The equations $x^n + y^n = z^2$ and $x^n + y^n = z^3$

Henri Darmon

September 9, 2007

In this note we consider the equations

$$x^n + y^n = z^2, \quad \text{and} \quad x^n + y^n = z^3,$$

which have a long history in connection with Fermat's last theorem. Call an integral solution (x, y, z) to one of the above equations *proper* if $\gcd(x, y, z) = 1$, and say that it is *non-trivial* if $xyz \neq 0$. We propose the following conjecture:

Conjecture.

1. *The equation $x^n + y^n = z^2$ has no non-trivial proper solutions when $n \geq 4$.*

2. *The equation $x^n + y^n = z^3$ has no non-trivial proper solutions when $n \geq 3$.*

Remarks:
1. It is easy to see that the equations $x^2 + y^2 = z^2$ and $x^3 + y^3 = z^2$ have infinitely many proper solutions. Fermat himself showed that the equation $x^4 + y^4 = z^2$ has no non-trivial integer solutions, probably the most elementary and widely quoted example of his method of descent, and thus derived a proof of Fermat's last theorem for exponent 4.

2. The statement that $x^n + y^n = z^3$ has no non-trivial proper solution when $n = 3$ is the statement of Fermat's Last Theorem for exponent 3, which was proved by Euler in 1753.

Andrew Wiles recently stunned the mathematical world by announcing the proof of Fermat's Last Theorem. In fact, he proved (for semi-stable elliptic curves) the celebrated conjecture of Shimura and Taniyama which asserts that every elliptic curve defined over $\mathbf{Q}$ is the quotient of a modular curve $X_0(N)$. Earlier work of Frey [Fr], Serre [Sr2], and Ribet [Ri] had shown that this conjecture implied Fermat's last theorem.

Our main result is the following:

Theorem A. *Let $p > 13$ be prime. If the Shimura-Taniyama conjecture is true, then*

1. *The equation $x^p + y^p = z^2$ has no non-trivial proper solutions when $p \equiv 1 \pmod 4$.*

2. *The equation $x^p + y^p = z^3$ has no non-trivial proper solutions when $p \equiv 1 \pmod 3$, and p is not a Mersenne[1] prime.*

The proof of theorem A is based on a variant of Frey's beautiful trick [Fr], combined with the deep work of Serre and Ribet. An essential new ingredient is a result of Kamienny on the finiteness of certain Eisenstein quotients over imaginary quadratic fields.

The assumption that the Shimura-Taniyama conjecture is true cannot be removed yet from the statement of the theorem, because, as we shall see, the elliptic curves that arise in the proof are not semi-stable. However, this assumption is not as formidable as it used to be! Indeed, one expects that very soon the conjecture of Shimura and Taniyama will be established for all elliptic curves, making theorem A unconditional.

Section 1 collects the basic facts on elliptic curves, modular Galois representations and Serre's conjectures that are used in the proofs. Section k ($2 \leq k \leq 3$) gives the proof of theorem A for the equation $x^p + y^p = z^k$. Finally section 4 comments on the method and its limitations.

[1]Recall that a *Mersenne prime* is a prime of the form $2^m - 1$.

Acknowledgements:
This work grew out of a joint project with Andrew Granville on the generalized Fermat equation $x^p+y^q = z^r$ [DG]. I am grateful to him for a stimulating collaboration. I also wish to thank Dan Abramovich and Noam Elkies for many helpful and enlightening conversations. This research was partially supported by NSF grant #DMS-8703372 and an NSERC post-doctoral fellowship.

1 Preliminaries

Let E be an elliptic curve defined over $\mathbf{Q}$, let N be its arithmetic conductor, Δ its discriminant, and j its j-invariant.

Let p be a prime which is greater or equal to 5. By considering the action of $G_{\mathbf{Q}} = \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on the p-division points of E, one obtains a Galois representation

$$\rho : G_{\mathbf{Q}} \longrightarrow \mathbf{GL}_2(\mathbf{F}_p).$$

Let $N(\rho)$ be the Artin conductor of ρ, defined as for a representation in characteristic zero, except that one ignores the contribution of the prime p (cf. [Sr2], p. 180). We will need the following facts about $N(\rho)$:

Lemma 1.1 .

