The Group Law on Weierstrass Elliptic Curves

An Elementary Formal Proof in Any Characteristic

David Kurniadi Angdinata [1] Junyan Xu [2]

[1] **London School of Geometry and Number Theory, UK**

[2] Cancer Data Science Laboratory, National Cancer Institute, Bethesda, MD, USA

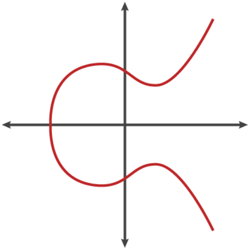
Fourteenth International Conference on Interactive Theorem Proving

Wednesday, 2 August 2023

Elliptic curves

An **elliptic curve** over a field F is a pair $(E, \mathcal{O})$:

- E is a *smooth projective curve* of *genus one* defined over F
- $\mathcal{O}$ is a distinguished point on E defined over F



Applications:

- computational mathematics
  - primality testing, integer factorisation, public-key cryptography
- algebraic geometry and number theory
  - Fermat's last theorem, the Birch and Swinnerton-Dyer conjecture

Weierstrass equations

Theorem (corollary of *Riemann–Roch*)

Any elliptic curve E over F can be given by $E(X, Y) = 0$, where

$$E(X, Y) := Y^2 + a_1 XY + a_3 Y - (X^3 + a_2 X^2 + a_4 X + a_6),$$

for some $a_i \in F$ such that $\Delta(a_i) \neq 0$, with $\mathcal{O}$ the point at infinity.

This is the **Weierstrass model** for E, but E has other models.

- If $\mathrm{char}(F) \neq 2, 3$, then E has a **short Weierstrass model**

  $$E(X, Y) := Y^2 - (X^3 + aX + b), \qquad a, b \in F,$$

  where $\Delta(a, b) = -16(4a^3 + 27b^2)$.
- If $\mathrm{char}(F) \neq 2$, then E has an **Edwards model**

  $$E(X, Y) := X^2 + Y^2 - (1 + dX^2Y^2), \qquad d \in F \setminus \{0, 1\},$$

  with $\mathcal{O} := (1, 0)$.

Weierstrass equations

Theorem (corollary of *Riemann–Roch*)

Any elliptic curve E over F can be given by $E(X, Y) = 0$, where

$$E(X, Y) := Y^2 + a_1XY + a_3Y - (X^3 + a_2X^2 + a_4X + a_6),$$

for some $a_i \in F$ such that $\Delta(a_i) \neq 0$, with $\mathcal{O}$ the point at infinity.

In the Weierstrass model, an **elliptic curve** over F is the data of:

- five coefficients $a_1, a_2, a_3, a_4, a_6 \in F$, and
- a proof that $\Delta(a_1, a_2, a_3, a_4, a_6) \neq 0$.

```
structure weierstrass_curve (F : Type) := (a₁ a₂ a₃ a₄ a₆ : F)

def weierstrass_curve.Δ {F : Type} [comm_ring F] (W : weierstrass_curve F) : F :=
  -(E.a₁^2 + 4*E.a₂)*(E.a₁^2*E.a₆ + 4*E.a₂*E.a₆ - E.a₁*E.a₃*E.a₄ + E.a₂*E.a₃^2 - E.a₄^2)
   - 8*(2*E.a₄ + E.a₁*E.a₃)^3 - 27*(E.a₃^2 + 4*E.a₆)^2
   + 9*(E.a₁^2 + 4*E.a₂)*(2*E.a₄ + E.a₁*E.a₃)*(E.a₃^2 + 4*E.a₆)

structure elliptic_curve (F : Type) [comm_ring F] extends weierstrass_curve F :=
  (Δ' : units F) (coe_Δ' : ↑Δ' = to_weierstrass_curve.Δ)
```

Weierstrass equations

Theorem (corollary of *Riemann–Roch*)

Any elliptic curve E over F can be given by $E(X, Y) = 0$, where

$$E(X, Y) := Y^2 + a_1XY + a_3Y - (X^3 + a_2X^2 + a_4X + a_6),$$

for some $a_i \in F$ such that $\Delta(a_i) \neq 0$, with $\mathcal{O}$ the point at infinity.

In the Weierstrass model, a **point** on E is either:

- the point at infinity $\mathcal{O}$, or
- two affine coordinates $x, y \in F$ and a proof that $(x, y) \in E$.

```
variables {F : Type} [field F] (E : elliptic_curve F)

def polynomial : F[X][Y] :=
 Y^2 + C (C E.a₁*X + C E.a₃)*Y - C (X^3 + C E.a₂*X^2 + C E.a₄*X + C E.a₆)

def equation (x y : F) : Prop := (E.polynomial.eval (C y)).eval x = 0

inductive point
  | zero
  | some {x y : F} (h : E.equation x y)
```

Theorem (the group law)

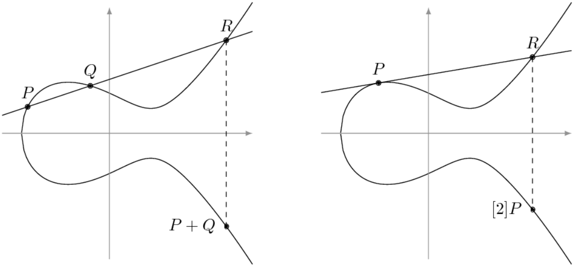
The points of E form an abelian group under a geometric addition law.

