

COMPUTATIONAL NUMBER THEORY

AMIN WITNO

Preface

Written at Philadelphia University, Jordan for Math 472, these notes¹ were used first time in the Fall 2006 semester. They have since been revised² and shall be revised again as often as the author teaches the course. Outline notes are more like a revision. No student is expected to fully benefit from these notes unless they have regularly attended the lectures.

1 The RSA Cryptosystem

Sensitive messages, when transferred over the internet, need to be encrypted, that is changed into a secret code in such a way that only the intended receiver who has the secret key is able to decrypt it. It is common that alphabetical characters are converted to their numerical ASCII equivalents before they are encrypted, hence the coded message will look like integer strings. The RSA Algorithm is an encryption-decryption process which is widely employed today. In practice the encryption key may be made public and doing so will not risk the security of the system. This feature is a characteristic of the so-called public-key cryptosystem.

Alia selects two distinct primes p and q which are very large, over a hundred digits each. She computes $n = pq$, $\phi(n) = (p - 1)(q - 1)$, and determines a rather small number e which will serve as her encryption key, making sure that e has no common factor with $\phi(n)$. She then chooses another integer $d < n$ satisfying $de \equiv 1 \pmod{\phi(n)}$, as her decryption key. When all is ready Alia gives to Bob the pair (n, e) and keeps the rest secret. Now whenever Bob wants to send a message (integer) $m < n$ to Alia, he encrypts it to $s = m^e \pmod{n}$. Upon receiving s , Alia decrypts it back to $s^d \pmod{n} = m$, the intended message. But why does this work, and how safe is it?

1.1 Basics of Number Theory

The domain of number theory is the integers. The basic concepts of elementary number theory are those of divisibility and prime numbers. Upon these is built a systematic study of the theory of congruences, otherwise known as modular arithmetic.

¹Copyrighted under a Creative Commons License

²Last Revision: 17-02-2009

Definition. The number d divides m , or m is divisible by d , if the operation $m \div d$ yields an integer. This relation may be written $d \mid m$, or $d \nmid m$ if it is not true. For examples, $3 \mid 18$ and $5 \nmid 18$. When $d \mid m$, the number d is called a *divisor* or *factor* of m , and m a *multiple* of d .

Proposition 1.1. If $d \mid m$ and $d \mid n$ then $d \mid (am + bn)$ for any integers a and b .

Exercise 1.1. Prove this proposition.

Definition. A *prime* number is an integer $p > 1$ with no divisors between 1 and p . An integer $n > 1$ which is not a prime number is called *composite*. For examples, 13 and 17 are primes, but 21 is composite because $21 = 3 \times 7$.

Prime numbers are infinitely many, the earliest proof of which was given by Euclid, about 1500 BC. An important property of a prime p is this: If $p \mid mn$ then either $p \mid m$ or $p \mid n$. We will not prove this now, but we assume a much stronger and very fundamental theorem.

Theorem 1.2 (The Fundamental Theorem of Arithmetic). Every integer $n \geq 2$ is the product of a unique collection of prime numbers, counting repeated factors.

For example in factoring the number 700 into primes, one may find $700 = 2 \times 2 \times 5 \times 5 \times 7$ and another $700 = 7 \times 2 \times 5 \times 2 \times 5$, but it would not be possible to find another prime factor outside the multiset $\{2, 2, 5, 5, 7\}$.

Definition. Let m and n be factored into powers of distinct primes: $m = \prod p_i^{j_i}$ and $n = \prod p_i^{k_i}$, with $j_i, k_i \geq 0$. Define the *greatest common divisor* of m and n to be $\gcd(m, n) = \prod p_i^{e_i}$ where $e_i = \min\{j_i, k_i\}$. For example with $m = 2^2 \times 3 \times 5^3 \times 7^3$ and $n = 2^5 \times 3^2 \times 5 \times 7^2$ we have $\gcd(m, n) = 2^2 \times 3 \times 5 \times 7^2$. Note that $\gcd(m, n)$ is the largest divisor common to m and n .

Exercise 1.2. Find $\gcd(400, 720)$, $\gcd(19392, 29391)$, $\gcd(2^3 \times 3^8 \times 5^4 \times 7^5, 3^7 \times 5^2 \times 7^2)$.

Lemma 1.3 (Euclid's Lemma). If $d \mid mn$ and $\gcd(d, m) = 1$ then $d \mid n$.

Note that the condition $\gcd(d, m) = 1$ in the lemma cannot be omitted. For example, $6 \mid 24 = 3 \times 8$ but neither $6 \mid 3$ nor $6 \mid 8$ is true. When $\gcd(m, n) = 1$ we say that m and n are *relatively prime*, which means that they share no common factor other than 1.

Exercise 1.3. Assuming FTA, prove Euclid's lemma.

Definition. For every real number x , the notation $\lfloor x \rfloor$ denotes the greatest integer $\leq x$. For example $\lfloor 3.14 \rfloor = 3$ and $\lfloor 2 \rfloor = 2$. Then with $n > 0$, define the *residue of m mod n* by $m \% n = m - \lfloor m/n \rfloor n$. Here the symbol (%) is read "mod".

For example $18 \% 5 = 3$ and $18 \% 3 = 0$. Note that $m \% n$ is really the remainder upon dividing m by n and it lies in the range $0 \leq m \% n \leq n - 1$. In particular $m \% n = 0$ if and only if $n \mid m$.

Theorem 1.4. For integers m and $n \neq 0$, we have $\gcd(m, n) = \gcd(n, m \% n)$.

Example (The Euclidean Algorithm). This theorem gives us a very fast way to evaluate gcd. For example, to evaluate $\gcd(216, 78)$ we repeatedly apply Theorem 1.1 to obtain

$\gcd(216, 78) = \gcd(78, 60) = \gcd(60, 18) = \gcd(18, 6) = \gcd(6, 0) = 6$. In each step we could have computed $m \% n$ the old fashioned way as follows.

$$\begin{aligned} 216 &= 2(78) + 60 \\ 78 &= 1(60) + 18 \\ 60 &= 3(18) + 6 \\ 18 &= 3(6) + 0 \end{aligned}$$

Exercise 1.4. Repeat the last exercise, this time using the Euclidean algorithm.

Moreover it can be shown that $\gcd(m, n)$ is actually an integral linear combination of m and n , that is $\gcd(m, n) = am + bn$. The algorithm involved in finding these integers a and b is called the *extended Euclidean algorithm* (EEA). To illustrate we continue the previous example by rewriting each equation in order to express each remainder as a linear combination of $m = 216$ and $n = 78$.

$$\begin{aligned} 60 &= 1(216) - 2(78) \\ 18 &= 1(78) - 1(60) \\ &= 1(78) - 1\{1(216) - 2(78)\} \\ &= -1(216) + 3(78) \\ 6 &= 1(60) - 3(18) \\ &= 1\{1(216) - 2(78)\} - 3\{-1(216) + 3(78)\} \\ &= 4(216) - 11(78) \end{aligned}$$

Next simplify the appearance of the above algorithm by not writing the m and n in each row. For convenience we add two extra rows at the top, corresponding to the equations $216 = 1(216) + 0(78)$ and $78 = 0(216) + 1(78)$, in this order.

$$\begin{array}{rcc} 216 & 1 & 0 \\ 78 & 0 & 1 \\ 60 & 1 & -2 \\ 18 & -1 & 3 \\ 6 & 4 & -11 \end{array}$$

Note again that the last row gives the desired result $\gcd(216, 78) = 6 = 4(216) - 11(78)$.

Exercise 1.5. Prove why EEA works and then apply it to the previous exercise.

Definition. Two integers a and b are *congruent modulo* $n > 0$ if $a \% n = b \% n$, in which case we write $a \equiv b \pmod{n}$. For example $19 \equiv 4 \pmod{3}$ since $19 \% 3 = 4 \% 3 = 1$. Equivalently we may define $a \equiv b \pmod{n}$ if and only if $n \mid (a - b)$.

The best way to understand congruence is to view it as an equivalence relation over the integers with n *congruence classes*, namely the classes of integers whose residues mod n are $0, 1, 2, \dots$, up to $n - 1$. Hence $a \equiv b \pmod{n}$ if and only if a, b belong to the same class, which is $a \% n = b \% n$. In particular $a \equiv a \% n \pmod{n}$.

In the language of algebra, these congruence classes are closed under addition and multiplication. What it means is that if $a_1 \equiv b_1 \pmod{n}$ and $a_2 \equiv b_2 \pmod{n}$ then $a_1 + a_2 \equiv b_1 + b_2 \pmod{n}$ and $a_1 a_2 \equiv b_1 b_2 \pmod{n}$. Division, however, is allowed only by a relatively prime factor:

Proposition 1.5. Assume that $\gcd(m, n) = 1$. Then $ma \equiv mb \pmod{n}$ if and only if $a \equiv b \pmod{n}$.

Exercise 1.6. Prove this proposition.

Theorem 1.6 (Chinese Remainder Theorem). Suppose again $\gcd(m, n) = 1$. Then $a \equiv b \pmod{mn}$ if and only if $a \equiv b \pmod{m}$ and $a \equiv b \pmod{n}$.

Exercise 1.7. Prove CRT using FTA.

Definition. Two integers a and b are *inverses* of each other modulo n if $ab \equiv 1 \pmod{n}$. For example 3 and 5 are inverses modulo 7 since $3 \times 5 = 15 \equiv 1 \pmod{7}$. Similarly the congruence $5^2 \equiv 1 \pmod{12}$ implies that 5 is its own inverse modulo 12.

It is not hard to see that the number m has an inverse modulo n if and only if $\gcd(m, n) = 1$, in which case its inverse can be computed using EEA, for if $\gcd(m, n) = 1 = am + bn$, then we have $am \equiv 1 \pmod{n}$.

Exercise 1.8. Find $7^{-1} \pmod{12}$, $35^{-1} \pmod{42}$, $27^{-1} \pmod{209}$.

Theorem 1.7 (Wilson's Theorem). If p is prime then $(p - 1)! \equiv -1 \pmod{p}$.

To prove Wilson's theorem, look at the numbers $1, 2, \dots, p - 1$. We can show why none of these is self-inverse, except 1 and $p - 1$. On the other hand, we can show that all of them have an inverse mod p and that each inverse is in this very list. Hence their product gives $p - 1 \pmod{p}$.

Exercise 1.9. Prove the converse of Wilson's theorem.

An efficient method for computing $a^k \% n$ for large integer k is the so-called *Successive Squaring Algorithm* (SSA), as follows.

