# ON THE TORSION POINTS OF ELLIPTIC CURVES & MODULAR ABELIAN VARIETIES

SETH KLEINERMAN

Contents

## 1. Introduction and Background

Age cannot wither her, nor custom stale
Her infinite variety.
—*Antony and Cleopatra*, 2.2.271-2.

The study of abelian varieties is of great interest in modern number theory. In this paper we will try to understand some of the basics of the varieties' torsion subgroups. That they are finite at all is a matter of some concern, and that will be the main thrust of the theory we develop. We will often specialize to results on elliptic curves, which are better understood.

1.1. **Elliptic Curve Definitions.** An elliptic curve $E$ over a field $K$ is the locus of points $(x, y)$ satisfying the *Weierstrass equation*

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

with coefficients $a_i \in K$. A notion of summation of points turns this set into an abelian group. We won't give the group law algorithm here explicitly, but refer the reader to [17], Ch. III §2. The motivation for the composition law is that Bezout's theorem indicates that a line intersects this degree three curve in exactly three points (counting multiplicity in the special case of tangency). Therefore there is a unique third point $R$ on both the curve and the line connecting any given points $P$ and $Q$. Also, given a base point $\mathcal{O}$, there is a unique point $R'$ on the curve that is on the line connecting $\mathcal{O}$ and $R$. Define $R' = P \oplus Q$. This binary operation gives $E$ the structure of an abelian group. (All the group laws are straightforward to observe except associativity, which one can demonstrate in a technical argument with the explicit equations for the curve and lines, but which we happily avoid here.)

This characterization is already an abuse of notation, since technically the curve ought to be in projective space for the discussion above to make any sense at all. However, the affine curve has a unique, well-defined projective closure, so rigor is not lost. This aside is by way of excusing the fact that sometimes we'll be lax and refer to the point $(x, y)$ on a curve instead of the point $[X : Y : Z]$ on the projectivized curve.

To further simplify things for us, if the field is not of characteristic 2 or 3, we can make a rather messy substitution

$$x \longmapsto \frac{1}{36}(x - 3a_1^2 - 12a_2) \tag{1.1}$$

$$y \longmapsto \frac{1}{216}(y - a_1x - a_3) \tag{1.2}$$

to reduce the equation to the form $y^2 = x^3 + Ax + B$, where $A$ and $B$ hide the ugly expressions

$$\begin{aligned} A =& -27a_1^4 - 216a_1^2a_2 - 432a_2^2 + 648a_1a_3 + 1296a_4, \\ B =& 54a_1^6 + 648a_1^4a_2 + 2592a_1^2a_2^2 - 1944a_1^3a_3 + 3456a_2^3 - 3888a_1^2a_4 - 7776a_1a_2a_3 \\ &+ 11664a_3^2 - 15552a_2a_4 + 46656a_6. \end{aligned}$$

When we need an explicit Weierstrass equation, we'll often use the simplified version for convenience, keeping in mind that the arguments would need to be modified to fit the full six-coefficient Weierstrass form if we are going to be able to apply them to curves over fields with characteristic 2 or 3. In the simplified case, however,

define the *discriminant* of the Weierstrass equation to be $\Delta = -16(4A^3 + 27B^2)$. The general form of the discriminant is not enlightening for our purposes, but note that it is not just $\Delta$ with the reverse substitutions for $A$ and $B$ (because then it would always be a multiple of 16). The discriminant of an elliptic curve is 0 if and only if it is singular. The group law still applies to the non-singular points of a given singular curve, however.

Finally, we remark that elliptic curves are abelian varieties of genus 1.

## 2. Formal Groups

2.1. **The Basics.** The following machinery is given in Silverman [17], chapter IV. We try to be a little clearer in the explanations, sometimes filling in arguments that he elides. This is evident particularly when it comes to facts about the inverse map, which are assumed in the source without any mention of what exactly is being assumed. But they aren't, in fact, terribly tedious or difficult to prove, and so we prove them here.

**Definition 2.1.** A *formal group* $\mathcal{F}$ over a ring $R$ is a two-variable power series $F(X, Y) \in R[[X, Y]]$ satisfying the following five properties:

(1) $F(X, Y) = X + Y +$ [terms of higher degree].
(2) $F(X, F(Y, Z)) = F(F(X, Y), Z)$.
(3) $F(X, Y) = F(Y, X)$.
(4) $F(T, i(T)) = 0$ has a unique solution for $i(T) \in R[[T]]$.
(5) $F(X, 0) = X$, $F(0, Y) = Y$.

Observe that these five properties provide an abelian group structure without any explicit elements. (The second, third, and fourth properties are associativity, commutativity, and inverses, respectively, and the fifth provides an identity.) Instead of elements, there is only the power series $F$ which provides the recipe for the group operation: hence the name "formal group."

**Definition 2.2.** A *homomorphism* between two formal groups $f : \mathcal{F} \to \mathcal{G}$ defined over a ring $R$ is a one-variable power series $f(T) \in R[[T]]$ with no constant term such that $f(F(X, Y)) = G(f(X), f(Y))$. The formal groups $\mathcal{F}$ and $\mathcal{G}$ are *isomorphic* if there exists a homomorphism $f$ as above and a reverse homomorphism $g$ such that $f(g(T)) = g(f(T)) = T$.

For example, define the map $[0] : \mathcal{F} \to \mathcal{F}$ by $[0](T) = 0$. It is a homomorphism because

$$[0](F(X, Y)) = 0 = F(0, 0) = F([0](X), [0](Y)).$$

That is an entirely trivial example, but it is important: we can use it to define a map $[m] : \mathcal{F} \to \mathcal{F}$ inductively for any integer $m$ by

$$[m](T) = F([m - 1](T), T) \qquad (m > 0)$$
$$[m](T) = F([m + 1](T), i(T)) \qquad (m < 0)$$

We'll show briefly that these maps, which are (perhaps confusingly at first) called *multiplication by m*, are all homomorphisms (we'll say that the case $m > 0$ will do to show this; the other case is similar).

*Proof.* Assume that $[m-1]$ is known to be a homomorphism. Then

$$\begin{aligned}
[m](F(X,Y)) &= F([m-1](F(X,Y)), F(X,Y)) \\
&= F(F([m-1]X, [m-1]Y), F(X,Y)) \\
&= F([m-1]X, F([m-1]Y, F(X,Y))) \\
&= F([m-1]X, F(F([m-1]Y, Y), X)) \\
&= F([m-1]X, F([m]Y, X)) \\
&= F(F([m-1]X, X), [m]Y) \\
&= F([m]X, [m]Y)
\end{aligned}$$

where we've used the commutativity and associativity properties of the formal group a few times. In particular, $[m]$ is a homomorphism as we'd wanted.

We should also note that in the downward induction we need to use the fact that $i : \mathcal{F} \to \mathcal{F}$, the unique inverse map associated to the formal group $\mathcal{F}$, is also a homomorphism. To see this, consider the expression

$$F(F(X, iX), F(Y, iY)) = F(F(X,Y), F(iX, iY))$$

obtained by repeated application of commutativity and associativity of the formal power series. The left hand side is trivially 0. On the right, remember that $iF(X,Y)$ is the unique power series for which $F(F(X,Y), iF(X,Y)) = 0$. So $iF(X,Y) = F(iX, iY)$ and $i : \mathcal{F} \to \mathcal{F}$ is a homomorphism. The second induction therefore follows exactly as the first did. □

We will now prove two facts about the multiplication-by-$m$ map.

**Proposition 2.3.** For any integer $m$, the map $[m](T) = mT +$ [higher order terms].

*Proof.* We first induce to obtain the result for non-negative $m$. It is clear that $[0](T) = 0$ works. Now assume the result about the map $[m-1]$. Using the recursive definition and then property 1 of formal groups, we have

$$\begin{aligned}
[m](T) &= F([m-1](T), T) \\
&= [m-1](T) + T + [\text{higher order terms}] \\
&= (m-1)T + T \mod T^2
\end{aligned}$$

which is exactly what we needed.

For the downward induction, we'll need to establish a minor result about $i(T)$ again, namely that it has no constant term. This follows from property 5 of formal power series:

$$i(0) = F(0, i(0)) = F(T, i(T))|_{T=0} = 0.$$

Now, notice that

$$0 = F(T, i(T)) = T + i(T) + \dots$$

where the $[\dots]$ are higher order combinations of $T$ and $i(T)$. Rearrange to find

$$i(T) = -T + \dots$$

We showed that $i(T)$ has no constant term, so everything on the right hand side after the $-T$ vanishes modulo $T^2$. Now the downward induction is clear. □

For the second proposition, we have to first prove a standard result about formal power series.

