Math 7359 (Elliptic Curves and Modular Forms)

Lecture #15 of 24 $\sim$ October 30, 2023

Dual Isogenies

- Dual Isogenies
- Properties of Dual Isogenies
- Applications to Elliptic Curves over $\mathbb{F}_q$.

Dual Isogenies, I

Our goal now is to show that “being isogenous” is an equivalence relation on elliptic curves.

- Since being isogenous is reflexive and transitive as we have already noted, it remains to show that every nonzero isogeny $\varphi : E_1 \to E_2$ induces some other nonzero isogeny $\hat{\varphi} : E_2 \to E_1$.

Dual Isogenies, I

Our goal now is to show that “being isogenous” is an equivalence relation on elliptic curves.

- Since being isogenous is reflexive and transitive as we have already noted, it remains to show that every nonzero isogeny $\varphi : E_1 \to E_2$ induces some other nonzero isogeny $\hat{\varphi} : E_2 \to E_1$.
- To see that this “dual isogeny” exists, we exploit the contravariant nature of the map $\varphi^* : \mathrm{Div}(E_2) \to \mathrm{Div}(E_1)$.
- Specifically, because φ^* scales degrees by $\deg \varphi$, as we showed earlier, it maps $\mathrm{Div}^0(E_2)$ into $\mathrm{Div}^0(E_1)$, and therefore it descends onto a well-defined map $\varphi^* : \mathrm{Pic}^0(E_2) \to \mathrm{Pic}^0(E_1)$.
- But as we also showed, the group operation in $\mathrm{Pic}^0(E)$ is isomorphic to the group law on E (namely, via the map sending a point $P \in E$ to the divisor class $[P] - [O]$), and so by composing these isomorphisms appropriately, we obtain a group homomorphism $\hat{\varphi} : E_2 \to E_1$.

Dual Isogenies, II

To see that this "dual isogeny" exists, we exploit the contravariant nature of the map $\varphi^* : \mathrm{Div}(E_2) \to \mathrm{Div}(E_1)$.

- Specifically, because φ^* scales degrees by $\deg \varphi$, as we showed earlier, it maps $\mathrm{Div}^0(E_2)$ into $\mathrm{Div}^0(E_1)$, and therefore it descends onto a well-defined map $\varphi^* : \mathrm{Pic}^0(E_2) \to \mathrm{Pic}^0(E_1)$.
- But as we also showed, the group operation in $\mathrm{Pic}^0(E)$ is isomorphic to the group law on E (namely, via the map sending a point $P \in E$ to the divisor class $[P] - [O]$), and so by composing these isomorphisms appropriately, we obtain a group homomorphism $\hat{\varphi} : E_2 \to E_1$.

Of course, it is not at all obvious that this group homomorphism $\hat{\varphi}$ is actually an isogeny, since there are very many possible homomorphisms between the point groups, most of which will not be defined by rational functions.

Dual Isogenies, III

Let's work out exactly what this map does to a point $Q \in E_2$:

Dual Isogenies, III

Let's work out exactly what this map does to a point $Q \in E_2$:

- First, we map Q to the divisor class $[Q] - [O]$.
- Then we apply φ^* and (3) to obtain $\deg_i \varphi \left(\sum_{P \in \varphi^{-1}(Q)} [P] - \sum_{R \in \varphi^{-1}(O)} [R] \right)$.
- Finally we must resolve this sum to write it in the form $[S] - [O]$: the result is then S.
- By our results from equivalence of divisors, we can just sum everything using the group law: this yields $S = \deg_i \varphi \left(\sum_{P \in \varphi^{-1}(Q)} P - \sum_{R \in \varphi^{-1}(O)} R \right)$.
- Since $\varphi^{-1}(Q) = \{P + R : R \in \varphi^{-1}(O)\}$ for any fixed $P \in \varphi^{-1}(Q)$, the difference is simply $[\deg_i \varphi \cdot \#\varphi^{-1}(Q)]P$.

Dual Isogenies, IV

So, to summarize, this map $\hat{\varphi} : E_2 \to E_1$ maps a point $Q \in E_2$ to $[\deg \varphi]P$ where P is any point in $\varphi^{-1}(Q)$.

- Note that this description of $\hat{\varphi}$ is well posed: regardless of which representative $P \in \varphi^{-1}(Q)$ is chosen, since the difference between any of these representatives lies in $\varphi^{-1}(O) = \ker \varphi$.

Dual Isogenies, IV

So, to summarize, this map $\hat{\varphi} : E_2 \to E_1$ maps a point $Q \in E_2$ to $[\deg \varphi]P$ where P is any point in $\varphi^{-1}(Q)$.

- Note that this description of $\hat{\varphi}$ is well posed: regardless of which representative $P \in \varphi^{-1}(Q)$ is chosen, since the difference between any of these representatives lies in $\varphi^{-1}(O) = \ker \varphi$.
- Equivalently, this says $\hat{\varphi}(\varphi(P)) = [\deg \varphi]P$ for all $P \in E_1$, meaning that the composition $\hat{\varphi} \circ \varphi$ is simply the multiplication-by-$[\deg \varphi]$ map on E_1.

The whole point of this calculation (aside from giving an explicit description of what this map would look like) is that this last description actually provides us with a way to prove that $\hat{\varphi}$ actually is an isogeny: we can use the universal property (9) of isogenies.

Dual Isogenies, V

So, let's go through the details:

Theorem (Existence of Dual Isogenies)

Let $\varphi : E_1 \to E_2$ be a nonconstant isogeny.

