Chapter 5
Prime Test

In the RSA algorithm in the previous chapter, we see that the decomposition of large prime factors constitutes the basis of RSA cryptosystem security. Theoretically, this security should not be questioned, because there is only the definition of prime in mathematics, and there is no general method to detect prime. The main purpose of this chapter is to introduce some basic prime test methods, including Fermat test, Euler test, Monte Carlo method, continued fraction method, etc., understanding the content of this chapter requires some special number theory knowledge.

5.1 Fermat Test

According to Fermat's congruence theorem (commonly known as Fermat's small theorem, which is a special case of Euler congruence theorem), if n is a prime number, the following congruence formula holds for all integers b, $(b, n) = 1$,

$$b^{n-1} \equiv 1 (\bmod n). \tag{5.1}$$

The above formula is an important characteristic of prime numbers. Although n satisfying the above formula is not necessarily prime, it can be used as an important basis for detecting prime numbers, because we can conclude that n not satisfying the above formula is definitely not a prime number. Using Formula (5.1) as the standard to detect prime numbers is called Fermat test.

Definition 5.1 An odd number n, assuming that n is a compound number (not a prime number) and there is a positive integer b, $(b, n) = 1$, satisfying

$$b^{n-1} \equiv 1 (\bmod n),$$

Z. Zheng, *Modern Cryptography Volume 1*, Financial Mathematics and Fintech,
https://doi.org/10.1007/978-981-19-0920-7_5

the compound number n is called a Fermat pseudo prime under base b.

The basic properties of pseudo prime numbers are discussed. Our working platform is a finite Abel group $\mathbb{Z}_n^*$, define as

$$\mathbb{Z}_n^* = \{\bar{a} | 1 \leq a \leq n, (a, n) = 1\}, \ n > 1, \tag{5.2}$$

where $\bar{a}$ is a congruence class of mod n represented by a. The multiplication of two congruence classes is defined as $\bar{a} \cdot \bar{b} = \overline{ab}$; obviously, $\mathbb{Z}_n^*$ forms an Abel group of order $\varphi(n)$ under multiplication, in a finite group G, the order of a group element $g \in G$ is defined as

$$o(g) = \min\{m : g^m = 1, 1 \leq m \leq |G|\}.$$

$o(g) = 1$ if and only if g is the unit element of group G. By the definition of $o(g)$, obviously,

$$g^t = 1 \Leftrightarrow o(g)|t. \tag{5.3}$$

The following two lemmas are the basic conclusions about the order of group element g.

Lemma 5.1 *G is a finite group, $g \in G$, $k \in \mathbb{Z}$ is an integer, then*

$$o(g^k) = \frac{o(g)}{(k, o(g))}, \tag{5.4}$$

where the denominator is the greatest common divisor of k and $o(g)$.

Proof Let $o(g) = m$, $o(g^k) = t$, obviously, $(g^k)^m = 1$, in particular,

$$g^{\frac{k \cdot m}{(k,m)}} = 1, \Longrightarrow t \left| \frac{m}{(k, m)} \right. .$$

On the other hand, by $g^{kt} = 1$, there is $m|kt$, thus

$$\frac{m}{(k, m)} \left| \frac{k}{(k, m)} t, \Longrightarrow \frac{m}{(k, m)} \right| t.$$

So we have $t = \frac{m}{(k,m)}$, the Lemma holds.

Lemma 5.2 *Suppose G is a finite Abel group, $a, b \in G$, $(o(a), o(b)) = 1$, then*

$$o(ab) = o(a)o(b).$$

Proof Let $o(a) = m_1, o(b) = m_2$, then $(m_1, m_2) = 1$. Let $o(ab) = t$, by $(ab)^{m_1 m_2} = a^{m_1 m_2} b^{m_1 m_2} = 1$, there is $t | m_1 m_2$, on the other hand, $(ab)^t = 1$, then $(ab)^{t m_1} = 1$, thus

$b^{tm_1} = 1, m_2|m_1t, m_2|t$. By the same reason, there is $m_1|t$, thus $m_1m_2|t, t = m_1m_2$. The Lemma holds.

Back to the finite group $\mathbb{Z}_n^*$, any integer $a \in \mathbb{Z}$, $(a, n) = 1$, then $\bar{a} \in \mathbb{Z}_n^*$, we denote $o(\bar{a})$ with $o(a)$, a is called the order mod n, obviously, $o(a) = o(b)$, if $a \equiv b(\text{mod}\, n)$. A basic problem in number theory is the existence of primitive roots of mod n. equivalently, is $\mathbb{Z}_n^*$ a cyclic group? If there is a positive integer a, $(a, n) = 1$, $o(\bar{a}) = |\mathbb{Z}_n^*| = \varphi(n)$, then $\mathbb{Z}_n^*$ is a cyclic group of order $\varphi(n)$, so that the primitive root of mod n exists and a is the primitive root of mod n.

Lemma 5.3 (Existence of primitive root) *If and only if* $n = 2, 4, p^\alpha(\alpha \geq 1)$ *and* $a = 2p^\alpha(\alpha \geq 1)$ *four cases, the primitive root of* mod n *exists, where* $p > 2$ *is an odd prime.*

Proof If $n = 2, 4$, then the lemma holds. If $n = p$, then $\mathbb{Z}_n = \mathbb{F}_p$, $\mathbb{Z}_n^* = \mathbb{F}_p^*$, by Lemma 4.7 of Chap. 4, it can be seen that $\mathbb{F}_p^*$ is a cyclic group of order $(p - 1)$, so mod p has primitive roots. Now, we need to prove for all positive integer α, the primitive root of mod p^α also exists. Therefore, let a be a primitive root of mod p, that is, the order of a mod p is $p - 1$. If the order of a mod p^α is denoted by $o(a)$, then

$$a^{o(a)} \equiv 1(\text{mod}\, p^\alpha), \Longrightarrow a^{o(a)} \equiv 1(\text{mod}\, p),$$

so there is $p - 1|o(a)$. And the number of elements of $\mathbb{Z}_{p^\alpha}^*$ is $\varphi(p^\alpha) = p^{\alpha-1}(p - 1)$, obviously, $o(a)|p^{\alpha-1}(p - 1)$, thus, $o(a) = p^i(p - 1), 0 \leq i \leq \alpha - 1$.

We might as well let $o(a) = p - 1$, if $o(a) = p^i(p - 1)$,$1 \leq i$, then replace a with a^{p^i}. By Lemma 5.1,

$$o(a^{p^i}) = \frac{p^i(p - 1)}{(p^i, p^i(p - 1))}.$$

Therefore, without losing generality, let $o(a) = p - 1$, then by Sylow theorem, when $\alpha > 1$, $p^{\alpha-a}|\varphi(p^\alpha)$, there is an integer b, $(b, n) = 1$, b is $o(n) = p^{\alpha-1}$ in the order of mod p^α, because of $(o(a), o(b)) = 1$, then by Lemma 5.2, there is

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

So the primitive root of mod p^α exists.

When $n = 2p^\alpha$, $p > 2$ is odd prime, then $\varphi(n) = \varphi(p^\alpha)$. Thus, the primitive root a of mod p^α is also an primitive root of mod $2p^\alpha$. The Lemma holds.

Lemma 5.4 *Let n be an odd compound number, then*

(i) $b \geq 1$ *is a positive integer,* $(b, n) = 1$, *n is Fermat pseudo prime under base b if and only if* $o(b)|n - 1$.
(ii) *n is Fermat pseudo prime under bases* b_1 *and* b_2, *then it is Fermat pseudo prime under bases* b_1b_2 *and* $b_1b_2^{-1}$, *where* b_2^{-1} *is the multiplicative inverse of* b_2 mod n.

(iii) *If exist one $b \in \mathbb{Z}_n^*$ does not satisfy Eq. (5.1), at least half of a, $b \in \mathbb{Z}_n^*$ do not satisfy Eq. (5.1).*

Proof (i) and (ii) are trivial. (i) can be obtained by (5.3). And $b_1, b_2 \in \mathbb{Z}_n^*$,

$$\begin{cases} b_1^{n-1} \equiv 1 (\bmod n), b_2^{n-1} \equiv 1 (\bmod n). \Longrightarrow (b_1 b_2)^{n-1} \equiv 1 (\bmod n). \\ b^{n-1} \equiv 1 (\bmod n), \Longrightarrow (b^{-1})^{n-1} \equiv 1 (\bmod n). \end{cases}$$

So there is (ii). To prove (iii). Let n not be Fermat pseudo prime to base b, if n is Fermat pseudo prime to base a, then n is not Fermat pseudo prime to base ab. By (ii), therefore, if there is a base to make n a Fermat pseudo prime number, there must be a base to make n not a Fermat pseudo prime number, so more than half of the base b must make n not a Fermat pseudo prime number. The Lemma holds.

By Lemma 5.3, if there is a base b so that n is not Fermat pseudo prime, detect a, $1 \leq a \leq n$, $(a, n) = 1$ in sequence, whether $a^{n-1} \equiv 1 (\bmod n)$; that is, there is more than 50% chance that find the exact b such that $b^{n-1} \not\equiv 1 (\bmod n)$, this proves that n is not a prime number. Is it possible that all a, $1 \leq a \leq n$, $(a, n) = 1$, n is Fermat pseudo prime to base a The answer is yes, such a number n is called Carmichael number.

Definition 5.2 A Carmichael number n is an odd compound number, and for $\forall$ $b \in \mathbb{Z}_n^*$, there is

$$b^{n-1} \equiv 1 (\bmod n).$$

For Carmichael number, we have the following engraving.

Theorem 5.1 *Let n be a compound number, then*

(i) *If there is an integer $a > 1$, $a^2 | n$, then n is not a Carmichael number.*
(ii) *Assuming that n is a square free number, then n is a Carmichael number $\Leftrightarrow$ for all prime p, $p|n$, there is $p - 1|n - 1$.*
(iii) *A Carmichael number is the product of at least three different prime numbers.*

Proof Let's prove (i) first. Let $p^2|n$, p be a prime number, by Lemma 5.3, mod p^2 has primitive roots. Let g be an original root of mod p^2, that is $o(g) = p(p - 1)$, let

$$n' = \prod_{p'|n, p' \neq p} p', \ \ p' \text{ is a prime number.}$$

According to the Chinese remainder theorem, there is a positive integer b such that

$$\begin{cases} b \equiv g (\bmod p^2), \\ b \equiv 1 (\bmod n'). \end{cases}$$

Then b is an primitive root of mod p^2, and $(b, n) = 1$. We assert that n to base b is not a Fermat pseudo prime. If n to base b is a Fermat pseudo prime, then

$$b^{n-1} \equiv 1(\bmod n), \Longrightarrow b^{n-1} \equiv 1(\bmod p^2), \Longrightarrow o(b)|n-1.$$

That is $p(p-1)|n-1$, but $p|n$ is contradict with $p|n-1$. So $b^{n-1} \not\equiv 1(\bmod n)$, n is not Carmichael number, (i) holds.