**Lemma 2.4.** *Let $R$ be a ring and $f(T) \in R[[T]]$ a one-variable power series, with $f(T) = aT +$ [higher order terms]. If $a \in R^*$, then there is a unique power series $g(T) \in R[[T]]$ that composes with $f(T)$ so that $f(g(T)) = T$. This power series happens to also satisfy $g(f(T)) = T$.*

*Proof.* The goal is to produce a power series that satisfies $f(g(T)) \equiv T$ modulo $T^n$ for any $n > 1$. We'll do this by explicitly constructing the polynomials $g_n(T)$ which we define to be $g(T)$ modulo $T^{n+1}$; since they'll be inductively well-defined, that will give us a precise definition of the formal power series $g(T)$.

First, take $g_1(T) = a^{-1}T$. Clearly $f(g_1(T)) = a(a^{-1}T) + \ldots \equiv T \mod T^2$, as we'd wanted.

Next, assume $f(g_{k-1}(T)) \equiv T \mod T^k$. We'll have $g_k = g_{k-1} + rT^k$ for some $r \in R$. Observe that

$$\begin{aligned} f(g_k(T)) &= f(g_{k-1}(T) + rT^k) \\ &\equiv f(g_{k-1}(T)) + arT^k \mod T^{k+1} \end{aligned}$$

since all cross-terms contributed by the $rT^k$ monomial have degree $> k$. Remember by our initial assumption that

$$\begin{aligned} f(g_{k-1}(T)) &\equiv T \mod T^k \\ &\equiv T + bT^k \mod T^{k+1} \end{aligned}$$

for some element $b \in R$. Then put it all together to get

$$f(g_k(T)) \equiv T + bT^k + arT^k \mod T^{k+1}.$$

So there's a unique choice $r = -b/a$ at each stage, well-defined because $a$ is a unit, such that $f(g_k(T)) \equiv T \mod T^{k+1}$, and the formal power series $g(T)$ therefore exists.

Now that we know $f(g(T)) = T$, we also have $g(f(g(T)) = g(T)$, which is a formal identity in $R[[g(T)]]$. In particular, under the formal map of variables $g(T) \mapsto T$, we find that $g(f(T)) = T$ as we'd claimed.

Finally, to show uniqueness of $g$, assume there's another power series $h(T)$ such that $(f \circ h)(T) = T$; then

$$g(T) = g \circ (f \circ h)(T) = (g \circ f) \circ h(T) = h(T).$$

□

Now we can prove the second fact, which quite naturally follows from the above.

**Proposition 2.5.** If the integer $m$ is a unit in $R^*$, then $[m] : \mathcal{F} \to \mathcal{F}$ is an isomorphism.

*Proof.* ¿From Proposition 2.3, $[m](T) = mT + \ldots$; knowing that $m \in R^*$ and applying Lemma 2.4, there is a power series $[m]^{-1}(T)$ such that $[m]^{-1}([m](T)) = [m]([m]^{-1}(T)) = T$. The only thing remaining is to show that $[m]^{-1}$ is a homomorphism $\mathcal{F} \to \mathcal{F}$. We know $[m]$ is a homomorphism, so

$$\begin{aligned} [m]^{-1}F(X,Y) &= [m]^{-1}F([m][m]^{-1}(X), [m][m]^{-1}(Y)) \\ &= [m]^{-1}[m]F([m]^{-1}(X), [m]^{-1}(Y)) \\ &= F([m]^{-1}(X), [m]^{-1}(Y)) \end{aligned}$$

gives us what we need. □

2.2. **So where are the elements?** The natural thing to do after defining a formal group is to think about supplying elements that can follow this recipe for their group structure.

Recall (Atiyah-Macdonald [1], ch. 10) that a complete local ring $R$ with maximal ideal $M$ is defined as a ring that is equal to its natural inverse limit:

$$R \cong \varprojlim R/M^n.$$

If we have a power series $f$ and an element $\alpha \in M$, we can use these to produce a unique element $\alpha_n \in M^n$, since truncating the power series at the degree $n$ term is fine (all higher terms will be congruent to zero modulo $M^n$). The completeness of the ring indicates that this sequence of elements $(a_1, \ldots, a_n, \ldots)$ is under isomorphism an element of $R$. In particular, this argument extends to show that with $R, M$ as above, a power series in two variables $F(X, Y)$ given as the group law for a formal group will converge for $X, Y \in M$. The set $M$ is thus made into a group, denoted $\mathcal{F}(M)$, with the group composition $x \oplus y = F(x, y)$ and inverse $\ominus x = i(x)$. The closure of the group is shown by property 1 of formal power series (Definition 2.1): the equation tells us that $x \oplus y \equiv 0$ modulo $M$. Finally, $\mathcal{F}(M^n)$ is a subgroup of $\mathcal{F}(M)$, since it is closed under composition by precisely the same argument.

*Remark* 2.6. Now we see exactly why the multiplication-by-$m$ map was so called; the map on the formal group induces a map on the associated group $\mathcal{F}(M)$ which acts by composing an element with itself $m$ times.

We now prove a major result about torsion in groups associated to formal groups.

**Proposition 2.7.** Let $\mathcal{F}$ be a formal group over a complete local ring $R$ with maximal ideal $M$, and let $k$ be the residue field $R/M$. Let the characteristic of $k$ be $p$. (Zero characteristic is allowed.) Then every element of finite order in $\mathcal{F}(M)$ has order a power of $p$.

*Proof.* For positive $p$, if there were a point $r \in M$ that served as a counterexample, it would have order $p^k m$ for some $m$ relatively prime to $p$. There would then also be a point of order $m$, namely $[p^k]r$. Therefore it is enough to prove that there are no points of order $m$ with $p$ and $m$ relatively prime. (For $p = 0$, we have to prove that there are no points of order $m > 0$ at all.)

Then let $m$ be a positive integer chosen as above, and let $x$ be a point of order $m$, so $[m](x) = 0$. Observe that $m$ cannot be an element of the ideal $M$, because then $R/M$ would have characteristic dividing $m$, which is not permitted. Therefore $m \in R^*$, and applying Proposition 2.5, the map $[m] : \mathcal{F} \to \mathcal{F}$ is an isomorphism. That's all we need, because it induces an isomorphism of groups $[m] : \mathcal{F}(M) \to \mathcal{F}(M)$, whose kernel consists only of the point 0, so $x$ must be trivial. □

2.3. **More Theory of Formal Groups.** If the ring is equipped with a discrete valuation $v$, we can say something further about the torsion in the group $F(M)$.

**Proposition 2.8.** Let $R$ be a complete local ring with maximal ideal $M$, and with a discrete valuation $v$. Let $p$ be the characteristic of the field $R/M$. Let $\mathcal{F}$ be a formal group over $R$, with associated group $\mathcal{F}(M)$. Given an element $x \in \mathcal{F}(M)$ of prime power order $p^n$, we have

$$v(x) \leq \frac{v(p)}{p^n - p^{n-1}}.$$

This proposition relies on the notion of invariant differentials for formal groups, the introduction of which may take us a bit too far afield. It is proved quite concisely in [17], ch. IV, §4 and §6.

## 3. Application to Elliptic Curves: the Finitude of $E(K)_{\text{TOR}}$

In this section we will demonstrate that torsion subgroups of elliptic curves over number fields are finite, as a consequence of a demonstrably injective map into a finite group. This result also follows from the Mordell-Weil theorem (which states that the group $E(K)$ of an elliptic curve over a number field is finitely generated). However, the argument that we will present here is independent and easier to prove. Additionally, it gives an effective bound (since one can compute the number of points in the reduction of the curve easily), where the Mordell-Weil theorem does not.

3.1. **An Elliptic Curve's Formal Group.** In what follows, we will use the simplified form of the elliptic curve, so the full generality of characteristic 2 and 3 is denied us; however, [17], IV §1, does that rather well. In choosing to work with the simplified curve, we hope to trade full generality for some added intuition, as the algebraic contortions with the five-parameter form are not particularly enlightening. The following section is mostly sketches to get a grasp of what's going on; the formulas themselves will be elided most of the time.