1. *The integer $N(\rho)$ divides N. In particular, if E has good reduction at l, then l does not divide $N(\rho)$.*

2. *If E has multiplicative reduction at $l \neq p$, then*

$$ord_l(N(\rho)) = \begin{cases} 0, & \text{if } ord_l(j) \equiv 0 \pmod{p}; \\ 1, & \text{otherwise.} \end{cases}$$

These properties are explained in [Sr2], p. 207.

If l is a prime not dividing $pN(\rho)$, then ρ is unramified at l. Let $\rho(\mathrm{Frob}_l)$ denote the image of the Frobenius conjugacy class at l in $\mathbf{GL}_2(\mathbf{F}_p)$.

Making a choice of a prime $\bar{p}$ of $\bar{\mathbf{Q}}$ above p gives an extension of the normalized valuation $\mathrm{ord}_p : \mathbf{Q}^* \longrightarrow \mathbf{Z}$ to $\mathrm{ord}_{\bar{p}} : \bar{\mathbf{Q}}^* \longrightarrow \mathbf{Q}$. Define $\mathrm{ord}_{\bar{p}}(0) = \infty > 0$. Given $a, b \in \bar{\mathbf{Q}}$, say that a and b are congruent (mod $\bar{p}$),

$$a \equiv b \pmod{\bar{p}}, \text{ if } \mathrm{ord}_{\bar{p}}(a-b) > 0.$$

Given an integer $M > 0$, we will be interested in holomorphic cusp forms f of weight 2 on $X_0(M)$ which are eigenvalues of the Hecke operators and the Atkin-Lehner involutions. We assume such a form has been normalized so that its first Fourier coefficient is equal to 1. We can then expand f in a Fourier expansion about the cusp ∞, by

$$f = \sum_{n=1}^{\infty} a_n q^n,$$

where the a_n are algebraic integers. We will use the expression "eigenform of level M" to denote a normalized newform of weight 2 on $X_0(M)$, since these are the only types of modular forms that will occur.

In [Sr2], conj. $3.2.3_?$ and $3.2.4_?$, J-P. Serre makes a general conjecture relating Galois representations such as ρ to eigenforms:

Conjecture 1.2 ($\mathbf{Sr}^?$) *Suppose that*

1. *The representation ρ is irreducible;*
2. *The curve E has either good reduction or multiplicative reduction at p, and $ord_p(j) \equiv 0 \pmod{p}$.*

Then there exists an eigenform f of level $N(\rho)$ with Fourier coefficients $\{a_n\}_{n \in \mathbf{Z}}$ satisfying

$$a_l = trace(\rho(frob_l)) \pmod{\bar{p}} \qquad \text{for all } l \not| pN(\rho).$$

Remark: The conjecture $3.2.3_?$ of [Sr2] is a good deal more general than the version given here; we have contented ourselves with stating only what we will need. The condition 2 in the statement of the conjecture, which may appear artificial, is necessary to ensure that the representation ρ is *finite* at p. Otherwise, Serre's conjecture associates to ρ a modular form of weight $p+1$ and level $N(\rho)$, satisfying the same conclusions as above.

The conjecture of Shimura and Taniyama, combined with the theory of Eichler and Shimura, can be stated as follows.

Conjecture 1.3 ($\mathbf{ST}^?$) *Let E and ρ be as above. Then there exists an eigenform f of level N with Fourier expansion given by*

$$f = \sum_{n=1}^{\infty} a_n q^n, \qquad a_1 = 1, a_n \in \mathbf{Z},$$

satisfying

$$a_l = trace(\rho(frob_l)) \pmod p \qquad \text{for all } l \not| pN.$$

A deep theorem of Ribet [Ri] shows that there are good reasons to believe conjecture $\mathrm{Sr}^?$.

Proposition 1.4 (Ribet) *Conjecture $ST^?$ implies conjecture $Sr^?$.*

In the proof of thm. A we will be assuming conjecture $\mathrm{ST}^?$ only in order to apply conjecture $\mathrm{Sr}^?$.

To check the hypotheses of conjecture $\mathrm{Sr}^?$, one needs to know when a representation ρ arising from elliptic curves is irreducible. The following powerful result of Mazur provides a good grip for handling such questions.

Proposition 1.5 (Mazur) *If $p > 13$, and ρ is reducible, then $j(E)$ belongs to $\mathbf{Z}[\frac{1}{2}]$.*

Proof: Theorem 7.1 of [Mz] lists all the possible elliptic curves defined over $\mathbf{Q}$ having a rational subgroup of order p, with $p > 13$. There are only finitely many, and it can be checked that they all have integral j-invariant, except for the pair of non-CM curves related by a 17-isogeny, whose j-invariant belongs to $\mathbf{Z}[\frac{1}{2}]$.