Now to prove (ii). If $\forall\ p,\ p|n$, there is $p-1|n-1$, then $\forall\ b \in \mathbb{Z}_n^*$,

$$b^{n-1} = (b^{\frac{n-1}{p-1}})^{p-1} \equiv 1(\bmod p), \forall\ p|n.$$

Because n is a square free number, so

$$b^{n-1} \equiv 1(\bmod n),\ \forall\ b \in \mathbb{Z}_n^*.$$

Therefore, n is the Carmichael number. Conversely, if there is a prime number p, $p|n$, but $p-1 \nmid n-1$, Let g be a primitive root of mod p, which is given by the Chinese remainder theorem,

$$\begin{cases} b \equiv g(\bmod p), \\ b \equiv 1\left(\bmod \dfrac{n}{p}\right). \end{cases}$$

Then $(b, n) = 1$, and

$$b^{p-1} \equiv g^{p-1} \equiv 1(\bmod p).$$

By $p-1 \nmid n-1$, then $g^{n-1} \not\equiv 1(\bmod p)$, so there is $b^{n-1} \not\equiv 1(\bmod n)$, this contradicts with the assumption that n is the Carmichael number. So (ii) holds.

To prove (iii), we just need to exclude that n is the product of two prime numbers. By (ii), let $n = pq,\ p < q$, if n is a Carmichael number, then $q-1 \mid n-1$, but $n-1 = p(q-1+1)-1 = p(q-1)+p-1$, then

$$n-1 \equiv p-1(\bmod q-1),$$

this contradicts with $n-1 \equiv 0(\bmod q-1)$, so $n = pq$ must not be a Carmichael number, the Theorem holds.

Below we give some examples of Carmichael numbers, from property (ii) in Theorem 5.1, we can easily verify whether a square free number is Carmichael number.

Example 5.1 The following positive integers n are Carmichael numbers,

$$n = 1105 = 5 \cdot 13 \cdot 7, n = 1729 = 1 \cdot 13 \cdot 19, n = 2465 = 5 \cdot 17 \cdot 29,$$
$$n = 2821 = 7 \cdot 13 \cdot 31, n = 6601 = 7 \cdot 23 \cdot 41.$$

Example 5.2 The positive integer $561 = 3 \cdot 11 \cdot 17$ is the smallest Carmichael number.

Proof Defined by, the Carmichael number is odd and compound, so the minimum Carmichael number is

$$n = 3 \cdot p \cdot q, \text{ where } p - 1 | n - 1, q - 1 | n - 1, \ p < q \text{ is a prime.}$$

Let $p = 5$, $p = 7$, the congruence equation

$$3 \cdot p \cdot q \equiv 1 (\operatorname{mod} q - 1), q > p$$

has no prime solution q, when $p = 11$, the above formula has a minimum solution $q = 17$, so $n = 3 \cdot 11 \cdot 17$ is the smallest Carmichael number.

Example 5.3 For given prime number $r \geq 3$, then the congruence equations

$$\begin{cases} rpq \equiv 1 (\operatorname{mod} p - 1) \\ rpq \equiv 1 (\operatorname{mod} q - 1) \end{cases}$$

has only finite different prime solutions p, q. Let's leave this conclusion for reflection.

5.2 Euler Test

Let $p > 2$ be an odd prime, Euler test uses the Euler criterion in the quadratic residue of mod p to detect whether a positive integer n is prime. Like Fermat's test, it is obvious that the n that passes the test cannot be determined as prime, but the n that fails the test is certainly not prime. We know that when the positive integers a and n are given ($n > 1$), the solution of the quadratic congruence equation $x^2 \equiv a (\operatorname{mod} n)$ is a famous "NP complete" problem. We can't find a general solution in an effective time. However, in the special case where $n = p > 2$ is an odd prime number, we have rich theoretical knowledge to discuss the quadratic residue of mod p, these knowledge include the famous Gauss quadratic reciprocal law and Euler criterion, which constitute the core knowledge system of elementary number theory. First, we introduce Legendre sign and let $p > 2$ be a given odd prime number.

$\mathbb{Z}_p^*$ is a $(p - 1)$-order cyclic group, $a \in \mathbb{Z}_p^*$ (i.e., $(a, p) = 1$), we define the Legendre symbolic function as

$$\left(\frac{a}{p}\right) = \begin{cases} 1, & \text{when } x^2 \equiv a (\operatorname{mod} p) \text{ is solvable} \\ -1, & \text{when } x^2 \equiv a (\operatorname{mod} p) \text{ is unsolvable} \end{cases}$$

If $(a, p) > 1$, that is $p \mid a$, we let $(\frac{a}{p}) = 0$, for $\forall\, a \in \mathbb{Z}$, Legendre symbolic function $(\frac{a}{p})$ is all defined, and it is a completely integral function of $\mathbb{Z} \to \{1, -1, 0\}$.

$$\left(\frac{ab}{p}\right)=\left(\frac{a}{p}\right)\left(\frac{b}{p}\right), \forall\, a, b \in \mathbb{Z}$$

and

$$\left(\frac{a}{p}\right)=\left(\frac{b}{p}\right), \text{ if } a \equiv b\,(\bmod\, p)\,.$$

If $(\frac{a}{p})=1$, then $x^2 \equiv a(\bmod\, p)$ is solvable, a is called a quadratic residue of mod p, if $(\frac{a}{p})=-1$, then $x^2 \equiv a(\bmod\, p)$ is unsolvable, a is called a quadratic nonresidue of mod p.

Lemma 5.5 *$a \in \mathbb{Z}$, $p \nmid a$, then the necessary and sufficient condition for a to be the quadratic residue of* mod p *is*

$$a^{\frac{p-1}{2}} \equiv 1(\bmod\, p).$$

Proof $\mathbb{Z}_p^*$ is a $p-1$-order cyclic group, let g be a primitive root of mod p, that is $\bar{g}$ is the generator of $\mathbb{Z}_p^*$, that is $\forall a \in \mathbb{Z}$, $(a, p)=1$, we have

$$a \equiv g^t(\bmod\, p), \text{ where } 1 \le t \le p-1.$$

Obviously, a is the quadratic residue of mod $p \Leftrightarrow t$ is even. Therefore, if t is even, then

$$a^{\frac{p-1}{2}} \equiv g^{\frac{t(p-1)}{2}} \equiv (g^{\frac{t}{2}})^{p-1} \equiv 1(\bmod\, p).$$

Conversely, if $a^{\frac{p-1}{2}} \equiv 1(\bmod\, p)$, then $o(a) \mid \frac{p-1}{2}$, and by Lemma 5.1, can calculate

$$o(a)=o(g^t)=\frac{p-1}{(t, p-1)}.$$

So

$$o(a) \mid \frac{p-1}{2} \Leftrightarrow 2|(t, p-1) \Leftrightarrow 2|t,$$

that is t is even, thus, a is a quadratic residue of mod p, the Lemma holds.

Lemma 5.6 (Euler criterion). *For $\forall\, a \in \mathbb{Z}$, we have*

$$a^{\frac{p-1}{2}} \equiv \left(\frac{a}{p}\right)(\bmod\, p). \tag{5.5}$$

Proof If $(a, p)>1$, that is $p|a$, the above formula holds. Might as well let $p \nmid a$. By Fermat congruence theorem $a^{p-1} \equiv 1(\bmod\, p)$, there is

$$(a^{\frac{p-1}{2}}+1)(a^{\frac{p-1}{2}}-1) \equiv 0(\bmod\, p).$$

Thus

$$a^{\frac{p-1}{2}} \equiv \pm 1 (\bmod p).$$

If $a^{\frac{p-1}{2}} \equiv 1 (\bmod p)$, by Lemma 5.5, then $(\frac{a}{p}) = 1$. If $a^{\frac{p-1}{2}} \equiv -1 (\bmod p)$, then $(\frac{a}{p}) = -1$. So (5.5) holds.

Definition 5.3 Suppose n is an odd compound number, if there is an integer b, $(b, n) = 1$, it satisfies

$$b^{\frac{n-1}{2}} \equiv \left(\frac{b}{n}\right) (\bmod n), \tag{5.6}$$

Call n an Euler pseudo prime under base b. Where $(\frac{b}{n})$ is Jacobi symbol, define as

$$\left(\frac{b}{n}\right) = \left(\frac{b}{p_1}\right)^{\alpha_1} \left(\frac{b}{p_2}\right)^{\alpha_2} \cdots \left(\frac{b}{p_s}\right)^{\alpha_s}, \text{ if } n = p_1^{\alpha_1} \cdots p_s^{\alpha_s}. \tag{5.7}$$

From the definition, we obviously have a corollary: if n is Euler pseudo prime under basis b, then n is Fermat pseudo prime under basis b. This conclusion can be proved by squaring both sides of Eq. (5.6) at the same time.

The following example shows that the inverse of inference is not tenable; that is, if n is Fermat pseudo prime under basis b, but not Euler pseudo prime.

Example 5.4 $n = 91$ is Fermat pseudo prime under basis $b = 3$, but not Euler pseudo prime. In fact, it's easy to calculate $3^6 \equiv 1 (\bmod 91)$, thus $3^{90} \equiv 1 (\bmod 91)$. From $3^6 \equiv 1 (\bmod 91)$, we have

$$3^{42} \equiv 1 (\bmod 91), \Longrightarrow 3^{45} \equiv 9 (\bmod 91).$$

So 91 to base 3 is not an Euler pseudo prime.

Example 5.5 $n = 91$ to base $b = 10$ is an Euler pseudo prime. Because

$$10^{45} \equiv 10^3 \equiv -1 (\bmod 91),$$

calculate Legendre symbols

$$\left(\frac{10}{91}\right) = \left(\frac{2}{91}\right) \cdot \left(\frac{5}{91}\right) = -1,$$

so $n = 91$ to base $b = 10$ is an Euler pseudo prime.

From the Euler criterion of Lemma 5.6, we can easily calculate the Legendre symbols of -1 and 2.

Lemma 5.7 *Let $p > 2$ be an odd prime, then we have*

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}, \ \left(\frac{2}{p}\right) = (-1)^{\frac{1}{8}(p^2-1)}. \tag{5.8}$$

Proof By Lemma 5.6,

$$(-1)^{\frac{p-1}{2}} \equiv \left(\frac{-1}{p}\right) (\text{mod } p),$$

Since both sides of the congruence are ±1, $p > 2$, there is $(\frac{-1}{p}) = (-1)^{\frac{p-1}{2}}$. To calculate the Legendre sign for 2, we notice that

$$\begin{cases} p - 1 \equiv (-1)^1 (\text{mod } p) \\ 2 \equiv 2 \cdot (-1)^2 (\text{mod } p) \\ p - 3 \equiv 3 \cdot (-1)^3 (\text{mod } p) \\ \vdots \\ r \equiv \dfrac{p-1}{2} \cdot (-1)^{\frac{p-1}{2}} (\text{mod } p), \end{cases}$$

where $r = \frac{p-1}{2}$, if $\frac{p-1}{2}$ is a even; $r = p - \frac{p-1}{2}$, if $\frac{p-1}{2}$ is an odd. There is

$$2 \cdot 4 \cdot 6 \cdots (p-1) \equiv \left(\frac{p-1}{2}\right)! (-1)^{\frac{1}{8}(p^2-1)} (\text{mod } p),$$

that is

$$2^{\frac{p-1}{2}} \equiv (-1)^{\frac{1}{8}(p^2-1)} (\text{mod } p),$$

by Lemma 5.6,

$$\left(\frac{2}{p}\right) \equiv (-1)^{\frac{1}{8}(p^2-1)} (\text{mod } p),$$

there is

$$\left(\frac{2}{p}\right) = (-1)^{\frac{1}{8}(p^2-1)},$$

Lemma 5.7 holds.

