ELLIPTIC CURVE CRYPTOGRAPHY

Winter term 2009/10

Michael Nüsken

February 1, 2010

Contents

1. Introduction

We consider elliptic curves from three different sides:

- Mathematics view: defines and analyzes their structure and properties.
- Computer science view: asks for efficient implementation of operations and properties.
- Cryptography view: requires some things to be intractable.

It is an interplay between the three areas and we will have to consider them all again and again.

1.1. Cryptography. As our main interest is cryptography, let's start here with two examples.

EXAMPLE 1.1 (ElGamal type signatures).

- *Global setup*: We fix a group G, an element P of finite order ℓ, a hash function $\mathrm{hash}\colon \{0,1\}^* \to \mathbb{Z}_\ell$, and a structureless type cast $^*\colon G \to \mathbb{Z}_\ell$.
- *User setup*: Each user chooses a private key $\alpha \in \mathbb{Z}_\ell$ and computes a public key $A = \alpha P$ from it. Each signer is henceforth identified by its public key A.
- *Signature verification*: A pair $(B, \gamma) \in G \times \mathbb{Z}_\ell^\times$ is *an ElGamal type signature by the signer* A iff the signature verification equation

$$B^* A + \gamma B = \mathrm{hash}(m) P \qquad \text{in } G \tag{1.2}$$

  holds.
- *Signature generation*: The signer A can generate such a signature as follows:
  (i) Choose a temporary secret $\beta \in \mathbb{Z}_\ell$ at random.
  (ii) Compute $B := \gamma P$.
  (iii) Solve the $\mathbb{Z}_\ell$-linear signature generation equation

$$B^* \alpha + \gamma \beta = \mathrm{hash}(m) \qquad \text{in } \mathbb{Z}_\ell \tag{1.3}$$

  for $\gamma \in \mathbb{Z}_\ell$.
  (iv) Return (B, γ).

Well, so far this is merely a set of protocols (even algorithms). To make it valuable we need to learn more:

- Does it work as wanted? (Mathematics)
- Does it perform fast? (Computer science)
- Is it secure? And what does that actually mean? (Cryptography) ◊

EXAMPLE 1.4 (ElGamal encryption).

- *Global setup*: We fix a group G, an element P of finite order ℓ.
- *User setup*: Each user chooses a private key $\alpha \in \mathbb{Z}_\ell$ and computes a public key $A = \alpha P$ from it. Each recipient is henceforth identified by its public key A.
- *Encryption*: The sender wants to encrypt a plain-text message $M \in G$ for recipient A.
  (i) The sender B chooses a temporary secret $\tau \in \mathbb{Z}_\ell$.
  (ii) He computes $T := \tau P$ and $C := M + \tau A$.
  (iii) Return (T, C).
- *Decryption*: The recipient A wants to decrypt a message (T, C) encrypted for her.
  (i) She computes $M' := C - \alpha T$

Again, this is a set of algorithms and we need to learn more:

- Does it work as wanted? (Mathematics)
- Does it perform fast? (Computer science)
- Is it secure? And what does that actually mean? (Cryptography) ◊

As you learn in cryptography these two examples are the most prominent ways to sign and to encrypt messages. Though classically only encryption was important, signatures are inevitable to authenticate in a world where you possibly do not see your partner.

Both of these schemes need a group. And a major consequence of the security question is that at least it should be difficult to find the private key α of a user A. That problem is called the *discrete logarithm problem with respect to P in the group G*. And it launches us towards elliptic curves. Since P shall be of finite order, we can immediately ignore any infinite group.

The simplest finite group is a cyclic groups $\mathbb{Z}_\ell^+$. It consist of the integers 0 through $\ell - 1$ and addition is performed by adding the integers and taking the remainder modulo ℓ. Take $P = 1$. Now, given $\alpha \in \mathbb{Z}_\ell$ it is of course easy to compute $A = \alpha P = \alpha$. But also to find α from P is trivial. And this situation does not change essentially when you take a different P in $\mathbb{Z}_\ell$.

The next best examples are the unit groups $\mathbb{Z}_p^\times$ of a finite field $\mathbb{Z}_p$. By mathematical theory this is a cyclic group, so up to changing names we have nothing new, right? Well, no. The transition from α to $A = \alpha P$ is more complex now. [You might prefer to write P^α since the operation is called multiplication, but that doesn't matter.] Though computing A from α has still polynomial runtime, the best known algorithm for finding α from A has runtime $L^{1/3}_{\mathcal{O}(1)}(n)$, where n denotes the bit length of p. [We also need n bits to represent a group element.]

Side remark: We define the notation

$$L^e_c(n) = \mathcal{O}\left(2^{cn^e(\log_2 n)^{1-e}}\right). \tag{1.5}$$

[1]

When e varies from 0 to 1 we pass from polynomial to simply exponential:

- $L^0_{\mathcal{O}(1)}(n) = n^{\mathcal{O}(1)}$ means polynomially bounded, and

- $L_c^1(n) = \mathcal{O}(2^{cn})$ is simply exponentially bounded.

This will ease our life in interpreting run-times between polynomial and exponential.

The third type of examples are elliptic curves. Despite their name they are not only curves but also additively written groups. If we consider them over a finite base field each elliptic curve is a finite group. Still mathematically we only consider a cyclic group, namely all multiples of a fixed group element P, but the transform $\alpha \mapsto A = \alpha P$ is again something new. And the structure of these elliptic curves is so beautifully weird that, while still being able to compute A given α easily, nobody has yet found a way to solve the discrete logarithm problem in an arbitrary elliptic curve faster than what is possible in any group. And these fastest generic algorithms need runtime $L_{\frac{1}{2}}^1(n) = \mathcal{O}\left(2^{\frac{n}{2}}\right)$.

Group	Exponentiation runtime	Best known discrete logarithm runtime
$\mathbb{Z}_\ell^+$	$\mathcal{O}^\sim(n)$	$\mathcal{O}^\sim(n) \subset L_{1+o(1)}^0(n)$
$\mathbb{Z}_p^\times$	$\mathcal{O}^\sim\left(n^2\right)$	$L_{c+o(1)}^{\frac{1}{3}}(n)$, $c = \left(\frac{64}{9}\right)^{\frac{1}{3}}$
Elliptic curve over $\mathbb{F}_q$	$\mathcal{O}^\sim\left(n^2\right)$	$L_{\frac{1}{2}}^1(n)$

Figure 1.1: Major available groups. Here $\ell \in \mathbb{N}_{\geq 2}$, p is a prime, q a prime power. In each case n is the bit length of the group size or the number of bits required to store a group element. So $n = \Theta(\log_2 \ell)$, $n = \Theta(\log_2 p)$, $n = \Theta(\log_2 q)$.

If we use Figure 1.1 to derive the bit size needed to reach a security level of 112 bits, that is, to have at least runtime 2^{112}, we learn (after running tests for determining missing $\mathcal{O}$-constants) that we need roughly

- $n = 2^{112}$ for $\mathbb{Z}_\ell^+$ (which makes all operations intractable),
- $n = 2048$ for $\mathbb{Z}_p^\times$, and
- $n = 224$ for an elliptic curve over $\mathbb{F}_q$.

Thus — as long as nobody invents better discrete logarithm algorithms — elliptic curves provide the shortest signatures at a given security level.

1.2. Books. There are many book on elliptic curves, and quite a bunch of really good ones.

- Menezes (1993)

  This book is a reference of relevant definitions and results. It provides only few proofs.

- Hankerson, Menezes & Vanstone (2004)

  An introductory book covering the most important aspects: Arithmetic, cryptographic protocols, and implementation.

- Washington (2003)

  This is *the* introduction to elliptic curves. The presentation only touches briefly the cryptographic situation but covers all mathematics.

- Blake, Seroussi & Smart (1999)

  This book provides a steep introduction to elliptic curves and all important aspects for cryptography.

- Blake, Seroussi & Smart (2005)

  This extends Blake *et al.* (1999) in many directions and covers important recent results.

- Silverman (1986)

  This is *the* bible. Any detail that you could not find elsewhere, here there's the way to it. However, this is the deepest and most mathematical of all books on this list and sometimes requires to look up other sources.

Much of these notes will follow Washington (2003).

1.3. Planned schedule.

- Elliptic curves over $\mathbb{F}_q$.
  - Definition and smoothness.
  * Projective space and the point at infinity.
  - Weierstrass form.
  - Transformation and other curve representation forms.
  - Singular cubics: structure.
  - Group law. Associativity.
  - Isomorphisms.
  - j-invariant.
  - Endomorphisms including scalar multiplication and the Frobenius.
  - Trace.
  - Size restriction (Hasse, Waterhouse).
  - Group structure. Supersingularity.
  - Torsion points.
  - Division polynomials.
  - Divisors.
- Elliptic curves over $\mathbb{Z}_N$ with, say, $N = p \cdot q$.

? Discrete logarithm problem.

  - Baby-step giant-step.
  - Pollard-ϱ.
  - Pohlig & Hellman.
  - Index calculus.
  - Generic groups and a lower bound.

  - Index calculus for elliptic curves?

- Elliptic curve discrete logarithm problem and related.
  - Singular cubics have easy DLP.
  - Reducing ECDLP to a discrete logarithm in a finite field.
  - Weil pairing. Millers algorithm.
  - (GHS and Weil descent.)
  - Elliptic curve Diffie-Hellman problem.
  - Decisional elliptic curve Diffie-Hellman problem.
- Counting points.
  - Subfield curves and their size (Hasse-Weil).
  - Koblitz curves: subfield curves $\mathbb{F}_2$ over $\mathbb{F}_{2^m}$.
  - Schoof's algorithm.
  - (SEA.)
- Selecting curves.
  - ... at random. Verifiably?
  - Generate parameters. Validate!
  - Generate with prescribed size (complex multiplication method).

? Implementation.
  - (Field arithmetic.)
  - Projective coordinates. Point representation. Other coordinates.
  - Point multiplication.
  - Montgomery (x-only).
  - Curves with endomorphisms.
  - Koblitz curves and use of the Frobenius in speeding up point scaling.

+ Fast addition and fast pairings.
  - Edwards curves.
  - Toric cubic classification, addition complexity.
  - Pairing lattices.
  - Construction of elliptic curves with small embedding degree.
  - Tricks to speed up pairings.

2. Elliptic curves over $\mathbb{F}_q$

We usually work over the field $\mathbb{F}_q$ with q elements. [This exists exactly if q is a prime power.] Sometimes, we use the field $\mathbb{R}$ of real numbers, the field $\mathbb{Q}$ of rational numbers, or the field $\mathbb{C}$ of complex numbers.

2.1. Geometry.

DEFINITION 2.1 (Tentative version). *An cubic curve over a field k (for example $k = \mathbb{F}_q$) is the zero set of a cubic polynomial f in two variables including 'points at infinity'.*

An elliptic curve over a field k is a smooth cubic curve with...

Consider a few examples:

1. $y - x^3$ over $\mathbb{F}_{27}$.
2. $y^2 - x^3$ over $\mathbb{R}$. This is Neil's parabola.
3. $y^2x + y^3 + x^3 - 3xy + x - 10$ over $\mathbb{F}_{13}$.
4. $x^3 + y^3 + 1$ over $\mathbb{Q}$. The fact that this curve has only two points is the case $n = 3$ of Fermat's Last Theorem, which states that $a^n + b^n = c^n$ has only trivial solutions (with $abc = 0$).
5. $y^2 = x^3 - x$ over $\mathbb{R}$

In each case the affine part of the curve is given by $\mathcal{C}_0 = \{(x, y) \in k^2 \mid f(x, y) = 0\}$. Points at infinity are still missing...

A general cubic polynomial $f(x, y) = \sum_{i+j\leq 3} \alpha_{ij} x^i y^j$ in two variables has ten coefficients. If you look for a curve passing through 9 given points this results in 9 linear equations for the 10 coefficients α_{ij}. There is thus a nontrivial solution and, in the generic case, only one up to scalar. For short: given 9 points in 'generic position' there is exactly one cubic curve that passes through these points.

To understand what happens at infinity we look at the *affine plane* k^2 in a different way. Embed k^2 in k^3 with third coordinate equal to 1:

$$\begin{array}{ccc} k^2 & \longrightarrow & k^3, \\ (x, y) & \longmapsto & (x, y, 1) \end{array}.$$

Next, a point (x, y) in the affine plane k^2 corresponds to the point $(x, y, 1)$ which in turn defines and is given by a line through the origin of k^3 and this point $(x, y, 1)$. Observe that some lines do not correspond to a point, actually exactly those that are parallel to the plane $z = 1$. These represent the points at infinity! See Figure 2.1.

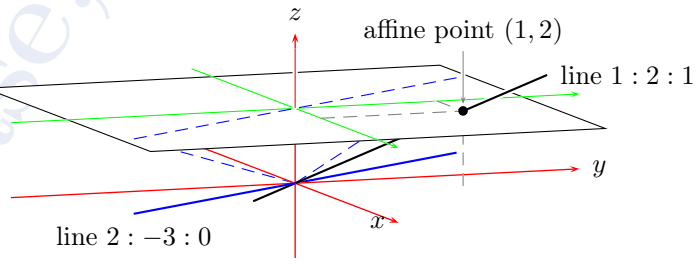

Figure 2.1: How to extend the affine plane with 'points at infinity'. The line $1 : 2 : 1$ defines the affine point $(1, 2)$. It is the projective point $[1 : 2 : 1]$. The line $3 : 2 : 0$ defines no affine point. We think of it as a direction in the affine plane. It is the projective point $[3 : 2 : 0]$.

Yet, now we have to rethink our cubic equation. How can we translate it such that a projective point $X : Y : Z$ solves the equation $F(X, Y, Z) = 0$ and the solutions still describe the same affine object? Note that we need that also $F(\alpha X, \alpha Y, \alpha Z) = 0$. This is granted if F is homogeneous, that is, all monomials in the polynomial F occur with same degree, say X^2Y is of degree 3 can be combined with YZ^2 or XYZ but not with XZ^3 or Z^2. Fortunately, there is a simple way to homogenize a polynomial: let $F(X, Y, Z) = f(X/Z, Y/Z) \cdot Z^d$ where $d = \operatorname{degree} f$. For example, we obtain:

- $F = YZ^2 - X^3$ for $f = y - x^3$.
- $F = Y^2Z - X^3$ for $f = y^2 - x^3$.
- $F = Y^2X + Y^3 + X^3 - 3XYZ + XZ^2 - 10Z^3$ for $f = y^2x + y^3 + x^3 - 3xy + x - 10$.

Now every solution of $F = 0$ defines an entire line of solutions, which we consider as a projective point. Now the curve is defined by

$$\mathcal{C} = \left\{ X : Y : Z \in \mathbb{P}^2(k) \,\middle|\, F(X, Y, Z) = 0 \right\}.$$

As the picture showed the difference is that we have a few more 'points' than before. It turns out that this is important...

Let's try to find the points at infinity in a few of our examples.

- Consider $f = y - x^3$. The homogeneous version is given by $F = YZ^2 - X^3$. And the points at infinity are those lines with $Z = 0$: $F(X, Y, 0) = -X^3$, thus we find $X = 0$ and since $(0, 0, 0)$ does not define a line we must choose $Y \neq 0$. However, it does not matter which value we take as we anyways take the line through that point. Thus the only point at infinity of the curve $f = 0$ is $0 : 1 : 0$. This is in direction of the y-axis.
- We have $F = Y^2Z - X^3$ for $f = y^2 - x^3$. Fixing $Z = 0$ implies $X^3 = 0$ and so again $0 : 1 : 0$ is the only point at infinity on this curve.
- $f = y^2x + y^3 + x^3 - 3xy + x - 10$. With $Z = 0$ we find $F(X, Y, 0) = Y^2X + Y^3 + X^3$. As $Y = 0$ forces all coordinates 0 we have $Y \neq 0$ and choose $Y = 1$. Thus we have to solve $F(X, 1, 0) == 1 + X + X^3$. Over $\mathbb{F}_{13}$ this has the only solution $X = -6$: $-6 : 1 : 0$ is the only (rational) point at infinity. You may expect two further solutions. They do exist but they are only in $\mathbb{F}_{13^2}$. [That's similar to $x^2 + 1$ over the reals. It has no solution in the real numbers but two in the complex numbers.]

You can easily change coordinates in many ways. Any invertible matrix $A \in k^{3\times 3}$ defines a projective coordinate change:

$$\begin{aligned} \mathbb{P}^2k &\longrightarrow \mathbb{P}^2k, \\ X : Y : Z &\longmapsto U : V : W, \\ \text{where } \begin{bmatrix} U \\ V \\ W \end{bmatrix} &= A \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \end{aligned}$$

Scaling X, Y, Z will scale U, V, W and thus the projective point $X : Y : Z$ will map to the projective point $U : V : W$ [this is fine also if you consider a projective point as a line in k^3 or a set of points in k^3]. Such a map is determined by the image of three points, for example,

the x-direction $1:0:0$, the y-direction $0:1:0$, and the affine origin $0:0:1$. Thus if you want the line at infinity, consisting of all points $X:Y:0$, to map to the line through $1:0:1$ and $0:1:1$, then simply use

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

Thus the point $X:Y:Z$ is mapped to the point $X:Y:X+Y+Z$. Or, in affine coordinates, the point (x,y) is mapped to the point $(\frac{x}{x+y+1}, \frac{y}{x+y+1})$. In this affine map we find quotients of linear polynomials; this is typical. Further, this map is of course not defined on the line $x+y=-1$. The reason is that the projective map maps this line to the line at infinity.

The next question takes up the word 'smooth' from the definition. To explain this we have first to define tangents to a plane curve. In school you consider the graph of a function g and a point (x_0,y_0) on it, that is, $y_0 = g(x_0)$. The tangent in (x_0,y_0) is the graph of the function $t_{x_0}\colon k \to k,\ x \mapsto y_0 + g'(x_0)(x-x_0)$ [recall the Taylor series]. If we describe the graph of g as a plane curve then it is the set of solutions of $f(x,y) = y - g(x)$. Rewriting the tangent as a set we obtain

$$T_{x_0} = \left\{ (x,y) \in k^2 \;\middle|\; \underbrace{\begin{bmatrix} f_x(x_0,y_0) & f_y(x_0,y_0) \end{bmatrix}}_{=:(\operatorname{grad} f)^{\mathrm{T}}(x_0,y_0)} \cdot \begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix} = 0 \right\}.$$

The gradient $\operatorname{grad} f(x_0,y_0)$ is also called normal vector to the curve $f=0$ for obvious reasons. We can now notice that the set T_{x_0} is a line iff the normal vector is nonzero. A point of the curve is called *non-smooth* or *singular* iff this normal vector vanishes, all other points are called *smooth* or *regular*. The curve is *smooth* iff all its points are regular. Equivalently, the affine curve is smooth iff the set $\{(x,y) \in k^2 \mid f(x,y)=0, f_x(x,y)=0, f_y(x,y)=0\}$ (of singular points of the curve $f=0$) is empty. Let's check our examples:

- $f = y - x^3$ over $\mathbb{F}_{27}$. We obtain $f_x = -3x^2$ and $f_y = 1$. Since f_y is always nonzero, all points of $\{(x,y) \in K^2 \mid f(x,y)=0\}$ are smooth.

  However, this curve has a point at infinity which is somehow out of view. To bring it back in view we can use a projective coordinate change which moves the line at infinity in view. Use the coordinate change mapping $1:0:0$ to $1:0:0$, $0:1:0$ to $0:0:1$, and $0:0:1$ to $0:1:0$. [The affine coordinate change is $(x,y) \mapsto (x/y, 1/y)$.] We obtain the new affine equation $h = y^2 - x^3$ which is Neil's parabola. The only point of $f=0$ at infinity was $0:1:0$ which is now the affine point $0:0:1$. That's our next example:

- $f = y^2 - x^3$ over $\mathbb{R}$. Here, $f_x = -3x^2$, $f_y = 2y$ and so the only critical point is $(x,y) = (0,0)$ and this is indeed on the curve and thus singular. This curve is not smooth. Since the first example is the 'same' projective curve we also observe that the point at infinity is regular, and so the only singular point of Neil's parabola is $(0,0)$ and the only singular point of the first example is its point $0:1:0$ at infinity.

- $f = y^2x + y^3 + x^3 - 3xy + x - 10$ over $\mathbb{F}_{13}$. We find $f_x = y^2 + 3x^2 - 3y + 1$, $f_y = 2yx + 3y^2 - 3x$. Searching for singular points is cumbersome here. Trying all $(x,y) \in \mathbb{F}_{13}$ shows that all (affine) points are regular. A little further work shows that also the points at infinity are regular.

Moreover, one can check that the polynomials f, f_x, f_y cannot have a common solution over any extension field of $\mathbb{F}_{13}$, for example by exhibiting a relation $af + bf_x + cf_y = 1$ with some polynomials $a, b, c \in \mathbb{F}_{13}[x, y]$.

- $f = x^3 + y^3 + 1$ over $\mathbb{Q}$. We find $f_x = 3x^2$ and $f_y = 3y^2$. The only possible singular point is $(x, y) = (0, 0)$, but it's not on the curve. So the affine part of the curve is smooth.

- $f = y^2 - x^3 + x$ over $\mathbb{R}$. We compute $f_x = -3x^2 + 1$ and $f_y = 2y$. Thus any singular point must have $y = 0$, and $f(x, 0) = -x(x - 1)(x + 1) = 0$. But neither $x = 0$, $x = 1$, nor $x = -1$ make f_x vanish. Thus also for this curve there is no affine singular point.

It would be helpful to be able to check smoothness also at infinity. We have seen one way to do so: change coordinates and check again. However, that's somehow cumbersome. Instead we'd like to translate our condition to the projective language.

Claim 2.2. *Given a polynomial $f \in k[x_1, \ldots, x_n]$, its homogenization*

$$F(\underline{X}) = X_{n+1}^{\deg f} f\left(\frac{X_1}{X_{n+1}}, \ldots, \frac{X_n}{X_{n+1}}\right) \in k[X_1, \ldots, X_{n+1}],$$

and a point $\underline{c}$ on $\{f = 0\}$, that is with $f(\underline{c}) = 0$. Let i be such that $c_i \neq 0$. Then the following are equivalent:

(i) The point $\underline{c}$ is a singular point.

(ii) All derivatives of f vanish at $\underline{c}$.

(iii) All derivatives of F vanish at $(c_1, \ldots, c_n, 1)$.

(iv) All derivatives of F vanish at $(\beta c_1, \ldots, \beta c_n, \beta)$ for some $\beta \in k^\times$.

(v) All derivatives of $g(\underline{x}) := F(x_1, \ldots, x_{i-1}, 1, x_i, \ldots, x_n)$ vanish at the point $(\frac{c_1}{c_i}, \ldots, \frac{c_{i-1}}{c_i}, \frac{c_{i+1}}{c_i}, \ldots, \frac{c_n}{c_i}, \frac{1}{c_i})$.

