Integrality Properties in the Moduli Space of Elliptic Curves

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch–Naturwissenschaftlichen Fakultät der Universität Basel

von

Stefan Schmid

aus

Deutschland

Basel, 2019

Genehmigt von der Philosophisch–Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Philipp Habegger
Prof. Dr. Yuri Bilu

Basel, den 19. Februar 2019.

Prof. Dr. Martin Spiess
Dekan

This page
intentionally
left blank.

Acknowledgements

Hereby, I would like to thank all people that accompanied me during my undergraduate and graduate studies, inside and outside the university.

A huge thank you goes out to my PhD advisor Professor Dr. Philipp Habegger. He offered me the PhD position, introduced me to this very interesting topic and kept me motivated throughout the years. I am very grateful for the time he dedicated for our meetings and the fertile discussions. His thorough read–through of the drafts of the work in hand has helped me tremendously. I am also thankful for the oppurtinity I got to move to Basel.

I would also like to thank Professor Dr. Yuri Bilu from the Université de Bordeaux for his helpful comments on this manuscript and for taking the time to referee my thesis.

Special thanks also to Gabriel Dill and Teresa Schmid for proof–reading one of the first drafts, and for their many helpful comments. Also thanks to Gabriel for the discussions and comments throughout the last years. This thank you also goes to Fabrizio Barroero and Francesco Veneziano.

Also thank you to the participiants of the Thematic Program on "Unlikely Intersections, Heights, and Efficient Congruencing" at the Fields Institute in Toronto, Canada.

Thank you to all people at the mathematics and computer science department of the University of Basel and Technical University of Darmstadt. I would also like to take the opportunity to thank my family for all the support over the years. This includes Fabienne Teysseire for her strong support during the last months before finishing.

Summary

In the thesis at hand we discuss two problems of integral points in the moduli space of elliptic curves. The first problem can be described as follows. We fix an algebraic number α that is the j–invariant of an elliptic curve without complex multiplication. We prove that the number of j–invariants with complex multiplication such that $j - \alpha$ is an algebraic unit can be bounded by a computable number.

The second problem is of similar nature. For this we fix j_0 the j–invariant of an elliptic curve without complex multiplication defined over some number field. We show that there are only finitely many algebraic units j such that elliptic curves with j–invariants j and j_0 are isogenous. A slight modification shows that only finitely j–invariants exists such that j and j_0 are isogenous and such that $j - \alpha$ is a unit, where α is an arbitrary but fixed j–invariant of an elliptic curve with complex multiplication.

Contents

Introduction

One aim of number theory is to describe the integral solutions of algebraic equations. One very important class of examples are diophantine equations, named after Diophantus of Alexandria, that are polynomial equations with integer coefficients. They have been studied since the ancient Greeks. General questions are if there are any solutions, and if, are there infinitely many? A first example are the linear diophantine equations. A general equation is given by $aX + bY = c$, where $a, b \neq 0$ and c are integers. It is well–known that such an equation has a solution if and only if the greatest common divisor of a and b divides c. Moreover, if there is a single solution, then there must be infinitely many.

Now that we have fully classified the linear case, we can consider polynomials in two variables of degree 2. Those are given by $aX^2 + bXY + cY^2 + dX + eY + f = 0$. Again the coefficients $a, b, c, d, e,$ and f are integers. An example of such an equation would be Pell's equation $X^2 - nY^2 = 1$, where n is a positive integer. Obviously, $x = \pm 1$ and $y = 0$ is a trivial solution. Amongst others, this has been studied by Fermat and Lagrange, who proved that there are x and $y > 0$ satisfying the equation if n is not a perfect square. In addition, there are infinitely many solutions. If n is a square, then there is only the trivial solution.

Next up would be integral equations in degree 3, e.g. Fermat equations $X^3+Y^3-Z^3 = 0$ or more general $X^n+Y^n-Z^n = 0$ for $n \geq 3$. Fermat's last theorem says, that there are no integer solutions other then $X = Y = Z = 0$. This was proven by Andrew Wiles and others in the 1990's and was an open problem for over 300 years. One very important ingredient to the proof is a conjecture by Gerhard Frey. It involves the so called Frey curves given by $y^2 = x(x-a^n)(x+b^n)$. He conjectured that given a non–trivial solution to a Fermat equation would mean that the associated Frey curve is not modular. This was proved by Ribet. Later results show that Fermat's last theorem follows from the Shimura–Taniyama conjecture, today known as the Modularity Theorem. The Frey curves are a special type of *elliptic curves*, which more generally are given by equations of the form $y^2 = x^3 + Ax + B$ with $-4A^3 - 27B^2 \neq 0$.

Elliptic Curves

Elliptic curves play a very important role in modern mathematics. They have been studied for over a century. In the modern world of technology, these curves are omnipresent and are for example used in cryptography. But they also played a very important role in the proof of Fermat's last theorem, monstrous moonshine and also are involved in the Birch and Swinnerton–Dyer conjecture. The interesting thing about these geometric objects is, that is possible to define a group structure by identifying the distinguished point at "infinty" with the neutral element of a group. This gives them a rich structure. The integers act on the points of an elliptic curve since they have a group structure. Most of the time the endomorphism ring of an elliptic curve is just $\mathbb{Z}$. Sometimes we have more complex endomorphisms and in that case an elliptic curve is said to have *complex multiplication*. Given an elliptic curve $E\colon y^2 = x^3 + Ax + B$ we have $-16(4A^3 + 27B^2) \neq 0$. To such an equation we can associate $j = 1728\frac{4A^3}{4A^3+27B^2}$. This is an invariant of the elliptic curve and is thus called the j–invariant. The j–invariants of elliptic curves with complex multiplication are called *singular moduli*.

A classical result by Kronecker states that singular moduli are algebraic integers. So the next natural question to ask is when are singular moduli algebraic units, i.e. units in the ring of algebraic integers. Indeed, David Masser asked at the AIM workshop on unlikely intersections in algebraic groups and Shimura varieties in Pisa in 2011, if there are only finitely many singular moduli that are algebraic units. His question was motivated by [BMZ13]. In 2014, Philipp Habegger gave an answer in [Hab15] to this question by proving

Theorem. *At most finitely many singular moduli are algebraic units.*

Thus, the next question to ask would be if there are any singular moduli that are algebraic units. His proof relies on Duke's equidistribution theorem which is not known to be effective. Hence, no bounds for the number of singular units were known. In his paper he also proved that there are only finitely many singular moduli j such that $j+1$ is a unit. An example of such a j would be the j–invariant of the curve $y^2 = x^3 + 1$. In 2018, Yuri Bilu, Philipp Habegger, and Lars Kühne used different methods to prove

Theorem. *There are no singular moduli that are algebraic units.*

The idea of the proof is as follows. They give lower and upper bounds for the height of such singular moduli. The height of an algebraic number basically measures its complexity. The height of an algebraic number α is defined by

$$h(\alpha)(\alpha) = \frac{1}{[K:\mathbb{Q}]} \sum_{\nu \in M_K} [K_\nu : \mathbb{Q}_\nu] \log \max\{1, |\alpha|_\nu\}.$$

Here K is any field containing α and M_K is a set of representatives of non–trivial absolute values extending the p–adic absolute values and the usual absolute value. Using the

height they prove that the absolute value of discriminant associated to the elliptic curve with j–invariant j is bound from above by 10^{15}. The rest of the discriminants can be checked by refining the arguments and by computer calculations.

In the work at hand we are going to investigate a similar problem. We have previously given an example such that $j+1$ is an algebraic unit. More generally one could look at $j-\alpha$ being a unit, and ask if there are finitely many singular moduli j satisfying this. Recently, Yingkun Li proved in [Li18] that if j and α are singular moduli with coprime fundamental discriminants, then $j-\alpha$ can not be a unit. His proof relies on different techniques based on the work of Gross–Zagier ([GZ84], [GZ86]) and Gross–Kohnen–Zagier ([GKZ87]).

We give a partial answer to the question when α is fixed as in the following theorem. We say Δ is the discriminant of a singular modulus if the endomorphism ring of the elliptic curve associated to the singular modulus has discriminant Δ.

Theorem 2.1 *Let j be a singular modulus and let Δ be its discriminant. Let α be an algebraic number that is the j–invariant of an elliptic curve without complex multiplication. If we assume that $j-\alpha$ is an algebraic unit, then $|\Delta|$ is bounded from above by a computable constant that can be found on page 36. In particular, the set of singular moduli j such that $j-\alpha$ is an algebraic unit is effectively computable.*

The ideas come from [BHK18]. The sketch of the proof is as follows. Write $j(\xi)=\alpha\in\bar{\mathbb{Q}}$ where the elliptic curve associated to α does not have complex multiplication. Here $j(\xi)$ denotes Klein's j–function evaluated at ξ. Also we can assume that ξ is in the fundamental domain $\mathcal{F}$ of the standard upper half–plane. To a singular modulus j we have attached an elliptic curve with complex multiplication. The endomorphism ring of this elliptic curve is an order of discriminant Δ. We can write $\Delta=Df^2$ where f is the conductor of the endomorphism ring in the full ring of integers of $\mathbb{Q}(\sqrt{\Delta})$, and D is the discrimiant of that field. The Galois conjugates of j form a full orbit of length the class number $\mathcal{C}(\Delta)$. We write $\mathcal{C}(\Delta;\xi;\varepsilon)$ for the number of singular moduli that can be written in the form $j(\tau)$ with $\tau\in\mathcal{F}$ and such that $|\tau-\xi|<\varepsilon$. We prove an explicit bound on $\mathcal{C}(\Delta;\xi;\varepsilon)$ which is given by

$$\mathcal{C}(\Delta;\xi;\varepsilon)\leq F(\Delta)\left(32|\Delta|^{1/2}\varepsilon^2\log\log(|\Delta|^{1/2})+11|\Delta|^{1/2}\varepsilon+2\right) \tag{1}$$

for $|\Delta|\geq 10^{14}$ and $0<\varepsilon<1/2$. Here

$$F(\Delta)=\max\left\{2^{\omega(a)};a\leq|\Delta|^{1/2}\right\},$$

and $\omega(n)$ is the number of distinct prime divisors of n. Now if $j-\alpha$ is an algebraic unit, the height can be bounded as

$$h(j-\alpha)\ll\frac{\mathcal{C}(\Delta;\xi';\varepsilon)}{\mathcal{C}(\Delta)}(\log|\Delta|)^4-\log\varepsilon \tag{2}$$

for some $\xi' \in \mathcal{F}$ associated to ξ. The constant in the inequality depends on α. We put $E(\Delta) = F(\log|\Delta|)^4$ and roughly choose ε to be

$$\varepsilon = \frac{\mathcal{C}(\Delta)}{E(\Delta)|\Delta|^{1/2}}.$$

If we substitute this and (1) into (2) and use estimates for $\omega(n)$ by Robin [Rob83] we get

$$h(j-\alpha) \ll \frac{E(\Delta)}{\mathcal{C}(\Delta)} + \log\frac{E(\Delta)|\Delta|^{1/2}}{\mathcal{C}(\Delta)}. \tag{3}$$

To bound $|\Delta|$ from above we need lower bounds for the height of $j-\alpha$. One can prove

$$h(j-\alpha) \gg \log|\Delta|$$

and

$$h(j-\alpha) \gg \frac{|\Delta|^{1/2}}{\mathcal{C}(\Delta)}.$$

The first inequality is due to Colmez [Col98] and Nakkajima–Taguchi [NT91], and the second inequality is elementary. Again the bounds depend on α. Combining the lower bounds with the upper bound from (3) we obtain

$$\max\left\{\frac{|\Delta|^{1/2}}{\mathcal{C}(\Delta)}, \log|\Delta|\right\} \ll \frac{E(\Delta)}{\mathcal{C}(\Delta)} + \log\frac{|\Delta|^{1/2}}{\mathcal{C}(\Delta)} + \log E(\Delta)$$

for large $|\Delta|$. Further analysis shows that $E(\Delta)|\Delta|^{-1/2} = |\Delta|^{o(1)}$ and $\log E(\Delta)/\log|\Delta| = o(1)$. Thus the inequality can not hold for large values of $|\Delta|$. All constants in the above deductions can be made explicit, but some are very large.

We started this introduction with algebraic equations and integral solutions. Curves are special algebraic equations and integral solutions of these equations correspond to points with integral coordinates on the curves. Now the *modular curve* $Y(1)$ is defined as the quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. This is nothing more than the affine line. The j–function gives a bijection from $Y(1)$ to the moduli space of isomorphism classes of complex elliptic curves, i.e. the space of isomorphism classes of complex elliptic curves parameterizes the modular curve. We can compactify this curve by adding a point at infinity to get $X(1)$, the projective line. This is a geometrically irreducible projective smooth curve defined over $\mathbb{Q}$.

The notion of integral points generalizes as follows. See [Ser89] for more details. Let X be a geometrically irreducible projective smooth curve defined over a number field L. Let $C \subseteq X(\bar{L})$ be a finite set of $\bar{L}$–rational points on X and write $\bar{L}[X \setminus C]$ for the set of rational functions that are regular on $X \setminus C$. A set $M \subseteq X(\bar{L}) \setminus C$ is called *quasi–algebraic–integral with respect to* C if for every $f \in \bar{L}[X \setminus C]$ there is a $\beta \in \bar{L}^\times$

such that $f(M) \subseteq \beta\mathcal{O}_{\bar{L}}$, where $\mathcal{O}_{\bar{L}}$ is the ring of algebraic integers in $\bar{L}$. One can think of this as the coordinates of M having a common denominator. It is always possible to add a finite set of rational points to a quasi–algebraic–integral set without loosing this property. So it is not a very useful concept to aim for effective results. The theorem of Habegger above can be reformulated in these terms: If $M \subseteq Y(1)(\bar{\mathbb{Q}})$ is a set of singular moduli that is quasi–algebraic–integral with respect to $\{0, \infty\}$, then M is finite.

Now assume that j is the j–invariant of an elliptic curve. Assume that j is an algebraic unit. Then by the work of Bilu–Habegger–Kühne the elliptic curve does not have complex multiplication. Thus there are infinitely many j–invariants of elliptic curves without complex multiplication that are algebraic units. This is because we can construct an elliptic curve with a given j–invariant. To prove a similar result we must find a replacement for the absence of the discriminant. For this we fix the j–invariant j_0 of an elliptic curve without complex multiplication. We say that j is isogenous to j_0 if the elliptic curves with j–invariant j and j_0, respectively, are isogenous. We look at the set of all j that are isogenous to j_0 such that j is an algebraic unit. This set is finite and this is one of our main theorems.

Theorem 3.21 *Let j_0 be the j–invariant of an elliptic curve without complex multiplication. Then there are at most finitely many j–invariants j of elliptic curves that are isogenous to an elliptic curve corresponding to j_0 and such that j is an algebraic unit.*

This problem can again be reformulated in a problem in the moduli space. Let $j_0 \in Y(1)(\bar{\mathbb{Q}})$ be fixed, but assume that j_0 is not a singular modulus. Since the modular j function is a surjective we can choose τ_0 in the Poincaré upper half–plane with $j(\tau_0) = j_0$. If $M \subseteq Y(1)(\bar{\mathbb{Q}})$ is a set of points of the form $j\left(\left(\begin{smallmatrix} m & l \\ 0 & n \end{smallmatrix}\right).\tau_0\right)$ with $\gcd(m, n, l) = 1$, $0 \leq l < m$, that is quasi–algebraic–integral with respect to $\{0, \infty\}$, then M is finite. In other words, M contains j–invariants isogenous to j_0.

The problem with this formulation is, that we can not give explicit bounds like the ones we will see shortly. We say that an isogeny between two elliptic curves is minimal if its degree is minimal amongst the isogenies between the curves. Assume we have fixed a model E_0 of an elliptic curve with j–invariant j_0 and that it is defined over K. Further assume $j(\tau_0) = j_0$. For an elliptic curve E_0 defined over K we let E_0^σ for an embedding $\sigma\colon K \hookrightarrow \mathbb{C}$ be the elliptic curve obtained by conjugating the coefficients. Note that in the following τ_0^σ is closely related to τ_0. Attached to the elliptic curve E_0 we have a representation ρ_∞ from the absolute Galois group $G_K = \mathrm{Gal}(\bar{K}/K)$ of K to $\mathrm{GL}_2(\hat{\mathbb{Z}})$. With this information the theorem above can be made explicit again and we have

Theorem 3.22 *Let $E_0\colon y^2 = 4x^3 - g_2x - g_3$ be an elliptic curve without complex multiplication defined over a number field K of degree D. Let j_0 be its j–invariant with $j(\tau_0) = j_0$ and $\tau_0 \in \mathcal{F}$. We choose ω_1 and ω_2 with $\omega_2/\omega_1 = \tau_0$ and $E_0(\mathbb{C}) \simeq \mathbb{C}/(\omega_1\mathbb{Z} + \omega_2\mathbb{Z})$ and similarly for E_0^σ, $\sigma\colon K \hookrightarrow \mathbb{C}$. Define $h = \max\{1, h(1, g_2, g_3), h(j_0)\}$.*

If j is the j–invariant of an elliptic curve isogenous to E_0 such that j is a unit, then the degree of the minimal isogeny between j_0 and j is bounded by

$$\max\left\{10^{180}(Cc_1)^{20},(Cc_2)^{10},e^{Cc_1+Cc_2+c_3},e^{120^2[\mathrm{GL}_2(\hat{\mathbb{Z}}):\rho_\infty(G_K)]^2},e^{18\pi h},D\right\},$$

where the constants are given by

$$\begin{aligned}
C &= 6\cdot 10^7\cdot h\cdot D[\mathrm{GL}_2(\hat{\mathbb{Z}}):\rho_\infty(G_K)],\\
c_1 &= 2\cdot 10^{51}D^6\cdot\max\{h,6\pi|\tau_0|\}^2\geq 1,\\
c_2 &= 14+3\log\left(\max_\sigma\{1,|\omega_{0,1}^\sigma|,|\omega_{0,2}^\sigma|\}\right)\ \textit{and}\\
c_3 &= 20-h(j_0)+6\log(1+h(j_0)).
\end{aligned}$$

For a fixed singular modulus α we can again look at the set of j–invariants j such that $j-\alpha$ is a unit and such that j is isogenous to j_0. In this case similar results hold and the statement goes as follows.

Theorem 3.25 *Assume α is the j–invariant of an elliptic curve with CM. Let j_0 be the j–invariant of an elliptic curve without CM. Then there are at most finitely many j–invariants j of elliptic curves that are isogenous to an elliptic curve corresponding to j_0 and such that $j-\alpha$ is an algebraic unit.*

A similar bound for the minimal isogeny as in the case when $\alpha=0$ can be obtained. The idea of the proof is related to the previous one. Given a minimal isogeny of degree N between j and j_0 we give lower and upper bounds for the height of $j-\alpha$ in terms of N. The bounds contradict each other for large values of N.

We discuss a couple of differences. Instead of counting CM–points in the fundamental domain we count points isogenous to a fixed point that lie in a neighborhood of a point in the fundamental domain. To do this we look at all the isogenous points of the fixed one and then translate them to the fundamental domain. If a resulting point is close to a fixed point ξ in the fundamental domain, then the entries of the matrix in $\mathrm{SL}_2(\mathbb{Z})$ must be close to an ellipse. We then count the possible matrices using a result of Davenport. We can then compare the isogeny orbit to the Galois orbit to estimate the number of embeddings contributing most to the height of $j-\alpha$. To get an explicit lower bound for $h(j-\alpha)$ we use a result by Autissier [Aut03] together with a classical result that relates the j–invariant of an elliptic curve with the Faltings height. One of the main ingredients for the upper bound are a result on linear forms in logarithms by Sinnou David [Dav95] and a result by Lombardo that gives an explicit upper bound for Serre's open image theorem. This bound is very big so that our result only gives big bounds.

Outline

A reader is supposed to know basic algebraic number theory, abstract algebra and complex analysis usually taught in undergraduate level. Moreover, some knowledge on quadratic forms is useful. A good understanding of elliptic curves and Klein's j–function are a plus. Nevertheless, we shall give a short introduction to those topics but we will omit one or two proofs. References can be found in the introductory chapter. We will state the necessary definitions and theorems in the first chapter. This will also include a section on heights in the projective space.

The second chapter is devoted to the proof of Theorem 2.1. We will workout all the details and the constants to get explicit bounds. One can also consult [BHK18] for more details or different perspectives.

The third chapter is a bit longer and includes a proof of Theorem 3.25 and Theorem 3.26. The chapter is split into two parts. We proof the case when $\alpha = 0$ for the sake of simplicity first and then do the general case. The first section contains some computations on the isogenous points. The argument reduces to counting lattice points.

Notation and terminology

We introduce some basic notation. For two sets A and B we write $A \subseteq B$ if A is a subset of B. We denote by $\mathbb{Z}$ the set of (rational) integers, and by $\mathbb{N} := \{1, 2, 3, \dots\}$ the positive integers. The sets $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are the fields of rational, real and complex numbers, respectively. The elements of $\mathbb{C}$ will be of the form $x + iy$ with $x, y \in \mathbb{R}$ and $i^2 = -1$. Throughout this exposition ζ will be the complex number $e^{2\pi i/6}$. For $z \in \mathbb{C}$ we will denote its complex conjugate by $\bar{z}$. We use $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ to indicate the real and imaginary part, respectively, of a complex number z. The letter $\mathbb{H}$ is reserved for the complex upper half–plane, i.e. all complex numbers satisfying $\mathrm{Im}(z) > 0$. We fix once and for all one embedding of $\bar{\mathbb{Q}}$ into the complex numbers. For any field K the n–dimensional projective space will be denoted by $\mathbb{P}^n_K$, and an element in there will be written as $[x_0 : \dots : x_n]$.

For a finite field extension L over K we denote by $[L : K]$ the degree of the extension. By $N_{L/K}$ we denote the norm of L over K.

For a finite set M we denote the number of elements in M by $\#M$. For integers a and b we write $a|b$ if a divides b, and we denote the greatest common divisor of a and b by $\gcd(a, b)$ or (a, b). The number $\sqrt{a}$, $a \geq 0$, is the unique positive solution of the polynomial $X^2 - a$. Moreover, $\sqrt{-a}$ is defined to be $i\sqrt{a}$. The function log denotes the logarithm defined for positive real numbers with $\log e = 1$. Throughout the text, we will use $q := e^{2\pi i\tau}$ for τ in the complex upper half–plane. As usual, Γ denotes the gamma function.

If k is a positive integer and R is a ring, we define the set of invertible k–by–k matrices with entries in R by $\mathrm{GL}_k(R)$, and write $\mathrm{SL}_k(R)$ for the subset of matrices with

determinant 1. For a group or module M we write $\mathrm{Aut}(M)$ for the set of automorphisms on M. For $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and $\tau \in \mathbb{H}$ we write $\gamma.\tau$ or $\gamma\tau$ for the usual action of $\mathrm{SL}_2(\mathbb{Z})$ on the upper half–plane by möbius transformations, also known as fractional linear transformations. For $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ the action is given by $(a\tau + b)/(c\tau + d)$. If G is a group and H is a subgroup of G, then the index of H in G is denoted by $[G : H]$. If R is a ring, then we denote by $R^\times$ the set of invertible elements.

1 Preliminaries

The aim of this chapter is to give some preliminaries for later chapters. Getting all of the details goes beyond the scope of this work and we will thus refer to different text books. See each section for more details. For the basics of algebraic number theory we refer to [Neu06] or [Rib01]. In this text K will be a field extension of $\mathbb{Q}$.

1.1 Quadratic forms

We want to list some of the important properties of quadratic forms since they are closely related to elliptic curves with complex multiplication. Details can be found in [Cox11].

Let $Q(x, y) = ax^2 + bxy + cy^2$ be a (binary) quadratic form. We call Q *primitive* if the coefficients a, b, c are coprime.

Two quadratic forms $Q(x, y)$ and $Q'(x, y)$ are said to be *equivalent* if there exists $\gamma \in \mathrm{GL}_2(\mathbb{Z})$ with

$$Q(x, y) = Q'((x, y)\gamma),$$

where $(x, y)\gamma$ denotes the usual matrix multiplication. This is obviously an equivalence relation.

The *discriminant* of a quadratic form $Q(x, y) = ax^2 + bxy + cy^2$ is defined by $\Delta = \Delta_Q = b^2 - 4ac$. If Q and Q' are equivalent as before, then it is easy to see that $\Delta_Q = (\det \gamma)^2 \Delta_{Q'} = \Delta_{Q'}$.

A quadratic form $Q(x, y)$ is called *positive definite* if $Q(x, y) > 0$ for all $(x, y) \neq 0$. It is called *negative definite* if $Q(x, y) < 0$. If Q is positive definite, then the polynomial $Q(x, 1)$ does not have any roots in $\mathbb{R}$, so that we must have $\Delta < 0$.

We say that a primitive positive definite quadratic form $Q(x, y) = ax^2 + bxy + cy^2$ is *reduced* if

$$-a < b \leq a < c \text{ or } 0 \leq b \leq a = c. \tag{1.1}$$

It was Gauß who showed in *Disquisitiones Arithmeticae* that every primitive positive definite form is equivalent to exactly one reduced form. There is a simple algorithm to determine this representative.

Let $Q(x, y)$ be a reduced positive definite quadratic form. By α_Q we denote the unique (complex) solution in the upper half–plane of the polynomial $Q(x, -1)$. We claim that

α_Q is in the fundamental domain

$$\mathcal{F} = \{z \in \mathbb{H}; |z| > 1 \text{ and } |\operatorname{Re}(z)| < 1/2\} \cup \{z \in \mathbb{H}; |z| \geq 1 \text{ and } 0 \leq \operatorname{Re}(z) \leq 1/2\}.$$

Lemma 1.1 *We have $\alpha_Q \in \mathcal{F}$. Moreover, $a \leq |\Delta/3|^{1/2}$.*

Proof. By the inequalities in (1.1) we have $|b| \leq a$ and thus

$$|\operatorname{Re}(\alpha_Q)| = \frac{|b|}{2a} \leq \frac{1}{2},$$

with equality only if $|b| = a$ and hence $b = a > 0$ by (1.1). The last condition means $\operatorname{Re}(\alpha_Q) = 1/2$ since $\operatorname{Re}(\alpha_Q) = b/(2a)$. In addition, the modulus of α_Q is given by

$$|\alpha_Q| = \sqrt{\frac{b^2}{4a^2} + \frac{|\Delta|}{4a^2}} = \sqrt{\frac{c}{a}}.$$

This is equal to 1 if and only if $c = a$ which again implies $b \geq 0$ and $\operatorname{Re}(\alpha_Q) \geq 0$.
The second claim also easily follows from (1.1) as

$$a = \sqrt{(4a^2 - a^2)/3} \leq \sqrt{(4ac - b^2)/3} = \sqrt{|\Delta|/3}. \qquad \square$$

Define the set $\mathcal{Q}(\Delta)$ to be the set of equivalence classes of primitive positive definite quadratic forms. The previous lemma shows that for fixed Δ the set $\mathcal{Q}(\Delta)$ is finite. This is because $a \leq |\Delta/3|^{1/2}$ and b satisfies $-a < b \leq a$ by equation (1.1). Thus, there are only finitely many values for a and b but c is determined by a and b, since $\Delta = b^2 - 4ac$. We will denote the number of elements in $\mathcal{Q}(\Delta)$ by $\mathcal{C}(\Delta)$ and call it the *class number*. Note that we do not use the usual notation $h(\Delta)$ for the class number, because h will be reserved for the height.
We finish with an example. If $\Delta = -4$, then $a = 1$ which again implies $b = 0$ and $c = 1$. Thus, $\mathcal{C}(-4) = 1$ and $x^2 + y^2$ is a representative of that class.

1.2 Heights

In this section we want to give some background on heights on projective varieties. We will skip most of the proofs and refer to [BG07] as a good source for the material covered. Another reference is [Wal13].

1.2.1 Absolute values

We want to give a brief overview on the theory of valuations on number fields. This is also discussed in [Neu06] or in [Rib12]. In this section, K will be a number field.

Definition. *By an **absolute value** on K we mean a multiplicative function $|\cdot|\colon K \to \mathbb{R}_{\geq 0}$ satisfying:*
(i) $|a| = 0$ if and only if $a = 0$.
(ii) $|ab| = |a||b|$ for all $a, b \in K$.
(iii) $|a + b| \leq |a| + |b|$ for all $a, b \in K$.
Furthermore, if it satisfies the ultrametric triangle inequality
(iv) $|a + b| \leq \max\{|a|, |b|\}$ for all $a, b \in K$,
*then then it is called **non–archimedean** absolute value. Otherwise, we say it is **archimedean** or **infinite**.*

Any rational number $a \neq 0$ can be written as a product of primes $\pm p_1^{\nu_{p_1}(a)} \cdots p_n^{\nu_{p_n}(a)}$, where the $\nu_{p_i}(a)$ are uniquely determined integers. Fixing a prime p, and defining $|a|_p = p^{-\nu_p(a)}$ defines an absolute value on $\mathbb{Q}$ called the p–adic absolute value. We say two absolute values $|\cdot|_1$ and $|\cdot|_2$ are equivalent, if and only if there is a positive real number s such that

$$|x|_1 = |x|_2^s$$

for all $x \in K$. Obviously, for two different prime numbers p and q the absolute values are not equivalent. Moreover, none of the p–adic absolute values is equivalent to the standard absolute value on $\mathbb{Q}$. We will also write $|\cdot|_\infty$ for the standard absolute value. The trivial absolute value is equal to 1 except at 0.

By a *place* ν of K we mean an equivalence class of non–trivial absolute values on K. We define $M_\mathbb{Q}$ to be the set of representatives of all non–trivial places on $\mathbb{Q}$ given by $\{|\cdot|_p; p \text{ rational prime}\} \cup \{|\cdot|_\infty\}$. Let L be a finite extension of K and ω, ν be places of L and K, respectively. We write $\omega|\nu$ and say ω divides ν, if the restriction of any element in ω is equal to some element of ν. We can also say that ω extends ν or that ω lies above ν. Any ν of K can be extended to a place ω of L. The set M_K will denote a set of representatives of places on K that restrict to an element of $M_\mathbb{Q}$. The set M_K^∞ is the subset of M_K that contains all the infinite places.

The p–adic numbers $\mathbb{Q}_p$ are the completion of $\mathbb{Q}$ with respect to the p–adic absolute value. If we have $\omega|\nu$ then we can look at the completions L_ω and K_ν. The degree d_ν of L_ω over K_ν is called the *local degree* of L/K in ω. We have the following lemma.

Lemma 1.2 *The degree $[L_\omega : K_\nu]$ is finite and we have*

$$\sum_{\omega|\nu} [L_\omega : K_\nu] = [L : K]$$

where the sum runs over all $\omega \in M_L$ extending ν.

The infinite places can be described well. We have the following result

Lemma 1.3 *Assume L/K is a Galois extension and $\mathrm{Gal}(L/K)$ its Galois group. Let $|\cdot|_{\omega_0}$ and $|\cdot|_{\omega}$ be absolute values of L extending ν. Then there is a $\sigma \in \mathrm{Gal}(L/K)$ with*

$$|x|_{\omega_0} = |\sigma(x)|_{\omega}$$

for all $x \in L$.

In fact, if we look at $K/\mathbb{Q}$, every infinite place of K is represented by $|\sigma(\cdot)|$ for some $\sigma\colon K \hookrightarrow \mathbb{C}$. For a fixed σ with $\sigma(K) \not\subseteq \mathbb{R}$, σ and the complex conjugate $\bar{\sigma}$ define the same absolute value.

The set M_K defined above satisfies the *product formula* given by

$$\prod_{\nu \in \mathcal{M}_K} |x|_{\nu}^{d_\nu} = 1$$

for all $x \in K^{\times}$. This will later be important for the definition of the height. We will in particular be interested in the equality obtained by taking the log.

If $\mathcal{O}_K$ is the ring of integers of K, then the non–zero prime ideals are in bijection with the non–archimedean absolute values on K. We define the *valuation ring* of ν as

$$R_\nu := \{x \in K; |x|_\nu \leq 1\}.$$

We have the following equality

$$\mathcal{O}_K = \bigcap_{\nu \text{ finite}} R_\nu, \tag{1.2}$$

where the intersection runs over all finite places of K.

1.2.2 Heights of algebraic numbers

We are now able to define the height of an algebraic number. The height of an algebraic number basically measures its arithmetic complexity. We first start with the simplest of them, i.e. the elements of $\mathbb{Q}$. Let $\frac{a}{b} \in \mathbb{Q}$ be a rational number with coprime integers a and b different from 0. The height is defined by

$$H\left(\frac{a}{b}\right) = \max\left\{|a|, |b|\right\}.$$

The height of 0 is defined to be 1. One can immediately see, that there are only finitely many rational numbers of bounded height.

