algorithm. As most algorithms for testing or proving primality, the AKS algorithm is also based on a variant of Fermat's little theorem. Its initial point is the following result. Let a and n be integers, $n \geq 2$ and $\gcd(a, n) = 1$. Then the number n is prime if and only if

$$(X + a)^n \equiv X^n + a \pmod{n}, \tag{15.42}$$

i.e. if and only if the corresponding coefficients of the polynomials on the left and right-hand side of congruence (15.42) are congruent modulo n.

15.10 Elliptic curve factorization method

If a positive integer n does not pass some of primality tests, then we know that n is certainly composite. However, those tests usually do not give any non-trivial factor of n. Therefore, the question is, how to find a non-trivial factor of a large composite number. This is considered to be a hard problem, and on its hardness, some of the most important public-key cryptosystems are based.

The methods of factorization can be divided into general and special. With the general methods, the expected number of operations depends only on the size of the number n, while with the special methods, it also depends on properties of the factors of n.

The naïve method of factorization is dividing n by all prime numbers $\leq \sqrt{n}$. The number of needed divisions is in the worst case around $\frac{2\sqrt{n}}{\ln n}$, so the complexity of this method is $O(\sqrt{n} \ln n)$. This method is very inefficient for large n's. However, it is good to use it in combination with better methods of factorization, for removing possible small factors of n.

Pollard's $p - 1$ *method* from 1974 belongs to the special factorization methods. Its initial point is again Fermat's little theorem. Let n be a composite number which we want to factorize and let p be one of its prime factors. Then $a^{p-1} \equiv 1 \pmod{p}$ for $\gcd(a, p) = 1$. Moreover, $a^m \equiv 1 \pmod{p}$ for any multiple m of $p - 1$. If we find m, then $\gcd(a^m - 1, n)$ gives us a factor (we hope a non-trivial one) of n. However, the question is how to find a multiple of $p - 1$ when we do not know p. This can be efficiently done in the case when $p - 1$ has only small prime factors. We say that a positive integer is *B-smooth* if all of its prime factors are $\leq B$. Let us assume in addition that all prime powers which divide $p - 1$ are $\leq B$. Then for m we can take the least common multiple of the numbers $1, 2, \ldots, B$.

Example 15.18. Let $n = 843281$. Let us choose $B = 8$ and $a = 2$. Then $m = 2^3 \cdot 3 \cdot 5 \cdot 7 = 840$. We have $2^{840} \bmod n = 652764$ and $\gcd(652763, n) = 281$. Indeed, $n = 281 \cdot 3001$. $\diamondsuit$

The success of Pollard's $p-1$ method directly depends on the smoothness of the number $p - 1$. In the worst case, which is when the number $\frac{p-1}{2}$ is prime, this method is no better than the naïve method of division. There are variants of this method which use the smoothness of one of the numbers $p + 1$, p^2+p+1, p^2+1 or p^2-p+1. However, the most significant modification of Pollard's $p - 1$ method is Lenstra's method of factorization which uses elliptic curves. In this method, again, the group $\mathbb{F}_p^*$ of order $p - 1$ is replaced by the group $E(\mathbb{F}_p)$ whose order varies within the interval $(p + 1 - 2\sqrt{p}, p + 1 + 2\sqrt{p})$, so we can hope that we will find an elliptic curve over $\mathbb{F}_p$ with sufficiently smooth order.

In 1987, H. W. Lenstra [272] proposed a modification of Pollard's $p - 1$ method which uses elliptic curves. As a result, he obtained a subexponential time algorithm which is even today, one of the most efficient algorithms for factorization.

Similarly to the method of primality proving using elliptic curves, we will again deal with elliptic curves over the ring $\mathbb{Z}/n\mathbb{Z}$. While with the primality proving, there was a (very small) possibility that n is composite (i.e. that $\mathbb{Z}/n\mathbb{Z}$ is not a field), here, we will be sure that n is composite. We will assume that $\gcd(n, 6) = 1$ and consider elliptic curves of the form

$$E_{a,b} : \quad y^2 = x^3 + ax + b,$$

where $\gcd(4a^3 + 27b^2, n) = 1$. When n is prime, then on an elliptic curve there exists only one projective point which does not correspond to an affine point (the point at infinity). In the case when n is composite, there can be more such points.

