Topics in Number Theory
Small Heights
University of Copenhagen

Linda Frey

October 26, 2020

Contents

1 Heights

Consider the following numbers.

$$\begin{aligned} 0 \\ 1 \\ 1.1 \\ 0.9 \\ 0.99999 \end{aligned}$$

The latter four seem close, but if we look at them from an arithmetic point of view, the number 0.99999 seems more complicated than 0 or 1. We want to define a measure for that.

Definition 1 (Height I)**.** Let $x = \frac{a}{b} \in \mathbb{Q}$ be a rational number in lowest terms with positive b. We define the *(logarithmic Weil) height* as follows:

$$h(x) = \log \max(|a|, b).$$

Example 2. We have

$$\begin{aligned} h(0) = h\left(\frac{0}{1}\right) &= \log \max(|0|, 1) = 0, \\ h(1) = h\left(\frac{1}{1}\right) &= \log \max(|1|, 1) = 0, \\ h(1.1) = h\left(\frac{11}{10}\right) &= \log \max(|11|, 10) = \log 11, \\ h(0.9) = h\left(\frac{9}{10}\right) &= \log \max(|9|, 10) = \log 9, \\ h(0.99999) = h\left(\frac{99999}{100000}\right) &= \log \max(|99999|, 100000) = \log 100000. \end{aligned}$$

Lemma 3. *Let $x \in \mathbb{Q}$. Then the following properties hold.*

1. $h(x) \geq 0$.
2. $h(x) = 0 \Leftrightarrow x = 0$ *or* $x = \pm 1$.
3. *For $x \neq 0$ and $k \in \mathbb{N}$ we have $h(x^k) = kh(x)$.*
4. *For $x \neq 0$ we have $h(x^{-1}) = h(x)$.*
5. *For $x \neq 0$ and $k \in \mathbb{Z}$ we have $h(x^k) = |k|h(x)$.*

6. $h(xy) \leq h(x) + h(y)$ *and* $h(\pm x) = h(x)$.

7. $h(x + y) \leq h(x) + h(y) + \log 2$.

8. *For* $B \in \mathbb{R}$ *the set* $\{x \in \mathbb{Q} | h(x) \leq B\}$ *is finite.*

Proof. Exercise □

The definition of the height can also be rewritten in terms of p-adic absolute values. Let $0 \neq x = \frac{a}{b} \in \mathbb{Q}$ in lowest terms with positive b. Then we can factorize both a and b which will give us a connection between the p-adic absolute value and the height.

Let now $b = p_1^{e_1} \cdots p_g^{e_g}$ with primes $p_1 < \ldots < p_g$ and $e_1, \ldots, e_g \in \mathbb{N}$. Let now p be a prime and $|\cdot|_p$ the p-adic absolute value on $\mathbb{Q}$. With this notation we have

$$|x|_{p_i} = \left|\frac{a}{b}\right|_{p_i} = |b|_{p_i}^{-1} = p_i^{e_i}$$

since a and b are coprime. For $p \notin \{p_1, \ldots, p_g\}$ we have $|b|_p = 1$ and hence $|x|_p = |\frac{a}{b}|_p = |a|_p \leq 1$. Hence for all primes p we have

$$\max(1, |x|_p) = |b|_p^{-1}.$$

So we can write

$$b = p_1^{e_1} \cdots p_g^{e_g} = |b|_{p_1}^{-1} \cdots |b|_{p_g}^{-1} = \prod_p |b|_p^{-1} = \prod_p \max(1, |x|_p).$$

Since only finitely many factors in the product are not equal to 1, there is no problem with the infinite product.

Now we can write the height as follows.

$$\begin{aligned} h(x) &= \log \max(|a|, b) = \log(b \max(1, |x|)) = \log b + \log \max(1, |x|) \\ &= \log \prod_p \max(1, |x|_p) + \log \max(1, |x|). \end{aligned}$$

We will use this later to generalize the definition for number fields. First, we will give a definition through the minimal polynomial.

Definition 4 (Height II and integral minimal polynomial)**.** Let $x \in \overline{\mathbb{Q}}$ be an algebraic number. After multiplying its minimal polynomial over $\mathbb{Q}$ with an appropriate integer, we get a unique polynomial $P \in \mathbb{Z}[X] \setminus \{0\}$ such that

- $P(x) = 0$,
- P is irreducible as an element of $\mathbb{Q}[X]$,
- the coefficients are coprime.

Over $\mathbb{Q}$ we can factorize $P(X) = a_d(X - x_1) \cdots (X - x_d)$. We call the x_i the *conjugates* of x and P the *integral minimal polynomial* of x. We define the *(logarithmic Weil) height* of x as

$$h(x) = \frac{1}{d} \log \left(a_d \prod_{i=1}^{d} \max(1, |x_i|) \right).$$

Example 5.

- This definition coincides with the definition before. The integral minimal polynomial of a rational number $\frac{a}{b}$ is $bX - a = b(X - \frac{a}{b})$, hence $h(\frac{a}{b}) = \frac{1}{1}\log(b\max(1, |\frac{a}{b}|)) = \log(\max(|b|, |a|))$.
- Consider $\alpha = \sqrt{2}$ with integral minimal polynomial $P(X) = X^2 - 2 = (X - \sqrt{2}) \cdot (X + \sqrt{2})$. Then

$$\begin{aligned} h(\sqrt{2}) &= \frac{1}{2}\log\left(|1|(\max(1, |\sqrt{2}|) \cdot \max(1, |-\sqrt{2}|)\right) \\ &= \frac{1}{2}\log(\sqrt{2} \cdot \sqrt{2}) \\ &= \frac{\log 2}{2}. \end{aligned}$$

To give an alternative definition of the height, we have to generalize the p-adic absoulte values to number fields.

Definition 6. Let K be a number field and $x \in K$. For $x \neq 0$ and $P \subset \mathbb{Z}_K$ a non-zero prime ideal we can write the fractional ideal $x\mathbb{Z}_K$ as $P^e I$ where $e \in \mathbb{Z}$ and $I \subset \mathbb{Z}_K$ is an ideal not contained in P. We define $v_P(x) = e$. Let $p\mathbb{Z} = P \cap \mathbb{Z}$ and let $e(P)$ be the ramification index of the ideal P (that is, its multiplicity in the prime ideal factorization of $p\mathbb{Z}_K$), then we define

$$|x|_P = p^{-\frac{v_P(x)}{e(P)}}.$$

With this normalization we get $|p|_P = p^{-1}$. Define $|0|_P = 0$. We call P a *finite place* of K. We call the set of finite places $M^0(K)$.
Let $\sigma : K \to \mathbb{C}$ be a field embedding. Define $|x|_\sigma = |\sigma(x)|$, where $|\cdot|$ is the standard absolute value on $\mathbb{C}$. We call $|\cdot|_\sigma$ an *infinite place* of K. We call the set of infinite places $M^\infty(K)$. Furthermore we let $M(K) = M^\infty(K) \cup M^0(K)$.

Remark 7. The absolute values $|\cdot|_P$ and $|\cdot|_\sigma$ restricted fo $\mathbb{Q}$ give the p-adic and the standard absolute value, respectively.

Lemma 8. *For any absolut value $|\cdot|_v$ as in Definition 6 and x, y elements of a number field K, the following is true:*

- $|x|_v \geq 0$ *and* $|x|_v = 0 \Leftrightarrow x = 0$.
- $|xy|_v = |x|_v|y|_v$.
- $|x + y|_v \leq |x|_v + |y|_v$ *for infinite places v and* $|x + y|_v \leq \max(|x|_v, |y|_v)$ *for finite places v.*

Proof. Exercise. □

Remark 9. For a number field K and an embedding $\sigma : K \to \mathbb{C}$ we get another embedding $\overline{\sigma} : K \to \mathbb{C}$ by $\overline{\sigma}(x) = \overline{\sigma(x)}$ for all $x \in K$. For $\sigma(K) \not\subset \mathbb{R}$, we have $\overline{\sigma} \neq \sigma$, but they always induce the same absolute value on K: $|\cdot|_{\overline{\sigma}} = |\cdot|_\sigma$.

Be aware that in $M^\infty(K)$ we have the absolute values $|\cdot|_\sigma$, so the two embeddings σ and $\overline{\sigma}$ will result in one absolute value in $M^\infty(K)$.

Lemma 10. *For a number field K and $x \in K^* = K \setminus \{0\}$ we have only finitely many $v \in M(K)$ with $|x|_v \neq 1$.*

Proof. We know from algebraic number theory that the set of embeddings $K \to \mathbb{C}$ is finite, so we only have to prove the statement for $v \in M^0(K)$ instead of $M(K)$. There are only finitely many prime ideals in the prime ideal factorization of $x\mathbb{Z}_K$. Hence, only finitely many of these prime ideals give a place v with $|x|_v \neq 1$. □

Now we can define the height also in terms of absolute values.

Definition 11 (Height III)**.** Let K be a number field and $x \in K$. We define the *(logarithmic Weil) height*

$$h_K = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M(K)} d_v \log \max(1, |x|_v),$$

where for finite places v we let $d_v = f(P)e(P)$ (where $e(P)$ is the ramification index of P and $f(P)$ is the residue degree of P) and for infinite places v we let $d_v = 1$ for $\sigma(K) \subset \mathbb{R}$ and $d_v = 2$ otherwise.

We will not prove the following remarks.

Remark 12. The definition of h_K is independent of the choice of K.

Remark 13. For a number field K and $x \in K^*$, we have

$$\prod_{v \in M(K)} |x|_v^{d_v} = 1.$$

Remark 14. The two Definitions 4 and 11 are equivalent.

