Arithmetic of elliptic curves and diophantine equations

Loïc Merel

Introduction and background

In 1952, P. Dénes, from Budapest [1], conjectured that three non-zero distinct n-th powers can not be in arithmetic progression when $n > 2$ [15], *i.e.* that the equation

$$x^n + y^n = 2z^n$$

has no solution in integers x, y, z, n with $x \neq \pm y$, and $n > 2$. One cannot fail to notice that it is a variant of the Fermat-Wiles theorem. We would like to present the ideas which led H. Darmon and the author to the solution of Dénes' problem in [13]. Many of them are those (due to Y. Hellegouarch, G. Frey, J.-P. Serre, B. Mazur, K. Ribet, A. Wiles, R. Taylor, ...) which led to the celebrated proof of Fermat's last theorem. Others originate in earlier work of Darmon (and Ribet).

The proof of Fermat's last theorem can be understood as an advance in the direction of the *abc* conjecture of D. Masser and J. Oesterlé. We view the solution to Dénes' conjecture as a modest further step. We would like to explain the additional techniques involved into this solution (and try to avoid too much overlap with the numerous other surveys on closely related topics).

A current approach to diophantine problems can be roughly described as follows. One is interested in problems of the following type.

I. Determination of the solutions to a diophantine equation of the form $a + b + c = 0$.

Using a machinery proposed by Hellegouarch, Frey and Serre, and established by theorems of Mazur, Ribet and Wiles (and later refinements by F. Diamond and K. Kramer) the problem **I** might be reduced to problems in the following theme.

[1] Ce texte est la version écrite de l'exposé donné lors du deuxième colloque de la société mathématique européenne à Budapest. À la grande surprise de l'auteur, il ne figurait pas dans les volumes publiés à cette occasion par Birkhaüser Verlag, lorsque ceux-ci sont apparus sur les rayons des librairies. Suite à mon exposé sur le même thème aux Journées Arithmétiques, les éditeurs du présent volume ont accepté de publier l'article en dépit de de la nature plus spécialisée du colloque.

II. The action of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on a few torsion points of an elliptic curve over $\mathbf{Q}$ characterizes the isogeny class of the curve.

or alternately

II'. same problem restricted to a certain class of elliptic curves (see section 1.1) over $\mathbf{Q}$ often called Frey curves.

The problem **II** and **II'** can be reformulated in terms of inexistence of rational points on modular curves (or generalized modular objects). This inexistence can sometimes be established using some algebraic geometry and, roughly speaking,

III. a diophantine argument coming from the study of the Galois cohomology of some modular object.

We illustrate this process by examples arranged in increasing order of difficulty.

Fermat's last Theorem

According to our picture, the proof of Wiles' theorem decomposes as follows. Fermat's Last Theorem is our problem of type **I**:

Theorem 0.1 (Wiles) *The equation $x^p + y^p = z^p$ has no solution in integers x, y and z with p prime number > 3 and $xyz \neq 0$.*

The modular machinery reduces it to a problem of type **II**:

Theorem 0.2 (Mazur) *An elliptic curve over $\mathbf{Q}$ has no $\mathbf{Q}$-rational subgroup of prime order $p > 163$.*

Weaker versions of type **II'** of the previous theorem are sufficient for the application to Fermat's Last Theorem:

Theorem 0.3 (Mazur) *A Frey curve has no $\mathbf{Q}$-rational subgroup of prime order $p > 3$ (the inexistence of $\mathbf{Q}$-rational points of prime order $p > 3$ on Frey curves is sufficient).*

The previous two theorems of Mazur rely in an essential way on the following result of type **III**:

Theorem 0.4 (Mazur) *The Eisenstein quotient of the jacobian of the modular curve $X_0(p)$ has finitely many $\mathbf{Q}$-rational points.*

Dénes' equation

Dénes' conjecture has been proved for $n = 4$ by Euler, for $n = 3$ by Legendre, for $n = 5$ by Dirichlet, and for $n = p$ prime with $p < 31$ by Dénes himself using the methods of Kummer.

The problem was reconsidered by Ribet [63] recently in the light of the proof of Fermat's last theorem: Dénes' equation can not have any non-trivial solution when $n = p$ is a prime congruent to 1 modulo 4 (see below). Darmon and the author proved the following theorem of type **I**.

Theorem 0.5 *The equation $x^p + y^p = 2z^p$ has no solution in integers x, y and z with p prime number > 2, and $x \neq \pm y$.*

Using the modular machinery we reduced first the problem to a positice answer to the following problem of type **II**.

Problem 0.6 (Serre) *Does there exists a number $B > 0$ such that for every every elliptic curve E over $\mathbf{Q}$ without complex multiplication and every prime number $p > B$ any automorphism of the group of points of p-division of E is given by the action of an element of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$?*

We proved a result of type **II'** in the direction of a positive answer to that question. This result is sufficient for the application to Dénes' conjecture.

Theorem 0.7 *For every Frey curve E except the one which has complex multiplication and every prime number $p > 3$ any automorphism of the group of points of p-division of E is given by the action of an element of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$.*

We obtain this result as a consequence of a theorem of type **III**, which is a special case of the conjecture of Birch and Swinnerton-Dyer (see section 2.3.).

Theorem 0.8 (Kolyvagin and others) *Any quotient defined over $\mathbf{Q}$ of the jacobian of a modular curve whose L-function does not vanish at the point 1 has finitely many $\mathbf{Q}$-rational points.*

Concerning Serre's problem in full generality we can offer only wishful thinking for reasons explained in section 2.4: One might hope to use as argument **III** more results in the direction of the conjecture of Birch and Swinnerton-Dyer such as the Gross-Zagier formula and the theorem of Kolyvagin for elliptic curves of rank one.

The Generalized Fermat equation

Fermat's Last Theorem and Dénes' conjecture are essentially special cases of the following problem of type **II**.

Conjecture 0.9 *Let a, b and c be three non-zero integers. One has the equality $ax^n + by^n + cz^n = 0$ for only finitely many values of (x^n, y^n, z^n) where x, y, z, and $n > 3$ are integers and x, y, and z are coprime.*

See [26] for more background on this problem. Observe that in the case where $a + b + c = 0$, the equation has solution in (x, y, z) for any exponent n. This difficulty appears already in Dénes' equation.

Mazur proved that conjecture 0.9 holds when $a = b = 1$ and c is a power of an odd prime number l which is not a Fermat prime or a Mersenne prime [66]. A. Kraus seems to have found recently explicit bounds for the solutions in terms of l [40].

Frey proved that the conjecture 0.9 is a consequence of a conjecture of type **II**. (Questions of this type were raised by Mazur in [50].)

Conjecture 0.10 (Frey) *Let E be an elliptic curve over $\mathbf{Q}$. There exists finitely many pairs (E', p) where E' is an elliptic curve over $\mathbf{Q}$ and p a prime number > 5 such that the sets of p-division points of E and E' are isomorphic as* $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$*-modules.*

Frey even shows ([24], [25]) that the conjecture 0.9 is equivalent to the restriction to Frey curves of the conjecture 0.10.

Darmon conjectures that there exist only finitely many triples (E, E', p) as in conjecture 0.10 with E and E' non-isogenous and $p > 5$.

No argument of type **III** seems to be available to solve the conjecture 0.10. One can only hope that the Birch and Swinnerton-Dyer conjecture is still relevant to the study of the arithmetic of twisted modular curves and their jacobian.

The Generalized Fermat-Catalan equation

Here is another well-known generalization of Fermat's equation.

Conjecture 0.11 *Let a, b and c be three non-zero integers. One has the equality $ax^r + by^s + cz^t = 0$ for only finitely many values of (x^r, y^s, z^t) with x, y, z, r, s, and t integers, x, y, z coprime, $xyz \neq 0$, and $\frac{1}{r} + \frac{1}{s} + \frac{1}{t} < 1$.*

The *abc* conjecture easily implies conjecture 0.11. For $a = b = -c = 1$, the conjecture 0.11 is sometimes called the Fermat-Catalan conjecture since it combines Fermat's theorem with the Catalan conjecture; The ten known triple (x^r, y^s, z^t) which satisfy the equality $x^r + y^s = z^t$ are listed in [3]. In the direction of the Generalized Fermat-Catalan conjecture, we have the following two results. The first is obtained as an application of Faltings' theorem [12]:

Theorem 0.12 (Darmon, Granville) *Let r, s and t be three positive integers such that $\frac{1}{r} + \frac{1}{s} + \frac{1}{t} < 1$. Let a, b, c be three non-zero integers. Then the equation $ax^r + by^s = cz^t$ has finitely many solutions in coprime integers x, y and z.*

Moreover the techniques developped by Darmon and the author to study Dénes' conjecture combined with earlier work by Darmon [8] led us to the following [13]:

Theorem 0.13 *1) Let n be an integer > 3. The equation $x^n + y^n = z^2$ has no solution in coprime and non-zero integers x, y, z.*

2) Suppose that every elliptic curve over $\mathbf{Q}$ is modular. Let n be an integer > 2. Then the equation $x^n + y^n = z^3$ has no solution in coprime and non-zero integers x, y, z.

The study of these equations has some history. For instance the case $n = 4$ in the first equation was solved by Fermat and the case $n = 3$ in the second equation was solved by Euler. We relied on work of B. Poonen to solve these equations by elementary methods for a few small values of n [62].

There is a considerable bibliography on the subject of diophantine equations of this type. We are very far from being complete here. The interested reader might consult [10], [12], [18], and [53].

The *abc* conjecture

We recall its statement for the convenience of the reader [59].

Conjecture 0.14 (Masser-Oesterlé) *For every real number $\epsilon > 0$ there exists a real $K_\epsilon > 0$ such that for every triple of coprime non-zero integers (a, b, c) satisfying $a + b + c = 0$ we have the following inequality:*

$$\mathrm{Sup}(|a|, |b|, |c|) < K_\epsilon(\mathrm{Rad}(abc))^{1+\epsilon},$$

where $\mathrm{Rad}(n)$ for any integer n is the product of the prime numbers dividing n.

The *abc* conjecture illustrates (and maybe concentrates) the difficulty of understanding how the additive and multiplicative structures of the integers relate.

Its resolution would affect our understanding of many problems in number theory ([51], [58], [76]).

As explained in [25] (see also [44], [73] and [74]), it is now a consequence of the following *degree conjecture*, which is a problem of type **II**:

Conjecture 0.15 *Let E be a modular elliptic curve over $\mathbf{Q}$ of conductor N. Let δ_E be the degree of the corresponding minimal modular parametrization $X_0(N) \longrightarrow E$. Let ϵ be a real number > 0. There exists a number T_ϵ independent of E such that one has*

$$\delta_E < T_\epsilon N^{2+\epsilon}.$$

Remark that this is no longer a uniformity statement since the degree δ_E is expected to be bounded by a number depending on the conductor of E. This problem is apparently related to the existence of rational points on certain Hilbert modular varieties.

This paper is organized to be accessible to a wide audience in its first part which explains the process of going from problem **I** to problem **II**. The second part might contain results of interest to the specialist (see sections 2.4 and 2.5) and deals mainly with Serre's problem. In the third part we explain how the degree conjecture can be studied from an elementary point of view.

Acknowledgement: I would like to thank H. Darmon, B. Edixhoven, I. Chen, and K. Ribet for comments useful in the course of writing this paper.

