BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 81, Number 1, January 1975

DIOPHANTINE EQUATIONS AND MODULAR FORMS

BY A. P. OGG

An *elliptic curve* over a field K may be defined to be a nonsingular projective plane cubic curve in standard form, which for characteristic $\neq 2, 3$ is

$$E: y^2 = 4x^3 - g_2x - g_3, \tag{1}$$

where $g_2, g_3 \in K$; that E is nonsingular means that the discriminant $\Delta = g_2^3 - 27g_3^2$ is not 0. (A slightly modified cubic equation is required in characteristic 2 or 3.) E has a natural group law, written additively, with the unique point at infinity, $0 = (\infty, \infty)$, as zero, defined by the rule that three points on E add up to 0 if and only if they are collinear. E is then an abelian variety of dimension 1 defined over K. Let $E(K)$ denote the group of points of E with coordinates in K.

Over $K = \boldsymbol{C}$, the field of complex numbers, if we are given a complex torus $\boldsymbol{C}/L$, where $L = \boldsymbol{Z}\omega_1 \oplus \boldsymbol{Z}\omega_2$ is a lattice, then we have an analytic isomorphism

$$\begin{aligned} \boldsymbol{C}/L &\xrightarrow{\sim} E: y^2 = 4x^3 - g_2x - g_3, \\ u &\mapsto (x, y) = (\mathfrak{P}(u), \mathfrak{P}'(u)) \end{aligned} \tag{2}$$

defined by the Weierstrass $\mathfrak{P}$-function. Here g_2 and g_3 depend on the lattice L. The isomorphism carries the natural group law on $\boldsymbol{C}/L$ onto the above geometrically defined group law on E, by the addition theorem for the $\mathfrak{P}$-function. Viewing E as $\boldsymbol{C}/L$, it is clear that the group of N-division points, $E_N = \{P \in E: N \cdot P = 0\}$, is isomorphic to $C_N \times C_N$, where C_N is the cyclic group of order N.

If $K = \boldsymbol{Q}$ is the field of rational numbers, then, by a celebrated theorem of Mordell, the group $E(\boldsymbol{Q})$ is finitely generated:

$$E(\boldsymbol{Q}) = \boldsymbol{Z}^r \times F, \tag{3}$$

where $F = E(\boldsymbol{Q})^t$, the torsion subgroup of $E(\boldsymbol{Q})$, is finite. In practice, for a given elliptic curve E, one can determine the torsion subgroup F rather

---

A talk presented to the American Mathematical Society at its November, 1973 meeting in Tucson. This paper was written in June, 1974, and contains some things which were not known at the time of the meeting; received by the editors June 28, 1974.

AMS (MOS) subject classifications (1970). Primary 10B02, 10D10, 14G05.

Key words and phrases. Elliptic curve, modular curve, rational torsion, descent, Jacobian, isogeny, involution.

easily, and the rank r, with somewhat more difficulty, by some kind of descent argument. On the present occasion, we are concerned with the *modular* version of the first problem, i.e.

Problem 1. As E varies over all elliptic curves defined over $\boldsymbol{Q}$, what torsion subgroups $F=E(\boldsymbol{Q})^t$ occur?

For trivial reasons, F is either cyclic or the product of a cyclic group with a group of order 2. ($E(\boldsymbol{R})$ has at most two components since E is cubic.) Hence, Problem 1 is not much different from the problem of finding what cyclic groups F can contain, i.e. determining for what positive integers N there exists an elliptic curve E defined over $\boldsymbol{Q}$ containing a rational point of order N.

CONJECTURE 1. *Some elliptic curve over $\boldsymbol{Q}$ has a rational point of order N for the eleven values $N=1$–10, 12 and for no other values of N.*

As explained below, these are the values of N for which the relevant modular curve is of genus 0. At the moment, Conjecture 1 has been verified (at least) for all $N<151$, and for all $N<250$ with at most four exceptions. If Conjecture 1 is true, then from known results it is easy to see that there are exactly fifteen possible isomorphy types for F: the eleven cyclic groups mentioned, and the product of a group of order 2 with a cyclic group of order 2, 4, 6, or 8. Conjecture 1 is a precise version, for $K=\boldsymbol{Q}$, of the "folklore conjecture" mentioned by Cassels [1] to the effect that $E(K)^t$ is bounded for any fixed number-field K; Manin has proved that at least the p-torsion is bounded, for any prime p.

A related problem, which has received more attention, although it is in some respects more difficult, and which is the principal subject of this report, is the following. Let F be a finite subgroup of E, an elliptic curve defined over $\boldsymbol{Q}$. F is a *rational group* if $s(F)=F$ for all $s\in \mathrm{Gal}(\overline{\boldsymbol{Q}}/\boldsymbol{Q})$; of course F is rational if $F\subset E(\boldsymbol{Q})$, but the converse is not necessarily true. F is a rational group if and only if F is the kernel of a rational *isogeny*, i.e. the kernel of a homomorphism $E\to E'$, defined over $\boldsymbol{Q}$, of elliptic curves defined over $\boldsymbol{Q}$. In this case we can assume that F is cyclic with no essential loss of generality, and we ask:

Problem 2. For which values of N does there exist a cyclic rational isogeny of elliptic curves, of degree N?

