Preface

This book grew out of the lecture notes for a course on "Elliptic Curves, Modular Forms and L-functions" that the author taught at an undergraduate summer school as part of the 2009 Park City Mathematics Institute. These notes are an *introductory survey* of the theory of elliptic curves, modular forms and their L-functions, with an emphasis on examples rather than proofs. The main goal is to provide the reader with a *big picture* of the surprising connections among these three types of mathematical objects, which are seemingly so distinct. In that vein, one of the themes of the book is to explain the statement of the modularity theorem (Theorem 5.4.6), previously known as the Taniyama-Shimura-Weil conjecture (Conjecture 5.4.5). In order to underscore the importance of the modularity theorem, we also discuss in some detail one of its most renowned consequences: Fermat's last theorem (Example 1.1.5 and Section 5.5).

It would be impossible to give the proofs of the main theorems on elliptic curves and modular forms in one single course, and the proofs would be outside the scope of the undergraduate curriculum. However, the definitions, the statements of the main theorems and their corollaries can be easily understood by students with some standard undergraduate background (calculus, linear algebra, elementary number theory and a first course in abstract algebra). Proofs that are accessible to a student are left to the reader and proposed as exercises

at the end of each chapter. The reader should be warned, though, that there are multiple references to mathematical objects and results that we will not have enough space to discuss in full, and the student will have to take these items on faith (we will provide references to other texts, however, for those students who wish to deepen their understanding). Some other objects and theorems are mentioned in previous chapters but only explained fully in later chapters. To avoid any confusion, we always try to clarify in the text which objects or results the student should take on faith, which ones we expect the student to be familiar with, and which will be explained in later chapters (by providing references to later sections of the book).

The book begins with some motivating problems, such as the congruent number problem, Fermat's last theorem, and the representations of integers as sums of squares. Chapter 2 is a survey of the algebraic theory of elliptic curves. In Section 2.9, we give a proof of the weak Mordell-Weil theorem for elliptic curves with rational 2-torsion and explain the method of 2-descent. The goal of Chapter 3 is to motivate the connection between elliptic curves and modular forms. To that end, we discuss complex lattices, tori, modular curves and how these objects relate to elliptic curves over the complex numbers. Chapter 4 introduces the spaces of modular forms for $\mathrm{SL}(2,\mathbb{Z})$ and other congruence subgroups (e.g., $\Gamma_0(N)$). In Chapter 5 we define the L-functions attached to elliptic curves and modular forms. We briefly discuss the Birch and Swinnerton-Dyer conjecture and other related conjectures. Finally, in Section 5.4, we justify the statement of the original conjecture of Taniyama-Shimura-Weil (which we usually refer to as the modularity theorem, since it was proved in 1999); i.e., we explain the surprising connection between elliptic curves and certain modular forms, and justify which modular forms correspond to elliptic curves.

In order to make this book as self-contained as possible, I have also included five appendices with concise introductions to topics that some students may not have encountered in their classes yet. Appendix A is a quick reference guide to two popular software packages: PARI and Sage. Throughout the book, we strongly recommend that the reader tries to find examples and do calculations using one of these

two packages. Appendix B is a brief summary of complex analysis. Due to space limitations we only include definitions, a few examples, and a list of the main theorems in complex analysis; for a full treatment see [**Ahl79**], for instance. In Appendix C we introduce the projective line and the projective plane. The p-adic integers and the p-adic numbers are treated in Appendix D (for a complete reference, see [**Gou97**]). Finally, in Appendix E we list infinite families of elliptic curves over $\mathbb{Q}$, one family for each of the possible torsion subgroups over $\mathbb{Q}$.

I would like to emphasize once again that this book is, by no means, a thorough treatment of elliptic curves and modular forms. The theory is far too vast to be covered in one single volume, and the proofs are far too technical for an undergraduate student. Therefore, the humble goals of this text are to provide a *big picture* of the vast and fast-growing theory, and to be an "advertisement" for undergraduates of these very active and exciting areas of number theory. The author's only hope is that, after reading this text, students will feel compelled to study elliptic curves and modular forms in depth, and in all their full glory.

There are many excellent references that I would recommend to the students, and that I have frequently consulted in the preparation of this book:

(1) There are not that many books on these subjects at the *undergraduate level*. However, Silverman and Tate's book [**SiT92**] is an excellent introduction to elliptic curves for undergraduates. Washington's book [**Was08**] is also accessible for undergraduates and emphasizes the cryptography applications of elliptic curves. Stein's book [**Ste08**] also has an interesting chapter on elliptic curves.

(2) There are several *graduate-level* texts on elliptic curves. Silverman's book [**Sil86**] is the standard reference, but Milne's [**Mil06**] is also an excellent introduction to the theory of elliptic curves (and also includes a chapter on modular forms). Before reading Silverman or Milne, the reader would benefit

from studying some algebraic geometry and algebraic number theory. (Milne's book does not require as much algebraic geometry as Silverman's.)

