Elliptic Functions and Elliptic Curves

ETH Zürich

Dr. Markus Schwagenscheidt

Spring Term 2022 - Last update: May 30, 2022

Contents

1 Introduction

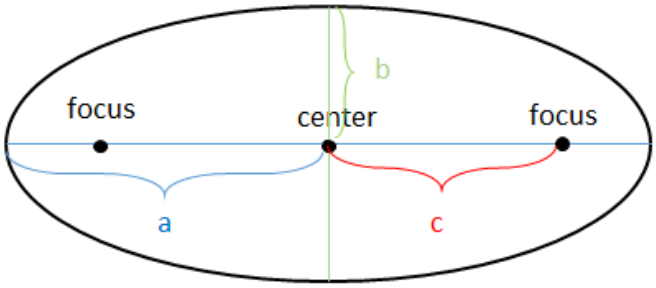
1.1 Elliptic integrals

The theory of elliptic functions historically emerged from the study of elliptic integrals. These integrals appear in the computation of the arc lenght of an ellipse and similar curves.

Example 1.1.1 (Arc length of an ellipse)**.** For $a, b > 0$ with $a \geq b$ we consider the ellipse given by all $(x, y) \in \mathbb{R}^2$ with

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

The points $(\pm c, 0)$ with $c := \sqrt{a^2 - b^2}$ are its *foci*, that is, for any point on the ellipse, the sum of the two distances to the foci is a constant. The ellipse is centered at the origin, with width $2a$ and height $2b$.



We would like to compute the arc length of the ellipse. For example, if $r = a = b$, then the ellipse is just a circle and it is well-known that its arc length is $2\pi r$. However, for a general ellipse there is no such simple closed formula.

Due to the symmetry of the ellipse, its arc length is 4 times the arc length of the part of the ellipse in the first quadrant. Recall that the arc length of a smooth curve C is given by $\int_a^b |\gamma'(t)|dt$, where $\gamma : [a, b] \to C$ is a parametrization of C. we can take the parametrization $\gamma(\varphi) = (a\cos(\varphi), b\sin(\varphi))$ with $\varphi \in [0, \pi/2]$ for the arc of the ellipse in the first quadrant, so the arc length of the ellipse is given by

$$4\int_0^{\pi/2} \sqrt{a^2\sin(\varphi)^2 + b^2\cos(\varphi)^2}d\varphi.$$

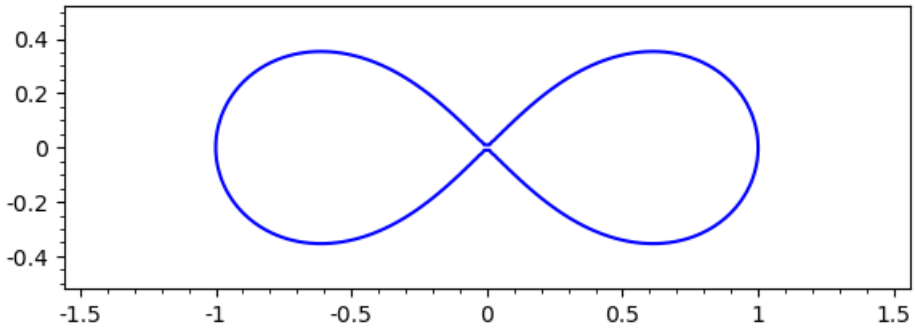
Replacing $t = \sin(\varphi)$, we obtain the arc length

$$4\int_0^1 \frac{\sqrt{a^2t^2 + b^2(1-t^2)}}{\sqrt{1-t^2}}dt,$$

which is an integral of a rational function in square roots of polynomials. This is an example of an *elliptic integrals* (are more precise definition will be given soon), which typically does not have a nice closed form.

Example 1.1.2 (Arc length of the lemniscate)**.** The lemniscate is given by the equation

$$(x^2 + y^2)^2 = x^2 - y^2.$$



Again, by symmetry its arc length is 4 times the arc length of its part in the first quadrant. This piece is parametrized by

$$\gamma(\varphi) = \left(\frac{\cos(\varphi)}{1 + \sin(\varphi)^2}, \frac{\sin(\varphi)\cos(\varphi)}{1 + \sin(\varphi)^2} \right), \quad \varphi \in [0, \pi/2].$$

A short computation now shows that the arc length of the lemniscate is given by

$$4 \int_0^1 \frac{1}{\sqrt{1 - t^4}} dt,$$

which is again an integral of a rational function in square roots of polynomials. This integral can in fact be written in terms of special values of the Gamma function. However, it would be desirable to also compute the arc length of a piece of the lemniscate. For example, the arc length from the origin to a point $\gamma(x)$ with $x \in [0, 1]$ is given by

$$\int_0^x \frac{1}{\sqrt{1 - t^4}} dt,$$

and this function does not have a closed expression in terms of simpler functions.

These examples lead to the following definition.

Definition 1.1.3. An *elliptic integral* is an integral of the form

$$\int R\left(x, \sqrt{p(x)}\right) dx$$

where $R(x, y)$ is a rational function of two variables, and $p(x)$ is a polynomial of degree 3 or 4 without multiple roots.

1.2 Fagnano's elliptic integral

For $x \in [0, 1]$ put

$$F(x) = \int_0^x \frac{1}{\sqrt{1 - t^4}} dt,$$

which measures the arc length the of the lemniscate from the origin to the point $\gamma(x)$, compare Example 1.1.2. This elliptic integral has some unexpected properties, which were first observed by Fagnano around 1750.

Theorem 1.2.1 (Fagnano ~ 1750)**.** *For all sufficiently small $x \geq 0$ we have*

$$2F(x) = F\left(2x \cdot \frac{\sqrt{1-x^4}}{1+x^4}\right).$$

Put $\sigma := F(1)$. The function $F : [0,1] \to [0,\sigma]$ is strictly increasing and continuous, and hence has a strictly increasing inverse function $G : [0,\sigma] \to [0,1]$. It satisfies the differential equation

$$G'^2 = 1 - G^4, \quad G(0) = 0, \quad G'(0) = 1. \tag{1.2.1}$$

Fagnano's Theorem can be restated for G as follows.

Theorem 1.2.2. *For all u small enough we have*

$$G(2u) = \frac{2G(u)G'(u)}{1+G^4(u)}.$$

Euler generalized Fagnano's duplication formula and obtained the following addition law for $F(x)$.

Theorem 1.2.3 (Euler 1761)**.** *For sufficiently small $x, y \geq 0$ we have*

$$F(x) + F(y) = F\left(\frac{x\sqrt{1-y^4} + y\sqrt{1-x^4}}{1+x^2y^2}\right).$$

This translates into an addition law for $G(u)$ as follows.

Theorem 1.2.4. *For sufficiently small $u, v \geq 0$ we have*

$$G(u+v) = \frac{G(u)G'(v) + G(v)G'(u)}{1+G^2(u)G^2(v)}. \tag{1.2.2}$$

Although the function G was only defined on the interval $[0,\sigma]$, one may extend it to a meromorphic function on $\mathbb{C}$, by taking G as the unique solution to the differential equation (1.2.1). The differential equation also implies that

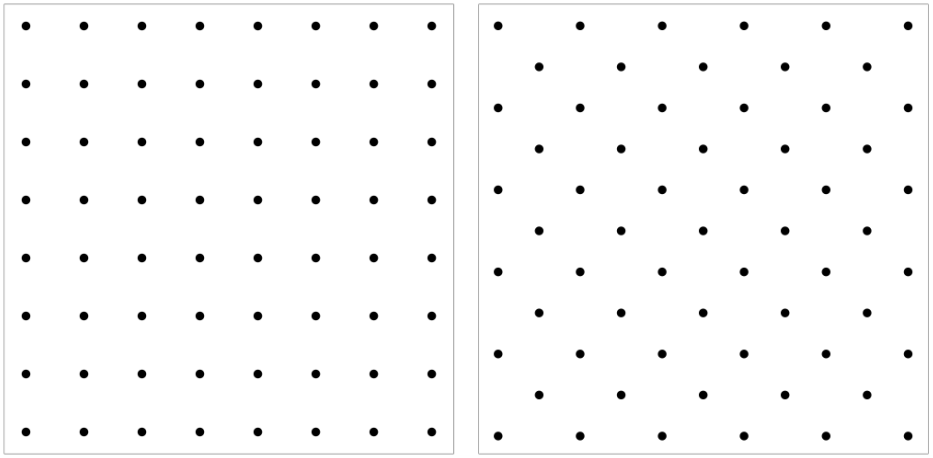
$$G(iu) = iG(u). \tag{1.2.3}$$

Moreover, Euler's Theorem (1.2.2) for $G(u)$ shows that G is *doubly periodic*. Indeed, for $v \in \mathbb{C}$ with $G'(v) = 1$ we have $G(v) = 0$ by the differential equation (1.2.1), hence $G(u+v) = G(u)$ for any $u \in \mathbb{C}$ by Euler's addition law (1.2.2). Moreover, by (1.2.3), for any such v, we will also have $G(u+iv) = G(u)$ for any $u \in \mathbb{C}$.

This led to the definition of elliptic functions.

Definition 1.2.5. A meromorphic function f on $\mathbb{C}$ is called *elliptic* (or *doubly periodic*) if there are $w_1, w_2 \in \mathbb{C}$ which are linearly independent over $\mathbb{R}$, such that $f(u+w_1) = f(u)$ and $f(u+w_2) = f(u)$ for all $u \in \mathbb{C}$.

It follows that an elliptic function satisfies the periodicity $f(u+w) = f(u)$ for all $w \in \mathbb{Z}w_1 + \mathbb{Z}w_2$. The set $\mathbb{Z}w_1 + \mathbb{Z}w_2$ is called a *lattice* in $\mathbb{C}$. They typically look as follows:



In this lecture, we will closely study elliptic functions, their period lattices, and the connection to elliptic curves.

2 Periods and lattices

We will closely follow the first part of the book *Elliptische Funktionen und Modulformen* by Koecher and Krieg.

2.1 Periods of meromorphic functions

Definition 2.1.1. A function $f : \mathbb{H} \to \mathbb{C}$ is *meromorphic* on $\mathbb{C}$ if there exists a closed, discrete subset $D_f \subset \mathbb{C}$ such that

1. $f : \mathbb{C} \setminus D_f \to \mathbb{C}$ is holomorphic, and
2. f has poles at the points of D_f.

Recall that a closed subset $D \subset \mathbb{C}$ is called discrete if every $c \in D$ has a neighbourhood U such that $D \cap U$ is finite.

If $f \neq 0$ is meromorphic on $\mathbb{C}$, then for each point $c \in \mathbb{C}$ there exists a neighbourhood U of c such that f has a Laurent expansion

$$f(z) = \sum_{n=n_0}^{\infty} a_{f,c}(n)(z-c)^n$$

on $U \setminus \{c\}$, with $a_{f,c}(n_0) \neq 0$. We call

$$\mathrm{ord}_c(f) := n_0$$

the *order of f at c*. If n_0 is positive, we say that f has a root of order n_0 at c, and if n_0 is negative, we say that f has a pole of order $|n_0|$ at c. The *residue* of f at c is defined by

$$\mathrm{res}_c(f) := a_{f,c}(-1).$$

The meromorphic functions on $\mathbb{C}$ form a field. Moreover, every meromorphic function f on $\mathbb{C}$ can be written as a quotient $f(z) = g(z)/h(z)$ with holomorphic functions g, h on $\mathbb{C}$.

For $w \in \mathbb{C}$ and $D \subset \mathbb{C}$ we write

$$D + w = \{d + w : d \in D\}.$$

Definition 2.1.2. Let f be a meromorphic function on $\mathbb{C}$ with set of poles D_f. Then $w \in \mathbb{C}$ is called a *period* of f if

1. $D_f + w = D_f$, and
2. $f(z + w) = f(z)$ for all $z \in \mathbb{C} \setminus D_f$.

We denote by $\mathrm{Per}(f)$ the set of all periods of f.

Note that 0 is always a period of f. Moreover, the sum of two periods of f is again a period of f, so $\mathrm{Per}(f)$ is a subgroup of the additive group $(\mathbb{C}, +)$. For constant f we have $\mathrm{Per}(f) = \mathbb{C}$.

Lemma 2.1.3. *If f is a non-constant meromorphic function, then $\mathrm{Per}(f)$ is a closed, discrete subgroup of $\mathbb{C}$.*

Proof. If $\mathrm{Per}(f)$ is not discrete or not closed, then there is a sequence $w_n \in \mathrm{Per}(f)$ of pairwise different complex numbers such that $w = \lim_{n\to\infty} w_n$ exists. Since D_f is closed we have $D_f + w = D_f$. Hence, if f is holomorphic in c, then f is holomorphic in $c + w$ as well, and we have $f(c) = f(c + w_n)$ for all n. By the identity theorem f must be constant. □

Now we can describe all possible sets of periods.

Lemma 2.1.4. *If f is a non-constant meromorphic function, then precisely one of the following three cases occurs:*

1. $\mathrm{Per}(f) = 0$.
2. *The exists a (uniquely determined up to sign) $w_f \in \mathbb{C} \setminus \{0\}$ such that*
$$\mathrm{Per}(f) = \mathbb{Z}w_f = \{mw_f : m \in \mathbb{Z}\}.$$
3. *There exist $w_1, w_2 \in \mathbb{C} \setminus \{0\}$ with the following properties:*
   a) $\mathrm{Per}(f) = \mathbb{Z}w_1 + \mathbb{Z}w_2 = \{m_1w_1 + m_2w_2 : m_1, m_2 \in \mathbb{Z}\}$,
   b) *w_1, w_2 are linearly independent over $\mathbb{R}$,*
   c) *$\tau = w_1/w_2$ satisfies $\mathrm{Im}(\tau) > 0, |\mathrm{Re}(\tau)| \le \frac{1}{2}$ and $|\tau| \ge 1$.*

Proof. Let $\mathrm{Per}(f) \neq \{0\}$. Since $\mathrm{Per}(f)$ is closed and discrete, there exists some $w_f \in \mathrm{Per}(f)$ with

$$0 < |w_f| = \inf\{|w| : 0 \neq w \in \mathrm{Per}(f)\}. \tag{2.1.1}$$

We first investigate the periods on the line $\mathbb{R}w_f$. We claim that

$$\mathrm{Per}(f) \cap \mathbb{R}w_f = \mathbb{Z}w_f. \tag{2.1.2}$$

We obviously have $\mathbb{Z}w_f \subset \mathrm{Per}(f) \cap \mathbb{R}w_f$. Conversely, for $w \in \mathrm{Per}(f) \cap \mathbb{R}w_f$ we have $w = \alpha w_f$ for some $\alpha \in \mathbb{R}$. We choose $m \in \mathbb{Z}$ with $|\alpha - m| < 1$ and obtain

$$|w - mw_f| = |\alpha - m| \cdot |w_f| < |w_f|.$$

Since $w - mw_f$ belongs to $\mathrm{Per}(f) \cap \mathbb{R}w_f$, and by (2.1.1) we must have $w = mw_f$, which implies $\mathrm{Per}(f) \cap \mathbb{R}w_f \subset \mathbb{Z}w_f$.

If $\mathrm{Per}(f)$ lies on a line through 0, that is, on $\mathbb{R}w_f$, then we find $\mathrm{Per}(f) = \mathbb{Z}w_f$, and we are in part 2. of the lemma.

Let us now assume $\mathrm{Per}(f) \neq \mathbb{Z}w_f$. Then there exists an element $w_1 \in \mathrm{Per}(f) \setminus \mathbb{Z}w_f$ with

$$|w_1| = \inf\{|w| : w \in \mathrm{Per}(f) \setminus \mathbb{Z}w_f\}. \tag{2.1.3}$$

Put $w_2 = w_f$. Then we have $\tau = w_1/w_2 \notin \mathbb{R}$ since we assumed that $\mathrm{Per}(f)$ does not lie on a line through the origin. In particular, w_1, w_2 are linearly independent over $\mathbb{R}$. Replacing w_1 with $-w_1$ if necessary, we can assume that $\mathrm{Im}(\tau) > 0$. Now (2.1.1) implies

$$|w_1| \geq |w_2|, \qquad \text{i.e. } |\tau| \geq 1,$$

and (2.1.3) yields

$$|w_1 \pm w_2| \geq |w_1|, \qquad \text{i.e. } |\tau \pm 1| \geq |\tau|,$$

and hence $|\mathrm{Re}(\tau)| \leq 1/2$.

It is clear that $\mathbb{Z}w_1 + \mathbb{Z}w_2 \subset \mathrm{Per}(f)$. Conversely, let $w \in \mathrm{Per}(f)$. Since w_1, w_2 form an $\mathbb{R}$-basis of $\mathbb{C}$, we can write $w = \alpha_1 w_1 + \alpha_2 w_2$ with $\alpha_1, \alpha_2 \in \mathbb{R}$. We choose $m_j \in \mathbb{Z}$ such that $\beta_j = \alpha_j - m_j$ satisfy $|\beta_j| \leq 1/2$ for $j = 1, 2$. Then we have

$$w' = w - m_1 w_1 - m_2 w_2 = \beta_1 w_1 + \beta_2 w_2 \in \mathrm{Per}(f).$$

If $\beta_1 = 0$ then $w' = 0$ (i.e. $w = m_1 w_1 + m_2 w_2$) follows from (2.1.2). If $\beta_1 \neq 0$, then we have $w' \in \mathrm{Per}(f) \setminus \mathbb{Z}w_f$ and

$$\begin{aligned} |w'|^2 = |\beta_1 w_1 + \beta_2 w_2|^2 &= (\beta_1^2 |\tau|^2 + 2\beta_1\beta_2 \mathrm{Re}(\tau) + \beta_2^2) \cdot |w_2|^2 \\ &\leq (\beta_1^2 + |\beta_1||\beta_2| + \beta_2^2) \cdot |\tau|^2 \cdot |w_2|^2 \leq \frac{3}{4}|w_1|^2, \end{aligned}$$

where we used that $\tau = w_1/w_2$ satisfies $|\tau| \geq 1$ and $|\mathrm{Re}(\tau)| \leq 1/2$. If follows from (2.1.3) that $w' = 0$, i.e. $w = m_1 w_1 + m_2 w_2$. This finishes the proof that $\mathrm{Per}(f) = \mathbb{Z}w_1 + \mathbb{Z}w_2$. □

2.2 Lattices in $\mathbb{C}$

Let V be a real vector space of dimension $n \geq 1$, e.g. $V = \mathbb{R}^n$. A subset $\Omega \subset V$ is called a *lattice* in V if there exists an $\mathbb{R}$-basis $(w_1, \ldots, w_n)$ of V such that

$$\Omega = \mathbb{Z}w_1 + \cdots + \mathbb{Z}w_n.$$

We also call $(w_1, \ldots, w_n)$ as basis of Ω. Note that for $0 \neq \lambda \in \mathbb{C}$ the set $\lambda\Omega$ is again a lattice.

We see that in case 3. of Lemma 2.1.4 the set $\mathrm{Per}(f)$ is a lattice in $\mathbb{C} \cong \mathbb{R}^2$. Moreover, we have seen that $\mathrm{Per}(f)$ is a closed and discrete subgroup of $(\mathbb{C}, +)$. More generally, we have the following result.

Lemma 2.2.1. *Every lattice Ω in $\mathbb{C}$ is closed and discrete in $\mathbb{C}$.*

Proof. Let (w_1, w_2) be a basis for Ω, that is, $\Omega = \mathbb{Z}w_1 + \mathbb{Z}w_2$, and w_1, w_2 are linearly independent over $\mathbb{R}$. Replacing Ω with $\frac{1}{|w_2|}\Omega$ and w_1 with $-w_1$ if necessary, we can assume that $\Omega = \mathbb{Z}\tau + \mathbb{Z}$ with $\tau = x + iy \in \mathbb{C}, y > 0$.

For $\rho > 0$ we put $M_\rho = \{w \in \Omega : |w| \leq \rho\}$. We want to show that M_ρ is finite. Indeed, let $w = m\tau + n \in M_\rho$ with $m, n \in \mathbb{Z}$, then we have

$$\rho^2 \geq |m\tau + n|^2 = (mx + n)^2 + m^2 y^2 \geq m^2 y^2$$

which implies $|m| \leq \rho/y$. Moreover, we have

$$\rho \geq |mx + n| \geq |n| - |mx|$$

which shows $|n| \leq \rho(1 + |x|/y)$. This show that M_ρ is finite, and finishes the proof of the lemma. □

Next, we would like to descibe the possible change-of-basis matrices of lattices Ω in $\mathbb{C}$. To this end, we consider the set

$$\mathrm{Mat}_2(\mathbb{Z}) = \left\{ U = \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{Z} \right\}$$

of integral 2 by 2 matrices. It is a ring under matrix addition and multiplication, with unit element $E = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. The group of units of $\mathrm{Mat}_2(\mathbb{Z})$ is given by the general linear group

$$\mathrm{GL}_2(\mathbb{Z}) = \{U \in \mathrm{Mat}_2(\mathbb{Z}) : \text{there is } V \in \mathrm{Mat}_2(\mathbb{Z}) \text{ with } UV = VU = E\}.$$

Lemma 2.2.2. *For $U \in \mathrm{Mat}_2(\mathbb{Z})$ the following are equivalent.*

1. *$U \in \mathrm{GL}_2(\mathbb{Z})$.*
2. *$\det(U) = \pm 1$.*
3. *U is invertible over $\mathbb{Q}$ and $U^{-1} \in \mathrm{Mat}_2(\mathbb{Z})$.*
4. *The map $U : \mathbb{Z}^2 \to \mathbb{Z}^2, x \mapsto Ux$ is bijective.*
5. *The map $U : \mathbb{Z}^2 \to \mathbb{Z}^2, x \mapsto Ux$ is surjective.*

Proof. Exercise. □

We will also consider the special linear group

$$\mathrm{SL}_2(\mathbb{Z}) = \{U \in \mathrm{GL}_2(\mathbb{Z}) : \det(U) = 1\}.$$

Lemma 2.2.3. *If $c, d \in \mathbb{Z}$ are coprime, then there is a matrix*

$$U = \begin{pmatrix} * & * \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

Moreover, U is determined uniquely up to multiplication from the left by a factor of the form $\left(\begin{smallmatrix} 1 & k \\ 0 & 1 \end{smallmatrix}\right)$ with $k \in \mathbb{Z}$

Proof. Exercise. □

Lemma 2.2.4. *Let Ω be a lattice in $\mathbb{C}$ and let (w_1, w_2) be a basis of Ω. Let $w_1', w_2' \in \mathbb{C}$. Then we have $w_1', w_2' \in \Omega$ if and only if there is $U \in \mathrm{Mat}_2(\mathbb{Z})$ with*

$$\begin{pmatrix} w_1' \\ w_2' \end{pmatrix} = U \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

Moreover, (w_1', w_2') is a basis of Ω if and only if $U \in \mathrm{GL}_2(\mathbb{Z})$.

Proof. Exercise. □

Let Ω be a lattice in $\mathbb{C}$ and let (w_1, w_2) be a basis of Ω. For $u \in \mathbb{C}$ we define the *fundamental parallelogram* w.r.t to (w_1, w_2) and base point u by

$$P(u; w_1, w_2) = \{u + \alpha w_1 + \alpha w_2 : 0 \leq \alpha_1 < 1, 0 \leq \alpha_2 < 1\}.$$

For $u = 0$ we also write $P(w_1, w_2) = P(0; w_1, w_2)$. The following result is clear from the definition.

Proposition 2.2.5. *Let P be a fundamental parallelogram for Ω. For each $z \in \mathbb{C}$ there is a unique $w \in \Omega$ such that $z + w \in P$. In particular, if z and $z + \omega$ with $\omega \in \Omega$ both belong to P, then $w = 0$.*

There are many different bases and hence different period parallelograms for Ω, but their volume is an invariant of Ω, called the volume of Ω.

