CONGRUENT NUMBERS, ELLIPTIC CURVES, AND MODULAR FORMS

GUY HENNIART

1. Invitation to the voyage

An integer $n \geq 1$ is called *congruent* if it is the area of a right triangle whose three sides have rational lenghts. Equivalently, n is congruent if there exists a rational number x such that $x^2 - n$ and $x^2 + n$ are both squares of rational numbers.

The very old problem[1] of determining all congruent integers has only recently been answered satisfactorily, and nowadays we know all congruent numbers less than 1000 (Table 1).[2] In fact, thanks to J.B. Tunnell (1983) we are now in possession of a very simple criterion for verifying that a given integer is not congruent. Justifying this criterion will take us on a journey into the forest of results and conjectures in the modern arithmetic theory of elliptic curves and modular forms that we have already met in the preceding articles.

Tunnell's criterion can be formulated as follows: Let $n \geq 1$ be an odd squarefree integer (that is, not divisible by the square of an integer greater than 1). Consider the following conditions:

- A1 n is congruent.
- B1 The number of integral triplets (x, y, z) satisfying $2x^2 + y^2 + 8z^2 = n$ equals twice the number of integral triplets (x, y, z) satisfying $2x^2 + y^2 + 32z^2 = n$.
- A2 $2n$ is congruent.
- B2 The number of integral triplets (x, y, z) satisfying $4x^2 + y^2 + 8z^2 = n$ equals twice the number of integral triplets (x, y, z) satisfying $4x^2 + y^2 + 32z^2 = n$.

Then A1 implies B1 and A2 implies B2. Moreover, it is conjectured that the conditions A1 and B1 (as well as A2 and B2) are actually equivalent. In numerous concrete examples this can in fact be proved: this is what allows us to determine the congruent numbers up to 1000.

[1] It first appears in a Chinese manuscript around the year 980; cf. Dickson.

[2] Not reproduced here. See F.R. Nemenzo [*All congruent numbers less than* 40000, Proc. Japan Acad. **74** (1998), 29–31] for computations up to $40,000$.

We also remark that there is a double interest in Tunnell's criterion. First, given an odd integer n one can easily (and quickly) determine the number of representations of n in $B1$ or $B2$. If the condition $B1$ is not satisfied, one deduces immediately that n is not congruent. It is then futile to search for rational sides of a hypothetical right triangle of area n. In fact, even if we know that n is congruent, we do not have a general algorithm at our disposal that permits us to find these sides. For example, the prime $p = 157$ is congruent, but the sides of the simplest right triangle of area 157 have more than 20 digits both in numerator and denominator (Table 2).

Nevertheless, the real interest of the congruent number problem does not lie so much in the criterion announced above than in the techniques used for its proof and the domains of arithmetic that one has to cross in order to establish this. After some elementary preliminaries we will look at the arithmetic of elliptic curves, then of modular forms of integral and finally of half integral weight. We remark that modular forms appear in numerous branches of mathematics and more and more even in theoretical physics.

2. Gentle Departure: Elementary Manipulations

An integer $n \geq 1$ is congruent if the system of equations

$$a^2 + b^2 = c^2, \quad ab = 2n \tag{1}$$

has a solution (a, b, c) in nonzero rational numbers. Let $m \geq 1$ be an integer: it is clear that if n is congruent, then so is m^2‘n, and vice versa. For determining congruent numbers it is therefore sufficient to consider squarefree integers $n \geq 1$, that is, integers not divisible by the square of a prime number. The rational solutions of the equation $a^2 + b^2 = c^2$ have been known for a long time. The integral solutions are obtained by writing $a^2 = (b - c)(b + c)$ and by studying the prime factorization of each of these factors. This gives the solutions in the form

$$a = \lambda(r^2 - s^2), \quad b = \lambda(2rs), \tag{2}$$

where r and s are nonzero coprime imtegers of different parity, and where λ is an arbitrary nonzero integer. The other solutions are obtained by interchanging a and b in these formulas above. The nonzero rational solutions are found by allowing λ to assume arbitrary nonzero rational values.

The equation $ab = 2n$ can be written as follows, using the parametrization just given:

$$\lambda^2 rs(r^2 - s^2) = n.$$

Putting $y = n^2/\lambda r^2$ and $x = -ns/r$ we obtain two nonzero rational solutions x and y satisfying

$$y^2 = x^3 - n^2 x. \tag{3}$$

Conversely, if we have rational solutions x and $y \neq 0$ of (2), then the rational numbers

$$a = |2nx/y|,\ b = |(n^2 - x^2)/y|,\ c = |(n^2 + x^2)/y|$$

are the sides of a right triangle of area n.

