Algebraising foundations of elliptic curves

Formalisation of Mathematics with Interactive Theorem Provers

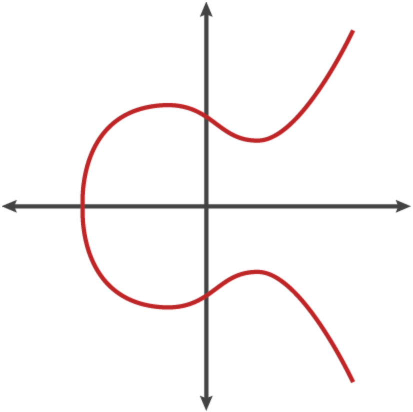
David Kurniadi Angdinata (with Junyan Xu)

London School of Geometry and Number Theory

Thursday, 13 February 2025

Introduction

Elliptic curves are algebraic curves given by cubic equations.



Their set of points can be endowed with a *group law*.

Motivation

Why should we care about elliptic curves?

They are prevalent in modern number theory.

- Wiles proved *Fermat's last theorem* by drawing a correspondence between certain elliptic curves and certain *modular forms*.
- The *Birch and Swinnerton-Dyer conjecture* predicts the arithmetic behaviour of elliptic curves based on their *L-functions*.

They see many computational applications.

- Intractability of the *discrete logarithm problem* for elliptic curves forms the basis behind many public key cryptographic protocols.
- The *Atkin–Morain primality test* and *Lenstra's factorisation method* use elliptic curves and are two of the fastest known algorithms.

Formalising the theory of elliptic curves would be great!

History

There is much previous work in various interactive theorem provers.

- Anthony Fox, Mike Gordon, and Joe Hurd (2006) formalised a definition of an elliptic curve in HOL4 over an arbitrary field F.
- Laurent Théry (2007) formalised a *direct* proof of the group law on an elliptic curve in Coq, assuming $\mathrm{char}(F) \neq 2, 3$.
- Evmorfia-Iro Bartzia and Pierre-Yves Strub (2014) formalised a *conceptual* proof of the group law in Coq, assuming $\mathrm{char}(F) \neq 2, 3$.
- Thomas Hales and Rodrigo Raya (2020) formalised a *direct* proof of the group law in Isabelle, assuming $\mathrm{char}(F) \neq 2$.
- Junyan Xu and I (2023) formalised a *novel conceptual* proof of the group law in Lean, *with no assumptions on* $\mathrm{char}(F)$.
- Junyan Xu and I (2024) discovered gaps in the standard proof of the *multiplication-by-n formula* on an elliptic curve, filled them in with *novel* arguments, and formalised the entire proof in Lean.

Elliptic curves

An **elliptic curve** over a field F is a smooth projective curve E over F of genus one, equipped with a distinguished point $\mathcal{O}$ defined over F.

These are all notions from modern algebraic geometry.

- A **curve** is a variety [1] of dimension one as a topological space.
- **Projective** means there is a closed immersion $E \hookrightarrow \mathrm{Proj}(F[X_i])$.
- **Smooth** essentially means all $\mathcal{O}_{E,\overline{x}}$ are regular local rings.
- **Genus** is the dimension of $H^1(E, \mathcal{O}_E)$ as an F-vector space.

In `mathlib`, we have schemes (Aug 2020), integral schemes (Dec 2021), projective schemes (Apr 2022), finite type morphisms (Oct 2022), Krull dimensions (May 2023), separated morphisms (Jun 2024), smooth morphisms (Jul 2024), and sheaf cohomology (Jul 2024).

Thanks to the work of Andrew Yang, Christian Merten, Joël Riou, and others, we can *almost* formalise the definition of elliptic curves in Lean!

