ELLIPTIC CURVES AND FACTORISATION

ALEKSANDER HORAWA

These are the notes for a talk at the Undergraduate Colloquium at Imperial College London.

The aim of the talk is to present a method of factoring numbers using elliptic curves due to Lenstra [Len87]. It is mostly based on [Kob94, Ch. VI]

1. Elliptic Curves

We first present some basic ideas related to elliptic curves. For a detailed introduction, see [ST92].

Definition 1. An *elliptic curve* over $\mathbb{R}$ is the set of solution $(x, y) \in \mathbb{R}^2$ of

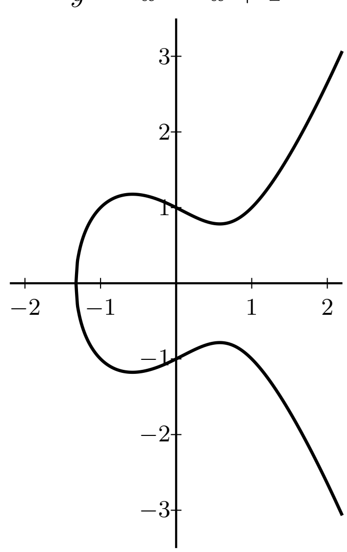
$$y^2 = x^3 + ax + b$$

for $a, b \in \mathbb{R}$ such that $27b^2 + 4a^3 \neq 0$, together with a point O called the *point at infinity*.

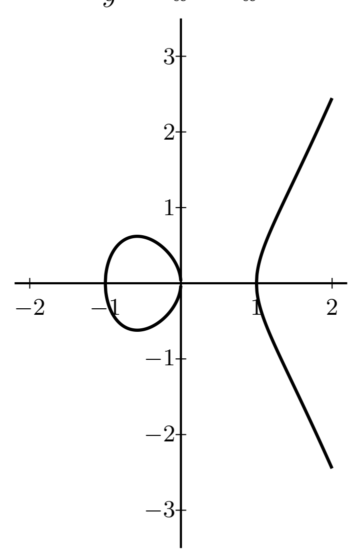
Why do we assume that $27b^2 + 4a^3 \neq 0$? This means that $x^3 + ax + b$ has no repeated roots, so we can define tangents at every point of the curve (or, as an algebraic geometer would say, the curve is non-singular).

Examples 2. The following curves are examples of elliptic curves. Note that the graphs are smooth everywhere.

$y^2 = x^3 - x + 1$

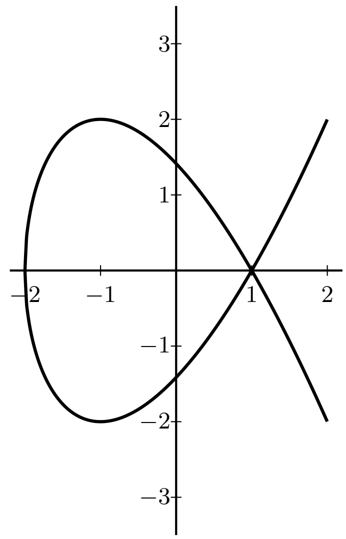


$y^2 = x^3 - x$

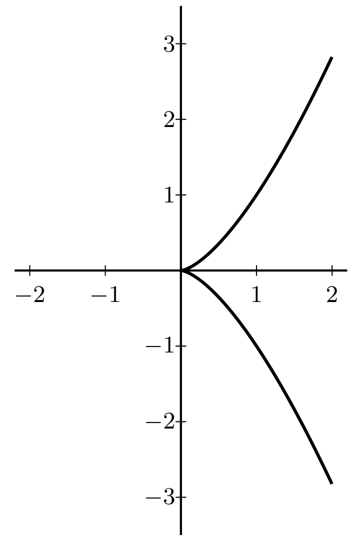


However, the following curve are not elliptic curves. Clearly, for both of them $27b^2 + 4a^3 = 0$.

$y^2 = x^3 - 3x + 2$



$y^2 = x^3$



In the first one, we cannot define a tangent at the point $(1, 0)$. In the second one we cannot define a tangent at $(0, 0)$.

What is the point at infinity, O? This point does not belong to the plane but we think of it as the *direction upwards*. That is, if we wish to draw a line through O and any given point P on the plane, we would simply draw a vertical line through P.

