Elliptic Curve Cryptography

Before we start talking about Elliptic Curve Cryptography it would be useful to explain the basics of Cryptography. Cryptography is used to securely transmit sensitive information over public channels. Lets propose a situation, Alice is trying to securely communicate with Bob while the malevolent Evelyn overhears everything sent between these two. Alice's goal is to be able to communicate information with Bob without Evelyn ever learning the meaning behind the messages sent between them. In order to achieve this Alice needs to send Bob some function he can use to encrypt his messages.

Assumptions within this hypothetical:

- Every message Alice sends to Bob and Bob sends to Alice will be intercepted by Evelyn
- Evelyn stronger computing power than Alice or Bob but not infinite(in the 100x to 1000x range), so we can't rely on superior computing power or luck.

Here are a few qualifications for a successful encryption scheme.

The function sent to be invertible, If Bob uses the function to encrypt his message but there is no way to restore it to its original state Alice won't be able to understand the message that Bob has sent no mater how hard she tries.

The function needs to be significantly easier for Alice to invert the function than Evelyn. If its equally as difficult for Alice and Evelyn to invert the function Evelyn will always decrypt the function first due to her superior computing power. Ideally it would take astronomically longer for Evelyn to decrypt the information than Alice making it not even worth it for Evelyn to try.

The function needs to be easy for Bob to pass information through. This isn't the most important aspect but it would be annoying for Bob if every time he had encrypt an message it would take a day of processing, especially if there is an urgent message.

How do we plan on achieving all these qualifications

If we release some public function known as our public key that allows our correspondents to encrypt their information, but have some hidden associated function known as a private key or trapdoor function. This trapdoor function allows the sender to more easily decrypt the message that anyone who has to decrypt it with the public key.

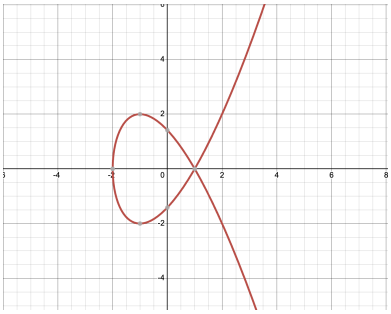
So if we can successfully generate some public key that is a function that is easy to encrypt but extremely hard to decrypt, with some associated private key that allows quick decryption of the information that is kept private, we have a successful encryption scheme.

What is an Elliptic Curve

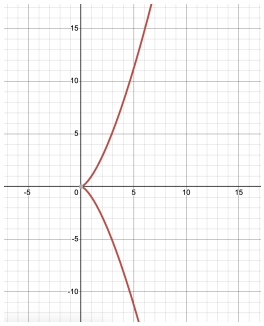
An Elliptic Curve E for some Field F are the points that satisfy the equation $y^2 = x^3 + ax + b$ and a "point at infinity" denoted as $\mathcal{O}$. Additionally the curve must be non-singular meaning that at no point do its partial derivatives vanish simultaneously.

Detecting Singular Elliptic Curves

Below are two examples of curves that are of the form $x^3 + ax + b$ but are singular. The first graph is the equation $y^2 = x^3 - 3x + 2$ and the second graph is the equation $y^2 = x^3$. Notice how it would be impossible to take a derivative at specific points on these graphs.



A Elliptic Curve with a “node” this is where the line intersects itself.



A Elliptic Curve with “Cusp” this is a point where the curve abruptly changes direction

You might be wondering if there are specific conditions that cause these curves to be singular. Rather than going through every case where it isn’t singular lets take a look at when it always is non-singular.

Proof that the partial derivatives of a elliptic curve with three distinct roots will never vanish simultaneously.

Let our elliptic curve be of the form $y^2 = x^3 + ax + b$ where a, b are real numbers, and having three distinct roots r_1, r_2, r_3. Taking the partial derivatives of this equation results in the following.

