Universiteit Utrecht

Bachelor Thesis

Mahler's Measure and Möbius Transformations

Jan-Willem van Ittersum

Supervisor:
Prof. dr. Gunther Cornelissen

July 27, 2015

Contents

Chapter 1

Introduction

In 1933, D. H. Lehmer wrote a paper [12] in which he describes a method of manufacturing large primes ('large' in a time when computers were not as well developed as now). Starting with a monic polynomial f with roots $\xi_i \in \mathbb{C}$ such that

$$f(z) = \prod_{i=1}^{n} (z - \xi_i),$$

he defines for $k \in \mathbb{N}$

$$\Delta_k(f) = \prod_{i=1}^{n} (\xi_i^k - 1).$$

For example, if we take $f(x) = x - 2$ we find the Mersenne numbers

$$\Delta_k(f) = 2^k - 1.$$

Since 1992, all largest known prime numbers are Mersenne numbers. At the time of writing

$$2^{57885161} - 1$$

is the largest known prime.

On the other hand, Lehmer used $f(x) = x^3 - x - 1$ and showed that

$$\Delta_{113}(f) = 63,088,004,325,217 \quad \text{and} \quad \Delta_{127}(f) = 3,233,514,251,032,733$$

are primes. Lehmer found that $\Delta_k(f)$ is more likely to be a prime number if the ratio of successive terms

$$\left| \frac{\Delta_{k+1}(f)}{\Delta_k(f)} \right|$$

is small. It can easily be shown that if all roots ξ_i of f satisfy $|\xi_i| \neq 1$ it holds that

$$\lim_{k \to \infty} \left| \frac{\Delta_{k+1}(f)}{\Delta_k(f)} \right| = \prod_{i=1}^{n} \max(1, |\xi_i|).$$

The right hand side of this equation is the quantity we are interested in in this thesis. It appears in different branches of mathematics, for example in the so-called Weil height in algebraic

number theory and in the topological entropy in algebraic dynamical systems. Moreover, it is generalised to polynomials in multiple variables and elliptic curves.

In this thesis, all these appearances of what we will later call the Mahler measure will not come up. We will study this quantity itself by investigating the relation between symmetries in polynomials and a lower bound of the Mahler measure. The starting point for this is a paper by Zagier which gives an elementary proof for an inequality by Zhang [19]. Dresden extended these ideas by proving another inequality involving the Mahler measure and some group. He ends his paper by asking whether his results can be generalised to other groups. Such a generalization is what I did and the results of this generalization will be covered in Chapter 6.

Contents In the second chapter we will introduce Mahler's measure. Together with the third chapter this is the required knowledge to understand the fifth chapter about the two papers of Zagier and Dresden. In Chapter 4, we will introduce Möbius transformations (the symmetries of polynomials) and search for finite groups of Möbius transformations. In some chapters, there is an intermezzo where basic mathematical knowledge which is not covered during the bachelor mathematics in Utrecht, is explained. The last chapter consists of my own generalizations of these two papers.

Acknowledgement I would like to thank Thijs van der Gugten, Lois van der Meijden, Merlijn Staps and Rik Voorhaar for reading my thesis and giving me a lot of useful comments. Also, for providing me with the subject of this thesis and giving me helpful suggestions how to proceed during our biweekly meetings, I thank my supervisor prof. dr. Gunther Cornelissen.

Chapter 2

Mahler's Measure

2.1 Mahler's Papers

In 1960 and 1962 Kurt Mahler published two papers [13] [14] in which he described a way to assign a real number to a polynomial. For a given polynomial he called this real number the *measure* of that polynomial. Other known ways to assign a real number to a polynomial are called the *height* and *length*. He defined this measure to give a new proof of the so-called Gelfond inequality [7] which establishes a lower bound for the height of a product of polynomials in terms of the heights of the factors. This inequality is frequently used in the theory of transcendental numbers. He also compared this measure with the height and length, as we will see in Section 2.4.

Intuitively we can interpret this measure, height and length as different ways to attach a 'size' to a polynomial. Following the ideas of his two papers we will introduce this so-called Mahler measure in this section. Although Mahler defined his measure for polynomials in many variables, for the sake of simplicity we will only study measures of polynomials in one variable.

Definition 2.1.1. Let $f(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_0 = \sum_{i=0}^{n} a_i z^i$ be a non-zero polynomial in $\mathbb{C}[z]$. Define the *Mahler measure* of f to be

$$M(f) = \exp \int_0^1 \log \left| f(e^{2\pi i \theta}) \right| \, \mathrm{d}\theta.$$

Also, define

$$m(f) = \log M(f) = \int_0^1 \log \left| f(e^{2\pi i \theta}) \right| \, \mathrm{d}\theta$$

as the *logarithmic Mahler measure*.

Remark. The Mahler measure should not be confused with the measure in the sense of set-theory. However, the logarithmic Mahler measure does satisfy comparable properties to a measure in measure theory. Recall (see, for example, page 22 of [17]) that for a σ-algebra $\mathcal{A}$ a measure μ is a map $\mu : \mathcal{A} \to \mathbb{R}$ satisfying

- $\mu(X) \geq 0$ for all $X \in \mathcal{A}$;
- $\mu(\emptyset) = 0$;

- $\mu\left(\biguplus_{j\in\mathbb{N}} A_j\right) = \sum_{j\in\mathbb{N}} \mu(A_j)$ for any countable family of pairwise disjoint sets $(A_j)_{j\in\mathbb{N}} \subset \mathcal{A}$.

In comparison, for the logarithmic Mahler measure of a non-zero polynomial f in $\mathbb{C}[z]$ we have

- $m(f) \geq 0$ if for the leading coefficient a_n we have that $|a_n| \geq 1$, which is the case when we for example have $f \in \mathbb{Z}[x]$;
- $m(1) = 0$;
- $m\left(\prod_{j=1}^{N} f_j\right) = \sum_{j=1}^{N} m(f_j)$ for polynomials $f_1, \ldots, f_N$ in $\mathbb{C}[z]$.

These three properties are a consequence of Theorem 2.1.3 below, which we will deduce later in this chapter.

Example 2.1.2. The Mahler measure of a constant polynomial $f(z) = a_0 \neq 0$ is given by

$$M(f) = \exp \int_0^1 \log |a_0| \, \mathrm{d}\theta = |a_0| \quad \text{and} \quad m(f) = \log M(f) = \log |a_0|.$$

Before we calculate the measure of non-constant polynomials, it is worth observing that if there exists a z_0 in $\mathbb{C}$ such that $f(z_0) = 0$, then $\log |f(z)|$ has a singularity at z_0. Therefore, using a theorem similar to Cauchy's residue theorem we will find that this measure depends on the zeros of f. In fact, Mahler observed that the measure equals the leading coefficient times the zeros of f *outside* the unit circle.

Theorem 2.1.3 (Mahler)**.** *For any $f \in \mathbb{C}[z]$ with leading coefficient $a_n \neq 0$ and zeros $\xi_1, \ldots, \xi_n$, we have*

$$M(f) = |a_n| \cdot \prod_{i=1}^{n} \max(1, |\xi_i|).$$

Note that an empty product is considered to be 1, so for a non-zero constant polynomial $f(z) = a_0$ the right-hand side equals $|a_0|$ in accordance with Example 2.1.2. This theorem is a consequence of Jensen's formula [8]. Therefore, before proving Mahler's theorem we will state and prove Jensen's formula.

2.2 Intermezzo: Jensen's Formula

Let f be a meromorphic complex function, that is f is analytic on $\mathbb{C}$ except at a discrete set of points S which are poles. Jensen's formula relates $\int_0^1 \log |f(e^{2\pi i\theta})| \, \mathrm{d}\theta$ to the zeros and poles of f. We will follow chapter XII, section 1 of Lang's book on complex analysis [11] combined with lemma 1.9 of [6] where part of the proof is formulated more elegantly. Recall that $\operatorname{ord}_a f$, the order of a function f in a, equals l if a is a zero with multiplicity l and equals $-m$ if f has a pole of order m in a. Otherwise, $\operatorname{ord}_a f$ is zero.

Lemma 2.2.1 (Jensen's formula)**.** *Let f be a meromorphic function which is not constant on the closed disc $\overline{D}_R$ of radius R and with power series expansion at 0 written as:*

$$f(z) = c_m z^m + c_{m+1} z^{m+1} + \ldots.$$

for some $m \in \mathbb{Z}$. Then

$$\int_0^1 \log|f(Re^{2\pi i\theta})|\,\mathrm{d}\theta + \sum_{\substack{a\in D_R\\ a\neq 0}} \mathrm{ord}_a(f)\log\frac{|a|}{R} + \mathrm{ord}_0(f)\log\frac{1}{R} = \log|c_m|.$$

Proof. We start by proving Jensen's formula in a few specific cases. Assume first of all that f has no zeros or poles in $\overline{D}_R$. Because in particular $f(0) \neq 0$ we have that $c_m = c_0 = f(0)$. Moreover, $\log f(z)$ is analytic on this disc (see page 123 of [11]), so

$$\log c_m = \log f(0) = \frac{1}{2\pi i}\int_{\partial D_R} \frac{\log f(z)}{z}\,\mathrm{d}z = \int_0^1 \log(f(Re^{2\pi i\theta}))\,\mathrm{d}\theta$$

by Cauchy's formula. This case is proven by taking the real part of the equality.

Secondly, suppose $f(z) = z$. Then $\log|f(Re^{2\pi i\theta})| = \log(R)$, $\mathrm{ord}_0 f = 1$ and $\mathrm{ord}_a f = 0$ for $a \neq 0$. Hence, the left-hand side of Jensen's formula equals $\log(R) + \log(\frac{1}{R}) = 0$ and the right-hand side equals $\log|c_m| = \log|c_1| = 0$, which proves the formula in this case.