1 The modular machinery

1.1 The curves of Frey and Hellegouarch

In order to study equations of the type $a+b+c=0$, with a, b and c coprime non-zero integers, Frey, following earlier work by Hellegouarch about Fermat's equation ([30], [31]), proposed to consider the elliptic curve $E_{a,b,c}$ over $\mathbf{Q}$ given by the cubic equation

$$y^2 = x(x-a)(x+b).$$

We can assume without loss of generality that b is even and that a is congruent to -1 modulo 4. We call elliptic curves given by cubic equations of this form *Frey curves.* Elementary facts concerning them can be found in various places of the literature ([22], [23], [67], [59]). The modular invariant of $E_{a,b,c}$ is given by the formula

$$j(E_{a,b,c}) = 2^8 \frac{(c^2-ab)^3}{(abc)^2}.$$

In particular $j(E_{a,b,c})$ is not integral (and $E_{a,b,c}$ has no complex multiplication) except in the case $(a,b,c)=(-1,2,-1)$.

All the 2-division points of $E_{a,b,c}$ are $\mathbf{Q}$-rational. The exact formula for the conductor $N_{a,b,c}$ of $E_{a,b,c}$ has been calculated by Diamond and Kramer [17]. One has

$$N_{a,b,c} = \mathrm{Rad}(abc)\epsilon_2(b),$$

where $\epsilon_2(b)=1$ if $32|b$, $\epsilon_2(b)=\frac{1}{2}$ if $16|b$ and $32 \not| b$, $\epsilon_2(b)=4$ if $4|b$ and $16 \not| b$, $\epsilon_2(b)=16$ if $4 \not| b$. In particular $E_{a,b,c}$ is a semi-stable elliptic curve if and only if $16|b$.

1.2 Wiles' theorem

The work of Wiles and the notion of modular elliptic curve has been explained in many expository articles recently. For a different perspective we explain the meaning of modularity for an elliptic curve in a non-standard way. The following formulation has the advantage of bringing directly to mind the quadratic reciprocity law of Gauss (the function Φ below is analogous to the Legendre symbol). This point of view was introduced by Y. Manin (and maybe B. Birch) more than twenty years ago [45], [46], [52].

Let E be an elliptic curve over $\mathbf{Q}$. Let l be a prime number. Let $|E(\mathbf{F}_l)|$ be the number of points of the reduction modulo l of a minimal Weierstrass model of E. Let $a_l(E) = l + 1 - |E(\mathbf{F}_l)|$.

Let Λ_N be the set of functions

$$\Phi : (\mathbf{Z}/N\mathbf{Z})^2 \longrightarrow \mathbf{Z}$$

satisfying for any $(u, v) \in (\mathbf{Z}/N\mathbf{Z})^2$ the equations (the *Manin relations*)

$$\Phi(u, v) + \Phi(-v, u) = 0,$$

$$\Phi(u, v) + \Phi(-u - v, u) + \Phi(v, -u - v) = 0.$$

Let Λ_E be the set of elements Φ of Λ_N satisfying for every prime number l an equation which can take several forms, the most simple that I know being the $a_l(E)$*-modular relations* (we abuse the notations here since the relations refers to *two* numbers: a_l and l):

$$a_l(E)\Phi(u, v) = \sum_{a>b\geq 0, d>c\geq 0, ad-bc=l} \Phi(au + cv, bu + dv),$$

where $(a, b, c, d) \in \mathbf{Z}^4$.

The curve E is said to be *modular of level* N if there is a non-zero element in Λ_E. Any element of Λ_E must be homogeneous, *i.e.* it satisfies the equality $\Phi(\lambda u, \lambda v) = \Phi(u, v)$ $(\lambda \in (\mathbf{Z}/N\mathbf{Z})^*)$.

Remark: The existence of a non-zero element of Λ_E is equivalent to the existence of a normalized eigenform of weight 2 for $\Gamma_1(N')$ $(N'|N)$ whose l-th Fourier coefficient is $a_l(E)$, hence the connection with the standard notion of modularity (see [55]). This can be proved by remarking that there is a bijection between the set of elements of Λ_N which vanish on non-primitive elements of $(\mathbf{Z}/N\mathbf{Z})^2$ and $\mathrm{H}_1(Y_1(N)(\mathbf{C}), \mathbf{Z})$. The functions of Λ_N satisfying the $a_l(E)$-modular relations correspond bijectively to the elements of $\mathrm{H}_1(Y_1(N)(\mathbf{C}), \mathbf{Z})$ which are eigenvalue of the Hecke operator T_l with the eigenvalue $a_l(E)$ see [55].

The concept of modularity is captured in the last family of equations. It is the expression of a law ruling all the reductions modulo l of the elliptic curve

E. Remark that the integers a, b, c, d involved in the sum depend on l but not on E or Φ.

The curve E is said to be *modular* if it is modular of some level. If E is a modular elliptic curve, then a theorem of H. Carayol ([5]) implies that it is modular of level the conductor of E. The set of functions Φ which establish the modularity at that level is a free $\mathbf{Z}$-module of rank 2. The recent work of Wiles [77] (completed by a joint work with Taylor [75] and subsequent work of Diamond and Kramer [17]) leads amongst other things to the following result, as explained in [17] (K. Rubin and A. Silverberg also remarked in [65] that the next theorem can be deduced from another work of Diamond [16]).

Theorem 1.1 (Wiles, Taylor-Wiles, Diamond-Kramer) *The elliptic curve $E_{a,b,c}$ is modular.*

1.3 Serre's conjectures

Let p be a prime number. Consider an irreducible Galois representation

$$\rho : \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \mathbf{GL}_2(\bar{\mathbf{F}}_p),$$

which is odd, *i.e.* the determinant of the image of a complex conjugation is -1.

There is a notion of modularity for such a representation. For every prime number $l \neq p$ at which the representation ρ is unramified, denote by $a_l(\rho)$ (resp. $b_l(\rho)$) the trace (resp. the determinant) of the image in $\mathbf{GL}_2(\bar{\mathbf{F}}_p)$ of a Frobenius endomorphism at the prime l. Once again we give a non-standard formulation of modularity as above.

Let $\Lambda_{N,p}$ be the set of functions

$$\Phi_p : (\mathbf{Z}/N\mathbf{Z})^2 \longrightarrow \bar{\mathbf{F}}_p$$

satisfying for any $(u, v) \in (\mathbf{Z}/N\mathbf{Z})^2$ the Manin relations

$$\Phi_p(u, v) + \Phi_p(-v, u) = 0,$$

$$\Phi_p(u, v) + \Phi_p(-u - v, u) + \Phi_p(v, -u - v) = 0.$$

Let Λ_ρ be the set of elements of $\Lambda_{N,p}$ satisfying for every prime number $l \not| pN$ the *$a_l(\rho)$-modular relations*:

$$a_l(\rho)\Phi_p(u, v) = \sum_{a>b\geq 0, d>c\geq 0, ad-bc=l} \Phi_p(au + cv, bu + dv),$$

where a, b, c and d are integers, and

$$b_l(\rho)\Phi_p(u, v) = l\Phi_p(lu, lv).$$

We say that the representation ρ is *modular of level N*, if there exists a non-zero function in Λ_ρ.

The representation ρ is said to be *modular* if it is modular of some level.

The fact that ρ is odd implies (by class field theory) that Φ_p is even (*i.e.* that $\Phi_p(-u,-v) = \Phi(u,v)$).

Remark: This definition of modularity can be shown to be equivalent to the one used originally by Serre. One can first reformulate Serre's conjecture using modular forms of weight 2 only by using results of Ash and Stevens [2]. Making use of the interpretation of functions satisfying the Manin relations in terms of the homology of modular curves as we did in the previous section (*i.e.* one shows that the existence of an element of Λ_ρ is equivalent to the existence of a non-zero modular form f of weight 2 for $\Gamma_1(N)$ in characteristic p which satisfies $T_l f = a_l(\rho) f$ (l prime $\not| Np$, T_l is the l-th Hecke operator). Take note that we allow the prime p to divide N here.

Serre conjectures that any odd, irreducible representation is modular. Moreover he predicts the minimal level of modularity of such a representation [67].

Let E be an elliptic curve over $\mathbf{Q}$. The set $E[p]$ of $\bar{\mathbf{Q}}$-rational p-division points of E is a $\mathbf{F}_p$-vector space of dimension 2. Let us choose, once and for all, an identification between this set and $\mathbf{F}_p^2$. This defines a Galois representation $\rho_{E,p}$ of the type considered above. The coefficients $a_l(\rho_{E,p})$ and $b_l(\rho_{E,p})$ are the reductions modulo p of $a_l(E)$ and l respectively.

If the elliptic curve E is modular of level N then the representation $\rho_{E,p}$ is modular of level N, as one can see by reducing a function Φ attached to E modulo p. By choosing an appropriate multiple of Φ one can insure that the reduction is non-zero. Therefore Φ_p is homogeneous. But it might occur that $\rho_{E,p}$ is modular of level strictly smaller than N. This is for instance the case if the reduction modulo p of Φ is non-zero and factorizes through $(\mathbf{Z}/N'\mathbf{Z})^2$, where N' is strictly smaller than N.

1.4 The theorems of Ribet and Mazur

Suppose that $p > 2$. For our purpose we need only the following recipe predicted by Serre and proved by Ribet for the level of modularity of the representation $\rho_{a,b,c,[p]} = \rho_{E_{a,b,c},[p]}$.

For any integer $n > 0$, denote by n_p the largest p-th power dividing n and let $\mathrm{Rad}_p(n) = \mathrm{Rad}(n/n_p)$. Let $N_{a,b,c}(p) = \mathrm{Rad}_p(abc)\epsilon_2(b)$.

Theorem 1.2 (Ribet) *Suppose that $p > 2$ and that the representation $\rho_{a,b,c,[p]}$ is modular and absolutely irreducible. Then $\rho_{a,b,c,[p]}$ is modular of level $N_{a,b,c}(p)$.*

At this point one needs the following theorem of B. Mazur. We do not insist on it for the moment since results of this type will be the subject of the second part of the paper.

Theorem 1.3 (Mazur) *Let p be a prime number. Let E be an elliptic curve over $\mathbf{Q}$ having all its 2-division points defined over $\mathbf{Q}$. Then $\rho_{E,p}$ is absolutely irreducible when $p > 3$.*

This theorem implies that we can not have $a_l(\rho_{a,b,c,[p]}) = 1 + l$ for almost all prime l when $p > 3$. (If $a_l(\rho_{a,b,c,[p]}) = 1 + l$ for almost all primes l, then 1 is an eigenvalue of almost every image of a Frobenius endomorphism. By Chebotarev's density theorem, the number 1 is an eigenvalue of every element in the image of $\rho_{a,b,c,[p]}$. This implies that the representation $\rho_{a,b,c,[p]}$ is reducible.)

1.5 Dénes' conjecture

We apply now the machinery described above to Dénes' equation. We merely repeat in essence here what is contained in [13], [64], and [67].

In view of the results mentioned in the background, we need only to prove Dénes' conjecture for prime exponents > 31. Let us consider a solution (x, y, z) to the equation $x^p + y^p = 2z^p$ (p prime number > 31, $xyz \neq 0$). We may suppose that x, y, and z are coprime and that x is congruent to -1 modulo 4. We want to show that we have $x = y = z = 1$.

Proposition 1.4 *Suppose that p is a prime > 3. The Galois representation $\rho_{x^p,-2z^p,y^p,[p]}$ is not surjective.*

Let us consider the Frey curve $E_{x^p,-2z^p,y^p}$.