$\frac{\partial F}{\partial y} = 2y$

$\frac{\partial F}{\partial x} = -3x^2 + a$

The only y value that results $\frac{\partial F}{\partial y} = 0$ is $y = 0$. Only at the roots of $x^3 + ax + b$ does $y = 0$, so the only points where $\frac{\partial F}{\partial y} = 0$ is at the points $(r_1, 0), (r_2, 0), (r_3, 0)$. If $\frac{\partial F}{\partial x} = 0$ at $x = r_1, r_2, r_3$ then both the function $y^2 = x^3 + ax + b$ and its derivative would share a root. A function and its derivative share a root only when there is a repeated root, thus this would contradict the fact that it is composed of three distinct roots. Therefore for an elliptic curve with three distinct roots, whenever $y = 0$(or equivalently $\frac{\partial F}{\partial y} = 0$) implies that $\frac{\partial F}{\partial x} \neq 0$. Thus proving that the partial derivatives never simultaneously vanish for an Elliptic curve with three distinct roots.

Notice that this whole proof hinges on the fact that the polynomial has three distinct roots. So if we had a polynomial with repeated roots there will be some point where $\frac{\partial F}{\partial y} = 0$ and $\frac{\partial F}{\partial x} = 0$ thus having a singularity. So if we want our curve to be non-singular we need to check if its composed out of three distinct roots.

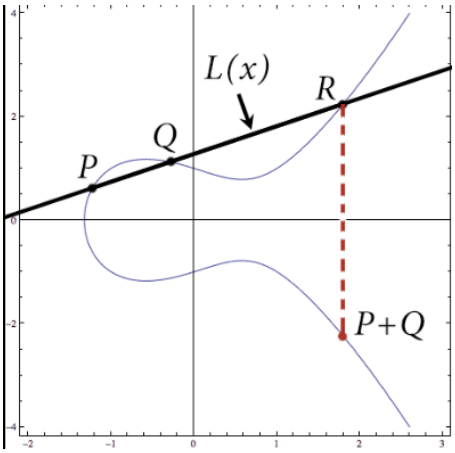
While we could factor it into its roots or use the cubic equation, but as both computer scientists and mathematicians we strive for the fastest and most concise answer possible. Since we don’t need to know what the roots of the equation are, just that repeated roots exist we should strive for a formula that efficiently detects this.

There if a function called the resultant which takes in two polynomials and returns 0 if and only if the polynomials share a root. But the question is what two polynomials to compare using this. If you were to compare a polynomial and itself it would always return 0 no matter what, because every root would be shared. So we rely on the the theorem that if a polynomial and a derivative share a root then that root is repeated. So if we take an arbitrary polynomial of the form $x^3 + ax + b$ and its derivative $3x^2 + a$, then input both of these into the resultant it will return $-16(4a^3 + 27b^2)$.

Therefore for some equation $y^2 = x^3 + ax + b$ instead of checking wether or not the roots factor into distinct roots, we can just check if the equation $-16(4a^3 + 27b^2)$. If this equals zero then the curve is singular and we should pick different a and b values.

Groups within Elliptic Curves

You might be wondering why we are so invested in making that that our curve is differentiable at all points on it. This is actually because you can form a group from the points on the curve using either tangent lines or the line intersecting two points. To "add" to points on the elliptic curve you draw a line intersecting with both of the points and then find where it next intercepts with the curve(beneath is a proof that this point will always exist). Then you flip it over the x-axis and that is your resulting point.