Mazur's result sufficed in showing that conjecture $\mathrm{ST}^?$ implies Fermat's last theorem (see [Sr2], §4.2, or [Fr]). A key ingredient in the proof of theorem A is the following result of S. Kamienny:

Proposition 1.6 (Kamienny) *Let K be a quadratic imaginary field, and let p be a prime which is split in K and q a prime which is not. Suppose that there is a prime n dividing the numerator of $\frac{(p+1)(q-1)}{24}$ but not $q(q-1)$, and that n does not divide the class number of K. Then any elliptic curve E over K having a subgroup of order pq defined over K has potentially good reduction at all primes not dividing 6.*

Proof: By [Ka], prop. 2.1., the n-Eisenstein quotient $A(pq)$ associated to $X_0(pq)$ gives a non-trivial optimal quotient of the new part of $J_0(pq)$ which has finite Mordell-Weil group over K. The result follows from cor. 4.3. of [Mz].

We will apply Kamienny's result only in the cases where the quadratic field K is $\mathbf{Q}(i)$ or $\mathbf{Q}(\sqrt{-3})$, and $q = 2$ or 3:

Corollary 1.7 *Let p be a prime with $p > 13$.*

1. *Suppose that $p \equiv 1 \pmod 4$. If E is an elliptic curve over $\mathbf{Q}(i)$, having a $\mathbf{Q}(i)$-rational subgroup of order $2p$, then $j(E)$ belongs to $\mathbf{Z}[i][\frac{1}{6}]$.*

2. *Suppose that $p \equiv 1 \pmod 3$ and that p is not a Mersenne prime. If E is an elliptic curve over $\mathbf{Q}(\sqrt{-3})$, having a $\mathbf{Q}(\sqrt{-3})$-rational subgroup of order $3p$, then $j(E)$ belongs to $\mathbf{Z}[\sqrt{-3}][\frac{1}{6}]$.*

2 The equation $x^p + y^p = z^2$

From now on, we assume that p is a prime which is > 13.
Let $a^p + b^p = c^2$ be a non-trivial proper solution to the equation

$$x^p + y^p = z^2.$$

Suppose that (a, b, c) satisfies one of the following congruences:

Case 1: $a \equiv 0 \pmod 2$, $\quad c \equiv 1 \pmod 4$.

Case 2: $a \equiv -1 \pmod 4$, $\quad c \equiv 0 \pmod 2$.

One can always make (a, b, c) satisfy one of the above systems of congruences, by interchanging a and b and repacing c by $-c$ if necessary. Let $E_{a,b,c}$ be the elliptic curve given by the following equations:

Case 1: $Y^2 = X^3 + cX^2 + a^p/4X$;

Case 2: $Y^2 = X^3 + 2cX^2 + a^pX$.

The curve $E_{a,b,c}$ has j-invariant

$$j = 2^6 \frac{(a^p + 4b^p)^3}{(a^{2p}b^p)}.$$

Its discriminant, Δ, is equal to $(a^2b)^p$ in case 1, and to $2^6(a^2b)^p$ in case 2.

Using Tate's algorithm (cf. [Ta]), we compute the arithmetic conductor of $E_{a,b,c}$.

Lemma 2.1 .

1. *In case* 1*, the curve* $E_{a,b,c}$ *has multiplicative reduction at* 2*, and hence its conductor over* $\mathbf{Q}_2$ *is* 2*.*

2. *In case* 2*, the curve* $E_{a,b,c}$ *has additive reduction at* 2*, and its conductor over* $\mathbf{Q}_2$ *is* 2^5*.*

3. *If* $l \neq 2$*, then* $E_{a,b,c}$ *has either good reduction at* l*, or multiplicative reduction.*

