Proof of the Shafarevich conjecture

Rebecca Bellovin

We have an isogeny of degree ℓ^h $\phi : B_1 \to B_2$ of abelian varieties over K isogenous to A. We wish to show that $h(B_1) = h(B_2)$. By filtering the kernel of ϕ, we may assume that ℓ annhilates the kernel of ϕ, which we call G. We define the following four representations over $\mathbf{Z}/\ell\mathbf{Z}$:

$$\begin{aligned} V_\ell &= T_\ell(B_1)/\ell \cdot T_\ell(B_1) \cong B_1[\ell](\overline{K}) \\ \tilde{V}_\ell &= \mathrm{Ind}_\pi^{\tilde{\pi}}(V_\ell) \\ W_\ell &= G(\overline{K}) \subseteq V_\ell \\ \tilde{W}_\ell &= \mathrm{Ind}_\pi^{\tilde{\pi}}(W_\ell) \subseteq \tilde{V}_\ell \end{aligned}$$

Note that the isogeny $\phi : B_1 \to B_2$ extends to an isogeny $\mathcal{B}_1^0 \to \mathcal{B}_2^0$ between the connected components of their Néron models, so G extends to a finite flat group scheme $\mathcal{G}$ over $\mathcal{O}_K$.

Since W_ℓ is an h-dimensional representation with $\mathbf{Z}/\ell\mathbf{Z}$ coefficients and $[K : \mathbf{Q}] = m$, $\wedge^{mh}(\tilde{W}_\ell) \subseteq \wedge^{mh}(\tilde{V}_\ell)$ is a one-dimensional representation of $\tilde{\pi} = G_{\mathbf{Q}}$. We will call this character $\chi : G_{\mathbf{Q}} \to (\mathbf{Z}/\ell\mathbf{Z})^\times$.

Because our abelian varieties have semistable reduction over K, the G_K-representation V_ℓ is unipotent at places of K away from ℓ, so $\wedge^h$ is unramified away from ℓ. The determinant of the induced representation V_ℓ will pick up inertia, though, so we need to twist by ε^h, where $\varepsilon : G_{\mathbf{Q}} \to \{\pm 1\} \subseteq (\mathbf{Z}/\ell\mathbf{Z})^\times$ is the character arising from the determinant of induction of the trivial representation: $\wedge^m \mathrm{Ind}_\pi^{\tilde{\pi}}(\mathbf{Z})$. This is because $\mathrm{Ind}_\pi^{\tilde{\pi}}(\mathbf{Z}^{\oplus h}) \cong \mathrm{Ind}_\pi^{\tilde{\pi}}(\mathbf{Z})^{\oplus h}$

Now we have a character of $G_{\mathbf{Q}}$, so we can use concrete computations in class field theory to study it. Let L be the finite extension of $\mathbf{Q}$ defined by $\ker(\chi \cdot \varepsilon^h)$. Class field theory (more precisely, the Kronecker-Weber theorem) tells us that every abelian extension of $\mathbf{Q}$ is contained in $\mathbf{Q}^{cyc} := \bigcup_p \cup_n \mathbf{Q}(\zeta_{p^n})$. Each tower $\cup_n \mathbf{Q}(\zeta_{p^n})$ is totally ramified at p and unramified elsewhere, so since

$L/\mathbf{Q}$ is unramified away from ℓ, $L \subset \cup_n \mathbf{Q}(\zeta_{\ell^n})$. By degree considerations ($[L : \mathbf{Q}] | \ell - 1$), $L \subset \mathbf{Q}(\zeta_\ell)$. And since $\mathbf{Q}(\zeta_\ell)/\mathbf{Q}$ is a cyclic extension, this implies that $\chi \cdot \varepsilon^h$ is some power of the cyclotomic character.

We can use Raynaud's results on finite flat group schemes to compute d. In fact, we claim that if $\ell^d = \# s^* \Omega^1_{\mathcal{G}/\mathcal{O}_K}$, then $\chi \cdot \varepsilon^h = \chi_0^d$.

