 87, 88, 90.$$

The Fermat test gives the wrong answer if the random choice for a hits one the 34 F-liars other and 1 and 90, which has probabilty $34/88 = 17/44$. Observe that $17/34 < 1/2$. This is a general fact, provided that the odd composite n has some F-witness in $(\mathbb{Z}/n\mathbb{Z})^*$. This follows from the interesting fact that the set L_n^F of F-liars is a subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$.

Proposition 5.4. *For any integer* $n \geq 2$*, the set* L_n^F *of* F*-liars is a subgroup of* $(\mathbb{Z}/n\mathbb{Z})^*$*. Furthermore, if* n *is an odd composite and if* n *possesses at least some* F*-witness in* $(\mathbb{Z}/n\mathbb{Z})^*$*, then the probability that the Fermat test gives the wrong answer, more precisely the probability that any* $a \in \{2, \ldots, n-2\}$ *is an* F*-liar for* n*, is at most* $1/2$*.*

Proof. Since $1 \equiv 1 \pmod{n}$, we have $1 \in L_n^F$. Since $(\mathbb{Z}/n\mathbb{Z})^*$ is a finite group, to show that L_n^F is a subgroup, it suffices to show closure under multiplication. If $a^{n-1} \equiv 1 \pmod{n}$ and $b^{n-1} \equiv 1 \pmod{n}$, then $(ab)^{n-1} \equiv a^{n-1}b^{n-1} \equiv 1 \pmod{n}$, as desired.

By Lagrange's theorem, the order $|L_n^F|$ of L_n^F divides the order $\varphi(n)$ of $(\mathbb{Z}/n\mathbb{Z})^*$. If there is some F-witness in $(\mathbb{Z}/n\mathbb{Z})^*$, then L_n^F is a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$. Since n is a composite and since L_n^F is a proper subgroup, we deduce that $\varphi(n) < n-1$ and that $|L_n^F|$ is a proper divisor of $\varphi(n)$, which implies that

$$|L_n^F| \leq (n-2)/2.$$

Thus, the probability that some a chosen in $\{2, \ldots, n-2\}$ belongs to $L_n^F - \{1, n-1\}$ is bounded by

$$\frac{(n-2)/2 - 2}{n-3} = \frac{n-6}{2(n-3)} < \frac{1}{2},$$

since $2n - 12 < 2n - 6$. □

The good news about Proposition 5.4 is that if n is an odd composite and if n has some F-witness in $(\mathbb{Z}/n\mathbb{Z})^*$, then the probability that the Fermat test gives the wrong answer is less than $1/2$. By repeating the test ℓ times, each time choosing randomly and independently some a in $\{2, \ldots, n-2\}$, we can make the probability of failure less than $(1/2)^\ell$.[2]

5.3 Carmichael Numbers

The bad news is that there exist odd composites n such that $L_n^F = (\mathbb{Z}/n\mathbb{Z})^*$; that is, n has no F-witness in $(\mathbb{Z}/n\mathbb{Z})^*$. The smallest such number is $561 = 3 \times 11 \times 17$. This number is a pseudoprime in any base relatively prime to 561. Such "nasty" numbers were first discovered by R. Carmichael in 1910, and motivates the following definition.

Definition 5.3. An integer $n \geq 3$ for which $a^{n-1} \equiv 1 \pmod{n}$ for all $a \in \{2, \ldots, n-1\}$, with $\gcd(a, n) = 1$, is called a *probable prime*. A composite integer $n \geq 3$ which is a probable prime is called a *Carmichael number*.

If $n \geq 4$ is even, we observed that $n - 1$ is an F-witness for n, so a Carmichael number must be odd.

[2]We have to be careful about which probability we are talking about. In this case, we are considering the conditional probability that the algorithm lies ℓ times (fails to report that n is composite), given that n is composite. However, as a user of the algorithm, it is more useful to know the conditional probability that n is composite, given that the algorithm runs ℓ times and each time fails to report that n is composite. The two conditional probabilities are related by Bayes's formula. The second conditional probability involves the density of primes. A computation shows that the probability $(1/2)^\ell$ must be (approximately) multiplied by $\ln n$. We will come back to this point later on.

Unfortunately for primality testing, there are infinitely many Carmichael numbers. This fact was proved in 1994 by Alford, Granville and Pomerance. They also proved that there is some integer $x_0 > 0$, such that for all $x \geq x_0$, the number $C(x)$ of Carmichael numbers not exceeding x satisfies $C(x) > x^{2/7}$.

Remark: The sufficiently large x_0 is not known explicitly, but it is conjectured that it is the 96th Carmichael number: 8719309.

Other authors define a Carmichael number as a composite integer $n \geq 3$ for which

$$a^n \equiv a \pmod{n} \quad \text{for all } a \in \mathbb{N}.$$

This second definition implies the first (Definition 5.3), because if $a^n \equiv a \pmod{n}$ and if $\gcd(a, n) = 1$, then we can divide by a and we obtain $a^{n-1} \equiv 1 \pmod{n}$. Definition 5.3 implies the second definition, but this requires a little work. We can use of a criterion due to A. Korselt. This criterion was found in 1899, eleven years before Carmichael actually produced the first example. Presumably Korselt believed that such numbers did not exist, and he developed a criterion as a first step in proving this.

Theorem 5.5. *(Korselt criterion) An integer $n \geq 2$ is a Carmichael number iff the following two conditions hold.*

(1) The number n is composite and not divisible by the square of any prime (it is square-free).

(2) For every prime p, if p divides n then $p - 1$ divides $n - 1$.

Proof. First, let n be a Carmichael number.

(1) Assume that n is divisible by the square of some prime p. Since n must be odd, we can write $n = p^k m$, where $p \geq 3$ is a prime, $k \geq 2$, and p does not divide m. We produce an F-witness in $(\mathbb{Z}/n\mathbb{Z})^*$ for n as follows:

Case 1. If $m = 1$, let $a = p + 1$. Clearly, $\gcd(p + 1, p^k) = 1$. We claim that $a^{n-1} \not\equiv 1 \pmod{n}$. We proceed by contradiction. If $a^{n-1} \equiv 1 \pmod{n}$, then since p^2 divides n, we have $a^{n-1} \equiv 1 \pmod{p^2}$. However, by the binomial formula, we have

$$a^{n-1} \equiv (1 + p)^{n-1} \equiv 1 + (n - 1)p + \sum_{i=2}^{n-1} \binom{n-1}{i} p^i \equiv 1 + (n - 1)p \pmod{p^2}.$$

Since $a^{n-1} \equiv 1 \pmod{p^2}$, we deduce that $(n - 1)p \equiv 0 \pmod{p^2}$, which means that p^2 divides $(n - 1)p$, and since p is prime, p divides $n - 1$. However, $n - 1 = p^k - 1$ with $k \geq 2$, so p does not divide $n - 1$, a contradiction.

Case 2. If $m \geq 3$, then we use the Chinese remainder theorem to find some a with $1 \leq a < p^2 m \leq n$ so that

$$\begin{aligned} a &\equiv p + 1 \pmod{p^2} \\ a &\equiv 1 \pmod{m}. \end{aligned}$$

Since p^2 divides $a-(p+1)$, the prime p does not divide a, so $\gcd(a,p^k)=1$. Since $a \equiv 1 \pmod{m}$, we also have $\gcd(a,m)=1$. Because $\gcd(p^k,m)=1$ and $n=p^km$, we conclude that $\gcd(a,n)=1$. We claim that $a^{n-1} \not\equiv 1 \pmod{n}$. As in Case 1, we proceed by contradiction. Then, by the same reasoning, we deduce that p divides $n-1$. This time, $n-1=p^km-1$, and again p does not divide $n-1$, a contradiction.

(2) By (1), n is a product of distinct primes. Assume that the prime p divides n. Since p is prime, the group $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic (see Theorem 4.34) so pick a generator g (a primitive root modulo p). By the Chinese remainder theorem, we can find some b such that

$$\begin{aligned} b &\equiv g \pmod{p} \\ b &\equiv 1 \pmod{n/p}. \end{aligned}$$

Since n is a product of distinct primes, the numbers p and n/p have no common factor, so $\gcd(b,n)=1$. Since n is a Carmichael number, we have

$$b^{n-1} \equiv 1 \pmod{n},$$

and since p divides n, we get

$$g^{n-1} \equiv b^{n-1} \equiv 1 \pmod{p}.$$

Since g has order $p-1$, the number $p-1$ divides $n-1$.

Conversely, assume that (1) and (2) hold. Let $a \in \{1,\ldots,n-1\}$ be any integer such that $\gcd(a,n)=1$. For any prime p dividing n, we also have $\gcd(a,p)=1$, so by Fermat's little theorem,

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

Since $p-1$ divides $n-1$, we also have

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

Since the prime factors of n are all distinct, we deduce that

$$a^{n-1} \equiv 1 \pmod{n},$$

which shows that n is a Carmichael number. □

Suppose n is a Carmichael number. Then by Korselt's criterion (Proposition 5.5) Conditions (1) and (2) hold. If $\gcd(a,p)=1$, then the proof of (2) shows that $a^{n-1} \equiv 1 \pmod{p}$, thus $a^n \equiv a \pmod{p}$. If p divides a then $a^n \equiv a \pmod{p}$ holds trivially (since p divides n). Therefore $a^n \equiv a \pmod{p}$ for all a, and since the prime factors of n are all distinct, we conclude that

$$a^n \equiv a \pmod{n} \quad \text{for all } n \in \mathbb{N}.$$

We saw in Theorem 5.5 that every Carmichael numbers contains distinct prime factors. The number of distinct prime factors must be at least three.

Proposition 5.6. *Every Carmichael number contains at least three distinct prime factors.*

Proof. We make use of Theorem 5.5. Assume that some Carmichael number n is the product of two distinct primes p and q. We may suppose that $3 \leq p < q$. By Theorem 5.5, Property (2) says that $p-1$ and $q-1$ both divide $n-1$. But, $n-1 = pq-1 = p(q-1)+p-1$, so $n-1 \equiv p-1 \pmod{(q-1)}$, and $p-1 \not\equiv 0 \pmod{(q-1)}$ since $q > p \geq 3$, a contradiction. □

Here is a list of the first ten smallest Carmichael numbers; see Ribenboim [18] (Chapter 2, Section IX):

$$\begin{aligned}
651 &= 3 \cdot 11 \cdot 17 \\
1105 &= 5 \cdot 13 \cdot 17 \\
1729 &= 7 \cdot 13 \cdot 19 \\
2465 &= 5 \cdot 17 \cdot 29 \\
2821 &= 7 \cdot 13 \cdot 31 \\
6601 &= 7 \cdot 23 \cdot 41 \\
8911 &= 7 \cdot 19 \cdot 67 \\
10585 &= 5 \cdot 29 \cdot 73 \\
15841 &= 7 \cdot 31 \cdot 73 \\
29341 &= 13 \cdot 37 \cdot 61.
\end{aligned}$$

If n is a Carmichael number, then $L_n^F = (\mathbb{Z}/n\mathbb{Z})^*$, so the set $\{1, \ldots, n-1\} - (\mathbb{Z}/n\mathbb{Z})^*$ of F-witnesses is quite thin, and the probability that the Fermat test gives the wrong answer (n is prime) is

$$\frac{\varphi(n)-2}{n-3} > \frac{\varphi(n)}{n} = \prod_{\substack{p \text{ is prime} \\ p|n}} \left(1 - \frac{1}{p}\right).$$

This bound is annoyingly close to 1 if n has only few large prime factors. For example, if n is the Carmichael number

$$n = 651693055693681 = 72931 \times 87517 \times 102103,$$

we find that $\varphi(n)/n > 0.99996$. Repeating the test does not help, because if n has only 3 or 4 factors and if the smallest prime factor is p_0, then it is not hard to see that we would have to repeat the test a number of times proportional to p_0 to make the error probability less than $1/2$. Therefore, a new idea is necessary to break the curse of Carmichael numbers.

5.4 The Miller–Rabin Test; *MR*-Witnesses and *MR*-Liars

The new idea is to make use of the nontrivial square root of unity test. If $n \geq 3$ is an odd integer, we can factor the largest power of 2 in $n-1$; that is, we write

$$n - 1 = 2^k t,$$

where t is odd. The point is that if p is prime, then for any a which is not a multiple of p, the residues of a^t and $a^{2^i t}$ (with $0 \leq i \leq k-1$) modulo p must satisfy some special condition.

Proposition 5.7. *Let p be an odd prime, and write*

$$p - 1 = 2^k t, \quad \textit{with } t \textit{ odd and } k \geq 1.$$

For any natural number a which is not a multiple of p, one of the following two conditions must hold:

(1) either $a^t \equiv 1 \pmod{p}$,

(2) or $a^{2^i t} \equiv n - 1 \pmod{p}$, *for some i with* $0 \leq i \leq k-1$.

Proof. By Fermat's little theorem, we have

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

that is

$$a^{2^k t} \equiv 1 \pmod{p}.$$

This implies that if we consider the list

$$b_0 = a^t, b_1 = a^{2t}, b_2 = a^{2^2 t}, \ldots, b_{k-1} = a^{2^{k-1} t}, b_k = a^{2^k t} = a^{n-1},$$

the last number is congruent to 1 modulo p, and since

$$a^{2^{i+1} t} = \left(a^{2^i t}\right)^2,$$

we have $b_{i+1} = b_i^2$, for $i = 0, \ldots, k-1$. There are only two possibilities:

(i) We have $b_0 = a^t \equiv 1 \pmod{p}$.

(ii) There is some b_i such that $b_i \not\equiv 1 \pmod{p}$, but $b_i^2 \equiv 1 \pmod{p}$, for some i with $0 \leq i \leq k-1$. Because n is prime, we know that $b_i^2 \equiv 1 \pmod{p}$ implies that $b_i \equiv \pm 1 \pmod{p}$, and since $+1$ is ruled out, we must have $b_i \equiv -1 \equiv n - 1 \pmod{p}$.

Case (i) corresponds to Case (1) and Case (ii) corresponds to Case (2). □

Observe that Condition (2) for $i = 0$ says that $a^t \equiv n - 1 \pmod{n}$, or equivalently $a^t \equiv -1 \pmod{n}$. Proposition 5.7 implies that if we can find some natural number a such that

(a) $a^t \not\equiv \pm 1 \pmod{n}$, and

(b) $a^{2^i t} \not\equiv n - 1 \pmod{n}$, for all i with $1 \leq i \leq k - 1$,

then n must be a composite.

Indeed, if a satisfies both properties (a) and (b) above, then a is not a multiple of n, since otherwise we would have $a^t \equiv 0 \pmod{n}$ and $a^{2^i t} \equiv 0 \pmod{n}$ for all i with $1 \leq i \leq k - 1$, so (a) and (b) would not hold. Then, by the contrapositive of Proposition 5.7, the number n can't be prime.

Clearly, $a \neq 1$, but $a \neq n-1$ as well, since $(n-1)^t \equiv -1$, because t is odd. Furthermore, if $a \in \{2, \ldots, n-2\}$ then we may assume that $\gcd(a, n) = 1$, since otherwise $2 \leq \gcd(a, n) \leq n - 2$, and n is composite.

The above leads to to the following definition.

Definition 5.4. Let $n \geq 3$ be any odd integer, and write $n - 1 = 2^k t$, with $k \geq 1$ and t odd.

(1) A number a such that $2 \leq a \leq n - 2$ with $\gcd(a, n) = 1$ is a *Miller–Rabin witness*, for short a *MR-witness* for n, if the following two conditions hold:

   (a) $a^t \not\equiv \pm 1 \pmod{n}$, and

   (b) $a^{2^i t} \not\equiv n - 1 \pmod{n}$, for all i with $1 \leq i \leq k - 1$.

(2) If n is composite, then any integer a with $1 \leq a \leq n - 1$ is *Miller–Rabin liar*, for short an *MR-liar* for n, iff a is not an *MR*-witness for n. The set of *MR*-liars for n is denoted by L_n^{MR}, and we have

$$\begin{aligned} L_n^{MR} = \{a \in \{1, \ldots, n-1\}, &\text{either } a^t \equiv 1 \pmod{n}, \\ &\text{or } a^{2^i t} \equiv n - 1 \pmod{n}, \text{ for some } i \text{ with } 0 \leq i \leq k - 1\}. \end{aligned}$$

The numbers $a = 1$ and $a = n - 1$ are trivial *MR*-liars. Observe that every *MR*-liar is an *F*-liar: If $a^t \equiv 1 \pmod{n}$, then

$$a^{n-1} \equiv \left(a^t\right)^{2^k} \equiv (1)^{2^k} \equiv 1 \pmod{n},$$

and if $a^{2^i t} \equiv n - 1 \pmod{n}$, for some i with $0 \leq i \leq k - 1$, then

$$a^{n-1} \equiv \left(a^{2^i t}\right)^{2^{k-i}} \equiv (-1)^{2^{k-i}} \equiv 1 \pmod{n},$$

since $i \leq k - 1$.

Thus, $L_n^{MR} \subseteq L_n^F$, but unfortunately, L_n^{MR} is not a group. For example, if $n = 325 = 5^2 \times 13$, then $n - 1 = 2^2 \times 81$, and it is easy to verify that

$$\begin{aligned} 7^{2\times 81} &\equiv 324 \pmod{325} \\ 32^{2\times 81} &\equiv 324 \pmod{325} \\ 224^{81} &\equiv 274 \pmod{325} \\ 224^{2\times 81} &\equiv 1 \pmod{325} \\ 224^{2^2\times 81} &\equiv 1 \pmod{325}, \end{aligned}$$

so 7 and 32 are both *MR*-liars, but their product 224 is a *MR*-witness. When n is not a Carmichael number, L_n^{MR} is contained in L_n^F which is a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$, so the proportion of *MR*-liars is less than $1/2$, but when n is a Carmichael number, we need to find another proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$ containing L_n^{MR}. Fortunately, this is possible.

Definition 5.5. An odd composite number n such that a with $2 \leq a \leq n-2$ is an *MR*-liar for n is called a *strong pseudoprime base* a.

Because every *MR*-liar is an F-liar, every strong pseudoprime base a is a pseudoprime base a. The converse is false.

The number 91 is an example of a pseudoprime base 10 which is also a strong pseudoprime base 10. Indeed, $90 = 2 \times 45$, and $10^{45} \equiv 90 \pmod{91}$, which shows that 10 is an *MR*-liar.

We saw earlier that $n = 341$ is a pseudoprime base 2. But 341 is not a strong pseudoprime base 2, because $340 = 2^2 \times 85$, $2^{85} \equiv 32 \pmod{341}$, and $2^{2\times 85} \equiv 1 \pmod{341}$, so 2 is an *MR*-witness for 341. In fact, 32 is a nontrivial square root of unity modulo 341.

The Carmichael number $n = 561 = 3 \times 11 \times 17$ is a pseudoprime for every base relatively prime to 561, and $560 = 2^4 \times 35$. For $a = 2$, we obtain

$$\begin{aligned} 2^{35} &\equiv 263 && \pmod{561} \\ 2^{2\times 35} &\equiv 263^2 \equiv 166 && \pmod{561} \\ 2^{2^2\times 35} &\equiv 166^2 \equiv 67 && \pmod{561} \\ 2^{2^3\times 35} &\equiv 67^2 \equiv 1 && \pmod{561} \end{aligned}$$

Since $263 \not\equiv \pm 1 \pmod{561}$, and $166, 67, 1 \not\equiv 560 \pmod{561}$, the number 2 is an *MR*-witness for 561, which is not a strong pseudoprime base 2.

Here is another example from Hoffstein, Pipher and Silverman [8] (Chapter 3, Section 3.4). We leave it as an exercise to check that if $n = 172947529$, then $n - 1 = 2^3 \times 21618441$, and with $a = 17$, we get

$$17^{21618441} \equiv 1 \pmod{172947529},$$

so 17 is not an *MR*-witness for 172947529. With $a = 3$, we get

$$3^{21618441} \equiv -1 \pmod{172947529},$$

and 3 is not an *MR*-witness for 172947529 either. However, with $a = 23$ we have

$$\begin{aligned} 23^{21618441} &\equiv 40063806 && \pmod{172947529} \\ 23^{2\cdot 21618441} &\equiv 2257065 && \pmod{172947529} \\ 23^{4\cdot 21618441} &\equiv 1 && \pmod{172947529} \end{aligned}$$

so $a = 23$ is an *MR*-witness for 172947529, which happens to be a Carmichael number with factorization

$$172947529 = 307 \times 613 \times 919.$$

The reader should check that the above number is indeed a Carmichael number by using Korselt's criterion.

Ribenboim gives an interesting table of the numbers below $N = 25 \cdot 10^9$ which are strong pseudoprimes to the bases $2, 3, 5$ simultaneously; see [18], Chapter 2, Section VIII, Table 11.

In the table below, the abbrevation psp means pseudoprime, and the abbreviation spsp means strong pseudprime.

psp to bases				
Number	7	11	13	Factorization
25 326 001	no	no	no	$2251 \cdot 11251$
161 304 001	no	spsp	no	$7333 \cdot 21997$
960 946 321	no	no	no	$11717 \cdot 82013$
1 157 839 381	no	no	no	$24061 \cdot 48121$
3 215 031 751	spsp	psp	psp	$151 \cdot 751 \cdot 28351$
3 697 278 427	no	no	no	$30403 \cdot 121609$
5 764 643 587	no	no	spsp	$37963 \cdot 151849$
6 770 862 367	no	no	no	$41143 \cdot 164569$
14 386 156 093	psp	psp	psp	$397 \cdot 4357 \cdot 8317$
15 579 919 981	psp	spsp	no	$88261 \cdot 176521$
18 459 366 157	no	no	no	$67933 \cdot 271729$
19 887 974 881	psp	no	no	$81421 \cdot 244261$
21 276 028 621	no	psp	psp	$103141 \cdot 206281$

There are $\pi(N) = 1\,091\,987\,405$ primes and 2163 Carmichael numbers less that or equal to $N = 25 \cdot 10^9$. The number 3 215 031 751 has the remarkable property that it is a strong pseudoprime simulaneously to the bases $2, 3, 5, 7$. Consequently, if N is any positive integer such that $N < 25 \cdot 10^9$ and $N \neq 3\,215\,031\,751$, and if $2, 3, 5, 7$ are not *MR*-witnesses, then N is prime. This can be checked very quickly using repeated squaring. According to Ribenboim, this is also true up to $N < 118\,670\,087\,467$; see [18] (Chapter 2, Section XI) and Niven, Zuckerman, and Montgomery [16] (Section 2.4).