It seems that we have replaced one problem by another by passing from one diophantine problem to another; nevertheless, (3) is the equation of an elliptic curve: it is known how to generate a third solution from two known ones. This way the solutions of (3) are given the structure of a group, hence the structure is much richer for (3) than for the example with which we started.

Remark. 1) If n is the area of a right triangle with rational sides $a < b < c$, then it can be seen by putting $d = (c/2)^2$ that $d - n$, d and $d + n$ are squares of rational numbers. Conversely, if $d-n$, d and $d+n$ are such squares, then $a = \sqrt{d+n} - \sqrt{d-n}$, $b = \sqrt{d+n} + \sqrt{d-n}$ and $c = 2\sqrt{d}$ are the sides of a right triangle with area n.

2) Determining the right triangles of area n with *integral* sides is easy if one has the prime factorization of n: one uses the equation $ab = 2n$. The case of rational solutions is of a much more complex nature. It is possible to obtain congruent numbers starting from the formulas (2) by letting r, s run through the integers and choosing λ in such a way that the area is squarefree. Nevertheless, we have no way of knowing at which moment a squarefree integer n will appear in this procedure: we don't know how to bound r and s in terms of n.

3. First Stage: The Group Law on Elliptic Curves

For our needs, an elliptic curve over a field K of characteristic different from 2 is given by an equation

$$y^2 = x^3 + ax^2 = bx = c, \tag{4}$$

where the polynomial on the right hand side has coefficients in K and has three different roots in an algebraic closure $\overline{K}$ of K (its discriminant is a non-zero element of K). We consider the set of solutions (x, y) of this equation over an extension L of K. In fact, for geometric reasons, we are mostly interested in the associated homogeneous equation

$$Y^2 Z = X^3 + aX^2 Z + bXZ^2 + cZ^3. \tag{5}$$

If L is an extension of K and if (X, Y, Z) is a triple of L^3 and a solution of this last equation, then all proportional triples are also solutions, and the non-zero proportional triplets define a unique point in the projective plane $\mathbb{P}^2(L)$ (of which X, Y and Z are called coordinates). The set $E(L)$ of points on the elliptic curve is the subset of points in $\mathbb{P}^2(L)$ that are associated to triplets of solutions of (5). Similarly, lines in $\mathbb{P}^2(L)$ are sets of points associated to non-zero solutions of a linear homogeneous equation $AX + BY + CZ = 0$ (where A, B and C are elements of L that are not all zero). The projective plane can be regarded as the union of L^2 corresponding to points with coordinates $(x, y, 1)$ and a line, called line at infinity, with equation $Z = 0$. From this point of view, $E(L)$ consists of two types of points: the point at infinity with coordinates $(0, 1, 0)$, and the points $(x, y, 1)$ corresponding to solutions of (4). That we consider (5) rather than (4) is due to the following result: every line in $\mathbb{P}^2(\overline{K})$ intersects $E(\overline{K})$ in *three* points if they are counted with multiplicity. Moreover, the assumption that f has three distinct roots in $\overline{K}$ implies that E is nonsingular (an algebraic concept corresponding to the fact that, for $K = \mathbb{R}$, E does not have a double point or a cusp): in particular, E possesses in each point $P \in E(L)$ a unique tangent (that is, a line intersecting $E(L)$ in P with multiplicity ≥ 2: the points of inflection correspond to intersections of multiplicity 3, the point at ∞ being one of them).

For each extensions L of K, we can define an addition $(P_1, P_2) \longmapsto P_1 + P_2$ on $E(L)$ such that $E(L)$ becomes an abelian group whose neutral element is the point at ∞, which will be denoted by $\mathcal{O}_E$ or $\mathcal{O}$ for that reason. Geometrically, three points P_1, P_2 and P_3 on $E(L)$ add up to $\mathcal{O}$ exactly when they are points of intersection (counted with multiplicity) of a line with $E(L)$. If P_i has coordinates $(x_i, y_i, 1)$ for $i = 1, 2$, and if $P_1 \neq -P_2$, then $P_4 = P_1 + P_2$ has coordinates $(x_4, y_4, 1)$ with

$$\begin{aligned} x_4 &= -x_1 - x_2 - a + m^2 \quad \text{if} \quad P_1 \neq P_2, \quad \text{where } m = \tfrac{y_2 - y_1}{x_2 - x_1}, \\ x_4 &= -2x_1 + m^2 \qquad\quad \text{if} \quad P_1 = P_2, \quad \text{where } m = \tfrac{f'(x_1)}{2y_1}, \\ y_4 &= -y_1 + m(x_1 - x_4) \qquad\qquad \text{in both cases.} \end{aligned}$$