(vi) Let $A \in k^{n \times n}$ invertible. Define $G(\underline{U}) = F(A^{-1}\underline{U})$, and $\underline{B} = A(c_1, \ldots, c_n, 1)$. Clearly, $G(\underline{B}) = 0$. Then all derivatives of G vanish at $\underline{B}$.

The proof is a simple exercise. The claim shows that the definition is invariant under coordinate changes and it does not matter which view we take whether it's projective or affine.

For us: to check whether the (projective closure of the) curve given by a polynomial f is *smooth* we check whether the equations $F = 0$, $F_X = 0$, $F_Y = 0$, $F_Z = 0$ has no nontrivial solution [in other words: the only solution is $(0, 0, 0)$].

One final observation: the curve given by $f = x(2y - x^2)$ over the field $\mathbb{Q}$ has only regular points. Still we do not want to call it an elliptic curve, since it is the union of two curves. Curves that are not the proper union of two curves are called irreducible.

So we arrive at the following, now complete definition:

Definition 2.3. *An elliptic curve over a field k is a smooth, cubic, irreducible, projective curve with a rational (flex) point.*

Remark: a reducible curve is the union of two curves. For us that means $f = gh$ for some non constant polynomials g, h. Then $F = GH$, and then $F_X = G_X H + GH_X$. Thus, any point of the intersection is automatically singular. Since our field usually is not algebraically closed, we have no guarantee that the intersection is non-empty.

When we consider the tangents at points of a real curve [—to do—] a few points show a special behavior: the tangent passes from one side to the other, the sign of the curvature changes. Usually, a tangent intersects the curve twice [like $y = x^2$ at $(x, y) = (0, 0)$], here it intersects the curve three times [like $y = x^3$ at $(x, y) = (0, 0)$]. Intersection multiplicities can be considered over any field.

Maybe this is the point where we should remark that each line intersect the curve exactly three times, well, provided the field is algebraically closed and intersections are counted with multiplicity. Over an arbitrary field and ignoring multiplicities every line intersects at most three times and 'most' lines do intersect three times. Well, this is the historic definition of *degree of a curve*!

2.2. Weierstraß form. Let's reconsider the tangent at a flex. By definition it already does intersect the curve three times, so it cannot intersect the curve anywhere else. This makes it best suited for putting the curve in a simpler form by moving it to infinity and its tangent to the line at infinity. The result of this (rational) coordinate transformation is the generalized Weierstraß form [Exercise!]. We say that a cubic curve is in *generalized Weierstraß form* if it is given by an equation

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

To remember the numbering notice that each term has weighted degree six, if we consider x of degree 2, y of degree 3, and a_i of degree i. It can be proved that every cubic curve with a rational point can be transformed into this form by a birational isomorphism. The basic problem is, given a rational point, to exhibit a rational flex point. That this can always be done however either relies on deep mathematical results (Riemann-Roch) or on cumbersome calculations (Nagell algorithm, see Silverman (1986) [or Connell (1999)]).

Yet, it's rather simple to eliminate further coefficients from the generalized Weierstraß form. The coordinate change $x = u - \alpha$, $y = v$ will lead to a different presentation of the right hand side which when we choose α carefully will be simpler than before. Clearly, the left hand side will have the same form and the right hand side will be $u^3 + (-3\alpha + a_2)u^2 + \tilde{a}_4u + \tilde{a}_6$. Assume $3 \neq 0$ in our base field k, that is, $\operatorname{char} k \neq 3$. Then we can take $\alpha = \frac{a_2}{3}$ and so obtain a new equation with $\tilde{a}_2 = 0$. Actually, this is barely a special case of how to get rid of the second coefficient of a polynomial provided it's degree is not zero in the base field. So assuming $\operatorname{char} k \neq 2$ we can do that with respect to the variable y on the left hand side $y^2 + (a_1x + a_3)y$ of the Weierstraß equation and by setting $y = v - \frac{1}{2}(a_1x + a_3)$ the left hand side becomes $v^2 + \frac{1}{2^2}(a_1x + a_3)^2$ and we obtain the equation

$$\begin{aligned} v^2 &= x^3 + a_2x^2 + a_4x + a_6 - \frac{a_1^2x^2 + 2a_3a_1x + a_3^2}{4} \\ &= x^3 + \underbrace{\left(a_2 - \frac{1}{4}a_1^2\right)}_{=:\tilde{a}_2} x^2 + \underbrace{\left(a_4 - \frac{1}{2}a_3a_1\right)}_{=:\tilde{a}_4} x + \underbrace{\left(a_6 - \frac{a_3^2}{4}\right)}_{=:\tilde{a}_6}. \end{aligned}$$

Substituting also x according to the previous trick by $x = u - \frac{1}{3}\tilde{a}_2 = u - \frac{1}{3}\left(a_2 - \frac{1}{4}a_1^2\right)$ yields the equation

$$\begin{aligned} v^2 &= (u - \frac{1}{3}\tilde{a}_2)^3 + \left(a_2 - \frac{1}{4}a_1^2\right)(u - \frac{1}{3}\tilde{a}_2)^2 + \left(a_4 - \frac{1}{2}a_3a_1\right)(u - \frac{1}{3}\tilde{a}_2) + \left(a_6 - \frac{a_3^2}{4}\right) \\ &= u^3 + au + b, \end{aligned}$$

which is now in *Weierstraß form*. As we have seen whenever $\operatorname{char} k \neq 2, 3$ by a linear coordinate change we can bring every curve in general Weierstraß form to this Weierstraß form. Clearly, this equation determines a cubic curve also over a field of characteristic 2 or 3, however there are curves that cannot be transformed into them. [Side remark: in characteristic two, any elliptic curve can either be transformed into $y^2 + y = x^3 + ax + b$, this is the supersingular case, or into $y^2 + xy = x^3 + ax + b$. In characteristic three any elliptic curve can be transformed into $y^2 = x^3 + ax^2 + b$.]

This simplifies our life considerably, since we can now perform all operations in this particularly simple presentation involving only two parameters rather than having to deal with five parameters... So for the most of the rest of the course we assume that our curve is in Weierstraß form

$$y^2 = x^3 + ax + b.$$

Now let's ask our questions:

- What happens at infinity?
- When is the curve smooth?

Well, to see the points at infinity we homogenize and get $Y^2Z = X^3 + aX^2Z + bZ^3$. The line at infinity are the points with $Z = 0$, so we are interested in solutions of $0 = X^3$. Well, only $X = 0$ fulfills this. Thus Y must be nonzero and we obtain as the only solution the projective point $\mathcal{O} := 0 : 1 : 0$.

To check whether $\mathcal{O}$ is a regular point we restrict to $Y = 1$ and obtain the equation $0 = -z + x^3ax^3z + bz^3 =: g(x, z)$. That $\mathcal{O}$ lies on the curve is reflected by $g(0, 0) = 0$. The normal vector here is given by the gradient, since $g_x = 3x^3$ and $g_z = -1 + ax^3 + 3z^2$, we obtain $(0, -1)$ as normal vector. It is never zero and thus $\mathcal{O}$ is a regular point. Moreover, we can determine the tangent at $\mathcal{O}$ by $0 = 0(x - 0) - 1(z - 0) = -z$. This describes the line at infinity [with respect to the original affine coordinates]. And there is no further point of the curve on this line exhibiting that $\mathcal{O}$ is a (rational) flex point. [Exercise: This is also true for a curve in generalized Weierstraß form.]

To decide whether a curve is smooth thus depends only on the affine part and we have to look for points (x, y) on the curve with vanishing normal vector. So let $f = -y^2 + x^3 + ax + b$, and compute $f_x = 3x^2 + a$, $f_y = -2y$. Any point with zero normal vector must have $y = 0$, thus we look for simultaneous solutions of $r(x) = x^3 + ax + b$ and its derivative $r'(x) = 3x^2 + a$. A common zero of a polynomial r and its derivative however is equivalent to a double zero: Write $r(x) = (x - x_0)(x - x_1)(x - x_2)$. Then $r'(x_0) = (x_0 - x_1)(x_0 - x_2)$ vanishes iff (=if and only if) $x_0 = x_1$ or $x_0 = x_2$. We thus see that the curve is smooth iff the right hand side of the Weierstraß equation has no double zero.

Before we formulate this as a theorem, we want to translate the condition on r into a condition on a and b.

Let's see, given the roots we find $x_0 + x_1 + x_2 = 0$, $x_0x_1 + x_1x_2 + x_2x_0 = a$, $x_0x_1x_2 = b$. Hm...

Let's try to compute the greatest common divisor of r and r'. To avoid denominators we multiply r by 3, so $3r = (x)r' + r_2$ with remainder $r_2 = 2ax + 3b$. Next, $4a^2r' = (6ax - 9b)r_2 + r_3$ with remainder $r_3 = 4a^3 + 27b^2$. That's interesting! The Euclidean algorithm computes the greatest common divisor of r and r'. It is non-trivial, that is, of degree at least one, iff r and r' have a common root. Which is the case, as our computation shows, iff $4a^3 + 27b^2 = 0$. Now we can conclude

THEOREM 2.4. *A cubic curve over a field k of characteristic neither 2 nor 3 given by a polynomial Weierstraß equation*

$$y^2 = x^3 + ax + b$$

has exactly one point at infinity, namely $\mathcal{O} = 0 : 1 : 0$ the y-direction. The curve is smooth at this point. Let $r = x^3 + ax + b$. Moreover, the following conditions are equivalent

(i) The curve is smooth, that is, all its points are regular.

(ii) The right hand side r has no multiple root.

(iii) The polynomials r and r' have no common root.

(iv) The discriminant *$4a^3 + 27b^2$ is nonzero.* □

DEFINITION 2.5. *An elliptic curve in Weierstraß form over a field k of characteristic neither 2 nor 3is given by an equation*

$$y^2 = x^3 + ax + b$$

with nonzero discriminant $\Delta = 4a^3 + 27b^2$.

The preceding theorem tells us that this defines a smooth projective cubic curve over any extension of the field k. Thus it is also irreducible. And the point at infinity is always a rational flex.

2.3. The operation. We now reconsider the observation that every line intersects a smooth cubic curve three times. This paves the way for a binary operation, as there we have to relate any two points with a third one. Well, let's try to do it: Fix an elliptic curve and take any two points P and Q.

The line through P and Q must have a third point R on it with coordinates over the *same* field. The line through P and Q consists of all points of the form $L_\lambda := P + \lambda(Q - P)$. Plugging that into the curve equation results in a polynomial $h(\lambda) = f(L_\lambda)$ in λ of degree, well, three. [Since the curve equation is of degree three, it is clear that this polynomial can be of degree three at most. However, in projective coordinates the polynomial is homogeneous of degree 3. Now having smaller degree than 3 means that this polynomial is identically zero. This in turn means that the entire line is part of the curve. But that cannot be because we require that the curve is irreducible.] Actually, there may be a few exceptions to this but then a point at infinity will complete the picture.

Since P and Q are on the curve, $\lambda = 0$ and $\lambda = 1$ annihilate h, respectively. Thus division of h by $\lambda(\lambda - 1)$ yields $h = h_3\lambda(\lambda - 1)(\lambda - \lambda_2)$ for some $h_3, \lambda_2 \in k$. Thus $R := L_{\lambda_2}$ is the third point on the line through P and Q. It is tempting to consider the operation given by

$$P \boxplus Q := R.$$

However, this does not give a nice operation. Of course, it is proper (well-defined) and commutative. But a neutral element is a problem: the line through P and $\mathcal{N}$ must have P as third point, ie. this line would be a tangent at P. Thus $\mathcal{N}$ would have to be on every tangent. Also associativity is a problem. [Exercise: pick a concrete curve over $\mathbb{R}$ with three points P, Q, S and compare $(P \boxplus Q) \boxplus S$ to $P \boxplus (Q \boxplus S)$. Actually, associativity collides with the 'super-symmetry' of our definition: Given $R = P \boxplus Q$ we also have $P = Q \boxplus R$ and $Q = R \boxplus P$. This is a bit too much...

As a second try, we let ourself be guided by trying to construct an operation where $(P+Q)+R = 0$. To turn that into a definition for $P+Q$ we need a negation producing $-R$. In Weierstraß form changing the sign of the y-coordinate maps curve points to curve points. Assume $R = (x_3, y_3)$ is on the curve, ie. $y_3^2 = x_3^3 + ax_3 + b$. Then also $-R := (x_3, -y_3)$ is on the curve. Fortunately, now $-(-R) = R$, which is also necessary for a negation. Using that *negation operation* define the *point addition* by

$$P + Q := -R.$$

Getting back to our original idea, we now find that

$$(P + Q) + R = \mathcal{O}$$

for the three points P, Q, and R on a given line. That was easy but somehow hides the geometry. So let's reconsider the negation. Adding P and $-P$ should also be possible and should result in the neutral element. By definition the line through P and $-P$ is the vertical line $x = x_P$. The points on this line are P and $-P$ — and $\mathcal{O}$! Thus $P + (-P) = -\mathcal{O}$. Only, we never said what $-\mathcal{O}$ should be: take $-\mathcal{O} := \mathcal{O}$. We conclude that $-P$ is simply the third point on the line through P and $\mathcal{O}$: $-P = P \boxplus \mathcal{O}$. Notice that $-(-P) = P$ as it should be. Combining we obtain the geometric description

$$P + Q = (P \boxplus Q) \boxplus \mathcal{O}.$$

This is it.

However there are still a few gaps. Another look at our definition reveals that our definition is not yet complete. We have not considered what line we take if $P = Q$. Also if one of the points or both equal $\mathcal{O}$ has to be considered separately. Thus we still have to tell what $P + P$, $P + \mathcal{O}$, $\mathcal{O} + P$, and $\mathcal{O} + \mathcal{O}$ shall be.

- First, if $P = Q$ then we take the tangent at P which also geometrically has P as a double intersection point. Thus this line has a unique further point $P \boxplus P$. [If you do not see it for a specific point P, maybe P is even a triple intersection point, ie. a flex, thus $P \boxplus P = P$.]
- Second, if one of the points is the point $\mathcal{O}$, we simply follow the geometrical description. Assume $P \neq \mathcal{O}$. Then the third point $P \boxplus \mathcal{O}$ is $-P$ whose negation is again P. Thus we should define $P + \mathcal{O} = P$. Since our description is symmetric this also implies $\mathcal{O} + P = P$.
- Third, if both points are $\mathcal{O}$ we have again to consider the tangent at $\mathcal{O}$ which is the line at infinity whose third point is again $\mathcal{O}$. The negation of $\mathcal{O}$ is $\mathcal{O}$ itself and thus $\mathcal{O} + \mathcal{O} = \mathcal{O}$.

This completes the definition. Actually, the geometric definition can be used for an elliptic curve in an arbitrary description provided you also specify a flex $\mathcal{O}$ as a neutral element.

However, we still can not *compute* the sum $P + Q$ of two points. Our task is this:

Task 2.6. *Given an elliptic curve*

$$E\colon y^2 = x^3 + ax + b$$

and two points $P = (x_1, y_1) \neq \mathcal{O}$ and $Q = (x_2, y_2) \neq \mathcal{O}$ describe the coordinates of $P + Q$ as functions of the coordinates of P and Q.

Note that we do not need formulae in case that one of P or Q is $\mathcal{O}$. Actually, also the case $Q = -P$ is easy: Then $P + Q = P + (-P) = \mathcal{O}$. We are left with three cases: $P \neq \pm Q$, $P = Q \neq -P$, and $P = Q = -P$. As the line through two different points is easier to describe we start with that case.

Case $P \neq \pm Q$: The line through P and Q consists of all points (x, y) with $y = mx + c$. The slope m is $\frac{y_2 - y_1}{x_2 - x_1}$. Since $P \neq \pm Q$ we have $x_1 \neq x_2$ and this is well-defined. Thus we obtain the equation $y = m \cdot (x - x_1) + y_1$ for the line through P and Q. Plugging this description into the curve equation we obtain

$$(m \cdot (x - x_1) + y_1)^2 = x^3 + ax + b$$

for the points that are on the line and on the curve. We can rewrite this in the form

$$0 = f\left(x, m \cdot (x - x_1) + y_1\right) = x^3 - m^2 x^2 + \dots$$

where we use $f(x, y) = -y^2 + x^3 + ax + b$. Solving this equation is surprisingly easy since we already know the two solutions $x = x_1$ and $x = x_2$. Using x_3 for the third solution we must have $f\left(x, m \cdot (x - x_1) + y_1\right) = (x - x_1)(x - x_2)(x - x_3) = x^3 - (x_1 + x_2 + x_3)x^2 + \dots$. By comparison of the x^2-coefficient [Convince yourself that the dots do not interfere!] we obtain $m^2 = x_1 + x_2 + x_3$ and thus we obtain for the third point $R = (x_3, y_3)$ on the P-Q-line the values $x_3 = m^2 - x_1 - x_2$, $y_3 = m(x_3 - x_1) + y_1$. Negating this solves the task in the present case, the sum $P + Q$ has the coordinates

$$\begin{aligned} x_{P+Q} &= m^2 - x_1 - x_2, \\ y_{P+Q} &= -m(x_3 - x_1) - y_1, \end{aligned} \qquad \text{where } m = \frac{y_2 - y_1}{x_2 - x_1}.$$

Case $P = Q$, $P \neq -P$: We now have to find a line through P that intersects the curve twice. (Alternatively, find the tangent at P but let's redo this piece of theory…) Again the line is given by $y = m(x - x_1) + y_1$ and plugging this into the curve equation yields again

$$0 = f\left(x, m \cdot (x - x_1) + y_1\right) = x^3 + ax + b - (m \cdot (x - x_1) + y_1)^2.$$

We have to determine m such that x_1 is a double root. Thus also the derivative of the univariate polynomial $f\left(x, m \cdot (x - x_1) + y_1\right)$ must vanish at $x = x_1$, ie. $0 = f_x(x_1, y_1) \cdot 1 + f_y(x_1, y_1) \cdot m = 3x_1^2 + a - 2 \cdot y_1 \cdot m$. So $m = \frac{3x_1^2 + a}{2y_1}$ provided $y_1 \neq 0$. Finalizing as before yields

$$\begin{aligned} x_{P+P} &= m^2 - 2x_1, \\ y_{P+P} &= -m(x_3 - x_1) - y_1, \end{aligned} \qquad \text{where } m = \frac{3x_1^2 + a}{2y_1}.$$

Case $P = Q$, $P = -P$: If $y_1 = 0$ we seem to be in trouble because $3x_1^2 + a \neq 0$ since the curve is smooth. This trouble is caused by the assumption that the wanted line is not vertical, so try the line given by $x = x_1$. And indeed it intersects the curve twice at P:

$f(x_1, y) = y^2$ vanishes twice at $y = 0$. The third point on that line is $\mathcal{O}$, so we obtain $P + P = \mathcal{O}$ for the points with y-coordinate $y_P = 0$.

Notice that it was essential that $\mathcal{O}$ is a flex. A point P is a flex, ie. the tangent at P intersect the curve three times at P, in other words: the tangent at P does not intersect the curve anywhere else. The consequence is that $P+P = -P$, or $(P+P)+P = \mathcal{O}$. We conclude: A point P is a flex iff $(P + P) + P = \mathcal{O}$. It will turn out that every elliptic curve has 1, 3, or 9 such points. As we have seen before we could move each flex and its tangent to infinity and the line at infinity, resulting in possibly different generalized Weierstraß equations for the same curve.

2.4. Group. We have not yet discussed the properties of the new operation though we have styled it to be a group operation. Let's check the properties.

(P) The operation is a well-defined map $E \times E \to E$. That's the case.

(A) Actually, we do have $(P + Q) + R = P + (Q + R)$. But we defer further treatment for the moment.

(N) The point $\mathcal{O}$ is the neutral element. Just review the definitions.

(I) The negation provides an additive inverse $-P$ for every point P of the curve E. Just verify that we obtained $P + (-P) = \mathcal{O}$ above.

(C) As the sum is determined by the line through P and Q which is also the line through Q and P we obtain $P + Q = Q + P$.

So the construction made above does all we want and so the curve with this addition is a commutative algebraic group.

2.5. Associativity. Well, associativity is tricky.

- Solution 1: Prove it computationally.

  Take the formulae that we have obtained for the addition. Compute the coordinates of $(P + Q) + R$ assuming that no special situation occurs. Do the same for $P + (Q + R)$. Compare.

  It turns out that you can do that with the help of a computer algebra system. But even here you need to be careful, otherwise the intermediate expressions will swell so much that they do not fit into memory.

- Solution 2: Prove it geometrically.

  This can be done, but note that $P + Q = (P \boxplus Q) \boxplus \mathcal{O}$ and so $(P + Q) + R = (((P \boxplus Q) \boxplus \mathcal{O}) \boxplus R) \boxplus \mathcal{O}$. This involves quite a lot of points.

- Solution 3: Construct the group operation differently.

  Algebraist define the so-called divisor class group for any algebraic variety. This is a commutative group by construction. For elliptic curves one can define a map from its point set to its divisor class group. This map can be proved to be an isomorphism using the theorem of Riemann-Roch. [That theorem makes a statement about how many functions with certain properties live on the curve. Even Silverman (1986) only cites it.]

2.6. Singular cubics. Now, we have spent so much effort on easy forms and smoothness and the group operation, we have lost the singular cubics completely out of sight. Curves in Weierstraß form $y^2 = x^3 + ax + b$ are singular iff the right hand side polynomial $x^3 + ax + b$ has a multiple root. If it's a triple root move it to zero, so the curve is $y^2 = x^3$. [This shift is defined over k.] If it's a double root then we can shift x to obtain the form $y^2 = x^2(x + a)$. [This shift is defined over k: recall that a double root can be obtained from the gcd of the polynomial and its derivative.]

As we have seen earlier the curve $y^2 = x^3$ is isomorphic to $v = u^3$ if we map $(u, v) \mapsto (u/v, 1/v) = (x, y)$ and back $(x, y) \mapsto (u, v) = (x/y, 1/y)$. Plugging this in yields $v^{-2} = u^3 v^{-3}$, or $v = u^3$ and back $y^{-1} = x^3 y^{-3}$, or $y^2 = x^3$. Now, it is obvious that we can parametrize all non-singular points just by u. Let E_{ns} be the non-singular part, $E_{\text{ns}} = \left\{(x, y) \in k^2 \setminus \{(0, 0)\} \;\middle|\; y^2 = x^3\right\} \dot\cup \{\mathcal{O}\}$. It turns out that E_{ns} is isomorphic to k as an algebraic variety. And the world is even more stunning:

THEOREM 2.7. *The smooth part E_{ns} of the elliptic curve $E\colon y^2 = x^3$ over the field k is isomorphic to the field k. Moreover, the group operation on E_{ns} defined as in the smooth case corresponds exactly to the addition in k. More precisely, the map*

$$\tau\colon \begin{array}{ccc} E_{\text{ns}} & \longrightarrow & k, \\ (x, y) & \longmapsto & x/y, \\ \mathcal{O} & \longmapsto & 0 \end{array}$$

is an isomorphism of (algebraic) groups.

