Adding Points on Elliptic Curves

For our purposes, an elliptic curve is the set of all solutions to the equation

\[ y^2 = x^3 + px + q, \]

where \( x \) and \( y \) are complex numbers and the cubic polynomial \( x^3 + px + q \) has no repeated roots. Thus, \( y^2 = x^3 + 2 \) and \( y^2 = x^3 - 2x \) are elliptic curves, since \( x^3 + 2 \) has one real and two complex conjugate roots, and the roots of \( x^3 - 2x \) are 0 and \( \pm \sqrt{2} \). On the other hand, \( y^2 = x^3 + x^2 \) and \( y^2 = x^3 \) are not, since 0 is a double root of \( x^3 + x^2 \) and a triple root of \( x^3 \).

Due both to their connection with Fermat’s Last Theorem and to their usefulness in public key cryptography, elliptic curves have recently become popular objects of study. Many mathematicians had a hand in showing that the construction (called the chord–and–tangent method) can be used to add points on an elliptic curve, and that this addition turns the set of points on such a curve into a group.

Here’s how it works. Suppose \( P \) and \( Q \) are points on the elliptic curve \( E \). Join \( P \) and \( Q \) by the line \( l \). Now \( l \) meets \( E \) in a third point we’ll call \( P \ast Q \). The sum \( P + Q \) is defined to be the reflection of \( P \ast Q \) in the \( x \)-axis, not \( P \ast Q \) itself.

Let’s look at this algebraically. If \( P = (x_1, y_1) \) and \( Q = (x_2, y_2) \), then the line \( l \) has an equation of the form \( y = mx + b \); solving the simultaneous equations \( y = mx + b \) and \( y^2 = x^3 + px + q \) leads to the one–variable equation

\[
\begin{align*}
x^3 - m^2x^2 + (p - 2mb)x + q - b^2 &= 0. \tag{1}
\end{align*}
\]

This cubic polynomial has three roots, namely \( x_1, x_2, \) and the \( x \)-coordinate \( x_3 \) of \( P \ast Q = (x_3, y_3) \). Reflecting \( P \ast Q \) in the \( x \)-axis gives us \( P + Q = (x_3, -y_3) \). For example, let \( E_1 \) be the curve \( y^2 = x^3 - 2x \) (on the left in Figure 1), \( P = (0,0) \) and \( Q = (-1,1) \). Then \( l \) is the line \( y = -x \) and (1) becomes \( x^3 - x^2 - 2x = 0 \), whose roots are 0, -1 and 2. Then \( P \ast Q = (2, -2) \) and so \( P + Q = (2, 2) \).
Let $E_2$ be the curve $y^2 = x^3 + 2$ (on the right in Figure 1); let us add $R = (-1, 1)$ to itself. Then $l$ is the line $y = (3x + 5)/2$ tangent to $E_2$ at $R$ and (1) becomes $(x + 1)^2(x - (17/4)) = 0$, whose roots are $-1$ (a double root) and $17/4$. Then $R \ast R = (17/4, 71/8)$, and so $R + R = (17/4, -71/8)$.

If $P$ and $Q$ have rational coordinates, so do $P + Q$ and $P \ast Q$, because

$$x^3 - m^2x^2 + (p - 2mb)x + q - b^2 = (x - x_1)(x - x_2)(x - x_3).$$

(2)

Since $m^2 = x_1 + x_2 + x_3$ and since $m, x_1$ and $x_2$ are rational, so is $x_3$. Finally, $b = y_1 - mx_1$ is rational and so $y_3 = mx_3 + b$ is also rational. This always works because each line in the plane meets an elliptic curve in three points, provided you count correctly. “Counting correctly” means three things:

- We allow complex coordinates. Thus, you can verify that $y = 2$ meets $E_1$ in the three points $(2, 2), (-1 + i, 2)$ and $(-1 - i, 2)$. Note that in Figure 1, what you see are just the points on the curves with both coordinates real—if we include complex points, then we would have a four-dimensional graph!