Such rational isogenies are known for twenty-six values of N, namely $N=1$–19, 21, 25, 27, 37, 43, 67, 163; in each of these cases there is an explanation for the existence of the isogenies, as will be shown below. This may very well be a complete list; at any rate I am now inclined to believe that there will be only a finite number of such values. At the moment it seems to be possible, with a little bit of luck, to settle the question for any particular value of N, often with surprising ease; the crucial tool is the new

descent theory of Barry Mazur [6]. Once Problem 2 is settled for a given value of N, it is an easy matter to determine whether the rational group in question contains a rational point or not, so, in practice, Problem 2 takes care of Problem 1.

1. **Modular interpretation.** In the classical case (i.e. over $\boldsymbol{C}$) an elliptic curve is of the form $E=\boldsymbol{C}/L$, where $L=\boldsymbol{Z}\omega_1\oplus\boldsymbol{Z}\omega_2$ is a lattice with period ratio $\tau=\omega_1/\omega_2$ in the upper half-plane $\mathfrak{H}$. E is isomorphic to $E'=\boldsymbol{C}/L'$ if and only if the lattices are proportional, i.e. $L'=tL$, where $t\in\boldsymbol{C}$, i.e. τ and τ' are equivalent under the *modular group* $\Gamma=\mathrm{SL}(2,\boldsymbol{Z})/(\pm 1)$, i.e. $\tau'=(a\tau+b)/(c\tau+d)$ for some $\left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)\in\mathrm{SL}(2,\boldsymbol{Z})$. Thus the space of (isomorphism classes of) elliptic curves over $\boldsymbol{C}$ is parametrized by $Y=\Gamma\backslash\mathfrak{H}$. Let X denote the compactification of Y by adding the point at infinity (the *cusp*); X is of genus 0. Finally, the (*elliptic modular*) *invariant*

$$j=j(E)=j(\tau)=(12g_2)^3/\Delta \tag{4}$$

($\Delta=g_2^3-27g_3^2$) defines an isomorphism of Y onto $\boldsymbol{C}$. j also defines a structure of an algebraic curve defined over $\boldsymbol{Q}$ on Y, and an elliptic curve E is isomorphic to one defined over a field K (i.e. with $g_2, g_3\in K$) if and only if $j(E)\in K$.

Our interest here is not in elliptic curves alone, but in elliptic curves together with a *level N structure*, i.e. some data concerning the points of order N. There are many kinds of level N structures, but we shall discuss here only the two which are relevant to Problems 1 and 2 above. Let us consider pairs (E, Q), resp. (E, C), where E is an elliptic curve and Q is a point of order N on E, resp. C is a cyclic subgroup of E of order N. The given pairs are *isomorphic* to (E', Q'), resp. (E', C') if there exists an isomorphism ϕ of E onto E' with $\phi(Q)=Q'$, resp. $\phi(C)=C'$. In analytic terms, we can choose the basis (ω_1, ω_2) of the lattice so that $Q=\omega_2/N$, resp. C is the group generated by ω_2/N. In the first case, (E, Q) is isomorphic to (E', Q') if and only if τ is equivalent to τ' under the subgroup $\Gamma_1(N)$ of Γ defined by matrices $\left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)$ with $c\equiv 0$ and $a\equiv d\equiv 1 \pmod N$; the space of (isomorphism classes of) pairs (E, Q) is thus parametrized by $Y_1(N)=\Gamma_1(N)\backslash\mathfrak{H}$. Similarly, the space of (isomorphism classes of) pairs (E, C) is parametrized by $Y_0(N)=\Gamma_0(N)\backslash\mathfrak{H}$, where $\Gamma_0(N)$ is the subgroup of Γ defined by matrices $\left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)$ with $c\equiv 0 \pmod N$. Let $X_1(N)$, resp. $X_0(N)$, denote the compactifications of $Y_1(N)$, resp. $Y_0(N)$, by adding cusps (rational boundary points, or ∞). The X's and Y's are algebraic curves defined over $\boldsymbol{Q}$; cf. Shimura [12] for the rational structure and [10] for a description of the cusps, and their fields of rationality. For example, if N is prime then $X_0(N)$ has two cusps, 0 and ∞, both rational over $\boldsymbol{Q}$.

(REMARK. The function field of $X_0(N)$ is $\boldsymbol{C}(j, j_N)$, where $j_N(\tau)=j(N\tau)$. It is classical that j and j_N satisfy an equation with rational coefficients; this gives a plane curve over $\boldsymbol{Q}$ whose desingularization is $X_0(N)$.)

Finally, for any prime p not dividing N, $X_1(N)$ or $X_0(N)$ has a good reduction modulo p, by Igusa [**4**], and so we can speak of points on $X_1(N)$ or $X_0(N)$ rational over any field K whose characteristic does not divide N; the set of such points is denoted by $X_1(N)(K)$ or $X_0(N)(K)$. The description of the cusps [**10**] is the same in characteristic p as in characteristic 0.