Proof: 1. By setting $Y = 8y + 4x$, $X = 4x$, we get the following equation for $E_{a,b,c}$ in case 1:

$$y^2 + xy = x^3 + (\frac{c-1}{4})x^2 + \frac{a^p}{2^6}x.$$

Since $p > 6$ and $c \equiv 1 \pmod 4$, the coefficients in the above equation belong to $\mathbf{Z}_2$. Hence the curve $E_{a,b,c}$ has reduction of the form

$$y^2 + xy = x^3 \text{ if } c \equiv 1 \pmod 8,$$

$$y^2 + xy = x^3 + x^2 \text{ if } c \equiv 5 \pmod 8,$$

and these are both of multiplicative type.
2. By making the change of variable $X = x + 1$, the curve $E_{a,b,c}$ has the equation

$$y^2 = x^3 + (3 + 2c)x^2 + (3 + 4c + a^p)x + (1 + 2c + a^p).$$

By applying Tate's algorithm and using the fact that $1+2c+a^p \equiv 0 \pmod 4$ and that $3+4c+a^p \equiv 2 \pmod 4$, we find that the fiber in the Néron model at 2 has two connected components, and the conductor of Ogg[2] is 2^5.
3. If l divides Δ, then l divides ab, and hence l does not divide c, since the solution (a, b, c) is assumed to be proper. Therefore, the cubic equation defining $E_{a,b,c}$ has at most a double root, and the reduction is multiplicative.

[2]Here we are using Ogg's formula which computes the conductor in terms of the number of components in the singular fiber of the Néron model. The original proof of this formula does not work in mixed characteristic 2, but recent work of Paul Lockhart and Joseph Silverman[LS] shows that it holds for the cases we are interested in.

Lemma 2.2 *$j(E_{a,b,c})$ does not belong to $\mathbf{Z}[\frac{1}{2}]$.*

Proof: For if it did, then ab would be a power of 2. Assume without loss of generality that $a = 2^\alpha$, and $b = \pm 1$. The equation becomes

$$2^{p\alpha} = c^2 \pm 1,$$

whose only solutions are $\alpha = c = 0$, contradicting the assumption that (a, b, c) is a non-trivial solution.

Let

$$\rho_{a,b,c} : \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \mathbf{GL}_2(\mathbf{F}_p)$$

be the mod p Galois representation attached to $E_{a,b,c}$, and let $N(\rho_{a,b,c})$ be its Artin conductor.

Lemma 2.3 .

1. *In case* 1*, the Artin conductor $N(\rho_{a,b,c})$ divides* 2*.*
2. *In case* 2*, the Artin conductor $N(\rho_{a,b,c})$ divides* 32*.*

Proof: Since the j-invariant of $E_{a,b,c}$ is $2^6(a^p+4b^p)/(a^{2p}b^p)$, and $\gcd(a, b) = 1$, it follows that $\mathrm{ord}_l(j) \equiv 0 \pmod{p}$ for all odd primes l dividing Δ (and hence N). Hence by lemma 1.1, no odd primes divide $N(\rho_{a,b,c})$. The computation of the conductor of E at 2 in parts 1 and 2 of lemma 2.1 shows that $N(\rho_{a,b,c})$ is equal to 2 in case 1, and divides 32 in case 2.

Theorem 2.4 *Assume conjecture ST$^?$. Then the equation $x^p + y^p = z^2$ has no proper solutions except for the trivial ones $(1, 0, \pm 1)$, $(0, 1, \pm 1)$, and $(\pm 1, \mp 1, 0)$, when $p \equiv 1$ (mod 4).*

Proof: The representation $\rho_{a,b,c}$ is irreducible by prop. 1.5 combined with lemma 2.2. Also, we have $\mathrm{ord}_p(j) \equiv 0 \pmod{p}$. Hence $E_{a,b,c}$ and $\rho_{a,b,c}$ satisfy the hypotheses in conj. Sr$^?$. By conjecture ST$^?$ combined with Ribet's prop. 1.4, conjecture Sr$^?$ holds, and hence there is associated to $\rho_{a,b,c}$ an eigenform f of level $N(\rho_{a,b,c})$ satisfying the conclusion of conjecture Sr$^?$. This rules out solutions in case 1, even without the assumption $p \equiv 1 \pmod{4}$, since there are no eigenforms of level 2.

In case 2, the representation $\rho_{a,b,c}$ is associated to an eigenform of level dividing 32. Since $X_0(32)$ has genus 1, there is only one such eigenform, which corresponds to the curve

$$X_0(32) = E_{-1,1,0} : Y^2 = X^3 - X,$$

with complex multiplication by $\mathbf{Z}[i]$. By the Chebotarev density theorem, the representations $\rho_{a,b,c}$ and $\rho_{-1,1,0}$ are conjugate. From the theory of complex multiplication, one knows that $\rho_{-1,1,0}$ maps $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ onto the normalizer H of a Cartan subgroup of $\mathbf{GL}_2(\mathbf{F}_p)$. More precisely, there is an exact sequence

$$0 \longrightarrow C \longrightarrow H \longrightarrow \mathbf{Z}/2\mathbf{Z} \longrightarrow 0,$$

where C is a maximal commutative subgroup of $\mathbf{GL}_2(\mathbf{F}_p)$, and the field fixed by $\rho^{-1}(C)$ is $\mathbf{Q}(i)$.