Lemma 2.2.6. *The volume of any fundamental parallelogram $P(u; w_1, w_2)$ for Ω equals* $\mathrm{vol}(\Omega) := |\mathrm{Im}(w_1\overline{w_2})|$, *and is independent of the basis (w_1, w_2) and the base point u.*

Proof. An elementary consideration shows that the volume of $P(u; w_1, w_2)$ is given by

$$\left|\det \begin{pmatrix} \mathrm{Re}w_1 & \mathrm{Im}w_1 \\ \mathrm{Re}w_2 & \mathrm{Im}w_2 \end{pmatrix}\right| = |\mathrm{Im}(w_1\overline{w_2})|.$$

If (w_1', w_2') is a different basis of Ω, then there exists a matrix $U = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbb{Z})$ with $\left(\begin{smallmatrix} w_1' \\ w_2' \end{smallmatrix}\right) = U \left(\begin{smallmatrix} w_1 \\ w_2 \end{smallmatrix}\right)$, and hence

$$|\mathrm{Im}(w_1'\overline{w_2'})| = \left|\det \begin{pmatrix} \mathrm{Re}w_1' & \mathrm{Im}w_1' \\ \mathrm{Re}w_2' & \mathrm{Im}w_2' \end{pmatrix}\right| = \left|\det \left(U \cdot \begin{pmatrix} \mathrm{Re}w_1 & \mathrm{Im}w_1 \\ \mathrm{Re}w_2 & \mathrm{Im}w_2 \end{pmatrix}\right)\right| = |\det(U)| \cdot |\mathrm{Im}(w_1\overline{w_2})|.$$

Using $|\det(U)| = 1$ we see that the volume is independent of the basis. □

Let Ω be a lattice in $\mathbb{C}$. Since Ω is a subgroup of the abelian group $(\mathbb{C}, +)$, is is a normal subgroup, and the factor group

$$\mathbb{C}/\Omega = \{a + \Omega : a \in \mathbb{C}\}$$

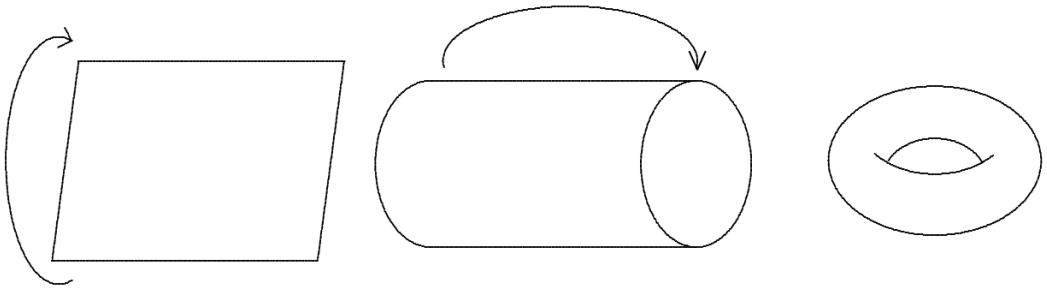
is an abelian group under the addition

$$(a + \Omega) + (b + \Omega) = (a + b) + \Omega.$$

Let $\pi : \mathbb{C} \to \mathbb{C}/\Omega$ be the canonical projection. By restricting it to a fundamental parallelogram $P = P(u; w_1, w_2)$, we obtain a bijection

$$\pi|_P : P \xrightarrow{\sim} \mathbb{C}/\Omega. \tag{2.2.1}$$

Since $\pi|_P$ identifies the opposite edges of the fundamental parallelogram P, we may view $\mathbb{C}/\Omega$ as a torus in $\mathbb{R}^3$.



2.3 Eisenstein series

Let Ω be a lattice in $\mathbb{C}$ with basis (w_1, w_2). We have already seen that the volume $\mathrm{vol}(\Omega) = \mathrm{vol}(P)$ of a fundamental parallelogram $P = P(u; w_1, w_2)$ for Ω is invariant under the choice of the basis of Ω. In this section we introduce other invariants of Ω, the so-called Eisenstein series. They will be defined as certain infinite series, and in order to study their convergence, we need some preparation.

We let

$$\delta = \delta(w_1, w_2) = \sup\{|z - w| : z, w \in P(w_1, w_2)\}$$

be the diameter of the fundamental parallelogram $P(w_1, w_2)$. For $\rho > 0$ we let

$$A_\rho(\Omega) = \#\{w \in \Omega : |w| \leq \rho\}$$

be the number of lattice points in a closed disc with radius ρ.

Lemma 2.3.1. *For $\rho \geq \delta$ we have*

$$\frac{\pi}{\mathrm{vol}(\Omega)}(\rho - \delta)^2 \leq A_\rho(\Omega) \leq \frac{\pi}{\mathrm{vol}(\Omega)}(\rho + \delta)^2.$$

Proof. We compare the two sets

$$K_\rho = \{z \in \mathbb{C} : |z| \leq \rho\}, \qquad M_\rho = \bigcup_{w \in \Omega, |w| \leq \rho} P(w; w_1, w_2).$$

Note that M_ρ is a disjoint union. Since $\rho \geq \delta$, we have

$$K_{\rho - \delta} \subset M_\rho \subset K_{\rho + \delta}.$$

Taking volumes yields the result, since $\mathrm{vol}(K_\rho) = \pi\rho^2$ and

$$\mathrm{vol}(M_\rho) = \sum_{w \in \Omega, |w| \leq \rho} \mathrm{vol}(P(w; w_1, w_2)) = \mathrm{vol}(\Omega) \cdot A_\rho(\Omega).$$

□

In the following, we will say that a multiple series

$$\sum_{g \in \mathbb{Z}^n} \alpha_g, \qquad (\alpha_g \in \mathbb{C}),$$

converges absolutely if there is some $C > 0$ such that $\sum_{g \in E} |\alpha_g| < C$ for every finite subset $E \subset \mathbb{Z}^n$. In this case, the series $\sum_{k \in \mathbb{N}} \alpha_{\varphi(k)}$ converges absolutely for every bijection $\varphi : \mathbb{N} \to \mathbb{Z}^n$ and is independent of the choice of φ. In particular, the value of the series does not change if we rearrange the terms.

Lemma 2.3.2. *Let $\alpha \in \mathbb{R}$. The series*

$$\sum_{0 \neq w \in \Omega} |w|^{-\alpha}$$

converges if and only if $\alpha > 2$.

Proof. Let $\alpha > 2$ and let $E \subset \Omega \setminus \{0\}$ with $E \neq \emptyset$ be a finite subset. Put $M = \max\{|w| : w \in E\}$. By the preceeding lemma, there is a constant $c_2 > 0$ such that

$$A_{n+1}(\Omega) - A_n(\Omega) \leq \frac{\pi}{\mathrm{vol}(\Omega)} \cdot ((n+1+\delta)^2 - (n-\delta)^2) \leq c_2 n$$

for all $n \geq \delta$. Define another consant

$$c_1 = \sum_{0 \neq w \in \Omega, |w| \leq \delta+1} |w|^{-\alpha}.$$

Then we find

$$\sum_{w \in E} |w|^{-\alpha} \leq c_1 + \sum_{n \in \mathbb{N}, \delta < n < M} (A_{n+1}(\Omega) - A_n(\Omega)) n^{-\alpha} \leq c_1 + c_2 \sum_{n=1}^{\infty} n^{1-\alpha} < \infty$$

since $\alpha > 2$.

Now let $\alpha \leq 2$. The series in question trivially diverges for $\alpha \leq 0$, so we may assume $0 < \alpha \leq 2$. Pick some $N \in \mathbb{N}$ with $N > 2\delta$. The preceeding lemma gives a constant $c_3 > 0$ such that

$$A_{kN}(\Omega) - A_{(k-1)N}(\Omega) \geq \frac{\pi}{\mathrm{vol}(\Omega)} ((kN-\delta)^2 - ((k-1)N+\delta)^2) \geq c_3 k$$

for all $k \in \mathbb{Z}$ with $k \geq 2$. Let $E_n = \{w \in \Omega : 0 < |w| \leq nN\}$. Then we have

$$\sum_{w \in E_n} |w|^{-\alpha} \geq \sum_{k=2}^{n} (A_{kN}(\Omega) - A_{(k-1)N}(\Omega)) \cdot (kN)^{-\alpha} \geq c_3 N^{-\alpha} \sum_{k=2}^{n} k^{1-\alpha}.$$

Since the series $\sum_{k>1} k^{1-\alpha}$ diverges for $\alpha \leq 2$, the series $\sum_{0 \neq w \in \Omega} |w|^{-\alpha}$ diverges for $\alpha \leq 2$, as well. This finishes the proof. □

The above lemma implies that the *Eisenstein series*

$$G_k = G_k(\Omega) = \sum_{0 \neq w \in \Omega} w^{-k}, \qquad k \in \mathbb{Z},$$

converges absolutely for $k \geq 3$. Note that $G_k(\Omega) = 0$ for odd $k \geq 3$ and any lattice Ω since the terms w^{-k} and $(-w)^{-k}$ cancel out in the sum. On the other hand, we will later see that $G_k(\Omega)$ is typically non-vanishing for even $k \geq 4$. The Eisenstein series will be important for us later since they appear in the Taylor expansions of elliptic functions.

Moreover, we will see the surprising fact that every G_k can be written as a polynomial over $\mathbb{Q}$ in G_4 and G_6. For example, we have $7G_8 = 3G_4^2$ and $11G_{10} = 5G_4G_6$. This also has some interesting number theoretical applications.

3 Elliptic functions

3.1 Basic definitions

We have seen in Lemma 2.1.4 that the set of periods $\mathrm{Per}(f)$ of a meromorphic function f on $\mathbb{C}$ is either $\{0\}$, or of the form $\mathbb{Z}w_f$ for some $w_f \in \mathbb{C}$, or a lattice in $\mathbb{C}$. In this section we will study elliptic functions and use them to show that for any lattice Ω in $\mathbb{C}$, there exists a meromorphic function f with $\mathrm{Per}(f) = \Omega$. Throughout, we let $\Omega = \mathbb{Z}w_1 + \mathbb{Z}w_2$ be a lattice in $\mathbb{C}$.

Definition 3.1.1. A meromorphic function f on $\mathbb{C}$ is called *elliptic* (or *doubly periodic*) with respect to Ω if $\Omega \subset \mathrm{Per}(f)$, that is, if

1. $D_f + w = D_f$ for all $w \in \Omega$, and
2. $f(z + w) = f(z)$ for all $w \in \Omega$ and $z \in \mathbb{C} \setminus D_f$.

We let $\mathcal{K}(\Omega)$ be the set of all elliptic functions with respect to Ω.

Note that it suffices to check the above two conditions for a basis of Ω. For an elliptic function $f \in \mathcal{K}(\Omega)$ we have

$$\mathrm{ord}_{c+w}(f) = \mathrm{ord}_c(f), \qquad \text{and} \qquad \mathrm{res}_{c+w}(f) = \mathrm{res}_c(f) \tag{3.1.1}$$

for all $w \in \Omega$. The following basic results are easy to prove from the definition of elliptic functions.

Proposition 3.1.2. *The elliptic functions $\mathcal{K}(\Omega)$ with respect to Ω form a subfield of the field of all meromorphic functions on $\mathbb{C}$ which contains the constant functions. Every $f \in \mathcal{K}(\Omega)$ only has finitely many poles in each fundamental parallelogram for Ω.*

Lemma 3.1.3. *Let $f \in \mathcal{K}(\Omega)$. Then we have $f' \in \mathcal{K}(\Omega)$ and $g(z) := f(nz + x) \in \mathcal{K}(\Omega)$ for every fixed $0 \neq n \in \mathbb{Z}$ and $x \in \mathbb{C}$.*

3.2 The four theorems of Liouville

In 1847 Liouville noticed that elliptic functions satisfy some strong conditions which are not obvious from the definition.

Theorem 3.2.1. *If $f \in \mathcal{K}(\Omega)$ is holomorphic on $\mathbb{C}$, then f is constant.*

Proof. Let P be a fundamental parallelogram for Ω. Since the closure of P is compact, f is bounded on P, i.e. there is some $C > 0$ with $|f(z)| \leq C$ for $z \in P$. For arbitrary $z \in \mathbb{C}$ there exists some $w \in \Omega$ such that $z + w \in P$. This implies

$$|f(z)| = |f(z + w)| \leq C,$$

so f is bounded on $\mathbb{C}$, hence constant. □

Theorem 3.2.2. *For $f \in \mathcal{K}(\Omega)$ and any fundamental parallelogram P for Ω we have*

$$\sum_{c \in P} \operatorname{res}_c(f) = 0.$$

Proof. Using (3.1.1) and Proposition 2.2.5 we see that the sum is finite and independent of the choice of the fundamental parallelogram P. Hence we may choose a base point $u \in \mathbb{C}$ such that the boundary ∂P of $P = P(u; w_1, w_2)$ does not contain any poles of f.

Now we integrate f over the boundary ∂P. By the residue theorem we have

$$\begin{aligned}\pm 2\pi i \sum_{c \in P} \operatorname{res}_c(f) &= \int_u^{u+w_1} f(z)dz + \int_{u+w_1}^{u+w_1+w_2} f(z)dz + \int_{u+w_1+w_2}^{u+w_2} f(z)dz + \int_{u+w_2}^{u} f(z)dz \\ &= \int_u^{u+w_1} (f(z) - f(z+w_2))dz + \int_{u+w_2}^{u} (f(z) - f(z+w_1))dz.\end{aligned}$$

where the sign $\pm$ depends on the orientation of ∂P. Note that the right-hand side vanishes for $f \in \mathcal{K}(\Omega)$, which finishes the proof. □

Theorem 3.2.3. *For $f \in \mathcal{K}(\Omega)$ non-constant, any fundamental parallelogram P for Ω, and any $x \in \mathbb{C}$ we have*

$$\sum_{c \in P} \operatorname{ord}_c(f - x) = 0. \tag{3.2.1}$$

Hence, if we count with multiplicities, we have

$$\begin{aligned}\textit{Number of poles of } f \textit{ in } P &= \textit{Number of zeros of } f \textit{ in } P \\ &= \textit{Number of } z \in P \textit{ with } f(z) = x.\end{aligned}$$

Moreover, every non-constant $f \in \mathcal{K}(\Omega)$ takes every value in P.

Proof. By Lemma 3.1.3 the function

$$g(z) = \frac{f'(z)}{f(z) - x}$$

is an elliptic function for Ω, and we have $\operatorname{res}_c(g) = \operatorname{ord}_c(f - x)$. Here we used that f is non-constant, hence $f(z) - x$ is not vanishing identically. Now the formula (3.2.1) follows from Theorem 3.2.2. In order to see that f takes any value x in P, note that $f(z) - x \in \mathcal{K}(\Omega)$ is non-constant and hence must have a pole in P by Theorem 3.2.1. Hence, by (3.2.1), $f(z) - x$ must also have a root in P. □

Theorem 3.2.4. *For $f \in \mathcal{K}(\Omega)$ and any fundamental parallelogram P for Ω we have*

$$\sum_{c \in P} (\operatorname{ord}_c(f)) \cdot c \in \Omega.$$

Proof. Similarly as in the proof of Theorem 3.2.2, we integrate the function $z\frac{f'(z)}{f(z)}$ over the boundary ∂P of a suitable fundamental parallelogram $P = P(u; w_1, w_2)$ of Ω. We obtain

$$\begin{aligned}&2\pi i \sum_{c \in P} \operatorname{ord}_c(f) \cdot c = \int_{\partial P} z\frac{f'(z)}{f(z)}dz \\ &= \pm\Big(\int_u^{u+w_1} z\frac{f'(z)}{f(z)} - (z+w_2)\frac{f'(z+w)}{f(z+w)}dz + \int_{u+w_2}^{u} z\frac{f'(z)}{f(z)} - (z+w_1)\frac{f'(z+w_1)}{f(z+w)}dz \Big) \\ &= \pm\Big(w_1 \int_u^{u+w_2} \frac{f'(z)}{f(z)}dz - w_2 \int_u^{u+w_1} \frac{f'(z)}{f(z)}dz \Big).\end{aligned}$$

Using $f(u) = f(u + w_j)$ we find

$$\int_u^{u+w_j} \frac{f'(z)}{f(z)} dz \in 2\pi i\mathbb{Z}$$

for $j = 1, 2$, which finishes the proof. □

Theorem 3.2.2 implies that there are no elliptic functions with only one first order pole in P. There must either be at least two poles of order one, or a pole of order two with residue zero.

If we count the zeros and poles of a non-constant elliptic function $f \in \mathcal{K}(\Omega)$ with multiplicities, then Theorem 3.2.3 says that there are points $a_1, \ldots, a_r$ and $b_1, \ldots, b_r \in P$, such the roots of f in P are precisely at the points $a_1, \ldots, a_r$ and the poles of f in P are at the points $b_1, \ldots, b_r$. Here the multiplicity of a root or pole is indicated by a repetition of the a_j or b_j. Now Theorem 3.2.4 can be written as

$$a_1 + \cdots + a_r \equiv b_1 + \cdots + b_r \pmod{\Omega}. \tag{3.2.2}$$

We call r the order of f. Theorems 3.2.1 and 3.2.2 say that every elliptic function of order 0 is constant, and that there are no elliptic functions of order 1. On the other hand, we will see that for $r \geq 2$ and $a_1, \ldots, a_r, b_1, \ldots, b_r \in P$ satisfying (3.2.2) there is a suitable elliptic function having roots in $a_1, \ldots, a_r$ and poles in $b_1, \ldots, b_r$.

3.3 First properties of the Weierstrass $\wp$-function

Theorem 3.3.1. *There exists an elliptic function $\wp = \wp_\Omega \in \mathcal{K}(\Omega)$, which has poles of order 2 precisely at the lattice points in Ω, and is holomorphic everywhere else. Its Laurent expansion at 0 is of the form*

$$\wp(z) = z^{-2} + a_1 z + \ldots. \tag{3.3.1}$$

The proof will be given in the next section, where we will explicitly construction $\wp$ as an infinite series. The function $\wp$ is called the *Weierstrass $\wp$ function* for Ω.

By (3.1.1) it is clear that $\wp$ has residue 0 at all poles (i.e. lattice points of Ω). Moreover, by Theorem 3.2.1 the elliptic function $\wp$ is uniquely determined by the above conditions.

Proposition 3.3.2. *1. $\wp$ is an even function, that is, $\wp(-z) = \wp(z)$. Hence we have $a_1 = 0$ in the Laurent expansion* (3.3.1).

2. $\wp'$ is an odd function which has poles of order 3 precisely at the lattice points of Ω and is holomorphic everywhere else.

Proof. For the first part, consider the elliptic function $f(z) = \wp(-z) - \wp(z)$. Using the Laurent expansion (3.3.1), we see that f is holomorphic on $\mathbb{C}$, hence constant by Theorem 3.2.1.

The second part immediately follows from the first part and Theorem 3.3.1. □

We can now already determine the roots of $\wp'$.

Lemma 3.3.3. *The roots of $\wp'$ are of order 1 and lie precisely at the points $w/2$ for which $w \in \Omega$ but $w/2 \notin \Omega$.*

Proof. Since $\wp'$ is an odd elliptic function, we have

$$\wp'(z+w) = \wp'(z) = -\wp'(-z)$$

for $w \in \Omega$. If $w/2 \notin \Omega$, then $w/2$ is not a pole of $\wp'$, and we can take $z = -w/2$ to get

$$\wp'(w/2) = -\wp'(w/2),$$

hence $\wp'(w/2) = 0$. Let w_1, w_2 be a basis of Ω and $P = P(w_1, w_2)$ the corresponding fundamental parallelogram. Then $\wp'$ has at least the tree roots

$$w_1/2, \quad w_2/2, \quad (w_1+w_2)/2$$

in P, and we want to show that those are all possible roots of $\wp'$ in P. From the Laurent expansion (3.3.1) we see that $\wp'$ has precisely one pole in P, which is of order 3. Hence, by Theorem 3.2.3 the three roots that we found above must have order 1, and there can be no other roots in P. For an arbitrary point $z \in \mathbb{C}$ we can find $w' \in \Omega$ such that $z - w' \in \Omega$. If z is a root of $\wp'$, then $z - w'$ is one of the three roots in P above, so z is of the form $w/2$ with $w \in \Omega$ but $w/2 \notin \Omega$. □

Lemma 3.3.4. *Let $P = P(w_1, w_2)$ be a fundamental parallelogram for Ω and put*

$$e_1 := \wp(w_1/2), \quad e_2 := \wp(w_2/2), \quad e_3 := \wp(w_3/2), \qquad w_3 := w_1 + w_2. \tag{3.3.2}$$

Then

$$\wp(z) - e_k \text{ has precisely one double root in } P, \text{ at } z = w_k/2, \tag{3.3.3}$$

for $k = 1, 2, 3$, and

$$\wp(z) - x \text{ has precisely two simple roots in } P \text{ if } x \neq e_1, e_2, e_3. \tag{3.3.4}$$

Proof. We apply Theorem 3.2.3 to $\wp$. Since $\wp$ has precisely one pole of order 2 in P (namely, at $z = 0$), the function $\wp(z) - x$ has two roots (counted with multiplicty) in P. Now there are two cases:

1. There is only one $u \in P$ with $\wp(u) = x$. Then $\wp(z) - x$ must have a double zero at u, so $\wp'(u) = 0$. By Lemma 3.3.3, this implies that $u \in \{w_1/2, w_2/2, w_3/2\}$.

2. There are two different $u, v \in P$ with $\wp(u) = \wp(v) = x$. Then $\wp(z) - x$ must have two simple roots at u and v, so $\wp'(u) \neq 0$ and $\wp'(v) \neq 0$, which by Lemma 3.3.3 means that $u, v \notin \{w_1/2, w_2/2, w_3/2\}$.

□

Since $w_1/2, w_2/2, w_3/2$ are pairwise different, we see from (3.3.3) that

$$e_1, e_2, e_3 \text{ are pairwise different.} \tag{3.3.5}$$

Note that taking a different basis for Ω only permutes the values e_1, e_2, e_3.

We obtain a first differential equation for the Weierstrass $\wp$-function.

Proposition 3.3.5. *For $z \in \mathbb{C} \setminus \Omega$ we have*

$$\wp'(z)^2 = 4 \cdot (\wp(z) - e_1) \cdot (\wp(z) - e_2) \cdot (\wp(z) - e_3).$$

Proof. We consider the elliptic function

$$f(z) = 4 \cdot (\wp(z) - e_1) \cdot (\wp(z) - e_2) \cdot (\wp(z) - e_3).$$

By (3.3.3) the function f has double roots precisely at the points $w_1/2, w_2/2, w_3/2$, and by Lemma 3.3.3 the same is true for $\wp'(z)^2$. Moreover, the only pole of f in P is a at 0, and is of order 6. From the Laurent expansions

$$\wp(z) = z^{-2} + \ldots, \qquad \wp'(z)^2 = 4z^{-6} + \ldots$$

(compare (3.3.1)) we see that $\wp'(z)^2$ also has a pole of order 6 at 0, and no other poles in P. We obtain that $\wp'(z)^2/f(z)$ is an elliptic function without any poles, hence constant by Theorem 3.2.1. Comparing the coefficients at z^{-6} in the Laurent expansions of $\wp'(z)^2$ and $f(z)$ around 0, we find that this constant is equal to 1. □

3.4 The field of elliptic functions

From the Laurent expansion (3.3.1) it is clear that in any polynomial in $\wp$, the poles cannot cancel out. In particular, $\wp$ is not algebraic, i.e. there is no non-zero polynomial $p(x)$ with $p(\wp) = 0$. Hence the field $\mathbb{C}(\wp)$ consisting of all rational functions in $\wp$ is isomorphic to the field of all rational functions over $\mathbb{C}$.

We can now describe the field $\mathcal{K}(\Omega)$ of elliptic function with respect to a lattice Ω in terms of the Weiestrass $\wp$-function and its derivative $\wp'$.

Proposition 3.4.1.

1. *The even elliptic functions in $\mathcal{K}(\Omega)$ are precisely the rational functions in $\wp$.*
2. *We have $\mathcal{K}(\Omega) = \mathbb{C}(\wp)[\wp']$.*
3. *The degree of the field extension of $\mathcal{K}(\Omega)$ over $\mathbb{C}(\wp)$ is 2.*

In other words, every $f \in \mathcal{K}(\Omega)$ can be written in a unique way as

$$f = R(\wp) + Q(\wp) \cdot \wp' \tag{3.4.1}$$

with rational functions R, Q over $\mathbb{C}$, and for even f we have $Q = 0$.

Proof. 1. For the proof, we will use the following helpful auxiliary result:

Claim: For each $m \in \mathbb{N}_0$ there exists a unique elliptic function $\wp_m \in \mathcal{K}(\Omega)$ which has poles of order $2m$ at the points in Ω and is holomorphic otherwise, and which has a Laurent expansion at $z = 0$ of the shape

$$\wp_m(z) = z^{-2m} + O(z^2).$$

Moreover, $\wp_m$ is a polynomial in $\wp$.