Definition. *The **logarithmic (Weil) height** of an algebraic number α is given by*

$$h(\alpha)(\alpha) = \frac{1}{[K:\mathbb{Q}]} \sum_{\nu \in M_K} d_\nu \log \max\{1, |\alpha|_\nu\},$$

where K is any field containing α and $d_\nu = [K_\nu : \mathbb{Q}_\nu]$ denotes the local degree.
If $x = [x_0 : \ldots : x_n] \in \mathbb{P}^n_{\bar{\mathbb{Q}}}$ and K is a number field containing all coordinates of x, then the height of x is defined by

$$h(x) = \frac{1}{[K:\mathbb{Q}]} \sum_{\nu \in M_K} d_\nu \log \max_i \{1, |x_i|_\nu\}.$$

One can show, that the definition is independent of K and together with the product formula, one can see that the definition is independent of the choice of representatives of x. Occasionally, we will write $\log^+ |\alpha|_\nu$ for $\log \max\{1, |\alpha|_\nu\}$. We will write $H(\alpha)$ for $H(\alpha) = e^{h(\alpha)}$. We have the following simple inequality for the height.

Lemma 1.4 *Let α, β be algebraic numbers. Then*

$$h(\alpha\beta) \leq h(\alpha) + h(\beta)$$

and

$$h(\alpha + \beta) \leq h(\alpha) + h(\beta) + \log 2.$$

Proof. Let $x, y \geq 1$. Then $x + y \leq xy + xy = 2xy$. Hence

$$\max\{1, a\} + \max\{1, b\} \leq 2 \max\{1, a\} \max\{1, b\} \tag{1.3}$$

for all $a, b \geq 0$. Assume K is a number field containing α and β, and let ν be a place of K. Then

$$\max\{1, |\alpha + \beta|_\nu\} \leq \max\{1, |\alpha|_\nu\} + \max\{1, |\beta|_\nu\} \leq 2 \max\{1, |\alpha|_\nu\} \max\{1, |\beta|_\nu\}.$$

Moreover, we have

$$|\alpha\beta|_\nu = |\alpha|_\nu |\beta|_\nu \leq \max\{1, |\alpha|_\nu\} \max\{1, |\beta|_\nu\}$$

and therefore

$$\max\{1, |\alpha\beta|_\nu\} \leq \max\{1, |\alpha|_\nu\} \max\{1, |\beta|_\nu\}. \tag{1.4}$$

Taking the logarithm on both sides of equations (1.3) and (1.4), multiplying by $d_\nu/[K : \mathbb{Q}]$ and taking the sum over all $v \in M_K$ shows the desired statement. □

Note that this can be generalized to r numbers by replacing $\log 2$ with $\log r$. Moreover, the bound is sharp. We will later need the following result.

Proposition 1.5 *Let α be an algebraic number and $\sigma \in \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. Then $h(\sigma(\alpha)) = h(\alpha)$.*

The same is true for points in the projective space. The following is a well–known theorem about algebraic numbers of height 0.

Theorem 1.6 (*Kronecker's theorem*)
For $\mu \in \bar{\mathbb{Q}}^\times$ we have $h(\mu) = 0$ if and only if μ is a root of unity.

The definition of the logarithmic height extends the definition of the height of a rational number when taking the logarithm. We have a similar well–known result.

Theorem 1.7 (*Northcott*)
Let K be a number field. Then there are only finitely many $\alpha \in K$ of bounded height $h(\alpha) \leq B$, $B \in \mathbb{R}$.

We can split the sum of the height into the finite and infinite places and obtain

$$h(\alpha) = \frac{1}{[K:\mathbb{Q}]} \left(\sum_{\sigma} \log \max\{1, |\sigma(\alpha)|\} + \sum_{\nu} d_\nu \log \max\{1, |\alpha|_\nu\} \right),$$

where σ runs over all field embeddings $\sigma\colon K \hookrightarrow \mathbb{C}$ and ν runs over all finite places of K.

By equation (1.2), if α is an algebraic integer, then $|\alpha|_\nu \leq 1$ for all finite places ν. Thus the height of an algebraic integer is given by

$$h(\alpha) = \frac{1}{[K:\mathbb{Q}]} \sum_{|\sigma(\alpha)|>1} \log |\sigma(\alpha)|. \tag{1.5}$$

The height also has the following symmetry. This will be important since we are going to use units in the ring of algebraic integers.

Lemma 1.8 *For any $\alpha \in K^\times$ we have $h(\alpha) = h(\alpha^{-1})$.*

Proof. We have $\log |\alpha|_\nu = \log^+ |\alpha|_\nu - \log^+ |\alpha^{-1}|_\nu$ for all ν. We multiply this by the local degree d_ν and take the sum over all ν. Then the left–hand side is 0 because of the product formula and the right–hand side is equal to $h(\alpha) - h(\alpha^{-1})$. □

Assume α is a unit in the ring of integers. Using (1.5) the height amounts to

$$h(\alpha) = h(\alpha^{-1}) = \frac{1}{[K:Q]} \sum_{|\sigma(\alpha^{-1})|>1} \log|\sigma(\alpha^{-1})| = -\frac{1}{[K:Q]} \sum_{|\sigma(\alpha)|<1} \log|\sigma(\alpha)|. \tag{1.6}$$

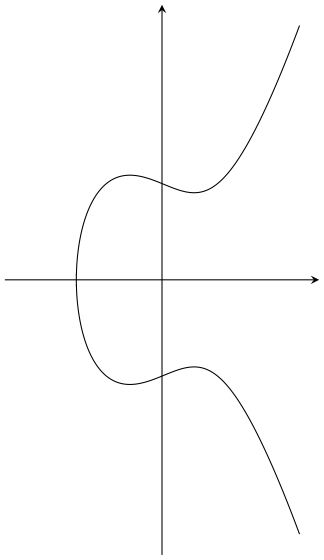
We will use this equality in later chapters.

1.3 Elliptic Curves

The theory of elliptic curves is explained in [Sil86b], [Sil86a] or [ST92]. Deeper results can be found in [Sil99]. Diamond and Shurman give a quick introduction of complex tori and modular curves in [DS05]. Let A, B be in K such that $-(A^3+27B^2) \neq 0$. Then the equation

$$E\colon y^2 = 4x^3 + Ax + B$$

defines an *elliptic curve*. This is a special case of a cubic curve. The defining equation is called *Weierstrass equation*. Adding a "point at infinity" it is possible to define a group structure. The following picture illustrates the example $y^2 = x^3 - 2x + 6$. Note that multiplying this equation by 4 and substituting $y' = 2y$ gives a Weierstrass equation $y^2 = 4x^3 - 8x + 24$.



By a lattice Λ in $\mathbb{C}$ we mean a discrete subgroup of rank 2. If $K \subseteq \mathbb{C}$ then we can see an elliptic curve E as a complex torus $\mathbb{C}/\Lambda$ by taking the quotient of the complex plane by a lattice Λ. This connection is via the *Weierstrass $\wp$–function*

$$\wp(z) = \frac{1}{z^2} + \sum_{\omega \in \Lambda \setminus \{0\}} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right), \; z \in \mathbb{C} \setminus \Lambda.$$

The sum converges absolutely and uniformly on all compact subsets not intersecting Λ, and hence the derivative is given by

$$\wp'(z) = -2 \sum_{\omega \in \Lambda} \frac{1}{(z-\omega)^3}.$$

It is well known that the Weierstrass $\wp$–function satisfies the differential equation

$$\wp'(z)^2 = 4\wp(z)^3 - g_2(\Lambda)\wp(z) - g_3(\Lambda). \tag{1.7}$$

Here $g_2(\Lambda) = 60G_4(\Lambda)$ and $g_3(\Lambda) = 140G_6(\Lambda)$ with $G_k(\Lambda) = \sum_{\omega\in\Lambda\setminus\{0\}} \omega^{-k}$. The functions G_k are so–called *Eisenstein series*. The affine algebraic curve defined by equation (1.7) gives the complex torus $\mathbb{C}/\Lambda$ the structure of an algebraic curve of genus 1 defined over $\mathbb{C}$.

The discriminant of a lattice can be defined as $\Delta(\Lambda) = (g_2(\Lambda))^3 - 27(g_3(\Lambda))^2$. We get a function $\Delta\colon \mathbb{H} \to \mathbb{C}$ by $\Delta(\tau) = \Delta(\Lambda_\tau)$ with $\Lambda_\tau = \mathbb{Z} + \tau\mathbb{Z}$. This function is called the *(modular) discriminant*.

Recall that $\mathrm{SL}_2(\mathbb{Z})$ acts on $\mathbb{H}$ by fractional linear transformations and the fundamental domain is defined by

$$\mathcal{F} = \{\tau \in \mathbb{H}; -1/2 < \mathrm{Re}(\tau) \leq 1/2 \text{ and } |\tau| > 1 \text{ or } |\tau| = 1 \text{ and } 0 \leq \mathrm{Re}(\tau) \leq 1/2\}.$$

This also allows us to define the *modular function*

$$j(\tau)\colon \mathbb{H} \to \mathbb{C},\ j(\tau) = 1728\frac{(g_2(\tau))^3}{\Delta(\tau)},$$

which is holomorphic on $\mathbb{H}$. Note that we could also define j for a lattice as g_2 and Δ are. This function is $\mathrm{SL}_2(\mathbb{Z})$–invariant, i.e.

$$j(\gamma(\tau)) = j(\tau)$$

for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and $\tau \in \mathbb{H}$. We therefore can restrict j to the fundamental domain $\mathcal{F}$. Writing $q = e^{2\pi i\tau}$ we can expand j as

$$j(\tau) = \frac{1}{q} + 744 + 196884q + 21493760q^4 + \cdots$$

with all coefficients non–negative integers as mentioned for example in [Leh42].

As shown above the j–function is connected to elliptic curves, but also has well known connections to class field theory and moonshine. We will shortly present the connection to the former.

1.3.1 Isogenies

The notion of an isogeny is very important for the rest of the work. We will write down important known facts which can be found in most books. In this section we are going to treat elliptic functions over the complex numbers.

Definition. *A nonzero morphism (of varieties) φ between two elliptic curves E_1 and E_2 is called **isogeny** if it maps the point at infinity of E_1 to the point at infinity of E_2, i.e. $\varphi(O) = O$. In this case, E_1 and E_2 are called **isogenous**.*

With this definition, it is possible to show that an isogeny is also a group homomorphism. It is well known that the kernel of such a map is finite, and that the map is surjective. The *degree* $\deg\varphi$ of an isogeny is defined to be the cardinality of the kernel.

When we think about E_1, E_2 being the quotient of the complex plane by lattices Λ_1, Λ_2, the definition amounts to saying that an isogeny is a nonzero holomorphic homomorphism between complex tori. Those are of the form $\varphi(z + \Lambda_1) = mz + \Lambda_2$ for some $m \in \mathbb{C} \setminus \{0\}$ with $m\Lambda_1 \subseteq \Lambda_2$. The degree is equal to the index $[\Lambda_2 : m\Lambda_1]$. The map φ is an isomorphism if and only if $m\Lambda_1 = \Lambda_2$. Isomorphic elliptic curves have the same j–invariant.

By this definition, the maps

$$[N]\colon \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda,\ z + \Lambda \mapsto Nz + \Lambda$$

define isogenies for non–zero integers N. The isogeny $[N]$ is called *multiplication–by–N map*. The elements of the kernel of $[N]$ are called the *N–torsion* points of the elliptic curve, and are denoted by $E[N]$.

Definition. *An isogeny is called **cyclic** if its kernel is a cyclic subgroup. We call a cyclic isogeny of degree N an N–**isogeny**.*

The following lemma can be found in [MW90]. We will need it in later sections, so we will give a short proof.

Lemma 1.9 *An isogeny of minimal degree between two elliptic curves over $\mathbb{C}$ is cyclic.*

Proof. Assume $\Phi\colon \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda', z + \Lambda \mapsto mz + \Lambda'$ with $m\Lambda \subseteq \Lambda'$ is non–cyclic. Write $\Lambda = \omega_1\mathbb{Z} + \omega_2\mathbb{Z}$. Let K be the kernel of Φ and N the order of K. Since K is a subgroup of the N–torsion points, the theory of finite abelian groups implies that K is isomorphic to $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/nn'\mathbb{Z}$ for positive integers n' and n. Note that $n > 1$ otherwise K would be a cyclic subgroup. We have $(m/n)\Lambda \subseteq \Lambda'$ because for $z = a\omega_1 + b\omega_2 \in \Lambda$ we see $\frac{m}{n}z = m\left(\frac{a}{n}\omega_1 + \frac{n'b}{nn'}\omega_2\right) \in m\tilde{K} \subseteq \Lambda'$ with $\tilde{K} = \cup_{z\in K} z$. A similar argument shows that

$$\varphi\colon \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda',\ z + \Lambda \mapsto \frac{m}{n}z + \Lambda'$$

is well–defined.

Now Φ factors as $\varphi \circ [n]$ where $[n]$ is the multiplication–by–n map on $\mathbb{C}/\Lambda$. (This is a special case of Corollary 4.11 in Chapter III of [Sil86b].) Now we have an isogeny φ from $\mathbb{C}/\Lambda$ to $\mathbb{C}/\Lambda'$ with degree less than the degree of Φ. Thus the degree of Φ was not minimal amongst the isogenies between $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda'$. This proves the claim. □

Note that the lemma is also true for two elliptic curves defined over a number field K when considering only isogenies defined over K. See Corollary 4.11 and page 74 in [Sil86b] or [MW90] for a detailed proof.

It is easy to describe an N–isogeny, by the action of integral 2–by–2 matrices on the lattice. Such an isogeny is related by a matrix of the form

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

with $0 \leq b < d$ and $N = ad$ with $\gcd(a, b, d) = 1$. The number of such matrices is given by the *Dedekind ψ–function*

$$\psi(N) = N \prod_{p|N} \left(1 + \frac{1}{p}\right)$$

running over all prime divisors of N. This is a multiplicative function. For details see for example Chapter 5 of [Lan87] or Lemma 11.24 in [Cox11].

For two elliptic curves E and E' we define

$$\mathrm{Hom}(E, E') = \{\text{isogenies from } E_1 \text{ to } E_2\} \cup \{[0]\}$$

and

$$\mathrm{End}(E) = \mathrm{Hom}(E, E)$$

called the endomorphism ring of E. As mentioned before, if we consider a lattice Λ then $\mathrm{End}(\mathbb{C}/\Lambda) = \{\alpha \in \mathbb{C} \setminus \{0\}; \alpha\Lambda \subseteq \Lambda\}$. We have already seen that $\mathbb{Z} \subseteq \mathrm{End}(E)$. The question now is, if there can be more endomorphisms. This brings us to the next part.

1.3.2 Complex Multiplication

Let $E\colon y^2 = 4x^3 - g_2x - g_3$ be an elliptic curve. We can associate a lattice to it which we assume is generated by ω_1 and ω_2. One can show that either $\mathrm{End}(E) = \mathbb{Z}$ or $\mathbb{Q}(\omega_1/\omega_2)$ is an imaginary quadratic extension of $\mathbb{Q}$ and $\mathrm{End}(E)$ is an order in that field. To see that $\mathrm{End}(E)$ lies in a quadratic field we pick $a, b, c, d \in \mathbb{Z}$ such that $\alpha\omega_1 = a\omega_1 + b\omega_2$ and $\alpha\omega_2 = c\omega_1 + d\omega_2$ where α represents an element of $\mathrm{End}(E)$. Hence α is a zero of the quadratic characteristic polynomial of $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Dividing $\alpha\omega_2$ by ω_1 we get $\alpha = c\tau + d$ with $\tau = \omega_1/\omega_2$. Since ω_1 and ω_2 span a lattice we can not have $\tau \in \mathbb{R}$. Thus if α is not an integer, then we must have $c \neq 0$ and $\mathbb{Q}(\alpha) = \mathbb{Q}(\tau)$. Conversely, if τ is imaginary quadratic and $\mathcal{O}$ an order in $\mathbb{Q}(\tau)$, then $E = \mathbb{C}/\mathcal{O}$ is an elliptic curve and $\mathrm{End}(E) = \mathcal{O}$.

Definition. *An elliptic curve E is said to have **complex multiplication**, or CM for short, if* $\mathrm{End}(E)$ *is strictly larger than* $\mathbb{Z}$.

One example is the elliptic curve associated to the lattice $\mathbb{Z}[i]$. The endomorphism ring is equal to $\mathbb{Z}[i]$. Note that $\mathbb{Z}[i] = i\mathbb{Z}[i]$ so that $g_3(Z[i]) = g_3(i\mathbb{Z}[i]) = i^6 g_3(\mathbb{Z}[i]) = -g_3(\mathbb{Z}[i])$ and hence $g_3(\mathbb{Z}[i]) = 0$. This implies $j(i) = j(\mathbb{Z}[i]) = 1728$. Giving the elliptic curve as an equation $E\colon y^2 = x^3 + x$ we see that it has complex multiplication because we can define

$$[i](x, y) = (-x, iy).$$

Another interesting example is given by $y^2 = x^3 + 1$. Here we have

$$[\zeta^2](x, y) = (\zeta^2 x, y),$$

where $\zeta = e^{2\pi i/6}$. This curve is isomorphic to the one corresponding to the lattice $\mathbb{Z}[\zeta]$ and a similar argument as before shows $g_2(\mathbb{Z}[\zeta]) = 0$. This implies $j(\zeta) = j(\mathbb{Z}[\zeta]) = 0$.

So the j–invariants of the elliptic curves we have seen are both in $\mathbb{Q}$. This of course can not be for all elliptic curves but we might ask for which elliptic curves the j–invariant is in $\bar{\mathbb{Q}}$.

Definition. *A **singular modulus** is the j–invariant of an elliptic curve with complex multiplication.*

Now the question is, whether or not, having complex multiplication changes under isogeny. If we have two isogenous elliptic curves $E_1 \simeq \mathbb{C}/\Lambda_1$ and $E_1 \simeq \mathbb{C}/\Lambda_1$, then E_1 has complex multiplication if and only if E_2 has complex multiplication as the following argument shows. Let $\varphi\colon \mathbb{C}/\Lambda_1 \to \mathbb{C}/\Lambda_2$ be an isogeny between the elliptic curves given by $z+\Lambda_1 \mapsto mz+\Lambda_2$. Then there exists an isogeny $\hat{\varphi}\colon \mathbb{C}/\Lambda_2 \to \mathbb{C}/\Lambda_1$ called the *dual–isogeny* and can be constructed as follows. We have $\Lambda_1 \subseteq \frac{1}{m}\Lambda_2$ by the definition of the isogeny. Write $\Lambda_2 = \omega_1\mathbb{Z} + \omega_2\mathbb{Z}$. By the theory of finite abelian groups there are positive integers n_1 and n_2 such that $\Lambda_1 = \frac{n_1}{m}\omega_1\mathbb{Z} + \frac{n_2}{m}\omega_2\mathbb{Z}$. Then $\frac{n_1 n_2}{m}\Lambda_2 = \frac{n_1}{m}\omega_1 n_2\mathbb{Z} + \frac{n_2}{m}\omega_2 n_1\mathbb{Z} \subseteq \Lambda_1$, and the dual isogeny is defined by $z+\Lambda_2 \mapsto \frac{n_1 n_2}{m}z+\Lambda_1$. Now if E_1 has complex multiplication and α is a complex number representing a non–integer in $\mathrm{End}(E_1)$, then

$$m\alpha\frac{n_1 n_2}{m}\Lambda_2 \subseteq m\alpha\Lambda_1 \subseteq m\Lambda_1 \subseteq \Lambda_2.$$

Thus, the map $\alpha \mapsto m\alpha\frac{n_1 n_2}{m}$ defines a group homomorphism from $\mathrm{End}(E_1)$ to $\mathrm{End}(E_2)$. This map is injective since the composition of isogenies is again an isogeny, i.e. a non–zero map.

Theorem 1.10 *Singular moduli are algebraic integers. If Δ is attached to the endomorphism ring of the elliptic curve associated to a singular modulus, then the degree of the singular modulus is the class number $\mathcal{C}(\Delta)$. Moreover, if $Q_1, \ldots, Q_{\mathcal{C}(\Delta)}$ is a full set of representatives of reduced positive definite forms, then $j(\alpha_{Q_1}), \ldots, j(\alpha_{Q_{\mathcal{C}(\Delta)}})$ is a full orbit of Galois conjugates.*

PROOF. See for example [Sil99, Chapter II, Theorem 4.3]. □

In other words, the j–invariants of elliptic curves with complex multiplication are not only algebraic but they are algebraic integers. In fact we have the following converse. If ω_1/ω_2 is in the upper half–plane and an algebraic number of degree at least 3, then the j–invariant of the elliptic curve associated to the lattice $\omega_1\mathbb{Z} + \omega_2\mathbb{Z}$ is transcendental. A proof of this can be found in Corollaire 3.2.4 of [Wal79].

1.3.3 Heights: Part II

For more details about this section see [HS13] for example. In the first section we introduced heights on the projective space. We can consider the points of $E(K)$ as points in $\mathbb{P}^2_K$. There is also a *canonical height* associated to points on elliptic curves. For $0 \neq P \in E(\bar{K})$ we define

$$h_x(P) = h([x(P) : 1]),$$

where $x(P) \in K$ denotes the x–coordinate of P and $h_x(O) = 0$.

Definition. *The Néron–Tate height of a point P on an elliptic curve E/K is defined as*

$$\hat{h}(P) := \lim_{n\to\infty} \frac{h_x([2^n]P)}{4^n},$$

where $[2^n]$ denotes the multiplication–by–2^n isogeny on E. Note that the limit exists. See for example [Mil06].

One can show that x may be replaced by any nonconstant even function f in the function field $K(E)$ after dividing the limit by $\deg(f)$. Clearly, the height of the neutral element of E is 0.

This gives a height for the points on an elliptic curve. We can also define the height of an elliptic curve. One way to define the height for an elliptic curve E with j–invariant j is $h(j)$ since it is an invariant. This is called the *modular height* of E. Another definition was given by Faltings in [Fal83]. Again, this is actually the height of a point when we treat elliptic curves as points on *modular curves*. In his paper he defined the nowadays called Faltings height for abelian varieties. Elliptic curves are the simplest example of abelian varieties.

For an elliptic curve E/K and $\sigma\colon K \hookrightarrow \mathbb{C}$ we define E^σ by applying σ on the coefficients of E. Then E^σ is defined over $\sigma(K)$. If j is the j–invariant of E, then the j–invariant of E^σ is $\sigma(j)$. We choose $\tau_\sigma \in \mathcal{F}$ such that $\sigma(j) = j(\tau_\sigma)$. Recall that every infinite place of M_K^∞ gives rise to a $\sigma\colon K \hookrightarrow \mathbb{C}$.

The minimal discriminant of E/K is defined as

$$\mathfrak{D}_{E/K} = \prod_{\mathfrak{p} \in M_K \setminus M_K^\infty} \mathfrak{p}^{\operatorname{ord}_\mathfrak{p} \Delta_\mathfrak{p}},$$

where $\Delta_\mathfrak{p}$ is the minimal discriminant of E at $\mathfrak{p}$.

Definition. *Let E/K be an elliptic curve. We assume that E has everywhere semi–stable reduction over K. (Such a number field exists.) We define the (stable)* ***Faltings height*** *of E by*

$$h(E) := \frac{1}{12[K:\mathbb{Q}]} \left(\log |N_{K/\mathbb{Q}}(\mathfrak{D}_{E/K})| - \sum_{\sigma\colon K \hookrightarrow \mathbb{C}} \log \left(|\Delta(\tau_\sigma)| \operatorname{Im}(\tau_\sigma)^6 \right) \right) + \frac{1}{2} \log \pi.$$

The definition is independent of K as long as E has everywhere semi–stable reduction. Note that in Faltings' original definition the term $\frac{1}{2} \log \pi$ was not present. We use the symbol h instead of h_F since there should be no ambiguity in the text. The Faltings height of an abelian variety can be defined as follows. Let A/K be an abelian variety of dimension g defined over a field K that admits semi–stable reduction and let $\mathcal{A} \to \operatorname{Spec} \mathcal{O}_K$ be the Neŕon model, where $\mathcal{O}_K$ is the ring of integers of K. Let $s\colon \operatorname{Spec} \mathcal{O}_K \to \mathcal{A}$ be the zero section. We denote by $\omega_{\mathcal{A}/\mathcal{O}_K}$ the pullback $s^* \Omega^g_{\mathcal{A}/\operatorname{Spec} \mathcal{O}_K}$ of the sheaf of differential g–forms on $\mathcal{A}$. This can be made into a metrized line bundle $\bar{\omega}_{\mathcal{A}/\mathcal{O}_K}$ and the Faltings height $h(A)$ of A is defined as the normalized Arakelov degree of $\bar{\omega}_{\mathcal{A}/\mathcal{O}_K}$. Again, the original definition by Faltings is $h(A) - \frac{g}{2} \log \pi$.

2 The CM case

We fix a singular modulus j. Elliptic curves with j–invariant j have the same endomorphism ring. Let Δ denote the discriminant of this ring. We want to prove the following result.

Theorem 2.1 *Let j be a singular modulus and let Δ be its discriminant. Let α be an algebraic number that is the j–invariant of an elliptic curve without complex multiplication. If we assume that $j - \alpha$ is an algebraic unit, then $|\Delta|$ is bounded from above by*

$$|\Delta| \leq e^{15C}.$$

Thus there are only finitely many singular moduli j such that $j - \alpha$ is an algebraic unit. Here C is a computable constant and can be found on page 36.

2.1 Bounding points in the fundamental domain

Let $\mathcal{Q}(\Delta)$ be the set ($\subseteq \mathbb{Z}^3$) of coefficients representing reduced primitive, positive definite quadratic forms with discriminant Δ and let $\mathcal{C}(\Delta)$ be the class number defined in Chapter 1. We will write $\Delta = Df^2$ throughout this exposition, where D is the discriminant of the imaginary quadratic field $\mathbb{Q}(\sqrt{\Delta})$ and $f \in \mathbb{N}$ is called the *conductor*. For $\xi \in \mathcal{F}$ and $\varepsilon > 0$ we define

$$\mathcal{C}(\Delta; \xi; \varepsilon) = \# \left\{ (a,b,c) \in \mathcal{Q}(\Delta); \left| \frac{-b + \Delta^{1/2}}{2a} - \xi \right| < \varepsilon \right\}.$$

We also define the function F of Δ by

$$F = F(\Delta) = \max \left\{ 2^{\omega(a)}; a \leq |\Delta|^{1/2} \right\},$$

where $\omega(n)$ is the number of distinct prime divisors of n. We also define the modified conductor by

$$\tilde{f} = \begin{cases} f & D \equiv 1 \bmod 4, \\ 2f & D \equiv 0 \bmod 4. \end{cases}$$

Then $\Delta/\tilde{f}^2$ is square–free.

Let $\sigma_k(n) = \sum_{d|n} d^k$. We are now ready to state the first lemma that gives a bound on the τ in a neighborhood of a fixed point such that $j(\tau)$ is a singular modulus of fixed discriminant. While this is a generalization of Theorem 2.1 in [BHK18], the constants are not as good as in that very paper.

Lemma 2.2 *Let Δ be a negative integer, $y = \mathrm{Im}(\xi) \geq \sqrt{3}/2$ and $0 < \varepsilon < 1/4$. Then*

$$\mathcal{C}(\Delta;\xi;\varepsilon) \leq F(\Delta)\left(32\frac{\sigma_1\left(\tilde{f}\right)}{\tilde{f}}\frac{|\Delta|^{1/2}}{4y^2-1}\varepsilon^2 + 8\sigma_0\left(\tilde{f}\right)\left|\frac{\Delta}{3}\right|^{1/4}\varepsilon + 8\frac{|\Delta|^{1/2}}{4y^2-1}\varepsilon + 2\right)$$

Proof. We start with $|\tau-\xi| < \varepsilon$. This implies that the real and imaginary parts satisfy

$$\mathrm{Im}(\tau) \in (\mathrm{Im}(\xi)-\varepsilon, \mathrm{Im}(\xi)+\varepsilon)$$
$$\mathrm{Re}(\tau) \in (\mathrm{Re}(\xi)-\varepsilon, \mathrm{Re}(\xi)+\varepsilon).$$

Now τ is of the form $(-b+\sqrt{\Delta})/2a$ and thus $\mathrm{Im}(\tau) = |\Delta|^{1/2}/2a$ and $\mathrm{Re}(\tau) = -b/2a$. This amounts to

$$y - \varepsilon < \frac{|\Delta|^{1/2}}{2a} < y + \varepsilon$$

or equivalently

$$a \in \left(\frac{|\Delta|^{1/2}}{2y+2\varepsilon}, \frac{|\Delta|^{1/2}}{2y-2\varepsilon}\right) =: I.$$

For b we obtain

$$2a(\mathrm{Re}(\xi)-\varepsilon) < -b < 2a(\mathrm{Re}(\xi)+\varepsilon), \tag{2.1}$$

so b lies in an interval of length $4a\varepsilon$. For two integers m and n we denote by $\gcd_2(m,n)$ the greatest common divisor d of m and n such that $d^2|m$ and $d^2|n$. We have $\Delta = b^2-4ac$, so in particular $b^2 \equiv \Delta \bmod a$. Thus, the residue classes modulo $a/\gcd_2(a,\Delta)$ of $b \in \mathbb{Z}$ satisfying $b^2 \equiv \Delta \bmod a$ is at most $2^{\omega(a/\gcd(a,\Delta))+1}$ by Lemma 2.4 in [BHK18]. Note that we have $\omega(a/\gcd(a,\Delta)) \leq \omega(a)$. But b also lies in the interval given in equation (2.1), so that by Lemma 2.5 of [BHK18] there are at most

$$\left(\frac{2a(\mathrm{Re}(\xi)+\varepsilon) - 2a(\mathrm{Re}(\xi)-\varepsilon)}{a/\gcd_2(a,\Delta)} + 1\right)2^{\omega(a)+1} = (4\varepsilon\gcd_2(a,\Delta)+1)2^{\omega(a)+1} \tag{2.2}$$

possible b's for any fixed a. Recall that $a \leq |\Delta/3|^{1/2}$ by Lemma 1.1, so that $2^{\omega(a)} \leq F$. Using the equality in (2.2) and applying Lemma 2.6 of [BHK18] in the second inequality

we get

$$\begin{aligned}\mathcal{C}(\Delta;\xi;\varepsilon) &\le 8\varepsilon \sum_{a\in I\cap\mathbb{Z}} \gcd_2(a,\Delta)2^{\omega(a)} + 2\sum_{a\in I\cap\mathbb{Z}} 2^{\omega(a)} \\ &\le 8\varepsilon F \sum_{a\in I\cap\mathbb{Z}} \gcd_2(a,\Delta) + 2F\#(I\cap\mathbb{Z}) \\ &\le 8\varepsilon F \sum_{\substack{d^2|\Delta \\ d\le|\Delta/3|^{1/4}}} d\cdot\#(I\cap d^2\mathbb{Z}) + 2F\#(I\cap\mathbb{Z}).\end{aligned}$$

Here we used Lemma 1.1 in the last step again. But since $\Delta/\tilde{f}^2$ is square–free we obtain

$$\mathcal{C}(\Delta;\xi;\varepsilon) \le 8\varepsilon F \sum_{\substack{d|\tilde{f} \\ d\le|\Delta/3|^{1/4}}} d\left(\frac{|I|}{d^2}+1\right) + 2F\left(|I|+1\right),$$

where $|I|$ is the length of I. This can be further simplified to

$$\begin{aligned}\mathcal{C}(\Delta;\xi;\varepsilon) &\le 8\varepsilon F|I| \sum_{\substack{d|\tilde{f} \\ d\le|\Delta/3|^{1/4}}} d^{-1} + 8\varepsilon F \sum_{\substack{d|\tilde{f} \\ d\le|\Delta/3|^{1/4}}} d + 2F\left(|I|+1\right) \\ &\le 8\varepsilon F|I|\frac{\sigma_1(\tilde{f})}{\tilde{f}} + 8\varepsilon F\left|\frac{\Delta}{3}\right|^{1/4}\sigma_0(\tilde{f}) + 2F\left(|I|+1\right).\end{aligned}$$

The length of I can be estimated by

$$\frac{|\Delta|^{1/2}}{2y-2\varepsilon} - \frac{|\Delta|^{1/2}}{2y+2\varepsilon} = |\Delta|^{1/2}\frac{2y+2\varepsilon-(2y-2\varepsilon)}{4y^2-4\varepsilon^2} \le |\Delta|^{1/2}\frac{4\varepsilon}{4y^2-1}.$$

This gives the desired inequality. □

The next corollary gives a bound on $\mathcal{C}(\Delta;\xi;\varepsilon)$ just in terms of Δ and ε.