Identity is given by $\mathcal{O}$.

```
instance : has_zero E.point := ⟨zero⟩
```

Negation and addition are characterised by

$$P + Q + R = 0 \quad \iff \quad P, Q, R \text{ are collinear.}$$



Group law

Theorem (the group law)

The points of E form an abelian group under a geometric addition law.

Negation is given by $-(x, y) := (x, \sigma(y))$, where

$$\sigma(Y) := -Y - a_1X - a_3.$$

```
def neg_Y (x y : F) : F := -y - E.a₁ * x + E.a₃

lemma equation_neg {x y : F} : E.equation x y → E.equation x (E.neg_Y x y) := ...

def neg : E.point → E.point
  | zero := zero
  | (some h) := some (equation_neg h)

instance : has_neg E.point := ⟨neg⟩
```

<u>**Note:**</u> in the **coordinate ring** $F[E] := F[X, Y]/(E(X, Y))$,

$$-(Y \cdot \sigma(Y)) = Y^2 + a_1XY + a_3Y \equiv X^3 + a_2X^2 + a_4X + a_6.$$

Group law

Theorem (the group law)

The points of E form an abelian group under a geometric addition law.

Addition is given by $(x_1, y_1) + (x_2, y_2) := -(x_3, y_3)$, where

$$x_3 := \lambda^2 + a_1\lambda - a_2 - x_1 - x_2, \qquad y_3 := \lambda(x_3 - x_1) + y_1.$$

```
def add : E.point → E.point → E.point
  | zero P := P
  | P zero := P
  | (some h₁) (some h₂) := some (equation_add h₁ h₂)

instance : has_add E.point := ⟨add⟩
```

Here,

$$\lambda := \begin{cases} \dfrac{y_1 - y_2}{x_1 - x_2} & \text{if } x_1 \neq x_2, \\ \dfrac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{y_1 - \sigma(y_1)} & \text{if } y_1 \neq \sigma(y_1), \\ \infty & \text{otherwise.} \end{cases}$$

Attempts at proof

One may attempt to prove the axioms directly.

```
instance : add_group E.point :=
 { zero         := zero,
   neg          := neg,
   add          := add,
   zero_add     := rfl,     -- by definition
   add_zero     := rfl,     -- by definition
   add_left_neg := ...,     -- by cases
   add_comm     := ...,     -- by cases
   add_assoc    := sorry }  -- seems impossible?
```

Associativity is a proof that

$$(P + Q) + R = P + (Q + R),$$

where each $+$ has five cases!

In the generic case, this is an equality of polynomials with 26,082 terms.

In contrast, the `ring` tactic in Lean can handle at most 1,000 terms.

Attempts at proof

Associativity is known to be mathematically difficult with many proofs.

Proof 1: just do it.

- elementary but slow
- several known formalisations
  - Théry (Coq, 2007): short Weierstrass model $Y^2 = X^3 + aX + b$
  - Hales, Raya (Isabelle, 2020): Edwards model $X^2 + Y^2 = 1 + dX^2Y^2$
  - Fox, Gordon, Hurd (HOL4, 2006): long Weierstrass model $Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6$ but no associativity

Proof 2: ad-hoc argument with projective geometry.

- only works generically via *Cayley–Bacharach*
- no known formalisations
  - our original attempt

Attempts at proof

One may instead identify the set of points $E(F)$ with a known group.

Proof 3: identify with a quotient of $\mathbb{C}$ by the *fundamental lattice* Λ_E.

- only works in characteristic zero via *uniformisation*
- no known formalisations
    - needs a lot of theory

Proof 4: identify with the *degree zero Weil divisor class group* $\mathrm{Pic}_F^0(E)$.

- algebro-geometric and usually uses *Riemann–Roch*
- one known formalisation
    - Bartzia, Strub (10,000 lines of Coq, 2014): short Weierstrass model

Proof 5: identify with the *ideal class group* $\mathrm{Cl}(F[E])$.

- purely algebraic and uses commutative algebra
- one known formalisation
    - our final proof (1,000 lines of Lean, 2023): long Weierstrass model

Sketch of proof

Proof of the group law.

1. Construct a function $E(F) \to \mathrm{Cl}(F[E])$.
2. Prove that $E(F) \to \mathrm{Cl}(F[E])$ respects addition.
3. Prove that $E(F) \to \mathrm{Cl}(F[E])$ is injective. $\square$

Here, the **ideal class group** $\mathrm{Cl}(R)$ of an integral domain R is the quotient group of *invertible fractional ideals* by *principal fractional ideals*.

