KERNVAK ALGEBRA

P. Stevenhagen & B. de Smit

Faculteit Wiskunde en Informatica
Universiteit van Amsterdam
1996-97

CONTENTS

VERZOEK: meld (type)fouten aan psh@fwi.uva.nl.

1. Elliptic integrals.

The subject of elliptic curves has its roots in the differential and integral calculus, which was developed in the 17th and 18th century and became the main subject of what is nowadays a 'basic mathematical education'. In calculus, one tries to integrate the *differentials* $f(t)dt$ associated with, say, a real-valued function f on the real line. As is well known, such integrals are related to the area of certain surfaces bounded by the graph of f. Explicit integration of the differential $f(t)dt$, which amounts to finding an anti-derivative F satisfying $dF/dt = f$, can only be performed for a very limited number of 'standard integrals'. These include the integrals of polynomial differentials $t^k dt$ with $k \in \mathbf{Z}_{\geq 0}$, rational differentials as $(t-\alpha)^{-k}$ with $k \in \mathbf{Z}_{>0}$ and a few 'exponential differentials' as $e^t dt$ and $\sin t\, dt$. Over the complex numbers, any rational differential can be written as a sum of elementary differentials.

Exercise 1. Show that every rational function $f \in \mathbf{C}(t)$ can be written as unique $\mathbf{C}$-linear combination of monomials t^k with $k \in \mathbf{Z}_{\geq 0}$ and fractions $(t-\alpha)^{-k}$ with $\alpha \in \mathbf{C}$ and $k \in \mathbf{Z}_{\geq 1}$. Use this representation to write $\int f(t)dt$ as a sum of elementary functions. [Hint: partial fraction expansion.]

Even if one restricts to polynomial or rational functions f, already the problem of computing the *length* of the graph of f, an old problem known as the 'rectification' of plane curves, leads to the non-elementary differential $\sqrt{1+f'(t)^2}dt$. If $R \in \mathbf{C}(x,y)$ is a rational function and $f \in \mathbf{C}[t]$ a polynomial that is not a square, the differential $R(t, \sqrt{f(t)}dt)$ is called *hyperelliptic*. We can and will always suppose that f is *separable*, i.e., it has no multiple roots. If f is of degree 1, one can transform $R(t, \sqrt{f(t)})dt$ into a rational differential by taking $\sqrt{f(t)}$ as a new variable. If f is quadratic, one can apply a linear transformation $t \mapsto at+b$ to reduce to the case $f(t) = 1-t^2$. We will see in a moment that the resulting integrals are closely related to the problem of computing lengths of circular arcs or, what amounts to the same thing, inverting trigonometric functions. If f is of degree 3 or 4 and squarefree, the differential $R(t, \sqrt{f(t)})dt$ is said to be *elliptic*.

Exercise 2. Show that the length of the ellipse with equation $y^2 = c^2(1-x^2)$ in $\mathbf{R}^2$ equals

$$2\int_{-1}^{1}\sqrt{\frac{1+(c^2-1)t^2}{1-t^2}}dt = 2\int_{-\pi/2}^{\pi/2}\sqrt{1+(c^2-1)\sin^2\phi}\,d\phi,$$

and that the differential $\sqrt{\frac{1+(c^2-1)t^2}{1-t^2}}dt$ is elliptic.

Elliptic differentials lead naturally to the study of elliptic functions and elliptic curves. In a similar way, the case of f of higher degree gives rise to hyperelliptic curves. More generally, it has gradually become clear during the 19th century that an algebraic differential $R(t,u)dt$, with R a rational function and t and u satisfying some polynomial relation $P(t,u) = 0$, should be studied as an object living on the plane algebraic curve defined by the equation $P(x,y) = 0$. For hyperelliptic differentials, this is the hyperelliptic curve given by the equation $y^2 = f(x)$.

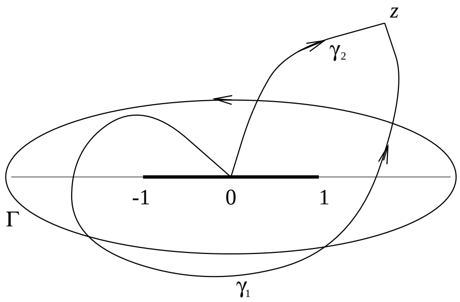
As an instructive example, we consider the differential $\omega = dt/\sqrt{1-t^2}$ related to the arc length of the unit circle. The reader can easily check that the graph of the function $f(t) = \sqrt{1-t^2}$ on the real interval $[-1,1]$ is a semicircle, and that we have $\omega = \sqrt{1+f'(t)^2}dt$. We attempt to define a map

$$\phi : z \longmapsto \int_0^z \omega = \int_0^z \frac{dt}{\sqrt{1-t^2}} \tag{1.1}$$

as a function on $\mathbf{C}$. Note that ω has integrable singularities at the points $t = \pm 1$.

There are two problems with the map ϕ. First of all, there is no canonical definition of a square root $\sqrt{1-t^2}$ for $t \in \mathbf{C}$. One can select a specific square root for $t \in [-1,1]$ or t on the imaginary axis, when $1-t^2$ is real and positive, but such extensions do not yield an obvious choice for, say, $t = \pm 2$. A rather uncanonical way out is the possibility of making a *branch cut*. This means that one defines ϕ not on $\mathbf{C}$, but on a subset of $\mathbf{C}$, such as $\mathbf{C} \setminus [-1,1]$, on which $\sqrt{1-t^2}$ admits a single-valued branch.

If one makes the proposed branch cut and chooses a branch of ω, a second problem arises: two different paths of integration can give rise to different values of $\phi(z)$, so the map ϕ is not well-defined.



The difference between any two values of $\phi(z)$ for the paths γ_1 and γ_2 in the picture is the value of the integral $\oint \omega$ along a simple closed curve Γ around the two singular points $t = \pm 1$ of ω. One can compute this contour integral in various ways.

Exercise 3. Apply the residue theorem to evaluate $\oint_\Gamma \omega$. [Answer: $\pm 2\pi$.]

As the value of the real integral $\int_{-1}^1 \omega$ is the length of a semicircle of radius 1, one easily sees that $\oint \omega$ has value $\pm 2\pi$, with the sign depending on the choice of the square root $\sqrt{1-t^2}$ along the path of integration. From the topology of $\mathbf{C} \setminus [-1,1]$, it is clear that the values of $\phi(z)$ computed along different paths always differ by a multiple of 2π.

There is a canonical reparation of the definition of ϕ that makes ϕ into a well-defined map on a 'natural domain' for ω. Rather than defining ϕ on $\mathbf{C}$ minus some branch cut, one considers the set

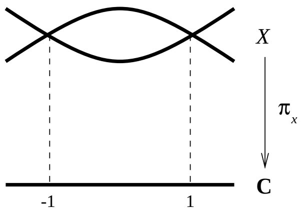
$$X = \{(x,y) \in \mathbf{C}^2 : y^2 = 1 - x^2\}.$$

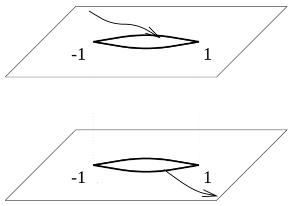
This set comes with a natural projection $\pi_x : X \to \mathbf{C}$ defined by $(x,y) \mapsto x$. Given any point $t \in \mathbf{C}$, the *fiber* $\pi_x^{-1}(t)$ consists of the points $(t,u) \in \mathbf{C}^2$ for which u is a square root of $1-t^2$. For $t \neq \pm 1$, there are exactly 2 such points, and one says that the projection

$\pi_x : X \to \mathbf{C}$ is *generically* 2-to-1. For the *branch points* $t = \pm 1$ there is only one point in the fiber.

As the complex curve X is a subset of $\mathbf{C}^2$, one cannot immediately picture X. There are two approximate solutions. The first consists of drawing $\mathbf{C}$ as a 1-dimensional object and representing π_X as in the picture below. One disadvantage of this method is that the points $(\pm 1, 0)$ on X appear to be of a special nature. The symmetry in x and y in the definition of X show that this cannot be the case.

Exercise 4. Draw the corresponding picture for the map $\pi_y : X \to \mathbf{C}$ sending (x, y) to y. Where are the points $(\pm 1, 0)$ in this picture?





Another, usually somewhat more complicated way to visualize X is to take two copies of $\mathbf{C}$ and 'glue them along a branch cut' as suggested in the picture. In the space obtained, paths passing through the branch cut in one copy of $\mathbf{C}$ emerge on the 'opposite side' of the branch cut in the *other* copy. A moment's reflection shows that, topologically, the resulting surface is homeomorphic to a cylinder. The path Γ becomes the simplest incontractible path on X. It is immediate from the picture that every path $0 \to z$ in $\mathbf{C}$ that does not pass through the branch points ± 1 can uniquely be lifted to a path $x_0 = (0, 1) \to (z, w)$, where w is a square root of $1 - z^2$ that is determined by the path $0 \to z$. The function $t \to \sqrt{1 - t^2}$, which has no natural definition on $\mathbf{C}$, has by construction a natural definition on X: it is the function $(t, u) \to u$. It is now also clear how one should integrate the differential dt/u, which we denote again by ω, along any path in X. We arrive at a definition of ϕ on X rather than $\mathbf{C}$, which is given by

$$\phi(x) = \int_{x_0}^{x} \omega = \int_{x_0}^{x} \frac{dt}{u} \qquad \text{for } x \in X \subset \mathbf{C}^2.$$

The integral is taken along X, and as we have a choice of paths its value is only determined up to multiples of 2π. This means that $\phi(z)$ is well defined as an element of the factor group $\mathbf{C}/2\pi\mathbf{Z}$ of $\mathbf{C}$. The elements of this group can be viewed as the complex numbers in the infinite strip $\{z : -\pi \leq \mathrm{Re}(z) \leq \pi\}$, where for any $r \in \mathbf{R}$, the elements $-\pi + ir$ and $\pi + ir$ on the boundary are identified. Topologically, one notes that just like X, the group $\mathbf{C}/2\pi\mathbf{Z}$ is a cylinder. The following theorem is therefore not so surprising.

1.2. Theorem. *The integration of the differential ω induces a bijection $\phi : X \xrightarrow{\sim} \mathbf{C}/2\pi\mathbf{Z}$.*

We leave it to the reader to give a complete proof of the theorem, as indicated in the exercises, and to show that ϕ is in a natural way a homeomorphism of topological spaces.

Theorem 1.2 has a number of interesting consequences. It shows that the set X, which is the *algebraic curve* in $\mathbf{C}^2$ defined by the equation $x^2+y^2=1$, is in a natural way a group. From the map ϕ, which is defined by means of an integral, it is not immediately clear what the sum of two points on X should be. However, in this case we know from calculus that integration of the real differential $\omega = dt/\sqrt{1-t^2}$ yields the function $\arcsin t$, a somewhat artificially constructed inverse to the sine function. In fact, our carefully constructed map ϕ has an *inverse* which is much easier to handle. From the observation that $\pi_x \circ \phi^{-1}$ is in fact the sine function and the identity $\phi^{-1}(0) = (0,1)$, the following theorem is now immediate.

1.3. Theorem. *The inverse $\phi^{-1} : \mathbf{C}/2\pi\mathbf{Z} \to X$ of the bijection ϕ in* 1.2 *is given by $\phi^{-1}(z) = (\sin z, \cos z)$.*

It follows from 1.3 that we may describe the natural addition on X by the formula $(\sin\alpha, \cos\alpha) + (\sin\beta, \cos\beta) = (\sin(\alpha+\beta), \cos(\alpha+\beta))$. From the addition formulas for the sine and cosine functions one deduces that the group law on X is in fact given by the simple polynomial formula

$$(1.4) \qquad (x_1, y_1) + (x_2, y_2) = (x_1y_2 + x_2y_1, y_1y_2 - x_1x_2).$$

The unit element of X is the point $(0,1)$, and the inverse of $(x,y) \in X$ is the point $(-x,y)$. This shows that X is in fact an *algebraic group*: for every subfield of $K \subset \mathbf{C}$, such as $\mathbf{Q}$ or $\mathbf{Q}(i)$, the set $X(K) \subset K^2$ of K-valued points of X is an abelian group. A picture of the real locus $X(\mathbf{R}) = \{(x,y) \in \mathbf{R}^2 : x^2+y^2=1\}$ explains why X is known as the *circle group*.

Exercise 5. Draw a picture of $X(\mathbf{R})$ and give a geometric description of the group law.

As we have constructed the circle group by analytic means, via the construction of ϕ, it is not immediately obvious that formula 1.4 defines a group structure on $X(K)$ for arbitrary fields K. Clearly, there is no 'analytical parametrization' ϕ^{-1} of X if we replace $\mathbf{C}$ by a field of positive characteristic, such as the finite field $\mathbf{F}_p$. Therefore, the following theorem does require a proof.

1.5. Theorem. *Let K be a field. Then formula* 1.4 *defines a group structure on the set $X(K) = \{(x,y) \in K^2 : x^2+y^2=1\}$.*

Proof. It is straightforward but unenlightening to check the group axioms from the definition. One can however observe that under the injective map $X(K) \to \mathrm{SL}_2(K)$ given by $(x,y) \mapsto \left(\begin{smallmatrix} y & -x \\ x & y \end{smallmatrix}\right)$, the operation given by 1.4 corresponds to the well known matrix multiplication. It follows that 1.4 defines a group structure on $X(K)$. □

Exercise 6. Let K be a field of characteristic 2. Show that the projection $\pi_x : X(K) \to K$ mapping (x,y) to x is a group isomorphism.

As is shown by the preceding exercise, one has to be careful when interpreting pictures over the complex numbers—such as that of the generically 2-to-1 projection $\pi_x : X(\mathbf{C}) \to \mathbf{C}$ above—in positive characteristic.

We now replace the differential $dt/\sqrt{1-t^2}$ in the preceding example by an elliptic differential $dt/\sqrt{f(t)}$ for some squarefree polynomial f of degree 3 or 4. We will see that the complex 'unit circle' $X = \{(x,y) : x^2+y^2=1\}$ gets replaced by the *elliptic curve* $E = \{(x,y) : y^2 = f(x)\}$, and the map $\phi^{-1} : z \to (\sin z, \cos z)$ by a map $z \mapsto (P(z), P'(z))$ for some *elliptic function* P. As in the case of the circle, the analytic parametrization by elliptic functions will equip E with a group structure. In the next section, we will give a geometric description of the group law and derive explicit algebraic addition formulas.

For a quadratic polynomial f a simple transformation $t \mapsto at+b$ suffices to map the roots of f to ±1, yielding a differential with $f(t) = 1-t^2$. In the elliptic case, the situation is more complicated. One can apply *Möbius transformations* $t \mapsto \frac{at+b}{ct+d}$ with $ad-bc \neq 0$, which act bijectively on the compactified complex plane $\mathbf{P}^1(\mathbf{C}) = \mathbf{C} \cup \{\infty\}$, commonly referred to as the *Riemann sphere*.

1.6. Lemma. *Under a Möbius transformation* $t \mapsto \frac{at+b}{ct+d}$, *elliptic differentials transform into elliptic differentials.*

Proof. It suffices to check this for a differential $\omega = dt/\sqrt{f(t)}$, with $f(t) = \sum_{k=0}^4 r_k t^k$ of degree 3 or 4. One finds that ω is transformed into

$$\omega^* = \frac{1}{\sqrt{f(\frac{at+b}{ct+d})}} d(\frac{at+b}{ct+d}) = \frac{(ad-bc)\ dt}{\sqrt{\sum_{k=0}^4 r_k (at+b)^k (ct+d)^{4-k}}}.$$

The polynomial $g(t) = \sum_{k=0}^4 r_k (at+b)^k (ct+d)^{4-k}$ is of degree at most 4. We leave it to the reader to verify that the degree is at least 3, so that ω^* is again elliptic. □

Exercise 7. Show that if the polynomial f in the preceding proof is of degree 4, the transformed differential has a polynomial g of degree 3 if and only if the Möbius transformation maps ∞ to a zero of f.

Möbius transformations can be used to map three of the roots of f to prescribed values in $\mathbf{P}^1(\mathbf{C})$. Different choices lead to different *normal forms* for elliptic differentials.

Exercise 8. Show that every elliptic differential $R(t, \sqrt{f(t)})$ can be transformed by a Möbius transformation into a differential for which f has one the following shapes:

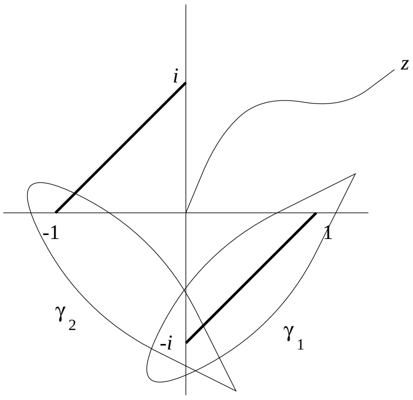
$$f(t) = t(t-1)(t-\lambda) \qquad f(t) = t^3 + at + b \qquad f(t) = (1-t^2)(1-k^2t^2).$$

[The corresponding normal forms are named after Legendre, Weierstrass and Jacobi.]

As an example of an elliptic differential, we consider the differential $\omega = dt/\sqrt{1-t^4}$ related to the rectification of the *lemniscate*. In order to find the analogue of 1.2 for ω, we start as in 1.1 and try to define a map

$$\psi : z \longmapsto \int_0^z \omega = \int_0^z \frac{dt}{\sqrt{1-t^4}}.$$

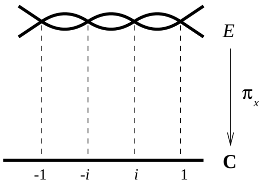
This time ω has integrable singularities in the 4th roots of unity, and it becomes single-valued if we make make branch cuts $[-1, i]$ and $[-i, 1]$. The picture in the complex plane is as follows.

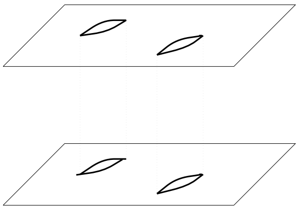


In order to obtain a natural domain for ω, we consider the algebraic curve E in $\mathbf{C}^2$ with equation

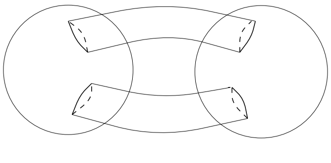
$$E = \{(x,y) : y^2 = 1 - x^4\} \subset \mathbf{C}^2.$$

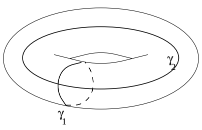
As in our previous example, the projection $\pi_x : E \to \mathbf{C}$ on the x-coordinate is generically 2-to-1 with branch points at ± 1 and $\pm i$. A topological model for E can be obtained by glueing two copies of $\mathbf{C}$ along our two branch cuts.





As $\psi(z)$ converges for $z \to \infty$, it makes sense to view ψ as a map on the Riemann sphere $\mathbf{P}^1(\mathbf{C})$. This means that we have to modify the picture above and add two 'points at infinity' to E, one coming from each copy of $\mathbf{C}$ in our topological picture. We write E again for the completed curve. We see from the picture that the glueing of two spheres along two branch cuts yields a doughnut-shaped surface known as a *torus*. On this surface, there are *two* independent incontractible paths. Under π_x, they are mapped to the paths γ_1 and γ_2 in our earlier picture. One can show that the homotopy classes of these paths generate the fundamental group $\pi(E) = \mathbf{Z} \times \mathbf{Z}$ of E.





Exercise 9. Show that $\gamma_1\gamma_2\gamma_1^{-1}\gamma_2^{-1}$ is a contractible path on E.

It follows that the values of ψ are uniquely determined as elements of the factor group $\mathbf{C}/(\mathbf{Z}\omega_1 + \mathbf{Z}\omega_2)$, where the *periods* ω_1 and ω_2 are defined as $\omega_i = \oint_{\gamma_i} \omega$ for $i = 1, 2$. From our initial picture we see that the path γ_1 maps to γ_2 onder multiplication by $-i$. As $1 - t^4$ is invariant under this transformation, we deduce that we have $\omega_2 = -i\omega_1$. The subgroup $\Lambda = \mathbf{Z}\omega_1 + \mathbf{Z}\omega_2$ is a rectangular *lattice* in $\mathbf{C}$, and the factor group $\mathbf{C}/\Lambda$ is therefore topologically a torus. We have the following analogue of 1.2.

1.7. Theorem. *The integration of the differential* $\omega = \frac{dx}{y}$ *along the completed curve* E *induces a bijection* $\psi : E \xrightarrow{\sim} \mathbf{C}/(\mathbf{Z}\omega_1 + \mathbf{Z}\omega_2)$.

For a complete proof of 1.7, and for similar results for other elliptic differentials, we refer to the exercises.

As a consequence of 1.7, we see that the *elliptic curve* E carries a natural group structure. Let the inverse function $\psi^{-1} : \mathbf{C}/(\mathbf{Z}\omega_1 + \mathbf{Z}\omega_2) \xrightarrow{\sim} E$ be given by $\psi^{-1}(z) = (P(z), Q(z))$. As the derivative of ψ in (x, y) with respect to x is by construction equal to $1/y$, the derivative of P in $z = \psi((x, y))$ equals $y = Q(z)$. We conclude that as in the previous example, the inverse of ψ is of the form $\psi^{-1}(z) = (P(z), P'(z))$ for some *elliptic function* P. As E has two points at infinity, the 'lemniscatic P-function' $P(z)$ has a pole in two values of z in $\mathbf{C}/(\mathbf{Z}\omega_1 + \mathbf{Z}\omega_2)$. At all other points, it is holomorphic. From the equation of E, it is clear that P is a solution to the differential equation

$$(P')^2 = 1 - P^4.$$

As a function on $\mathbf{C}$, it has even stronger periodicity properties than the sine function: it is *double-periodic* with independent periods ω_1 and ω_2.

Exercise 10. Define $p = \int_{-1}^{1} dt/\sqrt{1-t^4} \approx 2.622057556$, the elliptic analogue of $\pi = \int_{-1}^{1} dt/\sqrt{1-t^2}$. Show that we can take $\omega_1 = p + ip$ and $\omega_2 = p - ip$ in 1.7, and that the elliptic function P has poles in $\omega_1/2$ and $\omega_2/2$. Are these poles simple?

Just as the sine and cosine functions are more convenient to handle than the arcsine and arccosine functions arising from the integration of $dt/\sqrt{1-t^2}$, the functions P and P' constructed above are easier to study than the function $\psi(z) = \int^{z} dt/\sqrt{1-t^4}$. By clever substitutions in the integral defining ψ, one can prove Fagnano's duplication formula

$$P(2z) = \frac{2P(z)P'(z)}{1 + P(z)^4},$$

which dates back to 1718. Euler extended this result in 1751 and found the general addition formula

$$P(z_1) + P(z_2) = \frac{P(z_1)P'(z_2) + P'(z_1)P(z_2)}{1 + P(z_1)^2 P(z_2)^2}.$$

for the lemniscatic P-function.

The next section is devoted to the analysis of analytic functions on an arbitrary torus. We will show directly that *all* tori come with functions satisfying algebraic addition formulas.

Exercises.

11. Adapt the statement in exercise 1 for rational functions f with real coefficients and show that $\int f(t)dt$ can be expressed in terms of 'real elementary functions'.