Corollary 2.3 *For $|\Delta| \ge 10^{14}$ and $0 < \varepsilon < 1/4$ we have*

$$\mathcal{C}(\Delta;\xi;\varepsilon) \le F(\Delta)\left(32|\Delta|^{1/2}\varepsilon^2\log\log(|\Delta|^{1/2}) + 11|\Delta|^{1/2}\varepsilon + 2\right).$$

Proof. For $|\Delta| \ge 10^{14}$ we can find the following results as Lemma 2.8 in [BHK18]

$$\begin{aligned}\sigma_0(\tilde{f}) &\le |\Delta|^{0.192} \le |\Delta|^{1/4} \\ \sigma_1(\tilde{f})/\tilde{f} &\le 1.842\log\log(|\Delta|^{1/2}).\end{aligned}$$

Moreover, we have $y \geq \sqrt{3}/2$ and thus $4y^2 - 1 \geq 2$. Hence

$$\frac{\sigma_1\left(\tilde{f}\right)}{\tilde{f}} \frac{|\Delta|^{1/2}}{4y^2-1}\varepsilon^2 \leq |\Delta|^{1/2}\varepsilon^2 \log\log(|\Delta|^{1/2})$$

and

$$\begin{aligned} 8\sigma_0\left(\tilde{f}\right)\left|\frac{\Delta}{3}\right|^{1/4}\varepsilon + 8\frac{|\Delta|^{1/2}}{4y^2-1}\varepsilon &\leq \frac{8}{3^{1/4}}\,|\Delta|^{1/2}\,\varepsilon + 4|\Delta|^{1/2}\varepsilon \\ &\leq 7\,|\Delta|^{1/2}\,\varepsilon + 4|\Delta|^{1/2}\varepsilon, \end{aligned}$$

which gives the claimed statement. □

2.2 Height bounds

From now on α will be the j–invariant of an elliptic curve without complex multiplication. As a preparation we will start with some properties of the j–function.

Lemma 2.4 *The function $j(1/2+iy)$ as a function of y on the interval $[\sqrt{3}/2, \infty)$ is real and decreasing. The function $j(e^{i\theta})$ on the interval $[\pi/3, \pi/2]$ is real and increasing, and we have $j(e^{i\pi/2}) = j(i) = 1728$. The function $j(iy)$ on the interval $[1, \infty)$ is real and increasing.*

Proof. Recall $q = e^{2\pi i\tau}$. For $\tau = \frac{1}{2} + iy$ with $y \geq \sqrt{3}/2$ we have $q = e^{\pi i}e^{-2\pi y} = -e^{-2\pi y}$. Thus $j(\tau)$ is real since all non–zero coefficients of j are positive integers. We have $j(1/2 + i\sqrt{3}/2) = 0$ and from page 227 of [Cox11] we know $j(1/2 + \sqrt{-7}/2) = -15^3$. But the map $y \mapsto j(1/2 + iy)$ is continuous and injective because j is continuous and injective as a function on $\mathcal{F}$. Thus, it is monotonically decreasing.

Similarly, if $\tau = iy$ with $y \geq 1$, then $q = e^{-2\pi y}$. We know $j(i) = 1728 = 12^3$ and again from page 227 of [Cox11] we know $j(i\sqrt{2}) = 20^3$. The same argument as before shows the claim for the map $y \mapsto j(iy)$.

It remains to show that $j(e^{i\theta})$ is real because in that case $j(e^{i\pi/3}) = 0$ and $j(e^{i\pi/2}) = 1728$ imply the monotonicity. Write $\tau = e^{i\theta}$. We have $\bar{q} = e^{2\pi i(-\bar{\tau})}$ and

$$\overline{j(\tau)} = (\bar{q})^{-1} + \sum_{n=0}^{\infty} c_n(\bar{q})^n = j(-\bar{\tau}).$$

But j is $\mathrm{SL}_2(\mathbb{Z})$–invariant so that $\overline{j(\tau)} = j(-\bar{\tau}) = j(\tau)$ since $|\tau| = 1$. Therefore, $j(\tau)$ must be real. This completes the proof. □

The next two statements tell us something about the growth of $j(\tau)$ as $|\tau|$ goes to infinity.

Proposition 2.5 *If τ is in $\mathcal{F}$, then $\big||j(\tau)| - e^{2\pi \operatorname{Im}(\tau)}\big| \leq 2079$.*

This result can be found in Lemma 1 of [BMZ13]. We are going to prove the following result, which is of similar nature.

Lemma 2.6 *Let τ be complex with $\operatorname{Im}(\tau) \geq 1/2$. Then*

$$\big||j(\tau)| - e^{2\pi \operatorname{Im}(\tau)}\big| \leq 287473.$$

Proof. We have $j(\tau) = q^{-1} + c_0 + c_1 q + \cdots$, where as usual $q = e^{2\pi i \tau}$. Recall from Chapter 1 that the coefficients of the q–expansion of j are all non–negative integers. Then $||j| - |q^{-1}|| \leq \sum_{n=0}^{\infty} c_n |q|^n \leq \sum_{n=0}^{\infty} c_n q_0^n$ with $q_0 = e^{2\pi i \tau_0} = e^{-\pi}$ and $\tau_0 = i/2$. The right–hand side of the inequality is equal to $j(\tau_0) - q_0^{-1} = 66^3 - e^{\pi} \leq 287473$. Note that we have used $j(\tau_0) = j(-1/\tau_0)$ since the j–function is $\mathrm{SL}_2(\mathbb{Z})$–invariant, and that $j(-1/\tau_0) = j(2i) = 66^3$ by Table 12.20 in [Cox11]. □

In Lemma 2.4 we proved that the j–function is real on the vertical and unit circle geodesics of the fundamental domain and on the imaginary axis. We can even say that the j–function is not real outside of this set, as the following statement shows.

Corollary 2.7 *If $\tau \in \mathcal{F}$ with $\operatorname{Re}(\tau) \neq 0, \pm\frac{1}{2}$ and $|\tau| > 1$, then $\operatorname{Im}(j(\tau)) \neq 0$. Moreover, $\operatorname{Im}(j(\tau)) < 0$ for $0 < \operatorname{Re}(\tau) < 1/2$ and $\operatorname{Im}(j(\tau)) > 0$ for $-1/2 < \operatorname{Re}(\tau) < 0$.*

Proof. The proof is just an application of the intermediate value theorem. We use that j is injective on $\mathcal{F}$. Assume $j(\tau) = R$ real with $|\tau| > 1$ and $-1/2 < \operatorname{Re}(\tau) < 0$ or $0 < \operatorname{Re}(\tau) < 1/2$. If $0 \leq R \leq 1728$, then $j(e^{i\theta}) = R$ for some $\pi/3 \leq \theta \leq \pi/2$ by the intermediate value theorem applied to the real function $t \mapsto j(e^{it})$. This is a contradiction to the injectivity of j on $\mathcal{F}$ since $|\tau| > 1$.

Assume $R \geq 1728$. By Lemma 2.4 $j(iR) \geq 1728$ and applying Proposition 2.5 we have $j(iR) \geq e^{2\pi R} - 2079$. Thus $j(iR) \geq R$ and applying the intermediate value theorem again gives a $t \geq 1$ with $j(it) = R$. This is a contradiction since $0 < \tau < 1/2$. The case when $R < 0$ follows similarly.

To show $\operatorname{Im}(j(\tau)) < 0$ for $0 < \operatorname{Re}(\tau) < 1/2$ and $|\tau| > 1$ we assume we have τ_0, τ_1 in the interior of the fundamental domain $\mathcal{F}^\circ$ with positive real part and such that $\operatorname{Im}(j(\tau_0)) < 0$ and $\operatorname{Im}(j(\tau_1)) > 0$. Choose a path in $\mathcal{F}^\circ$, parametrized by $\gamma\colon [0,1] \to \mathcal{F}$, such that $\gamma(0) = \tau_0$, $\gamma(1) = \tau_1$ and such that every point in $\gamma([0,1])$ is in the interior of $\mathcal{F}$ and has positive real part. The function $t \mapsto \operatorname{Im}(j(\gamma(t)))$ is continuous and satisfies $\operatorname{Im}(j(\gamma(0))) = \operatorname{Im}(j(\tau_0)) < 0$ and $\operatorname{Im}(j(\gamma(1))) = \operatorname{Im}(j(\tau_1)) > 0$. By the intermediate value theorem we have a $0 < t < 1$ with $\operatorname{Im}(j(\gamma(t))) = 0$ which is impossible by the

choice of γ and the first claim of the corollary. So it suffices to give a value of $j(\tau)$ with $0 < \operatorname{Re}(\tau) < 1/2$ and $\operatorname{Im}(j(\tau)) < 0$. We have

$$j\left(\frac{1+5i}{4}\right) = -1728\left(\sqrt{5}-2\right)^{20}\left(3-2\sqrt{\sqrt{5}i}\right)^{6}\left(238\sqrt{5}+\frac{861}{2}+60\sqrt{\sqrt{5}i}\right)^{3}$$

by page 17 of [Adl14]. A computation with Sage shows $\operatorname{Im}\left(j\left(\frac{1+5i}{4}\right)\right) < 0$. We must have $\operatorname{Im}(j(\tau)) > 0$ for $-1/2 < \operatorname{Re}(\tau) < 0$ by the same argument and the fact that $j\colon \mathcal{F} \to \mathbb{C}$ is surjective. This completes the proof. □

The following two lemmas for $h(j)$ can be found in [BHK18]. The proofs follow directly from the statements in that very paper with the inequality $h(j-\alpha) \geq h(j) - h(\alpha) - \log 2$. For details see Proposition 4.1 and Proposition 4.3 in [BHK18].

Lemma 2.8 *We have* $[\mathbb{Q}(j) : \mathbb{Q}] = \mathcal{C}(\Delta)$*, and if* $|\Delta| \geq 16$*, then*

$$h(j-\alpha) \geq \frac{\pi|\Delta|^{1/2} - 0.01}{\mathcal{C}(\Delta)} - h(\alpha) - \log 2.$$

Proof. The first statement is a classical result, see for example Chapter 13 in [Cox11]. By Theorem 1.10 one of the conjugates of j is $j(\tau)$ with $\tau = \frac{\Delta + i\sqrt{|\Delta|}}{2}$. So one of the terms in the height of $h(j)$ is $\log\max\{1, |j(\tau)|\}$. We can forget about the other terms and get

$$h(j) \geq \frac{1}{\mathcal{C}(\Delta)} \log\max\{1, |j(\tau)|\}.$$

Proposition 2.5 implies $|j(z)| \geq -2079 + e^{2\pi \operatorname{Im}(z)} \geq 0.992 e^{2\pi \operatorname{Im}(z)}$ for all $z \in \mathcal{F}$ with $\operatorname{Im}(z) \geq 2$. Since $|\Delta| \geq 16$ this is the case for τ, and thus

$$\begin{aligned} h(j) &\geq \frac{1}{\mathcal{C}(\Delta)} \log\max\{1, |j(\tau)|\} \geq \frac{1}{\mathcal{C}(\Delta)} \log\left(0.992 e^{\pi\sqrt{|\Delta|}}\right) \\ &\geq \frac{\pi|\Delta|^{1/2} - 0.01}{\mathcal{C}(\Delta)}. \end{aligned}$$

The claim now follows from Lemma 1.4 since $h(j) = h(j - \alpha + \alpha) \leq h(j-\alpha) + h(\alpha) + \log 2$. □

Lemma 2.9 *We have*

$$h(j-\alpha) \geq \frac{3}{\sqrt{5}} \log|\Delta| - 9.79 - h(\alpha) - \log 2.$$

This lemma is more delicate and follows from work of Colmez [Col98] and Nakkajima–Taguchi [NT91]. The two previous lemmas bound the height of $j - \alpha$ from below. Next

we want to bound the height of $j - \alpha$ from above when $j - \alpha$ is an algebraic unit.

The following lemma says that if two points in the fundamental domain are close together, then the difference of the images under the j–function can be bound from below in terms of the difference of the points. Recall that $\zeta = e^{2\pi i/6}$.

Lemma 2.10 *Let $\zeta, \zeta^2 \neq \xi \in \bar{\mathcal{F}}$. Put $B = 4 \cdot 10^5 \max\{1, |j(\xi)|\}$ and $A = |j''(i)|$ in the case when $\xi = i$ and $A = |j'(\xi)|$ otherwise. For $|\tau - \xi| \leq \frac{A}{12A+108B} \leq \frac{1}{3}$ we have*

$$|j(\tau) - j(\xi)| \geq \frac{A}{4}|\tau - \xi|^2.$$

If $\xi \neq i$ we even have

$$|j(\tau) - j(\xi)| \geq \frac{A}{2}|\tau - \xi|$$

for $|\tau - \xi| \leq \frac{A}{6A+18B}$.

Note that we can write $j''(i) = -2 \cdot 3^4 \Gamma(1/4)^8/\pi^4$ as Kühne shows in the appendix of [Wüs14]. Since $\Gamma(1/4) = 3.6256\ldots > 3$ we could further estimate the first of the two bounds in the lemma by

$$|j(\tau) - 1728| \geq 12413|\tau - i|^2.$$

Proof. This is a special case of Lemma 2.4 in [BLP16]. We take $f(\tau) = j(\tau) - j(\xi)$. Assume $|\tau - \xi| \leq 1/3$. Then by Lemma 2.6

$$\begin{aligned} |f(\tau)| &\leq |j(\tau)| + |j(\xi)| \leq |j(\xi)| + e^{2\pi \operatorname{Im}(\tau)} + 287473 \\ &\leq |j(\xi)| + e^{2\pi(\operatorname{Im}(\xi)+1/3)} + 287473. \end{aligned}$$

We have $e^{2\pi/3} < 9$ so applying Proposition 2.5 we obtain

$$\begin{aligned} |f(\tau)| &\leq |j(\xi)| + e^{2\pi(\operatorname{Im}(\xi)+1/3)} + 287473 \\ &\leq |j(\xi)| + 9|j(\xi)| + 9 \cdot 2079 + 287473 \\ &\leq 10|j(\xi)| + 306184 \\ &\leq 4 \cdot 10^5 \max\{1, |j(\xi)|\}. \end{aligned}$$

We treat the two cases $\xi = i$ and $\xi \neq i$ separately and start with the latter. By [BLP16] we have

$$|j(\tau) - j(\xi)| - A|\tau - \xi| \geq -\frac{A/3 + B}{(1/3)^2}|\tau - \xi|^2 = -(3A + 9B)|\tau - \xi|^2,$$

where $A = |j'(\xi)|$. In the smaller disc $|\tau - \xi| \leq \frac{A}{2\cdot(3A+9B)}$ we then obtain

$$|j(\tau) - j(\xi)| \geq \frac{A}{2}|\tau - \xi|.$$

Now assume $\xi = i$ and put $A = |j''(i)|$. By [BLP16] we conclude

$$|j(\tau) - j(\xi)| - \frac{A}{2}|\tau - \xi|^2 \geq -\frac{A/9 + B}{(1/3)^3}|\tau - \xi|^3 = -(3A + 27B)\,|\tau - \xi|^3$$

and thus

$$|j(\tau) - j(\xi)| \geq \frac{A}{4}|\tau - \xi|^2.$$

in the smaller disc $|\tau - \xi| \leq \frac{A}{4\cdot(3A+27B)}$. □

We can also bound the j–invariants outside of a neighborhood of ξ as the following lemma shows. We write

$$\mathcal{F}_+ = \{\tau \in \mathcal{F}; 0 \leq \mathrm{Re}(\tau) \leq 1/2\}$$

and

$$\mathcal{F}_- = \{\tau \in \mathcal{F}; -1/2 \leq \mathrm{Re}(\tau) \leq 0\}.$$

Lemma 2.11 *Let $\zeta \neq \xi, \tau \in \mathcal{F}_+$ and put $A = |j''(i)|$ if $\xi = i$ and $A = |j'(\xi)|$ otherwise. Also define $B = 4 \cdot 10^5 \max\{1, |j(\xi)|\}$. Assume $|\tau - \xi| \geq \delta$, where δ is defined as the minimum of $\frac{A}{12A+108B}$ and half the (euclidean) distance of ξ to any geodesic of $\partial\mathcal{F}_+$ (i.e. the vertical line segments with real part 0 and $1/2$ and the part of the unit circle between those lines) not containing ξ. Then*

$$|j(\tau) - j(\xi)| \geq c(\xi)$$

where $c(\xi) > 0$ is an absolute constant depending on ξ and is given by

$$c(\xi) = \begin{cases} A\delta/2 & \text{if } \xi \in \partial\mathcal{F}_+ \setminus \{i\} \\ A\delta^2/4 & \text{if } \xi = i \\ \min\{|\,\mathrm{Im}(j(\xi))|, A\delta/2\} & \text{otherwise.} \end{cases}$$

Proof. We want to apply the maximum modulus principle. To do this we will give lower bounds on the boundary of $\mathcal{F}_+$, see Figure 2.1. By the previous lemma we have

$$|j(\tau) - j(\xi)| \geq \frac{A}{2}\delta$$

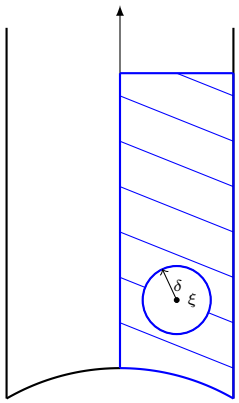


Figure 2.1: Application of the maximum modulus principle to the blue area.

or

$$|j(\tau) - j(i)| \geq \frac{A}{4}\delta^2$$

on the circle $|\tau - \xi| = \delta$. Also $\operatorname{Im}(\tau) \geq 6\operatorname{Im}(\xi)$ implies $\operatorname{Im}(\tau) \geq \operatorname{Im}(\xi) + 1$, so we obtain by applying Proposition 2.5 twice

$$\begin{aligned}
|j(\xi)| &\leq 2079 + e^{2\pi \operatorname{Im}(\xi)} = 2079 - 20e^{2\pi\frac{\sqrt{3}}{2}} + 20e^{2\pi\frac{\sqrt{3}}{2}} + e^{2\pi \operatorname{Im}(\xi)} \\
&< -2079 + 20e^{2\pi\frac{\sqrt{3}}{2}} + e^{2\pi \operatorname{Im}(\xi)} \leq -2079 + 20e^{2\pi \operatorname{Im}(\xi)} + e^{2\pi \operatorname{Im}(\xi)} \\
&\leq -2079 + 21e^{2\pi \operatorname{Im}(\xi)} < -2079 + e^{2\pi}e^{2\pi \operatorname{Im}(\xi)} \\
&\leq -2079 + e^{2\pi \operatorname{Im}(\tau)} \\
&\leq |j(\tau)|.
\end{aligned}$$

Thus, by using Proposition 2.5 twice we get

$$\begin{aligned}
|j(\tau) - j(\xi)| &\geq |j(\tau)| - |j(\xi)| \\
&\geq -2079 + e^{2\pi \operatorname{Im}(\tau)} - (2079 + e^{2\pi \operatorname{Im}(\xi)}) \\
&\geq -2 \cdot 2079 + e^{2\pi \operatorname{Im}(\tau)} - e^{2\pi \operatorname{Im}(\xi)}.
\end{aligned}$$

Now we use $\operatorname{Im}(\tau) \geq 6\operatorname{Im}(\xi)$ to get

$$|j(\tau) - j(\xi)| \geq -2 \cdot 2079 + e^{12\pi \operatorname{Im}(\xi)} - e^{2\pi \operatorname{Im}(\xi)}.$$

We have $e^{2x} \geq 2e^x$ for any $x \geq \log 2$ and $2\pi \operatorname{Im}(\xi) \geq 1$ so that

$$\begin{aligned} |j(\tau) - j(\xi)| &\geq -2 \cdot 2079 + e^{4\pi \operatorname{Im}(\xi)} + e^{2\pi \operatorname{Im}(\xi)} + e^{4\pi \operatorname{Im}(\xi)} - e^{2\pi \operatorname{Im}(\xi)} \\ &= -2 \cdot 2079 + e^{4\pi \operatorname{Im}(\xi)} + e^{4\pi \operatorname{Im}(\xi)}. \end{aligned}$$

But also because of $\operatorname{Im}(\xi) \geq \sqrt{3}/2$ we get

$$|j(\tau) - j(\xi)| \geq e^{4\pi \operatorname{Im}(\xi)}. \tag{2.3}$$

We have $\delta \leq 1/12$ by definition and Lemme 1 of [FP87] gives for $\xi \neq i$

$$\begin{aligned} \frac{A}{2}\delta \leq \frac{A}{24} &\leq \frac{8\pi}{24} e^{2(\pi+1)\max\{\operatorname{Im}(\xi), \operatorname{Im}(\xi)^{-1}\}} \\ &\leq \frac{8\pi}{24} e^{3\pi \max\{\operatorname{Im}(\xi), \operatorname{Im}(\xi)^{-1}\}}. \end{aligned}$$

So if $\operatorname{Im}(\xi) \geq 1$, then

$$|j(\tau) - j(\xi)| \geq \frac{A}{2}\delta.$$

If $\sqrt{3}/2 \leq \operatorname{Im}(\xi) \leq 1$, then

$$\frac{A}{2}\delta \leq \frac{8\pi}{24} e^{3\pi \operatorname{Im}(\xi)^{-1}} \leq \frac{8\pi}{24} e^{2\pi/\sqrt{3}} \leq e^{4\pi \cdot \sqrt{3}/2} \leq e^{4\pi \operatorname{Im}(\xi)}.$$

So in any case

$$|j(\tau) - j(\xi)| \geq \frac{A}{2}\delta$$

for $\xi \neq i$. If $\xi = i$, then as described after Lemma 2.10 we get

$$\frac{A}{4}\delta^2 \leq 12414\delta^2 \leq 87$$

which together with (2.3) gives

$$|j(\tau) - j(i)| \geq \frac{A}{4}\delta^2.$$

So we have treated all the τ with large imaginary part. Now we have to go through the different cases for ξ to bound the boundary. We start with $\operatorname{Re}(\xi) = 1/2$ so that we are in the case of Figure 2.2.

For the remainder of the proof we will be using that j is monotonically increasing or decreasing on the boundary as shown in Lemma 2.4. If τ is on the same boundary component as ξ, then $\operatorname{Im}(\tau) \geq \operatorname{Im}(\xi) + \delta$ or $\operatorname{Im}(\tau) \leq \operatorname{Im}(\xi) - \delta$. Therefore, by monotonicity

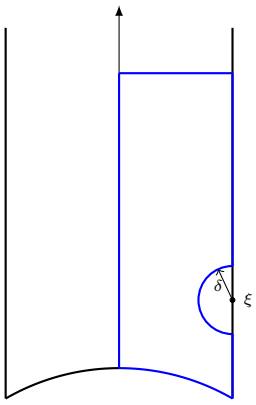


Figure 2.2: ξ with real part $1/2$.

either $|j(\tau)| > |j(\xi)|$ which implies

$$|j(\tau) - j(\xi)| \geq j(\xi) - j(\tau) \geq j(\xi) - j(\xi + i\delta) = |j(\xi) - j(\xi + i\delta)| \geq \frac{A}{2}\delta$$

or $|j(\tau)| < |j(\xi)|$ and

$$|j(\tau) - j(\xi)| \geq |j(\xi)| - |j(\tau)| \geq j(\xi - i\delta) - j(\xi) = |j(\xi) - j(\xi - i\delta)| \geq \frac{A}{2}\delta.$$

Note that the last inequality of both displays follows from Lemma 2.10 on the boundary $|\tau - \xi| = \delta$. If $|\tau| = 1$ or $\mathrm{Re}(\tau) = 0$, then $j(\tau) \geq 0$ and $j(\xi) < 0$, so that

$$\begin{aligned} |j(\tau) - j(\xi)| &= j(\tau) - j(\xi) \geq -j(\xi) \\ &\geq j(\xi - i\delta) - j(\xi) = |j(\xi) - j(\xi - i\delta)| \\ &\geq \frac{A}{2}\delta. \end{aligned}$$

Here we have used Lemma 2.10 for the last estimate. Altogether, if $\mathrm{Re}(\xi) = 1/2$ we get by the minimum modulus principle

$$|j(\tau) - j(\xi)| \geq \frac{A}{2}\delta$$

as desired.

The second case is when $\mathrm{Re}(\xi) = 0$ and $\xi \neq i$. We have $j(\tau) \leq 0$ if $\mathrm{Re}(\tau) = 1/2$, and

thus

$$\begin{aligned} |j(\xi) - j(\tau)| &\geq j(\xi) \geq j(\xi) - j(\xi - i\delta) \\ &= |j(\xi) - j(\xi - i\delta)| \\ &\geq \frac{A}{2}\delta. \end{aligned}$$

If $|\tau| = 1$, then again by monotonicity and Lemma 2.10

$$\begin{aligned} |j(\xi) - j(\tau)| &\geq j(\xi) - j(\tau) \geq j(\xi) - j(\xi - i\delta) \\ &= |j(\xi) - j(\xi - i\delta)| \\ &\geq \frac{A}{2}\delta. \end{aligned}$$

For the case when $\mathrm{Re}(\tau) = 0$ we have two subcases. When $\mathrm{Im}(\tau) \geq \mathrm{Im}(\xi) + \delta$, it follows $j(\tau) > j(\xi)$ and hence

$$|j(\xi) - j(\tau)| \geq j(\xi + i\delta) - j(\xi) \geq \frac{A}{2}\delta. \tag{2.4}$$

Or we have $\mathrm{Im}(\tau) \leq \mathrm{Im}(\xi) - \delta$ and we get

$$|j(\xi) - j(\tau)| \geq j(\xi) - j(\xi - i\delta) \geq \frac{A}{2}\delta.$$

In sum, by applying the minimum modulus principle we obtain

$$|j(\tau) - j(\xi)| \geq \frac{A}{2}\delta. \tag{2.5}$$

The third case is $|\xi| = 1$ and $\xi \neq i$. Write $\xi = e^{i\theta}$. Again we have three subcases. If $\mathrm{Re}(\tau) = 1/2$, then $j(\tau) \leq 0$ and

$$\begin{aligned} |j(\xi) - j(\tau)| &\geq j(\xi) - j(\tau) \geq j(\xi) \\ &\geq j(\xi) - j\left(e^{i(\theta - 2\arcsin(\delta/2))}\right) \\ &= |j(\xi) - j\left(e^{i(\theta - 2\arcsin(\delta/2))}\right)| \\ &\geq \frac{A}{2}\delta. \end{aligned} \tag{2.6}$$

Note that $e^{i(\theta - 2\arcsin(\delta/2))}$ is one of the two points where the circle of radius δ and the

unit circle intersect. If $\mathrm{Re}(\tau) = 0$, then $j(\tau) \geq 1728 > j\left(e^{i(\theta+2\arcsin(\delta/2))}\right) > j(\xi)$ and

$$\begin{aligned} |j(\xi) - j(\tau)| &\geq j(\tau) - j(\xi) \geq j\left(e^{i(\theta+2\arcsin(\delta/2))}\right) - j(\xi) \\ &= |j(\xi) - j\left(e^{i(\theta+2\arcsin(\delta/2))}\right)| \\ &\geq \frac{A}{2}\delta. \end{aligned}$$

If $|\tau| = 1$, then $j(\tau) < j(\xi)$ or $j(\tau) > j(\xi)$ and Lemma 2.10 tells us

$$|j(\xi) - j(\tau)| \geq \left|j(\xi) - j\left(e^{i(\theta\pm 2\arcsin(\delta/2))}\right)\right| \geq \frac{A}{2}\delta. \tag{2.7}$$

By applying the minimum modulus principle we obtain the same result as in equation (2.5).

If $\xi = i$, then the estimates in equations (2.4) and (2.7) hold with $A\delta/2$ replaced by $A\delta^2/4$. Also equation (2.6) holds with ξ replaced by i and $A\delta/2$ replaced by $A\delta^2/4$, and thus by the minimum modulus principle

$$|j(\tau) - j(i)| \geq \frac{A}{4}\delta^2.$$

The last case is $0 < \mathrm{Re}(\xi) < 1/2$ and $|\xi| > 1$. Let $\tau \in \partial\mathcal{F}_+$. Then $j(\tau)$ is real and we have

$$|j(\xi) - j(\tau)| \geq |\,\mathrm{Im}(j(\xi)) - \mathrm{Im}(j(\tau))| = |\,\mathrm{Im}(j(\xi))|.$$

This is the case shown in Figure 2.1. Applying the minimum modulus principle gives the desired result. □

Note that the same claim holds for $\mathcal{F}_-$ since we have the symmetry $\mathrm{Re}(j(x+iy)) = \mathrm{Re}(j(-x+iy))$ and $\mathrm{Im}(j(x+iy)) = -\mathrm{Im}(j(-x+iy))$ which directly follows from the q–expansion. If $j(\tau)$ and $j(\xi)$ are close we want $|\tau_\sigma - \xi_\sigma|$ to be small too, but we can not get this in general as Figure 2.3 shows.

Lemma 2.12 *In the same setting as in the previous lemma, if $|j(\tau) - j(\xi)| < c(\xi)$, then $|\tau - M\xi| < \delta$ with $M\xi \in \bar{\mathcal{F}}$ for some $M \in \mathcal{T}$ where $\mathcal{T} = \{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&\pm1\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)\}$.*

Proof. If ξ, τ are both in $\mathcal{F}_+$ or both in $\mathcal{F}_-$, then $|\tau - \xi| \leq \delta$ by Lemma 2.11. We now can assume without loss of generality $\mathrm{Re}(\xi) < 0$ and $\mathrm{Re}(\tau) > 0$. If ξ is on the boundary of $\mathcal{F}$, then $\mathrm{Re}(\xi) = -1/2$. Then we can apply Lemma 2.11 to τ and $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)\xi$ and use $j(\xi) = j(\xi+1)$. If $|\xi| = 1$ and $\mathrm{Re}(\xi) < 0$, then we can again apply Lemma 2.11 to τ and $\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)\xi$ and use $j(\xi) = j\left(\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)\xi\right)$. If $-1/2 < \mathrm{Re}(\xi) < 0$ and $\tau \in \mathcal{F}_+$, then by

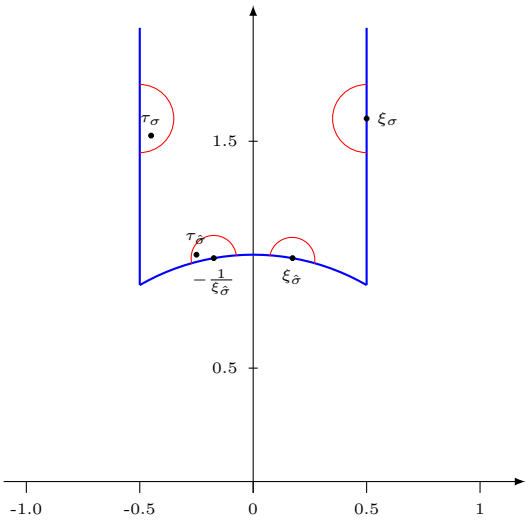


Figure 2.3: Neighborhoods of points on the boundary

Corollary 2.7

$$\begin{aligned}|j(\xi) - j(\tau)| &\geq |\operatorname{Im}(j(\xi)) - \operatorname{Im}(j(\tau))| = \operatorname{Im}(j(\xi)) - \operatorname{Im}(j(\tau)) \\ &= \operatorname{Im}(j(\xi)) + |\operatorname{Im}(j(\tau))| \\ &\geq \operatorname{Im}(j(\xi)) \\ &\geq c(\xi)\end{aligned}$$

contrary to the assumption. □

The following lemma can be found in [Hab15] as Lemma 5.