If z is even, then $32|2z^p$ and the Galois representation associated to p-division points of the Frey curve is modular of level 2. The space of homogeneous functions $(\mathbf{Z}/32\mathbf{Z})^2 \longrightarrow \mathbf{F}_p$ satisfying the Manin relations is of small dimension. All its elements satisfy $a_l(\rho)$ modular equations with $a_l(\rho) = l+1$ (l prime $\not| 2p$). This is impossible by Mazur's theorem.

If z is odd, the Galois representation associated to points of p-division is modular of level 32 by Ribet's theorem.

One can check by elementary but tedious computations of linear algebra that the $\mathbf{F}_p$-vector space of homogeneous functions of level 32 satisfying the Manin relations and the $a_l(\rho)$ modular relation with $a_l(\rho) \neq l + 1$ (for at least a prime $l \not| 2p$) is of dimension 2. Using the fact that $\rho_{-1,2,-1,[p]}$ is modular of level 32, one obtains that the function Φ_p attached to $\rho_{x^p,-2z^p,y^p,[p]}$ is proportional to one attached to $\rho_{-1,2,-1,[p]}$ and therefore that $a_l(\rho_{-1,2,-1,[p]}) = a_l(\rho_{x^p,-2z^p,y^p,[p]})$. This implies by Chebotarev's density theorem and Schur's lemma that the representations $\rho_{x^p,-2z^p,y^p,[p]}$ and $\rho_{-1,2,-1,[p]}$ are isomorphic.

Note that the Frey curve $E_{-1,2,-1}$ corresponds to the trivial solution $x = y = z = 1$ of Dénes' equation. It is an elliptic curve with complex multiplication by the ring $\mathbf{Z}[i]$ (*i.e.* its ring of endomorphism is isomorphic to $\mathbf{Z}[i]$).

The theory of complex multiplication informs us that the image of $\rho_{-1,2,-1,[p]}$ (and therefore of $\rho_{x^p,-2z^p,y^p,[p]}$) is strictly smaller than $\mathbf{GL}_2(\mathbf{F}_p)$. Specifically it is contained in the normalizer of a split (resp. non-split) Cartan subgroup of $\mathbf{GL}_2(\mathbf{F}_p)$ if p is congruent to 1 (resp. -1) modulo 4 (see below).

We now turn to the second part of our paper to explain why the image of a Galois representation of the type $\rho_{a,b,c,[p]}$ is surjective when $(a,b,c) \neq (-1,2,-1)$ and $p > 3$.

2 Rational points on modular curves

2.1 Serre's problem

In [66], Serre proved the following theorem.

Theorem 2.1 (Serre) *Let E be an elliptic curve over $\mathbf{Q}$ which does not have complex multiplication. There exists a number B_E such that for every prime number $p > B_E$ the map*

$$\rho_{E,p} : \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \mathbf{GL}(E[p]) \simeq \mathbf{GL}_2(\mathbf{F}_p)$$

is surjective.

The interested reader might consult [68] to see this theorem put in a larger theoretical context. Serre proposed the following problem:

*Can the number B_E be chosen uniformly (*i.e. *independently of E)?*

This problem is still unsolved. The smallest possible candidate for a uniform B_E, up to the current knowledge, is 37 [50].

There are formulas for B_E in terms of various invariants of E [47]. One might try to find an expression for B_E in terms of the conductor of E. Building on [70], A. Kraus proved the following result [40]:

Theorem 2.2 (Kraus) *Let E be a modular elliptic curve without complex multiplication. Then the representation $\rho_{E,p}$ is surjective when one has*

$$p > 68N_E(1 + \log\log N_E)^{1/2},$$

where N_E is the product of the prime numbers dividing the conductor of E.

Note that one can get stronger bounds assuming the Generalized Riemann Hypothesis [70]. In the case of a Frey curve $E_{a,b,c}$, $N_{E_{a,b,c}}$ is equal to $\epsilon_2(b)\mathrm{Rad}(abc)$. Kraus' theorem gives directly an estimate (left to the reader) for the size of hypothetic solutions to Dénes' equation.

To determine the image of $\rho_{E,p}$, we have to take into account that its determinant is $\mathbf{F}_p^*$. To show that $\rho_{E,p}$ is surjective it is enough to show that its image is not contained in any of the following proper maximal subgroups of $\mathbf{GL}_2(\mathbf{F}_p)$ [66]:

- A Borel subgroup, *i.e.* up to conjugacy the group of upper triangular matrices.

- The normalizer of a split Cartan subgroup of $\mathbf{GL}_2(\mathbf{F}_p)$, *i.e.* up to conjugacy the set of diagonal or antidiagonal matrices.

- The normalizer of a non-split Cartan subgroup of $\mathbf{GL}_2(\mathbf{F}_p)$, *i.e.* up to conjugacy the normalizer of the image of an embedding of $\mathbf{F}_{p^2}^*$ into $\mathbf{GL}_2(\mathbf{F}_p)$ (see below).

- An exceptional subgroup, *i.e.* a subgroup of $\mathbf{GL}_2(\mathbf{F}_p)$ whose image in $\mathbf{PGL}_2(\mathbf{F}_p)$ is isomorphic to the symmetric group S_4 or to one of the alternate groups A_4 or A_5. The two latter cases can not occur because of the surjectivity of the determinant. The case of S_4 can occur only if p is congruent to ± 3 modulo 8.

Serre showed that the image of $\rho_{E,p}$ can not be contained in an exceptional subgroup when $p > 13$ or $p = 7$ [66].

2.2 Mazur's theorems

Serre's problem can be translated in terms of rational points on modular curves, *i.e.* in diophantine terms.

Let N be an integer > 0. Let $Y(N)$ be the curve which is the moduli space of pairs $(E/S/\mathbf{Q}, \phi)$, where S is a $\mathbf{Q}$-scheme, E/S is an elliptic curve and ϕ is an isomorphism of group scheme between $(\mathbf{Z}/N\mathbf{Z})^2$ and the group $E[N]$ of N-division points of E. It is a curve defined over $\mathbf{Q}$ endowed with an action of $\mathbf{GL}_2(\mathbf{Z}/N\mathbf{Z})$ defined over $\mathbf{Q}$. Let K be a subgroup of $\mathbf{GL}_2(\mathbf{Z}/N\mathbf{Z})$. Let $Y_K = Y(N)/K$. This curve classifies elliptic curves over $\mathbf{Q}$ up to $\bar{\mathbf{Q}}$-isomorphism such that the image of the map

$$\rho_{E,N} : \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \mathbf{GL}(E[N]) \simeq \mathbf{GL}_2(\mathbf{Z}/N\mathbf{Z})$$

is contained in K.

A positive answer to Serre's problem is equivalent to show that none of these curves has a $\mathbf{Q}$-rational point which does not correspond to an elliptic curve with complex multiplication when $N = p$ is a large enough prime number and K is a proper subgroup of $\mathbf{GL}_2(\mathbf{Z}/p\mathbf{Z})$.

We describe briefly Mazur's approach to Serre's problem. Let X_K be the complete modular curve attached to the subgroup K of $\mathbf{GL}_2(\mathbf{Z}/N\mathbf{Z})$. Let J_K be the jacobian variety of X_K.

For any cusp Q of X_K, denote by $k_Q \subset \mathbf{Q}(\mu_N)$ its field of definition and by j_Q the morphism $X_K \longrightarrow J_K$ which sends P to the class of the divisor $(P) - (Q)$. Let $\mathcal{O} = \mathbf{Z}[\mu_N, \frac{1}{N}]$. Let l be a prime number which does not divides N. Let λ be a prime ideal of $\mathcal{O} \subset \mathbf{Z}[\mu_N]$ above l. Let $\mathcal{O}_\lambda$ be the completion of $\mathcal{O}$ at λ.

There is a canonical smooth model $\mathcal{X}_K$ of X_K over $\mathcal{O}$ [14]. Let A be an optimal quotient of J_K, *i.e.* an abelian variety defined over $\mathbf{Q}$ such that there exists a surjective homomorphism of abelian varieties over $\mathbf{Q}$ $J_K \longrightarrow A$ with connected kernel. Let $j_{Q,A}$ be the morphism $X_K \longrightarrow A$ obtained by composing the morphism $J_K \longrightarrow A$ with j_Q. By universal property of Néron model $j_{Q,A}$ extends to a morphism $\mathcal{X}_K^{\mathrm{smooth}} \longrightarrow \mathcal{A}$ defined over $\mathcal{O}$.

Proposition 2.3 *Let l be a prime number $\not| 2N$. Let λ be a prime of $\mathcal{O}$ dividing l. Suppose that there exists an abelian variety A defined over $\mathbf{Q}$ such that*

1) A is a non-trivial optimal quotient of J_K with finitely many $\mathbf{Q}$-rational points.

2) The morphism $j_{Q,A}$ is a formal immersion at Q in characteristic l for each cusp Q of X_K.

Then there is no elliptic curve E over $\mathbf{Q}$ such that l divides the numerator of $j(E)$ and such that the image of $\rho_{E,N}$ is contained in K.

Proof: Suppose that the conclusion of the theorem is false. Then the modular curve X_K has a $\mathbf{Q}$-rational point P which is not a cusp. Let us prove that $j_{Q,A}(P)$ is a torsion point of A for any cusp Q. Let D be a $\mathbf{Q}$-rational divisor of degree $n > 0$ with support on the cusps of X_K (There exists some). By the Drinfeld-Manin theorem [19] the class x of the divisor $D - n(Q)$ of degree 0 is torsion in J_K. The point $nj_Q(P) - x$ of J_K is $\mathbf{Q}$-rational. Its image in A is of finite order since $A(\mathbf{Q})$ is finite. Therefore $nj_{Q,A}(P)$ is of finite order. Therefore $j_{Q,A}(P)$ is a torsion point of A for any cusp Q of X_K.

Since l divides the denominator of $j(E)$, the sections $\mathrm{Spec}\mathcal{O} \longrightarrow \mathcal{X}_K$ defined by P and some cusp Q coincide at λ, for some prime divisor λ of l. This implies that $j_{Q,A}(P)$ crosses $j_{Q,A} = 0$ in the special fiber at λ of $\mathcal{A}$. Since $l > 2$ and since l is unramified in k_Q (because $l \not| N$) and since J_K (and therefore A) has good reduction at l, a standard specialization lemma can be applied to show that $j_{Q,A}(P) = 0$ (see [50]).

Therefore we have $j_{Q,A}(P) = j_{Q,A}(Q) = 0$ and the sections defined by P and Q coincide in characteristic l. This contradicts the fact that $j_{Q,A}$ is a formal immersion in characteristic l at the cusp Q.

Remarks: 1) Mazur's method can, to a certain extent, be generalized to algebraic number fields of higher degree ([1], [34], [54]).

2) Mazur showed how the formal immersion property can often be checked using the theory of Hecke operators and q-expansions of modular forms [50].

3) Central to the method is the existence of a non-trivial quotient abelian variety of the jacobian with finitely many rational points. In [49], Mazur introduced the Eisenstein quotient of $J_0(p)$ and proved the finiteness of the group of rational points of this quotient by a method which made use of the semi-stability of $J_0(p)$. Mazur's construction and proof have not been extended to non semi-stable jacobians of modular curves (see [27] for an attempt). The Eisenstein quotient was used in [50] as the auxiliary abelian variety A to prove the following result in the direction of Serre's problem (from which one deduces the theorem 1.3).