The idea to use the sequence $b_0, \ldots, b_k$ of Proposition 5.7 to design a test for compositeness was suggested around 1976 by J. Selfridge. Also around 1976, G. Miller designed

a deterministic test whose polynomial running time depends on the truth of the Extended Riemann Hypothesis (for short, ERH), a yet famous unproved number-theoretic conjecture. We will say a little more about it later. Some years later, around 1980, M. Rabin (and independently L. Monier) found a way of making Miller's test into a randomized algorithm. This algorithm is now know as the *Miller–Rabin* test. Here it is.

Miller–Rabin test

The input is an integer $n > 3$.

```
procedure miller-rabin(n)
begin
  Choose random integer a ∈ {2, ..., n − 2};
  If n is even or gcd(a, n) ≠ 1 then c := 1; return c; exit (* n is composite *)
  Decompose n as n − 1 = 2^k t, with t odd
  b := a^t mod n;
  if b = 1 or b = n − 1 then c := 0; return c; exit;
    (* n is a strong pseudoprime base a *)
  for i = 1 to k − 1 do
    b := b^2 mod n;
    if b = n − 1 then c := 0; return c; exit
    (* n is a strong pseudoprime base a *)
    if b = 1 then c := 1; return c; exit (* n is composite *)
  endfor ;
  c := 1; return c (* n is composite *)
end
```

We need to show that the algorithm behaves correctly; that is, we need to show that n is indeed composite when it returns the output $c = 1$ ("composite"). If n is even or if $\gcd(a, n) \neq 1$, then n is obviously composite. Othwerwise n is odd, and there are two ways that the algorithm returns the output $c = 1$. Let $b_0 = a^t \bmod n$ and $a_i = a^{2^i t} \bmod n$, for $i = 1, \ldots, k$.

(a) For some i, $1 \leq i \leq k-1$, the algorithm finds that $b = 1$. In order to reach this condition, it must be the case that $b_0, b_1, \ldots, b_{i-1} \notin \{1, n-1\}$, since otherwise the program would have stopped. As soon as $b_i = 1$, we also have $b_{i+1} = \ldots = b_k = 1$. But then, $b_0 \notin \{1, n-1\}$ and $b_i \neq n-1$ for $i = 1, \ldots, k-1$, so a is an *MR*-witness and n is indeed composite. Actually, $b_{i-1}^2 = 1$ and $b_i \notin \{1, n-1\}$, so n has a nontrivial square root and thus must be composite.

(b) The program goes through all $k-1$ rounds through the **for** loop and returns $c = 1$ ("composite"). In this case, all the tests (in the **if** statements) have failed, and we must have $b_i \notin \{1, n-1\}$ for $i = 0, \ldots, k-1$. Again a is an *MR*-witness and n is composite.

The computational complexity of this algorithm depends on what kind of fast algorithm is used to compute exponentiation modulo n. As explained in Dietzfelbinger [4] (Chapter 5, Section 5.2), it takes $O(\log_2 n)$ arithmetic operations and $O((\log_2 n)^3)$ bit operations. If a faster method is used for integer multiplication, then it takes $O^{\sim}((\log_2 n)^2)$ bit operations. Here, the notation $f = O^{\sim}(g)$ means that $f = O(g(\log_2(g))^k)$, for some $k \geq 0$; for details, see Dietzfelbinger [4] (Chapter 2, Sections 2.2 and 2.3). In brief, the Miller–Rabin test is polynomial in the bit length of the input n (of degree at most 3).

It remains to show that the probability that the Miller–Rabin test gives the wrong answer, "strong pseudoprime," when n is a composite, is less than $1/2$. Monier and Rabin proved that this probability is actually less than $1/4$, but for now, we show that this probability is less than $1/2$ because the proof is simpler. We follow the nice proof given in Dietzfelbinger [4] (Chapter 5).

We need to find an upper bound on $|L_n^{MR}|$. As we explained earlier, the set L_n^{MR} of *MR*-liars is contained in L_n^F, but it is not a subgroup.

If n is not a Carmichael number, then L_n^F is a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$, so the reasoning used in the proof of Proposition 5.4 applies, and the fraction of *MR*-liars in $\{2, \ldots, n-2\}$ is less than $1/2$.

If n is a Carmichael number, then we can find a proper subgroup B_n of $(\mathbb{Z}/n\mathbb{Z})^*$ that contains L_n^{MR} as follows. Write $n-1 = 2^k t$, with t odd. Since t is odd, we have $(n-1)^t \equiv n-1 \pmod{n}$, so there is a largest index $i \geq 0$ such that there is an *MR*-liar a_0 (recall that $a_0 \in \{1, \ldots, n-1\}$) with

$$a_0^{2^i t} \equiv n-1 \pmod{n}.$$

Denote this largest index by i_0. Since n is a Carmichael number, we have

$$a_0^{2^k t} \equiv a_0^{n-1} \equiv 1 \pmod{n},$$

hence $0 \leq i_0 \leq k-1$. Define B_n by

$$B_n = \{a \in \{1, \ldots, n-1\} \mid a^{2^{i_0} t} \bmod n \in \{1, n-1\}\}.$$

The following proposition is the key ingredient.

Proposition 5.8. *The set B_n defined above (for a Carmichael number n) has the following properties:*

(1) $L_n^{MR} \subseteq B_n$.

(2) The set B_n is a subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$.

(3) The group B_n is a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$.

Proof. (1) Pick any $a \in L_n^{MR}$.

Case 1: $a^t \equiv 1 \pmod{n}$. Then, $a^{2^{i_0}t} \equiv 1 \pmod{n}$ as well, and thus $a \in B_n$.

Case 2: $a^{2^i t} \equiv n-1 \pmod{n}$, for some i with $0 \leq i \leq k-1$. By the maximality of i_0, we have $i \leq i_0$. If $i = i_0$, then we get immediately that $a \in B_n$. If $i < i_0$, then

$$a^{2^{i_0}t} \equiv \left(a^{2^i t} \bmod n\right)^{2^{i_0-i}} \equiv 1 \pmod{n}$$

since $i_0 - i \geq 1$, and we conclude that $a \in B_n$.

(2) Obviously, $1 \in B_n$. Since $(\mathbb{Z}/n\mathbb{Z})^*$ is finite, we only need to check that B_n is closed under multiplication. Pick $a, b \in B_n$. This means that $a^{2^{i_0}t} \bmod n$ and $b^{2^{i_0}t} \bmod n$ belong to $\{1, n-1\}$. Now, we have

$$\begin{aligned} 1 \cdot 1 &\equiv 1 \pmod{n} \\ 1 \cdot (n-1) &\equiv n-1 \pmod{n} \\ (n-1) \cdot 1 &\equiv n-1 \pmod{n} \\ (n-1) \cdot (n-1) &\equiv 1 \pmod{n}, \end{aligned}$$

which implies that

$$(ab)^{2^{i_0}t} \equiv a^{2^{i_0}t} b^{2^{i_0}t} \equiv 1 \text{ or } n-1 \pmod{n};$$

that is, $ab \in B_n$, as required.

(3) We need to find some $a \in (\mathbb{Z}/n\mathbb{Z})^*$ which does not belong to B_n. We use the fact that a Carmichael number can be written as the product $n = n_1 n_2$ of two distinct odd numbers n_1 and n_2 such that $\gcd(n_1, n_2) = 1$, which is a consequence of the fact that a Carmichael number is a product of (at least 3) distinct primes.

Recall that a_0 is an *MR*-liar with $a_0^{2^{i_0}t} \equiv n-1 \pmod{n}$. Since n_1 and n_2 divide n, we also have $a_0^{2^{i_0}t} \equiv n-1 \pmod{n_1}$ and $a_0^{2^{i_0}t} \equiv n-1 \pmod{n_2}$. If we let $a_1 \equiv a_0 \bmod n_1$, then by the Chinese remainder theorem, there is a unique $a \in \{0, \ldots, n-1\}$ such that

$$\begin{aligned} a &\equiv a_1 \pmod{n_1} \\ a &\equiv 1 \pmod{n_2}. \end{aligned}$$

We claim that $a \in (\mathbb{Z}/n\mathbb{Z})^*$, but $a \notin B_n$. First, we show that $a \notin B_n$.

From $a_1 \equiv a_0 \pmod{n_1}$ and $a \equiv a_1 \pmod{n_1}$, we get $a \equiv a_0 \pmod{n_1}$, and since $a_0^{2^{i_0}t} \equiv n-1 \pmod{n_1}$, we get

$$a^{2^{i_0}t} \equiv n-1 \pmod{n_1}. \tag{$*$}$$

From $a \equiv 1 \pmod{n_2}$, we get

$$a^{2^{i_0}t} \equiv 1 \pmod{n_2}. \tag{$**$}$$

Now, since n_1 divides n, $(*)$ implies that

$$a^{2^{i_0}t} \not\equiv 1 \pmod{n},$$

and since n_2 divides n, $(**)$ implies that

$$a^{2^{i_0}t} \not\equiv n-1 \pmod{n}.$$

Consequently, $a^{2^{i_0}t} \bmod n \notin \{1, n-1\}$, and thus $a \notin B_n$. It remains to show that $a \in (\mathbb{Z}/n\mathbb{Z})^*$. By squaring $(*)$ and $(**)$, we get

$$a^{2^{i_0+1}t} \equiv 1 \pmod{n_1} \quad \text{and} \quad a^{2^{i_0+1}t} \equiv 1 \pmod{n_2},$$

and since $\gcd(n_1, n_2) = 1$, this yields

$$a^{2^{i_0+1}t} \equiv 1 \pmod{n},$$

which shows that $a \in (\mathbb{Z}/n\mathbb{Z})^*$ (by Proposition 5.3). □

Proposition 5.8 shows that if n is a Carmichael number, then L_n^{MR} is contained in the proper subgroup B_n of $(\mathbb{Z}/n\mathbb{Z})^*$, and by the reasoning used when L_n^{MR} is contained in the proper subgroup L_n^F of $(\mathbb{Z}/n\mathbb{Z})^*$, we conclude that the fraction of *MR*-liars in $\{2, \ldots, n-2\}$ is also less than $1/2$. In summary, we proved the following result.

Theorem 5.9. *If $n > 3$ is an odd composite, then the fraction of MR-liars in $\{2, \ldots, n-2\}$ is less than $1/2$. Consequently, the probability that the Miller–Rabin test gives the wrong answer $c = 0$, more precisely the probability that any $a \in \{2, \ldots, n-2\}$ is a MR-liar, is less than $1/2$.*

We will show below that if we repeat the Miller–Rabin test ℓ times, then the probability that the algorithm gives the answer $c = 0$ every time when n is composite is at most $(1/2)^\ell$. Here is the algorithm.

Algorithm Probable Prime

The input is an integer $n > 3$.

```
begin
  for i = 1 to ℓ do
    c = miller-rabin(n);
    if c = 1 then res = 1; return res; exit (* n is composite *)
  endfor ;
  res = 0; return res (* n is a probable prime *)
end
```

We can easily modify miller-rabin to avoid redecomposing n again and again when it is called.

If the algorithm stops with $res = 1$, then n is definitely composite. If n is prime, then the algorithm will runs through all ℓ steps and correctly return $res = 0$. If n is composite, the algorithm may return the wrong answer $res = 0$, but the probability that this happens is bounded by $(1/2)^\ell$.

We hinted earlier at the fact that the above probabilily is not really what we would like to know. To make this point clearer, let us define the events P, C, SP and SP_ℓ by

$$\begin{aligned} P &= \{n \mid n \geq 3,\ n \text{ is prime}\} \\ C &= \{n \mid n \geq 3,\ n \text{ is composite}\} \\ SP &= \{n \mid n \geq 3, \text{ the miller-rabin procedure returns } c = 0\} \\ SP_\ell &= \{n \mid n \geq 3, \text{ the miller-rabin procedure returns } c = 0 \ \ell \text{ times}\}. \end{aligned}$$

Observe that $P = \overline{C}$, the complement of C. Then, we have the three conditional probabilities

$$\begin{aligned} \mathsf{Pr}(P \mid SP_\ell) &= \mathsf{Pr}(n \text{ is not composite} \mid \text{miller-rabin returns } c = 0 \ \ell \text{ times}) \\ &= \mathsf{Pr}(n \text{ is prime} \mid \text{miller-rabin returns } c = 0 \ \ell \text{ times}), \\ \mathsf{Pr}(SP_\ell \mid P) &= \mathsf{Pr}(\text{miller-rabin returns } c = 0 \ \ell \text{ times} \mid n \text{ is not composite}) \\ &= \mathsf{Pr}(\text{miller-rabin returns } c = 0 \ \ell \text{ times} \mid n \text{ is prime}) \end{aligned}$$

and

$$\mathsf{Pr}(SP_\ell \mid C) = \mathsf{Pr}(\text{miller-rabin returns } c = 0 \ \ell \text{ times} \mid n \text{ is composite}).$$

The third probability $\mathsf{Pr}(SP_\ell \mid C)$ is the one we have been considering so far, but we should be more interested in the level of confidence that n is prime given that miller-rabin returns $res = 0$, and this is the first probability $\mathsf{Pr}(P \mid SP_\ell)$. This point is clearly articulated in Hoffstein, Pipher and Silverman [8]; most of the literature ignores it, and it is important to make it perfectly clear. Fortunately, $\mathsf{Pr}(P \mid SP_\ell)$ can be obtained using Bayes's formula:

$$\mathsf{Pr}(P \mid SP_\ell) = \frac{\mathsf{Pr}(SP_\ell \mid P)\mathsf{Pr}(P)}{\mathsf{Pr}(SP_\ell \mid P)\mathsf{Pr}(P) + \mathsf{Pr}(SP_\ell \mid \overline{P})\mathsf{Pr}(\overline{P})}.$$

In particular, we need to compute the probability $\mathsf{Pr}(P)$ that an integer $n \geq 3$ is prime. But there is a problem: the events under consideration are all infinite, so their probabilities are not well defined! A positive integer either is prime or is not prime; it cannot be 30% prime and 70% composite.

The problem with infinite events is pointed out by Hoffstein, Pipher and Silverman in [8] (Chapter 3, Section 3.4.1). The statement often found that a randomly chosen positive integer has probability $1/\ln n$ of being prime is meaningless. In order for all these probabilities to make sense, we need our events to be finite, and this can be achieved if we consider finite intervals.

We can make this idea precise following the hint given in Exercise 3.18 in Hoffstein, Pipher and Silverman in [8].

Let $c_1 < c_2$ be any two positive real numbers. For any positive integer n, let us find the probabilty $P(c_1, c_2; n)$ that an integer m such that $c_1 n \le m \le c_2 n$ is prime,

$$P(c_1, c_2; n) = \mathsf{Pr}(m \text{ is prime, } c_1 n \le m \le c_2 n).$$

The condition $c_1 n \le m \le c_2 n$ is really $\lceil c_1 n \rceil \le m \le \lfloor c_2 n \rfloor$. Since $c_1 < c_2$, there are $\lfloor c_2 n \rfloor - \lceil c_1 n \rceil + 1 \approx (c_2 - c_2)n$ integers in the interval $[\lceil c_1 n \rceil, \lfloor c_2 n \rfloor]$. There are $\pi(\lfloor c_2 n \rfloor) - \pi(\lceil c_1 n \rceil - 1)$ primes in the interval $[\lceil c_1 n \rceil, \lfloor c_2 n \rfloor]$, which by the Prime Number Theorem is approximately

$$\begin{aligned}
\pi(\lfloor c_2 n \rfloor) - \pi(\lceil c_1 n \rceil - 1) &\approx \frac{c_2 n}{\ln(c_1 n)} - \frac{c_1 n}{\ln(c_1 n)} \\
&= \frac{n(c_2 \ln(c_1 n) - c_1 \ln(c_2 n))}{\ln(c_1 n) \ln(c_2 n)} \\
&= \frac{n((c_2 - c_1) \ln n + c_2 \ln c_1 - c_1 \ln c_2)}{(\ln n + \ln c_1)(\ln n + \ln c_2)}.
\end{aligned}$$

Consequently, the probability $P(c_1, c_2; n)$ given by

$$P(c_1, c_2; n) = \frac{\pi(\lfloor c_2 n \rfloor) - \pi(\lceil c_1 n \rceil - 1)}{\lfloor c_2 n \rfloor - \lceil c_1 n \rceil + 1}$$

is approximately

$$P(c_1, c_2; n) = \frac{\left(1 + \dfrac{c_2 \ln(c_1) - c_1 \ln(c_2)}{(c_2 - c_1) \ln n}\right)}{\ln n \left(1 + \dfrac{\ln c_1}{\ln n}\right) \left(1 + \dfrac{\ln c_2}{\ln n}\right)}.$$

When n is large, this probability is approximately

$$P(c_1, c_2; n) = \frac{1}{\ln n},$$

independently of c_1 and c_2.

The above derivation justifies saying that for n large enough, the probability that an integer m in an interval $[\lceil c_1 n \rceil, \lfloor c_2 n \rfloor]$ is prime is approximately $1/\ln n$.

We can choose c_1 and c_2 so that $0 < c_1 < 1$ and $c_2 > 1$ (for example, $c_1 = \frac{1}{2}$, $c_2 = \frac{3}{2}$) so that $c_1 n \le n \le c_2 n$, and as long as n is large enough so that

$$\frac{c_2 \ln(c_1) - c_1 \ln(c_2)}{(c_2 - c_1) \ln n}, \quad \frac{\ln c_1}{\ln n}, \quad \text{and} \quad \frac{\ln c_2}{\ln n}$$

are very close to 0, the probability that some $m \in [\lceil c_1 n \rceil, \lfloor c_2 n \rfloor]$ is prime is approximately $\frac{1}{\ln n}$.

In view of all this, we revise the definitions of our events P, C, SP and SP_ℓ as follows. We pick some real numbers c_1 and c_2 such that $0 < c_1 < 1$ and $c_2 > 1$; for example, $c_1 = \frac{1}{2}$ and $c_2 = \frac{3}{2}$. For any positive integer n, let $\mathcal{I}(c_1, c_2; n)$ be the interval

$$\mathcal{I}(c_1, c_2; n) = \{m \in \mathbb{N} \mid \lceil c_1 n \rceil \leq m \leq \lfloor c_2 n \rfloor\}.$$

We may assume that n is large enough and that c_1 is chosen so that $\lceil c_1 n \rceil \geq 3$. For example, $c_1 = \frac{1}{2}$ and $n \geq 6$ will do. Then we define the following events:

$$\begin{aligned}
P(c_1, c_2; n) &= \{m \mid m \in \mathcal{I}(c_1, c_2; n),\ m \text{ is prime}\} \\
C(c_1, c_2; n) &= \{m \mid m \in \mathcal{I}(c_1, c_2; n),\ m \text{ is composite}\} \\
SP(c_1, c_2; n) &= \{m \mid m \in \mathcal{I}(c_1, c_2; n),\ \text{the miller-rabin procedure returns } c = 0\} \\
SP_\ell(c_1, c_2; n) &= \{m \mid m \in \mathcal{I}(c_1, c_2; n),\ \text{the miller-rabin procedure returns } c = 0 \ \ell \text{ times}\},
\end{aligned}$$

To simplify notation, we write $P(n), C(n), SP(n)$, $SP_\ell(n)$ instead of $P(c_1, c_2; n), C(c_1, c_2; n)$, $SP(c_1, c_2; n)$, $SP_\ell(c_1, c_2; n)$, and we also write $\mathcal{I}(n)$ instead of $\mathcal{I}(c_1, c_2; n)$. Then we define the three conditional probabilities

$$\begin{aligned}
&\mathsf{Pr}(P(n) \mid SP_\ell(n)) = \mathsf{Pr}(m \in \mathcal{I}(n) \text{ is prime} \mid \text{miller-rabin returns } c = 0 \ \ell \text{ times}), \\
&\mathsf{Pr}(SP_\ell(n) \mid P(n)) = \mathsf{Pr}(\text{miller-rabin returns } c = 0 \ \ell \text{ times} \mid m \in \mathcal{I}(n) \text{ is prime})
\end{aligned}$$

and

$$\mathsf{Pr}(SP_\ell(n) \mid C(n)) = \mathsf{Pr}(\text{miller-rabin returns } c = 0 \ \ell \text{ times} \mid m \in \mathcal{I}(n) \text{ is composite}).$$

To compute the probabilities on the righthand side, we use the fact that our Miller–Rabin algorithm (the procedure miller-rabin, not the algorithm Probable Prime) is a Monte Carlo algorithm, which means the following:

(1) If miller-rabin returns $c = 1$, then $m \in \mathcal{I}(n)$ definitely is composite (*i.e.* has Property $C(n)$). This is expressed by

$$\mathsf{Pr}(m \in \mathcal{I}(n) \text{ is composite} \mid \text{miller-rabin returns } c = 1) = 1,$$

or more concisely as

$$\mathsf{Pr}(C(n) \mid \overline{SP(n)}) = 1.$$

(2) If $m \in \mathcal{I}(n)$ is composite (has property $C(n)$), then miller-rabin returns $c = 1$ for at least $1/2$ of the number of choices for a. This is expressed by

$$\mathsf{Pr}(\text{miller-rabin returns } c = 1 \mid m \in \mathcal{I}(n) \text{ is composite}) \geq \frac{1}{2},$$

or more concisely as

$$\mathsf{Pr}(\overline{SP(n)} \mid C(n)) \geq \frac{1}{2}.$$

From Property (1) of a Monte Carlo algorithm, by contrapositive, we see that if m is not composite, then the algorithm always returns $c = 0$; that is,

$$\mathsf{Pr}(\text{miller-rabin returns } c = 0 \mid m \in \mathcal{I}(n) \text{ is prime}) = 1,$$

or more concisely as

$$\mathsf{Pr}(SP(n) \mid P(n)) = 1.$$

It follows that

$$\mathsf{Pr}(SP_\ell(n) \mid P(n)) = \mathsf{Pr}(SP(n) \mid P(n))^\ell = 1.$$

To evaluate $\mathsf{Pr}(SP_\ell \mid \overline{P(n)}) = \mathsf{Pr}(SP_\ell \mid C(n))$ we make use of the assumption that miller-rabin is run ℓ independent times and that by Property (2) of a Monte Carlo algorithm,

$$\mathsf{Pr}(\overline{SP(n)} \mid C(n)) \geq \frac{1}{2},$$

so we have

$$\begin{aligned}
\mathsf{Pr}(SP_\ell(n) \mid C(n)) &= \mathsf{Pr}(SP(n) \mid C(n))^\ell \\
&= (1 - \mathsf{Pr}(\overline{SP(n)} \mid C(n)))^\ell \\
&\leq \left(1 - \frac{1}{2}\right)^\ell \\
&= \left(\frac{1}{2}\right)^\ell.
\end{aligned}$$

The above derivation shows rigorously what we have been claiming: the probability that the algorithm says ℓ times that $m \in \mathcal{I}(n)$ is not a composite when in fact it is, is very small. Indeed,

$$\mathsf{Pr}(SP_\ell(n) \mid C(n)) \leq \left(\frac{1}{2}\right)^\ell.$$

As we said earlier, the probability we really want to know is $\mathsf{Pr}(P(n) \mid SP_\ell(n))$. We have all the ingredients to compute it, since we showed earlier that $\mathsf{Pr}(P(n))$ is approximately

$$\mathsf{Pr}(P(n)) = \frac{1}{\ln n}.$$

Then we compute $\mathsf{Pr}(P(n) \mid SP_\ell)$ using Bayes'rule, and we have

$$\begin{aligned}
\mathsf{Pr}(P(n) \mid SP_\ell(n)) &= \frac{\mathsf{Pr}(SP_\ell(n) \mid P(n))\mathsf{Pr}(P(n))}{\mathsf{Pr}(SP_\ell(n) \mid P(n))\mathsf{Pr}(P(n)) + \mathsf{Pr}(SP_\ell(n) \mid \overline{P(n)})\mathsf{Pr}(\overline{P(n)})} \\
&\geq \frac{1 \cdot \frac{1}{\ln(n)}}{1 \cdot \frac{1}{\ln(n)} + 2^{-\ell}\left(1 - \frac{1}{\ln(n)}\right)} \\
&= \frac{1}{1 + 2^{-\ell}(\ln(n) - 1)} \\
&= 1 - \frac{\ln(n) - 1}{2^\ell + \ln(n) - 1} > 1 - \frac{\ln(n)}{2^\ell}.
\end{aligned}$$

Therefore, if ℓ is large enough so that $2^\ell > \ln(n)$

$$\mathsf{Pr}(P(n) \mid SP_\ell(n)) > 1 - \frac{\ln(n)}{2^\ell},$$

(approximately), and we see that it is necessary to add a correction term approximately equal to $\ln(n)$, but this correction term is quickly offset by making ℓ a little bigger.

Our Monte Carlo algorithm does not give a definite answer. However, if ℓ is large enough, say $\ell = 100$, then the conditional probability that the number being tested is prime given that the test is negative ℓ times, is very close to 1 (note that $2^{-100} < 10^{-30}$). In other words, the degree of confidence that the number being tested is prime is very high.

Actually, Rabin and Monier independently proved in 1980 that if m is composite then the Miller–Rabin procedure returns $c = 1$ for at least 3/4 of the number of choices for a (provided that $m > 9$); this implies that

$$\mathsf{Pr}(\overline{SP(n)} \mid C(n)) \geq \frac{3}{4}.$$

Therefore, $\mathsf{Pr}(SP(n) \mid C(n)) \leq 1/4$, and for ℓ large enough, we have

$$\mathsf{Pr}(P(n) \mid SP_\ell(n)) > 1 - \frac{\ln(n)}{4^\ell}.$$

This shows that the Miller–Rabin procedure provides a higher degree of confidence than we originally found, in the sense that if the test is negative ℓ times for ℓ large enough, say $\ell = 100$, then $\mathsf{Pr}(P(n) \mid SP_\ell(n))$ is much closer to 1 than we originally determined.

For example, since $(1/4)^{100} < 2^{-60}$ and since $\ln(10) \approx 2.303$, for $n = 10^{10^{20}}$ (a very large number with 10^{20} digits), we have

$$\mathsf{Pr}(P(n) \mid SP_{100}(n)) > 1 - \frac{\ln(n)}{4^\ell} > 1 - 2.31 \cdot 10^{20} \cdot 10^{-60} > 1 - 10^{-39}.$$

This is a very small probability.

In order to prove that $\mathsf{Pr}(SP(n) \mid C(n)) \leq 1/4$, Rabin and Monier proved that if $n > 9$ is an odd composite, then

$$|L_n^{MR}| \leq \frac{\varphi(n)}{4}.$$

The proof is harder than the proof of Proposition 5.8, but it is not out of reach given a little bit of number theory. The general strategy is also to find a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$ that contains L_n^{MR} and to estimate its order, to show that it is bounded by $\varphi(n)/4$. The proof given in Crandall and Pomerance [3] (Chapter 3, Section 5) is presented in the next section. This proof mixes combinatorial and number theoretic ideas in a beautiful and clever way, but it can be omitted without causing a major gap in the understanding of the Miller–Rabin test. The probability $\mathsf{Pr}(SP(n) \mid C(n)) \leq 1/2$ is good enough to prove that the Miller–Rabin test can be trusted with a high degree of confidence.

5.5 The Monier–Rabin Bound on the Size of the Set of *MR*-liars

Let $n \geq 2$ be any odd integer and suppose that its prime factorization is

$$n = p_1^{j_1} p_2^{j_2} \cdots p_k^{j_k}, \quad j_i \geq 1, i = 1, \ldots, k.$$

Since n is odd, all the p_i are odd. Write $\omega(n) = k$ for the number of distinct prime factors in n. The key point is that L_n^{MR} is a subset of a group $\overline{\mathcal{S}}(n)$ of the form

$$\overline{\mathcal{S}}(n) = \{a \in (\mathbb{Z}/n\mathbb{Z})^* \mid a^m \equiv \pm 1 \pmod{n}\},$$

for some suitable m (depending on n), such that m divides $n-1$. Thus, to estimate the order of this group, we need to find the number of solutions $a \pmod{n}$ to the congruence

$$a^m \equiv \pm 1 \pmod{n}.$$

We will see that the second congruence (the case -1) reduces to the first (the case $+1$), so we are reduced to the problem of counting the number of solutions $a \pmod{n}$ to the congruence

$$a^m \equiv 1 \pmod{n}. \tag{$*$}$$

This is where some number theory comes in handy. Firstly, since the $p_i^{j_i}$ are relatively prime, $a \in \mathbb{Z}$ is a solution of $(*)$ iff a is a solution of the k congruences

$$a^m \equiv 1 \pmod{p_i^{j_i}}, \quad i = 1, \ldots, k. \tag{$**$}$$

Now, because p_i is an odd prime, the group of units $(\mathbb{Z}/p_i^{j_i}\mathbb{Z})^*$ of the ring $\mathbb{Z}/p_i^{j_i}\mathbb{Z}$ is cyclic (see Theorem 4.42). This means that there is some $g \in (\mathbb{Z}/p_i^{j_i}\mathbb{Z})^*$ (called a primitive root modulo $p_i^{j_i}$) such that

$$g, g^2, \ldots, g^{\varphi(p_i^{j_i})-1}, g^{\varphi(p_i^{j_i})} = 1$$

is a list of all elements in $(\mathbb{Z}/p_i^{j_i}\mathbb{Z})^*$. Then, we can easily determine when an element $a \in (\mathbb{Z}/p_i^{j_i}\mathbb{Z})^*$ is a solution of

$$a^m \equiv 1 \pmod{p_i^{j_i}}. \tag{$\dagger$}$$

If we write $a = g^k$, for some k with $1 \leq k \leq \varphi(p_i^{j_i})$, then we must have

$$g^{km} \equiv 1 \pmod{p_i^{j_i}}.$$

Now, $g \in (\mathbb{Z}/p_i^{j_i}\mathbb{Z})^*$ has order $\varphi(p_i^{j_i})$ (the smallest integer r such that $g^r \equiv 1 \pmod{p_i^{j_i}}$), so $\varphi(p_i^{j_i})$ must divide km. If $d = \gcd(m, \varphi(p_i^{j_i}))$ and if we write $m = dm_1$ and $\varphi(p_i^{j_i}) = dn_1$, then $\gcd(m_1, n_1) = 1$ and $\varphi(p_i^{j_i})$ must divide km iff n_1 divides km_1. Since $\gcd(m_1, n_1) = 1$, the number n_1 must divide k, and we find d solutions for k:

$$n_1, 2n_1, \ldots, (d-1)n_1, dn_1 = \varphi(p_i^{j_i}).$$

Therefore, we proved that equation (†) has

$$\gcd(m, \varphi(p_i^{j_i}))$$

solutions modulo $p_i^{j_i}$. Since m divides $n-1$, it is not divisible by p_i (because p_i divides n and $p_i \geq 3$), and since $\varphi(p_i^{j_i}) = p_i^{j_i}(p_i - 1)$, we get

$$\gcd(m, \varphi(p_i^{j_i})) = \gcd(m, p_i^{j_i}(p_i - 1)) = \gcd(m, p_i - 1).$$

By the Chinese remainder theorem, the solutions of $(*)$ modulo n are in bijection with the k-tuples $(a_1, \ldots, a_k)$, where each a_i is a solution of (†) modulo $p_i^{j_i}$. It follows that the number of solutions modulo n of the congruence $a^m \equiv 1 \pmod{n}$ is

$$\prod_{i=1}^{k} \gcd(m, p_i - 1).$$

In summary, we proved the following result.

Proposition 5.10. *Let $n \geq 2$ be any odd integer and suppose that its prime factorization is*

$$n = p_1^{j_1} p_2^{j_2} \cdots p_k^{j_k}, \quad j_i \geq 1, i = 1, \ldots, k.$$

For any integer $m \geq 1$ such that m divides $n-1$, the number of solutions modulo n of the congruence

$$a^m \equiv 1 \pmod{n}$$

is

$$\prod_{i=1}^{k} \gcd(m, p_i - 1).$$

An interesting corollary of Proposition 5.10 obtained by setting $m = n - 1$ is that every odd composite number n is a pseudoprime for at least two nontrivial bases $a \neq \pm 1$ modulo n, unless n is a power of 3.

We now return to the definition of the group $\overline{\mathcal{S}}(n)$.

Definition 5.6. For any odd composite n, if $n - 1 = 2^s t$, with t odd, then let $\nu(n)$ be the largest integer such that $2^{\nu(n)}$ divides $p - 1$ for every prime factor p of n. Then let

$$\overline{\mathcal{S}}(n) = \{a \in (\mathbb{Z}/n\mathbb{Z})^* \mid a^{2^{\nu(n)-1}t} \equiv \pm 1 \pmod{n}\}.$$

The following proposition shows why $\overline{\mathcal{S}}(n)$ is relevant.

Proposition 5.11. *For any odd composite integer n, the following properties hold:*

(1) The set of MR-liars L_n^{MR} is contained in $\overline{\mathcal{S}}(n)$.

(2) The set $\overline{\mathcal{S}}(n)$ is a subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$.

Proof. (1) Pick any $a \in L_n^{MR}$. There are two cases.

(i) If $a^t \equiv 1 \pmod{n}$, then obviously $a^{2^{\nu(n)-1}t} \equiv 1 \pmod{n}$, so $a \in \overline{\mathcal{S}}(n)$.

(ii) There is some smallest index i with $0 \leq i \leq s-1$ such that $a^{2^i t} \equiv n-1 \equiv -1 \pmod{n}$. For any prime p dividing n, we also have $a^{2^i t} \equiv -1 \pmod{p}$, and so

$$a^{2^{i+1}t} \equiv 1 \pmod{p}.$$

If k is the order of a in $(\mathbb{Z}/p\mathbb{Z})^*$, then k divides $2^{i+1}t$ but k does not divide $2^i t$. This is because if k divides $2^i t$, then we can write $2^i t = kq$, and then

$$a^{2^i t} \equiv a^{kq} \equiv (a^k)^q \equiv 1^q \equiv 1 \pmod{p}$$

since $a^k \equiv 1 \pmod{p}$ as k is the order of a modulo p, contradicting $a^{2^i t} \equiv -1 \pmod{p}$. It follows that 2^{i+1} is the exact power of 2 in the prime factorization of k so we can write $k = 2^{i+1}t_1$ for some t_1. Since by Fermat's little theorem, $a^{p-1} \equiv 1 \pmod{p}$, we see that $k = 2^{i+1}t_1$ divides $p-1$, and so 2^{i+1} divides $p-1$. Since this holds for every prime p dividing n, we have $i+1 \leq \nu(n)$. Since $a^{2^i t} \equiv -1 \pmod{n}$, if $i+1 < \nu(n)$, then by squaring we obtain $a^{2^{i+1}t} \equiv 1 \pmod{n}$, so $a^{2^{\nu(n)-1}t} \equiv 1 \pmod{n}$, else if $i+1 = \nu(n)$, then $a^{2^{\nu(n)-1}t} \equiv -1 \pmod{n}$. In both cases, $a \in \overline{\mathcal{S}}(n)$.

(2) The proof that $\overline{\mathcal{S}}(n)$ is a group is very similar to the fact that B_n is a group (see the proof of Proposition 5.8) and is left an an exercise. □

The next proposition gives a formula for the order of the group $\overline{\mathcal{S}}(n)$.

Proposition 5.12. *For any odd composite integer n, if we denote the number of distinct prime factors of n by $\omega(n)$ and if $n-1 = 2^s t$ with s, t and $\nu(n)$ as in Definition 5.6, then the order of the group $\overline{\mathcal{S}}(n)$ is given by*

$$|\overline{\mathcal{S}}(n)| = 2 \cdot 2^{(\nu(n)-1)\omega(n)} \prod_{p|n} \gcd(t, p-1).$$

Proof. Since

$$\overline{\mathcal{S}}(n) = \{a \in \mathbb{Z}/n\mathbb{Z} \mid a^{2^{\nu(n)-1}t} \equiv \pm 1 \pmod{n}\},$$

we need to count the number s_1 of solutions of the congruence $a^{2^{\nu(n)-1}t} \equiv 1 \pmod{n}$ and the number s_2 of solutions of $a^{2^{\nu(n)-1}t} \equiv -1 \pmod{n}$. As to the first congruence, if we let $m = 2^{\nu(n)-1}t$, by definition of $\nu(n)$, we know that $2^{\nu(n)}$ divides $p-1$ for every prime p dividing n. Let $n = p_1^{j_1} \cdots p_k^{j_k}$ be a prime factorization of n, where $p_1, \ldots, p_k$ are the distinct prime factors of n (with $k = \omega(n)$), which must be odd since n is odd. We have $p_i - 1 = 2^{\nu(n)} t_i$ for some t_i. If we had $\nu(n) > s$, then since

$$p_i - 1 = 2^{\nu(n)} t_i = 2^{s+1} 2^{\nu(n)-s-1} t_i$$

we could write $p_i = 1 + 2^{s+1}u_i$ for some integers $u_i = 2^{\nu(n)-s-1}t_i$, and then $n-1 = 2^s t$ would yield

$$(2^{s+1}u_1 + 1)\cdots(2^{s+1}u_k + 1) - 1 = 2^s t,$$

which would imply that

$$2^{s+1}u = 2^s t,$$

for some integer u. Since t is odd, this is impossible, and thus $\nu(n) \leq s$. Consequently, $m = 2^{\nu(n)-1}t$ divides $n-1 = 2^s t$. By Proposition 5.10, we have

$$s_1 = \prod_{i=1}^{\omega(n)} \gcd(m, p_i - 1).$$

But $m = 2^{\nu(n)-1}t$, t is odd and $2^{\nu(n)}$ divides each $p_i - 1$, so

$$\gcd(m, p_i - 1) = \gcd(2^{\nu(n)-1}t, p_i - 1) = 2^{\nu(n)-1}\gcd(t, p_i - 1),$$

so we get

$$s_1 = 2^{(\nu(n)-1)\omega(n)} \prod_{i=1}^{\omega(n)} \gcd(t, p_i - 1).$$

We now show that because n is odd the congruence $a^m \equiv -1 \pmod{n}$ has the same number of solutions as the congruence $a^m \equiv 1 \pmod{n}$.

For every odd prime p_i as above, we claim that $a^m \equiv -1 \pmod{p_i^{j_i}}$ iff $a^{2m} \equiv 1 \pmod{p_i^{j_i}}$ and $a^m \not\equiv 1 \pmod{p_i^{j_i}}$.[3] Obviously, if $a^m \equiv -1 \pmod{p_i^{j_i}}$, then $(a^m)^2 \equiv a^{2m} \equiv 1 \pmod{p_i^{j_i}}$ and $a^m \not\equiv 1 \pmod{p_i^{j_i}}$. Conversely, by Proposition 5.1, if p_i is an odd prime then there are exactly two square roots of unity mod $p_i^{j_i}$, namely $+1$ and -1. So if $(a^m)^2 \equiv a^{2m} \equiv 1 \pmod{p_i^{j_i}}$, then either $a^m \equiv 1 \pmod{p_i^{j_i}}$ or $a^m \equiv -1 \pmod{p_i^{j_i}}$, but since $a^m \not\equiv 1 \pmod{p_i^{j_i}}$, we mut have $a^m \equiv -1 \pmod{p_i^{j_i}}$.

We observed earlier that $2^{\nu(n)}$ divides $p_i - 1$, and it follows as above that the number of solutions of the equation $a^m \equiv -1 \pmod{p_i^{j_i}}$ is

$$2^{\nu(n)}\gcd(t, p_i - 1) - 2^{\nu(n)-1}\gcd(t, p_i - 1) = 2^{\nu(n)-1}\gcd(t, p_i - 1).$$

It follows that the number s_2 of solutions of $a^m \equiv -1 \pmod{n}$ is

$$s_2 = s_1.$$

Therefore, the order of the group $\overline{\mathcal{S}}(n)$ is indeed

$$|\overline{\mathcal{S}}(n)| = 2 \cdot 2^{(\nu(n)-1)\omega(n)} \prod_{p|n} \gcd(t, p-1),$$

as claimed. □

[3] I thank Tian Bai for pointing out an error in my previous proof, which claimed that the above fact held for any n rather that for each $p_i^{j_i}$ with p_i an odd prime.

Proposition 5.12 yields the main result of this section.

Theorem 5.13. *(Monier–Rabin) For every odd composite $n > 9$, we have*

$$|L_n^{MR}| \leq \frac{\varphi(n)}{4} \leq \frac{n-1}{4}.$$

Proof. As usual, write $n - 1 = 2^s t$, with t odd. By Proposition 5.11, $L_n^{MR} \subseteq \overline{\mathcal{S}}(n)$, so it suffices to prove that $|\overline{\mathcal{S}}(n)|/\varphi(n) \leq \frac{1}{4}$. If $n = p_1^{j_1} p_2^{j_2} \cdots p_k^{j_k}$ is the prime factorization of n, recall that

$$\varphi(n) = p_1^{j_1-1}(p_1 - 1)p_2^{j_2-1}(p_2 - 1) \cdots p_k^{j_k-1}(p_k - 1).$$

By Proposition 5.12,

$$|\overline{\mathcal{S}}(n)| = 2 \cdot 2^{(\nu(n)-1)\omega(n)} \prod_{p|n} \gcd(t, p-1),$$

so we need to prove that

$$\frac{\varphi(n)}{|\overline{\mathcal{S}}(n)|} = \frac{1}{2} \prod_{p^k \| n} p^{k-1} \frac{p-1}{2^{\nu(n)-1}\gcd(t, p-1)} \geq 4,$$

where the notation $p^k \| n$ means that p^k is the exact power of the prime p in the prime factorization of n. Each factor

$$\frac{p-1}{2^{\nu(n)-1}\gcd(t, p-1)} \qquad (*)$$

is an even integer. There are several cases.

Case 1: $\omega(n) \geq 3$. In this case, at least three of the factors $(*)$ are equal to 2, so $\varphi(n)/|\overline{\mathcal{S}}(n)| \geq 4$.

Case 2: $\omega(n) = 2$ and n is not squarefree. Then, some exponent $k - 1$ is at least 1, and since all the primes p are odd, the product of the p^{k-1} is at least 3, each factor $(*)$ is at least 2, so $\varphi(n)/|\overline{\mathcal{S}}(n)| \geq 6$.

Case 3: $n = pq$, for two distinct primes, p, q, with $p < q$. If $2^{\nu(n)+1}$ divides $q - 1$, then $q - 1 = 2^{\nu(n)+1}u$, and t is odd, we get

$$2^{\nu(n)-1}\gcd(t, q-1) = 2^{\nu(n)-1}\gcd(t, 2^{\nu(n)+1}u) = 2^{\nu(n)-1}\gcd(t, u) \leq 2^{\nu(n)-1}u = (q-1)/4,$$

and since the other fraction involving $p - 1$ is at least 2, we get $\varphi(n)/|\overline{\mathcal{S}}(n)| \geq 4$.