THEOREM 1. *Any point of $Y_1(N)$ or $Y_0(N)$, rational over a field K (of characteristic not dividing N), is represented by a K-rational pair (i.e. E is defined over K, and Q is rational over K, or C is a group rational over K), and conversely.*

This was proved by Serre in 1971, for characteristic $\neq 2$, 3, and for characteristic $\neq 0$ (essentially a question of finite fields) by Milne and me (independently) about a year later. The theorem holds for any type of level N structure, by the same proof—cf. [**2**, pp. 132–133]. In view of Theorem 1, we can restate our problems as:

Problem 1′. For which N is $Y_1(N)(\boldsymbol{Q})$ nonempty?

Problem 2′. For which N is $Y_0(N)(\boldsymbol{Q})$ nonempty?

If the genus of $Y_1(N)$ or $Y_0(N)$ is 0, then the set of rational points is in one-one correspondence with $\boldsymbol{Q}$, so we have an infinite number of rational points, in fact a linear family. The genus of $Y_1(N)$ is 0 exactly for $N=1$–10, 12, and Conjecture 1 arose (in not very convincing fashion) from the observation that these were just the N for which rational pairs (E, Q) were known [**1**], [**9**]. The conclusion toward which we are tending seems to be that modular curves only have rational points for which there is a reason. (It may be well to emphasize at this point that the work dealt with in this report is definitely special to the field $\boldsymbol{Q}$. We are not working over $\boldsymbol{Q}$, rather than a general number-field, just for reasons of simplicity.) As to Problem 2, $Y_0(N)(\boldsymbol{Q})$ is infinite if the genus g is 0, i.e. if $N=1$–10, 12, 13, 16, 18, 25, and is finite and known if $g=1$ (cf. [**10**, p. 229]), so we may as well assume $g\geqq 2$ hereafter.

In order to clarify the meaning of Theorem 1, we must discuss the following mildly technical point. The automorphism group $A=\mathrm{Aut}(E)$ of an elliptic curve is finite, and $A=(\pm 1)$ if $j\neq 0$, 1728. If the characteristic is $\neq 2$, 3, then $A=\mu_6$, resp. μ_4, if $j=0$, resp. 1728, where μ_m is the group of mth roots of 1; in characteristic 2, resp. 3, A is a noncommutative group of order 24, resp. 12, if $j=0$. We also have the automorphism groups of pairs: $A(Q)=\mathrm{Aut}(E, Q)$, resp. $A(C)=\mathrm{Aut}(E, C)$, is the group of all $\phi \in A$ with $\phi Q=Q$, resp. $\phi C=C$.

PROPOSITION 1. *For $N \geqq 4$, $A(Q)$ is trivial, while $A(C)$ is cyclic of order* 2, 4, *or* 6.

This is an easy exercise once you know the basic facts about the endomorphism ring of an elliptic curve [**1**, pp. 215–220]. In modern language [**2**], the triviality of $A(Q)$ means that $Y_1(N)$ is a "fine moduli scheme" for $N \geqq 4$; the pair (E, Q) representing a point of $Y_1(N)$ is unique up to a unique isomorphism. On the other hand, $Y_0(N)$ is never fine. One consequence of this is that the K-rational pair (E, C) representing a point of $Y_0(N)(K)$, whose existence is asserted by Theorem 1, is never unique. However, we do get any other such K-rational pair (E', C') by "twisting" (E, C); the set of (K-isomorphism classes of) all such (E', C') is parametrized by a Galois cohomology group $H^1(K, A(C))$. At any rate, a statement such as "$Y_0(N)$ parametrizes cyclic isogenies of degree N" requires careful interpretation.

2. **Rational points on the Jacobian.** We first make some general remarks about finding rational points on curves. Let X be a (projective, nonsingular) curve defined over a number-field K, of genus $g \geqq 2$. Let J be its *Jacobian*: J is the group of divisor classes of degree 0 on X, and can be given the structure of a projective variety of dimension g, defined over K, so that the group law is defined by K-rational functions, i.e. J is an *abelian variety* defined over K. (A *divisor* on X is a formal finite sum $D = \sum_P n_P \cdot (P)$, where $P \in X$ and $n_P \in \boldsymbol{Z}$; its *degree* is $\deg(D) = \sum_P n_P$. Two divisors D, E are *equivalent* if $D - E = (f) = \sum_P \operatorname{ord}_P(f) \cdot (P)$ is the divisor of some function f on X; then $\deg(D) = \deg(E)$. J is the set of equivalence classes of degree 0. We can embed X in J by assigning to $P \in X$ the class of $(P) - (P_0)$, where P_0 is any fixed point of X; the embedding is defined over K if P_0 is rational over K.) By Weil's generalization of Mordell's theorem (3), the group of K-rational points of J is finitely generated:

$$J(K) = \boldsymbol{Z}^r \times F, \tag{5}$$

where F is finite. The general diophantine program is to determine $J(K)$, or at least get some quantitative information about it, and then to find the intersection $X(K) = X \cap J(K)$. Since X is a 1-dimensional subvariety of J, which has dimension $g \geqq 2$, it is reasonable to hope that $X(K)$ will be finite (*Mordell's conjecture*). Taking now $K = \boldsymbol{Q}$, for simplicity, we make two remarks of a practical nature:

(i) It is easier to control the torsion subgroup F than the rank r. For one thing, if p is a prime at which J has a good reduction $\tilde{J}$, then reduction modulo p is injective on F, up to p-torsion. Thus, if we know some subgroup F' of F, we can often conclude that $F' = F$ by calculating $\tilde{J}$ for a few primes p.