Remark 15. Since the set $M^\infty(K)$ of absolute values comes from embeddings $K \to \mathbb{C}$, we have the following equality which also explains the factor d_v:

$$\frac{1}{[K:\mathbb{Q}]} \sum_{v \in M^\infty(K)} d_v \log \max(1, |x|_v) = \frac{1}{[K:\mathbb{Q}]} \sum_{\sigma: K \to \mathbb{C}} \log \max(1, |\sigma(x)|).$$

Lemma 16. *The norm is invariant under field isomorphisms: Let F and K be number fields and $\phi : F \to K$ a field isomorphism. Then $h_K(x) = h_F(\phi(x))$ for all $x \in F$.*

In this proof we will use some facts from algebraic number theory.

Proof. For any prime ideal $P \subset \mathbb{Z}_F$, $\phi^{-1}(P)$ is a prime ideal of $\mathbb{Z}_K$. Furthermore, we have $N(P)) = N(\phi^{-1}(P))$, hence $f(P) = f(\phi^{-1}(P))$. Let $p\mathbb{Z} = P\mathbb{Z}_F \cap \mathbb{Z}$, then $e(P) = e(\phi^{-1}(P))$ by comparing the prime ideal factorization of $p\mathbb{Z}_K$ and $p\mathbb{Z}_F$. This proves the equiality of the finite parts of the height. For any embedding $\sigma : F \to \mathbb{C}$, we get an embedding $\sigma \circ \phi : K \to \mathbb{C}$ and vice versa. This proves the equality of the infinite parts of the height. □

We can now prove the extended version of Lemma 3.

Lemma 17. *Let $x \in \overline{\mathbb{Q}}$. Then the following properties hold.*

1. $h(x) \geq 0$.
2. *For* $x \neq 0$ *and* $k \in N$ *we have* $h(x^k) = kh(x)$.
3. *For* $x \neq 0$ *we have* $h(x^{-1}) = h(x)$.
4. *For* $x \neq 0$ *and* $k \in \mathbb{Z}$ *we have* $h(x^k) = |k|h(x)$.
5. $h(xy) \leq h(x) + h(y)$ *and* $h(\pm x) = h(x)$.
6. $h(x + y) \leq h(x) + h(y) + \log 2$.

Proof. Exercise □

With all this properties we can prove two important theorems of height theory.

Theorem 18 (Northcott)**.** *Subsets of the algebraic numbers of bounded height and degree are finite.*

Proof. Let $x \in \overline{\mathbb{Q}}$ with $h(x) \leq B$ for a fixed B and $[\mathbb{Q}(x) : \mathbb{Q}] \leq D$ for a fixed D. Let $A \in \mathbb{Z}[X]$ be the integral minimal polynomial of x. We can factorize A over $\mathbb{C}$ as $A(X) = a_d(X - x_1) \cdots (X - x_d)$. We will use Definition 4 of the height to find

$$a_d \prod_{i=1}^{d} \max(1, |x_i|) = e^{\deg(A)h(x)} = e^{[\mathbb{Q}(x):\mathbb{Q}]h(x)} \leq e^{DB}.$$

This shows that $1 \leq a_d \leq e^{DB}$. We want to write $A(X) = a_d X^d + \ldots + a_1 X + a_0$ with $a_i \in \mathbb{Z}$. By comparison of coefficients of this form and the linear factorization we find that the $\frac{a_{d-1}}{a_d}, \ldots, \frac{a_0}{a_d}$ are up to sign elementar symmetric polynomials in the x_i. So for $0 \leq k \leq d - 1$ we find

$$|a_k| = a_d \Big| \sum_{i_1 > \ldots > i_k} x_{i_1} \cdots x_{i_k} \Big| \leq a_d \max(1, |x_1|) \cdots \max(1, |x_d) \sum_{i_1 > \ldots > i_k} 1$$
$$\binom{d}{k} e^{[\mathbb{Q}(x):\mathbb{Q}]h(x)} \leq \binom{D}{k} e^{DB}.$$

Since $\binom{d}{k} \leq 2^d \leq 2^D$ we find $|a_k| \leq e^{DB} 2^D$ for all $0 \leq k \leq d$. The coefficients of A are bounded by functions depending on B and D. Since the coefficients are integers and the degree of A is also bounded, there are only finitely many possibilities for A and hence also only finitely many x_i. □

Theorem 19 (Kronecker)**.** *The height of an algebraic number is zero if and only if it is either zero itself or a root of unity.*

Proof. Let us start with an algebraic number $x \in \overline{\mathbb{Q}}$ of height zero. From Lemma 17 we know $h(x) = h(x^2) = h(x^3) = \ldots = 0$. On the other hand, all x^k with $k \geq 0$ lie in the number field $\mathbb{Q}(x)$ and hence also in the subset of $\overline{\mathbb{Q}}$ of height bounded by zero and degree bounded by $[\mathbb{Q}(x) : \mathbb{Q}]$, hence in a finite set by the above theorem of Northcott. So in the series x^k there have to be indices $i < j$ such that $x^i = x^j$. For $x \neq 0$ this means that x is a root of unity. This is one direction of the equivalence in the statement. The other one is an exercise. □

Exercises

Exercise 1 (Important)**.** Prove Lemma 3.

Exercise 2. Prove Remark 7.

Exercise 3. Prove Lemma 8

Exercise 4. Prove Remark 9

Exercise 5 (Important)**.** Prove Lemma 17

Exercise 6 (Important)**.** Prove that a root of unity has height zero.

2 Bilu's Theorem

Example 20. Einfuehrendes Beispiel (oder mehrere), zum Beispiel Konjugierte von $2^{\frac{1}{d}}$ als Bild oder ohne Bild, weil es einfacher ist.

So algebraic numbers of small height seem to be equidistributed around the unit circle. We have to find a mathematical term for equidistribution.

Definition 21. For any $n \in \mathbb{N}$ fix $d_n \in \mathbb{N}$ and a tuple $z_n = (z_{n,1}, \ldots, z_{n,d_n}) \in \mathbb{C}^{d_n}$. We call the series $(z_n)_{n\in\mathbb{N}}$ *equidistributed* (with respect to the unit circle) if the following holds: For any continuous function $f : \mathbb{C}^* \to \mathbb{R}$ with compact support we have

$$\frac{1}{d_n}\sum_{j=1}^{d_n} f(z_{n,j}) = \int_0^1 f(e^{2\pi it})dt.$$

So instead of considering the average sum of the values of the function, we can just integrate around the unit circle.

Bilu's Theorem implies that embeddings of a series of small height is equidistributed.

Theorem 22 (Bilu)**.** *Let $(\alpha_n)_{n\in\mathbb{N}}$ be a series of pairwise distinct algebraic numbers with $\lim_{n\to\infty} h(\alpha) = 0$. Let $f : \mathbb{C}^* \to \mathbb{R}$ be a continuous function with compact support. Then*

$$\lim_{n\to\infty} \frac{1}{[\mathbb{Q}(\alpha_n) : \mathbb{Q}]} \sum_{\sigma:\mathbb{Q}(\alpha_n)\to\mathbb{C}} f(\sigma(\alpha_n)) = \int_0^1 f(e^{2\pi iz})dt.$$

The goal of this chapter is proving Bilu's Theorem.

For the rest of this chapter consider an algebraic number $x \in \overline{\mathbb{Q}}$ of "small" height: $h(x) \leq \delta$ with $\delta \in (0, \frac{1}{2}]$. Furthermore, let P be the integral minimal polynomial as in Definition 4.

We consider the linear factorization of P: $P(X) = a_d(X - x_1)\cdots(X - x_d)$, where $d = [\mathbb{Q}(x) : \mathbb{Q}]$, $a_d \in \mathbb{Z}$ and $x_i \in \mathbb{C}$. We want to sort the x_i with respect to their height. Let k be such that

$$|\log|x_i|| \leq \sqrt{\delta} \text{ for all } i \leq k, \tag{2.1}$$

$$|\log|x_i|| > \sqrt{\delta} \text{ for all } i > k. \tag{2.2}$$

For small δ, the x_i are close to the unit circle. Furthermore, define $\theta_i \in [0, 2\pi)$ such that

$$x_i = |x_i|e^{\sqrt{-1}\theta_i}. \tag{2.3}$$

Lemma 23. *Let $x \in \overline{\mathbb{Q}}^*$, then*

$$h(x) = \frac{1}{2[\mathbb{Q}(x) : \mathbb{Q}]} \sum_{v\in M(K)} |d_v \log(|x|_v)|$$

Proof. With Lemma 17 we have $h(x) = h(x^{-1})$. By Definition 11 and $K = \mathbb{Q}(x)$ we have

$$h(x) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M(K)} |d_v \log(\max(1, |x|_v))|$$
$$h(x^{-1}) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M(K)} |d_v \log(\max(1, |x^{-1}|_v))|.$$

We want to consider the sum $2h(x) = h(x) + h(x^{-1})$ and get

$$\begin{aligned} 2h(x) &= \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M(K)} |d_v \log(\max(1, |x|_v))| + \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M(K)} |d_v \log(\max(1, |x^{-1}|_v))| \\ &= \frac{1}{[K:\mathbb{Q}]} \sum_{\substack{v \in M(K) \\ |x|_v \geq 1}} |d_v \log |x|_v| + \frac{1}{[K:\mathbb{Q}]} \sum_{\substack{v \in M(K) \\ |x^{-1}|_v \geq 1}} |d_v \log |x^{-1}|_v| \\ &= \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M(K)} |d_v \log(|x|_v)|. \end{aligned}$$

□

Lemma 24. *We have* $\frac{d-k}{d} \leq 2\sqrt{\delta}$ *and* $a_d \leq e^{d\delta}$.

The difference $d - k$ is the cardinality of the conjugates of x that are not "close" to the unit circle.

Proof. From algebraic number theory we know that the Galois group $\mathrm{Gal}(\mathbb{Q}(x)/\mathbb{Q})$ acts transitively on the conjugates of x. This exactly means that the x_i are the images $\{\sigma(x) | \sigma : \mathbb{Q}(x) \to \mathbb{C}\}$.