We have defined elliptic curves over $\mathbb{R}$ to have nice examples to draw. However, there is no reason to limit ourselves to $\mathbb{R}$. We can define ellptic curves over any field (e.g. $\mathbb{Q}$, $\mathbb{C}$, $\mathbb{F}_q$).

For example, an elliptic curve over $\mathbb{Q}$ is:

$$E(\mathbb{Q}) = \{(x, y) \in \mathbb{Q} \mid y^2 = x^3 + ax + b\} \cup \{O\}$$

where $a, b \in \mathbb{Q}$ and $27b^2 + 4a^3 \neq 0$.

2. Addition on Elliptic Curves

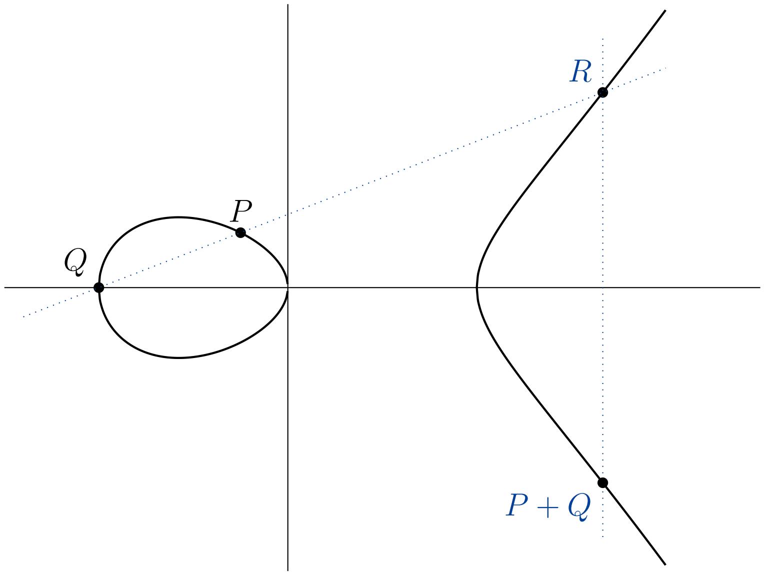
Why are elliptic curves so important and find so many applications? We can define a non-trivial *addition* on them!

Let E be an elliptic curve. Let us think how a line can intersect with the cubic. Using Bézout's theorem (i.e. counting the intersection multiplicity), we can show that:

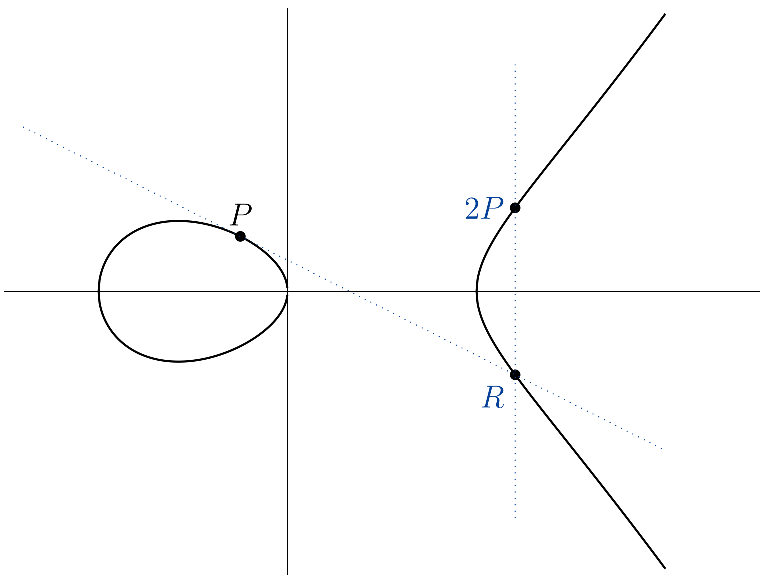
- any non-tangent line through two points on E intersects it at exactly one more point (this may be O);
- the tangent at O to E does not intersect E at any other point;
- any other line tangent to an elliptic curve intersects the curve at exactly one more point.

This allows us to naturally define the addition on E.