Lemma 2.13 *Let $\sigma\colon \bar{\mathbb{Q}} \hookrightarrow \mathbb{C}$ and let τ be imaginary quadratic. Let τ_σ satisfy $j(\tau_\sigma) = \sigma(j(\tau))$ with $\tau_\sigma \in \mathcal{F}$. If Δ is the discriminant of the endomorphism ring associated to $j(\tau)$, then τ_σ is imaginary quadratic and $h(\tau_\sigma) \leq \log \sqrt{|\Delta|}$.*

Proof. Write $\Delta = Df^2$ where f is the conductor and D a fundamental discriminant. The endomorphism ring of $j(\tau)$ can be described by $\mathbb{Z} + \omega f\mathbb{Z}$ with $\omega = (D + \sqrt{D})/2$. Thus ωf defines an endomorphism, i.e. $\omega f(\mathbb{Z} + \tau_\sigma\mathbb{Z}) \subseteq \mathbb{Z} + \tau_\sigma\mathbb{Z}$. Hence $\omega f = a + b\tau_\sigma$ and $\omega f\tau_\sigma = c + d\tau_\sigma$ with $a, b, c, d \in \mathbb{Z}$ and $b \neq 0$. Replacing ωf in the second equality gives

$$b\tau_\sigma^2 + (a - d)\tau_\sigma - c = 0,$$

so τ_σ is imaginary quadratic. The number ωf satisfies the polynomial equation $X^2 - (a + d)X + ad - bc = 0$. This polynomial has the same discriminant as the polynomial

$P = bX^2 + (a-d)X - c$. If we compute the discriminant of $X^2 - (a+d)X + ad - bc$ we get $(a+d)^2 - 4(ad-bc) = (\omega f - \overline{\omega f})^2 = (\omega - \bar{\omega})^2 f^2 = \Delta$. Since $P(\tau_\sigma) = 0$ we must have

$$\tau_\sigma = \frac{-(a-d) \pm i\sqrt{|\Delta|}}{2b}.$$

This implies $|\operatorname{Re}(\tau_\sigma)| \leq 1/2$ and thus $|a-d| \leq |b|$. Then $|\tau_\sigma|^2 = ((a-d)^2 + |\Delta|)/(2b)^2 \leq (b^2 + |\Delta|)/(2b)^2$. Proposition 1.6.6 in [BG07] says that $2h(\tau_\sigma)$ is at most the logarithmic Mahler measure of P, i.e.

$$2h(\tau_\sigma) \leq \log(|b||\tau_\sigma|^2) \leq \log\left(\frac{|b|}{4} + \frac{|\Delta|}{4|b|}\right).$$

We have $\operatorname{Im}(\tau_\sigma) = \sqrt{|\Delta|}/(2|b|) \geq \sqrt{3}/2$ because $\tau_\sigma \in \mathcal{F}$. Thus $|b| \leq \sqrt{|\Delta|/3} \leq |\Delta|$. This implies

$$2h(\tau_\sigma) \leq \log\left(\frac{|\Delta|}{4} + \frac{|\Delta|}{4|b|}\right) \leq \log\left(\frac{|\Delta|}{4} + \frac{|\Delta|}{4}\right) \leq \log|\Delta|.$$

This is the desired inequality. □

With these lemmas we are now able to bound the height from above using linear forms in logarithms. Specifically we are going to use a result by David Masser. For the following, if $\tau \in \mathcal{F}$ with $j(\tau)$ algebraic and a field embedding $\sigma\colon \mathbb{Q}(\alpha) \hookrightarrow \mathbb{C}$, $\alpha \in \bar{\mathbb{Q}}$, are given, then $\tau_\sigma \in \mathcal{F}$ is defined by $\sigma(j(\tau)) = j(\tau_\sigma)$. Note that for fixed $\xi \in \mathcal{F}$ different from ζ, ζ^2, i we have $j'(\xi_\sigma) \neq 0$. Suppose that $j(\xi) = \alpha$ is algebraic. We define the function

$$\mathcal{P}(\xi) = \log \max_\sigma \left\{1, c(\xi_\sigma)^{-1}\right\}, \tag{2.8}$$

where σ runs over all embeddings $\sigma\colon \mathbb{Q}(\alpha) \hookrightarrow \mathbb{C}$, and $c(\xi_\sigma)$ is defined as in Lemma 2.11 Note that the expression in the maximum is larger than 12 since $\delta_\sigma \leq 1/12$, and hence $\mathcal{P}(\xi) > 0$. The function is large when some ξ_σ is close to one of the three points i, ζ, ζ^2 or ξ_σ is close to the boundary of $\mathcal{F}$ or to the vertical imaginary axis.

Proposition 2.14 *Assume j is a singular modulus. Let $\alpha = j(\xi)$, $\xi \in \mathcal{F}$, be an algebraic number that is the j–invariant of an elliptic curve without complex multiplication. Each embedding $\sigma\colon \mathbb{Q}(\alpha) \hookrightarrow \mathbb{C}$ gives a δ_σ as in Lemma 2.11. Assume that $j - \alpha$ is an algebraic unit and let $0 < \varepsilon < 1/4$. We can bound the height by*

$$h(j-\alpha) \leq c_2 \frac{\sum_{\sigma\colon \mathbb{Q}(j,\alpha)\to\mathbb{C}} \sum_{M\in\mathcal{T}} \mathcal{C}(\Delta; M\xi_\sigma; \varepsilon)}{16 \cdot \mathcal{C}(\Delta)} (\log|\Delta|)^4 + 5\mathcal{P}(\xi) + |\log\varepsilon|,$$

where $\mathcal{T} = \{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&\pm1\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)\}$. The constant $c_2 \geq 1$ is the one appearing in [Mas06, Theorem I] and it only depends on α.

PROOF. Let $j(\xi) = \alpha$. Since $j - \alpha$ is an algebraic unit the height can be computed by

$$h(j-\alpha) = \frac{1}{[\mathbb{Q}(j,\alpha):\mathbb{Q}]} \sum_{\sigma} \log \max\left\{1, |\sigma(j-\alpha)^{-1}|\right\}, \tag{2.9}$$

where σ runs over all field embeddings $\sigma\colon \mathbb{Q}(j,\alpha) \to \mathbb{C}$.

Let

$$\varepsilon_0 = \varepsilon \cdot \min_{\sigma\colon \mathbb{Q}(\alpha)\to\mathbb{C}} \{1, c(\xi_\sigma)\}.$$

We split the sum (2.9) into terms for which $|\sigma(j-\alpha)| < \varepsilon_0$ and those for which $1 \geq |\sigma(j-\alpha)| \geq \varepsilon_0$.
Assume $|\sigma(j-\alpha)| < \varepsilon_0$. Thus $|\sigma(j-\alpha)| < c(\xi_\sigma)$. We want to show $|\tau_\sigma - M\xi_\sigma| < \varepsilon$ for some $M \in \mathcal{T}$. We can apply Lemma 2.12 to get $|\tau_\sigma - M\xi_\sigma| < \delta_\sigma$ for some $M \in \mathcal{T}$. By definition of δ_σ we have $\delta_\sigma \leq \frac{|j'(\xi_\sigma)|}{6|j'(\xi_\sigma)|+72\cdot 10^5 \max\{1,|j(\xi_\sigma)|\}}$, so that we can apply Lemma 2.10 and $j(\xi_\sigma) = j(M\xi_\sigma)$ to get

$$\frac{|j'(\xi_\sigma)|}{2}|\tau_\sigma - M\xi_\sigma| \leq |j(\tau_\sigma) - j(\xi_\sigma)| < \varepsilon_0 \leq \varepsilon c(\xi_\sigma) \leq \varepsilon\frac{|j'(\xi_\sigma)|}{2}\delta_\sigma \leq \varepsilon\frac{|j'(\xi_\sigma)|}{2},$$

which implies $|\tau_\sigma - M\xi_\sigma| < \varepsilon$. But we also get from this

$$|j(\tau_\sigma) - j(\xi_\sigma)| \geq \frac{|j'(\xi_\sigma)|}{2} \min\{|\tau_\sigma - M\xi_\sigma|; M \in \mathcal{T}\}.$$

Note that the right–hand side is not 0 since $j(\xi_\sigma)$ does not have complex multiplication but $j(\tau_\sigma)$ does. The same argument tells us $j'(\xi_\sigma) \neq 0$. By [Mas06, Theorem I] we obtain

$$\log|j(\tau_\sigma) - j(\xi_\sigma)| \geq \log\frac{|j'(\xi_\sigma)|}{2} - c_2 h'(\tau_\sigma)^4$$

where $h'(\tau_\sigma)$ denotes the maximum of 1 and the height of τ_σ. Here $c_2 \geq 1$ is the constant from [Mas06] that depends on ξ (and the ξ_σ.) We now use Lemma 2.13 to get

$$\begin{aligned}\log|\sigma(j-\alpha)| = \log|j(\tau_\sigma) - j(\xi_\sigma)| &\geq \log\frac{|j'(\xi_\sigma)|}{2} - c_2\left(\log|\Delta|^{1/2}\right)^4 \\ &\geq \log\frac{|j'(\xi_\sigma)|}{2} - \frac{c_2}{16}\left(\log|\Delta|\right)^4. \end{aligned}\tag{2.10}$$

Since with $|\sigma(j-\alpha)| < \varepsilon_0$ we also have $|\tau_\sigma - M\xi_\sigma| < \varepsilon$ for some matrix $M \in \mathcal{T}$ as mentioned before, we obtain that τ_σ corresponds to a form in $\mathcal{C}(\Delta; M\xi_\sigma; \varepsilon)$. We have

$$c(\xi_\sigma) \leq \frac{|j'(\xi_\sigma)|}{2}\delta \leq \frac{|j'(\xi_\sigma)|}{2}$$

so that

$$\log \max\left\{1, \frac{2}{|j'(\xi_\sigma)|}\right\} \leq \mathcal{P}(\xi).$$

Thus, if use (2.10) we get

$$\begin{aligned} h(j-\alpha) \leq & -\frac{1}{[\mathbb{Q}(j,\alpha):\mathbb{Q}]} \sum_{|\sigma(j-\alpha)|<\varepsilon_0} \log|\sigma(j-\alpha)| + |\log \varepsilon_0| \\ \leq & c_2 \frac{\sum_{\sigma\colon \mathbb{Q}(\alpha)\hookrightarrow\mathbb{C}} \sum_{M\in\mathcal{T}} \mathcal{C}(\Delta; M\xi_\sigma; \varepsilon)}{16\cdot[\mathbb{Q}(j,\alpha):\mathbb{Q}]} (\log|\Delta|)^4 + 4\mathcal{P}(\xi) + |\log\varepsilon_0|. \end{aligned}$$

If we plug in the definition for ε_0 we obtain

$$\begin{aligned} h(j-\alpha) \leq & c_2 \frac{\sum_{\sigma\colon \mathbb{Q}(\alpha)\hookrightarrow\mathbb{C}} \sum_{M\in\mathcal{T}} \mathcal{C}(\Delta; M\xi_\sigma; \varepsilon)}{16\cdot\mathcal{C}(\Delta)} (\log|\Delta|)^4 + 4\mathcal{P}(\xi) \\ & - \log \min_\sigma \{1, c(\xi_\sigma)\} + |\log\varepsilon| \\ = & c_2 \frac{\sum_{\sigma\colon \mathbb{Q}(\alpha)\hookrightarrow\mathbb{C}} \sum_{M\in\mathcal{T}} \mathcal{C}(\Delta; M\xi_\sigma; \varepsilon)}{16\cdot\mathcal{C}(\Delta)} (\log|\Delta|)^4 + 5\mathcal{P}(\xi) + |\log\varepsilon|, \end{aligned}$$

where we also used the first claim of Lemma 2.8 for the inequality. □

The following lemmas can be found in Section 3 of [BHK18] as Lemmas 3.5 and 3.6.

Lemma 2.15 *Assume that $|\Delta| \geq 10^{14}$. Then we have $F(\Delta) \geq |\Delta|^{0.34/\log\log(|\Delta|^{1/2})}$ and $F(\Delta) \geq 18\log\log(|\Delta|^{1/2})$.*

Lemma 2.16 *For $\Delta \neq -3, -4$ we have $\mathcal{C}(\Delta) \leq \pi^{-1}|\Delta|^{1/2}(2+\log|\Delta|)$.*

Proof. Theorem 10.1 in [Hua12] says $\mathcal{C}(\Delta) \leq \frac{\omega|\Delta|^{1/2}}{2\pi} K(d)$ where $K(d)$ can be bounded by $2+\log|\Delta|$ according to Theorem 14.3 in [Hua12] and ω is the number of roots of unity in the imaginary quadratic order of discriminant Δ. But since $\Delta \neq -3, -4$ we have $\omega = 2$ and the result follows. □

We define $E := E(\Delta) = F(\Delta)(\log|\Delta|)^4$.

Corollary 2.17 *Assume $j-\alpha$ is a unit and $\alpha = j(\xi)$ with $\xi \in \mathcal{F}$ and such that α is algebraic but not a singular modulus. For $|\Delta| \geq 10^{14}$ we have*

$$h(j-\alpha) \leq c_2[\mathbb{Q}(\alpha):\mathbb{Q}]\frac{E(\Delta)}{2\mathcal{C}(\Delta)} + \log\frac{E(\Delta)|\Delta|^{1/2}}{\mathcal{C}(\Delta)} + C'$$

where C' is a constant depending on α and is given by

$$C' = 4[\mathbb{Q}(\alpha):\mathbb{Q}]c_2 + 5\mathcal{P}(\xi).$$

PROOF. We can use the previous results together with the bound for $\mathcal{C}(\Delta;\xi_\sigma;\varepsilon)$ to bound the height $h(j-\alpha)$. Put $\varepsilon = \frac{\mathcal{C}(\Delta)}{F(\Delta)(\log|\Delta|)^4|\Delta|^{1/2}}$. Since we assume that $|\Delta| \geq 10^{14}$ we have $F(\Delta) \geq 256$ and we obtain together with Lemma 2.16

$$\varepsilon \leq \frac{\pi^{-1}|\Delta|^{1/2}(2+\log|\Delta|)}{256(\log|\Delta|)^4|\Delta|^{1/2}} \leq \frac{1}{256\pi}\frac{2+\log 10^{14}}{\log 10^{14}} < 2\cdot 10^{-3}.$$

Thus we can apply Proposition 2.14 together with Corollary 2.3 and obtain

$$\begin{aligned}
h(j-\alpha) \leq & [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{4F(\Delta)\left(32|\Delta|^{1/2}\varepsilon^2\log\log(|\Delta|^{1/2})+11|\Delta|^{1/2}\varepsilon+2\right)}{16\cdot\mathcal{C}(\Delta)}(\log|\Delta|)^4 \\
& + 5\mathcal{P}(\xi) + |\log\varepsilon| \\
= & [\mathbb{Q}(\alpha):\mathbb{Q}]c_2E\frac{128|\Delta|^{1/2}\log\log(|\Delta|^{1/2})}{16\cdot\mathcal{C}(\Delta)}\left(\frac{\mathcal{C}(\Delta)}{F(\Delta)(\log|\Delta|)^4|\Delta|^{1/2}}\right)^2 \\
& + [\mathbb{Q}(\alpha):\mathbb{Q}]c_2E\frac{44|\Delta|^{1/2}}{16\cdot\mathcal{C}(\Delta)}\frac{\mathcal{C}(\Delta)}{F(\Delta)(\log|\Delta|)^4|\Delta|^{1/2}} \\
& + [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{E}{2\mathcal{C}(\Delta)} \\
& + 5\mathcal{P}(\xi) + \log\left(\frac{F(\Delta)(\log|\Delta|)^4|\Delta|^{1/2}}{\mathcal{C}(\Delta)}\right).
\end{aligned}$$

We continue the estimate by simplifying the terms to get

$$\begin{aligned}
h(j-\alpha) \leq & [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{8\log\log(|\Delta|^{1/2})}{F(\Delta)}\frac{\mathcal{C}(\Delta)}{(\log|\Delta|)^4|\Delta|^{1/2}} \\
& + 3[\mathbb{Q}(\alpha):\mathbb{Q}]c_2 + [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{E}{2\mathcal{C}(\Delta)} \\
& + 5\mathcal{P}(\xi) + \log\left(\frac{E|\Delta|^{1/2}}{\mathcal{C}(\Delta)}\right).
\end{aligned}$$

Now we apply Lemma 2.15 to see

$$\begin{aligned}
h(j-\alpha) \leq & [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{1}{2}\frac{\mathcal{C}(\Delta)}{(\log|\Delta|)^4|\Delta|^{1/2}} \\
& + 3[\mathbb{Q}(\alpha):\mathbb{Q}]c_2 + [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{E}{2\mathcal{C}(\Delta)} \\
& + 5\mathcal{P}(\xi) + \log\left(\frac{E|\Delta|^{1/2}}{\mathcal{C}(\Delta)}\right).
\end{aligned}$$

We continue the estimate using Lemma 2.16

$$\begin{aligned} h(j-\alpha) \leq & [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{1}{2\pi}\frac{|\Delta|^{1/2}(2+\log|\Delta|)}{(\log|\Delta|)^4|\Delta|^{1/2}} \\ & + 3[\mathbb{Q}(\alpha):\mathbb{Q}]c_2 + [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{E}{2\mathcal{C}(\Delta)} \\ & + 5\mathcal{P}(\xi) + \log\left(\frac{E|\Delta|^{1/2}}{\mathcal{C}(\Delta)}\right). \end{aligned}$$

Simplifying again results into

$$\begin{aligned} h(j-\alpha) \leq & [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{1}{2\pi}\frac{(2+\log|\Delta|)}{(\log|\Delta|)^4} \\ & + 3[\mathbb{Q}(\alpha):\mathbb{Q}]c_2 + [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{E}{2\mathcal{C}(\Delta)} \\ & + 5\mathcal{P}(\xi) + \log\left(\frac{E|\Delta|^{1/2}}{\mathcal{C}(\Delta)}\right) \end{aligned}$$

The function $x \mapsto \frac{2+\log x}{(\log x)^4}$ is decreasing for $x > 1$. Thus we can substitute $x = 10^{14}$ to continue the bound and get

$$\begin{aligned} h(j-\alpha) \leq & [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{1}{2\pi}\frac{35}{32^4} \\ & + 3[\mathbb{Q}(\alpha):\mathbb{Q}]c_2 + [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{E}{2\mathcal{C}(\Delta)} \\ & + 5\mathcal{P}(\xi) + \log\left(\frac{E|\Delta|^{1/2}}{\mathcal{C}(\Delta)}\right). \end{aligned}$$

This gives the desired inequality. □

2.3 Proof of the main theorem in the CM case

We now want to bound Δ to complete the main proof. We will do this by using the lower and upper bounds we derived in the last section. Throughout this section we assume $|\Delta| \geq 10^{30}$.

Put

$$C = C' + h(\alpha) + \log 2 + 0.01. \tag{2.11}$$

Combining the lower bounds for $h(j-\alpha)$ from Lemmas 2.8 and 2.9 with the upper

bound from Corollary 2.17 we obtain the inequality

$$L := \max\left\{\pi\frac{|\Delta|^{1/2}}{\mathcal{C}(\Delta)}, \frac{3}{\sqrt{5}}\log|\Delta| - 10\right\} \leq [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{E(\Delta)}{2\mathcal{C}(\Delta)} + \log\frac{E(\Delta)|\Delta|^{1/2}}{\mathcal{C}(\Delta)} + C$$

or equivalently

$$1 \leq [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{E}{2L\cdot\mathcal{C}(\Delta)} + \frac{\log E + C}{L} + \frac{\log(|\Delta|^{1/2}/\mathcal{C}(\Delta))}{L}.$$

For the remainder we assume that $|\Delta|$ is large enough so that $\log E + C \geq 0$. By Lemma 2.15 this is the case when $|\Delta| \geq e^{e^{e^{-C}/18}}$. This in turn is true whenever $|\Delta| \geq 3$. Since $\frac{3}{\sqrt{5}}\log|\Delta| - 10 > 0$ for $|\Delta| \geq 10^{14}$ this allows us to replace L by $\frac{3}{\sqrt{5}}\log|\Delta| - 10$ in the middle term. Similarly we can replace L in the first term by $\pi|\Delta|^{1/2}/\mathcal{C}(\Delta)$ and obtain

$$\begin{aligned} 1 &\leq [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{E}{2\pi|\Delta|^{1/2}} + \frac{\log E + C}{\frac{3}{\sqrt{5}}\log|\Delta| - 10} + \frac{\log(|\Delta|^{1/2}/\mathcal{C}(\Delta))}{L} \\ &\leq [\mathbb{Q}(\alpha):\mathbb{Q}]c_2\frac{E}{2\pi|\Delta|^{1/2}} + \frac{\log E}{\frac{3}{\sqrt{5}}\log|\Delta| - 10} + \frac{\log(\pi^{-1}L)}{L} + \frac{C}{\frac{3}{\sqrt{5}}\log|\Delta| - 10} \end{aligned} \tag{2.12}$$

We want to show that the right–hand side is less than 1 for large enough $|\Delta|$. Before we start, we want to give a bound on $E(\Delta) = F(\Delta)(\log|\Delta|)^4$. To do this, we are going to bound $\log F(\Delta)$ and $\log E(\Delta)$. A bound for $\log F(\Delta)$ can be found in equation 5.7 of [BHK18] and is given by

$$\frac{\log F(\Delta)}{\log 2} \leq \frac{1}{2}\frac{\log|\Delta|}{\log\log|\Delta| - c_1 - \log 2},$$

where $c_1 < 1.1713142$ is defined by

$$\omega(N_1) = \frac{\log N_1}{\log\log N_1 - c_1}$$

and $N_1 = 2\cdot 3\cdot 5\cdots 1129$ is the product of the first 189 prime numbers. Then the bound on $\log E(\Delta)$ is given by

$$\log E(\Delta) \leq \frac{\log 2}{2}\frac{\log|\Delta|}{\log\log|\Delta| - c_1 - \log 2} + 4\log\log|\Delta|.$$

Now we want to bound $E|\Delta|^{-1/2}$. Since the function

$$u_0(x) = \frac{\log 2}{2}\frac{1}{\log\log x - c_1 - \log 2} + \frac{4\log\log x}{\log x} - \frac{1}{2}$$

is decreasing for $x \geq 10^{10}$ we obtain

$$\frac{\log(E|\Delta|^{-1/2})}{\log|\Delta|} \leq u_0(10^{30}) < -0.1085$$

for $|\Delta| \geq 10^{30}$. This in turn implies

$$E|\Delta|^{-1/2} < |\Delta|^{-0.1085}. \tag{2.13}$$

The next step is to bound the second term of (2.12). The functions

$$u_1(x) = \log 2\frac{1}{\log\log x - c_1 - \log 2} + 4\frac{\log\log x}{\log x}$$

and

$$u_2(x) = \left(\frac{3}{\sqrt{5}} - \frac{10}{\log x}\right)^{-1}$$

are decreasing for $x \geq 10^{10}$. We have

$$\frac{\log E}{\frac{3}{\sqrt{5}}\log|\Delta| - 10} \leq u_1(|\Delta|)u_2(|\Delta|) \leq u_1(10^{30})u_2(10^{30}) \leq 0.4896 \tag{2.14}$$

for $|\Delta| \geq 10^{30}$.

To bound the third term of (2.12) we remark that the function $x \mapsto x^{-1}\log(\pi^{-1}x)$ is decreasing for $x \geq e/\pi$. We have $L \geq \frac{3}{\sqrt{5}}\log|\Delta| - 10 \geq e/\pi$ for $\Delta \geq 10^{15}$ and therefore

$$\frac{\log(\pi^{-1}L)}{L} \leq \frac{\log\left(\pi^{-1}\left(\frac{3}{\sqrt{5}}\log|\Delta| - 10\right)\right)}{\frac{3}{\sqrt{5}}\log|\Delta| - 10}.$$

The function

$$u_3(x) := \frac{\log\left(\pi^{-1}\left(\frac{3}{\sqrt{5}}\log x - 10\right)\right)}{\frac{3}{\sqrt{5}}\log x - 10}$$

is decreasing for $x \geq 10^{15}$. Thus we obtain

$$\frac{\log(\pi^{-1}L)}{L} \leq u_3(|\Delta|) \leq u_3(10^{15}) < 0.0674 < \frac{1}{10} \tag{2.15}$$

for $|\Delta| \geq 10^{30}$.

For $|\Delta| \geq e^{10\frac{\sqrt{5}}{3}(C+1)}$ we have

$$\frac{C}{\frac{3}{\sqrt{5}}\log|\Delta| - 10} \leq \frac{1}{10}. \tag{2.16}$$

By equation (2.13) we can bound the first term of (2.12) by

$$\left(\frac{[\mathbb{Q}(\alpha):\mathbb{Q}]c_2}{2\pi}\right)|\Delta|^{-0.1085} \leq \frac{1}{10}$$

for $|\Delta| \geq (10[\mathbb{Q}(\alpha):\mathbb{Q}]c_2/(2\pi))^{10}$.

Using these two inequalities together with (2.14) and (2.15) we obtain that (2.12) is less than 1 for all

$$|\Delta| \geq \max\left\{10^{30}, e^{10\frac{\sqrt{5}}{3}(C+1)}, (10[\mathbb{Q}(\alpha):\mathbb{Q}]c_2/(4\pi))^{10}\right\}. \tag{2.17}$$

This contradicts the lower bound of (2.12).

The lower bound on $|\Delta|$ can be simplified. In equation (2.11) we have put $C = 4[\mathbb{Q}(\alpha):\mathbb{Q}]c_2 + 5\mathcal{P}(\xi) + h(\alpha) + \log 2 + 0.01$. Recall that $c_2 \geq 1$. This implies $C \geq 4.7$ and hence

$$e^{15C} \geq e^{70} \geq 10^{30}.$$

Moreover, we have

$$\begin{aligned} 15C \geq 10\frac{\sqrt{5}}{3}(C+1) &\geq 10\frac{\sqrt{5}}{3}2[\mathbb{Q}(\alpha):\mathbb{Q}]c_2 \geq 10\frac{5}{4\pi}2[\mathbb{Q}(\alpha):\mathbb{Q}]c_2 \\ &\geq 10\log\left(\frac{5}{4\pi}2[\mathbb{Q}(\alpha):\mathbb{Q}]c_2\right). \end{aligned}$$

Therefore, the bound on $|\Delta|$ from equation (2.17) simplifies to

$$|\Delta| \geq e^{15C}, \tag{2.18}$$

where $C > 0$ is a computable constant

$$C = 2[\mathbb{Q}(\alpha):\mathbb{Q}]c_2 + 6\mathcal{P}(\xi) + h(\alpha) + \log 2 + 0.01.$$

3 The non–CM case

We fix an elliptic curve without complex multiplication, and denote by j_0 its j–invariant. Assume that the curve is defined over a number field K contained in $\mathbb{C}$. Our aim is to prove the following result.

Theorem 3.1 *Let j_0 be the j–invariant of an elliptic curve without complex multiplication. Then there are at most finitely many j–invariants j of elliptic curves that are isogenous to an elliptic curve corresponding to j_0 and such that j is an algebraic unit.*

Assume $j(\tau_0) = j_0$ for $\tau_0 \in \mathcal{F}$. For any embedding $\sigma\colon K \hookrightarrow \mathbb{C}$ there is a $\tau_0^\sigma \in \mathcal{F}$ such that $j(\tau_0^\sigma) = \sigma(j_0)$.

For $\xi \in \bar{\mathcal{F}}$ and $\tau \in \mathbb{H}$ such that the curves corresponding to j_0 and $j(\tau)$ are isogenous we define the sets

$$\Sigma(\xi, \varepsilon) = \{\tau \in \mathcal{F}; |j(\tau) - j(\xi)| < \varepsilon\}$$

and

$$\Gamma(\xi, \varepsilon) = \{\sigma\colon K \to \mathbb{C}; \tau^\sigma \in \Sigma(\xi, \varepsilon)\}\,.$$

We will write Σ_ε and Γ_ε for $\Sigma(\zeta, \varepsilon)$ and $\Gamma(\zeta, \varepsilon)$, respectively, where $\zeta = e^{2\pi i/6}$.
The set Σ_ε is sketched in Figure 3.1.

To estimate the number of elements in Γ_ε one can use of the following theorem, which is Théorème 1 in [Ric13]. Recall the hyperbolic probability measure $\mathrm{d}\mu = \frac{3}{\pi}\frac{\mathrm{d}x\mathrm{d}y}{y^2}$.

Theorem 3.2 *Let $E_1, E_2, \ldots$ be pairwise isogenous elliptic curves without complex multiplication. Let j_n be the j–invariant of E_n and assume that the j_n are pairwise distinct. Then the sequence of Dirac meassures $\left(\delta_{\mathrm{Aut}(\mathbb{C}/\mathbb{Q})\cdot j_n}\right)$ converges to the hyperbolic probability measure μ. This amounts to saying*

$$\frac{1}{[\mathbb{Q}(j_n) : \mathbb{Q}]} \sum_{z \in \mathrm{Gal}(\mathbb{Q}(j_n)/\mathbb{Q})\cdot j_n} f(z) \to \int_{\mathcal{F}} f \mathrm{d}\mu$$

as $n \to \infty$ for any bounded and continuous $f\colon \mathcal{F} \to \mathbb{R}$.

Lemma 3.3 *We have $\#\Gamma_\varepsilon \leq c\varepsilon^{2/3} D_n$ for sufficiently large D_n depending on ε and some $c > 0$, where $D_n = [\mathbb{Q}(j_n) : \mathbb{Q}]$. Here Γ_ε corresponds to the τ_n with $j(\tau_n) = j_n$.*

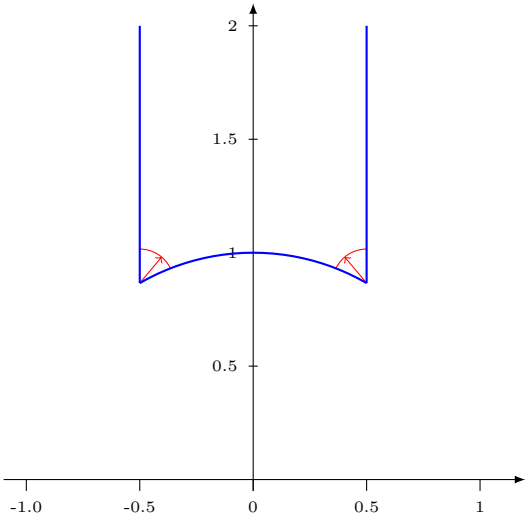


Figure 3.1: Neighborhood of ζ.

PROOF. By the last theorem we have $|\#\Gamma_\varepsilon/D_n - \mu(\Sigma_\varepsilon)| \to 0$ as $n \to \infty$. The proof of Lemma 2 in [Hab15] gives the estimate $\mu(\Sigma_\varepsilon) \leq c\varepsilon^{2/3}$. The claim follows. □

3.1 Isogenous points in the fundamental domain

We want to give an explicit bound for the number of elements in the Galois orbit of j_0 satisfying the condition above. First, we will bound the number of points in the Hecke orbit, and then use a result of Lombardo to estimate the total number. Two (equivalence classes of isomorphic) elliptic curves are in the same Hecke orbit if they are isogenous.

The following lemma is of similar nature to Lemma 2.2. We translate points in the upper half–plane into the fundamental domain with matrices in $\mathrm{SL}_2(\mathbb{Z})$, and thus get restrictions on then entries of the matrices.

Lemma 3.4 *Let $\xi \in \bar{\mathcal{F}}$ and $\varepsilon \in (0, \frac{\sqrt{3}}{3|\xi|+2}]$. Let $\tau \in \mathbb{H}$ satisfy $|\tilde{\tau} - \xi| \leq \varepsilon$, where $\tilde{\tau} \in \mathcal{F}$ is in the $\mathrm{SL}_2(\mathbb{Z})$–orbit of τ. Pick*

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$

such that $\gamma\tau = \tilde{\tau}$. Then there exist $\nu \in \{\pm 1\}$ such that

$$\left| a^2 + \nu 2|\operatorname{Re}(\xi)|ac + |\xi|^2 c^2 - \frac{\operatorname{Im}(\xi)}{\operatorname{Im}(\tau)} \right| \leq 7\frac{4|\xi|+1}{\sqrt{3}}|\xi|^2 \frac{\varepsilon^{1/2}}{\operatorname{Im}(\tau)}, \tag{3.1}$$

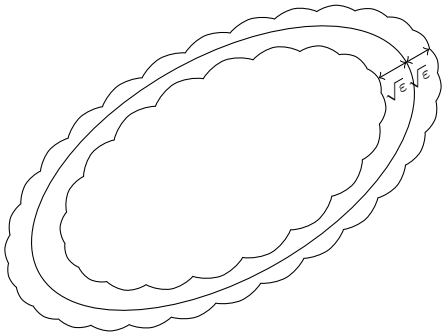


Figure 3.2: Neighborhood of ellipses

and

$$\max\left\{a^2, c^2\right\} \leq \frac{4|\xi|+1}{\sqrt{3}} \frac{1}{\operatorname{Im}(\tau)}. \tag{3.2}$$

Moreover, we have

$$|d| \leq |c||\operatorname{Re}(\tau)| + \frac{4|\xi|+1}{\sqrt{3}}$$

and

$$|b| \leq |a||\operatorname{Re}(\tau)| + \frac{4|\xi|+1}{\sqrt{3}}.$$

The lemma tells us, that the first column of γ, considered as a point in the plane, is close to a conic section. Since

$$(2\nu|\operatorname{Re}(\xi)|)^2 - 4|\xi|^2 = 4(\operatorname{Re}(\xi)^2 - \operatorname{Re}(\xi)^2 - \operatorname{Im}(\xi)^2) = -4\operatorname{Im}(\xi) < 0$$

the equation actually defines an ellipse as pictured in Figure 3.2. The ellipse is defined in terms of ξ and τ.