- We count multiplicities correctly. Thus, $y = (3x + 5)/2$ meets $E_2$ doubly at $R = (-1, 1)$—since $-1$ is a double root of $(x + 1)^2(x - (17/4))$—and singly at $(17/4, 71/8)$.

As for the line $x = 2$, there is a third point. Look at it this way: the line through $P + Q = (2, 2)$ and $S = (1.999, -1.9975\ldots)$ has equation $y = 3997.5x - 7993$, which is almost vertical. It turns out that

$$(P + Q) \ast S = (15980002.25\ldots, 6388049500.313\ldots).$$

This third point is very far from $P + Q$ and $S$, but it is on the curve. Moving $S$ closer to $(2, -2)$ moves $(P + Q) \ast S$ farther away; passing to the limit, if $S = (2, -2) = P \ast Q$, then $(P + Q) \ast S$ is “infinitely far away.” This last point does not have finite coordinates.

We call it the point at infinity, label it $O$, and include it as a point on every elliptic curve. The third rule for counting correctly is:

- We count the point at infinity, if necessary. Thus, $x = 2$ meets $E_1$ in $(2, 2), (2, -2)$ and $O$; $x = 0$ meets $E_1$ doubly at $(0, 0)$ and singly at $O$; and $x = 1$ meets $E_1$ at $(1, i), (1, -i)$ and $O$.

Note that if $P$ is a point, then the line through $P$ and the reflection of $P$ in the $x$–axis passes through $O$. Using this, we may now tell how to add points on an elliptic curve so as to include the counting rules and the point at infinity. Suppose $P$ and $Q$ are points on the elliptic curve $E$. To find $P + Q$, draw the line $l$ through $P$ and $Q$; if $P = Q$, then $l$ is the tangent line to $E$ at $P$. Locate $P \ast Q$, the third point at which
\textit{l} meets \textit{E}—counting correctly. Draw \textit{l}', the line through \textbf{O} and \( P \ast Q \); \( P + Q \) is the third point at which \textit{l}' meets \textit{E}.

As with addition of numbers, we write \( 2P \) for \( P + P \), \( 3P \) for \( P + P + P \), etc.

We can do this algebraically, too. If \( E : y^2 = x^3 + px + q \) is an elliptic curve, then we can express the sum \( P_1 + P_2 \) of points \( P_1 \) and \( P_2 \) on \( E \) by means of the following formulas. Let \( P_1 = (x_1, y_1) \), \( P_2 = (x_2, y_2) \) and \( P_1 + P_2 = (x_3, y_3) \).

If \( x_1 = x_2 \) and either \( y_1 \neq y_2 \) or \( y_1 = y_2 = 0 \), then \( P_1 + P_2 = O \), and we say that \( P_2 = -P_1 \).

Otherwise, the slope \( m \) of the line \( l \) through \( P_1 \) and \( P_2 \) is given by

\[
m = \begin{cases} 
\frac{y_2 - y_1}{x_2 - x_1}, & \text{if } x_1 \neq x_2; \\
\frac{3x_1^2 + p}{2y_1}, & \text{if } x_1 = x_2.
\end{cases}
\]

Finally, it follows from the discussion following (2), the fact that \( P_1 \ast P_2 = (x_3, -y_3) \), and a little algebra, that

\[
\begin{align*}
x_3 &= m^2 - x_1 - x_2, \\
y_3 &= -(y_1 + m(x_3 - x_1)).
\end{align*}
\]

For example, if \( E \) has equation \( y^2 = x^3 - x + 4 \), \( P_1 = (0, 2) \) and \( P_2 = (-1, -2) \), then you can check that \( (0, 2) + (-1, -2) = (17, -70) \). Sure enough, \((-70)^2 = 4900 = 17^3 - 17 + 4\), so \((17, -70)\) is on \( E \). To test these formulas, note that the point \((15, 58)\) is also on this curve. Find \((17, -70) + (15, 58)\) for yourself; surprised?