(ii) Determining the rank r is a question of descent; the larger F is, the easier the descent will be, and the greater the likelihood that r is small.

Now take $K=\boldsymbol{Q}$, $X=X_0(N)$, $J=J_0(N)$, the Jacobian of $X_0(N)$; for simplicity, let us also suppose that N is a prime >3. The genus is then

$$g = (N - \nu)/12, \tag{6}$$

where $N\equiv\nu \pmod{12}$, and $\nu=13, 5, 7, -1$. $X_0(N)$ has two cusps, 0 and ∞, both rational, and the divisor class of $(0)-(\infty)$ defines a point of $J_0(N)(\boldsymbol{Q})$ of finite order n, where

$$n = \text{num}((N - 1)/12) \tag{7}$$

is the numerator of $(N-1)/12$ (cf. [**10**]); for example $n=23$ if $N=47$. That n is reasonably large is to be considered helpful, as remarked above. A certain amount of numerical experimentation led to:

CONJECTURE 2. *The cyclic group of order n generated by the class of* $(0)-(\infty)$ *is the full torsion subgroup of* $J_0(N)(\boldsymbol{Q})$.

Conjecture 2 has been verified for all $N<250$ (except $N=227$), by reduction modulo suitable primes p (not dividing $n \cdot N$), using the excellent tables of Hecke operators compiled by Wada. (Cf. [**13**], or rather the computer print-out it reports on; Wada's machine apparently broke down at $N=227$.) More recently, Conjecture 2 has been proved up to 2-torsion by Mazur [**6**]; quite possibly he will have settled the problem of the 2-torsion as well by the time that this article has appeared.

(REMARK. Since isogenous abelian varieties over finite fields have the same number of rational points, it is an experimental fact that the torsion subgroup of $J_0(N)(\boldsymbol{Q})$ is maximal among the abelian varieties $\boldsymbol{Q}$-isogenous to $J_0(N)$.)

Equally important is another kind of torsion on $J_0(N)$. By Kummer theory, Shimura discovered another cyclic subgroup of $J_0(N)$ of order given by (7) which is a rational group, and on which Galois groups have the same effect as on the group μ_n of nth roots of 1 (cf. [**10**, p. 230]). We call this group the *Shimura group*, and say that its elements are *μ-rational.*

CONJECTURE 2 (TWISTED). *The Shimura group is the maximal finite μ-rational subgroup of* $J_0(N)$.

For an abelian variety over a finite field $\boldsymbol{F}_q$, it is easy to check that we have the same number of rational and μ-rational points of order prime to the characteristic. (One is given by the kernel of $\pi-1$, the other by the kernel of $\pi-q$, where π is the Frobenius endomorphism.) Hence the same calculations that verified Conjecture 2 for $N<250$, $N\neq227$, also verify Conjecture 2 (twisted) for those values of N.

The final basic fact about $X_0(N)$ is that it possesses an involution $w=w_N$. In terms of matrices, w is defined by $\left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right)$, which normalizes $\Gamma_0(N)$; in terms of pairs, w sends (E, C) to $(E/C, E_N/C)$. w is a rational involution of $X_0(N)$, so the quotient $X_0^+(N)$ is again defined over $\boldsymbol{Q}$. The number $\nu(N)$ of fixed points of w is given by Fricke's formula:

$$\begin{aligned} \nu(N) &= h(-N) + h(-4N) \qquad (\text{if } N \equiv 3 \pmod 4), \\ &= h(-4N) \qquad\qquad\quad (\text{otherwise}), \end{aligned} \tag{8}$$

where $h(-d)$ is the class-number (of quadratic forms of discriminant $-d$); the genus of $X_0^+(N)$ is

$$g^+ = (g+1)/2 - \nu(N)/4. \tag{9}$$

w also defines an involution w of $J_0(N)$, which we then break up into $+1$ and -1 eigenspaces, so to speak: $J_0^+(N)$, resp. $J_0^-(N)$, is the connected component of the kernel of $(w-1)$, resp. $(w+1)$, on $J_0(N)$; the dimension is g^+, resp. $g^-=g-g^+$.

It is clear that w acts as -1 on the rational group of order n on $(0)-(\infty)$, since w interchanges 0 and ∞; it is true that $w=-1$ on the Shimura group as well. Thus our known torsion lies in the minus part of $J_0(N)$. It is not true that $J_0^-(N)(\boldsymbol{Q})$ is always finite; however, Mazur's descent theory [6] has allowed him to prove that $J_0^-(N)$ has a nontrivial factor, which he calls the *chosen* factor, which has only finitely many rational points. (At least this is so if n is divisible by a prime $p \geqq 5$; if not, a condition of "regularity of the Eisenstein ideal" must be verified. I leave the description of Mazur's result rather vague, since it is undergoing constant improvement; anyway, we will not have to wait too long before the full account [6] is available.) This implies the Mordell conjecture (finiteness of $X_0(N)(\boldsymbol{Q})$) for most—probably all—of the curves $X_0(N)$.

It remains to make this result quantitative, i.e. to find $X_0(N)(\boldsymbol{Q})$ exactly (Problem 2). Three approaches, in the order in which they come to mind, are:

(i) Exploit the geometry of $X_0(N)$, e.g. the Riemann-Roch theorem, knowledge of Weierstrass points, etc.