(3) The theory of modular forms and L-functions is definitely a *graduate topic*, and the reader will need a strong background in algebra to understand all the fine details. Diamond and Shurman's book [**DS05**] contains a neat, modern and thorough account of the theory of modular forms (including much information about the modularity theorem). Koblitz's book [**Kob93**] is also a very nice introduction to the theory of elliptic curves and modular forms (and includes a lot of information about the congruent number problem). Chapter 5 in Milne's book [**Mil06**] contains a good, concise overview of the subject. Serre's little book [**Ser77**] is always worth reading and also contains an introduction to modular forms. Miyake's book [**Miy06**] is a very useful reference.

(4) Finally, if the reader is interested in computations, we recommend Cremona's [**Cre97**] or Stein's [**Ste07**] book. If the reader wants to play with fundamental domains of modular curves, try Helena Verrill's applet [**Ver05**].

I would like to thank the organizers of the undergraduate summer school at PCMI, Aaron Bertram and Andrew Bernoff, for giving me the opportunity to lecture in such an exciting program. Also, I would like to thank Ander Steele and Aaron Wood for numerous corrections and comments of an early draft. Last, but not least, I would like to express my gratitude to Keith Conrad, Fernando Gouvêa, David Pollack and William Stein, whose abundant comments and suggestions have improved this manuscript much more than it would be safe to admit.

Álvaro Lozano-Robledo

Chapter 1

Introduction

Notation:

$\mathbb{N} = \{1, 2, 3, \ldots\}$ is the set of natural numbers.

$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ is the ring of integers.

$\mathbb{Q} = \{\frac{m}{n} : m, n \in \mathbb{Z}, n \neq 0\}$ is the field of rational numbers.

$\mathbb{R}$ is the field of real numbers.

$\mathbb{C} = \{a + bi : a, b \in \mathbb{R},\ i^2 = -1\}$ is the field of complex numbers.

In this chapter, we introduce elliptic curves, modular forms and L-functions through examples that motivate the definitions.

1.1. Elliptic curves

For the time being, we define an elliptic curve to be any equation of the form

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

with $a, b, c \in \mathbb{Z}$ and such that the polynomial $x^3 + ax^2 + bx + c$ does not have repeated roots. See Section 2.2 for a precise definition.

Example 1.1.1. *Are there three consecutive integers whose product is a perfect square?*

There are some trivial examples that involve the number zero, for example, $0, 1$ and 2, whose product equals $0 \cdot 1 \cdot 2 = 0 = 0^2$, a square.

Are there any non-trivial examples? If we try to assign variables to our problem, we see that we are trying to find solutions to

$$y^2 = x(x+1)(x+2) \tag{1.1}$$

with $x, y \in \mathbb{Z}$ and $y \neq 0$. Equation (1.1) defines an elliptic curve. It turns out that there are no integral solutions other than the trivial ones (see Exercise 1.4.1). Are there rational solutions, i.e., are there solutions with $x, y \in \mathbb{Q}$? This is a more delicate question, but the answer is still no (we will prove it in Example 2.7.6). Here is a similar question, with a very different answer:

- *Are there three integers that differ by* 5*, i.e.,* x*,* $x+5$ *and* $x+10$*, and whose product is a perfect square?*

In this case, we are trying to find solutions to $y^2 = x(x+5)(x+10)$ with $x, y \in \mathbb{Z}$. As in the previous example, there are trivial solutions (those which involve 0) but in this case, there are non-trivial solutions as well:

$$\begin{aligned} (-9)\cdot(-9+5)\cdot(-9+10) = (-9)\cdot(-4)\cdot 1 = 36 &= 6^2 \\ 40\cdot(40+5)\cdot(40+10) = 40\cdot 45\cdot 50 = 90000 &= 300^2. \end{aligned}$$

Moreover, there are also *rational* solutions, which are far from obvious:

$$\begin{aligned} \left(\frac{5}{4}\right)\cdot\left(\frac{5}{4}+5\right)\cdot\left(\frac{5}{4}+10\right) &= \left(\frac{75}{8}\right)^2 \\ \left(-\frac{50}{9}\right)\cdot\left(-\frac{50}{9}+5\right)\cdot\left(-\frac{50}{9}+10\right) &= \left(\frac{100}{27}\right)^2 \end{aligned}$$

and, in fact, there are infinitely many *rational* solutions! Here are some of the x-coordinates that work:

$$x = -9,\ 40,\ \frac{5}{4},\ \frac{-50}{9},\ \frac{961}{144},\ \frac{7200}{961},\ -\frac{12005}{1681},\ -\frac{16810}{2401},\ -\frac{27910089}{5094049}, \ldots$$

In Sections 2.9 and 2.10 we will explain a method to find rational points on elliptic curves and, in Exercise 2.12.23, the reader will calculate all the rational points of $y^2 = x(x+5)(x+10)$. ■

Example 1.1.2 (The Congruent Number Problem)**.** *We say that* $n \geq 1$ *is a congruent number if there exists a right triangle whose sides are rational numbers and whose area equals* n*. What natural numbers are congruent?*

For instance, the number 6 is congruent, because the right triangle with sides of length $(a, b, c) = (3, 4, 5)$ has area equal to $\frac{3\cdot 4}{2} = 6$. Similarly, the number 30 is the area of the right triangle with sides $(5, 12, 13)$; thus, 30 is a congruent number.