Proof. We can clearly take $\wp_0 = 1$ and $\wp_1 = \wp$. For $m = 2$ we consider the Laurent expansion of $\wp^2$ at $z = 0$,

$$\wp^2(z) = (z^{-2} + a_2 z^2 + O(z^4)) \cdot (z^{-2} + a_2 z^2 + O(z^4)) = z^{-4} + 2a_2 + O(z^2).$$

Hence we may take $\wp_2(z) = \wp^2(z) - 2a_2$. We can continue like this and recursively define $\wp_m$ by subtracting from $\wp^m$ suitable multiples of $\wp_0, \wp_1, \ldots, \wp_{m-1}$. This shows that $\wp_m$ can be constructed as a polynomial in $\wp$.

The uniqueness of $\wp_m$ follows from Theorem 3.2.1, since the difference of two elliptic functions with the above Laurent expansions would be entire and vanishing at $z = 0$, hence equal to 0. □

Now we come back to the proof of Proposition 3.4.1. Let $c_1, \ldots, c_k$ be the poles of f in a fundamental parallelogram P for Ω which do not already lie in Ω. Then the function

$$g(z) := \prod_{j=1}^{k} (\wp(z) - \wp(c_j))^{-\mathrm{ord}_{c_j}(f)} \cdot f(z)$$

is an even elliptic function with poles only at the lattice points in Ω. Now $g(z)$ has a Laurent expansion at $z = 0$ of the form

$$g(z) = a_{-2d} z^{-2d} + a_{-2d+2} z^{-2d+2} + \cdots + a_{-2} z^{-2} + a_0 + O(z^2),$$

with constants $a_j \in \mathbb{C}$, and $-2d = \mathrm{ord}_0(g)$. Here we used that g is even. Hence, by Theorem 3.2.1 we obtain

$$g(z) = \sum_{m=0}^{d} a_{-2m} \wp_m(z),$$

since the difference of both sides is an entire elliptic function which vanishes at 0. Putting everything together, we find

$$f(z) = \prod_{j=1}^{k} (\wp(z) - \wp(c_j))^{\mathrm{ord}_{c_j}(f)} \cdot \left(\sum_{m=0}^{d} a_{-2m} \wp_m(z) \right).$$

Recall that each $\wp_m$ is a polynomial in $\wp$, so f is a rational function in $\wp$. This finishes the proof of part 1.

2. For $f \in \mathcal{K}(\Omega)$ we may write

$$f = g + h\wp', \qquad \text{where} \quad g(z) = \frac{1}{2}(f(z) + f(-z)), \quad h(z) = \frac{1}{2\wp'(z)}(f(z) - f(-z)).$$

Then $g, h \in \mathcal{K}(\Omega)$ are even elliptic functions, and hence are rational functions in $\wp$ by the first part of the proposition.

3. Since $\wp'$ is an odd function, we have $\wp' \notin \mathbb{C}(\wp)$, so the degree of $\mathcal{K}(\Omega)$ over $\mathbb{C}(\wp)$ is at least 2. By Proposition 3.3.5 we have $\wp'^2 \in \mathbb{C}(\wp)$, so the degree is equal to 2.

□

4 The Weierstrass $\wp$-function

Throughout this chapter we let $\Omega = \mathbb{Z}w_1 + \mathbb{Z}w_2$ be a lattice in $\mathbb{C}$.

4.1 Construction of the $\wp$-function

In order to prove Theorem 3.3.1 we will now construct the $\wp$-function explicitly as an infinite series. Since $\wp$ should be elliptic with respect to Ω, and should have poles of second order at lattice points in Ω, it is tempting to take as a candidate the series

$$\sum_{w\in\Omega} (z-w)^{-2}.$$

Unfortunately, by Lemma 2.3.2, this series does not converge absolutely. To overcome this problem, the summation needs to be modified.

Theorem 4.1.1. *The* Weierstrass $\wp$-function

$$\wp(z) = \wp_\Omega(z) = z^{-2} + \sum_{0\neq w\in\Omega} \left((z-w)^{-2} - w^{-2}\right), \quad z \in \mathbb{C}\setminus\Omega,$$

converges absolutely and uniformly in every compact subset of $\mathbb{C}\setminus\Omega$. *It is an even elliptic function with respect to* Ω *and has poles of second order with residue* 0 *in every lattice point of* Ω. *The Laurent expansion at* 0 *has the form*

$$\wp(z) = z^{-2} + a_2 z^2 + \dots$$

This result also implies Theorem 3.3.1. Before we come to the proof, we remark that one can show the convergence of the following series in a similar way:

Lemma 4.1.2. *For* $k \in \mathbb{N}$ *with* $k \geq 3$ *the series*

$$\sum_{w\in\Omega} (z-w)^{-k}$$

converges absolutely and uniformly on every compact subset of $\mathbb{C}\setminus\Omega$.

Proof of Theorem 4.1.1. The proof consists of four steps. For brevity, we set

$$f_w(z) = (z-w)^{-2} - w^{-2}$$

for $0 \neq w \in \Omega$, and $K_\rho = \{z \in \mathbb{C} : |z| \leq \rho\}$ for $\rho > 0$.

1. *Convergence:* Let $K \subset \mathbb{C} \setminus \Omega$ be a compact set, and let $\rho > 0$ big enough such that $K \subset K_\rho$. The finite sum over the terms with $|w| < \rho + 1$ converges absolutely and locally uniformly, so we can assume $|w| \geq \rho + 1$ in the following. Then we can estimate

$$\left|\frac{1}{(z-w)^2} - \frac{1}{w^2}\right| = \left|\frac{2zw - z^2}{w^2(z-w^2)}\right| = \left|\frac{2 - z/w}{(1-z/w)^2}\right| \cdot \frac{|z|}{|w^3|}$$
$$\leq \frac{2 + \rho/(\rho+1)}{(1-\rho/(\rho+1))^2} \cdot \frac{\rho}{|w|^3}.$$

   The convergence now follows from the convergence of the Eisenstein series G_3, see Lemma 2.3.2

2. *Claim: The series defining ℘ is meromorphic in $\mathbb{C}$ and has poles precisely at the lattice points Ω, which are of order 2 and have residue 0.*
   Proof: Let $\rho > 0$, and write

$$\wp(z) = z^{-2} + \sum_{|w|<\rho+1} f_w(z) + \sum_{|w|\geq\rho+1} f_w(z).$$

   The first sum is meromorphic on K_ρ with poles of second order and residue 0 at lattice points in K_ρ, and the second series is holomorphic on K_ρ since it converges absolutely and locally uniformly.

3. *Claim: ℘ is an even function and has the Laurent expansion $\wp(z) = z^{-2} + a_2 z^2 + \dots$.*
   Proof: We replace w by $-w$ in the sum and use the absolute convergence of the series to see that $\wp(-z) = \wp(z)$. Above we have shown that $\wp$ has a pole of second order with residue 0 at $z = 0$, so it has the Laurent expansion $\wp(z) = z^{-2} + a_0 + a_2 z^2 + \dots$. But since $f_w(0) = 0$ for $w \neq 0$, we have $a_0 = 0$.

4. *Claim: We have $\wp(z + w) = \wp(z)$ for all $w \in \Omega$ and $z \in \mathbb{C} \setminus \Omega$.*
   Proof: By the absolute and locally uniform convergence of the series defining $\wp$, and Lemma 4.1.2, we can differentiate $\wp(z)$ termwise to get the absolutely convergent series representation

$$\wp'(z) = -2 \sum_{w\in\Omega} (z-w)^{-3}$$

   for $z \in \mathbb{C} \setminus \Omega$. We see that $\wp'(z + w) = \wp'(z)$ for $w \in \Omega$. Hence, we have $\wp(z + w) = \wp(z) + C_w$ for some constant C_w (possibly depending on w, but not on z). Setting $z = -w/2$ we see that $C_w = 0$ since $\wp$ is even. This shows that $\wp$ is elliptic.

□

4.2 The Laurent expansion

Recall that for $k \in \mathbb{N}$ with $k \geq 3$ we defined the Eisenstein series

$$G_k := G_k(\Omega) := \sum_{0\neq w\in\Omega} w^{-k}.$$

By Lemma 2.3.2 it converges absolutely. Moreover, it vanishes identically for odd k. We let

$$\gamma := \gamma(\Omega) := \min\{|w| : 0 \neq w \in \Omega\}.$$

The Weierstrass ℘-function has the following Laurent expansion around $z = 0$:

Proposition 4.2.1. *For $z \in \mathbb{C}$ with $0 < |z| < \gamma(\Omega)$ we have*

$$\wp(z) = z^{-2} + \sum_{n=2}^{\infty} (2n-1) G_{2n} \cdot z^{2n-2} = z^{-2} + 3G_4 \cdot z^2 + 5G_6 \cdot z^4 + \dots \tag{4.2.1}$$

Proof. First note that we have

$$\frac{1}{(1-t)^2} = \frac{d}{dt}\left(\frac{1}{1-t}\right) = \frac{d}{dt} \sum_{m=0}^{\infty} t^m = \sum_{m=1}^{\infty} m t^{m-1}, \qquad (|t| < 1).$$

Hence, for $0 \neq w \in \Omega$ we may write

$$\frac{1}{(z-w)^2} - \frac{1}{w^2} = \frac{1}{w^2}\left(\frac{1}{(1-z/w)^2} - 1\right) = \sum_{m=2}^{\infty} m \frac{z^{m-1}}{w^{m+1}}, \qquad (|z| < \gamma),$$

and thus we get

$$\wp(z) = z^{-2} + \sum_{0 \neq w \in \Omega} \left(\sum_{m=2}^{\infty} m \frac{z^{m-1}}{w^{m+1}} \right), \qquad (0 < |z| < \gamma). \tag{4.2.2}$$

Since

$$\left| m \frac{z^{m-1}}{w^{m+1}} \right| \leq \gamma m \left(\frac{|z|}{\gamma} \right)^{m-1} |w|^{-3},$$

we see from Lemma 2.3.2 that the double series in (4.2.2) converges absolutely. Hence, we can change the order of summation and obtain

$$\wp(z) = z^{-2} + \sum_{m=2}^{\infty} m \left(\sum_{0 \neq w \in \Omega} \frac{1}{w^{m+1}} \right) \cdot z^{m-1} = z^{-2} + \sum_{m=2}^{\infty} m G_{m+1} \cdot z^{m-1}$$

for $0 < |z| < \gamma$. Recall that $G_k = 0$ for odd k, which gives the stated Laurent expansion. □

4.3 Eisenstein series, the discriminant, and the j-invariant

We have seen in Proposition 3.3.5 that the $\wp$-function satisfies the differential equation

$$\wp'(z)^2 = 4(\wp(z) - e_1)(\wp(z) - e_2)(\wp(z) - e_3), \qquad (z \in \mathbb{C} \setminus \Omega),$$

where $e_1 = \wp(w_1/2), e_2 = \wp(w_2/2)$, and $e_3 = \wp((w_1 + w_2)/2)$ for a basis (w_1, w_2) of Ω. Using the Laurent expansion (4.2.1) of $\wp$ in terms of Eisenstein series, we now derive a second differential equation.

Proposition 4.3.1. *The $\wp$-function satisfies the differential equation*

$$\wp'(z)^2 = 4\wp(z)^3 - g_2\, \wp(z) - g_3$$

with the Weierstrass invariants

$$g_2 := g_2(\Omega) := 60\, G_4(\Omega), \quad \textit{and} \quad g_3 := g_3(\Omega) := 140\, G_6(\Omega).$$

Proof. Starting from

$$\wp(z) = z^{-2} + 3G_4 \cdot z^2 + 5G_6 \cdot z^4 + O(z^6)$$

(see (4.2.1)) we compute

$$\begin{aligned}
\wp^2(z) &= z^{-4} + 6G_4 + 10G_6 \cdot z^2 + O(z^3),\\
\wp^3(z) &= z^{-6} + 9G_4 \cdot z^{-2} + 15G_6 + O(z),\\
\wp'(z) &= -2z^{-3} + 6G_4 \cdot z + 20G_6 \cdot z^3 + O(z^4),\\
\wp'^2(z) &= 4 \cdot z^{-6} - 24G_4 \cdot z^{-2} - 80G_6 + O(z).
\end{aligned}$$

This implies that

$$\wp'^2(z) - 4\wp^3(z) + g_2\wp(z) + g_3 = O(z). \tag{4.3.1}$$

The left-hand side is an elliptic function for Ω which can only have poles at the same points as $\wp$ and $\wp'$, i.e., at lattice points in Ω. But (4.3.1) show that the left-hand side is holomorphic at 0, hence holomorphic everywhere and thus a constant by Theorem 3.2.1. Again by (4.3.1), this constant is 0. □

Conversely, the $\wp$-function gives all solutions of the above differential equation.

Proposition 4.3.2. *Let Ω be a lattice in $\mathbb{C}$ with Weierstrass invariants $g_2 = 60G_4$ and $g_3 = 140G_6$. Then the non-constant meromorphic solutions (on some domain $G \subset \mathbb{C}$) of the differential equation*

$$f'^2 = 4f^3 - g_2 f - g_3.$$

are given by $f(z) = \wp(z + w), z \in G$, for $w \in \mathbb{C}$.

If, in addition, f is meromorphic on $\mathbb{C}$, then we have $\mathrm{Per}(f) = \Omega$. The lattice Ω is uniquely determined by $g_2(\Omega)$ and $g_3(\Omega)$.

Proof. Suppose that f is a non-constant meromorphic solution of the differential equation on a domain G. Pick some $u \in G$ and some disc $U \subset G$ around u such that f is holomorphic on U and f' is non-vanishing on U. Then f satisfies also satisfies the first order differential equation

$$f' = \sqrt{4f^3 - g_2 f - g_3}$$

on U, for an appropriate choice of the square root. By Lemma 3.3.4 there exists a $w \in \mathbb{C}$ such that $\wp(w + u) = f(u)$. By replacing w with $-w - 2u$ if necessary we can also assume $\wp'(w + u) = f'(u)$. Now the functions $f(z)$ and $g(z) = \wp(z + w)$ satisfy the same first order differential equation and agree at u, hence they agree for all $z \in U$ by the existence and uniqueness theorem for first order differential equations. The identity theorem then yields $f(z) = g(z)$ for all $z \in G$.

If f is meromorphic on $\mathbb{C}$, then we have $f(z) = \wp(z+w)$ for all z where neither f nor $\wp$ has a pole. By the identity theorem, $f(z)$ has poles precisely at the points $-w + \Omega$. Moreover, we see that $\mathrm{Per}(f) = \mathrm{Per}(\wp(\cdot + w)) = \Omega$. If two lattices Ω, Ω' have the same Weierstrass invariants g_2, g_3, then their corresponding Weierstrass $\wp$-functions $\wp_\Omega$ and $\wp_{\Omega'}$ satisfy the same differential equation, and we have $\wp_\Omega(z) = \wp_{\Omega'}(z+w)$ for some $w \in \mathbb{C}$ by the statement of the corollary. This implies $\Omega = \mathrm{Per}(\wp_\Omega) = \mathrm{Per}(\wp_{\Omega'}(\cdot + w)) = \Omega'$. □

Comparing the differential equations from Proposition 3.3.5 and Proposition 4.3.1, we obtain the identity

$$4\wp^3 - g_2\wp - g_3 = 4(\wp - e_1)(\wp - e_2)(\wp - e_3).$$

Since the $\wp$-function takes more than three different values, we obtain the following identity of polynomials:

Corollary 4.3.3. *We have*

$$4X^3 - g_2X - g_3 = 4(X - e_1)(X - e_2)(X - e_3).$$

In particular, we have

$$\begin{aligned} 0 &= e_1 + e_2 + e_3, \\ g_2 &= -4(e_1e_2 + e_2e_3 + e_3e_1), \\ g_3 &= 4e_1e_2e_3. \end{aligned}$$

Using these identities for e_1, e_2, e_3, we obtain the following relation.

Corollary 4.3.4. *We have*

$$g_2^3 - 27g_3^2 = 16(e_1 - e_2)^2(e_2 - e_3)^2(e_3 - e_1)^2 \neq 0.$$

We define the *discriminant* of Ω by

$$\Delta := \Delta(\Omega) := g_2^3 - 27g^2 = 16(e_1 - e_2)^2(e_2 - e_3)^2(e_3 - e_1)^2 \neq 0,$$

and the *j-invariant* of Ω by

$$j := j(\Omega) := (12g_2)^3/\Delta = -4 \cdot 12^3 \cdot \frac{(e_1e_2 + e_2e_3 + e_3e_1)^3}{(e_1 - e_2)^2(e_2 - e_3)^2(e_3 - e_1)^2}.$$

Corollary 4.3.5. *For $n \geq 4$ we have the recursion*

$$(n-3)(2n+1)(2n-1)G_{2n} = 3 \sum_{\substack{p\geq 2, q\geq 2 \\ p+q=n}} (2p-1)(2q-1)G_{2p}G_{2q}. \tag{4.3.2}$$

Proof. By differentiating the formula from Proposition 4.3.1 we obtain $\wp'' + 30G_4 = 6\wp^2$. If we plug in the Laurent expansion of $\wp$ given in (4.2.1) we get

$$\begin{aligned} &\sum_{n\geq 2}(2n-1)(2n-2)(2n-3)G_{2n}z^{2n-4} + 30G_4 \\ &= 12\sum_{n\geq 2}(2n-1)G_{2n}z^{2n-4} + 6\sum_{p\geq 2}\sum_{q\geq 2}(2p-1)(2q-1)G_{2p}G_{2q}z^{2p+2q-4}. \end{aligned}$$

Comparing coefficients at z^n gives the stated recursion. □

Example 4.3.6. We have the identities

$$7G_8 = 3G_4^2, \quad 11G_{10} = 5G_4G_6, \quad 143G_{12} = 42G_4G_8 + 25G_6^2 = 18G_4^3 + 25G_6^2.$$

It is easy to see from the above recursions that every G_k can be written as a polynomial over $\mathbb{Q}$ in G_4 and G_6.

Corollary 4.3.7. *We have* $G_k \in \mathbb{Q}[G_4, G_6]$.

This result yields another proof of the fact that the lattice Ω is already determined by its Weierstrass invariants g_2, g_3. Indeed, every G_k is a rational polynomial in g_2, g_3, hence the Laurent expansion (4.2.1) at $z = 0$ is determined by g_2, g_3. By the identity theorem, $\wp$ is determined by its Laurent expansion at 0, and Ω is uniquely determined by its Weierstrass $\wp$-function.

5 The dependence on the lattice

So far, we viewed the Eisenstein series G_k and the Weierstrass $\wp$-function as quantities attached to a fixed lattice Ω. In this chapter, we investigate the behaviour of $G_k(\Omega)$ and $\wp_\Omega$ when the lattice Ω varies.

5.1 Homogeneity and base change

If Ω is a lattice in $\mathbb{C}$, then $\lambda\Omega$ is a lattice for every $0 \neq \lambda \in \mathbb{C}$. From the series definitions of G_k and $\wp$ it is clear that we have

$$\wp_{\lambda\Omega}(\lambda z) = \lambda^{-2}\wp_\Omega(z), \quad \text{and} \quad G_k(\lambda\Omega) = \lambda^{-k}G_k(\Omega). \tag{5.1.1}$$

This also gives the identities

$$g_2(\lambda\Omega) = \lambda^{-4}g_2(\Omega), \tag{5.1.2}$$
$$g_3(\lambda\Omega) = \lambda^{-6}g_3(\Omega), \tag{5.1.3}$$
$$\Delta(\lambda\Omega) = \lambda^{-12}\Delta(\Omega), \tag{5.1.4}$$
$$j(\lambda\Omega) = j(\Omega). \tag{5.1.5}$$

Proposition 5.1.1. *For two lattices Ω and Ω' in $\mathbb{C}$, the following are equivalent.*

1. *We have $\Omega' = \lambda\Omega$ for some $0 \neq \lambda \in \mathbb{C}$.*
2. *$j(\Omega') = j(\Omega)$.*

Proof. We already observed above that $j(\lambda\Omega) = j(\Omega)$ for $\lambda \neq 0$. Conversely, suppose that $j(\Omega') = j(\Omega) \neq 0$. Then we have $g_2(\Omega) \neq 0$ and $g_2(\Omega') \neq 0$. Hence there is some $0 \neq \lambda \in \mathbb{C}$ such that

$$g_2(\Omega') = \lambda^{-4}g_2(\Omega) = g_2(\lambda\Omega).$$

Using $\Delta = g_2^3 - 27g_3^2$ and $\Delta(\lambda\Omega) = \lambda^{-12}\Delta(\Omega)$, we obtain

$$g_3(\Omega') = \pm\lambda^{-6}g_3(\Omega) = \pm g_3(\lambda\Omega).$$

Replacing λ with $i\lambda$ if necessary, we get $g_2(\Omega') = g_2(\lambda\Omega)$ and $g_3(\Omega') = g_3(\lambda\Omega)$. We have seen in Proposition 4.3.2 that g_2 and g_3 uniquely determine the lattice, so we obtain $\Omega' = \lambda\Omega$.

If $j(\Omega) = j(\Omega') = 0$, then $g_2(\Omega) = g_2(\Omega') = 0$, and it follows from Corollary 4.3.4 that $g_3(\Omega) \neq 0$ and $g_3(\Omega') \neq 0$. Now we can proceed in a similar way as before. □

Let (w_1, w_2) be a basis of Ω. Since w_1, w_2 are linearly independent over $\mathbb{R}$, we have $\tau := \frac{w_1}{w_2} \notin \mathbb{R}$. Replacing w_1 with $-w_1$ if necessary, we may assume that $\mathrm{Im}(\tau) > 0$. Hence, every lattice in $\mathbb{C}$ is of the form

$$\Omega = \lambda(\mathbb{Z}\tau + \mathbb{Z})$$

for some $\lambda \in \mathbb{C}$, and τ in the *upper half-plane*

$$\mathbb{H} = \{\tau \in \mathbb{C} : \mathrm{Im}(\tau) > 0\}.$$

Since $\wp$ and G_k are homogeneous in λ, it remains to study their behaviour on lattices $\Omega = \mathbb{Z}\tau + \mathbb{Z}$, as $\tau \in \mathbb{H}$ varies. Hence, we will now view $\wp$ and G_k as functions of $\tau \in \mathbb{H}$, that is, we define

$$\wp(z;\tau) := \wp_{\mathbb{Z}\tau+\mathbb{Z}}(z), \qquad G_k(\tau) := G_k(\mathbb{Z}\tau + \mathbb{Z}).$$

Proposition 5.1.2. *For* $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ *we have*

$$\wp\left(\frac{z}{c\tau+d}; \frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^2 \wp(z;\tau),$$

and

$$G_k\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^k G_k(\tau).$$

Proof. Let $\tau' = \frac{a\tau+b}{c\tau+d}$. Then we have

$$\mathbb{Z}\tau' + \mathbb{Z} = \mathbb{Z}\frac{a\tau+b}{c\tau+d} + \mathbb{Z} = (c\tau+d)^{-1}(\mathbb{Z}(a\tau+b) + \mathbb{Z}(c\tau+d)) = (c\tau+d)^{-1}(\mathbb{Z}\tau+\mathbb{Z}).$$

Here we used that the map $x \mapsto \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) x$ is a bijection on $\mathbb{Z}^2$. By the homogeneity of G_k we obtain

$$G_k(\tau') = G_k(\mathbb{Z}\tau' + \mathbb{Z}) = G_k((c\tau+d)^{-1}(\mathbb{Z}\tau+\mathbb{Z})) = (c\tau+d)^k G_k(\tau),$$

and similarly for $\wp$. □

Remark 5.1.3. The group $\mathrm{SL}_2(\mathbb{R})$ acts on $\mathbb{H}$ by *fractional linear transformations*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.\tau = \frac{a\tau+b}{c\tau+d}.$$

A holomorphic function $f : \mathbb{H} \to \mathbb{C}$ is called a *modular form* of weight $k \in \mathbb{Z}$ for $\mathrm{SL}_2(\mathbb{Z})$ if it satisfies the transformation law

$$f\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^k f(\tau)$$

for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and $\tau \in \mathbb{H}$, and if $f(\tau)$ remains bounded as $\mathrm{Im}(\tau) \to \infty$. If, in addition, $f(\tau)$ goes to 0 as $\mathrm{Im}(\tau) \to \infty$, then f is called a *cusp form*. We will show that $G_k(\tau)$ is a modular form of weight k, and $\Delta(\tau) = \Delta(\mathbb{Z}\tau+\mathbb{Z})$ is a cusp form of weight 12.