(ii) Obtain numerical information by reducing modulo convenient good primes p (i.e. $p \neq N$).

(iii) Reduce modulo the bad prime N (Deligne-Rapoport-Mazur).

A fair amount of ingenuity was expended on (i) before I realized that (ii) is usually easier; (ii) will be discussed, largely by example, in the next section. (iii) is the most powerful of all; it uses ideas which are much more sophisticated than those discussed so far in this article.

3. **Finite fields.** Let $K=\boldsymbol{F}_q$ be the finite field of q elements, and E an elliptic curve defined over K; let $\pi: E\to E$ be the Frobenius endomorphism, i.e. $\pi(x, y)=(x^q, y^q)$. By the theory of the endomorphism ring [1], π (or any other endomorphism) has a transpose $\bar{\pi}: E\to E$, such that $\pi+\bar{\pi}\in \boldsymbol{Z}$, and $\pi\cdot\bar{\pi}=\deg(\pi)=q$. Let $-d=(\pi-\bar{\pi})^2$; then d is an integer, $d\geqq 0$.

Let N be a positive integer relatively prime to q, and let C be a cyclic subgroup of E of order N. Clearly C is K-rational if and only if $\pi C=C$, i.e. $\pi-a$ vanishes on C for some integer a. From this we conclude easily that:

PROPOSITION 2. (i) *If $\pi\in \boldsymbol{Z}$, then all C's are K-rational.*
(ii) *If $\pi\notin \boldsymbol{Z}$, and N is relatively prime to $2d$, then there are*

$$\prod_q \left(1+\left(-\frac{d}{q}\right)\right)$$

subgroups C of E, cyclic of order N and rational over K; here q runs over the prime factors of N, and

$$\left(-\frac{d}{q}\right)$$

is the Legendre symbol.

Of course there are cases not covered by the proposition, but we will not state the general result here. The point is that it is easy to count the K-rational subgroups C of a given E; the more subtle question is to determine just how many points of $Y_0(N)(K)$ are thus represented (cf. remarks at the end of §1). In this direction we have the

LEMMA. *Let (E, C) and (E', C') be isomorphic pairs, where E and E' are defined over K, and $A(C)=\mathrm{Aut}(E, C)=A=\mathrm{Aut}(E)$. Then C is K-rational if and only if C' is.*

PROOF. Let ψ be an isomorphism of E onto E' with $\psi C=C'$. The conjugate ψ^π is also an isomorphism of E onto E', and $\phi=\psi^{-1}\circ\psi^\pi\in A=A(C)$. Now $\psi^\pi\circ\pi=\pi'\circ\psi$, where π' is the Frobenius endomorphism of E', i.e. $\psi\circ\phi\circ\pi=\pi'\circ\psi$. If $\pi C=C$, then

$$\pi' C' = \pi'\psi C = \psi\phi\pi C = \psi\phi C = \psi C = C'.$$

In particular, the Lemma applies if $j\neq 0$, 1728, when $A=\pm 1$, and we get

THEOREM 2. *Let $j\in K=\boldsymbol{F}_q$, $j\neq 0, 1728$. Choose any E defined over K with $j(E)=j$. Then the number of points of $Y_0(N)(K)$ over j is the same as the number of subgroups C of E which are K-rational and cyclic of order N.*

The following theorem gives a sample of what happens when there are additional automorphisms. Recall that in characteristic 2, resp. 3, the elliptic curve with $j=0$ has 24, resp. 12, automorphisms; however $A(C)$ is cyclic of order $m=2$, 4, or 6, if $N \geqq 4$, by Proposition 1.

THEOREM 3. *Let N be a prime >3. Then the points of $Y_0(N)$ over $j=0$, in characteristic 2 and 3, are as follows:*

(A) *In characteristic 2, we have $6 \cdot \nu(4)+4 \cdot \nu(6)+12 \cdot \nu(2)=N+1$, where $\nu(m)$ is the number of points of $Y_0(N)$ over $j=0$ represented by pairs (E, C) with $\mathrm{Aut}(E, C)$ of order m.*

(i) *If $N \equiv -1 \pmod 4$, then $\nu(4)=0$. If $N \equiv 1 \pmod 4$, then $\nu(4)=1$; the point is F_2-rational and fixed by the involution w.*

(ii) *If $N \equiv -1 \pmod 6$, then $\nu(6)=0$. If $N \equiv 1 \pmod 6$, then $\nu(6)=2$; the two points are rational over F_4, conjugate over F_2, and interchanged by w.*

(iii) *The points with $m=2$ are all rational over F_4. The number of them which are F_2-rational is 1 if*

$$\left(-\frac{8}{N}\right) = +1,$$

and otherwise 0.

(B) *In characteristic 3, $3 \cdot \nu(4)+2 \cdot \nu(6)+6 \cdot \nu(2)=N+1$.*

(i) *There is one point with $m=6$, F_3-rational and fixed by w, if $N \equiv 1 \pmod 6$, otherwise none.*

(ii) *There are two points with $m=4$, F_9-rational, conjugate over F_3, and interchanged by w, if $N \equiv 1 \pmod 4$, otherwise none.*

(iii) *The points with $m=2$ are all rational over F_9. The number of them which are F_3-rational is 1 if $N \equiv 1 \pmod 6$, and otherwise 0.*

Finally, one can count the number of values of $j \in F_q$ such that the corresponding E has a given endomorphism ring, by theorems of Deuring [3]. This, together with the generalizations of the theorems above, reduces the question of counting the number of points on $Y_0(N)$ over F_q to a computation involving class-numbers and Legendre symbols.