Next, let $f(z) = z - \xi$ for $\xi \in \overline{D}_R\backslash\{0\}$. Let $\beta = \frac{\xi}{R}$, then $|\beta| = |\frac{\xi}{R}| \leq 1$. Because $\log(1-\beta z)$ is analytic with radius of convergence $\frac{1}{|\beta|} \geq 1$, by Cauchy's theorem it follows that if $|\beta| < 1$ then

$$\int_0^1 \log(1 - e^{2\pi i\theta}\beta)\,\mathrm{d}\theta = \int_{|z|=1} \frac{\log(1-\beta z)}{z}\,\mathrm{d}z = 0.$$

Taking the real part and substituting $\theta \to -\theta$ we can rewrite this as

$$\int_0^1 \log|1 - e^{-2\pi i\theta}\beta|\,\mathrm{d}\theta = 0.$$

Multiplying with $\log|e^{2\pi i\theta}| = 1$ yields

$$\int_0^1 \log|e^{2\pi i\theta} - \beta|\,\mathrm{d}\theta = 0.$$

Adding $\log R$ we get that

$$\int_0^1 \log|Re^{2\pi i\theta} - \xi|\,\mathrm{d}\theta = \log R.$$

Because $f(z) = z - \xi$ implies that $\mathrm{ord}_\xi f = 1$, $\mathrm{ord}_a f = 0$ for all other $a \neq \xi$ and $c_f = -\xi$, we find that Jensen's formula holds in this case, namely

$$\int_0^1 \log|Re^{2\pi i\theta} - \xi|\,\mathrm{d}\theta + \log\frac{|\xi|}{R} = \log|-\xi|.$$

Next, assume $|\beta| = 1$. As in the case when $|\beta| < 1$, we want to calculate $\int_0^1 \log|1 - e^{2\pi i\theta}\beta|\,\mathrm{d}\theta$. This integral now becomes singular, so define

$$\int_0^1 \log|1 - \beta e^{2\pi i\theta}|\,\mathrm{d}\theta = \lim_{\varepsilon\to 0}\frac{1}{2\pi i}\int_{\Gamma(\beta,\varepsilon)} \frac{\log|1-\beta z|}{z}\,\mathrm{d}z$$

if this limit exists. In this equation $\Gamma(\beta,\varepsilon)$ is the contour indicated in Figure 2.1.

Now $\frac{\log|1-z\beta|}{z}$ is analytic on the closed unit disk except for $z = \beta^{-1}$. So, if $\gamma(\beta,\varepsilon)$ is the circle of radius ε around β^{-1}, it follows from Cauchy's theorem that

$$\frac{1}{2\pi i}\int_{\Gamma(\beta,\varepsilon)} \frac{\log|1-\beta z|}{z}\,\mathrm{d}z = \frac{1}{2\pi i}\int_{\gamma(\beta,\varepsilon)} \frac{\log|1-\beta z|}{z}\,\mathrm{d}z.$$

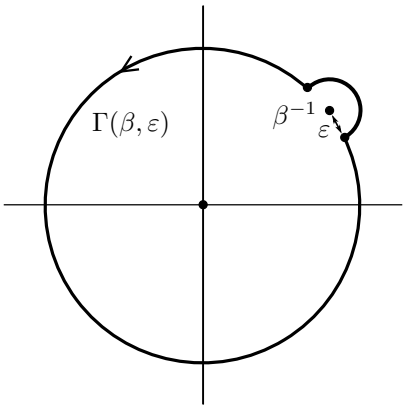

Figure 2.1: The contour $\Gamma(\beta,\varepsilon)$.

Parametrize $\gamma(\beta,\varepsilon)$ by $z=\beta^{-1}+\varepsilon e^{2\pi i\theta}$ for $\theta\in[0,1)$. Then

$$\frac{1}{2\pi i}\int_{\gamma(\beta,\varepsilon)}\frac{\log|1-\beta z|}{z}\,\mathrm{d}z=\int_0^1\frac{\log(\varepsilon e^{2\pi i\theta})}{\beta^{-1}+\varepsilon e^{2\pi i\theta}}\cdot\varepsilon e^{2\pi i\theta}\,\mathrm{d}\theta.$$

Because

$$\left|\frac{\varepsilon e^{2\pi i\theta}}{\beta^{-1}+\varepsilon e^{2\pi i\theta}}\right|\leq\frac{\varepsilon}{\beta^{-1}-\varepsilon}$$

is bounded for ε small enough, the absolute value of the integral is bounded by $C\varepsilon|\log(\varepsilon)|$ for some constant C. Hence, for $\varepsilon\to 0$ the integral vanishes. Similarly to the case where $|\beta|<1$, we can rewrite $\int_0^1\log|1-\beta e^{2\pi i\theta}|\,\mathrm{d}\theta=0$ to get Jensen's formula. So, also for $f(z)=z-\xi$ Jensen's equation is valid.

To deduce the general case of Jensen's formula, let

$$h(z)=f(z)\prod_{a\in\overline{D}_R}(z-a)^{-\operatorname{ord}_a f}$$

Because we multiply f by a function which by definition cancels all poles and zeros in $\overline{D}_R$, we conclude h has no poles or zeros in $\overline{D}_R$. Furthermore we can rewrite f as

$$f(z)=h(z)\prod_{a\in\overline{D}_R}(z-a)^{\operatorname{ord}_a f}.$$

Notice that because h is analytic on $\overline{D}_R$, Jensen's formula holds for h. In addition, we can go from h to f just by multiplication and division of factors for which we have proved Jensen's formula. Therefore, it is sufficient to prove that if Jensen's formula is valid for functions ϕ and ψ, then it is also valid for $\phi\psi$ and ϕ^{-1}. Let c_m, c_n and c_{mn} the leading coefficient of ϕ, ψ, respectively $\phi\psi$ and assume ϕ and ψ satisfy Jensen's formula. Because we have that

$$\begin{aligned}&\log|\phi\psi|=\log|\phi|+\log|\psi|,\\&\operatorname{ord}_a(\phi\psi)=\operatorname{ord}_a(\phi)+\operatorname{ord}_a(\psi),\\&c_{mn}=c_mc_n\text{ so that }\log|c_{mn}|=\log|c_m|+\log|c_n|,\end{aligned}$$

the product $\phi\psi$ satisfies Jensen's formula as well.

Now, let $c_{m'}$ and c_m be the leading coefficient of ϕ^{-1} respectively ϕ, and assume ϕ satisfies Jensen's formula. Then

$$\begin{aligned}
&\log|\phi^{-1}| = -\log|\phi|,\\
&\operatorname{ord}_a(\phi^{-1}) = -\operatorname{ord}_a(\phi),\\
&c_{m'} = c_m^{-1} \text{ so that } \log|c_{m'}| = -\log|c_m|.
\end{aligned}$$

So ϕ^{-1} satisfies Jensen's formula as well. Hence, by multiplying and dividing h successively by factors of the form $z-\xi$ we reach f. We conclude that Jensen's formula is valid for every meromorphic function. □

2.3 Mahler's Theorem

Using Jensen's formula we can now mimic Mahler's proof of Theorem 2.1.3.

Proof of Theorem 2.1.3. We can assume without loss of generality that the zeros of f have been numbered so that

$$\begin{aligned}
&0 = \xi_1 = \xi_2 = \ldots = \xi_M,\\
&0 < |\xi_{M+1}| \le \ldots \le |\xi_N| \le 1 < |\xi_{N+1}| \le \ldots \le |\xi_n|.
\end{aligned}$$

By putting $R = 1$ in Jensen's formula we find that

$$\begin{aligned}
M(f) &= \exp\int_0^1 \log|f(e^{2\pi i\theta})|\,\mathrm{d}\theta\\
&= \exp\left(\sum_{j=M+1}^{N}\log\frac{1}{|\xi_j|} + \log|c_m|\right)\\
&= \left|\frac{c_m}{\xi_{M+1}\xi_{M+2}\cdots\xi_N}\right|.
\end{aligned}$$

Expanding the product in $f(z) = a_n\prod_{j=1}^{M}(z-\xi_j)$ yields $c_m = \pm a_d\xi_{M+1}\xi_{M+2}\cdots\xi_n$. So this formula may also be written as

$$\begin{aligned}
M(f) &= |a_n\xi_{N+1}\xi_{N+2}\cdots\xi_n|\\
&= |a_n|\prod_{i=1}^{n}\max(1,|\xi_i|)
\end{aligned}$$

proving Theorem 2.1.3. □

Remark. As a consequence of this theorem we find that $M(f) \ge |a_n|$, so $M(f) \ge 1$ if f is monic. We also find that $M(1) = 1$ and $M(fg) = M(f)M(g)$. This shows that the properties of the logarithmic Mahler measure are comparable to a measure in measure theory as mentioned in Section 2.1.

Using this identity for the Mahler measure we have reduced the problem of calculating an integral to the problem of finding the roots of a polynomial. So, if all roots of a polynomial are given, we are able to calculate the Mahler measure. As a consequence, it is always possible to calculate the Mahler measure for a given polynomial of degree at most four. This can be done by using the formulae for the roots of linear, quadratic, cubic and quartic polynomials.

2.4 Inequalities Involving Other Measures

The Mahler measure is not the only useful real-valued function one can define on polynomials. In fact, there are more intuitive functions called the height and length which are defined by the coefficients of a polynomial. In his papers, Mahler compared the Mahler measure with these measures [13] [14]. We will reproduce some of his results here. Without loss of continuity this section can be skipped, as we will not use these results later.

Definition 2.4.1. For any polynomial $f(z) = \sum_{i=1}^{n} a_i z^i$ in $\mathbb{C}[z]$, define

$$H(f) = \max_{0 \le i \le n} (|a_i|), \qquad L(f) = \sum_{i=0}^{n} |a_i|$$

as the *height* respectively the *length* of f.

Example 2.4.2. Consider the polynomial $f(z) = (z-2)^n$. Using Theorem 2.1.3 we find that the Mahler measure equals $M(f) = 2^n$.
By Newton's binomial theorem

$$f(z) = \sum_{i=0}^{n} \binom{n}{i} z^n (-2)^{n-i}.$$

So

$$L(f) = \sum_{i=0}^{n} \binom{n}{i} 2^{n-i} = (1+2)^n = 3^n.$$

To calculate the height of f we have to find the coefficient that is maximal in absolute value. First we consider the difference between the absolute value of two successive coefficients:

$$\begin{aligned}
\binom{n}{i} 2^{n-i} - \binom{n}{i-1} 2^{n-i+1} &= 2^{n-i} \left(\frac{n!}{i!(n-i)!} - \frac{2 \cdot n!}{(i-1)!(n-i+1)!} \right) \\
&= 2^{n-i} \left(\frac{n! \cdot (n-i+1)}{i!(n-i+1)!} - \frac{2 \cdot n! \cdot i}{i!(n-i+1)!} \right) \\
&= \frac{2^{n-i} \cdot n!}{i!(n-i+1)!} (n - 3i + 1).
\end{aligned}$$

Note that this difference is positive if and only if $3i < n+1$. So the maximum of the coefficients is obtained for $i_{\max} = \left\lfloor \frac{n+1}{3} \right\rfloor$. Hence, the height of f is given by $\binom{n}{i_{\max}} 2^{n-i_{\max}}$.

In order to compare the Mahler measure with these other measures in general, we need the following lemma relating the coefficients of f to its Mahler measure.

Lemma 2.4.3. *For any non-zero polynomial $f(z) = \sum_{i=0}^{n} a_i z^i$ in $\mathbb{C}[z]$ we have*

$$|a_m| \le \binom{n}{m} M(f)$$

for all $m \in \{0, 1, \ldots, n\}$.

