Recall that $\mathcal{E}(\mathbb{Q})$ denotes the set of all rational points on an elliptic curve $\mathcal{E}$, and that by convention we're including the point at infinity (denotes ∞) as one of the elements of $\mathcal{E}$. As we saw, this makes $\mathcal{E}(\mathbb{Q})$ into an abelian group under the operation $*$, where the point $P * Q$ is defined by: (a) drawing a line passing through P and Q, (b) finding where it intersects the elliptic curve, and then (c) reflecting that point across the x-axis. One remarkable result I mentioned in passing (Mazur's theorem from 1977) is that either $\mathcal{E}(\mathbb{Q})$ is an infinite set or $1 \leq |\mathcal{E}(\mathbb{Q})| \leq 16$; moreover, $\mathcal{E}(\mathbb{Q}) \neq 11$. In fact, Mazur proved something rather more spectacular: he completely characterized the structure of the group $\mathcal{E}(\mathbb{Q})$. But before we can describe this, we must discuss the concept of *isomorphism*, one of the most important notions in mathematics. We approach this in an unexpected way, by playing a game!

Isomorphisms and the Game of 15

The game is called the *Game of 15*. It is a 2 player game. You begin with numbers $1, 2, \ldots, 9$ all available, and the players take turns selecting numbers (without replacement) from the list to add to their own individual collection of numbers. You win iff you have 3 (distinct) numbers in your collection that sum to 15.

Example Game 1: Leo vs Suzanna

Turn	L	S	L	S	L	S	L	
L:	5	5	5, 6	5, 6	5, 6, 1	5,6,1	**5**,6,**1**,**9**	win!
S:		7	7	7, 4	7, 4	7,4,8	7,4,8	

Note that in her final move, Suzanna picked 8 to prevent Leo from winning with 8,6,1. But Leo had set up a collection that would win with either an 8 or a 9, and Suzanna could not prevent both options!

Example Game 2:

Turn	S	L	S	L	S	L	S	
L:		5	5	5, 8	5, 8	5, 8, 7	5, 8, 7	
S:	6	6	6, 4	6, 4	6, 4, 2	6, 4, 2	6, **4**, **2**, **9**	win!

Several of you noted that this game *feels* like tic-tac-toe. Moreover, the conditions for winning the game (having three numbers between 1 and 9 that sum to 15) is reminiscent of a 3×3 magic square, in which each row, each column, and each of the two main diagonals consists of three numbers that sum to 15:

8	1	6
3	5	7
4	9	2

Combining these two ideas, we see that the game of 15 is like playing tic-tac-toe on a 3×3 magic square. How similar are these games? Some thought shows that they're *the same game* – just played with different notation. Thus we say the game of 15 and tic-tac-toe are *isomorphic*; they are two perspectives on the same game.

Two groups can also be isomorphic, e.g.
1) We checked that $\{1, i, i^2, i^3\}$ is a group under multiplication. This is isomorphic to the group $\mathbb{Z}_4$ under addition (mod 4). We denote this $\{1, i, i^2, i^3\} \cong \mathbb{Z}_4$. (Note that this notation suppresses the operation used to make each set into a group – the operation is implicit.)
2) Recall that we proved that $\mathbb{Z}_{mn}^\times$ has the same number of elements as $\mathbb{Z}_m^\times \times \mathbb{Z}_n^\times$. It turns out that more is true: $\mathbb{Z}_{mn}^\times \cong \mathbb{Z}_m^\times \times \mathbb{Z}_n^\times$ as groups. Here $\mathbb{Z}_{mn}^\times$ is a group under multiplication (mod mn), and $\mathbb{Z}_m^\times \times \mathbb{Z}_n^\times$ is a group under coordinate-wise multiplication (mod m) and (mod n) for the first and second coordinates, respectively; this is precisely the operation that was used in our proof of Quadratic Reciprocity.

Mazur's theorem

The concept of isomorphism was fundamental in the development of math in the 20th century, because it allows for a discussion of the behavior of an object rather than the specific language used to describe that object. A potent example of this is

Structure theorem for finitely generated abelian groups: Any finitely generated abelian group is isomorphic to a product of finitely many copies of $\mathbb{Z}$ with finitely many $\mathbb{Z}_n$'s.