Given an elliptic curve $E : y^2 = x^3 + Ax + B$, perform the change of coordinates $s = \frac{1}{y}$, $t = \frac{x}{y}$. The additive identity becomes $(0, 0)$ under the new $(s, t)$-coordinate system. The equation of the curve becomes

$$s = t^3 + Ats^2 + Bs^3.$$

What if we want $s$ as a function entirely of $t$? The natural thing to do is recursively substitute the equation for $s$ back into itself, as follows:

$$\begin{aligned}
s &= t^3 + Ats^2 + Bs^3 \\
&= t^3 + At(t^3 + Ats^2 + Bs^3)^2 + B(t^3 + Ats^2 + Bs^3)^3 \\
&= t^3 + At^7 + Bt^9 + s^2(2A^2t^5 + 3ABt^7) + s^3(2ABt^4 + 3B^2t^6) \\
&\quad + s^4(A^3t^3 + 3A^2Bt^5) + s^5(2A^2Bt^2 + 6AB^2t^4) \\
&\quad + s^6(AB^2t + 3B^3t^3 + A^3Bt^3) + s^7(A^2B^2t^2) + s^8(3AB^3t) + s^9(B^4).
\end{aligned}$$

This repeated substitution gets us a polynomial in $t$ followed by a polynomial in $s$ and $t$; by continuing inductively, the powers of $s$ get larger and larger, and thus contribute larger and larger powers of $t$, so the procedure feels like it should converge to an element of the ring of formal power series $R[[t]]$. In fact, it does; this is a consequence of Hensel's lemma (see Silverman [17], IV Lemma 1.2.).

In particular, there is a unique formal power series $s(t) \in R[[t]]$ such that $s(t) = t^3 + ats(t)^2 + Bs(t)^3$. Letting the expression $f(s, t) = t^3 + Ats^2 + Bs^3$, then $s(t)$ satisfies $f(s(t), t) = s(t)$. We know the first few terms as before, after a few self-substitutions:

$$\begin{aligned}
s(t) = t^3 + At^7 + Bt^9 + 2A^2t^{11} + 5ABt^{13} \\
+ (5A^3 + 3B^2)t^{15} + 21A^2Bt^{17} + \dots
\end{aligned} \tag{3.1}$$

Now substituting back for the standard coordinates, we have

$$x(t) = \frac{t}{s(t)} = \frac{1}{t^2} - At^2 - Bt^4 - A^2t^6 - 3ABt^8 + (-2A^3 - 2B^2)t^{10} - 10A^2Bt^{12} \dots$$
$$y(t) = \frac{1}{s(t)} = \frac{1}{t^3} - At - Bt^3 - A^2t^5 - 3ABt^7 + (-2A^3 - 2B^2)t^9 - 10A^2Bt^{11} \dots$$

Because of our discussion above on the convergence of formal power series, if $K$ is a complete local field with $R$ and $M$ as we've been using them above, then taking $t$ to be an element of $M$ gets us a point in $E(K)$. (The fact that these are Laurent series doesn't change anything about the convergence because we're only adding a finite number of terms to a series known to be convergent.) Furthermore, the map is injective, since $t = x/y$ is the inverse. We can think of $t$ as a parameter for all points on the curve with $t = x/y \in M$.

Finally, since there is a group law for the explicit addition of points $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2) \in E$, we can make a corresponding formula for the formal addition of indeterminates $t_1$ and $t_2$, which we eschew here, since we have made it a point to avoid the group law calculations above. However, the addition formula is a formal power series in two variables $F(t_1, t_2)$. It is constructed to obey all the properties that the usual addition on elliptic curves obeys; therefore it is the basis for a formal group.

**Definition 3.1.** The formal group $\hat{E}$ (over $R$) associated to the elliptic curve $E$ (over $K$) is the formal power series $F$ defined above, which gives the group law.

3.2. **Reduction on Elliptic Curves.** Let $K$ be a local field with discrete valuation $v$, and $R$ be the ring of integers of $K$. Recall that since $R = \{x \in K : v(x) \geq 0\}$, we can decompose $R$ into a group of units $R^*$ with valuation 0 and the set of elements with positive valuation; these last constitute an ideal in $R$, which is clearly the unique maximal ideal since it contains all non-units, so we'll call it $M$ as before. (The ring of integers of a local field is thus a local ring.)

Now let $M$ be $\pi R$ for some uniformizing element $\pi$, and for convention's sake let's say that the discrete valuation has been normalized so that $v(\pi) = 1$. Finally, let $k$ be the field $R/M$. In a bit we'll be able to think about moving a curve's base field from $K$ to $k$. But first we have to put the curve in canonical form.

**Definition 3.2.** Given an elliptic curve $E/K$ with Weierstrass equation $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$, we say the equation is *minimal* if $a_i \in R$ and $v(\Delta)$ is minimized over all possible changes of coordinates.

*Remark* 3.3. There is a minimal equation representing any curve, because the valuation $v$ is discrete.

*Remark* 3.4. The explicit change of coordinates allowed is given in [17], ch. III, §1. Observe that we can also minimize a simplified Weierstrass equation $y^2 = x^3 + Ax + B$ if we're not working over a local field with residue characteristic 2 or 3: certainly we can minimize the valuation of the determinant in the five-parameter form, then do the transformations 1.1 and 1.2 in order to put it in simplified form. The determinant changes by a factor of $1/6^{12}$, so it remains minimal with our valuation. Therefore in what follows we'll consider elliptic curves of the form $y^2 = x^3 + Ax + B$, with $A, B \in \mathbb{Z}$.

**Definition 3.5.** The natural *reduction map* is the map $R \to R/M$, which action on the elements of $R$ we denote by $r \mapsto \tilde{r}$. The reduction of an elliptic curve over a local field $K$ is the curve produced by the action of the reduction map on a minimal Weierstrass equation for $E$:

$$[E : y^2 = x^3 + Ax + B] \longmapsto [\tilde{E} : y^2 = x^3 + \tilde{A}x + \tilde{B}].$$

The reduction on a projective point $P = [x : y : z]$ is the reduction map on the coordinates of the point once it has been put into canonical form, that is, with all coordinates in the ring of integers $R$ and minimized so that one or more is a unit. This is required for the image of the point to be well-defined. Then $\tilde{P}$ is a point on $\tilde{E}/k$.

Since the curve's equation was chosen to be minimal, the map between curves is well-defined, up to a change of coordinates over the residue field $k$. It is a homomorphism because it maps lines in $\mathbb{P}^2(K)$ to lines in $\mathbb{P}^2(k)$, and therefore the group laws are preserved.

The reduced curve is not guaranteed to be non-singular. If the discriminant of a Weierstrass equation is 0, then its elliptic curve has singular points; however, it is a fact that the remaining non-singular points still form an abelian group. Let $\tilde{E}_{\text{ns}}(k)$ be the group of non-singular points of $\tilde{E}$. Then its inverse image in the set $E(K)$ is denoted $E_0(K)$, the points of non-singular reduction. Also define $E_1(K)$ to be the kernel of reduction, that is, the points that map to the identity on the reduced curve.

**Proposition 3.6.** The sequence

$$0 \to E_1(K) \to E_0(K) \to \tilde{E}_{\text{ns}}(K),$$

where the second map is inclusion and the last map is reduction, is exact.

*Proof.* This follows from the definitions. Since the map $E_1(K) \to E_0(K)$ is inclusion, as observed above, it is injective. Then the kernel of the reduction map is exactly $E_1(K)$ by definition. □

*Remark* 3.7. In fact, the reduction map is surjective (see [17], ch. VII, proposition 2.1—the proof uses Hensel's lemma for the lifting), so we can extend the above into a short exact sequence of groups, but the surjectivity is not essential to what follows.

Finally, we put together the reduction map and the formal group associated to the curve.

**Proposition 3.8.** Let $\hat{E}$ over $R$ be the formal group associated to $E$ over $K$, and let $s(t)$ be the power series as in equation 3.1. Then the map $\hat{E}(M) \to E_1(K)$ given by $t \longmapsto (\frac{t}{s(t)}, \frac{1}{s(t)})$ is an isomorphism. In this map, whose image is not explicitly composed of projective points, we assert that the point $z = 0$ is mapped to the additive identity $(0, \infty)$.

*Proof.* We already know that the image points satisfy the Weierstrass equation, and we have remarked that $t \in M$ makes the defining series converge to a point in $E(K)$ (it is in $E(K)$ and not $E(M)$ because the series is Laurent). To see that it is actually in $E_1(K)$, notice that after clearing denominators of elements of $M$ in the point $[\frac{t}{s(t)} : \frac{1}{s(t)} : 1]$, the first and third coordinates have positive valuation while

the second is a unit, and therefore the point reduces to $[0:1:0]$ on the curve $\tilde{E}(k)$. The explanation in [17] (ch. VII, §2) of this fact is not particularly clear.

It remains only to produce an inverse. We mentioned a putative one in our original remarks, namely $t = x/y$, so it suffices to show that given $(x, y) \in E_1(K)$, $x/y \in M$. But this is evident by inspection: if $[x : y : 1]$ is to reduce to $[0 : 1 : 0]$ as above, we must have $x$ and $y$ with negative valuation, and, further, $v(y) < v(x)$. Therefore $v(x/y) > 0$ and so $x/y \in M$. □

Now we have the material needed to prove the main result.