Proof. Number the zeros as in the proof of Theorem 2.1.3, such that $\xi_{N+1}, \xi_{N+2}, \ldots, \xi_n$ are the zeros outside the unit circle. Next let $I = \{i_1, i_2, \ldots, i_m\}$ be an arbitrary subset of $D =$

$\{1, 2, \ldots, n\}$, possibly empty or equal to D. From the numbering of the zeros it then follows that

$$|a_n \xi_{i_1} \xi_{i_2} \ldots \xi_{i_m}| \leq |a_n \xi_{N+1} \xi_{N+2} \cdots \xi_n|.$$

Note that each coefficient a_m of f equals, apart from a factor ± 1, the sum of $\binom{n}{m}$ terms of the form $a_n \xi_{i_1} \xi_{i_2} \ldots \xi_{i_m}$. This can be seen be expanding $f(z) = a_n \prod_{i=0}^{n} (z - \xi_i)$. Hence,

$$|a_m| \leq \binom{n}{m} |a_n \xi_{N+1} \xi_{N+2} \cdots \xi_n|.$$

Using Theorem 2.1.3 we find that $|a_m| \leq \binom{n}{m} M(f)$. □

Now we are ready to compare the Mahler measure with the height and length of a polynomial. It turns out that for two polynomials of the same degree these measures are comparable as follows.

Theorem 2.4.4. *Let $f \in \mathbb{C}[z]$ be a non-zero polynomial of degree n. Then*

$$2^{-n} L(f) \leq M(f) \leq L(f).$$

This inequality is best possible in the sense that equality on the left holds for example for $f(z) = (z-1)^n$ and equality on the right holds for example for $f(z) = z^n$.

Proof. Summing the equation in Lemma 2.4.3 over m from 0 to n, it follows that

$$L(f) \leq \sum_{m=0}^{n} \binom{n}{m} M(f) = 2^n M(f),$$

from which we deduce that $2^{-n} \leq M(f)$.
The other inequality with $L(f)$ follows because $|f(e^{2\pi i \theta})| \leq L(f)$, hence

$$M(f) = \exp \int_0^1 \log |f(e^{2\pi i \theta})| \, \mathrm{d}\theta \leq \exp \int_0^1 \log L(f) \, \mathrm{d}\theta = L(f).$$

For $f(z) = (z-1)^n$ we have that

$$L(f) = \sum_{i=0}^{n} \binom{n}{i} = 2^n.$$

Because $M(f) = 1$, we conclude that $2^{-n} L(f) = M(f)$.
On the other hand, for $f(z) = z^n$ we have that $L(f) = 1 = M(f)$, concluding the proof. □

Theorem 2.4.5. *Let f be a non-zero polynomial in $\mathbb{C}[z]$ of degree n. If $n \neq 0$, then*

$$2^{-n+1} H(f) \leq M(f) \leq \sqrt{n+1}\, H(f)$$

Proof. We will prove by induction to n that $\binom{n}{m} \leq 2^{n-1}$ for all $m \in \mathbb{N}$ with $m \leq n$. For $n = 1$ we have $\binom{n}{0} = \binom{n}{1} = 1 = 2^{1-1}$. Assuming we have proven the inequality for $n = k-1$, we have

$$\binom{k}{m} = \binom{k-1}{m} + \binom{k-1}{m-1} \leq 2^{k-2} + 2^{k-2} = 2^{k-1}.$$

Hence, Lemma 2.4.3 implies that $H(f) \leq 2^{n-1} M(f)$, so $2^{-n+1} H(f) \leq M(f)$.

To prove the last inequality we use Jensen's inequality below (see page 115 of [17] for a proof), which should not to be confused with Jensen's formula.

Lemma 2.4.6 (Jensen's inequality applied to the Lebesgue measure)**.** *Let $\Lambda : [0,\infty) \to [0,\infty)$ be a concave function. For any Lebesgue-integrable function $g : [a,b] \to [0,\infty)$ we have*

$$\frac{1}{b-a}\int_a^b \Lambda((b-a)g(x))\,\mathrm{d}x \leq \Lambda\left(\int_a^b g(x)\,\mathrm{d}x\right). \qquad \square$$

Let $g(x) = |f(e^{2\pi i x})|^2$ and note that $\Lambda(x) = \log\sqrt{x}$ is a concave function. Applying Jensen's inequality we obtain that

$$M(f) = \exp\int_0^1 \log|f(e^{2\pi i\theta})|\,\mathrm{d}\theta \leq \left(\int_0^1 |f(e^{2\pi i\theta})|^2\,\mathrm{d}\theta\right)^{\frac{1}{2}}.$$

Because the coefficients a_i are the Fourier coefficients of $f(e^{2\pi i\theta})$ we get by Parseval's formula (see vol I, page 37 of [21]) that

$$\int_0^1 |f(e^{2\pi i\theta})|^2\,\mathrm{d}\theta = \sum_{i=0}^{n} |a_i|^2 \leq (n+1)H(f)^2.$$

Therefore we conclude that $M(f) \leq \sqrt{n+1}\,H(f)$. $\square$

Remark. These inequalities for the height are not sharp. On the left we can only have equality if $n \leq 2$, as $\binom{3}{m} \neq 2^{3-1} = 4$ for all $m \in \mathbb{N}$ and hence by induction $\binom{n}{m} \neq 2^{n-1}$ for $n, m \in \mathbb{N}$ with $n \geq 3$. If n equals 1 or 2 equality holds for example for[1]

$$H(x-1) = 1 = M(x-1),$$
$$2^{-1}H((x-1)^2) = 1 = M((x-1)^2).$$

On the right equality never holds for $n > 0$. Namely, equality holds in Jensen's inequality if $\Lambda(x)$ is linear or g is constant. As $\log\sqrt{x}$ is not linear, equality on the right can only hold for constant polynomials.

[1] Mahler erroneously mentioned that equality can never hold on the left if the degree n exceeds 1.

Chapter 3

Integer Polynomials

3.1 Lehmer's Problem

In the previous chapter all polynomials had complex coefficients. From now on we will restrict ourselves to integer polynomials, i.e. polynomials with integer coefficients, as Lehmer did. About thirty years before Mahler, he already mentioned the Mahler measure in a paper discussing techniques for discovering large primes [12]. Given a small $\varepsilon > 0$, he wondered whether there exists a polynomial $f \in \mathbb{Z}[x]$ such that $1 < M(f) < 1 + \varepsilon$. Lehmer mentioned that the polynomial

$$x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1$$

had the smallest measure he could find for an integer polynomial. The logarithmic Mahler measure of this polynomial is approximately 0.162358. Because it is easier to talk about positive measure, from now on we will use the logarithmic Mahler measure.

Problem 3.1.1 (Lehmer)**.** *Does there exist a constant $D > 0$ such that for every non-zero integer polynomial f*

$$m(f) = 0 \quad \textit{or} \quad m(f) \geq D \quad ?$$

Remark. It suffices to solve this problem in the affirmative for irreducible f. Assume that a polynomial f with small non-zero logarithmic Mahler measure is reducible, that is $f = g \cdot h$. Because $m(f) = m(g) + m(h)$ and $m(\hat{f}) \geq 0$ for integer polynomials $\hat{f}$ as a consequence of Theorem 2.1.3, we have that

$$0 < \max(m(g), m(h)) \leq m(f).$$

Therefore we can assume without loss of generality that f is irreducible.

Notice that for $f(z) = z^n - 2$ we have $m(f) = \log(2)$, because f has as roots $\sqrt{2} \cdot \zeta_n^i$, where ζ_n is a primitive nth rooth of unity. So, in our search for polynomials with small logarithmic Mahler measure, we can assume that $m(f) < \log(2)$, that is

$$|a_n| \cdot \prod_{i=1}^{n} \max(1, |\xi_i|) < 2.$$

So, we can assume that the leading coefficient of such a polynomial is smaller than 2, hence we can assume it to be monic. Moreover, for a monic polynomial f we have that

$$\prod_{i=1}^{n} |\xi_i| \leq \prod_{i=1}^{n} \max(1, |\xi_i|) < 2.$$

As the lowest order non-zero coefficient of f equals the product of all non-zero roots of f, this coefficient is ± 1.

Example 3.1.2. In this example we will calculate the quadratic integer polynomial with the smallest Mahler measure greater than 1. From the remark above it follows that this polynomial is of the form $x^2 + ax + b$ for $a \in \mathbb{Z}$ and $b = \pm 1$. The roots of this polynomial are given by

$$x = \frac{-a \pm \sqrt{a^2 - 4b}}{2}.$$

If $a^2 - 4b < 0$ then the roots are complex conjugates of each other and therefore the complex norm of these roots is the same. Because the product of the roots is ± 1, we find that $M(f) = 1$ in this case.
Now, assume that $a^2 - 4b > 0$. Because the product of these roots is still ± 1 and $a \neq 0$, the norm of one of the roots is greater than or equal to 1, while the norm of the other root is smaller than or equal to 1 and hence does not contribute to the Mahler measure. So, if $a > 0$ we have that

$$M(f) = \left| \frac{-a - \sqrt{a^2 - 4b}}{2} \right|.$$

For $a = 1$, using $a^2 - 4b > 0$ and $b = \pm 1$, we find $b = -1$ and $M(f) = \frac{1+\sqrt{5}}{2}$. If $a = 2$, then we have again $b = -1$ and $M(f) = \frac{2+\sqrt{8}}{2}$. For $a > 2$ we have that $M(f) \geq \frac{3+\sqrt{5}}{2}$, where $b = 1$. For $a < 0$ we find that

$$M(f) = \left| \frac{-a + \sqrt{a^2 - 4b}}{2} \right|.$$

Similarly, we then find that $M(f) \geq (1 + \sqrt{5})/2$. So the smallest measure greater than 1 a quadratic polynomial can have is the golden ratio $(1 + \sqrt{5})/2$ for $f(z) = z^2 \pm z - 1$.

3.2 Intermezzo: Algebraic Integers

Before we move on to Lehmer's problem we will study the roots of integer polynomials. These roots are called algebraic numbers. Later we will require several results about algebraic numbers which we will prove in this section. We used the first chapter of the appendix of Everests and Wards book on heights of polynomials [6] to write this section.

Definition 3.2.1. A complex number α is an *algebraic number* if there is a non-zero polynomial $f \in \mathbb{Z}[x]$ for which $f(\alpha) = 0$. Moreover, α is an *algebraic integer* if there is a monic polynomial $f \in \mathbb{Z}[x]$ for which $f(\alpha) = 0$. Denote with A the set of algebraic integers.

Definition 3.2.2. Let K be a field extension of $\mathbb{Q}$ of finite degree. The set $K \cap A$ of *algebraic integers of* K is denoted by $\mathcal{O}_K$.