5.2 Eisenstein series

For even $k \geq 4$ we may write the Eisenstein series $G_k(\tau)$ for $\tau \in \mathbb{H}$ as the series

$$G_k(\tau) = G_k(\mathbb{Z}\tau+\mathbb{Z}) = \sum_{0 \neq w \in \mathbb{Z}\tau+\mathbb{Z}} w^{-k} = \sum_{m,n\in\mathbb{Z}}{}' (m\tau+n)^{-k}, \tag{5.2.1}$$

were the symbol $\sum'$ means that the summand for $(m,n) = (0,0)$ has to be omitted. Since the sum converges absolutely and locally uniformly, the Eisenstein series $G_k(\tau)$ defines a holomorphic function on $\mathbb{H}$.

Proposition 5.2.1. *For $\tau \in \mathbb{H}$ and even $k \geq 4$ we have the* Fourier expansion

$$G_k(\tau) = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!}\sum_{m=1}^{\infty}\sigma_{k-1}(m)e^{2\pi i m\tau},$$

where $\zeta(s) = \sum_{m=1}^{\infty} m^{-s}$, $(s \in \mathbb{C}, \mathrm{Re}(s) > 1)$, is the Riemann zeta function *and*

$$\sigma_s(m) = \sum_{d|m} d^s, \quad (s \in \mathbb{R}),$$

is a generalized divisor sum. *In particular, $G_k(\tau)$ is a modular form of weight k for* $\mathrm{SL}_2(\mathbb{Z})$.

Proof. Since the Eisenstein series converges absolutely, we may write the series in (5.2.1) as

$$G_k(\tau) = \sum_{n\neq 0} n^{-k} + \sum_{m\neq 0}\sum_{n\in\mathbb{Z}}(m\tau+n)^{-k} + 2\zeta(k) + 2\sum_{m=1}^{\infty}\sum_{n\in\mathbb{Z}}(m\tau+n)^{-k}.$$

We will use the so-called *Lipschitz formula*

$$\sum_{n\in\mathbb{Z}}(\tau+n)^{-k} = \frac{(-2\pi i)^k}{(k-1)!}\sum_{r=1}^{\infty} r^{k-1}e^{2\pi i r\tau}, \qquad (\tau \in \mathbb{H}, k \in \mathbb{N}, k \geq 3), \tag{5.2.2}$$

whose proof will omit here for brevity. Then we obtain

$$G_k(\tau) = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!}\sum_{s=1}^{\infty}\sum_{r=1}^{\infty} r^{k-1}e^{2\pi i rs\tau}.$$

By collecting the terms with $m = rs$ we obtain the stated Fourier expansion.

Since the Fourier expansion of $G_k(\tau)$ does not have terms of negative index, we have $\lim_{\mathrm{Im}(\tau)\to\infty} G_k(\tau) = 2\zeta(k)$, that is, $G_k(\tau)$ remains bounded as $\mathrm{Im}(\tau) \to \infty$. We have already seen above that $G_k(\tau)$ is holomorphic on $\mathbb{H}$ and satisfies the stated transformation law under $\mathrm{SL}_2(\mathbb{Z})$. Hence $G_k(\tau)$ is a modular form of weight k (but it is not a cusp form since $2\zeta(k) \neq 0$). □

Since $\zeta(k) \neq 0$ for even $k \geq 4$, we see that $G_k(\tau)$ does not vanish identically as a function of τ.

Example 5.2.2. Using the formulas $\zeta(4) = \frac{\pi^4}{90}$ and $\zeta(6) = \frac{\pi^6}{945}$ we obtain

$$G_4(\tau) = \frac{\pi^4}{45}\left(1 + 240\sum_{m=1}^{\infty}\sigma_3(m)e^{2\pi i m\tau}\right),$$
$$G_6(\tau) = \frac{2\pi^6}{945}\left(1 - 504\sum_{m=1}^{\infty}\sigma_5(m)e^{2\pi i m\tau}\right).$$

Using the identity $7G_8 = 3G_4^2$ one can now show that

$$7\zeta(8) = 6\zeta^2(4)$$

and

$$\sigma_7(m) = \sigma_3(m) + 120\sum_{\substack{r,s\in\mathbb{N}\\ r+s=m}}\sigma_3(r)\sigma_3(s)$$

for every $m \in \mathbb{N}$. We leave this as an exercise for the reader.

5.3 The discriminant

Recall the definition of the discriminant,

$$\Delta = g_2^3 - 27g_3^2, \qquad \text{where } g_2 = 60G_4, \quad g_3 = 140G_6.$$

Again, we may view Δ as a function on $\mathbb{H}$ by setting

$$\Delta(\tau) = \Delta(\mathbb{Z}\tau + \mathbb{Z}).$$

Proposition 5.3.1. *For $\tau \in \mathbb{H}$ the discriminant $\Delta(\tau)$ has a Fourier expansion of the shape*

$$\Delta(\tau) = (2\pi)^{12} \sum_{m=1}^{\infty} \tau(m) e^{2\pi i m\tau} \tag{5.3.1}$$

with coefficients $\tau(m) \in \mathbb{Z}$ and $\tau(1) = 1$. The discriminant $\Delta(\tau)$ defines a holomorphic function on $\mathbb{H}$ with $\Delta(\tau) \neq 0$ for all $\tau \in \mathbb{H}$. Moreover, it satisfies the functional equation

$$\Delta\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^{12}\Delta(\tau)$$

for every $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. In particular, $\Delta(\tau)$ is a cusp form of weight 12.

Proof. The holomorphicity of $\Delta(\tau)$ on $\mathbb{H}$ and the transformation law immediately follow from the corresponding properties of the Eisenstein series, see Proposition 5.2.1. We have also already seen in Corollary 4.3.4 that $\Delta(\Omega) \neq 0$ for every lattice Ω in $\mathbb{C}$, which implies $\Delta(\tau) \neq 0$ for every $\tau \in \mathbb{H}$. Hence it remains to show that $\Delta(\tau)$ has a Fourier expansion as stated above.

We abbreviate

$$A = \sum_{m=1}^{\infty} \sigma_3(m) e^{2\pi i m\tau}, \qquad B = \sum_{m=1}^{\infty} \sigma_5(m) e^{2\pi i m\tau},$$

which are (up to the missing constant terms) multiples of the Fourier expansions of the Eisenstein series $G_4(\tau)$ and $G_6(\tau)$, see Example 5.2.2. Hence, we find

$$\Delta(\tau) = \frac{(2\pi)^{12}}{1728} \cdot ((1 + 240A)^3 - (1 - 504B)^2) = (2\pi)^{12} \sum_{m=1}^{\infty} \tau(m) e^{2\pi i m}, \tag{5.3.2}$$

where $\tau(1) = 1$, and coefficients $\tau(m) \in \frac{1}{1728}\mathbb{Z}$. It remains to show that the coefficients $\tau(m)$ are integers, that is, the denominator 1728 cancels out. To see this, note that $d^3 \equiv d^5 \pmod{12}$ for $d \in \mathbb{Z}$, and hence $\sigma_3(m) \equiv \sigma_5(m) \pmod{12}$ for $m \in N$. This implies $A \equiv B \pmod{12}$ coefficient-wise. If we work modulo $1728 = 12^3$, we find

$$(1 + 240A)^3 - (1 - 504B)^2 \equiv 12^2(5A + 7B) \equiv 0 \pmod{12^3},$$

which means that the denominator 1728 is cancelled in each coefficient in (5.3.2), that is, $\tau(m) \in \mathbb{Z}$ for every $m \in \mathbb{N}$.

From the shape of the Fourier expansion we see that $\lim_{\mathrm{Im}(\tau)\to\infty} \Delta(\tau) = 0$, so $\Delta(\tau)$ is a cusp form. This finishes the proof. □

Remark 5.3.2. One can show that there are no non-zero cusp forms of weight less than 12, and that every cusp form of weight 12 is a constant multiple of Δ.

Remark 5.3.3. Set $q := e^{2\pi i\tau}$ for brevity. Then the first few coefficients of the Fourier expansion of $(2\pi)^{-12}\Delta(\tau)$ are given by

$$q - 24q^2 + 252q^3 - 1472q^4 + 4830q^5 - 6048q^6 - 16744q^7 + 84480q^8 - 113643q^9 - 115920q^{10} + \dots$$

The discriminant $\Delta(\tau)$ is sometimes referred to as the *Ramanujan Δ-function*, and its coefficients $\tau(m)$ are called *Ramanujan's τ-function*. Using the theory of modular forms one can show that the $\tau(m)$ are multiplicative, that is, they satisfy $\tau(m)\tau(n) = \tau(mn)$ if $\gcd(m,n) = 1$. They have many other interesting and deep properties, some of which are still only conjectured to be true. For example, Ramanujan conjectured in 1916 that the $\tau(p)$ for primes p satisfy the estimate $|\tau(p)| \leq 2p^{11/2}$, which was proved in 1974 by Deligne as a corollary to his celebrated proof of the Riemann hypothesis for zeta functions of algebraic varieties over finite fields. Moreover, Lehmer conjectured in 1947 that $\tau(m) \neq 0$ for all $m \in \mathbb{N}$, which is still an open problem.

5.4 The j-invariant

Finally, we can also view the j-invariant as a function on $\mathbb{H}$,

$$j(\tau) = \frac{(12g_2(\tau))^3}{\Delta(\tau)} = \frac{(720G_4(\tau))^3}{\Delta(\tau)}.$$

Proposition 5.4.1. *For $\tau \in \mathbb{H}$ the j-invariant $j(\tau)$ has a Fourier expansion of the shape*

$$j(\tau) = e^{-2\pi i\tau} + \sum_{m=0}^{\infty} j_m e^{2\pi i m\tau} \tag{5.4.1}$$

with coefficients $j_m \in \mathbb{Z}$. The j-invariant $j(\tau)$ defines a holomorphic function on $\mathbb{H}$ and satisfies the functional equation

$$j\left(\frac{a\tau + b}{c\tau + d}\right) = j(\tau) \tag{5.4.2}$$

for every $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$.

Proof. The holomorphicity and the transformation law of $j(\tau)$ follow from the corresponding properties of $G_4(\tau)$ and $\Delta(\tau)$, and the fact that $\Delta(\tau) \neq 0$ for all $\tau \in \mathbb{H}$. The shape of the Fourier expansion and the integrality of the coefficients j_m can easily be derived using the following general principle: if $f(q) = \sum_{n\geq 0} a_n q^n$ and $g(q) = \sum_{n\geq 0} b_n q^n$ with $a_n, b_n \in \mathbb{Z}$ are convergent power series for $|q| < 1$, with $b_0 = 1$ and $g(q) \neq 0$ for all $|q| < 1$, then $f(q)/g(q)$ is given by a convergent power series $\sum_{n\geq 0} c_n q^n$ for $|q| < 1$ with coefficients $c_n \in \mathbb{Z}$ and $c_0 = a_0$. The proof is easy and will be left as an exercise to the reader. □

Remark 5.4.2. Since the Fourier expansion of $j(\tau)$ has the term $e^{-2\pi i\tau}$, it does not remain bounded as $\mathrm{Im}(\tau) \to \infty$, so strictly speaking it is not a modular form of weight 0. However, it is holomorphic on $\mathbb{H}$, transforms like a modular form of weight 0 under $\mathrm{SL}_2(\mathbb{Z})$, and its Fourier expansion only has finitely many terms of negative index. Such a function is called a *weakly holomorphic modular form of weight* 0. In a similar way as we showed that every elliptic function is a rational function in $\wp$ and $\wp'$, one can show that every weakly holomorphic modular form of weight 0 is a polynomial in $j(\tau)$.

Remark 5.4.3. Let $q = e^{2\pi i\tau}$. The first few coefficients of the Fourier expansion of the j-invariant $j(\tau)$ are given by

$$j(\tau) = q^{-1} + 744 + 196884q + 21493760q^2 + 864299970q^3 + \dots$$

One can show that all the coefficients j_m are positive. Also note that, in comparison to the coefficients $\tau(m)$ of $\Delta(\tau)$, the coefficients j_m of $j(\tau)$ seem to grow much faster. Indeed, using the theory of modular forms one can show that the $\tau(m)$ grow (at most) like m^6, but the j_m grow (at most) like $e^{C\sqrt{m}}$ for some $C > 0$, as $m \to \infty$. The coefficients j_m have a deep interpretation as (linear combinations of) dimensions of irreducible represenations of the monster group. This result, known as *monstrous moonshine* and proved by Borcherds in 1992, is one of the most celebrated (relatively) recent results in number theory.

Proposition 5.4.4. *If τ, τ' satisfy $j(\tau) = j(\tau')$, then there exists a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ such that*

$$\tau' = \frac{a\tau + b}{c\tau + d}.$$

Proof. The assumption $j(\tau) = j(\tau')$ means that $j(\mathbb{Z}\tau + \mathbb{Z}) = j(\mathbb{Z}\tau' + \mathbb{Z})$. Proposition 5.1.1 shows that $\mathbb{Z}\tau' + \mathbb{Z} = \lambda(\mathbb{Z}\tau + \mathbb{Z})$ for some $0 \neq \lambda \in \mathbb{C}$. Hence $(\tau', 1)$ and $(\lambda\tau, \lambda)$ are two bases of the lattice $\mathbb{Z}\tau' + \mathbb{Z}$. By Lemma 2.2.4 there is some matrix $U = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbb{Z})$ such that $\tau' = a\lambda\tau + b\lambda$ and $1 = c\lambda\tau + d\lambda$, so $\tau' = \frac{a\tau+b}{c\tau+d}$. Since τ and τ' both lie in $\mathbb{H}$, and we have $\mathrm{Im}\left(\frac{a\tau+b}{c\tau+d}\right) = \det(U)\frac{\mathrm{Im}(\tau)}{|c\tau+d|^2}$, so we must have $\det(U) > 0$, i.e. $U \in \mathrm{SL}_2(\mathbb{Z})$. □

Proposition 5.4.5. *For every $c \in \mathbb{C}$, there exists some $\tau \in \mathbb{H}$ with $j(\tau) = c$.*

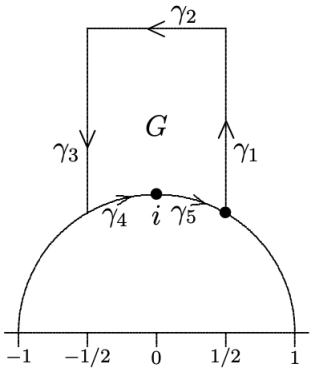
Proof. Suppose that $j(\tau) \neq c$ for all $\tau \in \mathbb{H}$, for some fixed $c \in \mathbb{C}$. Then the function

$$F(\tau) = \frac{j'(\tau)}{j(\tau) - c}$$

is holomorphic on $\mathbb{C}$. We consider the integral

$$\int_\gamma F(\tau)d\tau, \qquad \gamma = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 + \gamma_5,$$

with the path $\gamma = \partial G$ as in the following picture.



It follows from the transformation law (5.4.2) of $j(\tau)$ that

$$F(\tau+1) = F(\tau), \quad F(-1/\tau) = \tau^2 F(\tau).$$

This implies

$$\int_{\gamma_1} F(\tau)d\tau + \int_{\gamma_3} F(\tau)d\tau = \int_{\gamma_4} F(\tau) + \int_{\gamma_5} F(\tau)d\tau = 0.$$

From the shape of the Fourier expansion of the j-invariant we see that $F(\tau)$ has a Fourier expansion of the form

$$F(\tau) = \sum_{m\geq 0} a_m e^{2\pi i m\tau}, \qquad a_0 = -2\pi i.$$

This yields $\int_{\gamma_2} F(\tau)d\tau = 2\pi i$. By the residue theorem we have

$$2\pi i \sum_{\tau\in G} \mathrm{ord}_\tau(j-c) = \int_\gamma F(\tau)d\tau = 2\pi i.$$

But this is a contradiction since $j(\tau) - c$ has no poles, and no zeros by our assumption. This finishes the proof. □

Remark 5.4.6. Recall that the group $\mathrm{SL}_2(\mathbb{Z})$ acts on $\mathbb{H}$ by fractional linear transformations $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right).\tau = \frac{a\tau+b}{c\tau+d}$. Since $j\left(\frac{a\tau+b}{c\tau+d}\right) = j(\tau)$, the j-function can be viewed as a function on the quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. Then Propositions 5.4.4 and 5.4.5 say that the map

$$j : \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H} \to \mathbb{C}$$

is a bijection.

Corollary 5.4.7. *For each $c_2, c_3 \in \mathbb{C}$ with $c_2^3 - 27c_3^2 \neq 0$, there exists precisely one lattice Ω such that*

$$c_2 = g_2(\Omega) \qquad \textit{and} \qquad c_3 = g_3(\Omega).$$

Proof. By Proposition 5.4.5 there exists a lattice Ω such that $j(\Omega) = \frac{(12c_2)^3}{c_2^3 - 27c_3^2}$. We distinguish two cases:

1. $c_2 = 0$. Then we have $j(\Omega) = 0$, hence $g_2(\Omega) = 0$ and $g_3(\Omega) \neq 0$. Choose some $0 \neq \lambda \in \mathbb{C}$ such that $g_3(\Omega) = \lambda^6 c_3$. The homogeneity of g_2, g_3 shows that

   $$g_3(\lambda\Omega) = \lambda^{-6} g_3(\Omega) = c_3 \quad \text{and} \quad g_2(\lambda\Omega) = \lambda^{-4} g_2(\Omega) = 0 = c_2,$$

   so $\lambda\Omega$ is the lattice we are looking for.

2. $c_2 \neq 0$. Then we have $j(\Omega) \neq 0$, hence $g_2(\Omega) \neq 0$. Choose $0 \neq \lambda \in \mathbb{C}$ such that $g_2(\Omega) = \lambda^4 c_2$. We find $g_2(\lambda\Omega) = c_2$, and from $j(\lambda\Omega) = j(\Omega)$ we get $c_3^2 = g_3^2(\lambda\Omega)$. If $c_3 = -g_3(\lambda\Omega)$, we can replace λ by $i\lambda$ to obtain the desired lattice.

The uniqueness of Ω follows from the fact that a lattice is uniquely determined by $g_2(\Omega)$ and $g_3(\Omega)$, compare Proposition 4.3.2. □

6 Product expansions

In this chapter we discuss the Weierstrass σ-function and ζ-function, in order to construct elliptic functions with prescribed zeros and poles. Moreover, we will discuss the Jacobi theta function, and prove Euler's pentagonal number theorem. Throughout, we let Ω be a lattice in $\mathbb{C}$.

6.1 The Weierstrass σ, ζ- and η-function

The Weierstrass σ-function is defined as an infinite product. Hence, e first recall some of the necessary facts about infinite products. Let $a_1, a_2, a_3, \dots$ be a sequence of complex numbers which converges to 0. Then there is some $N \in \mathbb{N}$ such that $|a_n| < 1$ for $n \geq N$. We define the infinite product of the sequence $(1 + a_n)_{n\in\mathbb{N}}$ as

$$\prod_{n=1}^{\infty}(1 + a_n) := (1 + a_1) \cdot \dots \cdot (1 + a_N) \cdot \exp\left(\sum_{n=N+1}^{\infty} \log(1 + a_n)\right),$$

where we choose the principal branch of the logarithm. We say that the infinite product *converges absolutely* if the series $\sum_{n=N+1}^{\infty} |a_n|$ converges. In this case, the sum $\sum_{n=N+1}^{\infty} \log(1+a_n)$ converges absolutely. Moreover, a convergent infinite product equals 0 if and only if one of the factors $1 + a_n$ equals 0.

If $f_1, f_2, \dots$ is a sequence of holomorphic functions on some domain $D \subset \mathbb{C}$, we say that the infinite product $\prod_{n=1}^{\infty}(1 + f_n)$ *converges absolutely and locally uniformly* if the series $\sum_{n=1}^{\infty} f_n$ converges absolutely and locally uniformly. In this case, the infinite product $\prod_{n=1}^{\infty}(1 + f_n)$ defines a holomorphic function on D.

Proposition 6.1.1. *For $z \in \mathbb{C}$ the* Weierstrass σ-function

$$\sigma(z) := \sigma(z; \Omega) := z \prod_{0\neq w\in\Omega} \left(1 - \frac{z}{w}\right) e^{\frac{z}{w} + \frac{1}{2}\left(\frac{z}{w}\right)^2}.$$

converges abolutely and uniformly on every compact subset of $\mathbb{C}$, and hence defines and entire function. It has zeros of first order precisely at the points in Ω. Moreover, $\sigma(z)$ is an odd function.

Proof. To prove the convergence, let $K \subset \mathbb{C}$ be a compact set. A short computation using $e^x = \sum_{n=0} \frac{x^n}{n!}$ gives

$$\left|1 - \left(1 - \frac{z}{w}\right) e^{\frac{z}{w} + \frac{1}{2}\left(\frac{z}{w}\right)^2}\right| = \left|\frac{1}{8}\left(\frac{z}{w}\right)^3 \left(\left(\frac{z}{w}\right)^2 + 3\frac{z}{w} + 4\right) + \sum_{n=3}^{\infty} \frac{\left(\frac{z}{w} + \frac{1}{2}\left(\frac{z}{w}\right)^2\right)^n}{n!}\right| \leq C_K |w|^{-3},$$

with some constant only depending on K. Since the series $\sum_{0\neq w} |w|^{-3}$ converges, we find that the infinite product converges absolutely and locally uniformly.

The σ-function vanishes if and only if $z = 0$ or one of the factors $\left(1 - \frac{z}{w}\right) e^{\frac{z}{w} + \frac{1}{2}\left(\frac{z}{w}\right)^2}$ for some $0 \neq w \in \Omega$ vanishes, which happens precisely at $z = w$. Hence σ has zeros of order 1 precisely at the points in Ω.

To see that $\sigma(z)$ is odd, replace z with $-z$ and then replace w with $-w$ in the infinite product. □

Proposition 6.1.2. *For $z \in \mathbb{C} \setminus \Omega$ the* Weierstrass ζ-function

$$\zeta(z) := \zeta(z; \Omega) := \frac{\sigma'(z)}{\sigma(z)} = \frac{1}{z} + \sum_{0 \neq w \in \Omega} \left(\frac{1}{z - w} + \frac{1}{w} + \frac{z}{w^2} \right)$$

converges absolutely and uniformly on every compact subset of $\mathbb{C} \setminus \Omega$. It has poles of first order and residue 1 precisely at the points in Ω. Moreover, $\zeta(z)$ is an odd function.

Proof. For z in a compact subset $K \subset \mathbb{C}$ we estimate

$$\left| \frac{1}{z - w} + \frac{1}{w} + \frac{z}{w^2} \right| = \left| \frac{z^2}{w^2(z - w)} \right| \leq C_K |w|^{-3},$$

where the constant C_K only depends on K. Now the absolute and uniform convergence of $\zeta(z)$ on compact subsets of $\mathbb{C} \setminus \Omega$ follows from the convergence lemma 2.3.2. The statement about the poles is clear, and the fact that $\zeta(z)$ is odd can either be seen from the series definition, of from the fact that $\sigma(z)$ is odd (hence $\sigma'(z)$ is even). □

The ζ-function is closely related to the $\wp$-function.

Corollary 6.1.3. *For $z \in \mathbb{C} \setminus \Omega$ we have*

$$\zeta'(z) = -\wp(z).$$

Proof. This follows by comparing the infinite series representation of $\wp$ from Proposition 4.1.1 with the infinite series obtained by differentiating ζ termwise. □

Corollary 6.1.4. *We have the Laurent expansion*

$$\zeta(z; \Omega) = \frac{1}{z} - \sum_{k=2}^{\infty} G_{2k}(\Omega) z^{2k-1}$$

around $z = 0$.

Proof. This can be proved in the same way as the analogous result for $\wp$, Proposition 4.2.1. □

The ζ-function is not elliptic. However, we have the following result.

Lemma 6.1.5. *For $w \in \Omega$ the* Weierstrass η-function

$$\eta(w) := \eta(w; \Omega) := \zeta(z + w) - \zeta(z)$$

is independent of the choice of $z \in \mathbb{C} \setminus \Omega$. In particular, we have

$$\eta(w + w') = \eta(w) + \eta(w'), \qquad w, w' \in \Omega,$$

that is, $\eta : \Omega \to \mathbb{C}$ is a group homomorphism.