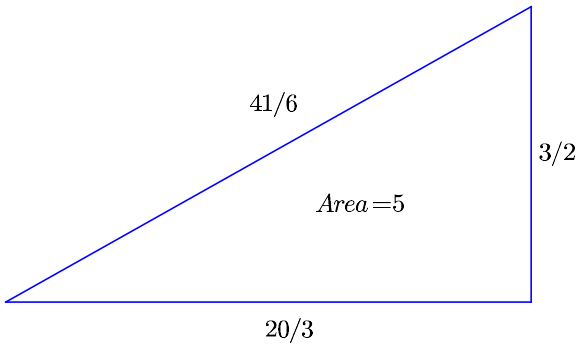

Figure 1. A right triangle of area 5 and rational sides.

The number 5 is congruent but there is no right triangle with integer sides and area equal to 5. However, our definition allowed *rational* sides, and the triangle with sides $\left(\frac{3}{2}, \frac{20}{3}, \frac{41}{6}\right)$ has area exactly 5. We do not allow, however, triangles with irrational sides even if the area is an integer. For example, the right triangle $(1, 2, \sqrt{5})$ has area 1, but that does not imply that 1 is a congruent number (in fact, 1 is *not* a congruent number, as we shall see below).

The congruent number problem is one of the oldest open problems in number theory. For more than a millennium, mathematicians have attempted to provide a characterization of all congruent numbers. The oldest written record of the problem dates back to the early Middle Ages, when it appeared in an Arab manuscript written before 972 (a later 10th century manuscript written by Mohammed Ben Alcohain would go as far as to claim that the principal object of the theory of rational right triangles is to find congruent numbers). It is known that Leonardo Pisano, a.k.a. *Fibonacci*, was challenged around 1220 by Johannes of Palermo to find a rational right triangle of area

$n = 5$, and Fibonacci found the triangle $(\frac{3}{2}, \frac{20}{3}, \frac{41}{6})$. We will explain a method to find this triangle below. In 1225, Fibonacci wrote a more general treatment about the congruent number problem, in which he stated (without proof) that if n is a perfect square, then n cannot be a congruent number. The proof of such a claim had to wait until Pierre de Fermat (1601-1665) settled that the number 1, and every square, are not congruent numbers (interestingly, his proof can be applied to prove the case $n = 4$ of Fermat's last theorem; see Example 1.1.5).

The connection between the congruent number problem and elliptic curves is as follows:

Proposition 1.1.3. *The number $n > 0$ is congruent if and only if the curve $y^2 = x^3 - n^2x$ has a point (x, y) with $x, y \in \mathbb{Q}$ and $y \neq 0$. More precisely, there is a one-to-one correspondence $C_n \longleftrightarrow E_n$ between the following two sets:*

$$\begin{aligned} C_n &= \{(a, b, c) : a^2 + b^2 = c^2,\ \frac{ab}{2} = n\} \\ E_n &= \{(x, y) : y^2 = x^3 - n^2x,\ y \neq 0\}. \end{aligned}$$

Mutually inverse correspondences $f : C_n \to E_n$ and $g : E_n \to C_n$ are given by

$$f((a, b, c)) = \left(\frac{nb}{c-a}, \frac{2n^2}{c-a}\right),\ g((x, y)) = \left(\frac{x^2 - n^2}{y}, \frac{2nx}{y}, \frac{x^2 + n^2}{y}\right).$$

The reader can provide a proof (see Exercise 1.4.3). For example, the curve $E : y^2 = x^3 - 25x$ has a point $(-4, 6)$ that corresponds to the triangle $(\frac{3}{2}, \frac{20}{3}, \frac{41}{6})$. But E has other points, such as $(\frac{1681}{144}, \frac{62279}{1728})$ that corresponds to the triangle

$$\left(\frac{1519}{492}, \frac{4920}{1519}, \frac{3344161}{747348}\right)$$

which also has area equal to 5. See Figure 2.

Today, there are partial results toward the solution of the congruent number problem, and strong results that rely heavily on famous (and widely accepted) conjectures, but we do not have a full answer yet. For instance, in 1975 (see [**Ste75**]), Stephens showed that the Birch and Swinnerton-Dyer conjecture (which we will discuss in Section 5.2) implies that any positive integer $n \equiv 5$, 6 or 7 mod 8 is a

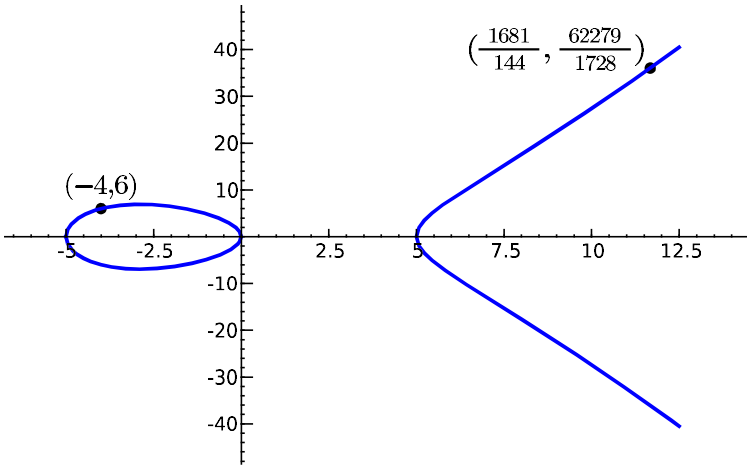

Figure 2. Two rational points on the curve $y^2 = x^3 - 25x$.