1. *If φ is separable, then there exists a unique isogeny $\hat{\varphi} : E_2 \to E_1$ such that $\hat{\varphi} \circ \varphi$ is multiplication by $\deg \varphi$ on E_1.*
2. *If $\mathrm{char}(k) = p > 0$ and Frob_p is the pth-power Frobenius morphism $\mathrm{Frob}_p : E \to E^{(p)}$, then there exists a unique isogeny $\widehat{\mathrm{Frob}_p} : E^{(p)} \to E$ such that $\widehat{\mathrm{Frob}_p} \circ \mathrm{Frob}_p$ is multiplication by $p = \deg(\mathrm{Frob}_p)$ on E.*
3. *There exists a unique isogeny $\hat{\varphi} : E_2 \to E_1$ such that $\hat{\varphi} \circ \varphi = [\deg \varphi]$ on E_1. This isogeny is called the dual isogeny of φ.*

Dual Isogenies, VI

1. If φ is separable, then there exists a unique isogeny $\hat{\varphi} : E_2 \to E_1$ such that $\hat{\varphi} \circ \varphi$ is multiplication by $\deg \varphi$ on E_1.

Proof:

Dual Isogenies, VI

1. If φ is separable, then there exists a unique isogeny $\hat{\varphi} : E_2 \to E_1$ such that $\hat{\varphi} \circ \varphi$ is multiplication by $\deg \varphi$ on E_1.

Proof:

- Let $\psi = [\deg \varphi]$ be the multiplication-by-$\deg \varphi$ map on E_1 and $E_3 = E_1$. Then since $\# \ker \varphi = \deg \varphi$, by Lagrange's theorem we see that $\ker \varphi \subseteq \ker \psi$.
- Now by the universal property (7) of separable isogenies, there exists a unique isogeny $\hat{\varphi} : E_2 \to E_1$ such that $\hat{\varphi} \circ \varphi = \psi = [\deg \varphi]$, as claimed.

Dual Isogenies, VII

2. If $\mathrm{char}(k) = p > 0$ and Frob_p is the pth-power Frobenius morphism $\mathrm{Frob}_p : E \to E^{(p)}$, then there exists a unique isogeny $\widehat{\mathrm{Frob}_p} : E^{(p)} \to E$ such that $\widehat{\mathrm{Frob}_p} \circ \mathrm{Frob}_p$ is multiplication by $p = \deg(\mathrm{Frob}_p)$ on E.

Proof:

Dual Isogenies, VII

2. If $\mathrm{char}(k) = p > 0$ and Frob_p is the pth-power Frobenius morphism $\mathrm{Frob}_p : E \to E^{(p)}$, then there exists a unique isogeny $\widehat{\mathrm{Frob}_p} : E^{(p)} \to E$ such that $\widehat{\mathrm{Frob}_p} \circ \mathrm{Frob}_p$ is multiplication by $p = \deg(\mathrm{Frob}_p)$ on E.

Proof:

- Let ω be the invariant differential on E.
- By property (13) of isogenies we see that $[p]^*\omega = p\omega = 0$, so $[p]$ is not separable since it is not injective on differentials.
- Hence by property (9) of isogenies, we may factor $[p]$ as $[p] = \alpha \circ \mathrm{Frob}_q$ where $q = \deg_i[p] = p^d$ for some integer $d \geq 1$ (note $d \geq 1$ because $[p]$ is not separable).
- Then since $\mathrm{Frob}_q = (\mathrm{Frob}_p)^d$ we see that $[p] = \alpha \circ (\mathrm{Frob}_p)^{d-1} \circ \mathrm{Frob}_p$.
- We can then take $\hat{\varphi} = \alpha \circ (\mathrm{Frob}_p)^{d-1}$.

Dual Isogenies, VIII

3. There exists a unique isogeny $\hat{\varphi} : E_2 \to E_1$ such that $\hat{\varphi} \circ \varphi = [\deg \varphi]$ on E_1.

Discussion:

Dual Isogenies, VIII

3. There exists a unique isogeny $\hat{\varphi} : E_2 \to E_1$ such that $\hat{\varphi} \circ \varphi = [\deg \varphi]$ on E_1.

Discussion:

- We emphasize here that this statement is equivalent to the one we worked out earlier as motivation for the construction for $\hat{\varphi}$: namely, for any $P \in C_1$, with $Q = \varphi(P)$ we have $\hat{\varphi}(Q) = [\deg \varphi]P$.
- So the point of this result is to complete the claim made at the end of that discussion: namely, that the map we constructed using the action of φ^* on divisor groups actually gives rise to an *isogeny* from E_2 to E_1, not just a group homomorphism.

Dual Isogenies, VIII

3. There exists a unique isogeny $\hat{\varphi} : E_2 \to E_1$ such that $\hat{\varphi} \circ \varphi = [\deg \varphi]$ on E_1.

Proof (existence):

Dual Isogenies, VIII

3. There exists a unique isogeny $\hat{\varphi} : E_2 \to E_1$ such that $\hat{\varphi} \circ \varphi = [\deg \varphi]$ on E_1.