By Lemma 23, we have

$$\begin{aligned} 2dh(x) &\geq \sum_{i=1}^{d} |\log |x_i|| \\ &\geq \sum_{i=k+1}^{d} |\log |x_i|| \\ &\geq (d-k)\sqrt{\delta}. \end{aligned}$$

Since we also have $2\delta \geq 2h(x)$, this means that $\frac{d-k}{d} d \leq d\sqrt{\delta}$ which is the first statement.

By the definition 4 of the height, we have $h(x) \geq \frac{\log a_0}{d}$, hence $a_0 \leq e^{d\delta}$ which is the second statement. □

Lemma 25. *For any* $n \in \mathbb{N}$*, we have*

$$\Big|\sum_{i=1}^{k} e^{\sqrt{-1} n \theta_i}\Big|^2 \leq (13d^2\sqrt{\delta} + d|\log \delta|) n \rho^n \text{ where } \rho = e^{2\sqrt{\delta}}.$$

Proof. We will use a fact about the discriminant of algebraic number theory as a black box:

$$0 \leq \log|\operatorname{Disc}(P)| = (2d-2)\log a_d + \sum_{\substack{1\leq i,j\leq d \\ i\neq j}} \log|x_i - x_j| \tag{2.4}$$

$$= (2d-2)\log a_d + S_1 + 2S_2 \tag{2.5}$$

where

$$S_1 = \sum_{\substack{1\leq i,j\leq k \\ i\neq j}} \log|x_i - x_j|,$$

$$S_2 = \sum_{\substack{1\leq i<j\leq d \\ k<j}} \log|x_i - x_j|.$$

We treat the sums independently and start with S_2. We use the following inequality which holds for all $z, w \in \mathbb{C}$ with $z \neq w$.

$$\begin{aligned}\log|z-w| &\leq \log(2\max(|z|,|w|) \\ &\leq \log 2 + \log\max(1,|z|) + \log\max(1,|w|) \\ &\leq 1 + \log\max(1,|z|) + \log\max(1,|w|).\end{aligned}$$

For fixed $j \leq d$ we get

$$\begin{aligned}\sum_{1\leq i<j} \log|x_i - x_j| &\leq d + d\log\max(1,|x_j|) + \sum_{1\leq i\leq d} \log\max(1,|x_i|) \\ &\leq d(1 + \log\max(1,|x_j|) + h(x)) \\ &\leq d(1 + \log\max(1,|x_j|) + \delta).\end{aligned}$$

Now we take the sum over all $j > k$:

$$\begin{aligned}S_2 &= \sum_{\substack{1\leq i<j\leq d \\ k<j}} \log|x_i - x_j| \\ &\leq \sum_{j=k+1}^{d} d(1 + \log\max(1,|x_j|) + \Delta) \\ &\leq d(d-k) + d\sum_{j=1}^{d} \log\max(1,|x_j|) + d(d-k)\delta \\ &\leq d(d-k)(1+\delta) + d^2\delta.\end{aligned}$$

By Lemma 24, we have $d - k \leq 2d\sqrt{\delta}$, hence (since $\delta < \frac{1}{2}$)

$$\begin{aligned}S_2 &\leq 2d^2\sqrt{\delta}(1+\delta) + d^2\delta \\ &\leq 3d^2\sqrt{\delta} + d^2\sqrt{\delta} \\ &= 4d^2\sqrt{\delta}.\end{aligned} \tag{2.6}$$

Now we turn to S_1. Let $i, j \in \{1, \dots, k\}$ with $i \neq j$. We want to bound $|x_i - x_j|$ and use without loss of generality that $|x_i| \leq |x_j|$. Then

$$\log|x_i - x_j| = \log|x_i| + \log\Big|1 - \frac{|x_j|}{|x_i|} e^{\sqrt{-1}(\theta_j - \theta_i)}\Big|.$$

By property (2.1) we have $|x_j| \leq e^{\sqrt{\delta}}$ and $|x_i^{-1}| \leq e^{\sqrt{\delta}}$, hence $1 \leq \frac{|x_j|}{|x_i|} \leq \rho = e^{2\sqrt{\delta}}$. From elementary geometry we know (or can prove easily) that for any $1 \leq r \leq \rho$ and $\theta \in \mathbb{R}$ we have

$$|1 - re^{\sqrt{-1}\theta}| \leq |1 - \rho e^{\sqrt{-1}\theta}|. \tag{2.7}$$

We apply this inequality with $r = |\frac{x_j}{x_i}|$ and $\theta = \theta_j - \theta_i$ and get

$$\begin{aligned}\log|x_i - x_j| &\leq \log|x_i| + \log\Big|1 - \frac{|x_j|}{|x_i|} e^{\sqrt{-1}(\theta_j - \theta_i)}\Big| \\ &\leq \sqrt{\delta} + \log|1 - \rho e^{\sqrt{-1}(\theta_j - \theta_i)}|\end{aligned}$$

This means that $\log|x_i - x_j| \leq 3\sqrt{\delta} + \log|1 - \rho^{-1} e^{\sqrt{-1}(\theta_j - \theta_i)}|$ which holds for all $1 \leq i, j \leq k$ with $i \neq j$. Taking the sum over those pairs, we find

$$S_1 \leq 3k^2\sqrt{\delta} + \sum_{\substack{1 \leq i,j \leq k \\ i \neq j}} \log\left|1 - \rho^{-1} e^{\sqrt{-1}(\theta_j - \theta_i)}\right|.$$

Since $\rho^{-1} \in (0, 1)$ we have $\log|1 - \rho^{-1}| = \log(1 - \rho^{-1})$ and the logarithm is well defined. We get

$$S_1 \leq 3k^2\sqrt{\delta} - k\log(1 - \rho^{-1}) + \sum_{1 \leq i,j \leq k} \log\left|1 - \rho^{-1} e^{\sqrt{-1}(\theta_j - \theta_i)}\right|.$$

Let for the moment $z = \rho^{-1} e^{\sqrt{-1}(\theta_j - \theta_i)}$, then $|z| = \rho^{-1} < 1$. The Taylor series $-\sum_{n \geq 1} \frac{z^n}{n}$ converges to $\log(1 - z)$. The real part of $\log(1 - z)$ is $\log|1 - z|$, so

$$\begin{aligned}S_1 &\leq 3k^2\sqrt{\delta} - k\log(1 - \rho^{-1}) + \Re\left(\sum_{1 \leq i,j \leq k} \log(1 - z)\right) \\ &\leq 3k^2\sqrt{\delta} - k\log(1 - \rho^{-1}) + \Re\left(\sum_{n \geq 1} \frac{\rho^{-n}}{n} \sum_{1 \leq i,j \leq k} e^{\sqrt{-1}(\theta_j - \theta_i)}\right).\end{aligned}$$

By expanding we find $\sum_{1 \leq i,j \leq k} e^{\sqrt{-1}(\theta_j - \theta_i)} = \big|\sum_{i=1}^n e^{\sqrt{-1} n \theta_i}\big|^2$, so

$$S_1 \leq \leq 3k^2\sqrt{\delta} - k\log(1 - \rho^{-1}) - \sum_{n \geq 1} \frac{\rho^{-n}}{n} \big|\sum_{i=1}^k e^{\sqrt{-1}\theta_i}\big|^2.$$

Now we have $k \leq d$ and $-\log(1 - \rho^{-1}) = l\log(1 - e^{-2\sqrt{\delta}})$. Elementary analysis gives $-\log(1 - e^{-2\sqrt{\delta}}) \leq \log\left(\frac{1}{\delta}\right)$ since $\delta \leq \frac{1}{2}$. Altogether we proved

$$S_1 \leq 3d^2\sqrt{\delta} + d\log\left(\frac{1}{\delta}\right) - \sum_{n \geq 1} \frac{\rho^{-n}}{n} \big|\sum_{i=1}^k e^{\sqrt{-1} n \theta_i}\big|^2.$$

We recall equation (2.5) and $a_d \leq ed\delta$ from Lemma 24. Together with the bound (2.6) for S_2 we find

$$0 \leq 2d(d-1)\delta + 8d^2\sqrt{\delta} + 3d^2\sqrt{\delta} + d\log\left(\frac{1}{\delta}\right) - \sum_{n\geq 1}\frac{\rho^{-n}}{n}\Big|\sum_{i=1}^{k} e^{\sqrt{-1}n\theta_i}\Big|^2.$$

The Proposition follows now with $\delta \leq 1$ and

$$\sum_{n\geq 1}\frac{\rho^{-n}}{n}\Big|\sum_{i=1}^{k} e^{\sqrt{-1}n\theta_i}\Big|^2 \leq 15d^2\sqrt{\delta} + d\log\left(\frac{1}{\delta}\right).$$

□

Now we want to sum over *all* i, not only over those where x_i is close to the unit circle. We get the following corollary.

Corollary 26. *We keep the notation from above. For any $n \in \mathbb{N}$ we have*

$$\frac{1}{d}\Big|\sum_{i=1}^{d} e^{\sqrt{-1}n\theta_i}\Big| \leq \left(15\sqrt{\delta}\frac{|\log\delta|}{d}\right)^{\frac{1}{2}} n^{\frac{1}{2}}e^{n\sqrt{\delta}} + 2\sqrt{\delta}.$$

Proof. Exercise. □

Example 27. We want to explain what the Corollary above has to do with equidistribution. We call an algebraic number x *totally real* if all $\sigma(x) \in \mathbb{R}$ for all field embeddings $\sigma : \mathbb{Q}(x) \to \mathbb{C}$. We keep the notation of $h(x) \leq \delta$ with $x \neq 0$ algebraic and $\delta \leq \frac{1}{2}$ small. Let furthermore $x_1, \ldots, x_d$ be the conjugates of x.

With Bilu's Theorem, the conjugates should be equidistributed around the unit circle for $\delta \to 0$. One consequence of that is that a totally real number cannot have arbitrarily small height. We will prove this using the above corollary.