Let $(\frac{a}{n})$ be a Jacobi symbol, defined by Eq. (5.6), then Lemma 5.7 can be extended to Jacobi symbol.

Lemma 5.8 *Let n be an odd, then we have*

$$\left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}}, \ \left(\frac{2}{n}\right) = (-1)^{\frac{1}{8}(n^2-1)}. \tag{5.9}$$

Proof The square of any odd number is congruent 1 under mod 8, that is $a^2 \equiv 1 \pmod 8$. Write $n = a^2 \cdot p_1 p_2 \cdots p_t$, where p_i are different prime numbers, then

$$n \equiv p_1 p_2 \cdots p_t \pmod 8.$$

Similarly, for $\forall\, n \in \mathbb{Z}$, by (5.7),

$$\left(\frac{b}{n}\right) = \left(\frac{b}{p_1}\right)\left(\frac{b}{p_2}\right)\cdots\left(\frac{b}{p_t}\right), \tag{5.10}$$

thus

$$\left(\frac{-1}{n}\right) = \left(\frac{-1}{p_1}\right)\left(\frac{-1}{p_2}\right)\cdots\left(\frac{-1}{p_t}\right) = (-1)^{\frac{p_1-1}{2}+\frac{p_2-1}{2}+\cdots+\frac{p_t-1}{2}} = (-1)^{\frac{n-1}{2}}. \tag{5.11}$$

The same can be proved $\left(\frac{2}{n}\right)$, the Lemma holds.

Corollary 5.1 *For all odd numbers n, they are Euler pseudo prime under the base* ± 1.

Proof It is trivial that n to 1 is an Euler pseudo prime number, and n to -1 is an Euler pseudo prime number, which is directly derived from Lemma 5.8.

Lemma 5.9 (Gauss.) *Let p and q be two different odd primes, then*

$$\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{\frac{1}{4}(p-1)(q-1)}.$$

Proof According to incomplete statistics, there are currently more than 270 methods to prove Gauss quadratic reciprocal law. In order to save space, we leave the proof to the readers, hoping that everyone can find their favorite proof method.

Next, we discuss the computational complexity of Fermat test and Euler test.

Lemma 5.10 *Let n be an odd,* $1 \le b < n$, $(b, n) = 1$, *then*

$$\begin{cases} \text{Time}(n \text{ to base } b'\text{s Fermat test}) = O(\log^3 n), \\ \text{Time}(n \text{ to base } b'\text{s Euler test}) = O(\log^4 n). \end{cases}$$

Proof By (5.1), the Fermat test of n to base b is actually an operation of b^{n-1} to mod n, by the Lemma 1.5 of Chap. 1, bit operations of $b^{n-1} \bmod n$,

$$\text{Time}(b^{n-1} \bmod n) = O(\log n \log^2 n) = O(\log^3 n).$$

Euler test of n to base b, by (5.6), the number of bit operations on the left is $O(\log^3 n)$. Find Jacobi symbol $(\frac{b}{n})$, from Eq. (5.7) and quadratic reciprocal law, the calculation can be transformed into the calculation of Legendre symbol. Each reciprocal law is actually a division, so we only consider the calculation of Legendre symbols. By Euler criterion,

$$\text{Time}\left(\text{calculate } \left(\frac{b}{p}\right)\right) = \text{Time}\left(b^{\frac{p-1}{2}} \bmod p\right) = O(\log^3 n).$$

The number of prime factors of each n has an estimated $O(\log \log n)$, so

$$\text{Time}\left(\text{calculate Jacobi symbol } \left(\frac{b}{n}\right)\right) = O(\log \log n \cdot \log^3 n) = O(\log^4 n).$$

We have completed the calculation of Lemma 5.10.

Solovay and Strassen proposed a probabilistic method to detect prime numbers by Euler test in 1977. When $n > 1$ is an odd number, k numbers are randomly selected, $b_1, b_2, \ldots, b_k$, where $1 < b_i < n$, $(b_i, n) = 1$. Use Eq. (5.6) to calculate both sides of each b in turn, and the required bit operation is $O(\log^4 n)$, if both sides of Eq. (5.6) are not equal, then n is not a prime number and the test is terminated. If k b pass the Euler test of Eq. (5.6), then n is the probability $< \frac{1}{2^k}$ of compound number, that is

$$P\{n \text{ is not prime}\} \le 2^{-k}.$$

The above formula is directly derived from Lemma 5.3. Let's introduce a better Miller–Rabin method than Solovay–Strassen method in a sense.

Definition 5.4 Let n be an odd compound number, write $n - 1 = 2^t \cdot m$, where $t \ge 1$, m is an odd. Let $b \in \mathbb{Z}_n^*$, if n and b satisfy one of the following conditions,

$$b^m \equiv 1 (\bmod n), \text{ or exists one } r, 0 \le r < t, \text{ such that } b^{2^r m} \equiv -1 (\bmod n). \tag{5.12}$$

Then n is called a strong pseudo prime under base b.

Lemma 5.11 *Suppose $n \equiv 3 (\bmod 4)$, then n is a strong Pseudoprime under base b if and only if n is an Euler Pseudoprime under base b.*

Proof Because $n \equiv 3 (\bmod 4)$, then $n - 1 = 2m$, that is $t = 1$, $m = \frac{1}{2}(n-1)$. By Definition 5.4, n is a strong pseudo prime under base b if and only if

$$b^m = b^{\frac{n-1}{2}} \equiv \pm 1 (\bmod n).$$

Therefore, if n is an Euler pseudo prime number under base b, the above formula holds, so it is also a strong pseudo prime number for base b. Conversely, if the above formula holds, because of $n \equiv 3 (\bmod 4)$, then $\frac{1}{2}(n-1)$ is an odd number, so $(\frac{-1}{n}) = -1$, and

$$\left(\frac{b}{n}\right) = \left(\frac{b}{n}\right)^{\frac{n-1}{2}} \equiv \left(\frac{b^{\frac{n-1}{2}}}{n}\right) \equiv b^{\frac{n-1}{2}} (\bmod n).$$

Therefore, n to base b is Euler pseudo prime. The Lemma holds.

Below we give the main results of this section.

Theorem 5.2 *Let n be an odd number, $b \in \mathbb{Z}_n^*$, then*

(i) If n to base b is a strong pseudo prime, then n to base b is an Euler pseudo prime.
(ii) Base b, which makes n a strong pseudo prime number, accounts for 25% *of* $1 \leq b < n$, $(b, n) = 1$ *at most.*

Before proving Theorem 5.2, let's introduce Miller–Rabin's test method, in order to test whether a large odd number n is a prime number, we write $n - 1 = 2^t \cdot m$, m is an odd number, $t \geq 1$, select one b at random, $1 \leq b < n$, $(b, n) = 1$. We first calculate $b^m \bmod n$, if we get the result is ± 1, then n passes the strong pseudo prime test (5.12). If $b^m \bmod n \neq \pm 1$, then we square $b^m \bmod n$ and find the minimum nonnegative residue of the squared number under $\bmod n$ to see if we get the result of -1 and perform r times. If we can't get -1, then n to base b fails to test Formula (5.12). Therefore, it is asserted that n to base b is not a strong pseudo prime number. If -1 is obtained by r squared, then n passes the test under base b.

In Miller–Rabin's test, if n to base b fails to pass the test Formula (5.12), then n must not be a prime number, if n to randomly selected k $b = \{b_1, b_2, \dots, b_k\}$ pass the test, by property (ii) of 5.2, each b_i accounts for no more than 25

$$P\{n \text{ not prime}\} \leq \frac{1}{4^k}. \tag{5.13}$$

Compared with the Solovay–Strassen method using Euler test, the Miller–Rabin method using strong pseudo prime test is more powerful.

To prove 5.2, we first prove the following two lemmas.

Lemma 5.12 *Let $G = \langle g \rangle$ be a finite group of order m, that is $o(g) = m$, then equation $x^k = 1$ has exactly d solutions in G, $d = (k, m)$.*

Proof $x \in G$, write $x = g^t$, then $x^k = g^{kt} = 1 \Leftrightarrow m|kt$, that is $\frac{m}{d}|\frac{k}{d} \cdot t$, thus $\frac{m}{d}|t$, let $t = \frac{m}{d} \cdot s$, then when $s = 1, 2, \dots, d$, $x = g^t$ has exactly d solutions. The Lemma holds.

Lemma 5.13 *Let p be an odd prime number, $p - 1 = 2^t m'$, $t \geq 1$, m' is prime, then*

$$x^{2^r m} \equiv -1 (\bmod p), m \text{ is odd} \tag{5.14}$$

The number of solutions N in $\mathbb{Z}_p^*$ satisfies

$$N = \begin{cases} 0, & \text{if } r \geq t; \\ 2^r(m, m'), & \text{if } r < t. \end{cases}$$

Proof Let g be a generator of $\mathbb{Z}_p^*$, write $x = g^j$, $1 \leq j \leq p-1$, because $o(g) = p-1$, so

$$g^{\frac{p-1}{2}} \equiv -1 (\bmod p).$$

Thus

$$x^{2^r m} \equiv -1 (\bmod p) \Leftrightarrow 2^r mj \equiv \frac{p-1}{2} (\bmod p-1).$$

Namely,

$$2^r mj \equiv 0 (\bmod p-1).$$

Because $p - 1 = 2^t m'$, the above formula is equivalent to

$$2^r mj \equiv 2^{t-1} m' (\bmod 2^t m'). \tag{5.15}$$

If $r > t - 1$, then the congruence has no solution to j, because m and m' are odd numbers, so when $r \geq t$, (5.14) is unsolvable. If $r < t$, let $d = (m, m')$, then

$$(2^r m, 2^t m') = 2^r d,$$

then Eq. (5.15) has exactly d solutions for j. Each j corresponds to one $x = g^j$, then the number of solutions of Eq. (5.14) to x is $N = 2^r d$, the Lemma holds.

With the above preparation, we now give the proof of Theorem 5.2.