Recall that if G is the generic fiber of a finite flat group scheme killed by p over a strictly henselian local ring R of mixed characteristic $(0, p)$, of low ramification, then the Galois action arising from G is $\tau_p^{v(\mathfrak{d}_{\mathcal{G}/R})/\# G}$. Here τ_p is the canonical tame character $\tau : I_t \to \mathbf{F}_p^\times$.

In our case, we start life with a (global) representation ρ of G_K, arising from a finite flat group scheme killed by p over a number ring $\mathcal{O}_K$, and we are interested in the determinant of its induction to $G_{\mathbf{Q}}$.

Note that the restriction of the cyclotomic character $\chi_0 : G_{\mathbf{Q}} \to (\mathbf{Z}/\ell\mathbf{Z})^\times$ to the inertia subgroup I_ℓ of $G_{\mathbf{Q}_\ell}$ is τ_ℓ, so to compute d it suffices to look at $\chi \cdot \varepsilon^h|_{I_\ell}$.

We started with a representation W_ℓ arising from the generic fiber of a finite flat group scheme $\mathcal{G}$ over $\mathcal{O}_K$. Inducing W_ℓ to $G_{\mathbf{Q}}$ corresponds to taking the Weil restriction $\mathcal{G}' := \operatorname{Res}_{\mathcal{O}_K/\mathbf{Z}} \mathcal{G}$, and restricting to I_ℓ corresponds to basechanging $\mathcal{G}$ to R, the strict henselization of $\mathbf{Z}_\ell$. The character ε dies upon being restricted to I_ℓ because $K/\mathbf{Q}$ is unramified at ℓ, so we will not have to worry about it. In other words, the character $\chi \cdot \varepsilon^h$ we are interested in is the determinant of the representation on the generic fiber of $\mathcal{G}'_R = \operatorname{Res}_{R^{\oplus m}/R}(\mathcal{G} \otimes_{\mathcal{O}_K} R^{\oplus m})$. We can write $\mathcal{G} \otimes_{\mathcal{O}_K} R^{\oplus m}$ as $\coprod_{i=1}^m \mathcal{G}_i$, where each $\mathcal{G}_i$ is a finite flat group scheme over R, each one corresponding to an embedding of $\mathcal{O}_K$ in R. Then $\mathcal{G}'_R = \mathcal{G}_1 \times_R \cdots \times_R \mathcal{G}_m$.

Raynaud's results then tell us that the determinant character is $\tau_\ell^{v(\mathfrak{d}_{\mathcal{G}'_R/R})/\ell^{mh}}$, so we need to compute $v(\mathfrak{d}_{\mathcal{G}'_R/R})/\ell^{mh}$ and compare it with $\# s^* \Omega^1_{\mathcal{G}/\mathcal{O}_K}$. Recall that the quantity $v(\mathfrak{d}_{\mathcal{G}/R})/\#\mathcal{G}$ is additive in exact sequences. Thus,

$$v(\mathfrak{d}_{\mathcal{G}'_R/R})/\ell^{mh} = \sum v(\mathfrak{d}_{\mathcal{G}_i/R})/\ell^h$$

On the other hand, we can use Brandon's corollary 2.6 to see that $|s^* \Omega^1_{\mathcal{G}/\mathcal{O}_K}| = \prod_{v|\ell} \# s^* \Omega^1_{\mathcal{G}_v/\mathcal{O}_{K,v}}$, and we can use corollary 2.10 to see that $|s^* \Omega^1_{\mathcal{G}_v/\mathcal{O}_{K,v}}|^{\#\mathcal{G}_v} =$

$|\mathcal{O}_{K,v}/\,\mathfrak{d}(\mathcal{G}_v/\mathcal{O}_{K,v})|$. Thus, if $|s^*\Omega^1_{\mathcal{G}/\mathcal{O}_K}| = \ell^d$,

$$d = \sum_v f_v v(\mathfrak{d}(\mathcal{G}_v/\mathcal{O}_{K,v}))/\#\mathcal{G}_v$$

But the valuation of the discriminant is preserved under unramified base change, so by going up to the strict henselisation R of $\mathbf{Z}_\ell$, we get

$$d = \sum_i v(\mathfrak{d}(\mathcal{G}_v/\mathcal{O}_{K,v}))$$

as desired.