Example 3.2.3. If $\alpha = \frac{p}{q} \in \mathbb{Q}$ (assuming $q \geq 1$), then we can take $f(x) = qx - p$ as a polynomial such that $f(\frac{p}{q}) = 0$. Because every integer polynomial with $\frac{p}{q}$ as one of its roots contains a factor $qx - p$ we find that $\frac{p}{q}$ is an algebraic integer if and only if it is an integer. Otherwise $q > 1$ and the polynomial would not be monic. So, $\mathcal{O}_{\mathbb{Q}} = \mathbb{Z}$. Hence, the algebraic integers generalize the integers to finite degree field extensions of $\mathbb{Q}$.

Example 3.2.4. We saw that the golden ratio $\frac{1+\sqrt{5}}{2}$ has $x^2 - x - 1$ as its minimal polynomial, hence it is an algebraic integer. So, for $K = \mathbb{Q}(\sqrt{5})$, $\frac{1+\sqrt{5}}{2}$ is in $\mathcal{O}_K$ and in fact $\mathcal{O}_K = \mathbb{Z}[\frac{1+\sqrt{5}}{2}]$.

Definition 3.2.5. An additive abelian group G is *finitely generated* if there exist finitely many elements $x_1, \ldots, x_s \in G$ such that every $g \in G$ can be witten as $g = n_1x_1 + n_2x_2 + \ldots + n_sx_s$ with integers $n_1, \ldots, n_s$.

Lemma 3.2.6. *A number $\alpha \in \mathbb{C}$ is an algebraic integer if and only if $\mathbb{Z}[\alpha]$ is finitely generated.*

Proof. Let α be an algebraic integer with monic polynomial $f \in \mathbb{Z}[x]$ of degree d and with coefficients a_i such that $f(\alpha) = 0$. Let G be the additive group generated by $1, \alpha, \ldots, \alpha^{d-1}$. Because f is monic we have that

$$\alpha^d + a_{d-1}\alpha^{d-1} + \ldots + a_0 = 0$$

So, $\alpha^d \in G$. Moreover, assuming $\alpha^k \in G$ for some $k \in \mathbb{Z}$ with $k \geq d$ we have that

$$\alpha^{k+1} = -\alpha_{d-1}\alpha^k - \ldots - a_0\alpha^{k+1-d}$$

which belongs to G. So, by induction all powers of α lie in G. Hence, $G = \mathbb{Z}[\alpha]$, which is finitely generated.

Conversely, assume that $\mathbb{Z}[\alpha]$ is finitely generated, with generators $\alpha_1, \ldots, \alpha_m$. Each α_i belongs to $\mathbb{Z}[\alpha]$, so we can find polynomials $f_i \in \mathbb{Z}[x]$ such that $\alpha_i = f_i(\alpha)$. Now, let $N > \deg f_i$ for all $i = 1, 2, \ldots, m$. We have that

$$\alpha^N = \sum_{j=1}^{m} a_j\alpha_j = \sum_{j=1}^{m} a_j f_j(\alpha)$$

for $a_j \in \mathbb{Z}$. Take

$$f(x) = x^N - \sum_{j=1}^{m} a_j f_j(x).$$

Clearly $f \in \mathbb{Z}[x]$ and it is monic because $N > \deg f_i$ for all $i = 1, 2, \ldots, m$. Finally, $f(\alpha) = 0$, from which we conclude that α is an algebraic integer. □

Theorem 3.2.7. *The set A of algebraic integers forms a ring.*

Proof. Let α and β be algebraic integers. By the previous lemma $\mathbb{Z}[\alpha]$ and $\mathbb{Z}[\beta]$ are finitely generated, thus so is $\mathbb{Z}[\alpha, \beta]$. Since $\mathbb{Z}[\alpha \pm \beta]$ and $\mathbb{Z}[\alpha\beta]$ are subgroups of $\mathbb{Z}[\alpha, \beta]$, they are finitely generated as well. Again by the previous lemma, we find that $\alpha \pm \beta$ and $\alpha\beta$ are algebraic integers. □

Corollary 3.2.8. *$\mathcal{O}_K$ forms a ring.*

Proof. $\mathcal{O}_K$ is the intersection of the ring A and the field K. □

Lemma 3.2.9. *Let $f, g \in \mathbb{Z}[x]$ be two non-zero polynomials of degree m, respectively n. Assume f is monic and let $\xi_1, \ldots, \xi_m$ be the roots of f. Then*

$$p = \prod_{i=1}^{m} g(\xi_i)$$

is an integer.

Proof. Every element σ of the Galois group of f permutes the roots of f. Observe that p is fixed under such a permutation, that is

$$\prod_{i=1}^{m} g(\xi_i) = \prod_{i=1}^{m} g(\xi_{\sigma(i)}).$$

Hence, p lies in the fixed field of f, which is $\mathbb{Q}$.
Moreover, each ξ_i is an algebraic integer. Because algebraic integers form a ring, p is an algebraic integer. We conclude that $p \in \mathcal{O}_{\mathbb{Q}} = \mathbb{Z}$. □

Observe that in the proof of this lemma we used algebraic integers, because f is monic. However, later we want to use this lemma when f is not monic. In that case, the above theory of algebraic integers will not help us. However, with the use of (elementary) symmetric polynomials we can generalize Lemma 3.2.9.

Lemma 3.2.10. *Let $f, g \in \mathbb{Z}[x]$ be two non-zero polynomials of degree m, respectively n. Let a be the leading coefficient of f and let $\xi_1, \ldots, \xi_m$ be the roots of f. Then*

$$p = a^n \prod_{i=1}^{m} g(\xi_i)$$

is an integer.

Proof. Observe that p is a symmetric function of $\xi_1, \ldots, \xi_d$, that is, for every $\sigma \in S_m$ we have

$$a^n \prod_{i=1}^{m} g(\xi_i) = a^n \prod_{i=1}^{m} g(\xi_{\sigma(i)}).$$

Hence, an element of the Galois group of f fixes p, so $p \in \mathbb{Q}$.
Moreover, by the Fundamental Theorem of Symmetric Polynomials (see, for example, page 140 of [16]) p is expressible as a polynomial of elementary symmetric functions in the roots ξ_i. Therefore, let $p = a^n \cdot q(s_1, \ldots, s_m)$ with s_i the ith elementary symmetric function and $q \in \mathbb{Z}[x]$. As g has degree n, q has degree n as well. Note that the elementary functions are - up to a sign and a factor a - the coefficients of f. Because of the factor a^n in front of q, we conclude that p is a polynomial in the coefficients of f. Hence, p is an integer. □

3.3 Vanishing Measure

Lehmer's problem is about small positive values of $m(f)$. In this section we will consider the situation when $m(f) = 0$, following Everest and Wards book [6]. This can be completely understood using Kronecker's lemma [10]. We first need the following definition.

Definition 3.3.1. Let α be an algebraic integer with minimal polynomial f. Denote the roots of f by $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_n$. These roots (along with α itself) are the *algebraic conjugates* of α.

Lemma 3.3.2 (Kronecker)**.** *Suppose that $\alpha \neq 0$ is an algebraic integer and the algebraic conjugates $\alpha_1 = \alpha, \ldots, \alpha_n$ of α all lie inside or on the unit circle, that is $|\alpha_i| \leq 1$. Then α is a root of unity.*

Proof. For $k \in \mathbb{N}$, consider the polynomial

$$F_k(z) = \prod_{i=1}^{n} (z - \alpha_i^k).$$

Note that F_1 is the minimal polynomial of α. The coefficients of F_k are symmetric functions in the powers α_i^k. Because the Galois group of F_1 permutes the roots of F_1, it follows that the coefficients of F_k lie in the fixed field of F_1, which is $\mathbb{Q}$. Because the algebraic integers form a ring, these coefficients are algebraic integers as well. Hence, the coefficients of F_k are integers. Moreover, because $|\alpha_i| \leq 1$, each of the coefficients is uniformly bounded when we vary k. Therefore the set $\{F_k\}_{k \in \mathbb{N}}$ must be finite. Hence there exist $l, m \in \mathbb{N}$ with $l > m$ such that $F_l = F_m$. So, the roots of F_l and F_m are the same, but they are possibly permuted. Let τ in the permutation group S_d be such that

$$\alpha_i^l = \alpha_{\tau(i)}^m.$$

If τ has order r in S_n, then

$$\alpha_i^{l^r} = \alpha_i^{m^r},$$

so

$$\alpha_i^{m^r} \left(\alpha_i^{l^r - m^r} - 1 \right) = 0.$$

Since $\alpha_i \neq 0$, this shows that α_i must be a root of unity. □

Recall that a polynomial in $\mathbb{Z}[x]$ is called *primitive* if the greatest common divisor of its coefficients equals 1.

Theorem 3.3.3. *Let f be a polynomial in $\mathbb{Z}[x]$ with $f(0) \neq 0$. Then $m(f) = 0$ if and only if f is primitive and all the roots of f are roots of unity.*

Proof. Assume f is primitive and all of its roots are roots of unity. Then the leading coefficient of f should be ± 1 since f divides $x^N - 1$ for some $N \geq 1$. So, $m(f) = 0$.
Conversely, if $m(f) = 0$ then the leading coefficient of f is ± 1, hence f is primitive. From this also follows that all the roots are algebraic integers and because $m(f) = 0$ they must lie inside or on the unit circle. By Kronecker's lemma it follows that all roots are roots of unity. □

Definition 3.3.4. A polynomial is *cyclotomic* if all zeros are roots of unity.

Remark. Many authors (see, for example page 293 of [5]) define for a positive integer n the n-th cyclotomic polynomial as

$$\Phi_n(x) = \prod_{\substack{1 \leq k \leq n \\ \gcd(k,n)=1}} \left(x - e^{2i\pi \frac{k}{n}} \right),$$

which is the minimal polynomial of $e^{2\pi i/n}$. However, we will use the word cyclotomic for primitive, possibly reducible, polynomials with all the zeros roots of unity. For example, we will call the reducible polynomials

$$(x-1)(x^2+1) \quad \text{and} \quad x^{15}-1$$

cyclotomic. With this definition we can restate the theorem by saying that for $f \in \mathbb{Z}[x]$, $m(f) = 0$ if and only if f is a monomial times a primitive cyclotomic polynomial.

Chapter 4

Möbius Transformations

4.1 Automorphisms of the Riemann Sphere

In this chapter we will study Möbius transformations, also known as fractional linear transformations, from the viewpoint of both complex analysis and group theory. These transformations are of interest as in the next chapter we will apply them to study Lehmer's problem. We will follow Lang's book on complex analysis [11].

Definition 4.1.1. A *Möbius transformation* $\sigma : \mathbb{C} \to \mathbb{C}$ is a function of the form

$$\sigma(z) = \frac{az + b}{cz + d}$$

where $a, b, c, d \in \mathbb{C}$ and $ad - bc \neq 0$.