- 1) Express k as the sum of powers of 2, say $k = \sum 2^{e_i}$.
- 2) Compute $a^2 \% n, a^4 \% n, a^8 \% n, \dots$ up to the highest exponent in Step 1.
- 3) Evaluate $a^k \% n = \prod a^{2^{e_i}}$.

Example. Let us compute $23^{106} \% 97$. We have $106 = 64 + 32 + 8 + 2 = 2^6 + 2^5 + 2^3 + 2^1$ so that $23^{106} = (23^{64})(23^{32})(23^8)(23^2)$. The successive squaring part goes as follows.

$$\begin{aligned} & (23)^2 \% 97 = 44 \\ 23^4 & \% 97 = (44)^2 \% 97 = 93 \\ 23^8 & \% 97 = (93)^2 \% 97 = 16 \\ 23^{16} & \% 97 = (16)^2 \% 97 = 62 \\ 23^{32} & \% 97 = (62)^2 \% 97 = 61 \\ 23^{64} & \% 97 = (61)^2 \% 97 = 35 \end{aligned}$$

Hence $23^{106} \% 97 = (35)(61)(16)(44) \% 97 = 25$.

Exercise 1.10. Compute $3^{57} \% 20, 47^{250} \% 100, 2^{1434} \% 309$.

Definition. The *Euler phi-function* $\phi(n)$ is the number of positive integers up to n which are relatively prime to n .

For example the positive integers relatively prime to 10 are 1, 3, 7, and 9, so $\phi(10) = 4$. Similarly $\phi(11) = 10$. Note that $\phi(p) = p - 1$ for prime p and $\phi(p^k) = p^k - p^{k-1}$. For n other than a prime power we have the following result to use.

Theorem 1.8. If $\gcd(m, n) = 1$ then $\phi(mn) = \phi(m)\phi(n)$.

Example. Hence if we know how to factor n into primes then we can easily evaluate $\phi(n)$. For example, given that $n = 2^5 \times 5^3 \times 17$, we have $\phi(n) = \phi(2^5)\phi(5^3)\phi(17) = (2^5 - 2^4)(5^3 - 5^2)(17 - 1) = (16)(100)(16) = 25600$.

Exercise 1.11. Evaluate $\phi(1024)$, $\phi(4800)$, $\phi(19392)$.

Theorem 1.9 (Euler's Theorem). If $\gcd(a, n) = 1$ then $a^{\phi(n)} \equiv 1 \pmod{n}$. In particular if p is prime then $a^{p-1} \equiv 1 \pmod{p}$.

For example $7^4 \equiv 1 \pmod{10}$ and $100^{10} \equiv 1 \pmod{11}$. The special case with prime modulus is due to Fermat and goes under its own name, *Fermat's Little Theorem*.

1.2 RSA Algorithm

Back to RSA, we summarize how the algorithm is done.

- 1) Alia selects two distinct primes p and q .
- 2) She determines the following quantities.
 - a) $n = pq$
 - b) $\phi(n) = (p - 1)(q - 1)$
 - c) encryption key e relatively prime to $\phi(n)$
 - d) decryption key d such that $de \% \phi(n) = 1$
- 3) Alia gives to Bob the numbers n and e only.
- 4) To send m , Bob encrypts it to $s = m^e \% n$.
- 5) Receiving s , Alia decrypts it back to $m = s^d \% n$.

So why does this work? First of all determining e and d are not too hard for Alia with the help of EEA. Neither it is hard encrypting $s = m^e \% n$ or decrypting $m = s^d \% n$ with SSA. We simply need to verify that the decryption process will indeed yield the intended message m . Since $de \% \phi(n) = 1$ we have $de = 1 + k\phi(n)$ for some integer k and therefore, by Euler's theorem.

$$s^d \equiv (m^e)^d = m^{de} = m^{1+k\phi(n)} = m(m^{\phi(n)})^k \equiv m \pmod{n}$$

Exercise 1.12. To illustrate, suppose Alia employs $n = 12319 = 97 \times 127$ and $e = 5969$.

- a) What is d ?
- b) If Bob wants to send $m = 8411$, what does he actually send to Alia?
- c) Verify that Alia gets m from (b) correctly.
- d) If Alia receives from Bob the encryption $s = 12160$, what is the real message m ?

Now, about security, what if a bad guy intercepts the secret message s , together with e and n ? Well, he will yet have to find d in order to read the message, and that in turn he also will need the factors p and q in order to compute $\phi(n)$. Woe to him, pq has over 200 digits and factoring a large integer this size will take a lifetime on today's best computer.

1.3 Remarks

- 1) RSA was named after its three inventors, Rivest, Shamir, and Adleman in 1977, a few years after the first concept of a public-key cryptosystem was suggested by Diffie and Hellman. It seems there was evident that such a concept had been secretly used years before by a British cryptographic agency, including a version of RSA written by Clifford Cocks.
- 2) In practice these days the size of p and q should be at least 100 digits each. Such primes are plenty, in fact there are roughly $\pi(10^{100}) - \pi(10^{99}) \approx 3.9 \times 10^{97}$ of them. (It is known that the number of primes up to x is estimated by $\pi(x) \approx x/\log x$. For instance, up to a million there are roughly $10^6/\log 10^6 \approx 72000$ primes.) Neither it is hard to find primes; we will discuss this in Chapter 3.
- 3) The decryption process $s^d \% n = m$ relies on Euler's theorem under the assumption that $\gcd(s, n) = 1$. It is not impossible that s turn out not relatively prime to n although such probability is extremely small. (Why?) Nevertheless it is an easy exercise to show that the claim remains true whether or not $\gcd(s, n) = 1$. (Hint: first show that $s^d \equiv m \pmod{p, q}$ independently and then apply CRT.)
- 4) The RSA works under a crucial assumption that it is hard to evaluate $\phi(n)$ without factoring $n = pq$. The problem of evaluating $\phi(n)$ is actually equivalent to that of factoring n in the sense that solving one solves the other as well: It is clear that knowing p and q yields $\phi(n) = (p-1)(q-1)$. Conversely knowing $\phi(n)$ will lead to the discovery of p and q as the roots of the quadratic polynomial

$$f(x) = x^2 + (\phi(n) - n - 1)x + n = x^2 - (p + q)x + pq = (x - p)(x - q)$$

Exercise 1.13. Suppose $n = 2747$ and we know $\phi(n) = 2640$. Factor n .

- 5) The RSA Laboratories used to offer factoring challenges with prizes ranging from \$10,000 to \$200,000. Here is one of the challenge numbers for \$50,000 called RSA-768, which has 232 decimal digits:

$$\begin{aligned} n = & 12301866845301177551304949583849627207728535695953 \\ & 34792197322452151726400507263657518745202199786469 \\ & 38995647494277406384592519255732630345373154826850 \\ & 79170261221429134616704292143116022212404792747377 \\ & 94080665351419597459856902143413 \end{aligned}$$

- 6) RSA may be too slow in practice when a massive amount of data is involved. In such a case it may be needed anyhow for exchanging a private key in order to establish another, faster type of cryptosystem.

1.4 Attacks

Over the years there have been various attempts to break the RSA cryptosystem. While none of these attacks is a serious blow to the system in general, it is worthwhile to be aware of certain circumstances under which a specific implementation of the RSA becomes vulnerable.

- 1) If p and q are quite close together, say of equal digit lengths, then it is not difficult to factor n using Fermat Factorization (Chapter 2). It is important therefore to select p and q of slightly different sizes.
- 2) If $p - 1$ factors into small primes then it is not hard to factor n by Pollard $p - 1$ method (Chapter 2). This should be avoided in practice by choosing another p if necessary.
- 3) If p/q is close to a rational number a/b where a and b are both small then n can be factored using continued fractions (Chapter 2). This too should be avoided.
- 4) Suppose n have N decimal digits. If the first or the last $N/4$ digits of p are known then it is not too hard to factor n . This applies when parts of p are predictable because it is obtained, for instance, by letting $p = M + k$ for some fixed large odd integer M and letting k runs from 2, 4, 6, ... until we hit a prime.
- 5) If an attacker somehow discovers d it is highly probable that he can factor n . The probabilistic algorithm for it runs as follows.
 - a) Let c be an integer such that $a^c \equiv 1 \pmod{n}$ for all a with $\gcd(a, n) = 1$, for example $c = de - 1$. Randomly select candidates for a , say $a = 2, 3, \dots, 20$.
 - b) Of these selection find one such that $a^{c/2} \not\equiv 1 \pmod{n}$. If none exists, replace c by $c/2$ and repeat this step.
 - c) Compute $g = \gcd(a^{c/2} - 1, n)$. If $g \neq 1$ then it is one of p, q , else try another a by resuming the previous step.

Example. Let $n = 323 = 17 \times 19$ with $e = 95$. A quick check shows that $d = 191$. Now let $c = 191 \times 95 - 1 = 18144$, so $c/2 = 9072$ and compute $2^{9072}, 3^{9072}, \dots, 20^{9072}$, all mod 323. It happens that all these yield 1, hence we start again with a new exponent $9072/2 = 4536$. This time we find $3^{4536} \% 323 = 305 \neq 1$ and evaluate $g = \gcd(304, 323) = 19$, and yes it is one of the factors of n . If by chance we had $g = 1$ we would have to continue searching with $4^{4536}, 5^{4536}, \dots$ and with the exponent $4536/2 = 2268$ in the next cycle, etc.

Exercise 1.14. Suppose $n = 209$, $e = 7$, and we know $d = 103$. Factor n .

- 6) Without knowing d it is possible to retrieve the message m via recursive exponentiation as follows. Let $s_0 = s$ and subsequently let $s_k = s_{k-1}^e \% n$. It can be shown that eventually this will lead to a term $s_K = s$ from which we conclude $s_{K-1} = m$. This algorithm is commonly called the *cycling attack* but fortunately this scheme is generally too slow to be effective and there are simple ways to make the system immune to it.

Example. Let $n = 299 = 13 \times 23$ and $e = 17$. Suppose the encrypted message is $s = 123$ so we start calculating $123^{17} \% 299 = 197$, $197^{17} \% 299 = 6$, $6^{17} \% 299 = 288$, and so on, generating the following sequence.

123, 197, 6, 288, 32, 210, 292, 119, 71, 41, 123

The last term $41^{17} \% 299 = 123 = s$ reveals that $m = 41$.

Exercise 1.15. Suppose $n = 319$, $e = 11$, and $s = 288$. Find m .

- 7) If $q < p < 2q$ and $d < \frac{1}{3}n^{1/4}$ then it is not too hard to find d . So in addition to the requirement that p and q should not be too close to each other, we should see that d is large enough, even though doing this would make the decryption process slower, unfortunately.
- 8) Paul Kocher in 1995 showed that it is possible to discover d by testing the system with a series of decryptions while carefully timing the computation times. The assumption is that we know the kind of hardware being used and there are ways to thwart this kind of attack.