Now we know that $\chi \cdot \varepsilon^h = \chi_0^d$, so $\chi_0^d(\mathrm{Frob}_p) = \pm p^d$ is a zero of $P_{mh}(T)$ modulo ℓ. But we chose N so large that no $\ell > N$ divides $P_i(\pm p^j)$ for

$$\begin{aligned} 0 \leq i &\leq 2gm \\ 0 \leq j &\leq gm \\ j &\neq \frac{1}{2}i \end{aligned}$$

so we must have $d = mh/2$.

Finally, Faltings's lemma 5 implies that

$$\begin{aligned} h(B_2) &= h(B_1) + \frac{1}{2}\log(\deg\phi) - \frac{1}{[K:\mathbf{Q}]}\log(\#s^*(\Omega^1_{\mathcal{G}/\mathcal{O}_K})) \\ &= h(B_1) + \frac{1}{2}h \cdot \log(\ell) - \frac{1}{m}d \cdot \log(\ell) \\ &= h(B_1) \end{aligned}$$

Now that we have proved the Shafarevich conjecture, we record the following important corollary.

Corollary 0.1. *Fix a number field K, a set of places S of K, and an integer g. Then there are finitely many isomorphism classes of complete curves C of genus g with good reduction outside of S.*

Proof. For such a curve C, we claim that the Jacobian of C is a principally polarized abelian variety with good reduction outside of S. Let v be a place of K not in S, and let $\mathcal{C}/\mathcal{O}_{K,v}$ be a smooth model of C. Then $\mathrm{Pic}_{\mathcal{C}/\mathcal{O}_{K,v}}$ is an abelian scheme over $\mathcal{O}_{K,v}$ whose generic fiber is the Jacobian of C. Thus, $\mathrm{Jac}(C)$ has good reduction outside S.

It is a classical fact that Jacobians of curves over an algebraically closed field are canonically principally polarized. The same fact for curves over more general bases is Proposition 6.9 in Mumford's GIT.

Now Torelli's theorem says that a curve over a perfect field is determined (up to isomorphism) by the isomorphism class of its Jacobian as a principally polarized abelian variety (together with the polarization!). Thus, we have an injective map

$$\{\text{Genus } g \text{ curves over } K \text{ with good reduction outside } S\} \to \{\text{Principally polarized abelian vari}$$

We would like to say that the right-hand side is finite, so that the left-hand side is finite as well. The Shafarevich conjecture implies that there are only finitely many abelian varieties with good reduction outside of S admitting a principal polarization, but we still need to know there are only finitely many pairs (A, ϕ), where A is such an abelian variety and ϕ is a principal polarization. This is provided by theorem 3.1 of Lecture 14 [ed: check this reference]. □

1 Parshin's Trick

Our strategy for deducing the Mordell conjecture from the Shafarevich conjecture is known as Parshin's trick. We fix a smooth projective curve X over a number field K, with good reduction outside S. We will find a finite extension K_1/K and a finite set of places S_1 of K_1. Then to each K-rational point x of X we will associate a smooth projective curve $Y(x)$ over K_1 with good reduction outside of S_1, together with a morphism $Y(x) \to X_{K_1}$ of bounded degree, branched only over x.

First of all, there is a non-trivial étale cover $p : X_1 \to X$ of degree $m > 2$, where X_1 has good reduction outside of S. This can be seen by embedding $X \hookrightarrow \mathrm{Jac}(X)$ (we assume X has a K-rational point; otherwise, we are done) and pulling back by $[2] : \mathrm{Jac}(X) \to \mathrm{Jac}(X)$ (we add the places over 2 to S). After making a finite extension of K, we may assume that the 2-torsion of $\mathrm{Jac}(X)$ is split, so that X_1/X is Galois.

Next, note that for any $x \in X(K)$, the fiber $f^{-1}(x)$ is split over a finite extension of K of degree at most m and unramified outside S. Take the compositum of all such fields and call it K_1. By Hermite's theorem, $[K_1 : K] < \infty$.