**Theorem 3.9.** *Let $E$ be an elliptic curve over $\mathbb{Q}$, and $p > 3$ a prime. Assume that $p$ does not divide the discriminant $\Delta$, so the reduced curve $\tilde{E}$ is non-singular. Then the induced map on torsion $E(\mathbb{Q})_{tor} \to \tilde{E}(\mathbb{F}_p)_{tor}$ is injective.*

The theorem is also true for $p = 3$, but false for $p = 2$. We will discuss this further below. This proof works verbatim for $p = 3$.

*Proof.* Consider the extension our curve $E(\mathbb{Q}) \subset E(\mathbb{Q}_p)$ to a local field in which it is embedded. We can now apply all the theory we've been assembling, with $K = \mathbb{Q}_p$. We have the standard $p$-adic valuation, and the maximal ideal is $M = (p)$.

First, with the aid of Proposition 2.8, we show that there cannot be any elements of order $p$ in the formal group $\hat{E}(M)$. Assume an element $x \in M$ has order $p$. Then by Proposition 2.8, we have $0 < v(x) \leq \frac{v(p)}{p-1}$. Since $v$ is the standard $p$-adic valuation on the rational numbers, $v(p) = 1$ and $v(x)$ is an integer. Since $p > 2$, we have $0 < v(x) \leq \frac{1}{p-1} < 1$, a contradiction.

Next, recall that there cannot be any elements of order relatively prime to $p$ in $\hat{E}(M)$, by Proposition 2.7. Hence we conclude that there are no torsion points of any kind in the group $\hat{E}(M)$.

By Proposition 3.8, $\hat{E}(M) \cong E_1(\mathbb{Q}_p)$ and thus there is no torsion in the group $E_1(\mathbb{Q}_p)$.

Finally, we avail ourselves of the exact sequence of Proposition 3.6, which tells us that the group $E_1(\mathbb{Q}_p)$ is injectively mapped into $E_0(\mathbb{Q}_p)$, thus has a torsion-free image under this map, and this image is exactly the kernel of the reduction map into $\tilde{E}_{\text{ns}}(\mathbb{F}_p)$. Since the reduced curve is non-singular, $\tilde{E}(\mathbb{F}_p) = \tilde{E}_{\text{ns}}(\mathbb{F}_p)$ and $E_0(\mathbb{Q}_p) = E(\mathbb{Q}_p)$. Thus we have an exact sequence

$$0 \to \hat{E}(M) \to E(\mathbb{Q}_p) \to \tilde{E}(\mathbb{F}_p).$$

In particular, all the points in $E(\mathbb{Q}_p)$ that reduce to the identity in $E(\mathbb{F}_p)$ are non-torsion. So the torsion in $E(\mathbb{Q}_p)$ and hence in $E(\mathbb{Q})$ must be mapped injectively into $\tilde{E}(\mathbb{F}_p)$.

□

**Corollary 3.10.** The order of a torsion subgroup of an elliptic curve over $\mathbb{Q}$ is finite.

*Proof.* There exist infinitely many primes of good reduction, since any prime that does not divide the discriminant is a prime of good reduction. Choose $p > 3$ and reduce. The reduction map is injective on torsion. □

*Remark* 3.11. A similar argument works on elliptic curves over finite extensions of $\mathbb{Q}$. It may become necessary to choose a higher prime so that $\frac{v(p)}{p-1} < 1$, since the

valuation of $p$ may not be 1 in the corresponding extension of $\mathbb{Q}_p$, but once a prime has been found that works, the torsion subgroup is automatically finite.

We should observe that if we assume the preliminary results for the five-parameter Weierstrass equations (not a big leap), Theorem 3.9 will also hold for $p = 3$ but not $p = 2$, since $v(2) = 1$ does not satisfy the condition $v(p) < p - 1$.

The elliptic curve $E : y^2 + xy + y = x^3 + x^2 - 10x - 10$ is an example of a curve with good reduction at $p = 2$ (since $\Delta = 3^4 5^4$) for which the corresponding reduction map is not injective on torsion. Checking the torsion by computer, we find that it is equal to $\mathbb{Z}/8\mathbb{Z}$; however, there are only 5 points that could possibly be in $E(\mathbb{F}_2)$. This is not a contradiction, however, because of the above observation.

To be explicit, we check the torsion over $\mathbb{F}_2$ as well, using the computer algebra package MAGMA:

```
> E := EllipticCurve([ 1, 1, 1, -10, -10 ]);
> F := ChangeRing(E,GF(2));
> AbelianGroup(F);
Abelian Group isomorphic to Z/4
Defined on 1 generator
Relations:
     4*$.1 = 0
```

A brief aside: how did we know that there could be at maximum five points on $E(\mathbb{F}_2)$? It's not hard for $\mathbb{F}_2$ because there are four points in $(\mathbb{F}_2)^2$ together with the point at infinity. But for larger primes, we can use the Hasse bound, proved in the 1930s; we state the result in a slightly weakened form here. The full result includes all finite fields.

**Theorem 3.12** (Hasse)**.** $\#E(\mathbb{F}_p) \leq p + 2\sqrt{p} + 1$.

Now apply this to the prime $p = 3$: it says that there at most seven points on an elliptic curve over $\mathbb{F}_3$. However, we've seen more torsion than that in an elliptic curve over $\mathbb{Q}$; just take the example above. However, the reason we can't apply the theorem is that the curve doesn't have good reduction at $p = 3$. This gives us an *a priori* bound on torsion. If the discriminant is not divisible by 3, then the torsion subgroup has seven points or fewer.

## 4. Statements of Two Theorems About Torsion

In 1977-8, B. Mazur published a paper [10] that completely answered the question of what torsion subgroups are possible for an elliptic curve over $\mathbb{Q}$. We cite his result here.

**Theorem 4.1** (Mazur)**.** *Let $E/\mathbb{Q}$ be an elliptic curve. Its torsion subgroup $E(\mathbb{Q})_{tor}$ must be one of the following fifteen groups: $\mathbb{Z}/n\mathbb{Z}$ for $n \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12\}$, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ for $n \in \{2, 4, 6, 8\}$. Further, all of these groups are represented as the torsion groups of some elliptic curve over $\mathbb{Q}$.*

In 1996, L. Merel proved a theorem that resolved a long-standing question about the torsion groups of elliptic curves over number fields $K$. The proof of it is also well beyond the scope of this paper (see [11]).

**Theorem 4.2** (Merel)**.** *Let $E/K$ be an elliptic curve, where the number field $K$ satisfies $[K : \mathbb{Q}] = d > 1$. Then any torsion point has order $n < d^{3d^2}$.*

In particular, Merel uses this theorem along with previous work of Faltings and Frey to demonstrate that there is a finite number of possible torsion groups of elliptic curves over a degree $d$ extension of $\mathbb{Q}$. This result is stunningly general. However, although one knows that the list is finite, and for a given $d$ one can take many elliptic curves, compute their torsion groups, and compile a list of the torsion that has appeared, it is not known how to prove that any such list would actually be complete. Mazur's method does not work on number fields.

## 5. Torsion Points on Elliptic Curves: Algorithmic Questions

5.1. **The Nagell-Lutz Theorem.** Now since $E(K)_{\text{tor}}$ is finite, the natural follow-up question is: is it easily computable? For elliptic curves, the answer is well known. Two papers published in the 1930s, one by E. Lutz and the other by T. Nagell, give a clear idea of what torsion points are possible. Their conclusion is presented in the following theorem:

**Theorem 5.1.** *Let $E$ be an elliptic curve over $\mathbb{Q}$ with Weierstrass equation $y^2 = x^3 + Ax + B$, with $A$, $B$ integers. Then the coordinates of a non-zero torsion point $P = (x_0, y_0) \in E(\mathbb{Q})$ are in $\mathbb{Z}$. Furthermore, $P$ is either of order 2, or else $y_0^2$ divides $4A^3 + 27B^2$.*

This characterization of torsion points on elliptic curves over $\mathbb{Q}$ limits the search space. Practically, points of order 2 are easy to find, because in that case $y_0 = 0$; and the second condition ensures that if $P$ is of higher order then there are only finitely many possibilities to check for the value of $y_0$. Therefore there are finitely many pairs $(x_0, y_0)$ to check, since each choice of $y_0$ produces a monic integral cubic equation for $x_0$, any integral solution of which will divide the constant term, $B - y_0^2$. We should remark that this is yet another proof of finiteness of the torsion subgroup, independent of either of the two methods previously mentioned.

The first step is to show that the coordinates of a torsion point are integers.

*Proof.* Let $m$ be the smallest integer for which $mP = \mathcal{O}$. If $m = 2$, then $P = -P = (x_0, -y_0)$. So $y_0 = 0$ and hence $x_0 \in \mathbb{Z}$, since it is the root of a monic polynomial with integral coefficients.