We let $n = 2$ and $x_i \in \mathbb{R}$. Then all θ_i as in (2.3) are either 0 or π. With $n = 2$ we have $e^{\sqrt{-1}n\theta_i} = 1$ and hence

$$\Big|\sum_{i=1}^{d} e^{\sqrt{-1}n\theta_i}\Big| = d.$$

The corollary implies

$$1 \leq (30\sqrt{\delta} + \frac{2}{d}\log\left(\frac{1}{\delta}\right))^{\frac{1}{2}}e^{2\sqrt{\delta}} + 2\sqrt{\delta}. \tag{2.8}$$

For $\delta = \frac{1}{2000}$, hence $h(x) \leq \frac{1}{2000}$ we get

$$(30\sqrt{\delta})^{\frac{1}{2}}e^{2\sqrt{\delta}} + 2\sqrt{\delta} < 1.$$

From inequality (2.8) we find that $d \leq 100$. By Northcott's Theorem there are only finitely many $x \in \overline{\mathbb{Q}}$ with $h(x) \leq \frac{1}{2000}$ and $[\mathbb{Q}(x) : \mathbb{Q}] \leq 100$. So there exists $\varepsilon > 0$ with $h(x) = 0$ or $h(x) \geq \varepsilon$. With Kronecker's Theorem and since $x_i \in \mathbb{R}$ for all i, we have $h(x) = 0$ if and only if $x \in \{0, \pm 1\}$. Now we proved the following statement:

Corollary 28. *There is an absolute constant $\varepsilon > 0$ such that for all totally real algebraic numbers that are not equal to either 0 or ± 1 we have $h(x) \geq \varepsilon$.*

First, we want to look at a certain class of functions we will call *admissible*.

Definition 29 (Admissible)**.** We call a function $f : \mathbb{C}^* \to \mathbb{C}$ *admissible*, if $f(z) = g(|z|)(\frac{z}{|z|})^n$ with $n \in \mathbb{Z}$ and $g : (0, \infty) \to \mathbb{C}$ a continuous function with compact support. This means that there is an $r \in (0, 1]$ with $g(t) = 0$ for all $t \in (0, r] \cup [\frac{1}{r}, \infty)$.

Writing $z = re^{\sqrt{-1}\theta}$ with $r > 0$ and $\theta \in \mathbb{R}$ in polar coordinates, an admissible function f with g as above has to fulfill the identity

$$f(z) = g(r)e^{\sqrt{-1}\theta n}.$$

Proposition 30. *Let f be an admissible function and n as in the definition above. Let $I = 0$ whenever $n \neq 0$ and $I = f(1)$ otherwise. For any $\varepsilon > 0$ there exists $\delta > 0, d_0 \geq 1$ such that for any algebraic $x \neq 0$ with $h(x) \leq \delta$ and $d = [\mathbb{Q}(x) : \mathbb{Q}] \geq d_0$ we have*

$$\left| \left(\frac{1}{d} \sum_{i=1}^{d} f(x_i) \right) - I \right| \leq \varepsilon, \tag{2.9}$$

where x_i are the conjugates of x.

Proof. Since f is an admissible function, we find $f(z) = g(|z|)(\frac{z}{|z|})^n$ with g as in the Definition 29. Since g is continuous with compact support, it is also bounded. So there exists $M \geq 0$ such that $g(t) \leq M$ for any $t \geq 0$. We will choose $\delta \in (0, \frac{1}{2}]$ to be small depending on f and ε. Furthermore, we will choose d_0 to be large depending on δ. Let $x_i \in \mathbb{C}$ be as above and sort them as in (2.1) and (2.2) where δ is still to be chosen.

We will treat two cases and start with $n = 0$. We have $f(z) = g(z)$ and $I = f(1) = g(1)$. On average we have

$$\begin{aligned} \left| \left(\frac{1}{d} \sum_{i=1}^{d} f(x_i) \right) - g(1) \right| &\leq \frac{1}{d} \sum_{i=1}^{k} |g(|x_i|) - g(1)| + \frac{1}{d} \sum_{i=k+1}^{d} |g(|x_i|) - g(1)| \\ &\leq \frac{1}{d} \sum_{i=1}^{k} |g(|x_i|) - g(1)| + \frac{1}{d} \sum_{i=k+1}^{d} 2M. \end{aligned}$$

For $i \leq k$ we have $|\log |x_i|| \leq \delta$, so $|x_i|$ is close to 1. For $\delta > 0$ small enough we have $|g(|x_i|) - g(1)| \leq \frac{\varepsilon}{2}$ (for $i \leq k$). We find that

$$\left| \left(\frac{1}{d} \sum_{i=1}^{d} f(x_i) \right) - g(1) \right| \leq \frac{\varepsilon}{2} + \frac{2M(d-k)}{d} \leq \frac{\varepsilon}{2} + 4M\sqrt{\delta}$$

by Lemma 24. Furthermore we can choose δ such that $4M\sqrt{\delta} \leq \frac{\varepsilon}{2}$ which proves the statement for $n = 0$.

We turn to the case $n \neq 0$. We have $I = 0$, so we only have to bound the absolute value of the sum. We have $f(x_i) = g(|x_i|)\frac{x_i}{|x_i|}^n = g(|x_i|)e^{\sqrt{-1}\theta_i n}$ in polar coordinates. Then we have

$$\begin{aligned}\frac{1}{d}\left|\sum_{i=1}^{d} f(x_i)\right| &= \frac{1}{d}\left|\sum_{i=1}^{d}(g(|x_i|) - g(1))e^{\sqrt{-1}\theta_i n} + g(1)e^{\sqrt{-1}\theta_i n}\right| \\ &\leq \frac{1}{d}\sum_{i=1}^{d}\left|g(|x_i|) - g(1)\right| + \frac{g(1)}{d}\left|\sum_{i=1}^{d} e^{\sqrt{-1}\theta_i n}\right|. \end{aligned} \tag{2.10}$$

Like in the first part, we may assume that $\delta > 0$ is so small that $|g(|x_i|) - 1| \leq \frac{\varepsilon}{5}$ for all $i \leq k$. For $i > k$ we have $|g(|x_i|) - g(1)| \leq 2M$ anyway because of the boundedness of g. We divide the sum into two parts and find

$$\frac{1}{d}\sum_{i=1}^{d}|g(|x_i|) - g(1)| \leq \frac{\varepsilon}{5} + 2M|g(1)|\frac{d-k}{d} \leq \frac{\varepsilon}{5} + 4M|g(1)|\sqrt{\delta} \leq \frac{\varepsilon}{2}$$

for δ small enough. This is the first sum in (2.10).

For the second sum we use Corollary 26. There we actually need $n \geq 1$. But since complex conjugation does not change our set of conjugates, the sum over $e^{\sqrt{-1}\theta_i n}$ is the same as the sum over $e^{-\sqrt{-1}\theta_i n}$ and we can assume $n \geq 1$:

$$\frac{1}{d}\left|\sum_{i=1}^{d} e^{\sqrt{-1}\theta_i n}\right| \leq (15\sqrt{\delta} + \frac{|\log\delta|}{d})^{\frac{1}{2}} n^{\frac{1}{2}} e^{n\sqrt{\delta}} + 2\sqrt{\delta}.$$

For $\delta > 0$ small enough, we have $|g(1)|(15\sqrt{\delta}n)^{\frac{1}{2}}e^{n\sqrt{\delta}} + 2\sqrt{\delta} \leq \frac{\varepsilon}{4}$. So for $d \geq d_0$ where d_0 is large enough depending on n, $|g(1)|$ and δ we have with the above inequality that

$$\frac{|g(1)|}{d}\left|\sum_{i=1}^{k} e^{\sqrt{-1}\theta_i n}\right| \leq \frac{\varepsilon}{2}.$$

After putting everything together, we get

$$\frac{1}{d}\left|\sum_{i=1}^{d} f(x_i)\right| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Which is what we wanted to prove for $n \neq 0$. □

Remark 31. With the notation from above, I is the expected limit of $\frac{1}{d}\sum_{i=1}^{d} f(x_i)$ as $i \to \infty$ since $\int_0^1 f(e^{2\pi\sqrt{-1}\theta})d\theta = I$.

Corollary 32. *Bilu's Theorem holds for all admissible functions $f : \mathbb{C}^* \to \mathbb{C}$.*

We know need two more statements on how to use Bilu's Theorem for admissible functions to prove it for all continuous functions with compact support. We will state two statements without proof. The proof of Fejér's Theorem can be found in the literature.

Theorem 33 (Fejér). *Let $r_0 \in (0,1]$ and for any $r \in [r_0, \frac{1}{r_0}]$ let $f_r : \mathbb{R} \to \mathbb{C}$ given with $f(\theta) = f(\theta + 2\pi)$ for all θ. Assume that $(r,\theta) \mapsto f_r(\theta)$ is continuous. The Fourier coefficients of f are*

$$c_{r,k} = \frac{1}{2\pi} \int_0^{2\pi} f(re^{\sqrt{-1}\theta}) e^{\sqrt{-1}\theta} d\theta \tag{2.11}$$

for all $k \in \mathbb{Z}$. Define $s_{r,n}(\theta) = \sum_{k=-n}^{n} c_{r,k} e^{\sqrt{-1}k\theta}$. For $\varepsilon > 0$ there exists $N_0 = N_0(\varepsilon)$ such that

$$\left| f_r(\theta) - \frac{1}{N} \sum_{n=0}^{N-1} s_{r,n}(\theta) \right| \leq \varepsilon$$

for all $\theta \in \mathbb{R}, r \in [r_0, \frac{1}{r_0}]$ and all $N \geq N_0$.

Corollary 34. *Let $f : \mathbb{C}^* \to \mathbb{C}$ be a continuous function with compact support and $\varepsilon > 0$. Then there exists a finite $\mathbb{C}$-linear combination of admissible functions $\tilde{f}$ such that*

$$|f(z) - \tilde{f}(z)| \leq \varepsilon$$

for all $z \in \mathbb{C}^*$.