12. Show that the map ϕ in 1.1 is well-defined as a map on the upper half plane $\mathcal{H} = \{z \in \mathbf{C} : \mathrm{Im} z > 0\}$, provided that we fix a branch of $\sqrt{1-t^2}$ on $\mathcal{H}$. Show that for the branch that is positive on $i\mathbf{R}_{>0}$, we obtain a bijective map $\phi : \mathcal{H} \to S$ to the semi-infinite strip $S = \{z \in \mathbf{C} : \mathrm{Re} z > 0 \text{ and } -\pi/2 < \mathrm{Im} z < \pi/2\}$. Derive theorem 1.2 from this statement. [Hint: determine the image of the real axis under ϕ.]

*13. Show that the map ϕ in 1.2 is an isomorphism of complex analytic spaces, i.e., a biholomorphic map between open Riemann surfaces.

14. A *lemniscate of Bernoulli* is the set L of points X in the Euclidean plane for which the product of the distances XP_1 and XP_2, with P_1 and P_2 given points at distance $P_1P_2 = 2d > 0$, is equal to d^2.
    a. Show that for a suitable choice of coordinates, the equation for L is $(x^2+y^2)^2 = x^2 - y^2$ or, in polar coordinates, $r^2 = \cos 2\phi$. Sketch this curve.
    b. Show that the arclength of the 'unit lemniscate' in (a) equals $2p$, with p defined as in exercise 10. [Note the similarity with the arclength of the unit circle, which equals 2π.]

15. This exercise gives a 'proof by algebraic manipulation' of Fagnano's duplication formula for the lemniscatic P-function. See also exercises 5.12 and 5.13.
    a. Show that the substitution $t = 2v/(1+v^2)$ transforms the differential $dt/\sqrt{1-t^2}$ to the rational differential $2dv/(1+v^2)$.
    b. Show that the substitution $t^2 = 2v^2/(1+v^4)$ transforms the differential $dt/\sqrt{1-t^4}$ to the differential $\sqrt{2}dv/(1+v^4)$, and that the subsequent substitution $v^2 = 2w^2/(1+w^4)$ leads to the differential $2dw/(1-w^4)$.
    c. Derive the relation $t = 2w\sqrt{1-w^4}/(1+w^4)$ for variables in (b), and use this to prove Fagnano's formula.

16. On the complex upper half plane $\mathcal{H}$, we can uniquely define a function

$$\phi(z) = \int_0^z \frac{dt}{\sqrt{(1-t^2)(4-t^2)}}$$

by integrating along paths in $\mathcal{H}$. We use the branch of $\sqrt{(1-t^2)(4-t^2)}$ that is positive on $i\mathbf{R}_{>0}$. Define $A, B \in \mathbf{C}$ by $A = \lim_{z\to 1}\phi(z)$ and $A + B = \lim_{z\to 2}\phi(z)$.
    a. Show that A is real and B purely imaginary, and that we have $\lim_{z\to\infty}\phi(z) = B$.
    b. Show that the map ϕ extends to a bijection between the completion of the elliptic curve $y^2 = (1-x^2)(4-x^2)$ and the torus $\mathbf{C}/\Lambda$ with $\Lambda = \mathbf{Z}\cdot 4A + \mathbf{Z}\cdot 2B$.

17. Prove theorem 1.7. [Hint: imitate the previous exercise.]

*18. Show that the map ψ in 1.7 is an isomorphism of complex analytic spaces, i.e., a biholomorphic map between compact Riemann surfaces.

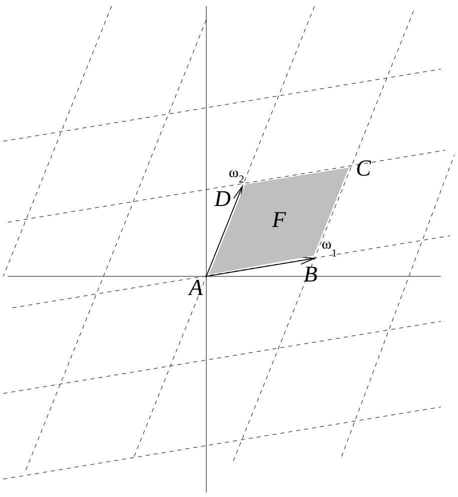
2. Elliptic functions

In this section, we will develop the basic theory of double-periodic functions encountered in the previous section.

A *lattice* in $\mathbf{C}$ is a discrete subgroup of $\mathbf{C}$ of rank 2. It has the form $\Lambda = \mathbf{Z}\omega_1 + \mathbf{Z}\omega_2$ for some $\mathbf{R}$-basis $\{\omega_1, \omega_2\}$ of $\mathbf{C}$. One often writes $\Lambda = [\omega_1, \omega_2]$. The factor group $T = \mathbf{C}/\Lambda$ is called a *complex torus*. A *fundamental domain* for T is a connected subset $F \subset \mathbf{C}$ for which every $z \in \mathbf{C}$ can uniquely be written as $z = f + \omega$ with $f \in F$ and $\omega \in \Lambda$. Note that any translate of a fundamental domain is again a fundamental domain. For every choice $\{\omega_1, \omega_2\}$ of a $\mathbf{Z}$-basis of Λ, the set

$$F = \{r_1\omega_1 + r_2\omega_2 : r_1, r_2 \in \mathbf{R}, \quad 0 \leq r_1, r_2 < 1\} \subset \mathbf{C}$$

is a fundamental domain for T.



An *elliptic function* with respect to Λ is a meromorphic function f on $\mathbf{C}$ that satisfies $f(z + \omega) = f(z)$ for all $\omega \in \Lambda$. Such a function is uniquely determined by its values on a fundamental domain. An elliptic function factors as $f : \mathbf{C} \to \mathbf{C}/\Lambda = T \to \mathbf{P}^1(\mathbf{C})$, so we can identify the set of elliptic functions with respect to Λ with the set $\mathcal{M}(T)$ of meromorphic functions on $T = \mathbf{C}/\Lambda$. Sums and quotients of meromorphic functions are again meromorphic, so the set $\mathcal{M}(T)$ is actually a field, the *elliptic function field* corresponding to T.

As T is compact, any holomorphic function $f \in \mathcal{M}(T)$ is bounded on T. This means that f comes from a bounded holomorphic function on $\mathbf{C}$, so by Liouville's theorem f is constant. We conclude that any non-constant elliptic function has at least one pole.

Exercise 1. Show that for any non-constant $f \in \mathcal{M}(T)$, the map $f : T \to \mathbf{P}^1(\mathbf{C})$ is surjective.

The most convenient way to describe the zeroes and poles of a function $f \in \mathcal{M}(T)$ is to define its associated *divisor*. The *divisor group* $\mathrm{Div}(T)$ is the free abelian group generated by the points of T. Equivalently, a divisor

$$D = \sum_{w\in T} n_w[w] \in \mathrm{Div}(T) = \oplus_{w\in T}\mathbf{Z}$$

is a *finite* formal sum of points of T with integer coefficients.

The divisor group $\mathrm{Div}(T)$ comes with canonical surjective homomorphisms to T and $\mathbf{Z}$. The *summation map* $\Sigma : \mathrm{Div}(T) \to T$ sends $\sum_{w\in T} n_w[w]$ to $\sum_{w\in T} n_w w$. The *degree map* $\deg : \mathrm{Div}(T) \to \mathbf{Z}$ sends $\sum_{w\in T} n_w[w]$ to $\sum_{w\in T} n_w$. The kernel of the degree map is the subgroup $\mathrm{Div}^0(T) \subset \mathrm{Div}(T)$ of divisors of degree zero.

The *order* $\mathrm{ord}_w(f) \in \mathbf{Z}$ of a non-zero function $f \in \mathcal{M}(T)^*$ at a point $w \in T$ is the minimum of all k for which the coefficient c_k in the Laurent expansion $f(z) = \sum_k c_k(z-w)^k$ of f around w is non-zero. If we view poles as zeroes of negative order, $\mathrm{ord}_w(f) \in \mathbf{Z}$ is simply the order of the zero of f in w.

A meromorphic function $f \in \mathcal{M}(T)^*$ has only finitely many zeroes and poles on the compact torus T, so the divisor map

$$\begin{aligned}\mathrm{div} : \mathcal{M}(T)^* &\longrightarrow \mathrm{Div}(T)\\ f &\longmapsto (f) = \sum_{w\in T} \mathrm{ord}_w(f)[w]\end{aligned}$$

is a well-defined homomorphism. The divisors in $\mathrm{Div}(T)$ coming from elliptic functions are called *principal divisors*. We will prove that a divisor is principal if and only if it is in the kernel of both Σ and deg.

2.1. Theorem. *Let $T = \mathbf{C}/\Lambda$ be a torus. Then there is an exact sequence*

$$1 \longrightarrow \mathbf{C}^* \longrightarrow \mathcal{M}(T)^* \xrightarrow{\mathrm{div}} \mathrm{Div}^0(T) \xrightarrow{\Sigma} T \longrightarrow 1.$$

As only constant functions on T are without zeroes and poles, the sequence is exact at $\mathcal{M}(T)^*$. The proof of the exactness at $\mathrm{Div}^0(T)$ consists of two parts. We first prove that principal divisors are of degree zero and in the kernel of the summation map. These are exactly the statements (ii) and (iii) of the lemma below.

2.2. Lemma. *Let f be a non-zero elliptic function on T. Then the following holds.*

(i) $\sum_{w\in T} \mathrm{res}_w(f) = 0$.

(ii) $\sum_{w\in T} \mathrm{ord}_w(f) = 0$.

(iii) $\sum_{w\in T} \mathrm{ord}_w(f)\cdot w = 0 \in T$.

Proof. Let F be a fundamental domain for T, and suppose—after translating F when necessary—that none of the zeroes and poles of f lies on the boundary ∂F of F. Then the expressions of the lemma are the values of the contour integrals

$$\frac{1}{2\pi i}\oint_{\partial F} f(z)dz, \qquad \frac{1}{2\pi i}\oint_{\partial F} \frac{f'(z)}{f(z)}dz, \qquad \frac{1}{2\pi i}\oint_{\partial F} z\frac{f'(z)}{f(z)}dz.$$

The first two integrals vanish since, by the periodicity of f and f'/f, the integrals along opposite sides of the parallellogram F coincide; as these sides are traversed in opposite directions, their contributions to the integral cancel.

The function $z\frac{f'(z)}{f(z)}$ is not periodic, but we can still compute the contribution to the integral coming from opposite sides AB and $DC = \{z + \omega_2 : z \in AB\}$ of F, as indicated in the earlier picture. We find

$$\int_{AB} z\tfrac{f'(z)}{f(z)}dz + \int_{CD} z\tfrac{f'(z)}{f(z)}dz = \int_{AB} z\tfrac{f'(z)}{f(z)}dz - \int_{AB}(z+\omega_2)\tfrac{f'(z)}{f(z)}dz = -\omega_2 \int_{AB} \tfrac{f'(z)}{f(z)}dz.$$

As the integral $\frac{1}{2\pi i}\int_{AB}\frac{f'(z)}{f(z)}dz$ is the winding number of the *closed* path described by $\frac{f'(z)}{f(z)}$ if z ranges from A to B along ∂F, $\frac{1}{2\pi i}$ times the value of the displayed integral is an integral multiple of ω_2, hence in Λ. The same holds for the other half $\int_{BC} + \int_{DA}$ of the integral, which yields an integral multiple of ω_1. The complete integral now assumes a value in Λ, and (iii) follows.

Assertion (ii) of the lemma shows that an elliptic function has as many zeroes as it has poles on T, if we count multiplicities. The number of zeroes (or, equivalently, poles) of an elliptic function f, counted with multiplicity, is called the *order* of f. Equivalently, it is the degree of the *polar divisor* $\sum_w \max(0, -\mathrm{ord}_w(f)) \cdot (w)$ of f. It follows from (i) that the order of an elliptic function cannot be equal to 1.

Exercise 2. Define the order of a meromorphic function on $\mathbf{P}^1(\mathbf{C})$, and show that functions of arbitrary order exist.

In order to complete the proof of 2.1, we need to show that a divisor of degree zero that is in the kernel of Σ actually corresponds to a function on T. This means that we somehow have to construct these functions.

Function theory provides us with two methods to construct meromorphic functions with prescribed zeroes or poles. An additive method consists in writing down a series expansion for the 'simplest elliptic function' associated to the lattice Λ, the *Weierstrass-$\wp$-function* $\wp_\Lambda(z)$. This is an even function of order 2 on T, which has a double pole at $0 \in T$. It is given by

$$\wp(z) = \wp_\Lambda(z) = \frac{1}{z^2} + \sum_{\omega\in\Lambda\setminus\{0\}} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}\right). \tag{2.3}$$

In order to show that the defining series converges uniformly on compact subsets of $\mathbf{C} \setminus \Lambda$, one uses the following basic lemma.

2.4. Lemma. *The Eisenstein series $G_k(\Lambda) = \sum_{\omega\in\Lambda\setminus\{0\}} \omega^{-k}$ is absolutely convergent for every integer $k > 2$.*

The proof of 2.4 is elementary. One can estimate the number of lattice points in an annulus $\{z \in \mathbf{C} : N \le z \le N+1\}$. Note that the values $G_k(\Lambda)$ equal zero if $k > 2$ is odd, since then the terms for ω and $-\omega$ cancel. □

Exercise 3. Prove lemma 2.4.

From the lemma, one deduces that $\wp_\Lambda$ is a well-defined meromorphic function on $\mathbf{C}$ with double poles at the elements of Λ. Some elementary calculus leads to the Laurent expansion

$$\wp(z) = \frac{1}{z^2} + \sum_{n=1}^{\infty}(2n+1)G_{2n+2}(\Lambda)z^{2n} \tag{2.5}$$

for $\wp(z)$ around the origin. In order to show that $\wp_\Lambda$ is periodic modulo Λ, one notes first that the derivative $\wp'(z) = \sum_{\omega\in\Lambda}(z-\omega)^{-3}$ is clearly periodic modulo Λ. For $\wp$ itself, it follows that for $\omega \in \Lambda$ we have $\wp(z+\omega) = \wp(z) + c_\omega$. As $\wp$ is an *even* function, we can take $z = \omega/2$ to find $c_\omega = 0$, so $\wp$ is also periodic modulo Λ.

A second method to construct periodic functions proceeds multiplicatively, by writing down a convergent Weierstrass product

$$\sigma(z) = \sigma_\Lambda(z) = z \prod_{\omega\in\Lambda\setminus\{0\}} (1 - \frac{z}{\omega})e^{(z/\omega)+\frac{1}{2}(z/\omega)^2}$$

for a function having simple zeroes at the points in Λ.

Exercise 4. Show that the product expansion for the σ-function converges uniformly on compact subsets of $\mathbf{C}$. [Hint: pass to the logarithm and use 2.3.]

By 2.3, termwise differentiation of the logarithmic derivative

$$\frac{d\log\sigma(z)}{dz} = \frac{\sigma'(z)}{\sigma(z)} = \frac{1}{z} + \sum_{\omega\in\Lambda\setminus\{0\}} \left(\frac{1}{z-\omega} + \frac{1}{\omega} + \frac{z}{\omega^2}\right) \tag{2.6}$$

yields the relation $\frac{d^2\log\sigma(z)}{dz^2} = -\wp(z)$. As $\wp(z)$ is periodic, we can find $a_\omega, b_\omega \in \mathbf{C}$ for each $\omega \in \Lambda$ such that we have $\sigma(z+\omega) = e^{a_\omega z + b_\omega}\sigma(z)$ for all $z \in \mathbf{C}$. One sometimes says that $\sigma(z)$ is a *theta function* with respect to the lattice Λ.

We are now in a position to finish the proof of 2.1. We still need to show that every divisor $D = \sum_w n_w[w]$ that is of degree 0 and in the kernel of the summation map is the divisor of an elliptic function. Write $\Sigma(D) = \sum_w n_w w = \omega \in \Lambda$. If ω is non-zero, we add the trivial divisor $[0] - [\omega]$ to D to obtain a divisor satisfying $\sum_w n_w w = 0$. Now the function $f_D = \prod_w \sigma(z-w)^{n_w}$ has divisor $\sum_w n_w[w]$, and for any $\omega \in \Lambda$ we find

$$\sigma(z+\omega) = e^{a_\omega \sum_w n_w w + b_\omega \sum_w n_w}\sigma(z) = \sigma(z).$$

Therefore f_D is in $\mathcal{M}(T)^*$. This finishes the proof of 2.1.

The factor group $\mathrm{Jac}(T) = \mathrm{Div}^0(T)/\mathrm{div}[\mathcal{M}(T)^*]$ of divisor classes of degree zero is the *Jacobian* of T. The content of theorem 2.1 may be summarized by the statement that T is canonically isomorphic to its Jacobian. We will return to this important property in 4.?.

The actual construction of elliptic functions in the proof of 2.1 shows that the field $\mathcal{M}(T)$ can be given explicitly in terms of functions related to the $\wp$-function. The precise statement is as follows.

2.7. Theorem. *The elliptic function field corresponding to $T = \mathbf{C}/\Lambda$ equals*

$$\mathcal{M}(T) = \mathbf{C}(\wp_\Lambda, \wp'_\Lambda).$$

This is a quadratic extension of the field $\mathbf{C}(\wp_\Lambda)$ of even elliptic functions.

Proof. Any elliptic function f is the sum $f(z) = \frac{f(z)+f(-z)}{2} + \frac{f(z)-f(-z)}{2}$ of an even and an odd elliptic function, and for odd f the function $f\wp'$ is even. It follows that $\wp'$ generates $\mathcal{M}(T)$ over the field of even elliptic functions, and that this extension is quadratic.

Let $f \in \mathcal{M}(T)^*$ be even. We need to show that f is a rational expression in $\wp = \wp_\Lambda$. We note first that $\mathrm{ord}_w(f)$ is even at '2-torsion points' w satisfying $w = -w \in T$: this follows from the fact that the derivatives of odd order of f are odd elliptic functions, and such functions have non-zero order at a point $w = -w \in T$. We can therefore write

$$(f) = \sum\nolimits_{w\in T} c_w([w] + [-w]) = \sum\nolimits_{w\in T} c_w([w] + [-w] - 2[0]).$$

We can assume that no term with $w = 0$ occurs in the last sum. As the functions f and $\prod_w (\wp(z) - \wp(w))^{n_w}$ have the same divisor, their quotient is a constant. □

Exercise 5. Let $f \in \mathcal{M}(T)$ have polar divisor $2 \cdot (0)$. Prove: $f = c_1\wp + c_2$ for certain $c_1, c_2 \in \mathbf{C}$.

The function $\wp'$ is an odd elliptic function with polar divisor $3 \cdot (0)$, so it is of order 3. Its 3 zeroes are the 3 points $\omega_1/2$, $\omega_2/2$ and $\omega_3/2 = (\omega_1 + \omega_2)/2$ of order 2 in $T = \mathbf{C}/\Lambda$. The even function $(\wp')^2$ has divisor $\sum_{i=1}^{3}[2 \cdot (\omega_i/2) - 2 \cdot (0)]$, so the preceding proof and a look at the first term $4z^{-6}$ of the Laurent expansion of $(\wp')^2$ around 0 show that we have a differential equation

$$(\wp'(z))^2 = 4\prod_{i=1}^{3}(\wp(z) - \wp(\omega_i/2)). \tag{2.8}$$

The coefficients of the cubic polynomial in 2.8 depend on the lattice Λ in the following explicit way.

2.9. Theorem. *The $\wp$-function for Λ satisfies a Weierstrass differential equation*

$$(\wp'_\Lambda)^2 = 4\wp_\Lambda^3 - g_2\wp_\Lambda - g_3$$

with coefficients $g_2 = g_2(\Lambda) = 60G_4(\Lambda)$ and $g_3 = g_3(\Lambda) = 140G_6(\Lambda)$. The discriminant $\Delta(\Lambda) = g_2(\Lambda)^3 - 27g_3(\Lambda)^2$ does not vanish.

Proof. The derivation of the differential equation is a matter of careful administration based on the Laurent expansion around $z = 0$ in (2.5). From the local expansions $\wp(z) = z^{-2} + 3G_4z^2 + O(z^4)$ and $\wp'(z) = -2z^{-3} + 6G_4z + 20G_6z^3 + O(z^5)$ one easily finds

$$\begin{aligned}(\wp'(z))^2 &= 4z^{-6} - 24G_4z^{-2} - 80G_6 + O(z^2)\\ 4\wp(z)^3 &= 4z^{-6} + 36G_4z^{-2} + 60G_6 + O(z^2).\end{aligned}$$

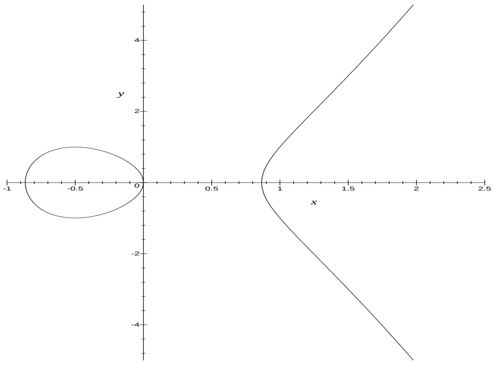
It follows that $(\wp'(z))^2 - 4\wp^3 + 60G_4\wp + 140G_6$ is a holomorphic elliptic function that vanishes at the origin, so it is identically zero. For the non-vanishing of the discriminant

$$\begin{aligned}\Delta(\Lambda) &= g_2(\Lambda)^3 - 27g_3(\Lambda)^2\\ &= 16\cdot(\wp(\tfrac{\omega_1}{2}) - \wp(\tfrac{\omega_2}{2}))^2\cdot(\wp(\tfrac{\omega_1}{2}) - \wp(\tfrac{\omega_3}{2}))^2\cdot(\wp(\tfrac{\omega_2}{2}) - \wp(\tfrac{\omega_3}{2}))^2,\end{aligned}$$

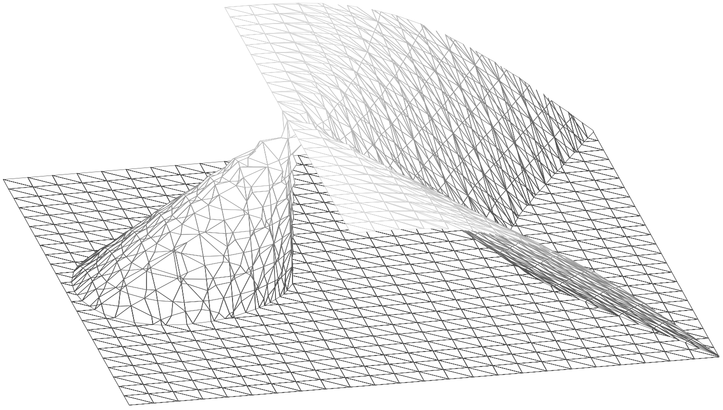
one observes that the function $\wp(z) - \wp(\omega_i/2)$ is elliptic of order 2 with a double zero at $\omega_i/2$, so it cannot vanish at $\omega_j/2$ for $j \neq i$. □

Exercise 6. Show that the non-constant solutions to the differential equation $(y')^2 = 4y^3 - g_2y - g_3$ corresponding to a lattice Λ are the functions $\wp_\Lambda(z - z_0)$ with $z_0 \in \mathbf{C}$. What are the constant solutions?

It follows from 2.9 that the map $W : z \mapsto (\wp(z), \wp'(z))$ maps T to a complex curve in $\mathbf{C}^2$ with equation $y^2 = 4x^3 - g_2x - g_3$. This is exactly the kind of map we have been considering in section 1. If g_2 and g_3 are real, one can sketch the curve in $\mathbf{R}^2$. For a Weierstrass polynomial having three real roots the picture looks as follows.