Proof. Let $R = |\xi|$ and $A = \operatorname{Im}(\xi)$. Moreover write $\tau = x + iy$. We have

$$\operatorname{Im}(\gamma\tau) = \frac{\operatorname{Im}\tau}{(cx+d)^2 + c^2y^2} \geq A - \varepsilon \geq A - \frac{\sqrt{3}}{3R+2} \geq \frac{\sqrt{3}}{4R+1}$$

by definition of ε and $A \geq \frac{\sqrt{3}}{2}$. Define $\delta_1 := \operatorname{Im}(\gamma\tau)^{-1}$. Then $\delta_1 \leq 1/(A-\varepsilon)$ and

$$(cx+d)^2 + c^2y^2 = \delta_1 y. \tag{3.3}$$

This yields $c^2 \le \delta_1/y \le (A-\varepsilon)^{-1}y^{-1}$, which implies the bound on c^2, and

$$|cx+d||c|y \le \frac{1}{2}\left((cx+d)^2+c^2y^2\right) = \frac{1}{2}\delta_1 y \le \frac{1}{2(A-\varepsilon)}y.$$

Further we get

$$|cx+d||c| \le \frac{1}{2(A-\varepsilon)} \le \frac{4R+1}{\sqrt{3}} \tag{3.4}$$

and hence

$$|d| \le |c||\operatorname{Re}(\tau)| + \frac{4R+1}{\sqrt{3}}$$

if $c \neq 0$. Thus the inequality for d in the statement is true in the case $c \neq 0$. But if $c = 0$, then $d = \pm 1$, and $|d| = 1 \le \frac{5}{\sqrt{3}} \le \frac{4R+1}{\sqrt{3}}$. Thus, the inequality for d holds in both cases.

Put $\gamma' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\gamma$. Then $\gamma'\tau = -\frac{1}{\tilde{\tau}}$. Define $\delta_2 := \operatorname{Im}(\gamma'\tau)^{-1}$, i.e.

$$(ax+b)^2 + a^2y^2 = \delta_2 y. \tag{3.5}$$

We put $r = |\tilde{\tau}|$ and $B = \operatorname{Im}(\tilde{\tau})$. Now by the general rule of transformation of the imaginary part under fractional linear transformations

$$\delta_2 = \operatorname{Im}(\gamma'\tau)^{-1} = \operatorname{Im}\left(-\frac{1}{\tilde{\tau}}\right)^{-1} = \left(\frac{\operatorname{Im}(\tilde{\tau})}{|\tilde{\tau}|^2}\right)^{-1} = \frac{r^2}{B}.$$

We remark that $B/r = \operatorname{Im}(\tilde{\tau}/|\tilde{\tau}|) \ge \sqrt{3}/2$ since $\tilde{\tau}/r \in \bar{\mathcal{F}}$, and similarly $A/R \ge \sqrt{3}/2$. This implies

$$\delta_2 \le \frac{2}{\sqrt{3}}r \le \frac{2}{\sqrt{3}}(R+\varepsilon) \le \frac{2R+1}{\sqrt{3}}.$$

We proceed as before with the bound on d and c^2. From (3.5) we obtain $a^2 \le \delta_2/y \le (3R+1)/(\sqrt{3}y)$, which is the desired inequality of the statement. Moreover, we obtain

$$|ax+b||a| \le \delta_2/2 \le (R+1)/\sqrt{3} \tag{3.6}$$

and hence

$$|b| \le |a||\operatorname{Re}(\tau)| + \frac{R+1}{\sqrt{3}},$$

whenever $a \neq 0$. Again, if $a = 0$, then $|b| = 1 \le \frac{4R+1}{\sqrt{3}}$, as claimed.

It remains to prove (3.1). We deal with the case $c = 0$ first. Then $a = d = \pm 1$ and

$y = \delta_1^{-1} = \operatorname{Im}\tilde{\tau} = \operatorname{Im}\tau$, and thus $|y - A| \leq |\tilde{\tau} - \xi| \leq \varepsilon$. This implies

$$\left|\frac{1}{A} - \frac{1}{y}\right| \leq \frac{\varepsilon}{Ay} \leq \frac{1}{A}\frac{\varepsilon^{1/2}}{y}.$$

Multiplying by A shows that Equation (3.1) is true for any value of ν. Now assume $c \neq 0$. We want to prove $|\delta_2 - \frac{|\xi|^2}{\operatorname{Im}\xi}| = |\delta_2 - \frac{|\xi|^2}{A}| \ll \varepsilon$. We compute

$$\begin{aligned}
\left|\delta_2 - \frac{R^2}{A}\right| &= \left|\frac{r^2}{B} - \frac{R^2}{A}\right| = \left|\frac{R^2B - r^2A}{AB}\right| \\
&= \left|\frac{R^2B - BRr + BRr - ARr + ARr - r^2A}{AB}\right| \\
&\leq RB\left|\frac{r-R}{AB}\right| + Rr\left|\frac{A-B}{AB}\right| + rA\left|\frac{r-R}{AB}\right| \qquad (3.7)\\
&\leq \frac{2}{\sqrt{3}}\varepsilon + \frac{4}{3}\varepsilon + \frac{2}{\sqrt{3}}\varepsilon \\
&\leq 4\varepsilon,
\end{aligned}$$

where we have used $|R - r| = ||\xi| - |\tilde{\tau}|| \leq |\xi - \tilde{\tau}| \leq \varepsilon$ and $|A - B| = |\operatorname{Im}(\xi) - \operatorname{Im}(\tilde{\tau})| \leq |\xi - \tilde{\tau}| \leq \varepsilon$ in the second last inequality.

Suppose $a = 0$ for now. Then $b = -c = \pm 1$ and $y = \delta_2^{-1}$. Multiplying (3.7) by $\operatorname{Im}(\xi) = A$ shows (3.1) as the following argument shows. We have $\delta_2^{-1} = \operatorname{Im}(\gamma'\tau) = \frac{\operatorname{Im}(\tau)}{|b|^2} = \operatorname{Im}(\tau)$ by the usual transformation formula for the imaginary part of the action of $\mathrm{SL}_2(\mathbb{Z})$ by fractional linear transformations. Thus

$$|A\delta_2 - R^2| = \left|R^2 - \frac{A}{\operatorname{Im}(\tau)}\right| = \left||\xi|^2 - \frac{\operatorname{Im}(\xi)}{\operatorname{Im}(\tau)}\right| \leq 4A\varepsilon \leq 4R\varepsilon^{1/2}.$$

We have $\operatorname{Im}(\tau) \leq 2/\sqrt{3}$ since $a = 0$ and γ translates τ into the fundamental domain. Therefore, the inequality remains true after multiplying the right–hand side by $2/\sqrt{3}\operatorname{Im}(\tau)^{-1}$. This shows equation (3.1).

Finally, assume $ac \neq 0$. Put $X := x + d/c$ and $Y := x + b/a$. Consider the difference of the two

$$X - Y = \left(x + \frac{d}{c}\right) - \left(x + \frac{b}{a}\right) = \frac{1}{ac}.$$

If we divide (3.3) by c^2 and rewrite the result in terms of Y we get

$$0 = X^2 + y^2 - \frac{\delta_1 y}{c^2} = \left(Y + \frac{1}{ac}\right)^2 + y^2 - \frac{\delta_1 y}{c^2} = Y^2 + \frac{2}{ac}Y + \frac{1}{(ac)^2} + y^2 - \frac{\delta_1 y}{c^2}.$$

Similarly, if we divide (3.5) by a^2 we find

$$0 = Y^2 + y^2 - \frac{\delta_2 y}{a^2}$$

Computing the resultant of the last two displays as polynomials in Y, and multiplying the result by $(ac)^4$ to kill the denominators, gives us the expression

$$a^4y^2\delta_1^2 - 2a^2c^2y^2\delta_1\delta_2 + c^4y^2\delta_2^2 + 4a^2c^2y^2 - 2a^2y\delta_1 - 2c^2y\delta_2 + 1 = 0. \tag{3.8}$$

Now write $\delta_1 = \frac{1}{A} + \varepsilon_1$ and $\delta_2 = \frac{R^2}{A} + \varepsilon_2$. Then

$$|\varepsilon_1| = \left|\delta_1 - \frac{1}{A}\right| = \left|\frac{A - \operatorname{Im}(\tilde{\tau})}{A \operatorname{Im}(\tilde{\tau})}\right| \leq \frac{2}{\sqrt{3}A}\varepsilon$$

since $\operatorname{Im}(\tilde{\tau}) \geq \sqrt{3}/2$ and $|\operatorname{Im}(\xi) - \operatorname{Im}(\tilde{\tau})| \leq |\xi - \tilde{\tau}| \leq \varepsilon$. Also $|\varepsilon_2| \leq 4\varepsilon$ by (3.7). Put $\sigma = \operatorname{Re}(\xi)$. If we substitute these expressions for δ_1 and δ_2 in (3.8) we obtain

$$\begin{aligned} 0 =& a^4y^2\left(\frac{1}{A} + \varepsilon_1\right)^2 - 2a^2c^2y^2\left(\frac{1}{A} + \varepsilon_1\right)\left(\frac{R^2}{A} + \varepsilon_2\right) \\ &+ c^4y^2\left(\frac{R^2}{A} + \varepsilon_2\right)^2 + 4a^2c^2y^2 - 2a^2y\left(\frac{1}{A} + \varepsilon_1\right) - 2c^2y\left(\frac{R^2}{A} + \varepsilon_2\right) + 1. \end{aligned} \tag{3.9}$$

After multiplying the equation by A^2/y^2 the terms that do not include ε_1 and ε_2 are given by

$$\begin{aligned} a^4 - 2a^2c^2A\frac{R^2}{A} + c^4A^2\frac{R^4}{A^2} + 4a^2c^2A^2 - 2a^2\frac{A}{y} - 2c^2\frac{R^2}{A}\frac{A^2}{y} + \frac{A^2}{y^2} & \\ = a^4 - 2a^2c^2R^2 + c^4R^4 + 4a^2c^2A^2 - 2a^2\frac{A}{y} - 2c^2R^2\frac{A}{y} + \frac{A^2}{y^2} & \\ = a^4 - 2a^2c^2R^2 + c^4R^4 + 4a^2c^2(R^2 - \sigma^2) - 2a^2\frac{A}{y} - 2c^2R^2\frac{A}{y} + \frac{A^2}{y^2} & \\ = \left(a^2 - 2\sigma ac + R^2c^2 - \frac{A}{y}\right)\left(a^2 + 2\sigma ac + R^2c^2 - \frac{A}{y}\right). & \end{aligned}$$

The terms that involve ε_1 and ε_2 in (3.9) after multiplying it by A^2/y^2 are given by

$$\begin{aligned} A^2 \Bigg(\left(a^4 \frac{2}{A} - 2a^2c^2\frac{R^2}{A} - 2a^2\frac{1}{y} \right) \varepsilon_1 + \left(-2a^2c^2\frac{1}{A} + 2c^4\frac{R^2}{A} - 2c^2\frac{1}{y} \right) \varepsilon_2 \\ + \left(a^4\varepsilon_1^2 - 2a^2c^2\varepsilon_1\varepsilon_2 + c^4\varepsilon_2^2 \right) \Bigg) \\ = A^2 \Bigg(a^2 \left(a^2\frac{2}{A} - 2c^2\frac{R^2}{A} - 2\frac{1}{y} \right) \varepsilon_1 + c^2 \left(-2a^2\frac{1}{A} + 2c^2\frac{R^2}{A} - 2\frac{1}{y} \right) \varepsilon_2 \\ + \left(a^2\varepsilon_1 - c^2\varepsilon_2 \right)^2 \Bigg). \end{aligned}$$

Putting everything together in one equation again we obtain

$$\begin{aligned} \left(a^2 - 2\sigma ac + R^2c^2 - \frac{A}{y} \right) \left(a^2 + 2\sigma ac + R^2c^2 - \frac{A}{y} \right) \\ = -A^2 \Bigg(2a^2 \left(a^2\frac{1}{A} - c^2\frac{R^2}{A} - \frac{1}{y} \right) \varepsilon_1 + 2c^2 \left(-a^2\frac{1}{A} + c^2\frac{R^2}{A} - \frac{1}{y} \right) \varepsilon_2 \\ + \left(a^2\varepsilon_1 - c^2\varepsilon_2 \right)^2 \Bigg). \end{aligned}$$

We are now ready to prove (3.1). Choose $\nu \in \{\pm 1\}$ such that

$$\left| a^2 + 2\nu|\sigma|ac + R^2c^2 - \frac{A}{y} \right| \leq \left| a^2 - 2\nu|\sigma|ac + R^2c^2 - \frac{A}{y} \right|.$$

Then

$$\begin{aligned} \left| a^2 + 2\nu|\sigma|ac + R^2c^2 - \frac{A}{y} \right|^2 \leq & \left| a^2 - 2\sigma ac + R^2c^2 - \frac{A}{y} \right| \left| a^2 + 2\sigma ac + R^2c^2 - \frac{A}{y} \right| \\ \leq & A^2 \max\left\{a^2, c^2\right\} \Bigg(2 \left(a^2\frac{1}{A} + c^2\frac{R^2}{A} + \frac{1}{y} \right) |\varepsilon_1| \\ & + 2 \left(a^2\frac{1}{A} + c^2\frac{R^2}{A} + \frac{1}{y} \right) |\varepsilon_2| \\ & + \max\{a^2, c^2\} \left(|\varepsilon_1| + |\varepsilon_2| \right)^2 \Bigg). \end{aligned} \tag{3.10}$$

Note that $1/A \leq 2/\sqrt{3}$ and $R^2/A \leq 2R/\sqrt{3}$ as remarked on page 40. We also have acquired a bound for $\max\{a^2, c^2\}$ in the beginning of the proof displayed in (3.2). There-

fore,

$$\begin{aligned} a^2\frac{1}{A}+c^2\frac{R^2}{A}+\frac{1}{y} &\le \frac{4R+1}{\sqrt{3}}\frac{1}{y}\frac{1}{A}+\frac{4R+1}{\sqrt{3}}\frac{1}{y}\frac{R^2}{A}+\frac{1}{y} \\ &\le \frac{4R+1}{\sqrt{3}}\frac{1}{y}\left(\frac{1}{A}+\frac{R^2}{A}+\frac{2}{\sqrt{3}}\right) \\ &\le \frac{4R+1}{\sqrt{3}}\frac{6R}{\sqrt{3}}\frac{1}{y}. \end{aligned}$$

Using the bounds for ε_1 and ε_2 we get

$$\begin{aligned} 2\left(a^2\frac{1}{A}+c^2\frac{R^2}{A}+\frac{1}{y}\right)|\varepsilon_1| &\le 2\frac{4R+1}{\sqrt{3}}\frac{6R}{\sqrt{3}}\frac{1}{y}\frac{2}{\sqrt{3}A}\varepsilon \le 10\frac{4R+1}{\sqrt{3}}\frac{\varepsilon}{y}, \\ 2\left(a^2\frac{1}{A}+c^2\frac{R^2}{A}+\frac{1}{y}\right)|\varepsilon_2| &\le 2\frac{4R+1}{\sqrt{3}}\frac{6R}{\sqrt{3}}\frac{1}{y}4\varepsilon \le 28R\frac{4R+1}{\sqrt{3}}\frac{\varepsilon}{y} \end{aligned}$$

and

$$(|\varepsilon_1|+|\varepsilon_2|)^2 \le \left(\frac{1}{A}\frac{2}{\sqrt{3}}\varepsilon+4\varepsilon\right)^2 \le \varepsilon^2\left(\frac{1}{A}\frac{2}{\sqrt{3}}+4\right)^2 \le \varepsilon^2\left(\frac{4}{3}+4\right)^2 \le 29\varepsilon^2 \le 11\varepsilon$$

since $\varepsilon \le \sqrt{3}/5$. Using these inequalities for (3.10) and applying (3.2) again we obtain

$$\begin{aligned} \left|a^2+2\nu|\sigma|ac+R^2c^2-\frac{A}{y}\right|^2 &\le A^2\max\left\{a^2,c^2\right\}\left(38R\frac{4R+1}{\sqrt{3}}\frac{\varepsilon}{y}+11\max\{a^2,c^2\}\varepsilon\right) \\ &\le 49A^2R\left(\frac{4R+1}{\sqrt{3}}\right)^2\frac{\varepsilon}{y^2}. \end{aligned}$$

Taking the square–root on both sides gets us

$$\left|a^2+2\nu|\sigma|ac+R^2c^2-\frac{A}{y}\right| \le 7AR^{1/2}\left(\frac{4R+1}{\sqrt{3}}\right)\frac{\varepsilon^{1/2}}{y}.$$

Using $A \le R$ and $y = \mathrm{Im}(\tau)$ we get

$$\left|a^2+2\nu|\sigma|ac+R^2c^2-\frac{A}{y}\right| \le 7R^2\left(\frac{4R+1}{\sqrt{3}}\right)\frac{\varepsilon^{1/2}}{\mathrm{Im}(\tau)}.$$

This proves (3.1). □

Note that the estimates might be improved slightly, especially when $\xi = \zeta$ or $\xi = \zeta^2$ with $\zeta = e^{2\pi i/6}$. We are going to look at those special cases eventually, so that $\mathrm{Im}(\zeta) =$

$\sqrt{3}/2$.
We want to use the last lemma to prove the following proposition.

Proposition 3.5 *Let N be an integer, and let E_0 be an elliptic curve, and $\xi \in \bar{\mathcal{F}}$. Further, assume that $0 \le \varepsilon \le (100^{-1}|\xi|^{-3}\operatorname{Im}(\xi))^2$. Then the number of $\tau \in \bar{\mathcal{F}}$ with $|\xi - \tau| \le \varepsilon$ and such that E_0 is N–isogenous to a curve corresponding to $j(\tau)$ is bounded by*

$$10^7|\tau_0||\xi|^5\left(\sqrt{N}\sigma_0(N) + \sqrt{\varepsilon}\psi(N)\right).$$

For the remainder of the section we are going to prove this proposition.

For fixed $\tau \in \mathbb{H}$ with bounded real part we want to bound the number of matrices that satisfy the conditions in the lemma. For this we define

$$\mathcal{M}(\xi; x; y; \varepsilon) = \#\{\gamma \in \mathrm{SL}_2(\mathbb{Z}); \exists \tau = \tilde{x} + iy, |\tilde{x}| \le |x|, |\gamma\tau - \xi| \le \varepsilon \text{ and } \gamma\tau \in \bar{\mathcal{F}}\}.$$

Note that the last lemma tells us that all τ on horizontal lines in the upper half–plane satisfy the same equation for (a, c). Thus, if we look at horizontal line segments the number $\mathcal{M}(\xi; x; y; \varepsilon)$ can be bounded independent in terms of x.

If γ is as in the last lemma, then the first column (a, c) is close to one of the two ellipses

$$X^2 \pm 2|\operatorname{Re}(\xi)|XY + |\xi|^2Y^2 = \frac{\operatorname{Im}(\xi)}{\operatorname{Im}(\tau)}.$$

More precisely, we have

$$\left|\lambda - \frac{\operatorname{Im}(\xi)}{\operatorname{Im}(\tau)}\right| \le 50|\xi|^3\frac{\varepsilon^{1/2}}{\operatorname{Im}(\tau)}, \quad \text{where} \quad \lambda = a^2 \pm 2|\operatorname{Re}(\xi)|ac + |\xi|^2c^2. \tag{3.11}$$

We need an upper bound for the number $N(\operatorname{Im}(\tau), \varepsilon)$ of lattice points $(a, c) \in \mathbb{Z}^2$ that satisfy (3.11). Each of these points lies in a neighborhood of an ellipse defined above. We are going to use a result by Davenport [Dav51]. The following theorem is a special case of the result of Davenport.

Theorem 3.6 *Let $\mathcal{R}$ be a region in the two–dimensional plane with smooth boundary. If $V(\mathcal{R})$ denotes the volume of $\mathcal{R}$ and $N(\mathcal{R})$ the number of points with integral coordinates in $\mathcal{R}$, then*

$$|N(\mathcal{R}) - V(\mathcal{R})| < 4(L + 1),$$

where L is the length of the boundary of $\mathcal{R}$.

Thus, we need to compute the volume and the circumference of the ellipses that bound the given neighborhood. Let us assume that

$$\varepsilon \le \left(\frac{\operatorname{Im}(\xi)}{100|\xi|^3}\right)^2$$

is small enough. We consider the case when $\nu = 1$. The ellipses are then given by

$$E_\pm \colon Aa^2 + Bab + Cb^2 = 1$$

with

$$A = \frac{\operatorname{Im}(\tau)}{\operatorname{Im}(\xi) \pm 50|\xi|^3\varepsilon^{1/2}}, \quad B = \frac{2|\operatorname{Re}(\xi)|\operatorname{Im}(\tau)}{\operatorname{Im}(\xi) \pm 50|\xi|^3\varepsilon^{1/2}}, \quad C = \frac{|\xi|^2\operatorname{Im}(\tau)}{\operatorname{Im}(\xi) \pm 50|\xi|^3\varepsilon^{1/2}}.$$

The area of the bigger ellipse is then given by

$$\operatorname{vol}(E_+) = \frac{2\pi}{\sqrt{4AC - B^2}} = \pi\frac{\operatorname{Im}(\xi) + 50|\xi|^3\varepsilon^{1/2}}{\operatorname{Im}(\tau)\sqrt{|\xi|^2 - \operatorname{Re}(\xi)^2}} = \pi\frac{\operatorname{Im}(\xi) + 50|\xi|^3\varepsilon^{1/2}}{\operatorname{Im}(\tau)\operatorname{Im}(\xi)}.$$

Similarly, we have

$$\operatorname{vol}(E_-) = \pi\frac{\operatorname{Im}(\xi) - 50|\xi|^3\varepsilon^{1/2}}{\operatorname{Im}(\tau)\operatorname{Im}(\xi)}$$

for the smaller ellipse.
We now want to bound the circumference of $E_\pm$. For this we will use the following lemma.

Lemma 3.7 *Let E be an ellipse given by $Aa^2+Bac+Cc^2 = 1$. Then the circumference L of E is bounded by*

$$L \leq \sqrt{2(A + C)}\operatorname{vol}(E).$$

Proof. To prove this we rotate the ellipse, so that the new equation becomes

$$A'a^2 + C'b^2 = 1. \tag{3.12}$$

The coefficients are given by

$$A' = \frac{A + C}{2} + \frac{A - C}{2}\cos(2\theta) - \frac{B}{2}\sin(2\theta)$$

and

$$C' = \frac{A + C}{2} - \frac{A - C}{2}\cos(2\theta) + \frac{B}{2}\sin(2\theta),$$

where θ satisfies $\cot 2\theta = \frac{A-C}{B}$ or $\tan 2\theta = \frac{B}{A-C}$. Note if $B = 0$, we have $\theta = 0$, so that $A' = A$ and $C' = C$. Now the circumference of an ellipse in the form of (3.12) can be estimated by

$$L \leq \sqrt{2}\pi\sqrt{\frac{1}{A'} + \frac{1}{C'}} \leq \sqrt{2}\pi\sqrt{\frac{A' + C'}{A'C'}} = 2\sqrt{2}\pi\sqrt{\frac{A' + C'}{4A'C'}}.$$

But if we put $B' = 0$, then $A' + C' = A + C$ and $4A'C' = 4A'C' - B'^2 = 4AC - B^2$ since the discriminant is an invariant. Thus

$$L \leq \sqrt{2}\sqrt{A+C}\frac{2\pi}{4AC - B^2} = \sqrt{2}\sqrt{A+C}\,\mathrm{vol}(E),$$

as desired. □

If L_+ denotes the circumference of E_+, then we have by the previous lemma

$$\begin{aligned} L_+ &\leq \sqrt{2}\pi\sqrt{\frac{\mathrm{Im}(\tau) + |\xi|^2\,\mathrm{Im}(\tau)}{\mathrm{Im}(\xi) + 50|\xi|^3\varepsilon^{1/2}}}\frac{\mathrm{Im}(\xi) + 50|\xi|^3\varepsilon^{1/2}}{\mathrm{Im}(\tau)\,\mathrm{Im}(\xi)} \\ &\leq \sqrt{2}\pi\frac{\sqrt{1+|\xi|^2}}{\mathrm{Im}(\xi)}\sqrt{\frac{\mathrm{Im}(\xi) + 50|\xi|^3\varepsilon^{1/2}}{\mathrm{Im}(\tau)}}. \end{aligned}$$

Now we use the bound on ε to get

$$\begin{aligned} L_+ &\leq \sqrt{2}\pi\frac{\sqrt{1+|\xi|^2}}{\mathrm{Im}(\xi)}\sqrt{\frac{\mathrm{Im}(\xi) + \frac{1}{2}\,\mathrm{Im}(\xi)}{\mathrm{Im}(\tau)}} \\ &= \sqrt{3}\pi\sqrt{\frac{1+|\xi|^2}{\mathrm{Im}(\xi)}}\frac{1}{\sqrt{\mathrm{Im}(\tau)}}. \end{aligned}$$

We have $\frac{|\xi|}{\mathrm{Im}(\xi)} \leq \frac{2}{\sqrt{3}}$ since ξ is in the fundamental domain. Hence $\frac{|\xi|^2}{\mathrm{Im}(\xi)} \leq \frac{4}{3}\,\mathrm{Im}(\xi)$ and therefore

$$\frac{1+|\xi|^2}{\mathrm{Im}(\xi)} = \frac{1}{\mathrm{Im}(\xi)} + \frac{|\xi|^2}{\mathrm{Im}(\xi)} \leq \frac{2}{\sqrt{3}} + \frac{4}{3}\,\mathrm{Im}(\xi) \leq \frac{8}{3}\,\mathrm{Im}(\xi).$$

Using this for the bound of L_+ yields

$$L_+ \leq 2\pi\frac{\sqrt{2\,\mathrm{Im}(\xi)}}{\sqrt{\mathrm{Im}(\tau)}}.$$

Similarly, we obtain

$$L_- \leq \sqrt{2}\pi\frac{\sqrt{1+|\xi|^2}}{\mathrm{Im}(\xi)}\sqrt{\frac{\mathrm{Im}(\xi) - 50|\xi|^3\varepsilon^{1/2}}{\mathrm{Im}(\tau)}}.$$

where L_- is the circumference of E_- and thus

$$\begin{aligned} L_- &\leq \sqrt{2}\pi \frac{\sqrt{1+|\xi|^2}}{\operatorname{Im}(\xi)} \sqrt{\frac{\operatorname{Im}(\xi)}{\operatorname{Im}(\tau)}} \\ &= \sqrt{2}\pi \sqrt{\frac{1+|\xi|^2}{\operatorname{Im}(\xi)}} \frac{1}{\sqrt{\operatorname{Im}(\tau)}} \\ &\leq 2\pi \frac{\sqrt{2\operatorname{Im}(\xi)}}{\sqrt{\operatorname{Im}(\tau)}}. \end{aligned}$$

Clearly this bound holds since L_- is the circumference of the smaller ellipse.

Let $N(E_\pm)$ denote the number of lattice points contained in $E_\pm$ as defined in Theorem 3.6. By this same theorem, the number of points contained in the elliptical annulus can be estimated by

$$\begin{aligned} N(E_+) - N(E_-) &= N(E_+) - \operatorname{vol}(E_+) - (N(E_-) - \operatorname{vol}(E_-)) + \operatorname{vol}(E_+) - \operatorname{vol}(E_-) \\ &\leq 4(L_+ + 1) + 4(L_- + 1) + \operatorname{vol}(E_+) - \operatorname{vol}(E_-) \\ &\leq 8(L_+ + 1) + \frac{100\pi|\xi|^3\varepsilon^{1/2}}{\operatorname{Im}(\tau)\operatorname{Im}(\xi)} \\ &\leq 16\pi \frac{\sqrt{2\operatorname{Im}(\xi)} + \sqrt{\operatorname{Im}(\tau)}}{\sqrt{\operatorname{Im}(\tau)}} + \frac{100\pi|\xi|^3\varepsilon^{1/2}}{\operatorname{Im}(\tau)\operatorname{Im}(\xi)}. \end{aligned}$$

Therefore, a bound for $N(\operatorname{Im}(\tau), \varepsilon)$ is given by twice this number since the ellipse for $\nu = -1$ gives the same bound.

To obtain a bound for the number of matrices satisfying the conditions in Lemma 3.4, we need to estimate the possible pairs (b, d) when (a, c) is fixed. Let (a, c) be fixed, and assume that (b, d) and (b', d') satisfy $ad - bc = 1$ and $ad' - b'c = 1$, respectively. Then $(b - b', d - d') = (ak, ck)$ for some integer k. Lemma 3.4 now implies

$$|k| \leq 2|\operatorname{Re}(\tau)| + 2\frac{4|\xi| + 1}{\sqrt{3}} \leq 2|\operatorname{Re}(\tau)| + 6|\xi|.$$

Thus, $\mathcal{M}(\xi; x; y; \varepsilon)$ is bounded by

$$\begin{aligned} N(y, \varepsilon) \cdot (2 \cdot (2x + 6|\xi|) + 1) &\leq N(y, \varepsilon) \cdot (4x + 13|\xi|) \\ &\leq 2\left(16\pi \frac{\sqrt{2\operatorname{Im}(\xi)} + \sqrt{y}}{\sqrt{y}} + \frac{100\pi|\xi|^3\varepsilon^{1/2}}{y\operatorname{Im}(\xi)}\right)(4x + 13|\xi|)\,. \end{aligned} \tag{3.13}$$

We now want to apply this result to estimate the number of points close to a fixed

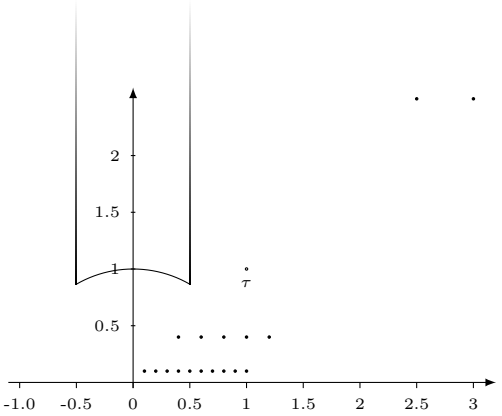


Figure 3.3: τ_0 and all except one τ_M for $N = 10$

point which are given by a cyclic isogeny of degree N. Let $\tau_0 \in \mathbb{H}$ be fixed. Let $N \in \mathbb{N}$. We will be working with matrices M of the form $\left(\begin{smallmatrix} m & l \\ 0 & n \end{smallmatrix}\right)$ with $N = mn$ and $0 \le l < n$. We will denote $M.\tau_0$ by τ_M. We want to bound the number of points τ_M satisfying $|\tilde{\tau}_M - \xi| \le \varepsilon$ with $\tilde{\tau}_M$ in the $\mathrm{SL}_2(\mathbb{Z})$–orbit of τ_M and in $\bar{\mathcal{F}}$. For this we momentarily fix a divisor n of N with $n \ge \sqrt{N}$ and a matrix M with $M = \left(\begin{smallmatrix} m & l \\ 0 & n \end{smallmatrix}\right)$ and $0 \le l < n$. Then $y := \operatorname{Im}(\tau_M) = \frac{m}{n}\operatorname{Im}(\tau_0)$ for any $0 \le l < n$. Figure 3.1 shows an example with $\tau = 1 + i$ and $N = 10$.

