Chapter 1

Three Motivating Problems

The goal of this book is to explain how the set of points on an elliptic curve can be given the structure of an abelian group, and how the arithmetic of elliptic curves over finite fields can be used as a powerful tool in cryptography and cryptanalysis. Perhaps to put it another way, the goal of this book is to help the reader understand the first sentence.

To motivate our study of elliptic curves, we consider three problems in number theory and geometry whose solutions use elliptic curves in an essential, if sometimes subtle, manner. Two of these problems, Fermat's Last Theorem and the congruent number problem, are problems whose statements are completely elementary. They are classical in feel, perhaps almost playful in tone. In contrast the third problem, applications of the theory of elliptic curves to cryptography, is a quite modern subject whose practical importance has grown enormously of late. In 2011 Koblitz et al. [**KKM11**] wrote that "over a period of sixteen years, elliptic curve cryptography went from being an approach that many people mistrusted or misunderstood to being a public key technology that enjoys almost unquestioned acceptance." A key question for us to address is, What are these elliptic

curves which have had such an impact, and how have they proven to be so important?

At first blush, elliptic curves appear to be nothing special. For us they will just be certain curves given by polynomials of degree 3, but before we even talk about cubics, let's back up a bit. In secondary school it is common to study the properties of lines and conics in the plane as well as the principles of Euclidean geometry. For example, if you were to consider the quadratic equation $y = x^2 + 2$, you might think of the set of all the points (x, y) in the plane which satisfy that equation as a geometric object, the parabola with line of symmetry the y-axis, vertex $(0, 2)$, and opening upward. It might also be interesting to consider what happens when we consider the points (x, y) which satisfy that equation when we restrict or allow the coordinates to come from different domains. For example, we could ask for the *rational points*, that is those points (x, y) in which both $x, y \in \mathbb{Q}$ (are rational numbers) such as $(0, 2)$ or $(1/3, 19/9)$, though you might question why that would be a useful thing to do. We shall answer that question later in the book, but for now let's simply observe that some but not all points on the curve have rational coordinates, e.g., $(\pm\sqrt{2}, 4)$, or $(\sqrt[3]{2}, 2 + \sqrt[3]{4})$ are real, but not rational points. Analogously, we might ask about complex points (points with complex coordinates). Again, why would we do that? Well, at least here, your previous experience in mathematics affords you some insight. If we asked what are the roots of $x^2 + 2$, you would either have said there are none or they are complex (imaginary). Said another way, $(\pm i\sqrt{2}, 0)$ are two complex points on the curve $y = x^2 + 2$ which are not on the real locus, the curve we draw. While we cannot fully appreciate this comment so early in the book, it is perhaps not too surprising that an interest in rational points should somehow be connected to number theory, since rational numbers are the quotients of integers, the domain of number theory. It is also the case that the complex points are often more interesting than the real points since there are more of them and since the set of complex solutions may have an even more interesting structure than we might first think. Indeed, to use the word "structure" in the context of the set of points on a curve is quite intentional, and in some sense it represents the origins of most of the applications we shall discuss.

For us an elliptic curve will be a the set of points (x, y) which satisfy an equation of the form

$$y^2 = x^3 + ax^2 + bx + c,$$

where the cubic, $x^3 + ax^2 + bx + c$, is nonsingular (has distinct roots). Our interests will include a consideration of the set of solutions (x, y) where x, y are restricted to different domains, and indeed we will not only be interested in points whose coordinates are rational, real, or complex numbers, but also those whose coordinates lie in so-called *finite fields*, which we have yet to define. But even we have to admit that despite the build up we have given to elliptic curves so far, the definition seems quite lackluster; in particular, the definition seems to shed no light at all on why elliptic curves should play such a pivotal role in number theory. Yet be assured that they do, and as well provide some of the best schemes for public-key cryptography. Certainly we shall take a closer look at all these things.

Among the three motivating problems, we shall look only briefly at the Fermat problem, slightly more deeply at the congruent number problem, and most deeply at developing an understanding of how elliptic curves are of critical use in cryptography. This emphasis is deliberate, in part because the role elliptic curves play in the solutions of the Fermat and the congruent number problems is more subtle and much more sophisticated, and in part because such an omission provides the opportunity for the reader to do some investigation on her own.

1.1. Fermat's Last Theorem

This theorem, conjectured by Fermat in 1635, states simply that for $n > 2$ the equation $x^n + y^n = z^n$ has no solutions in the integers except when one of the variables is zero. We note that this contrasts sharply with the case of $n = 2$ for which solutions (Pythagorean triples) abound. In fact, Pythagorean triples will play an integral role in the congruent number problem, but before leaping to make that connection, we need to say a few more words about the Fermat theorem.

In 1640, Fermat himself proved that the conjecture was true for $n = 4$ and noted that if $n = km$, the existence of a nontrivial solution to $x^n + y^n = z^n$ implied the existence of nontrivial solutions to $(x^k)^m + (y^k)^m = (z^k)^m$, that is to a Fermat problem whose exponent is a divisor of the original. Fermat's $n = 4$ result and this observation reduces the proof of the Fermat conjecture to showing there are no nontrivial solutions to $x^p + y^p = z^p$, where p is an odd prime.