In order to deal with the poles of the map W, we pass to the *projective completion* of our curve in $\mathbf{P}^2(\mathbf{C})$. This is by definition the zero set in $\mathbf{P}^2(\mathbf{C})$ of the homogenized equation $Y^2Z = 4X^3 - g_2XZ^2 - g_3Z^3$; it consists of the 'affine points' $(x : y : 1)$ coming from the original curve and the 'point at infinity' $(0 : 1 : 0)$. One can view the lines through the origin in $\mathbf{R}^3$ as the points of the real projective plane $\mathbf{P}^2(\mathbf{R})$, and draw the following picture of the completed curve. The point at infinity in this picture is the single line in the plane $Z = 0$.



2.10. Theorem. *Let $\Lambda \subset \mathbf{C}$ be a lattice. Then the Weierstrass map*

$$W: z \longmapsto \begin{cases} (\wp(z) : \wp'(z) : 1) & \text{for } z \neq 0 \\ (0:1:0) & \text{for } z = 0 \end{cases}$$

induces a bijection between the torus $T = \mathbf{C}/\Lambda$ and the complex elliptic curve E_Λ with projective Weierstrass equation

$$E_\Lambda : \quad Y^2Z = 4X^3 - g_2(\Lambda)XZ^2 - g_3(\Lambda)Z^3.$$

Proof. By 2.9, the torus $\mathbf{C}/\Lambda$ is mapped to the curve E_Λ. We have to show that every affine point $P = (x, y)$ on E_Λ is the image of a unique point $z \in T \setminus \{0\}$. The divisor of the function $\wp(z) - x$ is of the form $(w) + (-w) - 2(0)$ for some $w \in T$ that is determined up to sign. For $w = -w \in T$ we have $y = 0$, and $z = w$ is the unique point mapping to P. Otherwise, we have $\wp'(w) = \pm y \neq 0$, and exactly one of w and $-w$ maps to P. □

2.11. Corollary. *The Weierstrass parametrization* 2.10 *induces a group structure on the set $E_\Lambda(\mathbf{C})$ of points of the elliptic curve E_Λ. The zero element of $E_\Lambda(\mathbf{C})$ is the 'point at infinity' $O_E = (0 : 1 : 0)$, and the inverse of the point $(X : Y : Z)$ is $(X : -Y : Z)$. Any three distinct points in $E_\Lambda(\mathbf{C})$ that are collinear in $\mathbf{P}^2(\mathbf{C})$ have sum O_E.*

Proof. It is clear that $W(0) = O_E$ is the zero element for the induced group structure on $E_\Lambda(\mathbf{C})$, and that the inverse of the point $(\wp(z) : \wp'(z) : 1)$ is $(\wp(-z) : \wp'(-z) : 1) = (\wp(z) : -\wp'(z) : 1)$. It remains to show that three collinear points in $E_\Lambda(\mathbf{C})$ have sum zero. Let $L : aX + bY + cZ = 0$ be the line passing through three such points, and consider the associated elliptic function $f = a\wp + b\wp' + c$. If b is non-zero, the divisor of f is of the form $(f) = (w_1) + (w_2) + (w_3) - 3(0)$ for certain $w_i \in T$. We have $w_1 + w_2 + w_3 = 0 \in T$ by 2.1 (iii), and since the Weierstrass parametrization W maps the w_i to the three points of intersection of L and E_Λ, these points have sum O_E. For $b = 0$ and $a \neq 0$, we are in the case of a 'vertical line' with affine equation $x = -c/a$. The point O_E is on this line. The function $f = a\wp + c$ now has divisor $(f) = (w_1) + (w_2) - 2(0)$, and the same argument as above shows that the 2 affine points of intersection of L and E_Λ are inverse to each other. The case $a = b = 0$ does not occur since then the line L is the line at infinity $Z = 0$, which intersects E_Λ only in O_E. □

Exercise 7. Define multiplicities for the points of intersection of E_Λ with an arbitrary line L, and show that with these multiplicities the 'sum of the points in $L \cap E_\Lambda$' is always equal to O_E.

Corollary 2.11 shows that the group law on $E_\Lambda(\mathbf{C})$ has a simple geometric interpretation. In order to find the sum of 2 points P and Q in $E_\Lambda(\mathbf{C})$, one finds the third point $R = (a, b)$ of intersection of the line through P and Q with E. One than has $P + Q = -R$, so the sum of P and Q equals $(a, -b)$.

From the geometric description, one can derive an explicit addition formula for the points on E_Λ or, equivalently, addition formulas for the functions $\wp$ and $\wp'$. Let $P = (\wp(z_1), \wp'(z_1))$ and $Q = (\wp(z_2), \wp'(z_2))$ be points on E_Λ. If P and Q are inverse to each other, we have $z_1 = -z_2 \bmod \Lambda$ and $P+Q$ is the infinite point O_E. Otherwise, the affine line through P and Q is of the form $y = \lambda x + \mu$ with

$$\lambda = \frac{\wp'(z_1) - \wp'(z_2)}{\wp(z_1) - \wp(z_2)} = \frac{4\wp(z_1)^2 + 4\wp(z_1)\wp(z_2) + 4\wp(z_2)^2 - g_2}{\wp'(z_1) + \wp'(z_2)}.$$

The second expression, which is obtained by multiplication of numerator and denominator of the first expression and applying 2.9, is also well-defined for $P = Q$; in this case it yields the slope of the tangent line in P. As the cubic equation $4x^3 - g_2x - g_3 - (\lambda x + \mu)^2$ has roots $\wp(z_1)$, $\wp(z_2)$ and $\wp(z_1 + z_2)$, we find the x coordinate of $P+Q$ to be

$$(2.12) \qquad \wp(z_1 + z_2) = -\wp(z_1) - \wp(z_2) + \frac{1}{4}\left(\frac{\wp'(z_1) - \wp'(z_2)}{\wp(z_1) - \wp(z_2)}\right)^2 \qquad (z_1 \neq \pm z_2 \bmod \Lambda).$$

In the case $P = Q$, one can use the second expression for Λ to find the x-coordinate $\wp(2z_1)$ of $2P$ as a rational function in $\wp(z_1)$.

Exercise 8. Write $\wp(2z)$ as a rational function in $\wp(z)$. Show that this duplication formula for the $\wp$-function also follows from the limit form

$$\wp(2z) = -2\wp(z) + \frac{1}{4}\left(\frac{\wp''(z)}{\wp'(z)}\right)^2$$

of 2.12 and the differential equation $\wp'' = 6\wp^2 - \frac{1}{2}g_2$, which is obtained by differentiating 2.9.

As in the previous section, we find that the addition formulas on the elliptic curve E_Λ are *algebraic formulas* involving the coefficients g_2 and g_3 of the defining Weierstrass equation. We say that an elliptic curve E with Weierstrass equation $y^2 = 4x^3 - g_2x - g_3$ is *defined over* a subfield $K \subset \mathbf{C}$ if g_2 and g_3 are in K. If E is defined over a field $K \subset \mathbf{C}$, the set $E(K)$ of K-valued points is a subgroup of $E(\mathbf{C})$. We will especially be interested in the case where K is the field of rational numbers. When working over $\mathbf{Q}$, it is often convenient to choose variables $X = 4x$ and $Y = 4y$ satisfying the equation $Y^2 = X^3 - 4g_2 - 16g_3$.

In this case the determination of the group $E(\mathbf{Q})$ is a highly non-trivial problem that has its roots in antiquity. The observation that two (not necessarily distinct) points on a cubic curve can be used to find a third point already goes back to Diophantus. His method, which is basically a method for adding points, is known as the *chord-tangent method.* We will give a similar description of the group structure on the set of points of an arbitrary plane cubic curve in section 4.

Exercises.

9. Let f be a non-constant meromorphic function on $\mathbf{C}$. A number $\omega \in \mathbf{C}$ is said to be a *period* of f if $f(z+\omega) = f(z)$ for all $z \in \mathbf{C}$. Let Λ be the set of periods of f.
   a. Prove that Λ is a discrete subgroup of $\mathbf{C}$.
   b. Deduce that Λ is of one of the three following forms:

$$\Lambda = \{0\} \qquad \Lambda = \mathbf{Z}\omega \quad (\omega \neq 0) \qquad \Lambda = \mathbf{Z}\omega_1 \oplus \mathbf{Z}\omega_2 \quad (\text{with } \mathbf{C} = \mathbf{R}\omega_1 + \mathbf{R}\omega_2)$$

10. Let $\wp$ be the $\wp$-function associated to Λ. Show that the function $z \mapsto e^{\wp(z)}$ is holomorphic on $\mathbf{C} \setminus \Lambda$ and periodic modulo Λ, but not elliptic.

11. Let f be a meromorphic function with non-zero period ω and define $q = q(z) = e^{2\pi i z/\omega}$. Prove that there exists a meromorphic function $\hat{f}$ on $\mathbf{C}^*$ satisfying $f(z) = \hat{f}(q)$, and show that we have $\mathrm{ord}_q(\hat{f}) = \mathrm{ord}_z(f)$ for all $z \in \mathbf{C}$.

12. Let Λ be lattice and $\wp$ and σ the associated complex functions. Prove the identity

$$\wp(z) - \wp(a) = -\frac{\sigma(z-a)\sigma(z+a)}{\sigma(a)^2\sigma(z)^2} \qquad (a \notin \Lambda).$$

13. *(Degeneracy of the $\wp$-function.)* Let ω be an element in $\mathbf{C} \setminus \mathbf{R}$ and t a real number.
   a. Prove the identities

$$\lim_{t\to\infty} \wp_{[t,\omega t]}(z) = \frac{1}{z^2} \qquad \text{and} \qquad \lim_{t\to\infty} \wp_{[1,\omega t]}(z) = \frac{1}{\sin^2(\pi z)} + \frac{3}{\pi^2}$$

   for $z \in \mathbf{C}^*$ and $z \in \mathbf{C} \setminus \mathbf{Z}$, respectively.
   b. What are the degenerate forms of the function $\sigma(z)$ corresponding to the two cases above, and which identities replace the one in the previous exercise?
   c. Find the degenerate analogues of 2.10, and explain why these two forms of degeneracy are called *additive* and *multiplicative*, respectively.

14. Determine the general solution of the Weierstrass differential equation $(y')^2 = 4y^3 - g_2 y - g_3$ in the degenerate case $g_2^3 = 27g_3^2$.

15. Show that the derivative of the $\wp$-function satisfies $\wp'(z) = -\frac{\sigma(2z)}{\sigma(z)^4}$.

16. Let $\Lambda = [\omega_1, \omega_2]$ be a lattice with associated Weierstrass function $\wp$, and consider the Weierstrass functions $\wp_1$ and $\wp_2$ associated to the lattices $\Lambda_1 = \frac{1}{2}\Lambda$ and $\Lambda_2 = [\frac{1}{2}\omega_1, \omega_2]$. Prove the identities

$$\wp_1(z) = 4\wp(2z) \qquad \text{and} \qquad \wp_2(z) = \wp(z) + \wp(z + \tfrac{1}{2}\omega_1) - \wp(\tfrac{1}{2}\omega_1).$$

   What are the corresponding identities for $\wp_1'$ and $\wp_2'$?

17. Prove: $4\wp(2z) = \wp(z) + \wp(z + \frac{1}{2}\omega_1) + \wp(z + \frac{1}{2}\omega_2) + \wp(z + \frac{1}{2}\omega_3)$.

18. Define the *Weierstrass ζ-function* for the lattice $\Lambda = \mathbf{Z}\omega_1 + \mathbf{Z}\omega_2$ in $\mathbf{C}$ as in (2.6) by $\zeta(z) = \frac{d}{dz}\log\sigma(z)$.

a. Show that there exists a linear function $\eta : \Lambda \to \mathbf{C}$ such that $\zeta(z+\omega) = \zeta(z) + \eta(\omega)$ for $\omega \in \Lambda$ and $z \in \mathbf{C}$, and that $\eta(\omega) = 2\zeta(\omega/2)$ if $\omega \neq 2\Lambda$.

The numbers $\eta_i = \eta(\omega_i)$ $(i = 1, 2)$ are the *quasi-periods* of $\zeta(z)$.

b. Prove the *Legendre relation* $\eta_1\omega_2 - \eta_2\omega_1 = \pm 2\pi i$.
[Hint: the right hand side equals $\oint \zeta(z)dz$ around a fundamental parallelogram.]

c. Prove: $\sigma(z+\omega) = \pm e^{\eta(\omega)(z+\omega/2)}\sigma(z)$.

19. *(Weil reciprocity law.)* For an elliptic function f and a divisor $D = \sum_{w \in T} n_w \cdot (w) \in \mathrm{Div}(T)$ on the complex torus T, we let $f(D) = \prod_w f(w)^{n_w} \in \mathbf{C}$. Prove that for any two elliptic functions f and g with disjoint divisors, we have

$$f((g)) = g((f)).$$

[Hint: write f and g as products of σ-functions.]

20. Let $G_k = \sum_{\omega \in \Lambda'} \omega^{-k}$ be the Eisenstein series of order k, and define $G_2 = G_1 = 0$ and $G_0 = -1$.

a. Show that $(k-1)(k-2)(k-3)G_k = 6\sum_{j=0}^{k}(j-1)(k-j-1)G_jG_{k-j}$ for all $k \geq 6$.
[Hint: $\wp'' = 6\wp^2 - 30G_4$.]

b. Show that $G_8 = \frac{3}{7}G_4^2$, $G_{10} = \frac{5}{11}G_4G_6$ and $G_{12} = \frac{25}{143}G_6^2 + \frac{18}{143}G_4^3$ and that, more generally, every Eisenstein series can be computed recursively from G_4 and G_6 by the formula

$$(k^2-1)(k-6)G_k = 6\sum_{j=4}^{k-4}(j-1)(k-j-1)G_jG_{k-j}.$$

21. Let Λ be a lattice for which $g_2(\Lambda)$ and $g_3(\Lambda)$ are real. Prove that Λ is either a rectangular lattice spanned by a real and a totally imaginary number, or a rhombic lattice spanned by a real number ω_1 and number ω_2 satisfying $\omega_2 + \overline{\omega_2} = \omega_1$. Show that these cases can be distinguished by the sign of $\Delta(\Lambda)$, and that we have group isomorphisms

$$E_\Lambda(\mathbf{R}) \cong \begin{cases} \mathbf{R}/\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z} & \text{for } \Delta(\Lambda) > 0; \\ \mathbf{R}/\mathbf{Z} & \text{for } \Delta(\Lambda) < 0. \end{cases}$$

22. Let $L(nO)$ be the vector space of meromorphic functions on the torus $T = \mathbf{C}/\Lambda$ having a pole of order at most n in O. Prove:

$$\dim_{\mathbf{C}}(L(nO)) = \begin{cases} n & \text{for } n > 0; \\ 1 & \text{for } n = 0. \end{cases}$$

23. *(Riemann-Roch for the torus.)* For a divisor D on the torus T, let $L(D)$ be the vector space consisting of $f = 0$ and the meromorphic functions $f \neq 0$ on T for which the divisor $(f) + D$ is without polar part. Prove:

$$\dim_{\mathbf{C}}(L(D)) = \begin{cases} \deg(D) & \text{for } \deg(D) > 0; \\ 0 & \text{for } \deg(D) < 0. \end{cases}$$

What can you say if D is of degree 0?

3. Complex elliptic curves

We have seen in the previous section that every complex torus $T = \mathbf{C}/\Lambda$ is 'isomorphic' to the elliptic curve E_Λ with Weierstrass equation $y^2 = 4x^3 - g_2(\Lambda)x - g_3(\Lambda)$. The *uniformization theorem* 3.8 in this section states that conversely, every complex Weierstrass equation $y^2 = 4x^3 - g_2x - g_3$ of non-zero discriminant $\Delta = g_2^3 - 27g_3^2$ comes from a torus. This correspondence is actually an *equivalence of categories.* In order to make this into a meaningful statement, we have to define the maps in the categories of complex tori and complex elliptic curves, respectively.

We will first define a set $\mathrm{Hom}(T_1, T_2)$ of maps between complex tori called *isogenies* and study its structure. At the end of this section, we will describe the corresponding algebraic maps between complex Weierstrass curves, which are again called isogenies. These maps will turn out be an important tool in studying the arithmetic of elliptic curves over $\mathbf{Q}$.

3.1. Lemma. *Let $\psi : \mathbf{C}/\Lambda_1 \to \mathbf{C}/\Lambda_2$ be a continuous map between complex tori. Then there exists a continuous map $\phi : \mathbf{C} \to \mathbf{C}$ such that the diagram*

$$\begin{tikzcd}
\mathbf{C} \arrow[r, "\phi"] \arrow[d, "\mathrm{can}"] & \mathbf{C} \arrow[d, "\mathrm{can}"] \\
\mathbf{C}/\Lambda_1 \arrow[r, "\psi"] & \mathbf{C}/\Lambda_2
\end{tikzcd}$$

commutes. The map ϕ is uniquely determined up to an additive constant in Λ_2.

Proof. Choose $\phi(0)$ such that the diagram commutes for $z = 0$. If $z \in \mathbf{C}$ is arbitrary, choose a path $\gamma : 0 \to z$ in $\mathbf{C}$. Let $\overline{\gamma} : \psi(\bar{0}) \to \psi(\bar{z})$ be the path in $\mathbf{C}/\Lambda_2$ obtained by reducing modulo Λ_1 and applying ψ. As the natural map $\mathbf{C} \to \mathbf{C}/\Lambda_2$ is a covering map, $\overline{\gamma}$ can uniquely be lifted under this map to a path in $\mathbf{C}$ starting in $\phi(0)$, and we define $\phi(z)$ as the end point of this map. The value $\phi(z)$ is independent of the choice of the path γ since $\mathbf{C}$ is simply connected, and it is clear that ϕ is continuous. If ϕ' is another map for which the diagram commutes, then their difference $\phi - \phi'$ is a continuous map $\mathbf{C} \to \Lambda_2$, so it is constant. □

If the map ϕ in lemma 3.1 is a holomorphic function, we call ψ an *analytic map* between the tori. An analytic map $\psi : \mathbf{C}/\Lambda_1 \to \mathbf{C}/\Lambda_2$ is called an *isogeny* if it satisfies $\psi(0) = 0$. An analytic map ψ is the composition of the isogeny $\psi - \psi(0)$ with a translation over $\psi(0)$.

3.2. Theorem. *Let $\psi : \mathbf{C}/\Lambda_1 \to \mathbf{C}/\Lambda_2$ be an isogeny. Then there exists $\alpha \in \mathbf{C}$ such that we have*

$$\psi(z \bmod \Lambda_1) = \alpha z \bmod \Lambda_1 \qquad \textit{and} \qquad \alpha\Lambda_1 \subset \Lambda_2.$$

Conversely, every $\alpha \in \mathbf{C}$ satisfying $\alpha\Lambda_1 \subset \Lambda_2$ gives rise to an isogeny $\mathbf{C}/\Lambda_1 \to \mathbf{C}/\Lambda_2$.

Proof. Let $\phi : \mathbf{C} \to \mathbf{C}$ be the lift of ψ satisfying $\phi(0) = 0$. For every $\omega_1 \in \Lambda_1$, the holomorphic function $\phi(z) - \phi(z + \omega_1)$ has values in Λ_2, so it is constant. It follows that $\phi'(z)$ is a holomorphic function with period lattice Λ_1, so by Liouville's theorem it is constant. For ϕ itself we find $\phi(z) = \alpha z$ for some $\alpha \in \mathbf{Z}$. As Λ_1 maps to zero in $\mathbf{C}/\Lambda_2$, we have $\alpha\Lambda_1 \subset \Lambda_2$. Conversely, it is clear that any α of this form induces an isogeny. □

3.3. Corollary. *Every complex isogeny is a homomorphism on the group of points. The set of isogenies* $\mathrm{Hom}(\mathbf{C}/\Lambda_1, \mathbf{C}/\Lambda_2)$ *carries a natural group structure.* □

We say that two complex tori are *isogenous* if there exists a non-zero isogeny between them. Note that a non-zero isogeny is always surjective.

Exercise 1. Take $\Lambda_1 = \mathbf{Z} + \mathbf{Z}\,i$ and $\Lambda_2 = \mathbf{Z} + \mathbf{Z}\,i\pi$. Prove: $\mathrm{Hom}(\mathbf{C}/\Lambda_1, \mathbf{C}/\Lambda_2) = 0$.

For a non-zero isogeny $\psi : T_1 = \mathbf{C}/\Lambda_1 \to T_2 = \mathbf{C}/\Lambda_2$, we define the *degree* of ψ as

$$\deg(\psi) = \#\ker\psi = \#[(\alpha^{-1}\Lambda_2) \bmod \Lambda_1] = [\Lambda_2 : \alpha\Lambda_1].$$

The degree of the zero isogeny is by definition equal to 0.

For ψ of degree $n > 0$, we have inclusions of lattices $n\Lambda_2 \overset{n}{\subset} \alpha\Lambda_1 \overset{n}{\subset} \Lambda_2$. This shows that multiplication by n/α maps Λ_2 to a lattice of index n in Λ_1. The corresponding isogeny $\widehat{\psi} : T_2 \to T_1$ is the *dual isogeny* corresponding to ψ. Note that $\widehat{\psi} \circ \psi$ and $\psi \circ \widehat{\psi}$ are multiplication by n on T_1 and T_2, respectively.

Exercise 2. Show that being isogenous is an equivalence relation on the set of complex tori, and that there are uncountably many isogeny classes of complex tori.

Two complex tori $\mathbf{C}/\Lambda_1$ and $\mathbf{C}/\Lambda_2$ are isomorphic if there is an invertible isogeny between them, i.e., an isogeny of degree 1. This happens if and only if $\Lambda_2 = \alpha\Lambda_1$ for some $\alpha \in \mathbf{C}^*$. In that case we say that Λ_1 and Λ_2 are isomorphic or *homothetic*. For homothetic lattices Λ_1 and Λ_2 we have $g_2(\Lambda_2) = \alpha^{-4}g_2(\Lambda_1)$ and $g_3(\Lambda_2) = \alpha^{-6}g_3(\Lambda_1)$ for some α, so the *j-invariant*

$$j(\Lambda) = 1728\frac{g_2(\Lambda)^3}{g_2(\Lambda)^3 - 27g_3(\Lambda)^2} = 1728\frac{g_2(\Lambda)^3}{\Delta(\Lambda)}$$

of a lattice is defined on isomorphism classes of lattices. Note that $j(\Lambda)$ is well-defined since $\Delta(\Lambda)$ does not vanish. The factor $1728 = 12^3$ is traditional; it is related to the Fourier expansion of the j-function.

3.4. Lemma. *Two lattices are homothetic if and only if their j-invariants coincide.*

Proof. We still need to show that the equality $j(\Lambda_1) = j(\Lambda_2)$ implies that Λ_1 and Λ_2 are homothetic. From the equality $j(\Lambda_1) = j(\Lambda_2)$ we easily derive that there exists $\alpha \in \mathbf{C}^*$ such that we have $g_2(\Lambda_2) = \alpha^{-4}g_2(\Lambda_1)$ and $g_3(\Lambda_2) = \alpha^{-6}g_3(\Lambda_1)$. Then Λ_2 and $\alpha\Lambda_1$ have the same values of g_2 and g_3, so the $\wp$-functions $\wp_{\Lambda_2}$ and $\wp_{\alpha\Lambda_1}$ coincide. In particular, their sets of poles Λ_2 and $\alpha\Lambda_1$ coincide. □

Every lattice $\Lambda = [\omega_1, \omega_2]$ is homothetic to a lattice $[1, z]$ with $z = \omega_2/\omega_1$ in the complex upper half plane, so we can view j as a function $j : \mathcal{H} \to \mathbf{C}$. The Eisenstein series $G_k(z) = G_k([1, z])$ are holomorphic on $\mathcal{H}$ by 2.4, so j is again a holomorphic function on $\mathcal{H}$.