PROOF. Check that

$$\psi\colon \begin{array}{ccc} k & \longrightarrow & E_{\text{ns}}, \\ t & \longmapsto & (t^{-2}, t^{-3}), \\ 0 & \longmapsto & \mathcal{O}, \end{array}$$

is the inverse of the given map and thus the map is an isomorphism of algebraic curves. [Actually, projectively $\tau(X : Y : Z) = X : Y$ and $\psi(\alpha : \beta) = \alpha\beta^2 : \beta^3 : \alpha^3$.]

So it remains to check that if $t_1 + t_2 = t_3$ then $\psi(t_1) + \psi(t_2) = \psi(t_3)$. Let $P_i = \psi(t_i) = (x_i, y_i)$. The obtained formulae also work here, so —in the generic case—

$$\begin{aligned} x_{P_1+P_2} &= m^2 - x_1 - x_2, \\ y_{P_1+P_2} &= -m(x_3 - x_1) - y_1, \end{aligned} \qquad \text{where } m = \frac{y_2 - y_1}{x_2 - x_1}. \tag{2.8}$$

Plugging in we get

$$m = \frac{t_2^{-3} - t_1^{-3}}{t_2^{-2} - t_1^{-2}} = \frac{t_1^3 - t_2^3}{(t_1^2 - t_2^2)t_1 t_2} = \frac{t_1^2 + t_1 t_2 + t_2^2}{t_1 t_2 t_3}$$

and

$$x_{P_1+P_2} = \frac{(t_1^2 + t_1 t_2 + t_2^2)^2}{t_1^2 t_2^2 t_3^2} - t_1^{-2} - t_2^{-2} = t_3^{-2} = x_3.$$

For the other coordinate we obtain

$$y_{P_1+P_2} = -m(t_3^{-2} - t_1^{-2}) - t_1^{-3} = t_3^{-3} = y_3.$$

In principle, we would have to consider the non-generic case where $P_1 = P_2$ or $P_1 = -P_2$, however these pairs (t_1, t_2) are defined by polynomial equations whereas $\psi(t_1) + \psi(t_2) = \psi(t_1 + t_2)$ already holds on an open set, and thus on its closure which is everything. □

In a similar fashion we obtain a characterization of the structure of the smooth part $\mathrm{E}_{\mathrm{ns}} = \left\{(x, y) \in k \,\middle|\, y^2 = x^2(x + a), x \neq 0\right\} \dot\cup \{\mathcal{O}\}$ of the curve $y^2 = x^2(x + a)$ with $0 \neq a \in k$. Notice that $(0, 0)$ is the only singular point and $y = \pm\sqrt{a}x$ are the tangents to the two 'parts' of the curve at $(0, 0)$.

THEOREM 2.9. *Assume $a \in k$, $a \neq 0$, let α be a root of a in some extension field of k, and $K = k(\alpha)$. The smooth part E_{ns} of the elliptic curve $E\colon y^2 = x^2(x + a)$ over the field k is isomorphic to*

- *the multiplicative group $P = k^\times$ if a is a square in k, or*
- *the multiplicate subgroup $P = \left\{u + \alpha v \in K \,\middle|\, u, v \in k, u^2 - av^2 = 1\right\}$ of K if a is not a square in k.*

by the maps

$$\tau\colon \begin{array}{rcl} E_{\mathrm{ns}} & \longrightarrow & P, \\ (x, y) & \longmapsto & \frac{y+\alpha x}{y-\alpha x}, \\ \mathcal{O} & \longmapsto & 1, \end{array} \qquad \psi\colon \begin{array}{rcl} P & \longrightarrow & E_{\mathrm{ns}}, \\ t & \longmapsto & \left(\frac{4\alpha^2 t}{(t-1)^2}, \frac{4\alpha^3 t(t+1)}{(t-1)^3}\right), \\ 1 & \longmapsto & \mathcal{O}. \end{array}$$ □

[2]

2.7. Other equations. Once we have observed that we can bring most elliptic curves in Weierstraß form we try to improve on it...[3] Assume that e_0, e_1, and e_2 are the zeros of $x^3 + ax + b$. By a linear map we can easily map $e_0 \mapsto 0$ and $e_1 \mapsto 1$. Replacing $v = \frac{x-e_0}{e_1-e_0}$ and $v = (e_1 - e_0)^{-\frac{3}{2}} y$ we obtain the Legendre form

$$v^2 = u(u - 1)(u - \lambda)$$

with $\lambda = \frac{e_2-e_0}{e_1-e_0}$. Of course, we can only do this over a large enough field. Thus every elliptic curve, say over an algebraically closed field, is isomorphic to a curve in Legendre form. Unfortunately, λ is not uniquely defined. Actually, we have picked e_0 out of three roots and e_1 out of the remaining two; there are six ways of doing that. We obtain up to six Legendre forms with a parameter within $\left\{\lambda, \frac{1}{\lambda}, \frac{\lambda-1}{\lambda}, \frac{\lambda}{\lambda-1}, \frac{1}{1-\lambda}, 1 - \lambda\right\}$. This set only collapses for if $\lambda = -1$, $\lambda = \frac{1}{2}$, $\lambda = 2$, or $\lambda^2 - \lambda + 1 = 0$.

The discriminant turns out to be $\Delta = -\lambda^2(\lambda - 1)^2$, which reflects again that λ can be neither 0 nor 1. [For later reference: $j = \frac{2^8(\lambda^2-\lambda+1)^3}{\lambda^2(\lambda-1)^2}$.]

2.8. Isomorphisms and the j-invariant. After the preceding discussion we soon come up with the following definition for isomorphisms: We only consider maps that are somehow given by polynomials. In the projective world thus any map of the form

$$\beta\colon \begin{array}{rcl} \mathbb{P}^2 k & \longrightarrow & \mathbb{P}^2 k, \\ X_0 : X_1 : X_2 & \longmapsto & F_0(X) : F_1(X) : F_2(X) \end{array}$$

where F_0, F_1, and F_2 are homogeneous polynomials of the same degree. Or in the affine picture we consider maps of the form

$$\alpha\colon \begin{array}{ccc} k^2 & \longrightarrow & k^2, \\ (x_0, x_1) & \longmapsto & (f_0(x), f_1(x)) \end{array}$$

where f_0 and f_1 are quolynomials. Clearly, α is only defined where the denominators of f_0 and f_1 do not vanish. Similarly, β is, at first, only defined outside $V(F_0, F_1, F_2) = \{X_0 : X_1 : X_2 \,|\, F_0(X) = 0, F_1(X) = 0, F_2(X) = 0\}$. Then a morphism from a curve E to another curve F is given by such a map restricted to $E \subset \mathbb{P}^2 k$ whose images are all in $F \subset \mathbb{P}^2 k$. Finally, an isomorphism is any such map which has an inverse. However, we know a lot more than just 'we deal with a curve'.

Consider an elliptic curve $E\colon y^2 = x^3 + ax + b$. Let $u = \mu^2 x$ and $v = \mu^3 y$ with $\mu \in \overline{k}^\times$. Then we land on the curve

$$F\colon v^2 = u^3 + cu + d$$

with $c = \mu^4 a$, $d = \mu^6 b$. Moreover, the map $E \to F$, $(x, y) \mapsto (u, v)$ is an isomorphism. So finding isomorphisms for curves in Weierstraß form is easy in this case. But what about the general question? Given two curves, are they isomorphic? The tricky part is to make this decision and eventually to prove that they are not isomorphic. It turns out that given two curves in Weierstraß form, the previous ones are all the isomorphisms that we need to know:

Any isomorphism for two curves in Weierstraß form is of the form

$$\begin{array}{ccc} E & \longrightarrow & F, \\ (x, y) & \longmapsto & (\mu^2 x, \mu^3 y). \end{array}$$

More general, Silverman (1986) Proposition 3.1(b) states that any isomorphism for two curves in generalized Weierstraß form are of the form $E \to F$, $(x, y) \mapsto (u^2 x + r, u^3 y + su^2 x + t)$ for some parameters $u, r, s, t \in k$, $u \neq 0$. (This is again a consequence of the Riemann-Roch Theorem.)

We define the j-invariant $j(E)$ of an elliptic curve $E\colon y^2 = x^3 + ax + b$ in Weierstraß form by

$$j(E) := 1728 \frac{4a^3}{4a^3 + 27b^2}.$$

Its denominator is the discriminant, so never zero for an elliptic curve. The prior change of variables leaves this value unchanged: $1728 \frac{4(\mu^4 a)^3}{4(\mu^4 a)^3 + 27(\mu^6 b)^2} = 1728 \frac{4a^3}{4a^3 + 27b^2}$. More important, also the converse is true: if the j-invariants of two curves in Weierstraß form is equal then the curves are isomorphic in this way.

THEOREM 2.10. *Let $E\colon y^2 = x^3 + ax + b$ and $F\colon v^2 = u^3 + cu + d$ be two elliptic curves with same j-invariant. Then there exists a $\mu \in \overline{k}$ such that*

$$c = \mu^4 a, \qquad\qquad d = \mu^6 b.$$

The transformation $u = \mu^2 x$, $v = \mu^3 y$ provides an isomorphism from E to F.

PROOF. This is now surprisingly simple: we have

$$1728 \frac{4a^3}{4a^3 + 27b^2} = j = 1728 \frac{4c^3}{4c^3 + 27d^2}$$

Assume $a \neq 0$. Then also $c \neq 0$. Choose μ such that $c = \mu^4 a$. Now the equality of j-invariants implies that $\frac{27b^2}{4a^3} = \frac{27d^2}{4c^3}$ and so $d^2 = \left(\mu^6 b\right)^2$. Therefore $d = \pm\mu^6 b$. If $d = \mu^6 b$ we are done. Otherwise replace μ with $i\mu$ where $i^2 = -1$. We still have $c = (i\mu)^4 a$ but now also $d = (i\mu)^6 b$.

If $a = 0$ then $j = 0$ and $c = 0$. Since $\Delta(E) \neq 0$ we have $b \neq 0$, and similarly $d \neq 0$. So choose μ such that $d = \mu^6 b$. □

Actually by the above citation a bit more holds: two curves in Weierstraß form over an algebraically closed field are isomorphic iff their j-invariants are equal.

There are two special types of curves:

1. $j = 0$. This reflects that $a = 0$: $y^2 = x^3 + b$.

   Starting with the Fermat curve $x^3 + y^3 + 1 = 0$ we can transform it into Weierstraß form by moving the point at infinity to $0 : 1 : 0$ and its tangent to the line at infinity. Additional scaling and shifting in total moves the flex point $(-1, 0)$ to $(12, 36)$ and $(0, -1)$ to $(12, -36)$ and makes all curve coefficients integers. We obtain:

   $$y^2 = x^3 - 432.$$

   Over $\mathbb{Q}$ this curve has only three rational points: $(12, 36)$, $(12, -36)$ and $\mathcal{O}$. As the transformation is defined over $\mathbb{Q}$ this implies that also $x^3 + y^3 + 1 = 0$ has only three rational points, namely $(-1, 0)$, $(0, -1)$ and $1 : -1 : 0$. This in turn proves that $a^3 + b^3 = c^3$ has no integer solution with $abc \neq 0$.

2. $j = 1728$. Here $y^2 = x^3 + ax$.

   Examples are $a = -4$ or $a = -25$. By the above the two curves are of course isomorphic. However, over $\mathbb{Q}$ they are definitely different: The elliptic curve $y^2 = x^3 - 25x$ has infinitely many $\mathbb{Q}$-rational points, for example, all integer multiples of $(-4, 6)$. But the curve $y^2 = x^3 - 4x$ has only the rational points $(2, 0)$, $(-2, 0)$, $(0, 0)$ and $\mathcal{O}$.

We say that two different elliptic curves over a field k are *twists* of each other iff they are isomorphic over the algebraic closure $\overline{k}$, ie. there j-invariants are equal.

Finally, note that provided $j \neq 0, 1728$ in k the j-invariant of

$$y^2 = x^3 + \frac{3j}{1728 - j}x + \frac{2j}{1728 - j}$$

is j. As $j(y^2 = x^3 + 1) = 0$ and $j(y^2 = x^3 + x) = 1728$ we can realize any j-invariant. In other words: Every value in k is the j-invariant of some elliptic curve, which is even defined over k. The situation for fields of characteristic 2 or 3 is similar, however, the precise formulae are different.

2.9. Endomorphism. An endomorphism is a morphism of some elliptic curve to itself. Other than an isomorphism it does not need to have an inverse.

What endomorphism are obvious?

- The identity $\mathrm{id}\colon E \to E,\ P \mapsto P$.
- The negation $-\colon E \to E,\ P \mapsto -P$.

- The zero map $[0]\colon E \to E,\ P \mapsto \mathcal{O}$.
- The point doubling map $[2]\colon E \to E,\ P \mapsto 2P$. This is the first example where we should spent a though on whether this map is algebraic (ie. given by quolynomials). However, given $P = (x, y)$ with $y \neq 0$ we already know that $2P = (m^2 - 2x, -m(m^2 - 3x) - y)$ with $m = \frac{3x^2+a}{2y}$. Thus

$$2P = \left(\frac{(3x^2 + a)^2}{4y^2} - 2x, -\frac{3x^2 + a}{2y} \left(\frac{(3x^2 + a)^2}{4y^2} - 3x \right) - y \right).$$

- Generalizing the previous we obtain the *scalar multiplication* map:

$$[n]\colon \begin{array}{ccc} E & \longrightarrow & E, \\ P & \longmapsto & nP \end{array}.$$

  Using the point addition formulae we find by induction that also nP is given by quolynomials in the coordinates x and y of P.

- Knowing a bit about field isomorphisms leads to further automorphisms if the curve is defined over a small field. Assume $\alpha\colon k \to k$ is a field automorphism. If now $P = (x, y)$ is on the curve $E\colon y^2 = x^3 + ax + b$ then we have $\alpha(y)^2 = \alpha(x)^3 + \alpha(a)\alpha(x) + \alpha(b)$. Thus if $\alpha(a) = a$ and $\alpha(b) = b$ then also $(\alpha(x), \alpha(y))$ is a point on E. As the set of points fixed by such a field automorphism is always a subfield this just means that a and b do not span the field. Say E is defined over $\mathbb{F}_q$. [This is the first time we really depend on finite characteristic.] Then we can consider the Frobenius automorphism φ of $\mathbb{F}_q$: $\varphi_q\colon \overline{\mathbb{F}_q} \to \overline{\mathbb{F}_q},\ x \mapsto x^q$:

$$\varphi_q\colon \begin{array}{ccc} E & \longrightarrow & E, \\ (x, y) & \longmapsto & (x^q, y^q) \end{array}.$$

  This will be the identity on the $\mathbb{F}_q$-rational points $E(\mathbb{F}_q)$ of E but for all others will do something. Clearly, this map is given by quolynomials. Stunningly, though it is bijective it is never an isomorphism. [The inverse map is not given by quolynomials! Actually, for every fixed point $(x, y) \in E(\mathbb{F}_{q^m})$ we have $\varphi^m(P) = P$ (Remember Lagrange.) and thus can compute that inverse. However, to cover the entire algebraic closure $\overline{k}$ we need m arbitrarily large...]

We have seen that — after a bit extra massaging —

$$2P = \left(\frac{(3x^2 + a)^2 - 8xy^2}{4y^2}, \frac{-(3x^2 + a)^3 + (3x^2 + a)3x4y^2 - 8y^4}{8y^3} \right)$$

for a point $P = (x, y)$ on the curve $E\colon y^2 = x^3 + ax + b$. Hm... if the point is on the curve then we could replace y^2 in this formula with $x^3 + ax + b$. This would eliminate y completely apart from a single y in the denominator in $(2P)_y$. Or, if we prefer in its numerator: just expand that fraction with y and replace y^2 in the denominator again. We get

$$2P = \left(\frac{x^4 - 2ax^2 - 8bx + a^2}{4\left(x^3 + ax + b\right)}, \frac{x^6 + 5ax^4 + 20bx^3 - 5a^2x^2 - 4abx - a^3 - 8b^2}{8\left(x^3 + ax + b\right)^2} y \right).$$

So for the doubling endomorphism we obtain a quite nice shape: $(x, y) \mapsto (r_1(x), r_2(x)y)$ with some quolynomials $r_1, r_2 \in k(x)$. It turns out that we can always get such a form: Suppose α is an endomorphism of E. Then $\alpha(x, y) = (R_1(x, y), R_2(x, y))$ for some quolynomials $R_1, R_2 \in k(x, y)$. Clearly, we can replace every occurence of y^2 with $x^3 + ax + b$ since $P = (x, y)$ is on the curve. Thus $R_1(x, y) = \frac{p_1(x)+p_2(x)y}{p_3(x)+p_4(x)y}$ with polynomials $p_i \in k[x]$. By multiplying numerator and denominator with $p_3 - yp_4$ and again eliminating y^2 we obtain $R_1(x, y) = r_1(x) + r_3(x)y$. Similarly, we rewrite R_2 as $R_2(x, y) = r_4(x) + r_2(x)y$. Finally, observe that since our endomorphism α is a group morphism we have $\alpha(-P) = -\alpha(P)$, that is, $R_1(x, -y) = R_1(x, y)$ and $R_2(x, -y) = -R_2(x, y)$. This implies that $r_3(x) = 0$ and $r_4(x) = 0$ and we thus obtain the normal form

$$\begin{aligned}\alpha(x, y) &= (r_1(x), r_2(x)y) \\ &= \left(\frac{p(x)}{q(x)}, \frac{s(x)}{t(x)}y\right)\end{aligned}$$

with $r_1, r_2 \in k(x)$, $p, q, s, t \in k[x]$, $\gcd(p, q) = 1$, $\gcd(s, t) = 1$.

In an arbitrary representation of an endomorphism it is not immediately clear what happens to points where the quolynomials R_1, R_2 cannot be evaluated because some denominator vanishes. After transforming α to the normal form however, we can do this: if $q(x) = 0$ then $\alpha(x, y) = \mathcal{O}$. If $q(x) \neq 0$ then also $r_2(x)$ is well-defined.

To further understand endomorphism we define the degree of α:

Definition 2.11 (Degree of an endomorphism). *The degree of a nontrivial endomorphism* $\alpha\colon E \to E,\ (x, y) \mapsto \left(\frac{p(x)}{q(x)}, \frac{s(x)}{t(x)}y\right)$ *given by polynomials* $p, q, s, t \in k[x]$ *with* $\gcd(p, q) = 1$, $\gcd(s, t) = 1$ *is defined by*

$$\deg \alpha := \max\{\deg p, \deg q\}.$$

For $\alpha = 0$ *let* $\deg \alpha = 0$.

As the definition of morphism of curves greatly simplified for curves in Weierstraß form, also this definition actually is a special case of a much more general definition of degree of algebraic maps. Feel guided by the degree of a polynomial in the form that the inverse image of almost any value under the polynomial map is equal to its degree. Again: degree = number of "intersections". Actually, it turns out that not all endomorphisms do follow that feeling:

Definition 2.12 (Separable). *An endomorphism* $\alpha\colon E \to E,\ (x, y) \mapsto (r_1(x), r_2(x)y)$ *is* separable *iff* $r_1' \neq 0$ *(as a quolynomial).*

Equivalently, α is separable iff $p' \neq 0$ or $q' \neq 0$.

Example 2.13. Consider $\alpha = [2]$ as an example on a curve in Weierstraß form over a field of characteristic different from 2 (and 3). By the above $p(x) = x^4 - 2ax^2 - 8bx + a^2$ and $q(x) = 4(x^3 + ax + b)$. Obviously, $\deg[2] = 4$ here. And unless $\operatorname{char} k = 2$ we find that $p'(x) = 4x^3 - 4ax - 8b$ is nonzero and also that $q'(x) = 12x^2 + 4a$ is nonzero. Thus [2] is separable over any field of characteristic different from 2. ◊

Example 2.14. Over a finite field k the Frobenius φ_q is also an endomorphism. We have $p(x) = x^q$, $q(x) = 1$. So $\deg \varphi_q = q$. Thus $p'(x) = qx^{q-1} = 0$ and $q'(x) = 0$, and so φ is not separable. ◊

The above feeling is now made precise:

Theorem 2.15. *Let $\alpha \neq 0$ be a separable endomorphism of an elliptic curve E. Then*

$$\#\ker \alpha = \deg \alpha$$

where $\ker \alpha = \left\{P \in E\left(\overline{k}\right) \;\middle|\; \alpha(P) = \mathcal{O}\right\}$ is the kernel of the endomorphism $\alpha\colon E\left(\overline{k}\right) \to E\left(\overline{k}\right)$.

If $\alpha \neq 0$ is not separable then

$$\#\ker \alpha < \deg \alpha.$$

Thus as with polynomials the number of zeros is always at most the degree, and usually equal. A similar theorem actually holds for polynomials if you recall that a polynomial is called separable iff it does not have multiple roots (in the algebraic closure) or, equivalently iff it is coprime to its derivative.

Consequently, we expect the theorem to fail when replacing the algebraic closure with a too small field.

Before entering the proof note that the statement of the theorem carries much further since α is a morphism. Namely, any nonempty fiber $\alpha^{-1}(Q_0)$ has the same size: If $Q_0 = \alpha(P_0)$ is a point in the image then

$$\#\alpha^{-1}(Q_0) = \#\ker \alpha.$$

Namely, $\alpha(P) = Q_0 = \alpha(P_0)$ iff $P - P_0 \in \ker \alpha$.

Proof. Write $\alpha(x) = (r_1(x), r_2(x)y)$ with $r_1(x) = \frac{p(x)}{q(x)}$ with coprime $p(x)$, $q(x)$ as above. By assumption $r_1' \neq 0$ thus $p'q - pq'$ is not the zero polynomial.

Instead of considering the kernel we consider an arbitrary fiber. This makes our life easier as we can stay in the affine picture. Actually, we choose $(u, v) = \alpha(P)$ generically. We make that precise: Let S be the set of $x \in \overline{k}$ such that $(p'q - pq')(x)q(x) = 0$. Note that S is finite and thus not all of $\overline{k}$. Now choose a point $(u, v) \in E\left(\overline{k}\right)$ such that

- $u \neq 0$, $v \neq 0$, $(u, v) \neq \mathcal{O}$.
- $\deg(p - aq) \overset{!}{=} \max\{\deg p, \deg q\} = \deg \alpha$.
- $u \notin r_1(S)$.
- $(u, v) = \alpha(P)$ for some $P \in E\left(\overline{k}\right)$.

Since we allow $\overline{k}$ the image of α is infinite, whereas the other requirements only forbid finitely many values. Thus there are enough values to choose from.