In characteristic 2, for example, $X_0(N)$ has 2 cusps, F_2-rational, and points over $j=0$, F_2- and F_4-rational as described by Theorem 3; in addition we have points over $j \neq 0$, ∞ as follows. As above, put $-d=(\pi-\bar{\pi})^2=(\pi+\bar{\pi})^2-4q$; d will be odd and positive for $j \neq 0$. Then we have

$$\left(1+\left(-\frac{7}{N}\right)\right)$$

points over $j=1$ on $X_0(N)(F_2)$, and

$$2\left(1+\left(-\frac{15}{N}\right)\right)$$

points over F_4, and

$$3\left(1+\left(-\frac{31}{N}\right)\right)+3\left(1+\left(-\frac{23}{N}\right)\right)$$

points over F_8, and so forth.

EXAMPLE. Let c_i be the number of points on $X_0(47)(F_{2^i})$. We have 2 cusps, F_2-rational, and 4 points over $j=0$, F_4-rational only. Since

$$\left(-\frac{7}{47}\right)=-1, \qquad c_1=2;$$

since

$$\left(-\frac{15}{47}\right)=+1, \qquad c_2=2+4+4=10.$$

Continuing, we find $c_3=14$, $c_4=18$.

The formulas connected with the zeta-functions of curves over finite fields (Weil [**14**, pp. 60–72]) give us the number of points on the Jacobian, rational over a finite field, as follows. Let X be a curve of genus $g\geqq 1$ over F_q, and let J be its Jacobian. Let h be the order of $J(F_q)$, and let c_i be the order of $X(F_{q^i})$. We have $2g$ complex numbers $\pi_1, \pi_2, \cdots, \pi_{2g}$ (the eigenvalues of the Frobenius endomorphism on the first cohomology group) satisfying the *Lefschetz formula*

$$1+q^i-c_i=s_i=\pi_1^i+\pi_2^i+\cdots+\pi_{2g}^i. \tag{10}$$

We assume that $c_1, c_2, \cdots, c_g$ are known, and hence $s_1, s_2, \cdots, s_g$ are known. Put

$$f(x)=\prod_{i=1}^{2g}(x-\pi_i)=x^{2g}+a_1x^{2g-1}+\cdots+a_{2g-1}x+a_{2g}. \tag{11}$$

Here a_i is the ith elementary symmetric function in $-\pi_1, -\pi_2, \cdots, -\pi_{2g}$. We can order the π_i so that $\bar{\pi}_i=\pi_{g+i}$, for $1\leqq i\leqq g$; note that $a_{2g-i}=q^{g-i}a_i$. Putting $b_i=\pi_i+\bar{\pi}_i$ we get

$$f(x)=\prod_{i=1}^{g}(x-\pi_i)(x-\bar{\pi}_i)=\prod_{i=1}^{g}(x^2-b_ix+q).$$

With $y=x+q/x$ we have

$$\begin{aligned} F(y)=x^{-g}f(x)&=\prod_{i=1}^{g}(y-b_i)\\ &=(x^g+(q/x)^g)+a_1(x^{g-1}+(q/x)^{g-1})\\ &\quad+\cdots+a_{g-1}(x+q/x)+a_g\\ &=a_g+a_{g-1}y+a_{g-2}(y^2-2q)+a_{g-3}(y^3-3qy)+\cdots. \end{aligned} \tag{12}$$

Finally, $a_1, a_2, \cdots, a_g$ are calculated from $s_1, s_2, \cdots, s_g$ by the formulas (from any book on symmetric functions)

$$\begin{aligned} -a_1 &= s_1, \\ -2a_2 &= a_1 s_1 + s_2, \\ -3a_3 &= a_2 s_1 + a_1 s_2 + s_3, \\ &\cdots \\ -ga_g &= a_{g-1} s_1 + \cdots + a_1 s_{g-1} + s_g. \end{aligned} \tag{13}$$

The formula for $h=h_q=|J(F_q)|$ is

$$h = f(1) = F(q+1). \tag{14}$$

Let us now analyze $X_0(47)$. We have $g=4$, and (0)–(∞) defines a rational point of order $n=23$ (cf. (7)) on $J_0(47)$. Taking $q=2$, we found above $c_1, \cdots, c_4$, so by the scheme above we find

$$\begin{aligned} c_1 &= 2 & s_1 &= 1 & a_1 &= -1 \\ c_2 &= 10 & s_2 &= -5 & a_2 &= 3 \\ c_3 &= 14 & s_3 &= -5 & a_3 &= -1 \\ c_4 &= 18 & s_4 &= -1 & a_4 &= 3 \end{aligned}$$

and

$$\begin{aligned} F_2(y) &= 3 - y + 3(y^2 - 4) - (y^3 - 6y) + (y^4 - 8y^2 + 8) \\ &= y^4 - y^3 - 5y^2 + 5y - 1. \end{aligned}$$

Then $h_2=F_2(3)=23$. Similarly, we find $h_3=F_3(4)=7 \cdot 23$, which is enough to check Conjecture 2 for this case. $F_2(y)$ is irreducible; since it is the characteristic polynomial of the Hecke operator T_2 on the cusp forms of weight 2 for $\Gamma_0(47)$, it follows that $J_0(47)$ is irreducible over $\boldsymbol{Q}$. Then Mazur's "chosen factor" is all of $J_0(47)$, so $J_0(47)(\boldsymbol{Q})$ is the cyclic group of order 23 on (0)–(∞).