Theorem 2.4 (Mazur) *Let p be a prime number. Let E be an elliptic curve over $\mathbf{Q}$ such that the image of $\rho_{E,p}$ is contained in a Borel subgroup of $\mathbf{GL}_2(\mathbf{F}_p)$. Then one has $p \leq 163$. If one supposes that all the points of 2-division of E are rational, then one has $p \leq 3$.*

By the same method, and still using the Eisenstein quotient, F. Momose completed an earlier result of Mazur to obtain the following theorem [56].

Theorem 2.5 (Momose) *Let p be an odd prime number. Let E be an elliptic curve over $\mathbf{Q}$ such that $j(E) \notin \mathbf{Z}[\frac{1}{2p}]$ and such that the image of $\rho_{E,p}$ is contained in the normalizer of a split Cartan subgroup of $\mathbf{GL}_2(\mathbf{F}_p)$. Then one has $p = 13$ or $p \leq 7$.*

2.3 Winding quotients

The conjecture of Birch and Swinnerton-Dyer provides a criterion to decide whether an abelian variety over $\mathbf{Q}$ has finitely many rational points: one has to check whether the L-function of the abelian variety vanishes at the point 1. By a *winding quotient*, we mean a maximal quotient abelian variety of the jacobian of a modular curve (or alternately of the new part of the jacobian) whose L-function does not vanish at the point 1. This terminology has its origin in [48] and [49].

Let N be an integer > 0. Let $J_1(N)$ be the jacobian of the modular curve $X_1(N)$. Let $\mathbf{T}$ be the subring of $\mathrm{End}(J_1(N))$ generated by the Hecke operator T_l ($l \not| N$). It operates on the new-quotient $J_1^{\mathrm{new}}(N)$ of $J_1(N)$. Let I be an ideal of $\mathbf{T}$. Let J_I be the quotient abelian variety $J_1^{\mathrm{new}}(N)/IJ_1^{\mathrm{new}}(N)$. A classical theorem of Eichler, Shimura, Igusa and Carayol asserts, after an easy reformulation, that

$$L(J_I, 1) = \prod_{f, If=0} L(f, 1),$$

where f runs through the newforms (*i.e.* the normalized eigenforms of weight 2 which are new for $\Gamma_1(N)$), and where

$$L(f, 1) = 2\pi \int_0^\infty f(iy)\, dy.$$

In the direction of the conjecture of Birch and Swinnerton-Dyer, K. Kato has announced that, amongst other things, he has obtained the following theorem by an original method [36], [37], [38].

Theorem 2.6 (Kato) *Suppose that $L(J_I, 1) \neq 0$, then the group of rational points of J_I is finite.*

This theorem should be sufficient to decide whether the jacobian of any modular curve has a non-trivial quotient with finitely many rational points.

Let us mention the earlier work of V. Kolyvagin and D. Logachev [39] which treated the case of quotients of jacobians of the form $J_0(N)$. They built on earlier work by Kolyvagin, on a formula of B. Gross and D. Zagier [28], and on a result established independently by K. Murty and R. Murty [57] and by D. Bump, S. Friedberg, and J. Hoffstein [4]. This last case is sufficient for the practical applications which have been considered up to now. Indeed all three families of modular curves considered to solve Serre's problem are related to jacobian varieties of the type $J_0(N)$ (in the Borel and normalizer of a split Cartan cases this is evident, in the normalizer of a non-split Cartan case it is a theorem of Chen explained below). We turn to that case in more detail.

Let N be an integer > 0. Let d be a square-free integer dividing N (in particular one might have $d = 1$). Denote by w_d the Fricke involution $X_0(N)$. Denote by $J_0(N)_d$ the jacobian of the quotient curve $X_0(N)/w_d$ Let us describe a winding quotient of the jacobian $J_0(N)_d$ of the curve $X_0(N)$. (A finer version is described in P. Parent's forthcoming thesis.) Let $\mathbf{T}$ be the subring of $\mathrm{End}(J_0(N)_d)$ generated by the Hecke operators T_n ($n \geq 1$, $(N, n) = 1$). It operates on the singular homology group $\mathrm{H}_1(X_0(N)/w_d(\mathbf{C}), \mathbf{Q})$. Let e be the unique element of $\mathrm{H}_1(X_0(N)/w_d(\mathbf{C}), \mathbf{R})$ such that the integral of any holomorphic differential form ω on $X_0(N)/w_d(\mathbf{C})$ along a cycle of class e coincides with the integral along the path from 0 to $i\infty$ in the upper half-plane of the pullback of ω. Mazur has called e the *winding element* (see [48] for the origin of this terminology). Let I_e be the annihilator of e in $\mathbf{T}$. Let $J_0^{\mathrm{old}}(N)_d$ be the old abelian subvariety of $J_0(N)_d$. Let $J_e^{\mathrm{new}}(N)_d$ be the quotient abelian variety $J_0(N)_d/(I_e J_0(N)_d + J_0^{\mathrm{old}}(N)_d)$. The following proposition can be proved by mimicking what is in [55] (the case N prime and $d = 1$) and in [13] (the case $N = 2p^2$ and $d = p$).

Proposition 2.7 *The L-function of $J_e^{\mathrm{new}}(N)_d$ does not vanish at the point* 1.

By application of the theorem of Kolyvagin and Logachev one obtains.

Theorem 2.8 (Kolyvagin-Logachev) *The group of $\mathbf{Q}$-rational points of the abelian variety $J_e^{\mathrm{new}}(N)_d$ is finite.*

Therefore one obtains the following criterion which is *a priori* curious since it involves only the complex structure of the modular curve.

Corollary 2.9 *If e does not belong to the old part of $\mathrm{H}_1(X_0(N)_d(\mathbf{C}), \mathbf{Q})$, then $J_0^{\mathrm{new}}(N)_d$ has a non trivial quotient with finitely many rational points.*

To check such a criterion one might have to make calculations on modular symbols using Manin's presentation of the homology of $X_0(N)$ ([45], [54], [61]). The verification amounts then to a study of a graph with vertices indexed by $\mathbf{P}^1(\mathbf{Z}/N\mathbf{Z})$ (see [54], [61]) which is essentially the "dessin d'enfant" associated to the curve $X_0(N)$ in the sense of [29]. However in [13], this was done by a simpler argument that we do not describe here.

2.4 Chen's isogeny

In [13] a result of I. Chen was used to obtain our main result [6].

Theorem 2.10 (Chen) *Let K_{ns} (resp. K_{s}) be the normalizer of a non-split (resp. split) Cartan subgroup of $\mathbf{GL}_2(\mathbf{F}_p)$. Let X_{ns} and X_{s} be the corresponding modular curves. Let J_{ns} and J_{s} be the jacobians of X_{ns} and X_{s}. There is an isogeny of abelian varieties defined over $\mathbf{Q}$ between J_{s} and $J_{\mathrm{ns}} \times J_0(p)$.*

Chen uses trace formulas to prove that the two abelian varieties have the same L-function. Making use of a theorem of Faltings he concludes that the abelian varieties are isogenous. The problem was reconsidered by B. Edixhoven in [20], who gave a more elementary proof of Chen's theorem based on the theory of representation of $\mathbf{GL}_2(\mathbf{F}_p)$. Neither of these two proofs gave an explicit description of the isogeny, nor is the short proof that we give below.

Since J_{s} is isomorphic to the jacobian of $X_0(p^2)/w_p$ (where w_p is the Fricke involution), an elementary argument of sign of functional equation of L-function leads to the following consequence.

Corollary 2.11 *The L-function of any non-trivial quotient abelian variety of J_{ns} vanishes at* 1*.*

If one believes in the conjecture of Birch and Swinnerton-Dyer, no non-trivial quotient abelian variety of J_{ns} has finitely many rational points. And we are embarrassed to apply Mazur's method to X_{ns}.

One is tempted now to consider other results in the direction of the conjecture of Birch and Swinnerton-Dyer, such as the formula of Gross and Zagier [28] and the result of Kolyvagin on elliptic curves of analytic rank one to understand the arithmetic of J_{ns}.

We expect that Chen's isogeny can be deduced from from an explicit correspondence between X_{s} and X_{ns}.

Let $H_p = (\mathbf{P}^1(\mathbf{F}_{p^2}) - \mathbf{P}^1(\mathbf{F}_p))/\mathrm{Gal}(\mathbf{F}_{p^2}/\mathbf{F}_p)$ ("The upper half-plane over $\mathbf{F}_p$"). It is a finite set of cardinality $p(p-1)/2$ endowed with a transitive action of $\mathbf{GL}_2(\mathbf{F}_p)$ deduced from the homographies on $\mathbf{P}^1(\mathbf{F}_{p^2})$. Let $\lambda = \{t, -t\} \in H_p$ such that $t^2 \in \mathbf{F}_p$. Let K_λ be the stabilizer of λ in $\mathbf{GL}_2(\mathbf{F}_p)$. It is the normalizer of a non-split Cartan subgroup of $\mathbf{GL}_2(\mathbf{F}_p)$.

Let $c_\lambda = \mathbf{F}_p^* \lambda \subset H_p$. To any $g \in \mathbf{GL}_2(\mathbf{F}_p)$ we associate a subset gc_λ of H_p (the support of a "geodesic path" of H_p) and a pair of elements of $\mathbf{P}^1(\mathbf{F}_p)$ the set $\{g0, g\infty\}$ (two "cusps").

For every pair $\{P_1, P_2\}$ of distinct elements of $\mathbf{P}^1(\mathbf{F}_p)$ there is an element $g \in \mathbf{GL}_2(\mathbf{F}_p)$ such that $\{g0, g\infty\} = \{P_1, P_2\}$. This element is well defined up to multiplication by a diagonal or antidiagonal matrix. This implies that the set gc_λ depends only on $\{P_1, P_2\}$. Let us denote it by $< P_1, P_2 >$. Let C_p be the set of pairs of distinct elements of $\mathbf{P}^1(\mathbf{F}_p)$.

The group of diagonal or antidiagonal matrices is the normalizer K_{s} of the split Cartan subgroup of $\mathbf{GL}_2(\mathbf{F}_p)$ formed by diagonal matrices. The map $\mathbf{GL}_2(\mathbf{F}_p) \longrightarrow C_p$ which to g associates $\{g0, g\infty\}$ defines a bijection between $\mathbf{GL}_2(\mathbf{F}_p)/K_{\mathrm{s}}$ and C_p.

Let

$$f_p^0 : \mathbf{Z}[C_p] \longrightarrow \mathbf{Z}[H_p]$$

be the group homomorphism which associates to $[\{P_1, P_2\}]$ the sum of elements of $< P_1, P_2 >$ (the map which to the support of a "geodesic path" associates the sum of its points). Let

$$g_p^0 : \mathbf{Z}[\mathbf{P}^1(\mathbf{F}_p)] \longrightarrow \mathbf{Z}[C_p]$$

be the group homomorphism which to $[P]$ associates the sum of elements of C_p containing P.

The group homomorphism

$$\mathbf{Z}[\mathbf{GL}_2(\mathbf{F}_p)/K_{\mathrm{s}}] \longrightarrow \mathbf{Z}[\mathbf{GL}_2(\mathbf{F}_p)/K_{\mathrm{ns}}]$$

which to gK_{s} associates $\sum_{h \in \mathbf{GL}_2(\mathbf{F}_p)/(K_{\mathrm{s}} \cap K_{\mathrm{ns}})} ghK_{\mathrm{ns}}$ coincides with f_p^0 if one identifies H_p with $\mathbf{GL}_2(\mathbf{F}_p)/K_{\mathrm{ns}}$ and C_p with $\mathbf{GL}_2(\mathbf{F}_p)/K_{\mathrm{s}}$ as above.