We claim that there are exactly $\deg \alpha$ points $(x_1, y_1) \in E\left(\overline{k}\right)$ such that $\alpha(x_1, y_1) = (u, v)$. For any such point we have

$$\frac{p(x_1)}{q(x_1)} = u, \qquad y_1 r_2(x) = v.$$

Since $(u, v) \neq =$ we have $q(x_1 \neq 0)$ and $r_2(x_1)$ is defined. Next, noting $v \neq 0$ the second equation leaves exactly one choice for y_1: $y_1 = v/r_2(x)$. Thus we only need to count values for x_1.

By assumption $\deg(p - uq) = \deg \alpha$ and so the first equation $p(x_1) - uq(x_1) = 0$ has exactly $\deg \alpha$ roots counting multiplicities. In case there are no multiple roots we are done. So consider

$$p(x_1) - uq(x_1) = 0, \qquad p'(x_1) - uq'(x_1) = 0.$$

Combining them we obtain $up(x_1)q'(x_1) = up'(x_1)q(x_1)$ Since $u \neq 0$ we would have $x_1 \in S$ and thus $u = r_1(x_1) \in r_1(S)$ contrary to our choice. Therefore, all roots are single roots and we are done.

If α is not separable then we can all of the above but $p' - uq'$ is the zero polynomial and thus all root of $p - uq$ are multiple roots and so the fiber size is smaller than the degree. □

THEOREM 2.16. *If $\alpha\colon E\left(\overline{k}\right) \to E\left(\overline{k}\right)$ is a endomorphism then it is either zero or surjective.*

PROOF. Consider $(u, v) \in E\left(\overline{k}\right)$. We are looking for $x_1 \in \overline{k}$ such that $p(x_1) - uq(x_1) = 0$.

First, note that p or q is non-constant and p is nonzero. Otherwise the image of α could have at most three points, namely $\mathcal{O}$ and $(\frac{p}{q}, \pm v)$ for some v, and thus their fibers could not all be finite contradicting the previous result. Thus there is at most one value u such that $p - uq$ is constant.

In case $p - uq$ is non-constant, take a root x_1 of $p - uq$. Necessarily, $q(x_1) \neq 0$. Otherwise also $p(x_1) = 0$ and x_1 would be a common root contradicting the assumption that $\gcd(p, q) = 1$. Thus $u = \frac{p(x_1)}{q(x_1)}$ now. Let y_1 be a solution of $y^2 = x_1^3 + ux_1 + b$. Then $\alpha(x_1, y_1) = (u, \hat{v})$ and $\hat{v}^2 = u^3 + au + b = v^2$. Thus $\hat{v} = \pm v$. If $\hat{v} = v$ we are done, if $\hat{v} = -v$ then $\alpha(x_1, -y_1) = (u, v)$.

In case $p - uq$ is constant, choose any point (u_1, v_1) with $u \neq u_1$ such that $(u, v) + (u_1, v_1) \neq (u, \pm v)$. By the previous there are points such that $\alpha(P_1) = (u_1, v_1)$ and $\alpha(P_2) = (u, v) + (u_1, v_1)$. Then $\alpha(P_2 - P_1) = (u, v)$ and we are done. □

2.9.1. Separability. Our next aim is a criterion to decide separability of most endomorphism. It turns out that given an endomorphism α in the above form, the expression

$$c_\alpha := \frac{r_1'}{r_2}$$

is always constant. Moreover, it behaves nicely with respect to addition and composition of endomorphisms.

First, note that translations behave particularly nice:

LEMMA 2.17. *Let $E\colon y^2 = x^3 + ax + b$ be an elliptic curve and $(u, v) \in E$ any (nonzero) point. Write*

$$(x, y) + (u, v) = (f(x, y), g(x, y))$$

with quolynomials $f, g \in k[x, y]$. Then

$$\frac{\frac{\mathrm{d}}{\mathrm{d}x} f(x, y)}{g(x, y)} = \frac{1}{y}.$$

PROOF. We have $f(x,y) = m(x,y)^2 - x - u$, $g(x,y) = -m(x,y)(f(x,y)-u) - v$ with $m(x,y) = \frac{y-v}{x-u}$ for 'most' $(x,y) \in E$. Now compute

$$y\frac{\mathrm{d}}{\mathrm{d}x}f(x,y) - g(x,y) = y\left(f_x(x,y) + f_y(x,y)y'\right) - g(x,y).$$

Using $2y'y = 3x^2 + a$, $y^2 = x^3 + ax + b$ and $v^2 = u^3 + au + b$ this simplifies to zero and we are done. □

The statement made by Lemma 2.17 is that the differential $\frac{\mathrm{d}x}{y}$ is translation invariant. Additionally, one can prove that up to scalar it is the only translation invariant differential on E.

LEMMA 2.18. *Let α_1, α_2, α_{1+2}, and $\alpha_{1\circ 2}$ be endomorphisms of an elliptic curve E with $\alpha_{1+2} = \alpha_1 + \alpha_2$ and $\alpha_{1\circ 2} = \alpha_1 \circ \alpha_2$. Assume that c_{α_1} and c_{α_2} are both constant. Then*

$$c_{\alpha_{1+2}} = c_{\alpha_1} + c_{\alpha_2},$$
$$c_{\alpha_{1\circ 2}} = c_{\alpha_1} \cdot c_{\alpha_2}.$$

In particular, $c_{\alpha_{1+2}}$ and $c_{\alpha_{1\circ 2}}$ are also constant.

The previous unproven remark carries even further: the assumption of Lemma 2.18 is actually true for every endomorphism. Since an endomorphism α is a group morphism also $\frac{\mathrm{d}x}{y} \circ \alpha$ is a translation invariant differential and thus by the uniqueness must be a constant multiple of $\frac{\mathrm{d}x}{y}$. This constant is c_α. [However, we do not prove that uniqueness.]

PROOF. Write $\alpha_j(x,y) = (r_{1j}(x), yr_{2j}(x))$. Let $(x_j(x,y), y_j(x,y)) = \alpha_j(x,y)$, so $\alpha_{1+2}(x,y) = (x_1,y_1) + (x_2,y_2)$. We want to compute $c_{\alpha_{1+2}} = \frac{x'_{1+2}(x)}{y_{1+2}(x)/y}$. Well, considering x_{1+2} as a function composed from first applying (α_1, α_2) and then addition we obtain

$$x'_{1+2} = \frac{\partial x_{1+2}}{\partial x_1}\frac{\partial x_1}{\partial x} + \frac{\partial x_{1+2}}{\partial x_2}\frac{\partial x_2}{\partial x}.$$

Now, by Lemma 2.17 we know that $\frac{\partial x_{1+2}}{\partial x_1} = \frac{y_{1+2}}{y_1}$ by pretending that (x_2,y_2) does not vary. Similarly, $\frac{\partial x_{1+2}}{\partial x_2} = \frac{y_{1+2}}{y_2}$. By assumption $\frac{\partial x_1}{\partial x} = c_{\alpha_1}\frac{y_1}{y}$ and $\frac{\partial x_2}{\partial x} = c_{\alpha_2}\frac{y_2}{y}$. Putting everything together we obtain

$$x'_{1+2} = (c_{\alpha_1} + c_{\alpha_2})\frac{y_3}{y}$$

and so the first claim.

For $x_{1\circ 2}$ we have $x_{1\circ 2} = x_1 \circ x_2$ and the chain rule gives

$$x'_{1\circ 2} = \frac{\partial x_{1\circ 2}}{\partial x_2}\frac{\partial x_2}{\partial x}$$

We find $\frac{\partial x_{1\circ 2}}{\partial x_2} = \frac{\partial x_1}{\partial x}\big|_{x=x_2,y=y_2} = c_{\alpha_1}(x_2,y_2)\frac{y_1(x_2,y_2)}{y} = c_{\alpha_1}\frac{y_{1\circ 2})}{y}$ since c_{α_1} is assumed to be constant. As we also have $\frac{\partial x_2}{\partial x} = c_{\alpha_2}\frac{y_2}{y}$ we obtain

$$x'_{1\circ 2} = c_{\alpha_1}(x_2,y_2)\cdot c_{\alpha_2}\frac{y_{1\circ 2}}{y_2}\frac{y_2}{y}.$$

Consequently, $c_{1\circ 2} = c_1 \cdot c_2$. □

This enables us to compute the constant c_α for all endomorphisms

$$r\varphi_q + s\colon \begin{array}{ccc} E & \longrightarrow & E, \\ P & \longmapsto & r\varphi_q(P) + sP \end{array}.$$

We only need to compute c_{id} and c_{φ_q} and apply Lemma 2.18. Obviously, $c_{[1]} = \frac{x'}{1} = 1$. By induction this implies $c_{[n]} = n$. [Use $nP = (n-1)P + 1P$.] Next, $\varphi_q(x, y) = (x^q, y^q)$ and so $c_{\varphi_q} = \frac{qx^{q-1}}{y} = 0$. Thus we obtain

Lemma 2.19. *Let E be an elliptic curve over a field of characteristic p and $r, s \in \mathbb{Z}$. Then*

- *$r\varphi_q + s$ is an endomorphism of E, and it's nonzero iff $(r, s) \neq (0, 0)$.*
- *$c_{r\varphi_q + s} = s$.*
- *$r\varphi_q + s$ is separable iff $p \nmid s$.*

Proof. The first two statements are done. For the last statement simply observe that $r_1' = c_{r\varphi+s} r_2 = s r_2$ and this is nonzero iff $s \cdot 1 \neq 0$ in the field. □

We will later see that $\varphi_q^2 = \varphi_q \circ \varphi_q$ can be expressed as one of the above endomorphisms. As it is obvious that $\varphi_q \circ [n] = [n] \circ \varphi_q$ Lemma 2.19 already considers all endomorphisms obtainable from the identity and the Frobenius φ_q by adding, negating and composing. It turns out that for many elliptic curves over finite fields there are no other endomorphisms.

After introducing the Weil pairing we will also be able to give a formula for computing the degree of $r\varphi_q + s$. That however depends on the structure of the torsion (and how endomorphisms act there).

2.10. Torsion. To learn more about the group structure we now study the scalar multiplication maps in more detail. Namely, we now consider the kernels of the scalar multiplication maps: Let E be an elliptic curve (defined) over a field k as usual, and let n be a positive integers. Then we define the *n-torsion $E[n]$ of E* by

$$E[n] := \left\{ P \in E(\overline{k}) \,\middle|\, nP = \mathcal{O} \right\}.$$

Clearly, this is a subgroup of E. To calculate it we need to know more about the scalar multiplication map $[n]$. Let's consider a few small n to get an impression.

2.10.1. 1-torsion. For $n = 1$ we have the identity and thus $E[1] = \{\mathcal{O}\}$.

2.10.2. 2-torsion. That was easy, so let's try $n = 2$. We have already calculated $[2]$ and seen that its degree is 4 and it is also separable if the characteristic is not 2. The x-coordinates denominator is $q(x) = 4(x^3 + ax + b)$, so $[2]P = \mathcal{O}$ if either $P = \mathcal{O}$ or $P = (x, 0)$ with $q(x) = 0$. Thus if $\operatorname{char} k \neq 2$ we obtain that $E[2]$ is a four element group whose elements all have order 1 or 2, that is,

$$E[2] = \left\{ (x, 0) \in \overline{k}^2 \,\middle|\, x^3 + ax + b = 0 \right\} \dot\cup \{\mathcal{O}\} \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2.$$

The situation in characteristic 3 is similar as here we can bring every elliptic curve in the form $y^2 = x^3 + a_2 x^2 + a_4 x + a_6$, so that again negation is simply changing the y-sign

and so the elements of $E[2]$ the points $(x,0)$ and $\mathcal{O}$. For characteristic 2 the situation is different. First of all, there are no elliptic curves in Weierstraß form; all those cubic curves are non-smooth. However, any curve can be transformed into $y^2+xy=x^3+a_2x^2+a_6$ or $y^2+a_3y=x^3+a_4x+a_6$. [Start with a curve in general Weierstraß form, and replace $x=a_1^2u+a_3a_1^{-1}$, $y=a_1^3v+a_1^{-3}(a_1^2a_4+a_3^2)$ if $a_1\neq 0$ or $x=u+a_2$, $y=v$ otherwise.] In the first form one finds $[2](x,y)=\left(\frac{x^4+a_6}{x^2},\frac{x^4+a_6}{x^2}+\frac{y+x^2}{x}\left(\frac{x^4+a_6}{x^2}-x\right)+y\right)$. Thus if $2P=\mathcal{O}$ then either $P=\mathcal{O}$ or $P=(0,\sqrt{a_6})$. So here we obtain two points and $E[2]\cong\mathbb{Z}_2$. And in the second form $[2](x,y)=\left(\frac{x^4+a_4^2}{a_3^2},a_3+\frac{x^2+a_4}{a_3}\left(\frac{x^4+a_4^2}{a_3^2}-x\right)+y\right)$, and $2P=\mathcal{O}$ only for $P=\mathcal{O}$. So we obtain only a single points and $E[2]=\{\mathcal{O}\}$.

Proposition 2.20. *Let E be an elliptic curve over a field k. Then*

(i) if $\operatorname{char} k=2$ *we have* $E[2]\cong 0$ *or* $E[2]\cong\mathbb{Z}_2$.

(ii) if $\operatorname{char} k\neq 2$ *we have* $E[2]\cong\mathbb{Z}_2\oplus\mathbb{Z}_2$. □

2.10.3. 3-torsion. Let's try to do one more step and consider $n=3$. As $3P=\mathcal{O}$ iff $2P=-P$ we do not need to derive formulae for point tripling [3]. Moreover, $2P=P$ implies $P=\mathcal{O}$, so we actually only need to consider the x-coordinates: $m^2-2x=x$, where $m=\frac{3x^2+a}{2y}$ is the tangent's slope as usual. The y-coordinate of $2P$ can then only be $-y$. Rewriting this we obtain $-(3x^2+a)^2+12x(x^3+ax+b)=0$ or $3x^4+6ax^2+12bx-a^2=0$. Unless $\operatorname{char} k=3$ this is a degree 4 equation and all its solutions are indeed different: its discriminant is $-6912\Delta^2$ and thus nonzero. [You can also run the Euclidean algorithm on the polynomial and its derivative to see that.] Thus we find four values of x. Each of these four values of x produces two points (x,y) since x^3+ax+b has no multiple roots and so cannot vanish at these x-values. So including $\mathcal{O}$ we find 9 points in $E[3]$, each of order 1 or 3. In characteristic 3 we may assume E in the form $y^2=x^3+a_2x^2+a_4x+a_6$. When solving $2P=-P$ we now get $m^2-a_2-2x=x$ with $m=\frac{2a_2x+a_4}{2y}$. This simplifies to $a_2x^3+a_2a_6-a_4^2=0$. As $x^3+a_2x^2+a_4x+a_6$ has no multiple roots we have $a_2\neq 0$ or $a_4\neq 0$ because $x^3+a^3=(x+a)^3$ in characteristic 3. In case $a_2\neq 0$ the equation is of the form $a_2(x^3+a)=0$ and thus has a single triple zero. Corresponding to that x we find two values of y, again using that the right hand side polynomial has no multiple root. So we obtain three points in $E[3]$. Finally, if $a_2=0$ we are left with the equation $-a_4^2=0$ which is never true, and so $E[2]=\{\mathcal{O}\}$.

Proposition 2.21. *Let E be an elliptic curve over a field k. Then*

(i) if $\operatorname{char} k=3$ *we have* $E[3]\cong 0$ *or* $E[3]\cong\mathbb{Z}_3$.

(ii) if $\operatorname{char} k\neq 3$ *we have* $E[3]\cong\mathbb{Z}_3\oplus\mathbb{Z}_3$. □

2.10.4. 4-torsion. The situation for $n=4$ can still be studied along the lines of the case $n=3$. There is a faster way to get there by considering $E[4]$ as preimage of $E[2]$ under point doubling [2].

2.10.5. General case. For n a prime, you may have guessed the general situation already. To nicely state the result we need a further disctinction.

Definition 2.22. *We call an elliptic curve*

- ordinary *iff* $E[p] \cong \mathbb{Z}_p$, *and*
- supersingular *iff* $E[p] = 0$.

The names *ordinary* and *supersingular* go back to the study of the endomorphism rings and the j-invariants. In characteristic zero, the endomorphism ring of an elliptic curve usually is $\mathbb{Z}$, correspondingly we speak of *ordinary j-invariants*. The *singular j-invariants* are those corresponding to curves with larger endomorphism rings than usual, namely an order in a quadratic extension of $\mathbb{Q}$. (An order in $K|\mathbb{Q}$ is a subring of K and also a $\mathbb{Z}$-module of rank $K : \mathbb{Q}$.) In finite characteristic, the endomorphism ring can be even larger, namely, non-commutative and rank 4 rather than commutative and rank 1 or 2, and so the corresponding j-invariants were called *supersingular*.

The following theorem claims that every elliptic curve is either supersingular or regular and gives a complete description of the n-torsion of an elliptic curve (over the algebraic closure).

Theorem 2.23. *Consider an elliptic curve E over a field k of characteristic p and a positive integer n.*

(i) If $p \nmid n$ then we obtain the n-torsion

$$E[n] \cong \mathbb{Z}_n \oplus \mathbb{Z}_n.$$

(ii) If n is a p-power then

$$E[n] \cong \begin{cases} 0 & \text{if } E \text{ is ordinary,} \\ \mathbb{Z}_n & \text{if } E \text{ is supersingular.} \end{cases}$$

Combining, write $n = p^r n'$ with $p \nmid n'$. For ordinary curves we have $E[n] \cong \mathbb{Z}_n \oplus \mathbb{Z}_{n'}$ whereas for supersingular curves we have $E[n] \cong \mathbb{Z}_{n'} \oplus \mathbb{Z}_{n'}$.

The proof will occupy us for some time.

2.10.6. Induced torsion endomorphisms. An important consequence of this characterization is that we can study the behaviour of endomorphism (and also arbitrary group endomorphisms) restricted to the n-torsion. Fix n coprime to the characteristic. Then the previous theorem implies that $E[n] \cong \mathbb{Z}_n \oplus \mathbb{Z}_n$. To make this explicit choose $\beta_1, \beta_2 \in E[n]$ such that every n-torsion point can be expressed uniquely as $m_1\beta_1 + m_2\beta_2$ with $m_1, m_2 \in \mathbb{Z}_n$. If now $\alpha\colon E(\overline{k}) \to E(\overline{k})$ is a group morphism (or even an endomorphism) then α is $\mathbb{Z}_n$-linear on $E[n]$: $\alpha(m_1\beta_1 + m_2\beta_2) = m_1\alpha(\beta_1) + m_2\alpha(\beta_2)$. By expressing $\alpha(\beta_1) = a\beta_1 + c\beta_2$ and $\alpha(\beta_2) = b\beta_1 + d\beta_2$ we can describe the effect of α on the n-torsion by a matrix:

$$\alpha_n := \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Now, $\alpha(m_1\beta_1 + m_2\beta_2) = n_1\beta_1 + n_2\beta_2$ with $\begin{bmatrix} n_2 \\ n_1 \end{bmatrix} = \alpha_n \cdot \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}$. Many important properties do not depend on the choice of the basis β_1, β_2 and these give meaningful information about the endomorphism α.

We will use that construction to study certain endomorphisms of elliptic curves (see Section 2.12). But one can also use it to construct representations of Galois groups: If α is a field automorphism of $\overline{k}$ fixing k then we obtain a matrix α_n on the n-torsion. This defines a representation, namely a map $\mathrm{Gal}(\overline{k}|k) \to \mathrm{GL}_2(\mathbb{Z}_n)$. See Washington (2003), Example 3.1 for an illustration.

2.11. Division polynomials. Our next purpose is to prove Theorem 2.23. To that end we will determine enough information on the scalar multiplication endomorphism $[n]$. To that end we introduce division polynomials. There are three sequences: Writing $nP = \left(\frac{p_1(x)}{q_1(x)}, \frac{y \cdot p_2(x)}{q_2(x)}\right)$ the first family will describe q_1 and q_2, and the other two will give p_1 and yp_2.

Definition 2.24 (Division polynomials). *We define polynomials in $\mathbb{Z}[x, y, a, b]$ as follows:*

$$\psi_n := \begin{cases} 0 & \text{if } n = 0, \\ 1 & \text{if } n = 1, \\ 2y & \text{if } n = 2, \\ 3x^4 + 6ax^2 + 12bx - a^2 & \text{if } n = 3, \\ 4y(x^6 + 5ax^4 + 20bx^3 - 5a^2x^2 - 4abx - 8b^2 - a^3) & \text{if } n = 4, \\ \psi_{m+2}\psi_m^3 - \psi_{m-1}\psi_{m+1}^3 & \text{if } n = 2m+1 \geq 5,\ m \geq 2, \\ \frac{1}{2y}(\psi_m)(\psi_{m+2}\psi_{m-1}^2 - \psi_{m-2}\psi_{m+1}^2) & \text{if } n = 2m \geq 6,\ m \geq 3. \end{cases}$$

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

$$\omega_n := \frac{1}{4y}(\psi_{n+2}\psi_{n-1}^2 - \psi_{n-2}\psi_{n+1}^2).$$

Note that ψ_n depends only on other polynomials ψ_ℓ with $\ell < n$ and so it is well defined. Despite the few divisions all are polynomials, as we will check in the next lemma. You should recognize ψ_3 [from Section 2.10.3] and maybe also ψ_4. Moreover, they describe exactly the scalar multiplications $[n]$ of elliptic curves in Weierstraß form: $[n](x, y) = \left(\frac{\phi_n}{\psi_n^2}, \frac{\omega_n}{\psi_n^3}\right)$, which explains — after some rearrangements — their definition. We will now derive some essential properties.

Lemma 2.25. *Abbreviate $R = \mathbb{Z}[x, y^2, a, b]$. For any n we have:*

(i) $\psi_n \in \begin{cases} 2yR & \text{if } n \equiv_2 0, \\ R & \text{if } n \equiv_2 1. \end{cases}$

(ii) The weighted degree of ψ_n is $n^2 - 1$ where the weight of x is 2 and the weight of y is 3 and the corresponding leading coefficient is n. Moreover, the corresponding leading coefficient after possible replacements of y^2 by $x^3 + ax + b$ is given by

$$\operatorname{lt} \psi_n = \begin{cases} nyx^{\frac{n^2-4}{2}} & \text{if } n \equiv_2 0, \\ nx^{\frac{n^2-1}{2}} & \text{if } n \equiv_2 1. \end{cases}$$

(iii) $\psi_n \in \begin{cases} 0+4yR & \text{if } n \equiv_4 0, \\ (x^2+a)^{\frac{n^2-1}{4}} + 2R & \text{if } n \equiv_4 1, \\ 2y(x^2+a)^{\frac{n^2-4}{4}} + 4yR & \text{if } n \equiv_4 2, \\ (x^2+a)^{\frac{n^2-1}{4}} + 2R & \text{if } n \equiv_4 3, \end{cases}$

(iv) $\phi_n \in R$.