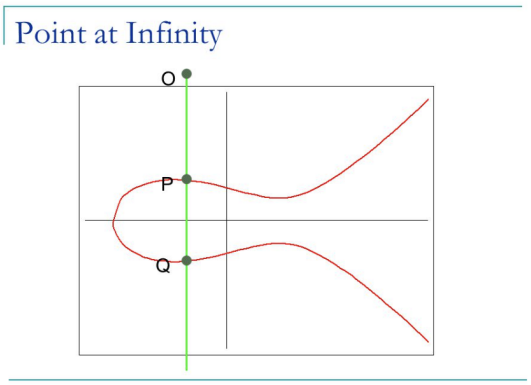


This is a diagram showing the process of adding point P and point Q

Proof that if a line intersects the curve at two points it intersects at a third

Let (x_1, y_1) and (x_2, y_2) be points on our curve E, where E is defined as $y^2 = x^3 + ax + b$ with $-16(4a^3 + 27b^2) \neq 0$. The line intercepting both of these points can be written as $y = mx + t$. The substituting this into our equation we find that $(mx + t)^2 = x^3 + ax + b$ can be rewritten into the form $x^3 - m^2x^2 + (a + 2tm)x + t^2$ which is a cubic equation. Cubic equations have three complex roots, and we already know two real roots x_1 and x_2. The complex conjugate root theorem(sited below) states they if $a + bi$ is a root to a polynomial then $a - bi$ is a root as well. Due to the fact that cubics only have three roots and we have already found two of them, that means that there has to exist a third rational root(or else there would be 2 rational roots and 2 complex roots). Thus it intersects with a third point on the elliptic curve.

The points that this process doesn't work for is the point at infinity $\mathcal{O}$ and points that are reflections of each other across the x-axis. The point at infinity $\mathcal{O}$ can be though of as a point that is infinitely far away, so the line between it and any other point can be interpreted as completely vertical line. This is why the reflection of the point is important because if we didn't reflect it $(a, b) + \mathcal{O}$ would equal $(a, -b)$ but instead with the reflection $(a, b) + \mathcal{O} = (a, b)$ allowing it to act as our identity element. Additionally because $\mathcal{O}$ is infinitely far away if you have $(a, b) + (a, -b)$ it equals $\mathcal{O}$.



The final thing we need to define before point addition is fully working is what happens when you add a point to itself. In this case we just take the tangent line of the elliptic curve at that point and see where it next intersects and then flip that over the x-axis. You might be worried that this line will not intersect with the curve again or intersect with it twice leading to this function not being well defined, but actually this will never be the case. The proof of this is below

Proof a tangent line of a point only intersects with two points

Let (x_0, y_0) be a point on a elliptic curve. The tangent line of this point equals

$$y = \frac{3x_0^2+a}{2y_0}(x - x_0) + y_0$$

Then substituting this into our elliptic curve we get
$(\frac{3x_0^2+a}{2y_0}(x - x_0) + y_0)^2 = x^3 + ax + b$ which then simplifies down into the a cubic equation. This cubic equation has a root with multiplicity of two at x_0, due to the complex conjugate root theorem there must be another rational root to this equation. This other rational root is the x coordinate of the other point of intersection.

Now that we have worked out all the quirks, lets formally define point addition.

Formal definition of point addition

Let E be an elliptic curve of the form $y^2 = x^3 + sx + t$, and let $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ be non-$\mathcal{O}$ points on curve E. We define $P_1 + P_2$ as the following

$$\text{If } P_1 = -P_2 \text{ then } P_1 + P_2 = \mathcal{O}$$
$$\text{Otherwise } P_1 + P_2 = (x_3, y_3) \text{ where } x_3 = \lambda^2 - x_1 - x_2 \text{ and } y_3 = \lambda(x_1 - x_3) - y_1$$
$$\lambda = \begin{cases} \frac{y_2-y_1}{x_2-x_1}, & \text{if } P_1 \neq P_2 \\ \frac{3x_1^2+s}{2y_1}, & \text{if } P_1 = P_2 \end{cases}$$

Additionally note that $-(a, b) = (a, -b)$ because this is the inverse of the point (a, b).

Proof that the set of points on an elliptic curve under point addition$(+)$ is a Group

Identity: For all points (a, b) on the curve, $(a, b) + \mathcal{O} = (a, b)$ by the definition of $\mathcal{O}$.

Associativity: This is true but the proof is pretty long you can prove it algebraically but it takes a long time to write. There is also a book linked below that has a good proof proving this and also elaborating on why it is important to flip over the x-axis within the context of associativity.

Closure: This is proven via the two pervious proofs that prove if you have a tangent line then it will intersect with only one other point, and if you have a line that intersects with two points it will intersect with a third point.