Since y does not depend on l and $|\operatorname{Re}(\tau_M)| \le |\operatorname{Re}(\tau_0)|+1$, the bound for $\mathcal{M}(\xi; |\operatorname{Re}(\tau_0)|+1; y; \varepsilon)$ is independent of l. This number does not estimate all the τ_M that are translated close to ξ as we will see later. The bound in (3.13) translates to

$$\mathcal{M}(\xi; |\operatorname{Re}(\tau_0)| + 1; y; \varepsilon) \le 8 \left(4\pi \frac{\sqrt{2\operatorname{Im}(\xi)} + \sqrt{\frac{m}{n}\operatorname{Im}(\tau_0)}}{\sqrt{\operatorname{Im}(\tau_0)}} \sqrt{\frac{n}{m}} + \frac{25\pi|\xi|^3\varepsilon^{1/2}}{\operatorname{Im}(\tau_0)\operatorname{Im}(\xi)} \frac{n}{m} \right) \cdot (4|\operatorname{Re}(\tau_0)| + 13|\xi| + 4)\,.$$

But $\frac{m}{n} \le 1$ since $n \ge \sqrt{N}$ and hence

$$\begin{aligned} &\mathcal{M}(\xi; |\operatorname{Re}(\tau_0)| + 1; y; \varepsilon) \\ &\quad \le 8\,(4|\operatorname{Re}(\tau_0)| + 17|\xi|) \left(4\pi \frac{\sqrt{2\operatorname{Im}(\xi)} + \sqrt{\operatorname{Im}(\tau_0)}}{\sqrt{\operatorname{Im}(\tau_0)}} \sqrt{\frac{n}{m}} + \frac{25\pi|\xi|^3\varepsilon^{1/2}}{\operatorname{Im}(\tau_0)\operatorname{Im}(\xi)} \frac{n}{m} \right) \\ &\quad \le 8\,(4|\operatorname{Re}(\tau_0)| + 17|\xi|) \left(8\pi \frac{\sqrt{2\operatorname{Im}(\xi)} + \sqrt{\operatorname{Im}(\tau_0)}}{\sqrt{\operatorname{Im}(\tau_0)}} \sqrt{\frac{n}{m}} + \frac{35\pi|\xi|^2\varepsilon^{1/2}}{\operatorname{Im}(\tau_0)} \frac{n}{m} \right) \end{aligned}$$

if we also apply $|\xi|^2/\operatorname{Im}(\xi) \le 4|\xi|/3$. Further we get

$$\begin{aligned}\mathcal{M}(\xi; |\operatorname{Re}(\tau_0)| + 1; y; \varepsilon) \le 64\pi\,(4|\operatorname{Re}(\tau_0)| + 17|\xi|) \\ \cdot \max\left\{ \frac{\sqrt{2\operatorname{Im}(\xi)} + \sqrt{\operatorname{Im}(\tau_0)}}{\sqrt{\operatorname{Im}(\tau_0)}}, 5\frac{|\xi|^2}{\operatorname{Im}(\tau_0)} \right\} \\ \cdot \left(\sqrt{\frac{n}{m}} + \frac{n}{m}\varepsilon^{1/2} \right).\end{aligned} \tag{3.14}$$

Now different τ_M (entry l different) can be translated close to ξ by the same matrix, so we have to restrict those. So if τ_M is translated into the disc around ξ by a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ then the real part x of τ_M satisfies

$$|cx + d|\,|c| \le 3|\xi| \quad \text{and} \quad |ax + b|\,|a| \le 3|\xi|$$

by (3.4) and (3.6). Assume that $c \neq 0$. Then $|x + d/c| \le 3|\xi|c^{-2}$, so that x lies in an interval I with center $-d/c$ and of length bounded by $6|\xi|c^{-2}$. This implies

$$|\{l \in \{0, \dots, n-1\} : (m\operatorname{Re}(\tau_0) + l)/n \in I\}| \le n|I| + 1 \le 6|\xi|\frac{n}{c^2} + 1.$$

A similar result is obtained if $a \neq 0$. So in any case

$$|\{l \in \{0, \dots, n-1\} : (m\operatorname{Re}(\tau_0) + l)/n \in I\}| \le 6|\xi|\frac{n}{\max\{|a|, |c|\}^2} + 1 \tag{3.15}$$

independent of whether the interval is centered around $-d/c$ or $-b/a$. Moreover, $\max\{|a|, |c|\}^2$ can be bounded by

$$\begin{aligned}3|\xi|^2 \max\{|a|, |c|\}^2 &\ge a^2 + \nu 2|\operatorname{Re}(\xi)|ac + |\xi|^2c^2 \\ &\ge \frac{\operatorname{Im}(\xi) - 50|\xi|^3\varepsilon^{1/2}}{y} = \frac{\operatorname{Im}(\xi) - 50|\xi|^3\varepsilon^{1/2}}{\operatorname{Im}(\tau_0)}\frac{n}{m}\,,\end{aligned}$$

where the last inequality follows from Equation (3.1). Using the upper bound on ε we obtain

$$3|\xi|^2 \max\{|a|, |c|\}^2 \ge \frac{\operatorname{Im}(\xi)}{2\operatorname{Im}(\tau_0)}\frac{n}{m}\,,$$

and hence

$$6|\xi|\frac{n}{\max\{|a|, |c|\}^2} \le 36|\xi|^3\frac{\operatorname{Im}(\tau_0)}{\operatorname{Im}(\xi)}m \le 50|\xi|^2\operatorname{Im}(\tau_0)m \tag{3.16}$$

since $|\xi|/\operatorname{Im}(\xi) \le 2/\sqrt{3}$.

Recall that the matrix M is of the form $\left(\begin{smallmatrix} m & l \\ 0 & n \end{smallmatrix}\right)$ with $N = mn$ and $0 \le l < n$ and

$\tau_M = M\tau_0$. As before, $\tilde{\tau}_M$ is in the $\mathrm{SL}_2(\mathbb{Z})$–orbit of τ_M and in $\bar{\mathcal{F}}$.

Let $\Lambda(\tau_0; N; \varepsilon)$ be the set of τ_M satisfying $|\tilde{\tau}_M - \xi| \leq \varepsilon$, where τ_M is as before. The number of elements in $\Lambda(\tau_0; N; \varepsilon)$ is surely bounded by the number of matrices M with lower right entry greater than $\sqrt{N}$ satisfying the condition plus the total number of matrices with $n \leq \sqrt{N}$. The latter is bounded by

$$\sum_{\substack{n|N \\ 0<n\leq\sqrt{N}}} n \leq \sqrt{N} \sum_{n|N} 1 = \sqrt{N}\sigma_0(N).$$

For $n \leq \sqrt{N}$ we are going to count $\tilde{\tau}_M$ independent of whether $|\tilde{\tau}_M - \xi| \leq \varepsilon$ or not since the number $\sqrt{N}\sigma_0(N)$ does not grow too fast. Now by the arguments we just made, we can bound the number of τ_M and thus the total number of points in $\Lambda(\tau_M; N; \varepsilon)$ as follows. Recall that $N = mn$.

$$\begin{aligned}|\Lambda(\tau_0; N; \varepsilon)| \leq \sum_{\substack{n|N \\ n\geq\sqrt{N}}} \mathcal{M}\left(\xi; |\operatorname{Re}(\tau_0)| + 1; \frac{N}{n^2}\operatorname{Im}(\tau_0); \varepsilon\right)\left(6|\xi|\frac{n}{\max\{|a|,|c|\}^2} + 1\right) \\ + \sqrt{N}\sigma_0(N).\end{aligned}$$

Here, for fixed n the number $\mathcal{M}\left(\xi; |\operatorname{Re}(\tau_0)| + 1; \frac{N}{n^2}\operatorname{Im}(\tau_0); \varepsilon\right)$ bounds the matrices that translate any τ_M of the form $\left(\begin{smallmatrix} m & l \\ 0 & n \end{smallmatrix}\right)$ with varying l close to ξ. But since different τ_M can be translated into the disc around ξ by the same matrix we have to compensate this with the inequality in (3.15). This in turn can be estimated as displayed in (3.16) so that

$$\begin{aligned}|\Lambda(\tau_0; N; \varepsilon)| \leq \sum_{\substack{n|N \\ n\geq\sqrt{N}}} \mathcal{M}\left(\xi; |\operatorname{Re}(\tau_0)| + 1; \frac{N}{n^2}\operatorname{Im}(\tau_0); \varepsilon\right)\left(50|\xi|^2\operatorname{Im}(\tau_0)\frac{N}{n} + 1\right) \\ + \sqrt{N}\sigma_0(N).\end{aligned}$$

By the inequality for $\mathcal{M}\left(\xi; |\operatorname{Re}(\tau_0)| + 1; \frac{N}{n^2}\operatorname{Im}(\tau_0); \varepsilon\right)$ in (3.14) and $1 \leq m$ we get

$$|\Lambda(\tau_0; N; \varepsilon)| \leq \sqrt{N}\sigma_0(N) + \sum_{\substack{m|N \\ m\leq\sqrt{N}}} m \cdot \mathcal{I}(\tau_0, \xi) \cdot \left(\sqrt{\frac{N}{m^2}} + \frac{N}{m^2}\varepsilon^{1/2}\right)$$

where

$$\mathcal{I}(\tau_0,\xi) = 64\pi\left(4|\operatorname{Re}(\tau_0)| + 17|\xi|\right)\left(50|\xi|^2\operatorname{Im}(\tau_0)+1\right) \cdot \max\left\{\frac{\sqrt{2\operatorname{Im}(\xi)}+\sqrt{\operatorname{Im}(\tau_0)}}{\sqrt{\operatorname{Im}(\tau_0)}}, 5\frac{|\xi|^2}{\operatorname{Im}(\tau_0)}\right\}.$$

We can continue the estimate

$$\begin{aligned}|\Lambda(\tau_0;N;\varepsilon)| &\le \sqrt{N}\sigma_0(N) + \mathcal{I}(\tau_0,\xi)\sum_{\substack{m|N\\ m\le\sqrt{N}}} m\left(\frac{\sqrt{N}}{m}+\frac{N}{m^2}\varepsilon^{1/2}\right)\\ &= \sqrt{N}\sigma_0(N) + \mathcal{I}(\tau_0,\xi)\sum_{\substack{m|N\\ m\le\sqrt{N}}}\left(\sqrt{N}+\frac{N}{m}\varepsilon^{1/2}\right).\end{aligned}$$

We split the sum to get

$$\begin{aligned}|\Lambda(\tau_0;N;\varepsilon)| &\le \sqrt{N}\sigma_0(N) + \mathcal{I}(\tau_0,\xi)\left(\sum_{\substack{m|N\\ m\le\sqrt{N}}}\sqrt{N} + \sum_{\substack{m|N\\ m\le\sqrt{N}}}\frac{N}{m}\varepsilon^{1/2}\right)\\ &= \sqrt{N}\sigma_0(N)(1+\mathcal{I}(\tau_0,\xi)) + \mathcal{I}(\tau_0,\xi)\varepsilon^{1/2}\sum_{\substack{n|N\\ n\ge\sqrt{N}}} n\\ &= \sqrt{N}\sigma_0(N)(1+\mathcal{I}(\tau_0,\xi)) + \mathcal{I}(\tau_0,\xi)\varepsilon^{1/2}\sigma_1(N). \qquad (3.17)\end{aligned}$$

Lemma 3.8 *Let N be a positive integer. Then $\sigma_1(N) \le \frac{\pi^2}{6}\psi(N)$.*

Proof. It is well–known that σ_1 is multiplicative. The function ψ is also multiplicative, see page 53 of [Lan87]. We have $\sigma_1(p^k) = \frac{p^{k+1}-1}{p-1}$ and $\psi(p^k) = p^{k-1}(p+1)$. Thus

$$\frac{\sigma_1(p^k)}{\psi(p^k)} = \frac{p^{k+1}-1}{(p-1)(p+1)p^{k-1}} \le \frac{p^2}{p^2-1} = \frac{1}{1-p^{-2}}.$$

For general N this yields

$$\frac{\sigma_1(N)}{\psi(N)} \le \prod_{p^k\|N}\frac{1}{1-p^{-2}} \le \prod_p \frac{1}{1-p^{-2}} = \zeta(2),$$

where ζ denotes the Riemann zeta function. This proves the claim. □

This proves the following claim.

Lemma 3.9 *Fix $\tau_0 \in \mathbb{H}$ and $\xi \in \bar{\mathcal{F}}$. Let $0 < \varepsilon \leq \left(\frac{\operatorname{Im}(\xi)}{100|\xi|^3}\right)^2$. Let $\Lambda(\tau_0; N; \varepsilon)$ be the number of $M\tau_0$, $M = \left(\begin{smallmatrix} m & l \\ 0 & n \end{smallmatrix}\right)$ with $N = mn$ and $0 \leq l < n$, that satisfy $|\gamma M\tau_0 - \xi| \leq \varepsilon$ for some $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Then*

$$|\Lambda(\tau_0; N; \varepsilon)| \leq \sqrt{N}\sigma_0(N)(1 + \mathcal{I}(\tau_0, \xi)) + \frac{\pi^2}{6}\mathcal{I}(\tau_0, \xi)\varepsilon^{1/2}\psi(N).$$

with

$$\mathcal{I}(\tau_0, \xi) = 64\pi\left(4|\operatorname{Re}(\tau_0)| + 17|\xi|\right)\left(50|\xi|^2 \operatorname{Im}(\tau_0) + 1\right) \cdot \max\left\{\frac{\sqrt{2\operatorname{Im}(\xi)} + \sqrt{\operatorname{Im}(\tau_0)}}{\sqrt{\operatorname{Im}(\tau_0)}}, 5\frac{|\xi|^2}{\operatorname{Im}(\tau_0)}\right\}.$$

To complete the proof of Proposition 3.5 we restrict τ_0 to the fundamental domain. Then $\operatorname{Im}(\tau_0) \geq \sqrt{3}/2$ and $|\operatorname{Re}(\tau_0)| \leq 1/2$. Therefore we get

$$\frac{\sqrt{2\operatorname{Im}(\xi)} + \sqrt{\operatorname{Im}(\tau_0)}}{\sqrt{\operatorname{Im}(\tau_0)}} = 1 + \frac{\sqrt{2\operatorname{Im}(\xi)}}{\sqrt{\operatorname{Im}(\tau_0)}} \leq 1 + 2\sqrt{\operatorname{Im}(\xi)} \leq 6|\xi|^2$$

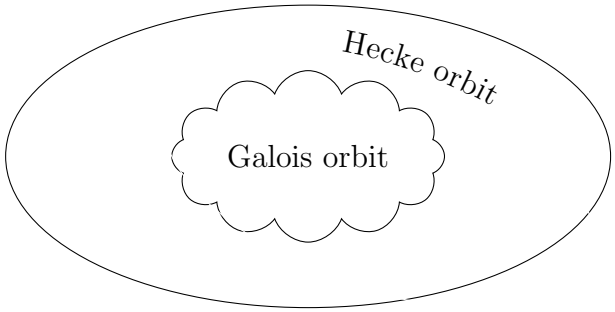
and hence

$$\max\left\{1 + \mathcal{I}(\tau_0, \xi), \frac{\pi^2}{6}\mathcal{I}(\tau_0, \xi)\right\} \leq 64\pi\left(2 + 17|\xi|\right) \cdot 60|\xi|^2|\tau_0| \cdot 6|\xi|^2 \leq 10^7|\tau_0||\xi|^5.$$

Recall from Chapter 1 that an N–isogeny is related to a matrix of the form $M = \left(\begin{smallmatrix} m & l \\ 0 & n \end{smallmatrix}\right)$ with $N = mn$, $0 \leq l < n$ and $\gcd(m, n, l) = 1$. We have considered such matrices without a condition on the greatest common divisor. Therefore we are done with the proof of Proposition 3.5.

3.2 Bounding the height

Recall that we have fixed an elliptic curve without complex multiplication defined over a number field K and j_0 is its j–invariant. Two points in the fundamental domain are in the same Hecke orbit if there exists an isogeny between them. We are going to compare the Galois orbit of j_0 to the Hecke orbit of all conjugates of E_0. We now want to bound the number of elements in $\Gamma(\xi, \varepsilon)$. For this we use the connection between the isogeny orbit and the Galois orbit of Serre's open image theorem. See Théorème 3 in §4 of [Ser72].



More precisely, we will be using a version proved by Lombardo [Lom15], that gives us an explicit bound. Let $G_K = \mathrm{Gal}(\bar{K}/K)$ be the absolute Galois group of K. Recall that G_K acts on the N–torsion points of N, and we thus get a representation

$$\rho_N \colon G_K \to \mathrm{Aut}(E[N]).$$

The group $\mathrm{Aut}(E[N])$ is isomorphic to $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$. It is possible to choose a suitable basis of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ so that we obtain a representation

$$\rho_\infty \colon G_K \to \mathrm{GL}_2(\hat{\mathbb{Z}})$$

after taking the inverse limit (over N.) Serre proved in [Ser72] that $[\mathrm{GL}_2(\hat{\mathbb{Z}}) : \rho_\infty(G_K)]$ is finite. The result by Lombardo implies

$$\left[\mathrm{GL}_2(\hat{\mathbb{Z}}) : \rho_\infty(G_K)\right] < \gamma_1 \cdot [K:\mathbb{Q}]^{\gamma_2} \cdot \max\left\{1, h(E), \log[K:\mathbb{Q}]\right\}^{2\gamma_2} \tag{3.18}$$

where $\gamma_1 = \exp\left(10^{21483}\right)$ and $\gamma_2 = 2.4 \cdot 10^{10}$. In particular, we obtain

$$[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(G_K)] < \gamma_1 \cdot [K:\mathbb{Q}]^{\gamma_2} \cdot \max\left\{1, h(E), \log[K:\mathbb{Q}]\right\}^{2\gamma_2}.$$

Note that Lombardo's result actually uses the original definition of the Faltings height. This information was acquired through a private conversation with the author. Since the original definition of the Faltings height is smaller than the height of E we defined in Chapter 1, we can just substitute $h(E)$ into Lombardo's result.

The cyclic isogenies of degree N correspond in a one–to–one fashion to the cyclic subgroups of order N in $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z} \simeq E[N]$. The action of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ on these subgroups is transitive as the next lemma states. We start with some group theory. We denote by φ Euler's totient function given by $\varphi(N) = \#(\mathbb{Z}/N\mathbb{Z})^\times = N\prod_{p|N}(1-1/p)$, where the product runs over all primes p dividing N. Recall $\psi(N) = N\prod_{p|N}(1+1/p)$.

Lemma 3.10 *The cardinality of* $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ *is equal to* $\varphi(N)^2\psi(N)N$. *Let* $\Delta \subseteq \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ *denote the subgroup of upper triangular matrices. Then* $\#\Delta = N\varphi(N)^2$.

There are $\psi(N)$ cyclic subgroups of $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$. The group $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ acts transitively on the cyclic subgroups of order N in $(\mathbb{Z}/N\mathbb{Z})^2$.

Lemma 3.11 *Let E/K be an elliptic curve, N an integer, and $\Phi \subseteq E[N]$ a cyclic subgroup of N–torsion points. Put $B = |\{\sigma(\Phi) : \sigma \in \mathrm{Gal}(\bar{K}/K)\}|$. Then we have*

$$\frac{\psi(N)}{B} \leq [\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(G_K)]\,.$$

Proof. Suppose Φ is generated by $P \in E[N]$. After choosing a basis, we may assume that P corresponds to $(1,0)$ in $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$. For any $\sigma \in \mathrm{Gal}(\bar{K}/K)$, the group $\sigma(\Phi)$ is generated by a point $(a,c) \in Z/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$ where $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is the image of σ under ρ_N. Let Δ be the subgroup of upper triangular matrices of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$. The equality $\sigma(\Phi) = \Phi$ holds if and only if σ is mapped into Δ under ρ_N. We thus have

$$B = \frac{\#\operatorname{im}\rho_N}{\#(\Delta \cap \operatorname{im}\rho_N)} \geq \frac{\#\operatorname{im}\rho_N}{\#\Delta} = \frac{\#\operatorname{im}\rho_N}{N\varphi(N)^2}.$$

This implies

$$\begin{aligned}\frac{\psi(N)}{B} \leq \frac{\psi(N)\varphi(N)^2 N}{\#\operatorname{im}\rho_N} &= \frac{\psi(N)\varphi(N)^2 N}{\#\,\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})}\,[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(G_K)] \\ &= [\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(G_K)]\,,\end{aligned}$$

as desired. □

We want to estimate a Mertens' type of sum. In fact, we are going to use a result by Mertens.

Lemma 3.12 *Let $n \geq 4$ be a positive integer. Then*

$$\sum_{p|n} \frac{\log p}{p} \leq 5.25 \log\log n,$$

where the sum runs over all prime divisors of n.

Proof. The function $\log x/x$ is decreasing on (e,∞). Note that $(\log 2)/2 < (\log 3)/3$. So let $n = p^a$ be a prime power with $p \neq 2$. Then

$$\frac{\log p}{p} \leq \frac{\log 3}{3} \leq 5 \log\log 3 \leq 5 \log\log p$$

and the claim holds. If $n = 2^a$ with $a \geq 2$, then

$$\frac{\log 2}{2} < 1 \leq 5 \log\log(4).$$

Now let $n = p^a q^b$ with different primes p, q and $a, b \geq 1$. We have $(\log 5)/5 < (\log 2)/2$ and $(\log p)/p < 0.5$. Thus

$$\frac{\log p}{p} + \frac{\log q}{q} \leq \frac{\log 2}{2} + \frac{\log 3}{3} \leq 1 < 5 \log\log 6 \leq 5 \log\log n.$$

So the claim is true for all $4 \leq n \leq 29$. let us now assume that n is composite with $\omega(n) \geq 3$. We can bound the sum by looking at the first $\omega(n)$ primes

$$\sum_{p|n} \frac{\log p}{p} \leq \frac{\log p_1}{p_1} + \frac{\log p_2}{p_2} + \cdots + \frac{\log p_{\omega(n)}}{p_{\omega(n)}}.$$

Note that $(\log 2)/2 < (\log 3)/3$, so that if 3 occurs in the prime decomposition of n and 2 does not, we can just estimate the largest prime divisor of n by $(\log 2)/2$ and get the same inequality. It is a well–known result by Cipolla in [Cip02], that the n–th prime p_n is bounded from above by $n(\log n + \log\log n)$ for sufficiently large n. Indeed Rosser proved in Theorem 2 of [Ros39] that $p_n \leq n(\log n + 2\log\log n)$ for all $n \geq 4$. Also compare to the bound in [RS62]. Hence $p_n \leq 2n \log n$ for all $n \geq 3$ since this bound also holds for $p_3 = 5$. Since we have $\omega(n) \geq 3$ we can apply this to the last inequality to obtain

$$\sum_{p|n} \frac{\log p}{p} \leq \sum_{p \leq 2\omega(n)\log\omega(n)} \frac{\log p}{p}.$$

By Mertens' Theorem (see [Mer74]) the sum on the right–hand side is bounded by

$$\sum_{p \leq 2\omega(n)\log\omega(n)} \frac{\log p}{p} \leq 2\log(2\omega(n)\log\omega(n))$$

for all $n \geq 1$ composite of at least 3 distinct primes. We have the trivial inequality

$$\omega(n) \leq \frac{\log n}{\log 2}.$$

This gives us

$$\begin{aligned}\sum_{p|n} \frac{\log p}{p} &\leq 2\log\left(2\frac{\log n}{\log 2}\log\frac{\log n}{\log 2}\right) \\ &\leq 2\log\log n + 2\log\left(\log\log n - \log\log 2\right) + 2.12\end{aligned}$$

and if $n \geq 5$ this gets us

$$\begin{aligned}\sum_{p|n} \frac{\log p}{p} &\leq 2\log\log n + 2\log(2\log\log n) + 2.12 \\ &\leq 2\log\log n + 2\log\log\log n + 3.51.\end{aligned}$$

But we have $\log\log n \leq \frac{36}{100}\log n$ since $x \mapsto (\log\log x)/\log x$ is decreasing for $x \geq 16$ and $(\log\log 30)/(\log 30) < 0.36$. Because of $3.51 + 2\log 0.36 < 1.25\log\log 30$ we obtain

$$\sum_{p \leq 2\omega(n)\log\omega(n)} \frac{\log p}{p} \leq 5.25\log\log n,$$

as desired. □

Proposition 3.13 *Let $E/\bar{\mathbb{Q}}$ and E_0/K be elliptic curves without CM such that there exists a cyclic isogeny of degree N from E_0 to E. Let ρ_N be the Galois representation associated to E_0. If $N \geq 4$, we have*

$$h(E) \geq h(E_0) + \frac{1}{2}\log N - 7 \cdot [\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(G_K)]\log\log N.$$

Proof. We denote by $h(E)$ and $h(E_0)$ the stable Faltings height of E and E_0, respectively. (The stated inequality does not depend on the normalization of the Faltings height.) We consider the action of $\mathrm{Gal}(\bar{\mathbb{Q}}/K)$ on the set of $\bar{\mathbb{Q}}$–isomorphism classes of elliptic curves. Let $E = E_1, \ldots, E_{\psi(N)}$ be representatives of elliptic curves that are N–isogenous to E_0. Note that the group $\mathrm{Gal}(\bar{\mathbb{Q}}/K)$ acts on the set $\{E_1, \ldots, E_{\psi(N)}\}$. By Corollaire 3.3 in [Aut03] we have

$$\frac{1}{\psi(N)}\sum_{i=1}^{\psi(N)} h(E_i) = h(E_0) + \frac{1}{2}\log N - \lambda_N$$

where $N = p_1^{\alpha_1}\cdots p_r^{\alpha_r}$ and $\lambda_N = \sum_{i=1}^{r} \frac{p_i^{\alpha_i}-1}{(p_i^2-1)p_i^{\alpha_i-1}}\log p_i$. Rearranging and using $|h(E_0) - h(E_i)| \leq 1/2\log N$ (e.g. [Ray85, Corollaire 2.1.4, page 207]) we obtain

$$\begin{aligned}\frac{n_1}{\psi(N)}h(E_1) &\geq h(E_0) + \frac{1}{2}\log N - \lambda_N - \sum_j \frac{n_j}{2\psi(N)}\log N - \sum_j \frac{n_j}{\psi(N)}h(E_0) \\ &= h(E_0) + \frac{1}{2}\log N - \lambda_N - \frac{\psi(N)-n_1}{2\psi(N)}\log N - \frac{\psi(N)-n_1}{\psi(N)}h(E_0) \\ &= \frac{n_1}{\psi(N)}h(E_0) + \frac{n_1}{2\psi(N)}\log N - \lambda_N,\end{aligned}$$

where we have grouped curves into $\mathrm{Gal}(\bar{\mathbb{Q}}/K)$–orbits, each of size n_j. The number n_1 is the number of elliptic curves up to $\bar{\mathbb{Q}}$–isomorphism that are in the $\mathrm{Gal}(\bar{\mathbb{Q}}/K)$–orbit of E_1. This implies

$$h(E_1) \geq h(E_0) + \frac{1}{2}\log N - \frac{\psi(N)}{n_1}\lambda_N.$$

We have

$$\frac{p_i^{\alpha_i}-1}{(p_i^2-1)p_i^{\alpha_i-1}} \leq \frac{p_i^{\alpha_i}}{(p_i^2-1)p_i^{\alpha_i-1}} = \frac{p_i}{p_i^2-1} \leq \frac{4}{3p_i}.$$

It follows from the last lemma and Lemma 3.11 that

$$h(E_1) \geq h(E_0) + \frac{1}{2}\log N - \frac{21}{3}\left[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(G_K)\right]\log\log N,$$

as desired. □

Corollary 3.14 *In the setting of the previous proposition, let j_0 and j be the j–invariants of E_0 and E, respectively. We have*

$$\begin{aligned} h(j_0) - 6\log(1+h(j_0)) + 6\log N - 84\left[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(G_K)\right]\log\log N \\ \leq h(j) + 16.212 \end{aligned}$$

Proof. Compare the proof to Proposition 2.1 in [Sil86a]. Using Proposition 3.2 of [Löb17] in the first step and Lemme 7.8 of [GR11] on the third we obtain

$$\begin{aligned} \frac{1}{12}h(j_0) - \frac{1}{2}\log(1+h(j_0)) - 2.071 + \frac{1}{2}\log N - 7\left[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(G_K)\right]\log\log N \\ \leq h(E_0) + \frac{1}{2}\log N - 7\left[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(G_K)\right]\log\log N + \frac{1}{2}\log\pi \\ \leq h(E) + \frac{1}{2}\log\pi \\ \leq \frac{1}{12}h(j) - \frac{1}{2}\log(1+h(j)) - 0.72 \\ \leq \frac{1}{12}h(j) - 0.72. \end{aligned}$$

Note that the authors of both cited papers use the normalization of the Faltings height of Deligne. Multiplying the inequality by 12 and rearranging the terms yields the desired inequality. □

In the proof of the next lemma we will use the function

$$D(z) = \max\{1, |\mathrm{Re}(z)|, \mathrm{Im}(z)^{-1}\}, \qquad \text{for all } z \in \mathbb{H}.$$

It appears in [HP12]. Note that if z is in $\bar{\mathcal{F}}$, then $D(z) \leq 2/\sqrt{3}$. The height of an

element in $\mathrm{Mat}_2(\mathbb{Q})$ will be the height of that element when regarded as a member of $\mathbb{Q}^4$. The following result can be found in [HP12].

Lemma 3.15 *If $z \in \mathbb{H}$, then there is a $\rho \in \mathrm{SL}_2(\mathbb{Z})$ with $H(\rho) \leq 264 D(z)^9$ and $\rho z \in \mathcal{F}$.*

Proof. Write $z = x + iy$ with $x, y \in \mathbb{R}$ and $y > 0$. We assume $y \leq 1/2$ for now and define $Q = y^{-1/2}$. We can find integers $c, d \in \mathbb{Z}$ with $1 \leq c < Q$ and $|cx + d| \leq 1/Q$ by page 1 of [Cas57]. There is no restriction in assuming that c and d are coprime. Also we can find a, b with $ad - bc = 1$ and $\max\{|a|, |b|\} \leq \max\{|c|, |d|\}$. Put $\beta = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Then $H(\beta) = \max\{|c|, |d|\}$. By definition we have $|c| < y^{-1/2} \leq D(z)^{1/2}$ and $|d| \leq |c||x| + Q^{-1} \leq D(z)^{1/2}|x| + 1 \leq 2D(z)^{3/2}$. This implies

$$H(\beta) \leq 2D(z)^{3/2}. \tag{3.19}$$

If $c \neq 0$ then $|cz + d| \geq |\operatorname{Im}(cz + d)| = |c| \operatorname{Im}(z) \geq D(z)^{-1}$. But if $c = 0$, then $d = \pm 1$, so

$$|cz + d| \geq D(z)^{-1} \tag{3.20}$$

holds in any case. Now since $y \leq 1/2$ we have

$$|z|^2 \leq x^2 + 1/4 \leq 2D(z)^2. \tag{3.21}$$

This implies $|az + b| \leq |a||z| + |b| \leq 2|a|D(z) + |b| \leq 3H(\beta)D(z)$. Using this and (3.19) we obtain

$$|\beta z| = \frac{|az + b|}{|cz + d|} \leq 3H(\beta)D(z)^2 \leq 6D(z)^{7/2}.$$

By the usual transformation law of the imaginary part under fractional linear transformations we obtain $\operatorname{Im}(\beta z) = \operatorname{Im}(z)|cz+d|^{-2} = y((cx+d)^2+c^2y^2)^{-1} \geq y(Q^{-2}+Q^2y^2)^{-1} = y(y + y)^{-1}$, i.e.

$$\operatorname{Im}(\beta z) \geq \frac{1}{2}. \tag{3.22}$$

If we apply (3.20) we obtain for the real part of βz

$$|\operatorname{Re}(\beta z)| = \frac{|ac|z|^2 + bd + x(ad + bc)|}{|cz + d|^2} \leq 4H(\beta)^2 D(z)^2 \max\{1, |x|, |z|^2\}. \tag{3.23}$$

But $|x| \leq |z|$ so that by (3.21) we get $|\operatorname{Re}(\beta z)| \leq 8H(\beta)^2 D(z)^4$. In total we obtain together with (3.19)

$$|\operatorname{Re}(\beta z)| \leq 32 D(z)^7. \tag{3.24}$$

We have modified z so that $\operatorname{Im}(\beta z) \geq 1/2$. If $y > 1/2$, then equations (3.19), (3.22) and (3.24) are true with β the identity matrix. There exists a matrix $\gamma = \left(\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ that translates βz into the vertical strip $\{\tau \in \mathbb{H}; -1/2 < \operatorname{Re}(\tau) \leq 1/2\}$. The matrix γ does

not change the imaginary part of βz. We have $t \leq \mathrm{Re}(\beta z)| + 1/2 \leq 33D(z)^7$ by applying (3.24). This implies $H(\gamma) \leq \max\{1, |t|\} \leq 33D(z)^7$. Now since $|\,\mathrm{Re}(\gamma\beta z)| \leq 1/2$ and $\mathrm{Im}(\gamma\beta z) \geq 1/2$ we get $\delta\gamma\beta z \in \mathcal{F}$ for some

$$\delta \in \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} \pm 1 & -1 \\ 1 & 0 \end{pmatrix} \right\}.$$

By construction $\rho = \delta\gamma\beta \in \mathrm{SL}_2(\mathbb{Z})$ and $\rho z \in \mathcal{F}$. Also since $H(\delta) = 1$ we obtain

$$H(\rho) \leq 2H(\delta)H(\gamma\beta) \leq 4H(\gamma)H(\beta) \leq 4 \cdot 33D(z)^7 \cdot 2D(z)^{3/2} \leq 264D(z)^{17/2}$$

by (3.19) and the height estimate for γ. □

Let N, m, n, l be integers satisfying $1 < N = mn$ and $0 \leq l < n$. Then we have $|\mathrm{Re}\left(\frac{m\tau+l}{n}\right)| \leq N\left(|\mathrm{Re}(\tau)| + 1\right) \leq N(D(\tau) + 1)$, and we can similarly bound the inverse of the imaginary part by $N(D(\tau) + 1)$. Thus

$$D\left(\frac{m\tau + l}{n}\right) \leq N(D(\tau) + 1). \tag{3.25}$$

Lemma 3.16 *Let $E_0\colon y^2 = 4x^3 - g_2x - g_3$ be the Weierstrass form of an elliptic curve without complex multiplication defined over a number field K. Let j_0 be the j–invariant of E_0 and put $h = \max\{1, h(1, g_2, g_3), h(j_0)\}$. Let ω_1 and ω_2 be periods of the elliptic curve such that $\tau_0 = \omega_2/\omega_1$ is in $\mathcal{F}$. Suppose that ξ is an algebraic number of degree 2. Let N, m, n, l be integers satisfying*

$$N \geq \left(\max\{e^{6\pi|\tau_0|/[K:\mathbb{Q}]}, e^{e \cdot h}, [K : \mathbb{Q}], (4 \cdot 10^{11}H(\xi))^{20}\}\right)^{1/20} =: \mathcal{N}(E_0, \xi),$$

$N = mn$ and $0 \leq l < n$. Let $\rho \in \mathrm{SL}_2(\mathbb{Z})$ satisfy $\rho\left(\begin{smallmatrix} m & l \\ 0 & n \end{smallmatrix}\right).\tau_0 \in \bar{\mathcal{F}}$. Write $\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right) = \rho\left(\begin{smallmatrix} m & l \\ 0 & n \end{smallmatrix}\right)$. Then there exists an explicit constant $c_1' \geq 1$ such that

$$\log|(\alpha - \xi\gamma)\omega_2 + (\beta - \xi\delta)\omega_1| \geq -c_1' \cdot (\log N)^4.$$

The constant c_1' depends on the elliptic curve E_0.