Proof. Using that $\wp$ is elliptic, we find

$$(\zeta(z+w)-\zeta(z))' = -\wp(z+w)-\wp(z) = 0,$$

so $\zeta(z+w)-\zeta(z)$ is independent of z. We can now compute

$$\begin{aligned}\eta(w+w') &= \zeta(z+w+w')-\zeta(z) = \zeta(z+w)-\zeta(z+w)+\zeta(z+w+w')-\zeta(z)\\ &= (\zeta(z+w)-\zeta(z))+(\zeta((z+w)+w')-\zeta(z+w)) = \eta(w)+\eta(w').\end{aligned}$$

This finishes the proof. □

The homomorphism η satisfies the so-called *Legendre relation*:

Proposition 6.1.6. *Let $\Omega = \mathbb{Z}w_1 + \mathbb{Z}w_2$ with $\mathrm{Im}(w_1/w_2) > 0$. Then we have*

$$\eta(w_2)w_1 - \eta(w_1)w_2 = 2\pi i.$$

In particular, for $w, w' \in \Omega$ we have

$$\eta(w)w' - \eta(w')w \in 2\pi i\mathbb{Z}.$$

Proof. Let $P = P(u; w_1, w_2)$ be a fundamental parallelogram, where the base point $u \in \mathbb{C}$ is chosen such that 0 lies in the interior of P. We integrate $\zeta(z)$ over the positively oriented boundary of P and use the residue theorem to get

$$\int_{\partial P} \zeta(z)dz = 2\pi i,$$

since $\zeta(z)$ has precisely one pole of first order and residue 1 in P, namely at $z = 0$. On the oter hand, we have

$$\begin{aligned}\int_{\partial P}\zeta(z)dz &= \left(\int_u^{u+w_2} + \int_{u+w_2}^{u+w_1+w_2} + \int_{u+w_1+w_2}^{u+w_1} + \int_{u+w_1}^{u}\right)\zeta(z)dz\\ &= \int_u^{u+w_2}(\zeta(z)-\zeta(z+w_1))dz + \int_{u+w_1}^{u}(\zeta(z)-\zeta(z+w_2))dz\\ &= \eta(w_1)w_2 - \eta(w_2)w_1,\end{aligned}$$

where we used that the parallelogram $(0, w_2, w_1+w_2, w_1)$ is positively oriented. Using that $\eta : \Omega \to \mathbb{C}$ is a homomorphism we also get $\eta(w)w' - \eta(w')w \in 2\pi i\mathbb{Z}$ for $w, w' \in \Omega$. □

Finally, note that for $0 \neq \lambda \in \mathbb{C}$ we have

$$\begin{aligned}\sigma(\lambda z; \lambda\Omega) &= \lambda\sigma(z;\Omega),\\ \zeta(\lambda z; \lambda\Omega) &= \frac{1}{\lambda}\zeta(z;\Omega),\\ \eta(\lambda w; \lambda\Omega) &= \frac{1}{\lambda}\eta(w;\Omega).\end{aligned}$$

6.2 The transformation law for σ

Since $\sigma(z;\Omega)$ is entire, it cannot be an elliptic function for Ω. However, it satsifies an interesting transformation law under $z \mapsto z + w$ for $w \in \Omega$.

Theorem 6.2.1. *For $w \in \Omega$ and $z \in \mathbb{C}$ we have*

$$\sigma(z+w) = \chi(w) e^{\eta(w)(z+w/2)} \sigma(z),$$

where

$$\chi(w) = \begin{cases} 1, & \text{if } w/2 \in \Omega, \\ -1, & \text{if } w/2 \notin \Omega. \end{cases}$$

Proof. Since σ vanishes at points in Ω, both sides of the above equation vanish for $z \in \Omega$, so we can assume $z \notin \Omega$ for the rest of the proof. In this case, we have $\sigma'(z) = \sigma(z)\zeta(z)$ by definition of the ζ-function. Moreover, using the definition of $\eta(w)$ we obtain

$$\begin{aligned} \frac{d}{dz}\left(\frac{\sigma(z+w)}{\sigma(z)}\right) &= \frac{\sigma'(z+w)\sigma(z) - \sigma(z+w)\sigma'(z)}{\sigma(z)^2} \\ &= \frac{\sigma(z+w)\zeta(z+w)\sigma(z) - \sigma(z+w)\sigma(z)\zeta(z)}{\sigma(z)^2} \\ &= \frac{\sigma(z+w)}{\sigma(z)}\eta(w). \end{aligned}$$

Hence, the value

$$\psi(w) := \frac{\sigma(z+w)}{\sigma(z)} e^{-\eta(w)(z+w/2)}$$

is independent of z. We want to show that it equals $\chi(w)$. If $w/2 \notin \Omega$, we may choose $z = -w/2$ and use that σ is odd, to obtain $\psi(w) = -1 = \chi(w)$. If $0 \neq w/2 \in \Omega$, we first write (for any $w \in \Omega$)

$$\psi(2w) = \frac{\sigma(z+2w)\sigma(z+w)}{\sigma(z+w)\sigma(z)} e^{-2\eta(w)(z+w)} = \psi(w)^2, \tag{6.2.1}$$

where we used that η is a homomorphism. Since Ω is discrete, there is some natural number $n \geq 1$ such that $w' = 2^{-n}w \in \Omega$ but $\frac{1}{2}w' = 2^{-n-1}w \notin \Omega$. Above we have seen that $\psi(w') = -1$. We obtain from (6.2.1)

$$\psi(w) = \psi(2^n w') = \psi(w')^{2^n} = (-1)^{2^n} = 1,$$

since $n \geq 1$. This shows $\psi = \chi$, and finishes the proof. □

Corollary 6.2.2. *If we put $f(z) = \frac{\sigma(z-a)}{\sigma(z-b)}$ with $a, b \in \mathbb{C}$, then we have for $w \in \Omega$ and $z \in \mathbb{C}, z \notin b + \Omega$ the transformation law*

$$f(z+w) = e^{\eta(w)(b-a)} f(z).$$

6.3 Existence of elliptic functions with prescribed zeros and poles

Let f be a non-constant elliptic function for a lattice Ω with fundamental parallelogram P. If we list the zeros $a_1, \dots, a_r$ and poles $b_1, \dots, b_r$ of f in P (with repetitions to account for multiplicities; compare Theorem 3.2.3), then Abel's relation (Theorem 3.2.4) says that

$$a_1 + \cdots + a_r \equiv b_1 + \cdots + b_r \pmod{\Omega}.$$

Conversely, we have the following existence theorem.

Theorem 6.3.1. *Let $a_1, \dots, a_r$ and $b_1, \dots, b_r$ be two finite sequences in $\mathbb{C}$, such that the sets $\{a_1 + \Omega, \dots, a_r + \Omega\}$ and $\{b_1 + \Omega, \dots, b_r + \Omega\}$ are disjoint, and such that*

$$w_0 := (b_1 + \cdots + b_r) - (a_1 + \cdots + a_r)$$

lies in Ω. Then

$$f(z) := e^{-\eta(w_0)z} \frac{\sigma(z - a_1) \cdots \sigma(z - a_r)}{\sigma(z - b_1) \cdots \sigma(z - b_r)}$$

is an elliptic function which has zeros precisely at the points $a_1 + \Omega, \dots, a_r + \Omega$ and poles precisely at the points $b_1 + \Omega, \dots, b_r + \Omega$ (where the order at such a point is given by the number of repetitions). Moreover, every elliptic function with zeros at $a_1, \dots, a_r$ and poles at $b_1, \dots, b_r$ is a constant multiple $f(z)$.

Proof. If we put $f_{a,b}(z) = \frac{\sigma(z-a)}{\sigma(z-b)}$, then we may write

$$f(z) = e^{-\eta(w_0)z} \prod_{j=1}^{r} f_{a_j, b_j}(z),$$

and it follows from Corollary 6.2.2 that

$$\begin{aligned} f(z + w) &= e^{-\eta(w_0)z} e^{-\eta(w_0)w} \prod_{j=1}^{r} e^{\eta(w)(b_j - a_j)} \prod_{j=1}^{r} f_{a_j, b_j}(z) \\ &= e^{-\eta(w_0)w + \eta(w)w_0} f(z) = f(z), \end{aligned}$$

where we used that $-\eta(w_0)w + \eta(w)w_0 \in 2\pi i\mathbb{Z}$ by the Legendre relation, Proposition 6.1.6. The statement about the order immediately follows from the fact that $\sigma(z)$ is entire and has zeros of order 1 precisely at the points in Ω.

Since two elliptic functions with the same zeros and poles (and the same multiplicities) only differ by a constant factor (compare Theorem 3.2.1), every elliptic function with the same zeros and poles as $f(z)$ is a constant multiple of $f(z)$ as defined in the theorem. □

6.4 The Jacobi theta function and the pentagonal number theorem

The *Jacobi theta function* for the lattice $\Omega = \mathbb{Z}\tau + \mathbb{Z}$ for $\tau \in \mathbb{H}$ is defined by

$$\vartheta(z; \tau) := \sum_{n \in \mathbb{Z}} e^{\pi i n^2 \tau + 2\pi i n z}.$$

Lemma 6.4.1. *The Jacobi theta function converges absolutely and locally uniformly on $\mathbb{C}\times\mathbb{H}$. For every fixed $\tau \in \mathbb{H}$ it defines an entire function in z, which has zeros (at least) at the points $\frac{\tau+1}{2} + \Omega$. Moreover, it satisfies the transformation laws*

$$\vartheta(z+1;\tau) = \vartheta(z;\tau) \qquad \text{and} \qquad \vartheta(z+\tau;\tau) = e^{-\pi i\tau - 2\pi iz}\vartheta(z;\tau).$$

Proof. Let $K \subset \mathbb{C}\times\mathbb{H}$ be compact. Then there exists some $\varepsilon > 0$ such that $|\mathrm{Im}(z)| \leq 1/\varepsilon$ and $\mathrm{Im}(\tau) > \varepsilon$ for all $(z,\tau) \in K$. We find

$$\sum_{n\in\mathbb{Z}} |e^{\pi in^2\tau + 2\pi inz}| = \sum_{n\in\mathbb{Z}} e^{-\pi n^2 \mathrm{Im}(\tau) - 2\pi n \mathrm{Im}(z)} \leq 1 + 2\sum_{n=1}^{\infty} e^{-\pi n^2\varepsilon + 2\pi n/\varepsilon}$$

for all $(z,\tau) \in K$. The series on the right-hand side can be estimated by $\sum_{n=1}^{\infty} e^{-Cn^2}$ for a suitable $C > 0$, and this series converges as a subseries of the geometric series. In particular, $\vartheta(z;\tau)$ defines an entire function in z, and a holomorphic function in $\tau \in \mathbb{H}$.

The rule $\vartheta(z+1;\tau) = \vartheta(z;\tau)$ is clear from the definition, and we have

$$\vartheta(z+\tau;\tau) = \sum_{n\in\mathbb{Z}} e^{\pi in^2\tau + 2\pi in(z+\tau)} = e^{-\pi i\tau - 2\pi iz} \sum_{n\in\mathbb{Z}} e^{\pi i(n+1)^2\tau + 2\pi i(n+1)z} = e^{-\pi i\tau - 2\pi iz}\vartheta(z;\tau).$$

Moreover, we compute

$$\vartheta(\tfrac{\tau+1}{2};\tau) = \sum_{n\in\mathbb{Z}} (-1)^n e^{\pi in(n+1)\tau} = \sum_{m\in\mathbb{Z}} (-1)^{-m-1} e^{\pi i(-m-1)(-m)\tau} = -\vartheta(\tfrac{\tau+1}{2};\tau),$$

hence $\vartheta(\frac{\tau+1}{2};\tau) = 0$. From the transformation law we get that $\vartheta(z;\tau)$ has zeros at $\frac{\tau+1}{2}+\Omega$. □

We prove the *Jacobi triple product identity*, which is an infinite product expansion for the Jacobi theta function.

Theorem 6.4.2. *We put $q = e^{2\pi i\tau}$ and $\xi = e^{2\pi iz}$ with $\tau \in \mathbb{H}$ and $z \in \mathbb{C}$. Then we have*

$$\vartheta(z;\tau) = \prod_{m=1}^{\infty} (1-q^m)(1+\xi q^{m-1/2})(1+\xi^{-1}q^{m-1/2}).$$

In particular, $z \mapsto \vartheta(z;\tau)$ has roots of first order precisely at the points $\frac{\tau+1}{2} + \Omega$.

Proof. The left-hand side is just the Jacobi theta function $\vartheta(z;\tau)$. We denote the right-hand side by $g(z;\tau)$. It is an entire function which has simple zeros in the points $z \in \frac{\tau+1}{2} + \Omega$. Moreover, one may check directly that we have

$$g(z+1;\tau) = g(z;\tau)$$

and

$$\frac{g(z+\tau;\tau)}{g(z;\tau)} = \frac{1+\xi^{-1}q^{-1/2}}{1+\xi q^{1/2}} = e^{-\pi i\tau - 2\pi iz},$$

which means that

$$g(z+\tau;\tau) = e^{-\pi i\tau - 2\pi iz} g(z;\tau).$$

Comparing this with the transformation rules of the Jacobi theta function from the last lemma, we see that for each fixed $\tau \in \mathbb{H}$ the function $\varphi(z;\tau) := \vartheta(z;\tau)/g(z;\tau)$ is an entire

elliptic function in z, and hence constant in z. Thus we may just write $\varphi(\tau)$ for $\varphi(z;\tau)$. It is now useful to view φ as a function of $q \in \mathbb{C}$ with $0 < |q| < 1$, by setting

$$\varphi(q) = \frac{\sum_{n\in\mathbb{Z}} \xi^n q^{n^2/2}}{\prod_{m=1}^{\infty}(1-q^m)(1+\xi q^{m-1/2})(1+\xi^{-1}q^{m-1/2})}.$$

The right-hand side is holomorphic in q, and extends to a holomorphic function at $q = 0$, with value $\varphi(0) = 1$. Moreover, a direct but tedious computation (which we leave as an exercise) shows that

$$\frac{\vartheta(1/4;\tau)}{g(1/4;\tau)} = \frac{\vartheta(1/2;4\tau)}{g(1/2;4\tau)},$$

which translates into

$$\varphi(q) = \varphi(q^4),$$

and inductively to

$$\varphi(q) = \varphi(q^{4k}), \qquad \text{for all } k \in \mathbb{N}.$$

Since $|q| < 1$, we see that q^{4k} tends to 0 for $k \to \infty$. Hence the identity theorem gives $\varphi(q) \equiv 1$ for all $q \in \mathbb{C}$ with $|q| < 1$, which shows that $\vartheta(z;\tau) = g(z;\tau)$. □

We can use the Jacobi theta function to construct elliptic functions.

Corollary 6.4.3. *For $a_1, \ldots, a_r, b_1, \ldots, b_r \in \mathbb{C}$ with $(a_1 + \cdots + a_r) - (b_1 + \cdots + b_r) \in \mathbb{Z}$ the function*

$$f(z) := \frac{\vartheta(z-a_1;\tau)\cdots\vartheta(z-a_r;\tau)}{\vartheta(z-b_1;\tau)\cdots\vartheta(z-b_r;\tau)}$$

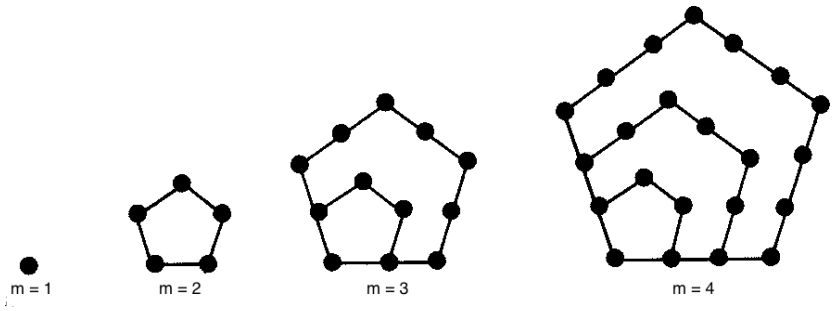
is an elliptic function for the lattice $\Omega = \mathbb{Z}\tau + \mathbb{Z}$. If the sets $\{a_1 + \Omega, \ldots, a_r + \Omega\}$ and $\{b_1 + \Omega, \ldots, b_r + \Omega\}$ are disjoint, then $f(z)$ has zeros at the points $z \in \frac{\tau+1}{2} + a_j + \Omega$ and poles in the points $z \in \frac{\tau+1}{2} + b_j + \Omega$ for $1 \le j \le r$, where the order is given by the number of repetitions of the a_j and b_j.

Replacing τ by $3\tau/2$ and setting $z = (\tau+2)/4$ in the Jacobi triple product identity, we obtain *Euler's pentagonal number theorem.*

Theorem 6.4.4. *Put $q = e^{2\pi i\tau}$ for $\tau \in \mathbb{H}$. Then we have*

$$\prod_{m=1}^{\infty}(1-q^m) = \sum_{n\in\mathbb{Z}}(-1)^n q^{(3n^2-n)/2}.$$

Remark 6.4.5. 1. The m-th *pentagonal number* is the number of distinct edges of m regular pentagons of side lengths $0, 1, 2, \ldots, (m-1)$, which are overlaid as in the following pictures:



The pentagonal numbers are obtained by the formula $(3m^2-m)/2$ with $m = 1, 2, 3, \ldots$. The first few pentagonal numbers are $1, 5, 12, 22, \ldots$. If we allow all $m \in \mathbb{Z}$, we obtain the *generalized pentagonal numbers* $(3m^2 - m)/2$, the first few of which are given by $0, 1, 2, 5, 7, 12, \ldots$. These are exactly the exponents at q in the series

$$\sum_{m\in\mathbb{Z}} (-1)^m q^{(3m^2-m)/2} = 1 - q - q^2 + q^5 + q^7 - q^{12} + \ldots$$

appearing in Euler's pentagonal theorem above.

2. A *partition* of $m \in \mathbb{N}$ is a tuple $(\lambda_1, \ldots, \lambda_k)$ of positive integers with $\lambda_1 \leq \cdots \leq \lambda_k$ such that $m = \lambda_1 + \cdots + \lambda_k$. The λ_j are called the *parts* of the partition. For example, the partitions of $m = 4$ are $(4), (1,3), (2,2), (1,1,2), (1,1,1,1)$. Its partitions into *distinct parts* are $(4), (1,3)$.

   Let $p^+(m)$ and $p^-(m)$ denote the number of partitions of m into distinct parts with an even and odd number of parts, respectively. For example, for $m = 4$ the partitions into different parts are (4) and $(1,3)$, so we have $p^+(4) = 1$ and $p^-(4) = 1$. By multiplying out

   $$\prod_{m=1}^{\infty} (1 - q^m) = (1-q)(1-q^2)(1-q^3)\cdots = \sum_{m=0}^{\infty} (p^+(m) - p^-(m))q^m,$$

   we see that Euler's pentagonal theorem is equivalent to the combinatorial identity

   $$p^+(m) - p^-(m) = \begin{cases} (-1)^n & \text{if } m = (3n^2-n)/2 \text{ for some } n \in \mathbb{Z}, \\ 0, & \text{otherwise.} \end{cases}$$

3. Let $p(n)$ denote the number of *all* partitions of n (possibly with repeating parts), and let $p(0) = 1$. For example, we have seen above that $p(4) = 5$. Using the geometric series it is not hard to show that

   $$\prod_{m=1}^{\infty} (1 - q^m)^{-1} = \sum_{n=0}^{\infty} p(n)q^n.$$

6.4.1 The Four Squares Theorem

In this section, we briefly sketch a proof of Lagrange's Four Squares Theorem, using the modularity properties of the Jacobi theta function.

Theorem 6.4.6 (Lagrange 1770)**.** *Every natural number $n \in \mathbb{N}$ can be written as a sum of four squares.*

In other words, if we let

$$r_4(n) := \{(a,b,c,d) \in \mathbb{Z}^4 : a^2 + b^2 + c^2 + d^2 = n\}$$

denote the number of ways to write n as a sum of four squares, then Lagrange's Theorem says that $r_4(n) \geq 1$ for every $n \in \mathbb{N}$. The so-called *representation number* $r_4(n)$ is related to the *Theta Nullwert*

$$\vartheta(\tau) := \vartheta(0;\tau) = \sum_{n\in\mathbb{Z}} e^{\pi i n^2 \tau}$$

by

$$\vartheta(\tau)^4 = \sum_{a,b,c,d\in\mathbb{Z}} e^{\pi i(a^2+b^2+c^2+d^2)\tau} = \sum_{n=0}^{\infty} r_4(n)e^{\pi in\tau},$$

that is, $r_4(n)$ is the n-th Fourier coefficient of $\vartheta(\tau)^4$. The first important ingredient for the proof of the Four Squares Theorem is the *theta transformation formula*.

Proposition 6.4.7. *For $\tau \in \mathbb{H}$ we have*

$$\vartheta\left(-\frac{1}{\tau}\right) = \sqrt{\frac{\tau}{i}}\vartheta(\tau),$$

where we take the principal branch of the square root.

The proof uses the *Poisson summation formula*, but we will skip it for brevity. The important observation is that the theta transformation law resembles the transformation law of a modular form of weight $\frac{1}{2}$ under the matrix $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. More precisely, using the theta transformation formula, one can show:

Corollary 6.4.8. *The function $\vartheta(2\tau)^4$ is a modular form of weight 2 for the group*

$$\Gamma_0(4) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : c \equiv 0 \pmod 4 \right\}.$$

Using the general theory of modular forms (for subgroups of $\mathrm{SL}_2(\mathbb{Z})$) one can show that the space of all modular forms of weight 2 for $\Gamma_0(4)$ is finite-dimensional (in fact, it has dimension 2). Moreover, one can construct an explicit basis using certain Eisenstein series. In this case, one can show that

$$\vartheta(2\tau)^4 = \frac{1}{3}(4E_2(4\tau) - E_2(\tau))$$

with the (non-modular) Eisenstein series

$$E_2(\tau) = 1 - 24\sum_{n=1}^{\infty} \sigma_1(n)e^{2\pi in\tau}.$$

A short computation now yields the Fourier expansion

$$\vartheta(2\tau)^4 = 1 + \sum_{n=1}^{\infty} \left(8 \sum_{\substack{d|n \\ 4\nmid d}} d\right) e^{2\pi in\tau}.$$

If we recall from above that the n-th Fourier coefficient of $\vartheta(2\tau)^4$ is given by $r_4(n)$, we obtain a refined version of Lagrange's Four Squares Theorem.

Theorem 6.4.9. *For every $n \in \mathbb{N}$ we have*

$$r_4(n) = 8 \sum_{\substack{d|n \\ 4\nmid d}} d$$

In particular, every $n \in \mathbb{N}$ can be written as a sum of four squares.

By looking at $\vartheta(\tau)^k$, one can get explicit formulas for $r_k(n)$, the number of ways to write n as a sum of k squares, for many values of k.

7 Elliptic curves and the addition theorem for the $\wp$-function

In this chapter we will discuss the connection between lattices in $\mathbb{C}$ and elliptic curves over the complex numbers.

7.1 The addition theorem for the $\wp$-function

Let $\Omega = \mathbb{Z}w_1 + \mathbb{Z}w_2$ be a lattice in $\mathbb{C}$, and let $\wp(z) = \wp_\Omega(z)$ be its Weierstrass-$\wp$ function.

Theorem 7.1.1 (Addition Theorem)**.** *For $z, w \in \mathbb{C}$ with $z, w, z \pm w \notin \Omega$ we have*

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

For the proof of the addition theorem, we collect some properties of the function on the right-hand side.

Proposition 7.1.2. *For fixed $w \in \mathbb{C} \setminus \frac{1}{2}\Omega$ the function*

$$f(z) := f(z; w) := \frac{1}{2}\frac{\wp'(z) - \wp'(w)}{\wp(z) - \wp(w)}$$

is an elliptic function with respect to Ω with poles of first order precisely at the points $z \in \Omega$ and $z \in -w + \Omega$, with Laurent expansions

$$f(z; w) = -\frac{1}{z} - \wp(w)z + O(z^2), \quad \text{around } z = 0, \tag{7.1.1}$$

$$f(z; w) = \frac{1}{z+w} + c(w) + O(z+w), \quad \text{around } z = -w, \tag{7.1.2}$$

with a constant $c(w) \in \mathbb{C}$ (we will see in the proof of the addition theorem that we have $c(w) = 0$).