Exercise 1.16. Suppose two companies are using RSA with $n_1 = 8051$ and $n_2 = 11371$ and you know that they share a common prime factor. Find a way to factor n_1, n_2 .

2 Factorization

The most basic factoring algorithm is the *trial division* where we experimentally divide the number n by the primes $p = 2, 3, 5, 7, \dots$ up to \sqrt{n} . If this fails to produce a factor then n is a prime number. This trial and error technique is obviously slow and in most programming application it is usually the first method tried, say up to $p < 10,000$ before it is abandoned in favor of other more advanced factoring techniques.

2.1 Divisibility Tests

In the absence of a calculator the following tests can be performed by hand to find small prime factors of n .

- 1) $2 \mid n$ if and only if the unit digit of n is even.
- 2) $3 \mid n$ if and only if the digit sum of n is a multiple of 3. For example for $n = 200612345$ the digit sum is $2 + 0 + 0 + 6 + 1 + 2 + 3 + 4 + 5 = 23$, not a multiple of 3, hence $3 \nmid n$. If needed we compute the digit sum of the digit sum of n until the sum is conveniently small enough. Similar criterion holds for divisibility by 9.
- 3) $5 \mid n$ if and only if the unit digit of n is either 0 or 5.
- 4) Suppose the number n consists of $3k$ digits. Then $7, 11, 13 \mid n$ if and only if the alternating sum of the k consecutive 3-digit blocks of n is divisible by 7, 11, or 13, respectively. To illustrate this let $n = 007656103$, where the two leading zeros have been added to make the number of digits a multiple of 3. We have $007 - 656 + 103 = -546 = -2 \times 3 \times 7 \times 13$, meaning that $7 \mid n$ and $13 \mid n$, but $11 \nmid n$.
- 5) Similarly $37 \mid n$ if and only if 37 divides the sum of the 3-digit blocks of n . For example with $n = 035487477$ we have $35 + 487 + 477 = 999$. Since $37 \mid 999$ then $37 \mid n$.
- 6) $11 \mid n$ if and only if the alternating sum of its digits is divisible by 11. For example $n = 7656103$ and $7 - 6 + 5 - 6 + 1 - 0 + 3 = 4$, not divisible by 11 and so $11 \nmid n$.

- 7) Given an integer n , remove the unit digit, say u , and denote what remains by t . Then $17 \mid n$ if and only if $17 \mid t - 5u$. For example $n = 23562$, where $u = 2$ and $t = 2356$. We have $t - 5u = 2346$. Repeating, with $n = 2346$, $t - 5u = 204$ and next $t - 5u = 0$. Since $17 \mid 0$ we conclude $17 \mid n$.
- 8) Keeping the same notations, $19 \mid n$ if and only if $19 \mid t + 2u$. With $n = 23562$, $t + 2u = 2360$, next 236, and next 35, which is enough to see that $19 \nmid n$.
- 9) Divisibility by 7 and by 13 can also be tested using a similar technique employing u and t , although in practice this method is inferior to the one we have already seen: $7 \mid n$ if and only if $7 \mid t - 2u$ and $13 \mid n$ if and only if $13 \mid t + 4u$.

Exercise 2.1. Prove the claims (1) to (9) above.

2.2 Fermat Factorization

If $n = x^2 - y^2$ then it factors to $n = (x + y)(x - y)$. This fact is the simple idea behind the method of Fermat Factorization. We seek a factor of n by calculating the numbers $y^2 = x^2 - n$ for each integer $x \geq \sqrt{n}$ until we find a perfect square. For example with $n = 4277$ we first calculate $\sqrt{4277} \approx 65.39$ so we start with $x = 66$.

$$\begin{aligned} 66^2 - 4277 &= 79 \\ 67^2 - 4277 &= 212 \\ 68^2 - 4277 &= 347 \\ 69^2 - 4277 &= 484 = 22^2 \end{aligned}$$

The result is $4277 = 69^2 - 22^2 = (69 + 22)(69 - 22) = 91 \times 47$.

Exercise 2.2. Factor the numbers 5963, 16781, and 70027.

Remark. Fermat Factorization always works when n is odd because if $n = ab$ with both a, b odd then $n = x^2 - y^2$ with $x = (a + b)/2$ and $y = (a - b)/2$. Moreover this shows that we should terminate the process when we reach $x = (n + 1)/2$, in which case $n = n \times 1$ is a prime. However, from \sqrt{n} until we reach $(n + 1)/2$ can be extremely long, unless the factors a, b are close together so that the small $y = (a - b)/2$ can be quickly discovered.

Suppose now that a, b are not that close together, say b is about 5 times as big as a . Then $5a$ and b would be close together, hence Fermat Factorization would be ideal for factoring the number $5n = 5a \times b$. For example, $n = 15963 = 51 \times 313$. We deliberately choose this number with one factor about 5 times the other. Not knowing the factors, applying Fermat Factorization would require 56 iterations. If instead we work with $5n = 79815$, where $\sqrt{79815} \approx 282.51$, then we start with $x = 283$,

$$\begin{aligned} 283^2 - 79815 &= 274 \\ 284^2 - 79815 &= 841 = 29^2 \end{aligned}$$

and at only the second iteration find that $79815 = 284^2 - 29^2 = 255 \times 313$. This then easily leads to the factors of n . In practice, of course, we do not know the values of a, b ahead of time, hence it remains a trial-and-error experiment.

2.3 Pollard Rho Method

Suppose n is a large composite and p is its smallest prime factor. If we randomly choose more than p numbers from 0 to $n - 1$ then by the Pigeonhole Principle two of them must satisfy the relation $x_1 \equiv x_2 \pmod{p}$ while $x_1 \not\equiv x_2 \pmod{n}$. In that case $\gcd(x_1 - x_2, n)$ gives a nontrivial factor of n .

In practice these “random” numbers can be generated by letting $x_0 = 2$ and recursively $x_k = (x_{k-1}^2 + 1) \% n$. It has been experimentally proven that the sequence generated in this way works well enough for our purposes. Note that if $x_i \equiv x_j \pmod{p}$ then $x_{i+1} \equiv x_{j+1} \pmod{p}$ hence the sequence $\{x_k\}$ is periodic, say of length d . Now rather than computing $\gcd(x_i - x_j, n)$ for many pairs x_i, x_j it is more efficient to consider only $i = 2j$ for $x_{2j} \equiv x_j \pmod{p}$ will hold whenever $d \mid j$.

Example. Let $n = 3107$. We start our calculation with $x_0 = 2$ as follows.

| k | $x_k \% n$ | $\gcd(x_{2k} - x_k, n)$ |
|-----|--------------------------|---|
| 1 | $2^2 + 1 = 5$ | |
| 2 | $5^2 + 1 = 26$ | $\gcd(26 - 5, 3107) = 1$ |
| 3 | $26^2 + 1 = 677$ | |
| 4 | $677^2 + 1 \equiv 1601$ | $\gcd(1601 - 26, 3107) = \gcd(1575, 3107) = 1$ |
| 5 | $1601^2 + 1 \equiv 3034$ | |
| 6 | $3034^2 + 1 \equiv 2223$ | $\gcd(2223 - 677, 3107) = \gcd(2156, 3107) = 1$ |
| 7 | $2223^2 + 1 \equiv 1600$ | |
| 8 | $1600^2 + 1 \equiv 2940$ | $\gcd(2940 - 1601, 3107) = \gcd(1339, 3107) = 13$ |

and we find that $3107 = 13 \times 239$.

Exercise 2.3. Apply Pollard rho method to factor the numbers 133, 1703, and 11357. Alternate your effort with $x_0 = 3$ instead of 2 and/or with the recurrence relation $x_k = x_{k-1}^2 - 1$.

Pollard rho method is guaranteed to work as long as we are sure that n is composite. Probabilistically it is an excellent factoring technique when n has a relatively small factor as compared to \sqrt{n} . Pollard rho method has also proved to be practical for factoring medium size integer up to 10^{15} .

2.4 Pollard $p - 1$ Method

Suppose n is a composite and p is a prime factor of n . We let $x_k = 2^{k!} \% n$ for $k = 1, 2, 3, \dots$ and simultaneously compute $\gcd(x_k - 1, n)$ in hope that we find a nontrivial factor of n . By Fermat’s Little Theorem this works when $(p - 1) \mid k!$ because then $2^{k!} = (2^{p-1})^{\frac{k!}{p-1}} \equiv 1 \pmod{p}$, hence $p \mid x_k - 1$ and $p \leq \gcd(x_k - 1, n) < n$. The exception would be if $x_k = 1$ and $\gcd(x_k - 1, n) = n$ which can occur when all the prime factors of n behave like p , i.e. $(p - 1) \mid k!$ but this is highly unlikely.

Example. Computing x_k can be done efficiently through the recurrence relation $x_k = x_{k-1}^2$. For example we choose $n = 57983$ and start with $x_1 = 2$.

| k | $x_k \% n$ | $\gcd(x_k - 1, n)$ |
|-----|------------------------|-----------------------------|
| 1 | 2 | $\gcd(1, 57983) = 1$ |
| 2 | $2^2 = 4$ | $\gcd(3, 57983) = 1$ |
| 3 | $4^3 = 64$ | $\gcd(63, 57983) = 1$ |
| 4 | $64^4 \equiv 20129$ | $\gcd(20128, 57983) = 1$ |
| 5 | $20129^5 \equiv 50290$ | $\gcd(50289, 57983) = 1$ |
| 6 | $50290^6 \equiv 24711$ | $\gcd(24710, 57983) = 1$ |
| 7 | $24711^7 \equiv 37816$ | $\gcd(37815, 57983) = 2521$ |
| 8 | $37816^8 \equiv 42858$ | $\gcd(42857, 57983) = 2521$ |

We find that $57983 = 2521 \times 23$.

Exercise 2.4. Use Pollard $p-1$ method to factor the numbers 689, 16637, and 315391, say with a bound up to $k = 10$. Try also with base $x_1 = 3$ instead of 2 for a change.

Pollard $p-1$ method needs the assumption that n is composite with a prime factor p such that $p-1$ factors into small primes, to ensure that it divides $k!$ for some relatively small value of k . Since there is no specific reason we use the initial term $x_1 = 2$, another value can be employed following an unsuccessful attempt.