Let us now describe the basic steps in *Lenstra's algorithm for factorization* (*Elliptic Curve Method* – ECM).

1. *The choice of an elliptic curve*: There are several ways of choosing a suitable elliptic curve. For example, we can randomly choose elements $a, x, y \in \mathbb{Z}_n$, and then calculate $b = (y^2 - x^3 - ax) \bmod n$. Let $g = \gcd(4a^3 + 27b^2, n)$. If $1 < g < n$, then we found a non-trivial factor of n. If $g = n$, then we choose new a, x, y. If $g = 1$, then we found an elliptic curve $E_{a,b}$ over $\mathbb{Z}/n\mathbb{Z}$ and the point $P = (x, y)$ on it.

2. Let k be the least common multiple of the numbers $1, 2, \dots, B$, for a suitably chosen bound B. In practice, we can first take say $B = 10000$, and then increase the bound if needed.

3. We calculate $[k]P \in E_{a,b}(\mathbb{Z}_n)$ using formulas for the addition of points:

   $$(x_3, y_3) = (\lambda^2 - x_1 - x_2 \bmod n,\ \lambda(x_1 - x_3) - y_1 \bmod n),$$

   where $\lambda = (3x_1^2 + a) \cdot (2y_1)^{-1} \bmod n$ if the points are equal and $\lambda = (y_1 - y_2)(x_1 - x_2)^{-1} \bmod n$, otherwise.

4. If in the calculation of $[k]P$, it happens that the sum of points cannot be calculated because for a number, let us denote it by d, we cannot compute the inverse d^{-1} because d does not have an inverse modulo n, then we compute $g = \gcd(d, n)$. If $g \neq n$, then we found a non-trivial factor of n.

5. In the case of failure, we can choose a new elliptic curve or increase the bound B.

Example 15.19. *Let us factorize the number $n = 629$.*

Solution: Let $B = 3$, so $k = 6$. Let us choose the elliptic curve $y^2 = x^3 + 8x + 9$ and an obvious point $P = (0, 3)$ on it. We calculate $[6]P = [2](P + [2]P)$. First, we calculate $[2]P$. The corresponding λ is $8 \cdot 6^{-1} = 211 \bmod 629$, so we obtain $[2]P = (491, 181)$. Then we calculate $[3]P = P + [2]P$. The corresponding λ is $178 \cdot 491^{-1} = 336 \bmod 629$, so $[3]P = (443, 222)$. Finally, we calculate $[6]P = [2]([3]P)$. The corresponding λ is $11 \cdot 444^{-1}$. When computing the inverse of 444 modulo 629, we find that the inverse does not exist because $\gcd(444, 629) = 37$. From this, we conclude that 37 is a factor of 629. Indeed, $629 = 17 \cdot 37$. ♢

What does the success of this algorithm depend on? Similarly to Pollard's $p - 1$ method, here k also should be a multiple of the order of the corresponding group. In this case, k should be a multiple of $|E(\mathbb{F}_p)|$, where p is a prime factor of n. Indeed, in that case, when calculating $[k]P$, the corresponding denominator will be divisible by p, so it will not be invertible modulo n. Namely, we will have $[k]P = \mathcal{O}$ in $E(\mathbb{F}_p)$.

When estimating the complexity of this algorithm, the crucial question is how to choose the bound B optimally. It can be shown that the optimal choice is

$$B = e^{(\sqrt{2}/2 + o(1))\sqrt{\ln p \ln \ln p}},$$

while the complexity of the algorithm is

$$e^{(\sqrt{2}+o(1))\sqrt{\ln p \ln\ln p}}.$$

In the worst case (when $p = O(\sqrt{n})$), the complexity of the elliptic curve factorization method is $e^{O(\sqrt{\ln n \ln\ln n})}$. Hence, that is a subexponential time algorithm.