Proof (*The proof of Theorem* 5.2). Let's first prove that (i), that is, n and b satisfy Eq. (5.12), we want to prove that formula (5.6) is satisfied; that is, if n to base b is a strong pseudo prime number, then n to base b is an Euler pseudo prime number, write $n - 1 = 2^t m$, m is prime, we prove the property (i) of Theorem 5.2 in three cases.

(1) $b^m \equiv 1 (\bmod n)$. In this case, it is obvious that $b^{\frac{n-1}{2}} \equiv 1 (\bmod n)$. Let's prove $(\frac{b}{n}) = 1$, in fact,

$$1 = \left(\frac{1}{p}\right) = \left(\frac{b^m}{p}\right) = \left(\frac{b}{p}\right)^m = 1.$$

There is

$$b^{\frac{n-1}{2}} \equiv \left(\frac{b}{n}\right) \equiv 1 (\bmod n).$$

That is n to base b is an Euler pseudo prime number.

(2) $b^{\frac{n-1}{2}} \equiv -1 (\bmod n)$. In this case, we have to prove $(\frac{b}{n}) = -1$, let $p | n$ be any prime factor of n, write $p - 1 = 2^{t_1} m_1$, where $t_1 \geq 1$, m_1 is an odd number.

Let’s calculate the Legendre symbol $(\frac{b}{p})$, in fact, $t_1 \geq t$, and

$$\left(\frac{b}{p}\right) = \begin{cases} -1, & \text{if } t_1 = t; \\ 1, & \text{if } t_1 > t. \end{cases} \tag{5.16}$$

Because

$$b^{\frac{n-1}{2}} = b^{2^{t-1}m} \equiv -1(\text{mod}\, n), \Longrightarrow b^{2^{t-1}mm_1} \equiv -1(\text{mod}\, n),$$

by $p|n$, we have

$$b^{2^{t-1}mm_1} \equiv -1(\text{mod}\, p). \tag{5.17}$$

If $t_1 < t$, from the above formula, there is

$$b^{2^{t_1}m_1} \equiv -1(\text{mod}\, p), \Longrightarrow b^{p-1} \equiv -1(\text{mod}\, p).$$

This contradicts Fermat’s congruence theorem, so we always have $t_1 \geq t$. If $t_1 = t$, by (5.17), then

$$\left(\frac{b}{p}\right) \equiv b^{\frac{p-1}{2}} = b^{2^{t-1}m} \equiv -1(\text{mod}\, p).$$

Because if the above formula is 1, both sides will be m power at the same time, which will contradict Formula (5.17). If $t_1 > t$, put both sides of Eq. (5.17) to the power of 2^{t_1-t} at the same time, then $(\frac{b}{p}) = 1$, so we have (5.16).
We now complete the proof of case (2) under the conclusion of Eq. (5.16), write $n = \prod_{p|n} p$, p does not require different, define the positive integer k as

$$k = \#\{p \mid p|n,\, p-1 = 2^{t_1}m_1,\, m_1 \text{ is odd},\, t_1 = t\}.$$

By (5.16), then

$$\left(\frac{b}{n}\right) = \prod \left(\frac{b}{p}\right) = (-1)^k. \tag{5.18}$$

Let’s prove that k is an odd number, because $t_1 \geq t$, $p-1 = 2^{t_1}m_1$, $n-1 = 2^t m$, under mod 2^{t+1}, we have

$$p \equiv \begin{cases} 1(\text{mod}\, 2^{t+1}), & \text{if } t_1 > t; \\ 1 + 2^t(\text{mod}\, 2^{t+1}), & \text{if } t_1 = t. \end{cases}$$

Because $n = 1 + 2^t(\text{mod}\, 2^{t+1})$, so

$$n \equiv 1 + 2^t \equiv 1 + k \cdot 2^t(\text{mod}\, 2^{t+1}),$$

So k must be odd, by (5.18), then $(\frac{b}{n}) = -1$. Case (2) is proved.

(3) $b^{2^{r-1}\cdot m} \equiv -1(\bmod n)$, where $1 \leq r \leq t$, $n-1 = 2^t \cdot m$.
In this case, we replace r of Eq. (5.12) with $r-1$. Because $r-1 \leq t-1$, so $b^{\frac{n-1}{2}} \equiv 1(\bmod n)$. To prove property (i) of Theorem 5.2, we have to prove $(\frac{b}{n}) = 1$, as in case (2), we let $p|n$, write $p-1 = 2^{t_1} \cdot m_1$, m_1 is odd, then we have $t_1 \geq r$, and

$$\left(\frac{b}{p}\right) = \begin{cases} -1, & \text{if } t_1 = r; \\ 1, & \text{if } t_1 > r. \end{cases} \tag{5.19}$$

The proof of Formula (5.19) is the same as that of case (2), write $n = \prod p$, p is not required to be a different prime, define positive integer k_1:

$$k_1 = \#\{p \mid p|n,\ p-1 = 2^{t_1}m_1, m_1 \text{ is odd}, t_1 = r\}.$$

as in case (2), we have $(\frac{b}{n}) = (-1)^{k_1}$, similarly, under mod 2^{r+1}, it can be proved that k_1 must be even. Thus $(\frac{b}{n}) = 1$, we have completed all the proofs of property (i) in Theorem 5.2.
Next, we prove property (ii) in Theorem 5.2. It is also discussed in three cases.

(1) n can be divided by a square number; that is, there is a prime number p, $p^\alpha || n$, $\alpha \geq 2$.
In this case, we prove that there are at least $\frac{1}{4}(n-1)$ b, $b \in \mathbb{Z}_n^*$, n to base b is not Fermat prime number, let alone a strong pseudo prime. First, suppose $b^{n-1} \equiv 1(\bmod n)$, then there is a prime p, $p^2|n$, thus $b^{n-1} \equiv 1(\bmod p^2)$. Because $\mathbb{Z}_{p^2}^*$ is a $p(p-1)$-order cyclic group (see Theorem 5.3), let g be a generator of $\mathbb{Z}_{p^2}^*$, then

$$\mathbb{Z}_{p^2}^* = \{g, g^2, \ldots, g^{p(p-1)}\}.$$

By Lemma 5.12, the number of b satisfying $b^{n-1} \equiv 1(\bmod p^2)$ is d,

$$d = (n-1, p(p-1)) = (n-1, p-1).$$

Because $p|n$, so $p \nmid n-1$, and $p \nmid d$; therefore, the maximum possibility of d is $p-1$; therefore, the proportion of b in $b^{n-1} \equiv 1(\bmod p^2)$ in $1 \leq b < n$ shall not exceed

$$\frac{p-1}{p^2-1} = \frac{1}{p+1} \leq \frac{1}{4}.$$

Therefore, there is at most b in the proportion of $\frac{1}{4}$, so that n to base b is Fermat prime, in case (1), we prove the property (ii) of Theorem 5.2.

(2) $n = pq$ are two different prime numbers.
In this case, let $p-1 = 2^{t_1}m_1$, $q-1 = 2^{t_2}m_2$, m_1, m_2 to be odd. Without losing generality, you can let $t_1 \leq t_2$. Let $b \in \mathbb{Z}_n^*$, in order for n to base b to be a strong pseudo prime number, it is necessary to satisfy

$$b^m \equiv 1(\bmod p),\ b^m \equiv 1(\bmod q) \tag{5.20}$$

or

$$b^{2^r m} \equiv -1(\bmod p),\ b^{2^r m} \equiv -1(\bmod q), 0 \le r < t. \tag{5.21}$$

By Lemma 5.12, the number of b satisfied (5.20) is $\le (m, m_1)(m, m_2) \le m_1 m_2$. By Lemma 5.13, for each r, $0 \le r < \min(t_1, t_2) = t_1$, the number of b satisfying $b^{2^r m} \equiv -1(\bmod n)$ is $2^r(m, m_1) \cdot 2^r(m, m_2) < 4^r m_1 m_2$. Because $n = pq$, then $\varphi(n) = (p-1)(q-1), \Longrightarrow n - 1 > \varphi(n) = 2^{t_1+t_2}$, therefore, the proportion of b of the strong pseudo prime of n to base b does not exceed

$$\frac{m_1m_2 + m_1m_2 + 4m_1m_2 + \cdots + 4^{t_1-1}m_1m_2}{2^{t_1+t_2}m_1m_2} = 2^{-t_1-t_2}\left(1 + \frac{4^{t_1}-1}{4-1}\right) \tag{5.22}$$

in $1 \le b < n$, $(b, n) = 1$.
If $t_1 < t_2$, then the above formula shall not exceed

$$2^{-2t_1-1}\left(\frac{2}{3} + \frac{4^{t_1}}{3}\right) \le 2^{-3} \cdot \frac{2}{3} + \frac{1}{6} = \frac{1}{4}.$$

If $t_1 = t_2$, then $m_1 \ne m_2$, so $(m, m_1) \le m_1$ and $(m, m_2) \le m_2$, one must be strictly less than. The reason is that if they are equal, then $m_1 | m$, $m_2 | m$, $n - 1 = 2^t m, \Longrightarrow n - 1 = 2^t m = pq - 1 \equiv q - 1(\bmod m_1)$, thus $m_1 | n - 1, \Longrightarrow m_1 | q - 1 = 2^{t_2} m_2, \Longrightarrow m_1 | m_2$, this is a contradiction. So $(m, m_1) \le m_1$ and $(m, m_2) \le m_2$ must have a strict less than 0. We have

$$(m, m_1) \cdot (m, m_2) \le \frac{1}{3} m_1 m_2.$$

If $m_1 m_2$ is substituted for $\frac{1}{3} m_1 m_2$ in Eq. (5.22), the proportion of n to b whose base b is a strong pseudo prime number does not exceed

$$\frac{1}{3} 2^{-2t_1}\left(\frac{2}{3} + \frac{4^{t_1}}{3}\right) \le \frac{1}{18} + \frac{1}{9} = \frac{1}{6} < \frac{1}{4}.$$

We complete the proof of property (ii) of Theorem 5.2 in case (2).

(3) Finally, suppose $n = p_1 p_2 \cdots p_k$, $k \ge 3$ is the product of different prime factors. In this case, write $p_i - 1 = 2^{t_i} m_i$, m_i as an odd number. As in case (2), without losing generality, it can make $t_1 \le t_j (1 \le j \le k)$. Similarly to the proof of formula (5.22), the proportion of b satisfying that n is a strong pseudo prime number for base b does not exceed

$$\begin{aligned}2^{-t_1-t_2-\cdots-t_k}\left(1+\frac{2^{k+1}-1}{2^k-1}\right) &\le 2^{-kt_1}\left(\frac{2^k-2}{2^k-1}+\frac{2^{kt_1}}{2^k-1}\right)\\ &= 2^{-kt_1}\cdot\frac{2^k-2}{2^k-1}+\frac{1}{2^k-1}\\ &\le 2^{-k}\frac{2^k-2}{2^k-1}+\frac{1}{2^k-1}\\ &= 2^{1-k}\\ &\le \frac{1}{4},\end{aligned}$$

because $k \ge 3$, in this way, we have completed all the proofs of Theorem 5.2.