To construct $Y(x)$, we fix some $y \in X_1(K_1)$ in the fiber over x and let $D = p^{-1}(x) - y$, a divisor of degree $m-1$. Then let A/K_1 be the generalized Jacobian of (X_1, D). That is, it is Pic^0 of the curve obtained from X_1 by scrunching D to a single K_1-rational point. Using y, we can embed $X_1 - D$ in A and pull back by multiplication by 2 on A to get an étale morphism $Y'(x) \to X_1 - D$. Then there is a smooth proper curve $Y(x) \to X_1$, étale over X_1 away from D.

Now we have a smooth curve $Y(x)$ over K_1, with a map down to X_1 such that the map down to X_{K_1} is branched only at x. On the other hand, if S_1 is a set of places of K_1 (including the places over 2) such that X and X_1 have proper smooth models over $\mathcal{O}_{K_1,S_1}$ and the morphism between them extends to an étale morphism, we can run the same construction using the generalized Jacobian for a model $\mathcal{X}_1$ of X_1 over $\mathcal{O}_{K_1,S_1}$ to get a curve $\mathcal{Y}'(x)$ over $\mathcal{O}_{K_1,S_1}$ which has an étale map down to $\mathcal{X}_1 \setminus D$.

Proposition 1.1. *The multiplication by* 2 *map in the generalized Jacobian* $[2] : \mathrm{Pic}^0_{X_1'/R} \to \mathrm{Pic}^0_{X_1'/R}$ *is finite.*

Proof. There is a lemma of Deligne and Rapoport [ed. insert actual reference] which says that to prove finiteness of a flat quasi-finite morphism, it is enough to check that it has constant fibral degree.

In our case, note that we constructed the curve $\mathcal{C}$ by scrunching a horizontal divisor of degree $m-1$, and our divisor is actually split. Thus, by [ref Néron Models], $\mathrm{Pic}^0_{\mathcal{C}/\mathcal{O}_{K_1,S_1}}$ on both the generic fiber and the special fiber is an extension of an abelian variety of by an affine piece; the scrunched curve is semistable, so there is no unipotent piece. The dimension of the abelian variety and the rank of the affine piece depend only on the divisor being scrunched.

Given a commutative algebraic group G over a perfect field, if we write the Chevalley decomposition

$$0 \to T \times U \to G \to A \to 0$$

then multiplication by 2 on G respects this decomposition. Then it is easy to see that $[2] : G \to G$ is a morphism of degree $2^{\dim T} \cdot 2^{\dim A}$.

In our case, since the pieces of $\mathrm{Pic}_{\mathcal{C}/\mathcal{O}_{K_1,S_1}}$ have the same dimension for each fiber, we are done. □

Now that we know $\mathcal{Y}'(x)$ is finite étale over $\mathcal{X}_1 - \mathcal{D}$, we know that $\mathcal{Y}'(x)$ is open in its normalization $\mathcal{Y}$ over all of $\mathcal{X}_1$. Then any bad points of $\mathcal{Y}$ are codimension 2 and happen over $\mathcal{D}$.

Proposition 1.2. *For $\mathcal{X}_1$ and $\mathcal{Y}$ as above, $\mathcal{Y}$ is smooth over $\mathcal{O}_{K_1,S_1}$*

Proof. First we note that we may assume that $\mathcal{Y}$ is an abelian Galois cover of $\mathcal{X}_1$. This is because the 2-torsion of Pic is split over a base field extension of degree at most $2^{\dim T} \cdot 2^{\dim A}$. Furthermore, this extension is unramified, because $2 \not| v$. There are only finitely many such extensions of K_1, so we can retroactively build them into K_1 without affecting our finiteness statements.

Let x_1 be a closed point of $\mathcal{D}$, and let S_{x_1} be the local ring of $\mathcal{X}_1$ at that point. Then S_{x_1} is a 2-dimensional regular local ring, and we can pull back the entire set-up to S_{x_1}. We write $\mathcal{Y}_{r_{x_1}} = \operatorname{Spec} T$, and since normalization commutes with localization, we are in a position to apply Abhyankar's lemma.