Proof. This is a special case of Théorème 2.1 in [Dav95]. We set $D = [K : \mathbb{Q}]$. Also see [DH09] for a similar result with a computable constant. We put $\mathcal{L}(z_0, z_1, z_2) = (\alpha - \xi\gamma)z_1 + (\beta - \xi\delta)z_2$. Our elliptic curve and the coefficients are in a number field of degree at most $2D$ since ξ is quadratic. Note that $(\alpha - \xi\gamma)\omega_2 + (\beta - \xi\delta)\omega_1 \neq 0$ otherwise we would have

$$\tau_0 = \frac{\xi\delta - \beta}{\alpha - \xi\gamma} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}.\xi$$

i.e. there is a isogeny of degree N between elliptic curves with j–invariant $j(\tau_0)$ and

$j(\xi)$. But one has complex multiplication and the other does not, so this is impossible. We choose the variables u_1, u_2 in the theorem to be ω_2 and ω_1, respectively. Then $\gamma_1 = \gamma_2 = (0, 0, 1)$ and $v = (1, \omega_2, \omega_1)$. We have to estimate the height of the coefficients of the linear form. For this, let H denote the multiplicative height. Let $\rho = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Then $\alpha = ma$ and $\gamma = cm$, and by Lemma 1.4 we obtain

$$\begin{aligned} H(\alpha - \xi\gamma) &\leq 2H(\alpha)H(-\xi\gamma) \\ &\leq 2H(\alpha)H(\xi)H(\gamma) \\ &= 2H(am)H(\xi)H(cm) \\ &\leq 2H(a)H(m)H(\xi)H(c)H(m) \\ &= 2H(a)H(\xi)H(c)m^2. \end{aligned}$$

Now $m \leq N$ and $H(a), H(c) \leq H(\rho)$ so that

$$H(\alpha - \xi\gamma) \leq 2H(\rho)^2 H(\xi)N^2.$$

Note that $\left(\begin{smallmatrix} m & l \\ 0 & n \end{smallmatrix}\right).\tau_0$ does not have CM and is thus not an elliptic point for $\mathrm{SL}_2(\mathbb{Z})$. This means that if $\rho' \in \mathrm{SL}_2(\mathbb{Z})$ transfers the point to the same points as ρ does, then $\rho' = \pm\rho$. Since $H(\rho) = H(-\rho)$ we can use Lemma 3.15 together with (3.25) to obtain

$$H(\rho) \leq 264D\left(\frac{m\tau_0 + l}{n}\right)^9 \leq 264(D(\tau_0) + 1)^9 N^9 \leq 2 \cdot 10^5 N^9, \tag{3.26}$$

because $\tau_0 \in \mathcal{F}$. Altogether we have

$$H(\alpha - \xi\gamma) \leq 8 \cdot 10^{10} H(\xi)N^{20}.$$

We have $\beta = al + bn$ and $\delta = cl + dn$. Recall $0 \leq l < n \leq N$ and thus

$$H(\beta) = |al + bn| \leq |al| + |bn| \leq (|a| + |b|)N \leq 2H(\rho)N$$

and

$$H(\delta) = |cl + dn| \leq |cl| + |dn| \leq (|c| + |d|)N \leq 2H(\rho)N.$$

For the height of $H(\beta - \xi\delta)$ we obtain

$$\begin{aligned} H(\beta - \xi\delta) &\leq 2H(\beta)H(\xi)H(\delta) \\ &\leq 8H(\rho)^2 H(\xi)N^2. \end{aligned}$$

Using (3.26) again this yields

$$H(\beta - \xi\delta) \leq 4 \cdot 10^{11} H(\xi)N^{20}.$$

Put $V_1 = V_2 = e^{\max\{h, 6\pi|\tau_0|\}}$. We have

$$\frac{3\pi|u_1|^2}{|\omega_1|^2 \operatorname{Im}(\tau_0) 2D} \leq \frac{3\pi|u_1|^2}{|\omega_1|^2 \operatorname{Im}(\tau_0) D} \leq \frac{3\pi|\tau_0|^2}{\operatorname{Im}(\tau_0) D} \leq 6\pi|\tau_0|$$

and also

$$\frac{3\pi|u_2|^2}{|\omega_1|^2 \operatorname{Im}(\tau_0) D} \leq \frac{3\pi}{\operatorname{Im}(\tau_0) D} \leq 6\pi$$

since $|\tau_0|^2/\operatorname{Im}(\tau_0) \leq 2|\tau_0|/\sqrt{3}$ and $\operatorname{Im}(\tau_0) \geq \sqrt{3}/2$ for $\tau_0 \in \bar{\mathcal{F}}$. Therefore, equation (3) of Théorème 2.1 in [Dav95] is satisfied independently of whether ξ is in K or not. Assume

$$N \geq \left(\max\{e^{6\pi|\tau_0|/D}, e^{e \cdot h}, D, (4 \cdot 10^{11} H(\xi))^{20}\}\right)^{1/20}.$$

Define

$$B = N^{21}.$$

We picked N large enough so that

$$B \geq 4 \cdot 10^{11} H(\xi) N^{20} \geq \max\{e^{eh}, e^{6\pi|\tau_0|/D}, H(\alpha - \xi\gamma), H(\beta - \xi\delta)\}.$$

This implies $B \geq V_1^{1/D} = V_2^{1/D}$. Thus, equations (1) and (2) of the theorem in [Dav95] are satisfied, and we are in the situation of the theorem to obtain as a result the lower bound

$$\begin{aligned} \log|\mathcal{L}(v)| &\geq -C \cdot 2^6 \cdot D^6 (\log B + \log(2D)) \cdot (\log\log B + h + \log(2D))^3 \log V_1 \log V_2 \\ &\geq -C \cdot 2^6 \cdot D^6 \cdot 54 \cdot \max\{h, 6\pi|\tau_0|\}^2 \cdot (\log B)^4 \end{aligned}$$

since $h \leq \log B$ and $\log(2D) \leq \log B$. If we substitute B and take C from [Dav95] we get

$$\begin{aligned} \log|\mathcal{L}(v)| &\geq -C \cdot 2^6 \cdot D^6 \cdot 54 \cdot \max\{h, 6\pi|\tau_0|\}^2 \cdot 21^4 \cdot (\log N)^4 \\ &\geq -4 \cdot 10^{50} \cdot D^6 \cdot \max\{h, 6\pi|\tau_0|\}^2 \cdot (\log N)^4. \end{aligned}$$

This gives the desired inequality of the lemma. □

Recall the definitions

$$\mathcal{F}_+ = \{\tau \in \mathcal{F}; 0 \leq \operatorname{Re}(\tau) \leq 1/2\}$$

and

$$\mathcal{F}_- = \{\tau \in \mathcal{F}; -1/2 \leq \operatorname{Re}(\tau) \leq 0\}.$$

The following lemma can be found in [BLP16]. Also see Lemma 2.10.

Lemma 3.17 *For $\tau \in \mathcal{F}_+$ we either have $|\tau - \zeta| \geq 10^{-3}$ and $|j(\tau)| \geq 4.4 \cdot 10^{-5}$ or $|\tau - \zeta| \leq 10^{-3}$ and*

$$44000|\tau - \zeta|^3 \leq |j(\tau)| \leq 47000|\tau - \zeta|^3.$$

For $\tau \in \mathcal{F}_-$ we either have $|\tau - \zeta^2| \geq 10^{-3}$ and $|j(\tau)| \geq 4.4 \cdot 10^{-5}$ or $|\tau - \zeta^2| \leq 10^{-3}$ and

$$44000|\tau - \zeta|^3 \leq |j(\tau)| \leq 47000|\tau - \zeta|^3.$$

We fix E_0 given by a Weierstrass equation $y^2 = 4x^3 - g_2x - g_3$, and assume it is defined over a number field K. Let j_0 be its j–invariant and pick $\tau_0 \in \mathcal{F}$ with $j(\tau_0) = j_0$. Let E be an elliptic curve with j–invariant j that is N–isogenous to E_0. As before, we set $j(\tau_0^\sigma) = \sigma(j(\tau_0))$ with $\tau_0^\sigma \in \mathcal{F}$ for any field embedding $\sigma\colon K \to \mathbb{C}$. By E_0^σ and E^σ we denote the Galois conjugates of E_0 and E, respectively.

Lemma 3.18 *Let $N \geq \mathcal{N}(E_0^\sigma, \zeta)$. We have $\log|\sigma(j)| \geq -c_1 \cdot (\log N)^6 - c_2$ for any $\mathbb{Q}$–homomorphism $\sigma\colon K \to \mathbb{C}$, where the constants are explicit and only depend on the fixed elliptic curve E_0. We have $c_1 \geq 1$ and we can have $c_2 \geq 0$.*

Proof. We assume $|\sigma(j)| \leq 10^{-3}$ for now. We have an N–isogeny between E_0^σ and E^σ since E_0 and E are N–isogenous. Let $E_0^\sigma(\mathbb{C}) \simeq \mathbb{C}/(\omega_{0,1}^\sigma\mathbb{Z} + \omega_{0,2}^\sigma\mathbb{Z})$ with $\tau_0^\sigma = \omega_{0,2}^\sigma/\omega_{0,1}^\sigma$ in the fundamental domain. Similarly, let τ^σ correspond to $E^\sigma(\mathbb{C})$. We can choose ω_1^σ and ω_2^σ such that $\tau^\sigma = \rho\left(\begin{smallmatrix} m & l \\ 0 & n \end{smallmatrix}\right)\tau_0^\sigma$ and such that τ^σ is in the fundamental domain $\mathcal{F}$. Write $\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right) = \rho\left(\begin{smallmatrix} m & l \\ 0 & n \end{smallmatrix}\right)$. A similar estimate to the one in the proof of Lemma 3.16 shows

$$\begin{aligned} |\gamma\omega_{0,2}^\sigma + \delta\omega_{0,1}^\sigma| &\leq 3N \max\{|\omega_{0,1}^\sigma|, |\omega_{0,2}^\sigma|\}H(\rho) \\ &\leq 792(D(\tau_0^\sigma) + 1)^9 \max\{|\omega_{0,1}^\sigma|, |\omega_{0,2}^\sigma|\}N^{10} \\ &\leq 10^6 \max\{|\omega_{0,1}^\sigma|, |\omega_{0,2}^\sigma|\}N^{10}, \end{aligned} \tag{3.27}$$

since $D(\tau_0^\sigma) \leq 2/\sqrt{3}$. Note that we have $\tau^\sigma \neq \zeta$ since E does not have CM. We have

$$\begin{aligned} \log|\tau^\sigma - \zeta| &= \log\left|\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right).\tau_0^\sigma - \zeta\right| \\ &= \log\left|\frac{\alpha\tau_0^\sigma + \beta}{\gamma\tau_0^\sigma + \delta} - \zeta\right| \\ &= \log\left|\frac{\alpha\omega_{0,2}^\sigma + \beta\omega_{0,1}^\sigma}{\gamma\omega_{0,2}^\sigma + \delta\omega_{0,1}^\sigma} - \zeta\right| \\ &= \log\left(\frac{1}{|\gamma\omega_{0,2}^\sigma + \delta\omega_{0,1}^\sigma|}|\alpha\omega_{0,2}^\sigma + \beta\omega_{0,1}^\sigma - \zeta(\gamma\omega_{0,2}^\sigma + \delta\omega_{0,1}^\sigma)|\right) \\ &= -\log|\gamma\omega_{0,2}^\sigma + \delta\omega_{0,1}^\sigma| + \log|(\alpha - \zeta\gamma)\omega_{0,2}^\sigma + (\beta - \zeta\delta)\omega_{0,1}^\sigma|. \end{aligned} \tag{3.28}$$

We can use (3.27) in the first step and Lemma 3.16 the second to get

$$\begin{aligned}\log|\tau^\sigma-\zeta| &\geq -\log\left(10^6\max\{|\omega_{0,1}^\sigma|,|\omega_{0,2}^\sigma|\}N^{10}\right)+\log|(\alpha-\zeta\gamma)\omega_{0,2}^\sigma+(\beta-\zeta\delta)\omega_{0,1}^\sigma|\\ &\geq -\log\left(10^6\max\{|\omega_{0,1}^\sigma|,|\omega_{0,2}^\sigma|\}N^{10}\right)-c_1'\cdot(\log N)^6,\end{aligned}$$

where c_1' is the constant from Lemma 3.16. The same bound holds for ζ replaced by ζ^2 since $\mathcal{N}(E,\zeta)=\mathcal{N}(E,\zeta^2)$. Assuming that τ^σ is closer to ζ, Lemma 3.17 says

$$|\sigma(j)|=|j(\tau^\sigma)|\geq 44000|\tau^\sigma-\zeta|^3.$$

This implies

$$\begin{aligned}\log|\sigma(j)| =&\log|j(\tau^\sigma)|\geq\log 44000+\log|\tau^\sigma-\zeta|^3\\ \geq&\log 44000-3\log\left(10^6\max\{|\omega_{0,1}^\sigma|,|\omega_{0,2}^\sigma|\}\right)\\ &-10\log N-3c_1'\cdot(\log N)^6\\ \geq&-14-3\log\left(\max\{|\omega_{0,1}^\sigma|,|\omega_{0,2}^\sigma|\}\right)\\ &-2\cdot 10^{51}D^6\cdot\max\{h,6\pi|\tau_0|\}^2\cdot(\log N)^6.\end{aligned}$$

So we can put $c_1=2\cdot 10^{51}D^6\cdot\max\{h,6\pi|\tau_0|\}^2\geq 1$. Since $|\omega_{0,1}^\sigma|$ and $|\omega_{0,2}^\sigma|$ can be small we put $c_2=14+3\log\left(\max\{1,|\omega_{0,1}^\sigma|,|\omega_{0,2}^\sigma|\}\right)$.

If $|\sigma(j)|\geq 10^{-3}$, then

$$\log(|\sigma(j)|)\geq\log(10^{-3})\geq -7>-c_2,$$

so the bound is true. □

Lemma 3.19 *We have* $|\tau_0^\sigma|/[K:\mathbb{Q}]\leq 3\max\{1,h(j_0)\}$.

Proof. Put $D=[K:\mathbb{Q}]$. We have

$$|\tau_0^\sigma|\leq\frac{3}{2}\log\max\{e,|j(\tau_0^\sigma)|\}$$

by Lemme 1 item (iv) in [FP87]. Thus

$$\begin{aligned}\frac{|\tau_0^\sigma|}{D} &\leq \frac{3}{2}\frac{1}{D}\log\max\{e,|j(\tau_0^\sigma)|\}\\ &\leq \frac{3}{2}\frac{1}{D}\left(1+\log\max\{1,|j(\tau_0^\sigma)|\}\right)\\ &= \frac{3}{2}\frac{1}{D}\left(1+\log\max\{1,|\sigma(j(\tau_0))|\}\right)\\ &\leq \frac{3}{2}\frac{1}{D}\left(1+\sum_{\nu\in M_K} d_\nu \log\max\{1,|j(\tau_0)|_\nu\}\right)\\ &\leq \frac{3}{2}+\frac{3}{2}h(j_0).\end{aligned}$$

This gives the desired inequality. □

Proposition 3.20 *Let j_0 and j be j–invariants of elliptic curves, where j_0 is associated to the elliptic curve E_0/K given by $E_0\colon y^2 = 4x^3 - g_2x - g_3$. Put $h = \max\{1, h(1,g_2,g_3), h(j_0)\}$ and $j(\tau_0) = j_0$ with $\tau_0 \in \mathcal{F}$. Assume we have a cyclic isogeny of degree N between E_0 and an elliptic curve corresponding to j. Further assume that j is an algebraic unit. If*

$$N \geq \max\left\{4\cdot 10^{11}, [K:\mathbb{Q}], e^{3h}\right\},$$

then the height of j can be estimated by

$$\begin{aligned}h(j) \leq & 6\cdot 10^7 h[K:\mathbb{Q}][\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}):\rho_N(G_K)]\left(N^{-1/10}+\sqrt{\varepsilon}\right)\left(c_1(\log N)^6+c_2\right)\\ &+3|\log\varepsilon|\end{aligned}$$

where $0<\varepsilon<10^{-5}$ is arbitrary.

Proof. Let E be an elliptic curve corresponding to j, so that there is a cyclic isogeny of degree N between E_0 and E. Let $\Phi \subseteq E_0[N]$ be the kernel of the given isogeny $E_0 \to E$. Put $G = \{\sigma \in \mathrm{Gal}(K(E_0[N])/K); \sigma(\Phi) = \Phi\}$, and let $K^\Phi = \{\alpha \in K(E_0[N]); \sigma(\alpha) = \alpha \text{ for all } \sigma \in G\}$ be the fixed field of Φ. By basic Galois theory we have $G = \mathrm{Gal}(K(E_0[N])/K^\Phi)$, and hence $\sigma(\Phi) = \Phi$ for all $\sigma \in \mathrm{Gal}(\overline{K^\Phi}/K^\Phi)$. This implies

$$\begin{aligned}B &= |\{\sigma(\Phi) : \sigma \in \mathrm{Gal}(K(E_0[N])/K)\}|\\ &= \frac{|\mathrm{Gal}(K(E_0[N])/K)|}{|G|}\\ &= |\mathrm{Gal}(K^\Phi/K)| = [K^\Phi : K].\end{aligned}$$

By Remark III.4.13.2 in [Sil86b] the elliptic curve E is defined over K^Φ, and hence

$j \in K^\Phi$. Let D be the degree of K^Φ over $\mathbb{Q}$. Put $\varepsilon_0 = 44000\varepsilon^3$. By (1.6) we have

$$\begin{aligned} h(j) &= -\frac{1}{D}\left(\sum_{|\sigma(j)|<\varepsilon_0} \log|\sigma(j)| + \sum_{\varepsilon_0 \le |\sigma(j)|<1} \log|\sigma(j)|\right) \\ &\le -\frac{1}{D}\sum_{|\sigma(j)|<\varepsilon_0} \log|\sigma(j)| + |\log \varepsilon_0|. \end{aligned} \tag{3.29}$$

Recall the definitions

$$\Gamma_\varepsilon = \{\sigma\colon K \to \mathbb{C}; \tau^\sigma \in \Sigma_\varepsilon\}$$

with

$$\Sigma_\varepsilon = \{\tau \in \mathcal{F}; |j(\tau)| < \varepsilon\}.$$

If $|\sigma(j)| = |j(\tau^\sigma)| < \varepsilon_0 \le 10^{-3}$ and $\tau^\sigma \in \mathcal{F}_+$, then by Lemma 3.17

$$|\tau^\sigma - \zeta|^3 \le \frac{|j(\tau^\sigma)|}{44000} < \frac{\varepsilon_0}{44000} = \varepsilon^3, \tag{3.30}$$

i.e. $|\tau^\sigma - \zeta| < \varepsilon$. If τ^σ is not in $\mathcal{F}_+$ but in $\mathcal{F}_-$, then $|\tau^\sigma - \zeta^2| < \varepsilon$ also follows from $|\sigma(j)| < \varepsilon_0$ and Lemma 3.17. We have $\sigma \in \Gamma_\varepsilon$ since

$$|\sigma(j)| = |j(\tau^\sigma)| < \varepsilon_0 = 47000\varepsilon^3 \le 47000 \cdot 10^{-10}\varepsilon \le \varepsilon.$$

Continuing the estimate of (3.29) this gives

$$\begin{aligned} h(j) &\le -\frac{1}{D}\sum_{|\sigma(j)|<\varepsilon_0} \log|\sigma(j)| + |\log \varepsilon_0| \\ &\le \frac{\#\Gamma_{\varepsilon_0}}{D} \max_{|\sigma(j)|<\varepsilon_0} \left\{\log\left|\sigma(j)^{-1}\right|\right\} + 3|\log\varepsilon| - \log 44000 \\ &\le \frac{\#\Gamma_{\varepsilon_0}}{D} \max_{|\sigma(j)|<\varepsilon_0} \left\{\log\left|\sigma(j)^{-1}\right|\right\} + 3|\log\varepsilon|. \end{aligned} \tag{3.31}$$

Since $\varepsilon \le 10^{-5} < 3/200^2$ we can apply Proposition 3.5 to each pair (E_0^σ, ζ) and (E_0^σ, ζ^2) where σ runs over all embeddings $\sigma\colon K \hookrightarrow \mathbb{C}$ as follows. For each $\sigma \in \Gamma_{\varepsilon_0}$ the number τ^σ is close to either ζ or ζ^2 as we have seen in (3.30) and gives an N–isogeny from E_0^σ to E^σ. Thus we can bound $\#\Gamma_{\varepsilon_0}$ by

$$\#\Gamma_{\varepsilon_0} \le 6 \cdot 10^7 h[K:\mathbb{Q}]^2 \left(\sqrt{N}\sigma_0(N) + \sqrt{\varepsilon}\psi(N)\right).$$

after applying Lemma 3.19. We also have $\varepsilon_0 \le 10^{-3}$ and $N \ge \mathcal{N}(E_0^\sigma, \zeta)$ by assumption

and the previous lemma, so we can apply Lemma 3.18 to get

$$\max_{|\sigma(j)|<\varepsilon_0} \left\{\log\left|\sigma(j)^{-1}\right|\right\} \leq c_1(\log N)^6 + c_2.$$

Using the last two inequalities for (3.31) we obtain

$$h(j) \leq \frac{6\cdot 10^7 h[K:\mathbb{Q}]\left(\sqrt{N}\sigma_0(N)+\sqrt{\varepsilon}\psi(N)\right)}{B}\left(c_1(\log N)^6+c_2\right)+3|\log\varepsilon|$$

since we have $D = [K^\Phi : \mathbb{Q}] = B\cdot[K:\mathbb{Q}]$. Nicolas shows on page 229 in [Nic87] that $\sigma_0(N) \leq N^{2/5}$ for $N \geq 10^7$. Moreover, we have

$$\begin{aligned}\sqrt{N}\sigma_0(N)\cdot B^{-1} \leq N^{9/10}\cdot B^{-1} &\leq N^{9/10}\cdot\psi(N)^{-1}\cdot[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}):\rho_N(G_K)]\\ &\leq N^{-1/10}\cdot[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}):\rho_N(G_K)].\end{aligned}$$

by Lemma 3.11. Using this in the inequality for the height above, and Lemma 3.11 again, we get

$$\begin{aligned}h(j) \leq & 6\cdot 10^7\cdot h[K:\mathbb{Q}][\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}):\rho_N(G_K)]\left(N^{-1/10}+\sqrt{\varepsilon}\right)\left(c_1(\log N)^6+c_2\right)\\ &+3|\log\varepsilon|,\end{aligned}$$

as desired. □

Theorem 3.21 *Let j_0 be the j–invariant of an elliptic curve without complex multiplication. Then there are at most finitely many j–invariants j of elliptic curves that are isogenous to an elliptic curve corresponding to j_0 and such that j is an algebraic unit.*

Proof. Let E_0 and E be elliptic curves with j–invariants j_0 and j, respectively. Suppose that there is an isogeny of degree N between them. We may assume that N is minimal. By Lemma 6.2 in [MW90] the isogeny is cyclic. If N is large enough, then Corollary 3.14 gives a lower bound for the height of j

$$\begin{aligned}h(j) \geq\ & h(j_0) - 6\log(1+h(j_0)) + 6\log N\\ & - 84\,[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}):\rho_N(G_K)]\log\log N - 16.212\end{aligned} \tag{3.32}$$

Moreover, if j is an algebraic unit and N is large as in Proposition 3.20, that proposition yields the upper bound

$$\begin{aligned}h(j) \leq & 6\cdot 10^7\cdot h[K:\mathbb{Q}][\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}):\rho_N(G_K)]\left(N^{-1/10}+\sqrt{\varepsilon}\right)\left(c_1(\log N)^6+c_2\right)\\ &+3|\log\varepsilon|,\end{aligned}$$

For large enough N, the preconceived restrictions on ε are met if we take $\varepsilon = 1/(\log N)^{12}$

since $N \geq 10^7$ and thus $\varepsilon < 10^{-5}$. Therefore, we have

$$\begin{aligned} h(j) \leq 6 \cdot 10^7 \cdot h[K:\mathbb{Q}][\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(G_K)] \left(N^{-1/10} + \frac{1}{(\log N)^6}\right) \\ \cdot \left(c_1 (\log N)^6 + c_2\right) \\ + 36 \log\log N. \end{aligned} \tag{3.33}$$

Recall that Serre proved that $[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(G_K)]$ is uniformly bounded in N. Also Lombardo gives an explicit bound in [Lom15]. As we have seen in Corollary 3.14 the lower bound for $h(j)$ grows as $\log N$ and the upper bound as $\log\log N$. This clearly gives a contradiction for large enough N, which leaves us with only finitely many N, and hence finitely many isogenies. □

The next proposition bounds the number of j satisfying the conditions in the theorem. Note that the index $[\mathrm{GL}_2(\hat{\mathbb{Z}}) : \rho_\infty(G_K)]$ can be bounded explicitly by the result of Lombardo. See [Lom15] or page 54.

Proposition 3.22 *Let $E_0 \colon y^2 = 4x^3 - g_2 x - g_3$ be an elliptic curve without complex multiplication defined over a number field K of degree D. Let j_0 be its j–invariant with $j(\tau_0) = j_0$ and $\tau_0 \in \mathcal{F}$. We choose ω_1 and ω_2 with $\omega_2/\omega_1 = \tau_0$ and $E_0(\mathbb{C}) \simeq \mathbb{C}/(\omega_1\mathbb{Z} + \omega_2\mathbb{Z})$ and similarly for E_0^σ, $\sigma \colon K \hookrightarrow \mathbb{C}$. Define $h = \max\{1, h(1, g_2, g_3), h(j_0)\}$. If j is the j–invariant of an elliptic curve isogenous to E_0 such that j is a unit, then the degree of the minimal isogeny between j_0 and j is bounded by*

$$\max\left\{10^{180}(Cc_1)^{20}, (Cc_2)^{10}, e^{Cc_1+Cc_2+c_3}, e^{120^2[\mathrm{GL}_2(\hat{\mathbb{Z}}):\rho_\infty(G_K)]^2}, e^{18\pi h}, D\right\},$$

where the constants are given by

$$\begin{aligned} C &= 6 \cdot 10^7 \cdot h \cdot D[\mathrm{GL}_2(\hat{\mathbb{Z}}) : \rho_\infty(G_K)], \\ c_1 &= 2 \cdot 10^{51} D^6 \cdot \max\{h, 6\pi|\tau_0|\}^2 \geq 1, \\ c_2 &= 14 + 3\log\left(\max_\sigma\{1, |\omega_{0,1}^\sigma|, |\omega_{0,2}^\sigma|\}\right) \text{ and} \\ c_3 &= 20 - h(j_0) + 6\log(1 + h(j_0)). \end{aligned}$$

Note that $c_3 < 26$ since $-x + 6\log(1+x)$ has a maximum at 5 and $-5 + 6\log(6) < 6$.

Proof. We proceed as in the proof of the theorem. The inequalities (3.32) and (3.33)

in the proof of the theorem give

$$\begin{aligned}6\log N \leq & C\left(N^{-1/10}+\frac{1}{(\log N)^6}\right)\left(c_1(\log N)^6+c_2\right)\\ & +36\log\log N+84\left[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}):\rho_N(G_K)\right]\log\log N\\ & -h(j_0)+6\log(1+h(j_0))+16.212\end{aligned}$$

and thus

$$\begin{aligned}6 \leq & \frac{1}{\log N}\left(Cc_1N^{-1/10}(\log N)^6+Cc_2N^{-1/10}+Cc_1+\frac{Cc_2}{(\log N)^6}\right.\\ & \left.+(84[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}):\rho_N(G_K)]+36)\log\log N+c_3\right)\\ \leq & \left(Cc_1N^{-1/10}(\log N)^5+Cc_2\frac{N^{-1/10}}{\log N}+\frac{Cc_1+Cc_2+c_3}{\log N}\right.\\ & \left.+120[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}):\rho_N(G_K)]\frac{\log\log N}{\log N}\right)\end{aligned} \tag{3.34}$$

We are going to bound each term by 1 individually. This will give a contradiction to the lower bound 6. We will work our way from the back to the front. We have $\log\log x < (\log x)^{1/2}$ for all $x \geq 10$. Thus

$$120[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}):\rho_N(G_K)] \leq \log N/\log\log N$$

follows from

$$120[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}):\rho_N(G_K)] \leq (\log N)^{1/2} \leq \log N/\log\log N,$$

which is true for all $N \geq e^{(120[\mathrm{GL}_2(\hat{\mathbb{Z}}):\rho_\infty(G_K)])^2}$.
The next term is $(Cc_1+Cc_2+c_3)/\log N$. This is bounded by 1 for all $N \geq e^{Cc_1+Cc_2+c_3}$.

The second term is less than 1 if $Cc_2 \leq N^{1/10}$ is satisfied and $N \geq 3$. This is true for all $N \geq \max\{(Cc_2)^{10}, 3\}$.

For the first term we need

$$Cc_1 \leq \frac{N^{1/10}}{(\log N)^5}.$$

We have $\log x \leq 40x^{1/100}$ for all $x \geq 10^{45}$. Thus the bound holds if

$$Cc_1 \leq 10^{-9}N^{1/20} = 10^{-9}40^5\frac{N^{1/10}}{(40N^{1/100})^5} \leq \frac{N^{1/10}}{(\log N)^5}.$$

This is true for $N \geq 10^{180}(Cc_1)^{20}$.

All those assumptions on N together with the constraint

$$N \geq \max\left\{4 \cdot 10^{11}, [K : \mathbb{Q}], e^{3h}\right\}$$

we made in the previous proposition gives the desired bound. □

This finishes the case $j - \alpha$ when α is zero. In the next section we are going to discuss the case when α is different from 0.

3.3 Translates

Fix $\alpha \in \bar{\mathbb{Q}}$ the j–invariant of an elliptic curve with complex multiplication, and let j_0 be the j–invariant of an elliptic curve without complex multiplication. We further assume $\alpha \neq 0$ since this is the case discussed in the last section. We now want to bound the j–invariants j such that the corresponding elliptic curve is isogenous to the elliptic curve E_0, and such that $j - \alpha$ is an algebraic unit. Note that the previous case is a special case of this where $\alpha = 0$. Let ξ be imaginary quadratic with $j(\xi) = \alpha$. We proceed as before, i.e. we want to give lower and upper bounds of $h(j - \alpha)$ that contradict each other.
On the one hand we have

$$h(j - \alpha) \geq h(j) - h(\alpha) - \log 2 \tag{3.35}$$

by Lemma 1.4. So if there is a cyclic N–isogeny between the curves corresponding to j and j_0, then Corollary 3.14 yields

$$\begin{aligned} h(j - \alpha) \geq h(j_0) &- 6\log(1 + h(j_0)) + 6\log N \\ &- 84\,[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(G_K)] \log\log N - 20 - h(\alpha). \end{aligned}$$

Now we want to bound the height from above. We need a similar statement to the one in Lemma 3.18. First, recall the definition of $c(\xi)$ in Lemma 2.11

$$c(\xi) = \begin{cases} |j'(\xi)|\delta/2 & \text{if } \xi \in \partial\mathcal{F}_+ \setminus \{i\} \\ |j''(i)|\delta^2/4 & \text{if } \xi = i \\ \min\{|\operatorname{Im}(j(\xi))|, |j'(\xi)|\delta/2\} & \text{otherwise} \end{cases}$$

where δ is defined as the minimum of $\frac{A}{12A+108B}$ and half the distance of ξ to any geodesic of $\partial\mathcal{F}_+$ not containing ξ, B is defined as $4 \cdot 10^5 \max\{1, |j(\xi)|\}$ and $A = |j''(i)|$ if $\xi = i$ and $A = |j'(\xi)|$ otherwise. We also assumed $\xi \neq \zeta, \zeta^2$.