Remark. In practice it may not be necessary to check $\gcd(x_k - 1, n)$ for each k because if it is nontrivial for some value k it will be for $k+1$ as well, like $k = 7$ and $k = 8$ in our example. For this reason and since the technique is not guaranteed to bring success, we might put a bound for k , say up to M then check only $\gcd(x_M, n)$ and stop the process if it fails to find a factor, or run it again with a different x_1 value. The setback for this is that sometimes M may turn out too big and we miss what we are after. For instance with $n = 57983$ above, 23 is the only other factor of n and $23 - 1 = 2 \times 11$ hence $x_k = 1$ for all $k \geq 11$. Nevertheless for n large this is not likely the case.

Both the rho method and the $p-1$ method were invented by J. M. Pollard in 1974. These days Pollard $p-1$ method is still commonly used for factoring medium size integers and is the basis for its more powerful generalization called the Elliptic Curve Method, which will not be discussed here.

2.5 Exponent Factorization

The next three factorization techniques rely upon the following principle which is really another form of Euclid's lemma.

Lemma 2.1. Suppose $x^2 \equiv y^2 \pmod{n}$ but $x \not\equiv \pm y \pmod{n}$. Then both $\gcd(x \pm y, n)$ give a nontrivial factor of n .

Exercise 2.5. Prove this lemma.

Exercise 2.6. Factor the number 45113 given that $11999^2 \equiv 29174^2 \pmod{45113}$.

Now suppose we find two integers a, b such that $a^b \equiv 1 \pmod{n}$. Factor $b = 2^c \times d$ such that d is odd, then recursively define $x_0 = a^d \% n$ and $x_k = x_{k-1}^2 \% n$ for $k = 1, 2, \dots, c$. If there is in this sequence $x_k \neq n-1$ such that $x_{k+1} = 1$ then by Lemma 3.1 $\gcd(x_k - 1, n)$ is a nontrivial factor of n .

Example. Let $n = 15677$ and suppose we know that $7^{3840} \equiv 1 \pmod{n}$. We factor $3840 = 2^8 \times 15$ and start computing with $a = 7$, $c = 8$, and $d = 15$.

$$\begin{aligned} 7^{15} \% 15677 &= 1597 \\ 1597^2 \% 15677 &= 10735 \\ 10735^2 \% 15677 &= 14275 \\ 14275^2 \% 15677 &= 5979 \\ 5979^2 \% 15677 &= 4881 \\ 4881^2 \% 15677 &= 10798 \\ 10798^2 \% 15677 &= 6955 \\ 6955^2 \% 15677 &= 8480 \\ 8480^2 \% 15677 &= 1 \end{aligned}$$

Hence we now evaluate $\gcd(8480 - 1, 15677) = 61$ and the result is $15677 = 61 \times 257$.

Note that we will always have $x_c = 1$ since $x_c \equiv a^b \pmod{n}$. In fact we should terminate this procedure as soon as we encounter the term 1, in which case this procedure fails or succeeds depending on whether or not the previous term is $n - 1$, respectively.

Exercise 2.7. Factor the number 23797 given that $2^{11648} \equiv 1 \pmod{23797}$. Similarly again with the congruence $5^{14848} \equiv 1 \pmod{59881}$.

Exponent Factorization method, so it is called, does not always work and evidently it would take another good technique to find the numbers a, b . However the method can be used to supplement others, for instance the Pollard $p - 1$ method, which fails when $2^{k!} \equiv x_{k-1}^k \equiv 1 \pmod{n}$ but gives us $a = x_{k-1}$ and $b = k$ to be used with Exponent Factorization method.

2.6 Quadratic Sieve

Suppose $n = 91027$ and we find the following congruences.

$$\begin{aligned} 427^2 &\equiv 5^2 \times 11 && \pmod{91027} \\ 523^2 &\equiv 2^6 \times 7 && \pmod{91027} \\ 675^2 &\equiv 2 \times 5 \times 7^2 && \pmod{91027} \\ 1091^2 &\equiv 2 \times 3^2 \times 5 \times 7 \times 11 && \pmod{91027} \end{aligned}$$

Multiplying them all results in a congruence with squares on both sides,

$$(427 \times 523 \times 675 \times 1091)^2 \equiv (2^4 \times 3 \times 5^2 \times 7^2 \times 11)^2 \pmod{91027}$$

which, after taking residues mod 91027, becomes $49336^2 \equiv 9611^2 \pmod{91027}$. Now use Lemma 3.1 to obtain a factor of n from $\gcd(49336 - 9611, 91027) = \gcd(39725, 91027) = 227$, and indeed $91027 = 227 \times 401$.

But how did we discover those congruences? We look in numbers that are just a bit larger than \sqrt{kn} so that their squares mod n are small, and select those having prime factors only up to 11. (There is no specific reason for choosing 11, but for our purposes it seems ideal to allow primes up to 19.) In fact we used in our example $427 = \lfloor 2n \rfloor + 1$, $523 = \lfloor 3n \rfloor + 1$, $675 = \lfloor 5n \rfloor + 1$, and $1091 = \lfloor 13n \rfloor + 4$. With three more selections like these we organize them in a table.

| | 427^2 | 523^2 | 675^2 | 854^2 | 1001^2 | 1046^2 | 1091^2 |
|----|---------|---------|---------|---------|----------|----------|----------|
| 2 | – | 6 | 1 | 2 | 6 | 8 | 1 |
| 3 | – | – | – | – | – | – | 2 |
| 5 | 2 | – | 1 | 2 | – | – | 1 |
| 7 | – | 1 | 2 | – | – | 1 | 1 |
| 11 | 1 | – | – | 1 | 1 | – | 1 |

Thus we are looking for a linear dependence among the columns modulo 2. Linear algebra guarantees such dependencies when the number of columns exceeds that of rows. In fact there are a few more from this table, for instance,

$$854^2 \times 1001^2 \equiv (2^4 \times 5 \times 11)^2 \pmod{91027}$$

which produces $35611^2 \equiv 880^2 \pmod{91027}$ and similarly $\gcd(35611 - 880, 91027) = 227$. Or another instance,

$$523^2 \times 1046^2 \equiv (2^7 \times 7)^2 \pmod{91027}$$

which turns out useless: $896^2 \equiv 896^2 \pmod{91027}$.

Exercise 2.8. Follow the example to factor the numbers 4897, 21733, and 95321, say allowing prime factors up to 19.

The technique described above is a simplified version of the so-called Quadratic Sieve, first introduced by Carl Pomerance in 1981. Historically it was the first method successful in factoring an arbitrary integer over 100 digits including one of the challenge numbers RSA-129. For integers over 200 digits a more powerful generalization of the quadratic sieve, called the General Number Field Sieve, is more effective.

2.7 Continued Fractions

A *continued fraction* is an expression involving a sequence of numbers, all except the first must be positive, in the form

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{\ddots}}}}$$

which for convenience shall be symbolically written $[a_0, a_1, a_2, a_3, \dots]$ whether or not the length is finite. For us a continued fraction is always understood with integer entries, which in the literature is distinguished by the name *simple* continued fraction.

Continued fractions is a vast subject and it is not our plan to study them in depth. We will be interested only in illustrating the main results which will lead to at least one factorization technique, based on Lemma 3.1, first studied by Lehmer and Powers in 1931.

Theorem 2.2. Every finite continued fraction represents a rational number. Conversely, every rational number can be represented by a finite continued fraction.

For examples $[1, 3] = 1 + 1/3 = 4/3$ and $[2, 1, 3] = 2 + 1/(4/3) = 11/4$. You can see that the first claim can be easily verified by induction.

Exercise 2.9. Evaluate the continued fraction $[1, 2, 3, 4]$.

To illustrate the converse we choose the rational number $253/17$ and proceed as follows.

$$\begin{aligned} 253/17 &= 14 + 15/17 \\ 17/15 &= 1 + 2/15 \\ 15/2 &= 7 + 1/2 \\ 2/1 &= 2 + 0 \end{aligned}$$

Hence $253/17 = [14, 1, 7, 2]$ and note that this procedure is really the Euclidean algorithm, which we have seen always terminates after a finite number of steps. Moreover we are given a hint that such a representation is uniquely determined by the above algorithm with the exception that in the last equation $2 + 0$ can well be expressed as $1 + 1/1$, giving an alternate tail $253/17 = [14, 1, 7, 1, 1]$.

Exercise 2.10. Represent the rational numbers $7/11$, $-19/9$, and $333/99$ using finite continued fractions.

Theorem 2.3. Suppose a_0, a_1, a_2, \dots is a sequence of positive integers except perhaps $a_0 \leq 0$. Then $[a_0, a_1, a_2, \dots, a_k] = p_k/q_k$ where p_k, q_k are obtained recursively as follow.

$$\begin{aligned} p_0 &= a_0 & q_0 &= 1 \\ p_1 &= a_0 a_1 + 1 & q_1 &= a_1 \\ p_k &= a_k p_{k-1} + p_{k-2} & q_k &= a_k q_{k-1} + q_{k-2} \end{aligned}$$

For example with the continued fraction $[3, 6, 1, 7]$ we have

$$\begin{array}{lll} p_0 = 3 & q_0 = 1 & [3] = 3 \\ p_1 = 3 \cdot 6 + 1 = 19 & q_1 = 6 & [3, 6] = 19/6 \\ p_2 = 1 \cdot 19 + 3 = 22 & q_2 = 1 \cdot 6 + 1 = 7 & [3, 6, 1] = 22/7 \\ p_3 = 7 \cdot 22 + 19 = 173 & q_3 = 7 \cdot 7 + 6 = 55 & [3, 6, 1, 7] = 173/55 \end{array}$$

Exercise 2.11. Repeat the example using the sequence $1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1$.

Keeping the same notations, we now define an infinite continued fraction, written $[a_0, a_1, a_2, \dots]$, to be the limit of the sequence p_k/q_k . This sequence, it can be proved, always converges to an irrational number and moreover the terms p_k/q_k provide an excellent rational number approximation to its value. It is left to the student to read about these results as an independent assignment.

Theorem 2.4. Every infinite continued fraction represents an irrational number. Furthermore it is periodic if and only if it represents a *quadratic irrational*, that is an irrational root of a quadratic polynomial with rational coefficients.

For example consider the periodic continued fraction $[3, 1, 2, 1, 2, 1, 2, \dots] = [3, \overline{1, 2}]$. We find what it represents as follows. Let $x = [\overline{1, 2}]$ then

$$x = 1 + \frac{1}{2 + \frac{1}{x}} = 1 + \frac{x}{2x + 1} = \frac{3x + 1}{2x + 1}$$

which yields the quadratic equation $2x^2 - 2x - 1 = 0$, whose positive root is $x = (1 + \sqrt{3})/2$. Therefore

$$[3, \overline{1, 2}] = 3 + \frac{2}{1 + \sqrt{3}} = 3 + \frac{2(1 - \sqrt{3})}{(1 + \sqrt{3})(1 - \sqrt{3})} = \frac{-6 + 2 - 2\sqrt{3}}{-2} = 2 + \sqrt{3}$$

Exercise 2.12. Find the quadratic irrationals represented by the periodic continued fractions $[\overline{1, 2}]$, $[1, \overline{2, 3}]$, and $[1, 2, \overline{3}]$. Write your answers in the form $(P + \sqrt{n})/Q$.