Suppose $Y_0(47)$ has a rational point P. Then (P)–(∞) is equivalent to $m((0)$–$(\infty))$, where m is an integer modulo 23, and $m \not\equiv 0, 1$. Reducing modulo 2, which is injective on our group of order 23, we have the same equation for the reduced points, which is not possible since we saw above that $Y_0(47)(\boldsymbol{F}_2)$ is empty. Hence $Y_0(47)(\boldsymbol{Q})$ is empty, i.e. no elliptic curve over $\boldsymbol{Q}$ admits a rational isogeny of degree 47.

Similar calculations show that $Y_0(N)(\boldsymbol{Q})$ is empty for $N=29, 31, 41, 59, 71$; for $N=23$, a different argument was required, using an equation for the curve, but we still have that $Y_0(23)(\boldsymbol{Q})$ is empty. These are the primes N for which $g^+=0$ (cf. (9)), i.e. $X_0(N)$ is hyperelliptic with hyperelliptic involution w (cf. [**11**]); as such these cases are atypically easy. (Our knowledge is chiefly about the minus part of $J_0(N)$, so it simplifies matters if the plus part vanishes.)

4. **Rational points on $X_0(N)$ in the general case.** Let N be a prime with $g \geqq 2$; assume that $N \neq 23, 29, 31, 41, 47, 59, 71$, where we know that $Y_0(N)(\boldsymbol{Q})$ is empty, i.e. assume that $g^+ > 0$.

There are two cases where rational points on $Y_0(N)$ are known. The first case is that of the rational fixed points of the canonical involution w. Such a point is represented by an elliptic curve over $\boldsymbol{Q}$ with complex multiplication by $\sqrt{-N}$. Since the curve is defined over $\boldsymbol{Q}$, the class-number of $\boldsymbol{Q}(\sqrt{-N})$ is 1, and so $N = 43, 67, 163$ by the theorem of Heegner-Stark-Baker. Conversely, for these three values, there is exactly one rational fixed point of w. (*Whimsical remark.* Perhaps there will be a proof someday that $Y_0(N)(\boldsymbol{Q})$ is empty for N sufficiently large, thus giving another solution to the problem of class-number 1.)

More interesting is the very special case of $N = 37$, exhaustively analyzed in [7]. $X_0(37)$ is hyperelliptic, since $g = 2$, i.e. $X_0(37)$ divided by a certain involution v is of genus 0. However $v \neq w$, since $g^+ = 1$. Putting $u = v \cdot w$ (also an involution), $J_0^-(37)$ is the same as $X_0(37)$ divided by u; it is of genus 1 and has exactly $3 = \text{num}((37-1)/12)$ rational points. Lifting back, $X_0(37)$ has at most six rational points; actually there are four: the two rational cusps, and their images under v. Thus $Y_0(37)(\boldsymbol{Q})$ consists of two points, interchanged by w.

It is shown in [**11**] that $N = 37$ is the only case where $X_0(N)$ is hyperelliptic with an "exceptional" hyperelliptic involution v. We assume henceforth that $N \neq 37$ as well as $g^+ > 0$ (already assumed above). We are then in the general situation: $X_0(N)$ is not hyperelliptic, and the plus-factor of $J_0(N)$ is not trivial. In this case the mapping

$$X_0(N) \to J_0^-(N), \tag{15}$$

carrying P to the class of $(P)-(wP)$, is injective on the complement of the set of fixed points of w. In many cases, this allows us to proceed with the same facility as in the cases at the end of the previous section, where the plus-factor was trivial. For example, for $N = 83$, $g = 7$ and $g^+ = 1$; Mazur's descent theory shows that $J_0^-(83)(\boldsymbol{Q})$ is cyclic of order $n = 41$, and by reduction modulo 2 we find that $Y_0(83)(\boldsymbol{Q})$ is empty.

(*Details.* $X_0(83)(\boldsymbol{F}_2)$ consists of 3 points, the two cusps, and one point over $j = 0$, with $m = 2$, necessarily fixed by w. As before, a rational point of $Y_0(83)$ would then be fixed by w, which we know is not the case.) The same sort of method shows that $Y_0(N)(\boldsymbol{Q})$ is empty for $N = 61, 79, 89$, and many (but not all!) other values of N.