[1] integral separated scheme of finite type over $\mathrm{Spec}(F)$

Elliptic curves in `mathlib`

Corollary (of the Riemann–Roch theorem)

The set of points of an elliptic curve over F is the vanishing locus of

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

for some $a_i \in F$ such that $\Delta \neq 0$, [2] *with an extra point at infinity $\mathcal{O}$.*

In other words, there is an equivalence of categories

$$\{\text{elliptic curves over } F\} \cong \{(a_1, a_2, a_3, a_4, a_6) \in F^5 \text{ such that } \Delta \neq 0\}.$$

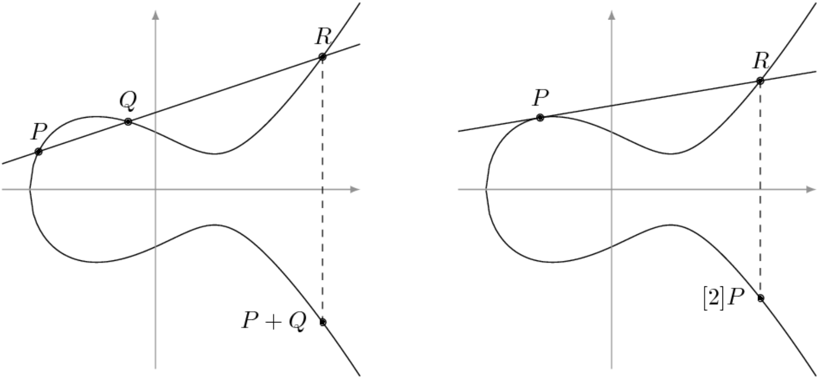
In `mathlib`, an **elliptic curve** E over a ring R is the data of a tuple $(a_1, a_2, a_3, a_4, a_6) \in R^5$ and a proof that $\Delta \in R^\times$. A **point** on E is then a sum type of $\mathcal{O}$ and **affine points** $(x, y) \in R^2$ such that $\mathcal{E}(x, y) = 0$.

The arithmetic can be formalised independently of the algebraic geometry.

[2] $\Delta := -(a_1^2+4a_2)^2(a_1^2a_6+4a_2a_6-a_1a_3a_4+a_2a_3^2-a_4^2)-8(2a_4+a_1a_3)^3-27(a_3^2+4a_6)^2+9(a_1^2+4a_2)(2a_4+a_1a_3)(a_3^2+4a_6)$

The group law

The set of points $E(F)$ can be endowed with a geometric addition law.



Theorem (the group law)

This addition law makes $E(F)$ an abelian group with identity $\mathcal{O}$.

The group law in `mathlib`

In `mathlib`, the addition law is given by explicit rational functions.

For instance, $(x_1, y_1) + (x_2, y_2) := (x_3, y_3)$, where

$$\begin{aligned} x_3 &:= \lambda^2 + a_1\lambda - a_2 - x_1 - x_2, \\ y_3 &:= -\lambda(x_3 - x_1) - y_1 - a_1x_3 - a_3. \end{aligned}$$

Here, the slope λ is given by

$$\lambda := \begin{cases} \dfrac{y_1 - y_2}{x_1 - x_2} & \text{if } x_1 \neq x_2, \\ \dfrac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x + a_3} & \text{if } y_1 \neq -y_1 - a_1x - a_3, \\ \infty & \text{otherwise.} \end{cases}$$

All of the axioms for an abelian group are easy *except for associativity*.

Associativity

Associativity is the statement that, for all $P, Q, R \in E(F)$,

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

In the generic case, [3] checking that their X-coordinates are equal is an equality of multivariate polynomials with 26,082 terms!

When $\mathrm{char}(F) \neq 2, 3$, a linear change of variables reduces $\mathcal{E}$ to

$$\mathcal{E}' := Y^2 - (X^3 + aX + b),$$

for some $a, b \in F$ such that $-16(4a^3 + 27b^2) \neq 0$.

This computation reduces to an equality of polynomials with 2,636 terms.

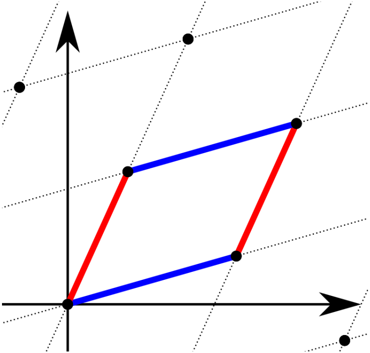
Automation in an interactive theorem prover enables manipulation of multivariate polynomials with at most 5,000 terms.