Proof. Note that f is not defined at the points $z \in \Omega$ and $z \in -w + \Omega$, but also at the points $z \in w + \Omega$. But since

$$\lim_{z \to w} f(z; w) = \frac{1}{2}\lim_{z \to w}\frac{(\wp'(z) - \wp'(w))/(z-w)}{(\wp(z) - \wp(w))/(z-w)} = \frac{1}{2}\frac{\wp''(w)}{\wp'(w)},$$

the points $z \in w + \Omega$ are removable singularities, that is, $f(z)$ is holomorphic there (here we used that $w \notin \frac{1}{2}\Omega$, hence $\wp'(w) \neq 0$). In order to determine the shape of the Laurent exansion of $f(z)$ around $z = 0$, we can use the Laurent expansions $\wp(z) = z^{-2} + O(z^2)$ and $\wp'(z) = -2z^{-3} + O(z)$. Moreover, $\wp(z) - \wp(w)$ has a simple root at every $z \in -w + \Omega$, and $\wp'(z) - \wp'(w) = -2\wp'(w) \neq 0$ (since $\wp$ is even, $\wp'$ is odd, and $w \notin \frac{1}{2}\Omega$; compare Lemma 3.3.3). Hence $f(z)$ has a simple pole at every $z \in -w + \Omega$. Since the sum of the residues of f in a fundamental parallelogram is 0 (compare Theorem 3.2.2), and the residue at $z = 0$ equals -1, the residues at points $z \in -w + \Omega$ must be 1. This gives the stated Laurent expansions. □

Proof of Theorem 7.1.1. We consider the elliptic function

$$g(z) = f(z;w)^2 - \wp(z+w) - \wp(z) - \wp(w), \qquad w \in \mathbb{C} \setminus \tfrac{1}{2}\Omega.$$

Then $g(z)$ is holomorphic apart from possible poles at the points $z \in \Omega$ and $z \in -w + \Omega$. At $z = 0$ we have

$$g(z) = (z^{-2} + 2\wp(w)) - \wp(w) - z^{-2} - \wp(w) + O(z) = O(z),$$

and at $z = -w$ we have

$$g(z) = \frac{1}{(z-w)^2} + \frac{2c(w)}{z+w} - \frac{1}{(z+w)^2} + O(1) = \frac{2c(w)}{z+w} + O(1).$$

If $c(w) \neq 0$ then $g(z)$ would have simple poles only at the points $z \in -w + \Omega$, which is impossible since the sum of the residues in a fundamental domain must be 0 by Theorem 3.2.2. This implies $c(w) = 0$. Hence $g(z)$ is holomorphic everywhere, and thus constant by Theorem 3.2.1. From the Laurent expansion $g(z) = O(z)$ at $z = 0$ we find $g(z) = 0$, which finishes the proof of the addition theorem in the case that $w \notin \frac{1}{2}\Omega$. But both sides of the addition theorem are holomorphic near $w \in \frac{1}{2}\Omega \setminus \Omega$, so for these points w the addition theorem follows by continuity. □

As a special case, we obtain the following *duplication formula* for the $\wp$-function.

Corollary 7.1.3. *For $z \in \mathbb{C} \setminus \frac{1}{2}\Omega$ we have*

$$\wp(2z) = \frac{1}{4}\left(\frac{12\wp(z)^2 - g_2}{2\wp'(z)}\right)^2 - 2\wp(z).$$

Proof. First, by letting $w \to z$ in the addition theorem, we obtain

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

Now the duplication formula follows from the differential equation

$$2\wp''(z) = 12\wp(z)^2 - g_2,$$

which in turn follows by differentiating the differential equation $\wp'(z)^2 = 4\wp(z)^3 - g_2\wp(z) - g_3$ from Proposition 4.3.1. □

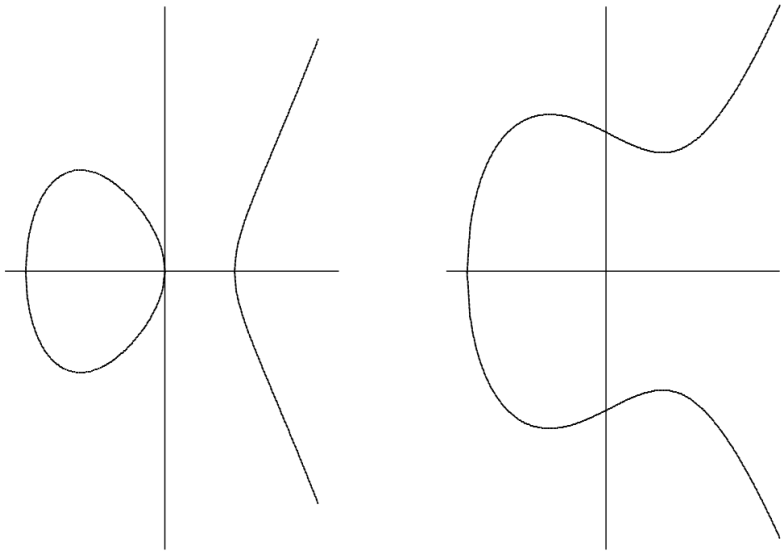
7.2 Elliptic curves over $\mathbb{C}$

Let $\Omega = \mathbb{Z}w_1 + \mathbb{Z}w_2$ be a lattice in $\mathbb{C}$ with Weierstrass invariants $g_2 = g_2(\Omega)$ and $g_3 = g_3(\Omega)$. The subset

$$E := E(\Omega) := \{(X, Y) \in \mathbb{C} \times \mathbb{C} : Y^2 = 4X^3 - g_2X - g_3\}$$

of $\mathbb{C} \times \mathbb{C}$ is called the (affine) *elliptic curve* associated to Ω.

Example 7.2.1. Let us suppose that g_2, g_3 are *real* numbers (recall from Corollary 5.4.7 that every pair (g_2, g_3) of complex numbers with $g_2^3 - 27g_3^2 \neq 0$ appears as the Weierstrass invariants of a unique lattice in $\mathbb{C}$). Then we may look at the real points on E, that is, the real solutions $(X, Y) \in \mathbb{R}^2$ of the equation $Y^2 = 4X^3 - g_2X - g_3$. The right-hand side has either three or one real root. Typical examples of such curves are $Y^2 = 4X^3 - 4X$ (with real roots $X = 0, \pm 1$) and $Y^2 = 4X^3 + 4$ (with real root $X = -1$). Their real points look as follows.



Using the Weierstrass $\wp$-function, we obtain a parametrization of E.

Proposition 7.2.2. *The map*

$$\Phi : (\mathbb{C}/\Omega) \setminus \{\Omega\} \to E(\Omega), \quad \Phi(z + \Omega) = (\wp(z), \wp'(z))$$

is a bijection.

Proof. First note that Φ is well-defined since $\wp$ and $\wp'$ are elliptic. Moreover, the differential equation $\wp'(z)^2 = 4\wp(z)^2 - g_2\wp(z) - g_3$ from Proposition 3.3.5 shows that the image of Φ is indeed contained in E.

For $(X, Y) \in E$ we choose some $z \in \mathbb{C}$ with $\wp(z) = X$, compare Lemma 3.3.4. Then we have

$$Y^2 = 4X^3 - g_2X - g_3 = 4\wp(z)^3 - g_2\wp(z) - g_3 = \wp'(z)^2,$$

where we again used the differential equation for $\wp(z)$. Hence we either have $Y = \wp'(z)$ or $Y = -\wp'(z)$. Since $\wp(z)$ is even and $\wp'(z)$ is odd, we may assume that $Y = \wp'(z)$ by replacing z with $-z$ if necessary. This shows that (X, Y) lies in the image of Φ, so Φ is surjective.

Now suppose that there are $z_1, z_2 \in \mathbb{C} \setminus \Omega$ with $(\wp(z_1), \wp'(z_1)) = (\wp(z_2), \wp'(z_2))$. The identity $\wp(z_1) = \wp(z_2)$ implies $z_1 \equiv \pm z_2 \pmod{\Omega}$ by Lemma 3.3.4. If $\wp'(z_1) \neq 0$, then $z_1 \not\equiv -z_1 \pmod{\Omega}$ since $\wp'(z)$ is odd, so we must have $z_1 \equiv z_2 \pmod{\Omega}$. If $\wp'(z_1) = 0$ (hence $\wp'(z_2) = 0$), then each of z_1 and z_2 is equivalent to one of $w_1/2, w_2/2, w_3/2 \pmod{\Omega}$ by Lemma 3.3.3. But since the values $\wp(w_k) = e_k$ are pairwise different by (3.3.5), we must have $z_1 \equiv z_2 \pmod{\Omega}$. This shows that Φ is injective. □

It is somewhat inconvenient that the trivial coset Ω in $\mathbb{C}/\Omega$ does not correspond to a point on E via Φ. To solve this issue, we add a 'point at infinity' $\mathcal{O}$ to the elliptic curve E,

$$\overline{E} := \overline{E}(\Omega) := E \cup \{\mathcal{O}\},$$

We will think of $\mathcal{O}$ as the 'point' (∞, ∞). Then we may extend the map Φ to a bijection

$$\Phi : \mathbb{C}/\Omega \to \overline{E}(\Omega), \qquad \Phi(z+\Omega) = \begin{cases} (\wp(z), \wp(z')), & \text{if } z \notin \Omega, \\ \mathcal{O}, & \text{if } z \in \Omega, \end{cases}$$

Recall from (2.2.1) that we may identify $\mathbb{C}/\Omega$ with a fundamental parallelogram for Ω, which can in turn be thought of as a torus. Hence, via the map Φ we may think of the elliptic curve $\overline{E}$ as a torus.

Using the bijection Φ, we can now carry over the natural group structure of $\mathbb{C}/\Omega$ to the elliptic curve $\overline{E}$. For $P, Q \in \overline{E}$ we define their sum as

$$P + Q := \Phi(\Phi^{-1}(P) + \Phi^{-1}(Q)), \tag{7.2.1}$$

where the addition on $\mathbb{C}/\Omega$ is given by $(u+\Omega) + (v+\Omega) := (u+v) + \Omega$. The following proposition is then clear.

Proposition 7.2.3. *Under the addition* (7.2.1)*, the elliptic curve $\overline{E}(\Omega)$ is an abelian group with unit element $\mathcal{O}$, and $\Phi : \mathbb{C}/\Omega \to \overline{E}(\Omega)$ is a group isomorphism. Moreover, for $z \in \mathbb{C} \setminus \Omega$ the inverse element can be computed as*

$$-(\wp(z), \wp'(z)) = (\wp(-z), \wp'(-z)) = (\wp(z), -\wp'(z)), \tag{7.2.2}$$

and for $u, v \in \mathbb{C}$ with $u, v, u+v \notin \Omega$ the addition can be computed as

$$(\wp(u), \wp'(u)) + (\wp(v), \wp'(v)) = (\wp(u+v), \wp'(u+v)). \tag{7.2.3}$$

Note that (7.2.2) tells us that the negative $-P$ of a point $P = (X, Y)$ on E is just given by

$$-P = (X, -Y).$$

However, the addition law (7.2.3) does not yet tell us how the components of $P + Q$ could be expressed in terms of the components of P and Q. Such an explicit formula will be derived in the next section, using a geometric interpretation of the addition law.

7.3 The addition law, geometrically

In this section we define an addition law on $\overline{E}$ by a geometric approach. To distuish it from the group law introduced above, we will denote the geometric addition law by $P \bullet Q$ (although we will see below that it essentially defines the same addition law). Since $\mathcal{O} = (\infty, \infty)$ should be the neutral element, we define $P \bullet \mathcal{O} = \mathcal{O} \bullet P = P$ for any $P \in \overline{E}$. Hence it remains to define $P \bullet Q$ for $P, Q \in E$.

We have to distinguish three cases concerning the position of P and Q on E. Thoughout we will write $P = (X_P, Y_P)$ for points $P \in \mathbb{C} \times \mathbb{C}$.

7.3.1 $P, Q \in E$ with $X_P \neq X_Q$

For $P, Q \in E$ with $X_P \neq X_Q$ we consider the complex line $\Gamma = \Gamma_{P,Q}$ through P and Q. It is given by the equation

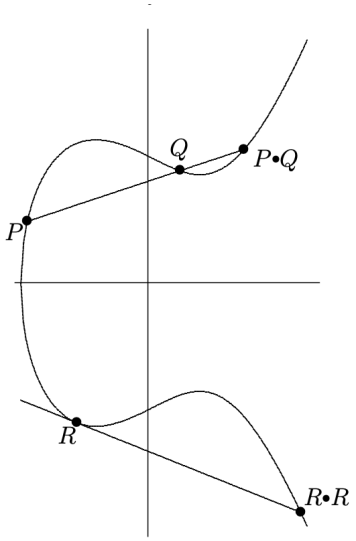
$$Y = a_{P,Q}X + b_{P,Q}$$

where

$$a_{P,Q} := \frac{Y_P - Y_Q}{X_P - X_Q}, \tag{7.3.1}$$

$$b_{P,Q} := Y_P - a_{P,Q}X_P = \frac{X_P Y_Q - X_Q Y_P}{X_P - X_Q}. \tag{7.3.2}$$

If we look at the real image of E in some examples (if g_2, g_3 are real) we see that the line Γ typically intersects E in a third point, which we denote by $P \bullet Q$.



To make this idea rigorous, we define $P \bullet Q$ as the point with coordinates

$$\begin{aligned} X_{P\bullet Q} &:= \frac{1}{4}a_{P,Q}^2 - X_P - X_Q, \\ Y_{P\bullet Q} &:= a_{P,Q}X_{P\bullet Q} + b_{P,Q}, \end{aligned} \tag{7.3.3}$$

and show that it is indeed the third intersection point of the line Γ with the elliptic curve E.

It is clear from the definition that $P \bullet Q$ lies on the line Γ.

Lemma 7.3.1. *For $X \in \mathbb{C}$ we have*

$$4X^3 - g_2X - g_3 = 4(X - X_P)(X - X_Q)(X - X_{P\bullet Q}) + (a_{P,Q}X + b_{P,Q})^2. \tag{7.3.4}$$

Proof. It is clear that the coefficients at X^3 on both sides of (7.3.4) agree, and it follows from (7.3.3) that the coefficients at X^2 agree, as well. Hence the difference of the two sides in (7.3.4) is linear in X. But this difference vanishes at the two distinct points $X = X_P$ and $X = X_Q$ since $P, Q \in \Gamma \cap E$, so it vanishes identically. □

Corollary 7.3.2. *For $P, Q \in E$ with $X_P \neq X_Q$ we have $P \bullet Q \in E$.*

In particular, the point $P \bullet Q$ is the third intersection point of the line Γ with E. The formulas in (7.3.3) are also called *intersection formulas*.

7.3.2 $P \neq Q$ with $X_P = X_Q$

For $P \neq Q$ with $X_P = X_Q$ the defining equation for E implies that we have $Y_P^2 = Y_Q^2$, that is, $Y_P = \pm Y_Q$. Since $P \neq Q$, we must have $Y_P = -Y_Q$. In the the real image of E (if g_2, g_3 are real), the line through P and Q is a vertical line, which 'intersects E at ∞'. Motivated by this, we define

$$P \bullet Q := \mathcal{O}.$$

7.3.3 $P = Q$

The idea is similar as in the case that $X_P \neq X_Q$, but now we consider the tangent line at P, and define $P \bullet P$ as the other intersection point of this tangent line with E.

If $Y_P = 0$, then the tangent line at P in the real image of E will be a vertical line. Hence we define

$$P \bullet P = \mathcal{O}$$

in this case.

If $Y_P \neq 0$, then the complex tangent line $\Gamma = \Gamma_P$ at P is given by

$$Y = a_P X + b_P$$

where

$$a_P = \frac{12X_P^2 - g_2}{2Y_P}, \qquad b_P = Y_P - a_P X_P.$$

We define the point $P \bullet P$ by

$$X_{P\bullet P} := \frac{1}{4}a_P^2 - 2X_P,$$
$$Y_{P\bullet P} := a_P X_{P\bullet P} + b_P.$$

Again, it is clear that $P \bullet P$ lies on the tangent line Γ. Similarly as above, one proves the following results:

Lemma 7.3.3. *For $X \in \mathbb{C}$ we have*

$$4X^3 - g_2 X - g_3 = 4(X - X_P)^2(X - X_{P\bullet P}) + (a_P X + b_P)^2$$

Corollary 7.3.4. *For $P \in E$ with $Y_P \neq 0$ we have $P \bullet P \in E$.*

7.3.4 Comparison of the addition laws

Recall that we defined

$$P \bullet \mathcal{O} = \mathcal{O} \bullet P = P$$

for $P \in \overline{E}$. Moreover, for $P \neq Q \in E$ we defined

$$P \bullet Q = \begin{cases} \text{third intersection point of the line through } P, Q \text{ with } E, & \text{if } X_P \neq X_Q, \\ \mathcal{O}, & \text{if } X_P = X_Q, \end{cases}$$

and for $P \in E$ we defined

$$P \bullet P = \begin{cases} \text{second intersection point of the tangent line through } P \text{ with } E, & \text{if } Y_P \neq 0, \\ \mathcal{O}, & \text{if } Y_P = 0, \end{cases}$$

Remark 7.3.5. The geometric addition law $P \bullet Q$ defined above is not associative! In particular, it does not give a group structure on $\overline{E}$.

Lemma 7.3.6. *Let $u, v, w \in \mathbb{C} \setminus \Omega$ such that $u + v + w \in \Omega$, and such that $w + \Omega$ is not one of the points $u + \Omega, v + \Omega$. Then the corresponding points on E satisfy*

$$\Phi(u) \bullet \Phi(v) = \Phi(w).$$

Proof. Recall that $\Phi(z) = (\wp(z), \wp'(z))$. For brevity, we put

$$P = \Phi(u), \qquad Q = \Phi(v), \qquad R = \Phi(w).$$

Let us first assume that $u + \Omega \neq v + \Omega$, such that $P \neq Q$. Then we have $X_P \neq X_Q$ since otherwise $X_P = \wp(u) = \wp(v) = X_Q$ would imply that $u + v \in \Omega$ (by Theorem 3.2.4), contradicting our assumption that $u + v + w \in \Omega$ but $u, v, w \notin \Omega$. Hence we are in the case of Section 7.3.1 of the geometric addition law.

In order to show that $P \bullet Q = R$, we have to check that R is the third intersection point of the line $\Gamma_{P,Q}$ through P and Q. Consider the elliptic function

$$f(z) := \wp'(z) - (a_{P,Q}\wp(z) + b_{P,Q}).$$

It has a third order pole in 0 and no other poles in $\mathbb{C}/\Omega$, so ist must have three roots in $\mathbb{C}/\Omega$. We have $f(u) = f(v) = 0$ since $P = \Phi(u)$ and $Q = \Phi(v)$ lie on the line $\Gamma_{P,Q}$ which is defined by $Y = a_{P,Q}X + b_{P,Q}$. Since $f(z)$ has precisely three roots in $\mathbb{C}/\Omega$ whose sum is in Ω by Theorem 3.2.4, we must have $f(w) = 0$. Since P, Q, R are pairwise different, this means that $R = \Phi(w)$ is the third intersection point of E and the line $\Gamma_{P,Q}$, so we have $P \bullet Q = R$.

It remains to consider the case that $u + \Omega = v + \Omega$, that is, $P = Q$, which is very similar. Now the assumption that $u + v + w = 2u + w \in \Omega$ and $u, w \notin \Omega$ implies that $2u \notin \Omega$, i.e. $\wp'(u) \neq 0$, which means that we are in the case of Section 7.3.3 with $Y_P \neq 0$. Hence, we consider the elliptic function

$$f(z) = \wp'(z) - (a_P\wp(z) + b_P).$$

Again, it has a third order pole and three roots in $\mathbb{C}/\Omega$. We have $f(u) = 0$ since $P = \Phi(u)$ lies on the tangent through P. Moreover, using the definition of a_P and the (derivative of) the differential equation $\wp'^2 = 4\wp^3 - \wp g_2 - g_3$ one can check that $f(z)$ has a double root at $z = u$. The third root mod Ω must be w by Theorem 3.2.4, which implies that $\Phi(w)$ is the third intersection point of E and the tangent through P, hence $P \bullet P = R$ as claimed. □

Remark 7.3.7. Lemma 7.3.6 can also be proved using the addition and duplication laws for the $\wp$-function, Theorem 7.1.1 and Corollary 7.1.3. Conversely, we can use Lemma 7.3.6 to obtain a new proof of the addition theorem, Theorem 7.1.1. Indeed, assume first that $u, v, w \in \mathbb{C} \setminus \Omega$ satisfy $u + v + w = 0$, and $u + \Omega, v + \Omega, w + \Omega$ are pairwise different. Then Lemma 7.3.6, together with (7.3.1) and (7.3.3), implies that

$$\begin{aligned}\wp(u + v) = \wp(-w) = \wp(w) &= \frac{1}{4}a_{P,Q}^2 - X_P - X_Q \\ &= \frac{1}{4}\left(\frac{\wp'(u) - \wp'(v)}{\wp(u) - \wp(v)}\right)^2 - \wp(u) - \wp(v).\end{aligned}$$

For general $u, v \in \mathbb{C} \setminus \Omega$ the addition theorem follows by continuity.

For $P = (X_P, Y_P) \in \mathbb{C} \times \mathbb{C}$ we put

$$P^* = (X_P, -Y_P).$$

We can now show that the addition law $P+Q$ on E defined via Φ and the geometric addition law $P \bullet Q$ on E (essentially) agree.

Proposition 7.3.8. *The addition $(P, Q) \mapsto P + Q$ on E defined via the bijection $\Phi : (\mathbb{C} \setminus \Omega)/\Omega \to E, \Phi(z + \Omega) = (\wp(z), \wp'(z))$, is given by*

$$P + Q = (P \bullet Q)^*, \qquad \text{if } X_P \neq X_Q,$$

and

$$2P = (P \bullet P)^*, \quad \text{if } Y_P \neq 0.$$

In particular, we have the formulas

$$X_{P+Q} = \frac{1}{4}a_{P,Q}^2 - X_P - X_Q, \qquad Y_{P+Q} = -a_{P,Q}X_{P+Q} - b_{P,Q}, \qquad \text{if } X_P \neq X_Q,$$

and

$$X_{2P} = \frac{1}{4}a_P^2 - 2X_P, \quad Y_{2P} = -a_P X_{2P} - b_P, \qquad \text{if } Y_P \neq 0.$$

Moreover, we have

$$-P = P^* = (X_P, -Y_P). \tag{7.3.5}$$

Proof. Let $P = \Phi(u), Q = \Phi(v)$, and put $w = -u - v$. Then by Lemma 7.3.6 we have

$$\begin{aligned} P + Q &= \Phi(\Phi^{-1}(P) + \Phi^{-1}(Q)) = \Phi(u + v) = \Phi(-w) \\ &= (\wp(-w), \wp'(-w)) = (\wp(w), -\wp'(w)) = (\Phi(w))^* = (P \bullet Q)^*. \end{aligned}$$

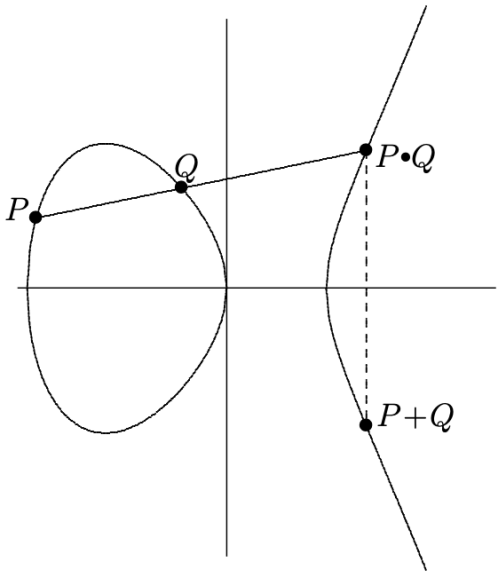
The explicit formulas for the addition law defined via Φ now follow from the explicit formulas for the geometric addition law. □

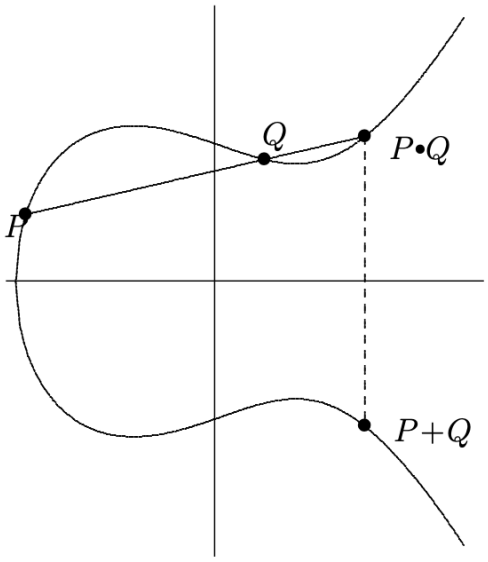
Remark 7.3.9. If we put $\mathcal{O}^* = \mathcal{O}$ then it follows from the above proposition that we have

$$P + Q = (P \bullet Q)^*$$

for all $P, Q \in \overline{E}$. Indeed, the remaining cases can be checked directly: if $P \neq Q$ with $X_P = X_Q$ then we have $Y_P = -Y_Q$, and $P+Q = \mathcal{O} = (P \bullet Q)^*$, where the first identity is (7.3.5), and the second identity holds by definition. Similarly, if $P = Q$ with $Y_P = 0$, then $2P = \mathcal{O} = (P \bullet P)^*$ by (7.3.5) and definition, respectively.