Theorem 2.5. Suppose we have a quadratic irrational in the form $\alpha = (P_0 + \sqrt{n})/Q_0$ such that $Q_0 \mid (n - P_0^2)$. Then $\alpha = [a_0, a_1, a_2, \dots]$ where the integers a_k for $k = 0, 1, 2, \dots$ can be obtained via the following recursive algorithm.

$$\begin{aligned} \alpha_k &= (P_k + \sqrt{n})/Q_k \\ a_k &= \lfloor \alpha_k \rfloor \\ P_{k+1} &= a_k Q_k - P_k \\ Q_{k+1} &= (n - P_{k+1}^2)/Q_k \end{aligned}$$

Example. First note that for every quadratic irrational the condition $Q_0 \mid (n - P_0^2)$, if not already true, can be realized by multiplying each term by $|Q_0|$. For example $\alpha = (4 + \sqrt{3})/2$ does not satisfy this condition but we have $\alpha = \frac{4 \cdot 2 + \sqrt{3 \cdot 2^2}}{2 \cdot 2} = \frac{8 + \sqrt{12}}{4}$ where $4 \mid (12 - 8^2) = -52$. Then we start with $P_0 = 8, Q_0 = 4$, and $n = 12$.

| k | P_k | Q_k | α_k | a_k |
|-----|---------------------|--------------------|----------------------------------|-------|
| 0 | 8 | 4 | $(8 + \sqrt{12})/4 \approx 2.86$ | 2 |
| 1 | $2 \cdot 4 - 8 = 0$ | $(12 - 0^2)/4 = 3$ | $(0 + \sqrt{12})/3 \approx 1.15$ | 1 |
| 2 | $1 \cdot 3 - 0 = 3$ | $(12 - 3^2)/3 = 1$ | $(3 + \sqrt{12})/1 \approx 6.46$ | 6 |
| 3 | $6 \cdot 1 - 3 = 3$ | $(12 - 3^2)/1 = 3$ | $(3 + \sqrt{12})/3 \approx 2.15$ | 2 |
| 4 | $2 \cdot 3 - 3 = 3$ | $(12 - 3^2)/3 = 1$ | $(3 + \sqrt{12})/1 \approx 6.46$ | 6 |

Since $(P_4, Q_4) = (P_2, Q_2)$ we stop at $k = 4$ and obtain the periodic continued fraction $(4 + \sqrt{3})/2 = [2, 1, \overline{6, 2}]$.

Exercise 2.13. Repeat the example with the numbers $\sqrt{3}$, the golden ratio $(1 + \sqrt{5})/2$, and $(5 - \sqrt{7})/4$, to represent them using periodic continued fractions.

Given a composite n , possibly very large, we find the continued fraction representation of $\sqrt{n} = [a_0, a_1, a_2, \dots]$ following Theorem 2.5, with $P_0 = 0, Q_0 = 1, \alpha_0 = \sqrt{n}$, and $a_0 = \lfloor \sqrt{n} \rfloor$. Conversely, the convergents (think of partial sums) $\frac{p_k}{q_k} = [a_0, a_1, a_2, \dots, a_k]$ can be evaluated as in Theorem 2.3.

The key to factoring n is the following identity.

$$p_k^2 - nq_k^2 = (-1)^{k+1} Q_{k+1}$$

It can be shown that this quantity is small, independently from k . In fact, $Q_k < 2\sqrt{n}$. Hence, the sequence p_k provides suitable trial numbers x to be used in a factor base technique similar to that in quadratic sieve, where

$$x^2 = p_k^2 \equiv (-1)^{k+1} Q_{k+1} \pmod{n}$$

Example. Let $n = 65363$. The two theorems generate the following table.

| k | P_k | Q_k | a_k | p_k | q_k | $p_k^2 - nq_k^2$ |
|-----|-------|-------|-------|-----------|--------|------------------|
| 0 | 0 | 1 | 255 | 255 | 1 | -338 |
| 1 | 255 | 338 | 1 | 256 | 1 | 173 |
| 2 | 83 | 173 | 1 | 511 | 2 | -331 |
| 3 | 90 | 331 | 1 | 767 | 3 | 22 |
| 4 | 241 | 22 | 22 | 17385 | 68 | -287 |
| 5 | 243 | 287 | 1 | 18152 | 71 | 221 |
| 6 | 44 | 221 | 1 | 35537 | 139 | -154 |
| 7 | 177 | 154 | 2 | 89226 | 349 | 313 |
| 8 | 131 | 313 | 1 | 124763 | 488 | -103 |
| 9 | 182 | 103 | 4 | 588278 | 2301 | 121 |
| 10 | 230 | 121 | 4 | 2477875 | 9692 | -7 |
| 11 | 254 | 7 | 72 | 178995278 | 700125 | 409 |
| 12 | 250 | 409 | 1 | 181473153 | 709817 | -98 |

Compare the last column to that of Q_k , where each entry is bounded by $2\sqrt{65363} < 512$. These will serve as the trial numbers $p_k^2 \pmod{65363}$, whose factorizations are given in the next table. We discard the trial number if it has a prime factor larger than, say, 17. We will also include the *prime* -1 in order to accomodate the plus/minus sign.

| $p_k^2 \equiv (-1)^{k+1}Q_{k+1}$ | -1 | 2 | 3 | 5 | 7 | 11 | 13 | 17 |
|----------------------------------|----|---|---|---|---|----|----|----|
| $255^2 \equiv -338$ | 1 | 1 | - | - | - | - | 2 | - |
| $767^2 \equiv 22$ | - | 1 | - | - | - | 1 | - | - |
| $18152^2 \equiv 221$ | - | - | - | - | - | - | 1 | 1 |
| $35537^2 \equiv -154$ | 1 | 1 | - | - | 1 | 1 | - | - |
| $588278^2 \equiv 121$ | - | - | - | - | - | 2 | - | - |
| $2477875^2 \equiv -7$ | 1 | - | - | - | 1 | - | - | - |
| $181473153^2 \equiv -98$ | 1 | 1 | - | - | 2 | - | - | - |

We quickly notice the single row of square—only to be disappointed:

$$\begin{aligned} 588278^2 &\equiv 11^2 && \pmod{65363} \\ 11^2 &\equiv 11^2 && \pmod{65363} \end{aligned}$$

The next attempt is again futile:

$$\begin{aligned} (767 \times 35537 \times 2477875)^2 &\equiv (-1 \times 2 \times 7 \times 11)^2 && \pmod{65363} \\ (-154)^2 &\equiv 154^2 && \pmod{65363} \end{aligned}$$

But at last we find a row combination that does the trick:

$$\begin{aligned} (255 \times 181473153)^2 &\equiv (-1 \times 2 \times 7 \times 13)^2 && \pmod{65363} \\ 22638^2 &\equiv 182^2 && \pmod{65363} \end{aligned}$$

This leads us to the final step of computing gcd,

$$\gcd(22638 - 182, 65363) = \gcd(22456, 65363) = 401$$

and the factorization of n ,

$$65363 = 163 \times 401$$

It is an unfortunate fact, as we have seen in this example, that the congruence $x^2 \equiv y^2 \pmod{n}$ may turn trivial. However, it is not hard to prove that such bad luck has only a 50% chance or less.

Exercise 2.14. Follow the example to factor the number 112529.

3 Primality Testing

Recall Wilson's theorem of Chapter 1; it provides a deterministic test to prove whether a given integer is prime or composite. This particular test is unfortunately far too slow to be useful in practical application. Nevertheless it serves as an illustration that primality testing and factorization are not identical problems. It is possible, for instance, to show an integer is composite without factoring it. The congruence $5! = 120 \equiv 0 \pmod{6}$ violates Wilson's theorem unless we conclude that 6 is composite.

3.1 More Number Theory

Definition. Suppose a and $n > 0$ are relatively prime. The *order* of a modulo n is the smallest positive integer k such that $a^k \% n = 1$. We denote this quantity by $|a|_n$ or simply $|a|$ when there is no ambiguity. For example $|2|_7 = 3$ because $k = 3$ is the smallest positive solution of the congruence $2^k \equiv 1 \pmod{7}$.

Exercise 3.1. Find $|3|_7, |5|_{12}, |4|_{25}$.

Theorem 3.1. We have $a^k \equiv 1 \pmod{n}$ if and only if $|a|_n \mid k$. In particular $|a|_n \mid \phi(n)$.

Definition. In the special case where $|a|_n = \phi(n)$, we call a a *primitive root* modulo n .

If a is a primitive root modulo n then every integer relatively prime to n is congruent to $a^k \pmod{n}$ for some exponent k . This is a very useful characteristic of a primitive root. However, for a given modulus n , primitive roots may not exist.

Theorem 3.2 (Primitive Root Theorem). Primitive roots modulo n exist if and only if $n = 1, 2, 4, p^k$, or $2p^k$ for any prime $p > 2$ and $k > 0$.

Definition. With a prime $p > 2$, the *Legendre symbol* is defined as follows.

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } x^2 \equiv a \pmod{p} \text{ has a solution} \\ -1 & \text{if } x^2 \equiv a \pmod{p} \text{ has no solution} \\ 0 & \text{if } p \mid a \end{cases}$$

Theorem 3.3 (Euler's Criterion). $\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod{p}$

We generalize the Legendre symbol by allowing the denominator to be composite. This will also provide a very practical way to compute the symbol. Let $n = p_1 p_2 \cdots p_k$ be the product of odd prime numbers, not necessarily distinct.

Definition. Define the *Jacobi symbol*

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right) \left(\frac{a}{p_2}\right) \cdots \left(\frac{a}{p_k}\right) \quad \text{and let} \quad \left(\frac{a}{1}\right) = 1$$

Theorem 3.4. Let n, m denote odd positive numbers.

- 1) $\left(\frac{ab}{n}\right) = \left(\frac{a}{n}\right) \left(\frac{b}{n}\right)$
- 2) $\left(\frac{a}{mn}\right) = \left(\frac{a}{m}\right) \left(\frac{a}{n}\right)$
- 3) $\left(\frac{-1}{n}\right) = (-1)^{(n-1)/2}$

$$4) \binom{2}{n} = (-1)^{(n^2-1)/8}$$

Theorem 3.5 (The Law of Quadratic Reciprocity). $\binom{m}{n} = \binom{n}{m} (-1)^{(m-1)(n-1)/4}$

Example. We use the properties of the Jacobi symbol to evaluate

$$\binom{22}{221} = \binom{2}{221} \binom{11}{221} = (-1)^{(221^2-1)/8} \binom{221}{11} (-1)^{(10)(220)/4} = - \binom{1}{11} (1) = -1$$

Exercise 3.2. Evaluate $\binom{37}{83}$, $\binom{-816}{239}$, $\binom{1414}{2063}$.