In particular, recalling Mordell's theorem that $\mathcal{E}(\mathbb{Q})$ is a finitely-generated abelian group[1], we see that

$$\mathcal{E}(\mathbb{Q}) \cong \mathbb{Z}^r \times \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \cdots \times \mathbb{Z}_{n_k}.$$

Mazur was able to characterize what the finite part of this looks like:

Theorem (Mazur, 1977): For any elliptic curve $\mathcal{E}/\mathbb{Q}$, either
- $\mathcal{E}(\mathbb{Q}) \cong \mathbb{Z}^r \times \mathbb{Z}_n$ for some $n \in \{1, 2, \ldots, 10, 12\}$, or
- $\mathcal{E}(\mathbb{Q}) \cong \mathbb{Z}^r \times \mathbb{Z}_2 \times \mathbb{Z}_{2n}$ for some $n \in \{1, 2, 3, 4\}$.

[1] We stated Mordell's theorem about $\mathcal{E}(\mathbb{Q})$ with respect to $\oplus$, but it applies equally well to $\mathcal{E}(\mathbb{Q})$ with respect to $*$.

This theorem justifies our earlier assertion about the number of rational points on an elliptic curve: if $r \geq 1$ there are infinitely many points, while if $r = 0$ there must be $1, 2, \cdots, 10, 12$, or 16 rational points on $\mathcal{E}$. But Mazur's theorem says much more: it tells us about the relationship of these points to one another. For example, if $\mathcal{E}(\mathbb{Q}) \cong \mathbb{Z}_n$ then there exists a rational point P that generates *all* the other points: they are ∞, P, $P * P$, $(P * P) * P$, etc. By contrast, if $\mathcal{E}(\mathbb{Q}) \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ then *every* rational point is its own inverse, which means that (apart from the point at infinity) $\mathcal{E}$ has three rational points, and the line tangent to $\mathcal{E}$ at each of these is vertical.

If $r \geq 1$ in Mazur's theorem, then $\mathcal{E}(\mathbb{Q})$ has infinitely many rational points. The number r, called the *rank* of $\mathcal{E}$, remains mysterious despite decades of intense research. Here are some open questions and theorems concerning the rank:

1. Folklore (perhaps due to Tate?) Conjecture: The rank is unbounded, i.e. for any $r \geq 0$ there exists some $\mathcal{E}$ with rank $\geq r$. All that's currently known is that there exist infinitely many elliptic curves with rank ≥ 19, and at least one elliptic curve with rank ≥ 28; both of these results are due to Elkies. This seems like tenuous evidence for the conjecture, and indeed, there are good reasons to disbelieve it!

2. Conjecture (Park-Poonen-Voight-Matchett Wood, 2019): There are at most finitely many elliptic curves (up to isomorphism) of rank ≥ 22. This conjecture follows from a careful analysis of a probabilistic model.

3. Conjecture (Katz-Sarnak): The average rank is $\frac{1}{2}$ (meaning take first n elliptic curves and average their ranks, where we order the curves by "naive height" using a and b). More precisely, the conjecture is the 50% of curves have rank 0 and 50% have rank 1, and while there do exist curves with higher ranks they're extremely rare. An older conjecture in the same vein, due to Goldfeld, asserts that for any given elliptic curve $\mathcal{E}$ the family of its 'quadratic twists' have rank 0 and 1 equally often, and rank ≥ 2 occurs 0% of the time.

4. Theorem (Bhargava-Shankar, 2015): The average rank of an elliptic curve is $\leq \frac{7}{6}$, and a positive proportion of elliptic curves have rank 0.

5. Theorem (Alex Smith, 2017): The Birch–Swinnerton-Dyer conjecture (a widely-believed conjecture that we'll describe in the next section) implies Goldfeld's conjecture.

Local solutions and the Sato-Tate conjecture

Thus far we've been focusing on $\mathcal{E}(\mathbb{Q})$, i.e. the points $(x,y) \in \mathbb{Q}^2$ satisfying the equation of $\mathcal{E}$. This is a difficult problem, because it's not obvious how to find *any* rational points, much less find all of them. Instead, let's consider a much simpler problem: finding the points in $\mathcal{E}(\mathbb{Z}_p)$, i.e. the points $(x,y) \in \mathbb{Z}_p^2$ satisfying the equation of $\mathcal{E}$. This is easier because for any particular prime p, the set $\mathcal{E}(\mathbb{Z}_p)$ is finite and can be found by brute force! For example, if $\mathcal{E}$ is the elliptic curve $y^2 = x^3 - 2$, you should verify that $\mathcal{E}(\mathbb{Z}_5)$ consists of five points:

$$\mathcal{E}(\mathbb{Z}_5) = \{(1, \pm 2), (2, \pm 1), (3, 0)\}.$$

Remark. In the literature it's common to include a 'point at infinity' in the set $\mathcal{E}(\mathbb{Z}_p)$, by analogy with the case of rational points. Thus, for the example above, many authors would say that $\mathcal{E}(\mathbb{Z}_5)$ contains 6 points. We will *not* follow this convention – it adds confusion and only becomes important when generalizing to other contexts – but do keep this in mind if you look at other sources.