congruent number. For example, $n = 157 \equiv 5 \mod 8$ must be a congruent number and, indeed, Don Zagier has exhibited a right triangle (a, b, c) whose area equals 157. The hypotenuse of the simplest such triangle is:

$$c = \frac{224403517704336969924557513090667486316094847204 1}{8912332268928859588025535178967163570016480830}.$$

In Example 5.2.7 we will see an application of the conjecture of Birch and Swinnerton-Dyer to find a rational point P on $y^2 = x^3 - 157^2 x$, which corresponds to a right triangle of area 157 via the correspondence in Proposition 1.1.3.

The best known result on the congruent number problem is due to J. Tunnell:

Theorem 1.1.4 (Tunnell, 1983, [Tun83]). *If n is an odd square-free positive integer and n is the area of a right triangle with rational sides, then the following numbers are equal:*

$$\begin{aligned} &\#\{(x, y, z) \in \mathbb{Z}^3 : n = 2x^2 + y^2 + 32z^2\} \\ &= \frac{1}{2}\left(\#\left\{(x, y, z) \in \mathbb{Z}^3 : n = 2x^2 + y^2 + 8z^2\right\}\right) \end{aligned}$$

and, if n is even,

$$\begin{aligned}&\#\{(x,y,z)\in\mathbb{Z}^3 : \frac{n}{2} = 4x^2+y^2+32z^2\}\\ &= \frac{1}{2}\left(\#\{(x,y,z)\in\mathbb{Z}^3 : \frac{n}{2} = 4x^2+y^2+8z^2\}\right).\end{aligned}$$

Moreover, if the Birch and Swinnerton-Dyer conjecture is true, then, conversely, these equalities imply that n is a congruent number.

For example, for $n = 2$ we have $\frac{n}{2} = 1 = 4x^2 + y^2 + 32z^2$ if and only if $x = z = 0$ and $y = \pm 1$, so the left-hand side of the appropriate equation in Tunnell's theorem is equal to 2. However, the right-hand side is equal to 1 and the equality does not hold. Hence, 2 is *not* a congruent number.

For a complete historical overview of the congruent number problem, see [**Dic05**], Ch. XVI. The book [**Kob93**] contains a thorough modern treatment of the problem. The reader may also find useful an expository paper [**Con08**] on the congruent number problem, written by Keith Conrad. Another neat exposition, more computational in nature (using Sage), appears in [**Ste08**], Section 6.5.3. ■

Figure 3. Pierre de Fermat (1601-1665).

Example 1.1.5 (Fermat's last theorem)**.** *Let $n \geq 3$. Are there any solutions to $x^n + y^n = z^n$ in integers x, y, z with $xyz \neq 0$?* The

answer is *no*. In 1637, Pierre de Fermat wrote in the margin of a book (Diophantus' *Arithmetica*; see Figure 9 in Section 5.5) that he had found a marvelous proof, but the margin was too small to contain it. Since then, many mathematicians tried in vain to demonstrate (or disprove!) this claim. A proof was finally found in 1995 by Andrew Wiles ([**Wil95**]). We shall discuss the proof in some more detail in Section 5.5. For now, we will outline the basic structure of the argument.

First, it is easy to show that, to prove the theorem, it suffices to show the cases $n = 4$ and $n = p \geq 3$, a prime. It is not difficult to show that $x^4 + y^4 = z^4$ has no non-trivial solutions in $\mathbb{Z}$ (this was first shown by Fermat). Now, suppose that $p \geq 3$ and a, b, c are integers with $abc \neq 0$ and $a^p + b^p = c^p$. Gerhard Frey conjectured that if such a triple of integers exists, then the elliptic curve

$$E : y^2 = x(x - a^p)(x + b^p)$$

would have some unexpected properties that would contradict a well-known conjecture that Taniyama, Shimura and Weil had formulated in the 1950's. Their conjecture spelled out a strong connection between elliptic curves and modular forms, which we will describe in Section 5.4. Ken Ribet proved that, indeed, such a curve would contradict the Taniyama-Shimura-Weil (TSW) conjecture. Finally, Andrew Wiles was able to prove the TSW conjecture in a special case that would cover the hypothetical curve E. Therefore, E cannot exist and the triple (a, b, c) cannot exist, either.