Now assume $m > 2$. Our goal will be to show that for every embedding $E(\mathbb{Q}) \hookrightarrow E(\mathbb{Q}_p)$, the coordinates of our point have nonnegative valuation in the local field. Let $M$ be the maximal ideal $(p)$ to avoid notation confusion.

First, if $v(x_0) \geq 0$, then the Weierstrass equation tells us that $v(y_0) \geq 0$ as well. So we'll assume $v(x_0) < 0$ and find a contradiction.

We observe that the curve has a minimal model, by Remark 3.3, so the reduction map is well defined. By applying the valuation to both sides of the Weierstrass equation again, $v(y_0) < v(x_0) < 0$, and so the projective point $[x_0 : y_0 : 1]$ reduces to the point $[0 : 1 : 0]$ under the map $E(\mathbb{Q}_p) \to \tilde{E}(\mathbb{F}_p)$ (we've used this reasoning before, in the proof of Proposition 3.8). So $(x_0, y_0)$ is a member of $E_1(\mathbb{Q}_p) \cong \hat{E}(M)$.

We can now apply our usual propositions about torsion in formal groups. Proposition 2.7 tells us that if $m$ is relatively prime to $p$ we already have our contradiction (there can't be such a torsion point), and Proposition 2.8 tells us that if $m = p^n$ is a power of $p$, then $0 < v(x/y) \leq \frac{1}{p^n - p^{n-1}}$. But for $m > 2$ this right hand side is less than 1, so we have our contradiction.

Because the coordinates, when thought of as elements of $\mathbb{Q}_p$, are in the ring of integers $\mathbb{Z}_p$ for every $p$, they must be in $\mathbb{Z} \subset \mathbb{Q}$.

□

It remains only to verify that $y_0^2$ divides $4A^3 + 27B^2$ when the torsion point is not of order 2. This is essentially a computation using the duplication formula for a point under the formal addition law on the elliptic curve, following [17].

*Proof.* Assume $2P \neq \mathcal{O}$. Write $2P = (x_1, y_1)$. Since it is also a torsion point, $x_1, y_1 \in \mathbb{Z}$. We can use the duplication formula for a point to express $x_1$ in terms of $x_0$:

$$x_1 = \frac{x_0^4 - 2Ax_0^2 - 8Bx_0 + A^2}{4(x_0^3 + Ax_0 + B)}.$$

Notice that the denominator is just $y_0^2$. Now we produce a polynomial identity by using the Euclidean algorithm on the numerator and denominator of the fraction above. Since the fraction is reduced, we'll get some linear combination of the two polynomials equal to a constant term.

$$(3X^2 + 4A)(X^4 - 2AX^2 - 8BX + A^2) -$$
$$(3X^3 - 5AX - 27B)(X^3 + AX + B) = 4A^3 + 27B^2.$$

Observe that the constant term works out to exactly the value we had above. So all we have to do is plug $X = x(P)$ into this polynomial identity, and simplify using the relations for $x_1$ and $y_0^2$:

$$(3x_0^2 + 4A)(x_1 \cdot 4y_0^2) -$$
$$(3x_0^3 - 5Ax_0 - 27B)(y_0^2) = 4A^3 + 27B^2.$$

All of the variables in this equation are integers. Now it is clear that $y_0^2$ divides the left hand side; therefore it must divide the right. □

5.2. **Doud's Algorithm.** The naive approach to finding the torsion on elliptic curves over $\mathbb{Q}$, then, would be to check each of the finitely many pairs of coordinates $(x, y)$ suggested by the Nagell-Lutz theorem, knowing by Mazur's Theorem (4.1) that a torsion point cannot have order greater than 12. In fact, this approach was widely implemented until fairly recently. J.W. Cremona's 1997 book [5] of elliptic curve algorithms gave it as the standard recipe for the computation of torsion. However, Darrin Doud [6] points out two disadvantages to this method: first, it requires factorization of the discriminant, which may be large; second, for each square divisor of the discriminant, a cubic equation must be solved, and there may be many such divisors. (Remember that the number of divisors grows exponentially in the number of prime factors of the discriminant.)

In its place, Doud suggests using the Weierstrass $\wp$ function associated to a complex lattice $\Lambda \in \mathbb{C}$. Given the lattice, we associate to it the doubly periodic function

$$\wp(z) = \frac{1}{z^2} + \sum_{\alpha \in \Lambda, \alpha \neq 0} \left( \frac{1}{(z-\alpha)^2} - \frac{1}{\alpha^2} \right).$$

It is a standard result (see [17], chapter VI, §3-5) that for the curve associated to the lattice, there is an isomorphism of groups $\phi : \mathbb{C}/\Lambda \to E(\mathbb{C})$ given by $\phi(z) = [\wp(z), \wp'(z), 1]$. (Here the association of curve and lattice means that the curve $E$ has Weierstrass coefficients that come from the Laurent series expansion of the function $\wp$, which depends on $\Lambda$. The precise algebraic way that the coefficients derive from the Laurent expansion is not essential to these remarks.) It is also known that the

reverse association can be made: given an elliptic curve $E(\mathbb{C})$, there is a lattice $\Lambda$ such that the Laurent expansion of $\wp(z)$ produces the Weierstrass coefficients for $E$.

This suggests a natural way of checking for rational torsion in $E$, namely, consider the torsion points in $\mathbb{C}/\Lambda$ and check to see whether their images are rational under the map $\phi$. Since the $n$-torsion points of $\mathbb{C}/\Lambda$ are just $\frac{1}{n}\Lambda/\Lambda$, there isn't much to compute. The essential steps are therefore producing the lattice from the curve in the first place, which is quick relative to the rest of the algorithm, and then following the $n$-torsion from the lattice into the curve, checking to see whether they're integral points. Doud uses reduction (to determine a multiple of the torsion subgroup), in conjunction with Mazur's Theorem, to determine which values of $n$ need to be observed.

So Doud's method is rather quick. The only possible implementational pitfall to the algorithm is its analyticity; one needs to take care that one uses enough terms of the $\wp$ function. Doud proved the required minimal precision for the algorithm in his paper, so the theory is fine. However, if the error bounds are not carefully implemented, one might falsely truncate the series for $\wp(z)$ too early and produce a point on the curve that's "close" to an integral point but not the one it's actually converging to. In that case the algorithm would return less torsion than actually present in the elliptic curve. In fact, this mistake was made in the original MAGMA implementation of Doud's method. In 2001, David Yeung observed that MAGMA's implementation of the algorithm was returning $\mathbb{Z}/8\mathbb{Z}$ for the torsion subgroup of a curve he tested, while PARI produced $\mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. After a bit of confusion, the precision of the analytic method was increased, Doud's algorithm produced the right answer, and there have been no known further problems. The algorithm represents a significant speed increase over the Nagell-Lutz approach.

## 6. Background on Modular Abelian Varieties

**Definition 6.1.** An *algebraic group* $V$ is an algebraic variety that is also a group, in that there are two regular maps, an inverse map $i : V \to V$ and a composition map $m : V \times V \to V$, satisfying the usual group axioms on the points of $V$. An *abelian variety* $A$ is a projective, irreducible algebraic group.

*Remark* 6.2. In order to show that the algebraic variety with the above properties is abelian, one needs to use the fact that $A$ is projective. The proof is found in Shafarevich [15], 4.3.

### 6.1. Modular Groups & Modularity.

Good references for the following definitions include [13], [14], and [16].

Denote the open upper half of the complex plane by $\mathfrak{h} \subset \mathbb{C}$. If we also include the points in $\mathbb{Q}$ and a point $\{\infty\}$ at infinity, we'll call it $\mathfrak{h}^*$. (There's a slightly peculiar topology around the added points (known as "cusps"), as there would have to be: see [16], §1.3, for a reference on this, the Poincaré metric.) The group $\mathrm{SL}_2(\mathbb{R})$ acts on $\mathfrak{h}^*$ by linear fractional transformations, as follows: given an element $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $ad - bc = 1$, and $z \in \mathfrak{h}^*$, let

$$gz = \frac{az + b}{cz + d}.$$

It is not too hard to check that $\mathfrak{h}^*$ is stable under the action of $\mathrm{SL}_2(\mathbb{R})$, since

$$\mathrm{Im}\,(gz) = \mathrm{Im}\left(\frac{(az+b)(c\bar{z}+d)}{|cz+d|^2}\right) = \frac{(ad-bc)\mathrm{Im}\,(z)}{|cz+d|^2},$$

which is positive if the original imaginary part was positive. Notice that the element $\left(\begin{smallmatrix}-1 & 0\\ 0 & -1\end{smallmatrix}\right)$ is the only non-identity element that acts trivially on all of $\mathfrak{h}^*$. So in what follows, we take care to mod out all the groups by $\pm 1$. Thus $\mathrm{PSL}_2(\mathbb{R})$ is a group that acts faithfully on $\mathfrak{h}^*$.