Example 4.1.2. The transformations $\sigma(z) = -z + 1 = \frac{-z+1}{0z+1}$ and $\sigma(z) = 1 - \frac{1}{z} = \frac{z-1}{z}$ are examples of Möbius transformations with coefficients in $\mathbb{Q}$.

With Definition 4.1.1, Möbius transformations are not defined for $z = -\frac{d}{c}$ when $c \neq 0$. Therefore, we extend our definition to the Riemann sphere $\hat{\mathbb{C}}$, that is the one-point compactification of $\mathbb{C}$ obtained by adding the point ∞. Define

$$\begin{cases} \sigma(\infty) = a/c \quad \text{and} \quad \sigma(-d/c) = \infty & \text{if } c \neq 0 \\ \sigma(\infty) = \infty & \text{if } c = 0 \end{cases}$$

These definitions are natural in the sense that if we take the appropriate limits, these limits coincide with our definitions. Hence σ is a continuous function on the Riemann sphere.

Recall the following well-known theorems about Möbius transformations. Proofs of these theorems can be found in [11].

Theorem 4.1.3. *Given any three distinct points z_1, z_2, z_3 on $\hat{\mathbb{C}}$ and any three distinct points w_1, w_2, w_3 (also on $\hat{\mathbb{C}}$), there exist a unique Möbius transformation σ such that*

$$\sigma(z_i) = w_i \quad \textit{for } i = 1, 2, 3.$$

Theorem 4.1.4. *A Möbius transformation maps straight lines and circles onto straight lines and circles.*

Remark. A straight line on the Riemann sphere is an ordinary line on $\mathbb{C}$ together with ∞. By adding ∞ to the line on $\mathbb{C}$, it becomes a circle on the Riemann sphere. Therefore, if we would have used another coordinate system on the sphere, we would not be able to distinguish between a straight line and a circle on the Riemann sphere. Hence, Theorem 4.1.4 could be reformulated as 'A Möbius transformation maps circles onto circles'.

Until now, it is unclear why we should be interested in specifically this kind of transformations and why they possess these kind of properties. In fact, Möbius transformations are the automorphisms of the Riemann sphere. To prove this, we first need to extend the notion of being meromorphic to the Riemann sphere.

Definition 4.1.5. Let $f : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ be a function on the Riemann sphere. Define $g(z) = f(1/z)$ for $z \neq 0, \infty$. Then f *is meromorphic at infinity* if g is meromorphic at 0. We say that f is *meromorphic on* $\hat{\mathbb{C}}$ if $f|_{\mathbb{C}}$ is meromorphic on $\mathbb{C}$ and f is also meromorphic at infinity.

For stating and proving the next lemma and theorem we use page 11 and 12 of [18].

Lemma 4.1.6. *The meromorphic functions on* $\hat{\mathbb{C}}$ *are precisely the rational functions.*

Proof. Let f be a given meromorphic function on $\hat{\mathbb{C}}$. Let h be the polynomial which is zero in all the poles of f counting multiplicity, that is, let $h(z) = \prod_{a\in\mathbb{C}}(z-a)^{\min(0,-\operatorname{ord}_a(f))}$. The polynomial h is well-defined because f has only finitely many poles. Consider the function $g(z) = f(z)h(z)$. This function is meromorphic because polynomials are meromorphic on $\hat{\mathbb{C}}$ and the product of two meromorphic functions is meromorphic. By construction, g has no poles on $\mathbb{C}$, hence it has a power series representation $g(z) = \sum_{i=0}^{\infty} a_n z^n$ for all $z \in \mathbb{C}$. Since g is meromorphic at ∞, the sum can contain only finitely many terms. So g is a polynomial and $f = g/h$ is rational function. □

An automorphism of the Riemann sphere is a bijective meromorphic function on $\hat{\mathbb{C}}$.

Theorem 4.1.7. *A function* $f : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ *is an automorphism of the Riemann sphere if and only if it is a Möbius transformation.*

Proof. First, assume that f is an automorphism of the Riemann sphere. By the previous lemma, we can find polynomials g and h such that $f = g/h$. Without loss of generality, assume that g and h have no common roots. Because $f(z) = 0$ has a unique solution, the degree of g is at most 1. Moreover, because $f(z) = \infty$ has a unique solution as well, the degree of h is at most 1. Hence, we can write f in the form of a Möbius transformation:

$$f(z) = \frac{az+b}{cz+d}.$$

If $ad = bc$, then g and h have the same root, which we assumed not to be the case. If $ad \neq bc$ then $\tilde{f} = \frac{dz-b}{-cz+a}$ is an inverse for f, hence f is an automorphism. Conversely, Möbius transformations are clearly meromorphic and invertible on $\hat{\mathbb{C}}$, hence they are automorphisms of the Riemann sphere. □

4.2 Projective Linear Group

The Möbius transformations form a group with composition as operation. Moreover, each two by two matrix can be identified with a Möbius transformation by the map

$$\phi : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \frac{az+b}{cz+d}.$$

Surprisingly, matrix multiplication corresponds exactly to the composition of transformations, as can easily be checked. Let K be a field. Then, ϕ is a surjective homomorphism from the general linear group $\mathrm{GL}(2, K)$, the group of all invertible 2×2 matrices. We will now find the kernel of this map to deduce an isomorphism.

Definition 4.2.1. Let K be a field. The *projective linear group* $\mathrm{PGL}(n, K)$ is the quotient of the general linear group $\mathrm{GL}(n, K)$ by its centre (which are the diagonal matrices).

Lemma 4.2.2. *The group of all Möbius transformations on the Riemann sphere with coefficients in a field extension K of $\mathbb{Q}$ is isomorphic to* $\mathrm{PGL}(2, K)$.

Proof. From our discussion at the beginning of the section it follows that the group of all Möbius transformations on the Riemann sphere is isomorphic to the quotient of $\mathrm{GL}(2, K)$ by the kernel of ϕ. Note that

$$\frac{az+b}{cz+d} = z \quad \text{for all} \quad z \in \hat{\mathbb{C}}$$

if and only if $a = d$ and $b = c = 0$. So the kernel ϕ is given by all matrices of the form λI where $\lambda \in K^*$ and I is the identity matrix. This is the center of $\mathrm{GL}(2, K)$. □

Remark. From now on we view the group of automorphisms of the Riemann sphere, the group of Möbius transformations and $\mathrm{PGL}(2, \mathbb{C})$ as identical, not merely isomorphic. We will freely use matrices to denote mappings and vice versa.

4.3 Finite Subgroups

In this section we are going to investigate the finite subgroups of $\mathrm{PGL}(2, K)$, where K is one of the fields $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$. A well-known result is the following one.

Theorem 4.3.1. *A finite subgroup of* $\mathrm{PGL}(2, \mathbb{C})$ *is isomorphic to a cyclic group, a dihedral group or a rotational symmetry group of one of the regular solids.*

We will omit the proof. However, the idea of the proof of this theorem is the following. First, one shows that a finite subgroup of $\mathrm{PGL}(2, \mathbb{C})$ is conjugate to a finite rotation group of $\hat{\mathbb{C}}$, which can be identified with $\mathrm{SO}(3, \mathbb{R})$. Thereafter one observes that a rotation which is not the identity has exactly two fixed points, corresponding to the intersection of the axis of rotation and the sphere. By applying (not) Burnside's lemma[1] (see page 98 of [1]) for the

[1] As Neumann pointed out [15], 'Burnside's Lemma' is not due to Burnside. Consequently, this lemma is sometimes referred to as 'the lemma that is not Burnside's' or 'the not Burnside lemma'. This result was already found, though with a rather unimportant restriction, in a paper by Cauchy in 1845. Without this restriction it can be traced back to F. G. Frobenius in 1887.

action of a finite group of rotations to the Riemann sphere, one can prove this theorem. This proof is due to Felix Klein in his famous *Lectures on the icosahedron* [9] and rewritten in the notation of modern mathematics by Shurman (see chapter 2 of [18]).

To find the finite subgroups of $\mathrm{PGL}(2,K)$ for the fields $K=\mathbb{Q}$ and $K=\mathbb{R}$ we follow Dresden [4]. We start with the following two lemmas.

Lemma 4.3.2. *A linear transformation $\sigma(z)=az+b$ has order n in $\mathrm{PGL}(2,\mathbb{C})$ if and only if a is a primitive nth root of unity.*

Proof. This follows directly from the observation that

$$\sigma^n(z)=a^nz+b(a^{n-1}+a^{n-2}\ldots+1). \quad \square$$

Lemma 4.3.3. *Suppose $\sigma\in\mathrm{PGL}(2,\mathbb{Q})$ has finite order. Then σ has order $1,2,3,4$ or 6 and in the last three cases, $\sigma(z)$ is conjugate to*

$$\frac{-1}{z+1},\frac{z-1}{z+1}, \quad \textit{resp.} \quad \frac{2z-1}{z+1}.$$

Remark. The order-2 maps are not all conjugate in $\mathrm{PGL}(2,\mathbb{Q})$. In particular, $\sigma(x)=-\frac{2}{z}$ is not conjugate to $\tau(x)=-z+1$. Namely, in that case there would be $a,b,c,d,\lambda\in\mathbb{Q}$ with $\lambda\neq 0$ such that

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}0&-2\\1&0\end{pmatrix}\begin{pmatrix}d&-b\\-c&a\end{pmatrix}=\lambda\begin{pmatrix}-1&1\\0&1\end{pmatrix}.$$

After matrix multiplication we find among other things that $bd+2ac=-\lambda$ and $d^2=2c^2$. The second equation is known to have to no solution in natural numbers, because 2 is not a square. Hence, the only rational solutions are $d=c=0$. This is in contradiction with the first equation.

In $\mathrm{PGL}(2,\mathbb{C})$ however, the order-2 maps are conjugate. It would be an interesting question to determine all order-2 maps in $\mathrm{PGL}(2,\mathbb{Q})$ up to conjugacy. Still there is a simple way to characterize order-2 maps. By matrix multiplication we find that

$$\sigma=\begin{pmatrix}a&b\\c&d\end{pmatrix} \quad \text{yields} \quad \sigma^2=\begin{pmatrix}a^2+bc&b(a+d)\\c(a+d)&bc+d^2\end{pmatrix}.$$

If $c=0$ we find $\sigma(z)=a/d\cdot z+b/d$. Only if $d=-a$ we find a map of order 2, namely

$$\sigma(z)=-z-b/a$$

If $c\neq 0$ we can scale our coefficients such that $c=1$. Moreover, because the component of the matrix σ^2 in the bottom left should be 0 we find that $d=-a$. So,

$$\sigma(z)=\frac{az+b}{z-a}$$

and one can easily check this indeed gives a Möbius transformation of order 2.

Proof of Lemma 4.3.3. One can easily verify that the three given maps in Lemma 4.3.3 have indeed order 3, 4, respectively 6. The map $\sigma(z)=-z+1$ has order 2 and of course the identity map has order 1.