Proof (existence):

- By property (9) of isogenies, we may decompose $\varphi = \alpha \circ \mathrm{Frob}_q = \alpha \circ (\mathrm{Frob}_p)^d$ where α is separable.
- By (1) there exists an isogeny $\hat{\alpha}$ with $\hat{\alpha} \circ \alpha = [\deg \alpha]$ and by (2) there exists an isogeny $\widehat{\mathrm{Frob}_p}$ with $\widehat{\mathrm{Frob}_p} \circ \mathrm{Frob}_p = [\deg \mathrm{Frob}_p]$.
- Then for $\hat{\varphi} = (\widehat{\mathrm{Frob}_p})^d \circ \hat{\alpha}$ we have $\hat{\varphi} \circ \varphi = (\widehat{\mathrm{Frob}_p})^d \circ \hat{\alpha} \circ \alpha \circ (\mathrm{Frob}_p)^d = (\widehat{\mathrm{Frob}_p})^d \circ [\deg \alpha] \circ (\mathrm{Frob}_p)^d = [\deg \alpha] \circ (\widehat{\mathrm{Frob}_p})^d \circ (\mathrm{Frob}_p)^d = [\deg \alpha][\deg \mathrm{Frob}_p]^d = [\deg \varphi]$.

Dual Isogenies, IX

3. There exists a unique isogeny $\hat{\varphi} : E_2 \to E_1$ such that $\hat{\varphi} \circ \varphi = [\deg \varphi]$ on E_1.

Proof (uniqueness):

- For uniqueness, suppose $\tilde{\varphi} \circ \varphi = [\deg \varphi] = \hat{\varphi} \circ \varphi$.
- Then $(\tilde{\varphi} - \hat{\varphi}) \circ \varphi = 0$.
- Taking degrees yields $\deg(\tilde{\varphi} - \hat{\varphi}) \deg \varphi = 0$, so since $\deg \varphi \neq 0$ that means $\deg(\tilde{\varphi} - \hat{\varphi}) = 0$ whence $\tilde{\varphi} = \hat{\varphi}$.

More Dual Isogenies, I

We will now establish some additional properties of dual isogenies, which will allow us in particular to understand the kernel of the multiplication-by-m map on an elliptic curve E more explicitly.

- Since the kernel of $[m]$ is just the group of m-torsion points, this will represent substantial progress in our understanding of the group structure of E, since the torsion subgroup of E is simply the union of the m-torsion subgroups for $m \geq 1$.

Notational Interlude

So far, we have been writing and referring to $[m]$ as "the multiplication-by-m map", and I haven't heard any objections.

- However, this is a mild abuse of notation, because the multiplication-by-m map is different on each elliptic curve.
- Morally, all of these maps behave exactly the same way, and because isogenies are all group homomorphisms, it's kosher to view all of them as equivalent.

Notational Interlude

So far, we have been writing and referring to $[m]$ as "the multiplication-by-m map", and I haven't heard any objections.

- However, this is a mild abuse of notation, because the multiplication-by-m map is different on each elliptic curve.
- Morally, all of these maps behave exactly the same way, and because isogenies are all group homomorphisms, it's kosher to view all of them as equivalent. But just to be safe, do this exercise, then explain why it justifies our abuse of notation:

Exercise: Show that for any integer m and any isogeny $\varphi : E_1 \to E_2$, we have $[m]_{E_2} \circ \varphi = \varphi \circ [m]_{E_1}$.

We have also been calling all of the Frobenius maps the same name (though that's even more justifiable because they're all literally the same function: the pth-power map). You should also convince yourself that this is acceptable.

More Dual Isogenies, I

Theorem (More With Dual Isogenies)

Let $\varphi : E_1 \to E_2$ be a nonconstant isogeny and $\hat{\varphi} : E_2 \to E_1$ be its dual isogeny.

4. *We have $\varphi \circ \hat{\varphi} = [\deg \varphi]$ on E_2.*
5. *For any isogenies $\varphi : E_1 \to E_2$ and $\psi : E_2 \to E_3$ we have $\widehat{\psi \circ \varphi} = \hat{\varphi} \circ \hat{\psi}$.*
6. *For any isogenies $\varphi, \psi : E_1 \to E_2$ we have $\widehat{\psi + \varphi} = \hat{\varphi} + \hat{\psi}$.*
7. *For any nonzero integer m we have $\widehat{[m]} = [m]$ and $\deg[m] = m^2$.*
8. *We have $\deg \hat{\varphi} = \deg \varphi$ and $\hat{\hat{\varphi}} = \varphi$.*

More Dual Isogenies, II

4. We have $\varphi \circ \hat{\varphi} = [\deg \varphi]$ on E_2.

Proof:

More Dual Isogenies, II

4. We have $\varphi \circ \hat{\varphi} = [\deg \varphi]$ on E_2.

Proof:

- Notice that $\hat{\varphi} \circ \varphi \circ \hat{\varphi} = [\deg \varphi] \circ \hat{\varphi} = \hat{\varphi} \circ [\deg \varphi]$ by (3) and the fact that the multiplication-by-m maps "commute" with all isogenies per the exercise earlier.
- Thus, $\hat{\varphi} \circ (\varphi \circ \hat{\varphi} - [\deg \varphi]) = 0$.
- So by taking degrees as usual we see that since $\hat{\varphi} \neq 0$ we have $\varphi \circ \hat{\varphi} = [\deg \varphi]$.

More Dual Isogenies, III

5. For any isogenies $\varphi : E_1 \to E_2$ and $\psi : E_2 \to E_3$ we have $\widehat{\psi \circ \varphi} = \hat{\varphi} \circ \hat{\psi}$.

<u>Proof</u>:

More Dual Isogenies, III

5. For any isogenies $\varphi : E_1 \to E_2$ and $\psi : E_2 \to E_3$ we have $\widehat{\psi \circ \varphi} = \hat{\varphi} \circ \hat{\psi}$.