Now consider the discrete subgroup $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$ of matrices with coefficients in $\mathbb{Z}$, again modulo $\pm 1$. It is called *the modular group*, and inherits an action on $\mathfrak{h}^*$. For any integer $N$, we can produce three subgroups of the modular group $\Gamma$ which also act on $\mathfrak{h}^*$:

$$\Gamma_0(N) = \left\{\begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathrm{PSL}_2(\mathbb{Z}) \text{ s.t. } \begin{pmatrix}a & b\\ c & d\end{pmatrix} \equiv \begin{pmatrix}* & *\\ 0 & *\end{pmatrix} \mod N\right\}$$

$$\Gamma_1(N) = \left\{\begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathrm{PSL}_2(\mathbb{Z}) \text{ s.t. } \begin{pmatrix}a & b\\ c & d\end{pmatrix} \equiv \begin{pmatrix}1 & *\\ 0 & 1\end{pmatrix} \mod N\right\}$$

$$\Gamma(N) = \left\{\begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathrm{PSL}_2(\mathbb{Z}) \text{ s.t. } \begin{pmatrix}a & b\\ c & d\end{pmatrix} \equiv \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \mod N\right\}$$

(Here the $*$s can be anything.) Now if we quotient $\mathfrak{h}^*$ by the action of these groups, we get Riemann surfaces ([16], §1.5).

**Definition 6.3.** For any positive integer $N$, the Riemann surfaces $X_0(N)$, $X_1(N)$, and $X(N)$ are the spaces $\Gamma_0(N)\backslash\mathfrak{h}^*$, $\Gamma_1(N)\backslash\mathfrak{h}^*$, $\Gamma(N)\backslash\mathfrak{h}^*$ respectively.

*Remark* 6.4. The quotient is traditionally written on the left because the group action is written on the left. Also, because we have the inclusion of groups $\Gamma(N) \subset \Gamma_1(N) \subset \Gamma_0(N)$, the Riemann surfaces have corresponding relationships, in reverse order: $X(N)$ is the "biggest," and the other two are progressively quotients of it. Finally, these are algebraic curves over $\mathbb{Q}$ (this is not obvious, and is proved in [16], §6.7). For example, [20] notes that $X_1(13)$ is the desingularization of the projective closure of the affine curve $y^2 = x^6 + 2x^5 + x^4 + 2x^3 + 6x^2 + 4x + 1$.

Given a compact Riemann surface $X$ of genus $g$, it has a Jacobian $J(X)$ associated to it, which is a complex torus $\mathbb{C}^g/\Lambda$ for a certain lattice $\Lambda$; we can also think about it as the space of divisors of degree zero on $X$ up to linear equivalence. (Recall that a *divisor* on $X$ is a formal sum of points $p$ in $X$ with integer coefficients,

$$D = \sum_{p\in X} n_p p,\ n \in \mathbb{Z},$$

its *degree* is the sum of those coefficients,

$$\deg(D) = \sum_{p\in X} n_p,$$

and any meromorphic function $f : X \to \mathbb{C}$ has a naturally associated divisor, namely

$$(f) = \sum_{p\in X} (\mathrm{ord}_p(f))p.$$

Then $J(X)$ is the space of divisors modulo the divisors of meromorphic functions.) It is not terribly obvious that the space of divisor classes of degree 0 has the structure of a complex torus; the result is due to Abel and Jacobi. For details of the

isomorphism at a more or less elementary level, including what the lattice $\Lambda$ is explicitly, see the author's junior paper [9].

We have defined three compact Riemann surfaces above, and briefly touched on the notion of the Jacobian. The next definition should therefore be intuitive.

**Definition 6.5.** Let $J_0(N)$ and $J_1(N)$ be the Jacobians of the Riemann surfaces $X_0(N)$ and $X_1(N)$, respectively.

And now we can see, at least etymologically, where the following definition comes from:

**Definition 6.6.** An abelian variety $A$ is *modular* if it is the quotient of $J_1(N)$ for some positive integer $N$, that is, if there is a surjective map $J_1(N) \twoheadrightarrow A$. The smallest integer $N$ for which this holds is referred to as the *level* of the variety.

*Remark* 6.7. Every elliptic curve is modular. This result was conjectured by Taniyama and Shimura and has been established as an extension of the celebrated work of Wiles and Taylor-Wiles (see [2]).

*Remark* 6.8. The formal groups machinery that we established earlier for elliptic curves works on modular abelian varieties as well, although there is no general formula to work with. But one can establish that the reduction map acting on the torsion points $A(\mathbb{Q})_{\text{tor}} \hookrightarrow A(\mathbb{F}_p)_{\text{tor}}$ is injective as well, using similar methods. We do not give the details here.

*Remark* 6.9. Since $J_1(N)$ has quite large dimension, we will work entirely with $J_0(N)$ in the following machinery and, later, in the computations. This is purely a pragmatic consideration, since all abelian varieties that are quotients of $J_0(N)$ are also quotients of $J_1(N)$, and so are modular. We should be aware, however, that there exist abelian varieties that are quotients of the latter but not the former; these will remain untreated in what follows.

*Remark* 6.10. Although we are only going to consider modular abelian varieties, there are many abelian varieties that are not modular. It is conjectured (in [12]) that an abelian variety is modular if and only if it is of "$\mathrm{GL}_2$-type". Briefly, an abelian variety $A/\mathbb{Q}$ is said to be of $\mathrm{GL}_2$-type if there exists $K \subset \mathbb{Q} \otimes \mathrm{End}_{\mathbb{Q}}(A)$ for which $[K : \mathbb{Q}] = \dim(A)$. Here, $\mathrm{End}_{\mathbb{Q}}(A)$ is the ring of endomorphisms of $A$ over $\mathbb{Q}$. This characterization is a generalization of the Taniyama-Shimura conjecture, since an elliptic curve is a dimension 1 abelian variety, and taking $K = \mathbb{Q}$ in the above definition we see that it is automatically of $\mathrm{GL}_2$-type.

6.2. **Hecke Operators and Modular Symbols.** ¿From here on, the theory will explicitly involve the modular subgroup $\Gamma_0(N)$ and its associated $X_0(N)$ and $J_0(N)$, because of a preceding remark (6.9).

**Definition 6.11.** The space of *modular symbols* is a shorthand for writing elements of the homology group. We write $\{\alpha, \beta\} \in H_1(X_0(N), \mathbb{Q})$ for the homology class represented by integration from $\alpha$ to $\beta$, both elements of the set of cusps $\mathbb{P}^1(\mathbb{Q})$.

*Remark* 6.12. For any pair of elements $\alpha, \beta \in \mathbb{P}^1(\mathbb{Q})$, the element $\{\alpha, \beta\}$ is in fact an element of the rational homology, as mentioned above; this is a theorem of Manin and Drinfeld.

Since elements of $\Gamma_0(N)$ act on $\mathfrak{h}^*$, they should also act on the modular symbols by $g(\{\alpha,\beta\}) = \{g\alpha, g\beta\}$. However, for $g \in \Gamma_0(N)$, this does nothing to the homology class, because it amounts only to a change of variables in the integration. This suggests that we should formalize the behavior of modular symbols, using what we already know of the structure of $H_1(X_0(N),\mathbb{Q})$.

**Proposition 6.13.** The modular symbols $\{\alpha,\beta\} \in H_1(X_0(N),\mathbb{Q})$ inherit the following properties from properties of integration:

(1) $\{\alpha,\alpha\} = 0$.
(2) $\{\alpha,\beta\} = -\{\beta,\alpha\}$.
(3) $\{\alpha,\beta\} + \{\beta,\gamma\} = \{\alpha,\gamma\}$.
(4) As observed above, for any $g \in \Gamma_0(N)$, $\{g(\alpha), g(\beta)\} = \{\alpha,\beta\}$.

**Definition 6.14.** Given a prime $p$ not dividing the level $N$, define the *Hecke operator* $T_p$ acting on $H_1(X_1(N),\mathbb{Q})$ by

$$T_p(\{\alpha,\beta\}) = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}(\{\alpha,\beta\}) + \sum_{k=0}^{p-1} \begin{pmatrix} 1 & k \\ 0 & p \end{pmatrix}(\{\alpha,\beta\}).$$

These operators commute. The ring of Hecke operators is denoted by $\mathbb{T} = \mathbb{Z}[T_1, T_2, \ldots]$, the set of $\mathbb{Z}$-linear combinations of the individual operators.