3.2 Carmichael Numbers

Suppose p is a prime. By FLT $a^{p-1} \equiv 1 \pmod{p}$ for every integer a not divisible by p . This statement is equivalent to having $a^p \equiv a \pmod{p}$ for any integer a . So here is another test for compositeness. For example we compute $2^{398617} \% 398617 = 291108$, hence $2^{398617} \not\equiv 2 \pmod{398617}$ and so 398617 cannot be a prime.

FLT, however, does not work in the other direction. We have $29^{35} \equiv 29 \pmod{35}$ even though 35 is not a prime. In this case we might want to test another value of a , for example $12^{35} \equiv 3 \not\equiv 12 \pmod{35}$, which confirms that 35 is composite.

Definition. But is it possible that for some composite n we have $a^n \equiv a \pmod{n}$ for every integer a ? The answer is yes and such n is called a *Carmichael number*.

Example. The smallest Carmichael number is $n = 561 = 3 \times 11 \times 17$. To verify that $a^{561} \equiv a \pmod{561}$ it suffices by CRT to justify the congruences $a^{561} \equiv a \pmod{3, 11, 17}$ independently. Now by FLT $a^2 \equiv 1 \pmod{3}$ when $3 \nmid a$ hence $a^{560} = (a^2)^{280} \equiv 1 \pmod{3}$ and it follows that $a^{561} \equiv a \pmod{3}$ for any integer a . For the other two moduli 11 and 17 it can be done in a similar way.

Exercise 3.3. Prove that 1729 is a Carmichael number by showing $a^{1729} \equiv a \pmod{1729}$ for every integer a .

Theorem 3.6 (Korselt's Criterion). A composite number n is a Carmichael number if and only if it factors into distinct primes such that $(p-1) \mid (n-1)$ for each factor p .

For example $1105 = 5 \times 13 \times 17$ where each factor is distinct. We check that $4 \mid 1104$ and $12 \mid 1104$ and $16 \mid 1104$. Hence 1105 is another Carmichael number.

Exercise 3.4. Apply Korselt's criterion on the numbers 10659, 19747, and 62745 to see if they are Carmichael numbers.

Carmichael numbers, in a way, are composites that masquerade themselves as primes. Obviously even numbers have no such chance to fool us and indeed they are already ruled out by the theorem:

Exercise 3.5. Use Korselt's criterion to show that Carmichael numbers are all odd, and that each must have at least three prime factors.

A number with only distinct prime factors is called a *square-free* number since it is not divisible by any square. A Carmichael number, in other word, must be composite, square-free, plus the divisibility conditions stated above. First conjectured in the year 1910, it was now proven in 1984 that there exist infinitely many Carmichael numbers. The first seven Carmichael numbers are 561, 1105, 1729, 2465, 2821, 6601, and 8911.

Definition. Define a composite number n to be a *Fermat pseudoprime* to the base a if it passes Fermat test: $a^n \equiv a \pmod{n}$, which is equivalent to the congruence $a^{n-1} \equiv 1 \pmod{n}$ when $\gcd(a, n) = 1$. Earlier we have seen that 35 is a Fermat pseudoprime to the base 29 but not to the base 12. We can now say that Carmichael numbers compose the intersection of all Fermat pseudoprimes to different bases.

Exercise 3.6. Find a Fermat pseudoprime less than 100 to the base 3.

3.3 Strong Pseudoprimes

The existence of Carmichael numbers is unfortunate as far as primality test using FLT is concerned. We need stronger, if not deterministic, primality tests which are reasonably easy to implement and fast.

Corollary 1.3 tells us another property of prime numbers: If $x^2 \equiv 1 \pmod{p}$ where p is prime then $x \equiv \pm 1 \pmod{p}$. Thus in an attempt to catch a Fermat pseudoprime whereby $a^{n-1} \equiv 1 \pmod{n}$ we “take square root” to obtain a congruence in the form $a^{\frac{n-1}{2}} \equiv \pm 1 \pmod{n}$. If the congruence does not hold then we know n is composite but if it does unfortunately no conclusion can be drawn, unless $a^{\frac{n-1}{2}} \equiv 1 \pmod{n}$ in which case we may iterate the process by taking another square root and again for as long as the exponent is an even number. This explains the following primality test.

Theorem 3.7 (Miller-Rabin Primality Test). Let $n - 1 = 2^c \times d$ with d odd and select an integer a as our base number. If there is a term $n - 1$ in the sequence

$$a^d \% n, a^{2d} \% n, a^{4d} \% n, a^{8d} \% n, \dots, a^{2^{c-1}d} \% n = a^{\frac{n-1}{2}} \% n$$

or if they are all 1, then we say that n passes the test. Otherwise n is composite.

Note that the sequence is generated by successive squaring, which is much faster than the reverse order of taking successive square roots. In particular the terms are uniformly all 1 if and only if $a^d \% n = 1$.

Example. Let $n = 561$, the first Carmichael number, we have $560 = 2^4 \times 35$. Choosing the base $a = 2$ we find in what follows that n fails MRT and hence 561 is composite.

$$\begin{aligned} 2^{35} &\equiv 263 \not\equiv \pm 1 \pmod{561} \\ 263^2 &\equiv 166 \not\equiv -1 \pmod{561} \\ 166^2 &\equiv 67 \not\equiv -1 \pmod{561} \\ 67^2 &\equiv 1 \not\equiv -1 \pmod{561} \end{aligned}$$

Definition. In view of Theorem 4.2, we call a composite odd number a *strong pseudoprime* to the base a if it passes MRT. The smallest example of a strong pseudoprime to the base 2 is $n = 2047$. One can verify that it is composite and, since $2046 = 2 \times 1023$, that $2^{1023} \equiv 1 \pmod{2047}$, thereby passing the test.

Exercise 3.7. Apply MRT on the numbers 3281, 4681, and 6673 using the base numbers 2, 3, and 5. Is any of these a strong pseudoprime? Can you conclude the primality for each number?

Note that $a^{n-1} \% n$ would have been the next square in the above sequence. Hence if a number n passes this test then it passes Fermat test $a^{n-1} \equiv 1 \pmod{n}$ as well. This fact shows that MRT is stronger than FLT in their roles of primality testing, stated as follows.

Theorem 3.8. To the same base, every strong pseudoprime is a Fermat pseudoprime.

What is more, there is no analog of Carmichael numbers to watch for here. It means that a strong pseudoprime will eventually fail MRT for some choice of a base number a . However to prove that n is prime by MRT may not be feasible since we would have to see that it pass the test for every base a . Nevertheless it has been shown that if n is composite then at least 75% of the base numbers a selected between 1 and $n - 1$ will fail it. Hence if n is composite the probability that it will pass the test for a few randomly chosen bases will be extremely small.

Theorem 3.9 (Rabin’s Probabilistic Primality Test). Given an odd integer n , we select k positive integers less than n as base numbers for the Miller-Rabin test. The probability that n is a strong pseudoprime to all k bases is less than $1/4^k$.

For example testing n using base numbers 2, 3, 5, 7, and 11 will prove either n is composite or else (possibly) prime with probability over $1 - (1/4)^5 = 0.9990234375$ of being correct. To convince ourselves even further with statistical facts, up to $n = 25,000,000,000$ there are

| | |
|-----------------|--|
| > 1,000,000,000 | primes |
| 21853 | Fermat pseudoprimes to base 2 |
| 4842 | strong pseudoprimes to base 2 |
| 2163 | Carmichael numbers |
| 184 | strong pseudoprimes to bases 2 and 3 |
| 13 | strong pseudoprimes to bases 2, 3, and 5 |
| 1 | strong pseudoprime to bases 2, 3, 5, and 7 |
| 0 | strong pseudoprime to bases 2, 3, 5, 7, and 11 |

In addition, the smallest strong pseudoprime to the bases 2 and 3 is 1,373,653 while to the bases 2, 3, and 5 is 25,326,001. The single pseudoprime to the bases 2, 3, 5, and 7 is known to be the number 3,215,031,751. Hence for the rest of the 25 billion numbers minus one, it suffices to test for primality up to these four bases. (In fact it is still true for all $n < 118,670,087,467$, which is the next strong pseudoprime to the four bases!)

3.4 Euler Pseudoprimes

Recall Euler’s criterion, which states that the Legendre symbol $\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod{p}$ for odd prime $p \nmid a$. This is the basis for our next primality test, the Euler test. For example with $n = 341$ and $a = 2$ we compute $\left(\frac{2}{341}\right) = -1$ contrary to $2^{170} \% 341 = 1$. In violation of Euler’s criterion, it is conclusive that 341 is composite. Note that we have replaced the Legendre symbol in Euler’s criterion by the Jacobi symbol $\left(\frac{2}{341}\right)$ since Legendre symbol only works with primes. This, however, does not affect our claim.

Definition. Similarly again, we define an *Euler pseudoprime* to the base a to be a composite odd number n for which the Jacobi symbol satisfies $\left(\frac{a}{n}\right) \equiv a^{(n-1)/2} \pmod{n}$. For example it is not hard to check that 1105 is an Euler pseudoprime to the base 2 since it is composite and $\left(\frac{2}{1105}\right) \equiv 2^{552} \equiv 1 \pmod{1105}$.

The next proposition states that the strength of Euler test in primality testing is somewhat in between FLT and MRT.

Theorem 3.10. To the same base, every strong pseudoprime is an Euler pseudoprime and every Euler pseudoprime is a Fermat pseudoprime.

To the same base a , conversely, an Euler pseudoprime n can sometimes be a strong pseudoprime, in particular when $n \% 4 = 3$ or when $\left(\frac{a}{n}\right) = -1$.

Exercise 3.8. Try to prove these two claims.

Similar to MRT, Euler test can be used as a deterministic primality test, for it can be proved that if n is composite then it is guaranteed to fail Euler's criterion for some base number a . This result is nevertheless not very practical for large n . Theoretically, what has been shown is that only up to half of the base numbers need to be tested in order to prove primality, and this fact supports the following probabilistic primality test.

Theorem 3.11 (Solovay-Strassen's Probabilistic Primality Test). Given an odd integer n , we randomly select k positive integers less than n as base numbers for Euler test. The probability that n is an Euler pseudoprime to all k bases is less than $1/2^k$.