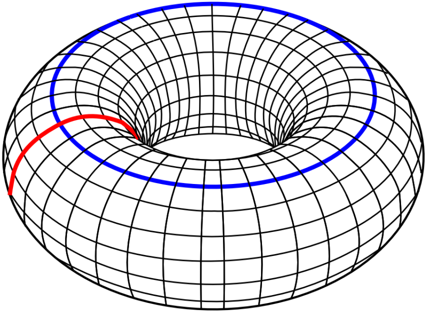
[3] P, Q, R, $P + Q$, $P + R$, and $Q + R$ are not $\mathcal{O}$ and have distinct X-coordinates

A complex uniformisation

Why should there be a group law in the first place?

Over $F = \mathbb{C}$, an elliptic curve is just a complex torus $\mathbb{C}/\Lambda_E$.





There is an explicit bijection from $E(\mathbb{C})$ to $\mathbb{C}/\Lambda_E$ that preserves the addition law, so the group law on $\mathbb{C}/\Lambda_E$ can be pulled back to $E(\mathbb{C})$.

An algebraic variant

In general, Riemann–Roch gives an explicit bijection from $E(F)$ to the degree-zero divisor class group $\mathrm{Pic}_F^0(E)$ that preserves the addition law.

While `mathlib` does not have divisors, it has ideals of integral domains D and the ideal class group [4] $\mathrm{Cl}(D)$, which are *purely commutative algebra*.

The map $E(F) \to \mathrm{Pic}_F^0(E)$ translates to the map

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

where D is the integral domain $F[X, Y]/(\mathcal{E})$.

Theorem (Xu)

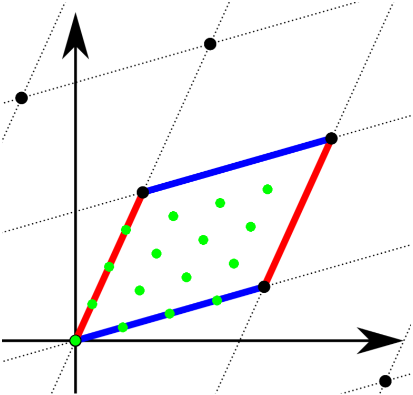
Proving that this map is injective only needs linear algebra.

[4] group of invertible fractional ideals of D

Multiplication by n

For each $n \in \mathbb{Z}$ and each point $P \in E(F)$, define $[n](P) := \underbrace{P + \cdots + P}_{n}$.

How many points $P \in E(\mathbb{C})$ are there such that $[n](P) = \mathcal{O}$?



The n-torsion subgroup

For each $n \in \mathbb{Z}$, define $E_F[n] := \{P \in E(F) : [n](P) = \mathcal{O}\}$.

Theorem (the n-torsion subgroup structure)

If $\operatorname{char}(F) \nmid n$*, then* $E_{\overline{F}}[n]$ *is isomorphic to* $(\mathbb{Z}/n)^2$*.*

When $\operatorname{char}(F) \neq p$, the p**-adic Tate module** $T_p E_{\overline{F}}$ sits in the diagram

$$\begin{tikzcd}
T_p E_{\overline{F}} \arrow[d, "\sim"] & := \varprojlim \Big(\cdots \arrow[r, "{[p]}"] & E_{\overline{F}}[p^3] \arrow[r, "{[p]}"] \arrow[d, "\sim"] & E_{\overline{F}}[p^2] \arrow[r, "{[p]}"] \arrow[d, "\sim"] & E_{\overline{F}}[p] \Big) \arrow[d, "\sim"] \\
\mathbb{Z}_p^2 & := \varprojlim \Big(\cdots \arrow[r, "\bmod p^3"] & (\mathbb{Z}/p^3)^2 \arrow[r, "\bmod p^2"] & (\mathbb{Z}/p^2)^2 \arrow[r, "\bmod p"] & (\mathbb{Z}/p)^2 \Big).
\end{tikzcd}$$

This makes $T_p E_{\overline{F}}$ a *two-dimensional p-adic Galois representation.* [5]

[5] crucial in the Mordell–Weil theorem, Tate's isogeny theorem, Serre's open image theorem, Wiles's modularity theorem, the Birch and Swinnerton-Dyer conjecture, etc.

An infamous exercise

The Arithmetic of Elliptic Curves by Silverman gives a formula for $[n](P)$.