We write the proof here, but omit it in the lecture.

Proof. Let $r_0 \in (0,1]$ with $f(z) = 0$ flor $|z| \leq r_0$ or $|z| \geq \frac{1}{r_0}$. For any $r > 0$ and any $n \in \mathbb{Z}$ we define

$$c_{r,k} = \frac{1}{2\pi} \int_0^{2\pi} f(re^{\sqrt{-1}\theta}) e^{\sqrt{-1}\theta} d\theta.$$

Let $s_{r,n}(re^{\sqrt{-1}\theta}) = \sum_{k=-n}^{n} c_{r,k} e^{\sqrt{-1}\theta}$ as in the Theorem of Fejér. Then $f_N(re^{\sqrt{-1}\theta}) = \frac{1}{N} \sum_{n=0}^{N-1} s_{r,n}(re^{\sqrt{-1}\theta})$ converges gleichmassig to $f(re^{\sqrt{-1}\theta})$ for $r \in [\frac{1}{r_0}, r_o]$ and $\theta \in \mathbb{R}$. Furthermore we have $s_n(z) = 0$ for $|z| \leq r_0$ or $|z| \geq \frac{1}{r_0}$ and the same holds for f_N. So f_N is a finite $\mathbb{C}$-lineare combination of admissible functions and the statement follows. □

Now we have collected all technicalities to prove the actual Theorem of Bilu.

Proof. We start with a series $\alpha_1, \alpha_2, \ldots$ non-zero, pairwise distinct algebraic numbers with $\lim_{n\to\infty} h(\alpha_n) = 0$. With Northcott's Theorem 18 we get that $[\mathbb{Q}(\alpha_n) : \mathbb{Q}] \to \infty$ as $n \to \infty$. So let $f : \mathbb{C}^* \to \mathbb{C}$ be a continuous function with compact support and let $\varepsilon > 0$. With the above corollary we find admissible functions $f_1, \ldots, f_N$ and complex numbers $\lambda_1, \ldots, \lambda_N$ such that $|f(z) - \tilde{f}(z)| \leq \frac{\varepsilon}{3}$ for all $z \in \mathbb{C}^*$, where $\tilde{f}(z) = (\lambda_1 f_1 + \ldots + \lambda_N f_N)(z)$. Now we use Proposition 30 with $\frac{\varepsilon}{3A}$ where $A = |\lambda_1| + \ldots + |\lambda_N|$ and get $\delta > 0$ and $d_0 \geq 1$. The statement of the proposition then holds for α_n whenever $h(\alpha_n) \leq \delta$ and $[\mathbb{Q}(\alpha_n) : \mathbb{Q}] \geq d_0$. Since the limit of the height of α_n tends to zero and the degrees $[\mathbb{Q}(\alpha_n) : \mathbb{Q}]$ tend to infinity, both conditions apply for n suitable large. For such n we get

$$\left| \frac{1}{[\mathbb{Q}(\alpha_n) : \mathbb{Q}]} \sum_{\sigma : \mathbb{Q}(\alpha_n) \to \mathbb{C}} f_i(\sigma(\alpha_n)) - \int_0^1 f_i(e^{2\pi\sqrt{-1}\theta}) d\theta \right| \leq \frac{\varepsilon}{3A}$$

for all $i \leq N$. We multiply by $|\lambda_i|$ and sum over $i \leq N$ to get

$$\left| \frac{1}{[\mathbb{Q}(\alpha_n) : \mathbb{Q}]} \sum_{\sigma : \mathbb{Q}(\alpha_n) \to \mathbb{C}} f_i(\sigma(\alpha_n)) - \int_0^1 f_i(e^{2\pi\sqrt{-1}\theta}) d\theta \right| \leq \frac{\varepsilon}{3A}(|\lambda_1| + \ldots + |\lambda_N|) \leq \frac{\varepsilon}{3}. \tag{2.12}$$

Now we can use the fact that $|f(z) - \tilde{f}(z)| \leq \frac{\varepsilon}{3}$ for all $z \in \mathbb{C}^*$ and find

$$\left| \frac{1}{[\mathbb{Q}(\alpha_n) : \mathbb{Q}]} \sum_{\sigma : \mathbb{Q}(\alpha_n) \to \mathbb{C}} \tilde{f}_i(\sigma(\alpha_n)) - \frac{1}{[\mathbb{Q}(\alpha_n) : \mathbb{Q}]} \sum_{\sigma : \mathbb{Q}(\alpha_n) \to \mathbb{C}} f_i(\sigma(\alpha_n)) \right| \leq \frac{\varepsilon}{3}$$

and

$$\left| \int_0^1 \tilde{f}_i(e^{2\pi\sqrt{-1}\theta}) d\theta - \int_0^1 f_i(e^{2\pi\sqrt{-1}\theta}) d\theta \right| \leq \frac{\varepsilon}{3}.$$

We use there inequalities together with inequality (2.12) and find

$$\left| \frac{1}{[\mathbb{Q}(\alpha_n) : \mathbb{Q}]} \sum_{\sigma : \mathbb{Q}(\alpha_n) \to \mathbb{C}} f(\sigma(\alpha_n)) - \int_0^1 f(e^{2\pi\sqrt{-1}\theta} d\theta \right| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \leq \varepsilon$$

for all suitable large n depending on ε. So it follows that

$$\lim_{n \to \infty} \frac{1}{[\mathbb{Q}(\alpha_n) : \mathbb{Q}]} \sum_{\sigma : \mathbb{Q}(\alpha_n) \to \mathbb{C}} f(\sigma(\alpha_n)) = \int_0^1 f(e^{2\pi\sqrt{-1}\theta} d\theta$$

which is the statement we wanted to show. □

We can generalize Bilu's Theorem to continuous and bounded functions $f : \mathbb{Q}' \to \mathbb{C}$. But further generalizations usually lead to counterexamples:

Example 35. Bilu's Theorem does not hold for unbounded continuous functions f. It is already wrong for the identity function $f(z) = z$: Let d be a suitably large natural number. Then we have $\pi(d) \leq d$ where π is the prime number counting function. Furthermore, $\pi(2d^2) \geq d^2 \frac{\log 2}{\log(2d^2)}$. So for $d \geq 2$ suitably large, we have $\pi(d) < \pi(2d^2)$ and there exists a prime p such that $d < p \leq 2d^2$.

We define the polynomial $P = X^d + pX^{d-1} + p \in \mathbb{Z}[X]$. With Eisenstein's criterion we find that P is irreducible as an element of $\mathbb{Q}[X]$. Let $P(x) = 0$ with x algebraic, then we have $x \neq 0$ and $y = \frac{1}{x}$ satisfies the equation $py^d + py + 1 = 0$. With Lemma 17 we get

$$h(py^d) = h(-py - 1) \leq \log 2 + h(py) \leq \log(2p) + h(y)$$

and (exercise)

$$h(py^d) \geq dh(y) - \log p.$$

Hence

$$h(y) \leq \frac{\log(2p^2)}{d-1} \leq \frac{\log(8d^4)}{d-1}.$$

So we also have $h(x) = h(x^{-1}) = h(y) \leq \frac{\log(8d^4)}{d-1}$. For $d \to \infty$ we get a series of algebraic numbers of small height. We want to average the conjugates and use $f(z) = z$. Using the formula for the trace of an algebraic number, we get

$$\frac{1}{d}\left|\sum_{\sigma:\mathbb{Q}(x)\to\mathbb{C}} \sigma(x)\right| = \frac{p}{d} > 1. \tag{2.13}$$

On the other hand

$$\int_0^1 f(e^{2\pi\sqrt{-1}t})dt = \int_0^1 e^{2\pi\sqrt{-1}t}dt = 0.$$

So the expression in (2.13) cannot converge to the integral above for $d \to \infty$.

Exercises

Exercise 7. "Prove" inequality (2.7) by drawing a picture.

Exercise 8 (Important)**.** Prove that for $y \in \overline{\mathbb{Q}}$, $d \in \mathbb{N}$ and p a prime, we have $h(py^d) \geq dh(y) - \log p$.

Exercise 9. Prove Corollary 26.

Exercise 10 (Important)**.** Let $S = \{z \in \mathbb{C}^* | z = re^{\sqrt{-1}\theta}\}$ with $r > 0$ and $|\theta| < \frac{\pi}{3}\}$. For any algebraic number x let $p = \frac{1}{d}|\{1 \leq i \leq d|\ x_i \in S\}|$. Then for any ε we can find $\delta \leq \frac{1}{2}$ such that for x with $h(x) \leq \delta$ we have $p \leq \frac{2(1+\varepsilon)}{3}$.

Prove the statement using Corollary 26. What does it mean? What would you expect p actually to be?

Exercise 11. Complete the proof of Proposition 30.

Exercise 12. Prove Corollary 32.

Exercise 13 (Important)**.** Let $S = \{z \in \mathbb{C}^* | z = re^{\sqrt{-1}\theta}\}$ with $r > 0$ and $|\theta| < \frac{\pi}{3}\}$. For any algebraic number x let $p = \frac{1}{d}|\{1 \leq i \leq d|\ x_i \in S\}|$. Then for any ε we can find $\delta \leq \frac{1}{2}$ such that for x with $h(x) \leq \delta$ we have $p \leq \frac{2(1+\varepsilon)}{3}$.

Prove the statement using Bilu's Theorem.

Exercise 14 (Important)**.** Also logarithmic singularities cannot be covered by Bilu's Theorem. Let $f(z) = \log\min(1, |z-2|)$ for $z \in \mathbb{C}^* \setminus \{2\}$ and $f(2)$ arbitrary, then let

$$F = \frac{1}{[\mathbb{Q}(x):\mathbb{Q}]}\sum_{\sigma:\mathbb{Q}(x)\to\mathbb{C}} f(\sigma(x)).$$

Show that F cannot converge to $\log 2$ if $h(x) \to 0$ and $[\mathbb{Q}(x) : \mathbb{Q}] \to \infty$. Show that $\int_0^1 f(e^{2\pi\sqrt{-1}t})dt = \log 2$.