The Taniyama-Shimura-Weil conjecture (Conjecture 5.4.5), i.e., the modularity theorem 5.4.6, was fully proved by Christophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor in their article [**BCDT01**]. ■

1.2. Modular forms

Let $\mathbb{C}$ be the complex plane and let $\mathbb{H}$ be the upper half of the complex plane, i.e., $\mathbb{H} = \{a+bi : a, b \in \mathbb{R}, b > 0\}$. A *modular form* is a function $f : \mathbb{H} \to \mathbb{C}$ that has several relations among its values (which we will specify in Definitions 4.1.3 and 4.2.1). In particular, the values of the function f satisfy several types of periodicity relations. For example,

the modular forms for $\mathrm{SL}(2,\mathbb{Z})$ satisfy, among other properties, the following:

- $f(z) = f(z+1)$ for all $z \in \mathbb{H}$, and
- $f\left(\frac{-1}{z}\right) = z^k f(z)$ for all $z \in \mathbb{H}$. The number k is an integer called the *weight* of the modular form.

We will describe modular forms in detail in Chapter 4. Let us see some examples that motivate our interest in these functions.

Example 1.2.1 (Representations of integers as sums of squares)**.** *Is the number $n > 0$ a sum of two (integer) squares?* In other words, are there $a, b \in \mathbb{Z}$ such that $n = a^2 + b^2$? And if so, in how many different ways can you represent n as a sum of two squares?

For instance, the number $n = 3$ cannot be represented as a sum of two squares but the number $n = 5$ has 8 distinct representations:

$$5 = (\pm 1)^2 + (\pm 2)^2 = (\pm 2)^2 + (\pm 1)^2.$$

Notice that here we consider $(-1)^2 + 2^2$, $1^2 + 2^2$ and $2^2 + 1$ as *distinct* representations of 5. A general formula for the number of representations of an integer n as a sum of 2 squares, due to Lagrange, Gauss and Jacobi, is given by

$$S_2(n) = 2\left(1 + \left(\frac{-1}{n}\right)\right) \sum_{d|n} \left(\frac{-1}{d}\right), \tag{1.2}$$

where $\left(\frac{m}{n}\right)$ is the Jacobi symbol and $\sum_{d|n}$ is a sum over all positive divisors of n (including 1 and n). Here we just need the easiest values $\left(\frac{-1}{n}\right) = (-1)^{(n-1)/2}$ of the Jacobi symbol. Let us see that the formula works:

$$S_2(3) = 2\left(1 + \left(\frac{-1}{3}\right)\right) \sum_{d|3} \left(\frac{-1}{d}\right) = 2(1 + (-1))(1 + (-1)) = 0,$$

$$S_2(5) = 2\left(1 + \left(\frac{-1}{5}\right)\right) \sum_{d|5} \left(\frac{-1}{d}\right) = 2(1 + 1)(1 + 1) = 8,$$

and $S_2(9) = 4$. Indeed, the number nine has 4 different representations: $9 = (\pm 3)^2 + 0^2 = 0^2 + (\pm 3)^2$. Let us explore other similar questions.

Let $n > 0$ and $k \geq 2$. Is the number $n > 0$ a sum of k (integer) squares? In other words, are there $a_1, \ldots, a_k \in \mathbb{Z}$ such that $n = a_1^2 + \cdots + a_k^2$? And if so, in how many different ways can you represent n as a sum of k squares? Lagrange showed that *every* natural number can be represented as a sum of $k \geq 4$ squares, but how many different representations are there?

Let $S_k(n)$ be the number of representations of n as a sum of k squares. Determining exact formulas for $S_k(n)$ is a classical problem in number theory. There are exact formulas known in a number of cases (e.g. Eq. 1.2). The formulas for $k = 4, 6$ and 8 are due to Jacobi and Siegel. We write $n = 2^\nu g$, with $\nu \geq 0$ and odd $g > 0$:

$$\begin{aligned} S_4(n) &= 8 \sum_{d|n,\ 4\nmid d} d, \\ S_6(n) &= \left(\left(\frac{-1}{g} \right) 2^{2\nu+4} - 4 \right) \sum_{d|g} \left(\frac{-1}{d} \right) d^2, \\ S_8(n) &= 16 \cdot \begin{cases} \sum_{d|n} d^3 & \text{if } n \text{ is odd,} \\ \sum_{d|n} d^3 - 2 \sum_{d|g} d^3 & \text{if } n \text{ is even.} \end{cases} \end{aligned}$$

For example, $S_4(4) = 8(1 + 2) = 24$ and, indeed

$$\begin{aligned} 4 &= (\pm 1)^2 + (\pm 1)^2 + (\pm 1)^2 + (\pm 1)^2 = (\pm 2)^2 + 0 + 0 + 0 \\ &= 0 + (\pm 2)^2 + 0 + 0 = 0 + 0 + (\pm 2)^2 + 0 = 0 + 0 + 0 + (\pm 2)^2. \end{aligned}$$

So there are $16 + 2 + 2 + 2 + 2 = 24$ possible representations of the number 4 as a sum of 4 squares. Notice that $S_4(2) = S_4(4)$. In how many ways can 4 be represented as a sum of 6 squares? We write $4 = 2^2 \cdot 1$, so $\nu = 2$ and $g = 1$, and thus,

$$S_6(4) = \left(\left(\frac{-1}{1} \right) 2^{2 \cdot 2 + 4} - 4 \right) \left(\left(\frac{-1}{1} \right) \cdot 1^2 \right) = (2^8 - 4) \cdot 1 = 252.$$