The points of $E(L)$ satisfying $2P = \mathcal{O}$ are, apart from $\mathcal{O}$, the points $(x, 0)$, where x is a root of f in L. In particular, our curve $E_n : y^2 = x^3 - n^2x$ has four points of order 2 in every extension of $\mathbb{Q}$, and these form a subgroup isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

We have thus placed our problem in a broader perspective: determining the structure of the group of rational points of an elliptic curve defined over $\mathbb{Q}$. The fundamental result in this domain is due to Mordell and Weil: if K is a number field (a finite extension of $\mathbb{Q}$) and if E is

an elliptic curve over K, then $E(K)$ is an abelian group of finite type: in fact, there exist integers e_1, $e_2 \geq 1$ with $e_1 \mid e_2$ and an integer $r \geq 0$ called the rank of $E(K)$ such that $E(K) \simeq \mathbb{Z}/e_1\mathbb{Z} \times \mathbb{Z}/e_2\mathbb{Z} \times \mathbb{Z}^r$.

For the curve $E_n(\mathbb{Q})$ it can be shown that $e_1 = e_2 = 2$; in other words, the subgroup of points of finite order of $E_n(\mathbb{Q})$ consists of points of order at most 2. The method of proof is interesting and uses techniques that are useful later on.

Let p be a prime. An equation

$$y^2 = f(x) = x^3 + ax^2 + bx + c \tag{6}$$

with integral coefficients can be reduced modulo p: one considers the equation $y^2 = \widetilde{f}(x)$, where $\widetilde{f}$ is the polyomial in $(\mathbb{Z}/p\mathbb{Z})[x]$ whose coefficients are the images of those of f; if p does not divide the discriminant of the polynomial f (in particular if p is large enough), then one gets an equation defining an elliptic curve over the finite field $\mathbb{Z}/p\mathbb{Z}$. In fact, the same thing is possible if f has coefficients in $\mathbb{Q}$ as long as the primes p do not divide the denominators of a, b or c. Moreover, a rational solution (x, y) of (6) reduces, if p is large enough, to a solution of the equation over $\mathbb{Z}/p\mathbb{Z}$; if P has finite order m, then the reduced point $\widetilde{P}$ has also order m on the curve $\widetilde{E}(\mathbb{Z}/p\mathbb{Z})$. More generally, let G be a finite subgroup of $E(\mathbb{Q})$: if p is large enough, then the reduction gives a group homomorphism $G \longrightarrow \widetilde{E}(\mathbb{Z}/p\mathbb{Z})$ which is *injective*; in particular, if G is the subgroup of points of finite order of $E(\mathbb{Q})$, then the order of G divides the order of $\widetilde{E}(\mathbb{Z}/p\mathbb{Z})$ for each sufficiently large prime p.

Consider the case of the curve E_n and let us take a prime p of the form $4k + 3$ not dividing $2n$. An elementary exercise, using that -1 is not a square in $\mathbb{Z}/p\mathbb{Z}$, shows that $\widetilde{E}(\mathbb{Z}/p\mathbb{Z})$ has exactly $p + 1$ points. Using Dirichlet's theorem on primes in arithmetic progressions one concludes that the number of points of finite order in $E_n(\mathbb{Q})$ divides 4, hence these points are exactly those of order dividing 2.

Recalling that n is congruent if and only if the equation $y^2 = x^3 - n^2x$ has a solution in nonzero rational numbers, we get the following criterium: n is congruent if and only if $E_n(\mathbb{Q})$ has a point of inifinite order (that is, if and only if its rank is ≥ 1).

4. Vista: Elliptic Curves over Finite Fields

Now we continue studying points on elliptic curves defined over finite fields $\mathbb{Z}/p\mathbb{Z}$ (and their finite extensions). The idea which we will make more precise below is that if E is an elliptic curve over $\mathbb{Q}$ with larg rank, then there will be "many" points over $E(\mathbb{Z}/p\mathbb{Z})$ for large primes

p, if only by reducing points from $E(\mathbb{Q})$. It turns out that with the help of Gauss and Jacobi sums one can compute, for each prime p not dividing $2n$, the number of points of E_n not only over $\mathbb{Z}/p\mathbb{Z}$ but also over any of its finite extensions. In fact, there is a general result due to Hasse and Weil: for every elliptic curve E over $\mathbb{Z}/p\mathbb{Z}$ there exists an algebraic integer α in some complex quadratic number field with absolute value $\sqrt{p}$ such that the cardinality N of $E(\mathbb{Z}/p\mathbb{Z})$ is given by $N = p + 1 - \alpha - \overline{\alpha}$. More generally, over an extension of degree r of $\mathbb{Z}/p\mathbb{Z}$, E has exactly N_r points, where