Two lattices $[1, z_1]$ and $[1, z_2]$ are homothetic if and only if we have $z_2 = \frac{az_1+b}{cz_1+d}$ for some matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbf{Z})$. The identity

$$\text{(3.5)} \qquad \mathrm{Im}\left(\frac{az+b}{cz+d}\right) = (ad-bc)\frac{\mathrm{Im}(z)}{|cz+d|^2}$$

shows that only the matrices in $\mathrm{SL}_2(\mathbf{Z})$ map $\mathcal{H}$ to itself. We conclude that $j : \mathcal{H} \to \mathbf{C}$ is constant on $\mathrm{SL}_2(\mathbf{Z})$-orbits, and that the induced function $j : \mathrm{SL}_2(\mathbf{Z}) \setminus \mathcal{H} \to \mathbf{C}$ on the orbit space is injective.

3.6. Theorem. *The map $j : \mathrm{SL}_2(\mathbf{Z}) \setminus \mathcal{H} \to \mathbf{C}$ is a bijection.*

The main ingredient in the proof of 3.6 is the construction of a fundamental domain for the action of $\mathrm{SL}_2(\mathbf{Z})$ on $\mathcal{H}$. The following statement is sufficient for our purposes.

3.7. Lemma. *Every $\mathrm{SL}_2(\mathbf{Z})$-orbit in $\mathcal{H}$ has a representative in the set*

$$D = \{z \in \mathcal{H} : |z| \geq 1 \quad \text{and} \quad -1/2 \leq \mathrm{Re}(z) < 1/2\}.$$

Proof. Pick $z \in \mathcal{H}$. As the elements $cz + d$ with $c, d \in \mathbf{Z}$ form a lattice in $\mathbf{C}$, the numerator $|cz + d|^2$ in (3.5) is bounded from below, so there exists an element z_0 in the orbit of z for which $\mathrm{Im}(z)$ is maximal. Applying a translation matrix $\left(\begin{smallmatrix} 1 & k \\ 0 & 1 \end{smallmatrix}\right)$ mapping z_0 to $z_0 + k$ when necessary, we may assume that $\mathrm{Re}(z_0)$ is in $[-1/2, 1/2)$. From the inequality $\mathrm{Im}(-1/z_0) = |z_0|^{-2}\mathrm{Im}(z_0) \leq \mathrm{Im}(z_0)$ we find $|z_0| \geq 1$, so z_0 is in D. □

Exercise 3. Find a representative in D for the $\mathrm{SL}_2(\mathbf{Z})$ orbit of $\frac{1+2i}{100}$.

Proof of 3.6. It remains to show that the image $j[\mathcal{H}]$ of the j-function is all of $\mathbf{C}$. As j is a non-constant holomorphic function on $\mathcal{H}$, its image $j[\mathcal{H}]$ is open in $\mathbf{C}$. We will show that $j[\mathcal{H}]$ is also closed in $\mathbf{C}$. By the connectedness of $\mathbf{C}$, this proves what we want.

Let $j = \lim_{n\to\infty} j(z_n)$ be a limit point of $j[\mathcal{H}]$ in $\mathbf{C}$. By picking the z_n suitably inside their $\mathrm{SL}_2(\mathbf{Z})$-orbit, we may assume that all z_n lie in D. If the values of $\mathrm{Im}(z_n)$ remain bounded, the sequence $\{z_n\}_n$ lies in a bounded subset of D, and we can pick any limit point $z \in \mathcal{H}$ of the sequence to find $j(z) = j \in j[\mathcal{H}]$.

If the values of $\mathrm{Im}(z_n)$ are not bounded, we can pass to a subsequence and assume $\lim_{n\to\infty} \mathrm{Im}(z_n) = +\infty$. From the definition of g_2 and g_3 in theorem 2.9 we now find

$$\lim_{n\to\infty} g_2(z_n) = 60 \cdot 2 \sum_{m=1}^{\infty} \frac{1}{m^4} = \frac{4\pi^4}{3} \qquad \text{and} \qquad \lim_{n\to\infty} g_3(z_n) = 140 \cdot 2 \sum_{m=1}^{\infty} \frac{1}{m^6} = \frac{8\pi^6}{27},$$

so $\Delta(z_n) = g_2(z_n)^3 - 27g_3(z_n)^2$ tends to 0. This implies $\lim_{n\to\infty} |j(z_n)| = +\infty$, contradicting the assumption that $j(z_n)$ converges. □

The main corollary of 3.6 is the following theorem. It enables us to translate many statements over complex elliptic curves into analytic facts.

3.8. Uniformization theorem. *Given any two integers $g_2, g_3 \in \mathbf{C}$ with $g_2^3 - 27g_2^2 \neq 0$, there exists a lattice $\Lambda \subset \mathbf{C}$ with $g_2(\Lambda) = g_2$ and $g_3(\Lambda) = g_3$. In particular, every complex elliptic curve comes from a complex torus in the sense of* 2.7.

Proof. Pick a lattice Λ with j-invariant $j(\Lambda) = g_2^3/(g_2^3 - 27g_3^2)$. As in the proof of 3.4, we find that there exists $\alpha \in \mathbf{C}$ satisfying $g_2(\Lambda) = \alpha^4 g_2$ and $g_3(\Lambda) = \alpha^6 g_3$. Now the lattice $\alpha\Lambda$ does what we want. □

A complex elliptic curve E in Weierstrass form, or briefly *Weierstrass curve*, can be specified as a pair (g_2, g_3) of coefficients in the corresponding equation $y^2 = 4x^3 - g_2 x - g_3$. We require that the discriminant $\Delta(E) = g_2^3 - 27g_2^2$ does not vanish and define the j-invariant of E as $j(E) = 1728\, g_2^3/\Delta(E)$. Weierstrass curves are said to be *isomorphic* if their j-invariants coincide. As we have already seen, Weierstrass curves with coefficients (g_2, g_3) and (g_2', g_3') are isomorphic if and only if there exists $\alpha \in \mathbf{C}$ satisfying $g_2' = \alpha^4 g_2$ and $g_3' = \alpha^6 g_3$.

Exercise 4. Show that a Weierstrass curve E is isomorphic to a Weierstrass curve defined over $\mathbf{Q}(j(E))$.

An *isogeny* between Weierstrass curves is for us simply a map coming from an isogeny between the corresponding complex tori. Its degree is the degree of the corresponding isogeny between tori. With this definition, the categories of complex tori and the category of Weierstrass curves, each with the isogenies as their morphisms, become equivalent in view of 3.8.

Our definition of an isogeny $\psi : E \to \widetilde{E}$ between curves parametrized by $\mathbf{C}/\Lambda$ and $\mathbf{C}/\widetilde{\Lambda}$ means that ψ fits in a commutative diagram

$$\begin{tikzcd}
\mathbf{C}/\Lambda \arrow[r, "z \mapsto \alpha z"] \arrow[d, "W"] & \mathbf{C}/\widetilde{\Lambda} \arrow[d, "\widetilde{W}"] \\
E \arrow[r, "\psi"] & \widetilde{E}.
\end{tikzcd}$$

Here W and $\widetilde{W}$ denote the Weierstrass parametrizations, and $\alpha \in \mathbf{C}$ satisfies $\alpha\Lambda \subset \widetilde{\Lambda}$. We see that ψ can be described in terms of Weierstrass $\wp$-functions as

$$\psi : (\wp(z), \wp'(z)) \longmapsto (\widetilde{\wp}(\alpha z), \widetilde{\wp}'(\alpha z)).$$

As $z \mapsto \widetilde{\wp}(\alpha z)$ and $z \mapsto \widetilde{\wp}'(\alpha z)$ are elliptic functions on $\mathbf{C}/\Lambda$, they are rational expressions in $\wp(z)$ and $\wp'(z)$. Thus ψ is actually an *algebraic map* $E \to \widetilde{E}$ that is everywhere defined. It is a *morphism* of curves in the sense of algebraic geometry.

3.9. Theorem. *Let $\psi : E \to \widetilde{E}$ be an isogeny of degree $n > 0$ between Weierstrass curves. Then there exist $\alpha \in \mathbf{C}$ and monic coprime polynomials $A, B \in \mathbf{C}[X]$ of degree n and $n-1$, respectively, such that ψ is given on the affine points of E by the algebraic map*

$$\psi : (x, y) \longmapsto \left(\frac{A(x)}{\alpha^2 B(x)}, \frac{A'(x)B(x) - A(x)B'(x)}{\alpha^3 B(x)^2} y \right).$$

Proof. We may suppose that ψ corresponds to a diagram as above. As $\wp(\alpha z)$ is even and periodic modulo Λ, there exist $c \in \mathbf{C}$ and monic coprime polynomials $A, B \in \mathbf{C}[X]$, say of degree a and b, for which we have the identity

$$\widetilde{\wp}(\alpha z) = c\,\frac{A(\wp(z))}{B(\wp(z))}.$$

Comparison of the orders and leading coefficients of the poles of these functions in $z = 0$ yields equalities $c = \alpha^{-2}$ and $2 = 2a - 2b$; in particular we have $a = b + 1$. Now consider the commutative diagram

$$\begin{tikzcd} \mathbf{C}/\Lambda \arrow[r, "\psi"] \arrow[d, "\wp"] & \mathbf{C}/\widetilde{\Lambda} \arrow[d, "\widetilde{\wp}"] \\ \mathbf{P}^1(\mathbf{C}) \arrow[r, "\psi_x"] & \mathbf{P}^1(\mathbf{C}), \end{tikzcd}$$

in which we write ψ again for the isogeny between tori corresponding to ψ. By definition of the degree, ψ is n to 1. The vertical maps are *generically* 2 to 1, meaning that for all but finitely many $x \in \mathbf{P}^1(\mathbf{C})$, the fibers $\wp^{-1}(x)$ and $\widetilde{\wp}^{-1}(x)$ consist of 2 elements. This implies that the composition $\widetilde{\wp} \circ \psi$ is generically $2n$ to 1, and consequently the map $\psi_x : \mathbf{P}^1(\mathbf{C}) \to \mathbf{P}^1(\mathbf{C})$, which maps x to $A(x)/(\alpha^2 B(x))$, is generically n to 1. This easily yields $a = n$, as desired. Differentiation of the identity for $\widetilde{\wp}(\alpha z)$ yields the value of the y-coordinate of ψ. □

Exercise 5. Let $A, B \in \mathbf{C}[X]$ be coprime polynomials of degree a and b. Show that the map on $\mathbf{P}^1(\mathbf{C})$ defined by $x \mapsto A(x)/B(x)$ is generically $\max(a, b)$ to 1.

3.10. Example. Let $\Lambda = [\omega_1, \omega_2]$ be any lattice, and define $\widetilde{\Lambda} = [\frac{1}{2}\omega_1, \omega_2]$. Then Λ is of index 2 in $\widetilde{\Lambda}$, and the natural map $T = \mathbf{C}/\Lambda \to \widetilde{T} = \mathbf{C}/\widetilde{\Lambda}$ is an isogeny of degree 2. Its kernel is generated by the 2-torsion element $\frac{1}{2}\omega_1 \in \mathbf{C}/\Lambda$. On the associated Weierstrass curve $E : y^2 = 4x^3 - g_2 x - g_3$, this corresponds to a point of the form $(a, 0)$. The equation can be written correspondingly as $y^2 = (x - a)(4x^2 + 4ax + \frac{g_3}{a})$.

In order to find the polynomials A and B from 3.9 in this case, we have to express the Weierstrass function $\widetilde{\wp}(z)$ associated to $\widetilde{\Lambda}$ as a rational function in the Weierstrass function $\wp(z)$ associated to Λ. From exercise 2.16, we have the useful identity

$$\widetilde{\wp}(z) = \wp(z) + \wp(z + \tfrac{1}{2}\omega_1) - \wp(\tfrac{1}{2}\omega_1).$$

It is now straightforward from the addition formula (2.12) to evaluate

$$\widetilde{\wp} = -2a + \frac{\wp'^2}{4(\wp - a)^2} = -2a + \frac{\wp^2 + a\wp + \frac{g_3}{4a}}{\wp - a} = \frac{\wp^2 - a\wp + 2a^2 + \frac{g_3}{4a}}{\wp - a}.$$

As expected, A and B are monic of degrees 2 and 1. Rewriting $\frac{g_3}{4a} = a^2 - \frac{g_2}{4}$, we can write the complete isogeny in algebraic terms as

$$(x, y) \longmapsto (\widetilde{x}, \widetilde{y}) = \left(x + \frac{12a^2 - g_2}{4(x - a)}, \left(1 - \frac{12a^2 - g_2}{4(x - a)^2}\right)y\right).$$

We refer to the exercises for a proof that $(\widetilde{x}, \widetilde{y})$ is a point on the Weierstrass curve $\widetilde{E}$ with equation $y^2 = 4(x + 2a)(x^2 - 2ax + g_2 - 11a^2)$.

Exercise 6. Show that the isogeny in 3.10 is given by $(x, y) \mapsto (x + x_T - a, y + y_T)$, where $(x_T, y_T) = (x, y) + (a, 0)$ in the group $E(\mathbf{C})$.

Theorem 3.9 shows that isogenies between elliptic curves, which we defined originally as *analytic maps* between tori, turn out be *algebraic maps*, i.e., given by rational functions in the coordinates. Conversely, one can show that all algebraic maps between Weierstrass curves are analytic, so that algebraic and analytic maps come down to the same thing. This equivalence is a simple example of a 'GAGA-phenomenon', an abbreviation referring to a 1956 paper of Serre, *Géométrie algébrique et géométrie analytique*, which is devoted to similar equivalences.

An even simpler example of the phenomenon indicated above is the classification in theorem 2.7 of the meromorphic functions on a torus T. Such functions, which are by definition analytic maps $T \to \mathbf{P}^1(\mathbf{C})$, turn out to be rational functions in the coordinates when viewed as maps on the associated Weierstrass curve. The function field $\mathcal{M}(T) = \mathbf{C}(\wp, \wp')$ of T is therefore isomorphic to the *function field* $\mathcal{M}(E)$ of rational functions in the affine coordinates on E. This field is usually defined as the field of fractions of the *coordinate ring* $\mathbf{C}[x, y]/(y^2 - 4x^3 + g_2x + g_3)$, which is the ring of polynomial functions on the affine part of E. From an algebraic point of view, $\mathcal{M}(E)$ is a quadratic extension $\mathbf{C}(x, \sqrt{4x^3 - g_2x - g_3})$ of the rational function field $\mathbf{C}(x)$.

The function field $\mathcal{M}(\mathbf{P}^1(\mathbf{C}))$ of meromorphic functions on the Riemann sphere is also algebraic: it is the rational function field $\mathbf{C}(x)$.

Every isogeny $\psi : T \to \widetilde{T}$ between complex tori induces a map $\psi^* : \mathcal{M}(\widetilde{T}) \to \mathcal{M}(T)$ in the opposite direction mapping an elliptic function $f \in \mathcal{M}(\widetilde{T})$ to $f \circ \psi$. If ψ is non-zero, this is an injective homomorphism of fields.

3.11. Theorem. *Let $\psi : T \to \widetilde{T}$ be an isogeny of degree $n > 0$. Then the field extension $\psi^*[\mathcal{M}(\widetilde{T})] \subset \mathcal{M}(T)$ is an algebraic extension of degree n.*

Proof. As $\mathcal{M}(T)$ and $\mathcal{M}(\widetilde{T})$ are quadratic extensions of $\mathbf{C}(\wp)$ and $\mathbf{C}(\widetilde{\wp})$, respectively, it suffices to show that $\mathbf{C}(\wp)$ is algebraic of degree n over $\psi^*[\mathbf{C}(\widetilde{\wp})]$. In view of 3.9, this follows from the following lemma. □

3.12. Lemma. *Let $A, B \in \mathbf{C}[X]$ be coprime polynomials of degree a and b. If A and B are not both constant, then $\mathbf{C}(x)$ is an algebraic of degree $\max(a, b)$ of $\mathbf{C}(\frac{A(x)}{B(x)})$.*

Proof. Write $Y = \frac{A(x)}{B(x)}$, then x is a zero of the polynomial $F = A(X) - YB(X) \in \mathbf{C}[X, Y]$ of degree $\max(a, b)$ in X with coefficients in $\mathbf{C}(Y)$.It remains to show that F is irreducible. As F is of degree 1 in Y, it can only be reducible if there is a polynomial in $\mathbf{C}[X] \setminus \mathbf{C}$ dividing it; this is excluded by the coprimality assumption on A and B. □

It is a general fact from algebraic geometry that degrees of maps can be read off from the degrees of the corresponding function field extension. Over $\mathbf{C}$ or $\overline{\mathbf{Q}}$, the degree of a map is the cardinality of all but finitely many fibers.

Exercise 7. Check this fact for the projections π_x and π_y of a Weierstrass curve E on the axes. *Can you generalize the argument to arbitrary rational functions $E \to \mathbf{P}^1(\mathbf{C})$?

Exercises.

8. The *multiplicator ring* of a lattice Λ is defined as $\mathcal{O} = \mathcal{O}(\Lambda) = \{\alpha \in \mathbf{C} : \alpha\Lambda \subset \Lambda\}$. Show that $\mathcal{O}$ is a subring of $\mathbf{C}$ isomorphic to the endomorphism ring $\mathrm{End}(\mathbf{C}/\Lambda)$ of the torus $\mathbf{C}/\Lambda$. Show also that we have $\mathcal{O}(\Lambda) = \mathbf{Z}$ unless Λ is homothetic to a lattice of the form $[1, \omega]$, with $\omega \in \mathbf{C} \setminus \mathbf{R}$ the zero of an irreducible quadratic polynomial $aX^2 + bX + c \in \mathbf{Z}[X]$, and that in this exceptional case we have $\mathcal{O}(\Lambda) = \mathbf{Z}[\frac{D+\sqrt{D}}{2}]$ with $D = b^2 - 4ac < 0$.
[In the exceptional case, we say that $\mathbf{C}/\Lambda$ has *complex multiplication* by $\mathcal{O}$.]

9. Show that the subrings of $\mathbf{C}$ that are lattices correspond bijectively to the set of negative integers $D \equiv 0, 1 \bmod 4$ under the map $D \mapsto \mathcal{O}(D) = \mathbf{Z}[\frac{D+\sqrt{D}}{2}]$. Show that there exists a ring homomorphism $\mathcal{O}(D_1) \to \mathcal{O}(D_2)$ if and only if D_1/D_2 is a square in $\mathbf{Z}$.
[One calls $\mathcal{O}(D)$ the *quadratic order of discriminant* D.]

*10. Show that the isomorphism classes of complex tori with complex multiplication by $\mathcal{O}$ correspond bijectively to the elements of the Picard group $\mathrm{Pic}(\mathcal{O})$ of $\mathcal{O}$.

11. Show that the degree map $\deg : \mathrm{End}(\mathbf{C}/\Lambda) \to \mathbf{Z}$ is a multiplicative function, and that there is a commutative diagram

$$\begin{tikzcd}
\mathrm{End}(\mathbf{C}/\Lambda) \arrow[r, "\sim"] \arrow[d, "\deg"] & \mathcal{O}(\Lambda) \arrow[r, phantom, "\subset"] \arrow[d, "z\mapsto z\bar{z}"] & \mathbf{C} \\
\mathbf{Z} \arrow[r, "\mathrm{id}"] & \mathbf{Z} \arrow[r, phantom, "\subset"] & \mathbf{R}.
\end{tikzcd}$$

12. Compute the structure of the group $\mathrm{Hom}(\mathbf{C}/\Lambda_1, \mathbf{C}/\Lambda_2)$ for each of the following choices of Λ_1 and Λ_2:
   a. $\Lambda_1 = \Lambda_2 = \mathbf{Z} + \mathbf{Z}\,i$;
   b. $\Lambda_1 = \mathbf{Z} + \mathbf{Z}\,i$ and $\Lambda_2 = \mathbf{Z} + \mathbf{Z}\,2i$;
   b. $\Lambda_1 = \mathbf{Z} + \mathbf{Z}\,i$ and $\Lambda_2 = \mathbf{Z} + \mathbf{Z}\sqrt{-2}$.

13. Show that every group $\mathrm{Hom}(\mathbf{C}/\Lambda_1, \mathbf{C}/\Lambda_2)$ is a free abelian group of rank at most 2. Show that the rank is non-zero if $\mathbf{C}/\Lambda_1$ and $\mathbf{C}/\Lambda_2$ are isogenous, and that it is 2 if and only if $\mathbf{C}/\Lambda_1$ and $\mathbf{C}/\Lambda_2$ have complex multiplication by rings $\mathcal{O}_1$ and $\mathcal{O}_2$ having the same field of fractions.

14. A non-zero isogeny $\psi : T_1 \to T_2$ is said to be *cyclic* if $\ker\psi$ is a cyclic subgroup of T_1. Show that complex tori are isogenous if and only if there exists a cyclic isogeny between them. Show also that a torus admitting a cyclic endomorphism (different from the identity) has complex multiplication.

15. Show that the set $D \subset \mathcal{H}$ in 3.7 contains a *unique* representative of every $\mathrm{SL}_2(\mathbf{Z})$-orbit if we remove the elements on its boundary satisfying $|z| = 1$ and $\mathrm{Re}(z) > 0$.

*16. Let $f : E \to \widetilde{E}$ be a rational map between Weierstrass curves, i.e., a map of the form $(x, y) \mapsto (f_1(x, y), f_2(x, y))$ for functions $f_1, f_2 \in \mathcal{M}(E)$ with the property that the image of (x, y) lies in $\widetilde{E}(\mathbf{C})$ whenever it is defined. Show that f can be defined on all points of E, and that it corresponds to an analytic map of the corresponding tori.

17. Determine the Weierstrass polynomial $W(X)$ of the curve $\widetilde{E}$ in example 3.10 by proving the following statements.
    a. $W(X) = 4(X - \widetilde{\wp}(\frac{1}{2}\omega_2))(X - \widetilde{\wp}(\frac{1}{4}\omega_1))(X - \widetilde{\wp}(\frac{1}{4}\omega_1 + \frac{1}{2}\omega_2))$.
    b. We have $\widetilde{\wp}(\frac{1}{2}\omega_2)) = -2a$.
    c. The function $4(\wp(z) - a)(\wp(z + \frac{1}{2}\omega_1) - a)$ is constant with value $12a^2 - g_2$.
    d. We have $(\widetilde{\wp}(\frac{1}{4}\omega_1) - a)^2 = 4(\wp(\frac{1}{4}\omega_1) - a)^2 = 12a^2 - g_2$.
    e. $W(X) = 4(X + 2a)((X - a)^2 + g_2 - 12a^2) = 4(X + 2a)(X^2 - 2aX + g_2 - 11a^2)$.

18. Show that after a linear change of variables $X = 4(x - a)$ and $Y = 4y$, the equation of the Weierstrass curves E in 3.10 becomes $Y^2 = X(X^2 + \alpha X + \beta)$ with $\alpha = 12a$ and $\beta = 48a^2 - 4g_2$. Show that a similar change of variables then reduces the 2-isogenous curve to the form $Y^2 = X(X^2 - 2\alpha X + \alpha^2 - 4\beta)$.