Now we show that $1, 2, 3, 4$ and 6 are the only possible orders. Let $\sigma(z) = \frac{az+b}{cz+d} \in \mathrm{PGL}(2, \mathbb{Q})$ with finite order n. If $c = 0$ we can write $\sigma(z) = a/d \cdot z + b/d$. Because the only rational roots of unity are ± 1, by the previous lemma we find that σ has order 1 or 2.

If $c \neq 0$ we find at least one complex fixed point of σ by solving the equation

$$\frac{a\alpha + b}{c\alpha + d} = \alpha$$

for α. Because we have to solve a quadratic equation, this fixed point α has degree of at most 2 over $\mathbb{Q}$. Now, we will conjugate $\sigma(z)$ with

$$s(z) = \frac{1}{z - \alpha}$$

to get $\hat{\sigma}(z) = s \circ \sigma \circ s^{-1}(z)$. Since $s(\alpha) = \infty$, we have $s^{-1}(\infty) = \alpha$, and because σ fixes α we find that $\hat{\sigma}(\infty) = \infty$. From this it follows that $\hat{\sigma}$ is linear, thus there are $A, B \in \mathbb{Q}(\alpha)$ such that $\hat{\sigma}(z) = Az + B$. Because $\hat{\sigma}$ has the same order as σ, by the previous lemma we find that A is a primitive nth root of unity. Note that $\mathbb{Q}(\alpha)$ is at most quadratic over $\mathbb{Q}$ and $A \in \mathbb{Q}(\alpha)$. The only roots of unity for which the minimal polynomial over $\mathbb{Q}$ is at most quadratic are $\pm 1, \pm i$ and $\pm 1/2 \pm i\sqrt{3}/2$. Therefore, these are the only possibilities for A. So the order of σ is $1, 2, 3, 4$ or 6.

Finally, assume that $n = 3, 4$ or 6. We will show that $\sigma(z)$ is conjugate to one of the transformations above. Because σ has only a finite number of fixed points, there are three distinct numbers P, Q and R such that $\sigma : P \mapsto Q \mapsto R$. Let $s(z)$ be the unique Möbius transformation such that $s(P) = 0$, $s(Q) = -1$ and $s(R) = \infty$, which exists by Theorem 4.1.3, and note that s has rational coefficients. Then, for $\hat{\sigma}(z) = s \circ \sigma \circ s^{-1}(z)$, we have that $\sigma : 0 \mapsto -1 \mapsto \infty$. Therefore we have that

$$\hat{\sigma}(z) = \frac{\hat{a}z - 1}{z + 1},$$

where $\hat{a} = \hat{\sigma}(\infty)$. By calculating $\hat{\sigma}^n(z)$ for n equals $3, 4$ and 6, one can show that $n = 3, 4$, or 6 forces $\hat{a}$ to be $0, 1$, respectively 2. □

Theorem 4.3.4. *All finite subgroups of* $\mathrm{PGL}(2, \mathbb{R})$ *are isomorphic to a cyclic or a dihedral group.*

Proof. Observe that $\mathrm{PGL}(2, \mathbb{R})$ is a subgroup of $\mathrm{PGL}(2, \mathbb{C})$. Hence, by Theorem 4.3.1 we need only to show that the rotational symmetry groups of the regular solids are not subgroups of $\mathrm{PGL}(2, \mathbb{R})$. The rotational symmetry groups of the regular solids are the alternating groups A_4 and A_5 and the permutation group S_4 (see chapter 8 of [1]). Since $A_4 \subset S_4 \subset A_5$, it is enough to show that A_4 cannot be realized in $\mathrm{PGL}(2, \mathbb{R})$. We will use the fact that all the products of elements of order 3 and 2 in A_4 are of order 3, which can easily be checked by considering products as (abc) with $(ab)(cd)$ or writing down the multiplication table of A_4.

Assume that G is a finite subgroup of $\mathrm{PGL}(2, \mathbb{R})$ which is isomorphic to A_4. After an appropriate conjugation we can by Lemma 4.3.3 assume that an element of order 3 in G is given by

$$\frac{-1}{z + 1}.$$

However, every element of order 2 in $\mathrm{PGL}(2,\mathbb{R})$ is by the remark below Lemma 4.3.3 of the form $-z+b$ or $(az+b)/(z-a)$. Now, $-1/(z+1)$ composed with $-z+b$ is

$$\frac{1}{z-b-1}.$$

This has order 3 only for $b=-1\pm i$. And $-1/(z+1)$ composed with $(az+b)/(z-a)$ is

$$\frac{-z+a}{(a+1)z+(b-a)}.$$

This has order 3 only if $b=\frac{1}{2}(1+2a\pm\sqrt{-4a^2-4a-3})$, which is a complex number for all real values of a. This is a contradiction, so A_4 is not a subgroup of $\mathrm{PGL}(2,\mathbb{R})$. □

Now, we have enough information to determine all finite subgroups of $\mathrm{PGL}(2,\mathbb{Q})$. Let D_n be the dihedral group of order $2n$ with two non-commuting generators: one of order n and one of order 2.

Theorem 4.3.5. *A finite subgroup of* $\mathrm{PGL}(2,\mathbb{Q})$ *is isomorphic to a cyclic group* $\mathbb{Z}_n$ *or a dihedral group* D_n *where* n *equals* $1,2,3,4,$ *or* 6. *Moreover, for these values of* n *there indeed exists a subgroup of* $\mathrm{PGL}(2,\mathbb{Q})$ *isomorphic to* $\mathbb{Z}_n$ *respectively* D_n.

Proof. Because $\mathbb{Q}\subset\mathbb{R}$, by the previous theorem if follows that we only have to consider cyclic and dihedral groups. By Lemma 4.3.3 we can only have elements of order $1,2,3,4$ and 6, hence a finite subgroups of $\mathrm{PGL}(2,\mathbb{Q})$ is isomorphic to a cyclic group $\mathbb{Z}_n$ or a dihedral group D_n where n equals $1,2,3,4$, or 6.

The existence of subgroups of $\mathrm{PGL}(2,\mathbb{Q})$ isomorphic to a cyclic group for n is $1,2,3,4$ and 6 follows directly from the existence of elements in $\mathrm{PGL}(2,\mathbb{Q})$ of these orders n in Lemma 4.3.3. The following groups are finite dihedral subgroups of $\mathrm{PGL}(2,\mathbb{Q})$

$$\begin{aligned}
\left\{\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}0&1\\1&0\end{pmatrix}\right\} &\simeq D_1\simeq\mathbb{Z}_2,\\
\left\langle\begin{pmatrix}1&1\\1&-1\end{pmatrix},\begin{pmatrix}0&1\\1&0\end{pmatrix}\right\rangle &\simeq D_2,\\
\left\langle\begin{pmatrix}0&1\\-1&1\end{pmatrix},\begin{pmatrix}0&1\\1&0\end{pmatrix}\right\rangle &\simeq D_3,\\
\left\langle\begin{pmatrix}1&1\\-1&1\end{pmatrix},\begin{pmatrix}0&1\\1&0\end{pmatrix}\right\rangle &\simeq D_4,\\
\left\langle\begin{pmatrix}2&1\\-1&1\end{pmatrix},\begin{pmatrix}0&1\\1&0\end{pmatrix}\right\rangle &\simeq D_6,
\end{aligned}$$

concluding the theorem. □

Chapter 5

Theorems of Zhang and Dresden

5.1 Zhang's Theorem

Lehmer's problem is still unsolved. However, there is a Lehmer-like statement for the sum of the Mahler measure of $f(z)$ and $f(-z+1)$. Namely, letting n be the degree of f, Zhang found a constant $D > 0$, such that for $\sigma(z) = -z + 1$ [20]

$$m(f) + m(f \circ \sigma) = 0 \quad \text{or} \quad m(f) + m(f \circ \sigma) \geq nD$$

for all non-zero $f \in \mathbb{Z}[x]$. In particular, if f is a non-constant polynomial we have that

$$m(f) + m(f \circ \sigma) = 0 \quad \text{or} \quad m(f) + m(f \circ \sigma) \geq D.$$

This theorem was proven in the context of Mahler measure-like functions on algebraic integers. Namely, using the usual and so-called p-adic absolute values on an algebraic integer one can define a positive real-valued function h on every algebraic integer α, called the height of α. If the minimal polynomial f of α has degree n, then we have that

$$h(\alpha) = \frac{1}{n}\, m(f).$$

We will follow a more elementary proof of Zhang's theorem by Zagier [19], although phrased in terms of the Mahler measure (instead of the height of algebraic integers) as in [6]. Dresden has proven a similar result by using the transformation $\sigma(z) = 1 - \frac{1}{z}$ instead of $\sigma(z) = -z + 1$. Both are examples of sums of the Mahler measure under Möbius transformations. We will study these two examples before we present a general theorem of sums of Mahler measures under Möbius transformations.

Theorem 5.1.1 (Zhang, Zagier)**.** *Let $\omega = \frac{1}{2} + \frac{1}{2}i\sqrt{3}$ be a primitive 6th root of unity and let $\sigma(z) = 1 - z$. Suppose $f \in \mathbb{Z}[z]$ has degree n and 0,1 and ω are not roots of f. Then*

$$m(f) + m(f \circ \sigma) \geq \frac{n}{2} \log\left(\frac{1+\sqrt{5}}{2}\right)$$

and equality holds if and only if f or $f \circ \sigma$ is a power of $(z^2-z+1)^2+z(z-1)^2 = z^4-z^3+z^2-z+1$.

Remark. If we would not impose the condition that 0, 1 and ω are not roots of f, then z^n, $(1-z)^n$ or $(z^2-z+1)^n$ could divide $f(z)$ for every $n \in \mathbb{N}$. Hence $(1-z)^n, z^n$, respectively $((1-z)^2-(1-z)+1)^n = (z^2-z+1)^n$ would divide $f(1-z)$. The logarithmic Mahler measure of these polynomials then vanishes, because all roots are on the unit circle. Hence, by repeatedly multiplying f with one of these three polynomials, the right-hand side of Theorem 5.1.1 could be made arbitrary large, while the left-hand side remains unchanged.

This theorem will follow from Lemma 5.1.2 below. However, to prove this lemma we need to generalize the maximum modulus principle to so-called harmonic functions, which we will do in the next section. We denote $\log^+(z) = \log \max(z, 1)$.

Lemma 5.1.2. *For all $z \in \mathbb{C}\backslash\{0, 1, \omega, \overline{\omega}\}$ we have*

$$\frac{\sqrt{5}-1}{2\sqrt{5}} \log|z^2-z| + \frac{1}{2\sqrt{5}} \log|z^2-z+1| - \log^+|z| - \log^+|1-z| \leq -\frac{1}{2} \log\left(\frac{1+\sqrt{5}}{2}\right)$$

with equality precisely if z or $1-z$ equals a root of $z^4 - z^3 + z^2 - z + 1$.