(v) $\omega_n \in \begin{cases} R & \text{if } n \equiv_2 0, \\ yR & \text{if } n \equiv_2 1. \end{cases}$

For all n we have $\phi_n, \psi_n^2 \in R$. Replacing y^2 with x^3+ax+b we can interpret any polynomial in R as a polynomial in $\mathbb{Z}[x,a,b]$.

(vi) *For any field K and $a,b \in K$ and for any n the polynomials $\phi_n|_{y^2=x^3+ax+b}$ and $\psi_n^2|_{y^2=x^3+ax+b}$ are coprime.*

Note that (iii) is a stronger version of (i).

Proof. (i): For $n \le 4$ there is nothing to do. So assume $n > 4$. Let $m = \lfloor \frac{n}{2} \rfloor$, so that $n = 2m$ or $n = 2m+1$. Denote $R = \mathbb{Z}[x, y^2, a, b]$. We distinguish four cases according to $n \operatorname{rem} 4$:

$n \operatorname{rem} 4$	$m \operatorname{rem} 2$	ψ_m	$(\psi_{m+2}$	ψ_{m-1}^2	$-$	ψ_{m-2}	$\psi_{m+1}^2)$	$=$	$2y\psi_n$
0	0	$2yR$	$(2yR$	R^2	$-$	$2yR$	$R^2)$	$\subseteq$	$4y^2R$
2	1	R	$(R$	$4y^2R$	$-$	R	$4y^2R)$	$\subseteq$	$4y^2R$

$n \operatorname{rem} 4$	$m \operatorname{rem} 2$	ψ_{m+2}	ψ_m^3	$-$	ψ_{m-1}	ψ_{m+1}^3	$=$	ψ_n
1	0	$2yR$	$(2yR)^3$	$-$	R	R^3	$\subseteq$	R
3	1	R	R^3	$-$	$(2yR)$	$(2yR)^3$	$\subseteq$	R

The table shows that if n is even then $\psi_n \in 2yR$ and if n is odd in R.

(ii): We actually claim that the weighted leading term of ψ_n after possible replacements of y^2 with x^3 is given by

$$\operatorname{lt} \psi_n = \begin{cases} nyx^{\frac{n^2-4}{2}} & \text{if } n \equiv_2 0, \\ nx^{\frac{n^2-1}{2}} & \text{if } n \equiv_2 1. \end{cases}$$

The claim being true for $n \le 4$ we consider $n \ge 5$. Just writing down the leading terms and eventually replacing y^2 with $\operatorname{lt}(x^3+ax+b)$ we obtain:

$n \operatorname{rem} 4$	ψ_m	$\Big($	ψ_{m+2}	ψ_{m-1}^2	$-$	ψ_{m-2}	ψ_{m+1}^2	$\Big)$	$=$	$2y\psi_n$
0	$myx^{\frac{m^2-4}{2}}$	$\Big($	$(m+2)yx^{\frac{m^2+4m}{2}}$	$(m-1)^2x^{\frac{2m^2-4m}{2}}$	$-$	$(m-2)yx^{\frac{m^2-4m}{2}}$	$(m+1)^2x^{\frac{2m^2+4m}{2}}$	$\Big)$	$\subseteq$	$2y \cdot 2myx^{\frac{(2m)^2-4}{2}}$
2	$mx^{\frac{m^2-1}{2}}$	$\Big($	$(m+2)x^{\frac{m^2+4m+3}{2}}$	$(m-1)^2y^2x^{\frac{2m^2-4m-6}{2}}$	$-$	$(m-2)x^{\frac{m^2-4m+3}{2}}$	$(m+1)^2y^2x^{\frac{2m^2+4m-6}{2}}$	$\Big)$	$\subseteq$	$2y \cdot 2myx^{\frac{(2m)^2-4}{2}}$

$n \operatorname{rem} 4$	ψ_{m+2}	ψ_m^3	$-$	ψ_{m-1}	ψ_{m+1}^3	$=$	ψ_n
1	$(m+2)yx^{\frac{m^2+4m}{2}}$	$m^3y^3x^{\frac{3m^2-12}{2}}$	$-$	$(m-1)x^{\frac{m^2-2m}{2}}$	$(m+1)^3x^{\frac{3m^2+6m}{2}}$	$\subseteq$	$(2m+1)x^{\frac{(2m+1)^2-1}{2}}$
3	$(m+2)x^{\frac{m^2+4m+3}{2}}$	$m^3x^{\frac{3m^2-3}{2}}$	$-$	$(m-1)yx^{\frac{m^2-2m-3}{2}}$	$(m+1)^3y^3x^{\frac{3m^2+6m-9}{2}}$	$\subseteq$	$(2m+1)x^{\frac{(2m+1)^2-1}{2}}$

(iii): Let $m = \lfloor \frac{n}{2} \rfloor$, so that $n = 2m$ or $n = 2m+1$. We need to check the claim through the recursion. As the result varies with n mod 4 and the ingredients with m mod 4, we consider eight cases according to n mod 8.

n mod 8	m mod 4	ψ_m	$\left($	ψ_{m+2}	ψ_{m-1}^2	$-$	ψ_{m-2}	ψ_{m+1}^2	$\right)$	$=$	$2y\psi_n$
0	0	$4yR$	$\left($	$2yR$	R	$-$	$2yR$	R	$\right)$	$\subseteq$	$8y^2R$
2	1	$t^{\frac{m^2-1}{4}}+2R$	$\left($	R	$(4y)^2R$	$-$	$t^{\frac{(m-2)^2-1}{4}}+2R$	$(2y)^2t^{\frac{(m+1)^2-4}{2}}+4yR$	$\right)$	$\subseteq$	$4y^2t^{\frac{n^2-4}{4}}+8y^2R$
4	2	$2yR$	$\left($	$4yR$	R	$-$	$4yR$	R	$\right)$	$\subseteq$	$8y^2R$
6	3	$t^{\frac{m^2-1}{4}}+2R$	$\left($	$t^{\frac{(m+2)^2-1}{4}}+2R$	$(2y)^2t^{\frac{(m-1)^2-4}{2}}+4yR$	$-$	R	$(4y)^2R$	$\right)$	$\subseteq$	$4y^2t^{\frac{n^2-4}{4}}+8y^2R$

n mod 4	m mod 4	ψ_{m+2}	ψ_m^3	$-$	ψ_{m-1}	ψ_{m+1}^3	$=$	ψ_n
1	0	$2yR$	$(4y)^3R$	$-$	$t^{\frac{(m-1)^2-1}{4}}+2R$	$t^3\frac{(m+1)^2-1}{4}+2R$	$\subseteq$	$t^{\frac{n^2-1}{4}}+2R$
3	1	$t^{\frac{(m+2)^2-1}{4}}+2R$	$t^3\frac{m^2-1}{4}+2R$	$-$	$4yR$	$(2y)^3R$	$\subseteq$	$t^{\frac{n^2-1}{4}}+2R$
5	2	$4yR$	$(2y)^3R$	$-$	$t^{\frac{(m-1)^2-4}{4}}+2R$	$t^3\frac{(m+1)^2-4}{4}+2R$	$\subseteq$	$t^{\frac{n^2-1}{4}}+2R$
7	3	$t^{\frac{(m+2)^2-1}{4}}+2R$	$t^{\frac{m^2-1}{4}}+2R$	$-$	$2yR$	$(4y)^3R$	$\subseteq$	$t^{\frac{n^2-1}{4}}+2R$

(iv): We either have $\phi_n \in xR^2-(2yR)\cdot(2yR) \subseteq R$ if n is odd or $\phi_n \in x(2yR)^2 - R\cdot R \subseteq R$ if n is even. [Notice that $y \notin R$ but $y^2 \in R$.]

(v): Using (i) we find: If n is odd then $4y\omega_n \in R\cdot(2yR)^2 - R\cdot(2yR)^2 \subseteq 4y^2R$ and so $\omega_n \in yR$. If n is even then $4y\omega_n \in 2yR\cdot R^2 - 2yR\cdot R^2 \subseteq 2yR$ and so $2\omega_n \in R$.

To get rid of the remaining 2 we need to use (iii):

n rem 4	ψ_{n+2}	ψ_{n-1}^2	$-$	ψ_{n-2}	ψ_{n+1}^2	$=$	$14y\omega_n$
0	$2yt^{\frac{(n+2)^2-4}{4}}+4yR$	$t^{\frac{(n-1)^2-1}{2}}+2R$	$-$	$2yt^{\frac{(n-2)^2-4}{4}}+4yR$	$t^{\frac{(n+1)^2-1}{2}}+2R$	$\subseteq$	$4yR$
2	$4yR$	R	$-$	$4yR$	R	$\subseteq$	$4yR$

(vi): [—to do—] □

THEOREM 2.26. *Consider a base field k of characteristic different from 2, $a, b \in k$ and $n \in \mathbb{N}$. Then*

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

PROOF. [—to do—] □

2.11.1. The Weil pairing. Given that for n coprime to the characteristic, we now know that

$$E[n] \cong \mathbb{Z}_n \times \mathbb{Z}_n$$

we could define kind of a scalar product on $E[n]$ as follows. Fix a $\mathbb{Z}_n$-basis (T_1, T_2) of $E[n]$ and choose values for $e(T_i, T_j)$ in some appropriate group. Since we want e bilinear we then have $e(s_1T_1 + s_2T_2, t_1T_1 + t_2T_2) = \sum_{i,j} s_i e(T_i, T_j) t_j$. Actually, we want more: the pairing must also be non-degerenate, that is, if for all $T \in E[n]$ we have $e(S, T) = 0$ then we have $S = \mathcal{O}$, and also if for all $S \in E[n]$ we have $e(S, T) = 0$ then we have $T = \mathcal{O}$. We can grant this by requiring that the matrix $[e(T_i, T_j)]_{i,j}$ is invertible. All these things are now pairings on the n-torsion. However, we do not know anything about how to compute the pairing nor whether this is compatible with possible algebraic structures. In that light, it is only a minor complication to take a multiplicatively written group for the values: Let

$$\mu_n = \left\{ x \in \overline{k} \,\middle|\, x^n = 1 \right\}$$

be the group of nth roots of unity. Since n is coprime to the characteristic $\#\mu_n = n$ and so μ_n is a cyclic group of order n.

Theorem 2.27 (Weil pairing). *Let E be an elliptic curve defined over a field k and let n be a positive integer coprime to the characteristic of k. Then a* Weil pairing

$$e_n \colon E[n] \times E[n] \longrightarrow \mu_n$$

satisfying the following properties exists.

(i) e_n is bilinear, that is, for all $S, S_1, S_2, T, T_1, T_2 \in E[n]$

$$e_n(S_1 + S_2, T) = e_n(S_1, T) \cdot e_n(S_2, T),$$
$$e_n(S, T_1 + T_2) = e_n(S, T_1) \cdot e_n(S, T_2).$$

(ii) e_n is non-degenerate, that is, for all $T \in E[n]$

$$\forall S \in E[n] \colon e_n(S, T) = 1 \implies T = \mathcal{O},$$
$$\forall S \in E[n] \colon e_n(T, S) = 1 \implies T = \mathcal{O}.$$

(iii) e_n is antisymmetric, that is, for all T

$$e_n(T, T) = 1.$$

In particular, $e_n(T, S) = e_n(S, T)^{-1}$.

(iv) e_n is compatible with the Galois actions, that is, for every automorphism σ of $\overline{k}$ fixing k (in particular, for a curve in Weierstraß form this means that $\sigma(a) = a$ and $\sigma(b) = b$) we have

$$e_n(\sigma S, \sigma T) = \sigma(e_n(S, T)).$$

(v) For every endomorphism α of E we have

$$e_n(\alpha(S), \alpha(T)) = e_n(S, T)^{\deg \alpha}.$$

We do not give a proof here. Just a word on it. The Weil pairing can be defined after studying the behaviour of functions on an elliptic curve. One constructs an algebraic function g_T on E that has a single zero at each T' with $nT' = T$ and a single pole at each n-torsion point. Then we define $e_n(S,T) = g_T(P+S)/g_T(P)$ for some point P so that this is nicely defined. The tricky points are the existence of g_T and that the various choices do not influence the result.

[—to do—]

There is a different construction yielding the Tate pairing $\langle \cdot, \cdot \rangle_n$ which is slightly easier to compute and is not antisymmetric. The two are connected by a congruence of the form

$$e_n(S,T) \equiv \frac{\langle T, S \rangle_n}{\langle S, T \rangle_n}.$$

We now derive a few consequence of the existence of the Weil pairing.

COROLLARY 2.28. *Let $\{T_1, T_2\}$ be a $\mathbb{Z}_n$-basis of $E[n]$. Then $e_n(T_1, T_2)$ is a primitive nth root of unity.*

PROOF. This is a consequence of the non-degeneracy. Clearly, $\zeta = e_n(T_1, T_2)$ is an nth root of unity. Thus $\zeta^d = 1$ for some d dividing n. We have to show that $d = n$. Consider $T = dT_2$. Then $e_n(T_1, T) = e(T_1, T_2)^d = \zeta^d = 1$. And also $e_n(T_2, T) = e(T_2, T_2)^d = 1^d = 1$. But then $e_n(s_1T_1 + s_2T_2, T) = e_n(T_1, T)^{s_1} e_n(T_2, T)^{s_2} = 1$. Since $\{T_1, T_2\}$ generates $E[n]$ we obtain $dT_2 = T = \mathcal{O}$ by Theorem 2.27(ii). This implies $d = n$ and thus ζ is a primitve nth root of unity. □

Note that the Weil pairing is not uniquely defined by Theorem 2.27. If r is coprime to n then e_n^r also fulfills all the wanted properties. However, up to this transformation the Weil pairing is unique. To see this just note that e_n is defined by the value $e_n(T_1, T_2)$ for a basis $\{T_1, T_2\}$ of $E[n]$. And raising to a power coprime to n simply replaces one primitive nth root of unity with another.

COROLLARY 2.29. *If $E[n] \subseteq E(k)$ then $\mu_n \subseteq k$.*

Notice that you can apply this fact also for extensions of k, as we only require that E is defined over k, and then it is also defined over any extension of k.

PROOF. As before fix a $\mathbb{Z}_n$-basis $\{T_1, T_2\}$ of $E[n]$, and consider $\zeta = e_n(T_1, T_2)$. This is a primitive nth root of unity, that is, it generates μ_n. Thus it suffices to show that $\zeta \in k$. Notice that $k(\zeta)|k$ is a Galois extension since n is not a multiple of the characteristic. So let σ be any automorphism of $\overline{k}$ fixing k (extending any $\sigma \in \mathrm{Gal}(k(\zeta)|k)$). By assumption $T_1, T_2 \in E(k)$ thus we obtain

$$\zeta = e_n(T_1, T_2) = e_n(\sigma T_1, \sigma T_2) = \sigma(\zeta).$$

By the fundamental theorem of Galois theory ζ must be in k. □

This immediately implies that over the rationals $\mathbb{Q}$ we rarely have all n-torsion rational, since $\mathbb{Q}$ only contains μ_1 and μ_2 but no higher μ_n:

THEOREM 2.30. *Let E be an elliptic curve over $\mathbb{Q}$. Then for $n \geq 3$ we have $E[n] \not\subseteq E(\mathbb{Q})$.* □

When $n = 2$ we can have $E[2] \subseteq E(\mathbb{Q})$. This happens whenever the roots of the right hand side of the Weierstraß equation are all rationals. For example, $E : y^2 = x(x-1)(x-2)$ has $E[2] = \{\mathcal{O}, (0,0), (1,0), (2,0)\}$. Actually, for $n > 12$ or $n = 11$ there is *no* elliptic curve over $\mathbb{Q}$ with a rational point of order n.

Our next issue is to consider what the Weil pairing can tell about endomorphisms, in particular, their degree. This clearly has to involve Theorem 2.27(v). As already pointed out in Section 2.10.6 we can describe the restriction of an endomorphism α on the n-torsion by a $\mathbb{Z}_n$-matrix after choosing a $\mathbb{Z}_n$-basis $\{T_1, T_2\}$ of $E[n]$. Based on this, Theorem 2.27(v) allows us to characterize the degree.

PROPOSITION 2.31. *Let α be an endomorphism of an elliptic curve E defined over a field k, and n an integer coprime to the characteristic of k. Then*

$$\deg(\alpha) \equiv_n \det(\alpha_n).$$

PROOF. By Corollary 2.28 the value $\zeta = e_n(T_1, T_2)$ is a primitive nth root of unity on a $\mathbb{Z}_n$-basis $\{T_1, T_2\}$ of $E[n]$. Let $\alpha_n = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be the matrix representing α with respect to that basis. Let's compute the pairing on the image of the basis:

$$\begin{aligned} e_n(\alpha(T_1), \alpha(T_2)) &= e_n(aT_1 + cT_2, bT_1 + dT_2) \\ &= e_n(T_1, T_1)^{ab} e_n(T_1, T_2)^{ad} e_n(T_2, T_1)^{bc} e_n(T_2, T_2)^{cd} \\ &= e_n(T_1, T_2)^{ad-bc} = \zeta^{ad-bc}. \end{aligned}$$

On the other hand side, by Theorem 2.27(v), we obtain $e_n(\alpha(T_1), \alpha(T_2)) = e_n(T_1, T_2)^{\deg \alpha} = \zeta^{\deg \alpha}$. Since ζ is a primitive nth root of unity we obtain $\deg(\alpha) \equiv_n ad - bc = \det(\alpha_n)$. □

It seems that we have just reduced one difficult problem to another one. However, this connection now allows us to deduce a really simple tool to compute the degree of many endomorphisms without even trying to represent them by quolynomials:

PROPOSITION 2.32. *Let α, β be endomorphisms of an elliptic curve E defined over k, and $a, b \in \mathbb{Z}$. Then*

$$\deg(a\alpha + b\beta) = a^2 \deg \alpha + b^2 \deg \beta + ab\,(\deg(\alpha + \beta) - \deg \alpha - \deg \beta)\,.$$

In particular, we can now immediately tell the degree of an endomorphism $r\varphi_q + s$. This will be very fruitful.

PROOF. Fix a number n coprime to the characteristic. Clearly, the induced matrix $(a\alpha + b\beta)_n$ of $a\alpha + b\beta$ on the n-torsion is $a\alpha_n + b\beta_n$. By Proposition 2.31 we have $\deg(a\alpha + b\beta) \equiv_n \det(a\alpha_n + b\beta_n)$. Now, for 2×2-matrices we can check that

$$\det(a\alpha_n + b\beta_n) = a^2 \det \alpha_n + b^2 \det \beta_n + ab\,(\det(\alpha_n + \beta_n) - \det \alpha_n - \det \beta_n)\,.$$

Using Proposition 2.31 also for the determinants on the right we infer that

$$\deg(a\alpha + b\beta) \equiv_n a^2 \deg \alpha + b^2 \deg \beta + ab\,(\deg(\alpha + \beta) - \deg \alpha - \deg \beta)\,.$$

But this holds for all n coprime to the characteristic and thus must be an equality in $\mathbb{Z}$. □

2.12. All about Frobenius. We have already defined the Frobenius endomorphism for an elliptic curve E defined over $\mathbb{F}_q$ earlier:

$$\varphi_q\colon \begin{array}{ccc} E & \longrightarrow & E, \\ (x,y) & \longmapsto & (x^q, y^q) \end{array}.$$

We have seen that it is not separable and has degree q. Moreover, it helps to characterize the $\mathbb{F}_q$-rational points:

$$E(\mathbb{F}_q) = \ker(\varphi_q - 1).$$

By definition $\ker(\varphi_q - 1)$ are all points $P = (x, y) \in E$ such that $\mathcal{O} = (\varphi_q - 1)(P) = \varphi_q(P) - P$ or $\varphi_q(P) = P$. Thus we obtain $x^q = x$ and $y^q = y$. This implies that $x, y \in \mathbb{F}_q$ and so $P \in E(\mathbb{F}_q)$. Vice versa, every point P with coordinates in $\mathbb{F}_q$ is fixed by φ_q and thus $(\varphi_q - 1)(P) = \mathcal{O}$. Together with the results from the previous section we can now express the number of $\mathbb{F}_q$-rational points:

$$\#E(\mathbb{F}_q) = \deg(\varphi_q - 1)$$

by simply observing that $\varphi_q - 1$ is separable using Lemma 2.19.

When you elementary try to compute the size of an elliptic curve you find $0 \le \#E(\mathbb{F}_q) \le 2q + 1$. You may also write down the formula

$$\#E(\mathbb{F}_q) = q + 1 + \sum_{x \in \mathbb{F}_q} \left(\frac{x^3 + ax + b}{\mathbb{F}_q} \right)$$

where $\left(\frac{a}{\mathbb{F}_q}\right)$ is 0 if $a = 0$, $+1$ if $a \neq 0$ is a square and -1 if $a \neq 0$ is not a square. However, we can do much better using the presented theory:

THEOREM 2.33 (Hasse). *If E is an elliptic curve defined over $\mathbb{F}_q$ then*

$$\#E(\mathbb{F}_q) = q + 1 - t \qquad \text{with } |t| \le 2\sqrt{q}.$$

PROOF. Write $\deg(\varphi_q - 1) = \#E(\mathbb{F}_q) = q + 1 - t$. Recall $\deg \varphi_q = q$ and $\deg(-1) = 1$. By Proposition 2.32 we obtain

$$\begin{aligned} \deg(r\varphi_q - s) &= r^2 \deg \varphi_q + s^2 \deg(-1) + rs \underbrace{(\deg(\varphi_q - 1) - \deg \varphi_q - \deg(-1))}_{=q+1-t-q-1=-t} \\ &= r^2 q + s^2 - rst \\ &= r^2 \left(\frac{s^2}{r^2} - t\frac{s}{r} + q \right). \end{aligned}$$

Of course, $\deg(r\varphi_q - s) \ge 0$ for all $r, s \in \mathbb{Z}$. Thus the polynomial

$$X^2 - tX + q$$

must be nonnegative for all rational numbers $\frac{s}{r}$. Since the rational numbers (even if we restrict to $p \nmid s$) are dense in $\mathbb{R}$ the polynomial must be completely nonnegative. Thus it cannot have two real zeros and its discriminant $t^2 - 4q$ must be negative or zero. And $t^2 \le 4q$ is the claim. □

To ease the life Washington only proves Theorem 2.27(v) only for separable endomorphisms and the Frobenius endomorphism. Then we can use Proposition 2.32 only for endomorphisms where α, β and $a\alpha + b\beta$ are separable or a Frobenius. In the previous proof this means, that we need $r\varphi_q - s$, φ_q and -1 each be either separable or Frobenius. This is the case if $p \nmid s$. But even then we can continue: the fractions $\frac{s}{r}$ with $p \nmid s$, $r \in \mathbb{Z} \setminus \{0\}$ are still dense in $\mathbb{R}$. Thus the quadratic polynomial $X^2 - tX + q$ must still be nonnegative on the reals.