Inverses: By the definition of point addition $(+)$ we for every point (a, b) on the elliptic curve there exists some point $(a, -b)$ such that $(a, b) + (a, -b) = \mathcal{O}$

Calculating the order of an Elliptic curve under a finite field.

We if we have an Elliptic Curve E with points in some finite field F, We denote this Elliptic Curve as $E(F)$. For example an Elliptic Curve with points in $\mathbb{Z}/p$ where p is some prime would be denoted by $E(\mathbb{Z}/p)$. If you have an Elliptic Curve with inputs in a finite field, you will have a finite number of points(Because there is a limited number of possible x and y values). It turns out to be surprisingly easy to count the number of points on the curve, but before we explain how we need to establish what a Legendre Symbol is.

Legendre Symbols

These are denoted as $(\frac{a}{p})$ where p is an odd prime and a is some integer. The formal definition is below

$$(\frac{a}{p}) = \begin{cases} 0 & \text{if } p|a \\ 1 & \text{if } a \text{ is non-zero square} \\ -1 & \text{if } a \text{ is not a square mod p} \end{cases}$$

There are two properties that make computing Legendre symbols incredibly easy. The first of which is the fact that Legendre Symbols are multiplicative meaning that $(\frac{ab}{p}) = (\frac{a}{p})(\frac{b}{p})$, This allows us to break down an incredibly large numbers into a bunch of smaller prime numbers.

The second of which is the law of quadratic reciprocity which states that for Legendre symbol $(\frac{p}{q})$ where p and q are odd primes if p or $q \equiv 1 \bmod 4$ then $(\frac{p}{q}) = (\frac{q}{p})$. If both are p and $q \equiv 3 \bmod 4$

Utilizing both of these properties we can quickly calculate the value of a Legendre Symbol, by breaking a into its prime factors and then using quadratic reciprocity to flip it to work in a smaller finite field. An example of this is shown below

$$(\frac{105}{131}) = (\frac{5}{131})(\frac{7}{131})(\frac{3}{131}) = (\frac{131}{5})(-(\frac{131}{7}))(\frac{131}{3}) = (\frac{1}{5})(-(\frac{5}{7}))(-(\frac{2}{3})) = (1)(1)(1) = 1$$

Counting points on an elliptic curve using these symbols

If you have some value $n \in \mathbb{Z}/p$ you can quickly determine wether or not there is are any points on the Elliptic Curve at $x = n$. The condition for this is that there exists some $m \in \mathbb{Z}/p$ such that $(m)^2 = n^3 + an + b$. A super quick way to check this is by using Legendre Symbols because if $(\frac{n^3+an+b}{p}) = 1$ then there exists some element $m \in \mathbb{Z}/p$ such that $m^2 = n^3 + an + b$. So if $(\frac{n^3+an+b}{p}) = 1$ then we know there is at least one point on the curve at $x = n$. But we actually know there is two points on the curve because if $(m)^2 = n^3 + an + b$ then so does$(-m)^2 = n^3 + an + b$. Therefore if we iterate through all the values in $\mathbb{Z}/p$ and find the Legendre symbol is equal to 1 then we add 2 points, if we find that the Legendre symbol equals -1 then we don't add any points, and if we find that the Legendre symbol equals 0 then we add one point(this will only happen at $x = 0$). The equation for this is below

$$|E(\mathbb{F}_p)| = \Sigma_{a=0}^{p-1} 1 + (\frac{f(a)}{p}) = p + \Sigma_{a=0}^{p-1} (\frac{f(a)}{p})$$

Discrete Logarithm Problem

Let G be some group, let g be an element of G and some integer x. Given elements g and p where $p = g^x$ find the value of $x \bmod |g|$. Beneath we go over a few algorithms that provide us with a solution to this problem.