(1) The point at infinity O is defined to be the identity (i.e. $-O = O$ and $P + O = O + P = P$ for any point P).
(2) The negative $-P$ of $P = (x, y)$ is defined to be $(x, -y)$.
(3) If $P \neq Q$, then the line through P and Q intersects the curve at another point, say R. We then define $P + Q = -R$.



(4) If the line tangent to P intersects the curve at point R, then $2P = -R$.



Why do we not define $P + Q$ equal to R, the third point of intersection? There are several reasons for this. To name one, we want O to be the identity of the group, i.e. $P + 0 = P$. Since the line through P and O is the line through P pointing upwards, it intersects the cubic at $R = -P$. Therefore, we need $P + O = -R = P$.

One can check that this makes E into an abelian group. The only group axiom which is not obvious from the definition is associativity, which can be shown using projective geometry or Abel's Theorem (see [Kir92, Ch. 3]).

The above definition is geometric in its nature, making in rather involved computationally. Fortunately, the addition law can be expressed by explicit formulas. Suppose we have $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ on the curve, and we wish to find $P+Q = (x_3, y_3)$ and $2P = (x_4, y_4)$. By writing down the equation of the line passing through two points checking where it intersects the curve, one verifies that (see [Kob94, Ch. VI.1] for details):

$$\begin{aligned} x_3 &= \left(\tfrac{y_2-y_1}{x_2-x_1}\right)^2 - x_1 - x_2, \quad & y_3 &= -y_1 + \left(\tfrac{y_2-y_1}{x_2-x_1}\right)(x_1 - x_3), \\ x_4 &= \left(\tfrac{3x_1^2+a}{2y_1}\right)^2 - 2x_1, & y_4 &= -y_1 + \left(\tfrac{3x_1^2+a}{2y_1}\right)(x_1 - x_4). \end{aligned} \tag{1}$$

While these formulas may seem complicated, they are very easy to implement in an algorithm.

3. Lenstra's Factoring Algorithm

Suppose we want to factor n (and that 2 and 3 do not divide n).

The basic idea is to consider the curve E modulo n:

$$E \mod n = E(\mathbb{Z}/n\mathbb{Z}) = \{(x, y) \mid y^2 \equiv x^3 + ax + b \mod n\} \cup \{O\}$$

where $a, b \in \{0, 1, \dots, n-1\}$ and $\gcd(27b^2 + 4a^3) = 1$.

Addition can simply be defined by the formulas (1) used modulo n. However, they involve division, which is not always well-defined modulo n. In fact, division by d is well-defined if and only if $\gcd(d, n) = 1$.

For example, $7 \equiv 2 \mod 5$, so $7/2 \equiv 1 \mod 5$, we can divide by 2 modulo 5. However, $1 \not\equiv 4 \mod 6$. even though $2 \equiv 8 \mod 6$, we cannot divide by 2 modulo 6.

If you know some ring theory, you can immediately see that d has to be a unit in the ring $\mathbb{Z}/n\mathbb{Z}$.

The intuition behind the algorithm is the following. After we add two points P and Q in $E \mod n$ and we get d in the denominator, then $\gcd(d, n) \neq 1$ if and only if we have hit O, the point at infinity. However, $\gcd(d, n) | n$, so there is a chance we have found a divisor of n (unless $\gcd(d, n) = n$). The formal statement of this and the proof can be found in [Kob94, Prop. VI.3.1].

Below we present Lenstra's method for factorisation of integers following [Kob94, Ch. VI.3]. A detailed explanation can be found in [Len87, Sec. 2].

Algorithm 3 (Lenstra)**.**

(1) *Choose a curve* $E = \{y^2 = x^3 + ax + b\}$ *with* $a, b \in \mathbb{Z}$ *and a point* $P = (x, y)$ *on it.*
(2) *Let* $d = \gcd(4a^3 + 27b^2, n)$*. If* $1 < d < n$*, then we have found a proper divisor and we are done. If* $d = n$*, then go back to (1). Otherwise, proceed to (3).*
(3) *Choose a bound* B *and a bound* C*, and let* k *be the product of powers of primes not exceeding* B *which are less than* C*, that is*