$$N_r = p^r + 1 - \alpha^r - \overline{\alpha}^r.$$

In the case of the curve E_n, one finds $\alpha = \sqrt{-p}$ if p has the form $4k + 3$; if $p = 4k + 1$, then $p = a^2 + b^2$ with integers a and b such that a is odd: then $\alpha = a + bi$, where the sign of a is determined by demanding that

$$a \equiv \begin{cases} 1 \bmod 4 & \text{if } 4 \mid b \text{ and } n \text{ is a square modulo } p \\ 1 \bmod 4 & \text{if} 4 \nmid b \text{ and } n \text{ is not a square modulo } p \\ 3 \bmod 4 & \text{otherwise} \end{cases}$$

The preceding result can also be expressed in a different form: if we put $a = \alpha + \overline{\alpha}$ (this is an integer) and if we observe that $Z_p(E, T)$ is the rational fraction $\frac{1-aT+pT^2}{(1-T)(1-pT)}$, then $TZ'_p(E, T)/Z_p(E, T)$ can be developed into a formal power series of the form $\sum_{r\geq 1} N_r T^r$. The rational function $Z_p(E, T)$ is called the zeta function of the curve E over $\mathbb{Z}/p\mathbb{Z}$.

Remark. The results above can be generalized to nonsingular projective curves (Weil [16]) and even to nonsingular projective varieties of arbitrary dimension over finite fields (Deligne [3], [5]).

5. View on an Inaccessible Summit: BSD

Suppose that we are given an elliptic curve E over $\mathbb{Q}$ defined by $y^2 = f(x)$ where f has integral coefficients (this can be done without loss of generality as we are free to replace x and y by multiples). For each prime p not dividing the discriminant of f one has the elliptic curve reduced modulo p and its zeta function $Z_p(E, T)$ introduced above. For p defining the discriminant it is also possible to define in a natural way a factor of this kind (one has to change the equation of E a bit): for $E = E_n$ and p dividing $2n$, one gets

$$Z_p(E_n, T) = \frac{1}{(1 - T)(1 - pT)}.$$

For each p we define

$$L_p(E,T) = \frac{(1-T)(1-pT)}{Z_p(E,T)};$$

thus $L_p(E_n,T) = 1/(1-a_pT+pT^2$, with $a_p \in \mathbb{Z}$, if $p \nmid 2n$.

We also introduce the complex valued function $L_\infty(s) = (2\pi)^{-s}\Gamma(s)$, where Γ denotes Euler's Gamma function. We also consider the infinite product

$$\Lambda(E,s) = L_\infty(s) \prod_{p \text{ prime}} L_p(E,p^{-s}).$$

This infinite product which somehow collects the information on E coming from different primes converges for all s with real part $> \frac{3}{2}$: this follows from the inequality $|a_p| \leq 2\sqrt{p}$.

A conjecture known as the Weil conjecture, or the Weil-Taniyama conjecture, or Shimura-Taniyama conjecture,[3] says that an elliptic curve E over $\mathbb{Q}$ is "modular" (see Section 6); in particular, it is expected that $\Lambda(E,s)$ can be extended to the whole complex plane as a holomorphic function of s satisfying a functional equation

$$\Lambda(E,s) = wN^{1-s}\Lambda(E,2-s),$$

where w (the "constant" of the functional equation) is $+1$ or -1, and where N is a positive integer called the conductor of the curve: N is defined geometrically from E, and is divisible only by primes dividing the discriminant.

It can be proved that curves "with complex multiplication" such as E_n are modular. An elliptic curve over $\mathbb{Q}$ is said to have complex multiplication if the ring of endomorphisms of $E(\mathbb{C})$ that can be expressed by formulas that are polynomials in the projective coordinates is bigger than the ring of multiplications $P \longmapsto nP$ by integers n: this ring consists of integers in some complex quadratic number field, hence the name complex multiplication. For the curves E_n, this field is $\mathbb{Q}(i)$, where $i^2 = -1$, and an automorphism of E_n is given by $(x,y) \longmapsto (-x,iy)$. The conductor of E_n is $N = 32n^2$ if n is odd and $N = 16n^2$ if n is even. The sign of w in the functional equation is $+1$ for $n \equiv 1,2,3 \bmod 8$ and -1 for $n \equiv 5,6,7 \bmod 8$.