Now, when $p = 1 \pmod 4$, it is known that C is a *split* Cartan subgroup, which is the stabilizer of two one-dimensional $\mathbf{F}_p$-subspaces in the space of p-division points of $E_{a,b,c}$. Therefore, $E_{a,b,c}$ has two subgroups of order p which are rational over $\mathbf{Q}(i)$. In addition, it can be seen from the equations that $E_{a,b,c}$ has a rational subgroup of order 2. Part 1 of cor. 1.7 implies that ab is divisible only by 2 and 3; since we are in case 2, and ab is assumed to be odd, it follows that ab is a power of 3. Since $\gcd(a, b) = 1$, assume without loss of generality that $a = 3^\alpha$ and $b = \pm 1$. Then we have

$$3^{p\alpha} = c^2 \pm 1,$$

and at least one of $c \pm 1$ or $c \pm i$ is a unit. Hence $c = 0$, and the solution must be a trivial one.

Remarks:
1. In [Fr], Frey considered the family of elliptic curves

$$E^{frey}_{a,b,c} : y^2 = x(x - a^p)(x - b^p)$$

indexed by solutions $a^p - b^p = c^p$ of Fermat's equation. The 3 trivial solutions to the Fermat equation give rise to degenerate elliptic curves with nodal singularities. In the family we consider, one of the trivial solutions, namely $(-1, 1, 0)$, gives rise to the elliptic curve with complex multiplications by $\mathbf{Z}[i]$. This is the source of the extra difficulties, which make it necessary to invoke

prop. 1.6 and still prevent us from tackling the case $p \equiv -1 \pmod 4$. A similar difficulty occurs in the case of the equation $x^p + y^p = z^3$, with $\mathbf{Z}[i]$ replaced by $\mathbf{Z}[\frac{1+\sqrt{-3}}{2}]$.

2. One can consider variants of the form $Ax^p + By^p = Cz^2$ of the original equation which have no trivial solution with $z = 0$. For such equations, one can hope to prove more and with less effort, although the results one obtains are less natural. To illustrate this, we can show:

Proposition 2.5 *Assume $ST^?$. If p is a prime which is 11 or > 13, then the equation $x^p + 4y^p = z^2$ has no proper solutions except for the trivial ones $(1, 0, \pm 1)$ and $(0, 1, \pm 2)$.*

Proof: Let $a^p + 4b^p = c^2$ be a proper solution. We consider three seperate cases:

Case 1: a is even; $E_{a,b,c} : Y^2 = X^3 + cX^2 + a^p/4X$.

Case 2: a is odd, b is even; $E_{a,b,c} : Y^2 = X^3 + cX^2 + b^pX$.

Case 3: a and b are odd; $E_{a,b,c} : Y^2 = X^3 + cX^2 + b^pX$.

In cases 1 and 2 one finds as before that $E_{a,b,c}$ is semistable and that $N(\rho_{a,b,c})$ is equal to 2. From this one derives a contradiction, since there are no eigenforms of level 2. In case 3, one finds that $N(\rho_{a,b,c})$ divides 16. Since there are no modular forms of weight 2 and level dividing 16 (the curve $X_0(16)$ has genus 0) we conclude that $E_{a,b,c}$ cannot exist, as before. It follows that $E_{a,b,c}$ must be a degenerate elliptic curve, corresponding to $ab = 0$, leading to the trivial solutions above.

3 The equation $x^p + y^p = z^3$

Let (a, b, c) be a proper non-trivial solution to the equation

$$x^p - y^p = z^3.$$

Suppose that a, b, c satisfy one of the following congruences:

Case 1: b is even, and $c \equiv -1 \pmod 4$.

Case 2: ab is odd.

Define the elliptic curve $E_{a,b,c}$ as follows:

Case 1: $E_{a,b,c} : Y^2 = X^3 + 3cX^2 + 4b^p$,

Case 2: $E_{a,b,c} : Y^2 = X^3 - 3(9a^p - b^p)cX + 2(27a^{2p} - 18a^p b^p - b^{2p})$.