Proof:

- Observe that $(\hat{\varphi} \circ \hat{\psi}) \circ (\psi \circ \varphi) = \hat{\varphi} \circ [\hat{\psi} \circ \psi] \circ \varphi = \hat{\varphi} \circ [\deg \psi] \circ \varphi = [\deg \psi][\deg \varphi] = [\deg \psi \circ \varphi]$.
- Since the dual isogeny is unique by (3), we must have $\widehat{\psi \circ \varphi} = \hat{\varphi} \circ \hat{\psi}$.

More Dual Isogenies, IV

6. For any isogenies $\varphi, \psi : E_1 \to E_2$ we have $\widehat{\psi + \varphi} = \hat{\varphi} + \hat{\psi}$.

Discussion:

More Dual Isogenies, IV

6. For any isogenies $\varphi, \psi : E_1 \to E_2$ we have $\widehat{\psi + \varphi} = \hat{\varphi} + \hat{\psi}$.

Discussion:

- The proof of this result is rather involved.
- Why? The difficulty is that the definition of the dual isogeny $\hat{\varphi}$ does not play especially nicely with addition.
- Specifically, we have $\hat{\varphi}(Q) = [\deg \varphi]P$ for any $P \in \varphi^{-1}Q$, while $\hat{\psi}(Q) = [\deg \psi]P'$ for any $P' \in \psi^{-1}(Q)$ and $\widehat{\psi + \varphi}(Q) = [\deg(\varphi + \psi)](P'')$ for any $P'' \in (\varphi + \psi)^{-1}(Q)$.
- There is no obvious relation between P, P', and P'', since they are preimages of Q under three different isogenies.
- So how could we possibly show that $[\deg(\varphi + \psi)](P'') = [\deg \varphi]P + [\deg \psi]P'$?

More Dual Isogenies, V

6. For any isogenies $\varphi, \psi : E_1 \to E_2$ we have $\widehat{\psi + \varphi} = \hat{\varphi} + \hat{\psi}$.

Discussion, more:

More Dual Isogenies, V

6. For any isogenies $\varphi, \psi : E_1 \to E_2$ we have $\widehat{\psi + \varphi} = \hat{\varphi} + \hat{\psi}$.

<u>Discussion, more</u>:

- So how could we possibly show that $[\deg(\varphi + \psi)](P'') = [\deg \varphi]P + [\deg \psi]P'$?
- One approach is to work instead with divisors, and show that the difference between the two sides resolves to the identity inside the Picard group.
- Equivalently, we want to show that $[\deg(\varphi + \psi)](P'') - [\deg \varphi]P - [\deg \psi]P' + O$ is principal inside $\mathrm{Div}(E_1)$.
- This is better because it's the same as $(\varphi + \psi)^*(Q) - \varphi^*(Q) - \psi^*(Q) + O$, as a divisor.
- To find an f for which this equals $\mathrm{div}(f)$ we will use general coordinates for P and Q, and then specialize.

More Dual Isogenies, VI

6. For any isogenies $\varphi, \psi : E_1 \to E_2$ we have $\widehat{\psi + \varphi} = \hat{\varphi} + \hat{\psi}$.

Proof (part 1):

More Dual Isogenies, VI

6. For any isogenies $\varphi, \psi : E_1 \to E_2$ we have $\widehat{\psi + \varphi} = \hat{\varphi} + \hat{\psi}$.

Proof (part 1):

- If φ, ψ, or $\varphi + \psi$ is zero, the result is trivial, so assume all of them are nonzero.
- Let (x_1, y_1) and (x_2, y_2) be coordinates on E_1.
- Because φ, ψ, and $\varphi + \psi$ are all morphisms, $\varphi(x_1, y_1)$, $\psi(x_1, y_1)$, and $(\varphi + \psi)(x_1, y_1)$ are all elements of the function field $E_2(k(x_1, y_1))$ of E_2 over the field $k(x_1, y_1)$.
- Let D be the divisor $[(\varphi + \psi)(x_1, y_1)] - [\varphi(x_1, y_1)] - [\psi(x_1, y_1)] + [O]$ on E_2 over $k(x_1, y_1)$ – in other words, in $\mathrm{Div}_{k(x_1,y_1)}(E_2)$ – since it has degree 0 and the point sum resolves to the identity, it is the divisor of some function $f \in k(x_1, y_1)(E_2) = k(x_1, y_1, x_2, y_2)$.
- Now we switch emphasis on the coordinates.

More Dual Isogenies, VII

6. For any isogenies $\varphi, \psi : E_1 \to E_2$ we have $\widehat{\psi + \varphi} = \hat{\varphi} + \hat{\psi}$.

Proof (part 2):

- $\operatorname{div}(f) = [(\varphi + \psi)(x_1, y_1)] - [\varphi(x_1, y_1)] - [\psi(x_1, y_1)] + [O]$ on E_2 over $k(x_1, y_1)$.
- Now consider $\operatorname{div}(f)$ inside the divisor group $\operatorname{Div}_{k(x_2,y_2)}(E_2)$ – i.e., with x_2, y_2 constant and x_1, y_1 the variables. Let us compute the zeroes and poles (and their orders) of f.