Now suppose that $Y_0(N)$ has a rational point P. We assume that P is not fixed by w. (As shown above, there is exactly one such point for $N = 43, 67, 163$, and for no other N.) Suppose that $(P)-(wP)$ is in our

group of order n on (0)–(∞):

$$(P) - (wP) \sim m((0) - (\infty)), \tag{16}$$

where m is an integer modulo n, $m \neq 0, \pm 1$. ($m=0$ only if P is fixed by w, and $m = \pm 1$ only if P is a cusp.) This assumption will be satisfied if we know that $J_0^-(N)(\boldsymbol{Q})$ is finite, and cyclic of order n; by Mazur's descent theory, and by some recent work of A. Brumer and K. Kramer on the nonchosen minus-factors, this will be true for all $N<250$ with the possible exceptions of $N=151, 193, 199, 227$. By using the description of Deligne and Rapoport [**2**] of the Néron model of $X_0(N)$ at the bad prime N, Mazur [**6**] concludes that actually $m = \pm\frac{1}{3}$, say $m=\frac{1}{3}$, i.e. $N \equiv -1 \pmod 3$ and we have

$$3(P) + (\infty) \sim 3(wP) + (0). \tag{17}$$

From this Mazur quickly concludes that $N=53, 113, 137$ are the only possible values of N among $N<250$ ($N=23, 29, 41$ having been eliminated already). On the other hand, a trick I developed in [**11**] shows that (17) will not hold for any $N>250$. (Reducing modulo 2, we have a map of degree at most 4, defined over $\boldsymbol{F}_2$, of $X_0(N)$ onto the projective line. Hence $X_0(N)(\boldsymbol{F}_4)$ has at most $4(1+4)=20$ points. Since it has more than $N/12$ points over $j=0$ (cf. Theorem 3), we get $N<240$.) As it happens, for many values of N both this method and the more naive method described above of reducing modulo suitable primes $p \neq N$ are successful; unfortunately, both seem to fail for $N=53, 113, 137$. However, one can no doubt settle these cases by a closer examination (finding an equation for the curves if it comes to that), but for the moment we know only that $Y_0(N)(\boldsymbol{Q})$ has 0 or 2 points.

To sum up, all of these various methods have shown that if N is a prime $\geqq 23$, then $Y_0(N)(\boldsymbol{Q})$

(i) has exactly two points for $N=37$,
(ii) has exactly one point for $N=43, 67, 163$,
(iii) has 0 or two points for $N=53, 113, 137$,
(iv) is not known for $N=151, 193, 199, 227$, and
(v) is empty for all other $N<250$.

Looking ahead, $X_0(N)$ has not yet been analyzed for the composite values $N=65, 91, 125, 169$, but if we assume that that will be done and that the various gaps in our knowledge for a prime $N<250$ will be filled, the essential difficulty seems to be the following. For a large prime N it is expected that $J_0^-(N)$ will have a relatively large "chosen" part, to which Mazur's descent theory applies, and a smaller but nontrivial nonchosen part, which in some cases (e.g. $N=389$) will have an infinite number of rational points. It may well be that the mapping (15), followed

by the projection onto the chosen factor, is still injective, but even if that is so it will be hard to study the map systematically. Perhaps it is more realistic to hope that some entirely new idea will come to our rescue for large N.

References

1. J. W. S. Cassels, *Diophantine equations with special reference to elliptic curves*, J. London Math. Soc. **41** (1966), 193–291. MR **33** #7299.

2. P. Deligne and M. Rapoport, *Les schémas de modules de courbes elliptiques. Modular functions of one variable.* II, Lecture Notes in Math., vol. 349, Springer, Berlin and New York, 1973, pp. 143–316.

3. M. Deuring, *Die Typen der Multiplikatorenringe elliptischer Funktionenkörper.* Abh. Math. Sem. Univ. Hamburg **14** (1941), 197–272. MR **3**, 104.

4. J. Igusa, *Kroneckerian model of fields of elliptic modular functions*, Amer. J. Math. **81** (1959), 561–577. MR **21** #7214.

5. D. Kubert, *Universal bounds on the torsion of elliptic curves*, Preprint, Yale University, Department of Mathematics, November 1973.

6. B. Mazur, *Modular curves and the Eisenstein ideal* (in preparation).

7. B. Mazur and P. Swinnerton-Dyer, *Arithmetic of Weil curves*, Invent. Math. **25** (1974), 1–62.

8. B. Mazur and J. Tate, *Points of order* 13 *on elliptic curves*, Invent. Math. **22** (1973), 41–49.

9. A. Ogg, *Rational points of finite order on elliptic curves*, Invent. Math. **12** (1971), 105–111. MR **45** #178.

10. ———, *Rational points on certain elliptic modular curves*, Proc. Sympos. Pure Math., vol. 24, Amer. Math. Soc., Providence, R.I., 1973, pp. 221–231.

11. ———, *Hyperelliptic modular curves*, Bull. Math. Soc. France (to appear).

12. G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Kanô Memorial Lectures, no. 1, Publ. Math. Soc. Japan, no. 11, Shoten, Tokyo; Princeton Univ. Press, Princeton, N.J., 1971. MR **47** #3318.

13. H. Wada, *A table of Hecke operators*, Proc. Japan Acad. **49** (1973), 380–384.

14. A. Weil, *Sur les courbes algébriques et les variétés qui s'en déduisent*, Actualités Sci. Indust., no. 1041=Publ. Inst. Math. Univ. Strasbourg **7** (1945), Hermann, Paris, 1948. MR **10**, 262.

Department of Mathematics, University of California, Berkeley, California 94720