Although there are algorithms with better complexity (the number field sieve algorithm), an important property of the ECM is that its complexity depends on the smallest prime factor of n. Therefore, it is not the most suitable for factorization of moduli in the RSA cryptosystem, i.e. numbers of the form $n = pq$, where p and q are primes of similar size. However, in the factorization of "random" numbers, the ECM often provides better results than other methods because such numbers usually have a prime factor which is significantly smaller than $\sqrt{n}$. Even with the application of asymptotically better methods, within those algorithms, it is needed to factorize auxiliary numbers for which we can expect to behave as random numbers and then the ECM can be used as an auxiliary method.

Among the factorizations obtained using the ECM, let us mention finding the 33-digit factor of the Fermat number $2^{2^{15}} + 1$ (Crandall, van Halewyn, 1997), and finding the 49-digit factor of the Mersenne number $2^{2071} - 1$ (Zimmermann, 1998).

In order to increase the chance for the success of this algorithm, Montgomery [309] as well as Atkin and Morain [14] proposed the use of elliptic curves over $\mathbb{Q}$ with large torsion group and positive rank. Namely, we saw that Proposition 15.7 implies that $|E(\mathbb{F}_p)|$ is divisible by the order of the torsion group of $E(\mathbb{Q})$, so the larger torsion group of $E(\mathbb{Q})$ increases the likelihood that the order $|E(\mathbb{F}_p)|$ will have small prime factors. Furthermore, it was shown in [56, 149] that the additional advantage can be obtained by using elliptic curves which have even larger torsion group over a number field of small degree. Here, we encounter the question which torsion groups are possible, for instance, over quadratic or cubic fields.

Kenku and Momose [243] and Kamienny [240] proved that in quadratic fields there can occur precisely the following 26 groups:

$$\begin{aligned} \mathbb{Z}/k\mathbb{Z}, &\quad 1 \le k \le 18,\ k \ne 17, \\ \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2k\mathbb{Z}, &\quad 1 \le k \le 6, \\ \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3k\mathbb{Z}, &\quad k = 1, 2, \\ \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}. & \end{aligned} \tag{15.43}$$

For results concerning the possibilities for the torsion subgroup over a fixed quadratic field see [318, 401]. Similarly to Mazur's theorem, here also each of these 26 groups appears as the torsion group over a quadratic field for infinitely many non-isomorphic elliptic curves. With cubic fields, the situation is different. Namely, Najman [321] proved that there is a unique curve with torsion group $\mathbb{Z}/21\mathbb{Z}$ over a cubic field. Let us mention that for each of the 26 groups from (15.43), there is an elliptic curve over a quadratic field with that torsion group and positive rank, and for all groups except $\mathbb{Z}/15\mathbb{Z}$ also with the rank ≥ 2 (see [8, 56, 320]).

15.11 Exercises

1. Let $P = (-1, 2)$ and $Q = (1, -2)$ be points on the elliptic curve over $\mathbb{Q}$ given by the equation $y^2 = x^3 - x + 4$. Determine points $P + Q$, $P - Q$, $2P$ and $2Q$ on that curve.

2. Write the elliptic curve over $\mathbb{Q}$ given by the equation
$$y^2 + xy + y = x^3 - x^2 - 5x$$
in the short Weierstrass form.

3. Write the elliptic curve over $\mathbb{Q}$ given by the equation
$$x^3 + y^3 + x^2 + x + 1 = 0$$
in the short Weierstrass form.

4. Show that the curve
$$y^2 = x^3 + 4x^2 - 3x - 18$$
is singular. Determine its singular point and its rational parametrization.

5. Determine the j-invariant of the elliptic curve
$$y^2 + xy + y = x^3 - x^2 - 2.$$

6. The elliptic curve over $\mathbb{Q}$ is given by the equation
$$E \,:\, y^2 = x^3 + 1875x + 156250.$$
Determine its minimal Weierstrass equation.