5.2 Intermezzo: Harmonic Functions

In this section, we will follow chapter VIII on harmonic functions of Lang's book on complex analysis [11]. In the next section we will apply the tools developed in this section to the functions $\log|\cdot|$ and $\log^+|\cdot|$ to prove Lemma 5.1.2. We write $z \in \mathbb{C}$ as $x + iy$ for $x, y \in \mathbb{R}$.

Definition 5.2.1. A function $f : \mathbb{C} \to \mathbb{R}$ is called *harmonic* if it has continuous partial derivatives of order one and two and satisfies

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0.$$

Lemma 5.2.2. *The real part of an analytic function is harmonic.*

Proof. Let f be an analytic function. Its real and imaginary parts $u(x, y)$ and $v(x, y)$ are smooth. By the Cauchy-Riemann equations we have that

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

Taking the partial derivative of these equations to x respectively y and using that $\frac{\partial^2}{\partial x\, \partial y} = \frac{\partial^2}{\partial y\, \partial x}$ for smooth functions we find that

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

We conclude that u is harmonic. □

Example 5.2.3. For a given determination of the logarithm on $\mathbb{C}$ minus a half-line we have that $\log|z| = \Re(\log(z))$, the real part of $\log(z)$. By choosing two different branches of the logarithm we see that $\log|z|$ is harmonic on $\mathbb{C}^*$.
Moreover, $\log^+|z|$ is the real part of the zero function on the unit disk D_1 and equals $\log|z|$ outside $\overline{D}_1$. So $\log^+|z|$ is harmonic on $\mathbb{C}$ outside the unit circle $|z| = 1$.

The converse of Lemma 5.2.2 is also true and useful to deduce a maximum modulus theorem for harmonic functions below.

Lemma 5.2.4. *Let u be a harmonic function on a simply connected open set U. Then there exists an analytic function f on U such that $u = \Re f$.*

Proof. Consider

$$h = \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}.$$

Because

$$\frac{\partial^2 u}{\partial x^2} = -\frac{\partial^2 u}{\partial y^2} \quad \text{and} \quad \frac{\partial^2 u}{\partial x\, \partial y} = \frac{\partial^2 u}{\partial y\, \partial x}$$

it follows from the Cauchy-Riemann equations that h is analytic. As U is simply connected, h has a primitive f on U. Let $v = \Re f$. Then

$$h(z) = f'(z) = \frac{\partial v}{\partial x} - i\frac{\partial v}{\partial y}.$$

So the partial derivatives of u and v are the same. Now, let $g = u - v$. Then

$$\frac{\partial g}{\partial x} = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial x} = \Re h(z) - \Re h(z) = 0.$$

Analogously, we find that $\frac{\partial g}{\partial y} = 0$.
Now let $z_0 \in U$ and $\gamma : I \to U$ be a path from z_0 to a point z in U. This path exists because U is connected. By the chain rule it follows that

$$\frac{d}{dt} g(\gamma(t)) = \frac{\partial g}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial t} = 0.$$

Therefore $g(\gamma(t))$ is constant, so $g(z_0) = g(z)$. Hence, g is constant. So $u = v + C$ with $C \in \mathbb{C}$. From this it follows that $f - C$ is the desired analytic function on U with real part u. □

Theorem 5.2.5. *Let u be a harmonic function on a connected open set U. If u has a maximum at a point $z_0 \in U$, then u is constant.*

Proof. Let $D \subset U$ be an open disk containing x_0. Because a disk is simply connected, by Lemma 5.2.4 there is an analytic function f on D such that $u = \Re f$. Because the composition of analytic functions is analytic we also have that $e^{f(z)}$ is an analytic function and moreover

$$|e^{f(z)}| = e^{\Re f(z)} = e^{u(z)}.$$

Since the real exponent function is strictly increasing, the maximum z_0 for u is also a maximum for e^u and hence a maximum of $|e^f|$. By the maximum modulus principle for analytic functions, it follows that e^f is constant on the disk D, hence e^u is constant on D and finally u is constant on D.
We will now prove that u is not only constant on D, but also on U. Let S be the set of points $z \in U$ such that u is constant in an open neighbourhood of z with value $u(z_0)$. We give an open-closed argument to show that $S = U$. First of all S is not empty, because $z_0 \in S$. Secondly, by definition S is open. We will now show that S is closed in U. Let z_1 be in the closure of S in U. Because u is continuous and every neighbourhood of z_1 contains points of

S, we have that $u(z_1) = u(z_0)$. So u has a maximum at z_1 and by the first part of the proof we find that u is locally constant near z_1. Hence, $z_1 \in S$, from which it follows that S is closed. So, S is a non-empty connected component of U. Because U is connected, we find that $U = S$. □

Corollary 5.2.6. *Let u be a harmonic function on a connected open set U and continuous on its closure $\overline{U}$. If u is not constant on U, then any maximum of u on $\overline{U}$ occurs on the boundary ∂U.*

Proof. This follows immediately from Theorem 5.2.5, because this theorem implies that u does not have a maximum on U. □

5.3 Proof of Zhang's Theorem

Proof of Lemma 5.1.2 by Zagier. Let $L(z)$ be the left hand side of the proposed inequality, that is

$$L(z) = \frac{\sqrt{5}-1}{2\sqrt{5}} \log|z^2 - z| + \frac{1}{2\sqrt{5}} \log|z^2 - z + 1| - \log^+|z| - \log^+|1 - z|.$$

For $|z| > 2$ we can rewrite this as

$$\begin{aligned} L(z) &= \frac{\sqrt{5}-1+1}{2\sqrt{5}} \log|z^2 - z| + \frac{1}{2\sqrt{5}} \log\left|1 + \frac{1}{z^2 - z}\right| - \log|z| - \log|1 - z| \\ &= -\frac{1}{2} \log|z^2 - z| + \frac{1}{2\sqrt{5}} \log\left|1 + \frac{1}{z^2 - z}\right|. \end{aligned}$$

From this it is clear that $L(z) \to -\infty$ as $|z| \to \infty$. The function L is not defined for the points $0, 1, \omega, \overline{\omega}$, because these are the roots of $z^2 - z$ and $z^2 - z + 1$. However, in the limit where z goes to one of these points, $L(z) \to -\infty$. Moreover, L is continuous away from these points. So L attains a maximum.
Now, note that away from the circles $|z| = 1$ and $|1 - z| = 1$ the function L is harmonic, as we have seen in Example 5.2.3. By Corollary 5.2.6 the maximum of L on one of the connected components of $\{z \in \mathbb{C} \mid |z| \neq 1 \text{ and } |1 - z| \neq 1\}$ must be attained on these circles. Hence, L is maximized on one of these circles. Since L is symmetric under $z \mapsto 1 - z$ and $z \mapsto \overline{z}$, we may assume $z = e^{i\theta}$ for $0 \leq \theta \leq \pi$. Let

$$S = |z - 1|^2 = (e^{i\theta} - 1)(e^{-i\theta} - 1) = 4\sin^2\tfrac{\theta}{2},$$

and note that

$$|z^2 - z + 1|^2 = (e^{2i\theta} - e^{i\theta} + 1)(e^{-2i\theta} - e^{-i\theta} + 1) = (1 - 4\sin^2\tfrac{\theta}{2})^2 = (1 - S)^2.$$

We distinguish two cases. First, let $0 < \theta < \frac{\pi}{3}$, so that $0 < S < 1$. Note that $\theta = 0$ and $\theta = \frac{\pi}{3}$ correspond to $z = 0$ respectively $z = \omega$, which we excluded. Then

$$L(z) = \tfrac{\sqrt{5}-1}{4\sqrt{5}} \log S + \tfrac{1}{2\sqrt{5}} \log(1 - S).$$

Differentiating with respect to S yields that the only extremum found for $S \in (0,1)$ is $L = -\frac{1}{2}\log\left(\frac{1+\sqrt{5}}{2}\right)$ for $S = \frac{3-\sqrt{5}}{2}$ and $\theta = \frac{\pi}{5}$.
Similarly, for $\frac{\pi}{3} < \theta \leq 1$ and hence $1 < S \leq 4$ we have that

$$L(z) = \tfrac{-\sqrt{5}-1}{4\sqrt{5}}\log S + \tfrac{1}{2\sqrt{5}}\log(1-S).$$

The only extremum of L is found for $S = \frac{3+\sqrt{5}}{2}$ where $L = -\frac{n}{2}\log\left(\frac{1+\sqrt{5}}{2}\right)$ and $\theta = \frac{3\pi}{5}$.
We conclude that $L(z) \leq -\frac{1}{2}\log\left(\frac{1+\sqrt{5}}{2}\right)$ for all $z \in \mathbb{C}\backslash\{0,1,\omega,\overline{\omega}\}$. Note that equality holds if $z = e^{\frac{\pi}{5}i}$ or $z = e^{\frac{3\pi}{5}i}$ after applying $z \mapsto 1-z$ and/or $z \mapsto \overline{z}$ to z. Hence, we have equality if z or $1-z$ equals $e^{\pm\frac{\pi}{5}i}$ or $e^{\pm\frac{3\pi}{5}i}$. As $z^4 - z^3 + z^2 - z + 1$ has $e^{\pm\frac{\pi}{5}i}$ and $e^{\pm\frac{3\pi}{5}i}$ as its roots, we have equality if z or $1-z$ is a root of $z^4 - z^3 + z^2 - z + 1$. □

Proof of Theorem 5.1.1. Assume f has leading coefficient a. Let ξ_i be the zeros of f and for $z = \xi_i$, sum the inequality of Lemma 5.1.2 over all i to obtain

$$\frac{\sqrt{5}-1}{2\sqrt{5}}\log\left|\prod_{i=1}^{n}(\xi_i^2-\xi_i)\right| + \frac{1}{2\sqrt{5}}\log\left|\prod_{i=1}^{n}(\xi_i^2-\xi_i+1)\right| + \\ -\sum_{i=1}^{n}\log^+|\xi_i| - \sum_{i=1}^{n}\log^+|1-\xi_i| \leq -\frac{n}{2}\log\left(\frac{1+\sqrt{5}}{2}\right).$$

Note that $m(f) = \log|a| + \sum_{i=1}^{n}\log^+|\xi_i|$. Writing

$$2\log|a| = \frac{\sqrt{5}-1}{2\sqrt{5}}\log|a|^2 + \frac{1}{2\sqrt{5}}\log|a|^2 + \log|a|,$$

add and subtract $2\log|a|$ to the left hand side of this inequality to obtain

$$\frac{\sqrt{5}-1}{2\sqrt{5}}\log\left|a^2\prod_{i=1}^{n}(\xi_i^2-\xi_i)\right| + \frac{1}{2\sqrt{5}}\log\left|a^2\prod_{i=1}^{n}(\xi_i^2-\xi_i+1)\right| + \\ -m(f) - m(f\circ\sigma) + \log|a| \leq -\frac{n}{2}\log\left(\frac{1+\sqrt{5}}{2}\right).$$