Euler test and strong pseudo prime test require some complex quadratic residual techniques. We summarize the main conclusions of this section as follows:

(A) n to base b is a strong pseudo prime number $\Rightarrow$ n to base b is an Euler pseudo prime number $\Rightarrow$ n to base b is a Fermat pseudo prime number; therefore, the strong pseudo prime test is the best way to detect prime numbers.

(B) Although no test can successfully detect a prime number at present, the probability detection method of strong pseudo prime number test, that is, Miller–Rabin method, can obtain that the success probability (see (5.13)) of detecting whether any odd number n is a prime number can be infinitely close to 1. That is

$$P\{\text{detect whether odd } n \text{ is prime}\} > 1-\varepsilon, \forall\, \varepsilon > 0 \text{ given}.$$

Moreover, the computational complexity of the detection algorithm is polynomial.

5.3 Monte Carlo Method

Using all the prime number test methods introduced in the previous two sections, for a huge odd number n, even if we already know that n is not a prime number, we cannot successfully decompose n, because the prime number test does not provide prime factor decomposition information, A more direct method—like the sieve method—verifies whether the prime factor of n is for prime numbers not greater than $\sqrt{n}$, because a compound number n must have a prime factor p, $p \le \sqrt{n}$. Selected $p \le \sqrt{n}$, the bit operation required to divide n by p is $O(\log n)$, there are $O(\frac{\sqrt{n}}{\log n})$ prime numbers $p \le \sqrt{n}$ in total, therefore, the bit operation required for such a verification is $O(\sqrt{n})$. A more effective method was proposed by J. M. Pollard in 1975. We call it Monte Carlo method, or "rho" method.

First, find a convenient mapping f of $\mathbb{Z}_n \xrightarrow{f} \mathbb{Z}_n$; for example, $f(x)$ is an integer coefficient polynomial, such as $f(x) = x^2 + 1$; secondly, a prime number x_0 is randomly generated, let $x_1 = f(x_0)$, $x_2 = f(x_1)$, ..., $x_{j+1} = f(x_j)(j = 0, 1, 2, \cdots)$. In these x_j, we want to find two integers x_j and x_k, which are different elements in $\mathbb{Z}_n$, but there are some factors d of n, $d|n$, and x_j and x_k are the same elements in $\mathbb{Z}_d$, that is to say

$$x_j \not\equiv x_k (\bmod n),\ (x_j - x_k, n) > 1. \tag{5.23}$$

Once x_j and x_k are found, the algorithm is said to be completed.

Theorem 5.3 *Let S be a set of r elements, let $f : S \to S$ is a mapping, $x_0 \in S$, define $x_{j+1} = f(x_j)(j = 0, 1, 2, \ldots)$. Suppose λ is a positive real number, let $l = 1 + [\sqrt{2\lambda r}]$, then the condition $x_0, x_1, \ldots, x_l$ is the ratio $\leq e^{-\lambda}$ of the mapping f of elements in different s to the initial value x_0, (f, x_0), f in all mappings S and all $x_0 \in S$.*

Proof The total number of mappings f from $f : S \to S$ is r^r, because each $x \in S$, we can arrange r images for it, that is, $f(x)$ has r choices. The initial value x_0 has r choices, so the total number of (f, x_0) is r^{r+1}. The question is which of these (f, x_0) choices can satisfy the condition that $x_0, x_1, \ldots, x_l$ is a different element in S. we want to prove that the proportion of (f, x_0) satisfying the condition in r^{r+1} (f, x_0) is not greater than $\leq \mathrm{e}^{-\lambda}$.

When $x_0 \in S$ given, there are r x_0 choices, then $x_1 = f(x_0)$ has only $r - 1$ choices and $x_2 = f(x_1)$ has only $r - 2$ choices, this goes on until $x_l = f(x_{l-1})$, there are only $r - l$ options. The remaining $x \in S$ and f can be selected arbitrarily; that is, there are r^{r-l} choices. Therefore, when x_0 is given, there are N f to make (f, x_0) meet the required conditions, where

$$N = r^{r-l} \prod_{j=0}^{l} (r - j).$$

Divide N by r^{r+1}, and the proportion of (f, x_0) satisfying the condition is

$$\frac{N}{r^{r+1}} = r^{-l} \prod_{j=1}^{l} (r - j) = \prod_{j=1}^{l} \left(1 - \frac{j}{r}\right), \tag{5.24}$$

We notice that the real number $x \in (0, 1)$, then $\log(1 - x) < -x$. Take the logarithm to the right of the above formula, then

$$\sum_{j=1}^{l} \log\left(1 - \frac{j}{r}\right) < -\sum_{j=1}^{l} \frac{j}{r} = \frac{-l(l+1)}{2r} < -\frac{l^2}{2r}.$$

Because of $l = 1 + [\sqrt{2\lambda r}] > \sqrt{2\lambda r}$, from the above formula,

$$\sum_{j=1}^{l} \log\left(1-\frac{j}{r}\right) < -\lambda.$$

By (5.24), we have

$$\frac{N}{r^{r+1}} \leq \mathrm{e}^{-\lambda}.$$

We complete the proof of Theorem 5.3.

Monte Carlo method uses a polynomial $f(x) \in \mathbb{Z}[x]$, so that n is a positive integer, and the congruence equation of mod n is invariant to polynomial $f(x)$, that is

$$a \equiv b(\bmod n), \Longrightarrow f(a) \equiv f(b)(\bmod n). \tag{5.25}$$

$x_0 \in \mathbb{Z}_n$ given, $x_{j+1} = f(x_j)(j = 0, 1, \ldots)$, if you find an $x_{k_0} \in \mathbb{Z}_n$ that satisfies $x_{k_0} \equiv x_{j_0}(\bmod r)$, where $r|n$, $r > 1$, $k_0 > j_0$. By (5.25),

$$f(x_{k_0}) \equiv f(x_{j_0})(\bmod r), \Longrightarrow x_{k_0+1} \equiv x_{j_0+1}(\bmod r).$$

Thus for any $k > j$, if $k - j = k_0 - j_0$, there is $x_k \equiv x_j(\bmod r)$, this proves that a polynomial mapping $\mathbb{Z}_n \xrightarrow{f} \mathbb{Z}_n$ produces k_0 different residue classes under mod $r(r|n)$,

$$\{x_0, x_1, \ldots, x_{k_0-1}\}.$$

Therefore, there is the following Lemma 5.14.

Lemma 5.14 *$f(x) \in \mathbb{Z}[x]$ is a polynomial, $n > 1$ is an positive integer, let $x_0 \in \mathbb{Z}_n$, $x_j = f(x_{j-1})(j = 1, 2, \ldots)$, if k is the first subscript, there is a j, $0 \leq j < k$, such that*

$$(x_k - x_j, n) = r > 1.$$

Then $\{x_0, x_1, \ldots, x_{k-1}\}$ is k different residual classes under mod r*, so it is also k different residual classes under* mod n*. Moreover, Monte Carlo calculation defined by f can only produce k different residual classes.*

We call the polynomial f and the initial value x_0 described in Lemma 5.14 an average mapping. When the first subscript k is very large, the amount of calculation is very large. Here we give an improved Monte Carlo algorithm.

$f(x) \in \mathbb{Z}[x]$ given, Monte Carlo algorithm needs to continuously calculate $x_k(k = 1, 2, \ldots)$. Let $2^h \leq k < 2^{k+1}(h \geq 0)$, $j = 2^h - 1$; that is, k is an $(h + 1)$-bit number, j is the maximum h-bit number, compare x_k with x_j and calculate $(x_k - x_j, n)$, if $(x_k - x_j, n) > 1$, then the calculation is terminated, otherwise consider $k + 1$. The improved Monte Carlo algorithm only needs to calculate $(x_k - x_j, n)$ once for each k , $j = 2^h - 1$. There is no need to verify every j, $0 \leq j < k$, when k is very large, it reduces a lot of computation, but there is a disadvantage. It may miss the smallest subscript k satisfying the condition, but the error is controllable. In fact, we have the following error estimation.

Lemma 5.15 *$f(x) \in \mathbb{Z}[x]$, $n \geq 1$ given, $x_0 \in \mathbb{Z}_n$, $x_j = f(x_{j-1})(j = 1, 2, \ldots)$, let k_0 be the smallest subscript and satisfy $(x_{k_0} - x_{j_0}, n) > 1$, where $0 \leq j_0 < k_0$, assuming that k is the smallest positive integer satisfying $(x_k - x_j, n) > 1$ in the improved Monte Carlo algorithm, we have $k \leq 4k_0$.*

Proof Suppose k_0 has $(h + 1)$ bits. Let $j = 2^{h+1} - 1, k = j + (k_0 - j_0)$. By Lemma 5.14, then

$$(x_{k_0} - x_{j_0}, n) > 1, \Longrightarrow (x_k - x_j, n) > 1.$$

Obviously, j is the maximum number of $(h + 1)$ bits and k is the number of $(h + 2)$ bits, so k is the required subscript calculated by the improved Monte Carlo algorithm. Obviously,

$$k = j + (k_0 - j_0) \leq 2^{h+1} - 1 + 2^{h+1} < 4 \cdot 2^h \leq 4k_0.$$

Lemma 5.15 holds.

Example 5.6 Let $n = 91$, $f(x) = x^2 + 1$, $x_0 = 1$. By Monte Carlo algorithm, then $x_1 = 2, x_2 = 5, x_3 = 26$ and $x_4 = 40$ (because$26^2 + 1 \equiv 40 (\bmod 91)$). By the improved Monte Carlo algorithm, only $(x_4 - x_3, 91)$ needs to be detected to obtain

$$(x_4 - x_3, 91) = (14, 91) = 7.$$

Lemma 5.16 *Let n be an odd number and a compound number, and r be a factor of n, $r|n$, $1 < r < \sqrt{n}$. Let $f(x) \in \mathbb{Z}[x]$, $x_0 \in \mathbb{Z}_n$ given, then the computational complexity of finding r by Monte Carlo algorithm (f, x_0) is*

$$\text{Time}((f, x_0)) = O(\sqrt{n} \log^3 n) \text{ bits.} \tag{5.26}$$

Further, there is a normal number C, so that for any positive real number λ, the success probability of Monte Carlo algorithm (f, x_0) to find a nontrivial factor r of n is greater than $1 - \mathrm{e}^{-\lambda}$, that is

$$P\{(f, x_0)\text{find out } r|n, r > 1\} \geq 1 - \mathrm{e}^{-\lambda}. \tag{5.27}$$

The number of bit calculation operations required by the algorithm that depends on parameter λ (to ensure the success rate of the algorithm) is $O(\sqrt{\lambda}\sqrt[4]{n} \log^3 n)$.