Therefore one deduces from f_p^0 a correspondence f_p: $X_{\mathrm{s}} \longrightarrow X_{\mathrm{ns}}$ given by the canonical morphism $X_{K_{\mathrm{ns}} \cap K_{\mathrm{s}}} \longrightarrow X_{\mathrm{ns}} \times X_{\mathrm{s}}$ and of degree $\frac{p-1}{2}$.

Let B_0 be the Borel subgroup of $\mathbf{GL}_2(\mathbf{F}_p)$ consisting of upper triangular matrices. The map $\mathbf{GL}_2(\mathbf{F}_p) \longrightarrow \mathbf{P}^1(\mathbf{F}_p)$ which to to g associates $g\infty$ defines a bijection between $\mathbf{GL}_2(\mathbf{F}_p)/B_0$ and $\mathbf{P}^1(\mathbf{F}_p)$.

The group homomorphism

$$\mathbf{Z}[\mathbf{GL}_2(\mathbf{F}_p)/B_0] \longrightarrow \mathbf{Z}[\mathbf{GL}_2(\mathbf{F}_p)/K_{\mathrm{s}}]$$

which to gB_0 associates $\sum_{h \in \mathbf{GL}_2(\mathbf{F}_p)/(K_{\mathrm{s}} \cap B_0)} ghK_{\mathrm{s}}$ coincides with g_p^0 if one identifies $\mathbf{P}^1(\mathbf{F}_p)$ with $\mathbf{GL}_2(\mathbf{F}_p)/B_0$ and C_p with $\mathbf{GL}_2(\mathbf{F}_p)/K_{\mathrm{s}}$ as above.

Denote, as usual, by $X_0(p)$ the modular curve X_{B_0}. One deduces from g_p^0 a correspondence g_p: $X_0(p) \longrightarrow X_{\mathrm{s}}$ of degree p.

Let $\theta = \sum_{x \in H_p} [x] \in \mathbf{Z}[H_p]$. Let $P \in \mathbf{P}^1(\mathbf{F}_p)$. A straightforward calculations shows that $f_p^0 \circ g_p^0(P) = \theta$.

Problem 2.12 *What is the image of f_p^0?*

To our embarassment we could not give an answer to this question. It is likely that this image generates $\mathbf{Q}[H_p]$ as a vector space This was verified by Chen in a few non-trivial cases.

If this is true, the inverse image by f_p^0 of $\mathbf{Z}\theta$ is generated by the image of g_p^0. In particular the kernel of f_p^0 is generated by the image by g_p^0 of elements of degree 0.

Let us say that a divisor of X_{s} is *p-old* if it is the image by the correspondence $X_0(p) \longrightarrow X_{\mathrm{s}}$ and that a divisor of X_{ns} is *p-old* if it is the inverse image of a divisor of $X(1)$ by the morphism $X_{\mathrm{ns}} \longrightarrow X(1)$. Since $X(1)$ is a curve of genus 0, the class any p-old divisor of degree 0 is 0 in $J_{\mathrm{ns}}(p)$. The properties of f_p^0 translate immediately into properties of the correspondence f_p. The image by f_p of a p-old divisor of X_{s} is a p-old divisor of X_{ns}.

The correspondence f_p defines by Albanese and Picard functoriality two homomorphisms of abelian varieties f_{p*}: $J_{\mathrm{s}} \longrightarrow J_{\mathrm{ns}}$ and f_p^*: $J_{\mathrm{ns}} \longrightarrow J_{\mathrm{s}}$. The correspondence g_p defines similarly an homomorphism of abelian variety g_{p*}: $J_0(p) \longrightarrow J_{\mathrm{s}}$. Since one has $f_{p*} \circ g_{p*}(J_0(p)) = 0$, one can expect that f_{p*} is the explicit description of Chen's isogeny. To prove this it would suffice to show that the cokernel of f_p^0 is finite.

Remark: Using the techniques of [43], the determination of the image of f_p^0 should lead to the explicit description of the kernel of f_p^*. It would not be surprising if the image of f_p^0 contained $\frac{p-1}{2}\mathbf{Z}[H_p]$.

Let us give a quick proof, essentially suggested by Edixhoven, of Chen's theorem. Let E be a set with one element considered as a trivial $\mathbf{GL}_2(\mathbf{F}_p)$-set.

Lemma 2.13 *The sets* $\mathbf{P}^1(\mathbf{F}_p) \cup H_p$ *and* $C_p \cup E$ *are weakly isomorphic as* $\mathbf{GL}_2(\mathbf{F}_p)$*-sets,* i.e. *any element of* $\mathbf{GL}_2(\mathbf{F}_p)$ *has the same number of fixed points in each set.*

Proof: Let $g \in \mathbf{GL}_2(\mathbf{F}_p)$. The number of fixed points of g depends only on the conjugacy class of g. There are four types of conjugacy classes. In each case we indicate the number of fixed points of g in $\mathbf{P}^1(\mathbf{F}_p)$, H_p, C_p and E respectively. First case: g is a scalar matrix; The numbers of fixed points are $p+1$, $\frac{p(p-1)}{2}$, $\frac{p(p+1)}{2}$, and 1. Second case: g has two distinct eigenvalues in $\mathbf{F}_p$; We have then 2, 0, 1, and 1. Third case: g is not a scalar matrix and has one double root in its characteristic polynomial; The numbers of fixed points are 1, 0, 0, and 1. Fourth and final case: g has no eigenvalue in $\mathbf{F}_p$; One obtains 2, 1, 0, and 1.

In each case the sum of the first two numbers equals the sum of the remaining two.

As a consequence the $\mathbf{Q}[\mathbf{GL}_2(\mathbf{F}_p)]$-modules $\mathbf{Q}[\mathbf{P}^1(\mathbf{F}_p) \cup H_p]$ and $\mathbf{Q}[C_p \cup E]$ are isomorphic (see [72] exercise 5, chapter 13). Using the identifications of $\mathbf{P}^1(\mathbf{F}_p)$, H_p, C_p, and E as cosets of $\mathbf{GL}_2(\mathbf{F}_p)$ given above, such an isomorphism defines a correspondence from $X_0(1) \cup X_{\mathrm{s}}(p)$ to $X_0(p) \cup X_{\mathrm{ns}}(p)$, which defines in turn, by Albanese fonctoriality an homomorphism of abelian varieties

$$J_0(1) \times J_{\mathrm{s}}(p) \longrightarrow J_0(p) \times J_{\mathrm{ns}}(p).$$

Since $J_0(1)$ is trivial, the Chen's theorem follows.

One can deduce *mutatis mutandis* similar results by adding an extra and prime to p level structure. For the purpose of the study of Frey curves and Dénes' conjecture we needed the following variant.

Let K be the subgroup of $\mathbf{GL}_2(\mathbf{Z}/2p\mathbf{Z})$ such that the image of K in $\mathbf{GL}_2(\mathbf{F}_2)$ and $\mathbf{GL}_2(\mathbf{F}_p)$ are a Borel subgroup and the normalizer of a non-split Cartan subgroup respectively. The modular curve $X_K = X_{\mathrm{ns},2}$ associated is the product $X_{\mathrm{ns}} \times_{X(1)} X_0(2)$.

Corollary 2.14 *The p-new quotient of* $J_0(2p^2)/w_p$ *and* $J_{\mathrm{ns},2}$ *are isogenous abelian varieties.*

2.5 Serre's problem and Frey curves

We describe now a weak result in direction of Serre's problem relatively to the modular curve X_K quotient of $X(2p)$, where the image of K in $\mathbf{GL}_2(\mathbf{F}_2)$ and $\mathbf{GL}_2(\mathbf{F}_p)$ are a Borel subgroup and the normalizer of a Cartan subgroup respectively as above.

Darmon and I obtained the part concerning non-split Cartan subgroups of the following theorem. The part concerning split Cartan subgroup is a theorem of Momose except when $p \in \{5, 7, 13\}$ (see also [33]).

Theorem 2.15 *Let* E *be an elliptic curve over* $\mathbf{Q}$ *and* $p \geq 5$ *be a rational prime satisfying the following assumptions:*

1) E *has a* $\mathbf{Q}$*-rational torsion point of order* 2.

2) The image of $\rho_{E,p}$ *in* $\mathbf{GL}_2(\mathbf{F}_p)$ *is isomorphic to the normalizer of a Cartan subgroup.*

Then $j(E)$ *belongs to* $\mathbf{Z}[\frac{1}{2p}]$.

Proof: We apply Mazur's method to the modular curve X_K, where K is the subgroup of $\mathbf{GL}_2(\mathbf{Z}/2p\mathbf{Z})$ whose images in $\mathbf{GL}_2(\mathbf{F}_2)$ and $\mathbf{GL}_2(\mathbf{F}_p)$ are a Borel subgroup and the normalizer of a Cartan subgroup.

This was carried out in detail in [13] in the non-split Cartan case. A key point was to show that new part of the jacobian of the curve $X_0(2p^2)/w_p$ has a non-trivial winding quotient when $p > 3$. This was done by using the criterion of the corollary 2.9.

In the split case, the modular curve X_K we consider is isomorphic to $X_0(2p^2)/w_p$. Its jacobian has a non-trivial winding quotient A as we just mentioned. It remains to check the formal immersion criterion at all cusps in every characteristic l not equal to 2 or p. This is done as in [50], proposition 3.2 (resp. [56], proposition 2.5) at the two cusps above the cusp ∞ of $X_0(p^2)$ (resp. at the other cusps).

Remarks: 1) Our method shows easily that the theorem 2.15 still holds if the hypothesis 1) is replaced by:

E has a $\mathbf{Q}$-rational isogeny of order $r \neq p$, where $r = 3, 5, 7$ or 13 (the cases where $X_0(p)$ has genus 0).

2) A better understanding of the bad reduction of non semi-stable jacobians of modular curves might lead to a stronger version of the theorem with the conclusion $j(E) \in \mathbf{Z}[\frac{1}{2}]$.

3) When the image of $\rho_{E,p}$ is contained in the normalizer of a non-split Cartan subgroup the conclusion of the theorem might be improved to $j(E) \in \mathbf{Z}[\frac{1}{p}]$.

The theorem 2.15 has the following implication for Frey curves.

Corollary 2.16 *Let $(a, b, c) \neq (-1, 2, -1)$ be a as in section 1.1. Let p be a prime number > 3. Then the representation $\rho_{a,b,c,[p]}$ is surjective.*

Proof: Suppose that the elliptic curve $E_{a,b,c}$ is semistable (*i.e.* $16|b$), Serre proved that $\rho_{a,b,c,[p]}$ is surjective or of image contained in a Borel subgroup [69]. The latter case is impossible by Mazur's theorem since $p > 3$. Therefore we may suppose that $16 \not| b$.

We have to prove that the image of $\rho_{a,b,c,[p]}$ is not contained in any of the proper maximal subgroups of $\mathbf{GL}_2(\mathbf{F}_p)$ described at the beginning of section 2.1. Taking into account that the image of $\rho_{a,b,c,[p]}$ is not contained in Borel subgroup (Mazur's theorem), it is contained in the normalizer of a Cartan subgroup or in an exceptional subgroup. Th groups of the two latter types do not possess elements of order $p > 3$.