The formulas for $S_k(n)$ given above are derived using the theory of modular forms, as follows. We define a formal power series $\Theta(q)$ by

$$\Theta(q) = \sum_{j=-\infty}^{\infty} q^{j^2}$$

and, for $k \geq 2$, consider the power series expansion of the kth power of Θ:

$$\begin{aligned}(\Theta(q))^k &= \left(\sum_{j=-\infty}^{\infty} q^{j^2}\right)^k \\ &= \left(\sum_{a_1=-\infty}^{\infty} q^{a_1^2}\right)\cdots\left(\sum_{a_k=-\infty}^{\infty} q^{a_k^2}\right) = \sum_{n\geq 0} c_n q^n.\end{aligned}$$

What is the nth coefficient, c_n, of Θ^k? If the readers stare at the previous equation for a while, they will find that c_n is given by

$$c_n = \#\{(a_1, \ldots, a_k) \in \mathbb{Z}^k : a_1^2 + \cdots + a_k^2 = n\}.$$

Therefore, $c_n = S_k(n)$ and $(\Theta(q))^k = \sum_{n\geq 0} S_k(n)q^n$. In other words, Θ^k is a generating function for $S_k(n)$. But, how do we find closed formulas for $S_k(n)$? This is where the theory of modular forms becomes particularly useful, for it provides an alternative description of the coefficients of Θ^k.

It turns out that, for even $k \geq 2$, the function Θ^k is a modular form of weight $\frac{k}{2}$ (more precisely, it is a modular form for the group $\Gamma_1(4)$), and the space of all modular forms of weight $\frac{k}{2}$, denoted by $M_{\frac{k}{2}}(\Gamma_1(4))$, is finite dimensional (we will carefully define all these terms later). For instance, let $k = 4$. Then $M_2(\Gamma_1(4))$, the space of modular forms of weight $\frac{4}{2} = 2$ for $\Gamma_1(4)$, is a 2-dimensional $\mathbb{C}$-vector space and a basis is given by modular forms with q-expansions:

$$\begin{aligned}f(q) &= 1 + 24q^2 + 24q^4 + 96q^6 + 24q^8 + 144q^{10} + 96q^{12} + \cdots \\ g(q) &= q + 4q^3 + 6q^5 + 8q^7 + 13q^9 + 12q^{11} + 14q^{13} + \cdots .\end{aligned}$$

Therefore, $\Theta^4(q) = \lambda f(q) + \mu g(q)$ for some constants $\lambda, \mu \in \mathbb{C}$. We may compare q-expansions to find the values of λ and μ:

$$\begin{aligned}\Theta^4(q) = \sum_{n\geq 0} S_4(n)q^n &= 1 + 8q + 24q^2 + 32q^3 + 24q^4 + \cdots \\ \lambda f(q) + \mu g(q) &= \lambda + \mu q + 24\lambda q^2 + 4\mu q^3 + \cdots .\end{aligned}$$

Therefore, it is clear that $\lambda = 1$ and $\mu = 8$, so $\Theta^4 = f + 8g$. Since the expansions of f and g are easy to calculate (for example, using Sage; see Appendix A.2), we can easily calculate the coefficients of the q-expansion of Θ and, therefore, values of $S_4(n)$.

The exact formulas given above for $S_k(n)$, however, follow from some deeper facts. Here is a sketch of the ideas involved (the reader may skip these details for now and return here after reading Chapter 4): given $\Theta^4 = \sum c_n q^n$ and $F(q) = \sum(\sum_{d|n} d)q^n$, one can find an eigenvector $G(q) = \sum b_n q^n$ for a collection of linear maps T_n (the so-called Hecke operators, $T_n : M_2(\Gamma_1(4)) \to M_2(\Gamma_1(4))$ among spaces of modular forms, i.e., $T_n(G) = \lambda_n G$ for $n > 1$, and the eigenvalues $\lambda_n = b_n/b_1 = \sum_{d|n} d$. Moreover, the eigenvector G can be written explicitly as a combination of Θ^4 and F. Finally, one can show that the coefficients c_n must be given by the formula $c_n = 8\sum_{d|n,\ 4\nmid d} d$ (see [**Kob93**], III, §5, for more details). ■

1.3. L-functions

An L-function is a function $L(s)$, usually given as an infinite series of the form

$$L(s) = \sum_{n=1}^{\infty} a_n n^{-s} = \sum_{n=1}^{\infty} \frac{a_n}{n^s} = a_1 + \frac{a_2}{2^s} + \frac{a_3}{3^s} + \cdots$$

with some coefficients $a_n \in \mathbb{C}$. Typically, the function $L(s)$ converges for all complex numbers s in some half-plane (i.e., those s with real part larger than some constant), and in many cases $L(s)$ has an analytic or meromorphic continuation to the whole complex plane. Mathematicians are interested in L-functions because they are objects from analysis that, sometimes, capture very interesting algebraic information.