$$k = \prod_{l \leq B} l^{\alpha_l}$$

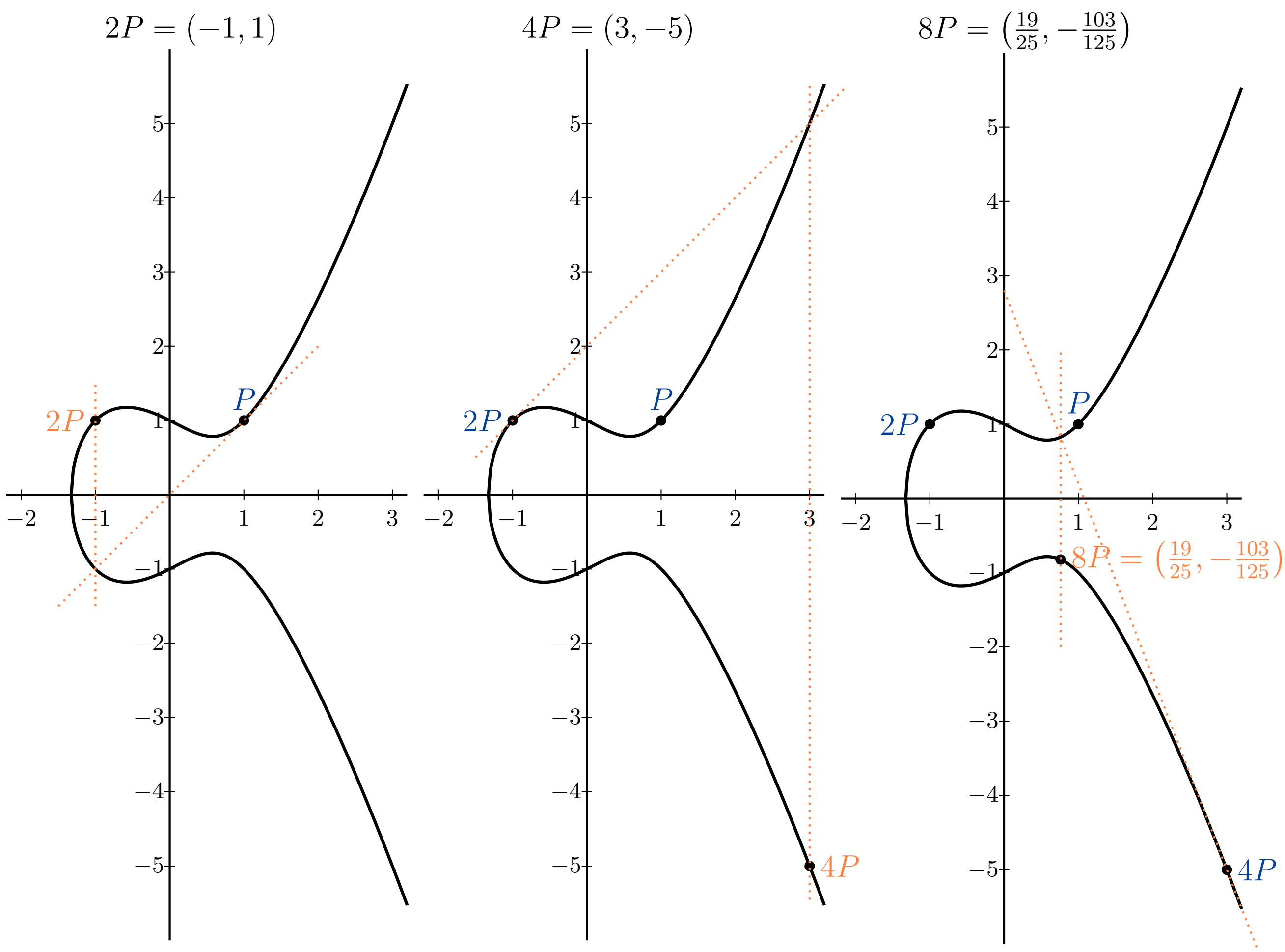
where l is prime and $l^{\alpha_l} \leq C$.