3.5 Extended Fermat Tests

Recall that Fermat pseudoprimes are an evidence that the converse of FLT is not necessarily true. There are, nevertheless, primality tests based on a partial converse of FLT. What it means is that if an odd integer n and some base number a satisfy the congruence $a^{n-1} \equiv 1 \pmod{n}$ plus some other conditions which are to be specified later, then n has to be a prime. In the next theorem we group together three such tests, none of which is deterministic.

Theorem 3.12. Suppose n is odd and $a^{n-1} \equiv 1 \pmod{n}$. Then n is a prime if in addition any one of the following conditions is satisfied.

- 1) (*Lucas' Primality Test*) $a^{(n-1)/q} \not\equiv 1 \pmod{n}$ for every prime divisor q of $n-1$.
- 2) (*Pocklington's Primality Test*) $\gcd(a^{(n-1)/q} - 1, n) = 1$ for every prime $q \mid F$ where $n-1 = FR$ with $F > R$ and $\gcd(F, R) = 1$.
- 3) (*Proth's Primality Test*) $a^{(n-1)/2} \equiv -1 \pmod{n}$ and $n-1 = 2^c \times d$ where $2^c > d$.

Example. To illustrate we apply Lucas' test with $n = 3329$. We try $a = 3$ and find that $3^{3328} \equiv 1 \pmod{3329}$. Since $3328 = 2^8 \times 13$ we proceed with the two prime divisors $q = 2$ and $q = 13$. Neither of these is congruent to one: $3^{3328/2} = 3^{1664} \equiv -1 \pmod{3329}$ and $3^{3328/13} = 3^{256} \equiv 2970 \pmod{3329}$. Hence 3329 is a prime.

Note that in this case if n is prime then $a^{(n-1)/2} \equiv \pm 1 \pmod{n}$ by Corollary 1.3. So in order for Lucas test to work it is necessary that $a^{(n-1)/2} \equiv -1 \pmod{n}$.

Example. For Pocklington's test let $n = 3001$. We have $3000 = 5^3 \times 24$ where $5^3 > 24$ and $\gcd(5^3, 24) = 1$. Choose $a = 2$, which satisfies the congruence $2^{3000} \equiv 1 \pmod{3001}$, and find that $2^{3000/5} = 2^{600} \equiv 1125 \pmod{3001}$. It happens that $\gcd(1125-1, 3001) = 1$ hence we conclude that 3001 is prime.

The advantage of this test is that we do need to complete the factorization of $n-1$, that is the R part. The idea is to factor out one or more primes off $n-1$, to the highest possible exponent, until $F > R$, where R is the unfactored part and hope that the gcd conditions come up allright.

Example. Lastly, with the same $n = 3329$ used in Lucas' test, Proth's test requires checking only the first congruence (with $q = 2$) because $3328 = 2^8 \times 13$ satisfies the condition $2^8 > 13$. It gives the same conclusion that 3329 is a prime. In fact due to its simplicity, this test has been used and with a degree of success to search for large primes of the form $2^c \times d + 1$ where $d < 2^c$.

Exercise 3.9. Illustrate the three tests, whichever applicable, using the prime numbers 1009, 2689, and 23801.

3.6 Fermat Numbers

In the remainder of this chapter we discuss numbers of special types which are of interest in relation to primality testing and prime number search or otherwise simply recreational.

Proposition 3.13. If $a^k + 1$ is an odd prime then a is even and $k = 2^n$.

Exercise 3.10. Prove this, then find a nontrivial factor of the number $2^{40} + 1$.

Definition. For each $n \geq 0$ we define the *Fermat number* $F_n = 2^{2^n} + 1$. The first few are $F_0 = 3, F_1 = 5, F_2 = 17, F_3 = 257, F_4 = 65537$, which happen to be all primes, thus called *Fermat primes*. However the next Fermat number, $F_5 = 4294967297$, is composite. In fact it is not known whether or not there are any more Fermat prime beyond F_4 and it is commonly believed there are not.

Proposition 3.14. Properties of Fermat numbers:

- 1) Every prime divisor of F_n is of the form $2^{n+1}k + 1$.
- 2) Fermat numbers are relatively prime one to another.
- 3) Fermat numbers satisfy the recurrence relation $F_n = F_0 F_1 F_2 \dots F_{n-1} + 2$.

Proposition (1) can be used to test the primality of F_n for small values of n . In fact a stronger statement, proved with the help of Euler test, holds with prime divisors of the form $2^{n+2}k + 1$.

Example. Take $F_4 = 65537$, the only possible prime divisors are of the form $64k + 1$ and there is only one of them up to $\sqrt{65537} < 257$, namely 193. It happens that $193 \nmid 65537$ hence F_4 is a Fermat prime.

Proposition (2) says that each successive term in the Fermat number sequence gives totally new prime factors. Thus another proof that the number of primes is infinite!

Theorem 3.15 (Pepin's Primality Test for Fermat Numbers). Other than F_0 , the Fermat number F_n is prime if and only if $3^{(F_n-1)/2} \equiv -1 \pmod{F_n}$.

Note that since $F_n - 1$ is a power of 2, the sufficiency in this theorem is an immediate result of Lucas' test (or Proth's test with $d = 1$). To prove necessity is a bit more involved and we will not give it here. For example $F_1 = 5$ is prime, and we check that $3^2 \equiv -1 \pmod{5}$.

Exercise 3.11. Illustrate Pepin's test for the Fermat numbers F_2, F_3, F_4 .

Pepin's primality test, though deterministic, is practically useless due to the enormous size of F_n . It might yet be reasonably computable by a calculator, for example, to use this theorem to show that F_5 is composite without factoring it.

3.7 Mersenne Primes

Proposition 3.16. If $a^k - 1$ is prime with $k \geq 2$ then $a = 2$ and k is prime.

Exercise 3.12. Prove this, then find two nontrivial factors of the number $2^{35} - 1$.

Definition. For each prime p we define the *Mersenne number* $M_p = 2^p - 1$. When M_p happens to be prime we call it a *Mersenne prime*.

The first few Mersenne numbers are indeed Mersenne primes: $M_2 = 3, M_3 = 7, M_5 = 31, M_7 = 127$. The next one, $M_{11} = 2047 = 23 \times 89$ is not. The following proposition can help in testing the primality of M_p for small values of p .

Proposition 3.17. The divisors of Mersenne numbers:

- 1) Every divisor of M_p is of the form $2pk + 1$.
- 2) Every divisor of M_p is also of the form $8k \pm 1$.
- 3) Suppose $p \% 4 = 3$. Then $2p + 1$ is prime if and only if it divides M_p .

Example. To illustrate Proposition (1) we consider $M_{11} = 2047$. The only possible prime divisors are in the form $22k + 1$ and there is only one of them up to $\sqrt{2047} < 46$, namely 23. It happens that $23 \mid 2047$, hence M_{11} is composite. Proposition (2) can also be used in combination with (1), thereby further reducing the number of possible candidate divisors.

Exercise 3.13. Determine the primality of the Mersenne numbers M_{13} and M_{17} .

Proposition (3) can be used to prove, for example, that M_{23} is composite for $23 \% 4 = 3$ and $2 \times 23 + 1 = 47$ is prime, hence $47 \mid M_{23}$. The previous fact that $23 \mid M_{11}$ could have been derived as well in a similar manner.

Remark. It is not known whether or not there are infinitely many Mersenne primes although it is commonly conjectured there are. The largest known Mersenne prime M_p at the time of this writing is when $p = 43112609$, discovered in August 2008. This is just the 45th Mersenne prime and its length is 12,978,189 decimal digits.

Theorem 3.18 (Lucas-Lehmer Primality Test for Mersenne Numbers). Let M_p be a Mersenne number. Consider the recursive sequence given by $x_n = (x_{n-1}^2 - 2) \% M_p$ with initial term $x_1 = 4$. Then M_p is prime if and only if $x_{p-1} = 0$.

Example. Consider the Mersenne number $M_5 = 31$, a prime.

$$\begin{array}{rcl} x_1 & = & 4 \qquad \% 31 = 4 \\ x_2 & = & 4^2 - 2 \qquad \% 31 = 14 \\ x_3 & = & 14^2 - 2 \qquad \% 31 = 8 \\ x_4 & = & 8^2 - 2 \qquad \% 31 = 0 \end{array}$$

Exercise 3.14. Repeat the previous exercise, this time using LLT.

In an earlier version of this test, Lucas used the initial term $x_1 = 3$ but it worked only when $p \% 4 = 3$. LLT is surprisingly simple and easily implemented. It is one of the algorithms employed by The Great Internet Mersenne Prime Search (GIMPS) at www.mersenne.org, a site dedicated to finding world record primes.

3.8 Perfect Numbers

Definition. Over the domain of positive integers let the function $\sigma(n)$ denote the sum of all the divisors of n . For example $\sigma(12) = 1 + 2 + 3 + 4 + 6 + 12 = 28$. We call n *perfect* when $\sigma(n) = 2n$, for instance $6 = 1 + 2 + 3 + 6 = 2 \times 6$. Equivalently a number is perfect if it equals the sum of its proper divisors, meaning those other than the number itself.

Proposition 3.19. If $\gcd(m, n) = 1$ then $\sigma(mn) = \sigma(m)\sigma(n)$.

Exercise 3.15. Illustrate this proposition with two examples and then prove it.

In number theory a function $f(n)$ with the property that $f(mn) = f(m)f(n)$ whenever $\gcd(m, n) = 1$ is called a *multiplicative* function. We have now seen two examples of multiplicative functions namely $\sigma(n)$ and $\phi(n)$.

Using the multiplicity of $\sigma(n)$ we can prove the next theorem which says that Mersenne primes are tightly connected with even perfect numbers. The first statement was demonstrated by Euclid whereas the second, two millennia later, was proved by Euler.

Theorem 3.20. If $M_p = 2^p - 1$ is a Mersenne prime then $2^{p-1}M_p$ is a perfect number. Conversely every even perfect number comes in this form.

Hence there is a one-to-one correspondence between Mersenne primes and even perfect numbers. The list of Mersenne primes 3, 7, 31, ... corresponds to the list of even perfect numbers 6, 28, 496, ...

Exercise 3.16. Find the first six even perfect numbers in this way.

Proposition 3.21. Properties of even perfect numbers:

- 1) The unit digit of any even perfect number must be 6 or 8.
- 2) Except 6, every even perfect number is of the form $9k + 1$.
- 3) In binary, an even perfect number reads p ones followed by $p - 1$ zeros.
- 4) Every even perfect number is a *triangular* number, that is, a number of the form $1 + 2 + 3 + \dots + k$, and in particular with $k = M_p$.