Example 1.3.1 (The Riemann zeta function)**.** The Riemann zeta function, usually denoted by $\zeta(s)$, is perhaps the most famous L-function:

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots .$$

The reader may already know some values of ζ. For example $\zeta(2) = \sum \frac{1}{n^2}$ is convergent by the p-series test, and its value is $\pi^2/6$ (this value can be computed using Fourier analysis and Parseval's equality). The connection between $\zeta(s)$ and number theory comes from the fact

that $\zeta(s)$ has an *Euler product*:

$$\begin{aligned}\zeta(s) &= \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \frac{1}{1-p^{-s}} \\ &= \left(\frac{1}{1-2^{-s}}\right) \cdot \left(\frac{1}{1-3^{-s}}\right) \cdot \left(\frac{1}{1-5^{-s}}\right) \cdots .\end{aligned}$$

This Euler product is not difficult to establish (Exercise 1.4.8) and has the very interesting consequence that any information on the distribution of the zeros of $\zeta(s)$ can be translated into information about the distribution of prime numbers among the natural numbers. ■

Example 1.3.2 (Dirichlet L-function)**.** *Let $a, N \in \mathbb{N}$ be relatively prime integers. Are there infinitely many primes p of the form $a+kN$ (i.e., $p \equiv a \mod N$) for $k \geq 0$?* The answer is *yes* and this fact, known as Dirichlet's theorem on primes in arithmetic progressions, was first proved by Dirichlet using a particular kind of L-function that we know today as a Dirichlet L-function.

Figure 4. Johann Peter Gustav Lejeune Dirichlet (1805-1859) and Georg Friedrich Bernhard Riemann (1826-1866).

Let $N > 0$. A *Dirichlet character* (modulo N) is a function $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$ that is a homomorphism of groups, i.e., $\chi(nm) = \chi(n)\chi(m)$ for all $n, m \in (\mathbb{Z}/N\mathbb{Z})^\times$. Notice that $\chi(n) \in \mathbb{C}$ and $\chi(n)^{\varphi(N)} = 1$ for all $\gcd(n, N) = 1$. Therefore, $\chi(n)$ must be a root of unity. We extend χ to $\mathbb{Z}$ as follows. Let $a \in \mathbb{Z}$. If $\gcd(a, N) = 1$, then $\chi(a) = \chi(a \bmod N)$. Otherwise, if $\gcd(a, N) \neq 1$, then $\chi(a) = 0$.

A Dirichlet L-function is a function of the form

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s},$$

where χ is a given Dirichlet character. For example, one can take χ_0 to be the trivial Dirichlet character, i.e., $\chi_0(n) = 1$ for all $n \geq 1$. Then $L(s, \chi_0)$ is the Riemann zeta function $\zeta(s)$. Dirichlet L-functions also have Euler products:

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_p \frac{1}{1 - \chi(p)p^{-s}}.$$

The idea of the proof of Dirichlet's theorem generalizes the following proof, due to Euler, of the infinitude of the primes. Consider $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \frac{1}{1-p^{-s}}$ and suppose there are only finitely many primes. Then the product over all primes is finite, and therefore its value at $s = 1$ would be finite (a rational number, in fact). However, $\zeta(1) = \sum_{n=1}^{\infty} 1/n$ is the harmonic series, which diverges! Therefore, there must be infinitely many prime numbers.

Dirichlet adapted this argument by looking instead at a different function:

$$\Psi_{a,N}(s) = \sum_{p \equiv a \bmod N} \frac{1}{p^s}.$$

He showed that (a) for every non-trivial Dirichlet character χ modulo N, we have $L(1, \chi) \neq 0$ or ∞, and (b) this implies that $\Psi_{a,N}(1)$ diverges to ∞. Part (b) follows from the equality

$$\log(\zeta(s)) + \sum_{\substack{\chi \bmod N \\ \chi \neq 1}} \chi(a)^{-1} \log(L(s, \chi))$$

$$= \phi(N) \left(\sum_{p \equiv a \bmod N} \frac{1}{p^s} \right) + g(s),$$

where $g(s)$ is a function with $g(1)$ finite, and ϕ is the Euler ϕ-function. Therefore, there cannot be a finite number of primes of the form $p \equiv a \bmod N$. ■

Example 1.3.3 (Representations of integers as sums of squares)**.** *Is the number $n > 0$ a sum of three integer squares?* In Subsection 1.2, we saw formulas for the number of representations of an integer as a sum of $k = 2$, 4, 6 and 8 integer squares, but we avoided the same question for odd k. The known formulas for $S_3(n)$, $S_5(n)$ and $S_7(n)$ involve values of Dirichlet L-functions.