Exercise (3.7(d))

Let $n \in \mathbb{Z}$. Prove that for any affine point $(x, y) \in E(F)$,

$$[n]((x,y)) = \left(\frac{\phi_n(x,y)}{\psi_n(x,y)^2}, \frac{\omega_n(x,y)}{\psi_n(x,y)^3} \right).$$

Silverman gives definitions for $\phi_n, \omega_n \in F[X, Y]$ in terms of certain *division polynomials* $\psi_n \in F[X, Y]$, which feature in Schoof's algorithm.

Conjecture

No one has done Exercise 3.7(d) purely algebraically.

This formula does not account for affine points $(x, y) \in E(F)$ such that $\psi_n(x, y) = 0$, which occurs precisely when $[n]((x, y)) = \mathcal{O}$.

Projective coordinates

In projective coordinates, the multiplication-by-n formula becomes

$$[n]((x,y)) = [(\phi_n(x,y)\psi_n(x,y), \omega_n(x,y), \psi_n(x,y)^3)].$$

In `mathlib`, a **projective point** is a class of $(x,y,z) \in F^3$ such that

$$y^2z + a_1xyz + a_3yz^2 = x^3 + a_2x^2z + a_4xz^2 + a_6z^3.$$

The point at infinity becomes $[(0,1,0)]$.

More naturally, in projective coordinates with weights $(2,3,1)$,

$$[n]((x,y)) = [(\phi_n(x,y), \omega_n(x,y), \psi_n(x,y))].$$

In `mathlib`, a **Jacobian point** is a class of $(x,y,z) \in F^3$ such that

$$y^2 + a_1xyz + a_3yz^3 = x^3 + a_2x^2z^2 + a_4xz^4 + a_6z^6.$$

The point at infinity becomes $[(1,1,0)]$.

The polynomials ψ_n

For any ring R, the *n*-**th division polynomial** $\psi_n \in R[X, Y]$ is given by

$$\begin{aligned}
\psi_0 &:= 0, \\
\psi_1 &:= 1, \\
\psi_2 &:= 2Y + a_1 X + a_3, \\
\psi_3 &:= 3X^4 + (a_1^2 + 4a_2)X^3 + 3(2a_4 + a_1 a_3)X^2 + 3(a_3^2 + 4a_6)X + (a_1^2 a_6 + 4a_2 a_6 - a_1 a_3 a_4 + a_2 a_3^2 - a_4^2), \\
\psi_4 &:= \psi_2 \left(\begin{array}{c} 2X^6 + (a_1^2 + 4a_2)X^5 + 5(2a_4 + a_1 a_3)X^4 + 10(a_3^2 + 4a_6)X^3 + 10(a_1^2 a_6 + 4a_2 a_6 - a_1 a_3 a_4 + a_2 a_3^2 - a_4^2)X^2 \\ + ((a_1^2 + 4a_2)(a_1^2 a_6 + 4a_2 a_6 - a_1 a_3 a_4 + a_2 a_3^2 - a_4^2) - (2a_4 + a_1 a_3)(a_3^2 + 4a_6))X \\ + ((2a_4 + a_1 a_3)(a_1^2 a_6 + 4a_2 a_6 - a_1 a_3 a_4 + a_2 a_3^2 - a_4^2) - (a_3^2 + 4a_6)^2) \end{array} \right), \\
\psi_{2n+1} &:= \psi_{n+2}\psi_n^3 - \psi_{n-1}\psi_{n+1}^3, \\
\psi_{2n} &:= \frac{\psi_{n-1}^2 \psi_n \psi_{n+2} - \psi_{n-2}\psi_n \psi_{n+1}^2}{\psi_2}, \\
\psi_{-n} &:= -\psi_n.
\end{aligned}$$

In `mathlib`, ψ_n is defined in terms of some polynomial $\Psi_n \in R[X]$ such that $\psi_n = \Psi_n$ when n is odd and $\psi_n = \psi_2 \Psi_n$ when n is even.

The polynomials ϕ_n

The polynomial $\phi_n \in R[X, Y]$ is given by

$$\phi_n := X\psi_n^2 - \psi_{n+1}\psi_{n-1}.$$

In `mathlib`, ϕ_n is defined in terms of some polynomial $\Phi_n \in R[X]$, since

$$\begin{aligned} \psi_2^2 &= (2Y + a_1X + a_3)^2 \\ &= 4(Y^2 + a_1XY + a_3Y) + a_1^2X^2 + 2a_1a_3X + a_3^2 \\ &\equiv 4X^3 + b_2X^2 + 2b_4X + b_6 \mod \mathcal{E}, \end{aligned}$$

so ψ_n^2 and $\psi_{n+1}\psi_{n-1}$ are congruent to polynomials in $R[X]$.