Since E is not semi-stable, its minimal discriminant Δ is equal to $16(abc)^2$ [59]. Let l be an odd prime number. The l-adic valuation of Δ is divisible by p, otherwise Tate's theory would produce a unipotent element in the image of $\rho_{a,b,c,[p]}$. Therefore abc is the product of a power of 2 by a p-th power; In other words, the triple (a, b, c) comes from a solution of the equation $x^p + 2^\alpha y^p + z^p = 0$. By the proof of Fermat's last theorem and its variants, one must have $\alpha = 1$ [64]. In that case the image of $\rho_{a,b,c,[p]}$ is contained in the normalizer of a Cartan subgroup. This is impossible by theorem 2.15.

Remark: Concerning the case $p = 3$, Diamond and Kramer showed that the image of $\rho_{a,b,c,[p]}$ is irreducible when $E_{a,b,c}$ is not semi-stable and $p \geq 3$ [17]. One might gain some insight from [21] to study the situation when $p = 3$.

Corollary 2.17 (Dénes' conjecture) *The equation $x^n + y^n = 2z^n$ has no solution in integers x, y, z, n with $n > 2$, $xyz \neq 0$ and $x \neq \pm y$.*

3 The degree of modular parametrizations

3.1 The pairing Δ

Let E be an elliptic curve over $\mathbf{Q}$ of conductor N. Let $\tilde{\Lambda}_0(N)$ be the set of functions

$$\Phi : \mathbf{P}^1(\mathbf{Z}/N\mathbf{Z}) \longrightarrow \mathbf{Z}$$

satisfying the Manin relations described in the section 1.2 and that we repeat now:

$$\Phi(u,v) + \Phi(-v,u) = 0$$

and

$$\Phi(u,v) + \Phi(-u-v,u) + \Phi(v,-u-v) = 0.$$

Let $\tilde{\Lambda}_E$ be the set of elements of $\tilde{\Lambda}_0(N)$ satifying the $a_l(E)$-modular relations for every prime number $l \not| N$:

$$a_l(E)\Phi(u,v) = \sum_{a>b\geq 0, d>c\geq 0, ad-bc=l} \Phi(au+cv, bu+dv),$$

where a, b, c, and d are integers.

Remark: The group $\tilde{\Lambda}_0(N)$ coincides with the group of functions $\Phi : \mathbf{P}^1(\mathbf{Z}/N\mathbf{Z}) \longrightarrow \mathbf{Z}$ satisfying the following property: For any $\sum_M u_M[M] \in \mathbf{Z}[\mathbf{SL}_2(\mathbf{Z})]$ such that $\sum_M u_M([M\infty]-[M0]) = 0$ in $\mathbf{Z}[\mathbf{P}^1(\mathbf{Q})]$, one has for any $(u,v) \in \mathbf{P}^1(\mathbf{Z}/N\mathbf{Z})$

$$\sum_M u_M \Phi((u,v)M) = 0.$$

The $a_l(E)$-modular relation can be reformulated in a similar way [55].

As we already mentioned $\tilde{\Lambda}_E$ is a free $\mathbf{Z}$-module of rank 2. We define an alternate $\mathbf{Z}$-valued pairing Δ on $\tilde{\Lambda}_0(N)$ as follows. Let Φ_1 and Φ_2 be elements of $\tilde{\Lambda}_0(N)$. We set

$$\Delta(\Phi_1,\Phi_2) = \frac{1}{2} \sum_{(u,v)\in\mathbf{P}^1(\mathbf{Z}/N\mathbf{Z})} \Phi_1(u,v)\Phi_2(v,-u-v) - \Phi_1(v,-u-v)\Phi_2(u,v).$$

We still denote by Δ the restriction of Δ to $\tilde{\Lambda}_E$. Let Δ_0 be a bilinear, alternate and surjective pairing $\tilde{\Lambda}_E \times \tilde{\Lambda}_E \longrightarrow \mathbf{Z}$. It exists and is unique up to sign. We have $\Delta = d_E\Delta_0$, with d_E integer > 0 if one makes the suitable choice of sign for Δ_0. In other words, $d_E\mathbf{Z}$ is the image of Δ.

3.2 Connection with the degree

Let E be a modular elliptic curve of conductor N. A non-constant morphism $X_0(N) \longrightarrow E$ which sends the cusp ∞ to 0 will be said to be *minimal* if it does not factorizes through an endomorphism of E which is not an automorphism. Such a morphism factorizes through the map $j_\infty : X_0(N) \longrightarrow J_0(N)$ which sends the cusp ∞ to 0. A modular elliptic curve E_0 is said to be *optimal* if the the induced morphism of abelian varieties $J_0(N) \longrightarrow E_0$ has a connected kernel. The corresponding minimal parametrization $X_0(N) \longrightarrow E_0$ is the *optimal parametrization.* It is unique up to automorphism of E_0.

Let C_N be the *cuspidal subgroup* of $J_0(N)$, *i.e.* the group generated by the classes of divisors of degree 0 with support on the cusps. It is finite by the Drinfeld-Manin theorem and is defined over $\mathbf{Q}$. Let C_{E_0} be its image by an optimal parametrization $X_0(N) \longrightarrow E_0$.

The elliptic curve E is said to be *cuspidal* if it fits in the exact sequence

$$0 \longrightarrow C_{E_0} \longrightarrow E_0 \longrightarrow E \longrightarrow 0,$$

where E_0 is an optimal elliptic curve. The morphism $X_0(N) \longrightarrow E$ obtained by composing the optimal parametrization with the isogeny $E_0 \longrightarrow E$ is said to be the *cuspidal parametrization.* It is not *a priori* clear that it is minimal.

The following proposition is similar to a formula contained in [55] and seems to be essentially the formula given by Zagier and Cremona ([78], [7]).

Proposition 3.1 *Let E be a cuspidal modular elliptic curve of conductor N. The degree of the cuspidal parametrization is equal to d_E.*

Proof: First we collect some information on the homology of modular curves. In what follows the homology groups of curves are the singular homology groups of the associated Riemann surfaces. Intersection products will be denoted by $\bullet$.

Let *cusps* be the set of cusps of the modular curve $X_0(N)$. We will consider the relative homology group $\mathrm{H}_1(X_0(N), \mathit{cusps}; \mathbf{Z})$. Integration of holomorphic differential forms defines a group homomorphism

$$R : \mathrm{H}_1(X_0(N), \mathit{cusps}; \mathbf{Z}) \longrightarrow \mathrm{H}_1(X_0(N); \mathbf{R}),$$

which to x associates the only element $R(x) \in \mathrm{H}_1(X_0(N); \mathbf{R})$ such that $\int_x \omega = \int_{R(x)} \omega$ $(\omega \in \mathrm{H}^0(X_0(N), \Omega^1))$. Let H^∂ be the image of R. Let D be a divisor of degree 0 with support on the cusps of $X_0(N)$. Let $x_D \in \mathrm{H}_1(X_0(N), \mathit{cusps}; \mathbf{Z})$ be the class of a cycle of boundary D. The image of $R(x_D)$ in $\mathrm{H}^\partial/\mathrm{H}_1(X_0(N); \mathbf{Z})$ depends only on the class of D in C_N. This defines a group isomorphism between C_N and $\mathrm{H}^\partial/\mathrm{H}_1(X_0(N); \mathbf{Z})$. The Drinfeld-Manin theorem asserts that the image H^∂ of R is contained in $\mathrm{H}_1(X_0(N); \mathbf{Q})$ (or equivalently that C_N is finite). There is a perfect duality between $\mathrm{H}_1(X_0(N), \mathit{cusps}; \mathbf{Z})$ and $\mathrm{H}_1(Y_0(N); \mathbf{Z})$

given by the intersection products. The dual of R defines a group homomorphism R^*: $\mathrm{H}_1(X_0(N);\mathbf{Z}) \longrightarrow \mathrm{H}_1(Y_0(N);\mathbf{Q})$. Let $\mathrm{H}^{\partial *}$ be the set of elements x of $\mathrm{H}_1(X_0(N);\mathbf{Z})$ such that $x \bullet y \in \mathbf{Z}$ $(y \in \mathrm{H}^{\partial})$.

Lemma 3.2 *The image of* $\mathrm{H}^{\partial *}$ *by* R^* *is a direct factor of* $\mathrm{H}_1(Y_0(N);\mathbf{Z})$.

Proof: Making use of duality, the statement of the lemma translates into the obvious statement that the image of R is H^{∂}.

Let π: $X_0(N) \longrightarrow E$ be the cuspidal parametrization.

Lemma 3.3 *The image of the map* π^**:* $\mathrm{H}_1(E,\mathbf{Z}) \longrightarrow \mathrm{H}_1(X_0(N);\mathbf{Z})$ *deduced from the cuspidal parametrization is a direct factor of* $\mathrm{H}^{\partial *}$.

Proof: Let $\mathbf{U} = \{z \in \mathbf{C}/|z| = 1\}$. Recall that there are canonical isomorphisms of complex Lie groups: $J_0(N)(\mathbf{C}) \simeq \mathrm{Hom}(\mathrm{H}_1(X_0(N);\mathbf{Z}),\mathbf{U})$, $E(\mathbf{C}) \simeq \mathrm{Hom}(\mathrm{H}_1(E,\mathbf{Z}),\mathbf{U})$, and $E_0(\mathbf{C}) \simeq \mathrm{Hom}(\mathrm{H}_1(E_0,\mathbf{Z}),\mathbf{U})$. With this identification C_N corresponds to the elements of $\mathrm{Hom}(\mathrm{H}_1(X_0(N);\mathbf{Z}),\mathbf{U})$ vanishing on $\mathrm{H}^{\partial *}$. The homorphisms of complex Lie groups $J_0(N)(\mathbf{C}) \longrightarrow E_0(\mathbf{C}) \longrightarrow E(\mathbf{C})$ coincide with those deduced from the group homomorphism $\mathrm{H}_1(E,\mathbf{Z}) \longrightarrow \mathrm{H}_1(E_0,\mathbf{Z}) \longrightarrow \mathrm{H}_1(X_0(N);\mathbf{Z})$. Let π_0: $X_0(N) \longrightarrow E_0$ be the optimal parametrization. The optimality property translates into the fact that the image of π_0^* is a direct factor of $\mathrm{H}_1(X_0(N);\mathbf{Z})$. The definition of E translates into the fact that the image of π^* is a direct factor of $\mathrm{H}^{\partial *}$.

As a consequence of the two previous lemmas, $R^* \circ \pi^*(\mathrm{H}_1(E,\mathbf{Z}))$ is a direct factor of $\mathrm{H}_1(Y_0(N);\mathbf{Z})$. It is the set of elements x of $\mathrm{H}_1(Y_0(N);\mathbf{Z})$ satisfying the equality $T_l x = a_l(E)x$, (l prime number not dividing N).