Until 1839, only the cases $p = 3, 5, 7$ had been resolved; this included Sophie Germain's important work which eventually allowed the conjecture to be proved for all odd primes less than 100. In the 1850s, Ernst Kummer developed techniques to prove the Fermat conjecture for all "regular" primes, which is believed to be an infinite family. Modern computing methods verified the conjecture for primes less than four million.

The first real breakthrough came in 1985 when Gerhard Frey suggested that if there were a counterexample to Fermat's conjecture, it could be used to create an elliptic curve having properties which would provide a counterexample to yet another unproved conjecture due to Taniyama and Shimura. While the Fermat conjecture enjoyed a reputation as a long-standing open problem in mathematics, the Taniyama–Shimura conjecture had deep implications for how the theory of modular forms and elliptic curves fit together. If this later conjecture had been false, it would have been quite disappointing.

In the period 1985–1986, Jean-Pierre Serre showed how the Taniyama–Shimura conjecture together with another smaller conjecture — termed the "epsilon conjecture" — would imply Fermat's theorem. Ken Ribet proved the epsilon conjecture in 1986 reducing the Fermat theorem to a proof of the Taniyama–Shimura conjecture for a special class of elliptic curves. In 1994, Andrew Wiles (after seven or more years of intense work, together with a last minute assist by Richard Taylor) succeeded in proving the required case of the Taniyama–Shimura conjecture, and hence proving Fermat's Last Theorem. And in case you were wondering, the full Taniyama–Shimura conjecture has now been proven as well.

1.2. The Congruent Number Problem

A positive integer is called a *congruent number* if it is the area of a right triangle whose sides all have rational length. For example, 6 is a congruent number since 6 is the area of a 3-4-5 right triangle.

It is also true that 5 is a congruent number, though this is something you might not guess right off. But indeed, 5 is the area of the right triangle with sides: 3/2, 20/3, 41/6. Any reasonable person would agree that one can check the result, but it is certainly mysterious how one would come up with a triangle having those sides. However, with an unexpected solution in hand, it now becomes a much more interesting question to ask which integers are congruent numbers. For example, later in the book we shall see that 157 is a congruent number. Surely it can't be that hard to check that such a small number is or is not a congruent number. But we shall see that the answer to this question is more elusive than it may first appear.

A key observation in characterizing congruent numbers is that if N is a congruent number, then Nt^2 is also for any positive integer t; indeed if N is the area of a triangle having rational sides a, b, c, then Nt^2 is the area of a triangle with sides at, bt, ct. Let's consider the triangle showing that 5 is a congruent number. It is easily seen that 6 is the common denominator of the rational numbers 3/2, 20/3, 41/6, and from our observation above, since 5 is a congruent number, $5 \cdot 6^2$ is also, being the area of a right triangle with sides $9 = 6 \cdot \frac{3}{2}$, $40 = 6 \cdot \frac{20}{3}$, and $41 = 6 \cdot \frac{41}{6}$. But of course this means that 9, 40, 41 is a Pythagorean triple! Conversely, suppose that A, B, C are a Pythagorean triple, and N is the area of the corresponding right triangle. Write $N = N_0 t^2$ where N_0 is square free (1 or the product of distinct primes), $t > 0$. Then N_0 is a congruent number, being the area of a right triangle with rational sides $A/t, B/t, C/t$.

So there is a clear relationship between congruent numbers and Pythagorean triples, which means if we had a way to list all Pythagorean triples, we would know which numbers were congruent numbers. In fact, we will show how to list all the Pythagorean triples! Unfortunately, the congruent numbers that come out of the list do not appear

in any particular order and are often repeated, so this procedure cannot definitely answer whether a given integer is a congruent number. Still it will provide a good deal of insight into the connections between algebra and geometry, so we will spend significant time with it.

As of this writing, the congruent number problem remains open, though many partial results are known. Jerrold Tunnell [**Tun83**] developed a condition based on the arithmetic of elliptic curves (and yet another open conjecture—the Birch and Swinnerton-Dyer conjecture) which provides a beautiful answer to this question. We discuss Tunnell's approach at the end of the book.

1.3. Cryptography

Cryptography is a subject that has a long and fascinating history and is a matter of critical importance to all of us in an age when so many transactions happen electronically. It is the subject around which essentially all the background material on number theory and algebra that we develop in this book will be focused.

There are many interesting questions of a practical nature which cryptography solves and which we shall examine, but as a teaser, we mention only a few in this introductory chapter. It is not terribly difficult to send private messages to a friend even over an insecure channel. What becomes trickier is when you want to do the same with someone you don't know. Why would you want to do that? Well, every time you order something online, you want to communicate securely with the vendor so that confidential information (e.g., a credit card number) is not revealed over the insecure web. But how can you (that is, your computer) and your vendor do this? On a different note, how can someone who has received an email from you prove to a third party that the message is indeed from you and not someone forging your address? Or, how can someone be sure a message has not been tampered with (e.g., when a bank receives a message to transfer funds from one account to another)?

All of these are vital questions that modern cryptography answers effectively, and elliptic curves figure prominently in the mix. Of course given any cryptographic system, there are many individuals who do their best to break it, so we will look at standard kinds of

cryptographic attacks on various systems, and how vulnerable each system is to different types of attacks. In the end, elliptic curve cryptography turns out to be among the best public-key cryptosystems currently in use.