Proof From the discussion of computational complexity in Chap. 1, finding the maximum common divisor of two integers and the addition, subtraction, multiplication and division in mod n are polynomial. Let C_1 satisfies

$$\text{Time}((y - z, n)) \leq C_1 \log^3 n, \text{ where } y, z \leq n.$$

C_2 satisfies

$$\text{Time}(f(x) \bmod n) \leq C_2 \log^3 n, x \in \mathbb{Z}_n.$$

If k_0 is (f, x_0), the first subscript in the calculation satisfies $(x_{k_0} - x_{j_0}, n) > 1$, by the improved Monte Carlo algorithm, we have $(x_k - x_j, n) > 1$, where $j = 2^h - 1$, $2^h \leq k < 2^{h+1}$. By Lemma 5.15, $k \leq 4k$. Thus

$$\text{Time}(\text{found by } (f, x_0)\ k) \leq 4k_0(C_1 \log^3 n + C_2 \log^3 n). \tag{5.28}$$

Let $(x_{k_0} - x_{j_0}, n) = r > 1, r < \sqrt{n}$, by Lemma 5.14, $k_0 \leq r$, so

$$\text{Time}(\text{find } r, r|n, r < \sqrt{n}) \leq 4\sqrt{n}(C_1 \log^3 n, C_2 \log^3 n).$$

Equation (5.26) proved. In the sense of probability, that is, on the premise of allowing certain errors, Eq. (5.26) can be further improved.

Let $\lambda > 0$ be any given real number, by Lemma 5.3, ratio of $k_0 \geq 1 + \sqrt{2\lambda r}$ $< \mathrm{e}^{-\lambda}$, in other words, the probability of successfully finding $r, r|n, r \leq \sqrt{n}$ is

$$P\{\text{find out } r, r|n, r < \sqrt{n}\} \geq 1 - \mathrm{e}^{-\lambda}.$$

In order to ensure the success rate, then $k_0 \leq 1 + \sqrt{2\lambda r}$. By (5.28), the number of bit operations required shall not be greater than

$$4(1 + \sqrt{2\lambda r})(C_1 \log^3 n + C_2 \log^3 n) = O(\sqrt{\lambda}\sqrt[4]{n} \log^3 n).$$

We have completed the proof of Lemma.

Remark 5.1 A basic assumption of Monte Carlo method is that the integer coefficient polynomial f can be used as an average mapping (see Lemma 5.14); this has not yet been proved.

5.4 Fermat Decomposition and Factor Basis Method

Lemma 5.17 *Suppose n is an odd number, there is a 1-1 correspondence between factorization $n = a \cdot b (a \geq b > 0)$ of n and expression $n = t^2 - s^2$ (t and s are nonnegative integers) of n. The corresponding $\sigma : (a, b) \to (t, s)$ can be written as $\sigma((a, b)) = (t, s)$, where*

$$\sigma((a, b)) = \left(\frac{a + b}{2}, \frac{a - b}{2}\right).$$

Inverse mapping is

$$\sigma^{-1}((t, s)) = (t + s, t - s).$$

Proof If $n = ab$, because both a and b are odd, then $n = (\frac{a+b}{2})^2 - (\frac{a-b}{2})^2$, so define

$$\sigma((a,b)) = \left(\frac{a+b}{2}, \frac{a-b}{2}\right).$$

Conversely, if $n = t^2 - s^2$, then $n = (t+s)(t-s)$. So define $\sigma^{-1}((t,s)) = (t+s, t-s)$, we prove $\sigma^{-1}\sigma = 1, \sigma\sigma^{-1} = 1$. By the definition,

$$\begin{cases} \sigma^{-1}\sigma((a,b)) = \sigma^{-1}\left(\frac{a+b}{2}, \frac{a-b}{2}\right) = (a,b), \\ \sigma(\sigma^{-1}((t,s))) = \sigma(t+s, t-s) = (t,s). \end{cases}$$

So σ is a 1-1 correspondence between the two decomposition $n = ab = t^2 - s^2$, the Lemma holds.

The above simple lemma provides us with a method of factor decomposition, called Fermat factor decomposition: if $n = ab$, a is very close to b, then $n = (\frac{a+b}{2})^2 + (\frac{a-b}{2})^2 = t^2 - s^2$, where s is very small and t is only a little larger than $\sqrt{n}$. Therefore, starting from $t = [\sqrt{n}] + 1$, we successively detect whether $t^2 - n$ is a complete square number. If not, we change it to $t = [\sqrt{n}] + 2$ for detection. In this way, until $t^2 - n = s^2$, we get $n = (t+s)(t-s)$ through Fermat factorization. This method is effective when $n = ab$, a and b are very close.

Fermat factor decomposition can be further expanded into a factor-based method to become a more effective factor decomposition method. Its basic idea is: in Fermat factorization, $t^2 - n^2$ is required to be a complete square, which is difficult to appear in practice, but $t^2 \equiv s^2 (\bmod n)$, $t \not\equiv \pm s (\bmod n)$ is easy to appear. Calculate the maximum common divisor $(t+s, n)$ and $(t-s, n)$, then we have factorization

$$n = (t+s, n)(t-s, n).$$

Definition 5.5 Let B be h different primes (maybe $p_1 = -1$), B is called a factor base. An integer b is called a B-number, if the minimum nonnegative residue of b^2 under $\bmod n$ can be expressed as the product of prime numbers in B, where n is the given positive integer.

Example 5.7 Let $n = 4633$, $B = \{-1, 2, 3\}$, then 67, 68, 69 are all B-number, because $67^2 \equiv -144 (\bmod 4633)$, $68^2 \equiv -9 (\bmod 4633)$, $69^2 \equiv 128 (\bmod 4633)$.

If b is a B-number, $b^2 \bmod n$ represents the minimum nonnegative residue of b^2 under $\bmod n$, by the definition,

$$b^2 \bmod n = \prod_{i=1}^{h} p_i^{\alpha_i}, \alpha_i \geq 0.$$

Let $e = \{e_1, e_2, \ldots, e_h\} \in \mathbb{F}_2^h$ be an h-dimensional binary vector, define

$$e_j = \begin{cases} 0, & \text{if } \alpha^j \text{ is even;} \\ 1, & \text{if } \alpha^j \text{ is odd.} \end{cases} \quad 1 \le j \le h.$$

e is called the binary vector corresponding to b if $\{b_i\} = A$ is a set of B-numbers. The binary vector corresponding to each b_i is denoted as $e_i = \{e_{i_1}, e_{i_2}, \dots, e_{i_h}\}$, denote $b_i^2 \bmod n$ with a_i. We have

$$\prod_{i \in A} a_i = \prod_{j=1}^{h} p_j^{\sum_{i \in A} \alpha_{ij}}, \text{ where } a_i = \prod_{j=1}^{h} p_j^{\alpha_{ij}},$$

Suppose $\sum_{i \in A} e_i = (0, 0, \dots, 0)$ is the zero vector in $\mathbb{F}_2^h$, then

$$\sum_{i \in A} \alpha_{ij} \equiv 0 (\bmod 2), \forall\, 1 \le j \le h.$$

That is, $\prod a_i$ is a square number. Let $r_j = \frac{1}{2} \sum_{i \in A} \alpha_{ij}$, then

$$\prod_{i \in A} a_i = \left(\prod_{j=1}^{h} p_j^{r_j} \right)^2, \text{ define } c = \prod_{j=1}^{h} p_j^{r_j}, \tag{5.29}$$

On the other hand, $b_i \bmod n$ represents the minimum nonnegative residue of b_i under $\bmod n$, let

$$b = \prod_{i \in A} (b_i \bmod n) = \prod_{i \in A} \delta_i, \tag{5.30}$$

where $\delta_i = b_i \bmod n$, that is $0 \le \delta_i < n$, and $b_i \equiv \delta_i (\bmod n)$, thus

$$\prod_{i \in A} b_i \equiv b (\bmod n).$$

Because of $a_i = b_i^2 \bmod n$, that is $0 \le a_i < n$, and $b_i^2 \equiv a_i (\bmod n)$. There is

$$\prod_{i \in A} b_i^2 = b^2 \equiv \prod_{i \in A} a_i = c^2 (\bmod n).$$

Two different integers b and c defined by Eqs. (5.29) and (5.30) satisfy $b^2 \equiv c^2 (\bmod n)$, We write the above analysis as the following lemma.

Lemma 5.18 *Let $A = \{b_1, b_2, \dots, b_i, \dots\}$ be a finite set of some B-numbers, let $e_i = (e_{i_1}, e_{i_2}, \dots, e_{i_h}) \in \mathbb{F}_2^h$ be the binary vector corresponding to b_i, $a_i = b_i^2 \bmod n$, $\delta_i = b_i \bmod n$. If $\sum_{i \in A} e_i = 0$ is the zero vector in $\mathbb{F}_2^h$, then $\prod_{i \in A} a_i$ is a square number. Write*

$$a_i = \prod_{j=1}^{h} p_j^{\alpha_{ij}}, \quad \prod_{i \in A} a_i = \prod_{j=1}^{h} p_j^{\sum_{i \in A} \alpha_{ij}} = c^2.$$

where

$$c = \prod_{j=1}^{h} p_j^{\frac{1}{2}\sum_{i \in A} \alpha_{ij}},$$

Further let $b = \delta_1 \delta_2 \cdots$, *we have* $b^2 \equiv c^2 (\bmod n)$.

From the above lemma, if $b^2 \equiv c^2 (\bmod n)$, $b \not\equiv \pm c (\bmod n)$. Then we will find a nontrivial factor $d = (b + c, n)$ of n. Now the question is, if $b^2 \equiv c^2 (\bmod n)$, how likely is $b \not\equiv \pm c (\bmod n)$? Might as well make $(b, n) = (c, n) = 1$, otherwise both sides are divided by $(b, n)^2$: by $b^2 \equiv c^2 (\bmod n)$, $\Longrightarrow (bc^{-1})^2 \equiv 1 (\bmod n)$. The problem is transformed into how many solutions x are in $x^2 \equiv 1 (\bmod n)$, $1 \le x < n$.

Lemma 5.19 *Let n be an odd number, then the number of solutions of* $x^2 \equiv 1 (\bmod n)$ *is* 2^r, *where r is the number of different prime factors of n.*

Proof If $r = 1$, then $n = p^\alpha (\alpha \ge 1)$, p is an odd prime, now $x^2 \equiv 1 (\bmod p^\alpha)$ has two solutions $x = \pm 1$, because let g be the original root of mod p^α, then $x = g^t (1 \le t \le p^{\alpha-1}(p-1))$, $x^2 = 1 \Leftrightarrow p^{\alpha-1}(p-1) | 2t$. So there are only two solutions $t = \frac{1}{2} p^{\alpha-1}(p-1)$ and $t = p^{\alpha-1}(p-1)$. So $x \equiv \pm 1 (\bmod p^\alpha)$. If $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, then the number of solutions of $x^2 \equiv 1 (\bmod n)$ deduced from the Chinese remainder theorem is 2^r. The Lemma holds!