More Dual Isogenies, VII

6. For any isogenies $\varphi, \psi : E_1 \to E_2$ we have $\widehat{\psi + \varphi} = \hat{\varphi} + \hat{\psi}$.

Proof (part 2):

- $\mathrm{div}(f) = [(\varphi + \psi)(x_1, y_1)] - [\varphi(x_1, y_1)] - [\psi(x_1, y_1)] + [O]$ on E_2 over $k(x_1, y_1)$.
- Now consider $\mathrm{div}(f)$ inside the divisor group $\mathrm{Div}_{k(x_2,y_2)}(E_2)$ – i.e., with x_2, y_2 constant and x_1, y_1 the variables. Let us compute the zeroes and poles (and their orders) of f.
- If $P \in E_1(\overline{k(x_2, y_2)})$ is a point with $\varphi(P) = (x_2, y_2)$, then since D has the term $-[\varphi(x_1, y_1)]$ in it, D has a pole at P of order $e_\varphi(P)$ by the definition of the ramification index.
- In the same way, if Q has $\psi(Q) = (x_2, y_2)$ then the term $-[\psi(x_1, y_1)]$ yields a pole of order $e_\psi(Q)$ at Q.
- Similarly, if R has $(\varphi + \psi)(R) = (x_2, y_2)$ then $[(\varphi + \psi)(x_1, y_1)]$ gives a zero of order $e_{\varphi+\psi}(R)$ at R.

More Dual Isogenies, VIII

6. For any isogenies $\varphi, \psi : E_1 \to E_2$ we have $\widehat{\psi + \varphi} = \hat{\varphi} + \hat{\psi}$.

Proof (finale):

More Dual Isogenies, VIII

6. For any isogenies $\varphi, \psi : E_1 \to E_2$ we have $\widehat{\psi + \varphi} = \hat{\varphi} + \hat{\psi}$.

Proof (finale):

- So that means the divisor of f inside $\mathrm{Div}_{k(x_2,y_2)}(E_2)$ has the form $(\varphi + \psi)^*[(x_2, y_2)] - \varphi^*[(x_2, y_2)] - \psi^*[(x_2, y_2)] + \sum n_i P_i$ for some "constants" $P_i \in E_1(k)$.
- Since this is the divisor of a function, the sum of all the points resolves to the identity.
- Since $\sum n_i P_i$ is constant and does not depend on (x_2, y_2) this means the sum $(\widehat{\varphi + \psi})(x_2, y_2) - \hat{\varphi}(x_2, y_2) - \hat{\psi}(x_2, y_2)$ is a constant.
- Since it is the identity when $(x_2, y_2) = O$, it is always the identity.

More Dual Isogenies, IX

7. For any nonzero integer m we have $\widehat{[m]} = [m]$ and $\deg[m] = m^2$.

Proof:

More Dual Isogenies, IX

7. For any nonzero integer m we have $\widehat{[m]} = [m]$ and $\deg[m] = m^2$.

<u>Proof</u>:

- We clearly have $\widehat{[1]} = [1]$. Then $\widehat{[m]} = [m]$ for positive m follows by a trivial induction from (6), and for negative m it follows by noting that $\widehat{[-1]} = [-1]$ and using (5).
- For the degree of $[m]$ we note that by definition of the dual isogeny we have $[\deg[m]] = \widehat{[m]} \circ [m] = [m] \circ [m] = [m^2]$, and so $\deg[m] = m^2$.

More Dual Isogenies, X

8. We have $\deg \hat{\varphi} = \deg \varphi$ and $\hat{\hat{\varphi}} = \varphi$.

Proof:

More Dual Isogenies, X

8. We have $\deg \hat{\varphi} = \deg \varphi$ and $\hat{\hat{\varphi}} = \varphi$.

Proof:

- For the first, taking degrees in $[\deg \varphi] = \hat{\varphi} \circ \varphi$ and using (7) yields $(\deg \varphi)^2 = (\deg \hat{\varphi})(\deg \varphi)$.
- Cancelling yields the desired $\deg \hat{\varphi} = \deg \varphi$.
- For the second, observe by definition that $\hat{\hat{\varphi}} \circ \hat{\varphi} = [\deg \hat{\varphi}] = [\deg \varphi] = \varphi \circ \hat{\varphi}$ on E_1.
- So since $\hat{\varphi}$ is nonzero, the usual degree argument shows that $\hat{\hat{\varphi}} = \varphi$.

The m-Torsion Subgroup of E, I

Now we can properly study the m-torsion subgroup $E[m]$, since it is simply the kernel of the multiplication-by-m map $[m]$:

Theorem (The m-Torsion Subgroup)

Let m be a nonzero integer and E be an elliptic curve over k.

9. *For any nonzero integer m, if $\mathrm{char}(k)$ does not divide m then the m-torsion subgroup $E[m]$ is isomorphic to $(\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})$.*
10. *If $\mathrm{char}(k) = p$, then either $E[p^d] = \{O\}$ for all $d \geq 1$, or $E[p^d]$ is isomorphic to $\mathbb{Z}/p^d\mathbb{Z}$ for all $d \geq 1$.*

So we see that, aside from some potential drop in the p-power torsion in characteristic p (which is perhaps unsurprisingly caused by inseparability), the m-torsion subgroup always has two generators.

The m-Torsion Subgroup of E, II

9. For any nonzero integer m, if $\mathrm{char}(k)$ does not divide m then the m-torsion subgroup $E[m]$ is isomorphic to $(\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})$.

<u>Proof</u>:

The m-Torsion Subgroup of E, II

9. For any nonzero integer m, if $\mathrm{char}(k)$ does not divide m then the m-torsion subgroup $E[m]$ is isomorphic to $(\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})$.