The remaining subcase is that $2^{\nu(n)}$ is the exact power of 2 in $q - 1$, and we can write $q - 1 = 2^{\nu(n)}u$, where u is odd. Because $n - 1 = pq - 1 = p(q - 1) + p - 1$ and $p < q$, we see that $q - 1$ does not divide $n - 1$. This implies that there is an odd prime q_1 dividing $q - 1$ to a higher power than it divides $n - 1$. Since $n - 1 = 2^s t$ and $q - 1 = 2^{\nu(n)}u$, we have $t = q_1^h t_1$ and $u = q_1^{h+1}u_1$ for some $h \geq 1$ and some t_1, u_1 with $\gcd(q_1, t_1) = 1$. It follows that

$$q - 1 = 2^{\nu(n)} q_1^{h+1} u_1$$

and

$$\begin{aligned} 2^{\nu(n)-1}\gcd(t, q-1) &= 2^{\nu(n)-1}\gcd(q_1^h t_1, 2^{\nu(n)} q_1^{h+1} u_1) \\ &= 2^{\nu(n)-1} q_1^h \gcd(t_1, q_1 u_1) \\ &= 2^{\nu(n)-1} q_1^h \gcd(t_1, u_1) \\ &\leq 2^{\nu(n)-1} q_1^h u_1 \\ &= \frac{(q-1)}{2q_1} \leq \frac{q-1}{6}. \end{aligned}$$

Since the other fraction is at least 2, we conclude that $\varphi(n)/|\overline{\mathcal{S}}(n)| \geq 6$.

Case 4: $n = p^k$, for some $k \geq 2$. In this case, $n - 1 = p^k - 1 = 2^s t$ with t odd. Since $k \geq 2$, we have

$$p^k - 1 = (p-1)(p^{k-1} + \cdots + p + 1) = 2^s t,$$

and since

$$p - 1 = 2^{\nu(n)} u$$

with u odd, we conclude that u is a divisor of t. Then

$$\gcd(t, p-1) = \gcd(t, 2^{\nu(n)} u) = \gcd(t, u) = u,$$

which implies that

$$\frac{p-1}{2^{\nu(n)-1}\gcd(t, p-1)} = \frac{2^{\nu(n)} u}{2^{\nu(n)-1} u} = 2,$$

and thus, $\varphi(n)/|\overline{\mathcal{S}}(n)| = p^{k-1}$. Therefore, unless $p^k = 9$, which means that $k = 2$ and $p = 3$, we have $\varphi(n)/|\overline{\mathcal{S}}(n)| \geq 5$ (since 5 is the next prime after 3). This last case finishes the proof. □

Remarks:

1. Another proof of Theorem 5.13 is given in Koblitz [10] (Chapter V, Proposition V.1.7).
2. The group $\overline{\mathcal{S}}(n)$ is actually the group generated by the set L_n^{MR} of *MR*-liars. This result due to Jim Haglund is Problem 3.16 in Crandall and Pomerance [3].

5.6 The Least *MR*-Witness for n

Theorem 5.13 shows that an odd composite n has at least $3n/4$ *MR*-witnesses in the set $\{2, \ldots, n-2\}$. A natural question then arises: what is the size of the smallest *MR*-witness, $W(n)$, for n? If, by luck, the size of $W(n)$ is bounded by a constant, or a slow-growing function, then there is hope that a practical deterministic algorithm (that is, not a randomized algorithm) can be found.

Unfortunately, there is no constant bound. Indeed, Alford, Granville and Pomerance showed that for infinitely many odd composite n, we have

$$W(n) > (\ln n)^{1/(3 \ln \ln \ln n)}.$$

In Crandall and Pomerance [3], it is also shown that $W(n) \geq 3$ for infinitely many n (with an explicit description). Around 1976, Gary Miller showed that $W(n) = O((\ln n)^2)$, assuming that the Extended Riemann Hypothesis (for short ERH) holds. Then, Bach (1985) proved that $W(n) < 2(\ln n)^2$; see Crandall and Pomerance [3] (Chapter 3, Section 3.5).

The ERH is a generalization of the Riemann Hypothesis (for short RH), one of the most famous conjectures of mathematics. Explaining what is the ERH would lead us to far, and we refer the reader to Crandall and Pomerance for an explanation [3] (Chapter 1, Section 1.4). However, we discuss briefly the RH.

The RH has to with the location of the zeros of the zeta function, ζ. For any real $s > 1$, the function $\zeta(s)$ is given by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

If we allow s to be a complex argument, then the above sum converges absolutely for $\mathrm{Re}(s) > 1$. It is also possible to extend ζ to the entire complex plane (by analytic continuation), so that $\zeta(s)$ is regular for every s except $s = 1$, where it has a simple pole with residue 1 (this means that $(s-1)\zeta(s)$ is holomorphic in $\mathbb{C}$, with value 1 at $s = 1$). Two good sources are Apostol [1] and Edwards [5]. Ribenboim's lovely book [18] (especially Chapter 4) is also highly recommended. The connection with prime numbers was noticed by Euler and is this:

Theorem 5.14. *(Euler) If $\mathcal{P}$ denotes the set of all primes, then for every $s \in \mathbb{C}$ such that $\mathrm{Re}(s) > 1$,*

$$\zeta(s) = \prod_{p \in \mathcal{P}} \frac{1}{1 - p^{-s}}.$$

The value $\zeta(s)$ of the zeta-function is known when s is an even integer, but $\zeta(s)$ is not known for not a single odd integer! Remarkably, the location of the zeros of ζ has crucial impact on the distribution of the primes. For example, the fact that $\zeta(s) \neq 0$ on the line $\mathrm{Re}(s) = 1$ leads to the Prime Number Theorem. The Riemann Hypothesis, stated in 1859 by Riemann in an eight-page memoir, says this:

Conjecture (*Riemann hypothesis* (RH))

All the zeros of ζ in the critical strip $0 < \mathrm{Re}(s) < 1$ lie on the vertical line $\mathrm{Re}(s) = 1/2$.

The RH has been verified by computer calculations for many values of n, but it still remains one of the central conjectures of mathematics. It is equivalent to various statements about the distribution of the primes. For example, von Koch proved in 1901 that the RH is equivalent to the following fact:

$$|\pi(x) - \mathrm{Li}(x)| < \sqrt{x} \cdot \ln(x), \quad \text{for all } x \geq 2.01,$$

where $\pi(x)$ is the number of primes not exceeding x, and the function Li (logarithmic integral) is given by

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\ln t}.$$

It is easy to see that

$$\mathrm{Li}(x) = \frac{x}{\ln x} + \int_2^x \frac{dt}{(\ln t)^2} - \frac{2}{\ln 2}.$$

It is amazing that Gauss conjectured in 1791 (at the age of fourteen) that $\pi(x) \sim \mathrm{Li}(x)$. We refer the reader to Crandall and Pomerance [3] (Chapter 1) for more on this topic.

The Extended Riemann Hypothesis (ERH) has to do with the zeros of the Dirichlet L-functions $L(s, \chi)$, which generalizes the ζ-function. Here, χ denotes a Dirichlet character. Apostol [1] is an excellent source to learn about L-functions. The ζ-function corresponds to the special case $\chi = 1$. The ERH says this:

Conjecture (*Extended Riemann hypothesis* (ERH))

For any Dirichlet character χ, all the zeros of $L(s, \chi)$ in the region $\mathrm{Re}(s) > 0$ lie on the vertical line $\mathrm{Re}(s) = 1/2$.

Assuming the ERH, Bach's result, that $W(n) < 2(\ln)^2$, yields a deterministic algorithm for testing for primality. Simply try the Miller–Rabin procedure for $a = 2, 3, \ldots, 2(\ln n)^2$. Besides the fact that the ERH is still not proved, in practice, the randomized version of the Miller–Rabin test is faster. As of now, if you want a reliable test, either you have to have faith in the ERH, or faith that an event that has probability less than 10^{-30} will never happen in our lifetime. This probability is much smaller than the probability of hardware or software failure anyway!

Chapter 6

The Solovay–Strassen Test

6.1 Quadratic Residues

The Solovay–Strassen primality test was published in 1977, and thus slightly predates the Miller–Rabin test. It is also a randomized algorithm of Monte Carlo type, and it gives the output "composite," given that the input n is composite, with probability greater than $1/2$. The Solovay–Strassen is based on a criterion due to Euler to test whether a number which is not a multiple of a prime p is a quadratic residue. This test involves the Jacobi symbol, which is a generalization of the Legendre symbol. Properties of the Jacobi symbol yield a fast method for checking Euler's criterion.

If p is a prime and m is an integer which is not a multiple of p, we can look for solutions x of the quadratic congruence

$$x^2 \equiv m \pmod{p}. \tag{$*$}$$

In other words, we are looking for a square root of m modulo p.

When $p = 2$, every odd integer is a quadratic residue and there are no quadratic nonresidues. This case is not very interesting, so from now on we assume that $p \geq 3$.

If p is an odd prime, note that the above congruence has at most two solutions. Indeed, $\mathbb{Z}/p\mathbb{Z}$ is a field, so the polynomial $x^2 - m$ has at most two roots. Moreover, if x is a solution of $(*)$, then so is $-x$, hence the number of solutions is either 0 or 2.

It is convenient to allow p to be any integer ≥ 3.

Definition 6.1. Given any integer $n \geq 3$, for any integer m such that $\gcd(m, n) = 1$, we say that m is a *quadratic residue* mod n (or a *square* mod n) if the congruence

$$x^2 \equiv m \pmod{n} \tag{$\dagger$}$$

has a solution. If $(\dagger)$ has no solution we say that m is a *quadratic nonresidue* mod n.

Observe that the integers m such that $\gcd(m, n) > 1$ are considered neither quadratic residues nor quadratic nonresidues.

Consider the example $n = 13$. The squares modulo 13 of the numbers in $\{1, 2, \ldots, 12\}$ are

$$1, 4, 9, 3, 12, 10, 10, 12, 3, 9, 4, 1,$$

and thus, there are $6 = 12/2$ quadratic residues:

$$1, 3, 4, 9, 10, 12.$$

The square roots of 12 modulo 13 are 5 and 8. For $n = 26$, the quadratic residues are

$$1, 3, 9, 17, 23, 25.$$

Because 26 is even, they must be odd. For $n = 27$, the quadratic residues are

$$1, 4, 7, 10, 13, 16, 19, 22, 25.$$

When n is prime, as in the case $n = 13$, there is the same number of quadratic residues and nonresidues. This is a general fact.

Proposition 6.1. *Let p be an odd prime. Then the set of quadratic residues is a subgroup of $(\mathbb{Z}/p\mathbb{Z})^*$ of order $(p-1)/2$. This subgroup consists of the residues modulo p of the numbers*

$$1^2, 2^2, 3^2, \ldots, \left(\frac{p-1}{2}\right)^2.$$

Proof. It is clear that 1 is a quadratic residue. If $x^2 \equiv a \pmod{p}$ and $y^2 \equiv b \pmod{p}$, then

$$(xy)^2 \equiv x^2 y^2 \equiv ab \pmod{p},$$

so ab is also a quadratic residue, and the quadratic residues form a group.

If $x^2 \equiv y^2 \pmod{p}$, then

$$(x - y)(x + y) \equiv 0 \pmod{p}.$$

Assume that $1 \leq x, y \leq (p-1)/2$. Since p is prime, either p divides $x - y$ or p divides $x + y$. But, $1 \leq x, y \leq (p-1)/2$, which implies $2 \leq x + y \leq p - 1$, so $x - y$ must be divisible by p, and thus $x = y$. Therefore, the residues modulo p of the square numbers listed in the proposition are all distinct. Since

$$(p - k)^2 \equiv k^2 \pmod{p},$$

every quadratic residue modulo p is congruent to a unique number in this list. □

When p is an odd prime, we know that $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic (see Theorem 4.34). If g is any primitive root for $(\mathbb{Z}/p\mathbb{Z})^*$, then $g^{p-1} \equiv 1 \pmod{p}$, so

$$(g^{(p-1)/2})^2 \equiv g^{p-1} \equiv 1 \pmod{p},$$

and $g^{(p-1)/2}$ is a square root of 1. Since $\mathbb{Z}/p\mathbb{Z}$ is a field, the only square roots of 1 are ± 1, so $g^{(p-1)/2} \equiv \pm 1 \pmod{p}$. But if $g^{(p-1)/2} \equiv 1 \pmod{p}$ then g is not a primitive root, so $g^{(p-1)/2} \equiv -1 \pmod{p}$.

Proposition 6.1 also shows the following fact.

Proposition 6.2. *If p is an odd prime, then for any primitive root $g \in (\mathbb{Z}/p\mathbb{Z})^*$, the quadratic residues are the even powers g^{2i}, and the quadratic nonresidues are the odd powers g^{2i+1}, with $0 \leq i \leq (p-3)/2$.*

We can use the above fact to find square roots modulo p for primes of the form $p = 4k+3$. Indeed, if $a = g^{2i}$ is any quadratic residue, then we claim that

$$x = a^{(p+1)/4} = a^{k+1}$$

is a square root of a modulo p.

Since $g^{(p-1)/2} \equiv -1 \pmod{p}$, we get

$$x \equiv a^{(p+1)/4} \equiv (g^{2i})^{(p+1)/4} \equiv g^{i(p+1)/2} \equiv g^{i(p-1)/2} g^i \equiv (g^{(p-1)/2})^i g^i \equiv (-1)^i g^i \pmod{p},$$

and thus,

$$x^2 \equiv (-1)^{2i} g^{2i} \equiv a \pmod{p}.$$

If p is a prime of the form $p = 4k+1$, it is (a lot!) harder to find square roots modulo p; see the end of Section 6.5, and Crandall and Pomerance [3] (Chapter 2, Section 3).

6.2 The Legendre Symbol

At this stage, it is convenient to introduce the Legendre symbol. The remarkable fact about the Legendre symbol is that it gives us an efficient method for testing whether a number m is a quadratic residue mod n without actually solving the congruence $x^2 \equiv m \pmod{n}$. The Jacobi symbol defined in Section 6.3 gives us an even more efficient method which avoids factoring.

Definition 6.2. Let p be an odd prime. For any integer m, the *Legendre symbol* $\left(\frac{m}{p}\right)$ is defined as follows:

$$\left(\frac{m}{p}\right) = \begin{cases} +1 & \text{if } m \text{ is a quadratic residue modulo } p \\ -1 & \text{if } m \text{ is a quadratic nonresidue modulo } p \\ 0 & \text{if } p \text{ divides } m. \end{cases}$$

Observe that $\left(\frac{1}{p}\right) = 1$ and $\left(\frac{m^2}{p}\right) = 1$ for every integer m which is not a multiple of p. For a numerical example, since the squares modulo 11 of the numbers in $\{1, 2, \ldots, 10\}$ are

$$1,\ 4,\ 9,\ 5,\ 3,\ 3,\ 5,\ 9,\ 4,\ 1,$$

we see that 7 is not a quadratic residue modulo 11, so $\left(\frac{7}{11}\right) = -1$. We saw earlier that the quadratic residues modulo 13 are $1, 3, 4, 9, 10, 1$ so $\left(\frac{3}{13}\right) = 1$.

If $m \equiv n \pmod{p}$, then clearly $\left(\frac{m}{p}\right) = \left(\frac{n}{p}\right)$, so the function $m \mapsto \left(\frac{m}{p}\right)$ is periodic with period p. We also have

$$\left(\frac{mn}{p}\right) = \left(\frac{m}{p}\right)\left(\frac{n}{p}\right),$$

a very useful property for evaluating the Legendre symbol. To prove this property easily, we will establish Euler's criterion. First, observe that for any m which is not a multiple of p, by Fermat's little theorem, we have

$$m^{p-1} \equiv 1 \pmod{p}.$$

If p is an odd prime, then we get

$$m^{p-1} - 1 \equiv (m^{(p-1)/2} - 1)(m^{(p-1)/2} + 1) \equiv 0 \pmod{p},$$

and it follows that

$$m^{(p-1)/2} \equiv \pm 1 \pmod{p}. \qquad (*_E)$$

Remarkably, $m^{(p-1)/2} \equiv 1 \pmod{p}$ iff m is a quadratic residue modulo p.

Theorem 6.3. *(Euler's criterion) If p is an odd prime, then for any integer m, we have*

$$\left(\frac{m}{p}\right) \equiv m^{(p-1)/2} \pmod{p}.$$

Proof. If $m \equiv 0 \pmod{p}$, then both sides of the equation are 0, so the equation is trivially true.

Suppose $\left(\frac{m}{p}\right) = 1$. In this case, there is some $x \in \{1, \ldots, p-1\}$ such that $x^2 \equiv m \pmod{p}$; by Fermat's little theorem $x^{p-1} \equiv 1 \pmod{p}$, and so,

$$m^{(p-1)/2} \equiv (x^2)^{(p-1)/2} \equiv x^{p-1} \equiv 1 \equiv \left(\frac{m}{p}\right) \pmod{p}.$$

This proves the formula if $\left(\frac{m}{p}\right) = 1$.

Finally, assume that $\left(\frac{m}{p}\right) = -1$. The polynomial $x^{(p-1)/2} - 1$ has degree $(p-1)/2$, and since $\mathbb{Z}/p\mathbb{Z}$ is a field (since p is prime), it has at most $(p-1)/2$ roots in $\mathbb{Z}/p\mathbb{Z}$. However, by Proposition 6.1, the $(p-1)/2$ quadratic residues a in $\mathbb{Z}/p\mathbb{Z}$ are the residues of the numbers

$$1^2, 2^2, 3^2, \ldots, \left(\frac{p-1}{2}\right)^2$$

modulo p, so $a \equiv y^2 \pmod{p}$ for $y \in \{1, 2, \ldots, (p-1)/2\}$. By Fermat's little theorem $y^{p-1} \equiv 1 \pmod{p}$, so we get $a^{(p-1)/2} \equiv (y^2)^{(p-1)/2} \equiv y^{p-1} \equiv 1 \pmod{p}$. Therefore, the $(p-1)/2$ quadratic residues are roots of $x^{(p-1)/2} - 1$, and the nonresidues are not, and since m is a nonresidue, we must have

$$m^{(p-1)/2} \not\equiv 1 \pmod{p}.$$

Since by $(*_E)$ we have $m^{(p-1)/2} \equiv \pm 1 \pmod{p}$, we conclude that

$$m^{(p-1)/2} \equiv -1 \equiv \left(\frac{m}{p}\right) \pmod{p},$$

which finishes the proof. □

Following Serre [20], another proof of Euler's criterion can be given using some algebra. This proof shows a result that will be used in the proof of the law of quadratic reciprocity so we record it as the following proposition.

Proposition 6.4. *Let p be an odd prime.*

(1) Pick any $m \in (\mathbb{Z}/p\mathbb{Z})^*$, and Let Ω be any field extension of $\mathbb{F}_p$ which contains a square root y of m (an algebraic closure of $\mathbb{Z}/p\mathbb{Z}$, will do). Then

$$\left(\frac{m}{p}\right) = m^{(p-1)/2} = y^{p-1} \quad \text{in } (\mathbb{Z}/p\mathbb{Z})^*.$$

(2) The set $(\mathbb{Z}/p\mathbb{Z})^{*2}$ of squares in $(\mathbb{Z}/p\mathbb{Z})^*$ is a a subgroup of order $(p-1)/2$. This subgroup is the kernel of the homomorphism $m \mapsto m^{(p-1)/2}$ from $(\mathbb{Z}/p\mathbb{Z})^*$ to $\{-1, +1\}$.

Proof. (1) Since $y \in \Omega$ is a square root of $m \in (\mathbb{Z}/p\mathbb{Z})^* \subseteq \Omega$, we have $m = y^2$. Then (in $\mathbb{Z}/p\mathbb{Z}$) we deduce that

$$y^{p-1} = (y^2)^{(p-1)/2} = m^{(p-1)/2} = \pm 1,$$

since $1 = m^{p-1} = (m^{(p-1)/2})^2 = (y^{p-1})^2 = 1$, and because p being an odd prime, the only roots of unity are $+1$ and -1.

We claim that m is a square in $(\mathbb{Z}/p\mathbb{Z})^*$ iff $y \in (\mathbb{Z}/p\mathbb{Z})^*$ iff $y^{p-1} = 1$. By definition of the Legendre symbol, this is equivalent to

$$\left(\frac{m}{p}\right) = m^{(p-1)/2} = y^{p-1} \quad \text{in } (\mathbb{Z}/p\mathbb{Z})^*,$$

as claimed.

If $y \in (\mathbb{Z}/p\mathbb{Z})^*$, then $m = y^2$ is a square in $(\mathbb{Z}/p\mathbb{Z})^*$. Conversely, assume that $m = a^2$ for some $a \in (\mathbb{Z}/p\mathbb{Z})^*$. Since we also have $m = y^2$, we get $y^2 = a^2$, that is, $y^2 \equiv a^2 \pmod{p}$, thus $(y-a)(y+a) \equiv 0 \pmod{p}$, and since p is prime either $y = a$ or $y = -a$, which shows that $y \in (\mathbb{Z}/p\mathbb{Z})^*$.

If $y \in (\mathbb{Z}/p\mathbb{Z})^*$, then by Fermat's little theorem $y^{p-1} \equiv 1 \pmod{p}$. For the other direction of the second equivalence, note that if $y^{p-1} = 1$ and $y \notin (\mathbb{Z}/p\mathbb{Z})^*$, then the equation $z^{p-1}-1 = 0$ has p roots in Ω, since the $p-1$ elements of $(\mathbb{Z}/p\mathbb{Z})^*$ are roots of $z^{p-1}-1 = 0$, a contradiction since Ω is a field.

(2) It is obvious that the map $m \mapsto m^{(p-1)/2}$ is a homomorphism from $(\mathbb{Z}/p\mathbb{Z})^*$ to $\{-1, +1\}$, and from the above discussion, its kernel is the set $(\mathbb{Z}/p\mathbb{Z})^{*2}$ of squares in $(\mathbb{Z}/p\mathbb{Z})^*$. Now, $(\mathbb{Z}/p\mathbb{Z})^*$ is a cyclic group of order $p-1$, so the above homomorphism must be surjective (otherwise, every element of $(\mathbb{Z}/p\mathbb{Z})^*$ would have order $(p-1)/2$). It follows that $(\mathbb{Z}/p\mathbb{Z})^{*2}$ is a subgroup of order $(p-1)/2$. □

It is now easy to establish the multiplicative property of the Legendre symbol.