Assuming that Λ can be extended to a holomorphic function in the complex plane, the value of Λ at $s = 1$. the "center" of the functional equation, becomes extremely interesting. Birch and Swinnerton-Dyer [1] have conjectured that the rank r of $E(\mathbb{Q})$ *equals* the order of the zero

[3]This conjecture was proved by Wiles & Taylor for semistable elliptic curves in 1994, and it was proved completely in 1999 by Christophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor.

of $\Lambda(E,s)$ at $s=1$: more precisely, they even predict the dominant term of the Taylor expansion of Λ at $s=1$. The weak form of the conjecture which is used for the problem of congruent numbers ($E=E_n$) is the following:

Conjecture. We have $L(E,1)=0$ if and only if the rank of $E(\mathbb{Q})$ is >0.

Remarks. 1. A very heuristic reasoning for motivating this conjecture goes as follows: if the product giving $L(E,s)$ would converge in $s=1$, then we would have

$$L(E,1) \;=\; \prod_p \frac{p}{p+1-a_p} \;=\; \prod \frac{p}{N(p)},$$

where $N(p)$ is the number of points of the curve modulo p. If $E(\mathbb{Q})$ is finite, one may hope that the a_p are distributed uniformly between $-2\sqrt{p}$ and $2\sqrt{p}$, which leads to a nonzero value for $L(E,1)$; if $E(\mathbb{Q})$ is infinite, then $N(p)$ will often be as large as possible, that is close to $2\sqrt{p}$, and this will give $L(E,1)$ the value 0.

2. Assuming the preceding conjectures plus another one (the generalized Riemann conjecture for Λ), Mestre has made the link between the rank of $E(\mathbb{Q})$ and the distances between a_p and $-2\sqrt{p}$ for primes p. Moreover, by choosing curves E such that a_p is close to $-2\sqrt{p}$ for smalle primes p he has obtained explicit curves of large rank. The present record is $r\geq 14$.

There are numerous results that corroborate the conjecture of Birch and Swinnerton-Dyer (see [12, App. §16]). In particular, in every case where it has been verified that a given integer is congruent or not, the weak form of the conjecture above is verified; numerous other examples have been treated. As for general results, there are the theorems of Coates and Wiles [2] and of Gross and Zagier [4].

Coates and Wiles have proved that if an elliptic curve E over $\mathbb{Q}$ with complex multiplication and a rational point of infinite order, then $L(E,1)=0$.

Gross and Zagier have shown that if E is a "modular" elliptic curve such that $L(E,1)=0$ and $L'(E,1)\neq 0$, then $E(\mathbb{Q})$ contains a rational point of infinite order.

We remark that there are rapidly converging series giving $L(E,1)$ (and $L'(E,1)$ if $L(E,1)=0$) which allow us, for given E, to determine whether $L(E,1)$ (or $L'(E,1)$) is 0 or not.

As for the curves E_n, we have the following criteria:

- if $L(E,1)\neq 0$, then n is not congruent;
- if $L(E,1)=0$ and $L'(E,1)\neq 0$, then n is congruent.

We remark that because of the fact that we know the sign of the functional equation for E_n, the first case can only occur if $n \equiv 1, 2, 3 \bmod 8$, the second only if $n \equiv 5, 6, 7 \bmod 8$. The conjecture of Birch and Swinnerton-Dyer says that if $n \equiv 1, 2, 3 \bmod 8$, then n is congruent if and only if $L(E_n, 1) = 0$, and that numbers $n \equiv 5, 6, 7 \bmod 8$ are always congruent. Apart from applying the preceding two criteria, the only way of verifying this conjecture is by trying to find a point of infinite order on E_n; in this way all $n \leq 1000$ have been handled. Nevertheless we do not have a method giving such a point in a time bounded in terms of n.

The method of Tunnell consists in expressing $L(E_n, 1)$ in simple arithmetic terms. In particular, one gets a criterium for non-congruence, and, if the conjecture of Birch and Swinnerton-Dyer is correct, also for congruence. For obtaining the criterium of Tunnell we first have to explore the properties of L-functions of elliptic curves and pass into the arithmetic-analytic domain of modular forms.

6. Another Valley: Modular Forms

Modular forms are functions on the upper half plane $\mathbb{H}$ of complex numbers with positive imaginary part. The group $\mathrm{GL}_2(\mathbb{R})^+$ of 2×2 square matrices with real entries and positive discriminant acts on $\mathbb{H}$ via

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{az + b}{cz + d}.$$

Let $k \geq 1$ be an integer: if $f : \mathbb{H} \longrightarrow \mathbb{C}$ is a function on $\mathbb{H}$ and $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ an element of $\mathrm{GL}_2(\mathbb{R})^+$, then $f|_k[\alpha]$ denotes the function $\mathbb{H} \longrightarrow \mathbb{C}$ defined by the formula

$$f|_k[\alpha](z) = (ad - bc)^{k/2} f(\alpha z)(cz + d)^{-k}.$$

Assume that f is a meromorphic funtion $\mathbb{H} \longrightarrow \mathbb{C}$. If f verifies $f(z+1) = f(z)$, then it can be written as a function of $q = e^{2\pi i z}$ in the puntured disc $|q| < 1$. If f can be extended to a holomorphic function on the disc $|q| < 1$, then f is called *holomorphic at* ∞. In this case, f admits a Fourier expansion at infinity

$$f(z) = \sum_{n \geq 0} a_n q^n,$$

where the a_n are called the Fourier coefficients of f. We say that f vanishes at ∞ if in addition $a_0 = 0$.