Lemma 5.20 *n is an odd number and is the product of the power of more than two different primes,* $B = \{p_1, p_2, \ldots, p_h\}$ *is a factor base. Randomly select two B-numbers b and c, then* $b^2 \equiv c^2 (\bmod n)$, $\Longrightarrow b \equiv \pm c (\bmod n)$*'s rate is* $\le \frac{1}{2}$.

Proof $x^2 \equiv 1 (\bmod n)$ has 2^r different solutions $(\bmod n)$, $r \ge 2$. The two solutions corresponding to $x \equiv \pm 1 (\bmod n)$ correspond to $b \equiv \pm c (\bmod n)$. Thus

$$b^2 \equiv c^2 (\bmod n), \Longrightarrow b \equiv \pm c (\bmod n)\text{'s rate } \le \frac{2}{2^r} \le \frac{1}{2},$$

Lemma 5.20 holds.

According to Lemma 5.20, b and c are selected by using factor basis, if $b \equiv \pm c (\bmod n)$, then select failure, and the probability of failure is $\le \frac{1}{2}$. If the selection fails, select another b_1 and c_1, in this way, we randomly select k b and c equally almost independently, and the probability of success of $b \not\equiv \pm c (\bmod n)$ is

$$P\{b^2 \equiv c^2 (\bmod n), b \not\equiv \pm c (\bmod n)\} \ge 1 - \frac{1}{2^k}. \tag{5.31}$$

In other words, the probability of finding a nontrivial factor $d = (b + c, n)$ of n by using the factor base can be infinitely close to 1. Below, we systematically summarize the factor base decomposition method as follows:

Factor-based method

Let n be a large odd number and y be an appropriately selected integer (e.g., $y \leq n^{\frac{1}{10}}$), let the factor base be

$$B = \{-1, p \mid p \text{ is prime}, \ p \leq y\}.$$

Select a certain number of B-number at random, $A_1 = \{b_1, b_2, \ldots, b_N\}$, usually $N \leq \pi(y) + 2$ will meet the needs. Each b_i is expressed as the product of prime numbers in B. Calculate the corresponding binary vector e_i, select a subset $A \subset A_1$ in A_1, such that $\sum_{i \in A} e_i = 0$, b_i corresponding to binary vector e_i, denote as $A = \{b_1, b_2, \ldots, b_i, \ldots\}$. Let

$$b = \prod_{i \in A} (b_i \bmod n) = \prod_{i \in A} \delta_i, \ \text{where } \delta_i = b_i \bmod n$$

and

$$c = \prod_{j \in B} p_j^{r_j} \bmod n, \ r_j = \frac{1}{2} \sum_{i \in A} \alpha_{ij}.$$

We have $b^2 \equiv c^2 (\bmod n)$, if $b \equiv \pm c (\bmod n)$, then reselect the subset A, Until finally $b \not\equiv \pm c (\bmod n)$, in this way, we find a nontrivial factor $d|n$ of n, $d = (b + c, n)$. Therefore, there is factorization $n = d \cdot \frac{n}{d}$.

Factor decomposition using factor-based method cannot guarantee the success rate of 100% because $b \not\equiv \pm c (\bmod n)$ cannot be deduced from $b^2 \equiv c^2 (\bmod n)$, however, the success probability of factorization for large odd n can be infinitely close to 1. Under the condition of success probability $\geq 1 - \frac{1}{2^k}$ (k is a given normal number), the computational complexity of factorization n of by factor-based method can be estimated as

$$\text{Time(factor-based method to } n \text{ factorization)} = O(\mathrm{e}^{c\sqrt{\log n \log \log n}}). \tag{5.32}$$

The proof of Formula (5.32) is relatively complex. No detailed proof is given here. Interested readers can refer to pages 136–141 of (Pomerance, 1982a) in reference 5. The exact value of C in (5.32) is unknown. It is generally guessed that $C = 1 + \varepsilon$, where $\varepsilon > 0$ is any small positive real number.

Let k be the number of bits of n, and the estimate on the right of (5.32) can be written as $O(\mathrm{e}^{c\sqrt{k \log k}})$. Therefore, the computational complexity of the factor-based method is sub-exponential. Compared with the Monte Carlo method introduced in the previous section (see (5.31)), its computational complexity is exponential, because

$$O(\sqrt{n}) = O(\mathrm{e}^{c_1 k}), \ \text{where } c_1 = \frac{1}{2} \log 2.$$

As we all know, the security of RSA public key cryptography is based on the prime factorization $n = pq$ of n. Although there is no general method to factorize any large odd n, although Monte Carlo method and factor-based method are probability calculation methods, the probability of successful factorization is very large, The disadvantage is that their computational complexity is exponential and sub exponential, which is the reason for choosing huge prime numbers p and q in RSA.

5.5 Continued Fraction Method

In the factor-based method introduced in the previous section, $b^2 \bmod n$ can be the residual of the minimum absolute value of b^2 under $\bmod n$, that is

$$b^2 \equiv b^2 \bmod n (\bmod n), \ \ |b^2 \bmod n| \le \frac{n}{2}.$$

In this way, $b^2 \bmod n$ can be decomposed into the product of some smaller prime numbers. The continued fraction method is the best method at present. How to find the integer b, so that $|b^2 \bmod n| < 2\sqrt{n}$, $b^2 \bmod n$ is more likely to be decomposed into the product of some small prime numbers. First, we introduce what is continued fraction and some basic properties.

Suppose $x \in \mathbb{R}$ is a real number, $[x]$ is the integer part of x, and $\{x\}$ is the decimal part of x. Let $a_0 = [x]$, if $\{x\} \neq 0$, and let $a_1 = [\frac{1}{\{x\}}]$, because of $x = [x] + \{x\}$, there is

$$x = a_0 + \frac{1}{\{x\}} = a_0 + \frac{1}{a_1 + \{\{x\}^{-1}\}}.$$

If $\{\{x\}^{-1}\} \neq 0$, write

$$a_2 = [\{\{x\}^{-1}\}^{-1}],$$

consider

$$\{\{\{x\}^{-1}\}^{-1}\}^{-1},$$

So we got

$$x = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}}.$$

The above formula is called the continued fraction expansion of real number x. To save space, write $x = [a_0, a_1, \ldots, a_n, \ldots]$, if and only if x is a rational number, the continued fraction of x is expanded to be finite, denote as

$$x = [a_0, a_1, \ldots, a_n], \ \text{where } a_n > 1.$$

It is called the standard expansion of rational number x.

Definition 5.6 $x = [a_0, a_1, \ldots, a_n, \ldots]$ is the continued fraction expansion of x, for $i \geq 0$, call $\frac{b_i}{c_i} = [a_0, a_1, \ldots, a_i]$ the ith asymptotic fraction of x, specially,

$$\frac{b_0}{c_0} = \frac{a_0}{1}, \ \frac{b_1}{c_1} = \frac{a_1 a_0 + 1}{a_1}.$$

The progressive fraction $\frac{b_i}{c_i}$ of the real number x is a reduced fraction, that is $(b_i, c_i) = 1$, and has the following properties.

Lemma 5.21 *$x = [a_0, a_1, \ldots, a_n, \cdots]$ is the continued fraction expansion of x, $\frac{b_i}{c_i}$ is the asymptotic fraction, then*

(i) when $i \geq 2$,

$$\frac{b_i}{c_i} = \frac{a_i b_{i-1} + b_{i-2}}{a_i c_{i-1} + c_{i-2}}. \tag{5.33}$$

(ii) If $i \geq 1$, then

$$b_i c_{i-1} - b_{i-1} c_i = (-1)^{i-1}. \tag{5.34}$$

Proof We prove that (i) by induction. Obviously, the proposition of $i = 2$ holds, that is

$$\frac{b_2}{c_2} = \frac{a_2 b_1 + b_0}{a_2 c_1 + c_0} = \frac{a_2(a_1 a_0 + 1) + a_0}{a_2 a_1 + 1}.$$

If the proposition holds for i, that is

$$\frac{b_i}{c_i} = \frac{a_i b_{i-1} + b_{i-2}}{a_i c_{i-1} + c_{i-2}}.$$

Then write $[a_0, a_1, \ldots, a_i, a_{i+1}] = [a_0, a_1, \ldots, a_i + \frac{1}{a_{i+1}}]$,

$$\frac{b_{i+1}}{c_{i+1}} = \frac{\left(a_i + \frac{1}{a_{i+1}}\right) b_{i-1} + b_{i-2}}{\left(a_i + \frac{1}{a_{i+1}}\right) c_{i-1} + c_{i-2}} = \frac{a_{i+1} b_i + b_{i-1}}{a_{i+1} c_i + c_{i-1}}.$$

So (i) holds.

We prove Formula (5.34) by induction, when $i = 1$,

$$b_1 c_0 - b_0 c_1 = a_1 a_0 + 1 - a_1 a_0 = 1 = (-1)^0.$$

So when $i = 1$, the proposition holds, and when i, the proposition holds, that is

$$b_i c_{i-1} - b_{i-1} c_i = (-1)^{i-1}.$$

Then

$$\begin{aligned} b_{i+1}c_i - b_i c_{i+1} &= (a_{i+1}b_i + b_{i-1})c_i - b_i(a_{i+1}c_i + c_{i-1}) \\ &= b_{i-1}c_i - b_i c_{i-1} \\ &= (-1)^i. \end{aligned}$$

Lemma 5.21 holds.

Continued fractions have many important applications in numbers, such as rational approximation of real numbers and rational approximation of algebraic numbers. Periodic continued fractions are an important special case in rational approximation of algebraic numbers. $x = [a_0, a_1, \ldots, a_n, \ldots]$. If these a_i occur in cycles of a certain length, they are called periodic continued fractions. The famous Lagrange theorem shows that the necessary and sufficient condition for the expansion of the continued fraction of x into a periodic continued fraction is that x is a quadratic real algebraic number. Here we do not discuss some profound properties of continued fractions, but only prove some properties we need.

Lemma 5.22 *Let $x > 1$ be a real number, $\frac{b_i}{c_i} (i \geq 0)$ is the asymptotic fraction of x, then*

$$|b_i{}^2 - x^2 c_i^2| < 2x, \forall\, i \geq 0.$$

Proof Because x is between progressive scores $\frac{b_i}{c_i}$ and $\frac{b_{i+1}}{c_{i+1}}$, by property (ii) of Lemma 5.21, there is

$$\left| \frac{b_{i+1}}{c_{i+1}} - \frac{b_i}{c_i} \right| = \frac{1}{c_i c_{i+1}}, i \geq 0.$$

Thus

$$\begin{aligned} |b_i{}^2 - x^2 c_i^2| &= c_i^2 \left| x - \frac{b_i}{c_i} \right| \left| x + \frac{b_i}{c_i} \right| \\ &< c_i^2 \cdot \frac{1}{c_i c_{i+1}} \left(x + \left(x + \frac{1}{c_i c_{i+1}} \right) \right). \end{aligned}$$

So

$$\begin{aligned} |b_i{}^2 - x^2 c_i^2| - 2x &< 2x \left(-1 + \frac{c_i}{c_{i+1}} + \frac{1}{2x c_{i+1}^2} \right) \\ &< 2x \left(-1 + \frac{c_i}{c_{i+1}} + \frac{1}{c_{i+1}} \right) \\ &< 2x \left(-1 + \frac{c_{i+1}}{c_{i+1}} \right) = 0. \end{aligned}$$

The Lemma holds.