The ingredients for Theorem 2.33 are on the one hand side the fact that $\varphi_q - 1$ is separable and describes the $\mathbb{F}_q$-rational points and on the other hand side the existence and properties of the Weil pairing.

And we can now also learn an important fact about the Frobenius endomorphism itself.

THEOREM 2.34 (Characteristic polynomial of the Frobenius). *Let E be an elliptic curve over $\mathbb{F}_q$. Define $t = q + 1 - \#E(\mathbb{F}_q)$ as before. Then*

$$\varphi_q^2 - t\varphi_q + q = 0$$

in the endomorphism ring of E. Moreover, if $\varphi_q^2 - k\varphi_q + q = 0$ then $k = t$.

Furthermore, $t \equiv_n \operatorname{trace}(\varphi_q)_n$ for all n coprime to the characteristic. This also uniquely determines t. Also, $q \equiv_n \det(\varphi_q)_n$ for all n coprime to the characteristic.

Because of the last statement we call t the trace of the Frobenius. We also call $X^2 - tX + q$ the characteristic polynomial of the Frobenius.

PROOF. Our aim is to prove that $\varphi_q^2 - t\varphi_q + q$ is the zero endomorphism. We have proved that any other endomorphism has a finite kernel. So we have to show that $\varphi_q^2 - t\varphi_q + q$ has an infinite kernel.

Again fix n coprime to the characteristic, choose a basis for the n-torsion $E[n]$ and express φ_q on it by a matrix

$$(\varphi_q)_n = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{Z}_n^{2\times 2}.$$

We want to find its characteristic polynomial χ. By definition $\chi(X) = (-1)^2 \det(\varphi_q - X)$ and it is a monic degree-2 polynomial. As we know by Proposition 2.31 that $\det(\alpha_n) \equiv_n \deg \alpha$ we consider enough values of χ. First, $\chi(0) = \det \varphi_q \equiv_n \deg \varphi_q = q$. Second, $\chi(1) = \det(\varphi_q - 1) \equiv \deg(\varphi_q - 1) = \#\ker(\varphi_1 - 1) = \#E(\mathbb{F}_q) = q + 1 - t$. Writing $\chi = X^2 - kX + q$ we see that $\chi(1) = q + 1 - k$ and $k = t$. Thus $X^2 - tX + q$ is the characteristic polynomial of the 2×2-matrix $(\varphi_q)_n$ and by Caley-Hamilton we obtain

$$(\varphi_q)_n^2 - t\,(\varphi_q)_n + q = 0 \qquad \text{in } \mathbb{Z}_n^{2\times 2}.$$

That is, $E[n]$ is in the kernel of $\varphi_q^2 - t\varphi_q + q$ for any n coprime to the characteristic. But this can only be true if this kernel is infinite and the map the zero endomorphism. In other words, we have

$$\varphi_q^2 - t\varphi_q + q = 0 \qquad \text{in } \operatorname{End} E.$$

Next, assume that $\varphi_q^2 - k\varphi_q + q = 0$. Then $k\varphi_q = t\varphi_q$ and thus $(k - t)\varphi_q = 0$. However, φ_q is bijective and thus $[k - t] = 0$. But again since only the zero endomorphism has an infinite kernel, we obtain $k = t$.

The statements about trace and determinant of $(\varphi_q)_n$ have been proved during this proof. □

This theorem tells us a lot about the structure of the endomorphism ring. Clearly, the scalar multiplications and the Frobenius are available. But already the square of the Frobenius can be expressed in terms of them. So instead of having integer polynomial expressions in the Frobenius we only need the linear ones. Moreover, in many cases (namely for ordinary curves), all endomorphisms are of the form $r\varphi_q + s$. Then

$$\operatorname{End} E \cong \mathbb{Z}[X]/\left\langle X^2 - tX + q \right\rangle.$$

is an order in $\mathbb{Q}[X]/\left\langle X^2 - tX + q \right\rangle$ (basically, an *order* of an extension K over $\mathbb{Q}$ is a subring of the extension field K and a $K : \mathbb{Q}$-dimensional $\mathbb{Z}$-module). In particular, the endomorphism ring can be embedded in a quadratic extension of the rationals, actually even in the ring of integers of such a quadratic extension.

In the light of Theorem 2.33 we see that usually φ_q cannot be equal to a scalar multiplication. Only in the (rare) case that $t^2 = 4q$, in particular, q must be a square, we observe that $(\varphi_q - 2\sqrt{q})^2 = 0$ and thus $\varphi_q = [2\sqrt{q}]$ using Theorem 2.16.

2.13. Structure. We have already seen the structure of the torsion subgroups of an elliptic curve. Actually, the structure of the $\mathbb{F}_q$-rational part is rather similar and we can see this similarly:

THEOREM 2.35. *Given an elliptic curve E over a finite field $\mathbb{F}_q$. Then*

- *$E(\mathbb{F}_q)$ is isomorphic to $\mathbb{Z}_m$ for some positive integer m, or*
- *$E(\mathbb{F}_q)$ is isomorphic to $\mathbb{Z}_{m_1} \oplus \mathbb{Z}_{m_2}$ for some positive integers m_1, m_2 with $m_2 \mid m_1$.*

PROOF. Clearly, $E(\mathbb{F}_q)$ is a finite abelian group and thus isomorphic to some group

$$\mathbb{Z}_{m_1} \oplus \mathbb{Z}_{m_2} \oplus \cdots \oplus \mathbb{Z}_{m_r}$$

with $m_{i+1} \mid m_i$ for $i \in \{1, \ldots, r-1\}$. Well, $\mathbb{Z}_{m_i}$ has m_r m_r-torsion elements. Thus the entire group has m_r^r m_r-torsion elements. But the m_r-torsion of E has at most m_r^2 elements. Thus $r \leq 2$ and we are done. □

One can show additionally that m_2 divides $q-1$ unless $t = \pm 2\sqrt{q}$, which can only happen if q is an even power. More precisely:

THEOREM 2.36. *Assume $N = m_1 m_2$ with $m_2 \mid m_1$, $m_1 = p^e m_1'$, $p \nmid m_1'$. Then there exists an elliptic curve E over $\mathbb{F}_q$ with $E(\mathbb{F}_q) \simeq \mathbb{Z}_{m_1} \oplus \mathbb{Z}_{m_2}$ iff*

- *$t = \pm\sqrt{q}$ (the extreme possibilities) and $m_1 = m_2'$, or*
- *$m_2 \mid q - 1$.* □

2.14. Determining size. Concerning the size of an elliptic curve we have already the best bounds. Recall that trivially $0 \leq \#E(\mathbb{F}_q) \leq 2q + 1$, and the theorem of Hasse even says $|\#E(\mathbb{F}_q) - (q+1)| \leq 2\sqrt{q}$. One can show that the latter is optimal. However, given a specific curve finding its size and structure is still a problem.

Counting the points of an elliptic curve can be done by brute force (which is no solution) taking time $\mathcal{O}\left(q^2\right)$, or a little better by trying each x and determining whether x^3+ax+b is a square. The latter corresponds to the formula

$$\#E(\mathbb{F}_q) = q + 1 + \sum_{x\in\mathbb{F}_q} \left(\frac{x^3+ax+b}{\mathbb{F}_q}\right)$$

mentioned earlier. Evaluating this takes time $\mathcal{O}^{\sim}(q)$.

2.14.1. Orders of points. Recall the theorem of Lagrange: given a finite group G and any subgroup H. Then the size of H divides the size of G. Clearly, the group $H = \langle P\rangle$ generated by an element $P \in G$ is a subgroup of G, it's size $\operatorname{ord} P$ is called the *order of* P. One can easily shows that

$$\operatorname{ord} P := \#\left\{iP \in G \,\middle|\, i \in \mathbb{Z}\right\} = \min\left\{i \in \mathbb{N}_{>0} \,\middle|\, iP = \mathcal{O}\right\}.$$

This order must divide the size of G by the cited theorem.

Combining this with the Hasse bound yields a way to determine the curve size as follows: Suppose we can find a point of order larger than $4\sqrt{q}$ in $E(\mathbb{F}_q)$. Then $\#E(\mathbb{F}_q) \equiv_{\operatorname{ord} P} 0$ and $|\#E(\mathbb{F}_q) - (q+1)| \le 2\sqrt{q}$ allows only a single value for $\#E(\mathbb{F}_q)$. Even if the point order is smaller than $4\sqrt{q}$ we often obtain only a short list of possiblities. We can also use several points.

Example 2.37. The curve $E\colon y^2 = x^3 + 2x + 1$ has the point $P = (0,1)$. Over $\mathbb{F}_{101}$ we find that $23P = \mathcal{O}$. Thus $\#E(\mathbb{F}_{101})$ is divisible by 23. The Hasse bound gives $|\#E(\mathbb{F}_{101}) - 102| \le 2\sqrt{101} \approx 20.1$, thus $82 \le \#E(\mathbb{F}_q) \le 122$. Thus $\#E(\mathbb{F}_{101})$ is either $92 = 4\cdot 23$ or $115 = 5\cdot 23$. In diesem Fall genügt es nun zu prüfen, ob es einen Punkt der Ordnung 2 gibt. Diese sind ja durch $y = 0$ charakterisiert, es zeigt sich, dass $Q = (-13, 0)$ auf der Kurve liegt und also muss $\#E(\mathbb{F}_{101}) = 92$ gelten.

Since we know that the 92-torsion is $\mathbb{Z}_{92} \oplus \mathbb{Z}_{92}$, the $\mathbb{F}_{101}$-rational part of the curve has 92 points, we can infer that $E(\mathbb{F}_{101})$ is either $\mathbb{Z}_{46} \oplus \mathbb{Z}_2$ or $\mathbb{Z}_{92}$. Checking the order 2 points again, we find that there is only one order 2 point (the mentioned one), and thus we find that $E(\mathbb{F}_{101})$ is cyclic of order 92. ◊

Example 2.38. The curve $E\colon y^2 = x^3 - 33x - 22$ over $\mathbb{F}_{101}$ has the point $P_1 = (36, -20)$ of order 11 and the point $P_2 = (32, -28)$ of order 9. Consequently, $\#E(\mathbb{F}_{101})$ is a multiple of $9\cdot 11$. The Hasse bound again gives $82 \le \#E(\mathbb{F}_q) \le 122$ and so $\#E(\mathbb{F}_q) = 99$.

The structure is obvious since 9 and 11 are coprime. Write $1 = s\cdot 11 + t\cdot 9$, then $P = sP_1 + tP_2$ has order 99, and so also this curve is cyclic. ◊

The question to be posed now is: how do we find the order of a point? Actually, we do this more or less by brute force.

Solution 1: Brute force. Just compute tP starting for $t = 0, 1, 2, 3, \ldots$ Runtime $\mathcal{O}(q)$.

Solution 2: Well, we know that the group size is in the Hasse interval $q + 1 - \left\lfloor 2\sqrt{q}\right\rfloor .. q + 1 + \left\lfloor 2\sqrt{q}\right\rfloor$. Can't we use that? Well, yes: Just run t only through that interval. This does not necessarily give the order but at least a multiple. Factoring the successful t and checking its divisors gives the order. Runtime $\mathcal{O}\left(q^{\frac{1}{2}}\right)$ plus the time for factoring t.

Solution 3: There is one further improvement, inspired by the baby-step giant-step algorithm for computing a discrete logarithm. Set $B := \lceil \sqrt[4]{q} \rceil$. Write $t = q + 1 + t_1 2B + t_0$ and try to find $t_0 \in 0..B-1$, $t_1 \in -(B-1)..B-1$ such that $(t - t_0)P = \pm t_0 P$ (by computing and comparing only the x-coordinates). Since we can tabulate $t_0 P$, we can independentely run through the options for t_1. Finally, factor t and check its divisors (well, cleverly). So we only need runtime $\mathcal{O}\left(q^{\frac{1}{4}}\right)$ plus the time for factoring t.

2.14.2. Subfield curves. Assume you want to determine the size of $E(\mathbb{F}_{q^m})$ for a curve that is defined over $\mathbb{F}_q$. Does knowledge on the size of $E(\mathbb{F}_q)$ help? Well, let's recap what we know:

- The characteristic polynomial of the Frobenis φ_q is $\chi = T^2 - tT + q$, then $\chi(\varphi_q) = 0$. Moreover, the polynomial is uniquely determined as a monic degree-2 polynomial by $\chi(0) = q$ and $\chi(\varphi_q) = 0$.
- The size of $E(\mathbb{F}_q)$ is $\chi(1)$.
- Moreover, the Frobenius of $\mathbb{F}_{q^m}$ is $\varphi_{q^m} = (\varphi_q)^m$.

Altogether this leads to

THEOREM 2.39. *Assume E is an elliptic curve defined over $\mathbb{F}_q$ with $1 - t + q$ $\mathbb{F}_q$-rational points. Let $\chi(T) = T^2 - tT + q$ be the characteristic polynomial of the Frobenius φ_q, and write $\chi = (T - \alpha)(T - \beta)$ with $\alpha, \beta \in \mathbb{C}$. Then for $n \geq 1$*

$$\#E(\mathbb{F}_{q^m}) = 1 - (\alpha^m + \beta^m) + q^m.$$

First, note that $s_n = \alpha^n + \beta^n$ is always an integer. [Either use Galois theory or note that $s_0 = 2$, $s_1 = \alpha + \beta$, $s_n = ts_{n-1} - qs_{n-2}$ for $n \geq 2$ (ie. $\chi(\Delta)(s) = 0$ when $\Delta(s) = (s_1, s_2, \dots)$).]

PROOF. We determine the characteristic polynomial of φ_{q^m}. Let $f = (T^m - \alpha^m)(T^m - \beta^m) = T^{2m} - (\alpha^m + \beta^m)T + q^m$. As noted f is an integer polynomial. Further, it is monic of degree 2, and $f(0) = q^n$. Further, χ divides f since $T - \alpha$ divides $T^m - \alpha^m$ and $T - \beta$ divides $T^m - \beta^m$. Thus $f(\varphi_q) = 0$ and $\varphi_q^{2m} - (\alpha^m + \beta^m)\varphi_q^m + q^m = 0$. As $\varphi_{q^m} = \varphi_q^m$ the polynomial $T^2 - (\alpha^m + \beta^m)T + q^m$ must be the characteristic polynomial of φ_{q^m}. Evaluating at 1 gives the claim. □

Unfortunately, subfield curves are rare and it could be that these are insecure for cryptographic applications.

2.14.3. Schoof's algorithm. Counting the size of $\mathbb{F}_q$-rational parts arbitrary curves with all the solutions so far can best be solved in time $\mathcal{O}\left(q^{\frac{1}{4}}\right)$. Only Schoof (1985) found a polynomial time algorithm. The idea actually is simple: We exploit the description of the ℓ-torsion and the characteristic polynomial of the Frobenius.

Assume you have a point (x, y) of order ℓ. Then we know that

$$\varphi_q^2(x, y) + q \cdot (x, y) = t \cdot \varphi_q(x, y).$$

Actually, we can compute all ingredients of this equation but t. And the equation determines t modulo the order ℓ of (x, y). Well, this is it: take a point of order ℓ compute $\varphi_q^2(x, y) + q \cdot (x, y)$

and $\varphi_q(x,y)$ and determine $t \bmod \ell$ by any discrete log algorithm. Do this for various ℓ whose product exceeds $4\sqrt{q}$, the information gathered then determines t and so we have $\#E(\mathbb{F}_q) = q + 1 - t$.

This could be the entire story. However, there are a few details to be resolved. First: it may happen that we do not have a point of order ℓ, in particular, such a point most of the time does not exist in $E(\mathbb{F}_q)$. So we would have to work in an extension. But even then, a point of order ℓ still had to be found. Instead, we work in the ring extension

$$R = \mathbb{F}_q[x,y] / \left\langle \psi_\ell(x), -y^2 + x^3 + ax + b \right\rangle.$$

Any order ℓ point has coordinates in this ring. However, we have to be careful not to overuse this. One problem is that this is only a ring, another that the degree of ψ_ℓ is roughly $\frac{1}{2}\ell^2$. This is polynomial in ℓ but large: a single multiplication in R will cost around $\mathcal{O}\left(\ell^4\right)$. Actually, we translate everything 'down to earth' and just work with univariate polynomials, usually reduced modulo ψ_ℓ. So let's do it.

First, choose a set S of primes such that $\prod_{\ell \in S} \ell > 4\sqrt{q}$. Make sure that $\max S$ is small: we need $\max S \in \mathcal{O}(\log q)$. We could simply take the smallest primes, that would guarantee this. Even, if you leave out about half of the primes this still is fine. Actually, we do want to skip at least the characteristic. Now, it suffices to determine $t \bmod \ell$ for $\ell \in S$ and use the Chinese Remainder theorem to determine $t \bmod \prod_{\ell \in S} \ell$. This determines $t \in [-2\sqrt{q}, 2\sqrt{q}]$.

Case $\ell = 2$. In order to determine $t \bmod 2$ we only need to determine whether there is a point of order 2. This is the same as checking whether $x^3 + ax + b$ has a zero in $\mathbb{F}_q$. To do that we determine $h := \gcd(x^q - x, x^3 + ax + b)$. This could be done by the Euclidean algorithm. However, the runtime will be $\mathcal{O}^\sim(q)$ which is intolerable. To remedy this we first compute $x_q := x^q \bmod x^3 + ax + b$. By square-and-multiply this needs only $\mathcal{O}(\log q)$ steps, each of which can be performed in time $\mathcal{O}\left(\log^2 q\right)$ (or better). Then running the Euclidean algorithm to compute $h := \gcd(x_q - x, x^3 + ax + b)$ is fast. If $h = 1$ then there is no point of order 2 and $\#E(\mathbb{F}_q)$ is odd, if h has degree 1 then there is exactly one point of order 2 and $\#E(\mathbb{F}_q)$ is divisible by 2 (and the $\mathbb{F}_q$-rational 2-torsion is isomorphic to $\mathbb{Z}_2$), if h has degree 3 then there are three points of order 2 and $\#E(\mathbb{F}_q)$ is a multiple of 4 (and the $\mathbb{F}_q$-rational 2-torsion is isomorphic to $\mathbb{Z}_2 \oplus \mathbb{Z}_2$).

Case ℓ odd, $\ell \neq p$ Recall that for odd ℓ the division polynomial ψ_ℓ is a polynomial in x after inserting a and b and reducing modulo $-y^2 + x^3 + ax + b$. Further, they characterize the (finite) ℓ-torsion points:

$$(x,y) \in E[\ell] \iff \psi_\ell(x) = 0.$$

Now, we consider the characteristic polynomial of the Frobenius applied to a (symbolic) ℓ-torsion point (x,y):

$$\varphi_q^2(x,y) + q \cdot (x,y) = t \cdot \varphi_q(x,y).$$

We first compute the left hand side and $\varphi_q(x,y)$. For the left hand side, note that since (x,y) is an ℓ-torsion point we can reduce q modulo ℓ: $q_\ell := q \operatorname{rem} \ell$. So compute $\varphi_q^2(x,y) = (x^{q^2}, y^{q^2})$ reduced modulo ψ_ℓ. Next, use the division polynomials to compute $q_\ell \cdot (x,y)$ reduced modulo ψ_ℓ. Then we have to add these two. This leads to three different cases:

1. The x-coordinates differ and we use the generic addition formula.

2. We have $\varphi_q^2(x,y) = q \cdot (x,y)$, and thus we have to use the doubling formula.
3. We have $\varphi_q^2(x,y) = -q \cdot (x,y)$, and so the left hand side turns out to be $\mathcal{O}$.

You may want to defer the computation of the y-coordinate until it is really needed.

Case 1: The left hand side is non-zero and we can already say that $t \not\equiv_\ell 0$. Denote the left hand side with (x', y') and let $(x_j, y_j) := j \cdot (x,y)$ with x_j, y_j polynomials reduced modulo ψ_ℓ and the curve equation, computed from the division polynomials. The x-coordinate of the left hand side then is [—to do—]

2.15. Parametrizations do not exist. The following result will be needed later to prove that the Picard group is isomorphic to the curve. But also in his own right this is an interesting result. Whenever we describe algebraic objects we have two options: first, we can always write down some equations, well, not explicit in every case but at least in principle, and, second, we can parametrize all or most points in question. For an elliptic curve, the second option is none, at least when we restrict ourselves to quolynomial functions.

LEMMA 2.40. *Consider an elliptic curve $E\colon y^2 = x^3 + ax + b$ over a field k (of characteristic different from 2). Assume $X, Y \in \overline{k}(t)$ are quolynomial functions in t such that $Y^2 = X^3 + aX + b$. Then X and Y are both constant.*

PROOF. Write $x^3 + ax + b = (x - e_0)(x - e_1)(x - e_2)$ with $e_0, e_1, e_2 \in \overline{k}$. Since the curve is smooth, the e_i are distinct. Write

$$X = \frac{X_n}{X_d}, \qquad Y = \frac{Y_n}{Y_d},$$

with $X_n, X_d, Y_n, Y_d \in \overline{k}[t]$ are polynomials, X_n, X_d coprime and Y_n, Y_d coprime. The relation for X, Y yields the following equation of polynomials in $\overline{k}[t]$:

$$Y_n^2 X_d^3 = Y_d^2 \left(X_n^3 + a X_n X_d^2 + b X_d^3 \right).$$

We will conclude from this equation that the four polynomials X_d, $X_n - e_0 X_d$, $X_n - e_1 X_d$ and $X_n - e_2 X_d$ are squares in $\overline{k}[t]$.

Since Y_n and Y_d are coprime X_d^3 must be a multiple of Y_d^2. Moreover, X_d and $X_n^3 + aX_nX_d^2 + bX_d^3$ are also coprime, since X_d and X_n are corpime. Thus also Y_d^2 must be a multiple of X_d^3. Rescaling X_n, X_d, Y_n, Y_d we may assume that

$$\begin{aligned} X_d^3 &= Y_d^2, \\ Y_n^2 &= X_n^3 + a X_n X_d + b X_d^3 \\ &= (X_n - e_0 X_d)(X_n - e_1 X_d)(X_n - e_2 X_d). \end{aligned}$$

Thus X_d must be a square. Moreover, the factors $X_n - e_i X_d$ are pairwise coprime: Otherwise, if $X_n - e_i X_d$ and $X_n - e_j X_d$ have a common root, then also $e_j (X_n - e_i X_d) - e_i (X_n - e_j X_d) = (e_j - e_i) X_n$ and $(X_n - e_i X_d) - (X_n - e_j X_d) = (e_j - e_i) X_d$ have this root contradicting the assumption that X_n and X_d are coprime. Consequently, each factor $X_n - e_i X_d$ must be a square.

That is too much unless X_n and X_d are constant as Lemma 2.41 shows. Well, now X is constant and by the original equation then also Y is constant and we are done. □

LEMMA 2.41. *Assume P_1 and P_2 are coprime polynomials in $\overline{k}[t]$, and there are four pairs $(a_i, b_i) \in \overline{k}^2$ such that*

- *any two pairs (a_i, b_i) are linearly independent in $\overline{k}^2$,*
- *each polynomial $a_i P_1 + b_i P_2$ is a square.*

Then P and Q are constant polynomials.