Let us first define the Dirichlet character that we shall use here. The reader should be familiar with the Legendre symbol $\left(\frac{n}{p}\right)$, which is equal to 0 if $p|n$, equal to 1 if n is a square mod p, and equal to -1 if n is not a square mod p. Let $m > 0$ be a natural number with prime factorization $m = \prod_i p_i$ (the primes are not necessarily distinct). First we define

$$\left(\frac{n}{2}\right) = \begin{cases} 0 & \text{if } n \text{ is even,} \\ 1 & \text{if } n \equiv \pm 1 \bmod 8, \\ -1 & \text{if } n \equiv \pm 3 \bmod 8. \end{cases}$$

Now we are ready to define the *Kronecker symbol* of n over $m > 0$ by

$$\left(\frac{n}{m}\right) = \prod_i \left(\frac{n}{p_i}\right).$$

For any $n > 0$, the symbol $\left(\frac{-n}{\cdot}\right)$ induces a Dirichlet character χ_n defined by $\chi_n(a) = \left(\frac{-n}{a}\right)$, and we can define the associated L-function by

$$L(s, \chi_n) = \sum_{a=1}^{\infty} \frac{\chi_n(a)}{a^s}.$$

We are ready to write down the formula for $S_3(n)$, due to Gauss, Dirichlet and Shimura (there are also formulas for $S_5(n)$, due to Eisenstein, Smith, Minkowski and Shimura, and a formula for $S_7(n)$, also due to Shimura). For simplicity, let us assume that n is odd and square free (for the utmost generality, please check [**Shi02**]):

$$S_3(n) = \begin{cases} 0 & \text{if } n \equiv 7 \bmod 8, \\ \frac{24\sqrt{n}}{\pi} L(1, \chi_n) & \text{otherwise.} \end{cases}$$

The reader is encouraged to investigate this problem further by attempting Exercises 1.4.6 and 1.4.7. ■

1.4. Exercises

Exercise 1.4.1. Use the divisibility properties of integers to show that the only solutions to $y^2 = x(x+1)(x+2)$ with $x, y \in \mathbb{Z}$ are $(0,0)$, $(-1,0)$ and $(-2,0)$. (Hint: If a and b are relatively prime and ab is a square, then a is a square and b is a square.)

Exercise 1.4.2. Find all the Pythagorean triples (a,b,c), i.e., $a,b,c \in \mathbb{Z}$ and $a^2+b^2=c^2$, such that $b^2+c^2=d^2$ for some $d \in \mathbb{Z}$. In other words, find all the integers a,b,c,d such that (a,b,c) and (b,c,d) are both Pythagorean triples. (Hint: You may assume that $y^2 = x(x+1)(x+2)$ has no rational points other than $(0,0)$, $(-1,0)$ and $(-2,0)$.)

Exercise 1.4.3. Prove Proposition 1.1.3; i.e., show that $f((a,b,c))$ is a point in E_n, that $g((x,y))$ is a triangle in C_n and that $f(g((x,y))) = (x,y)$ and $g(f((a,b,c))) = (a,b,c)$.

Exercise 1.4.4. Calculate $S_4(n)$, for $n = 1,3,5,6$, by hand, using Jacobi's formula and also by finding all possible ways of writing n as a sum of 4 squares.

Exercise 1.4.5. The goal of this problem is to find the q-expansion of $\Theta^6(q)$:

(1) Find by hand the values of $S_6(n)$, for $n = 0,1,2$; i.e., find all possible ways to write $n = 0,1,2$ as a sum of 6 squares.

(2) Using Sage, calculate the dimension of $M_{\frac{k}{2}}(\Gamma_1(4))$ (see Appendix A.2) and a basis of modular forms for $k = 6$.

(3) Write Θ^6 as a linear combination of the basis elements found in part 2.

(4) Use part 3 to write the q-expansion of Θ^6 up to $O(q^{20})$.

(5) Use the expansion of Θ^6 to verify that $S_6(4) = 252$. Also, calculate $S_6(19)$ using Jacobi's formula and verify that it coincides with the coefficient of Θ^6 in front of the q^{19} term.

Exercise 1.4.6. Show that any integer $n \equiv 7 \bmod 8$ cannot be represented as a sum of three integer squares.

Exercise 1.4.7. Find the number of representations of $n = 3$ as a sum of 3 squares. Then compare your result with the value of the formula given in Example 1.3.3; i.e., use a computer to approximate

$$S_3(3) = \frac{24\sqrt{3}}{\pi} L(1, \chi_3) = \frac{24\sqrt{3}}{\pi} \sum_{a=1}^{\infty} \frac{\left(\frac{-3}{a}\right)}{a}$$

by adding the first $10,000$ terms of $L(1, \chi_3)$. Do the same for $n = 5$ and $n = 11$. Does the formula seem to work for $n = 2$? (*Note: the command* `kronecker(-n,m)` *calculates the Kronecker symbol* $\left(\frac{-n}{m}\right)$ *in* Sage.)

Exercise 1.4.8. Prove that the Riemann zeta function $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ has an Euler product; i.e., prove the following formal equality of series

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \frac{1}{1 - p^{-s}}.$$

(Hint: There are two possible approaches:

Hint (a). Expand the right-hand side using the Fundamental Theorem of Arithmetic and the algebraic equality $\frac{1}{1+x} = \sum_{k=0}^{\infty} x^k$. [This approach helps build an intuition about what is going on, but may be hard to write into a rigorous proof]

Hint (b). Calculate $(1 - 1/2^s)\zeta(s)$ and $(1 - 1/3^s)(1 - 1/2^s)\zeta(s)$, etc.)