<u>Proof</u>:

- By (7), the degree of $[m]$ is m^2 and as we have previously noted using the action on differentials, $[m]$ is separable whenever $\mathrm{char}(k)$ does not divide m.
- Therefore, by our properties of isogenies, we see that $\#E[m] = \# \ker[m] = \deg[m] = m^2$.
- The specific structural statement then follows essentially immediately from the structure theorem for finite abelian groups and the Chinese remainder theorem.

The m-Torsion Subgroup of E, III

9. For any nonzero integer m, if $\mathrm{char}(k)$ does not divide m then the m-torsion subgroup $E[m]$ is isomorphic to $(\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})$.

Proof (continued):

The m-Torsion Subgroup of E, III

9. For any nonzero integer m, if $\mathrm{char}(k)$ does not divide m then the m-torsion subgroup $E[m]$ is isomorphic to $(\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})$.

Proof (continued):

- Explicitly: for each prime $p|m$, the group $E[p]$ is an elementary abelian p-group of order p^2, hence is isomorphic to $(\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/p\mathbb{Z})$.
- Then for each prime power $p^d|m$, the group $E[p^d]$ has at most two components in its decomposition each of which has order at most p^d, but since $E[p^d]$ has order p^{2d}, that means $E[p^d]$ is isomorphic to $(\mathbb{Z}/p^d\mathbb{Z}) \times (\mathbb{Z}/p^d\mathbb{Z})$.
- Finally, $E[m]$ is isomorphic to the product of its prime-power torsion subgroups, and the product of these is isomorphic to $(\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})$.

The m-Torsion Subgroup of E, IV

In characteristic zero, we can use the previous result to give an explicit structure for the torsion subgroup E_{tor}, which (we emphasize again) is for the set of torsion points of E over the algebraic closure:

The m-Torsion Subgroup of E, IV

In characteristic zero, we can use the previous result to give an explicit structure for the torsion subgroup E_{tor}, which (we emphasize again) is for the set of torsion points of E over the algebraic closure:

Exercise: Show that when $\mathrm{char}(k) = 0$, the group E_{tor} of all torsion points on E is isomorphic to $(\mathbb{Q}/\mathbb{Z}) \times (\mathbb{Q}/\mathbb{Z})$. [Hint: Note that E_{tor} is the inverse limit of $E[n!]$ as $n \to \infty$.]

We will return to this result later when we talk about elliptic curves over $\mathbb{C}$, where this result will become very geometrically natural.

The m-Torsion Subgroup of E, V

10. If $\mathrm{char}(k) = p$, then either $E[p^d] = \{O\}$ for all $d \geq 1$, or $E[p^d]$ is isomorphic to $\mathbb{Z}/p^d\mathbb{Z}$ for all $d \geq 1$.

Proof:

The m-Torsion Subgroup of E, V

10. If $\mathrm{char}(k) = p$, then either $E[p^d] = \{O\}$ for all $d \geq 1$, or $E[p^d]$ is isomorphic to $\mathbb{Z}/p^d\mathbb{Z}$ for all $d \geq 1$.

Proof:

- As in (9) we know that $\deg[p^d] = p^{2d}$, but now since $p | p^d$, the map $[p^d]$ is inseparable.
- If φ is the pth-power Frobenius map, then as we showed in (2), $\hat{\varphi} \circ \varphi = [p]$, so $(\hat{\varphi} \circ \varphi)^d = [p^d]$.
- By our properties of isogenies, we have $\# E[p^d] = \# \ker[p^d] = \deg_s [p^d] = \deg_s (\hat{\varphi} \circ \varphi)^d = \deg_s (\hat{\varphi})^d$ because $\deg_s \varphi = 1$ as φ is purely inseparable.
- Now, since $\deg \hat{\varphi} = \deg \varphi = p$ by (8), and $\deg_s \hat{\varphi} \deg_i \hat{\varphi} = p$, we either have $\deg_s \hat{\varphi} = 1$ or $\deg_s \hat{\varphi} = p$.
- In the first case $\# E[p^d] = 1$ for all d, whence $E[p^d] = \{O\}$.
- In the second case $\# E[p^d] = p^d$ for all d. Then as in (9), we see $E[p^d]$ is isomorphic to $\mathbb{Z}/p^d\mathbb{Z}$.

The Hasse Bound, I

We can apply these results very fruitfully to establish some structural statements about the group of points on an elliptic curve over a finite field. We will need a few facts about quadratic forms.

- Recall that if G is an abelian group, a function $d : G \to \mathbb{Z}$ is a quadratic form when
  1. $d(-g) = g$ for all $g \in G$, and
  2. When the pairing $\langle \cdot, \cdot \rangle : G \times G \to \mathbb{Z}$ with $\langle g, h \rangle = \frac{1}{2}[d(g+h) - d(g) - d(h)]$ is bilinear (i.e., is $\mathbb{Z}$-linear in both g and h).
- We also say that a quadratic form is positive-definite when $d(g) \geq 0$ for all $g \in G$, with equality if and only if $g = 0$.

The Hasse Bound, II

As might be expected, the Cauchy-Schwarz inequality holds for positive-definite quadratic forms: $\langle g, h\rangle^2 \leq d(g)\, d(h)$.

The Hasse Bound, II

As might be expected, the Cauchy-Schwarz inequality holds for positive-definite quadratic forms: $\langle g, h\rangle^2 \leq d(g)\, d(h)$.

- To prove this note that if $d(g) = 0$ the result is trivial, and for $d(g) > 0$, for all integers a, b we have $a^2 d(g) - 2ab \langle g, h\rangle + b^2 d(h) = \langle ag - bh, ag - bh\rangle = d(ag - bh) \geq 0$ by bilinearity and positive-definiteness.
- Setting $a = \langle g, h\rangle$ and $b = d(g)$ then yields $d(g)[d(g)^2 d(h) - \langle g, h\rangle^2] \geq 0$, and so upon dividing by $d(g)$ we obtain the desired inequality.