We describe now the bijection between $\mathrm{H}_1(Y_0(N);\mathbf{Z})$ and $\tilde{\Lambda}_0(N)$. Recall that the map $\Gamma_0(p)g \mapsto (0,1)g$ defines a bijection between $\Gamma_0(p)\backslash\mathbf{SL}_2(\mathbf{Z})$ and $\mathbf{P}^1(\mathbf{Z}/N\mathbf{Z})$. Let $\mathcal{E} = \mathbf{SL}_2(\mathbf{Z})i \cup \mathbf{SL}_2(\mathbf{Z})\rho$ where $i = e^{i\pi/2}$ and $\rho = e^{i\pi/3}$. Let c_0 be the geodesic path from i to ρ in the upper half-plane. For $g \in \mathbf{SL}_2(\mathbf{Z})$, consider the class in $\mathrm{H}_1(Y_0(N),\mathcal{E};\mathbf{Z})$ of the image in $Y_0(N)$ of the path gc_0. It depends only on $x = (0,1)g \in \mathbf{P}^1(\mathbf{Z}/N\mathbf{Z})$. Denote it by $[x]$. Let $\Phi \in \tilde{\Lambda}_0(N)$. In [55], we proved that $c_\Phi = \sum_{x\in\mathbf{P}^1(\mathbf{Z}/N\mathbf{Z})} \Phi(x)[x]$ belongs to $\mathrm{H}_1(Y_0(N);\mathbf{Z})$ identified with a subgroup of $\mathrm{H}_1(Y_0(N),E;\mathbf{Z})$. The map $\Phi \mapsto c_\Phi$ defines a group isomorphism i: $\tilde{\Lambda}_0(N) \longrightarrow \mathrm{H}_1(Y_0(N);\mathbf{Z})$. The action of Hecke operators translates as follows on $\tilde{\Lambda}_0(N)$ [55]:

$$i \circ T_l \circ i^{-1}(\phi)(u,v) = \sum_{a>b\geq 0, d>c\geq 0, ad-bc=l} \phi(au+cv, bu+dv),$$

where a,b, c, and d are integers and $l \not| N$ is a prime number.

The system of eigenvalue $a_l(E)$ appears exactly with multiplicity one in $\mathrm{H}_1(Y_0(N);\mathbf{Z})$ and therefore in $\tilde{\Lambda}_0(N)$. Since $\mathrm{H}_1(E,\mathbf{Z})$ is a direct factor of $\mathrm{H}_1(Y_0(N);\mathbf{Z})$, the isomorphism between $\mathrm{H}_1(Y_0(N);\mathbf{Z})$ and $\tilde{\Lambda}_0(N)$ induces an isomorphism between $\mathrm{H}_1(E,\mathbf{Z})$ and $\tilde{\Lambda}_E$.

The embedding $X_0(N) \subset Y_0(N)$ defines a surjective group homomorphism S: $\tilde{\Lambda}_0(N) \simeq \mathrm{H}_1(Y_0(N); \mathbf{Z}) \longrightarrow \mathrm{H}_1(X_0(N); \mathbf{Z})$.

Lemma 3.4 *The alternate bilinear pairing $\tilde{\Lambda}_0(N) \times \tilde{\Lambda}_0(N) \longrightarrow \mathbf{Z}$ which to (Φ_1, Φ_2) associates $S(c_{\Phi_1}) \bullet S(c_{\Phi_2})$ coincides with the pairing Δ.*

Proof: Let $(\Phi_1, \Phi_2) \in \Lambda_0(N)^2$. Let us calculate $S(c_{\Phi_1}) \bullet S(c_{\Phi_2})$. For $x = (0,1)g \in \mathbf{P}^1(\mathbf{Z}/N\mathbf{Z})$, denote by $\tilde{x}$ the image of gc_0 in $X_0(N)$. Let $(x, x') = ((u,v),(u',v')) = ((0,1)g,(0,1)g') \in \mathbf{P}^1(\mathbf{Z}/N\mathbf{Z})$. The hyperbolic triangle of the upper half-plane whose vertices are ∞, 0, and ρ is the boundary of a fundamental domain of $\mathbf{SL}_2(\mathbf{Z})$. It contains the support of the path c_0. Therefore the pathes $\tilde{x}$ and $\tilde{x}'$ do not meet if $x \neq x'$ except maybe at their extremities.

These pathes meet at their extremities if and only if $\Gamma_0(N)g = \Gamma_0(N)g'$, $\Gamma_0(N)g = \Gamma_0(N)g'\sigma$, $\Gamma_0(N)g = \Gamma_0(N)g'\tau$, or $\Gamma_0(N)g = \Gamma_0(N)g'\tau^2$, where σ and τ are generators of the stabilizers of i and ρ respectively in $\mathbf{SL}_2(\mathbf{Z})$. These cases correspond to the cases $(u,v) = (u',v')$, $(u,v) = (-v',u')$, $(u,v) = (v', -u'-v')$, or $(u,v) = (-u'-v', u')$ respectively.

Therefore to calculate our intersection product we have only to take care of the contributions at the elements of $\mathcal{E} \in X_0(N)$. An analysis of these contributions reveals that the pairing on $\mathrm{H}_1(Y_0(N); \mathbf{Z})$ which to (x,y) associates $S(x) \bullet S(y)$ is induced by the pairing

$$\mathrm{H}_1(Y_0(N), \mathcal{E}; \mathbf{Z}) \times \mathrm{H}_1(Y_0(N), \mathcal{E}; \mathbf{Z}) \longrightarrow \mathbf{Q}$$

which to the classes of $\tilde{x}'$ and $\tilde{x}$ associates 0 (resp. $1/2$, resp. $-1/2$) if $x \neq x'\tau$ and $x \neq x'\tau^2$ (resp. $x = x'\tau$, resp. $x = x'\tau^2$). This last pairing coincides with Δ when one identifies $\tilde{\Lambda}_0(N)$ with $\mathrm{H}_1(Y_0(N); \mathbf{Z})$.

We conclude the proof of the proposition 3.1. The degree $\mathrm{deg}\pi$ of π appears in the formula

$$\pi^*(x) \bullet \pi^*(y) = (\mathrm{deg}\pi)\, x \bullet y,$$

where x and y are elements of $\mathrm{H}_1(E, \mathbf{Z})$. In particular the image I of the pairing on $\mathrm{H}_1(E, \mathbf{Z})$ which to (x,y) associates $\pi^*(x) \bullet \pi^*(y)$ is equal to $\mathrm{deg}\pi\, \mathbf{Z}$.

The embedding of $\mathrm{H}_1(E, \mathbf{Z})$ as a direct factor of $\tilde{\Lambda}_0(N) \simeq \mathrm{H}_1(Y_0(N); \mathbf{Z})$ identifies it with $\tilde{\Lambda}_E$. By lemma 3.4, the pairing $(x,y) \mapsto \pi^*(x) \bullet \pi^*(y)$ coincides with Δ. Therefore I is equal to the image of Δ and d_E is the degree of the cuspidal parametrization.

Proposition 3.5 *Let E be an optimal modular elliptic curve. The order of C_E is bounded by a universal constant.*

Suppose that E is a Frey curve. Then the order of C_E divides 16 if E is not semi-stable; It divides 24 if E is semi-stable.

Proof: The group C_E is isomorphic to $\mathbf{Z}/n\mathbf{Z} \times \mathbf{Z}/m\mathbf{Z}$ with $n|m$. Therefore E has a $\mathbf{Q}$-rational cyclic subgroup of order m/n. Mazur's theorem implies that

$m/n \leq 163$, and $m/n \leq 12$ if E is a Frey curve. If E is a non-semi-stable Frey curve, it has no rational subgroup of order 3 [17], and m/n divides 4.

It remains to find a bound for n. Recall that the conductor N of E over $\mathbf{Q}$ divides $3^3.2^6$ times the square of a squarefree number [60] (in the case of a Frey curve N divides 16 times a squarefree number). Recall that the cusps of $X_0(N)$ are defined over the cyclotomic field $\mathbf{Q}(\mu_d)$, where d^2 is the largest square dividing N. In particular, any element of C_E is defined over $\mathbf{Q}(\mu_d)$ (over $\mathbf{Q}(i)$ when E is a Frey curve). Because of the Weil pairing, the field defined by the points of n-division of E contains $\mathbf{Q}(\mu_n)$. Therefore one has $n|d$ ($n|4$ in the case of a Frey curve and $n|2$ if E is semi-stable). Therefore n divides 24 times a square free number. We have only to bound its prime divisors.

Let p be a prime number dividing n. The action of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on the points of p-division of E factorizes through an abelian group since the cusps are defined over a cyclotomic field. Therefore $\rho_{E,p}$ is not surjective. Its image is contained in a maximal subgroup of $\mathbf{GL}_2(\mathbf{F}_p)$. It can not be contained in Borel subgroup when $p > 163$. It can not be contained in an exceptional subgroup when $p > 13$. When $p > 163$, it is contained in an abelian subgroup of the normalizer of a Cartan subgroup. This abelian subgroup must be contained either in the Cartan subgroup itself, which is impossible when $p > 37$ by a theorem of Serre ([70]) or into a group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$, which is also impossible because of the surjectivity of the Weil pairing. Therefore one has $p \leq 163$.

In the case of a Frey curve, we have proved that the order of C_E divides 48 or 32. Any element of C_E is defined over the field $\mathbf{Q}(i)$, which is a quadratic field. Kamienny, Kenku and Momose ([35], [32]) determined all the possible torsion subgroups of elliptic curves over quadratic fields. By looking at their list, one finds that that any such torsion subgroup is of order dividing 16 or 24.

Corollary 3.6 *Let E be a Frey curve. The degree δ_E of the corresponding minimal parametrization satisfies the following inequalities*

$$\frac{1}{24} d_E \leq \delta_E \leq 12 d_E.$$

Proof: Let E' be the optimal elliptic curve in the isogeny class of E. Let δ'_E be the degree of the corresponding modular optimal parametrization. Since E is a Frey curve it has no rational cyclic subgroup of order > 12. Therefore one has $\delta'_E \leq \delta_E \leq 12\delta'_E$. The corollary follows from the equality $|C_{E'}|\delta'_E = d_E$ (proposition 3.1) and from the proposition 3.5.

3.3 A problem

We conclude this article by raising a question which can be formulated in elementary terms. By a theorem of Faltings, an elliptic curve over $\mathbf{Q}$ is characterized by its conductor N and its coefficients $a_l(E)$ for l prime number

$< c(\log N)^2$, where c is a universal constant (see [71], note 6 page 632). We have only weaker results of this type for newforms: Newforms of weight two for $\Gamma_0(N)$ are characterized by their Fourier coefficients a_l for l prime number $\leq g_N$, where g_N is the genus of the curve $X_0(N)$.

The coefficients $a_l(E)$ satisfy the inequality $|a_l(E)| < 2\sqrt{l}$.

In view of the degree conjecture and of the corollary 3.6 the following problem arises.

Problem 3.7 *Given an integer $N > 0$ and a family of integers $a_l(E)$ (l prime number $\leq g_N$) of absolute value $< 2\sqrt{l}$, the $\mathbf{Z}$-module of functions $\mathbf{P}^1(\mathbf{Z}/\mathbf{NZ}) \longrightarrow \mathbf{Z}$ satisfying the Manin relations and the $a_l(E)$-modular relations is free of rank 0 or 2. Assuming that it is of rank 2, what bound depending on N only can be given for d_E?*

We view see this problem more as an indication of the difficulty of the degree conjecture rather than a promising way to tackle the *abc* conjecture.

References

[1] D. Abramovich, *Formal finiteness and the torsion conjecture on elliptic curves. A footnote to a paper: "Rational torsion of prime order in elliptic curves over number fields" by S. Kamienny and B. Mazur*, Astérisque **228** (1995), *Columbia University Number Theory Seminar (New-York, 1992)*, 5–17.

[2] A. Ash, G. Stevens, *Modular forms in characteristic l and special values of their L-functions*, Duke Math. J., **53** (1986), no.3, 849–868.

[3] F. Beukers, *The diophantine equation $Ax^p + By^q = Cz^r$*, preprint 1995.

[4] D. Bump, S. Friedberg, J. Hoffstein, *Nonvanishing theorems for L-functions of modular forms and their derivatives*, Invent. Math. **102** (1990), 543–618.