Lemma 5.23 *Let n be a positive integer and n not a complete square. Let $\{\frac{b_i}{c_i}\}_{i \geq 0}$ be the asymptotic fraction of the continued fraction expansion of $\sqrt{n}$, and $b_i^2 \bmod n$ be the residue of the minimum absolute value of b_i^2 under $\bmod n$, then we have*

$$b_i^2 \bmod n < 2\sqrt{n}, \ \forall\, i \geq 0.$$

Proof By Lemma 5.22, let $x = \sqrt{n}$, then

$$b_i^2 \equiv b_i^2 - nc_i^2 (\bmod n).$$

Because

$$|b_i^2 - nc_i^2| < 2\sqrt{n}, \Longrightarrow b_i^2 \bmod n < 2\sqrt{n}, \ \forall\, i \geq 0.$$

The Lemma holds.

Combining the above Lemma 5.23 with the factorization method, we obtain the continued fraction decomposition method.

Continued fraction decomposition method:

The operations of $\bmod n$ involved in this algorithm, except that it is specially pointed out, are the minimum nonnegative residue of $\bmod n$. If n is a large odd number, it is also a compound number, first let $b_{-1} = b$, $b_0 = a_0 = [\sqrt{n}]$, and $x_0 = \sqrt{n} - a_0 = \{\sqrt{n}\}$, calculate $b_0^2 \bmod n$, in fact, $b_0^2 \bmod n = b_0^2 - n$. Second, consider $i = 1, 2, \ldots$. To determine b_i, we proceed in several steps:

1. Let $a_i = [\frac{1}{x_{i-1}}]$, and $x_i = \frac{1}{x_{i-1}} - a_i (i \geq 1)$.
2. Let $b_i = a_i b_{i-1} + b_{i-2}$, the minimum nonnegative residual $b_i \bmod n$ of b_i under $\bmod n$ is still recorded as b_i.
3. calculate $b_i^2 \bmod n$.

By Lemma 5.23, $b_i^2 \bmod n < 2\sqrt{n}$, it can be decomposed into the product of some small prime numbers. If a prime number p appears in the decomposition of two or more $b_i^2 \bmod n$, or in the decomposition of an $b_i^2 \bmod n$, p appears to an even power, p is called a standard prime number, in other words, a standard prime p is

$$p | b_i^2 \bmod n, \ \ p | b_j^2 \bmod n, i \neq j.$$

Or

$$p^\alpha \| b_i^2 \bmod n, \ \alpha \text{ is even}.$$

We choose factor base B as

$$B = \{-1, \ \text{standard prime}\}.$$

In this way, all $b_i^2 \bmod n$ are B-numbers, and the corresponding binary vector is e_i. Select a subset $A = \{b_i\}, \Longrightarrow \sum_{i \in A} e_i = 0$. Let

$$b = \prod_{i \in A} (b_i \bmod n) = \prod_{i \in A} \delta_i$$

and $c = \prod_{j \in B} p_j^{r_j}$, where

$$r_j = \frac{1}{2}\sum_{i\in A}\alpha_{ij}, \forall j \in B.$$

If $b \not\equiv \pm c(\bmod n)$, then $(b+c, n)$ is a nontrivial factor of n, and we obtain the factorization of n. If $b \equiv \pm c(\bmod n)$, then another subset A is selected and repeated to complete the continued fraction factorization method.

Example 5.8 The continued fraction method is used to factorize $n = 9073$.

Solution: We calculate a_i, b_i and $b_i^2 \bmod n$ in turn, where $b_i = (a_i b_{i-1} + b_{i-2}) \bmod n$, the table is as follows:

i	0	1	2	3	4
a_i	95	3	1	26	2
b_i	95	286	381	1119	2619
$b_i^2 \bmod n$	−48	139	−7	87	−27

From the value of $b_i^2 \bmod n$, we can choose the factor base B as $B = \{-1, 2, 3, 7\}$. Then $b_i^2 \bmod n$ is the number of B-number, when $i = 0, 2, 4, \ldots$. The corresponding binary vector is

$$e_0 = (1, 4, 1, 0), e_2 = (1, 0, 0, 1), \text{and } e_4 = (1, 0, 3, 0).$$

Easy to calculate $e_0 + e_4 = (0, 0, 0, 0)$. Therefore, we choose

$$\begin{cases} b = 95 \cdot 2619 \equiv 3834 \bmod 9073; \\ c = 2^2 \cdot 3^2 = 36. \end{cases}$$

Because $b^2 \equiv c^2 (\bmod 9073)$, that is $3834^2 = 36^2 (\bmod 9073)$, but $3834 \not\equiv \pm 36$ $(\bmod 9073)$, so we get a nontrivial factor of $n = 9073, d = (3834 + 36, 9073) = 43$. Thus $9073 = 43 \cdot 211$, the factorization of 9073 is obtained.

Exercise 5

1. p is a prime, if and only if $b^{p-1} \equiv 1(\bmod p^2)$, p^2 to base b is a Fermat pseudo prime.
2. What is the minimum pseudo prime number with Fermat pseudo prime for base 5? What is the minimum Fermat pseudo prime number for base 2?
3. $n = pq, p \neq q$ are two primes, let $d = (p-1, q-1)$, it is proved that n to base b is Fermat pseudo prime number, if and only if $b^d \equiv 1(\bmod n)$, and calculate the number of bases b.
4. If $b \in \mathbb{Z}_n^*$, n to base b is Fermat pseudo prime, then n to base $-b$ and b are Fermat pseudo prime numbers.
5. If n to base 2 is Fermat pseudo prime, then $N = 2^n - 1$ is also Fermat pseudo prime.
6. If n to base b is Fermat pseudo prime, and $(b-1, n) = 1$, then $N = \frac{b^n-1}{b-1}$ to base b is also Fermat pseudo prime.

7. Prove that the following integers are Carmichael numbers:

$$1105 = 5 \cdot 13 \cdot 17, 1729 = 7 \cdot 13 \cdot 19, 2465 = 5 \cdot 17 \cdot 29, 2821 = 7 \cdot 13 \cdot 31,$$

$$6601 = 7 \cdot 23 \cdot 41, 29{,}341 = 13 \cdot 37 \cdot 61, 172{,}081 = 7 \cdot 13 \cdot 31 \cdot 61, 278{,}545 = 5 \cdot 17 \cdot 29 \cdot 113.$$

8. Find all Carmichael numbers of form $3pq$ and all Carmichael numbers of form $5pq$.
9. Prove that 561 is the minimum Carmichael number.
10. If n to base 2 is a Fermat pseudo prime, prove $N = 2^n - 1$ is a strong pseudo prime.
11. There are infinite Euler pseudo primes and strong pseudo primes for base 2.
12. If n to base b is a strong pseudo prime, then n to base b^k is also a strong pseudo prime for any integer k.
13. The Fermat factorization method is used to decompose the positive integer as follows:

$$n = 8633, n = 809{,}009, n = 92{,}296{,}873, n = 88{,}169{,}891.$$

14. The Fermat factorization method is used to decompose the positive integer as follows:

$$n = 68{,}987, n = 29{,}895{,}581, n = 19{,}578{,}079, n = 17{,}018{,}759.$$

15. Expand the rational number $x = \frac{45}{89}, x = \frac{55}{89}, x = 1.13$ into continued fractions.
16. Let a be a positive integer, $x = [a, a, a, \cdots]$, calculate $x =$?

References

Adelman, L. M., Pomerance, C., & Rumely, R. S. (1983). On distinguishing prime number from composite numbers. *Annals of Mathematics, 117,* 173–206.

Berent, R. P., & Pollared, J. M. (1981). Factorization of the eighth Fermat number. *Mathematics of Computation, 36*, 627–630.

Blair, W. D., Lacampague, C. B., & Selfridge, J. L. (1986). Factoring large numbers on a pocket calculator. *The American Mathematical Monthly, 93,* 802–808.

Brent, R. P. (1980). An improved Monte Carlo factorization algorithm. *BIT, 20,* 176–184.

Cohen, H., & Lenstra, H. W. (1984). Primality testing and Jacobi sums. *Mathematics of Computation, 142*, 297–330.

Dawonport, H. (1982). *The higher arithmetic*. Cambridge University Press.

Dickson, L. E. (1952). *History of the theory of number* (Vol. 1). Chelsea.

Dixon, J. D. (1984). Factorization and primality tests. *The American Mathematical Monthly, 91,* 333–352.

Guy, R. K. (1975). How to factor a number. In *Proceedings of the 5th Manitoba Conference on Numerical Mathematics* (pp. 49–89).

Kranakis, E. (1986). *Primality and cryptography*. Wiley.

Lehman, R. S. (1974). Factoring large number. *Mathematics of Computation, 28,* 637–646.

Lehmer, D. H., & Powers, R. E. (1931). On factoring large number. *Bulletin of the American Mathematical Society, 37,* 770–776.

Miller, G. L. Riemann's hypothesis and tests for primality. In *Proceedings of the 7th Annual ACM Symposium on the Theory of Computing* (pp. 234–239).

Morrison, M. A., & Brillhart, J. (1975). A method of factoring and the factorization of $\mathbb{F}_7$. *Mathematics of Computation, 29,* 183–205.

Pollard, J. M. (1975). A Monte Carlo method for factorization. *BIT, 15,* 331–334.

Pomerance, C. (1981). Recent development in primality testing. *The Mathematical Intelligencer,3,* 97–105.

Pomerance, C. (1982a). Analysis and comparison of some integer factoring algorithms. *Computation Methods in Number Theory, Part 1*.

Pomerance, C. (1982b). The search for prime number. *Scientific American,427,* 136–147.

Pomerance, C., & Wagstaff, S. S. (1983). Implementation of the continued fraction integer factoring algorithm. In *Proceedings of the 12th Winnipeg Conference on Numerical Methods and computing*.

Rabin, M. O. (1980). Probabilities algorithms for testing Primality. *Journal of Number Theory, 12,* 128–138.

Solovag, R., & Strassen, V. (1977). A fast Munte Carlo test for primality. *SIAM Journal for Computing, 6,* 84–85.

Wagon, S. (1986). Primality testing. *The Mathematical Intelligence, 8,* 58–61.

Wunderlich, M. C. (1979). A running time and analysis of Brillhart's continued fraction factoring method. *Number Theory, Carbondale, Springer Lecture Notes, 175*, 328–342.

Wunderlich, M. C. (1985). Implementing the continued fraction factoring algorithm on parallel machines. *Mathematics of Computation, 44,* 251–260.