*Remark* 6.15. The individual operations $\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)(\{\alpha,\beta\})$ and $\left(\begin{smallmatrix} 1 & k \\ 0 & p \end{smallmatrix}\right)(\{\alpha,\beta\})$ are not well-defined, in that they depend on a choice of coset representative $\{\alpha,\beta\}$. However, the sum in the definition of the Hecke operator does not. Also observe that the "action" of a $\mathrm{GL}_2(\mathbb{Z})$ matrix on a modular symbol, and therefore on its constituent elements in $\mathbb{P}^1(\mathbb{Q})$, has not yet been mentioned. By a slight abuse of confidence but not of definition, we can extend our concept of matrix action on $\mathbb{P}^1(\mathbb{Q})$ to include all elements of $\mathrm{GL}_2(\mathbb{Z})$, since the imaginary part of an element of $\mathbb{P}^1(\mathbb{Q})$ is zero and the set is therefore preserved by linear fractional transformations.

## 7. Computations with Modular Abelian Varieties

Here we will briefly mention an algorithm (mostly given in [5], ch. II) for finding a multiple of $\#A(\mathbb{Q})_{\mathrm{tor}}$ for modular abelian varieties where we explicitly know the smallest level $N$ of a surjective map $J_1(N) \to A$. This is essentially by way of giving background for the computations we perform in §8.

7.1. **Finding the Quotients of $J_0(N)$.** The first thing to do is compute the homology group $H_1(X_0(N),\mathbb{Q})$, as represented by the group of modular symbols. This group is isomorphic to a certain finite group $H(N)$ with relations that are relatively easy to work with (this is a result due to Manin, cited in [5], theorem 2.1.4), so the computations are easier than one would expect given the definition of the modular symbols space. The group $H(N)$ decomposes into simple $\mathbb{T}$-modules $M_1 \ldots M_{2r}$ under the action of the Hecke algebra, and there is a way in which these modules pair up canonically to associate with each of the $r$ abelian varieties of level $N$.

7.2. **The Torsion Multiple.** It is a deep fact that the characteristic polynomial of the Hecke operator $T_p$ acting on the $M_i$ associated to a variety, evaluated at $p+1$, gives the number of points in the reduction over $\mathbb{F}_p$. This is analogous to the situation with elliptic curves, although without a Weierstrass equation the Hecke

operators are the only way to find the size of the reduced variety. For primes of good reduction, the torsion subgroup is injectively reduced and its order is a divisor of charpoly$(T_p)(p+1)$.

So once we have the varieties associated to the chosen level represented as $\mathbb{T}$-modules $M_i$, we compute the charpoly of $T_p$ acting on $M_i$, and evaluate it at $p+1$, and then take the GCD of the resulting torsion multiple for various primes $p$ not dividing the level.

## 8. Examples

Using two computer workstations, MECCAH (the Mathematics Extreme Computation Cluster at Harvard) and Neron (named for André Néron), William Stein computed the characteristic polynomials of the first eight prime Hecke operators acting on all simple quotients of $J_0(N)$ up to $N = 5000$, with less comprehensive calculations up to $N = 7244$, as well as the dimensions of these varieties. We used this data to compute a multiple of the torsion group for each one, as above, using primes up to 19.

This first table shows the possible torsion multiples, sorted by dimension, for low dimensions. Full ranges are given by a dash.

| Dimension | Torsion Multiples Found |
|---|---|
| 1 | 1-10, 12, 16, 20* |
| 2 | 1-17, 19, 20, 22-24, 27, 28, 30-32, 36, 37, 44, 52, 56, 64 |
| 3 | 1-20, 22-24, 28, 31, 32, 36-38, 49, 58, 64, 72, 80, 83†, 92, 160 |
| 4 | 1-17, 19-29, 31, 32, 34-36, 38, 40, 44, 48, 49, 55, 56, 64, 74, 76, 96, 152, 176, 208, 256, 328† |
| 5 | 1-14, 16, 18-25, 27-29, 31, 32, 34-36, 38, 41†, 44, 48, 50, 54, 56, 58, 63, 64, 68, 72, 86†, 92, 104, 110, 128, 212†, 224, 272, 288, 320 |
| 6 | 1-18, 20, 22, 24, 25, 27, 28, 31, 32, 34, 36, 37, 40, 41†, 48, 49, 52, 53†, 56, 64, 70, 74, 76, 80, 96, 104, 120, 124, 128, 192, 212†, 248, 288, 296, 432 |

Two things jump out at us. First, the starred number in the above chart is the only case where the torsion multiple is immediately seen to be too large, by Theorem 4.1. The variety is the tenth in level 2310. Since $2310 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11$, it is likely that we merely didn't use enough primes (the computations could only take advantage of the primes 13, 17, and 19). A quick follow-up search in the modular forms database shows that our surmise was correct: using primes through 53, the GCD of the torsion multiples is reduced quite drastically, to 4, which is well within the bound imposed by Theorem 4.1.

Next, the daggers in the table represent torsion multiples divisible by a prime greater than 40. Since there isn't even any 11-torsion in dimension 1, it's interesting to investigate some examples of torsion points of large prime order in higher dimensions. We'll make a separate table of these †-ed varieties; where possible we use the modular forms database to see whether the torsion multiple gets smaller using primes up to 53.

| N (level) | Number within level | Dimension | Multiple of torsion | Multiple of torsion (using MFD) |
|---|---|---|---|---|
| 5010 | 16 | 3 | 83 | – |
| 498 | 7 | 4 | 328 | 82 |
| 830 | 11 | 5 | 41 | 41 |
| 1038 | 10 | 5 | 86 | 86 |
| 642 | 8 | 5 | 212 | 106 |
| 83 | 2 | 6 | 41 | $41^{\ddagger}$ |
| 326 | 5 | 6 | 41 | $41^{\ddagger}$ |
| 2158 | 10 | 6 | 41 | 41 |
| 1070 | 12 | 6 | 53 | 53 |
| 422 | 6 | 6 | 212 | $53^{\ddagger}$ |

Let us consider the second entry in the table: the 7th modular abelian variety of level 498; it has dimension 4. Using primes up to 19, we computed 328 as the torsion multiple and 40709 (a prime) as the multiple of the field discriminant. Since 498 is only divisible by two of the primes on our list (2 and 3), one may be inclined to "trust," more or less, the computation of the torsion multiple. In particular, 328 has a large prime factor, namely 41, and we'd like to know whether we can observe 41-torsion in the variety explicitly.

There is a series of maps

$$X_0(N) \hookrightarrow J_0(N) \twoheadrightarrow A$$

where the first map between points and divisors is $P \mapsto [P - \infty]$. The image of the cusps in $X_0(N)$ under this map can be used to generate a subgroup of the Jacobian called the cuspidal subgroup, $C$. It is a fact that $C(\mathbb{Q}) \subset J_0(N)(\mathbb{Q})_{\mathrm{tor}}$. Then under the quotient map into $A(\mathbb{Q})$ it will end up as a subgroup of the torsion on the variety. Thus this is a way of explicitly constructing a lower bound on the torsion, by actually finding a subgroup of it.

So the natural thing to do is to try that with this variety. We construct it using MAGMA, as follows:

```
> S := CuspidalSubspace(ModularSymbols(498,2,+1));
> D := TraceSortDecomposition(NewformDecomposition(S));
> A := D[15];
```

Then the command `RationalCuspidalSubgroup(A)` gives the subgroup generated by the image of the cuspidal subgroup under quotient. In this case, it returns `Abelian Group of order 1`.

(Note that in the varieties marked with a ‡ in the table above, this method produces a nontrivial subgroup and therefore confirms the high prime torsion explicitly.)

Perhaps unfortunately, in the example we were considering, the subgroup of the torsion group that this method produces is trivial. But there does seem to be 41-torsion in the variety. If we use primes up to 53 (computed in the modular forms database), the torsion multiple is refined to 82. All this is very suggestive, but we'd like to know whether 41 isn't just a red herring that keeps appearing.

To this end, we can turn to the Birch and Swinnerton-Dyer conjecture to make predictions. The conjecture supposes an equality

$$\frac{L(A,1)}{\Omega_A} = \frac{\prod_{p|N} c_p \cdot \#Ш(A)}{\#A(\mathbb{Q})_{\text{tor}} \cdot \#A^\vee(\mathbb{Q})_{\text{tor}}},$$

where the left hand side is an analytic invariant associated to the variety and the right hand side contains certain algebraic invariants. The ones we are particularly concerned with are the torsion group sizes in the denominator on the right, of course. In the example above with level 498, many of these quantities can be computed explicitly. Using the MAGMA commands `LRatio` and `TamagawaNumber` (the former predicts the left hand side of the equation up to a constant conjecturally equal to 1), we find that the BSD conjecture predicts that

$$56 = \frac{(2296 \cdot 164 \cdot 1)\#Ш(A)}{\#A(\mathbb{Q})_{\text{tor}} \cdot \#A^\vee(\mathbb{Q})_{\text{tor}}}.$$

All the remaining unknown quantities are integers, and the numerator on the right hand side is divisible by $41^2$, while the left hand side has no divisors of 41. Hence if BSD is correct, we should expect the variety and its dual to have 41-torsion. (We expect both of them to have it because neither of them can have two factors of 41, since the torsion multiple is an upper bound on all curves in the isogeny class.) Often we can use the BSD conjecture to make statements about the torsion subgroup.