4. Elliptic curves over fields

By analytic means we saw in the previous section that a complex torus is a plane cubic curve. It turned out that the group structure on the torus had the following interpretation on the curve: we have $P+Q+R=0$ if and only if P, Q, and R are collinear. We will show later that every plane cubic curve over $\mathbf{C}$ arises from such a torus, so that indeed every cubic curve has such a group structure. It turns out that this is a very general fact that has little to do with $\mathbf{C}$ or with analysis. For instance, one can show the same for plane cubic curves over a field of characteristic $p>0$. In order to state and prove precise results one needs to develop some algebraic geometry.

This is not a course on algebraic geometry, so it will not be our goal to define the various notions in a very general setting. Our curves will always be embedded in a plane, and we will postpone talking about morphisms for a while. The one book that dominates the subject of algebraic geometry is Hartshorne's *Algebraic Geometry* (Springer GTM 52). Silverman's book on elliptic curves starts with two chapters of basics from algebraic geometry, but he does not include proofs of all theorems.

4.1. Projective space. Let k be a field and let n be a non-negative integer. Affine n-space $\mathbf{A}^n(k)$ is the set k^n. Projective n-space $\mathbf{P}^n(k)$ is the set of 1-dimensional subspaces of the k-vector space k^{n+1}. More explicitly,

$$\mathbf{P}^n(k) = (k^{n+1}\backslash\{(0,\ldots,0)\})/\sim$$

where

$$x\sim y \iff \exists\lambda\in k^* \quad x=\lambda y.$$

We denote the equivalence class of a vector $(x_0,\ldots,x_{n+1})\in k^{n+1}$ by $(x_0:\cdots:x_{n+1})\in\mathbf{P}^n(k)$.

We have only defined $\mathbf{A}^n$ and $\mathbf{P}^n$ by giving a certain set for each field k. In fact, $\mathbf{A}^n$ and $\mathbf{P}^n$ are algebraic varieties. The precise meaning of this statement is in the subject of algebraic geometry; we will not define the category of algebraic varieties here, but we will work with ad-hoc definitions that are sufficient for our number theoretic purposes.

The group $\mathrm{GL}_{n+1}(k)$ of k-linear automorphisms of k^{n+1} acts on $\mathbf{P}^n(k)$. Of course the scalar matrices act trivially, so the projective linear group $\mathrm{PGL}_{n+1}(k)=\mathrm{GL}_{n+1}(k)/k^*$ acts on $\mathbf{P}^n(k)$.

We have an embedding $\mathbf{A}^n\to\mathbf{P}^n$ defined by $(x_1,\ldots,x_n)\mapsto(x_1:\cdots:x_n:1)$. Projective space thus consists of an affine part and a copy of $\mathbf{P}^{n-1}$ which is “at infinity,” namely the points whose last coordinate is 0.

4.2. Plane curves. For us, a plane curve C over a field k will be a non-zero homogeneous polynomial F of degree $d>0$ in X,Y,Z with coefficients in k. We then say that the curve is given by the equation

$$C\colon\qquad F(X,Y,Z)=0.$$

For two curves C and C' given by polynomials F and F' we say that $C' \subset C$ if F' divides F in $k[X, Y, Z]$. We say that $C' = C$ if F and F' differ by a non-zero scalar. By the set of points of C we mean

$$C(k) = \{(a\colon b\colon c) \in \mathbf{P}^2(k) : F(a, b, c) = 0\}.$$

We have $F(\lambda a, \lambda b, \lambda c) = \lambda^d F(a, b, c)$, so either all or none of the representatives in k^3 of a point in $\mathbf{P}^2(k)$ lie in the zero-set of F. In the next exercise it is shown that this equation characterizes homogeneous polynomials if k is not a finite field.

Exercise 1. Suppose that k is an infinite field and that d is a positive integer. Show that a polynomial $F \in k[X, Y, Z]$ is homogeneous of degree d if and only if for all $\lambda, a, b, c \in k$ we have $F(\lambda a, \lambda b, \lambda c) = \lambda^d F(a, b, c)$. [*Hint:* Show first that a polynomial in X, Y, Z which is zero as a function $k^3 \to k$ is zero as a polynomial.]

The affine part of the curve is given by the non-homogeneous equation $F(X, Y, 1) = 0$. Conversely, a non-homogeneous polynomial $f(X, Y)$ of degree d can be made homogeneous of degree d by sticking in the correct power of Z in each term. In other words, the homogeneous form of f is $F(X, Y, Z) = Z^d f(X/Z, Y/Z)$. When we write down a non-homogeneous equation of a curve, we mean the curve given by this homogenized polynomial in X, Y and Z.

Let us stress that a curve is much more then its set of points: it is the equation up to a scalar. A curve over a field is also a curve over every extension field. For instance, the curves $X^2 + Y^2 + 1 = 0$ and $X^4 + Y^4 + 1 = 0$ are both curves over $\mathbf{R}$ without any points, but their sets of $\mathbf{C}$-points are different (even topologically).

The easiest example of a curve is a line:

$$L\colon \qquad aX + bY + cZ = 0 \qquad\qquad (a, b, c) \neq (0, 0, 0).$$

The projective plane has the following properties: two distinct points have a unique line passing through them, and two distinct lines (in the sense of the last paragraph!) intersect in a unique point. The proof of these statements is basic linear algebra over k.

Exercise 2. Give this proof in detail.

More generally, **Bezout's theorem** tells us that two curves of degree n and m, for which the defining equations have no common factor, intersect in exactly nm points. There are however three subtleties to keep in mind:

- One has to work in the projective plane because some points may be "at infinity" (just as when we intersect two lines).
- Some intersection points might only be visible over a larger field, so we have to assume that k is algebraically closed.
- We have to count the points with a suitable "multiplicity".

For example, the curve $Y = X^2$ over $k = \mathbf{Q}$ intersects the line $Y = 4$ in $(2,4)$ and $(-2,4)$. But it intersects the line $Y = 2$ in $(\sqrt{2},2)$ and $(-\sqrt{2},2)$, and it intersects the line $Y = 0$ in $(0,0)$ with multiplicity 2.

Of course, one needs to give a precise definition of this intersection multiplicity. If the two curves are defined by the affine equations $f(X,Y) = 0$ and $g(X,Y) = 0$ and the point $(0,0)$ lies on both, then the intersection multiplicity at that point is the k-dimension of $k[[X,Y]]/(f,g)$. With projective linear transformations one can extend this definition to other points than $(0,0)$, but we will not pursue this further. Instead we will work with a definition that only works for intersecting a curve and a line, and for now we will only prove Bezout in that case. There are elementary expositions of the general case of Bezout's theorem in Silverman-Tate, and Knapp.

4.3. Divisor of a form on $\mathbf{P}^1$. Let k be an algebraically closed field, and let $F \in k[U,V]$ be a non-zero homogeneous polynomial of degree d. We claim that F is a product of d linear homogeneous polynomials: $F = \prod_{i=1}^{d}(a_iU + b_iV)$, and that the points $(a_1{:}\, b_1), \ldots, (a_d{:}\, b_d)$ in $\mathbf{P}^1(k)$ are unique up to order. To see this, write $F = V^kF_0$ with $V \nmid F_0$, and use that $F_0(U,1) \in k[U]$ factorizes uniquely as $c\prod_{i=1}^{d-k}(U - u_i)$. (It is here that we use that k is algebraically closed.) We now let $\operatorname{div}(F)$ be the formal sum of the points $(b_i{:}\, -a_i)$, which is an element of the free abelian group $\operatorname{Div}(\mathbf{P}^1(k))$ with $\mathbf{P}^1(k)$ as a generating set. We define the order $\operatorname{ord}_P(F) \in \mathbf{Z}$ of F at P by $\operatorname{div}(F) = \sum_P \operatorname{ord}_P(F)[P]$. We see that for $P = (x{:}\, y)$ we have $\operatorname{ord}_P(F) \geq 0$ and

$$\operatorname{ord}_P(F) \geq 1 \quad\Longleftrightarrow\quad F(x,y) = 0.$$

Thus, $\operatorname{div}(F)$ keeps track of the zeroes of F, and it also keeps track of multiplicities. The degree of $\operatorname{div}(F)$ is d. In other words, $\sum_{P\in\mathbf{P}^1(k)} \operatorname{ord}_P(F) = d$.

This construction is natural in the sense that a projective linear coordinate change on $\mathbf{P}^1(k)$ respects the construction of $\operatorname{div}(F)$. More precisely, for $\sigma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(k)$ we let $(\sigma^*F)(U,V) = F(aU + bV, cU + dV)$. Then we have $\sigma(\operatorname{div}(\sigma^*(F))) = \operatorname{div}(F)$. (Here we again write σ for the homomorphism $\operatorname{Div}(\mathbf{P}^1(k)) \to \operatorname{Div}(\mathbf{P}^1(k))$ that sends $[P]$ to $[\sigma P]$.)

4.4. The intersection divisor of a line and a curve. Let C be a curve given by a homogeneous equation $F(X,Y,Z) = 0$ of degree d. We still assume that the field k is algebraically closed. Let L be a line whose defining equation does not divide F. The line L comes from a 2-dimensional vector space in k^3, for which we choose a basis $(v_1,v_2,v_3), (w_1,w_2,w_3)$. We now have a bijection

$$\phi\colon\quad \mathbf{P}^1(k) \xrightarrow{\sim} L(k) \qquad\qquad (\lambda{:}\, \mu) \mapsto (\lambda v_1 + \mu w_1{:}\, \lambda v_2 + \mu w_2{:}\, \lambda v_3 + \mu w_3).$$

We let the "pull-back" of F to $\mathbf{P}^1(k)$ be the homogeneous polynomial $\phi^*F \in k[U,V]$ given by

$$(\phi^*F)(U,V) = F(Uv_1 + Vw_1, Uv_2 + Vw_2, Uv_3 + Vw_3).$$

Note that ϕ^*F is homogeneous of degree d. Our assumption that the equation of L does not divide F ensures that ϕ^*F is not zero. We now define the *intersection divisor* $L \cdot C$ to be the image under ϕ of $\operatorname{div}(\phi^*F)$. By considering the $\mathrm{PGL}_2(k)$-action on $\mathbf{P}^1(k)$ one sees that $L \cdot C$ does not depend on the choice of the basis $(v_1, v_2, v_3), (w_1, w_2, w_3)$.

Just as the notion of the order of a function on $\mathbf{P}^1(k)$ at a point P behaves well under the $\mathrm{PGL}_2(k)$-action, one can show the following property: if $\sigma \in \mathrm{PGL}_3(k)$ then $\sigma(L) \cdot \sigma(C) = \sigma(L \cdot C)$. Thus forming the intersection divisor is stable under projective linear transformations of the projective plane.

Exercise 3. Make this statement precise and prove it. What is the correct definition of $\sigma(C)$?

4.5. Intersection multiplicities. Now let k be any field, and again let C be a curve given by a homogeneous equation $F(X, Y, Z) = 0$ of degree d. For a line L and $P \in \mathbf{P}^2(k)$ we define the *intersection multiplicity* $i(L, C; P) \in \mathbf{Z}_{\geq 0} \cup \{\infty\}$ as follows. If $L \subset C$ then we put $i(L, C; P) = \infty$ for $P \in L(k)$ and $i(L, C; P) = 0$ for $P \notin L(k)$. If $L \not\subset C$ the we let $i(L, C; P)$ be the multiplicity with which P occurs in the intersection divisor $L \cdot C$. If L is the line through two distinct points $P = (p_0 \colon p_1 \colon p_2)$ and $Q = (q_0 \colon q_1 \colon q_2)$, then we have

$$i(L, C; P) = \operatorname{ord}_T F(p_0 + Tq_0, p_1 + Tq_1, p_2 + Tq_2),$$

where ord_T is the number of factors T in a polynomial in T ($\operatorname{ord}_T(0) = \infty$). We have

$$i(L, C; P) \geq 1 \quad \Longleftrightarrow \quad P \in L(k) \cap C(k).$$

If k is algebraically closed and $L \not\subset C$ then we see that the following version of **Bezout's theorem** holds:

$$\sum_{P \in \mathbf{P}^2(k)} i(L, C; P) = d;$$

This follows from the fact that the intersection divisor we constructed has the same degree as the equation defining C.

We say that L is a *tangent line* of C at P if $i(L, C; P) \geq 2$, and we say that L is a *flex* of C if $i(L, C; P) \geq 3$.

4.6. Lemma. *Let C be a curve over k given by $F(X, Y, Z) = 0$ and let $P = (p_0 \colon p_1 \colon p_2) \in C(k)$. If all partial derivatives F_X, F_Y, and F_Z vanish at P, then every line through P is a tangent line of C at P. Otherwise, there is a unique tangent line of P at C and it is given by the equation*

$$L_P : \qquad F_X(p_0, p_1, p_2)X + F_Y(p_0, p_1, p_2)Y + F_Z(p_0, p_1, p_2)Z = 0.$$

Note that the partial derivatives of polynomials can be defined purely algebraically, so the statements make sense over any field k.

Proof. We will also consider the equation L_P if all partials do vanish at P, in which case L_P is the whole $\mathbf{P}^2$. To see that $P \in L_P(k)$ note that $dF(X,Y,Z) = XF_X + YF_Y + ZF_Z$, where d is the degree of F. Suppose that $Q = (q_0{:}\, q_1{:}\, q_2) \in L(k)$ with $Q \neq P$. Then $i(L,C;P)$ is the number of factors T in the polynomial

$$f(T) = F(p_0 + Tq_0, p_1 + Tq_1, p_2 + Tq_2) \in k[T].$$

Expanding this as a polynomial in T one sees that

$$f(T) \equiv F(p_0,p_1,p_2) + (F_X(p_0,p_1,p_2)q_0 + F_Y(p_0,p_1,p_2)q_1 + F_Z(p_0,p_1,p_2)q_2)T \bmod (T^2)$$

Thus, L is a tangent to C at P if and only if $(q_0{:}\, q_1{:}\, q_2) \in L_P(k)$. If all partials vanish at P then we see that every L is a tangent. Now suppose that L_P is a line. We saw that L is a tangent if and only if $\{P,Q\} \subset L(k) \cap L_P(k)$, which can only happen if $L = L_P$. This proves the lemma.

We say that $P \in C(k)$ is a *non-singular point* of C if C has a unique tangent line at P. We say that C is *smooth* if $C(\overline{k})$ consists of only non-singular points. For instance, the curve $(Y - X^2)(Y - 2) = 0$ is defined over $\mathbf{Q}$, and it has only non-singular $\mathbf{Q}$-points, but it is not smooth.

Exercise 4. If a point $P \in \mathbf{P}^2(k)$ lies on two curves given by homogeneous equations $f_1 = 0$ and $f_2 = 0$, then P is a singular point of the curve given by the equation $f_1 f_2 = 0$.

Together with Bezout's theorem (in the strong form that we did not prove) the exercise above implies that a smooth curve is given by an irreducible equation.

Exercise 5. Sketch the affine real points of the curves $y = x^3$, $y^3 = x^4$, and $y^2 = x^3$. For which curves is $(0,0)$ a singular point?

Exercise 6. Show that a cubic curve which does not contain a line (i.e., the defining homogeneous polynomial of the curve has no linear factors) has at most one singular point.

Exercise 7. (Singular cubics.) Let k be a field whose characteristic is not 2, and let $a \in k$. Let C_a be the curve $y^2 = x^3 + ax^2$. Sketch C_0 and C_1 for $k = \mathbf{R}$. Show that for every line L through $P = (0,0) \in C(k)$ we have $L \cdot C = 2[P] + [Q]$ for unique $Q \in C(k)$. This gives an injection of $C(k)\backslash\{P\}$ to the set of lines through P. What is its image?

4.7. Elliptic curves. Let k be a field. By an elliptic curve E we will mean a smooth plane cubic curve with a specified point $0_E \in E(k)$. For every line L the divisor $L \cdot E$ is a sum of three points. We claim that for any $P, Q \in E(k)$ there is a unique line L such that $L \cdot C = [P] + [Q] + [R]$ for some $R \in E(k)$. To see this, note first that for any two distinct points in $\mathbf{P}^2(k)$ there is exactly one line containing both, so there is only one choice for L if $P \neq Q$. If $P = Q$ then the only choice for L is the tangent line of C at P, and it is unique because C is non-singular. Note that the uniqueness of L also implies the uniqueness of

R, and we will write $R = P * Q$. A priori we only have $R \in E(\overline{k})$. To see that $R \in E(k)$, note that we obtain R as the third zero of a polynomial over k of degree 3, whose other two roots are also in k. This shows the claim.

It is clear that $P*Q = Q*P$ and $P*(P*Q) = Q$ for all $P, Q \in E(K)$. We now define $P + Q$ to be $0_E * (P * Q)$. Note that the operation "$*$" does not depend on the choice of 0_E, but the operation "$+$" does.

Exercise 8. Check that $0_E + P = P$. Putting $-P = (0_E * 0_E) * P$ show that $P + (-P) = 0_E$. If $L \cdot C = [P] + [Q] + [R]$, what can you say about $P + Q + R$?

4.8. Theorem. *The operation "+" gives $E(k)$ the structure of an abelian group with identity element 0_E.*

In the exercise above it was shown that a unit element exists, and that inverses exist. The operation is also commutative, so only the axiom of associativity $P+(Q+R) = P+(Q+R)$ remains to be checked.

We will not give a complete proof of this fact in full generality. If $\mathrm{char}(K) = 0$ then we will deduce it from the complex analytic theory. If $\mathrm{char}(K)$ is not 2 or 3, then we will explain how to give a proof based solely on formulas, which can be written out with help of a computer.

However it is useful to know the essence of the proof that an algebraic geometer might give, so we sketch it here with one gap. We may assume that k is algebraically closed. We already know that $\mathrm{Div}(E)$ is the free abelian group on points of the elliptic curve. Let us write $\mathrm{Div}^0(E)$ for the subgroup of divisors for which the coefficients add up to 0. Let $\mathrm{Pr}(E)$ be the subgroup of $\mathrm{Div}^0(E)$ generated by all divisors of the form $L \cdot E - M \cdot E$ for lines L, M (neither contained in C). Let us now consider the map

$$\phi\colon \quad E(k) \to \mathrm{Div}^0(E)/\mathrm{Pr}(E) \qquad\qquad P \mapsto ([P] - [0_E] \bmod \mathrm{Pr}(E)).$$

If we take a line L through points P and Q and we take M through 0_E and $P * Q$ then we see that

$$[P] + [Q] + [P * Q] \equiv [0_E] + [P * Q] + [P + Q] \bmod \mathrm{Pr}(E),$$

so that $\phi(P + Q) = \phi(P) + \phi(Q)$. This also implies that ϕ is surjective. With some more algebraic geometry, notably the theorem of Riemann-Roch, one can show that ϕ is also injective. It then follows that the group law on E is associative. In fact, $\mathrm{Pr}(E)$ is the group of principal divisors, just like in the complex analytic case.

Note that $\mathrm{Div}^0(E)/\mathrm{Pr}(E)$ is a group in a natural way, but that the isomorphism ϕ depends of the choice of a point 0_E. Again this reflects that the group law on E depends on the choice of 0_E.

Exercise 9. Suppose we are given a second point $0'_E$ on an elliptic curve E. Express the group operation $+'$ on $E(k)$ with identity element $0'_E$ in terms of the operation $+$.

Exercise 10. The 2-torsion of an elliptic curve E is $E(k)[2] = \{P \in E(k)\colon P + P = O_E\}$. Suppose that $\#E(k)[2] = 4$. Show that 0_E is a flex of E if and only if the points $E(k)[2]\backslash\{0_E\}$ are collinear.

5. Elliptic curves in Weierstrass form

In this section we assume that the characteristic of k is not 2 or 3. A Weierstrass equation, or an elliptic curve in Weierstrass form, is an equation of the form

$$E: \quad Y^2 = X^3 + aX + b$$

for certain $a, b \in k$ with $4a^3 + 27b^2 \neq 0$. This last condition ensures that the polynomial $X^3 + aX + b$ has no double zero, so that the curve is smooth. As the distinguished point 0_E we take $0_E = (0:1:0)$. If we set $Z = 0$ in the homogeneous equation for E, then we get $X^3 = 0$, so the line at infinity is a flex line at the unique point 0_E of E at infinity. For $P = (x, y) \in E(k)$ one easily sees that $-P = (x, -y)$.

Exercise 1. How do we recognize 2-torsion points (points P with $P + P = 0_E$) on E? Show that the 3-torsion points on E are the flexes. How many 3-torsion points do you think exist when $k = \mathbf{C}$?

5.1. Addition formulas. Suppose that $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ are affine points on our curve E in Weierstrass form. Let us suppose that $Q \neq -P$. We will find a formula for $P + Q = (x_3, y_3)$.

The line L for which $L \cdot E = [P] + [Q] + [P * Q]$ is of the form $y = \lambda x + \nu$. Here $\lambda \in k$ is given by

$$\lambda = \frac{y_2 - y_1}{x_2 - x_1} = \frac{x_1^2 + x_1 x_2 + x_2^2 + a}{y_1 + y_2}$$

in the sense that each formula is correct if it is defined (i.e., has non-zero denominator), and that at least one is defined. We then have $\nu = y_1 - \lambda x_1$. Intersecting with E we see that we have an identity in $k[x]$.

$$x^3 + ax + b - (\lambda x + \nu)^2 = (x - x_1)(x - x_2)(x - x_3).$$

By looking at the coefficient of x^2 on both sides we see that

$$x_3 = \lambda^2 - x_1 - x_2; \qquad y_3 = -(\lambda x_3 + \nu).$$

With these formulas one can do explicit computations on elliptic curves.

Exercise 2. Show that the point $(2, 4)$ on the elliptic curve $y^2 = x^3 + 4x$ has order 4.

Exercise 3. Show that a point (x, y) of $E(k)$ has order 3 if and only if

$$x^4 + 2ax^2 + 4bx - a^2/3 = 0.$$

*Same question for order 4:

$$x^6 + 5ax^4 + 20bx^3 - 5a^2x^2 - 4abx - 8b^2 - a^3 = 0.$$

Could you have predicted that we should get a degree 6 equation?

5.2. Applications to algorithms. One of the most striking applications of elliptic curves is their use in factoring algorithms and primality tests. Older algorithms in this area that operate with the better known group $(\mathbf{Z}/p\mathbf{Z})^*$ for a prime number p, can often be adapted to work on the group of points of an elliptic curve over the field $k = \mathbf{Z}/p\mathbf{Z}$. This can drastically enhance the performance because the group $E(\mathbf{Z}/p\mathbf{Z})$ potentially has much better properties than $(\mathbf{Z}/p\mathbf{Z})^*$. A theorem of Hasse says that

$$p+1-2\sqrt{p} \;\le\; \#E(\mathbf{Z}/p\mathbf{Z}) \;\le\; p+1+2\sqrt{p},$$

so the groups $E(\mathbf{Z}/p\mathbf{Z})$ and $(\mathbf{Z}/p\mathbf{Z})^*$ have approximately the same order. Often the algeorithm works well if the group order is built up from small primes, and one has more of a chance that this happens if one can vary this order within Hasse's interval. We refer to Koblitz's recent book *A course in number theory and cryptography* (Springer GTM 114, 1987) for more details.