Proposition 6.5. *For any odd prime p and any integers m, n, we have*

$$\left(\frac{mn}{p}\right) = \left(\frac{m}{p}\right)\left(\frac{n}{p}\right).$$

Proof. If p divides m or p divides n, then p divides mn so $\left(\frac{mn}{p}\right) = 0$, and either $\left(\frac{m}{p}\right) = 0$ or $\left(\frac{n}{p}\right) = 0$; it follows that $0 = \left(\frac{mn}{p}\right) = \left(\frac{m}{p}\right)\left(\frac{n}{p}\right) = 0$.

If p divides neither m nor n, then p does not divide mn and by Euler's criterion, we have

$$\left(\frac{mn}{p}\right) \equiv (mn)^{(p-1)/2} \equiv m^{(p-1)/2} n^{(p-1)/2} \equiv \left(\frac{m}{p}\right)\left(\frac{n}{p}\right) \pmod{p}.$$

The symbols $\left(\frac{mn}{p}\right), \left(\frac{m}{p}\right), \left(\frac{n}{p}\right)$ are all equal to $+1$ or -1, so $\left(\frac{mn}{p}\right) - \left(\frac{m}{p}\right)\left(\frac{n}{p}\right)$ is either 0, $+2$, or -2, and since $\left(\frac{mn}{p}\right) - \left(\frac{m}{p}\right)\left(\frac{n}{p}\right)$ is divisible by $p \geq 3$, we must have $\left(\frac{mn}{p}\right) = \left(\frac{m}{p}\right)\left(\frac{n}{p}\right)$. □

The following properties are also useful for evaluating the Legendre symbol.

Proposition 6.6. *For any odd prime p, the following properties hold:*

(1) If $m \equiv n \pmod{p}$, then $\left(\frac{m}{p}\right) = \left(\frac{n}{p}\right)$.

(2) If $\gcd(m, p) = 1$, then $\left(\frac{m^2}{p}\right) = 1$ and $\left(\frac{m^2 n}{p}\right) = \left(\frac{n}{p}\right)$.

(3) $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$, or equivalently

$$\left(\frac{-1}{p}\right) = \begin{cases} 1 & \text{if } p \equiv 1 \bmod 4 \\ -1 & \text{if } p \equiv 3 \bmod 4. \end{cases}$$

(4) $\left(\dfrac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$, *or equivalently*

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & \text{if } p \equiv 1 \bmod 8 \text{ or } p \equiv 7 \bmod 8 \\ -1 & \text{if } p \equiv 3 \bmod 8 \text{ or } p \equiv 5 \bmod 8. \end{cases}$$

Proof. Part (1) is obvious from the definition of the Legendre symbol. Part (2) follows from Euler's criterion and Proposition 6.5. The details are left to the reader.

By Euler's criterion we have

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}.$$

Since p is an odd prime, either p is of the form $p = 4k+1$, in which case $(p-1)/2 = 2k$, so $(-1)^{(p-1)/2} = (-1)^{2k} = 1$, or p is of the form $p = 4k+3$, in which case $(p-1)/2 = 2k+1$, so $(-1)^{(p-1)/2} = (-1)^{2k+1} = -1$. This proves Part (3).

To prove (4), consider the $(p-1)/2$ congruences:

$$\begin{aligned} p-1 &\equiv 1(-1)^1 && \pmod p \\ 2 &\equiv 2(-1)^2 && \pmod p \\ p-3 &\equiv 3(-1)^3 && \pmod p \\ 4 &\equiv 4(-1)^4 && \pmod p \\ &\vdots \\ r &\equiv \frac{p-1}{2}(-1)^{(p-1)/2} && \pmod p, \end{aligned}$$

where $r = p-(p-1)/2$ or $r = (p-1)/2$. Multiply all these together, and observe that every integer on the left is even. We obtain

$$2 \cdot 4 \cdot 6 \cdots (p-1) \equiv \left(\frac{p-1}{2}\right)!(-1)^{1+2+\cdots+(p-1)/2} \pmod p,$$

an since

$$1 + 2 + \cdots + \frac{p-1}{2} = \frac{1}{2}\frac{(p-1)}{2}\left(\frac{p-1}{2} + 1\right) = \frac{p^2-1}{8},$$

we get

$$2^{(p-1)/2}\left(\frac{p-1}{2}\right)! \equiv \left(\frac{p-1}{2}\right)!(-1)^{(p^2-1)/8} \pmod p.$$

However, $((p-1)/2)!$ is not a multiple of p, so we get

$$\left(\frac{2}{p}\right) = 2^{(p-1)/2} \equiv (-1)^{(p^2-1)/8} \pmod p,$$

as claimed. □

Part (3) of Proposition 6.6 says that -1 (equivalently $p-1$) is a quadratic residue modulo p iff p is a prime of the form $p = 4k+1$, and a nonresidue iff p is of the form $p = 4k+3$. Part (4) says that 2 is quadratic residue modulo p iff p is of the form $p = 8k+1$ or $p = 8k+7$, and a nonresidue iff p is of the form $p = 8k+3$ or $p = 8k+5$.

Remark: Another proof of Part (4) can be given using a primitive eighth root of unity. Here is a slick proof due to Jean–Pierre Serre (see [20]). If p is an odd prime, then p is of the form $4k \pm 1$, so $p^2 - 1 \equiv 0 \pmod 8$. Since the multiplicative group of the finite field $\mathbb{F}_{p^2}$ is cyclic of order p^2-1, and since 8 divides p^2-1, if g generates $\mathbb{F}_{p^2}$ then $\alpha = g^{(p^2-1)/8} \in \mathbb{F}_{p^2}^*$ has order 8 (a primitive eigth root of unity), and let $y = \alpha + \alpha^{-1}$. Since α has order 8, we have $\alpha^4 = -1$, so $\alpha^2 + \alpha^{-2} = 0$, and thus $y^2 = (\alpha + \alpha^{-1})^2 = 2$. Since $\mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$ is a subfield of $\mathbb{F}_{p^2}$ and y is a square root of 2 in $\mathbb{F}_{p^2}$, from Proposition 6.4(1) and Euler's criterion,

$$\left(\frac{2}{p}\right) = y^{p-1}.$$

Since p is prime, we also have

$$y^p = (\alpha + \alpha^{-1})^p = \alpha^p + \alpha^{-p}.$$

If $p \equiv \pm 1 \pmod 8$, since $\alpha^8 = 1$, we get $y^p = \alpha + \alpha^{-1} = y$, and thus

$$\left(\frac{2}{p}\right) = y^{p-1} = 1.$$

If $p \equiv \pm 5 \pmod 8$, since $\alpha^8 = 1$ and $\alpha^4 = -1$, we get

$$y^p = \alpha^5 + \alpha^{-5} = -(\alpha + \alpha^{-1}) = -y.$$

This implies that

$$\left(\frac{2}{p}\right) = y^{p-1} = -1.$$

Using Propositions 6.5 and 6.6, we can evaluate $\left(\frac{m}{p}\right)$ provided that we know how to factor m. Actually, by extending $\left(\frac{m}{p}\right)$ to the Jacobi symbol and using the quadratic reciprocity law, it is possible to evaluate $\left(\frac{m}{p}\right)$ using Euclidean division, without knowing how to factor.

The distribution of the quadratic residues is a topic of great importance. In spite of intense research, the current state of knowledge is still rather incomplete. The following result from Niven, Zuckerman, and Montgomery [16] (Section 3.3, Theorem 3.9) tells us that the least positive quadratic nonresidue cannot be too large.

Proposition 6.7. *If p is an odd prime, then the least positive quadratic nonresidue n modulo p satisfies the inequality $n \leq \sqrt{p}$.*

Proof. Let m be the smallest positive integer such that $mn > p$, so that $(m-1)n \leq p$. Since $n \geq 2$ and p is an odd prime, we have $(m-1)n < p$. Thus

$$0 < mn - p < n.$$

Since n is the least positive nonresidue mod p, the number $mn - p$ must be a quadratic residue mod p, so

$$\left(\frac{mn-p}{p}\right) = 1.$$

Since $\left(\frac{n}{p}\right) = -1$, by Proposition 6.5 and Proposition 6.6, we get

$$\left(\frac{mn-p}{p}\right) = \left(\frac{mn}{p}\right) = \left(\frac{m}{p}\right)\left(\frac{n}{p}\right) = -\left(\frac{m}{p}\right),$$

so

$$\left(\frac{m}{p}\right) = -1.$$

Since n is the smallest positive quadratic nonresidue mod p, we must have $m \geq n$, and since $(m-1)n < p$, we get

$$(n-1)^2 < (n-1)n \leq (m-1)n < p.$$

Thus $n - 1 < \sqrt{p}$, that is, $n \leq \sqrt{p}$, as claimed. □

Euler's criterion has the following corollary which is the basis of the Solovay–Strassen test.

Proposition 6.8. *If p is an odd prime, then for any integer $m \in \{1, \ldots, p-1\}$, we have*

$$\left(\frac{m}{p}\right) m^{(p-1)/2} \equiv 1 \pmod{p}.$$

Proof. Since $m \in \{1, \ldots, p-1\}$, the Legendre symbol $\left(\frac{m}{p}\right)$ is not zero, and Euler's criterion tells us that $\left(\frac{m}{p}\right)$ and $m^{(p-1)/2}$ mod p are either both $+1$ or both -1, which implies that their product is 1 modulo p. □

By taking the contrapositive, it appears that we obtain a criterion for compositeness used in the Solovay–Strassen test:

If $n \geq 3$ is odd and if there is some $a \in \{2, \ldots, n-1\}$ such that

$$\left(\frac{a}{n}\right) a^{(n-1)/2} \not\equiv 1 \pmod{n},$$

then n is composite.

However, we haven't yet defined $\left(\frac{a}{n}\right)$ for a composite number n. This can be done by introducing the Jacobi symbol. Having made sense of $\left(\frac{a}{n}\right)$ where n is composite, two issues remain:

1. Proving that only a fraction of numbers in $\{2, \ldots, n-1\}$ are liars, that is, satisfy the condition of Proposition 6.8 even though n is composite.

2. Find an efficient method to evaluate $\left(\frac{a}{n}\right)a^{(n-1)/2}$ modulo n.

Fortunately, at most half of the integers in $\{2, \ldots, n-1\}$ are liars. For the second point, we make use of the famous quadratic reciprocity law.

6.3 The Jacobi Symbol

The definition of the Jacobi symbol favors the quadratic reciprocity law at the expense of the connection with quadratic residues. As a consequence, $\left(\frac{m}{n}\right) = 1$ does not necessarily imply that m is a quadratic residue modulo n. On the positive side, properties of the Jacobi symbol yield a more efficient algorithm for evaluating $\left(\frac{m}{n}\right)$.

Definition 6.3. Let $P \geq 3$ be a positive odd integer and let $P = p_1^{j_1} p_2^{j_2} \cdots p_k^{j_k}$ be the prime factorization of P. For any integer m, the *Jacobi symbol* $\left(\frac{m}{P}\right)$ is defined as follows:

$$\left(\frac{m}{P}\right) = \left(\frac{m}{p_1}\right)^{j_1}\left(\frac{m}{p_2}\right)^{j_2}\cdots\left(\frac{m}{p_k}\right)^{j_k}.$$

By convention, $\left(\frac{m}{1}\right) = 1$.

Clearly,

$$\left(\frac{1}{P}\right) = 1,$$

and the Jacobi symbol agrees with the Legendre symbol if P is prime. If $\gcd(m, P) > 1$, then m is a multiple of some prime factor p_i of P, so $\left(\frac{m}{P}\right) = 0$, and otherwise $\left(\frac{m}{P}\right) = \pm 1$.

Since the primes $p_1, \ldots, p_k$ are all distinct, m is a quadratic residue modulo P iff $\left(\frac{m}{p_i}\right) = 1$ for all p_i, but is is possible that $\left(\frac{m}{P}\right) = 1$ even though m is a quadratic nonresidue modulo P. For example, we have

$$\left(\frac{2}{15}\right) = \left(\frac{2}{3}\right)\left(\frac{2}{5}\right) = (-1)(-1) = 1,$$

but 2 is not a quadratic residue modulo 15. On the other hand, if $\left(\frac{m}{P}\right) \equiv -1 \pmod{p}$, then $\gcd(m, P) = 1$, and m is a quadratic nonresidue modulo P.

The Jacobi symbol satisfies the following properties which are very useful for evaluating it.

Proposition 6.9. *For any odd positive integers $m, n \geq 3$, and any integers a, b, the following properties hold:*

(1) $\left(\dfrac{ab}{m}\right) = \left(\dfrac{a}{m}\right)\left(\dfrac{b}{m}\right)$.

(2) $\left(\dfrac{a}{mn}\right) = \left(\dfrac{a}{m}\right)\left(\dfrac{a}{n}\right)$.

(3) If $a \equiv b \pmod{m}$, *then* $\left(\dfrac{a}{m}\right) = \left(\dfrac{b}{m}\right)$.

(4) If $\gcd(a, m) = 1$, *then* $\left(\dfrac{a^2 b}{m}\right) = \left(\dfrac{b}{m}\right)$.

(5) If $\gcd(a, m) = 1$, *then* $\left(\dfrac{a}{m^2 n}\right) = \left(\dfrac{a}{n}\right)$.

(6) $\left(\dfrac{2^{2k}a}{m}\right) = \left(\dfrac{a}{m}\right)$ *and* $\left(\dfrac{2^{2k+1}a}{m}\right) = \left(\dfrac{2}{m}\right)\left(\dfrac{a}{m}\right)$, *for all* $k \geq 1$.

(7) $\left(\dfrac{-1}{m}\right) = (-1)^{\frac{m-1}{2}}$, *or equivalently*

$$\left(\frac{-1}{m}\right) = \begin{cases} 1 & \text{if } m \equiv 1 \bmod 4 \\ -1 & \text{if } m \equiv 3 \bmod 4. \end{cases}$$

(8) $\left(\dfrac{2}{m}\right) = (-1)^{\frac{m^2-1}{8}}$, *or equivalently*

$$\left(\frac{2}{m}\right) = \begin{cases} 1 & \text{if } m \equiv 1 \bmod 8 \text{ or } m \equiv 7 \bmod 8 \\ -1 & \text{if } m \equiv 3 \bmod 8 \text{ or } m \equiv 5 \bmod 8. \end{cases}$$

Proof. Parts (1)–(4) follow easily from Propositions 6.5 and 6.6, and Definition 6.3. Part (5) follows from Part (2). For Part 6, observe that $\left(\frac{4}{m}\right) = \left(\frac{2}{m}\right)^2 = 1$, and then apply (4) repeatedly to eliminates factors of 4 in the "numerator." For Part (7), write $m = p_1 p_2 \cdots p_k$ as a product of odd prime factors p_i, not necessarily distinct. Then, we have

$$m = \prod_{i=1}^{k}(1 + p_i - 1) = 1 + \sum_{i=1}^{k}(p_i - 1) + R,$$

where R consists of a sum of products of at least two factors of the form $p_i - 1$, so that R is a multiple of 4. Hence,

$$m \equiv 1 + \sum_{i=1}^{k}(p_i - 1) \pmod{4},$$

or

$$\frac{m-1}{2} \equiv \sum_{i=1}^{k} \frac{(p_i - 1)}{2} \pmod 2.$$

Therefore,

$$\left(\frac{-1}{m}\right) = \prod_{i=1}^{k} \left(\frac{-1}{p_i}\right) = \prod_{i=1}^{k} (-1)^{(p_i-1)/2} = (-1)^{\sum_{i=1}^{k}(p_i-1)/2} = (-1)^{(m-1)/2},$$

as claimed.

For Part (8), write

$$m^2 = \prod_{i=1}^{k} (1 + p_i^2 - 1) = 1 + \sum_{i=1}^{k} (p_i^2 - 1) + R,$$

where R is a sum of products of at least two factors of the form $p_i^2 - 1$. Now, since $p_i \equiv \pm 1 \pmod 4$, we have $p_i^2 - 1 \equiv 0 \pmod 8$, and thus,

$$m^2 \equiv 1 + \sum_{i=1}^{k} (p_i^2 - 1) \pmod{64},$$

which yields

$$\frac{m^2-1}{8} \equiv \sum_{i=1}^{k} \frac{p_i^2 - 1}{8} \pmod 8.$$

The above congruence also holds modulo 2, so we get

$$\left(\frac{2}{m}\right) = \prod_{i=1}^{k} \left(\frac{2}{p_i}\right) = \prod_{i=1}^{k} (-1)^{(p_i^2-1)/8} = (-1)^{(m^2-1)/8},$$

as claimed. □

Remark: Proposition 6.9 holds trivally if $m = 1$ or $n = 1$.

If p is an odd prime, deciding whether some integer m is a quadratic residue mod p can be done by computing the Legendre symbol $\left(\frac{m}{p}\right)$. Using the Jacobi symbol, this computation can be performed in polynomial time (see Section 6.5). However, finding a square root of m modulo p is hard. So far, no polynomial-time algorithm is known. The inherent difficulty of finding square roots modulo p can be exploited in designing encryption schemes.

We observed that if P is not prime, the fact that $\left(\frac{m}{P}\right) = 1$ does not necessarily imply that m is quadratic residue mod P. However, if $P = pq$ is the product of two distinct odd primes, then we have the following result.

Proposition 6.10. *Let p and q be two distinct odd primes. If $\left(\frac{m}{p}\right) = 1$ and $\left(\frac{m}{q}\right) = 1$, then m is a quadratic residue modulo pq.*

Proof. By hypothesis there exist c_1 and c_2 such that $c_1^2 \equiv m \pmod{p}$ and $c_2^2 \equiv m \pmod{q}$. Since p and q are distinct primes, $\gcd(p, q) = 1$, so by the Chinese remainder theorem we can find c such that

$$\begin{aligned} c &\equiv c_1 \pmod{p} \\ c &\equiv c_2 \pmod{q}. \end{aligned}$$

We deduce that

$$c^2 \equiv c_1^2 \equiv m \pmod{p} \quad \text{and} \quad c^2 \equiv c_2^2 \equiv m \pmod{q},$$

and since $\gcd(p, q) = 1$, we conclude that $c^2 \equiv m \pmod{pq}$, which means that m is a quadratic residue modulo pq. □

The fact that if $N = pq$ is the product of two distinct odd primes kept secret, then it is generally difficult to decide whether some integer m is a quadratic residue modulo N (since computing the Jacobi symbol $\left(\frac{m}{N}\right)$ does not help), is the basis of an encryption scheme due to Goldwasser and Micali.

The Goldwasser–Micali encryption scheme is elegant and easy to prove correct but the problem is that this encryption scheme encodes one bit at a time, so it is not really practical because the length of the encoded message (the ciphertext) is equal to the length of the plaintext multiplied by $\log_2 N$. If N has 1000 bits, the expansion factor is 1000. We refer the interested reader to Hoffstein, Pipher and Silverman [8] (Chapter 3, Section 3.10) for a presentation of the Goldwasser–Micali encryption scheme.

6.4 The Solovay–Strassen Test; E-Witnesses and E-Liars

Now that we have the Jacobi symbol, $\left(\frac{a}{m}\right)$ makes sense if m is an odd positive integer, and we are we ready to present the Solovay–Strassen test. This test relies on the following fact:

If $n \geq 3$ is odd and if there is some $a \in \{2, \ldots, n-1\}$ such that

$$\left(\frac{a}{n}\right) a^{(n-1)/2} \not\equiv 1 \pmod{n},$$

then n is composite.

Definition 6.4. Let $n \geq 3$ be any odd integer.

(1) An integer a such that $2 \leq a \leq n-1$ is called an *Euler witness*, for short an *E-witness* for n, if

$$\left(\frac{a}{n}\right) a^{(n-1)/2} \not\equiv 1 \pmod{n}.$$

(2) If $n > 3$ is an odd composite, then an integer a with $1 \leq a \leq n-1$ is an *Euler liar*, for short an *E-liar* for n, if

$$\left(\frac{a}{n}\right) a^{(n-1)/2} \equiv 1 \pmod{n}.$$

The set of E-liars is denoted by L_n^E. An odd composite number n such that a with $2 \leq a \leq n-2$ is an E-liar for n is called an *Euler pseudoprime base* a.

Consider $n = 325$, a composite. For $a = 15$, we have $\gcd(15, 325) = 5$, hence $\left(\frac{15}{325}\right) = 0$, and 15 is an E-witness. For $a = 2$, we have $2^{162} \equiv 129 \pmod{325}$, so 2 is also an E-witness. For $a = 7$, we have $7^{162} \equiv 324 \pmod{325}$, and $\left(\frac{7}{325}\right) = -1$; consequently, 7 is an E-liar for 325.

The first fact to observe is that every E-liar is an F-liar.

Proposition 6.11. *For any odd composite $n > 3$, we have $L_n^E \subseteq L_n^F$.*

Proof. If a is an E-liar, then

$$\left(\frac{a}{n}\right) a^{(n-1)/2} \equiv 1 \pmod{n},$$

and since $\left(\frac{a}{n}\right) \in \{-1, 1\}$, we must have $a^{(n-1)/2} \equiv \pm 1 \pmod{n}$, which yields

$$a^{n-1} \equiv 1 \pmod{n},$$

showing that a is an F-liar. □

The second fact is that the number of E-liars is at most half of the number of elements in $(\mathbb{Z}/n\mathbb{Z})^*$. The reason is that L_n^E is a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$.