Recall the definition

$$\mathcal{N}(E_0,\xi) := \left(\max\{e^{6\pi|\tau_0|/[K:\mathbb{Q}]}, e^{e\cdot h}, [K:\mathbb{Q}], (4\cdot 10^{11}H(\xi))^{20}\}\right)^{1/20}.$$

Lemma 3.23 *Let $j(\tau)$ be N–isogenous to E_0 and $N \geq \mathcal{N}(E_0^\sigma, \xi^\sigma)$.*

$$\log|\sigma(j-\alpha)| \geq -c_1(\log N)^6 - c_2$$

for any embedding $\sigma\colon K \hookrightarrow \mathbb{C}$. Here the constants are effective and depend on the fixed elliptic curve E_0 and c_2 additionally depends on ξ. We also have $c_1 \geq 1$ and we can have $c_2 \geq 0$.

Proof. The setup is the same as in Lemma 3.18. Let E be an elliptic curve with j–invariant $j(\tau)$ and E^σ be the elliptic curve E conjugated by σ. Then there is an N–isogeny between $E_0 = E_0^\sigma$ and E^σ since E_0 and E are N–isogenous. Let $E_0^\sigma(\mathbb{C}) \simeq \mathbb{C}/(\omega_{0,1}^\sigma\mathbb{Z} + \omega_{0,2}^\sigma\mathbb{Z})$ with $\tau_0^\sigma = \omega_{0,2}^\sigma/\omega_{0,1}^\sigma$ in the fundamental domain. Similarly, let τ^σ correspond to $E^\sigma(\mathbb{C})$. We can choose ω_1^σ, ω_2^σ and $\rho \in \mathrm{SL}_2(\mathbb{Z})$ such that $\tau^\sigma = \rho\left(\begin{smallmatrix} m & l \\ 0 & n\end{smallmatrix}\right)\tau_0^\sigma$ and such that τ^σ is in the fundamental domain $\mathcal{F}$. Write $\left(\begin{smallmatrix}\alpha & \beta \\ \gamma & \delta\end{smallmatrix}\right) = \rho\left(\begin{smallmatrix} m & l \\ 0 & n\end{smallmatrix}\right)$.

Assume $|\sigma(j-\alpha)| < c(\xi^\sigma)$ for a moment. Put $A^\sigma = |j''(\xi^\sigma)|$ if $\xi^\sigma = i$ and $A^\sigma = |j'(\xi^\sigma)|$ otherwise. By Lemma 2.12 we obtain $|\tau^\sigma - M\xi^\sigma| < \delta^\sigma$ for some $M \in \mathcal{T}$ with $\mathcal{T} = \{\left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right), \left(\begin{smallmatrix}1 & \pm 1\\ 0 & 1\end{smallmatrix}\right), \left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)\}$. The number δ^σ is the δ stated above but associated to ξ^σ. Since δ^σ satisfies by definition $\delta^\sigma \leq \frac{A^\sigma}{12A^\sigma + 108B^\sigma}$, where $B^\sigma = 4\cdot 10^5 \max\{1, |j(\xi^\sigma)|\}$, we obtain by Lemma 2.10 the inequality

$$|j(\tau^\sigma) - j(\xi^\sigma)| \geq \frac{A^\sigma}{4}|\tau^\sigma - M\xi^\sigma|^2 \tag{3.36}$$

for some $M \in \mathcal{T}$.

Equation (3.27) says $|\gamma\omega_{0,2}^\sigma + \delta\omega_{0,1}^\sigma| \leq 10^6 \max\{|\omega_{0,1}^\sigma|, |\omega_{0,2}^\sigma|\}N^{10}$. Note that we also have $\tau^\sigma \neq M\xi^\sigma$ since ξ comes from a curve with complex multiplication. We can substitute $M\xi^\sigma$ for ζ in (3.28) to get the equality

$$\log|\tau^\sigma - M\xi^\sigma| = -\log|\gamma\omega_{0,2}^\sigma + \delta\omega_{0,1}^\sigma| + \log|(\alpha - M\xi^\sigma\gamma)\omega_{0,2}^\sigma + (\beta - M\xi^\sigma\delta)\omega_{0,1}^\sigma|.$$

Since ξ^σ is algebraic of degree two so is $M\xi^\sigma$. We have

$$\log|\tau^\sigma - M\xi^\sigma| \geq -\log\left(10^6\max\{|\omega_{0,1}^\sigma|, |\omega_{0,2}^\sigma|\}N^{10}\right) - c_1'\cdot(\log N)^6.$$

as in the proof of Lemma 3.18. Here c_1' is the constant from Lemma 3.16.

As before we obtain by (3.36)

$$\begin{aligned}\log|\sigma(j-\alpha)| &= \log|j(\tau^\sigma)-j(\xi^\sigma)| \geq \log\frac{A^\sigma}{4}+\log|\tau^\sigma - M\xi^\sigma|^2\\ &\geq \log\left(\frac{A^\sigma}{4}\right)-2\log\left(10^6\max\{|\omega_{0,1}^\sigma|,|\omega_{0,2}^\sigma|\}\right)\\ &\quad -10\log N-2c_1'\cdot(\log N)^6\\ &\geq \log\left(\frac{A^\sigma}{4}\right)-2\log\left(10^6\max\{|\omega_{0,1}^\sigma|,|\omega_{0,2}^\sigma|\}\right)\\ &\quad -3c_1'\cdot(\log N)^6\end{aligned}$$

where we have used the fact that $N \geq \mathcal{N}(E_0,\xi^\sigma) \geq 4\cdot 10^{11}$. If we put

$$c_2 = \log\max\left\{1, c(\xi^\sigma), \frac{4}{A^\sigma}10^{12}|\omega_{0,1}^\sigma|^2, \frac{4}{A^\sigma}10^{12}|\omega_{0,1}^\sigma|^2\right\}$$

the claim holds independently of whether $|\sigma(j-\alpha)| < c(\xi^\sigma)$ or not. □

We want to apply this lemma. Recall that $\alpha = j(\xi)$ is a singular modulus. Let Δ be the discriminant of the associated endomorphism ring. For any $\sigma\colon K \hookrightarrow \mathbb{C}$ the singular moduli $j(\xi^\sigma)$ have the same associated discriminant. Recall the definition of $\mathcal{P}(\xi)$ on page 29

$$\mathcal{P}(\xi) = \log\max_\sigma\left\{1, c(\xi^\sigma)^{-1}\right\}.$$

Proposition 3.24 *Let j_0 and j be j–invariants of elliptic curves, where j_0 is associated to the elliptic curve E_0/K defined by $E_0\colon y^2 = 4x^3 - g_2x - g_3$. Put $h = \max\{1, h(1,g_2,g_3), h(j_0)\}$. Assume we have a cyclic isogeny of degree N between E_0 and an elliptic curve corresponding to j. Further assume that j is an algebraic unit. If*

$$N \geq \max\{e^{3h}, [K:\mathbb{Q}], 4\cdot 10^{11}\sqrt{|\Delta|}\}$$

then the height of j can be estimated by

$$\begin{aligned}h(j-\alpha) \leq &\frac{10^8 h[K:\mathbb{Q}]^2|\Delta|^5[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}):\rho_N(G_K)]}{[\mathbb{Q}(\alpha):\mathbb{Q}]}(N^{-1/10}+\sqrt{\varepsilon})\left(c_1(\log N)^6+c_2\right)\\ &+\mathcal{P}(\xi)+2|\log\varepsilon|.\end{aligned}$$

where

$$0 < \varepsilon < 10^{-4}\min_{\sigma\colon K\hookrightarrow\mathbb{C}}\{|\xi^\sigma|^{-4}\}$$

is arbitrary.

PROOF. Recall from the proof of Proposition 3.20 the field K^Φ for which we have $j \in K^\Phi$. We also showed that $[K^\Phi : \mathbb{Q}] = B[K : \mathbb{Q}]$ and $B = [K^\Phi : K]$. Let

$$\varepsilon_0 = \varepsilon^2 \cdot \min_{\sigma\colon K \hookrightarrow \mathbb{C}} \{1, c(\xi^\sigma)\}.$$

Let $|\sigma(j-\alpha)| < \varepsilon_0 < c(\xi^\sigma)$. We have $N \geq \mathcal{N}(E_0^\sigma, \xi^\sigma)$ since we have $\sqrt{|\Delta|} \geq H(\xi^\sigma)$ by Lemma 2.13 and the statement of Lemma 3.19. So the previous lemma says

$$\log|\sigma(j-\alpha)| \geq -c_1(\log N)^6 - c_2,$$

where we now can take c_2 to be the maximum over all constants that we get from the lemma for each ξ^σ. We have $\sigma \in \Gamma(\xi^\sigma, \varepsilon)$ since we assumed $|\sigma(j-\alpha)| < \varepsilon_0 < \varepsilon$. The same argument as in the proof of Proposition 3.20 shows

$$\begin{aligned} h(j-\alpha) &\leq -\frac{1}{D} \sum_{|\sigma(j-\alpha)|<\varepsilon_0} \log|\sigma(j-\alpha)| + |\log \varepsilon_0| \\ &\leq \frac{\sum_{\sigma\colon K\hookrightarrow\mathbb{C}} \#\Gamma(\xi^\sigma, \varepsilon_0)}{D} \max_{|\sigma(j-\alpha)|<\varepsilon_0} \left\{\log\left|\sigma(j-\alpha)^{-1}\right|\right\} + |\log\varepsilon_0| \\ &\leq \frac{\sum_{\sigma\colon K\hookrightarrow\mathbb{C}} \#\Gamma(\xi^\sigma, \varepsilon_0)}{D} \left(c_1(\log N)^6 + c_2\right) + |\log\varepsilon_0|, \end{aligned} \tag{3.37}$$

where D is the degree of $K^\Phi(\alpha)$ over $\mathbb{Q}$. Now if $\rho \in \Gamma(\xi^\sigma, \varepsilon_0)$, then $|j(\tau^\rho) - j(\xi^\sigma)| < \varepsilon_0 \leq c(\xi^\sigma)$. With δ^σ as in Lemma 2.11 we get from Lemma 2.12

$$|\tau^\rho - M\xi^\sigma| < \delta^\sigma$$

for some $M \in \mathcal{T}$. As before we put $A^\sigma = |j''(i)|$ if $\xi^\sigma = i$ and $A^\sigma = |j'(\xi)|$ otherwise. Recall that $\delta^\sigma \leq 1$ so that $c(\xi^\sigma) \leq A^\sigma/2$ or $c(\xi^\sigma) \leq A^\sigma/4$. Lemma 2.10 then implies

$$\frac{A^\sigma}{2^k}|\tau^\rho - M\xi^\sigma|^2 \leq |j(\tau^\rho) - j(\xi^\sigma)| < \varepsilon_0 \leq c(\xi^\sigma)\varepsilon^2 \leq \frac{A^\sigma}{2^k}\varepsilon^2,$$

where $k \in \{1, 2\}$ depending on whether $M\xi^\sigma = i$ or not. Therefore we have $|\tau^\rho - M\xi^\sigma| < \varepsilon$. So every $\rho \in \Gamma(\xi^\sigma, \varepsilon_0)$ gives a point satisfying $|\tau^\rho - M\xi^\sigma| \leq \varepsilon$ and an N–isogeny between E_0^ρ and E^ρ. Note that M can only be different from the identity if ξ^σ lies on the boundary of $\mathcal{F}$. In any case since ξ (and all $M\xi^\sigma$) is imaginary quadratic, some conjugate lies on the imaginary axis and is the largest with respect to the absolute value. By the explicit description in Chapter 1 it is given by $i|\Delta|^{1/2}/2$. Moreover, ε satisfies the conditions of Proposition 3.5. We thus can apply Proposition 3.5 to bound

$$\#\Gamma(\xi^\sigma, \varepsilon_0) \leq 4 \cdot 10^7 [K : \mathbb{Q}]^2 |\Delta|^5 h \left(\sqrt{N}\sigma_0(N) + \sqrt{\varepsilon}\psi(N)\right).$$

Note that $\varepsilon \leq (100^{-1}|M\xi^\sigma|\operatorname{Im}(M\xi^\sigma))^2$ holds since if M is different from the identity, then ξ^σ lies on the boundary of the fundamental domain and we obtain $|M\xi^\sigma| = |\xi^\sigma|$ and $\operatorname{Im}(M\xi^\sigma) = \operatorname{Im}(\xi^\sigma)$. We can continue the height estimate in (3.37)

$$\begin{aligned} h(j-\alpha) \leq & [K:\mathbb{Q}]\frac{4[K:\mathbb{Q}]^2 10^7 |\Delta|^5 h\left(\sqrt{N}\sigma_0(N)+\sqrt{\varepsilon}\psi(N)\right)}{D}\left(c_1(\log N)^6+c_2\right) \\ & + |\log\varepsilon_0| \\ \leq & [K:\mathbb{Q}]\frac{4[K:\mathbb{Q}]^2 10^7 |\Delta|^5 h\left(\sqrt{N}\sigma_0(N)+\sqrt{\varepsilon}\psi(N)\right)}{[K^\Phi:\mathbb{Q}][\mathbb{Q}(\alpha):\mathbb{Q}]}\left(c_1(\log N)^6+c_2\right) \\ & + |\log\varepsilon_0| \\ \leq & \frac{4[K:\mathbb{Q}]^2 10^7 |\Delta|^5 h\left(\sqrt{N}\sigma_0(N)+\sqrt{\varepsilon}\psi(N)\right)}{B[\mathbb{Q}(\alpha):\mathbb{Q}]}\left(c_1(\log N)^6+c_2\right)+|\log\varepsilon_0|. \end{aligned}$$

We have bounded the term $\left(\sqrt{N}\sigma_0(N)+\sqrt{\varepsilon}\psi(N)\right)/B$ in Proposition 3.20, so that we obtain

$$\begin{aligned} h(j-\alpha) \leq & \frac{10^8[K:\mathbb{Q}]^2|\Delta|^5 h[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}):\rho_N(G_K)]}{[\mathbb{Q}(\alpha):\mathbb{Q}]}(N^{-1/10}+\sqrt{\varepsilon})\left(c_1(\log N)^6+c_2\right) \\ & + |\log\varepsilon_0|. \end{aligned}$$

Now

$$|\log\varepsilon_0| = 2|\log\varepsilon| + |\log\min_\sigma\{1, c(\xi^\sigma)\}| = 2|\log\varepsilon| + \mathcal{P}(\xi).$$

Replacing this into the height bound be obtain

$$\begin{aligned} h(j-\alpha) \leq & \frac{10^8[K:\mathbb{Q}]^2|\Delta|^5 h[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}):\rho_N(G_K)]}{[\mathbb{Q}(\alpha):\mathbb{Q}]}(N^{-1/10}+\sqrt{\varepsilon})\left(c_1(\log N)^6+c_2\right) \\ & + \mathcal{P}(\xi) + 2|\log\varepsilon|. \end{aligned}$$

□

Theorem 3.25 *Assume α is the j–invariant of an elliptic curve with CM. Let j_0 be the j–invariant of an elliptic curve without CM. Then there are at most finitely many j–invariants j of elliptic curves that are isogenous to an elliptic curve corresponding to j_0 and such that $j-\alpha$ is an algebraic unit.*

PROOF. In the same situation as before we get an additional $-\log 2 - h(\alpha)$ term from

(3.35) for the lower bound and obtain

$$h(j_0) - 6\log(1+h(j_0)) + 6\log N - 84\,[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(G_K)]\log\log N \\ - 20 - h(\alpha) \leq h(j-\alpha).$$

We want to pick $\varepsilon = 1/(\log N)^{12}$ again. Thus, N must be large enough so that

$$\varepsilon = 1/(\log N)^{12} \leq 10^{-4} \min_{\sigma}\{|\xi^{\sigma}|^{-4}\}$$

or equivalently

$$(\log N)^{12} \geq 10^4 \max_{\sigma}\{|\xi^{\sigma}|^4\}.$$

But as mentioned in the previous proof, ξ and $|\xi^{\sigma}|$ are imaginary quadratic and one of its conjugates is $i|\Delta|/2$ and has maximal modulus amongst them. Hence it suffices for N to satisfy $\log N \geq 3|\Delta|$, i.e.

$$N \geq e^{3|\Delta|}.$$

If N additionally satisfies the conditions of the previous proposition then

$$\begin{aligned} h(j-\alpha) \leq & \frac{10^8 h[K:\mathbb{Q}]^2|\Delta|^5[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(G_K)]}{[\mathbb{Q}(\alpha):\mathbb{Q}]} \\ & \cdot \left(N^{-1/10} + \frac{1}{(\log N)^6}\right)\left(c_1(\log N)^6 + c_2\right) \\ & + \mathcal{P}(\xi) + 24\log\log N. \end{aligned}$$

The growth of the bounds for $h(j-\alpha)$ is as before, and we get the same contradiction. □

In total we obtain the following result. We also recall that $c_3 < 26$.

Proposition 3.26 *Let $E_0 \colon y^2 = 4x^3 - g_2x - g_3$ be an elliptic curve without complex multiplication defined over a number field K of degree D. Let j_0 be its j–invariant with $j(\tau_0) = j_0$ and $\tau_0 \in \mathcal{F}$. Define $h = \max\{1, h(1, g_2, g_3), h(j_0)\}$. Let $\xi \in \mathcal{F}$ be imaginary quadratic and let Δ be the discriminant of the endomorphism ring. Put $\alpha = j(\xi)$. If j is the j–invariant of an elliptic curve isogenous to the elliptic curve E_0 and $j - \alpha$ is a unit, then the degree of the minimal isogeny is bounded by*

$$\max\Big\{10^{180}(\hat{C}c_1)^{20}, (\hat{C}c_2)^{10}, e^{\hat{C}c_1+\hat{C}c_2+c_3+\mathcal{P}(\xi)}, e^{120^2[\mathrm{GL}_2(\hat{\mathbb{Z}}):\rho_\infty(G_K)]^2}, \\ e^{3h}, [K:\mathbb{Q}], e^{3|\Delta|}, 4\cdot 10^{11}\sqrt{|\Delta|}\Big\},$$

where $\hat{C} = 10^8 h[K:\mathbb{Q}]^2|\Delta|^5[\mathrm{GL}_2(\hat{\mathbb{Z}}) : \rho_\infty(G_K)]/[\mathbb{Q}(\alpha):\mathbb{Q}]$ and $c_3 = 20 - h(j_0) + 6\log(1+h(j_0))$.

PROOF. The bounds from the previous proof give the same inequality as in (3.34) with C replaced by the new constant $\hat{C}$ and the third term becomes

$$\frac{\hat{C}c_1 + \hat{C}c_2 + c_3 + \mathcal{P}(\xi)}{\log N}.$$

Also we have the additional prerequisites $N \geq e^{3|\Delta|}$ and $N \geq 4 \cdot 10^{11}\sqrt{|\Delta|}$ from the proof of the last theorem. □

Bibliography

[Adl14] Semjon Adlaj: *Multiplication and division on elliptic curves, torsion points and roots of modular equations*. 2014.

[Aut03] Pascal Autissier: *Hauteur des correspondances de Hecke*. In: *Bulletin de la Société mathématique de France* 131.3 (2003), pp. 421–433.

[BG07] Enrico Bombieri and Walter Gubler: *Heights in Diophantine Geometry*. Cambridge University Press, 2007.

[BHK18] Yuri Bilu, Philipp Habegger, and Lars Kühne: *No singular modulus is a unit*. In: *International Mathematics Research Notices* (2018).

[BLP16] Yuri Bilu, Florian Luca, and Amalia Pizarro–Madariaga: *Rational products of singular moduli*. In: *Journal of Number Theory* 158 (2016), pp. 397–410.

[BMZ13] Yuri Bilu, David Masser, and Umberto Zannier: "An effective "Theorem of André" for CM–points on a plane curve". In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 154. 01. Cambridge Univ Press. 2013, pp. 145–152.

[Cas57] John William Scott Cassels: *An introduction to Diophantine approximation*. Vol. 1957. Cambridge University Press Cambridge, 1957.

[Cip02] Michele Cipolla: *La determinazione asintotica dell' nimo numero primo*. In: (1902), pp. 132–166.

[Col98] Pierre Colmez: *Sur la hauteur de Faltings des variétés abéliennes à multiplication complexe*. In: *Compositio Mathematica* 111.3 (1998), pp. 359–369.

[Cox11] David A Cox: *Primes of the form $x^2 + ny^2$: Fermat, class field theory, and complex multiplication*. Vol. 34. John Wiley & Sons, 2011.

[CS86] Gary Cornell and Joseph H Silverman: *Arithmetic geometry*. Springer verlag, 1986.

[Dav51] Harold Davenport: *On a principle of Lipschitz*. In: *J. London Math. Soc* 26 (1951), pp. 179–183.

[Dav95] Sinnou David: *Minorations de formes linéaires de logarithmes elliptiques*. In: *Mémoires de la Societé Mathématique de France* 62 (1995), pp. 1–143.

[DH09] Sinnou David and Noriko Hirata–Kohno: *Linear forms in elliptic logarithms*. In: *Journal für die reine und angewandte Mathematik (Crelles Journal)* 628 (2009), pp. 37–89.

[DS05] Fred Diamond and Jerry Shurman: *A First Course in Modular Forms*. Vol. 228. Springer, 2005.

[Fal83] Gerd Faltings: *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*. In: *Inventiones mathematicae* 73.3 (1983), pp. 349–366.

[FP87] Alain Faisant and Georges Philibert: *Quelques résultats de transcendance liés à l'invariant modulaire j*. In: *Journal of number theory* 25.2 (1987), pp. 184–200.

[GKZ87] Benedict Gross, Winfried Kohnen, and Don Zagier: *Heegner points and derivatives of L–series. II*. In: *Mathematische Annalen* 278.1-4 (1987), pp. 497–562.

[GR11] Éric Gaudron and Gaël Rémond: *Théorème des périodes et degrés minimaux d'isogénies*. In: (2011).

[GZ84] Benedict H. Gross and Don B. Zagier: *On singular moduli*. Universität Bonn. SFB 40. Theoretische Mathematik/Max–Planck–Institut für Mathematik, 1984.

[GZ86] Benedict H. Gross and Don B. Zagier: *Heegner points and derivatives of L–series*. In: *Inventiones mathematicae* 84.2 (1986), pp. 225–320.

[Hab15] Philipp Habegger: *Singular moduli that are algebraic units*. In: *Algebra & Number Theory* 9.7 (2015), pp. 1515–1524.

[HP12] Philipp Habegger and Jonathan Pila: *Some unlikely intersections beyond André–Oort*. In: *Compositio Mathematica* 148.01 (2012), pp. 1–27.

[HS13] Marc Hindry and Joseph H. Silverman: *Diophantine geometry: an introduction*. Vol. 201. Springer, 2013.

[Hua12] L.–K. Hua: *Introduction to number theory*. Springer Verlag, 2012.

[Lan02] Serge Lang: *Algebra*. Revised third edition. Springer, 2002.

[Lan87] Serge Lang: *Elliptic Functions*. Second. Springer, 1987.

[Leh42] D. H. Lehmer: *Properties of the coefficients of the modular invariant $J(\tau)$*. In: *American Journal of Mathematics* 64.1 (1942), pp. 488–502.

[Li18] Yingkun Li: *Singular Units and Isogenies Between CM Elliptic Curves*. In: *arXiv preprint arXiv:1810.13214* (2018).

[Löb17] Steffen Löbrich: *A gap in the spectrum of the Faltings height*. In: *Journal de Théorie des Nombres de Bordeaux* 29.1 (2017), pp. 289–305.

[Lom15] Davide Lombardo: *Bounds for Serre's open image theorem for elliptic curves over number fields*. In: *Algebra & Number Theory* 9.10 (2015), pp. 2347–2395.

[Mas06] David W. Masser: *Elliptic Functions and Transcendence*. Vol. 437. Springer, 2006.

[Mer74] Franz Mertens: *Ein Beitrag zur analytischen Zahlentheorie.* In: *Journal für die reine und angewandte Mathematik* 78 (1874), pp. 46–62.

[Mil06] James S. Milne: *Elliptic Curves*. BookSurge Publishers, 2006.

[MW90] David W. Masser and G. Wüstholz: *Estimating isogenies on elliptic curves.* In: *Inventiones mathematicae* 100.1 (1990), pp. 1–24.

[Neu06] Jürgen Neukirch: *Algebraische Zahlentheorie*. Springer, 2006.

[Neu99] Jürgen Neukirch: *Algebraic number theory*. Vol. 322. Springer Berlin, 1999.

[Nic87] Jean–Louis Nicolas: "On highly composite numbers". In: *Ramanujan Revisited, Proceedings of the Centenary Conference*. University of Illinois at Urbana–Champaign, 1987, pp. 215–244.

[NT91] Yukiyoshi Nakkajima and Yuichiro Taguchi: *A generalization of the Chowla–Selberg formula.* In: *J. reine angew. Math* 419 (1991), pp. 119–124.

[Ray85] Michel Raynaud: *Hauteurs et isogénies.* In: *Astérisque* 127 (1985), pp. 199–234.

[Rib01] Paulo Ribenboim: *Classical theory of algebraic numbers*. Springer, 2001.

[Rib12] Paulo Ribenboim: *The theory of classical valuations*. Springer Science & Business Media, 2012.

[Ric13] Rodolphe Richard: *Répartition galoisienne d'une classe d'isogénie de courbes elliptiques.* In: *International Journal of Number Theory* 9.02 (2013), pp. 517–543.

[Rob83] Guy Robin: *Estimation de la fonction de Tchebychef θ sur le k–ième nombre premier et grandes valeurs de la fonction ω (n) nombre de diviseurs premiers de n.* In: *Acta Arithmetica* 42.4 (1983), pp. 367–389.

[Ros39] Barkley Rosser: *The n–th Prime is Greater than $n \log n$.* In: *Proceedings of the London Mathematical Society* 2.1 (1939), pp. 21–44.

[RS62] Barkley Rosser and Lowell Schoenfeld: *Approximate formulas for some functions of prime numbers.* In: *Illinois Journal of Mathematics* 6.1 (1962), pp. 64–94.

[Ser72] Jean–Pierre Serre: *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques.* In: *Inventiones mathematicae* 15.4 (1972), pp. 259–331.

[Ser89] Jean–Pierre Serre: *Lectures on the Mordell–Weil theorem*. Vol. 15. Vieweg, 1989.

[Sil86a] Joseph H. Silverman: "Heights and elliptic curves". In: *Arithmetic geometry*. Springer, 1986, pp. 253–265.

[Sil86b] Joseph H. Silverman: *The Arithmetic of Elliptic Curves*. Vol. 106. Springer, 1986.

[Sil99] Joseph H. Silverman: *Advanced topics in the arithmetic of elliptic curves (corrected second printing)*. In: *Graduate Texts in Mathematics* 151 (1999).

[ST92] Joseph H. Silverman and John T. Tate: *Rational points on elliptic curves*. Vol. 9. Springer, 1992.

[Wal13] Michel Waldschmidt: *Diophantine approximation on linear algebraic groups: transcendence properties of the exponential function in several variables*. Vol. 326. Springer Science & Business Media, 2013.

[Wal79] Michel Waldschmidt: *Nombres transcendants et groupes algébriques*. Vol. 69. American Mathematical Society, 1979.

[Wüs14] Gisbert Wüstholz: *A note on the conjectures of André–Oort and Pink*. In: *Bull. Ins. Math., Acad. Sinica (New Series)* 9 (2014), pp. 735–779.

List of Symbols

$\mathrm{Aut}(E[N])$	Group of automorphisms of the group $E[N]$. 52
$\mathcal{C}$	The class number. 2
$c(\xi)$	A constant depending on ξ. 22, 68
$\mathcal{C}(\Delta;\xi;\varepsilon)$	Set of CM–points of given discriminant and close to ξ. 15
d_ν	The local degree of L/K at a place ν. 3
Δ	The discriminant of a quadratic form or a endomorphism ring. 1, 15
$\Delta(\tau)$	The modular discriminant. 8
$\partial\mathcal{F}_+$	Boundary components of $\mathcal{F}_+$. 22
$D(z)$	A measure of distance from z to the fundamental domain. 57
E	An elliptic curve. vi, 7
$E[N]$	The N–torsion points on the elliptic curve E. 9
$E(\Delta)$	Function $F(\Delta)$ multiplied by $(\log\lvert\Delta\rvert)^4$. 30
$\mathrm{End}(E)$	Endomorphism ring of the elliptic curve E. 10
$E_\pm$	Two ellipses depending on a sign. 45
E^σ	The Galois conjugate of E when conjugating all the coefficients. 61
$\mathcal{F}$	The usual fundamental domain of the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$. 2
f	Conductor of imaginary quadratic field extension. 15
$F(\Delta)$	A function depending on Δ. 15
$\bar{\mathcal{F}}$	The closure of $\mathcal{F}$. 37
$\tilde{f}$	Modified conductor. 15
$\mathcal{F}_-$	The elements of the usual fundamental domain with non–positive real part. 22, 61

$\mathcal{F}_+$	The elements of the usual fundamental domain with non–negative real part. 22, 61
$[G : H]$	The index of H in G. xii
Γ	The ususal gamma function extending the factorial function. 21
gcd	The greatest common divisor. ix
G_K	The absolute Galois group $\mathrm{Gal}(\bar{K}/K)$. 52, 66
Γ_ε	The set $\Gamma(\zeta, \varepsilon)$. 37
$\Gamma(\xi, \varepsilon)$	Set of field embeddings. 37
$\mathbb{H}$	The Poincaré upper half–plane. xi
$h(\alpha)$	The (logarithmic) Weil height of α. vi, 5
$h(E)$	The Faltings height of E. 13
H	The Weil height. 4
h	Usually the max of two different heights associated to an elliptic curve. 58
$\mathcal{I}(\tau_0, \xi)$	A constant depending on the input. 50
j	The j–invariant of an elliptic curve. vi
$j(\tau)$	Klein's modular j–function. 8
$[L : K]$	The degree of the field extension. xi
$\Lambda(\tau_0; N; \varepsilon)$	Subset of the points close to ξ and N–isogenous to τ_0. 49
L_-	Circumference of E_-. 47
L_+	Circumference of E_+. 46
M_K	Set of places of K. 3
M_K^∞	Set of infinite places of K. 3
$\mathcal{M}(\xi; x; y; \varepsilon)$	Subset of $\mathrm{SL}_2(\mathbb{Z})$ that translates τ close to ξ. 44
$[N]$	The multiplication–by–N isogeny. 9
$\mathcal{N}(E_0, \xi)$	A lower bound on the degree of an isogeny. 58, 68
$N(E_\pm)$	Number of lattice points contained in $E_\pm$. 47
$N(\mathrm{Im}(\tau), \varepsilon)$	Number of lattice points close to an ellipse. 44
$N(\mathcal{R})$	Number of lattice points in the region $\mathcal{R}$. 44
$\mathcal{O}_{\bar{L}}$	The ring of algebraic integers of $\bar{L}$. viii

$\mathbb{P}^n_K$	The n–dimensional projective space over K. xi
$\mathcal{P}$	A function depending on a point in the standard fundamental domain. 28, 70
$\wp$	The Weierstraß $\wp$–function. 7
Q	A binary quadratic form. 1
q	$q = e^{2\pi i\tau}$. xi
$\mathcal{Q}(\Delta)$	Set of equivalence classes of primitive positive definite quadratic forms. 2
ρ_∞	The Galois representation associated to the elliptic curve E_0. 52
ρ_N	The N–torsion Galois representation. 52
σ	Usually an embedding of a number field K into the complex numbers. 4
$\sigma_k(n)$	Sum of the k–th power of the distinct divisors of n. 16
Σ_ε	The set $\Sigma(\xi, \varepsilon)$. 37
$\Sigma(\xi, \varepsilon)$	A set of points in the fundamental domain. 37, 81
τ_σ	A point in the fundamental domain related to τ by a field embedding. 28
τ_M	The points $M.\tau_0$. 48
τ_0^σ	A point in the fundamental domain related to τ_0 by a field embedding. 37, 61
$\tilde{\tau}$	The represantive in the fundamental domain of the $\mathrm{SL}_2(\mathbb{Z})$–orbit of τ. 38
$\mathcal{T}$	Set of matrices in $\mathrm{SL}_2(\mathbb{Z})$. 27
φ	Euler's totient function. 53
$\mathrm{vol}(E_+)$	The area of the ellipse E_+. 45
$V(\mathcal{R})$	Volume of the region $\mathcal{R}$. 44
$[x_0 : \ldots : x_n]$	Point in the projective space. xi
ψ	Dedekind ψ–function. 10
$Y(1)$	The modular curve of level 1. viii

ζ	The 6–th root of unity $e^{2\pi i/6}$. 11, 20