5.3. Computational proof of associativity. One can also use the explicit formulas to check associativity of the group law, i.e., $P + (Q + R) = (P + Q) + R$. By hand, one checks the case that one of the three points is 0_E, and the case that the outcome of any addition is 0_E. Now we know that for each addition one of the two formulas will work. If one writes these formulas in projective form, then we get two triples of homogeneous forms in $x_1, y_1, z_1, x_2, y_2, z_2$ that either give coordinates of the sum, or $(0{:}\,0{:}\,0)$. Doing the addition of three points gives 2 choices for placing brackets and then 4 possibilities for picking formulas for the additions. This gives 8 triples of homogeneous forms in 9 variables $x_1, \ldots, z_3$, and we have to check that for each triple of points on E the formulas give the same projective point, or $(0{:}\,0{:}\,0)$. Thus, we need to check that all 2×2-determinants of 2×2-submatrices of our 8×3-matrix of forms lie in the ideal of $\mathbf{Z}[a, b][x_1, \ldots, z_3]$ generated by $\{y_i^2 - x_i^3 - ax_i - b : i = 1, 2, 3\}$. This would prove the result for any elliptic curve in Weierstrass form, in any characteristic not 2 or 3.

Exercises.

4. *(Porism of Diophantus.)* Let $d > 0$ be an integer that is a difference of two positive rational cubes. Show that d is also a *sum* of two positive rational cubes. Find such cubes for $d = 7 = 2^3 - 1^3$ and $d = 19 = 3^3 - 2^3$.

5. Determine the values of $\alpha \in \mathbf{Q}$ for which for which the cubic curve E with equation $X^3 + Y^3 + Z^3 - \alpha XYZ = 0$ is non-singular. For these α, let $O = (0 : -1 : 1)$ be the origin on E. Find a Weierstrass equation for E.

6. Determine the structure of the group $E(\mathbf{F}_5)$ when E is the elliptic curve with Weierstrass equation
$$y^2 = x^3 + x, \qquad y^2 = x^3 + 2x, \qquad y^2 = x^3 + 1.$$
Prove: $\#E(\mathbf{F}_5) \leq 11$ for every elliptic curve over $\mathbf{F}_5$. Can you improve this bound? Can you find examples where $\#E(\mathbf{F}_5)$ is close to your bound?

7. Show that the nodal cubic curve with singular Weierstrass equation $y^2 = x^3 - x^2$ is birationally equivalent to $\mathbf{P}^1(K)$ over every field K of characteristic different from 2.
[Hint: exercise 4.7].

8. Same question for the cuspidal cubic curve with singular Weierstrass equation $y^2 = x^3$.

9. Find a Weierstrass equation for the cubic curve with affine equation $y^3 - y^2 = x^3 - x$ and origin $O = (1, 0)$.

10. Find a Weierstrass equation for the cubic curve with equation $X^2Y - XY^2 - XZ^2 + Y^2Z = 0$ and origin $O = (1 : 1 : 1)$.

11. Let E be a smooth cubic curve with origin O, defined over an algebraically closed field K of characteristic different from 2. Show that the 2-torsion subgroup $E[2](K)$ consists of 4 points, and that the 3 non-trivial points in $E[2](K)$ are collinear if and only if O is a flex.

12. Let C be the curve with affine equation $Y^2 = 1 - X^4$ encountered in section 1. Determine the singular projective points of C, and show that the rational functions $x = 2(Y + 1)/X^2$ and $y = 4(Y + 1)/X^3$ yield a birational equivalence between C and the elliptic curve with Weierstrass equation $y^2 = x^3 + 4x$.

13. Derive Euler's addition formula for the lemniscatic P-function.

6. Weak Mordell-Weil theorem

In this section we are concerned with elliptic curves over $\mathbf{Q}$. Let us fix notation: E will be an elliptic curve given by an equation

$$E: \quad Y^2 = W(X), \qquad W(X) = X^3 + aX^2 + bX + c \in \mathbf{Q}[X].$$

The theorem of Mordell-Weil says that the group $E(\mathbf{Q})$ is finitely generated. By the structure of finitely generated abelian groups this means that

$$E(\mathbf{Q}) \;\cong\; T \oplus \mathbf{Z}^r,$$

for a finite abelian group T, the torsion group of $E(\mathbf{Q})$, and an integer $r \geq 0$, called the *rank* of E.

At this point all we know about the group $E(\mathbf{Q})$ is that it is a countable subgroup of $E(\mathbf{C})$, which is isomorphic as an abelian group to $(\mathbf{R}/\mathbf{Z}) \times (\mathbf{R}/\mathbf{Z})$. In exercise 1 below we will see that $(\mathbf{R}/\mathbf{Z}) \times (\mathbf{R}/\mathbf{Z})$ contains many countable subgroups which are not finitely generated.

The proof of the Mordell-Weil theorem proceeds in two steps. We first show the weak version, which says that $E(\mathbf{Q})/2E(\mathbf{Q})$ is finite. The second step is a descent argument, which will be given in the next section. For the weak version we need to work with the 2-torsion points of the elliptic curve. This forces us to not only do arithmetic in $\mathbf{Q}$ but also in the possibly quadratic or cubic number field that is generated by the x-coordinate of a 2-torsion point. In this section we will only write out the proof in the case that all arithmetic takes place in $\mathbf{Q}$. So we will prove the following proposition.

6.1. Proposition. *If $W(X)$ has three roots in $\mathbf{Q}$ then $E(\mathbf{Q})/2E(\mathbf{Q})$ is finite.*

Our proof extends to the general case, where no assumption is made on the roots of $W(X)$, if the reader is willing to take some algebraic number theory for granted. More precisely, the argument needs the fact that the class group is finite and that the unit group is finitely generated.

Using isogenies of degree 2, one can also do the case that $W(X)$ has at least one rational root, while keeping all arithmetic in $\mathbf{Q}$. We will describe this with explicit formulas from complex analysis that we saw already in Section 3. This method is often more suitable for the actual computation of $E(\mathbf{Q})/2E(\mathbf{Q})$ for specific E.

We start the proof of Proposition 6.1 by recalling the definition of the group operation on $E(\mathbf{Q})$. Let us write out what it means for three affine points (x_i, y_i) of $E(\mathbf{Q})$ to have sum zero. There should be a line $Y = lX + m$ whose intersection with E consists of the three points (counting multiplicity), so

$$(*) \qquad W(X) - (lX + m)^2 = (X - x_1)(X - x_2)(X - x_3).$$

Let us now consider the ring $R = \mathbf{Q}[X]/(W(X))$ and put $\overline{X} = (X \bmod W(X)) \in R$. Then we have

$$(**) \qquad (x_1 - \overline{X})(x_2 - \overline{X})(x_3 - \overline{X}) = (l\overline{X} + m)^2.$$

Note that for a polynomial $g(X) \in \mathbf{Q}[X]$ we have $g(\overline{X}) \in R^*$ if the roots in $\mathbf{C}$ of $W(X)$ are not roots of g. In particular, for a point $(x, y) \in E(\mathbf{Q})$ which is not 2-torsion, the element $x - \overline{X}$ is a unit in R. Thus, we have a map

$$E(\mathbf{Q})\backslash E(\mathbf{Q})[2] \xrightarrow{\varphi} R^*/R^{*2} \qquad (x, y) \mapsto x - \overline{X} \bmod R^{*2},$$

which has the property that for three points on the left with sum zero, the product of the images on the right is 1. In order to define φ on a 2-torsion point $(x, 0) \in E(\mathbf{Q})$ we add a correction term $W_x(\overline{X})$, where $W_x \in \mathbf{Q}[X]$ is the quadratic polynomial for which $W(X) = (X - x)W_x(X)$. Note that no complex root of $W(X)$ is a root of $x - X + W_x(X)$ so that $x - \overline{X} + W_x(\overline{X})$ is a unit in R. We now define $\varphi(P) \in R^*/R^{*2}$ for $P \in E(\mathbf{Q})$ by

$$\varphi(P) = \begin{cases} 1 & \text{if } P = 0_E; \\ x - \overline{X} \bmod R^{*2} & \text{if } P = (x, y) \text{ with } y \neq 0; \\ x - \overline{X} + W_x(\overline{X}) \bmod R^{*2} & \text{if } P = (x, 0) \in E[2]; \end{cases}$$

We will see in the exercises below why this correction term is the only reasonable thing to try.

Proposition 6.1 will follow from the next three lemmas, in which it is proved step by step that φ induces a bijection between $E(\mathbf{Q})/2E(\mathbf{Q})$ and a finite subgroup of R^*/R^{*2}.

6.2. Lemma. *The map φ: $E(\mathbf{Q}) \to R^*/R^{*2}$ is a homomorphism of groups.*

Proof. We first remark that $\varphi(-P) = \varphi(P) = \varphi(P)^{-1}$. Therefore, we only need to show that for points $P_1, P_2, P_3 \in E(\mathbf{Q})$ with sum 0_E we have $\varphi(P_1)\varphi(P_2)\varphi(P_3) = 1$. In the case that one of the P_i is 0_E this follows from the fact that $\varphi(-P) = \varphi(P)$. So let us assume all P_i are affine, and put $P_i = (x_i, y_i)$. Equation $(**)$ above gives $\varphi(P_1)\varphi(P_2)\varphi(P_3) = 1$ if no P_i is 2-torsion.

Suppose that there is exactly one 2-torsion point among the P_i, say $P_1 = (x_1, 0)$. Then we can write the line as $Y = l(X - x_1)$, and by taking out a factor $(X - x_1)$ in the equation $(*)$ and writing $T = W_{x_1} \in \mathbf{Q}[X]$ we obtain

$$T(X) - l^2(X - x_1) = (X - x_2)(X - x_3).$$

Since $T(\overline{X})(\overline{X} - x_1) = 0$ we have $T(\overline{X})(x_2 - \overline{X})(x_3 - \overline{X}) = T(\overline{X})^2$. With $(**)$ one sees that

$$(x - \overline{X} + T(\overline{X}))(x_2 - \overline{X})(x_3 - \overline{X}) = (l^2(\overline{X} - x_1)^2 + T(\overline{X})^2) = (l(\overline{X} - x_1) + T(\overline{X}))^2.$$

The only case left is the case that there are at least two 2-torsion points among the three affine points P_1, P_2, P_3 with sum 0_E. This can only happen if the P_i are three distinct 2-torsion points. Then $W(X) = (X-x_1)(X-x_2)(X-x_3)$ and by the Chinese Remainder Theorem we have an isomorphism of rings

$$R \xrightarrow{\sim} \mathbf{Q}\times\mathbf{Q}\times\mathbf{Q} \qquad \overline{X} \mapsto (x_1, x_2, x_3).$$

Let us write down the image of $\varphi(P_i)$ for $i = 1$, 2, 3:

$$\begin{array}{lccccccc}
\varphi(x_1,0) & \mapsto & (& (x_1-x_2)(x_1-x_3), & x_1-x_2, & x_1-x_3 &) \\
\varphi(x_2,0) & \mapsto & (& x_2-x_1, & (x_2-x_3)(x_2-x_1), & x_2-x_3 &) \\
\varphi(x_3,0) & \mapsto & (& x_3-x_1, & x_3-x_2, & (x_3-x_1)(x_3-x_2) &)
\end{array}$$

Since the product in each column is a square, we have proved Lemma 6.2. □

6.3. Lemma. *The kernel of φ is $2E(\mathbf{Q})$.*

Proof. Suppose that $P = (x,y)$ lies in the kernel of φ. We are looking for a point $Q \in E(\mathbf{Q})$ with $P = 2Q$. In view of equation $(**)$ this means that we would like to write $x - \overline{X}$ in R as the square of the quotient of two linear polynomials in $\overline{X}$. Let us first show that $x - \overline{X}$ is a square in R. This is just the property $\varphi(P) = 1$ if P is not 2-torsion. If P is 2-torsion, then still $x - \overline{X}$ is a square modulo $W_x(\overline{X})$, because the correction term vanishes modulo $W_x(\overline{X})$. Modulo $x - \overline{X}$ it is zero, which is also a square. Since $X - x$ is coprime to $W_x(X)$ the Chinese Remainder Theorem implies that $x - \overline{X}$ it is a square in R.

We can now write $x - \overline{X} = (p_2\overline{X}^2 + p_1\overline{X}^2 + p_0)^2$. Since $\overline{X}$ satisfies no polynomial relation of degree less than three, $x - \overline{X}$ is not the square of a constant or a linear polynomial in $\overline{X}$, so $p_2 \neq 0$. For $s, t \in \mathbf{Q}$ consider the element

$$(s\overline{X} + t)(p_2\overline{X}^2 + p_1\overline{X} + p_0) \in R$$

By using the equation $W(\overline{X}) = 0$ we can rewrite this expression in as a degree 2 polynomial in $\overline{X}$. For fixed p_0, p_1, p_2 the coefficient of $\overline{X}^2$ is a linear homogeneous expression in s and t. Thus, there exists a pair $(s,t) \neq (0,0)$ for which this coefficient vanishes. Since $p_2 \neq 0$ we have $s \neq 0$. Thus, we can take $s = -1$ and we obtain

$$(t - \overline{X})(p_2\overline{X}^2 + p_1\overline{X} + p_0) = l\overline{X} + m$$

for certain $l, m \in \mathbf{Q}$. Squaring gives

$$(t - \overline{X})^2(x - \overline{X}) = (l\overline{X} + m)^2.$$

But now the monic cubic polynomial $(lX + m)^2 - (t - X)^2(x - X)$ is divisible by $W(X)$ so it must be equal to $W(X)$. It follows that $Q = (t, lt + m)$ is an element of $E(\mathbf{Q})$, and that P is $2Q$ or $-2Q$. □

6.4. Lemma. *Suppose that $W(X) \in \mathbf{Z}[X]$ and that $e \in \mathbf{Z}$ is a root of $W(X)$. Let $R \to \mathbf{Q}$ be the ring homomorphism $R \to \mathbf{Q}$ sending $\overline{X}$ to e and let φ_e be the composite map*

$$\varphi_e: \quad E(\mathbf{Q}) \xrightarrow{\varphi} R^*/R^{*2} \longrightarrow \mathbf{Q}^*/\mathbf{Q}^{*2}$$

Then φ_e has finite image. More particularly, for $(\alpha \bmod \mathbf{Q}^{*2})$ in the image we have:

(1) *for all $p \nmid W_e(e)$ the number $\operatorname{ord}_p(\alpha)$ is even;*

(2) *if $W(X)$ has no real roots that are smaller than e, then $\alpha > 0$.*

Proof. Suppose that $\alpha = \varphi_e(P)$ with $P = (x, y) \in E(\mathbf{Q})$. If $P = (e, 0)$ then $\alpha = (W_e(e) \bmod \mathbf{Q}^{*2})$, so (1) holds. To see (2) in this case, note that $W_e(X)$ is a monic quadratic polynomial, so that $W_e(e) < 0$ implies that $W_e(X)$ has a real root smaller than e.

Now suppose that $x \neq e$ so that $\alpha = (x - e \bmod \mathbf{Q}^{*2})$. If d is the denominator of x then d^3 is the denoninator of $W(x)$ and since $y^2 = W_e(x)$ we see that $2 \mid \operatorname{ord}_p(d) = -\operatorname{ord}_p(x - e)$ for all prime numbers $p \mid d$. Now let p be a prime number with $n = \operatorname{ord}_p(x - e)$ odd. If $n < 0$ then $\operatorname{ord}_p(x) = n$ and $\operatorname{ord}_p(W(x)) = 3n$, so that $\operatorname{ord}_p(y^2)$ is odd. Therefore, $n > 0$, and since $y^2 = (x - e)W_e(x)$ we see that $\operatorname{ord}_p(W_e(x))$ is odd, and therefore positive. Reducing modulo p, and using that $x \equiv e \bmod p$ we see that $p \mid W_e(e)$. The condition in (2) means that $W_e(t) > 0$ for $t \in \mathbf{R}$ with $t < e$, so if $x < e$ then we would have $y^2 < 0$. This shows (1) and (2).

It follows that $\varphi_e(E(\mathbf{Q}))$ is finite because an element $\alpha \in \mathbf{Q}^*/\mathbf{Q}^{*2}$ is determined by its sign and, for every prime number p, the parity of $\operatorname{ord}_p(\alpha)$. □

6.5. Lemma. *If $W(X)$ has 3 rational roots then the map φ: $E(\mathbf{Q}) \to R^*/R^{*2}$ has finite image.*

Proof. By scaling x and y we may assume that $W(X) \in \mathbf{Z}[X]$ (see exercise 9). By Gauss's lemma this also implies that the roots e_1, e_2, e_3 of $W(X)$ are integers. If we identify R with $\mathbf{Q} \times \mathbf{Q} \times \mathbf{Q}$ as in the end of the proof of Lemma 6.2, then we see that the map φ consists of three components φ_{e_1}, φ_{e_2}, φ_{e_3}, each of which has finite image by 6.4. □

This completes our proof of Proposition 6.1. The proof that the image of φ is finite without conditions on the roots of $W(X)$ is very similar to the proof of Lemma 6.4 if one knows enough about arithmetic in number fields (see exercise 8).

Let us show how to make the upper bounds for $E(\mathbf{Q})/2E(\mathbf{Q})$ that one obtains from the finiteness proof a bit more explicit. If $W(X) \in \mathbf{Z}[X]$ then we say that a prime number p is "bad," or that E has *bad reduction* modulo p, if the reduction $\overline{W}(X)$ of $W(X)$ modulo p has a double root. For such a prime p the polynomial $\overline{W}(X) \in \mathbf{F}_p(X)$ has exactly 1 or 2 roots, in which case we say that p is "instable" or "semi-stable" respectively.

6.6. Corollary. *Suppose that $W(X) \in \mathbf{Z}[X]$ and that $W(X)$ has three roots in $\mathbf{Q}$. Let $n_{\rm ss}$ and $n_{\rm is}$ be the number of semi-stable and instable primes for $W(X)$. Then*

$$\dim_{\mathbf{F}_2}(E(\mathbf{Q})/2E(\mathbf{Q})) = 2 + r \qquad \text{with} \qquad r \leq n_{\rm ss} + 2n_{\rm is} - 1.$$

Proof. For any prime number p let $H_p \subset \mathbf{F}_2 \times \mathbf{F}_2 \times \mathbf{F}_2$ be the sub-$\mathbf{F}_2$-vector space of vectors (a_1, a_2, a_2) satisfying

(1) $a_1 + a_2 + a_3 = 0$;
(2) $a_i = 0$ if there is no $j \neq i$ with $e_i \equiv e_j \bmod p$.

We also define H_p for $p = \infty$ by replacing (2) with

(2′) $a_i = 0$ if e_i is the smallest real root of $W(X)$.

It follows from what we have proved already that φ induces an injective homomorphism

$$E(\mathbf{Q})/2E(\mathbf{Q}) \to H_\infty \times \prod_{p \text{ prime}} H_p.$$

The statement now follows from the fact that

$$\dim_{\mathbf{F}_2}(H_p) = \begin{cases} 0 & \text{if } p \text{ is ``good''}; \\ 1 & \text{if } p \text{ is semi-stable or } p = \infty; \\ 2 & \text{if } p \text{ is instable prime.} \end{cases}$$ □

Once we know that $E(\mathbf{Q})$ is finitely generated, we also know that the number r above is the rank of E. The rank itself is invariant under scaling of the elliptic curve, but the bound given above is not. Thus, the bound works best when applied to a "minimal" Weierstrass equation.

Let us consider again the case that $W(X)$ has at least one rational root. If $W(X)$ also has a quadratic irreducible factor then the argument we alluded to before requires arithmetic in a quadratic number field. One can avoid this by using isogenies of degree 2. The idea of this second method is that the multiplication-by-2-map $E \xrightarrow{2} E$ breaks up as a product of two isogenies $E \to E' \to E$ which are each of degree 2. One then shows that the maps $E(\mathbf{Q}) \to E'(\mathbf{Q})$ and $E'(\mathbf{Q}) \to E(\mathbf{Q})$ have finite cokernel. It turns out that this last cokernel is exactly the image of φ_e in Lemma 6.4, which we already know is finite.

One can make this argument precise by giving explicit formulas, which one recovers from complex analysis. We assume that $W(X)$ has a rational root. After translating we can assume that this root is 0:

$$E: \quad Y^2 = W(X), \qquad W(X) = X(X^2 + aX + b) \in \mathbf{Q}[X].$$

In order to find an equation for the isogenous curve E', consider the Weierstrass parametrization for E with period lattice $\Lambda = \mathbf{Z}\omega_1 + \mathbf{Z}\omega_2$ which maps the 2-torsion point $\omega_1/2 \in \mathbf{C}/\Lambda$ to $(0,0) \in E(\mathbf{C})$:

$$\mathbf{C}/\Lambda \to E(\mathbf{C}) \qquad z \mapsto (4\wp(z) - 4\wp(\omega_1/2)), 4\wp'(z)).$$

Put $\Lambda' = \mathbf{Z}\omega_1/2 + \mathbf{Z}\omega_2$, then we want $E'(\mathbf{C})$ to be $\mathbf{C}/\Lambda'$ and we need an isogeny ψ such that the diagram

$$\begin{tikzcd}
\mathbf{C}/\Lambda \arrow[r] \arrow[d] & \mathbf{C}/\Lambda' \arrow[d] \\
E(\mathbf{C}) \arrow[r, "\psi"] & E'(\mathbf{C}).
\end{tikzcd}$$

commutes. From Section 3 we see that E' can be taken to be the curve

$$E': \quad Y^2 = W(X), \qquad W(X) = X(X^2 - 2aX + a^2 - 4b) \in \mathbf{Q}[X]$$

and that ψ is given by

$$\psi: \quad E \to E' \qquad (x,y) \mapsto \left(\frac{x^2+ax+b}{x}, (1-b/x^2)y\right),$$

with $\psi(0_E) = \psi((0,0)) = 0_{E'}$. Since the formula for ψ is defined with rational coefficients we obtain a group homomorphism $E(\mathbf{Q}) \xrightarrow{\psi} E'(\mathbf{Q})$.

6.7. Lemma. *Let φ_0 be the map in 6.4 for the curve E'. Then the sequence*

$$E(\mathbf{Q}) \xrightarrow{\psi} E'(\mathbf{Q}) \xrightarrow{\varphi_0} \mathbf{Q}^*/\mathbf{Q}^{*2}$$

is exact

Proof. For $(x', y') \in E'(\mathbf{Q})$ with $x' \neq 0$ we have $(x'^2 + ax' + b)/x = (y'/x')^2$, so that indeed $\mathrm{Im}(\psi) \subset \mathrm{Ker}(\varphi_0)$. To show equality, suppose that $P' = (x', y') \in \mathrm{Ker}(\varphi_0) \subset \mathrm{E}'(\mathbf{Q})$. If $P' = (0,0)$ then $\varphi_0(P') = W_0(0) = a^2 - 4b$ is square, which means that $x^2 + ax + b$ has a rational root e, and $P = \psi(e, 0)$. If $P' = (x', y') \neq (0,0)$, then $x' = t^2$ with $t \in \mathbf{Q}^*$. To find $P = (x,y) \in E(\mathbf{Q})$ with $\psi(P) = P'$ substitute $y = tx$ and solve $(1 - \frac{b}{x^2})tx = y'$, i.e., solve $x^2 - (y'/t)x - b = 0$. This equation can be solved if $(y'/t)^2 + 4b$ is a square. But $(y'/t)^2 = x'^2 - 2ax' + a^2 - 4b = (x' - a)^2 - 4b$ so indeed $(y'/t)^2 + 4b$ is a square, and a point $P = (x,y)$ exists with $\psi(P) = P'$. □

We deduce that the homomorphism ψ: $E'(\mathbf{Q}) \to E(\mathbf{Q})$ has finite cokernel. Under the Weierstrass parametrization for E' the element $\omega_2/2 \in \mathbf{C}/\Lambda'$ maps to $(0,0) \in E'(\mathbf{C})$. If we apply the same process to the curve E' we find a curve E'' corresponing to $\mathbf{C}/\frac{1}{2}\Lambda$ which is just E scaled by a factor 2. One then checks that the following diagram commutes

$$\begin{tikzcd}
\mathbf{C}/\Lambda \arrow[r] \arrow[d] & \mathbf{C}/\Lambda' \arrow[r] \arrow[d] & \mathbf{C}/\frac{1}{2}\Lambda \arrow[r, "2"] \arrow[d] & \mathbf{C}/\Lambda \arrow[d] \\
E(\mathbf{C}) \arrow[r, "\psi"] & E'(\mathbf{C}) \arrow[r, "\psi'"] & E''(\mathbf{C}) \arrow[r, "\sim"] & E(\mathbf{C}).
\end{tikzcd}$$

The map $E''(\mathbf{C}) \to E(\mathbf{C})$ is just scaling $(x,y) \mapsto (x/4, y/8)$. All vertical maps are isomorphisms of groups, and we deduce that the $E(\mathbf{C}) \to E(\mathbf{C})$ is multiplication by 2.