Theorem 6.12. *If $n > 3$ is an odd composite, then the set L_n^E of E-liars is a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$.*

Proof. We already know that $L_n^E \subseteq L_n^F \subseteq (\mathbb{Z}/n\mathbb{Z})^*$. Obviously, $1 \in L_n^E$. If $a, b \in L_n^E$, then we have $\left(\frac{a}{n}\right) a^{(n-1)/2} \equiv 1 \pmod{n}$ and $\left(\frac{b}{n}\right) b^{(n-1)/2} \equiv 1 \pmod{n}$. By Property (1) of the Jacobi symbol (Proposition 6.9), we get

$$\left(\frac{ab}{n}\right)(ab)^{(n-1)/2} \equiv \left(\frac{a}{n}\right) a^{(n-1)/2} \left(\frac{b}{n}\right) b^{(n-1)/2} \equiv 1 \cdot 1 \equiv 1 \pmod{n}$$

which shows that $ab \in L_n^E$. Therefore, L_n^E is a subgroup of L_n^F.

It remains to show that there is some $a \in (\mathbb{Z}/n\mathbb{Z})^*$ which does not belong to L_n^E; that is, that there is some E-witness in $(\mathbb{Z}/n\mathbb{Z})^*$. There are two cases:

Case 1. The number n contains some square factor p^2, for some prime $p \geq 3$. In this case, when we proved (1) of Korselt's criterion (Theorem 5.5), we produced an F-witness a in $(\mathbb{Z}/n\mathbb{Z})^*$. By Proposition 6.11 (using its contrapositive), we conclude that a is an E-witness in $(\mathbb{Z}/n\mathbb{Z})^*$.

Case 2. The number n is squarefree, so we can write $n = pm$, for some odd prime p and some odd number $m \geq 3$ which is not a multiple of p.

Let $b \in \{1, \ldots, p-1\}$ be some quadratic nonresidue modulo p, so that $\left(\frac{b}{p}\right) = -1$. Using the Chinese remainder theorem, we find some a with $1 \leq a \leq n-1$ such that

$$\begin{aligned} a &\equiv b \pmod{p} \\ a &\equiv 1 \pmod{m}. \end{aligned}$$

We claim that a is an E-witness in $(\mathbb{Z}/n\mathbb{Z})^*$. Since b is a nonresidue modulo p, the prime p does not divide a, and $\gcd(a, m) = 1$, so $\gcd(a, n) = 1$ and $a \in (\mathbb{Z}/n\mathbb{Z})^*$.

Next, observe that

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p}\right)\left(\frac{a}{m}\right) = \left(\frac{b}{p}\right)\left(\frac{1}{m}\right) = (-1) \cdot 1 = -1.$$

If a were an E-liar, then the fact that $\left(\frac{a}{n}\right) = -1$ would imply that

$$a^{(n-1)/2} \equiv -1 \pmod{n}.$$

Since m divides n, we would have

$$a^{(n-1)/2} \equiv -1 \pmod{m},$$

contradicting the fact that $a \equiv 1 \pmod{m}$. Therefore, a is indeed an E-witness in $(\mathbb{Z}/n\mathbb{Z})^*$ for n. □

Theorem 6.12 implies that the order $|L_n^E|$ of the group L_n^E divides the order $\varphi(n)$ of $(\mathbb{Z}/n\mathbb{Z})^*$, and therefore,

$$|L_n^E| \leq \frac{\varphi(n)}{2} \leq \frac{n-1}{2}.$$

Here is the *Solovay–Strassen* primality test.

Solovay–Strassen test

The input is an odd integer $n > 3$.

procedure solovay-strassen(n)
begin
  Choose random integer $a \in \{2, \ldots, n-2\}$;
  if $\left(\frac{a}{n}\right) a^{(n-1)/2} \not\equiv 1 \pmod{n}$
    then $c := 1$; **return** c; **exit**; ($*$ n is a composite $*$)
  else $c := 0$; **return** c ($*$ n is a probable prime $*$)
end

The Solovay–Strassen test is a Monte-Carlo algorithm. If the procedure returns $c = 1$, then n is composite: that is,

$$\mathsf{Pr}(n \text{ composite} \mid \text{solovay strassen return } c = 1) = 1.$$

If n is composite, then the solovay-strassen test returns $c = 1$ for at least $1/2$ of the number of choices for a; that is,

$$\mathsf{Pr}(\text{solovay strassen return } c = 1 \mid n \text{ is composite}) \geq \frac{1}{2}.$$

If we repeat the Solovay–Strassen test ℓ times, as in the case of the Miller–Rabin test, we obtain the fact that

$$\mathsf{Pr}(\text{solovay strassen return } c = 0 \ \ell \text{ times} \mid n \text{ is composite}) \leq \left(\frac{1}{2}\right)^{\ell},$$

and that (approximately)

$$\mathsf{Pr}(n \text{ is prime} \mid \text{solovay strassen return } c = 0 \ \ell \text{ times}) \geq 1 - \frac{\ln(n)}{2^{\ell}}.$$

We still have to show how the Jacobi symbol can be evaluated quickly. For this we need the quadratic reciprocity law.

6.5 The Quadratic Reciprocity Law

The *quadratic reciprocity law*, first stated by Euler (in a complicated form) around 1744–1746, was rediscovered in 1785 by Legendre who gave a partial proof. Gauss discovered the quadratic reciprocity independently at the age of eighteen and was the first one to give a complete proof in 1796. In fact, according to Dirichlet–Dedekind [12] (Chapter 3, Section 42), Gauss gave six different proofs of the quadratic reciprocity law! A seventh proof was found in Gauss' Nachlass. Curiously, some authors (including Apostol) claim that Gauss gave eight different proofs.

The quadratic recipocity law states that if p and q are distinct odd primes, then $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$ unless $p \equiv q \equiv 3 \pmod{4}$, in which case $\left(\frac{p}{q}\right) = -\left(\frac{q}{p}\right)$. The extension of the quadratic reciprocity law to the Jacobi symbol for distinct odd integers $m, n \geq 3$ such that $\gcd(m, n) = 1$ is easy.

Theorem 6.13. *(Quadratic reciprocity law) If m and n are any odd integers $m, n \geq 3$ such that $\gcd(m, n) = 1$, then*

$$\left(\frac{m}{n}\right) = (-1)^{\frac{m-1}{2}\frac{n-1}{2}}\left(\frac{n}{m}\right).$$

Equivalently,

$$\left(\frac{m}{n}\right) = \begin{cases} \left(\frac{n}{m}\right) & \text{if } m \equiv 1 \bmod 4 \text{ or } n \equiv 1 \bmod 4 \\ -\left(\frac{n}{m}\right) & \text{if } m \equiv 3 \bmod 4 \text{ and } n \equiv 3 \bmod 4. \end{cases}$$

Furthermore,

$$\left(\frac{1}{m}\right) = 1$$
$$\left(\frac{-1}{m}\right) = (-1)^{\frac{m-1}{2}}$$
$$\left(\frac{2}{m}\right) = (-1)^{\frac{m^2-1}{8}},$$

or equivalently

$$\left(\frac{-1}{m}\right) = \begin{cases} 1 & \text{if } m \equiv 1 \bmod 4 \\ -1 & \text{if } m \equiv 3 \bmod 4, \end{cases}$$

and

$$\left(\frac{2}{m}\right) = \begin{cases} 1 & \text{if } m \equiv 1 \bmod 8 \text{ or } m \equiv 7 \bmod 8 \\ -1 & \text{if } m \equiv 3 \bmod 8 \text{ or } m \equiv 5 \bmod 8. \end{cases}$$

Observe that the quadratic reciprocity law holds trivially if $\gcd(m, n) > 1$, since in this case $\left(\frac{m}{n}\right) = \left(\frac{n}{m}\right) = 0$.

Remark: We could define the Legendre symbol for $n = 2$. Since every odd number $m = 2k + 1$ is a quadratic residue modulo 2 and there are no quadratic nonresides, we can set

$$\left(\frac{m}{2}\right) = \begin{cases} +1 & \text{if } m \text{ is odd} \\ 0 & \text{if } m \text{ is even.} \end{cases}$$

But then, the quadratic reciprocity law fails. Similarly, we could define the Jacobi symbol if n is even, but this is futile since the quadratic reciprocity law also fails. This is the reason why the Legendre symbol $\left(\frac{m}{p}\right)$ is only defined for an odd prime p and the Jacobi symbol $\left(\frac{m}{n}\right)$ for a positive odd integer n.

We prove the quadratic reciprocity law in Section 6.7, but first we show how it can be used togther with the properties stated in Proposition 6.9 to evaluate quickly the Jacobi symbol.

We use the following steps recursively to evaluate the Jacobi symbol $\left(\frac{a}{n}\right)$.

(1) Reduce modulo n. If $a \notin \{1, \ldots, n-1\}$, compute $\left(\frac{a \bmod n}{n}\right)$.

(2) If $a = 0$, then the result is 0.

(3) If $a = 1$, then the result is 1.

(4) Remove factors of 4 from the numerator. If 4 divides a, then compute $\left(\frac{a/4}{n}\right)$.

(5) Remove factors of 2 from the numerator. If 2 divides a, then if $n \equiv 1, 7 \pmod 8$, compute $\left(\frac{a/2}{n}\right)$, else if $n \equiv 3, 5 \pmod 8$, compute $-\left(\frac{a/2}{n}\right)$.

(6) Apply quadratic reciprocity, Case 1. If $n > a > 1$ and ($a \equiv 1 \pmod 4$ or $n \equiv 1 \pmod 4$), then compute $\left(\frac{n \bmod a}{a}\right)$.

(7) Apply quadratic reciprocity, Case 2. If $n > a$, $a \equiv 3 \pmod 4$ and $n \equiv 3 \pmod 4$, then compute $-\left(\frac{n \bmod a}{a}\right)$.

The rules for evaluating the Jacobi symbol are more powerful than the rules for evaluating the Legendre symbol because in (6) and (7) it is not necessary to assume that a is prime. Thus, there is no need for factoring a, which is a great advantage, because factoring is generally considered hard.

Here is an illustration of the use of the above rules to evaluate a Jacobi symbol. Consider

$$\left(\frac{773}{1373}\right).$$

In the present case, 773 and 1373 are prime, so we are in fact computing the Legendre symbol. We have

$$\left(\frac{773}{1373}\right) \overset{6}{=} \left(\frac{600}{773}\right) \overset{4}{=} \left(\frac{150}{773}\right) \overset{5}{=} -\left(\frac{75}{773}\right) \overset{6}{=} -\left(\frac{23}{75}\right) \overset{7}{=} \left(\frac{6}{23}\right) \overset{5}{=}$$
$$\left(\frac{3}{23}\right) \overset{6}{=} -\left(\frac{2}{3}\right) \overset{5}{=} \left(\frac{1}{3}\right) \overset{3}{=} 1.$$

Therefore, 773 is a quadratic residue modulo 1373. Another way to show this is to use the Euler criterion and to compute 773^{686} mod 1373 (we find that the result is indeed 1).

The following example taken from Hoffstein, Pipher and Silverman [8] shows the superiority of the Jacobi symbol. Consider computing

$$\left(\frac{228530738017}{9365449244297}\right),$$

where the two numbers involved are indeed prime (check it using Miller-Rabin or Solovay–Strassen!). By the quadratic reciprocity law, we get

$$\left(\frac{228530738017}{9365449244297}\right) = \left(\frac{9365449244297}{228530738017}\right),$$

and since $9365449244297 \equiv 224219723617 \pmod{228530738017}$, we have

$$\left(\frac{9365449244297}{228530738017}\right) = \left(\frac{224219723617}{228530738017}\right).$$

Now, although this is not obvious, 224219723617 is composite, so to proceed with the Legendre symbol we need to factor 224219723617, not an easy task. With the Jacobi symbol, we can apply the law of quadratic reciprocity and then reduce modulo the denominator and we get

$$\left(\frac{224219723617}{228530738017}\right) = \left(\frac{228530738017}{224219723617}\right) = \left(\frac{4311014400}{224219723617}\right).$$

Since $4311014400 = 2^{10} \cdot 4209975$, we get

$$\left(\frac{4311014400}{224219723617}\right) = \left(\frac{4209975}{224219723617}\right) = \left(\frac{224219723617}{4209975}\right) = \left(\frac{665092}{4209975}\right).$$

We will let the reader finish the computation and eventually find that the answer is -1.

Here is an iterative algorithm for evaluating the Jacobi symbol $\left(\frac{a}{n}\right)$ where $n \geq 3$ is an odd integer.

Evaluation of the Jacobi Symbol $\left(\frac{a}{n}\right)$

function jacobi(a, n)
  $b := a \bmod n$; $c := n$; $s := 1$;
  while $b \geq 2$ do
    while $4 \mid b$ **do**
      $b := b/4$
    endwhile;
    if $2 \mid b$ **then**
      if $c \equiv 3, 5 \pmod 8$ **then** $s := -s$ **endif**;
      $b := b/2$
    endif;
    if $b = 1$ **then return** s **exit endif**;
    if $b \equiv c \equiv 3 \pmod 4$ **then** $s := -s$ **endif**;
    $b := c \bmod b$; $c := b$
  endwhile;
  return sb
end

It is not hard to see that the invariant

$$s\left(\frac{b}{c}\right) = \left(\frac{a}{n}\right)$$

is maintained during execution of the program. Also, if $\gcd(a, n) > 1$, then at some point b becomes 0, so there is no need to compute $\gcd(a, n)$. We leave it as an exercise to prove that the above program computes the Jacobi symbol $\left(\frac{a}{n}\right)$; for help, consult Dietzfelbinger [4] (Section 6.3). It is also easy to prove that the number of iterations of the main **while** loop is $O(\log n)$ and that that the program runs in $O((\log n)^2)$ bit operations if $|a| < n$ (see Crandall and Pomerance [3], Chapter 2).

It is remarkable that *deciding* whether a is a quadratic residue modulo p (p prime) can be done quickly (in polynomial time in $\log p$), basically the same complexity as computing the gcd. However *finding* a square root in $\mathbb{Z}/p\mathbb{Z}$ is hard (with p prime). So far, no known deterministic polynomial-time algorithm is known. It is known that if the ERH holds, then there is a quadratic nonresidue $d < 2(\log p)^2$. From this, a square root can be found in polynomial time, if it exists.

6.6 A Randomized Algorithm to Find a Square Root mod p

If p is an odd prime, we saw in Section 6.1 that if p is of the form $p = 4k + 3$, then a square root of a quadratic residue a is easily found: $a^{(p+1)/4}$ is a square root of a mod p. We now show that if $p = 4k + 1$ is prime, then there is a randomized algorithm to find a square root of a quadratic residue a mod p.

First, we show that if $p = 8k + 3$ or if $p = 8k + 7$ is prime, then it is easy to find a square root of a quadratic residue a mod p. We can write $p - 1 = 2(4k + 1)$ or $p - 1 = 2(4k + 3)$, that is, $p - 1 = 2t$ where t is odd. Since a is a quadratic residue mod p, by Euler's criterion

$$a^{(p-1)/2} \equiv a^t \equiv 1 \pmod{p}.$$

Thus $a^{t+1} \equiv a \pmod{p}$, and since t is odd, $t + 1$ is even, so $y = a^{(t+1)/2} = a^{(p+1)/4} \bmod p$ is a square root of a.

If $p = 8k + 5$ is prime, then $p - 1 = 4(2k + 1)$, that is, $p - 1 = 4t = 2^s t$ with $s = 2$ and t odd, and if $p = 8k + 1$ is prime, then $p - 1 = 8k$. In the second case, $p - 1 = 2^s t$ with $s \geq 3$ and t odd. Since a is a quadratic residue mod p, by Euler's criterion we still have $a^{(p-1)/2} \equiv 1 \pmod{p}$, namely

$$a^{2^{s-1}t} \equiv 1 \pmod{p}.$$

We deduce that $a^{1+2^{s-1}t} \equiv a \pmod{p}$, but since $s \geq 2$, the number $1 + 2^{s-1}t$ is odd, so this does not help.

What we have though, is

$$(a^t)^{2^{s-1}} \equiv 1 \pmod{p}. \qquad (*)$$

This implies that the order of $A = a^t$ is a divisor of 2^{s-1}, and let this order be 2^h, with $0 \leq h \leq s - 1$.

Surprisingly, having a quadratic nonresidue d mod p helps. By Euler's criterion, since $(p-1)/2 = 2^{s-1}t$, we have

$$(d^t)^{2^{s-1}} \equiv d^{2^{s-1}t} \equiv d^{(p-1)/2} \equiv -1 \pmod{p}.$$

If we let $D \equiv d^t \pmod{p}$, then we see that

$$D^{2^{s-1}} \equiv -1 \pmod{p},$$

so D has order exactly 2^s, and so does D^{-1}. This is because since D is invertible, $D^k \equiv 1 \pmod{p}$ iff $1 \equiv (D^{-1})^k \pmod{p}$. Since $(\mathbb{Z}/p\mathbb{Z})^*$ is a cyclic group, by Proposition 4.18(e), it contains a single subgroup of order 2^s generated by D^{-1}, and since A has order at most 2^{s-1}, it must belong to this subgroup, so it must be some power of D^{-1}. Thus $A \equiv (D^{-1})^r \pmod{p}$ for some r such that $1 \leq r \leq 2^s$.

Since the order of A is 2^h we have

$$A^{2^h} \equiv ((D^{-1})^r)^{2^h} \equiv (D^{-1})^{r2^h} \equiv 1 \pmod{p},$$

and since D^{-1} has order 2^s, we see that 2^s divides $r2^h$. Therefore $r = r_1 2^{s-h}$, and since $r \leq 2^s$, we have $1 \leq r_1 \leq 2^h$ with $0 \leq h \leq s-1$. If r_1 is even, say $r_1 = 2r_2$, then $r = r_2 2^{s-h+1}$, and then

$$A^{2^{h-1}} \equiv ((D^{-1})^{r_2 2^{s-h+1}})^{2^{h-1}} \equiv (D^{-1})^{r_2 2^s} \equiv ((D^{-1})^{2^s})^{r_2} \equiv 1 \pmod{p},$$

contradicting the fact that A has order 2^h.

Therefore, we showed that $A \equiv (D^{-1})^r \pmod{p}$, with $r = r_1 2^{s-h}$, for some odd r_1 such that $1 \leq r_1 < 2^h$, and $0 \leq h \leq s-1$.

If $h = 0$, then $r_1 = 1$ and $r = 2^s$, in which case $A \equiv (D^{-1})^{2^s} \equiv 1 \pmod{p}$, so equivalently we set $r = 0$. Thus either $r = 0$, or $r = r_1 2^{s-h}$ for some odd r_1 with $1 \leq r_1 < 2^h$ and $1 \leq h \leq s-1$. Observe that if $r \neq 0$ then $r < 2^s$.

If we decompose $r = r_1 2^{s-h} \neq 0$ into a sum of powers of 2, which amounts to writing r in binary, we obtain

$$r = 2^{s-h} + 2^{i_2} + \cdots + 2^{i_k}$$

for some increasing sequence $s - h < i_2 < \cdots < i_k$. Since $r < 2^s$, we have $1 \leq k \leq s-1$, and since $h \leq s-1$, the number r is even. We let $\mu = r/2$, and because $r < 2^s$, we have $\mu < 2^{s-1}$.

Since $A \equiv (D^{-1})^r \equiv D^{-2\mu} \pmod{p}$, we have

$$AD^{2\mu} \equiv 1 \pmod{p}, \tag{$**$}$$

that is, $a^t D^{2\mu} \equiv 1 \pmod{p}$, and thus

$$a^{t+1} D^{2\mu} \equiv a \pmod{p}.$$

Since t is odd, $t+1$ is even, and so $y = a^{(t+1)/2}D^{\mu}$ is a square root of $a \bmod p$.

Observe that $(**)$ is equivalent to

$$D^{2\mu} \equiv A^{-1} \pmod{p},$$

that is, *we are trying to find a discrete logarithm for A^{-1} to the base D*. In general, finding a discrete logarithm is very hard. Luckily, in the present case, the discrete logarithm 2μ can be easily found. What happens is that the pieces 2^{i_j} in the binary decomposition of $r = 2\mu$ can be found using a simple test.

Let us examine the case $s = 2$ more closely. In this case, $p - 1 = 4t$ with t odd, that is $p = 4(2h+1)+1 = 8h+5$. By Theorem 6.13, we have

$$\left(\frac{2}{p}\right) = -1,$$

that is, 2 is a quadratic nonresidue. Then with $A \equiv a^t \pmod{p}$ and $D = 2^t \pmod{p}$, we have $AD^{2\mu} \equiv 1 \pmod{p}$ for some μ such that $0 \leq \mu < 2^{s-1} = 2^1 = 2$. Either $\mu = 0$ and then $a^{(t+1)/2}$ is a square root of a, or $a^{(t+1)/2}2^t$ is is a square root of a. Since $p = 4t+1$, we have $t = (p-1)/4$ and $(t+1)/2 = (p+3)/8$, so either $a^{(p+3)/8}$ is a square root of a, or $a^{(p+3)/8}2^{(p-1)/4}$ is a square root of a.

In summary, we have the following algorithm. Given an odd prime p, write $p - 1 = 2^s t$ with $s \geq 1$ and t odd.

(1) If $s = 1$, then $p \equiv 3, 7 \pmod{8}$, and $y = a^{(t+1)/2} = a^{(p+1)/4} \bmod p$ is a square root of $a \bmod p$.

(2) If $s = 2$, then $p \equiv 5 \pmod{8}$, and either $a^{(p+3)/8}$, or $a^{(p+3)/8}2^{(p-1)/4}$, is a square root of $a \bmod p$.

(3) If $s \geq 3$, then $p \equiv 1 \pmod{8}$. Choose at random some integer $d \in \{2, \ldots, p-1\}$ such that d is quadratic nonresidue mod p. We can check this by computing the Legendre symbol $\left(\frac{d}{p}\right)$ (using the fast method for computing Jacobi symbols).