If $N > 0$ is an integer, the group $\Gamma_0(N)$ is the group of square matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with coefficients in $\mathbb{Z}$, determinant 1, and such that c is a multiple of N.

A modular form of weight k for $\Gamma_0(N)$ is a meromorphic function $f : \mathbb{H} \longrightarrow \mathbb{C}$ such that
i) $f|_k[\alpha] = f$ for all $\alpha \in \Gamma_0(N)$;
ii) $f|_k[\alpha]$ is holomorphic at ∞ for all $\alpha \in \Gamma_0(N)$.
Moreover, f is called a cusp form if $f|_k[\alpha](\infty) = 0$.

Remarks. 1) One can regard modular forms as functions of lattices in $\mathbb{C}$ (these are subgroups of $\mathbb{C}$ generated by an $\mathbb{R}$-basis of $\mathbb{C}$). These lattices correspond in a simple way to *elliptic curves* over $\mathbb{C}$. For this connection between modular forms and elliptic curves, which shall not be discussed here, see [9] or [10].

2) Modular forms occur in various places in arithmetic (as well as in other parts of mathematics). For example, there are the Eisenstein series of weight $2k$ (for $\Gamma_0(1)$) whose development at ∞ is

$$G_k(z) = 2\zeta(2k) + (2\pi i)^k \frac{2}{2k-1} \sum_{n=1}^{\infty} \sigma_{2k-1}(n) q^n,$$

where $\sigma_\ell(n)$ is the sum of the ℓ-th powers of the divisors of n, and where $\zeta(s)$ is Riemann's zeta function $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$.

We also have the function Δ, a cusp form of weight 12 for $\Gamma_0(1)$, given by $\Delta = (2\pi)^{12} q \prod_{n=1}^{\infty} (1-q)^{24}$. In the situation of Remark 1), Δ corresponds to the discriminant of the elliptic curve attached to a lattice. Finally, modular forms can be given in terms of theta functions: for example, if ℓ is a multiple of 4, there is a modular form Θ_ℓ of weight $\ell/2$ for $\Gamma_0(4)$, whose development at ∞ is

$$\Theta_\ell(z) = \sum_{m \geq 0} r_\ell(m) q^m,$$

where $r_\ell(m)$ is the number of representations of m as a sum of ℓ squares.

The space of modular forms for $\Gamma_0(N)$ and given weight k has finite dimension as a vector space over $\mathbb{C}$. Moreover, we can define a subspace of primitive forms, not coming from forms for $\Gamma_0(M)$ with $M < N$ (this can be made more precise).

For each integer $n \geq 1$ there is a *Hecke operator* $T(n)$ that acts on the space of modular forms of weight k for $\Gamma_0(N)$; it fixes the space of cusp forms and the space of primitive forms. The space of primitive cusp forms has a basis consisting of eigenvectors for each $T(n)$. If $f = \sum_{n=1}^{\infty} a_n q^n$ is such an eigenvector, then $T(n)f = (a_n/a_1)f$ (we

normalize the eigenvectors by the condition $a_1 = 1$); moreover, we have $f|_k[\tau_N] = w\overline{f}$, where $\tau_N = \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix}$, w is a complex number of absolute value 1, and where $\overline{f}$ is the modular form $\overline{f}(z) = \sum_{n=1}^{\infty} \overline{a_n} q^n$, where $\overline{a}$ denotes complex conjugation.

To a cusp form f of weight k for $\Gamma_0(N)$ one associates its L-function via the formula $L(f,s) = \sum_{n=1}^{\infty} a_n n^{-s}$; it can be shown that this series converges for all s with real part $> 1 + \frac{k}{2}$, and that it can be extended to a holomorphic function in s with functional equation

$$\Lambda(f,s) = N^{k/2-s}\Lambda(f|_k[\tau_N], k-s),$$

where $\Lambda(f,s) = (2\pi)^{-s}\Gamma(s)L(f,s)$.

If the primitive cusp form f is an eigenvector for the Hecke operators, we therefore have

$$\Lambda(f,s) = wN^{k/2-s}\Lambda(\overline{f}, k-s),$$

where w has absolute value 1. If, in addition, the Fourier coefficients of f are rational, then we find

$$\Lambda(f,s) = wN^{k/2-s}\Lambda(f, k-s)$$

with $w^2 = 1$.