6.8. Proposition. *If $W(X)$ has a rational root then $E(\mathbf{Q})/2E(\mathbf{Q})$ is finite.*

Proof. The muliplication by 2 map on $E(\mathbf{Q})$ is a composition

$$E(\mathbf{Q}) \xrightarrow{\psi} E'(\mathbf{Q}) \xrightarrow{\psi'} E''(\mathbf{Q}) \xrightarrow{\sim} E(\mathbf{Q}).$$

For each map the index of the image is finite, so $E(\mathbf{Q})/2E(\mathbf{Q})$ is finite as well. □

Exercises.

1. Show that $(\mathbf{R}/\mathbf{Z})\times(\mathbf{R}/\mathbf{Z})$ has a countably infinite subgroup in which every element has finite order. Show it also has a countable subgroup A for which $A/2A$ is not finite. Show that it also has a countably infinite subgroup A for which $A=2A$, and every non-zero element has infinite order.

2. Let A be a commutative ring containing a field K, and suppose that A is of finite dimensional d as a vector space over K. We define a norm map $N_{A/K}\colon A\to K$ as follows: an element $a\in A$ is mapped to the determinant of the K-linear endomorphism $x\mapsto xa$ of A.

(1) Show that the norm induces a homomorphism $A^*\to K^*$.

Let K be a field and let $f\in K[X]$ be a non-constant polynomial. Then the ring $A=K[X]/(f)$ contains K and the dimension d of A over K is the degree of f. Over an algebraic closure $\overline{K}$ of K we can write $f=(X-e_1)(X-e_2)\cdots(X-e_d)$ with $e_1,e_2,\ldots,e_d\in\overline{K}$. For any $g\in K[X]$ we now claim that

$$N_{A/K}(g \bmod f)=g(e_1)g(e_2)\cdots g(e_d).$$

(2) Show this with the Chinese Remainder Theorem in the in the case that f has no double roots.

(3) Show that the image of φ is contained in the kernel of the map

$$R^*/R^{*2}\longrightarrow\mathbf{Q}^*/\mathbf{Q}^{*2}$$

induced by $N_{R/\mathbf{Q}}\colon R\to\mathbf{Q}$.

3. Suppose that $P=(x,0)\in E(\mathbf{Q})[2]$. Show that $R\cong\mathbf{Q}\times S$ for a ring S with $x-\overline{X}$ mapping to $(0,u)$ for some $u\in S^*$. Show that $W_x(\overline{X})$ is sent to $(N_{S/\mathbf{Q}}(u),0)$. Suppose that a candidate correction term $T\in R$ maps to $(t,0)$ with $t\in\mathbf{Q}^*$ and satisfies $N_{R/\mathbf{Q}}(x-\overline{X}+T)\in\mathbf{Q}^{*2}$. Show that $T=v^2W_x(\overline{X})$ for some $v\in\mathbf{Q}^*$.

4. Find the 2-power torsion and sets of representatives for $E(\mathbf{Q})/2E(\mathbf{Q})$ for the following elliptic curves E:

(1) $Y^2=X(X-3)(X+4)$;

(2) $Y^2=X(X-1)(X+3)$;

(3) $Y^2=X(X+1)(X-14)$.

5. Show that $E(\mathbf{Q})[4]$ is a group of order at most 8.

6. Let p be an odd prime number. Consider the elliptic curve E: $Y^2=X^3-p^2X$.

(1) Compute the image of the 2-torsion of E under φ.

(2) Show that for all $P\in E(\mathbf{Q})$ there is a 2-torsion point Q such that $\varphi(P-Q)=(d_1,d_2,d_3)$ with every d_i a divisor of 2.

(3) Show that the rank r of E is at most 2.

(4) *For primes p with $p\equiv 3 \bmod 8$ show that $r=0$.

7. Find the 2-power torsion and sets of representatives for $E(\mathbf{Q})/2E(\mathbf{Q})$ for the elliptic curve E: $Y^2=X(X^2+1)$. Try to do this both with arithmetic in $\mathbf{Q}(i)$ and with isogenies.

8. *Assuming that for each number field the class group is finite and the unit group of the ring of integers is finitely generated, show that Mordell's theorem also holds if the 2-torsion is not rational.

9. In the proof of Lemma 6.4 we reduced to the case that a and b are integers by "scaling". This exercise is intended to make this more precise. Let $u \in \mathbf{Q}^*$, and consider E': $Y^2 = X^3+u^2aX^2+u^4bX+cu^6$. Show that we have isomorphisms $E(\mathbf{Q}) \xrightarrow{\sim} E'(\mathbf{Q})$ and $R \xrightarrow{\sim} R' = \mathbf{Q}[X]/(X^3+u^2aX^2+u^4bX+cu^6)$ such that the diagram

$$\begin{tikzcd} E(\mathbf{Q}) \arrow[r, "\varphi"] \arrow[d] & R^*/R^{*2} \arrow[d] \\ E'(\mathbf{Q}) \arrow[r, "\varphi'"] & R'^*/R'^{*2} \end{tikzcd}$$

is commutative. How does the discriminant change if we pass from E to E'? Can you do the same for a map $X \mapsto uX+v$ rather than $X \mapsto uX$?

10. Give an analog of Corollary 6.6 in the case that $W(X)$ has exactly one root in $\mathbf{Q}$ by using the proof of Proposition 6.8.

7. Heights and the Mordell-Weil theorem

In this section we show how to deduce that $E(\mathbf{Q})$ is finitely generated if one knows that $E(\mathbf{Q})/2E(\mathbf{Q})$ is finite. In this section we suppose that E is given by a Weierstrass equation

$$E: \quad Y^2 = W(X), \qquad W(X) = X^3 + aX + b \in \mathbf{Q}[X].$$

We will develop the notion of the "height" of a point in $E(\mathbf{Q})$. This height will be such that for every $B \in \mathbf{R}$ there are only finitely many points P in $E(\mathbf{Q})$ of height less than B. Up to translation by one of finitely many points representing $E(\mathbf{Q})/2E(\mathbf{Q})$ we can now write a big point $P \in E(\mathbf{Q})$ as 2 times a point Q. We will show that for some $B \in \mathbf{R}$ and all P of height at least B the height of this Q is at most half of the height of P. Thus, large P can be rewritten as a combination of smaller points, and in the end we see that $E(\mathbf{Q})$ is generated by the points of height at most B and the representatives of $E(\mathbf{Q})/2E(\mathbf{Q})$.

7.1. Definition. *For $x = a/b \in \mathbf{Q}$ with $a, b \in \mathbf{Z}$ coprime, the height of x is the positive integer*

$$H(x) = \max\{|a|, |b|\}.$$

For convenience we will put $H(\infty) = H(1/0) = 1$. In fact, heights can be defined on projective spaces, and we only consider the case of a $\mathbf{P}^1$. We need a lemma about how the height of a rational number changes if we apply a rational function to it.

7.2. Lemma. *Let $f, g \in \mathbf{Q}[X]$ be coprime polynomials and let $n = \max\{\deg f, \deg g\}$. Then there are real numbers $C_1, C_2 > 0$ such that for all $x \in \mathbf{Q}$ we have*

$$C_1 H(x)^n \leq H(f(x)/g(x)) \leq C_2 H(x)^n.$$

Proof. We may assume that f and g have integer coefficients. Let M be the largest absolute value of a coefficient occurring in f or g. Suppose that $x = a/b$ with $a, b \in \mathbf{Z}$ coprime. We write $f(x)/g(x) = A/B$ with $A = b^n f(a/b)$ and $B = b^n g(a/b)$. Both A and B are $\mathbf{Z}$-linear combinations of the numbers a^n, $a^{n-1}b$, …, b^n with coefficients bounded by M. Since $|a|, |b| \leq H(x)$ we see that

$$H(f(x)/g(x)) \leq \max\{|A|, |B|\} \leq (n+1)MH(x)^n.$$

This shows the second inequality with $C_2 = (n+1)M$. The first inequality is more subtle, because we have to prove two things: we need $\max\{|A|, |B|\}$ to be big, and we need $\gcd(A, B)$ to be small.

We claim that there are polynomials f_2 and g_2 in $\mathbf{Z}[X]$ of degree at most $n-1$ such that $f_2 f + g_2 g = R$ for some non-zero $R \in \mathbf{Z}$. To see this, suppose that $\deg(f) \leq \deg(g)$ and note that we can invert (g mod f) in the ring $\mathbf{Q}[X]/(f)$. This gives $gg_1 = 1 - ff_1$ for $f_1, g_1 \in \mathbf{Q}[X]$ of degree at most $n-1$, and we get f_2, g_2 by multiplying out denominators of the coefficients. Now we plug in $x = a/b$ and multiply by b^{2n-1}, so that we get

$$(b^{n-1} f_2(a/b))A + (b^{n-1} g_2(a/b))B = b^{2n-1}R,$$

where A and B are as above. Since a $\mathbf{Z}$-linear combination of A and B is divisible by $\gcd(A,B)$ we see that the coprime-to-b-part of $\gcd(A,B)$ divides the constant R. The coefficients in front of A and B are at most $CH(x)^{n-1}$, for some $C>0$ depending only on f_2 and g_2. Thus, we have

$$|b|^{2n-1}R \le 2\max\{|A|,|B|\}CH(x)^{n-1},$$

which implies that

$$\max\{|A|,|B|\} \ge S|b|^{2n-1}/H(x)^{n-1} \quad \text{for some constant } S>0.$$

We now observe that the whole setup is symmetric in a and b, if we allow ourselves to change f and g. More precisely, if we let $f^*(X)=X^nf(1/X)$ and $g^*(X)=X^ng(1/X)$ be the reciprocal polynomials, then $A=a^nf^*(b/a)$ and $B=a^ng^*(b/a)$. Thus we see that coprime-to-a-part of $\gcd(A,B)$ also divides some fixed non-zero integer R^*, and that

$$\max\{|A|,|B|\} \ge S^*|a|^{2n-1}/H(x)^{n-1} \quad \text{for some constant } S^*>0.$$

Since $H(x)$ is equal to $|a|$ or to $|b|$ we conclude that

$$H(f(x)/g(x)) = H(A/B) = \frac{\max\{|A|,|B|\}}{\gcd(A,B)} \ge \frac{\min\{S,S^*\}}{|RR^*|}H(x)^n. \qquad \square$$

We now apply this lemma in the descent process—it implies that dividing a point by 2 on E significantly reduces the height of its x-coordinate. For the point 0_E we say that its x-coordinate is infinity, which has height 1.

7.3. Corollary. *There is a real number $C>0$ (depending on E) such that for all points $P,Q\in E(\mathbf{Q})$ with $P=2Q$ the heights of the x-coordinates satisfy*

$$H(x_Q) \le CH(x_P)^{1/4}.$$

Proof. We saw before that x_P can be expressed as a quotient of a degree 4 polynomial and a coprime degree 3 polynomial in x_Q. (See Exercise 8 in Section 1, or write it out with the formulas in section 5.1.) □

We need one more ingredient to do the descent argument. It is an estimate in the easy direction, but it involves both the x- and y-coordinate of a point.

7.4. Lemma. *For every $Q\in E(\mathbf{Q})$ there is a $C\in\mathbf{R}$ such that for all $P\in E(\mathbf{Q})$ we have*

$$H(x_{P+Q}) \le CH(x_P)^2.$$

Proof. We will use the addition formula from 5.1. For $Q = 0_E$ the statement is trivial, so assume that $Q = (x_0, y_0)$ is an affine point. We write $P = (x, y)$ and since we may exclude finitely many P we can assume that $x \neq x_0$. Then we have

$$x_{P+Q} = -x - x_0 + \left(\frac{y - y_0}{x - x_0}\right)^2 = \frac{-x^3 + y^2 - 2y_0 y + \text{degree 2 in } x}{\text{degree 2 in } x}$$

The term $-x^3 + y^2$ can be rewritten as a linear polynomial in x. Recall from the proof of Lemma 6.4 that we can write $(x, y) = (r/t^2, s/t^3)$ with r, s, $t \in \mathbf{Z}$ and $\gcd(r, t) = \gcd(s, t) = 1$. We obtain

$$x_{P+Q} = \frac{Nst + h_1(r, t^2)}{h_2(r, t^2)},$$

where N is an integer and h_1 and h_2 are homogeneous forms of degree 2 in 2 variables. Clearly $|h_i(r, t^2)|$ can be bounded by a constant times $H(x)^2$. We also know that $|t|$ is bounded by $H(x)^{1/2}$. It remains to show that $|s|$ is bounded by a constant times $H(x)^{3/2}$. This is another instance of Lemma 7.2:

$$|s^2| \leq H(y)^2 = H(y^2) = H(W(x)) \leq C_0 H(x)^3. \qquad \square$$

We are now ready to put the ingredients together and prove the following.

7.5. Proposition. *Assume that $E(\mathbf{Q})/2E(\mathbf{Q})$ is finite. Then $E(\mathbf{Q})$ is finitely generated as an abelian group.*

Proof. Choose a finite subset S of $E(\mathbf{Q})$ which represents all classes in $E(\mathbf{Q})/2E(\mathbf{Q})$. There is a constant C_1 so that for all $P \in E(\mathbf{Q})$ and $Q \in S$ we have $H(x_{P-Q}) \leq C_1 H(x_P)^2$. For each $P \in E(\mathbf{Q})$ there is a $Q \in S$ so that $P - Q = 2R$ for some $R \in E(\mathbf{Q})$. We then have

$$H(x_R) \leq CH(x_{P-Q})^{1/4} \leq CC_1^{1/2} H(x_P)^{1/2},$$

with C as in Corollary 7.3. If $H(x_R) > H(x_P)/2$ then it follows that $H(x_P) < B$, where $B = 4C_1C^2$. Let T be the set of points $P \in E(\mathbf{Q})$ for which $H(x_P) < B$. Since there are only finitely many rational numbers of height at most B and since there are at most two points in $E(\mathbf{Q})$ with a given x-coordinate, the set T is finite. We claim that the group generated by S and T contains every point $P \in E(\mathbf{Q})$. In order to show this with induction to the height of x_P one writes $P = Q + 2R$ for some $Q \in S$ and $R \in E(\mathbf{Q})$. Then either R lies in T, or we have $H(x_R) \leq H(x_P)/2$, in which case the induction hypothesis says that R lies in the group generated by S and T. □

Exercises.

1. For an integer $N \geq 0$ let R_N be the number of $x \in \mathbf{Q}$ with $H(x) \leq N$.
   (1) Show that $R_N \leq 2N(N+1)$.
   (2) *Show that
$$\lim_{n \to \infty} \frac{R_N}{N^2} = \frac{12}{\pi^2}.$$

2. For the curve $Y^2 = X^3 + 4$ give an explicit C for which the statement in Corollary 7.3 holds.

3. Let $f \in \mathbf{Q}[X]$ be a polynomial of degree at least 2. Show that there are only finitely many $x \in \mathbf{Q}$ for which $\{x, f(x), f(f(x)), \cdots\}$ is finite.

8. Reduction and torsion points

Let E be an elliptic curve over $\mathbf{Q}$ given by an equation

$$E: \quad Y^2 = W(X), \qquad W(X) = X^3 + aX + b \in \mathbf{Z}[X].$$

The discriminant of $W(X)$ is the non-zero integer $\Delta = -4a^3 - 27b^2$. In this section we show the theorem on Nagell-Lutz:

8.1. Theorem. *If $P = (x, y) \in E(\mathbf{Q})$ is an affine point of E of finite order, then $x, y \in \mathbf{Z}$ and either $y = 0$ or $y^2 \mid \Delta$.*

This gives us very explicit information on where to look for rational torsion points on E. The proof is based on the notion of reduction modulo a prime p.

8.2. The p-adic valuation on Q. Fix a prime number p. For an integer $x \in \mathbf{Z}$ we denote the number of factors p in x by $\text{ord}_p(x)$, with the convention that $\text{ord}_p(0) = +\infty$. For $x \in \mathbf{Q}$ we write $x = a/b$ with $a, b \in \mathbf{Z}$ and put $\text{ord}_p(x) = \text{ord}_p(a) - \text{ord}_p(b)$, which does not depend on the choice of a and b. We have the following properties, which express that $\text{ord}_p(\cdot)$ is a discrete valuation on $\mathbf{Q}$:

(1) $\text{ord}_p(xy) = \text{ord}_p(x) + \text{ord}_p(y)$ for all $x, y \in \mathbf{Q}$;
(2) $\text{ord}_p(x + y) \geq \min\{\text{ord}_p(x), \text{ord}_p(y)\}$ for all $x, y \in \mathbf{Q}$;
(3) $\text{ord}_p(x + y) = \min\{\text{ord}_p(x), \text{ord}_p(y)\}$ for all $x, y \in \mathbf{Q}$ with $\text{ord}_p(x) \neq \text{ord}_p(y)$.

It follows that the set $\mathbf{Z}_{(p)} = \{x \in \mathbf{Q}\colon\ \text{ord}_p(x) \geq 0\}$ is a subring of $\mathbf{Q}$ containing $\mathbf{Z}$. An element $x \in \mathbf{Z}_{(p)}$ is a unit if and only if $\text{ord}_p(x) = 0$. We now have a ring homomorphism $\mathbf{Z}_{(p)} \to \mathbf{F}_p$ called the reduction map that sends x to $\bar{x} = x \bmod p$.

8.3. The reduction map. We say that an n-tuple $(a_1, \ldots, a_n) \in \mathbf{Q}^n$ is primitive (with respect to p) if $\min\{\text{ord}_p(a_1), \ldots, \text{ord}_p(a_n)\} = 0$. We say that polynomial with coefficients in $\mathbf{Q}$ is primitive if its coefficients form a primitive tuple. We have a well-defined map $\mathbf{P}^n(\mathbf{Q}) \to \mathbf{P}^n(\mathbf{F}_p)$ which for primitive $(a_0, \ldots, a_n) \in \mathbf{Q}^{n+1}$ sends $(a_0 : \ldots : a_n)$ to $(\bar{a}_0 : \ldots : \bar{a}_n) \in \mathbf{P}^n(\mathbf{F}_p)$. Now suppose that $n = 2$ and that we have a curve C given by a homogeneous polynomial $F(X, Y, Z) \in \mathbf{Q}[X, Y, Z]$. We can multiply F by a non-zero scalar so that F becomes primitive. Then we let $\bar{C}$ be the curve over $\mathbf{F}_p$ given the equation that we get from F by reducing its coefficients modulo p. For example, the line L: $X + Y/2 + Z/2 = 0$ reduces to the line $\bar{L}$: $Y + Z = 0$ over $\mathbf{F}_2$.

For our elliptic curve this means that we have a map $E(\mathbf{Q}) \to \bar{E}(\mathbf{F}_p)$. In order to see how this map behaves with respect to the group operation on $E(\mathbf{Q})$ we need the following Lemma.

8.4. Lemma. *Suppose that C and L are a curve and a line over $\mathbf{Q}$, and suppose that $\bar{C}$ does not contain $\bar{L}$. Suppose also that $C \cdot L = [P_1] + \cdots + [P_d]$ for rational points P_i. Then we have $\bar{C} \cdot \bar{L} = [\bar{P}_1] + \cdots + [\bar{P}_d]$.*

Proof. We refer to Section 4.3 and 4.4 for terminology and notation. Choose two points $(v_1 : v_2 : v_3)$ and $(w_1 : w_2 : w_3)$, with (v_1, v_2, v_3) and (w_1, w_2, w_3) primitive, which lie on L and which do not have the same reduction modulo p. Then we can write $P_i = (a_i v_1 + b_i w_1 : a_i v_2 + b_i w_2 : a_i v_3 + b_i w_3)$ for primitive (a_i, b_i). Let C be given by a primitive form $F(X, Y, Z)$. Then we have

$$F(Uv_1 + Vw_1, Uv_2 + Vw_2, Uv_3 + Vw_3) = c\prod_{i=1}^{d}(b_i U - a_i V),$$

for some scalar $c \in \mathbf{Q}^*$. If $\mathrm{ord}_p(c) \geq 1$ then we would have $\bar{L} \subset \bar{C}$, which contradicts the assumption. If $\mathrm{ord}_p(c) \leq -1$ then the product of the non-zero polynomials $\bar{b}_i U - \bar{a}_i V$ would be zero in $\mathbf{F}_p[U, V]$, which contradicts the fact that $\mathbf{F}_p[U, V]$ is a domain. Thus, $\mathrm{ord}_p(c) = 0$, and reducing the whole equation modulo p we get the result. □

8.5. Corollary. *If $p \nmid 2\Delta$ then $\bar{E}$ is an elliptic curve over $\mathbf{F}_p$ and the reduction map $E(\mathbf{Q}) \to \bar{E}(\mathbf{F}_p)$ is a group homomorphism.*

Proof. For odd p the only non-smooth points of $\bar{E}$ are the points $(x, 0)$, where $x \in \mathbf{F}_p$ is a double root of $\bar{W}$. Such an x only exists when $p \mid \Delta(\bar{W}) = \bar{\Delta}$. So for $p \nmid 2\Delta$ the curve $\bar{E}$ is an elliptic curve, and $\bar{E}(\mathbf{F}_p)$ is a group. Lemma 8.4 implies that the reduction map is a homomomorphism. □

8.6. Kernel of reduction. For a prime number p and an integer $n \geq 1$ we let E_n be the "kernel of reduction" modulo p^n:

$$E_n = E_n^{(p)} = \{(x : y : z) \in E(\mathbf{Q}) : \quad y \in \mathbf{Z}_{(p)}^* \text{ and } x, z \in p^n \mathbf{Z}_{(p)}\}.$$

We saw that E_1 is the kernel of a group homomorphism if $p \nmid 2\Delta$. We now show that for arbitrary p and n the set E_n is a subgroup of $E(\mathbf{Q})$. We will use that $(0 : 1 : 0)$ is a smooth point of $\bar{E}$, even if $\bar{E}$ is not smooth.

8.7. Proposition. *Fix a prime number p and let $n \geq 1$. Then the following hold:*
(1) *for every $P = (x : y : z) \in E_1$ we have $y \neq 0$ and*

$$P \in E_n \iff \mathrm{ord}_p(x/y) \geq n \iff \mathrm{ord}_p(z/y) \geq 3n;$$

(2) *the subset E_n of $E(\mathbf{Q})$ is a subgroup;*
(3) *the map $E_n \to \mathbf{Z}_{(p)}$ given by $(x : y : z) \mapsto x/y$ induces a group homomorphism*

$$E_n/E_{n+2} \longrightarrow p^n \mathbf{Z}_{(p)}/p^{n+2}\mathbf{Z}_{(p)} \cong \mathbf{Z}/p^2\mathbf{Z}.$$

Proof. A rational point on E is always of the form $P = (rt : s : t^3)$ for r, s, $t \in \mathbf{Z}$ with $\gcd(r, t) = \gcd(s, t) = 1$. If $P \in E_1$ then $p \mid t$, so $p \nmid r$ and $p \nmid t$, and (1) follows.