The addition law (in the real picture) geometrically looks as follows.


P
Q
P•Q
P+Q


P
Q
P•Q
P+Q

8 Rational points on elliptic curves

In this chapter we follow the book *Rational points on elliptic curves* by Silverman and Tate.

8.1 Mordell's Theorem

In Section 7.2 we defined an elliptic curve over $\mathbb{C}$ to be the set of solutions of the equation

$$E : Y^2 = 4X^3 - g_2X - g_3,$$

where $g_2 = g_2(\Omega), g_3 = g_3(\Omega)$ are the Weierstrass invariants of a lattice Ω in $\mathbb{C}$, which satisfy $g_2^3 - 27g_3^2 \neq 0$. In order to simplify the notation, we would like to forget for now that the coefficients g_2, g_3 come from a lattice. By Corollary 5.4.7 for every $c_2, c_3 \in \mathbb{C}$ with $c_2^3 - 27c_3^2 \neq 0$ there exists a lattice Ω with Weierstrass invariants $c_2 = g_2(\Omega)$ and $c_3 = g_3(\Omega)$. Hence, after replacing Y by $2Y$ (which will be inessential for our application) and setting $a = -g_2/4$ and $b = -g_3/4$, every elliptic curve has the form

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

where $a, b \in \mathbb{C}$ satisfy $4a^3 + 27b^2 \neq 0$.

Definition 8.1.1. A *rational elliptic curve* is given by an equation of the form

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

with $a, b \in \mathbb{Q}$ satisfying $4a^3 + 27b^2 \neq 0$.

To simplify the notation, we will view the point at infinity $\mathcal{O}$ as a point on E, and no longer distinguish between E and $\overline{E} = E \cup \{\infty\}$.

Remark 8.1.2. The condition $4a^3 + 27b^2 \neq 0$ is equivalent to saying that the polynomial $f(X) = X^3 + aX + b$ does not have multiple roots.

Remark 8.1.3. Replacing Y by $2Y$ slighly changes the explicit formula for the addition law in Proposition 7.3.8. For example, for the X-coordinate of $P + Q$ we have

$$X_{P+Q} = \left(\frac{Y_P - Y_Q}{X_P - X_Q}\right)^2 - X_P - X_Q, \quad \text{if } X_P \neq X_Q,$$

and

$$X_{2P} = \left(\frac{3X_P^2 + a}{2Y_P}\right)^2 - 2X_P, \qquad \text{if } Y_P \neq 0.$$

For the rest of this chapter, E will denote a rational elliptic curve. We may ask for its *rational points*

$$E(\mathbb{Q}) = \{(X, Y) \in \mathbb{Q}^2 : Y^2 = X^3 + aX + b\} \cup \{\mathcal{O}\}.$$

Note that we view the point at infinity $\mathcal{O}$ as a rational point by definition. Using the explicit formulas for the group law on E given in Remark 8.1.3, we see that $E(\mathbb{Q})$ is a subgroup of E. The fundamental result about $E(\mathbb{Q})$ we want to prove is Mordell's Theorem:

Theorem 8.1.4 (Mordell)**.** *The group $E(\mathbb{Q})$ is finitely generated.*

Explicitly, this means that there exist points $P_1, \ldots, P_k \in E(\mathbb{Q})$ such that every point $P \in E(\mathbb{Q})$ can be written as $P = n_1P_1 + \cdots + n_kP_k$ for some integers $n_1, \ldots, n_k \in \mathbb{Z}$.

By the structure theorem for finitely generated abelian groups, we have an isomorphism

$$E(\mathbb{Q}) \cong \mathbb{Z}^r \times E(\mathbb{Q})_{\text{tors}}$$

for some integer $r \geq 0$, called the *rank* of E, and a finite group $E(\mathbb{Q})_{\text{tors}}$, which is called the *torsion subgroup* of $E(\mathbb{Q})$ and consists precisely of the elements of finite order in $E(\mathbb{Q})$.

In order to prove Mordell's Theorem, we will now study heights on rational elliptic curves and use a descent argument.

8.2 The Descent Theorem

We will use the following criterion in order to prove Mordell's Theorem.

Theorem 8.2.1 (Descent Theorem)**.** *Let Γ be an abelian group, and suppose that there is a function*

$$h : \Gamma \to [0, \infty)$$

with the following four properties:

(a) For every real number C, the set $\{P \in \Gamma : h(P) \leq C\}$ is finite.

(b) For every $P_0 \in \Gamma$ there is a constant κ_0 such that

$$h(P + P_0) \leq 2h(P) + \kappa_0 \quad \text{for all } P \in \Gamma.$$

(c) There is a constant κ such that

$$h(2P) \geq 4h(P) - \kappa \quad \text{for all } P \in \Gamma.$$

(d) The subgroup 2Γ has finite index in Γ.

Then Γ is finitely generated.

Proof. Since 2Γ has finite index in Γ by assumption (d), we can choose finitely many coset representatives $Q_1, \ldots, Q_n$ for 2Γ in Γ. This means that for any $P \in \Gamma$ there is an index i_1 such that

$$P - Q_{i_1} \in 2\Gamma,$$

so we can write

$$P - Q_{i_1} = 2P_1$$

for some $P_1 \in \Gamma$. Now we do the same thing with P_1. Continuing like this, we obtain

$$\begin{aligned} P - Q_{i_1} &= 2P_1, \\ P_1 - Q_{i_2} &= 2P_2, \\ P_2 - Q_{i_3} &= 2P_3, \\ &\vdots \\ P_{m-1} - Q_{i_m} &= 2P_m, \end{aligned} \tag{8.2.1}$$

where $Q_{i_1}, \dots, Q_{i_m}$ are chosen in a suitable way from the coset representatives $Q_1, \dots, Q_n$ and where $P_1, \dots, P_m$ are suitable elements in Γ. By substituting the equations (8.2.1) into each other recusively, we obtain

$$P = Q_{i_1} + 2P_1,$$

and then

$$P = Q_{i_1} + 2Q_{i_2} + 4P_2,$$

and, after some more steps,

$$P = Q_{i_1} + 2Q_{i_2} + 4Q_{i_3} + \cdots + 2^{m-1}Q_{i_m} + 2^m P_m.$$

In particular, this says that P is in the subgroup of Γ generated by the Q_i's and P_m. We are going to show that by choosing m large enough, we can force P_m to have height less than a certain fixed bound that does not depend on the initial point P. Then the finite set of points with height less than this bound, together with the Q_i's, will generate Γ.

The basic idea is that since P_i is more-or-less equal to $2P_{i+1}$, the height of P_{i+1} should be more-or-less one-fourth the height of P_i. So the sequence of points $P, P_1, P_2, \dots$ should have decreasing height, and eventually the point P_m should end up in a set of points of bounded height (this proof strategy is where the name *descent theorem* comes from). Now we turn this idea into a valid proof.

From assumption (b) applied to the points $Q_1, \dots, Q_n$ we get constants $\kappa_1, \dots, \kappa_n$ such that

$$h(P - Q_i) \le 2h(P) + \kappa_i \quad \text{for all } P \in \Gamma,$$

for each $1 \le i \le n$. Let

$$\kappa' = \max\{\kappa_1, \dots, \kappa_n\}.$$

Then

$$h(P - Q_i) \le 2h(P) + \kappa' \quad \text{for all } P \in \Gamma \text{ and all } 1 \le i \le n. \tag{8.2.2}$$

Here we used again the assumption (d), or in other words, that there are only finitely many coset representatives $Q_1, \dots, Q_n$ for 2Γ in Γ.

Let κ be the constant from item (c). Moreover, let P_j one of the elements in the sequence $P, P_1, P_2, \dots$. We want to show that $h(P_j)$ is considerably smaller than $h(P_{j-1})$. We calculate

$$\begin{aligned} 4h(P_j) &\le h(2P_j) + \kappa \\ &= h(P_{j-1} - Q_{i_j}) + \kappa \\ &\le 2h(P_{j-1}) + \kappa' + \kappa, \end{aligned}$$

where we used (8.2.2) in the last step. We write this as

$$\begin{aligned} h(P_j) &\leq \frac{1}{2}h(P_{j-1}) + \frac{\kappa' + \kappa}{4} \\ &= \frac{3}{4}h(P_{j-1}) - \frac{1}{4}\Big(h(P_j) - (\kappa' + \kappa)\Big). \end{aligned}$$

From this we see that if $h(P_{j-1}) \geq \kappa' + \kappa$, then

$$h(P_j) \leq \frac{3}{4}h(P_{j-1}).$$

Hence, in the sequence of points $P, P_1, P_2, \ldots$, as long as the point P_{j-1} satisfies $h(P_{j-1}) \geq \kappa' + \kappa$, then the next point in the sequence has much smaller height, namely $h(P_j) \leq \frac{3}{4}h(P_{j-1})$. But if you start with any number and keep multiplying it by $\frac{3}{4}$, it approaches zero. So eventually we will find an index m such that $h(P_m) \leq \kappa' + \kappa$.

In conclusion, we have shown that every element $P \in \Gamma$ can be written in the form

$$P = a_1Q_1 + a_2Q_2 + \cdots + a_nQ_n + 2^mR$$

for certain integers $a_1, \ldots, a_n$ and $m \in \mathbb{N}$, and some $R \in \Gamma$ with bounded height $h(R) \leq \kappa' + \kappa$. Hence the set

$$\{Q_1, Q_2, \ldots, Q_n\} \cup \{R \in \Gamma : h(R) \leq \kappa' + \kappa\}$$

generates Γ. From (a) and (d) this set is finite, which finishes the proof. □

We will apply the Descent Theorem with $\Gamma = E(\mathbb{Q})$, and the logarithmic height h on $E(\mathbb{Q})$ defined in the next section. To do that, has to prove that the logarithmic height has the four properties stated in the Descent Theorem. Here we will prove the properties (a),(b), and (c). The assection (d) that $2E(\mathbb{Q})$ has finite index in $E(\mathbb{Q})$ is much more difficult to prove, and we refer the reader to Section 3.4 in the book by Silverman and Tate .

8.3 Heights

Definition 8.3.1. Let $x = \frac{m}{n}$ be a rational number written in lowest terms. The *height* $H(x)$ of x is the minimum of the absolute values of the numerator and denominator of x,

$$H(x) = H\left(\frac{m}{n}\right) = \max\{|m|, |n|\} \in \mathbb{N}_0.$$

The height is a measure for the complexity of a rational number. For example, although 1 and $\frac{999}{1000}$ are close to each other in absolute value, we have $H(1) = 1$ and $H(\frac{999}{1000}) = 1000$.

Proposition 8.3.2. *The height on $\mathbb{Q}$ has the* finiteness property*: The set of all rational numbers whose height is less than some fixed number is a finite set.*

Proof. If $H(\frac{m}{n}) \leq C$ then we have $|m|, |n| \leq C$, which leaves only finitely many possibilities for $m, n \in \mathbb{Z}$. □

Definition 8.3.3. For a rational elliptic curve E we define the *height* of a point $\mathcal{O} \neq P = (X, Y) \in E(\mathbb{Q})$ by

$$H(P) = H(X) \in \mathbb{N}_0,$$

and we set $H(\mathcal{O}) = 1$. Moreover, we define the *logarithmic height* by

$$h(P) = \log H(P) \in \mathbb{R}_{\geq 0}.$$

Note that it makes sense to measure the complexity of a point P on $E(\mathbb{Q})$ only in terms of its X-coordinate, since its Y coordinate is determined by X up to a sign.

We now want to show that the logarithmic height $h(P)$ has the four properties required in the Descent Theorem 8.2.1. The finiteness property (a) is easy to prove:

Lemma 8.3.4. *For every real number C, the set $\{P \in E(\mathbb{Q}) : h(P) \leq M\}$ is finite.*

Proof. If $h(P) \leq C$ then $H(P) \leq e^C$. Since $H(P) = H(X)$ for $P = (X, Y)$, this leaves only finitely many possibilities for X, and since $Y^2 = X^3 + aX + b$, we have at most two possibilities for Y for every possible X. □

The items (b) and (c) in the Descent Theorem are more complicated, and will be discussed in the following two subsections. The proof of item (d), the finiteness of $2E(\mathbb{Q})$ in $E(\mathbb{Q})$, is considerably harder and will be discussed in a separate section afterwards.

8.3.1 The height of $P + P_0$

In this section we prove the following lemma, which asserts that the logarithmic height $h(P)$ satisfies assumption (b) of the Descent Theorem 8.2.1.

Lemma 8.3.5. *Let $P_0 \in E(\mathbb{Q})$ be fixed. There is a constant κ_0 (depending on P_0 and a, b), such that*

$$h(P + P_0) \leq 2h(P) + \kappa_0 \quad \text{for all } P \in E(\mathbb{Q}).$$

For the proof of the lemma, we will use the following observations:

Lemma 8.3.6. *Let $P = (X, Y) \in E(\mathbb{Q})$ with $P \neq \mathcal{O}$.*

1. *We can write the coordinates of P in lowest terms as*

$$X = \frac{m}{e^2}, \qquad Y = \frac{n}{e^3},$$

   with integers m, n, and e, with $e > 0$ and $\gcd(m, e) = \gcd(n, e) = 1$.

2. *There is a constant $K > 0$ (depending on a, b) such that*

$$|m| \leq H(P), \quad e^2 \leq H(P), \quad |n| \leq KH(P)^{3/2} \quad \text{for all } P = \left(\frac{m}{e^2}, \frac{n}{e^3}\right).$$

Proof. We leave this as an exercise to the reader. □

Proof of Lemma 8.3.5. For the proof, we will explicitly write out the formula for the sum of two points, compare Corollary 8.1.3.

First note that the lemma is trivial for $P_0 = \mathcal{O}$, so we may assume that $P_0 \neq \mathcal{O}$. Moreover, we note that in proving the existence of κ_0, it is enough to prove that the inequality holds for all P execept those a finite set. Hence it suffices to prove Lemma 8.3.5 for all $P \notin \{P_0, -P_0, \mathcal{O}\}$, which implies that $X \neq X_0$. This assumption saves us from doing a case distinction when applying the explicit formulas for the addition law.

Let $P = (X, Y)$ and write

$$P + P_0 = (\xi, \eta).$$

Then we have $h(P + P_0) = \log H(\xi)$, so we need a formula for ξ in terms of (X, Y) and (X_0, Y_0). Since $X \neq X_0$ by assumption, the formula from Corollary 8.1.3 says

$$\begin{aligned}\xi &= \frac{(Y - Y_0)^2}{(X - X_0)^2} - X - X_0 \\ &= \frac{(Y - Y_0)^2 - (X - X_0)^2(X + X_0)}{(X - X_0)^2}.\end{aligned}$$

If we multiply this out, we find that $Y^2 - X^3$ appears in the numerator. Since P is on the curve, we may replace $Y^2 - X^3$ with $aX + b$, so we end up with an expression

$$\xi = \frac{AY + BX^2 + CX + D}{EX^2 + FX + G}$$

where A, B, C, D, E, F, G are certain rational numbers that can be expressed in terms of a, b, and (X_0, Y_0). Further, by multiplying the numerator and the denominator by the least common denominator of $A, B, \ldots, G$, we may assume that $A, B, \ldots, G$ are all integers.

The important fact is that once the curve and the point P_0 are fixed, this expression is correct for all points P. So it will be allright if our constant κ_0 depends on $A, B, \ldots, G$, as long as it does not depend on X, Y.

Now substitute $X = m/e^2$ and $Y = n/e^3$ as in Lemma 8.3.6 above, and clear denominators by multiplying numerator and denominator by e^4. We find that

$$\xi = \frac{Ane + Bm^2 + Cme^2 + De^4}{Em^2 + Fme^2 + Ge^4}.$$

Notice that we now have an expression for ξ that is an integer divided by an integer. It need not be in lowest terms, but cancellation will only make the height smaller. Thus

$$H(\xi) \leq \max\{|Ane + Bm^2 + Cme^2 + De^4|, |Em^2 + Fme^2 + Ge^4|\}$$

In Lemma 8.3.6 we have noted that

$$e \leq H(P)^{1/2}, \quad n \leq KH(P)^{3/2}, \quad \text{and} \quad m \leq H(P),$$

where the constant K only depends on a, b, but not on P. Using these and the triangle inequality gives

$$\begin{aligned}|Ane + Bm^2 + Cme^2 + De^4| &\leq |Ane| + |Bm^2| + |Cme^2| + |De^4| \\ &\leq (|AK| + |B| + |C| + |D|)H(P)^2\end{aligned}$$

and

$$\begin{aligned}|Em^2 + Fme^2 + Ge^4| &\leq |Em^2| + |Fme^2| + |Ge^4| \\ &\leq (|E| + |F| + |G|)H(P)^2.\end{aligned}$$

Therefore

$$H(P + P_0) = H(\xi) \leq \max\{|AK| + |B| + |C| + |D|, |E| + |F| + |G|\}H(P)^2.$$

Taking logarithms of both sides gives

$$h(P + P_0) \leq 2h(P) + \kappa_0,$$

where the constant

$$\kappa_0 = \log \max\{|AK| + |B| + |C| + |D|, |E| + |F| + |G|\}$$

depends only on a, b and (X_0, Y_0), but not on $P = (X, Y)$. This finishes the proof. □

8.3.2 The height of $2P$

Lemma 8.3.7. *There is a constant κ (depending on a, b) such that*

$$h(2P) \geq 4h(P) - \kappa \quad \textit{for all } P \in E(\mathbb{Q}).$$

Proof. Again, it doesn't matter if we discard finitely many points P, and now it will be convenient to discard the finitely many points satisfying $2P = \mathcal{O}$ in order to avoid case distinctions when applying the explicit formulas for the addition law on $E(\mathbb{Q})$.

Let $P = (X, Y)$, write $2P = (\xi, \eta)$, and put

$$f(X) = X^3 + aX + b$$

for brevity. The formulas from Corollary 8.1.3 state that

$$\xi = \alpha^2 - 2X, \quad \text{where} \quad \alpha = \frac{f'(X)}{2Y}$$

Putting everything over a common denominator and using $Y^2 = X^3 + aX + b$ we find

$$\xi = \frac{f'(X)^2 - 8Xf(X)}{4f(X)} = \frac{X^4 + \ldots}{4X^3 + \ldots}$$

Note that the denominator is nonvanishing since $2P \neq \mathcal{O}$.

Thus $\xi = \frac{\phi(X)}{\psi(X)}$ is the a quotient of two polynomials $\phi(X), \psi(X)$ with integer coefficients. Since $f(X)$ does not have multiple roots, $f(X)$ and $f'(X)$ have no common roots, so the polynomials in the numerator and denominator of ξ have no common roots.

Since $h(P) = h(X)$ and $h(2P) = h(\xi) = h(\phi(X)/\psi(X))$, we want to prove that

$$h\left(\frac{\phi(X)}{\psi(X)}\right) \geq 4h(X) - \kappa,$$

for some constant κ (depending on a, b, or in other words, on the polynomials ϕ, ψ). This will follow from the next general lemma about heights (on $\mathbb{Q}$) and quotients of polynomials. □

Lemma 8.3.8. *Let ϕ and ψ be polynomials with integer coefficients a no common roots. Let d be the maximum of the degrees of ϕ and ψ.*

(a) There is an integer $R \geq 1$ (depending on ϕ and ψ), such that for all rational numbers m/n (written in lowest terms) we have

$$\gcd\left(n^d\phi\left(\frac{m}{n}\right), n^d\psi\left(\frac{m}{n}\right)\right) \quad \text{divides } R.$$

(b) There are constants κ_1 and κ_2 (depending on ϕ and ψ), such that for all rational numbers m/n that are not roots of ψ we have

$$dh\left(\frac{m}{n}\right) - \kappa_1 \leq h\left(\frac{\phi(m/n)}{\psi(m/n)}\right) \leq dh\left(\frac{m}{n}\right) + \kappa_2.$$

Proof. (a) For $d = 0$ the statement is trivial, so we will assume $d \geq 1$. Further, without loss of generality we may assume that $\deg(\phi) = d$ and $\deg(\psi) = e \leq d$. We write

$$\Phi(m,n) := n^d\phi\left(\frac{m}{n}\right) = a_0m^d + a_1m^{d-1}n + \cdots + a_{d-1}mn^{d-1} + a_dn^d,$$
$$\Psi(m,n) := n^d\psi\left(\frac{m}{n}\right) = b_0m^en^{d-e} + b_1m^{e-1}n^{d-e+1} + \cdots + b_en^d.$$

and we put

$$\gamma := \gamma(m,n) = \gcd(\Phi(m,n), \Psi(m,n)).$$

We want to find a bound for γ which is independent of m, n.

Since $\phi(X)$ and $\psi(X)$ have no common roots, they are relatively prime in the Euclidean ring $\mathbb{Q}[X]$, and we can find polynomials $F(X)$ and $G(X)$ with rational coefficients such that

$$F(X)\phi(X) + G(X)\psi(X) = 1. \tag{8.3.1}$$

Let $A \in \mathbb{Z}$ be such that $AF(X)$ and $AG(X)$ have integer coefficients, and let D be the maximum of the degrees of F and G. Note that A and D do not depend on m, n.

Now we evaluate (8.3.1) at $X = m/n$ and multiply by An^{D+d}:

$$\underbrace{n^DAF\left(\frac{m}{n}\right)}_{\in\mathbb{Z}} \cdot \underbrace{n^d\phi\left(\frac{m}{n}\right)}_{=\Phi(m,n)} + \underbrace{n^DAG\left(\frac{m}{n}\right)}_{\in\mathbb{Z}} \cdot \underbrace{n^d\psi\left(\frac{m}{n}\right)}_{=\Psi(m,n)} = An^{D+d},$$

From this we see that $\gamma = \gcd(\Phi(m,n), \Psi(m,n))$ divides An^{D+d}. This is not yet good enough, since we need to show that γ divides some fixed integer which does not depend on n. We will show that γ divides Aa_0^{D+d}, where a_0 is the leading coefficient of $\phi(X)$.

Note that γ divides $\Phi(m,n)$, so it also divides

$$An^{D+d-1}\Phi(m,n) = Aa_0m^dn^{D+d-1} + Aa_1m^{d-1}n^{D+d} + \cdots + Aa_dn^{D+2d-1}.$$

On the right-hand side, every summand after the first one contains An^{D+d}, which divisible by γ as we just showed above. Hence γ divides the first term $Aa_0m^dn^{D+d-1}$. Thus

$$\gamma \quad \text{divides} \quad \gcd(An^{D+d}, Aa_0m^dn^{D+d-1}).$$

Since m and n are coprime, we find that γ divides Aa_0n^{D+d-1}. If $D+d-1=0$, then we are done. If not, then n^{D+d-2} is an integer, and we use that γ divides $Aa_0n^{D+d-2}\Phi(m,n)$ to repeat the above argument and find that γ divides $Aa_0^2n^{D+d-2}$. We repeat this procedure until we find that γ divides Aa_0^{D+d}, which finishes the proof of (a).