Example

Any nonzero ideal $I \trianglelefteq R$ such that $I \cdot J$ is principal for some ideal $J \trianglelefteq R$ is an invertible fractional ideal of R.

Ideal class groups were formalised in Lean's mathematical library `mathlib` by Baanen, Dahmen, Narayanan, Nuccio (2021).

Key: the coordinate ring $F[E]$ is an integral domain.

Sketch of proof

Proof of the group law.

1. Construct a function $E(F) \to \mathrm{Cl}(F[E])$. ✓
2. Prove that $E(F) \to \mathrm{Cl}(F[E])$ respects addition. ✓
3. Prove that $E(F) \to \mathrm{Cl}(F[E])$ is injective. □

Consider the function `point.to_class` given by

$$\begin{array}{rcl} E(F) & \longrightarrow & \mathrm{Cl}(F[E]) \\ \mathcal{O} & \longmapsto & [(1)] \\ (x, y) & \longmapsto & [(X - x, Y - y)] \end{array}.$$

Note: $(X - x, Y - y)$ is invertible, since

$$(X - x, Y - y) \cdot (X - x, Y - \sigma(y)) = (X - x).$$

The function `point.to_class` respects addition, since

$$(X - x_1, Y - y_1) \cdot (X - x_2, Y - y_2) \cdot (X - x_3, Y - \sigma(y_3)) = (Y - \lambda(X - x_3) - y_3).$$

Proof of injectivity

Theorem (Xu, 2022)

The function `point.to_class` *is injective.*

Key: $F[E] = F[X,Y]/(E(X,Y))$ is free over $F[X]$ with basis $\{1, Y\}$, so it has a norm $\mathrm{Nm} : F[E] \to F[X]$ given by $\mathrm{Nm}(f) := \det([\cdot f])$.

Lemma (A)

If $f \in F[E]$, *then* $\deg(\mathrm{Nm}(f)) \neq 1$.

Proof of Lemma (A).

Let $f = p + qY$ for $p, q \in F[X]$. Then

$$\begin{aligned}\mathrm{Nm}(f) &\equiv \det \begin{pmatrix} p & q \\ q(X^3 + a_2X^2 + a_4X + a_6) & p - q(a_1X + a_3) \end{pmatrix} \\ &= p^2 - pq(a_1X + a_3) - q^2(X^3 + a_2X^2 + a_4X + a_6).\end{aligned}$$

Then $\deg(\mathrm{Nm}(f)) = \max(2\deg(p), 2\deg(q) + 3)$. $\square$

Proof of injectivity

Theorem (Xu, 2022)

The function `point.to_class` is injective.

Key: $F[E] = F[X,Y]/(E(X,Y))$ is free over $F[X]$ with basis $\{1, Y\}$, so it has a norm $\mathrm{Nm} : F[E] \to F[X]$ given by $\mathrm{Nm}(f) := \det([\cdot f])$.

Lemma (B)

If $f \in F[E]$, then $\deg(\mathrm{Nm}(f)) = \dim_F(F[E]/(f))$.

Proof of Lemma (B).

Multiplication by f has Smith normal form

$$[\cdot f] \sim \begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix}, \qquad p, q \in F[X].$$

- Taking determinants gives $\mathrm{Nm}(f) = pq$.
- Taking quotients gives $F[E]/(f) \cong F[X]/(p) \oplus F[X]/(q)$. $\square$

Proof of injectivity

Theorem (Xu, 2022)

The function `point.to_class` *is injective.*

Key: $F[E] = F[X, Y]/(E(X, Y))$ is free over $F[X]$ with basis $\{1, Y\}$, so it has a norm $\mathrm{Nm} : F[E] \to F[X]$ given by $\mathrm{Nm}(f) := \det([\cdot f])$.

Proof of Theorem.

Suffices to show if $(x, y) \in E(F)$, then $(X - x, Y - y)$ is not principal.

Suppose otherwise that $(X - x, Y - y) = (f)$ for some $f \in F[E]$. Then

$$F \overset{1^{\text{st}}\text{iso}}{\cong} F[X, Y]/(X - x, Y - y) \overset{3^{\text{rd}}\text{iso}}{\cong} F[E]/(X - x, Y - y) = F[E]/(f).$$

Taking dimensions gives

$$1 = \dim_F(F) = \dim_F(F[E]/(f)) \overset{(B)}{=} \deg(\mathrm{Nm}(f)) \overset{(A)}{\neq} 1.$$

Contradiction! $\square$

Concluding retrospectives

Some thoughts:

- proof works for nonsingular points of Weierstrass curves
- formalisation encouraged proof accessible to undergraduates
- heavy use of linear algebra and ring theory in `mathlib`
- fully integrated to `mathlib` and even ported to `mathlib4`

Some projects:

- division polynomials, torsion subgroups, and Tate modules
- elliptic curves over discrete valuation rings and the reduction map
- verification of computational algorithms and cryptographic protocols
- equivalence with scheme-theoretic definitions via Riemann–Roch
- elliptic curves over specific fields: finite fields, local fields, number fields, global function fields, complete fields