(4) *Attempt to compute $kP = \underbrace{P + P + \cdots + P}_{k \text{ times}}$ working modulo n. If you complete the calculation, go back to (1) and choose a different pair (E, P). If the calculation fails, it was impossible to find the inverse of $x_1 - x_2$ or $2y_1$ in one of the partial sums, i.e. we have a denominator x which is not coprime to n. Then $d = \gcd(x, n)$ is either a proper divisor of n (in which case we are done) or n itself (in which case we go back to (1) and choose a different pair (E, P)).*

For this algorithm to work effectively, we firstly need an efficient way of computing kP mod n. There are a few methods to approach this. For example, using the formulas (1), we can easily compute $(2^i)P \mod n = 2(2(\dots(2P)\dots)) \mod n$ and add points. Therefore, to compute kP, we just need to express k in binary. However, it will be faster to write $kP = \prod_{l \leq B} l^{\alpha_l} P$, and express each of the l^{α_l} in binary, and then do the computation.

The other issue is the choice of a, b, and a point P on E in step (1). In general, one could vary a, x, y, and set $b = y^2 - x^3 - ax$ to ensure P lies on E.

Example 4. We show how the algorithm works in practice by factorising $n = 35$. We choose $y^2 = x^3 - x + 1$ as the elliptic curve with the point $P = (1, 1)$ on it.



The point $8P = \left(\frac{19}{25}, -\frac{103}{125}\right)$ is not well-defined modulo 35. Therefore, we obtain $d = \gcd(25, 35) = 5$, a divisor of 35.

Remark 5. According to [Len87, Sec. 2], this is one of the fastest known factoring methods. However, Lenstra's method is substantially faster, if n has a prime factor much smaller than $\sqrt{n}$.

4. Comparison to Pollard's $p-1$ Method

While Lenstra's factorisation algorithm at first glance looks very surprising, it is actually a very natural idea. In fact, it is the analog of Pollard's $p-1$ method, a known factoring algorithm, with the group $\mathbb{Z}/p\mathbb{Z}$ replaced by E mod p.

We start by recalling the idea of Pollard's $p-1$ method. Fermat's Little Theorem says for a prime $p \nmid a$ and any $K \in \mathbb{N}$, we have

$$a^{K(p-1)} \equiv 1 \mod p.$$

Moreover, if p is a divisor of n and $x \equiv 1$ mod p, then

$$\gcd(x-1, n) = p.$$

To find a divisor of n, we choose an a, a large k, and compute $\gcd(a^k - 1, n)$. If for a divisor p, $p-1|k$, then $\gcd(a^k - 1, n) \neq 1$ will be a divisor of n.

Lenstra's method uses the same idea for the group E mod p instead of $\mathbb{Z}/p\mathbb{Z}$. The analog of Fermat's Little Theorem is, of course:

$$(Ka_p)P = 0 \mod p$$

where $a_p = |E \bmod p|$.

Note that Pollard's $p-1$ method will fail if for each prime divisor p of n, $p-1$ has a large prime factor. The reason Lenstra's algorithm avoids this problem is that a_p will vary for different choices of elliptic curves.

Theorem 6 (Hasse's Bound). *Let p be prime, $q = p^r$, and a_q be the number of $\mathbb{F}_q$-points on an elliptic curve defined over $\mathbb{F}_q$. Then*

$$a_q = q + 1 - t_q,$$

where $|t_q| \leq 2\sqrt{q}$.

If for some prime $p|n$, the number $p + 1 - t_p$ has no large prime factors, the method is likely to yield a divisor of n, and otherwise not. The advantage is that if for a chosen pair (E, P) the method fails, then we simply choose a different pair (E', P') and try again.

References

[Kir92] Frances Kirwan, *Complex algebraic curves*, London Mathematical Society Student Texts, vol. 23, Cambridge University Press, Cambridge, 1992, doi:10.1017/CBO9780511623929. MR 1159092 (93j:14025)

[Kob94] Neal Koblitz, *A course in number theory and cryptography*, second ed., Graduate Texts in Mathematics, vol. 114, Springer-Verlag, New York, 1994, doi:10.1007/978-1-4419-8592-7. MR 1302169 (95h:94023)

[Len87] Hendrik W. Lenstra, Jr., *Factoring integers with elliptic curves*, Ann. of Math. (2) **126** (1987), no. 3, 649–673, doi:10.2307/1971363. MR 916721 (89g:11125)

[ST92] Joseph H. Silverman and John Tate, *Rational points on elliptic curves*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1992, doi:10.1007/978-1-4757-4252-7. MR 1171452 (93g:11003)