Let L be a line with $L \cdot E = [P_1] + [P_2] + [P_3]$, where $P_i = (x_i : y_i : z_i) \in E(\mathbf{Q})$, and suppose that $P_1, P_2 \in E_n$. By Lemma 8.4 the line $\bar{L}$ is a tangent of $\bar{E}$ in $(0 : 1 : 0) \in \bar{E}(\mathbf{F}_p)$. But $(0 : 1 : 0)$ is a smooth point of $\bar{E}(\mathbf{F}_p)$ with flex line $Z = 0$, so $\bar{L}$ is the line $Z = 0$, and

$\bar{P}_3 = (0 : 1 : 0)$. It follows that $P_3 \in E_1$, and that L has a homogeneous equation of the form

$$L: \quad Z = \alpha X + \beta Y,$$

with α, $\beta \in p\mathbf{Z}_{(p)}$. Since $P_1 \in E_n$ we see with (1) that $\beta = z_1/y_1 - \alpha x_1/y_1 \in p^{n+2}\mathbf{Z}_{(p)}$. If we substitute the equation for L in the equation for E we get

$$Y^2(\alpha X + \beta Y) = X^3 + aX(\alpha X + \beta Y)^2 + b(\alpha X + \beta Y)^3.$$

Dividing by Y^3 and putting $T = X/Y$ we get a cubic equation

$$(1 + a\alpha^2 + b\alpha^3)T^3 + (2a\alpha\beta + 3b\alpha^2\beta)T^2 + \text{ lower order terms in } T = 0.$$

The roots of this equation are the numbers x_i/y_i, so

$$(2a\alpha\beta + 3b\alpha^2\beta) = (1 + a\alpha^2 + b\alpha^3)\left(\frac{x_1}{y_1} + \frac{x_2}{y_2} + \frac{x_3}{y_3}\right).$$

The left hand side lies in $p^{n+1}\mathbf{Z}_{(p)}$ and since $1 + a\alpha^2 + b\alpha^3 \in \mathbf{Z}_{(p)}^*$ we deduce that

$$(*) \qquad \frac{x_1}{y_1} + \frac{x_2}{y_2} + \frac{x_3}{y_3} \in p^{n+2}\mathbf{Z}_{(p)}.$$

By (1) we have x_1/y_1, $x_2/y_2 \in p^n\mathbf{Z}_{(p)}$, and it follows that $x_3/y_3 \in p^n\mathbf{Z}_{(p)}$. We knew already that $P_3 \in E_1$ so again with (1) we see that $P_3 \in E_n$. Thus, the point $P_1 + P_2 = -P_3$ also lies in E_n and (2) follows. Finally, $(*)$ implies that the map in (3) is a homomorphism. □

8.8. Corollary. *If $P \in E_1$ has finite order then $P = 0_E$.*

Proof. Suppose that $P \in E_1$ has prime order. There is an $n \geq 1$ so that $P \in E_n$ but $P \notin E_{n+1}$. Under the homomorphism in (3) above P maps to 0 or to an element of prime order. But the only elements in $p^n\mathbf{Z}_{(p)}/p^{n+2}\mathbf{Z}_{(p)}$ of prime order are in $p^{n+1}\mathbf{Z}_{(p)}/p^{n+2}\mathbf{Z}_{(p)}$, and by (1) above this means that $P \in E_{n+1}$: contradiction. Thus, E_1 has no elements of prime order, so its torsion subgroup must be trivial. □

8.9. Corollary. *If $p \nmid 2\Delta$ then the homomorphism $E(\mathbf{Q})_{\text{tor}} \to \bar{E}(\mathbf{F}_p)$ is injective.*

Proof. The kernel of the map is exactly $E_1 \cap E(\mathbf{Q})_{\text{tor}} = (E_1)_{\text{tor}} = \{0_E\}$. □

Proof of Theorem 8.1. Suppose that $P = (x, y) \in E(\mathbf{Q})$ is a torsion point. If x or y has a denominator which is divisible by p, then the reduction $\bar{P} \in \bar{E}(\mathbf{F}_p)$ lies on the line $Z = 0$. Since this line intersects $\bar{E}$ only in $(0{:}\,1{:}\,0) \in \bar{E}(\mathbf{F}_p)$, it follows that $P \in E_1$. But we just showed that E_1 is torsion free, so $x, y \in \mathbf{Z}$.

If $2P = 0_E$ then we have $y = 0$ and we are done. Otherwise, the point $2P = (x', y')$ is also an affine torsion point so $x', y' \in \mathbf{Z}$. The tangent line at P has slope $\lambda = W'(x)/2y$ where W' is the derivitive of $W(X) = X^3 + aX + b$. Recall that we have $x' + 2x = -\lambda^2$ so we see that $\lambda \in \mathbf{Z}$ and $y \mid W'(x)$ Thus the polynomial

$$f(T) = W(T + x) = W(x) + W'(x)T + \text{ higher order } = c_0 + c_1T + c_2T^2 + T^3$$

satisfies $y^2 \mid c_0$ and $y \mid c_1$. This implies that

$$y^2 \mid -27c_0^2 + 18c_0c_1c_2 - 4c_0c_2^3 - 4c_1^3 + c_1^2c_2^2 = \Delta(f) = \Delta(W). \qquad \square$$

Exercises.

1. Find the groups of rational torsion points for
   (1) $Y^2 = X^3 + 1$;
   (2) $Y^2 = X^3 - 43X + 166$;
   (3) $Y^2 = X^3 - 219X + 1654$;
   (4) $Y^2 = X(X-1)(X+2)$.

2. *Extend theorem 8.1 to elliptic curves E: $Y^2 = X^3 + aX^2 + bX + c$ with a, b, $c \in \mathbf{Z}$ but $a \neq 0$.

3. For $p = 2$, can it happen that $\bar{E}$ is smooth?

4. Fix a prime number p. Let $\bar{E}_{\mathrm{ns}}(\mathbf{F}_p)$ be the set of non-singular points of $\bar{E}(\mathbf{F}_p)$, and let $E_0 = \{P \in E(\mathbf{Q}) : \bar{P} \in \bar{E}_{\mathrm{ns}}(\mathbf{F}_p)\}$. Show that
   (1) $\bar{E}_{\mathrm{ns}}(\mathbf{F}_p)$ has a natural group structure;
   (2) E_0 is a subgroup of E;
   (3) the sequence $0 \to E_1 \to E_0 \to \bar{E}_{\mathrm{ns}}(\mathbf{F}_p)$ is exact.

5. Consider the elliptic curve E: $Y^2 = X^3 + a$ with $a \in \mathbf{Z}$.
   (1) Show that for $p \equiv 2 \bmod 3$ with $p \nmid a$ we have $\#\bar{E}(\mathbf{F}_p) = p + 1$.
   (2) Show that $\#E(\mathbf{Q})_{\mathrm{tor}} \mid 6$.

6. For $x, y \in \mathbf{Q}$ and let $d_p(x,y) = p^{-\operatorname{ord}_p(x-y)}$. Show that $\mathbf{Q}$ is a metric space with metric d_p. Show that the completion $\mathbf{Q}_p$ has a natural field-structure—it is called the field of p-adic numbers. Show that $\mathbf{Q}_p$ is locally compact. Show that $1 + p + p^2 + \cdots = 1/(1-p)$ in $\mathbf{Q}_p$.

Sinterklaasopgaven.

1. *(Fermat aan Mersenne, 1643)* Vind een rechthoekige driehoek met geheeltallige zijden zodat de som $a+b$ van de rechthoekszijden en de hypotenusa c kwadraten zijn.

   a. Laat zien dat een niet-triviale oplossing (a,b,c) aanleiding geeft tot een punt op de affiene kromme
   $$C : Y^2 = 2X^4 - 1.$$
   [Hint: neem $z = a - b$ en $x^2 = c$, dan geldt $2x^4 - y^4 = z^2$.]

   b. Geef een birationele equivalentie aan tussen C en de elliptische kromme $E : y^2 = x^3 + 8x$. [Hint: in nieuwe variabelen $u = \frac{1}{X-1}$ en $v = \frac{Y}{(X-1)^2}$ komt het punt $(1,1) \in C$ in oneindig te liggen, dus
   $$v^2 = u^4 + 8u^3 + 12u^2 + 8u + 2 = (u^2 + 4u - 2)^2 + 24u - 2.$$
   Neem nu $T = v + (u^2 + 4u - 2)$, vermenigvuldig met T en schrijf $Tu = S$ om de vergelijking $T^3 - 2S^2 - 8ST + 4T^2 - 24S + 2T = 0$ te krijgen.]

   c. Laat zien dat $P = (1,3) \in E(\mathbf{Q})$ oneindige orde heeft, en vind een punt op E dat aanleiding geeft tot een oplossing van Fermats probleem.
   [Hint: het punt P correspondeert met het triviale punt $(-1,1)$ op C, en $2P$ geeft een negatieve waarde voor b. Maar $4P$ werkt!]

2. *(Fermat voor exponent* 4*)* Bewijs: voor alle geheeltallige oplossingen van de vergelijking $x^4 + y^4 = z^4$ geldt $xyz = 0$.
   [Hint: voor $u = x/y$ en $v = z^2/y^2$ geldt $v^2 = u^4 + 1$. De coördinatentransformatie $u \mapsto u$ en $v \mapsto v + u^2$ geeft een kubische kromme $v^2 + 2u^2v = 1$ met flex in oneindig.]

3. Laat zien dat $P = (-2,8)$ oneindige orde heeft in $E(\mathbf{Q})$ voor $E : y^2 = x^3 - 36x$, en bepaal de structuur van $E(\mathbf{Q})$. Welk punt op E correspondeert met de Pythagoreïsche 3-4-5-driehoek die laat zien dat 6 een congruent getal is?

4. Bepaal de rang van $E : y^2 = x^3 - 49x$ en geef een Pythagoreïsche driehoek met oppervlakte 7.
   [Hint: $(25,120) \in E(\mathbf{Q})$.]

9. Function fields, local rings, and morphisms of curves

Sofar, we mainly dealt with elliptic curves over subfields of $\mathbf{C}$, and we were able to use complex analysis to show many properties. This section is devoted to some algebraic geometry that we need to deal with elliptic curves in characteristic $p > 0$. This section is much more sketchy than the last 8 sections: it is just an account of what was stated and what was proved in the lectures. We refer to Lang's *Algebra* for resultants, field theory (separable, inseparable and transcendental extensions) and Nakayama's lemma. In Silverman's *Arithmetic of Elliptic Curves* one finds a more extensive treatment of the material in this section, but he does often refer elsewhere for proofs.

9.1. Function fields and local rings. If a curve C over a field K is given by an irreducible homogeneous equation $F(X,Y,Z) = 0$ then the function field $K(C)$ consists of equivalence classes of pairs (f,g) with $f,g \in K[X,Y,Z]$ homogeneous of the same degree, and $F \nmid g$. Here we say that $(f,g) \sim (f',g')$ if $F \mid fg' - gf'$. The equivalence class of (f,g) is denoted by f/g. The field operations on $K(C)$ are defined in the obvious way, e.g., $f/g + f'/g' = (fg' + gf')/gg'$. We say that $\varphi \in K(C)$ has no pole at P if $\varphi = f/g$ with $g(P) \neq 0$, and we then put $\varphi(P) = f(P)/g(P) \in K$. The collection of such φ is called the local ring at P, and we denote it by $\mathcal{O}_P$. The evaluation map $\varphi \mapsto \varphi(P)$ is a ring homomorphism $\mathcal{O}_P \to K$. Its kernel $\mathfrak{m}_P$ is a maximal ideal, and $\mathcal{O}_P = \mathfrak{m}_P \cup \mathcal{O}_P^*$. We showed that $\mathfrak{m}_P$ can always be generated as an $\mathcal{O}_P$-ideal by two elements. With Nakayama's lemma we showed that it can be generated by one element if and only if P is a smooth point.

9.2. Resultants. For a commutative ring A and two polynomials $f,g \in A[X]$ of degree n and m we introduced the resultant $R = R_X(f,g) \in A$ as a determinant of size $n+m$. We showed three properties:

(1) $R \in (f,g)$;
(2) If A is a UFD and $R = 0$ then f and g have a factor in common which is not a unit.
(3) If $A = A_0[X_1,\ldots,X_n]$ and f and g are also homogeneous in $X_1,\ldots,X_n,X$ of degrees n and m, then R is homogeneous of degree nm in $X_1,\ldots,X_n$.

We deduced a number of corollaries:

(1) Bezout: suppose we are given two curves C_1, C_2 by homogeneous equations F_1, F_2 of degree n and m, and suppose that F_1 and F_2 have no non-constant factor in common. Then $C_1(\bar{K}) \cap C_2(\bar{K})$ is non-empty, and it has cardinality at most nm.

(2) A curve C over a field K which is irreducible over $\bar{K}$ only has finitely many non-smooth points over $\bar{K}$.

(3) If K is a field and $f,g \in K[x,y]$ have no non-constant factor in common, then $K[x,y]/(f,g)$ is finite dimensional as a vector space over K.

It also follows that a rational function f on an irreducible curve C over a field K has only finitely many poles, i.e., $f \in \mathcal{O}_P$ for all but finitely many $P \in C(K)$.

9.3. Discrete valuation rings. With property (3) above and Nakayama's lemma we deduced that for smooth $P \in C(K)$ the ring $\mathcal{O}_P$ is a DVR, i.e, there is a $\pi \in \mathcal{O}_P$ such that for all $f \in K(C)^*$ there is a unique $n \in \mathbf{Z}$ such that $f \in \pi^n \mathcal{O}_P^*$. This n is independent of the choice of π and we write $n = \mathrm{ord}_P(f)$.

9.4. Rational maps and morphisms. We now introduce rational maps between two irreducible curves C_1 and C_2 over K. We define the set of rational maps from C_1 to C_2 to be $\mathrm{Rat}(C_1, C_2) = C_2(K(C_1))$. A rational map $\varphi = (\varphi_1 \colon \varphi_1 \colon \varphi_2) \in C_2(K(C_1))$ is said to be *defined* at $P \in C_1(K)$ if there is a $\lambda \in K(C_1)$ such that the three functions $\lambda\varphi_i$ are all in $\mathcal{O}_P$, but not all in $\mathfrak{m}_P$. We then let $\varphi(P) = ((\lambda\varphi_0)(P) : (\lambda\varphi_1)(P) : (\lambda\varphi_2)(P))$. By the discrete valuation property, such a λ always exists if P is a smooth point. We say that φ is a morphism if it is defined at all $P \in C_1(\bar{K})$. In particular, every rational map $C_1 \to C_2$ is a morphism if C_1 is itself smooth.

Now assume that C_1 is irreducible over $\bar{K}$. For a nonconstant rational map φ from C_1 to C_2 we showed (up to exercise 9) that we have an induced K-linear field homomorphisms $\varphi^* \colon K(C_2) \to K(C_1)$. Conversely, every K-linear field homomorphism $K(C_2) \to K(C_1)$ occurs this way.

9.5. Facts from algebraic geometry. Let $\varphi \in \mathrm{Rat}(C_1, C_2)$ with C_1 and C_2 irreducible over $\bar{K}$. By general field theory which we did not go into (transcendence theory) the field extension $K(C_1)/\varphi^*(K(C_2))$ has finite degree $d = d_s d_i$, where d_s is the degree of the separable part. A general result about the algebraic geometry of curves says that for almost all points $Q \in C_2(\bar{K})$ the fiber $\varphi^{-1}(Q) = \{P \in C_1(\bar{K}) : \varphi$ defined at P and $\varphi(P) \in Q\}$ consists of exactly d_s points.

9.6. Morphisms between elliptic curves. Since elliptic curves are smooth, rational maps between them are morphisms. We say that a rational map $E_1 \to E_2$ between elliptic curves over a field K is an *isogeny* if $0 \mapsto 0$, and the set of isogenies is written $\mathrm{Hom}(E_1, E_2)$. We have the following three properties.

(1) For $P \in E(K)$ there is a translation map $t_P : E \to E$ sending Q to $P + Q$.

(2) The group structure on $E_2(K(E_1))$ gives $\mathrm{Rat}(E_1, E_2)$ the structure of an abelian group and for $P \in E_1(K)$ we have $(\varphi + \psi)(P) = \varphi(P) + \psi(P)$.

(3) For an isogeny $\varphi \colon E_1 \to E_2$ the map $E_1(K) \to E_2(K)$ is a group homomorphism.

The proof of (1) is in exercise 11.

The proof of (2) is exactly like the proof that the reduction map $E(\mathbf{Q}) \to E(\mathbf{F}_p)$ is a group homomorphism when E has good reduction at p. The role of the discrete valuation ring $\mathbf{Z}_{(p)}$ is now played by $\mathcal{O}_P$.

The proof of (3) is still a bit difficult for us. One can sketch it like this: if $R = P + Q$ in $E_1(K)$ then there is a function $f \in K(E_1)$ with divisor $[P] + [Q] - [R] - [0]$. Taking the field norm of f down to $K(E_2)$ (via φ^*) gives a function with divisor $[\varphi(P)] + [\varphi(Q)] - [\varphi(R)] - [0]$, so $\varphi(P) + \varphi(Q) = \varphi(R)$.

Exercises.

1. Laat K een lichaam zijn van karakteristiek $p > 0$. Stel dat $a \in K$ een element is dat geen kwadraat is in K. Laat $C \subset \mathbf{P}^2$ de kromme over K zijn die gegeven is door $X^2 - aY^2 = 0$.
   a. Bewijs dat C irreducibel is, en bepaal $C(K)$.
   b. Laat zien dat $x = X/Z \in K(C)$, en dat $K(C)$ een kwadratische uitbreiding is van het lichaam $K(x)$ (= quotientenlichaam van $K[x]$). Is deze uitbreiding separabel?
   c. Bewijs dat C niet irreducibel is over $\bar{K}$, de algebraische afsluiting van K.
   d. Bepaal de gladde punten van $C(\bar{K})$. (Onderscheid twee gevallen: karakteristiek 2 en niet-2.)

2. Een lichaam K van karakteristiek p heet *perfect* als $p = 0$ of $K^p = K$. Hier is $K^p = \{x^p \colon x \in K\}$.
   a. Bewijs dat een lichaam K perfect is dan en slechts dan als alle eindige uitbreidingen van K separabel zijn.
   b. Laat zien dat elk lichaam een perfecte afsluiting K_{pf} heeft met de eigenschap dat elk lichaamshomomorfisme $K \to L$ met L perfect, te schrijven is als compositie $K \subset K_{\mathrm{pf}} \xrightarrow{\varphi} L$ voor een unieke φ.
   c. Laat zien dat deze perfecte afsluiting uniek uniek is, dat wil zeggen uniek op uniek isomorfisme na.
   d. Voor welke lichamen K is de algebraische afsluiting van K uniek uniek?

3. Laat p een priemgetal zijn en laat $K = \mathbf{F}_p(X,Y)$ het quotientenlichaam zijn van de polynoomring $\mathbf{F}_p[X,Y]$.
   a. Bewijs dat K een uitbreiding is van $K^p = \mathbf{F}_p(X^p, Y^p)$ van graad p^2.
   b. Bewijs dat er oneindig veel tussenlichamen zijn van de uitbreiding $K^p \subset K$.

4. Laat K een lichaam zijn en $L = K(\alpha)$ een lichaamsuitbreiding van K van eindige graad. Een *derivatie* van L over K is een K-lineaire afbeelding $d : L \to L$ die voldoet aan $dxy = xdy + ydx$. Bewijs dat L/K separabel is dan en slechts dan als $d = 0$ de enige derivatie van L over K is.

5. Let K be a field and let C be the line over K in $\mathbf{P}^2$ given by $X = 0$.
   a. Show that $K(C)$ is isomorphic to the quotient field $K(t)$ of the polynomial ring $K[t]$ with $Y/Z \in K(C)$ mapping to t.
   b. For algebraically closed K show that the intersection within $K(C)$ of the rings $\mathcal{O}_P$, with P ranging over $C(K)$, is equal to K.

6. Let $f = (X - a_1) \cdots (X - a_n)$ and $g = (X - b_1) \cdots (X - b_m)$. Then $R(f,g)$ is a polynomial expression in the variables $a_1, \ldots, a_n, b_1, \ldots, b_m$.
   a. Show that this expression is homogeneous and compute its degree.
   b. Show that $R(f,g) = \prod_{i=1}^{n} \prod_{j=1}^{m} (a_i - b_j)$.
   c. If we define the discriminant of f as $\Delta(f) = \prod_{i<j} (a_i - a_j)^2$, then show that $\Delta(f) = \pm R(f, f')$, and determine the sign.

7. Let R be a noetherian domain which is not a field, and let $K = Q(R)$ be its quotient field. Show that R is a DVR if and only if for all $x \in K$ we have $x \in R$ or $x^{-1} \in R$.

8. Let C be an irreducible curve over K. Suppose that $C(K)$ contains a smooth point P. Let $f \in K(C)$ be algebraic over K.

a. Show that $f \in \mathcal{O}_P$.
b. Show that $f \in K$.

9. The object of this exercise is to prove that K is algebraically closed in $K(C)$ without assuming existence of a smooth K-valued point. Let C be a curve over K which is irreducible over K. Suppose that C is not the line at infinity, and that it is given by an affine equation $f(x, y) = 0$. For an algebraic extension L of K let $\mathcal{O}_L$ be the quotient ring $L[x, y]/(f)$. This is a domain, and its quotient field is the function field $L(C)$.
   a. For an ideal I in $K[x, y]$, let $\bar{I}$ be the ideal of $\bar{K}[x, y]$ generated by I. Show that $\bar{I} \cap K[x, y] = I$, and that a K-basis of $K[x, y]/I$ is a $\bar{K}$-basis of $\bar{K}[x, y]/\bar{I}$.
   b. Show that in $\bar{K}(C)$ we have $\mathcal{O}_{\bar{K}} \cap \bar{K} = K$.
   c. For $\varphi = a/b \in K(C) \cap \mathcal{O}_{\bar{K}}$ with $a, b \in \mathcal{O}_K$ show that $\varphi \in \mathcal{O}_K$ by showing that $a \in (b)$.
   d. Deduce that K is algebraically closed in $K(C)$.

10. In this exercise we show that rational functions and rational maps are determined by the induced maps on sets of $\bar{K}$-valued points.
   a. Let C be an irreducible curve over a field K. Show that for two distinct $f_1, f_2 \in K(C)$ there are infinitely many points $P \in C(\bar{K})$ with $f_1, f_2 \in \mathcal{O}_P$ and $f_1(P) \neq f_2(P)$.
   b. Let C_1, C_2 be irreducible curves over an algebraically closed field $\bar{K}$. Show that every element $\varphi \in \mathrm{Rat}(C_1, C_2)$ is determined by where it maps the smooth $\bar{K}$-valued points of C_1.

11. Let E: $y^2 = x^3 + ax + b$ be an elliptic curve over a field K, and let $P \in E(K)$.
   a. Deduce with the addition formulas that there is a unique morphism t_P: $E \to E$ sending $Q \in E(\bar{K})$ to $P + Q \in E(\bar{K})$ when $Q \neq -P$.
   b. Show that $t_P \circ t_Q = t_{P+Q}$ for all $P, Q \in E(K)$.
   c. Show that $t_P(-P) = 0_E$.