[5] H. Carayol,*Sur les représentations λ-adiques associées aux formes modulaires de Hilbert*, Ann. Sci. de l'ENS, **19** (1986), 409–468.

[6] I. Chen, *The Jacobian of non-split Cartan modular curves*, To appear in the Proceedings of the London Mathematical Society.

[7] J. Cremona, *Computing the degree of a modular parametrization*, in *Algorithmic number theory (Ithaca, NY, 1994)*, 134–142, Lecture Notes in Comput. Sci., **877**, Springer, Berlin, 1994.

[8] H. Darmon, *The equations $x^n + y^n = z^2$ and $x^n + y^n = z^3$*, Internat. Math. Res. Notices, **10** (1993), 263–274.

[9] H. Darmon, *Serre's conjecture*, in *Seminar on Fermat's last Theorem*, CMS Conference Proceedings **17**, American Mathematical Society, Providence, 135–155.

[10] H. Darmon, *Faltings plus epsilon, Wiles plus epsilon and the Generalized Fermat Equation*, preprint 1997.

[11] H. Darmon, *Faltings plus epsilon, Wiles plus epsilon, and the Generalized Fermat Equation*, preprint 1997.

[12] H. Darmon, A. Granville, *On the equations $x^p + y^q = z^r$ and $z^m = f(x, y)$*, Bulletin of the London Math. Society, no 129, **27** part 6, November 1995, 513–544.

[13] H. Darmon, L. Merel, *Winding quotients and some variants of Fermat's last theorem*, To appear in Crelle.

[14] P. Deligne, M. Rappoport, *Les schémas de module des courbes elliptiques*, in *Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972)*, 143–316, Lecture Notes in mathematics **349**, Springer, Berlin, 1975.

[15] P. Dénes, *Über die Diophantische Gleichung $x^\ell + y^\ell = cz^\ell$*, Acta Math. **88** (1952), 241–251.

[16] F. Diamond, *On deformation rings and Hecke rings*, Ann. of Math. (2), **144** (1996), no. 1, 137–166.

[17] F. Diamond, K. Kramer, *Modularity of a family of elliptic curves*, Math. Res. Letters, **2** (1995), 299–304.

[18] L.E. Dickson, *History of the theory of numbers*, Chelsea, New York, 1971.

[19] V. Drinfeld, *Two theorems on modulars curves*, Funct. anl. appl., **2** (1973), 155–156.

[20] B. Edixhoven, *On a result of Imin Chen*, preprint 1995. To appear in: Séminaire de théorie des nombres de Paris, 1995-96, Cambridge University Press.

[21] N. Elkies, *Wiles minus epsilon implies Fermat*, in *Elliptic curves, modular forms and Fermat's Last Theorem (Hong-Kong 1993)*, J. Coates, S-T. Yau, eds., Internat. Press, Cambridge, MA, 1995, 38–40.

[22] G. Frey, *Links between stable elliptic curves and certain diophantine equations*, Ann. Univ. Saraviensis, Ser. Math **1** (1986), 1–40.

[23] G. Frey, *Links between solutions of $A - B = C$ and elliptic curves*, Lect. Notes in Math. **1380** (1989), 31–62.

[24] G. Frey, *On elliptic curves with isomorphic torsion structures and corresponding curves of genus* 2, in *Elliptic curves, modular forms and Fermat's Last Theorem (Hong-Kong,1993)*, J. Coates, S-T. Yau, eds., Internat. Press, Cambridge, MA, 1995, 79–98.

[25] G. Frey, *On ternary relations of Fermat type and relations with elliptic curves*, preprint 1996.

[26] A. Granville, *On the number of solutions of the generalized Fermat equation*, in *Number Theory (Halifax, NS, 1994)*, 197–207, CMS Conf. Proc., **15**, Amer. Math. Soc., Providence, RI, (1994).

[27] B. Gross, G. Lubin, *The Eisenstein descent on $J_0(N)$*, Invent. Math., **83** (1986), 303–319.

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

[29] A. Grothendieck, *Esquisse d'un programme*, 1984.

[30] Y. Hellegouarch, *Points d'ordre $2p^h$ sur les courbes elliptiques.* Acta Arith., **26** (1974/75), no. 3, 253–263.

[31] Y. Hellegouarch, *Thèse*, Université de Besançon, 1972.

[32] M. Kenku, F. Momose, *Torsion points on elliptic curves defined over quadratic fields*, Nagoya Mathematical Journal, **109** (1988), 125–149.

[33] S. Kamienny, *Points on Shimura curves over fields of even degree*, Math. Ann. **286** (1990), 731–734.

[34] S. Kamienny, *Torsion points of elliptic curves over fields of higher degree*, International Mathematics Research Notices, **6** (1992), 129–133.

[35] S. Kamienny, *Torsion points on elliptic curves and q-coefficients of modular forms*, Invent. Math., **109** (1992), 221–229.

[36] K. Kato, *p-adic Hodge theory and special values of zeta functions of elliptic cusp forms*, to appear.

[37] K. Kato,*Euler systems, Iwasawa theory, and Selmer groups*, preprint.

[38] K. Kato,*Generalized explicit reciprocity laws*, preprint.

[39] V. A. Kolyvagin, D. Logachev, *Finiteness of the Shafarevich-Tate group and the group of rational points for some modular abelian varieties,* Leningrad Math. J., vol. **1** no. 5 (1990), 1229–1253.

[40] A. Kraus, *Une remarque sur les points de torsion des courbes elliptiques*, C. R. Acad. Sci. Paris, **321**, Série I (1995), 1143–1146.

[41] A. Kraus, *Sur certaines variantes de l'équation de Fermat*, preprint 1997.

[42] G. Ligozat, *Courbes modulaires de niveau* 11, in *Modular functions of one variable V*, Lecture Notes in Math. **601** (1977), 115–152.

[43] S. Ling, J. Oesterlé, *The Shimura subgroup of $J_0(N)$*, In *Courbes modulaires et courbes de Shimura*, Astérisque **196-197**, (1991), 171–203.

[44] L. Mai, R. Murty, *The Phragmen-Lindelof theorem and modular elliptic curves*, in *The Rademacher legacy to mathematics University Park, PA, 1992*, 335–340, Contemp. Math., **166**, Amer. Math. Soc., Providence, RI, 1994.

[45] Y. Manin, *Parabolic points and zeta functions on modular curves*, Math. USSR Izvestija, **6**, no. 1 (1972), 19–64.

[46] Y. Manin, *Modular forms and number theory*, In the proceedings of the international congress of mathematicians 1978, (1980), 177–186.

[47] D. Masser, G. Wüstholz, *Galois properties of division fields of elliptic curves*, Bull. London Math. Soc. **25** (1993), no. 3, 247–254.

[48] B. Mazur, H.P.F. Swinnerton-Dyer, *The arithmetic of Weil curves*, Invent. Math. **25** (1974), 1–61.

[49] B. Mazur, *Modular curves and the Eisenstein ideal*, Publ. Math. IHES, **47** (1977), 33–186.

[50] B. Mazur, *Rational isogenies of prime degree*, Invent. Math. **44** (1978), 129–162.

[51] B. Mazur, *Questions about number*, in *New Directions in Mathematics*, to appear.

[52] B. Mazur, *Courbes elliptiques et symboles modulaires*, Séminaire Bourbaki **414**, Lecture Notes in mathematics **317** (1973), 277–294.

[53] B. Mazur, Letter to J. Ellenberg.

[54] L. Merel, *Bornes pour la torsion des courbes elliptiques sur les corps de nombres*, Invent. Math. **124** (1996), no. 1–3, 437-449.

[55] L. Merel, *Homologie des courbes modulaires affines et paramétrisations modulaires*, in *Elliptic curves, modular forms, and Fermat's last theorem (Hong-Kong 1993)*, J. Coates, S.-T. Yau, eds, Internat. Press, Cambridge, MA, 1995, 110–130.

[56] F. Momose, *Rational points on the modular curves $X_{\mathrm{split}}(p)$*, Compositio Math. **52** (1984), 115–137.

[57] K. Murty, R. Murty, *Mean values of derivatives of L-series*, Ann. Math. **133** (1991), 447–475.

[58] A. Nitaj, *La conjecture abc*, Enseign. Math., **42** (1996), no. 1-2, 3–24.

[59] J. Oesterlé, *Nouvelles approches du "théorème" de Fermat*, Sém. Bourbaki **694**, Astérisque **161-162**, S.M.F. (1988), 165–186.

[60] I. Papadopoulos, *Sur la classification de Néron des courbes elliptiques*, J. Number Theory, **44** (1993), no.2, 119–152.

[61] P. Parent, *Bornes effectives pour la torsion des courbes elliptiques sur les corps de nombres*, prépublication 95-33, Institut de recherches mathématiques de Rennes (1995).

[62] B. Poonen, *Some diophantine equations of the form $x^n + y^n = z^m$*, to appear.

[63] K. Ribet, *On modular representations of* $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ *arising from modular forms*, Invent. Math. **100** (1990), 431–476.

[64] K. Ribet, *On the equation $a^p + 2^\alpha b^p + c^p = 0$*, Acta Arith., **79** (1997), no. 1, 7–16.

[65] K. Rubin, A. Silverberg, *A report on Wiles' Cambridge lecture*, Bull. Amer. Math. Soc. (N.S.), **31** (1994), no.1, 15–38.

[66] J.-P. Serre, *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques*, Invent. Math. **15** (1972), 259–331.

[67] J.-P. Serre, *Sur les représentations modulaires de degré* 2 *de* $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$, Duke Math. J. Vol. **54**, no. 1 (1987), 179–230.

[68] J.-P. Serre, *Propriétés conjecturales des groupes de Galois motiviques et des représentations l-adiques*, Proceedings of Symposia in Pure Mathematics, **55** (1994), Part 1, 377–400.

[69] J.-P. Serre, *Travaux de Wiles (et Taylor,...), Partie I*, Séminaire Bourbaki, **803**, Juin 1995, Astérisque **237** (1996), 5, 319–332.

[70] J.-P. Serre, *Quelques applications du théorème de densité de Chebotarev*, Pub. Math. I.H.E.S, **54** (1981), 123–201.

[71] J.-P. Serre, *Oeuvres*, vol. III, Springer-Verlag.

[72] J.-P. Serre, *Représentations linéaires des groupes finis*, Hermann, Paris, 1978.

[73] J. Silverman, *Heights and elliptic curves*, in *Arithmetic geometry (Storrs, Conn., 1984)*, 151–166, Springer, New-York, 1986.

[74] L. Szpiro, *Discriminants et conducteurs de courbes elliptiques*, in *Séminaire sur les pinceaux de courbes elliptiques (Paris, 1988)*, Astérisque, **183** (1990), 7–18.

[75] R. Taylor, A. Wiles, *Ring-theoretic properties of certain Hecke algebras*, Ann. of Math. **141** (1995), 553–572.

[76] P. Vojta, *Diophantine approximation and value distribution theory*, Lecture Notes in Mathematics, **1239**, Springer-Verlag, Berlin, 1987.

[77] A. Wiles, *Modular elliptic curves and Fermat's Last Theorem*, Ann. of Math. **141** (1995), 443–551.

[78] D. Zagier, *Modular parametrizations of elliptic curves*, Can. Math. Bull., **28** (1985), 372–384.

Loïc Merel (Miller Institute), Department of Mathematics, 970 Evans, University of California, Berkeley, CA 94720, USA. merel@math.berkeley.edu