(b) We will only prove the lower bound, which is the one we need to complete the proof of Lemma 8.3.7. Again, we may discard finitely many rational numbers, so we will assume that m/n is not a root of ϕ. Moreover, since $h(r)=h(1/r)$ for any non-zero rational number, we may interchange ϕ and ψ if necessary, and again assume that $\deg(\phi)=d$ and $\deg(\psi)=e\leq d$.

We want to estimate the height of

$$\xi=\frac{\phi\left(\frac{m}{n}\right)}{\psi\left(\frac{m}{n}\right)}=\frac{n^d\phi\left(\frac{m}{n}\right)}{n^d\psi\left(\frac{m}{n}\right)}=\frac{\Phi(m,n)}{\Psi(m,n)}.$$

Since $\Phi(m,n)$ and $\Psi(m,n)$ are integers whose greates common divisor γ is bounded by some R (independently of m,n) by part (a) we have

$$\begin{aligned}H(\xi)&=\frac{1}{\gamma}\max\{|\Phi(m,n)|,|\Psi(m,n)|\}\\&\geq\frac{1}{R}\max\left\{\left|n^d\phi\left(\frac{m}{n}\right)\right|,\left|n^d\psi\left(\frac{m}{n}\right)\right|\right\}\\&\geq\frac{1}{2R}\left(\left|n^d\phi\left(\frac{m}{n}\right)\right|+\left|n^d\psi\left(\frac{m}{n}\right)\right|\right).\end{aligned}$$

In the last line we used $\max\{a,b\}\geq\frac{1}{2}(a+b)$. We need to compare $H(\xi)$ to

$$H\left(\frac{m}{n}\right)^d=\max\{|m|^d,|n|^d\},$$

so we consider the quotient

$$\begin{aligned}\frac{H(\xi)}{H(m/n^d)}&\geq\frac{1}{2R}\cdot\frac{\left|n^d\phi\left(\frac{m}{n}\right)\right|+\left|n^d\psi\left(\frac{m}{n}\right)\right|}{\max\{|m|^d,|n|^d\}}\\&=\frac{1}{2R}\cdot\frac{\left|\phi\left(\frac{m}{n}\right)\right|+\left|\psi\left(\frac{m}{n}\right)\right|}{\max\{|\frac{m}{n}|^d,1\}}.\end{aligned}$$

To estimat this, we consider the function $p(t)$ on $\mathbb{R}$ defined by

$$p(t)=\frac{|\phi(t)|+|\psi(t)|}{\max\{|t|^d,1\}}.$$

Since ϕ has degree d and ψ has degree at most d, we see that $p(t)$ has non-zero limit as $|t|\to\infty$. The limit is $|a_0|$ if ψ has degree strictly less than d, and it is $|a_0|+|b_0|$ if ψ has degree equal to d. In any case, outside of some closed interval I the function $p(t)$ is bounded away from 0, that is, there is some constant $\varepsilon>0$ such that $p(t)\geq\varepsilon$ for all $t\in\mathbb{R}\setminus I$. Inside the interval I, the function $p(t)$ is continuous and non-vanishing since by assumption ϕ and ψ do not have common zeros. Hence $p(t)$ is bounded away from 0 on all of $\mathbb{R}$, so there is a constant $C_1>0$ such that $p(t)\geq C_1$ for all $t\in\mathbb{R}$.

We prove earlier that

$$\frac{H(\xi)}{H(m/n)^d}\geq\frac{1}{2R}p\left(\frac{m}{n}\right)$$

and using $p(t) \geq C_1$ for all t gives

$$H(\xi) \geq \frac{C_1}{2R} H\left(\frac{m}{n}\right)^d.$$

The constants C_1 and R depend on ϕ and ψ, but not on m and n, so taking logarithms gives the desired inequality

$$h(\xi) \geq dh\left(\frac{m}{n}\right) - \kappa_1 \qquad \text{with } \kappa_1 = \log(2R/C_1).$$

This finishes the proof of Lemma 8.3.8, and thereby also the proof of Lemma 8.3.7. □

8.4 Outlook: Points of finite order

Let

$$E : Y^2 = X^3 + aX + b, \qquad (a, b \in \mathbb{Q}, \quad \Delta := 4a^2 + 27b^2 \neq 0)$$

be a rational elliptic curve. By replacing (X, Y) with $(X/m^2, Y/m^3)$ for a suitable integer m, we may assume that $a, b \in \mathbb{Z}$, which we will do from now on.

A point $P \in E(\mathbb{Q})$ has *finite order* (or is a *torsion point*) if there exists an integer $n \in \mathbb{Z}$ such that $nP = \mathcal{O}$. By Mordell's Theorem, the *torsion subgroup* $E(\mathbb{Q})_{\text{tors}}$ of all points of finite order is a finite subgroup of $E(\mathbb{Q})$.

Theorem 8.4.1 (Nagell-Lutz)**.** *Let $P = (X, Y) \in E(\mathbb{Q})_{\text{tors}}$ be a rational point of finite order. Then*

1. *X and Y are integers, and*
2. *either $Y = 0$, in which case P has order 2, or Y^2 divides the discriminant $\Delta = 4a^3 + 27b^2$.*

The Nagell-Lutz Theorem yields a procedure to determine all points of finite order: First, we can form a finite list of possible torsion points by taking $Y = 0$ and all integers Y such that Y^2 divides Δ, and check whether the corresponding values for X with $(X, Y) \in E(\mathbb{Q})$ are integral. For each of the points P in this list of possible torsion points, we compute $P, 2P, 3P, \ldots$ until we either arrive at $\mathcal{O}$, in which case P is indeed a torsion point, or we arrive at a point nP which does not have integral coordinates, in which case nP, and hence P, cannot be a torsion point by the Nagell-Lutz Theorem. A priori, it could happen that *all* points $P, 2P, 3P, \ldots$ have integral coordinates, even if P is not a torsion point. This could only happen if E contains infinitely many integral points, which is not possible by Siegel's Theorem.

Theorem 8.4.2 (Siegel)**.** *A rational elliptic curve E contains only finitely many integral points.*

Now the question arises what points of finite order may appear. For example, one can write down examples of elliptic curves having points of order $2, 3, 4, 5, 6, 7, 8, 9, 10$ and 12, but it is impossible to find a point of order 11 or of order larger than 12. More precisely, it turns out that there are not many that possibilies for $E(\mathbb{Q})_{\text{tors}}$:

Theorem 8.4.3 (Mazur)**.** *The torsion group $E(\mathbb{Q})_{\text{tors}}$ is isomorphic to one of the following:*

- *$\mathbb{Z}/n\mathbb{Z}$ with $1 \leq n \leq 10$ or $n = 12$.*
- *$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2n\mathbb{Z}$ with $1 \leq n \leq 4$.*

9 The Birch and Swinnerton-Dyer Conjecture

9.1 The BSD Conjecture

The algebraic rank

Let

$$E:\ Y^2 = X^3 + aX + b \qquad (a, b \in \mathbb{Q}, \quad \Delta := 4a^3 + 27b^2 \neq 0)$$

be a rational elliptic curve. By replacing (X, Y) with $(X/m^2, Y/m^3)$ for a suitable integer m, we may assume that a, b are integers, which we will do from now on.

Mordell's Theorem tells us that the group $E(\mathbb{Q})$ of rational points of a rational elliptic curve is finitely generated, so by the structure theorem for finitely generated abelian groups we have

$$E(\mathbb{Q}) \cong \mathbb{Z}^r \times E(\mathbb{Q})_{\text{tors}}$$

for some $r \in \mathbb{N}_0$, called the *algebraic rank* (or *Mordell-Weil rank*) of E, and with the finite *torsion subgroup* $E(\mathbb{Q})_{\text{tors}}$ consisting of the elements of finite order in $E(\mathbb{Q})$. The Theorem of Lutz-Nagell gives an effective way to compute $E(\mathbb{Q})_{\text{tors}}$, and the Theorem of Mazur gives a classification of all possible torsion groups. It remains to consider the rank r, but unfortunately, it is usually difficult to determine and there is currently no known algorithm that will compute the rank for any given rational elliptic curve.

It is conjectured that the 'average rank' of an elliptic curve should be $\frac{1}{2}$. Roughly speaking, this means that if you pick an elliptic curve at random there will be a 50% chance that the curve has rank 0, and a 50% chance that it has rank 1. In particular, curves of rank ≥ 2 are 'rare'.

The highest known rank is $r = 20$ (found in 2020), and an elliptic curve with rank at least 28 is known (but its precise rank is unknown)[1]. It is not known whether there exist elliptic curves of any given rank, or even whether the rank can be arbitrarily big.

The Birch and Swinnerton-Dyer Conjecture connects the algebraic rank to an analytic quantity attached to E, which we will explain next.

The analytic rank

Let p be a prime number. Since the coefficients a, b are integers, it makes sense to reduce the equation $Y^2 = X^3 + aX + b \bmod p$ and ask for solutions $(X, Y) \in \mathbb{F}_p^2$. The rough idea is that the curve $E(\mathbb{F}_p)$ should be easier to understand than $E(\mathbb{Q})$, and that it might be possible to piece together information from all the $E(\mathbb{F}_p)$ for primes p to get new information about $E(\mathbb{Q})$. For example, $E(\mathbb{Q})$ is difficult to determine in general, but for a fixed prime p, the set $E(\mathbb{F}_p)$ can be determined in finitely many steps by just checking for every $(X, Y) \in \mathbb{F}_p^2$ whether it satisfies the equation of $E \bmod p$ or not.

[1] see `https://web.math.pmf.unizg.hr/~duje/tors/rankhist.html` for the current records

However, if $\Delta = 0 \pmod{p}$ then the reduction of E mod p is not an elliptic curve anymore, and in this case we say that E has *bad reduction at p*. Otherwise we say that E has *good redution at p*. Fortunately, E has bad reduction only at the finitely many primes p dividing Δ. However, there may be different integral Weierstrass equations for the same curve E, which have different discriminants, so in order to obtain a well-defined notion of good and bad reduction, we must actually choose a *minimal* integral Weierstrass equation for E, which means that $|\Delta|$ is minimal among all integral Weierstrass equations.

It is clear that $E(\mathbb{F}_p)$ has at most $2p+1$ points (for each $X \in \mathbb{F}_p$ there are at most 2 values of Y such that $(X, Y) \in E(\mathbb{F}_p)$, and $\mathcal{O}$ always lies on $E(\mathbb{F}_p)$). In fact, the number of points in $E(\mathbb{F}_p)$ is closer to $p+1$. A theorem of Hasse states that we have the bound

$$|\#E(\mathbb{F}_p) - p - 1| \leq 2\sqrt{p}.$$

We will be interested in the quantity

$$a_p := \#E(\mathbb{F}_p) - p - 1.$$

All the local informations a_p about the numbers of points on $E(\mathbb{F}_p)$ for primes p are collected in the *Hasse-Weil L-function* of E, defined by the Euler product

$$L(E, s) := \prod_{p \nmid \Delta} (1 - a_p p^{-s} + p^{1-2s})^{-1} \cdot \prod_{p | \Delta} (1 - a_p p^{-s})^{-1},$$

for $s \in \mathbb{C}$, whenever the product converges absolutely. It follows from the Hasse bound that $L(E, s)$ converges absolutely for $\mathrm{Re}(s) > \frac{3}{2}$, and by the general theory of L-functions is is expected that $L(E, s)$ should have an analytic continuation to the entire complex plane, and should satisfy a functional equation under $s \mapsto 2 - s$. However, it turned out that this is a very difficult problem. It was proved in 1941 by Deuring for a certain class of elliptic curves (those with *complex multiplication*) by relating $L(E, s)$ to the L-functions of Hecke Grossencharacters, which are easier to understand. The analytic continuation and functional equation of $L(E, s)$ for all rational elliptic curves was proves around 1999 as a part of Andrew Wiles' proof of Fermat's Last Theorem, by relating $L(E, s)$ to L-functions of modular forms. We will briefly discuss this result below.

Now we define the *analytic rank* of E as the *order of vanishing of* $L(E, s)$ *at* $s = 1$, or in symbols

$$r_{\mathrm{an}}(E) := \mathrm{ord}_{s=1} L(E, s).$$

Note that $s = 1$ is the center of the (expected) functional equation of $L(E, s)$ under $s \mapsto 2 - s$.

The BSD conjecture

We can now state the famous Birch and Swinnerton-Dyer conjecture.

Conjecture 9.1.1 (Birch and Swinnerton-Dyer 1965)**.** *Let E be a rational elliptic curve. Then the algebraic rank and the analytic rank are equal:*

$$r(E) = r_{an}(E) = \mathrm{ord}_{s=1} L(E, s).$$

This formulation is sometimes called the *weak version*, and there is also a stronger version which predicts the precise value of $L^{(r)}(E, 1)$ in terms of certain algebraic quantities attached to E.

Birch and Swinnerton-Dyer arrived at this conjecture through extensive numerical experiments and heuristics. A particularly remarkable aspect is the fact that, at the time the conjecture was made, it was not even clear that $L(E,s)$ has an analytic continuation to $s=1$. Moreover, in the conjectural explicit formula for $L^{(r)}(E,1)$ the order of the *Tate-Shafarevich group* $Ш(E/\mathbb{Q})$ appears, which is not known to be finite in general!

Although the conjecture is still largely open, there has been some progress for curves of rank 0 and 1. For example, it follows from celebrated works of Gross and Zagier, and Kolyvagin from 1989 that the BSD conjecture is true in the case that the elliptic curve has algebraic rank 0 or 1. In a recent breakthrough, Bhargavar and Shankar showed that a positive proportion of elliptic curves have rank 0 and hence satisfy the BSD conjecture.

9.2 Fermat's Last Theorem and the Taniyama-Shimura Conjecture

Around 1637 Fermat conjectured his famous *Last Theorem*, saying that for $n \geq 3$ the equation

$$a^n + b^n = c^n$$

has no positive integer solutions a, b, c. Although many cases of the conjecture (for special values of n) had been proved, a complete proof was found only in 1995 by Andrew Wiles and Richard Taylor, building on work of Serre, Ribet, and many others. The proof heavily uses the theory of elliptic curves and modular forms. In particular, a key element in the proof of Fermat's Last Theorem is the *Modularity Theorem*, previously the *Taniyama-Shimura Conjecture*. To state it, we recall that every modular form f of weight k has a *Fourier expansion* of the shape $f(\tau) = \sum_{n=0}^{\infty} a_n e^{2\pi i n\tau}$ with *Fourier coefficients* $a_n \in \mathbb{C}$. The L-function of f is now defined as

$$L(f,s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s},$$

and it is relatively easy to show that $L(f,s)$ converges absolutely for $\mathrm{Re}(s) > k$, has a meromorphic continuation to the entire complex plane, and satisfies a functional equation under $s \mapsto k-s$.

Theorem 9.2.1 (Modularity Theorem; Wiles, Taylor,... 1995)**.** *Every rational elliptic curve E is modular, which means that there exists a (unique) modular form (of weight 2 for a certain subgroup of $\mathrm{SL}_2(\mathbb{Z})$) such that $L(E,s) = L(f,s)$.*

As an immediate corollary, one obtains the meromorphic continuation and the functional equation for $L(E,s)$, which was impossible to prove directly. Moreover, this opens up new ways to study the BSD conjecture by studying L-functions of modular forms. This was done, for example, in the work of Gross and Zagier, which led to a proof of the BSD conjecture for rank 0 and rank 1 curves.

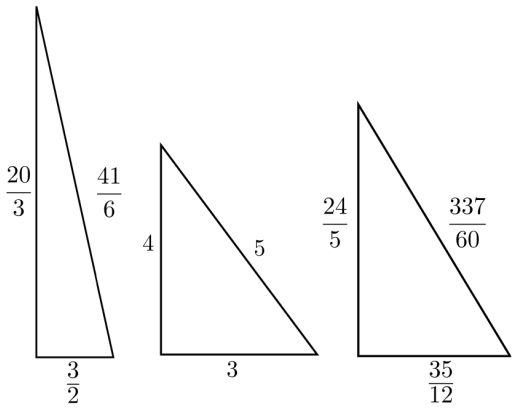
The connection between the Modularity Theorem and Fermat's last theorem was suggested by Frey in 1986, and proved shortly after by Serre and Ribet. They showed that if $a, b, c \in \mathbb{N}$ are a (hypothetical) solution to Fermat's equation $a^p + b^p = c^p$ with some prime $p \geq 5$, then the *Frey curve*

$$E_{a,b,c} : y^2 = x(x-a^p)(x+b^p)$$

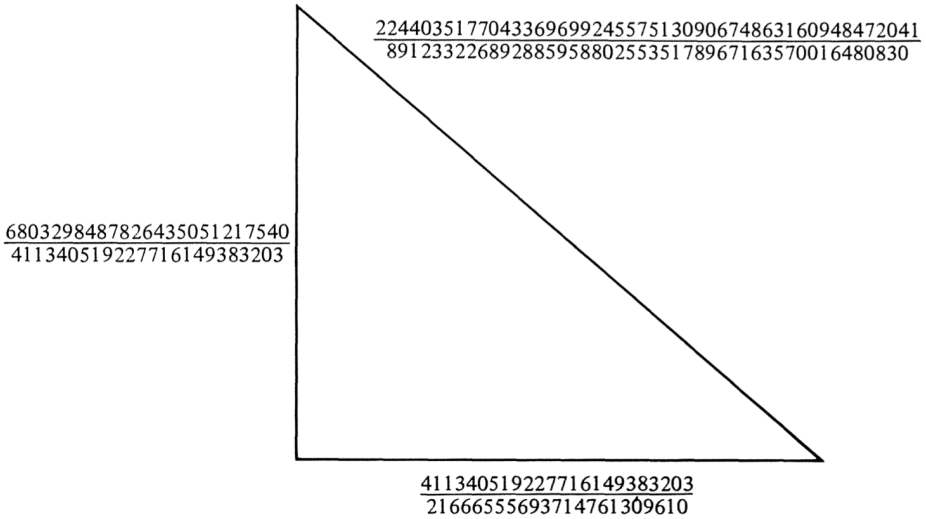
would not be modular, contradicting the Modularity Theorem.

9.3 Congruent numbers and Tunnell's Theorem

The BSD Conjecture has many applications in number theory, a very prominent one being a solution to the *Congruent Number Problem*. A natural number n is called a *congruent number* if it is the area of a right triangle with rational side lengths. For example, $5, 6, 7$ are congruent numbers as they the areas of the following right triangles:



An amusing example is the following (simplest!) rational right triangle with area 157, due to Zagier.



On the other hand, one can show by elementary (yet difficult) considerations that $1, 2, 3, 4$ are not congruent. The *Congruent Number Problem* asks for a (simple) description of all congruent numbers.

The above examples show that finding a suitable triangle for a congruent number can be difficult since the side lengths will usually be complicated rational numbers. Conversely, showing that a number is *not* a congruent number seems to be even more difficult since we need to show that there is *no* right triangle with rational side lengths having area n.

A partial (and conjecturally complete) solution to the Congruent Number Theorem is given by Tunnell's Theorem.

Theorem 9.3.1 (Tunnell 1983)**.** *Let n be a square free natural number. Define the* representation numbers

$$\begin{aligned} A_n &= \#\{(x,y,z) \in \mathbb{Z}^3 \ : n = 2x^2 + y^2 + 32z^2\}, \\ B_n &= \#\{(x,y,z) \in \mathbb{Z}^3 \ : n = 2x^2 + y^2 + 8z^2\}, \\ C_n &= \#\{(x,y,z) \in \mathbb{Z}^3 \ : n = 8x^2 + 2y^2 + 64z^2\}, \\ D_n &= \#\{(x,y,z) \in \mathbb{Z}^3 \ : n = 8x^2 + 2y^2 + 16z^2\}. \end{aligned}$$

Suppose that n is congruent. If n is odd, then $2A_n = B_n$, and if n is even, then $2C_n = D_n$.

Conversely, if the BSD conjecture is true for curves of the form $E_n : y^2 = x^3 - n^2x$, then these equalities are sufficient to conclude that n is a congruent number.

Note that the restriction to square free numbers is inesssential, since a natural number n is congruent if and only if n/d^2 is congruent for any $d \in \mathbb{Q}$. Indeed, if (a,b,c) is a rational right triangle for n, then $(a/d, b/d, c/d)$ is a rational right triangle for n/d^2.

Note that the numbers A_n, B_n, C_n, D_n can be computed quite easily by just trying all possible solutions (x,y,z) with $|x|, |y|, |z| \leq \sqrt{n}$. In particular, Tunnell's Theorem gives an effective procedure to find non-congruent numbers. For example, for $n = 1$ we have $4 = 2A_1 \neq B_1 = 2$ (the only solutions being $(x,y,z) = (0, \pm 1, 0)$ in both cases), so Tunnell's Theorem tells us that 1 is not congruent.

Conversely, if we apply Tunnell's Theorem with the congruent number $n = 5$, we see that $0 = 2A_5 = B_5 = 0$, so we may not conclude from the theorem that n is a congruent number. However, if $2A_n = B_n$ or $2C_n = D_n$ then this is a very strong hint that n should be a congruent number, and we can start looking for a suitable triangle, for example by a computer search.

Congruent numbers and elliptic curves - the idea of the proof of Tunnell's Theorem

The question whether a natural number n is congruent is closely related to the rank of a certain elliptic curve E_n. First notice that n is congruent if and only if there exist positive rational numbers a, b, c satisfying the equations

$$\begin{aligned} a^2 + b^2 &= c^2, \\ n &= \frac{1}{2}ab. \end{aligned}$$

Now if we set $x = n(a+c)/b$ and $y = 2n^2(a+c)/b^2$, we see after a short calculation that the point (x,y) lies on the elliptic curve

$$E_n : y^2 = x^3 - n^2x$$

and satisfies $y \neq 0$ (which means that (x,y) does not have order 2). Conversely, given a rational point $(x,y) \in E_n$ with $y \neq 0$ we set $a = (x^2 - n^2)/y, b = 2nx/y, c = (x^2 + n^2)/y$ to obtain a right triangle with rational side lengths and area n. Moreover, it is not hard to show that the only torsion points on the curve E_n are those with $y = 0$. In other words: the rational right triangles with area n correspond to the rational points $(x,y) \in E_n$ of infinite order. Hence, we obtain the following criterion.

Proposition 9.3.2. *A natural number n is congruent if and only if the elliptic curve $E_n : y^2 = x^3 - n^2x$ has rank > 0.*

The proposition tells us that we could decide wether n is congruent if we could check whether E_n has positive rank. By the BSD conjecture, this should be the case if the L-function $L(E, s)$ vanishes at $s = 1$. In our situation, we are lucky since the E_n are quite special: the curves E_n all have *complex multiplication*, and for such curves we have the following partial answer to the BSD conjecture:

Theorem 9.3.3 (Coates-Wiles 1976)**.** *Let E be a rational elliptic curve with complex multiplication. If E has positive rank, then $L(E, 1) = 0$.*

In particular, if n is congruent, then E_n has positive rank, so $L(E_n, 1) = 0$, and the converse would be true under the BSD conjecture. Now we need to decide wether $L(E_n, 1) = 0$, and here the Modularity Theorem comes into play. It tells us that there exists a modular form G_n of weight 2 such that $L(G_n, s) = L(E_n, s)$. So we need to decide when the L-function of a modular form vanishes at $s = 1$. Using the theory of modular forms, one can show that $L(G_n, 1)$ equals the n-th Fourier coefficient of a certain modular form f of weight $\frac{3}{2}$. The crucial insight is that this modular form f can be constructed very explicitly as a linear combination of *theta series* of the form

$$\theta_{[a,b,c]}(z) = \sum_{x,y,z \in \mathbb{Z}} q^{ax^2+by^2+cz^2}$$

for suitable integers a, b, c. More precisely, it turns out that the n-th coefficient of f is given by a non-zero multiple of $2A_n - B_n$ if n is odd and $2C_n - D_n$ if n is even. Summarizing: if n is congruent, then E_n has positive rank, hence $L(E_n, 1) = 0$ by Coates-Wiles, so $L(G_n, 1) = 0$ by the Modularity Theorem, and thus the n-th coefficient of f vanishes, which means $2A_n = B_n$ if n is odd and $2C_n = D_n$ if n is even. The converse would be true if we could show that $L(E_n, 1) = 0$ implies $r(E) > 0$, which would be a consequence of the BSD conjecture.