Compute $A \equiv a^t \pmod{p}$ and $D \equiv d^t \pmod{p}$. Search for μ such that $0 \leq \mu < 2^{s-1}$ and $AD^{2\mu} \equiv 1 \pmod{p}$ as follows:

Initialize m as $m := 0$. For $i = 1, \ldots, s-1$, check whether $(AD^m)^{2^{s-1-i}} \equiv -1 \pmod{p}$.

If yes, then increment m by 2^i; $m := m + 2^i$. Otherwise, increment i by 1; $i := i + 1$. When $i = s$, stop; we found $m = 2\mu$ such that $AD^m \equiv 1 \pmod{p}$, and $y = a^{(t+1)/2}D^{m/2} \bmod p$ is a square root of $a \bmod p$.

The correctness of Step 3 of the algorithm can be shown as follows. If A has order 1, then $r = 0$ and $A \equiv a^t \equiv 1 \pmod{p}$. For $m = 0$, since $AD^0 \equiv A \equiv 1 \pmod{p}$, the test $(AD^m)^{2^{s-1-i}} \equiv -1 \pmod{p}$ fails for $i = 1, \ldots, s-1$, so the program ends with $m = 0$, and $a^{(t+1)/2}$ is indeed a square root of $a \bmod p$.

If A has order 2^h greater than 1, then $1 \leq h \leq s-1$, $r = 2^{s-h}r_1$ with r_1 odd, so $A = (D^{-1})^{r_1 2^{s-h}}$, with $1 \leq r_1 < 2^h$. For $m = 0$,

$$(AD^0)^{2^{s-1-i}} \equiv ((D^{-1})^{r_1 2^{s-h}})^{2^{s-1-i}} \equiv (D^{-1})^{r_1 2^{s+s-h-i-1}} \pmod{p}.$$

For $i = s-h$, we have $s+s-h-i-1 = s-1$, and since r_1 is odd the test

$$(AD^0)^{2^{s-1-(s-h)}} \equiv (D^{-1})^{r_1 2^{s-1}} \equiv ((D^{-1})^{2^{s-1}})^{r_1} \equiv (D^{-1})^{2^{s-1}} \equiv -1 \pmod{p}$$

succeeds, but for $1 \leq i \leq s-h-1$, we have

$$(AD^0)^{2^{s-1-i}} \equiv (D^{-1})^{r_1 2^{s+s-h-i-1}} \equiv ((D^{-1})^{2^s})^{r_1 2^{s-h-i-1}} \equiv 1 \pmod{p},$$

so the test

$$(AD^0)^{2^{s-1-i}} \equiv -1 \pmod{p}$$

fails for $1 \leq i \leq s-h-1$. Thus we set $m = 2^{s-h}$. If $r_1 = 1$, we are done. Otherwise $r = 2^{s-h} + r_2 2^{i_2}$ with $i_2 > s-h$ and r_2 odd. Next we look at

$$(AD^{2^{s-h}})^{2^{s-1-i}} \equiv ((D^{-1})^{2^{s-h}+r_2 2^{i_2}} D^{2^{s-h}})^{2^{s-1-i}} \equiv (D^{-1})^{r_2 2^{s-1+i_2-i}} \pmod{p}.$$

Since $i_2 > s-h$, for $i_2 - i > 0$, that is, $i = s-h+1, \ldots, i_2 - 1$, we have

$$(AD^{2^{s-h}})^{2^{s-1-i}} \equiv (D^{-1})^{r_2 2^{s-1+i_2-i}} \equiv 1 \pmod{p},$$

so the test fails, but for $i = i_2$, we have

$$(AD^{2^{s-h}})^{2^{s-1-i_2}} \equiv (D^{-1})^{r_2 2^{s-1}} \equiv -1 \pmod{p},$$

and the test succeed, so we set $m = m + 2^{i_2}$. If $r_2 = 1$ we are done. Otherwise $r = 2^{s-h} + 2^{i_2} + r_3 2^{i_3}$ with $i_3 > i_2$ and r_3 odd. We repeat the test

$$(AD^m)^{s-1-i} \equiv -1 \pmod{p}$$

starting with $i = i_2 + 1$. As in the previous case we will have

$$(AD^m)^{s-1-i} \equiv 1 \pmod{p}$$

for $i = i_2 + 1, \ldots, i_3 - 1$, but

$$(AD^m)^{s-1-i_3} \equiv -1 \pmod{p},$$

so we set $m = m + 2^{i_3}$. What is going on is now clear, and by an argument by induction left to the reader, we can prove that when the algorithm stops with $i = s$, we have

$$m = r = 2^{s-h} + 2^{i_2} + \cdots + 2^{i_k}.$$

Then $a^{(t+1)/2} D^{m/2} \bmod p$ is a square root of $a \bmod p$.

It is informative to apply the above algorithm to $p = 17$. It is easily seen that the quadratic residues mod 17 are

$$1,\ 4,\ 9,\ 16,\ 8,\ 2,\ 15,\ 13.$$

Let us apply the algorithm to $a = 2$. We have $p - 1 = 16 = 2^4 \cdot 1$, so $s = 4$ and $t = 1$. It easy to see that 2 has order $8 = 2^3 \bmod 17$, and the inverse of 2 mod 17 is 9. Let us pick $d = 5$, which is indeed a quadratic nonresidue mod 17. We could have picked $d = 3$, but $d = 5$ is less trivial. Since $t = 1$, we have

$$A = a^1 = 2, \quad D = d^1 = 5.$$

It is easy to check that 5 has order $16 = 2^4 \bmod 17$. The inverse of 5 modulo 17 is 7, and we are trying to find r such that $2 \cdot 5^r \equiv 1 \pmod{17}$, which is equivalent to $2 \equiv 7^r \pmod{17}$, and to $5^r \equiv 9 \pmod{17}$. It turns out that $r = 10$ works. Let us see how the algorithm finds $r = 10 = 2^1 + 2^3$.

Recall that we are performing the test

$$(AD^m)^{2^{s-1-i}} \equiv -1 \pmod{p},$$

that is

$$(2 \cdot 5^m)^{2^{3-i}} \equiv -1 \pmod{17},$$

starting with $m = 0$, and from $i = 1, \ldots, s - 1$. The index i runs from $i = 1$ to $i = 3$, and so $3 - i$ runs from 2 to 0.

For $m = 0$ and $i = 1$, that is $3 - i = 2$, the algorithm finds that $2^{2^2} \equiv 16 \equiv -1 \pmod{17}$. Thus we set $m = 2^1 = 2$, and we test whether $(2 \cdot 5^2)^{2^{3-i}} \equiv -1 \pmod{17}$ for $i = 2, 3$. We find that $(2 \cdot 5^2)^2 \equiv 1 \pmod{17}$ for $i = 2$ and that $(2 \cdot 5^2)^1 \equiv -1 \pmod{17}$ for $i = 3$, so we set $m = m + 2^3 = 2 + 2^3 = 10$.

The algorithm found $m = 10$, so $y = 2 \cdot 5^5 \pmod{17}$ is a square root of 2 mod 17. We can check that $2 \cdot 5^5 \equiv 11 \pmod{17}$, and $11^2 \equiv 2 \pmod{17}$, so 11 is a square root of 2 mod 17.

As an exercise, the reader should check that with $d = 3$, again the order of 3 is $16 = 2^4$, the inverse of 3 mod 17 is 6, and $2 \equiv 6^2 \pmod{17}$. In this case $m = r = 2$. Since $2^{2^2} \equiv 2^4 \equiv -1 \pmod{17}$ for $i = 1$, we get $m = 2^1 = 2$. Since $2 \cdot 3^2 \equiv 18 \equiv 1 \pmod{17}$, the algorithm terminates with $m = 2$, and we find $2 \cdot 3 = 6$ as square root of 2 mod 17.

Here is another example from Koblitz [10] (Chapter II, Section 2). Let $p = 401$ and $a = 186$. By computing the Legendre symbol $\left(\frac{186}{401}\right)$, we find that a is a quadratic residue mod 401. We also find that 3 is a quadratic nonresidue mod 401 by computing the Legendre symbol $\left(\frac{3}{401}\right)$. We have $p - 1 = 401 - 1 = 400 = 2^4 \times 25$, so $s = 4$ and $t = 25$. We find that

$$A \equiv 186^{25} \equiv 98 \pmod{401}, \quad D \equiv 3^{25} \equiv 268 \pmod{401}.$$

We have $s - 1 - i = 3 - i$ and we are testing whether

$$(98 \cdot 268^m)^{2^{3-i}} \equiv -1 \pmod{401},$$

starting from $m = 0$, with $i = 1, 2, 3$.

For $i = 1$, we have $3 - i = 2$, and we check that

$$98^4 \equiv -1 \pmod{401},$$

so we set $m = 2^1 = 2$. Next $i = 2$, so $3 - i = 1$, and we compute

$$(98 \times 268^2)^2 \equiv (98^2 \bmod 401) \times (268^4 \bmod 401) \equiv 381 \times 20 \equiv 1 \pmod{401},$$

so we increment i to $i = 3$, and we compute

$$(98 \times 268^2)^1 \equiv 98 \times (268^2 \bmod 401) \equiv 98 \times 45 \equiv -1 \pmod{401},$$

so $m = m + 2^3 = 2 + 8 = 10$. The algorithm stops with $m = 10$.

We check that $268^{10} \equiv 356 \pmod{401}$, and that 356 is the inverse of 98 mod 401, so indeed

$$98 \times 268^{10} \equiv 1 \pmod{401}.$$

The square root $a^{(t+1)/2} D^{m/2} \bmod p$ is equal to

$$186^{23} \times 268^5 \bmod 401 = (186^{23} \bmod 401) \times (268^5 \bmod 401) = 103 \times 147 \bmod 401 = 304.$$

We easily confirm that $y = 304$ is a square root of 186 mod 401.

According to Crandall and Pomerance [3] (Chapter 2, Algorithm 2.3.8), the above algorithm was originally invented by Tonelli in 1891, based on ideas of Gauss. They also give another randomized algorithm using arithmetic in the finite field $\mathbb{F}_{p^2}$ due to Cipolla (1907).

If n is composite, there is no known fast method for computing square roots. In fact, it can be shown that doing so is essentially equivalent to factoring n in the following sense: if there is a polynomial time, possibly randomized, algorithm to compute square roots modulo any n, then there is a randomized polynomial time algorithm for factoring n. This is Theorem 14.21 in Motwani and Raghavan [15].

6.7 Proof of the Quadratic Reciprocity Law

At least 150 proofs of the quadratic reciprocity law have been published. Gauss himself gave seven different proofs. We follow a short proof using "Gauss sums" due to Jean-Pierre Serre [20]. This proof is not entirely elementary because it uses the fact that if p and q are distinct odd primes, then there is a field extension Ω of $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ in which there is an element of order q (a primitive qth root of unity). Any algebraic closure of $\mathbb{F}_p$ will do. Since $p^{q-1} \equiv 1 \pmod{q}$ (by Fermat's little theorem), q divides $p^{q-1} - 1$, so the finite field $\mathbb{F}_{p^{q-1}}$ also works since its multiplicative group is cyclic of order $p^{q-1} - 1$.

Theorem 6.14. *(Quadratic reciprocity law for primes (Gauss)) For any two distinct odd primes p, q, we have*

$$\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{q}{p}\right).$$

Proof. Let Ω be any field extension of $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ in which there is a primitive qth root of unity, for example $\Omega = \mathbb{F}_{p^{q-1}}$. If $w \in \Omega$ is a primitive qth root of unity, define the *Gauss sum* y by

$$y = \sum_{a \in \mathbb{F}_q} \left(\frac{a}{q}\right) w^a,$$

with $\left(\frac{0}{q}\right) = 0$. Then, we have two steps.

Step 1. We prove that

$$y^2 = (-1)^{(q-1)/2} q.$$

Step 2. We prove that

$$y^{p-1} = \left(\frac{p}{q}\right).$$

If we assume that Step 1 and Step 2 have been established, by Step 1, y is a square root of $(-1)^{(q-1)/2}q$, and by a Proposition 6.4(1)

$$\left(\frac{(-1)^{(q-1)/2}q}{p}\right) = y^{p-1},$$

so by Step 2

$$\left(\frac{(-1)^{(q-1)/2}q}{p}\right) = \left(\frac{p}{q}\right).$$

On the other hand, by Proposition 6.5 and Proposition 6.6, we have

$$\left(\frac{p}{q}\right) = \left(\frac{(-1)^{(q-1)/2}q}{p}\right) = \left(\frac{(-1)^{(q-1)/2}}{p}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}\left(\frac{q}{p}\right),$$

which proves the desired formula. □

Proof of Step 1. We have

$$y^2 = \Big(\sum_{a\in\mathbb{F}_q}\Big(\frac{a}{q}\Big)w^a\Big)\Big(\sum_{b\in\mathbb{F}_q}\Big(\frac{b}{q}\Big)w^b\Big) = \sum_{a,b\in\mathbb{F}_q}\Big(\frac{ab}{q}\Big)w^{a+b} = \sum_{c\in\mathbb{F}_q} w^c\Big(\sum_{a\in\mathbb{F}_q}\Big(\frac{a(c-a)}{q}\Big)\Big),$$

by making the change of variable $c = a+b$. Now, if $a \neq 0$, we have $a(c-a) = -a^2(1-ca^{-1})$, and since $\big(\frac{a}{q}\big) = \pm 1$ since $a \in \mathbb{F}_q$, we have $\big(\frac{a}{q}\big)^2 = 1$, so

$$\Big(\frac{a(c-a)}{q}\Big) = \Big(\frac{-a^2(1-ca^{-1})}{q}\Big) = \Big(\frac{a}{q}\Big)^2\Big(\frac{-1}{q}\Big)\Big(\frac{1-ca^{-1}}{q}\Big) = (-1)^{(q-1)/2}\Big(\frac{1-ca^{-1}}{q}\Big),$$

and thus,

$$(-1)^{(q-1)/2}y^2 = \sum_{c\in\mathbb{F}_q} S_c w^c,$$

with

$$S_c = \sum_{a\in\mathbb{F}_q^*}\Big(\frac{1-ca^{-1}}{q}\Big).$$

If $c = 0$, then

$$S_0 = \sum_{a\in\mathbb{F}_q^*}\Big(\frac{1}{q}\Big) = q-1.$$

Otherwise, $d = 1-ca^{-1}$ runs over $\mathbb{F}_q - \{1\}$, and we have

$$S_c = \sum_{a\in\mathbb{F}_q^*}\Big(\frac{1-ca^{-1}}{q}\Big) = \sum_{d\in\mathbb{F}_q}\Big(\frac{d}{q}\Big) - \Big(\frac{1}{q}\Big) = \sum_{d\in\mathbb{F}_q^*}\Big(\frac{d}{q}\Big) - \Big(\frac{1}{q}\Big) = -\Big(\frac{1}{q}\Big) = -1,$$

since in $\mathbb{F}_q^*$ there are as many squares as nonsquares (see Proposition 6.1). As a consequence.

$$\sum_{c\in\mathbb{F}_q} S_c w^c = S_0 + \sum_{c\in\mathbb{F}_q^*} S_c w^c = q - 1 - \sum_{c\in\mathbb{F}_q^*} w^c = q,$$

since $\sum_{c\in\mathbb{F}_q^*} w^c = -1$ (because w is a primitive qth root of unity, so $w^q = 1$,

$$0 = w^q - 1 = (w-1)(w^{q-1} + \cdots + w + 1),$$

which implies $w^{q-1} + \cdots + w + 1 = 0$, and thus, $\sum_{c\in\mathbb{F}_q^*} w^c = w^{q-1} + \cdots + w = -1$), which proves Step 1. □

Proof of Step 2. Since Ω is a field of characteristic p, we have

$$(x_1 + x_2)^p = x_1^p + x_2^p.$$

Then if $m \geq 3$ we have

$$(x_1 + \cdots + x_m)^p = (x_1 + \cdots + x_{m-1})^p + x_m^p,$$

and by induction

$$(x_1 + \cdots + x_m)^p = x_1^p + \cdots + x_m^p,$$

for all $x_1, \ldots, x_m \in \Omega$. Consequently, since p is odd $\left(\frac{a}{q}\right)^p = \left(\frac{a}{q}\right)$, we get

$$y^p = \sum_{a \in \mathbb{F}_q} \left(\frac{a}{q}\right)^p w^{ap} = \sum_{a \in \mathbb{F}_q} \left(\frac{a}{q}\right) w^{ap},$$

and by making the change of variable $b = ap$, we get

$$y^p = \sum_{b \in \mathbb{F}_q} \left(\frac{bp^{-1}}{q}\right) w^b = \sum_{b \in \mathbb{F}_q} \left(\frac{p}{q}\right)^2 \left(\frac{bp^{-1}}{q}\right) w^b = \sum_{b \in \mathbb{F}_q} \left(\frac{bp}{q}\right) w^b = \left(\frac{p}{q}\right) \sum_{b \in \mathbb{F}_q} \left(\frac{b}{q}\right) w^b = \left(\frac{p}{q}\right) y.$$

Therefore, $y^{p-1} = \left(\frac{p}{q}\right)$, as claimed. □

The proof of the quadratic reciprocity law for the Jacobi symbol is now easy to obtain. For the reader's convenience, we repeat the statement of the theorem.

Theorem 6.15. *(Quadratic reciprocity law for the Jacobi symbol) If m and n are any odd integers $m, n \geq 3$ such that $\gcd(m, n) = 1$, then*

$$\left(\frac{m}{n}\right) = (-1)^{\frac{m-1}{2}\frac{n-1}{2}} \left(\frac{n}{m}\right).$$

Equivalently,

$$\left(\frac{m}{n}\right) = \begin{cases} \left(\frac{n}{m}\right) & \text{if } m \equiv 1 \bmod 4 \text{ or } n \equiv 1 \bmod 4 \\ -\left(\frac{n}{m}\right) & \text{if } m \equiv 3 \bmod 4 \text{ and } n \equiv 3 \bmod 4. \end{cases}$$

Proof. Write the prime factorizations of m and n as $m = p_1 \cdots p_s$ and $n = q_1 \cdots q_t$, where the p_i and q_j are primes. Then, we have

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = \prod_{i=1}^{s} \prod_{j=1}^{t} \left(\frac{p_i}{q_j}\right)\left(\frac{q_j}{p_i}\right) = (-1)^r,$$

for some integer r. Applying the quadratic reciprocity law for primes to each factor, we find that

$$r = \sum_{i=1}^{s} \sum_{j=1}^{t} \frac{(p_i - 1)}{2} \frac{(q_j - 1)}{2} = \sum_{i=1}^{s} \frac{p_i - 1}{2} \sum_{j=1}^{t} \frac{q_j - 1}{2}.$$

During the proof of part (7) of Proposition 6.9, we showed that

$$\sum_{i=1}^{s} \frac{p_i - 1}{2} \equiv \frac{m-1}{2} \pmod 2$$
$$\sum_{j=1}^{t} \frac{q_j - 1}{2} \equiv \frac{n-1}{2} \pmod 2.$$

Therefore,

$$r \equiv \frac{(m-1)}{2}\frac{(n-1)}{2} \pmod 2,$$

which proves our formula. □

Another way of proving the law of quadratic reciprocity (for primes) is to use Gauss sets. Given any odd prime p, any subset S of $\mathbb{F}_p^* = \{1, 2, \ldots, p-1\}$ such that $S \cap (-S) = \emptyset$ and

$$\mathbb{F}_p^* = S \cup -S$$

is called a *Gauss set*. Observe that since $|S| = |-S|$, the fact that $\mathbb{F}_p^*$ is the disjoint union of S and $-S$ implies that $|S| = (p-1)/2$. In particular,

$$S = \left\{1, 2, \ldots, \frac{p-1}{2}\right\}$$

is a Gauss set. Then, for any $s \in S$ and any $a \in \mathbb{F}_p^*$, we can write

$$as = e_s(a)s_a,$$

for some $s_a \in S$ and with $e_s(a) = \pm 1$. (Of course, as is multiplication in $\mathbb{F}_p$, so $as \equiv e_s(a)s_a \pmod p$.)

Lemma 6.16. *(Gauss's lemma) For any odd prime p and any $a \in \mathbb{F}_p^*$, we have*

$$\left(\frac{a}{p}\right) = \prod_{s \in S} e_s(a).$$

Proof. First, observe that if $s \neq s'$, then $s_a \neq s'_a$, because if $s_a = s'_a$, since $as = e_s(a)s_a$, $as' = e_{s'}(a)s'_a = e_{s'}(a)s_a$, and $e_s(a), e_{s'}(a) \in \{-1, 1\}$, we would have $as = as'$, thus $s = \pm s'$, contradicting the fact that S and $-S$ disjoint. Therefore, the map $s \mapsto s_a$ is a bijection of S. If we multiply the equations

$$as = e_s(a)s_a$$

for all $s \in S$, we get

$$a^{(p-1)/2} \prod_{s \in S} s = \left(\prod_{s \in S} e_s(a)\right) \prod_{s \in S} s_a = \left(\prod_{s \in S} e_s(a)\right) \prod_{s \in S} s,$$

which implies that

$$a^{(p-1)/2} = \prod_{s \in S} e_s(a).$$

However, we know that

$$\left(\frac{a}{p}\right) = a^{(p-1)/2},$$

which proves the lemma. □

As an application of Lemma 6.16, the reader should reprove that

$$\left(\frac{2}{p}\right) = (-1)^{(p-1)/2},$$

by setting $a = 2$ and using the set S from above. It turns out that

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

where $n(p)$ is the number of integers s such that

$$\frac{p-1}{4} < s \leq \frac{p-1}{2}.$$

We can use the Gauss set

$$S = \left\{1, 2, \ldots, \frac{p-1}{2}\right\}$$

and Lemma 6.16 to compute the Legendre symbol $\left(\frac{p}{q}\right)$. Given any $a \in \{1, \ldots, p-1\}$, consider the residues mod p of the multiples of a,

$$\left\{a, 2a, \ldots, \frac{p-1}{2}a\right\},$$

say $r_1, \ldots, r_{(p-1)/2}$.