Let j and Δ denote the j-invariant and discriminant of the curve $E_{a,b,c}$. They are given by:

Case 1:

$$j = -2^4 3^3 \frac{(a^p - b^p)^2}{a^p b^p}, \qquad \Delta = -2^8 3^3 a^p b^p,$$

Case 2:

$$j = -3^3 \frac{(a^p - b^p)(9a^p - b^p)^3}{a^p b^p}, \qquad \Delta = -2^{12} 3^3 a^p b^{3p}.$$

Remark: The curves $E_{a,b,c}$ that are considered in cases 1 and 2 are not twists of one another as they were in section 2. Geometrically, the curve $E_{a,b,c}$ in case 1 arises as the pullback of a universal elliptic curve over the j-line to a covering of degree 2 which is ramified above $j = 0$ and $j = 1728$. The curve in case 2 arises from a universal family over $X_0(3)$. Since this last fact will be important in the proof later, we record it as a lemma:

Lemma 3.1 *In case* 2*, the curve $E_{a,b,c}$ has a rational subgroup of order* 3*.*

Proof: One can check that the point

$$P = (x, y) = (3c^2, 4\sqrt{b^p})$$

belongs to $E_{a,b,c}$ and generates a rational subgroup of $E_{a,b,c}$ of order 3.

Now we study the conductor of $E_{a,b,c}$, as before:

Lemma 3.2 .

1. *In case* 1*, the curve $E_{a,b,c}$ has multiplicative reduction at* 2*. The conductor of $E_{a,b,c}$ over $\mathbf{Q}_3$ divides* 27*.*
2. *In case* 2*, the curve $E_{a,b,c}$ has good reduction at* 2*. The conductor of $E_{a,b,c}$ over $\mathbf{Q}_3$ divides* 27*.*
3. *If $l \neq 2$ or* 3*, then $E_{a,b,c}$ has either good reduction at l, or multiplicative reduction.*

Proof: This is proved as for lemma 2.1, applying Tate's algorithm. The details are left to the reader.

As before, let $\rho_{a,b,c}$ be the Galois representation arising from the action of Galois on the p-division points of $E_{a,b,c}$.

Lemma 3.3 .

1. *In case 1, the conductor $N(\rho_{a,b,c})$ is even and divides* 54.

2. *In case 2, the conductor $N(\rho_{a,b,c})$ divides* 27.

Proof:
1. In case 1, by part 1 of lemma 3.2, the conductor of $E_{a,b,c}$ at 2 is equal to 2. Moreover, the representation $\rho_{a,b,c}$ is ramified at 2, and hence 2 divides $N(\rho_{a,b,c})$ exactly. Likewise, the conductor of $\rho_{a,b,c}$ at 3 divides 27. For all other primes l, the curve $E_{a,b,c}$ has multiplicative reduction, by part 3 of lemma 3.2, and $\mathrm{ord}_l(j) \equiv 0 \pmod{p}$. Hence l does not divide $N(\rho_{a,b,c})$, by lemma 1.1, and therefore $N(\rho_{a,b,c})$ divides 54.

2. This follows in the same way from parts 2 and 3 of lemma 3.2 combined with lemma 1.1.

Theorem 3.4 *Assume conjecture $ST^?$. Let $p \equiv 1$ (mod 3) be a prime which is not a Mersenne prime. Then the equation $x^p + y^p = z^3$ has no proper solutions except the trivial ones $(\pm 1, 0, \pm 1)$, $(0, \pm 1, \pm 1)$, and $(\pm 1, \mp 1, 0)$.*

Proof: As in the proof of th. 2.4, the representation $\rho_{a,b,c}$ is irreducible and $\mathrm{ord}_p(j) \equiv 0 \pmod{p}$. Hence by conjecture $\mathrm{Sr}^?$, $\rho_{a,b,c}$ gives rise to an eigenform of level $N(\rho_{a,b,c})$.
Case 1: Since $N(\rho_{a,b,c})$ is divisible by 2 and divides 54, and there are no eigenforms on $X_0(2)$, $X_0(6)$, and $X_0(18)$, the representation $\rho_{a,b,c}$ must come from an eigenform of level 54. Inspection of tables 5 and 3 in [MF] shows that there are two such forms, whose fifth Fourier coefficients are ± 3. On the other hand,
a) If 5 divides ab, then $E_{a,b,c}$ has multiplicative reduction at 5. If this reduction is split, then $E_{a,b,c}$ is isomorphic to a Tate curve over $\mathbf{Q}_5$, and the group scheme of p-torsion points of $E_{a,b,c}$ over $\mathbf{Q}_5$ is an extension of $\mathbf{Z}/p\mathbf{Z}$ by μ_p. Hence the eigenvalues of Frob_5 are 1 and 5. It follows that $a_5 = 6$. If the