Exercise (3.7(c))

Let $n \in \mathbb{Z}$. Prove that ϕ_n and ψ_n^2 have no common roots.

This *needs* Exercise 3.7(d) and the assumption that $\Delta \neq 0$.

The polynomials ω_n

The polynomial $\omega_n \in R[X, Y]$ is given by

$$\omega_n := \frac{1}{2}\left(\frac{\psi_{2n}}{\psi_n} - a_1\phi_n\psi_n - a_3\psi_n^3\right).$$

Lemma (Xu)

Let $n \in \mathbb{Z}$. Then $\psi_{2n}/\psi_n - a_1\phi_n\psi_n - a_3\psi_n^3$ is divisible by 2 in $\mathbb{Z}[a_i, X, Y]$.

Example ($a_1 = a_3 = 0$)

$$\omega_2 = \frac{2X^6+4a_2X^5+10a_4X^4+40a_6X^3+10(4a_2a_6-a_4^2)X^2+(4a_2(4a_2a_6-a_4^2)-8a_4a_6)X+(2a_4(4a_2a_6-a_4^2)-16a_6^2)}{2}.$$

Define ω_n as the image of the quotient under $\mathbb{Z}[a_i, X, Y] \to R[X, Y]$.

When $n = 4$, this quotient has 15,049 terms.

Elliptic divisibility sequences

Integrality relies on the fact that ψ_n is an **elliptic divisibility sequence**.

Exercise (3.7(g))

For all $n, m, r \in \mathbb{Z}$, prove that $\psi_n \mid \psi_{nm}$ and

$$\mathrm{ES}(n, m, r) : \psi_{n+m}\psi_{n-m}\psi_r^2 = \psi_{n+r}\psi_{n-r}\psi_m^2 - \psi_{m+r}\psi_{m-r}\psi_n^2.$$

Note that $\mathrm{ES}(n+1, n, 1)$ gives ψ_{2n+1} and $\mathrm{ES}(n+1, n-1, 1)$ gives ψ_{2n}.

Surprisingly, this needs the stronger result that ψ_n is an **elliptic net**.

Theorem (Xu)

Let $n, m, r, s \in \mathbb{Z}$. Then

$$\begin{aligned}\mathrm{EN}(n, m, r, s) : \psi_{n+m}\psi_{n-m}\psi_{r+s}\psi_{r-s} = {} & \psi_{n+r}\psi_{n-r}\psi_{m+s}\psi_{m-s} \\ & - \psi_{m+r}\psi_{m-r}\psi_{n+s}\psi_{n-s}.\end{aligned}$$

Xu gave an elegant proof of this on Math Stack Exchange.

Ellipticity of ψ_n

It suffices to prove $\mathrm{EN}(n,m,r,s)$ by strong induction on n assuming that $n,m,r,s \in \mathbb{N}$ [6] such that $n > m > r > s$. Firstly,

$$\mathrm{EN}(n,m,1,0) = \mathrm{EN}(\tfrac{n+m+1}{2}, \tfrac{n+m-1}{2}, \tfrac{n-m+1}{2}, \tfrac{n-m-1}{2}).$$

If $n = m+1$, then $\mathrm{EN}(m+1,m,1,0)$ holds by definition of ψ_{2n+1}. Otherwise $n > m+1$, then inductive hypothesis applies since $\frac{n+m+1}{2} < n$. This gives $\mathrm{EN}(n,m,1,0)$ for all $n,m > 1$. Furthermore,

$$\begin{aligned}
\mathrm{EN}(n,m,r,0) &= \psi_r^2 \cdot \mathrm{EN}(n,m,1,0) - \psi_m^2 \cdot \mathrm{EN}(n,r,1,0) + \psi_n^2 \cdot \mathrm{EN}(m,r,1,0),\\
\mathrm{EN}(n,m,r,1) &= \psi_{r+1}\psi_{r-1} \cdot \mathrm{EN}(n,m,1,0) - \psi_{m+1}\psi_{m-1} \cdot \mathrm{EN}(n,r,1,0) + \psi_{n+1}\psi_{n-1} \cdot \mathrm{EN}(m,r,1,0).
\end{aligned}$$