3 Elliptic Curves

In this chapter we will give the basic definitions in the theory of elliptic curves. Since the results in this section are basic and available in all standard books we will skip the proofs and refer to [Sil09] for deeper interest. We will also closely follow Silverman's notations and definitions.

Definition 36 (Elliptic curve)**.** An elliptic curve E over a field K is given by the equation $Y^2 = X^3 + AX + B$ with $A, B \in K$ and $4A^3 + 27B^2$ is non-zero. Then for any field L containing K we set $E(L) := \{(x, y) \in L^2 | y^2 = x^3 + Ax + B\} \cup \{\mathcal{O}\}$ where $\mathcal{O}$ is the point at infinity. There is a well-known group structure on E where two distinct points are added by taking the third intersection point of the line through the points and the elliptic curve and mirroring it on the x-axis. One can add a point to itself by taking the tangent line instead of the line through two distinct points. For $N \in \mathbb{N}$ we call $E(L)[N] := \{(x, y) \in E(L) | N \cdot (x, y) = \mathcal{O}\}$ the N-torsion points and $E_{\text{tor}} := \bigcup_{N \in \mathbb{N}} E(\overline{\mathbb{Q}})[N]$ the torsion points.

Definition 37 (Conductor, j-invariant)**.** We call $j_E := 1728 \frac{4A^3}{4A^3+27B^2}$ the j-invariant of E. For the precise definition of the conductor N of an elliptic curve E, we refer to §10 in [Sil94]. For us, the following facts will be sufficient:

- $\frac{\text{rad}(6N)}{6}$ is the product of all primes $p \geq 5$ such that the reduction of E mod p is a singular curve.
- For elliptic curves over $\mathbb{Q}$, the conductor is always at least 11 (see [Cre97], Table 1) .

Definition 38 (Complex multiplication)**.** Let E be an elliptic curve over $\mathbb{Q}$. We call $\text{End}(E)$ the set of $\overline{\mathbb{Q}}$-endmorphisms of E. Since we can add points to themselves, it will always contain $\mathbb{Z}$. In the case where $\text{End}(E)$ is strictly larger than $\mathbb{Z}$, we say that E *has complex multiplication*.

Remark 39. For an elliptic curve over $\mathbb{Q}$, the following is true. Whenever $\text{Aut}(E)$ is strictly larger than $\mathbb{Z}$, it will be of the form $\mathcal{O}_D = \mathbb{Z}[\frac{1}{2}(D + \sqrt{D})]$ for D congruent to 0 or 1 modulo 4. In that case we say that E has *complex multiplication by* $\mathcal{O}_D$.

Definition 40 (Supersingular prime)**.** Let E be an elliptic curve and p a prime. We call p supersingular for E if $E(\overline{\mathbb{F}}_p)[p] = \{\mathcal{O}\}$, which means that the point at infinity is the only p-torsion point in $E(\overline{\mathbb{F}}_p)$.

Elkies' result tells us that there are many of them.

Theorem 41 ([Elk87])**.** *Let E be an elliptic curve over $\mathbb{Q}$. Then there are infinitely many supersingular primes for E.*

Remark 42. In [FRM96] Fouvry and Ram Murty proved that the number of supersingular primes for an elliptic curve E that are smaller than a sufficiently large x is at least $c \log\log x$ for an absolute positive constant c but this result is not explicit.

Definition 43 (Surjective primes)**.** Let E be an elliptic curve defined over $\mathbb{Q}$ and let $p \in \mathbb{N}$ be a prime. We say that p is *surjective* (for E) if the Galois representation $\rho_p : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}\ E[p]$ is surjective.

Serre's Théoréme in [Ser72] states that there is a bound such that all primes greater than that bound are surjective if E does not have complex multiplication. Originally, his result contained no explicit bound. Explicit and effective estimates for this bound were developed later. One example is the result of Le Fourn which we will state in the chapter on small heights.

The proof of Elkies' Theorem gives an algorithm for finding supersingular primes. It requires finding primes in arithmetic progressions. Dirichlet's Theorem tells us that we can find such a prime and Linnik's Theorem tells us how big it is. Although many authors have improved the exponent in Linnik's Theorem not much has appeared in the literature on the multiplicative constant and only effective, but not explicit results are known there. Bennett, Martin, O'Bryant and Rechnitzer equip us with another result which is asymptotically weaker than Linnik's Theorem and its refinements but which is completely explicit.

We want to introduce two properties that we need later on.

Definition 44. Let $p \geq 5$ and let E be an elliptic curve over $\mathbb{Q}$. We say that p is suitable if p is a supersingular prime for E and p is surjective.

Exercises

Exercise 15 (Important)**.** Draw (or plot) some elliptic curves and add points on them using a ruler.

Exercise 16 (Important)**.** Which are the 2-torsion points of the elliptic curve?

4 Small Heights in Infinite Extensions

In this chapter we want to give an application of Bilu's Theorem in today's research. We fix for the rest of this chapter an elliptic curve E defined over $\mathbb{Q}$ with conductor N and j-invariant j_E.

Theorem 45 ([Hab13], Theorem 1)**.** *The infinite extension $\mathbb{Q}(\mathrm{E}_{tor})$ of $\mathbb{Q}$ has the Bogomolov property.*

Whenever E has complex multiplication, the statement is easy to prove.

Theorem 46. *Let E be an elliptic curve defined over $\mathbb{Q}$ with complex multiplication. Then for $\alpha \in \mathbb{Q}(E_{tor})^* \setminus \mu_\infty$ we have*

$$h(\alpha) \geq 3^{-14}.$$

Proof. If E has complex multiplication, then there exists a quadratic number field K such that $K(E_{\text{tor}})/K$ is abelian. We call this field the CM field of E. Theorem 1.2 of [AZ10] tells us that the height of $\alpha \in K(E_{\text{tor}})^* \setminus \mu_\infty$ is bounded from below by 3^{-14} which is always bigger than the bound in the theorem. □

So from now on let E be without complex multiplication. The following proposition is our starting point.

Proposition 47 ([Hab13], Proposition 6.1)**.** *Suppose E does not have complex multiplication. There exists a constant $c > 0$ depending only on E with the following property. If $\alpha \in \mathbb{Q}(E_{tor}) \setminus \mu_\infty$ is non-zero, there is a non-zero $\beta \in \overline{\mathbb{Q}} \setminus \mu_\infty$ with $h(\beta) \leq c^{-1}h(\alpha)$ and*

$$h(\alpha) + \max\{0, \frac{1}{[\mathbb{Q}(\beta) : \mathbb{Q}]} \sum_{\tau:\mathbb{Q}(\beta)\to\mathbb{C}} \log|\tau(\beta) - 1|\} \geq c.$$

We want to bound the sum in the above proposition from below. Our goal is to eventually show that this is negligible when compared to c. Originally, in Habegger's paper this was done by using Bilu's Theorem. We will cite Habegger's proof and then give an explicit version.

Proof. (of Theorem 45 in the non-CM case)

Our argument is by contradiction. We suppose that $\alpha_1, \alpha_2, \ldots$ is a sequence of non-zero elements of $\mathbb{Q}(\mathrm{E}_{\text{tor}}) \setminus \mu_\infty$ with $\lim_{k\to\infty} h(\alpha_k) = 0$.

For $m \in \mathbb{N}$ we define a continuous and bounded function $f_m : \mathbb{C} \setminus \{0\} \to \mathbb{R}$ by setting

$$f_m(z) = \min(m, \max(-m, \log|z - 1|))$$

for $z \neq 1$ and $f_m(1) = -m$.

The sequence of functions $s \mapsto f_m(e^{2\pi is})$ converges pointwise to $s \mapsto \log|e^{2\pi is} - 1|$ on $(0, 1)$ as $m \to \infty$. Clearly, $|f_m(e^{2\pi is})| \leq |\log|e^{2\pi is} - 1||$ and $\int_0^1 |\log|e^{2\pi is} - 1||ds < \infty$. So the Dominant Convergence Theorem from analysis implies

$$\lim_{m\to\infty} \int_0^1 f_m(e^{2\pi is})ds = \int_0^1 \log|e^{2\pi is} - 1|ds.$$

The latter integral is the logarithmic Mahler measure of the polynomial $X - 1$. As such, it vanishes by Jensen's Formula. So we may fix once and for all an m with

$$\int_0^1 f(e^{2\pi is})ds < \frac{c}{2} \quad \text{and} \quad \log(1 + 2e^{-m}) \leq \frac{c}{2} \tag{4.1}$$

where c is the positive constant from Proposition 47 and $f = f_m$.

The proposition also gives us a non-zero $\beta_k \in \overline{\mathbb{Q}} \setminus \mu_\infty$ for each α_k which satisfies

$$h(\alpha_k) + \max\left(0, \frac{1}{[\mathbb{Q}(\beta_k) : \mathbb{Q}]} \sum_{\tau:\mathbb{Q}(\beta_k)\to\mathbb{C}} \log|\tau(\beta_k) - 1|\right) \geq c. \tag{4.2}$$

and

$$h(\beta_k) \leq \frac{h(\alpha_k)}{c}. \tag{4.3}$$

We proceed by bounding the sum in (4.2) from above. Let $\tau : \mathbb{Q}(\beta_k) \to \mathbb{C}$ be an embedding. We write $z = \tau(\beta_k) \in \mathbb{C} \setminus \{0, 1\}$ and split up into cases depending on the size of $|z - 1|$.