reduction of $E_{a,b,c}$ is non-split, then $a_5 = -6$. But neither of these cases can occur, since $\pm 3 \equiv \pm 6 \pmod{p}$ would imply $p = 2$ or 3.
b) If 5 does not divide ab, then $E_{a,b,c}$ has good reduction at 5, and one can compute the trace of $\rho_{a,b,c}(\mathrm{Frob}_5)$ directly, by counting the number of points of $E_{a,b,c}$ over $\mathbf{F}_5$. Running over all possible values of a^p and b^p in $\mathbf{Z}/5\mathbf{Z}^*$, one finds that the possibilities for this trace are 0, ± 1, and ± 4. Since $p > 7$, there are no values of a and b mod 5 for which this trace is congruent to $\pm 3 \pmod{p}$.

Hence, in case 1, the curve $E_{a,b,c}$ cannot exist, and thus we have proved that the equation $x^p + y^p = z^3$ has no proper solutions with xy even and non-zero, assuming only conjecture $\mathrm{ST}^?$.

Case 2: Since there are no eigenforms of level 3 or 9, it follows from conjecture $\mathrm{Sr}^?$ that $\rho_{a,b,c}$ corresponds to an eigenform of level 27. There is only one such form, since $X_0(27)$ has genus 1; in fact, $X_0(27)$ is the elliptic curve

$$E_{1,1,0} : Y^2 = X^3 + 16,$$

which has complex multiplication by $\mathbf{Z}[\frac{1+\sqrt{-3}}{2}]$. By the Chebotarev density theorem, $\rho_{a,b,c}$ is conjugate to $\rho_{1,1,0}$. By the theory of complex multiplication, $\rho_{1,1,0}$ maps to the normalizer H of a Cartan subgroup of $\mathbf{GL}_2(\mathbf{F}_p)$. This Cartan subgroup is split if $p \equiv 1 \pmod{3}$, and is non-split otherwise, and the quadratic field cut out by the homomorphism $G_{\mathbf{Q}} \longrightarrow H \longrightarrow \mathbf{Z}/2\mathbf{Z}$ is the field $\mathbf{Q}(\sqrt{-3})$. Hence, if $p \equiv 1 \pmod{3}$, the curve $E_{a,b,c}$ has two subgroups of order p which are rational over $\mathbf{Q}(\sqrt{-3})$. In addition, $E_{a,b,c}$ has a subgroup of order 3 rational over $\mathbf{Q}$, by lemma 3.1. By part 2 of cor. 1.7, it follows that ab is divisible only by 2 and 3, and hence is a power of 3 since ab is odd. By an argument similar to the one of sec. 2, one concludes that $c = 0$ and hence the solution (a, b, c) is trivial.

This concludes the proof of thm. 3.4. Theorem A follows by combining thm. 2.4 and thm. 3.4.

4 Comments

1. It can be shown that the generalized Fermat equation

$$Ax^p + By^q = Cz^r, \qquad A, B, C \in \mathbf{Z}, \quad p, q, r \in \mathbf{N}, \quad ABC \neq 0,$$

has finitely many proper solutions if $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} < 1$ (see [DG]). The proof given in [DG] is based on a descent argument using unramified coverings of $\mathbf{P}_1 - \{0, 1, \infty\}$ of *signature* (p, q, r). (An algebraic covering map $X \longrightarrow \mathbf{P}_1$ is said to be of signature (p, q, r) if it is Galois, is unramified outside of 0, 1, and ∞, and if the ramification indices above these three points are p, q and r respectively.) The traditional descent methods used to attack Fermat's last theorem, based on factorizing $x^p - y^p$ over $\mathbf{Q}(\mu_p)$, can be viewed geometrically as exploiting a covering of signature (p, p, p) with solvable Galois group, corresponding to an unramified covering of the Fermat curve. Frey's beautiful idea exploits the covering $X(2p) \longrightarrow X(2)$ which is ramified over the 3 cusps of $X(2)$ and is of signature (p, p, p). The proof in this paper relied on the modular coverings $X(k, p) \longrightarrow X_0(k)$, where $X(k, p)$ is the modular curve classifying elliptic curves with a subgroup of order k and full level p structure; when $k = 2, 3$, this covering is of signature (k, p, p).