PROOF. Assume to the contrary that P_1, P_2 are a minimal counterexample with respect to $\max(\deg P_1, \deg P_2)$. In particular, $\max(\deg P_1, \deg P_2) > 0$ and there are no polynomials of smaller maximal degree such that the four linear combinations are squares. Next, the assumption that P_1, P_2 are coprime implies that they are $\overline{k}$-linearly independent. Further, by assumption we can write

$$a_i P_1 + b_i P_2 = R_i^2 \tag{2.42}$$

for some $R_i \in \overline{k}[t]$. Then R_i cannot have a common root with R_j for $i \neq j$ since otherwise P_1 and P_2 would have a common root. Also, R_i cannot be a multiple of R_j since otherwise P_1 and P_2 would be linearly dependent.

By assumption (a_1, b_1) and (a_2, b_2) are a basis of $\overline{k}^2$ and so we can write (a_3, b_3) and (a_4, b_4) as linear combinations of them. Combining with (2.42) we find constants $c_1, c_2, d_1, d_2 \in \overline{k}$ such that

$$\begin{aligned} R_3^2 &= c_1^2 R_1^2 - d_1^2 R_2^2 = (c_1 R_1 - d_1 R_2)(c_1 R_1 + d_1 R_2), \\ R_4^2 &= c_2^2 R_1^2 - d_2^2 R_2^2 = (c_2 R_1 - d_2 R_2)(c_2 R_1 + d_2 R_2). \end{aligned}$$

Now, it suffices to prove that

1. for $i \in \{1, 2\}$ the factors $c_i R_1 \pm d_i R_2$ occuring on the right hand sides are coprime and
2. each two of the four vectors $(c_i, \pm d_i)$ are linearly independent.

Because then each of the four must be a square. Moreover, $2\max(\deg R_1, \deg R_2)$ is at most $\max(\deg P_1, \deg P_2)$. Either this contradicts our assumption that P_1, P_2 is a *minimal* counterexample or R_1, R_2 are both constant. In the latter case, also P_1 and P_2 are constant since (a_1, b_1) and (a_2, b_2) are linearly independent. But this contradicts our assumption that $\max(\deg P_1, \deg P_2) > 0$. Thus this assumption must be wrong and there are no nonconstant polynomials fulfulling the conditions of the lemma.

So it remains to prove the two claims. Ad 1: If $c_i R_1 \pm d_i R_2$ had a common root then also R_1 and R_2, both expressable as linear combinations of the factors, would have a common root, but that is not the case as we already observed. Ad 2: If (c_1^2, d_1^2) is a constant multiple of (c_2^2, d_2^2) then R_3 would be a constant multiple of R_4. This is not the case and so we infer that $(c_1, \pm d_1)$ and $(c_2, \pm d_2)$ are linearly independent (which, in this low-dimensional case, means that neither is a multiple of the other). Also, both c_1 and d_1 must be non-zero since otherwise R_3 would be a multiple of R_1 or R_2. This shows that (c_1, d_1) and $(c_1, -d_1)$ are linearly independent.

This proves all claims. □

2.16. Orders, divisors and pairings. So far we have only briefly considered functions on an elliptic curve. As with morphism the first question is: which functions do we want to consider? It is evident that we want to stick to functions described by polynomials or quolynomials. However, as before we have to be a bit careful. A function f on an elliptic curve E shall be an algebraic morphism on an open subset. Concretely: any quolynomial on $\mathbb{P}^2(k)$ that is defined on an open subset of E induces a function f on E.

If you consider a quolynomial f on the line $\mathbb{C}$ then we can ask for its zeroes and poles. For example, $f = \frac{x(x-3)^3}{(x-2)^2}$ has a single zero at 0, a triple zero at 3 and a double pole at 2. So far, there is nothing miraculous. Now turn the question upside down: say, I need a function g with a five-tuple zero at 5 and a double pole at 2. Does there exist one? Is it unique? The answers are easy: yes, there is one: $g = \frac{17(x-5)^5}{(x-2)^2}$. Of course, instead of 17 you could put any other non-zero number. In other words: it is not unique. However, up to that scalar it is unique, any other solution $\widehat{g}$ is merely a non-zero multiple of g.

So it seems we understood this. But what about other curves? The simplest next object is the projective line $\mathbb{P}^1\mathbb{C}$. Again, any quolynomial defines a function, and for example the above f has a single zero at 1, a triple zero at 3 and a double pole at 2, as before, and additionally a double pole at ∞. Wait— what's there at infinity? The problem is that we do not really see how the function looks there. Only if we change the view point, we can see that: replacing $x = 1/y$ we find $f = \frac{1(1-3y)^3}{(1-2y)^2y^2}$. This new expression tells the same story at all points given by $x \neq 0$. Well, $y = 1/x$ is just another name for the same location. With the x-description we can see what happens at 0 and any location in $\mathbb{C} \setminus \{0\}$, with the y-description we can see what happens at ∞ and any location in $\mathbb{C} \setminus \{0\}$. Now, $y = 0$ corresponds to the point at infinity. And there f has a double pole. Next issue: can we find a function g on $\mathbb{P}^1\mathbb{C}$ with a five-tuple zero at 5 and a double pole at 2. Well, already over $\mathbb{C}$ we had only one choice up to scalar: $g = \frac{17(x-5)^5}{(x-2)^2}$. So how does that behave at infinity? With the same substitution as before we get $g = \frac{17(1-4y)^4}{(1-2y)^2y^2}$. This has a double pole at infinity, regardless of the scalar. But we did not want another pole, so there is no solution. Actually, it is easy to see that whenever we consider a quolynomial h given as a quotient $h = f/g$ of polynomials f, g, then h has zeroes where f vanishes and poles where g vanishes. And there is a $(\deg g - \deg f)$-fold zero or a $(\deg f - \deg g)$-fold pole at infinity: substituting gives

$$h = \frac{y^{\deg f} f|_{x=1/y}}{y^{\deg g} g|_{x=1/y}} y^{\deg g - \deg f}$$

where we spend enough factors to make numerator and denominator polynomials again. Well, as you certainly know f has $\deg f$ many zeros provided you count them with multiplicities. Thus h has $\deg f$ zeros and $\deg g$ poles different from ∞ and ∞ is a zero with multiplicity $\deg g - \deg f$. It seems that we should consider a pole as a zero with negative multiplicity. If we do so then the multiplicities now add up to zero: $\deg f - \deg g + (\deg g - \deg f) = 0$.

Actually, we want to do all that stuff over an elliptic curve. The easy part is which functions we should start with: we simply take all those functions that are given by some quolynomial in two variables x, y which can be evaluated at least one curve point and only evaluate this at the points of the curve. In particular, $\frac{1}{-y^2+x^3+ax+b}$ does *not* define a function. There a further non-immediate problem. Consider the curve $E\colon y^2 = x^3 - x$ over any field k, and the function given as $f = \frac{x}{y}$. We can immediately see that it can be evaluated at all finite curve points (x, y) with $y \neq 0$. We do not see how it behaves at infinity but that

happened already over $\mathbb{P}^1\mathbb{C}$. But $(0,0)$ is a point of the curve. And here numerator and denominator of f vanish. So what? Well, rewriting the curve equation gives

$$f = \frac{x}{y} = \frac{y}{x^2 - 1}.$$

Recall that we are only interested in values *on* the curve. Now, it is obvious that f has a zero at $(0,0)$ since the numerator of the new description vanishes and the denominator does not. This problem, namely that a quolynomials numerator and denominator vanish simultaneously, can always be repaired in such a way. Let (x,y) be a curve point and f a quolynomial. Then one can prove that numerator or denominator do not vanish at (x,y), or we can rewrite f using the curve equation such that this is the case. We can thus regard a function as having values in $\overline{k} \cup \{\infty\}$. To evaluate f simply plug in the point coordinates into numerator and denominator. If the denominator does not vanish, just divide. Otherwise, if the numerator does not vanish call the value ∞. In the remaining case, namely both denominator and numerator vanish, rewrite the function…

Let's ease our life by introducing some concepts.

Definition 2.43. *Let E be an elliptic curve. The* divisor group $\operatorname{Div}(E)$ *is the free abelian group generated by symbols $[P]$ for the points P of E, and its elements are called* divisors. *In other words, a divisor is a finite linear combination*

$$D = \sum_j a_j [P_j]$$

with $a_j \in \mathbb{Z}$. Further we define the degree of a divisor D *by* $\deg D = \sum_j a_j \in \mathbb{Z}$, *and the* sum of a divisor D *is* $\operatorname{sum} D = \sum_j a_j P_j \in E$. *The degree-0 divisors form a subgroup* $\operatorname{Div}^0(E)$ *of the divisor group.*

Noting that $\operatorname{sum}([P] - [\mathcal{O}]) = P$ we obtain that $\operatorname{sum}\colon \operatorname{Div}^0(E) \to E$ is a surjective homomorphism. The kernel of this map is also interesting as you will see soon.

Consider a function f on E. The function f has a *zero* at a point P of the curve if $f(P) = 0$, and a *pole* if $f(P) = \infty$. To get more detailed information we have to describe f like we do with polynomials: a polynomial g has a k-fold zero at a if $g = (x-a)^k h$ for some polynomial h which does not vanish. This concept can be carried over to functions on a curve. However, it is not clear how to replace $x - a$ as that is a specific polynomial depending on a somehow. Actually, we could also have taken $3(x-a)$ instead… It is thus not so clear how to specify the role of $x - a$ in a form that we can generalize to curves. Well, the following fact solves that point:

Fact 2.44. *Let E be an elliptic curve, and $P \in E$ some point. Then there exists a* uniformizer u_P at P *such that for every function f on E there exists an integer $r \in \mathbb{Z}$ and a function g such that*

$$f = (u_P)^r g, \qquad \text{and} \qquad g(P) \neq 0, \infty.$$

Moreover, the integer r is uniquely defined by this requirement (regardless of the used uniformizer). We define $\operatorname{ord}_P f := r$.

Notice that this fits to the following wanted property: if $f = f_1 f_2$ and both f_1 and f_2 vanish at P, then we should find an at least double zero for f at P. Following that thought a uniformizer at P is a function that has a zero at P but cannot be written as a product of two functions having a zero at P, so it is indecomposable in this sense. In other words, the order r is just the maximal number of factors vanishing at P that can occur in a factorization of f that has only factors defined at P. [If f has a pole at P then r will be negative.] Also we observe that the order is multiplicative:

LEMMA 2.45. *Given functions f_1, f_2 on a curve E and a point $P \in E$, we have* $\operatorname{ord}_P(f_1 f_2) = \operatorname{ord}_P f_1 \cdot \operatorname{ord}_P f_2$. □

Let's consider an example: on the curve $E\colon y^2 = x^3 - x$ we take the function $f = x$ and the point $P = (0, 0)$. Clearly, f has a zero at P. However, the line $x = 0$ is a tangent at the curve which we may take as a hint for a double zero. And indeed the curve equation implies

$$f = x = y^2 \frac{1}{x^2 - 1}.$$

As y also vanishes at P, f has a at least double zero. Since actually y is a unifomizer at P the function f actually has exactly a double zero at P. Consequently, $\operatorname{ord}_P x = 2$ and $\operatorname{ord}_P \frac{x}{y} = 1$.

CLAIM 2.46. *At any finite point P of an elliptic curve, the uniformizer u_P can be taken from the equation of a line that passes through P but is not tangent to E.*

Consequently, we can take the vertical line $u_P = x - x_0$ for any finite point $P = (x_0, y_0)$ with $y_0 \neq 0$, and the horizontal line $u_P = y$ if $y_0 = 0$. For the point at infinity, we may take $u_{\mathcal{O}} = \frac{x}{y}$: In projective coordinates $y^2 z = x^3 + ax^2 z + bz^3$ and $u_{\mathcal{O}} = \frac{x}{y}$. Specializing to the chart given by $y = 1$ we find $z = x^3 + ax^2 z + bz^3$. The tangent at $\mathcal{O}$ is now given by $z = 0$. However, the given uniformizer $u_P = x$ is the function from the line $x = 0$ which is *not* tangent at $\mathcal{O}$.

DEFINITION 2.47. *For a non-zero function f on an elliptic curve E we define the* divisor of f *by*

$$\operatorname{div}(f) := \sum_{P \in E} \operatorname{ord}_P(f) \cdot [P] \in \operatorname{Div}(E).$$

Any divisor of this form is called a principal divisor.

Note that the given sum is always a finite sum:

PROPOSITION 2.48. *Let E be an elliptic curve and f a non-zero function on E. Then*

(i) f has only finitely many zeros and poles.

(ii) We have $\deg(\operatorname{div}(f)) = 0$.

(iii) If f has no zeros or poles, ie. $\operatorname{div}(f) = 0$, *then f is a constant.*

Actually, this proposition holds for any smooth, irreducible, projective curve. In particular, it holds for the curve $\mathbb{P}^1\mathbb{C}$ that we discussed above. However, the affine line $\mathbb{C}$ is not projective and (ii) does not hold. And on a non-irreducible curve (iii) may be wrong: on $(y-x)(y+x) = 0$ the function $f = \frac{x}{y}$ has divisor 0. [Clearly, x has a single zero at the only critical point $P = (0,0)$ and also y has a single zero there. By Lemma 2.45 the degree of f at P must be zero and so its divisor vanishes.]

Consider the simplest possible function: a line $f = ax + by + c$. [By abuse of language we also call the function 'line', though strictly speaking the line is given by the solutions of $f = 0$.] Say, it passes through the points $P_1, P_2, P_3 \in E$. If $b \neq 0$ then the line does not pass through $\mathcal{O}$ and f has a triple pole there. We obtain

$$\operatorname{div}(ax + by + c) = [P_1] + [P_2] + [P_3] - 3[\mathcal{O}].$$

If $b = 0$ then the line passes through, say, $P_3 = (x_3, y_3)$, $-P_3 = (x_3, -y_3)$ and $\mathcal{O}$ and we find

$$\operatorname{div}(x - x_3) = [P_3] + [-P_3] - 2[\mathcal{O}].$$

Consequently, rewriting $P_3 = P_1 + P_2$,

$$\operatorname{div}\left(\frac{ax + by + c}{x - x_3}\right) = [P_1] + [P_2] - [P_1 + P_2] - [\mathcal{O}], \tag{2.49}$$

or

$$[P_1] + [P_2] = [P_1 + P_2] + [\mathcal{O}] + \operatorname{div}\left(\frac{ax + by + c}{x - x_3}\right).$$

This is related to the question which divisors are principal, ie. are divisors of a function. Since we can choose the line through any two given points $P_1, P_2 \in E$ we can replace a divisor $[P_1] + [P_2]$ with $[P_1 + P_2] + [\mathcal{O}]$ plus the divisor of some function g.

Theorem 2.50. *Consider an elliptic curve E and a divisor D. Then*

$$\exists f\colon D = \operatorname{div}(f)$$

iff

$$\operatorname{sum}(D) = \mathcal{O} \qquad \textit{and} \qquad \deg(D) = 0.$$

Before we enter the proof let us interpret what we obtain. Namely, we now get another proof of the associativity. Note that the sum operator maps divisors to points of the curve E, and $\operatorname{sum}(D_1 + D_2) = \operatorname{sum} D_1 + \operatorname{sum} D_2$. The theorem tells us important information about its kernel and even more. Let $\operatorname{Div}^0(E)$ be the group of divisors of degree zero and $\operatorname{Princ}(E)$ be the group of divisors given by functions.

Corollary 2.51. *The map*

$$\operatorname{sum}\colon \operatorname{Div}^0(E)/\operatorname{Princ}(E) \longrightarrow E(\overline{k})$$

is an isomorphism of groups.

Proof. The map $\operatorname{sum}\colon \operatorname{Div}^0(E \to E(\overline{k})$ is surjective since $\operatorname{sum}([P] - [\mathcal{O}]) = P$. Its kernel are precisely the principal divisors as Theorem 2.50 tells us. □

The surprise here actually is that this proof is valid even if we do not know that the operation of $E(\overline{k})$ is associative. In other words, [—to do—].

Proof (Theorem 2.50). We start with a few preliminary observations. First note that the above for any two given points $P_1, P_2 \in E$ yields a function g with

$$\operatorname{div}(g) = [P_1] + [P_2] - [P_1 + P_2] - [\mathcal{O}].$$

This allows us to replace $[P_1]+[P_2]$ in a divisor by $[P_1+P_2]+[\mathcal{O}]+\operatorname{div}(g)$. Used iteratively we can shrink to almost nothing. Before executing this note that the sum of g's divisor actually is $\mathcal{O}$: $\operatorname{sum}(\operatorname{div}(g)) = P_1 + P_2 - (P_1 + P_2) - \mathcal{O} = \mathcal{O}$. Now let's do the announce induction. Assume D is a divisor, say $D = \sum a_j[P_j]$. Replace iteratively the sum of all points with positive sign with $\left[\sum_{a_j>0} a_j P_j\right] - z_1[\mathcal{O}] + \operatorname{div}(h_1)$ and the sum of all points with negative sign with $-\left[\sum_{a_j<0} -a_j P_j\right] + z_2[\mathcal{O}] - \operatorname{div}(h_2)$. Therein h_1 and h_2 are products of functions g, and so clearly fulfill $\operatorname{sum}(h_i) = \mathcal{O}$. Combining this we find $D = [P] - [Q] + z[\mathcal{O}] + \operatorname{div}(h)$ where $z = -z_1 + z_2$, $h = \frac{h_1}{h_2}$. Since we know by Proposition 2.48(ii) that $\deg \operatorname{dev} h = 0$ we find that $\deg D = 1 - 1 + z + 0 = z$. Further, by construction $\operatorname{sum}\operatorname{div} h = \operatorname{sum}\operatorname{div} h_1 - \operatorname{sum}\operatorname{div} h_2 = \mathcal{O}$ and thus $\operatorname{sum} D = P - Q$. We arrive at the following: given any divisor D on E there exists a function h (with $\operatorname{sum}\operatorname{div} h = \mathcal{O}$) and points $P, Q \in E$ such that

- $D = [P] - [Q] + z[\mathcal{O}] + \operatorname{div}(h)$,
- $\operatorname{sum} D = P - Q$, $\deg D = z$.

Now, assume that $\operatorname{sum}(D) = \mathcal{O}$ and $\deg(D) = 0$. Then we find that $P = Q$ and $z = 0$, and consequently $D = \operatorname{div}(h)$ for some function h.

Conversely, assume $D = \operatorname{div}(f)$. Then again by Proposition 2.48(ii) that $\deg D = 0$. And by the previous we find that $D = [P] - [Q] + \operatorname{div}(h)$ with $\operatorname{sum}\operatorname{div} h = \mathcal{O}$ and consequently $[P] - [Q] = \operatorname{div}(fh^{-1})$. The following Lemma 2.53 proves that then $P = Q$ and so $\operatorname{sum} D = P - Q = \mathcal{O}$. □

This proof actually gives a way for finding a function with a given divisor.

Exercise 2.52. *Consider the elliptic curve $E\colon y^2 = x^3 + 4x$ over $\mathbb{F}_{11}$. Find a function with divisor*

$$D = [(0,0)] + [(2,4)] + [(4,5)] + [(6,3)] - 4[\mathcal{O}].$$

Lemma 2.53. *Let $P, Q \in E(\overline{k})$ and there is a function h on E with $[P] - [Q] = \operatorname{div}(h)$. Then*

$$P = Q.$$

The proof of this lemma either requires the Theorem of Riemann-Roch or a trick plus the parametrization impossibility from Lemma 2.40.

Proof. Suppose $P \neq Q$ and $[P] - [Q] = \operatorname{div}(h)$. This means that h has a pole at Q and so also $h - c$ also has a pole at Q for any constant $c \in \overline{k}$. By Proposition 2.48(ii) the function $h - c$ must have exactly one single zero and no other poles or zeros. We now show that every

function on E is a rational function of h. This will turn out to be impossible and so proof that our assumptions cannot hold, ie. the lemma is proved.

So take any function f on E. First, assume that f has neither a zero nor a pole at Q. Consider

$$g := \prod_{R \in E(\overline{k})} (h - h(R))^{\operatorname{ord}_R(f)}.$$

By the above remark $\operatorname{div}(h - h(R)) = [R] - [Q]$ and so $\operatorname{div}(g) = \sum_{R \in E(\overline{k})} \operatorname{ord}_R(f)\,([R] - [Q]) = \operatorname{div}(f) - \deg(f) \cdot [Q] = \operatorname{div}(f)$. Thus the divisor of g/f is zero and by Proposition 2.48(iii) $f = c \cdot g$. Observe that g is a rational expression of h, f is a rational function of h.

In the general case $\hat{f} := h^{\operatorname{ord}_Q f} \cdot f$ has order 0 at Q and so the previous case applies and $\hat{f}$ is a rational function of h. Since $f = h^{-\operatorname{ord}_Q f} \hat{f}$, this proves that f is a rational function of h.

In particular, x and y are rational functions of h. In other words: there is a way to parametrize a part of the elliptic curve with rational functions. But that is impossible as Lemma 2.40 shows. □

2.17. Pairings.

2.17.1. The Weil pairing. Our next goal is the construction of the Weil pairing including a complete proof for its properties as listed in Theorem 2.27. We consider an elliptic curve E over a field k and an integer n coprime to the characteristic of k. We want to construct a pairing

$$e_n \colon E[n] \times E[n] \longrightarrow \mu_n.$$

Actually, there are two equivalent constructions, both of which will be of importance to us.

2.17.2. Classical construction. We want to construct $e_n(S, T)$ for two n-torsion points $S, T \in E[n]$. First, we find a function g_T on E with

$$\operatorname{div}(g_T) = \widetilde{D} := \sum_{nT''=T} [T''] - \sum_{nR=\mathcal{O}} [R]. \tag{2.54}$$

As we can verify that $\deg(\widetilde{D}) = 0$ and $\operatorname{sum}(\widetilde{D}) = \mathcal{O}$ such a function does exists by Theorem 2.50. To that end note that $S_{\widetilde{D}} := \{T'' \in E \,|\, nT'' = T\} = \{T' + R \,|\, R \in E[n]\}$ if T' is a point with $nT' = T$. [If $nT'' = T$ then $n(T'' - T') = \mathcal{O}$ and $T'' = T' + R$ with $R = T'' - T' \in E[n]$. Conversely, $n(T' + R) = nT' + \mathcal{O} = T$.] Using $\#E[n] = n^2$ we obtain $\operatorname{sum}(\widetilde{D}) = \sum_{R \in E[n]} (T' + R - R) = n^2 T' = nT = \mathcal{O}$.