Example. We illustrate using the third perfect number, 496. First note that $496 \equiv 1 \pmod{9}$. (As seen in Section 2.1, basically this means that the iterative digit sums terminate with 1, that is $4 + 9 + 6 = 19$, next $1 + 9 = 10$, next $1 + 0 = 1$.) Then in binary $496 = 111110000_2$ —five ones (corresponding to M_5) and four zeros. Finally $496 = 1 + 2 + 3 + \dots + 31$, where again $31 = M_5$.

Exercise 3.17. Repeat the example using the perfect numbers from the last exercise.

Perfect numbers have a very fascinating history. They seem to possess certain mystical values to the ancient Greeks and have been studied to modern times by mathematicians, philosophers, poets, even theologians. While the list of Mersenne primes, hence even perfect numbers, seems to be endless, no one to date has ever seen an odd perfect number, if any exists.

4 Mathematica Projects

This chapter will be provided in a separate class handouts.

To Learn More

For those interested in a further study, the following books are well recommended.

- 1) David M. Bressoud, *Factorization and Primality Testing*, Springer 1989.
- 2) Harold M. Edwards, *Higher Arithmetic: An Algorithmic Introduction to Number Theory*, American Mathematical Society 2008.
- 3) M. Křížek, F. Luca, and L. Somer, *17 Lectures on Fermat Numbers: from Number Theory to Geometry*, Springer 2001.
- 4) A. Witno, *Theory of Numbers*, BookSurge 2008.

Pseudoprimes < 400,000

A list of Fermat pseudoprimes to the base 2. A dagger (†) indicates a Carmichael number and a double dagger (‡) indicates a strong pseudoprime to the base 2.

| | | | | | | | |
|----------|---------|----------|---------|--------|---------|---------|---------|
| 341 | †561 | 645 | †1105 | 1387 | †1729 | 1905 | ‡2047 |
| †2465 | 2701 | †2821 | ‡3277 | ‡4033 | 4369 | 4371 | ‡4681 |
| 5461 | †6601 | 7957 | ‡8321 | 8481 | †8911 | 10261 | †10585 |
| 11305 | 12801 | 13741 | 13747 | 13981 | 14491 | 15709 | ‡†15841 |
| 16705 | 18705 | 18721 | 19951 | 23001 | 23377 | 25761 | ‡†29341 |
| 30121 | 30889 | 31417 | 31609 | 31621 | 33153 | 34945 | 35333 |
| 39865 | †41041 | 41665 | ‡42799 | †46657 | ‡49141 | 49981 | ‡†52633 |
| 55245 | 57421 | 60701 | 60787 | †62745 | †63973 | 65077 | ‡65281 |
| 68101 | 72885 | ‡74665 | †75361 | ‡80581 | 83333 | 83665 | ‡85489 |
| 87249 | ‡88357 | 88561 | ‡90751 | 91001 | 93961 | †101101 | ‡104653 |
| 107185 | 113201 | †115921 | 121465 | 123251 | †126217 | 129889 | 129921 |
| ‡130561 | 137149 | 149281 | 150851 | 154101 | 157641 | 158369 | 162193 |
| †162401 | 164737 | †172081 | 176149 | 181901 | 188057 | †188461 | 194221 |
| 196021 | ‡196093 | 204001 | 206601 | 208465 | 212421 | 215265 | 215749 |
| 219781 | ‡220729 | 223345 | 226801 | 228241 | ‡233017 | 241001 | 249841 |
| ‡†252601 | ‡253241 | ‡256999 | 258511 | 264773 | 266305 | ‡271951 | 272251 |
| 275887 | 276013 | †278545 | ‡280601 | 282133 | 284581 | 285541 | 289941 |
| 294271 | †294409 | ‡†314821 | 318361 | 323713 | 332949 | †334153 | †340561 |
| 341497 | 348161 | ‡357761 | 367081 | 387731 | ‡390937 | 396271 | †399001 |

Primes < 4,000

| | | | | | | | | | | |
|------|------|------|------|------|------|------|------|------|------|------|
| 2 | 3 | 5 | 7 | 11 | 13 | 17 | 19 | 23 | 29 | 31 |
| 37 | 41 | 43 | 47 | 53 | 59 | 61 | 67 | 71 | 73 | 79 |
| 83 | 89 | 97 | 101 | 103 | 107 | 109 | 113 | 127 | 131 | 137 |
| 139 | 149 | 151 | 157 | 163 | 167 | 173 | 179 | 181 | 191 | 193 |
| 197 | 199 | 211 | 223 | 227 | 229 | 233 | 239 | 241 | 251 | 257 |
| 263 | 269 | 271 | 277 | 281 | 283 | 293 | 307 | 311 | 313 | 317 |
| 331 | 337 | 347 | 349 | 353 | 359 | 367 | 373 | 379 | 383 | 389 |
| 397 | 401 | 409 | 419 | 421 | 431 | 433 | 439 | 443 | 449 | 457 |
| 461 | 463 | 467 | 479 | 487 | 491 | 499 | 503 | 509 | 521 | 523 |
| 541 | 547 | 557 | 563 | 569 | 571 | 577 | 587 | 593 | 599 | 601 |
| 607 | 613 | 617 | 619 | 631 | 641 | 643 | 647 | 653 | 659 | 661 |
| 673 | 677 | 683 | 691 | 701 | 709 | 719 | 727 | 733 | 739 | 743 |
| 751 | 757 | 761 | 769 | 773 | 787 | 797 | 809 | 811 | 821 | 823 |
| 827 | 829 | 839 | 853 | 857 | 859 | 863 | 877 | 881 | 883 | 887 |
| 907 | 911 | 919 | 929 | 937 | 941 | 947 | 953 | 967 | 971 | 977 |
| 983 | 991 | 997 | 1009 | 1013 | 1019 | 1021 | 1031 | 1033 | 1039 | 1049 |
| 1051 | 1061 | 1063 | 1069 | 1087 | 1091 | 1093 | 1097 | 1103 | 1109 | 1117 |
| 1123 | 1129 | 1151 | 1153 | 1163 | 1171 | 1181 | 1187 | 1193 | 1201 | 1213 |
| 1217 | 1223 | 1229 | 1231 | 1237 | 1249 | 1259 | 1277 | 1279 | 1283 | 1289 |
| 1291 | 1297 | 1301 | 1303 | 1307 | 1319 | 1321 | 1327 | 1361 | 1367 | 1373 |
| 1381 | 1399 | 1409 | 1423 | 1427 | 1429 | 1433 | 1439 | 1447 | 1451 | 1453 |
| 1459 | 1471 | 1481 | 1483 | 1487 | 1489 | 1493 | 1499 | 1511 | 1523 | 1531 |
| 1543 | 1549 | 1553 | 1559 | 1567 | 1571 | 1579 | 1583 | 1597 | 1601 | 1607 |
| 1609 | 1613 | 1619 | 1621 | 1627 | 1637 | 1657 | 1663 | 1667 | 1669 | 1693 |
| 1697 | 1699 | 1709 | 1721 | 1723 | 1733 | 1741 | 1747 | 1753 | 1759 | 1777 |
| 1783 | 1787 | 1789 | 1801 | 1811 | 1823 | 1831 | 1847 | 1861 | 1867 | 1871 |
| 1873 | 1877 | 1879 | 1889 | 1901 | 1907 | 1913 | 1931 | 1933 | 1949 | 1951 |
| 1973 | 1979 | 1987 | 1993 | 1997 | 1999 | 2003 | 2011 | 2017 | 2027 | 2029 |
| 2039 | 2053 | 2063 | 2069 | 2081 | 2083 | 2087 | 2089 | 2099 | 2111 | 2113 |
| 2129 | 2131 | 2137 | 2141 | 2143 | 2153 | 2161 | 2179 | 2203 | 2207 | 2213 |
| 2221 | 2237 | 2239 | 2243 | 2251 | 2267 | 2269 | 2273 | 2281 | 2287 | 2293 |
| 2297 | 2309 | 2311 | 2333 | 2339 | 2341 | 2347 | 2351 | 2357 | 2371 | 2377 |
| 2381 | 2383 | 2389 | 2393 | 2399 | 2411 | 2417 | 2423 | 2437 | 2441 | 2447 |
| 2459 | 2467 | 2473 | 2477 | 2503 | 2521 | 2531 | 2539 | 2543 | 2549 | 2551 |
| 2557 | 2579 | 2591 | 2593 | 2609 | 2617 | 2621 | 2633 | 2647 | 2657 | 2659 |
| 2663 | 2671 | 2677 | 2683 | 2687 | 2689 | 2693 | 2699 | 2707 | 2711 | 2713 |
| 2719 | 2729 | 2731 | 2741 | 2749 | 2753 | 2767 | 2777 | 2789 | 2791 | 2797 |
| 2801 | 2803 | 2819 | 2833 | 2837 | 2843 | 2851 | 2857 | 2861 | 2879 | 2887 |
| 2897 | 2903 | 2909 | 2917 | 2927 | 2939 | 2953 | 2957 | 2963 | 2969 | 2971 |
| 2999 | 3001 | 3011 | 3019 | 3023 | 3037 | 3041 | 3049 | 3061 | 3067 | 3079 |
| 3083 | 3089 | 3109 | 3119 | 3121 | 3137 | 3163 | 3167 | 3169 | 3181 | 3187 |
| 3191 | 3203 | 3209 | 3217 | 3221 | 3229 | 3251 | 3253 | 3257 | 3259 | 3271 |
| 3299 | 3301 | 3307 | 3313 | 3319 | 3323 | 3329 | 3331 | 3343 | 3347 | 3359 |
| 3361 | 3371 | 3373 | 3389 | 3391 | 3407 | 3413 | 3433 | 3449 | 3457 | 3461 |
| 3463 | 3467 | 3469 | 3491 | 3499 | 3511 | 3517 | 3527 | 3529 | 3533 | 3539 |
| 3541 | 3547 | 3557 | 3559 | 3571 | 3581 | 3583 | 3593 | 3607 | 3613 | 3617 |
| 3623 | 3631 | 3637 | 3643 | 3659 | 3671 | 3673 | 3677 | 3691 | 3697 | 3701 |
| 3709 | 3719 | 3727 | 3733 | 3739 | 3761 | 3767 | 3769 | 3779 | 3793 | 3797 |
| 3803 | 3821 | 3823 | 3833 | 3847 | 3851 | 3853 | 3863 | 3877 | 3881 | 3889 |
| 3907 | 3911 | 3917 | 3919 | 3923 | 3929 | 3931 | 3943 | 3947 | 3967 | 3989 |