Since $\mathcal{E}(\mathbb{Z}_p)$ is always finite, we can immediately ask how big it is. Trivially we have

$$|\mathcal{E}(\mathbb{Z}_p)| \le p^2,$$

since there are p^2 points in $\mathbb{Z}_p^2$. A bit more thought shows that

$$|\mathcal{E}(\mathbb{Z}_p)| \le 2p$$

(you should prove this!). But it turns out that one can do better: in his thesis, Emil Artin conjectured that $|\mathcal{E}(\mathbb{Z}_p)| \sim p$, or in other words,

$$\lim_{p \to \infty} \frac{|\mathcal{E}(\mathbb{Z}_p)|}{p} = 1.$$

A decade later, Hasse proved this. In fact, he was able to measure not just the ratio but the difference:

Theorem (Hasse, 1933): Let $a_p := p - |\mathcal{E}(\mathbb{Z}_p)|$. Then $|a_p| \le 2\sqrt{p}$.

To fully appreciate the strength of this result, it might be helpful to recall the situation of the distribution of prime numbers: the prime number theorem asserts $\pi(x) \sim \int_2^x \frac{dt}{\log t}$, while the Riemann Hypothesis gives the error term $\pi(x) = \int_2^x \frac{dt}{\log t} + O(x^{1/2+\epsilon})$. Thus Hasse proves an analogue of the Riemann Hypothesis for the size of $\mathcal{E}(\mathbb{Z}_p)$. This is just an analogy, of course, but as we'll see shortly there's more to it than meets the eye.

Although the behavior of the error term a_p in Hasse's bound is complicated, computations led Sato and Tate to independently conjecture that a_p has a nice probabilistic distribution:

they guessed that the normalized error $\frac{a_p}{2\sqrt{p}}$ is distributed according to the *semicircle law*, i.e. the proportion of the time that $\frac{a_p}{2\sqrt{p}}$ is in some range is the area under the unit semicircle in that range. A decade ago, their conjecture was proved (by Clozel–Harris–Shepherd-Barron–Taylor in a special case, and subsequently by Barnet-Lamb–Geraghty–Harris–Taylor in general). Here's the formal statement:

Theorem (formerly the Sato-Tate Conjecture): For any $I \subseteq [-1, 1]$, we have

$$\frac{1}{\pi(N)} \left| \{p \le N : \frac{a_p}{2\sqrt{p}} \in I\} \right| \quad \xrightarrow{N\to\infty} \quad \frac{2}{\pi} \int_I \sqrt{1-t^2}\, dt.$$

(Actually this is only for elliptic curves without 'complex multiplication'.)

A stronger version of this conjecture that quantifies the rate at which the convergence takes place was stated in a 1999 paper of Akiyama and Tanigawa:

Conjecture (Akiyama-Tanigawa, 1999):

$$\frac{1}{\pi(N)} \left| \{p \le N : \frac{a_p}{2\sqrt{p}} \in I\} \right| = \frac{2}{\pi} \int_I \sqrt{1-t^2}\, dt + O\left(N^{-1/2+\epsilon}\right)$$

This conjecture, which is still open, is reminiscent of the Riemann Hypothesis. In fact, Akiyama-Tanigawa prove that their conjecture implies the Generalized Riemann Hypothesis, so presumably their conjecture is quite hard to prove.

The Birch–Swinnerton-Dyer conjecture

Recall that the error term given by Hasse's theorem is

$$a_p := p - |\mathcal{E}(\mathbb{Z}_p)|;$$

these numbers go by the fancy name *trace of Frobenius*. The Sato-Tate conjecture shows that the a_p behave like random fluctuations, but it turns out they have a lot of structure! To describe this, we first encode these numbers into an analogue of the Riemann zeta function called the L-function associated to the elliptic curve $\mathcal{E}$:

$$L(s, \mathcal{E}) := \prod_p \left(1 - \frac{a_p}{p^s} + \frac{1}{p^{2s-1}}\right)^{-1}$$

(the actual definition is a bit more technical – there are a finite number of primes for which the factor is different from what's written above – but the above is essentially correct). If we expand this product, we obtain a series of the following form:

$$L(s, \mathcal{E}) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}.$$

(It's a good exercise to figure out a_{12} in terms of the a_p's.) This looks like the Riemann ζ function but with more complicated coefficients a_n.