The Hasse Bound, III

Now we can apply these facts to the Frobenius map to count the number of points on an elliptic curve over a finite field.

Theorem (Points on Elliptic Curves over $\mathbb{F}_q$)

Let $q = p^d$ be a prime power and let $E/\mathbb{F}_q$ be an elliptic curve defined over $\mathbb{F}_q$.

1. *The degree map $\deg : \mathrm{Hom}(E_1, E_2) \to \mathbb{Z}$ is a positive-definite quadratic form.*
2. *The Frobenius map $\varphi = \mathrm{Frob}_q$ has the property that $1 - \varphi$ is separable.*
3. *(Hasse Bound) The number of points on $E(\mathbb{F}_q)$ satisfies $|\#E(\mathbb{F}_q) - q - 1| \leq 2\sqrt{q}$.*

The Hasse Bound, IV

1. The degree map $\deg : \mathrm{Hom}(E_1, E_2) \to \mathbb{Z}$ is a positive-definite quadratic form.

Proof:

The Hasse Bound, IV

1. The degree map $\deg : \mathrm{Hom}(E_1, E_2) \to \mathbb{Z}$ is a positive-definite quadratic form.

Proof:

- First, $\deg(-\varphi) = \deg([-1])\deg(\varphi) = \deg(\varphi)$.
- Second, the associated pairing $\langle\varphi, \psi\rangle = \deg(\varphi + \psi) - \deg(\varphi) - \deg(\psi)$ is bilinear, because
$[\langle\varphi, \psi\rangle] = [\deg(\varphi + \psi)] - [\deg(\varphi)] - [\deg(\psi)]$
$= \widehat{\varphi + \psi} \circ (\varphi + \psi) - \hat{\varphi} \circ \varphi - \hat{\psi} \circ \psi$
$= (\hat{\varphi} + \hat{\psi}) \circ (\varphi + \psi) - \hat{\varphi} \circ \varphi - \hat{\psi} \circ \psi = \hat{\psi} \circ \varphi + \hat{\varphi} + \psi$ using (6).
- But now this last expression is linear in both φ and ψ by (6), so the pairing is bilinear.
- Finally, the degree map is clearly positive-definite since $\deg(\varphi) \geq 0$ with equality if and only if $\varphi = 0$.

The Hasse Bound, V

The degree map as a quadratic form on elliptic curves only having multiplication-by-m maps is not very exciting: $\deg[m] = m^2$, so the quadratic form is just $q(x) = x^2$ (not too exciting).

Here's a more interesting example:

Exercise: On the elliptic curve $y^2 = x^3 - x$ with the isogeny $[i](x, y) = (-x, iy)$ discussed previously, for $\varphi = [a] + [b][i]$ with $a, b \in \mathbb{Z}$, calculate $\hat{\varphi}$. Use the result to find $\deg \varphi$ and compute the associated quadratic form.

The Hasse Bound, V

The degree map as a quadratic form on elliptic curves only having multiplication-by-m maps is not very exciting: $\deg[m] = m^2$, so the quadratic form is just $q(x) = x^2$ (not too exciting).

Here's a more interesting example:

Exercise: On the elliptic curve $y^2 = x^3 - x$ with the isogeny $[i](x, y) = (-x, iy)$ discussed previously, for $\varphi = [a] + [b][i]$ with $a, b \in \mathbb{Z}$, calculate $\hat{\varphi}$. Use the result to find $\deg \varphi$ and compute the associated quadratic form.

If you have any kind of sensible guess about what $\hat{\varphi}$ is, you're almost certainly right.

The Hasse Bound, VI

2. The Frobenius map $\varphi = \mathrm{Frob}_q$ has the property that $1 - \varphi$ is separable.

Proof:

The Hasse Bound, VI

2. The Frobenius map $\varphi = \mathrm{Frob}_q$ has the property that $1 - \varphi$ is separable.

Proof:

- By light abuse of notation, we write $1 - \varphi$ instead of $[1] - \varphi$, since $[1] - \varphi$ is much uglier to read.
- Let ω be the invariant differential on E.
- By additivity of inverse image maps on differentials, we have $(1 - \varphi)^*\omega = [1]^*\omega - \varphi^*\omega = \omega$ since $\varphi^*\omega = 0$ because φ is inseparable hence is trivial on differentials).
- But now since $1 - \varphi$ is nontrivial on differentials, it is separable.

The Hasse Bound, VI

2. The Frobenius map $\varphi = \mathrm{Frob}_q$ has the property that $1 - \varphi$ is separable.

Proof:

- By light abuse of notation, we write $1 - \varphi$ instead of $[1] - \varphi$, since $[1] - \varphi$ is much uglier to read.
- Let ω be the invariant differential on E.
- By additivity of inverse image maps on differentials, we have $(1 - \varphi)^* \omega = [1]^* \omega - \varphi^* \omega = \omega$ since $\varphi^* \omega = 0$ because φ is inseparable hence is trivial on differentials).
- But now since $1 - \varphi$ is nontrivial on differentials, it is separable.

Exercise: Show more generally that $a + b\varphi$ is separable if and only if the characteristic p does not divide a.

The Hasse Bound, VII

3. (Hasse Bound) The number of points on $E(\mathbb{F}_q)$ satisfies $|\#E(\mathbb{F}_q) - q - 1| \leq 2\sqrt{q}$.