Suppose for the moment that $|z - 1| \geq e^m$. Then $|z| \geq e^m - 1 \geq \frac{e^m}{2}$ since $m \geq 1$. So $\frac{|z-1|}{|z|} \leq 1 + \frac{1}{|z|} \leq 1 + 2e^{-m}$. Applying the logarithm and using (4.1) gives

$$\log|z - 1| \leq \log(1 + 2e^{-m}) + \log|z| \leq \frac{c}{2} + \log|z| \leq \frac{c}{2} + \log\max(1, |z|).$$

Because $f(z) = m \geq 0$ we conclude

$$\log|\tau(\beta_k) - 1| \leq \frac{c}{2} + \log\max(1, |\tau(\beta_k)|) + f(\tau(\beta_k)). \tag{4.4}$$

The second case is $|z - 1| < e^m$. Then $\log|z - 1| \leq \max(-m, \log|z - 1|) = f(z)$. So (4.4) holds as well.

Taking the sum over all field embeddings $\tau : \mathbb{Q}(\beta_k) \to \mathbb{C}$, applying (4.4), and dividing by the degree yields

$$\frac{1}{[\mathbb{Q}(\tau_k) : \mathbb{Q}]} \sum_\tau \log|\tau(\beta_k) - 1| \leq \frac{c}{2} + h(\beta_k) + \frac{1}{[\mathbb{Q}(\tau_k) : \mathbb{Q}]} \sum_\tau f(\tau(\beta_k)).$$

Hence (4.2) implies

$$h(\alpha_k) + \max\left(0, \frac{c}{2} + h(\beta_k) + \frac{1}{[\mathbb{Q}(\tau_k) : \mathbb{Q}]} \sum_\tau f(\tau(\beta_k))\right) \geq c. \tag{4.5}$$

The sequence $h(\alpha_1), h(\alpha_2), \ldots$ tends to zero, hence so does $h(\beta_1), h(\beta_2), \ldots$ by (4.3). We will apply Bilu's Theorem to $\beta_1, \beta_2, \ldots$ and the function f. On letting $k \to \infty$ the sum (4.5) over the τ converges to the integral $\int_0^1 f(e^{2\pi is})ds < \frac{c}{2}$ and both terms involving the height vanish. This is a contradiction. □

Lemma 48. *Let $\beta \in \overline{\mathbb{Q}}^*$ of degree $d \geq 2$ and let $0 < \varepsilon \leq \frac{1}{2}$. Then*

$$\frac{1}{[\mathbb{Q}(\beta):\mathbb{Q}]} \sum_{\tau:\mathbb{Q}(\beta)\hookrightarrow\mathbb{C}} \log|\tau(\beta)-1| \leq 2(\varepsilon|\log\varepsilon|+|\log(1-\varepsilon)|)+\frac{2}{\varepsilon d}\log d+\left(1+\frac{1}{\varepsilon}\right)h(\beta),$$

where τ runs over all embeddings of $\mathbb{Q}(\beta)$ into $\mathbb{C}$.

Proof. Let $F(x) = a_d x^d + \cdots + a_0 = a_d \cdot (x-\beta_1) \cdot \cdots \cdot (x-\beta_d)$ be integral minimal polynomial of β. Since

$$0 \neq |F(1)| = |a_d| \cdot \prod_{i=1}^{d} |\beta_i - 1|$$

we get

$$\begin{aligned}
\frac{1}{d}\log|F(1)| &= \frac{\log|a_d|}{d} + \frac{1}{d}\sum_{i=1}^{d}\log|\beta_i-1| \\
&\geq \frac{1}{d}\sum_{i=1}^{d}\log|\beta_i-1| \\
&= \frac{1}{[\mathbb{Q}(\beta):\mathbb{Q}]}\sum_{\tau:\mathbb{Q}(\beta)\hookrightarrow\mathbb{C}}\log|\tau(\beta)-1|. \qquad (4.6)
\end{aligned}$$

So it is enough to bound $|F(1)|$ in order to prove the Lemma.

For any polynomial $G = g_n x^n + \cdots + g_0 \in \mathbb{Z}[x]$ we define its height as $H(G) := \max_i |g_i|$. Furthermore, let $G_k := \frac{1}{k!}\frac{d^k G}{dx^k} = \sum_{i=k}^{n}\binom{i}{k} g_i x^{i-k} \in \mathbb{Z}[x]$ and $D \geq d$. We will fix D later in terms of ε and d. By Mignotte's Theorem B in [Mig89] we can find a polynomial $A(x) = \sum_{i=0}^{D-d} a_i x^i \in \mathbb{Z}[x] \setminus \{0\}$ of degree at most $D-d$ such that

$$H(A \cdot F) \leq ((D+1)^{\frac{d}{2}} H(\beta)^{Dd})^{\frac{1}{D+1-d}}. \qquad (4.7)$$

Let $k \in \mathbb{N}_0$ be the multiplicity of the zero at 1 of A. Since the degree of A is at most $D-d$ we have $k \leq D-d$. Then $A_{k-i}(1) = 0$ for all positive $i \leq k$ and $A_k(1) \neq 0$. As $A_k(1) \in \mathbb{Z}$ we find $|A_k(1)| \geq 1$ and thus by the Leibniz formula we get

$$\begin{aligned}
|F(1)| &\leq |A_k(1)||F(1)| \\
&= |(A\cdot F)_k(1)| \\
&\leq (D-k+1)H((A\cdot F)_k) \\
&\leq (D-k+1)\binom{D}{k}H(A\cdot F). \qquad (4.8)
\end{aligned}$$

By putting inequalities (4.6), (4.7) and (4.8) together we get

$$\begin{aligned}\frac{1}{[\mathbb{Q}(\beta):\mathbb{Q}]}\sum_{\tau:\mathbb{Q}(\beta)\hookrightarrow\mathbb{C}}\log|\tau(\beta)-1| &\le \frac{1}{d}\log|F(1)|\\ &\le \frac{1}{d}\log\left((D-k+1)\binom{D}{k}H(A\cdot F)\right)\\ &\le \frac{1}{d}\log\left((D-k+1)\binom{D}{k}((D+1)^{\frac{d}{2}}H(\beta)^{Dd})^{\frac{1}{D+1-d}}\right)\\ &\le \frac{1}{d}\log\left(\binom{D}{k}(D+1)^{\frac{d}{2(D+1-d)}+1}H(\beta)^{\frac{Dd}{D+1-d}}\right).\end{aligned}$$

The right hand side equals

$$\frac{1}{d}\log\binom{D}{k}+\left(\frac{1}{2(D+1-d)}+\frac{1}{d}\right)\log(D+1)+\frac{D}{D+1-d}h(\beta).$$

Note that $\varepsilon d\le\varepsilon[(1+\varepsilon)d]$ and so with $D:=[(1+\varepsilon)d]$ we have

$$k\le D-d\le\varepsilon d\le\varepsilon[(1+\varepsilon)d].$$

So we can apply Lemma 16.19 of [FG06] with $q=\varepsilon>0$ and $n=D$. We get $\binom{[(1+\varepsilon)d]}{k}\le 2^{-(1+\varepsilon)d(\varepsilon\log\varepsilon+(1-\varepsilon)\log(1-\varepsilon))}$. Since $\varepsilon<1$ we can write $|\log\varepsilon|$ instead of $-\log\varepsilon$ and $|\log(1-\varepsilon)|$ instead of $-\log(1-\varepsilon)$. So we can bound the above expression by

$$((1+\varepsilon)\varepsilon|\log\varepsilon|+(1-\varepsilon^2)|\log(1-\varepsilon)|)\log 2+\frac{1+2\varepsilon}{2\varepsilon d}\log((1+\varepsilon)d+1)+(1+\frac{1}{\varepsilon})h(\beta).$$

We start by bounding the first summand:

$$\begin{aligned}((1+\varepsilon)\varepsilon|\log\varepsilon|+(1-\varepsilon^2)|\log(1-\varepsilon)|)\log 2 &\le (\frac{3}{2}\varepsilon|\log\varepsilon|+|\log(1-\varepsilon)|)\log 2\\ &\le 2(\varepsilon|\log\varepsilon|+|\log(1-\varepsilon)|).\end{aligned}$$

The second summand can also be bounded further:

$$\begin{aligned}\frac{1+2\varepsilon}{2\varepsilon d}\log((1+\varepsilon)d+1) &\le \frac{2}{2\varepsilon d}\log(d^2)\\ &= \frac{2}{\varepsilon d}\log d.\end{aligned}$$

We put both bounds together and get

$$2(\varepsilon|\log\varepsilon|+|\log(1-\varepsilon)|)+\frac{2}{\varepsilon d}\log d+(1+\frac{1}{\varepsilon})h(\beta) \tag{4.9}$$

as an upper bound for $\frac{1}{[\mathbb{Q}(\beta):\mathbb{Q}]}\sum_{\tau:\mathbb{Q}(\beta)\hookrightarrow\mathbb{C}}\log|\tau(\beta)-1|$. □

Later, we will fix an ε and then get an explicit bound. But first, we want to look at the terms separately.

Lemma 49. *Let* $0 < x \leq \frac{1}{2}$. *Then*

$$-2(x\log x + \log(1-x)) \leq -(2 + \frac{4}{\log 2})x\log x.$$

Proof. Exercise. □

For our purpose the following corollary is sufficient.

Corollary 50. *Let* $0 < x \leq \frac{1}{2}$ *and* $0 < \gamma < 1$. *We have*

$$-2(x\log x + \log(1-x)) \leq 8\frac{1}{\gamma e}x^{1-\gamma}.$$

Proof. Exercise. □

We need a similar result for the second summand.

Lemma 51. *Let* $0 < \eta < 1$ *and* $d \geq 16$. *Then for every* $x > \frac{1}{4d}\left(\frac{\log\log d}{\log d}\right)^3$ *we have*

$$\frac{\log d}{d} \leq \frac{19}{\eta^4}x^{1-\eta}.$$

Remark 52. The constraint $d \geq 16$ guarantees that $\frac{\log\log d}{\log d}$ is a decreasing function.

Proof. Exercise. □

In the next lemma we combine all of the previous results of this section.