The conjecture of Shimura-Taniyama-Weil says that if E is an elliptic curve defined over $\mathbb{Q}$ with L-function $L(E,s) = \sum_{n=1}^{\infty} a_n n^{-s}$ and conductor N, then $f(z) = \sum_{n=1}^{\infty} a_n q^n$ defines a primitive cusp form of weight 2 for $\Gamma_0(N)$, which in addition is eigenvector for the Hecke operators. If this is true, we say that E is modular (or a Weil curve). This implies that $L(E,s)$ extends to a holomorphic function of s with functional equation (cf §5). [In fact, E will be modular if and only if the functions $L(E,\chi,s) = \sum_{n=1}^{\infty} a_n \chi(n) n^{-s}$ are holomorphic and satisfy an appropriate functional equation, for all Dirichlet characters $\chi : (\mathbb{Z}/r\mathbb{Z})^{\times} \longrightarrow \mathbb{C}^{\times}$ ([17]); this is how one can show that the curves with complex multiplication are modular.] If E is given, it can be checked in finitely many steps whether E is modular.

Conversely, if f is a primitive cusp form of weight 2 for $\Gamma_0(N)$ and a normalized eigenvector for the Hecke operators such that its Fourier coefficients are rational, then f is the modular form attached to some elliptic curve E_f of conductor N defined over $\mathbb{Q}$. The curve E_f can be determined geometrically from N ([10, Chap. 7]).

Since the curves E_n have complex multiplication, they are modular, and we have $L(E_n,s) = L(E_1,\chi_n,s)$, where χ_n is the quadratic Dirichlet character attached to the quadratic field $\mathbb{Q}(\sqrt{-n}\,)$.

7. Coronidis Loco: Modular forms of half-integral weight

Theta functions are of an exceptional importance in arithmetic. Considering Θ_4 as a modular form, for example, one can easily determine the number of representations of n as a sum of four squares. One would like to generalize this to sums of ℓ squares, where ℓ is not necessarily a multiple of 4, and consider

$$\Theta_\ell = \Theta_1^\ell = \sum_{n=0}^{\infty} r_\ell(n) q^n$$

as a special modular form of weight $\ell/2$. When ℓ is even, this is pretty easy: the notion of a modular form of integral weight can be generalized by introducing a Dirichlet character $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \longrightarrow \mathbb{C}^\times$ and replacing condition i) by

ii) $f|_k[\alpha] = \chi(d) f$ for all $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$.

Thus, for ℓ even, Θ_ℓ is a modular form of weight $\ell/2$ for $\Gamma_0(4)$ and for the character $\chi_{-1}^{\ell/2}$, where χ_{-1} is the non-trivial character of $(\mathbb{Z}/4\mathbb{Z})^\times$.

If ℓ is odd, the generalization is less direct: in fact, the defining formula for $f|_k[\alpha]$ contains $(cz+d)^k$, and for $k = \ell/2$ one has to characterize the square root of $cz+d$. The solution is to take the Θ_ℓ as models: a modular form of weight $\ell/2$ for $\Gamma_0(N)$ and character χ (where N is a multiple of 4 and χ a character $(\mathbb{Z}/N\mathbb{Z})^\times \longrightarrow \mathbb{C}^\times$) is a holomorphic function on $\mathbb{H}$ such that

$$\frac{f(\alpha, z)}{\Theta^\ell(\alpha, z)} = \frac{\chi(d) f(z)}{\Theta^\ell(z)} \quad \text{for } \alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$$

plus some other conditions on the holomorphy at ∞; one also has the notion of a cusp form.

Another difficulty that has prevented any substantial progress prior to the works of Shimura [11] in 1973 was the definition of Hecke operators $T(n)$ as above: a natural definition modeled after the one for integral weight gives $T(p) = 0$ for primes p; for integral weight, the $T(n)$ are functions of $T(p)$ for p prime. But actually, as Shimura has shown, for half-integral weight the basic operators are the $T(p^2)$ for primes p.

Shimura proves a really extraordinary correspondance between forms of half-integral weight and those of integral weight.

Theorem 1. *Let $\ell \geq$ be an odd integer, $k = (\ell-1)/2$, N a multiple of 4, and χ a character of $(\mathbb{Z}/N\mathbb{Z})^\times$. Let f be a modular form for $\Gamma_0(N)$ of weight $\ell/2$ with character χ, and assume that f is eigenvector of*

$T(p^2)$ with eigenvalue λ_p, where p is prime. Then there exists a form g of weight k for $\Gamma_0(N/2)$ with character χ^2, and g is eigenvector of $T(p)$ with eigenvalue λ_p for primes p. The form g is unique up to scalars. If $\ell \geq 5$, g is a cups form.