Proof:

The Hasse Bound, VII

3. (Hasse Bound) The number of points on $E(\mathbb{F}_q)$ satisfies $|\#E(\mathbb{F}_q) - q - 1| \leq 2\sqrt{q}$.

Proof:

- By basic Galois theory of finite fields, an element $x \in \overline{\mathbb{F}_q}$ lies in $\mathbb{F}_q$ if and only if $x^q = x$, which is to say, if and only if it is fixed by the qth-power Frobenius map φ.
- So now if we choose a Weierstrass equation for E over $\mathbb{F}_q$, since $E^{(q)} = E$ since the coefficients lie in $\mathbb{F}_q$ by hypothesis, we see a point $[X : Y : Z] \in E(\mathbb{F}_q)$ if and only if $\varphi(X : Y : Z) = [X : Y : Z]$, which is equivalent to saying that $[X : Y : Z] \in \ker(1 - \varphi)$.
- Hence $\#E(\mathbb{F}_q) = \#\ker(1 - \varphi)$.
- Let's now study this quantity.

The Hasse Bound, VIII

3. (Hasse Bound) The number of points on $E(\mathbb{F}_q)$ satisfies $|\#E(\mathbb{F}_q) - q - 1| \leq 2\sqrt{q}$.

Proof:

- We have $\#E(\mathbb{F}_q) = \#\ker(1 - \varphi)$.

The Hasse Bound, VIII

3. (Hasse Bound) The number of points on $E(\mathbb{F}_q)$ satisfies $|\#E(\mathbb{F}_q) - q - 1| \leq 2\sqrt{q}$.

Proof:

- We have $\#E(\mathbb{F}_q) = \#\ker(1-\varphi)$.
- By (2), the map $1-\varphi$ is separable, so $\#\ker(1-\varphi) = \deg(1-\varphi)$ by our results on isogenies.
- By (1), since the degree map is a positive-definite quadratic form, we may apply the Cauchy-Schwarz inequality to see that $\langle 1, -\varphi\rangle^2 \leq \deg[1]\deg(-\varphi) = q$ whence $|\langle 1, -\varphi\rangle| \leq \sqrt{q}$.
- Since $\langle 1, -\varphi\rangle = \frac{1}{2}[\deg(1-\varphi) - \deg(-\varphi) - \deg(1)] = \frac{1}{2}[\deg(1-\varphi) - q - 1]$, applying the results above yields $|\#E(\mathbb{F}_q) - q - 1| \leq 2\sqrt{q}$, as claimed.

The Hasse Bound, IX

We can give some intuition for why we might expect an inequality like the Hasse bound to hold.

- Assuming characteristic not equal to 2 for simplicity, consider a Weierstrass equation $y^2 = p(x)$ for E.

The Hasse Bound, IX

We can give some intuition for why we might expect an inequality like the Hasse bound to hold.

- Assuming characteristic not equal to 2 for simplicity, consider a Weierstrass equation $y^2 = p(x)$ for E.
- For each of the q possible finite values of x, there are either 2, 1, or 0 possible values of y, according to whether x is a nonzero square, zero, or a nonsquare.
- Since the squaring map $x \mapsto x^2$ is a homomorphism with kernel $\{\pm 1\}$ in $\mathbb{F}_q$, there are $(q-1)/2$ nonzero squares and $(q-1)/2$ nonsquares, so the expected number of values of y for any given x is equal to 1.
- Since there are q possible x, the expected number of finite points (x, y) is q, so together with the point at ∞, this gives an expected $q+1$ points on $E(\mathbb{F}_q)$.

The Hasse Bound, X

We trivially have the inequality $|\#E(\mathbb{F}_q) - q - 1| \leq q$ since the number of points is at least 1 and at most $2q + 1$.

- Hasse's bound is therefore a strengthening of the error term from this "trivial estimate" q to the estimate $2\sqrt{q}$.
- In fact we can give some statistical motivation for why this estimate on the deviation is somewhat reasonable:

Exercise: Suppose X is the sum of q independent random variables each of which takes the values 0 and 2 each with probability $1/2$. Show that the standard deviation of X is $\sqrt{q}$.

The Hasse Bound, XI

Exercise: Suppose X is the sum of q Bernoulli random variables each of which takes the values 0 and 2 each with probability 1/2. Show that the standard deviation of X is $\sqrt{q}$.

- If we approximate the point-count on E as the sum of q independent coin flips each of which yields 0 or 2 points, then by the exercise above, the standard deviation in the total number of points would be $\sqrt{q}$.
- The Hasse bound thus says our count will always be within 2 standard deviations of the mean.
- Of course, this is only a heuristic, since the actual variables themselves are not independent, but it's useful for seeing why the results come out near $\sqrt{q}$.

Winding Down

Although the Hasse bound may seem to be a rather pedestrian estimate on the number of points, it is actually quite deep.

- The error estimate in the Hasse bound is actually (more or less exactly) the Riemann hypothesis for elliptic curves.
- More precisely, there is a series of conjectures made by Weil, known as the Weil conjectures, which described various algebro-geometric results about the zeta functions of algebraic varieties over finite fields.
- The conjectures themselves were proven by Weil, Dwork, Grothendieck, and Deligne between the 1950s and 1970s.
- I will give the statements next time, and we can then show all of them hold for elliptic curves.

Summary

We introduced dual isogenies and established many of their properties.

We used dual isogenies to establish the Hasse bound.

Next lecture: The Weil conjectures, the Tate module.