Similarly, we can find a function f_T on E such that

$$\operatorname{div}(f_T) = nD, \qquad D = [T] - [\mathcal{O}]. \tag{2.55}$$

Noting that $\deg(nD) = 0$ and $\operatorname{sum}(nD) = nT = \mathcal{O}$ Theorem 2.50 implies the existence of f_T. The reason for considering f_T is that this yields a second description of g_T. The function $f_T \circ [n]$ has an n-fold zero at every point T'' that maps to T under the scalar multiplication $[n]$ and an n-fold pole at every point R that maps to $\mathcal{O}$ under $[n]$. So we find that

$$\begin{aligned} \operatorname{div}(f_T \circ [n]) &= \sum_{nT''=T} n[T''] - \sum_{nR=\mathcal{O}} n[R] \\ &= \operatorname{div}(g_T^n). \end{aligned}$$

By Proposition 2.48(iii) this implies that $f_T \circ [n]$ and g_T^n are scalar multiples of each other. Rescaling f_T we may assume that

$$f_T \circ [n] = g_T^n.$$

Now take any point $P \in E(\overline{k})$ where g_T does not vanish. Then

$$g_T(S+P)^n = f_T(n(S+P)) = f_T(nP) = g_T(P)^n.$$

This implies that

$$\frac{g_T(S+P)}{g_T(P)} \in \mu_n.$$

Thus the (apart at its poles) continuous function $P \mapsto \frac{g_T(S+P)}{g_T(P)}$ landing in a finite set must be constant since E is connected. This proves that the definition

$$e_n(S,T) = \frac{g_T(S+P)}{g_T(P)} \tag{2.56}$$

is independent of P. (We only need that $nP \neq T$ and $nP \neq \mathcal{O}$ to ensure that $g_T(P)$ is neither zero nor a pole.) Further, since g_T is determined up to a scalar that cancels, this definition is also independent of the choice of g_T. Now we can prove Theorem 2.27:

THEOREM 2.57 (Weil pairing). *Let E be an elliptic curve defined over a field k and let n be a positive integer coprime to the characteristic of k. Then the* Weil pairing

$$e_n \colon E[n] \times E[n] \longrightarrow \mu_n$$

satisfies the following properties.

(i) e_n is bilinear, that is, for all $S, S_1, S_2, T, T_1, T_2 \in E[n]$

$$e_n(S_1+S_2,T) = e_n(S_1,T) \cdot e_n(S_2,T),$$
$$e_n(S,T_1+T_2) = e_n(S,T_1) \cdot e_n(S,T_2).$$

(ii) e_n is non-degenerate, that is, for all $T \in E[n]$

$$\forall S \in E[n] \colon e_n(S,T) = 1 \implies T = \mathcal{O},$$
$$\forall S \in E[n] \colon e_n(T,S) = 1 \implies T = \mathcal{O}.$$

(iii) e_n is antisymmetric, that is, for all T

$$e_n(T,T) = 1.$$

In particular, $e_n(T,S) = e_n(S,T)^{-1}$.

(iv) e_n is compatible with the Galois actions, that is, for every automorphism σ of $\overline{k}$ fixing k (in particular, for a curve in Weierstraß form this means that $\sigma(a) = a$ and $\sigma(b) = b$) we have

$$e_n(\sigma S, \sigma T) = \sigma(e_n(S,T)).$$

(v) For every endomorphism α of E we have

$$e_n(\alpha(S), \alpha(T)) = e_n(S,T)^{\deg \alpha}.$$

As we observed earlier it is enough for our purposes when we only prove (v) in case α is separable.

Proof. (i): Linearity with respect to S is easy:

$$\begin{aligned} e_n(S_1,T)e_n(S_2,T) &= \frac{g_T(S_1+P)}{g_T(P)} \cdot \frac{g_T(S_2+S_1+P)}{g_T(S_1+P)} \\ &= \frac{g_T(S_1+S_2+P)}{g_T(P)} = e_n(S_1+S_2,T) \end{aligned}$$

provided P is chosen such that no zero or pole of g_T is hit. Linearity with respect to T is a bit more involved. Abbreviate $T_3 = T_1 + T_2$ and recall the functions f_{T_i}, g_{T_i} defined above. We constructed a function h such that $\operatorname{div}(h) = [T_1] + [T_2] - [T_3] - [\mathcal{O}]$. By the choice of f_{T_i} we obtain

$$\begin{aligned} \operatorname{div}\left(\frac{f_{T_1} f_{T_2}}{f_{T_3}}\right) &= \operatorname{div}(f_{T_1}) + \operatorname{div}(f_{T_2}) - \operatorname{div}(f_{T_3}) \\ &= n[T_1] + n[T_2] - n[T_3] - n[\mathcal{O}] \\ &= n \operatorname{div}(h) = \operatorname{div}(h^n). \end{aligned}$$

By Proposition 2.48(iii) there is a constant $c \in \overline{k}^\times$ with

$$f_{T_1} f_{T_2} = c h^n f_{T_3}.$$

Since $g_{T_i}^n = f_{T_i} \circ [n]$ we obtain

$$g_{T_1} g_{T_2} = d(h \circ [n]) g_{T_3}$$

for some $d \in \overline{k}^\times$ with $d^n = c$. Observe that $h(n(S+P)) = h(nP)$ since $nS = \mathcal{O}$. Now,

$$\begin{aligned} e_n(S,T_1)e_n(S,T_2) &= \frac{g_{T_1}(S+P)}{g_{T_2}(P)} \cdot \frac{g_{T_2}(S+P)}{g_{T_2}(P)} \\ &= \frac{g_{T_3}(S+P)}{g_{T_3}(P)} \cdot \frac{h(n(S+P))}{h(nP)} \\ &= e_n(S,T_3). \end{aligned}$$

This completes the linearity.

(ii): The complex picture suggests the here-needed following

Claim. *Assume g is a function on E with $g(P+S) = g(P)$ for any $P \in E$ and $S \in E[n]$, in other words, g is $E[n]$-periodic. Then there is a function h such that $g = h \circ [n]$.*

To prove non-degeneracy in T assume $e_n(S,T) = 0$ for all S. By construction we have $g_T(P+S) = g_T(P)$ for all $P \in E$ and $S \in E[n]$. Thus by the previous we find a function h such

that $g_T = h \circ [n]$. From the construction, we also have $f_T \circ [n] = g_T^n$. Now $f_T \circ [n] = (h \circ [n])^n$ and since $[n]$ is surjective this implies $f_T = h^n$. Now consider the corresponding divisor:

$$n \operatorname{div}(h) = \operatorname{div} h^n = \operatorname{div} f = n([T] - [\mathcal{O}]).$$

Thus the divisor of h must be $[T] - [\mathcal{O}]$. But that implies $T = \mathcal{O}$ by Lemma 2.53.

Non-degeneracy in S follows with (iii). [Watch that proof carefully concerning usage of this part.]

Proof of the claim: [—to do—]
(iii): [—to do—]
(iv): [—to do—]
(v): [—to do—] □

2.17.3. Tate pairing. [—to do—]

2.17.4. Symmetrical construction. There is a much more symmetrical way to construct the Weil pairing. This also leads to a much better handable algorithmic description.

THEOREM 2.58. *Let $S, T \in E[n]$ and $D_S, D_T \in \operatorname{Div}^0(E)$ with disjoint support such that*

$$\operatorname{sum}(D_S) = S, \qquad \operatorname{sum}(D_T) = T.$$

Further let f_S and f_T be functions on E such that

$$\operatorname{div}(f_S) = nD_S, \qquad \operatorname{div}(f_T) = nD_T.$$

Then the Weil pairing is given by

$$e_n(S,T) = \frac{f_T(D_S)}{f_S(D_T)},$$

where evaluation of a function f at a divisor $\sum a_P[P]$ is defined by $f(\sum a_P[P]) = \prod f(P)^{a_P}$.

If we choose $D_S = [S] - [\mathcal{O}]$ and $D_T = [T+R] - [R]$ for some arbitrary point R such that the divisors have disjoint support, then Theorem 2.58 tells us that

$$e_n(S,T) = \frac{f_S(R) f_T(S)}{f_S(T+R) f_T(\mathcal{O})}.$$

Now, to compute that value we only need to compute $\frac{f_T(S)}{f_T(\mathcal{O})}$ and $\frac{f_S(R)}{f_S(T+R)}$ where $\operatorname{div} f_S = n[S] - n[\mathcal{O}]$ and $\operatorname{div} f_T = n[T+R] - n[R]$.

Actually, this is motivated by the Tate-Lichtenbaum pairing which is given by

$$\langle S, T \rangle_n = f_S(D_T) \cdot \left(\mathbb{F}_q^\times\right)^n \in \mathbb{F}_q^\times / \left(\mathbb{F}_q^\times\right)^n,$$

or the reduced Tate-Lichtenbaum pairing

$$\tau_n(S,T) = \langle S, T \rangle_n^{\frac{q^k-1}{n}} \in \mu_n$$

for $S \in E(\mathbb{F}_q)[n]$ and $T \in E(\mathbb{F}_q)/nE(\mathbb{F}_q)$. Igoring a few details we have

$$e_n(S,T) = \frac{\langle T, S \rangle_n}{\langle S, T \rangle_n}$$

2.17.5. Miller's algorithm. Thus computing both types of pairing reduces to the following

TASK 2.59. *Let $P, Q \in E$ (possibly subject to additional conditions) and assume* $\operatorname{div} f_P = n[P+R]-n[R]$ *with $R \in E$ such that the divisor of f_P and the divisor $[Q_1]-[Q_2]$. Compute*

$$\frac{f_P(Q_1)}{f_P(Q_2)}.$$

We break this down by successively solving the following, slightly more complicated

TASK(j) 2.60. *Let $P, Q \in E$ (possibly subject to additional conditions) and assume*

$$\operatorname{div} f_j = D_j := j[P+R] - j[R] - [jP] + [\mathcal{O}]$$

with $R \in E$ such that the divisor of f_P and the divisor $D_Q = [Q_1] - [Q_2]$ with sum Q. Compute

$$\frac{f_j(Q_1)}{f_j(Q_2)}.$$

Assuming that Task(j) and Task(k) have been solved we want to derive a solution for task $j+k$. Let $\ell = ax + by + c$ be the line through jP and kP, and let $v = x + d$ be the vertical line trough $(j+k)P$. Then by (2.49) we have

$$\operatorname{div}\left(\frac{ax+by+c}{x+d}\right) = [jP] + [kP] - [(j+k)P] - [\mathcal{O}].$$

By assumption

$$\begin{aligned}\operatorname{div}(f_j) &= j[P+R] - j[R] - [jP] + [\mathcal{O}],\\ \operatorname{div}(f_k) &= k[P+R] - k[R] - [kP] + [\mathcal{O}].\end{aligned}$$

Multiplying the functions we obtain

$$\operatorname{div}\left(f_j f_k \frac{ax+by+c}{x+d}\right) = (j+k)[P+R] - (j+k)[R] - [(j+k)P] + [\mathcal{O}].$$

This is $D_{j+k} = \operatorname{div}(f_{j+k})$ so that $f_{j+k} = \gamma f_j f_k \frac{ax+by+c}{x+d}$ for some constant γ, and

$$\frac{f_{j+k}(Q_1)}{f_{j+k}(Q_2)} = \frac{f_j(Q_1)}{f_j(Q_2)} \cdot \frac{f_k(Q_1)}{f_k(Q_2)} \cdot \frac{\left.\frac{ax+by+c}{x+d}\right|_{(x,y)=Q_1}}{\left.\frac{ax+by+c}{x+d}\right|_{(x,y)=Q_2}} \tag{2.61}$$

now describes the value of f_{j+k} at D_Q. All we need are the values of f_j and f_k at D_Q, the points jP and kP. Performing the addition $jP + kP$ gives the point $(j+k)P$ and the function $\frac{ax+by+c}{x+d}$, evaluating at D_Q and then multiplying with the values of f_j and f_k at D_Q yields the desired value of f_{j+k} at D_Q along with the point $(j+k)P$.

If now $P \in E[n]$ then $nP = \mathcal{O}$. Thus solving Task(n) yields with $\operatorname{div}(f_n) = n[P+R] - n[R] - [\mathcal{O}] + [\mathcal{O}] = \operatorname{div}(f_P)$ the desired value

$$\frac{f_P(Q_1)}{f_P(Q_2)} = \frac{f_n(Q_1)}{f_n(Q_2)}.$$

Notice that Task(0) is trivial: $D_0 = 0$, so $f_0 = 1$. Also Task(1) is easy: $D_0 = [P+R] - [R] - [P] + [\mathcal{O}]$, so $f_1 = \frac{x+d}{ax+by+c}$ where $\ell = ax + by + c$ is the line through P and R and $v = x + d$ is the vertical line through $P + R$. Thus

$$\frac{f_1(Q_1)}{f_1(Q_2)} = \frac{\left.\frac{ax+by+c}{x+d}\right|_{(x,y)=Q_1}}{\left.\frac{ax+by+c}{x+d}\right|_{(x,y)=Q_2}}$$

Miller's algorithm now simply follows an addition chain for nP and performs point addition and point doublings along with multiplying the corresponding values of f_j. If we simply use add and double we obtain

Algorithm 2.62. Miller's algorithm.
Input: Points $P, R, Q_1, Q_2 \in E$, the desired index n.
Output: The value $\frac{f_P(Q_1)}{f_P(Q_2)}$ where $\operatorname{div} f_P = n[P+R] - n[R] - [nP] + [\mathcal{O}]$.

1. Compute $P + R$, the line $\ell = ax + by + c$ through P and R, the vertical line $v = x + d$ through $P + R$ and let $g \leftarrow \frac{\left.\frac{ax+by+c}{x+d}\right|_{(x,y)=Q_1}}{\left.\frac{ax+by+c}{x+d}\right|_{(x,y)=Q_2}}$.
2. Let $f \leftarrow g$, $J \leftarrow P$, $j \leftarrow 1$.
3. Write $n = (n_{r-1}, \ldots, n_1, n_0)$ in base 2.
4. For $i = r - 2$ down to 0 do 5–15
5. Let $\ell = ax + by + c$ be the tangent at J.
6. $S \leftarrow 2J$.
7. Let $v = x + d$ be the vertical line through S.
8. Let $f \leftarrow f^2 \cdot \frac{\ell}{v}\big|_{Q_1} \cdot \frac{v}{\ell}\big|_{Q_2}$.
9. $J \leftarrow S$, $j \leftarrow 2j$.
10. If $n_i = 1$ then
11. Let $\ell = ax + by + c$ be the line through J and P.
12. $S \leftarrow J + P$.
13. Let $v = x + d$ be the vertical line trough S.
14. Let $f \leftarrow f \cdot g \cdot \frac{\ell}{v}\big|_{Q_1} \cdot \frac{v}{\ell}\big|_{Q_2}$.
15. $J \leftarrow S$, $j \leftarrow j + 1$.
16. Return f.

As a consequence computing a pairing is only a constant factor slower than a scalar multiplication by n.

2.17.6. Properties and proofs. [—to do—]

2.18. All so simple using Riemann-Roch. In this section we consider a smooth, irreducible, projective curve C defined over a field k. The curve may be given as the roots of a

polynomial in $\mathbb{P}^2\overline{k}$, or as the set of common zeroes of several polynomials in $\mathbb{P}^r\overline{k}$ (provided this set is one-dimentsional). We assume that C cannot be written as the union of two smaller curves.

We can introduce functions and divisors on C as we did for elliptic curves, and Proposition 2.48 holds again. For two divisors $D_i = \sum a_{i,P}[P]$ we define

$$D_1 \geq D_2 \ :\Longleftrightarrow\ \forall P a_{1,P} \geq a_{2,P}.$$

and consider the space of functions

$$\mathcal{L}(D) := \{f \mid \operatorname{div}(f) + D \geq 0\} \cup \{0\}$$

whose divisor is at least $-D$, and its dimension $\ell(D) := \dim \mathcal{L}(D)$. The requirement $\operatorname{div}(f) + D \geq 0$ allows f to have poles where the coefficient in D is positive and requires zeroes where the coefficient in D is negative. For example, if $D = [P] - 3[\mathcal{O}]$ then the function is allowed to have a triple pole at infinity, it must have a single zero at P and at all other places it may have zeroes but never poles.

Next, we collect a few basic information about these spaces and their dimension.

Proposition 2.63. *Let C be a smooth, irreducible, projective curver over a field k, and let D, D_1, D_2 be divisors on C.*

(i) *If $\deg D < 0$ then $\mathcal{L}(D) = 0$.*

(ii) *If D_1, D_2 only differ by a principal divisor then $\mathcal{L}(D_1)$ is isomorphic to $\mathcal{L}(D_2)$.*

(iii) *$\mathcal{L}(0) = \overline{k}$.*

(iv) *$\ell(D) < \infty$. Moreover, $\ell(D) \leq \ell(D + [P]) \leq \ell(D) + 1$.*

(v) *If $\deg D = 0$ then $\ell(D) = 0$ or $\ell(D) = 1$.*

Proof. (i): Assume $\mathcal{L}(D) \neq 0$. Then there is a non-zero function f with $\operatorname{div}(f) + D \geq 0$, and so $\deg D = \deg \operatorname{div} f + \deg D \geq 0$ due to Proposition 2.48(ii).

(ii): Write $D_1 = D_2 + \operatorname{div}(g)$. Then $\mathcal{L}(D_1) \to \mathcal{L}(D_2),\ f \mapsto fg$ is the claimed isomorphism.

(iii): If $\operatorname{div}(f) \geq 0$ then f has no poles. By Proposition 2.48(ii) the degree of $\operatorname{div}(f)$ is zero, and so f also has no zeroes. Thus Proposition 2.48(iii) proves that f is constant.

(iv): It suffices to prove that $\ell(D + [P]) - \ell(D) \in \{0, 1\}$. Then the claim follows by induction on $\deg D_+$ (where $D = D_+ - D_-$ with $D_+, D_- \geq 0$). Clearly, $\mathcal{L}(D) \subseteq \mathcal{L}(D + [P])$, since this allows functions to have an additional pole at P. If the two spaces are different then there exist $g, h \in \mathcal{L}(D + [P]) \setminus \mathcal{L}(D)$. Then $\operatorname{ord}_P(g) = \operatorname{ord}_P(h) = -\operatorname{ord}_P D - 1 =: r$. [Since g is in $\mathcal{L}(D + [P])$ we have $\operatorname{org}_P g + \operatorname{ord}_P D + 1 \geq 0$. However, g is not in $\mathcal{L}(D)$ and thus $\operatorname{org}_P g + \operatorname{ord}_P D \ngeq 0$.] Let u be a uniformizer at P, and write

$$g = u^r \widehat{g}, \qquad h = u^r \widehat{h}$$

with $c := \widehat{g}(P) \neq 0, \infty$ and $d := \widehat{h}(P) \neq 0, \infty$. Now,

$$dg - ch = u^r (d\widehat{g} - c\widehat{h})$$

where $(d\widehat{g} - c\widehat{h})(P) = 0$. Thus, $dg - ch$ has order greater than r at P and $dg - ch \in \mathcal{L}(D)$. Thus g and h are linearly independent modulo $\mathcal{L}(D)$, and $\ell(D + [P]) - \ell(D) \in \{0, 1\}$.

(v): This follows from (iii) and (i) with the previous. □

Now we are ready to formulate the famous

THEOREM OF RIEMANN-ROCH 2.64. *Let C be a smooth, irrducible, projective curve. Then there exists an integer g, the* genus of C*, and a divisor $\mathcal{K}$, a* canonical divisor*, such that for all divisors D we have*

$$\ell(D) - \ell(\mathcal{K} - D) = \deg(D) - g + 1. \qquad \square$$

The canonical divisor is the divisor of a differential on C, it is unique up to principal divisors. Well, we do not prove this here. Instead we consider a couple of consequences.

COROLLARY 2.65. $\deg(\mathcal{K}) = 2g - 2$.

PROOF. Apply Theorem of Riemann-Roch 2.64 with $D = 0$ and $D = \mathcal{K}$. With the help of Proposition 2.63(iii) we get

$$\begin{aligned} 1 - \ell(\mathcal{K}) &= -g + 1, \\ \ell(\mathcal{K}) - 1 &= \deg(\mathcal{K}) - g + 1. \end{aligned}$$

Adding gives the claim. $\square$

COROLLARY 2.66. *If $\deg(D) > 2g - 2$ then $\ell(D) = \deg(D) - g + 1$.*

PROOF. Obvious with $\deg(\mathcal{K} - D) < 0$, Proposition 2.63(i) and Theorem of Riemann-Roch 2.64. $\square$

COROLLARY 2.67. *Let P, Q be points on C. If $g \geq 1$ and $[P] - [Q] = \operatorname{div}(f)$ for some function f on C then $P = Q$.*

PROOF. Assume $P \neq Q$. The function f^n has a pole of order n at Q. Functions with different pole orders are linearly independent. Thus the set

$$\left\{1, f, f^2, \ldots, f^{2g-1}\right\}$$

spans a subspace of $\mathcal{L}\left((2g-1)[Q]\right)$ of dimension $2g$, thus $\ell\left((2g-1)[Q]\right) \geq 2g$. Applying Corollary 2.66 yields $\ell\left((2g-1)[Q]\right) = (2g-1) - g + 1 = g$. Together we obtain $2g \leq g$, that is $g \leq 0$ contradicting the assumptions. $\square$

References

IAN BLAKE, GADIEL SEROUSSI & NIGEL SMART (1999). *Elliptic Curves in Cryptography.* Number 265 in London Mathematical Society Lecture Note Series. Cambridge University Press. ISBN 0-521-65374-6.

IAN BLAKE, GADIEL SEROUSSI & NIGEL SMART (editors) (2005). *Advances in Elliptic Curves in Cryptography.* Number 317 in London Mathematical Society Lecture Note Series. Cambridge University Press. ISBN 0-521-60415-X.

Ian Connell (1999). Elliptic Curve Handbook. URL `http://www.math.mcgill.ca/connell/`. A complete file is available for download under `www.ucm.es/BUCM/mat/doc8354.pdf`.

Daniel Hankerson, Alfred Menezes & Scott Vanstone (2004). *Guide to Elliptic Curve Cryptography*. Springer-Verlag, New York. ISBN 0-387-95273-X.

Alfred Menezes (1993). *Elliptic curve public key cryptosystems*. Kluwer Academic Publishers, Boston MA.

René J. Schoof (1985). Elliptic Curves over Finite Fields and the Computation of Square Roots mod p. *Mathematics of Computation* **44**(170), 483–494.

Joseph H. Silverman (1986). *The Arithmetic of Elliptic Curves*. Number 106 in Graduate Texts in Mathematics. Springer-Verlag, New York. ISBN 0-387-96203-4, 3-540-96203-4.

Lawrence C. Washington (2003). *Elliptic Curves — Number Theory and Cryptography*. Discrete Mathematics and its Applications. CRC Press, Boca Raton, FL, USA. ISBN 1-58488-365-0.