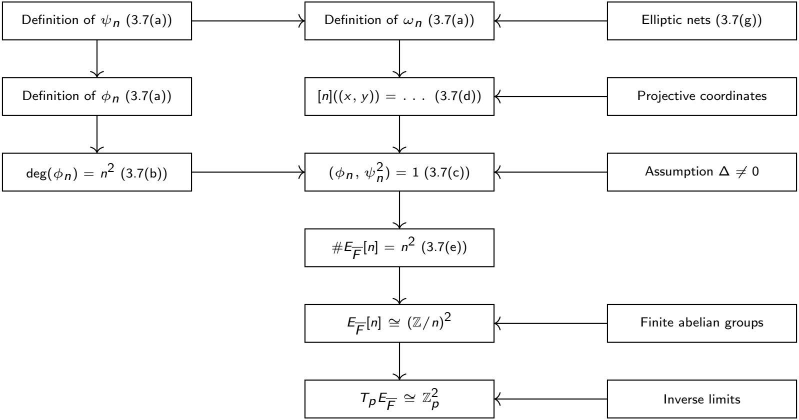
This gives $\mathrm{EN}(n,m,r,0)$ and $\mathrm{EN}(n,m,r,1)$ for all $n,m,r > 1$. Finally,

$$\begin{aligned}
\mathrm{EN}(n,m,r,s) = {} & \psi_m^2 \cdot \mathrm{EN}(n,r,s,1) + \psi_{m+1}\psi_{m-1} \cdot \mathrm{EN}(n,r,s,0) + \psi_{m+r}\psi_{m-r} \cdot \mathrm{EN}(n,s,1,0)\\
& - \psi_r^2 \cdot \mathrm{EN}(n,m,s,1) - \psi_{r+1}\psi_{r-1} \cdot \mathrm{EN}(n,m,s,0) - \psi_{m+s}\psi_{m-s} \cdot \mathrm{EN}(n,r,1,0)\\
& + \psi_s^2 \cdot \mathrm{EN}(n,m,r,1) + \psi_{s+1}\psi_{s-1} \cdot \mathrm{EN}(n,m,r,0) + \psi_{r+s}\psi_{r-s} \cdot \mathrm{EN}(n,m,1,0)\\
& - 2\psi_n^2 \cdot \mathrm{EN}(m,r,s,1) \quad .
\end{aligned}$$

This gives $\mathrm{EN}(n,m,r,s)$ for all $n,m,r,s > 1$.

[6] the complete proof also needs the case when $n,m,r,s \in \frac{1}{2}\mathbb{N} \setminus \mathbb{N}$

Blueprint for $T_p E_{\overline{F}}$



Definition of ψ_n (3.7(a))

Definition of ω_n (3.7(a))

Elliptic nets (3.7(g))

Definition of ϕ_n (3.7(a))

$[n]((x, y)) = \ldots$ (3.7(d))

Projective coordinates

$\deg(\phi_n) = n^2$ (3.7(b))

$(\phi_n, \psi_n^2) = 1$ (3.7(c))

Assumption $\Delta \neq 0$

$\# E_{\overline{F}}[n] = n^2$ (3.7(e))

$E_{\overline{F}}[n] \simeq (\mathbb{Z}/n)^2$

Finite abelian groups

$T_p E_{\overline{F}} \simeq \mathbb{Z}_p^2$

Inverse limits

Future projects

Projects without algebraic geometry:

- algorithms that only use the group law
- finite fields: the Hasse–Weil bound, the Weil conjectures
- local fields: the reduction homomorphism, Tate's algorithm, the Neron–Ogg–Shafarevich criterion, the Hasse–Weil L-function
- number fields: Neron–Tate heights, the Mordell–Weil theorem, Tate–Shafarevich groups, the Birch and Swinnerton-Dyer conjecture
- complete fields: complex uniformisation, p-adic uniformisation

Projects with algebraic geometry:

- elliptic curves over global function fields
- the projective scheme associated to an elliptic curve
- integral models and finite flat group schemes
- divisors on curves and the Riemann–Roch theorem
- modular curves and Mazur's theorem