It turns out that the function $L(s,\mathcal{E})$, which looks pretty randomly defined, enjoys some nice properties. For example, it converges for any complex number s with real part larger than $\frac{3}{2}$. Moreover, there exists a unique way to extend the range of definition to a function that's defined – and even differentiable – for all s in the complex plane; since this extension is unique, we abuse notation and refer to this extended function as $L(s,\mathcal{E})$. This extended function exhibits a remarkable symmetry: there's an explicit relationship one can write down between $L(s,\mathcal{E})$ and $L(2-s,\mathcal{E})$.

Perhaps the most surprising feature of this elliptic curve L-function is that, even though it's built out of information about $\mathcal{E}(\mathbb{Z}_p)$, it gives information about $\mathcal{E}(\mathbb{Q})$. This is the famous Birch–Swinerton-Dyer conjecture, widely considered one of the most important open problems today:

<u>Conjecture (Birch and Swinnerton-Dyer):</u> If $\mathcal{E}/\mathbb{Q}$ has rank r, then $L(s,\mathcal{E})$ vanishes with multiplicity r at $s=1$. In other words, $L(s,\mathcal{E})=(s-1)^r g(s)$ where $g(1)\neq 0$ or ∞.

Although this is wide open, we do know that it holds a lot of the time:

<u>Theorem (Bhargava-Shankar):</u> A positive proportion of elliptic curves satisfy BSD.

To recap: using the error terms a_p from Hasse's theorem about counting the points in $\mathcal{E}(\mathbb{Z}_p)$, we created a strange function $L(s,\mathcal{E})$ that has nice properties, which in turn seems to tell us about rational points on elliptic curves. This is one justification for saying that the a_p have structure. But wait... there's more!

Fermat's Last Theorem

Recall from above that the error terms a_p can be extended to a sequence of numbers (a_n) by expanding the initial product definition of $L(s,\mathcal{E})$ into a series. In the mid-20th century, Taniyama considered what would happen if we used the sequence (a_n) as the coefficients of a Fourier series, i.e. he considered the function

$$F(z) := \sum_{n=1}^{\infty} a_n e(nz).$$

He empirically discovered that $F(z)$ has some remarkable properties, and (together with Shimura) put forward the following guess:

Conjecture (Taniyama-Shimura): Given an elliptic curve $\mathcal{E}/\mathbb{Q}$, the associated fourier expansion

$$F(z) := \sum_{n=1}^{\infty} a_n e(nz)$$

is differentiable in the upper half $\mathbb{H}$ of the complex plane (i.e. for all complex z with positive imaginary part), and satisfies the functional equation

$$F(-1/z) = z^2 F(z)$$

for all $z \in \mathbb{H}$.

In fancy language, this conjecture asserts that F is a *modular form*, more precisely a *cusp form of weight 2*, and so any elliptic curve for which the Taniyama-Shimura conjecture holds is said to be *modular* (because its a_p's produce a modular form). If you've thought about modular forms before then this is really nifty, but even if you've never heard of modular forms, the conjecture is still easy to appreciate: it asserts that the strange sequence (a_n), itself derived from the sequence (a_p) of error terms in Hasse's theorem, always produces a Fourier series with a remarkable self-symmetry.

In 1993/4, Wiles proved that a large class of elliptic curves are modular, and shortly thereafter the full Taniyama-Shimura conjecture was proved by Breuil, Conrad, Diamond, and Taylor; it is now usually called the Modularity Theorem. These developments would have garnered interest in certain circles, but they really became a big deal because of an earlier observation due to Frey, subsequently made precise by Serre and Ribet:

Theorem: If $a^p + b^p = c^p$ has a nontrivial solution, then the elliptic curve $y^2 = x(x - a^p)(x + b^p)$ isn't modular.

Since the modularity theorem (a.k.a. the Taniyama-Shimura conjecture) implies that every elliptic curve is modular, we deduce that Fermat's Last Theorem must hold!