Lemma 53. *Let* $\delta < \frac{1}{2}$ *and let* $\beta \in \overline{\mathbb{Q}^*} \setminus \mu_\infty$ *be such that* $[\mathbb{Q}(\beta) : \mathbb{Q}] \geq 16$ *and* $h(\beta)^{\frac{1}{2}} \leq \frac{1}{2}$. *Then we have*

$$\frac{1}{[\mathbb{Q}(\beta):\mathbb{Q}]}\sum_{\tau:\mathbb{Q}(\beta)\hookrightarrow\mathbb{C}}\log|\tau(\beta)-1| \leq \frac{640}{\delta^4}h(\beta)^{\frac{1}{2}-\delta}. \tag{4.10}$$

Proof. Set $\varepsilon = h(\beta)^{\frac{1}{2}}$. Then Lemma 48 gives

$$\begin{aligned}
&\frac{1}{[\mathbb{Q}(\beta):\mathbb{Q}]}\sum_{\tau:\mathbb{Q}(\beta)\hookrightarrow\mathbb{C}}\log|\tau(\beta)-1| \\
&\leq -2(\varepsilon\log\varepsilon + \log(1-\varepsilon)) + \frac{2}{\varepsilon d}\log d + (1+\frac{1}{\varepsilon})h(\beta) \\
&\leq -2(h(\beta)^{\frac{1}{2}}\log h(\beta)^{\frac{1}{2}} + \log(1-h(\beta)^{\frac{1}{2}})) + \frac{2\log d}{h(\beta)^{\frac{1}{2}}d} + h(\beta)^{\frac{1}{2}} + h(\beta)^{\frac{1}{2}}.
\end{aligned}$$

Now since $h(\beta)^{\frac{1}{2}} \leq \frac{1}{2}$, we can apply Corollary 50 to the first term. By the main theorem of [Vou96] we also have $h(\beta) > \frac{1}{4d}\left(\frac{\log\log d}{\log d}\right)^3$ and so we can apply Lemma 51 to the second term and for any $0 < \gamma, \eta < 1$ we get:

$$\frac{1}{[\mathbb{Q}(\beta):\mathbb{Q}]}\sum_{\tau:\mathbb{Q}(\beta)\hookrightarrow\mathbb{C}}\log|\tau(\beta)-1| \leq \frac{8}{\gamma e}h(\beta)^{\frac{1}{2}(1-\gamma)} + \frac{38}{\eta^4}h(\beta)^{\frac{1}{2}(1-\eta=} + 2h(\beta)^{\frac{1}{2}}.$$

Now we set $\gamma := 2\delta$ and $\eta := 2\delta$ and get

$$\begin{aligned}\frac{8}{\gamma e}h(\beta)^{\frac{1}{2}(1-\gamma)} + \frac{38}{\eta^4}h(\beta)^{\frac{1}{2}-\eta} + 2h(\beta)^{\frac{1}{2}} &\leq \frac{1}{\delta^4}h(\beta)^{\frac{1}{2}-\delta}(\frac{8}{2e}\delta^3 + 38 + 2\delta^4) \\ &\leq \frac{640}{\delta^4}h(\beta)^{\frac{1}{2}-\delta},\end{aligned}$$

which is what we wanted to show. □

Theorem 54 ([LF16], Theorem 4.2)**.** *Let E be an elliptic curve over $\mathbb{Q}$ without complex multiplication and let j_E be the j-invariant of E. Then for*

$$p > 10^7 \max\{985, \frac{1}{12}h(j_E) + 3\}^2$$

the Galois representation ρ_p is surjective.

Theorem 55 ([Fre17], Theorem 2.7,)**.** *Let E be an elliptic curve with j-invariant j_E and conductor N. Let*

$$B_E = \begin{cases} \left(\frac{\log j_E}{2\pi}\right)^2 & \text{if } j_E > 0, \\ \left(\frac{\log |j_E|}{\pi} + 1\right)^2 & \text{if } j_E < 0, \\ 0 & \text{if } j_E = 0, \end{cases}$$

$M \in \mathbb{N}$ and $n = \max(11, M, B_E)$. Then there exists a supersingular prime p of E such that $p \geq n$ and

$$\log p \leq 4 \cdot 10^3 e^{\frac{1}{300}Ne^{\vartheta(n)}}(Ne^{\vartheta(n)})^2 h^*(j_E).$$

With the two theorems above we find an explicit upper bound - depending only on invariants of the elliptic curve - for a suitable prime.

We gathered all the results we need and are now able to prove the main theorem.

Theorem 56. *Let E be an elliptic curve defined over $\mathbb{Q}$ without complex multiplication and let $p \geq 5$ be a supersingular prime of E such that the Galois representation $Gal(\bar{\mathbb{Q}}/\mathbb{Q}) \to Aut\ E[p]$ is surjective. Then for $\alpha \in \mathbb{Q}(E_{tor})^* \setminus \mu_\infty$ we have*

$$h(\alpha) \geq \frac{(\log p)^5}{10^{21}p^{44}}.$$

Proof. Without loss of generality, assume that $h(\alpha) \leq \frac{1}{40p^4}$.

Proposition 47 gives us $\beta \in \overline{\mathbb{Q}}^* \setminus \mu_\infty$ with $h(\beta) \leq 10p^4 h(\alpha)$ and

$$h(\alpha) \geq \frac{1}{5}\left(\frac{\log p}{2p^8} - \max\Big\{0, \frac{1}{[\mathbb{Q}(\beta):\mathbb{Q}]}\sum_{\tau:\mathbb{Q}(\beta)\hookrightarrow\mathbb{C}} \log|\tau(\beta) - 1|\Big\}\right).$$

We want to distinguish two cases:

Case 1: $\deg \beta \geq 16$.

Here we can use Lemma 53 with $\delta = \frac{3}{10}$ and together with $h(\beta) \leq 10p^4 h(\alpha) \leq \frac{1}{4}$ we get

$$\begin{aligned} h(\alpha) &\geq \frac{1}{5}\left(\frac{\log p}{2p^8} - \frac{40}{(\frac{3}{10})^4}(10p^4 h(\alpha))^{\frac{1}{5}}\right) \\ &\geq \frac{1}{5}\left(\frac{\log p}{2p^8} - 4.94 \cdot 10^3 (10p^4 h(\alpha))^{\frac{1}{5}}\right). \end{aligned}$$

Since $h(\alpha) \leq 1$ we can make use of the fact that $h(\alpha)^{\frac{1}{5}} \geq h(\alpha)$. Then we find that

$$h(\alpha)^{\frac{1}{5}} + \frac{4.94}{5} 10^3 (10p^4 h(\alpha))^{\frac{1}{5}} \geq \frac{\log p}{10p^8},$$

which gives us

$$h(\alpha) \geq \left(\frac{1}{1 + \frac{4.94}{5} 10^3 (10p^4)^{\frac{1}{5}}} \frac{\log p}{10p^8}\right)^5.$$

We can simplify this and get

$$\begin{aligned} h(\alpha) &\geq \left(\frac{1}{1 + \frac{4.94}{5} 10^3 (10p^4)^{\frac{1}{5}}} \frac{\log p}{10p^8}\right)^5 \\ &\geq \left(\frac{1}{10^3 (10p^4)^{\frac{1}{5}}} \frac{\log p}{10p^8}\right)^5 \\ &\geq \frac{(\log p)^5}{10^{21} p^{44}}. \end{aligned}$$

Case 2: $d \leq 15$.

In this case we easily get an estimate with Corollary 2 of [Vou96]:

$$\begin{aligned} h(\beta) &> \frac{2d}{(\log(3d))^3} \\ &\geq \frac{1}{(\log(45))^3} \\ &\geq 0.02. \end{aligned}$$

This is always bigger than our bound from above so we proved the theorem. □

Exercises

Exercise 17. Prove Lemma 49.

Exercise 18. Prove Lemma 50.

Exercise 19. Prove Lemma 51.

Exercise 20 (Important)**.** Find (all) mistakes in these lecture notes.

5 Mordell-Weil

See [Sil09], chapter VIII.

Bibliography

[AZ10] F. Amoroso and U. Zannier. A uniform relative Dobrowolski's lower bound over abelian extensions. *Bull. Lond. Math. Soc.*, 42(3):489–498, 2010.

[Cre97] J. E. Cremona. *Algorithms for modular elliptic curves*. Cambridge University Press, Cambridge, second edition, 1997.

[Elk87] N. D. Elkies. The existence of infinitely many supersingular primes for every elliptic curve over $\mathbf{Q}$. *Invent. Math.*, 89(3):561–567, 1987.

[FG06] J. Flum and M. Grohe. *Parameterized complexity theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer-Verlag, Berlin, 2006.

[Fre17] L. Frey. Explicit Small Heights in Infinite Non-Abelian Extensions. *ArXiv e-prints*, December 2017.

[FRM96] E. Fouvry and M. Ram Murty. On the distribution of supersingular primes. *Canad. J. Math.*, 48(1):81–104, 1996.

[Hab13] P. Habegger. Small height and infinite nonabelian extensions. *Duke Math. J.*, 162(11):2027–2076, 2013.

[LF16] S. Le Fourn. Surjectivity of Galois representations associated with quadratic Q-curves. *Mathematische Annalen*, 365(1):173–214, 2016.

[Mig89] M. Mignotte. Sur un théorème de M. Langevin. *Acta Arith.*, 54(1):81–86, 1989.

[Ser72] J.-P. Serre. Propriétés galoisiennes des points d'ordre fini des courbes elliptiques. *Invent. Math.*, 15(4):259–331, 1972.

[Sil94] J. H. Silverman. *Advanced topics in the arithmetic of elliptic curves*, volume 151 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1994.

[Sil09] J. H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition, 2009.

[Vou96] P. Voutier. An effective lower bound for the height of algebraic numbers. *Acta Arith.*, 74(1):81–95, 1996.