This Shimura correspondence was studied by Waldspurger [14, 15] using the theory of automorphic representations. Using very deep techniques he studied the surjectivity and the fibers of the Shimura map. His fundamental discovery is a connection between the fibre above g and the values of the functions $L(g, \chi, 1)$ for the Dirichlet characters $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \longrightarrow \mathbb{C}^\times$. In our particular case, taking for g the modular form of weight 2 associated to the curve E_1, one obtains the existence of two modular forms of weight 3/2,

$$f(z) = \sum_{m=1}^{\infty} a_m q^m \quad \text{and} \quad f'(z) = \sum_{m=1}^{\infty} a'_m q^m$$

with integral coefficients a_m and a'_m; their image under the Shimura map is g, and we have

$$L(E_n, 1) = \begin{cases} a_n^2 \beta / 4\sqrt{n} & \text{if } n \text{ is odd,} \\ a'^2_{n/2} \beta / 2\sqrt{n} & \text{if } n \text{ is even,} \end{cases}$$

where $\beta = \int_1^\infty dx/\sqrt{x^3 - x} \approx 2.622\ldots$.

Putting $\Theta = \Theta_1$, Tunnell found that

$$\begin{aligned} f(z) &= (\Theta(z) - \Theta(4z))(\Theta(32z) - \Theta(8z)/2)\Theta(2z) \\ f'(z) &= (\Theta(z) - \Theta(4z))(\Theta(32z) - \Theta(8z)/2)\Theta(4z) \end{aligned}$$

Developing these products in a power series in q, one obtains expressions for a_n and a'_n in terms of the representations of n that occur in the criterion announced in the introduction.

Now we have reached the end of our journey, and have found that the theory of elliptic curves and modular forms of arbitrary weight is linked with the problem of congruent numbers and the problem of representing integers by quadratic forms of three variables. Tunnell's criterium allows us to find – or rediscover – classes of non-congruent numbers: for example, prime numbers congruent to 3 mod 8. Conjecturally, his criterium is also a criterium for congruence: it remains to prove the conjecture of Birch and Swinnerton-Dyer!

References

[1] B.J. Birch, H.P.F. Swinnerton-Dyer, *Notes on elliptic curves. I*, J. Reine Angew. Math. **212** (1963), 7–25: II, ibid. **218** (1965), 79–108 7

[2] J. Coates, A. Wiles, *On the conjecture of Birch and Swinnerton-Dyer*, Invent. Math. **39** (1977), 223–251 8

[3] P. Deligne, *La conjecture de Weil*, Publ. Math. IHES **43** (1974), 273–307 6

[4] B.H. Gross, D. Zagier, *Heegner points and derivatives of L-series*, Invent. Math. **84** (1986), 225–320 8

[5] N. Katz, *An overview of Deligne's proof of the Riemann hypothesis for varieties over finite field*, Proc. Symp. Pure Math. 28 (1976), 275–305 6

[6] N. Koblitz, *Introduction to elliptic curves and modular forms*, Springer-Verlag 1984

[7] J. Lagrange, *Nombres congruents et courbes elliptiques*, Sém. Delange-Pisot-Poitou 1974–75. no. 16

[8] J.-F. Mestre, *Construction d'une courbe elliptique de rang* ≥ 12, C. R. Acad. Sci. Paris **295** (1982), 643–644

[9] J.-P. Serre, *Cours d'Arithmétique*, P.U.F. 1970 10

[10] G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, P.U.F. 1971 10, 11

[11] G. Shimura, *On modular forms of half-integral weight*, Ann. Math. **97** (1973), 440–481; see also Lecture Notes Math. **320** (1973), 59–74 12

[12] J. Silverman, *The arithmetic of elliptic curves*, Springer-Verlag 1986 8

[13] J.B. Tunnell, *A classical diophantine problem and modular forms of weight* 3/2, Invent. Math. **72** (1983), 323–334

[14] J.-L. Waldspurger, *Correspondance de Shimura*, J. Math. Pures Appl. **59** (1980), 1–132 13

[15] J.-L. Waldspurger, *Sur les coefficients de Fourier des formes modulaires de poids demi-entier*, J. Math. Pures Appl. **60** (1981), 375–484 13

[16] A. Weil, *Number of solutions of equations in prime fields*, Bull. Amer. Math. Soc. **55** (1949), 497–508 6

[17] A. Weil, *Über die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen*, Math. Ann. **168** (1967), 149–156 11

Translation by Franz Lemmermeyer