If $r_s \in S$, then $as = sa \equiv r_s \pmod{p}$ so $e_s(a) = +1$, and if $r_s \notin S$, then $p - r_s \in S$, and $as = sa \equiv r_s \equiv -(p - r_s) \pmod{p}$, we have $e_s(a) = -1$. It follows that $\prod_{s \in S} e_a(s) = (-1)^{\mu}$, where μ is the number of residues r_s not in S.

Therefore, we proved that

$$\left(\frac{a}{p}\right) = (-1)^{\mu},$$

where μ is the number of residues $r_s \bmod p$ of the multiples of a,

$$\left\{a, 2a, \ldots, \frac{p-1}{2}a\right\},$$

such that r_s is not in S. This is how Lemma 6.16 is stated in Apostol [1] (Chapter 9, Theorem 9.6) and Niven, Zuckerman, and Montgomery [16] (Section 3.1, Theorem 3.2).

There is a formula that gives the parity of m in terms of a and p; see Apostol [1] (Chapter 9, Theorem 9.7) and Niven, Zuckerman, and Montgomery [16] (Section 3.1, Theorem 3.3). Using Lemma 6.16 and this formula, a short proof of the law of quadratic reciprocity can be obtained; see Apostol [1] (Chapter 9, Theorem 9.8), and Niven, Zuckerman, and Montgomery [16] (Section 3.2, Theorem 3.4). This is essentially Gauss' third proof; see Dirichlet–Dedekind [12] (Chapter 3, Sections 42 and 43). In some sense, this proof is more elementary than the one we gave. Generalizations of the law of quadratic reciprocity are discussed in Ireland and Rosen [9]. The historical notes found at the end of every chapter of Ireland and Rosen [9] are also very informative.

6.8 Eisenstein's Proof of the Quadratic Reciprocity Law

Another strikingly short proof due to Eisenstein (1845) using a trigonometric identity is given in Serre [20] (Chapter 1, Appendix). This proof is just too beautiful to be left aside, so here is Serre's proof of the quadratic reciprocity law after Eisenstein (1845).

Eisenstein's proof. The idea is to use Gauss's lemma, which says that

$$\left(\frac{a}{p}\right) = \prod_{s \in S} e_s(a),$$

where $e_a(s) = \pm 1$ is defined so that $as = e_s(a)s_a$, with s and s_a in the Gauss set

$$S = \left\{1, 2, \ldots, \frac{p-1}{2}\right\}.$$

The key ingredient (stroke of genius) of the proof is that if $as = e_s(a)s_a$, then

$$\sin\left(\frac{2\pi}{p}\, as\right) = e_s(a) \sin\left(\frac{2\pi}{p}\, s_a\right). \tag{$*_{\sin}$}$$

We simply used the fact that the sine function has the property that $\sin(-x) = -\sin x$. But ($*_{\sin}$) yields

$$e_a(a) = \frac{\sin\left(\frac{2\pi as}{p}\right)}{\sin\left(\frac{2\pi s_a}{p}\right)},$$

so we obtain

$$\left(\frac{a}{p}\right) = \prod_{s \in S} e_s(a) = \prod_{s \in S} \frac{\sin\left(\frac{2\pi as}{p}\right)}{\sin\left(\frac{2\pi s_a}{p}\right)} = \prod_{s \in S} \frac{\sin\left(\frac{2\pi as}{p}\right)}{\sin\left(\frac{2\pi s}{p}\right)},$$

since the map $s \mapsto s_a$ is bijective. Now we need to evaluate the expression involving the ratios of sines. This can be done using the following proposition:

Proposition 6.17. *For any positive odd integer $m \geq 3$, we have*

$$\frac{\sin mx}{\sin x} = (-4)^{(m-1)/2} \prod_{1 \leq j \leq \frac{m-1}{2}} \left(\sin^2 x - \sin^2 \left(\frac{2\pi j}{m} \right) \right).$$

Proof. We prove by induction that for any positive odd integer $m \geq 3$, $\cos mx/\cos x$ and $\sin mx/\sin x$ are polynomials of degree $(m-1)/2$ in $\sin^2 x$ whose leading coefficient is $(-4)^{(m-1)/2}$. For $m = 3$, we have the well known formulae

$$\begin{aligned} \cos 3x &= 4\cos^3 x - 3\cos x \\ \sin 3x &= -4\sin^3 x + 3\sin x, \end{aligned}$$

and we get

$$\begin{aligned} \frac{\cos 3x}{\cos x} &= 4\cos^2 x - 3 \\ &= 4(1 - \sin^2 x) - 3 \\ &= -4\sin^3 x + 1 \\ \frac{\sin 3x}{\sin x} &= -4\sin^2 x + 3, \end{aligned}$$

so the base case holds. For the induction step, we use the formulae

$$\begin{aligned} \cos(m+2)x &= \cos mx \cos 2x - \sin mx \sin 2x \\ \sin(m+2)x &= \sin mx \cos 2x + \cos mx \sin 2x, \end{aligned}$$

and

$$\begin{aligned} \cos 2x &= 1 - 2\sin^2 x \\ \sin 2x &= 2\sin x \cos x. \end{aligned}$$

Then we get

$$\begin{aligned} \frac{\cos(m+2)x}{\cos x} &= \frac{\cos mx}{\cos x} \cos 2x - \sin mx \frac{\sin 2x}{\cos x} \\ &= \frac{\cos mx}{\cos x}(1 - 2\sin^2 x) - \sin mx \frac{2\sin x \cos x}{\cos x} \\ &= \frac{\cos mx}{\cos x}(1 - 2\sin^2 x) - 2\frac{\sin mx}{\sin x} \sin^2 x. \end{aligned}$$

By the induction hypothesis, both $\cos mx/\cos x$ and $\sin mx/\sin x$ are polynomials of degree $(m-1)/2$ in $\sin^2 x$, so $\cos(m+2)x/\cos x$ is a polynomial of degree $(m+1)/2$ in $\sin^2 x$.

The leading terms of the polynomials in $\sin^2 x$ for $\cos mx/\cos x$ and $\sin mx/\sin x$ are both $(-4)^{(m-1)/2}$, so the leading term of $\cos(m+2)x/\cos x$ in $\sin^2 x$ is

$$-2(-4)^{(m-1)/2} - 2(-4)^{(m-1)/2} = (-4)^{(m+1)/2},$$

establishing the induction step. We also get

$$\begin{aligned}\frac{\sin(m+2)x}{\sin x} &= \frac{\sin mx}{\sin x}\cos 2x + \cos mx\frac{\sin 2x}{\sin x}\\ &= \frac{\sin mx}{\sin x}(1-2\sin^2 x) + \cos mx\frac{2\sin x\cos x}{\sin x}\\ &= \frac{\sin mx}{\sin x}(1-2\sin^2 x) + 2\frac{\cos mx}{\cos x}\cos^2 x\\ &= \frac{\sin mx}{\sin x}(1-2\sin^2 x) + 2\frac{\cos mx}{\cos x}(1-\sin^2 x).\end{aligned}$$

By the induction hypothesis, both $\cos mx/\cos x$ and $\sin mx/\sin x$ are polynomials of degree $(m-1)/2$ in $\sin^2 x$, so $\sin(m+2)x/\sin x$ is a polynomial of degree $(m+1)/2$ in $\sin^2 x$. The leading terms of the polynomials in $\sin^2 x$ for $\cos mx/\cos x$ and $\sin mx/\sin x$ are both $(-4)^{(m-1)/2}$, so the leading term of $\sin(m+2)x/\sin x$ in $\sin^2 x$ is

$$-2(-4)^{(m-1)/2} - 2(-4)^{(m-1)/2} = (-4)^{(m+1)/2},$$

establishing the induction step.

Finally, observe that $\sin mx/\sin x$ vanishes for $x = 2\pi j/m$ with $j = 1, \ldots, (m-1)/2$, so the polynomial in $\sin^2 x$ expressing $\sin mx/\sin x$ vanishes for $\sin^2 x = \sin^2\left(\frac{2\pi j}{m}\right)$, and since this polynomial has degree $(m-1)/2$, it is the product of the factors

$$\sin^2 x - \sin^2\left(\frac{2\pi j}{m}\right),$$

up to its leading coefficient. But we just showed that this leading coefficient is $(-4)^{(m-1)/2}$, which proves our formula. □

Let T be the set

$$T = \left\{1, 2, \ldots, \frac{a-1}{2}\right\}.$$

Using Proposition 6.17 with $m = a$ and $x = (2\pi s)/p$, we obtain

$$\begin{aligned}\left(\frac{a}{p}\right) = \prod_{s\in S} e_s(a) &= \prod_{s\in S}(-4)^{(a-1)/2}\prod_{t\in T}\left(\sin^2\left(\frac{2\pi s}{p}\right) - \sin^2\left(\frac{2\pi t}{a}\right)\right)\\ &= (-4)^{(a-1)(p-1)/4}\prod_{s\in S}\prod_{t\in T}\left(\sin^2\left(\frac{2\pi s}{p}\right) - \sin^2\left(\frac{2\pi t}{a}\right)\right),\end{aligned}$$

since S has $(p-1)/2$ elements. Thus we have

$$\left(\frac{a}{p}\right) = (-4)^{(a-1)(p-1)/4} \prod_{s\in S}\prod_{t\in T}\left(\sin^2\left(\frac{2\pi s}{p}\right) - \sin^2\left(\frac{2\pi t}{a}\right)\right). \qquad (*_1)$$

Exchanging the roles of a and p, we obtain

$$\left(\frac{p}{a}\right) = (-4)^{(a-1)(p-1)/4} \prod_{s\in S}\prod_{t\in T}\left(\sin^2\left(\frac{2\pi t}{a}\right) - \sin^2\left(\frac{2\pi s}{p}\right)\right). \qquad (*_2)$$

Comparing $(*_1)$ and $(*_2)$, we see that the factors are identical except for their sign. Since there are $(a-1)(p-1)/4$ factors, we deduce that

$$\left(\frac{p}{a}\right) = (-1)^{(a-1)(p-1)/4}\left(\frac{p}{a}\right),$$

which is indeed the law of quadratic reciprocity. □

6.9 Strong Pseudoprimes are Euler Pseudoprimes

We conclude this chapter by showing that every strong pseudoprime base a is also an Euler pseudoprime base a. This is another indication that the Miller–Rabin test is somewhat better than the Solovay–Strassen test (recall that the proportion of MR-liars is at most $1/4$, whereas the proportion of E-liars is at most $1/2$). We follow Koblitz's proof [10] (Chapter V).

We begin with an easy result, but first, observe that if a is an E-liar, then

$$\left(\frac{a}{n}\right) a^{(n-1)/2} \equiv 1 \pmod{n},$$

which implies that $\gcd(a, n) = 1$, and since $\left(\frac{a}{n}\right) \in \{-1, 1\}$, the above condition is equivalent to

$$\left(\frac{a}{n}\right) \equiv a^{(n-1)/2} \pmod{n}.$$

Proposition 6.18. *If n is any composite of the form $n = 4k + 3$, then n is a strong pseudoprime base a iff n is an Euler pseudoprime base a.*

Proof. Since $n = 4k + 3$, we have $n - 1 = 2(2k + 1)$, so $n - 1 = 2^s t$ with $s = 1$ and $t = (n-1)/2$. Thus, n is a strong pseudoprime base a iff $a^{(n-1)/2} \equiv \pm 1 \pmod{n}$. If n is an Euler pseudoprime, then the above congruence holds, by definition.

Conversely, assume that $a^{(n-1)/2} \equiv \pm 1 \pmod{n}$. Then, since $(n-3)/4 = k$ is an integer, we have

$$\left(\frac{a}{n}\right) = \left(\frac{a^{(n-3)/4}}{n}\right)^2 \left(\frac{a}{n}\right) = \left(\frac{a^{(n-3)/2}}{n}\right)\left(\frac{a}{n}\right) = \left(\frac{a^{(n-3)/2}\cdot a}{n}\right) = \left(\frac{a^{(n-1)/2}}{n}\right),$$

and because $(n-1)/2 = 2k+1$, we have

$$\left(\frac{-1}{n}\right) = -1,$$

which implies that

$$\left(\frac{a}{n}\right) \equiv a^{(n-1)/2} \pmod{n},$$

as desired. □

The case were $n = 4k+1$ is more involved.

Theorem 6.19. *If n is a strong pseudoprime base a, then n is an Euler pseudoprime base a.*

Proof. Write $n-1 = 2^s t$ with t odd and assume that a in an MR-liar for n, which means that either

(a) $a^t \equiv 1 \pmod{n}$, or

(b) $a^{2^i t} \equiv n-1 \pmod{n}$, for some i with $0 \leq i \leq s-1$.

We consider several cases.

Case 1. Assume that $a^t \equiv 1 \pmod{n}$. In this case, since $(n-1)/2 = 2^{s-1}t$, we have

$$a^{(n-1)/2} \equiv 1 \pmod{n}.$$

We also have

$$1 = \left(\frac{1}{n}\right) = \left(\frac{a^t}{n}\right) = \left(\frac{a}{n}\right)^t,$$

and since t is odd me must have $\left(\frac{a}{n}\right) = 1$, so $\left(\frac{a}{n}\right) \equiv a^{(n-1)/2} \pmod{n}$.

Case 2. Assume that (b) holds for $i = s-1$; that is, $a^{(n-1)/2} \equiv -1 \pmod{n}$. We must show that $\left(\frac{a}{n}\right) = -1$.

Let p be any prime divisor of n and write $p-1 = 2^{s'}t'$, with t' odd. We make the following claim:

Claim. We have $s' \geq s$ and

$$\left(\frac{a}{p}\right) = \begin{cases} -1 & \text{if } s' = s \\ 1 & \text{if } s' > s. \end{cases}$$

Proof of the claim. From

$$a^{(n-1)/2} = a^{2^{s-1}t} \equiv -1 \pmod{n},$$

raising both sides to the power t' we we obtain

$$\left(a^{2^{s-1}t}\right)^{t'} \equiv \left(a^{2^{s-1}t'}\right)^{t} \equiv -1 \pmod{n},$$

and since p divides n, we also have

$$\left(a^{2^{s-1}t'}\right)^{t} \equiv -1 \pmod{p}.$$

If we had $s' < s$, then we would not have

$$a^{p-1} \equiv a^{2^{s'}t'} \equiv 1 \pmod{p},$$

contradicting Fermat's little theorem. Thus, $s' \geq s$. If $s' = s$, then

$$\left(a^{2^{s-1}t'}\right)^{t} \equiv -1 \pmod{p},$$

and since t is odd this implies that

$$\left(\frac{a}{p}\right) \equiv \left(\frac{a}{p}\right)^{t} \equiv (a^{(p-1)/2})^{t} \equiv \left(a^{2^{s'-1}t'}\right)^{t} \equiv -1 \pmod{p}.$$

On the other hand, if $s' > s$, then the congruence

$$\left(a^{2^{s-1}t'}\right)^{t} \equiv -1 \pmod{p}$$

raised to the power $2^{s'-s}$ implies that $\left(a^{2^{s'-1}t'}\right)^{t} \equiv \left(a^{(p-1)/2}\right)^{t} \equiv 1 \pmod{p}$, and since t is odd, $\left(\frac{a}{p}\right)^{t} = \left(\frac{a}{p}\right) = 1$. □

Write n as a product of primes (not necessarily distinct), $n = p_1 p_2 \cdots p_m$, and let k be the number of primes p such that $s' = s$ when we write $p - 1 = 2^{s'}t'$ with t' odd, counting such a prime with its multiplicity. By the claim, $s' \geq s$ and

$$\left(\frac{a}{n}\right) = \prod_{i=1}^{m}\left(\frac{a}{p_i}\right) = (-1)^k.$$

On the other hand, working modulo 2^{s+1}, we see that $p \equiv 1 \pmod{2^{s+1}}$ unless p is one of the k primes for which $s' = s$, in which case $p \equiv 1 + 2^s \pmod{2^{s+1}}$. Since $n = 1 + 2^s t \equiv 1 + 2^s \pmod{2^{s+1}}$, we have

$$1 + 2^s \equiv p_1 \cdots p_m \equiv (1 + 2^s)^k \equiv 1 + k2^s \pmod{2^{s+1}},$$

using the binomial formula in the last step. The congruence

$$2^s \equiv k2^s \pmod{2^{s+1}}$$

implies that k is odd, hence

$$\left(\frac{a}{n}\right) = (-1)^k = -1,$$

as was to be proved.

Case 3. Assume that (b) holds for $i < s-1$; that is, $a^{2^i t} \equiv -1 \pmod{n}$. Raising this congruence to the power 2^{s-1-i}, we get $a^{(n-1)/2} \equiv 1 \pmod{n}$, so we have to prove that $\left(\frac{a}{n}\right) = 1$. As in Case 2, write $p-1 = s'^{s'}t'$ with t' odd for every prime factor p of n.

Claim. We have $s' \geq i+1$ and

$$\left(\frac{a}{p}\right) = \begin{cases} -1 & \text{if } s' = i+1 \\ 1 & \text{if } s' > i+1. \end{cases}$$

The proof of the above claim is identical to the proof of the claim in Case (2). Similarly to Case (2), let k be the number of primes (not necessarily distinct) such that $s' = i+1$. As in Case (2), we have

$$\left(\frac{a}{n}\right) = (-1)^k.$$

On the other hand, since $i < s-1$, we have $n = 1 + 2^s t \equiv 1 \pmod{2^{i+2}}$, and also

$$n \equiv p_1 \cdots p_m \equiv (1+2^{i+1})^k \equiv 1 + k2^{i+1} \pmod{2^{i+2}},$$

which implies

$$2^s t \equiv k2^{i+1} \pmod{2^{i+2}}.$$

Since $i+2 \leq s$, the number k must be even, and

$$\left(\frac{a}{n}\right) = (-1)^k = 1,$$

as desired. □

There are examples of composite numbers n such that n is an Euler pseudoprime base a but n is not a strong pseudoprime base a. This behavior is observed for numbers of the form $(6m+1)(12m+1)(18m+1)$, where each factor is prime and m is odd; see Exercise 17 in Section 1 of Chapter V of Koblitz [10].

Acknowlegments: I wish to thank Dan Guralnik for inspiring me to write up the review sections on groups. I learned about Theorem 4.22 from his wonderful's lectures in ESE 680-001. Not too surprisingly, I found that this theorem is used by J.P. Serre in his outstanding *Lectures in Arithmetic* [20]. I also thank Peter Freyd, Ron Donagi and Steve Shatz. Peter made a number of suggestions/corrections. In particular, he brought to my attention the facts about square roots of unity stated as Proposition 5.1 and Theorem 5.2. Ron and Steve pointed out that Theorem 4.44 implies that there are four square roots of unity when $n = 2^m$ with $m \geq 3$.

Bibliography

[1] Tom M. Apostol. *Introduction to Analytic Number Theory*. Undergraduate Texts in Mathematics. Springer, first edition, 1976.

[2] Nicolas Bourbaki. *Algèbre, Chapitres 4-7*. Eléments de Mathématiques. Masson, 1981.

[3] Richard Crandall and Carl Pomerance. *Prime Numbers. A Computational Perspective*. Springer, second edition, 2005.

[4] Martin Dietzfelbinger. *Primality Testing in Polynomial Time. From Randomized Algorithms to "Primes is in P"*. LNCS 3000. Springer, first edition, 2004.

[5] Harold M. Edwards. *Riemann's Zeta Function*. Dover, first edition, 2001.

[6] Jean H. Gallier. *Discrete Mathematics*. Universitext. Springer Verlag, first edition, 2011.

[7] Carl Friedrich Gauss. *Recherches Arithmétiques*. Edition Jacques Gabay, first edition, 1807. French Translation of the *Disquisitiones Arithmeticae*.

[8] Jeffrey H. Hoffstein, Jill Pipher, and Joseph H. Silverman. *An Introduction to Mathematical Cryptography*. Undergraduate Texts in Mathematics. Springer, first edition, 2008.

[9] Kenneth Ireland and Michael Rosen. *A Classical Introduction to Modern Number Theory*. GTM No. 84. Springer Verlag, second edition, 1998.

[10] Neal Koblitz. *A Course in Number Theory and Cryptography*. GTM No. 114. Springer Verlag, second edition, 1994.

[11] Serge Lang. *Algebra*. Addison Wesley, third edition, 1993.

[12] Peter Gustav Lejeune-Dirichlet. *Lectures on Number Theory*, volume 16 of *History of Mathematics*. AMS, first edition, 1999. Translation by John Stillwell of *Vorlesungen über Zahlentheorie* with supplements by Richard Dedekind, 1863.

[13] L. Lovász, J. Pelikán, and K. Vesztergombi. *Discrete Mathematics. Elementary and Beyond*. Undergraduate Texts in Mathematics. Springer, first edition, 2003.

[14] Alfred J. Menezes, Paul C. van Oorschot, and Scott A. Vanstone. *Handbook of Applied Cryptography*. CRC Press, fifth edition, 2001.

[15] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorihms*. Cambridge University Press, first edition, 1995.

[16] Ivan Niven, Herbert S. Zuckerman, and Hugh L. Montgomery. *An Introduction to the Theory of Numbers*. Wiley, fifth edition, 1991.

[17] Vaughan R. Pratt. Every prime has a succinct certificate. *SIAM Journal on Computing*, 4(3):214–220, 1975.

[18] Paulo Ribenboim. *The Little Book of Bigger Primes*. Springer-Verlag, second edition, 2004.

[19] R.L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978.

[20] Jean-Pierre Serre. *A Course in Arithmetic*. Graduate Text in Mathematics, No. 7. Springer, first edition, 1973.

[21] Victor Shoup. *A Computational Introduction to Number Theory and Algebra*. Cambridge University Press, second edition, 2009.

[22] Joseph H. Silverman. *A Friendly Introduction to Number Theory*. Prentice Hall, third edition, 2006.