Abhyankar's lemma (as given in Freitag-Kiehl, A I.11) tells us that if $\mathfrak{m}$ is a maximal ideal of T (corresponding to a putatively bad point of $\mathcal{Y}$) and t is a local parameter for x_1 on the base, then $T_{\mathfrak{m}}^{sh} = S_{x_1}^{sh}[\sqrt[e]{t}]$. Here e is the ramification degree. In our case, the ramification group must be a 2-group, hence prime to the residue characteristic. But then we can simply compute the module of relative differentials, and since smoothness at a point can be checked after passing to the strict henselizations of the local rings, we see that $\mathcal{Y}$ is smooth over $\mathcal{O}_{K_1,S_1}$. □

The upshot is that for each rational point x of X, we have constructed a smooth projective curve $Y(x)$ over X_{K_1} with good reduction outside S_1 (not depending on x), and the morphism $Y(x) \to X_{K_1}$ branched only over x. Moreover, the degree of this morphism is at most $2^{g(X_1)} \cdot m$, so the genus of $Y(x)$ is bounded. By Corollary 0.1, there are only finitely many isomorphism classes of such curves $Y(x)$, so it only remains to show that for specified curves X and Y of genus at least 2, there are only finitely many non-constant morphisms $Y \to X_{K_1}$. This is known as de Franchis's theorem, and we give a proof based on the Hilbert scheme.

Lemma 1.3. *Let X and Y be projective curves over a number field K, and let $C \subset X \times_K Y$ be the graph of a morphism $Y \to X$. Then the Hilbert polynomial of C is*

Proof. Let $\mathcal{L}_X, \mathcal{L}_Y$ be very ample line bundles embedding X and Y in projective space, respectively. Then $pr_1^*\mathcal{L}_X \otimes pr_2^*\mathcal{L}_Y$ gives an embedding of $X \times_K Y$ in projective space, and its restriction $\mathcal{L} := pr_1^*\mathcal{L}_X \otimes pr_2^*\mathcal{L}_Y|_C$ to C embeds C in projective space.

To compute the Hilbert polynomial of C as a closed subscheme of $X \times_K Y$, we need to know $\chi(n\mathcal{L})$. But Riemann-Roch tells us that $\chi(n\mathcal{L}) = n \cdot \deg(\mathcal{L}) - g + 1$, where g is the genus of C, which is the same as the genus of Y, so really we only need to compute the degree of $\mathcal{L}$. This is not difficult:

$$\deg(\mathcal{L}) = \deg(pr_1^*\mathcal{L}_X|_C) + \deg(pr_2^*\mathcal{L}_Y|_C) = d \cdot \deg \mathcal{L}_X + \deg(\mathcal{L}_Y)$$

The last equality follows because $pr_1|_C : C \to X$ is a degree d morphism and $pr_2|_C : C \to Y$ is a degree 1 morphism. □

Thus, the space of degree d morphisms $Y \to X_{K_1}$ is a particular component of the Hilbert scheme of closed subschemes of $X_{K_1} \times_{K_1} Y$, and therefore finite type over K_1. To conclude that there are actually only finitely many morphisms $Y \to X_{K_1}$, it suffices to show that the tangent space at a point of this space is zero-dimensional. Here is where we will use that the genus of X is at least 2.

Proposition 1.4. *For fixed projective curves X and Y, the space of degree d morphisms $Y \to X$ is zero-dimensional.*

Proof. Let $f : Y \to X$ be such a morphism. The tangent space at f is given by $\mathrm{Def}(f)$, the set of deformations of f to $Y \otimes K[\varepsilon] \to X \otimes K[\varepsilon]$. To give such a morphism f' is the same as giving a morphism $f'^{\#} : \mathcal{O}_X[\varepsilon] \to (f_*\mathcal{O}_Y)[\varepsilon]$ which agrees with $f^{\#}$ modulo ε. This, in turn, is the same as giving an $\mathcal{O}_X$ derivation $\mathcal{O}_X \to f_*\mathcal{O}_Y$, so we have

$$\begin{aligned} \mathrm{Def}(f) &= \mathrm{Der}_{\mathcal{O}_X}(\mathcal{O}_X, f_*\mathcal{O}_Y) = \mathrm{Hom}_{\mathcal{O}_X}(\Omega^1_{X/K}, f_*\mathcal{O}_Y) \\ &= H^0(X, T_{X/K} \otimes f_*\mathcal{O}_Y) = H^0(Y, f^*T_{X/K} \end{aligned}$$

Since X has genus at least 2, its tangent bundle has negative degree and there are no sections. □