The coverings of signature (p, q, r) arising from modular curves (i.e., as pullbacks of the covering $X(p) \longrightarrow X(1)$) can be classified; the possible signatures are $(2, 3, p)$, $(2, p, p)$, $(3, p, p)$, $(3, 3, p)$ and (p, p, p). Thus, Frey's approach might be used to study which powers can be expressed as sums of two relatively prime cubes, for example, but we have not attempted this.

2. The five triples of exponents listed in remark 1 show the limits of the method used. No family of elliptic curves could be used to study the equation $x^p + y^p = z^5$, for example. One might ask whether one can use other families (say, families of curves of genus 2) to shed light on such equations. Unfortunately, the analogue of Mazur's theorem is not known in this case (are there universal bounds on the torsion in Jacobians of genus 2 curves over $\mathbf{Q}$?), and the results obtained are thus bound to be weaker.

3. To remove the conditions $p \equiv 1 \pmod 4$ for the equation $x^p + y^p = z^2$, or the condition $p \equiv 1 \pmod 3$ for the equation $x^p + y^p = z^3$, it seems one would need to know more about the surjectivity of the Galois representations associated to elliptic curves. For example, it is believed that when p is larger than some explicit universal bound and E is an elliptic curve with no complex multiplications, then the image of the Galois representation in $\mathrm{Aut}\,(E_p)$ is surjective (cf. the discussion in [Sr1], p. 299, §4.3). This would be enough to imply that the equations $x^p + y^p = z^2$ and $x^p + y^p = z^3$ have no non-trivial proper solutions when p is large enough, assuming conjecture $\mathrm{ST}^?$.

More precisely, suppose $p \equiv -1 \pmod 4$ in case 1, or $p \equiv -1 \pmod 3$ in case 2. Let $X_{ns}(k,p)$ be the modular curve which classifies elliptic curves together with a rational subgroup of order k and whose Galois representation on p-division points maps to the normalizer of a non-split Cartan subgroup of $\mathbf{GL}_2(p)$. Then our curve $E_{a,b,c}$ constructed from a non-trivial proper solution to the Fermat-type equation gives rise to a non-cuspidal rational point on $X_{ns}(2,p)$ or $X_{ns}(3,p)$. As N. Elkies has remarked, when $k = 2, 3$, the Jacobian of $X_{ns}(k,p)$ is isogenous to a part of the Jacobian of the curve $X_0(kp^2)/w_{p^2}$, where w_{p^2} is the Atkin-Lehner involution. This latter Jacobian has a p-Eisenstein quotient in its minus part for the Fricke involution w_{kp^2} which is a good candidate for an optimal quotient of the Jacobian of $X_{ns}(k,p)$ having finite Mordell-Weil group. Even if the finiteness of this p-Eisenstein quotient is established, one still has to contend with the presence of other cusps on $X_{ns}(k,p)$ which are rational over $\mathbf{Q}(\mu_p)^+$ and may lie in the way of eliminating possible solutions by an Eisenstein descent argument.

References

[DG] Darmon, H., Granville, A., *On the equations $Ax^p + By^q = Cz^r$ and $z^m = F(x,y)$*, to appear.

[E] Elkies, N., Private communication.

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

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

[LS] Lockhart, P. and Silverman, J., to appear.

[MF] Modular Functions of One Variable, SLN 476, B. Birch and W. Kuyk eds.

[Mz] Mazur, B., *Rational isogenies of prime degree*, inv. Math. 44, 129-162 (1978)

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

[Sr1] Serre, J.-P., *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques*, Inv. Math. 15, 259-331 (1972).

[Sr2] Serre, J.-P., *Sur les représentations modulaires de degré* 2 *de* $Gal(\bar{\mathbf{Q}}/\mathbf{Q})$, Duke Math. J. Vol. 54, no. 1, 179-230 (1987).

[Ta] Tate, J., *Algorithm for determining the type of a singular fiber in an elliptic pencil*, in Modular Functions of One Variable, SLN 476, pp. 33–52.