Brute Force

Start with $i = 1$ then check if $g^i = p$. If $g^i \neq g^x$ then increment the value of i and check again. Do this until you find a value of i such that $g^i = g^x$. At this point $i \equiv x \bmod |g|$ and thus you have solved the discrete logarithm problem. This has a time complexity of $O(n)$ and the pseudo code is beneath.

```
def discrete_log_prob(g,g_x)
    i=0
    while True:
        if g**i==g_x:
            break
        i+=1
    return i
```

Pollard Rho Algorithm

This algorithm is a probabilistic algorithm that solves the discrete log problem. Have have two walkers one at location $(g^{c_1}p^{d_1})$ and the other at $(g^{c_2}p^{d_2})$ where c_1, c_2, d_1, d_1 are random integers. You then have these walkers walk pseudo randomly until you find that $(g^{c_1}p^{d_1}) = (g^{c_2}p^{d_2})$. This pseudo random walk is different within each group but it somehow needs to randomly increment the values of c_1, d_1, c_2, d_2. Once you have found $(g^{c_1}p^{d_1}) = (g^{c_2}p^{d_2})$ you can rewrite this as $(g^{c_1}(g^x)^{d_1}) = (g^{c_2}(g^x)^{d_2})$ which then simplifies to $g^{c_1+xd_1} = g^{c_2+xd_2}$. Then we know that $c_1 + xd_1 \equiv c_2 + xd_2 \bmod |g|$ which can further be rewritten as $c_1 - c_2 \equiv x(d_2 - d_1) \bmod |g|$. If $(d_2 - d_1)$ has a modular inverse we can solve and find that $(c_1 - c_2)(d_2 - d_1)^{-1} \equiv x \bmod |g|$. Thus you have solved the Discrete Logarithm Problem. This has an average time complexity of $O(\sqrt{n})$ making it significantly faster than the brute force approach.

Within specifically elliptic curves we don't know the value of $|g|$ so we instead find the value of $|E|$ and mod it by that instead(Using Legendre Symbols we can really quickly calculate $|E|$). We can do this because $|g|$ divides the order of $|G|$ due to Lagrange's theorem. This gives us some value x_0 such that $x_0 \equiv x \bmod |g|$. Which is all we need to intercept messages within Elliptic-Curve Cryptography.

Actual Elliptic Curve Cryptography

1. Bob generates some elliptic curve E and make sure $4a^3 + 27b^2 \neq 0$. Additionally they generate some large prime p and find some point $P \in E(\mathbb{Z}/p)$. They then make public E, P and p.
2. Bob then generates some random integer a and calculates $aP = Q$. Then he publishes Q as his public key. a is Bob's private key
3. Alice chooses some point M on the curve that somehow encodes her message and computes $C_1 = kP$ where k is some random integer. She then additionally computes $C_2 = M + kQ$. She then sends Bob (C_1, C_2).
4. Bob receives (C_1, C_2) and computes $C_2 + -aC_1 = (M + kQ) + (-a(kP)) = M + kaP - kaP = M$. Bob has received Alice's message M

If someone wants to decrypt a message that they have intercepted they would have to find Bob's private key. They could do this a variety of ways but could use either of the ways of solving the Discrete Logarithm Problem we have talked about previously, but due to the high time complexity of these algorithms.

Sources

General introduction

https://math.uchicago.edu/~may/REU2020/REUPapers/Shevchuk.pdf

Proves that if a polynomial and its derivative share a root its a repeated root

https://lpsa.swarthmore.edu/BackGround/RepeatRootPoly/RepeatedRoots.html

Contains proof that the binary operation on elliptic curves is associative

https://books.google.fr/books?id=mAJei2-JcE4C&pg=PR8&lpg=PR8&dq=tate+silverman&source=bl&ots=MvvlWHLtd6&sig=-sCJ_g-uLJvwagXuulamUFvd0KU&hl=fr&ei=WT2rTtauEuOI4gS2udDZDg&sa=X&oi=book_result&ct=result&resnum=4&ved=0CDsQ6

Complex conjugate root theorem

https://en.wikipedia.org/wiki/Complex_conjugate_root_theorem#:~:text=In mathematics%2C the complex conjugate,also a root of P.