8.1. **Bounds using Good Reduction.** If $A$ is of dimension 2 and has good reduction at the prime 3, then $\#A(\mathbb{Q})_{\text{tor}} \leq \#A(\mathbb{F}_3)$. It is known that the characteristic polynomial of the Hecke operator $T_3$ factors as $(x - a_3)(x - a_3^\sigma)$. Further, a result due to Deligne states that $|a_p|, |a_p^\sigma| \leq 2p^{\frac{k-1}{2}}$, where $k$ is the dimension. In this case, $a_3$ and its conjugate are bounded in absolute value by $2\sqrt{3}$. Writing these as elements of the ring of integers in the real quadratic field $\mathbb{Q}[\sqrt{D}]$, we have $|a + b\sqrt{D}|, |a - b\sqrt{D}| \leq 2\sqrt{3}$. We'd like to bound the possible value of the characteristic polynomial evaluated at 4=3+1, which would tell us the number of points over $\mathbb{F}_3$.

Plugging in $x = 4$ and multiplying out, we need to find the maximal value of $16 + (a^2 - b^2 D) - 8a$, which is at its largest when $b = 0$ and, if the absolute value of $a$ is bounded, when $a < 0$ is minimized within the range. We do some fiddling with Deligne bound and the triangle inequality to find this range:

$$\begin{aligned} |2a| &\leq |a + b\sqrt{D}| + |a - b\sqrt{D}| \\ |2a| &\leq 4\sqrt{3} \\ |a| &\leq 2\sqrt{3} \end{aligned}$$

Since $a$ can only take on discrete integer or possibly half-integer values, $a = -3$ is the least possible; and thus the number of torsion points on a dimension 2 modular abelian variety with good reduction at $p = 3$ is at most $(4 - (-3))^2 = 49$.

To confirm, we search our data for varieties with dimension 2 and a computed torsion multiple greater than 49:

| N (level) | Number within level | Dimension | Multiple of field discriminant | Multiple of torsion |
|---|---|---|---|---|
| 390 | 8 | 2 | 32 | 56 |
| 2574 | 26 | 2 | 164 | 64 |
| 3120 | 29 | 2 | 32 | 56 |
| 4368 | 32 | 2 | 17 | 52 |
| 5070 | 29 | 2 | 32 | 56 |
| 5850 | 64 | 2 | 32 | 56 |

Happily enough, all the levels here are divisible by 3, so we don't have to worry about the torsion multiple violating our bound. In fact, for a curve with good reduction at 3, the highest torsion multiple in the data we have computed so far is 22. The variety in which this putative torsion occurs is a quotient of $J_0(322)$. As in the previous section, for this variety `RationalCuspidalSubgroup` is trivial, but the BSD conjecture suggests that there's at least 11-torsion in the variety and its dual.

This table also serves to illustrate two contradictory effects, since the torsion multiple is high and many small primes divide the level. First, since many small primes divide the level, the torsion multiple may be too high, since we have fewer numbers in the GCD calculation. Second, if the torsion is high, it ought to have many small primes dividing the level, so as to have bad reduction at the small primes and therefore the capacity for higher torsion. It is not easy to tell which effect is dominant in a particular example.

8.2. **Conjectures.** Sifting through the data, we find strong evidence for the following conjectures.

**Conjecture 8.1.** (1) Given $p_1, p_2, p_3$ distinct primes, if the numerator of $\frac{p_i-1}{12}$, which we'll call $\ell_i$, is a prime for some $i$, then there is an $\ell_i$-torsion point in some modular abelian variety whose level is $p_1 \cdot p_2 \cdot p_3$.
(2) An analogous statement holds for any odd number of distinct primes $p_i$.
(3) The above statements hold for any prime $\ell_i$ dividing the numerator of $\frac{p_i-1}{12}$.

**Conjecture 8.2.** For any prime $p > 2$, there is a $p$-torsion point in some modular abelian variety whose level is $p^3$.

For an example of the first statement, recall the abelian variety with level $498 = 2 \cdot 3 \cdot 83$ above; the BSD computation suggested strongly that it had a 41-torsion point. Similar computations can be made with other varieties in that table: the one with level $642 = 2 \cdot 3 \cdot 107$ appeared to have a 53-torsion point. In fact, upon closer inspection, all the varieties with high torsion that we found seem to satisfy this conjecture.

This first observation is bound to be related to the following theorem, which is proved in [10]:

**Theorem 8.3** (Mazur)**.** *Let $p > 3$ be a prime, and let $n$ be the numerator of $\frac{p-1}{12}$. Then the torsion subgroup of the Mordell-Weil group of $J_0(p)$ is a cyclic subgroup of order $n$.*

Although this theorem addresses the case of prime level, it is not immediately clear why Conjecture 8.1 seems to hold when the level is the product of an odd

number of distinct primes. It is also worth noting that levels that are the product of an even number of distinct primes do not exhibit this behavior.

The second observation is a bit more puzzling still, and we do not suggest a reason for its truth here.

It is entirely possible that both observations have been made before or already proved; however, we do not at this point know of a source for them.

## 9. Acknowledgments

I would like to thank my adviser, Benjamin Peirce Assistant Professor William A. Stein, for his enthusiasm and explanations.

## References


[1] M.F. Atiyah and I.G. Macdonald. *Introduction to Commutative Algebra*. Addison-Wesley, 1969.

[2] C. Breuil, B. Conrad, F. Diamond, and R. Taylor. *On the Modularity of Elliptic Curves over $\mathbb{Q}$: Wild 3-adic Exercises*. Journal of the American Mathematical Society, 14, no. 4 (2001), 843-939.

[3] J.W.S. Cassels and E.V. Flynn. *Prolegomena to a Middlebrow Arithmetic of Curves of Genus 2*. Cambridge University Press, 1996.

[4] G. Cornell and J.H. Silverman, eds. *Arithmetic Geometry*. Springer-Verlag, 1986.

[5] J.E. Cremona. *Algorithms for Modular Elliptic Curves*, second edition. Cambridge University Press, 1997.

[6] D. Doud. *A Procedure to Calculate Torsion of Elliptic Curves Over $\mathbb{Q}$*. Manuscripta Mathematica, 95 (1998), 463-469.

[7] E.V. Flynn, F. Leprévost, E.F. Schaefer, W.A. Stein, M. Stoll, and J.L. Wetherell. *Empirical Evidence for the Birch and Swinnerton-Dyer Conjectures for Modular Jacobians of Genus 2 Curves*. Mathematics of Computation, 236 (2001), 1675-1697.

[8] N. Katz. *Galois Properties of Torsion Groups on Abelian Varieties*. Inventiones Mathematicae, 62 (1981), 481-502.

[9] S.J. Kleinerman. *The Jacobian, the Abel-Jacobi Map, and Abel's Theorem*. Submitted as junior paper, Harvard, 2003.

[10] B. Mazur. *Modular Curves and the Eisenstein Ideal*. Publications mathématiques de l'I.H.É.S., 47, no. 2 (1977), 33-186.

[11] L. Merel. *Bornes pour la Torsion des Courbes Elliptiques sur les Corps des Nombres*. Inventiones Mathematicae, 124 (1996), 437-449.

[12] K.A. Ribet. *Abelian Varieties over $\mathbb{Q}$ and Modular Forms*. Article found in *Modular Curves and Abelian Varieties*. Birkhaüser, 2004.

[13] K.A. Ribet and W.A. Stein. *Modular Forms, Hecke Operators, and Modular Abelian Varieties*. Preprint.

[14] J.-P. Serre. *A course in Arithmetic*. Springer-Verlag, 1973.

[15] I.R. Shafarevich. *Basic Algebraic Geometry 1*. Springer-Verlag, 1994.

[16] G. Shimura. *Introduction to the Arithmetic Theory of Automorphic Functions*. Princeton University Press, 1971.

[17] J.H. Silverman. *The Arithmetic of Elliptic Curves*. Springer-Verlag, 1986.

[18] ———. *Advanced Topics in the Arithmetic of Elliptic Curves*. Springer-Verlag, 1994.

[19] J.H. Silverman and J. Tate. *Rational Points on Elliptic Curves*. Springer-Verlag, 1992.

[20] W.A. Stein. *The Arithmetic of $J_1(p)$*. Notes for a talk delivered at Brandeis in October 2003.

[21] W.A. Stein. *Explicit Approaches to Modular Abelian Varieties*. Ph.D. thesis, University of California at Berkeley, 2000.