Now, let $\mathrm{Gal}(f)$ be the Galois group of f. An element of $\mathrm{Gal}(f)$ permutes the roots of f, so

$$\prod_{i=1}^{n}(\xi_i^2-\xi_i) \quad \text{and} \quad \prod_{i=1}^{n}(\xi_i^2-\xi_i+1)$$

lie in the fixed field of f, which is $\mathbb{Q}$. Moreover,

$$\left|a^2\prod_{i=1}^{n}(\xi_i^2-\xi_i)\right| = \left|a\prod_{i=1}^{n}(\xi_i)\cdot a\prod_{i=1}^{n}(1-\xi_i)\right| = |f(0)f(1)|$$

and

$$\left|a^2\prod_{i=1}^{n}(\xi_i^2-\xi_i+1)\right| = \left|a\prod_{i=1}^{n}(\omega-\xi_i)\cdot a\prod_{i=1}^{n}(\overline{\omega}-\xi_i)\right| = |f(\omega)f(\overline{\omega})|.$$

Now, note that $0, 1, \omega$ and $\overline{\omega}$ are algebraic integers. Because the algebraic integers form a ring as we saw in Theorem 3.2.7, these products are algebraic integers. Equivalently, we could have

used lemma 3.2.9 to prove that these product are algebraic integers. Because $0, 1$ and ω are not roots of f, these product are positive integers. Therefore the logarithm of these product is greater than zero. We also have that $|a| \geq 1$, hence $\log|a| \geq 0$. So

$$-m(f) - m(f \circ \sigma) \leq -\frac{n}{2} \log\left(\frac{1+\sqrt{5}}{2}\right).$$

We conclude that

$$m(f) + m(f \circ \sigma) \geq \frac{n}{2} \log\left(\frac{1+\sqrt{5}}{2}\right).$$

Equality holds precisely if for all the roots of f equality holds in Lemma 5.1.2 and f is monic. Hence, equality holds if and only if f or $f \circ \sigma$ is a power of $(z^2 - z + 1)^2 + z(z-1)^2 = z^4 - z^3 + z^2 - z + 1$. □

5.4 Dresden's Theorem

Dresden has done something similar to Zhang's theorem using the permutation $\sigma(z) = 1 - \frac{1}{z}$ instead of $\sigma(z) = 1 - z$ [3]. For $z \in \mathbb{C}\backslash\{0, 1\}$ we have for Dresden's permutation that

$$\sigma^2(z) = 1 - \frac{1}{1 - \frac{1}{z}} = \frac{1}{1-z} \quad \text{and} \quad \sigma^3(z) = \frac{1}{1 - 1 + \frac{1}{z}} = z.$$

We will prove his theorem in this section following his ideas, although phrased in terms of Mahler measures instead of heights.

A problem in this case is that do not have that $f \circ \sigma \in \mathbb{Z}[x]$ for all $f \in \mathbb{Z}[x]$. For example, for $f(x) = x - 2$ we have that

$$f \circ \sigma(x) = -\frac{1}{x} - 1.$$

However, we can still consider the polynomial with as its roots the image of the roots of f under σ. That is, if ξ_i are roots of f, then we let f_σ be the primitive polynomial with as roots $\sigma(\xi_i)$. In Lemma 6.1.2 we will prove that $f \in \mathbb{Z}[x]$ and that if f is an irreducible polynomial, then f_σ is irreducible.

Theorem 5.4.1 (Dresden)**.** *Let $\omega = \frac{1}{2} + \frac{1}{2}i\sqrt{3}$ and $\sigma(z) = 1 - \frac{1}{z}$. Let $f \in \mathbb{Z}[x]$ with degree n be given such that $0, 1$ and ω are not roots of f. Let α be the unique root of*

$$g(z) = (z^2 - z + 1)^3 - (z^2 - z)^2 = z^6 - 3z^5 + 5z^4 - 5z^3 + 5z^2 - 3z + 1$$

with the greatest absolute value and positive imaginary part. Then

$$m(f) + m(f_\sigma) + m(f_{\sigma^2}) \geq n \log|\alpha|.$$

Equality holds if f, f_σ or f_{σ^2} is power of $g(z)$.

The proof of this theorem resembles the proof of Zhang's theorem. Similarly, we first need an inequality involving the permutations of z under σ.

Lemma 5.4.2. *There exists a $B \in \mathbb{R}$ with $0 < B < \frac{1}{2}$ such that for all $z \in \mathbb{C}\backslash\{0, 1, \omega, \overline{\omega}\}$ we have*

$$B \log \left| \frac{(z^2 - z + 1)^3}{(z^2 - z)^2} \right| - \log^+ |z| - \log^+ \left| 1 - \frac{1}{z} \right| - \log^+ \left| \frac{1}{1 - z} \right| \leq - \log |\alpha|.$$

Equality holds for all roots of $g(z) = z^6 - 3z^5 + 5z^4 - 5z^3 + 5z^2 - 3z + 1$.

We follow Dresden's proof of this lemma and describe how the value of B can be constructed. By doing this we gain more insight into the general case, whereas in the proof Zhang's lemma the constants $\frac{\sqrt{5}-1}{2\sqrt{5}}$ and $\frac{1}{2\sqrt{5}}$ were already given.

Proof of Lemma 5.4.2. Let $L(z)$ be the left hand side of the inequality. For $z \neq 0, 1$ we have that

$$\log \left| \frac{(z^2 - z + 1)^3}{(z^2 - z)^2} \right| = \log |z^2| + \log \left| \frac{(z^2 - z + 1)^3}{z^2 (z^2 - z)^2} \right|.$$

Now, note that for $|z| \to \infty$ we have that

$$\frac{(z^2 - z + 1)^3}{z^2 (z^2 - z)^2} \to 1, \quad 1 - \frac{1}{z} \to 1 \quad \text{and} \quad \frac{1}{1 - z} \to 0.$$

Hence,

$$\lim_{|z| \to \infty} L(z) = \lim_{|z| \to \infty} \log |z^{2B}| - \log |z|.$$

Because $B < \frac{1}{2}$ it follows that $L \to -\infty$ for $|z| \to \infty$. Similarly we have that L goes to $-\infty$ for z near $0, 1, \omega$ and $\overline{\omega}$.
We wish to find the maximum value of L, since such a maximum will give precisely the inequality necessarily to prove the lemma. By the same argument as in the proof of Zhang's theorem the function L attains its maximum on one of the curves $|z| = 1, |1 - \frac{1}{z}| = 1$ and $|\frac{1}{1-z}| = 1$ by the maximum principle for harmonic functions. Because L is symmetric under σ it is enough to consider only one of these three curves. Hence, consider the straight line $|1 - \frac{1}{z}| = 1$ which is parametrized by $z = \frac{1}{2} + iy$. Because L is symmetric under complex conjugation we only consider $y \geq 0$. Let $S = |\frac{1}{2} + iy|^2 = y^2 + \frac{1}{4}$. Then we have that

$$\left| \frac{(z^2 - z + 1)^3}{z^2 (z^2 - z)^2} \right| = \frac{(\frac{3}{4} - y^2)^3}{(y^2 + \frac{1}{4})^2} = \frac{(S + 1)^3}{S^2}$$

and

$$\left| \frac{1}{1 - z} \right| = \left| \frac{\frac{1}{2} + iy}{\frac{1}{4} + y^2} \right| = \frac{1}{\sqrt{S}}, \quad \left| 1 - \frac{1}{z} \right| = \left| \frac{-\frac{1}{4} + y^2 - iy}{\frac{1}{4} + y^2} \right| = 1.$$

We now distinguish two cases. When $0 < y < \frac{\sqrt{3}}{2}$ and hence $\frac{1}{4} < S < 1$ we have that

$$L = 3B \log(1 - S) + (\tfrac{1}{2} - 2B) \log S.$$

By differentiating with respect to S we find that L attains its maximum for $S = \frac{-4B+1}{2B+1}$. By computing the second derivative test, we conclude that this is indeed a maximum for $B < \frac{1}{4}$. Substituting this value for S in L yields

$$L = \tfrac{1}{2} \left((6B) \log(6B) + (1 - 4B) \log(1 - 4B) - (1 - 2B) \log(1 - 2B) \right).$$

Minimizing this for B yields that B is the single real root of $184z^3 + 6z - 1$. Let $-D$ be the value of L for this value of S and B.
In the second case, when $\frac{\sqrt{3}}{2} < y$ and hence $1 < S$, we find that

$$L = 3B\log(S-1) - (\tfrac{1}{2} + 2B)\log(S)$$

which is maximal for $S = \frac{4B+1}{-2B+1}$. For the same B as in the first case we find that L is now bounded above by

$$\tfrac{1}{2}\left((6B)\log(6B) - (1-4B)\log(1-4B) + (1-2B)\log(1-2B)\right).$$

It can be checked that this value is smaller than $-D$. Thus, the maximum value of g is $-D$. To deduce when equality is holding, note that this maximum is attained at

$$S = \frac{-4B+1}{1+2B}, \quad \text{or equivalently} \quad B = \frac{S-1}{-2S-4}.$$

By expressing B in terms of S and because B is a root of $184x^3 + 6x - 1$ one can show that S satisfies

$$S^3 - 2S^2 + 3S - 1 = 0.$$

Recalling that $S = y^2 + \frac{1}{4}$ and $z = \frac{1}{2} + iy$ we see that S is attained for a root of the polynomial

$$g(z) = z^6 - 3z^5 + 5z^4 - 5z^3 + 5z^2 - 3z + 1.$$

The other five roots of this polynomial are also maxima of g. These roots reflect the symmetry of the inequality: they can be found by complex conjugation and applying σ to the root α. We thus find that $D = \log|\alpha|$. □

Proof of Theorem 5.4.1. Assume f, f_σ and f_{σ^2} have leading coefficient a, b, respectively c. Sum the inequality of Lemma 5.4.2 over all zeros ξ_i of f and add $nB\log|(abc)^2| - \log|a| - \log|b| - \log|c|$ to the left-hand side. Because $B < 1/2$ and $a, b, c \in \mathbb{Z}$ this is smaller than or equal to 0. We then obtain

$$nB\log\left|(abc)^2\prod_{i=1}^{n}\frac{(\xi_i^2-\xi_i+1)^3}{(\xi_i^2-\xi_i)^2}\right| - m(f) - m(f_\sigma) - m(f_{\sigma^2}) \leq -n\log|\alpha|.$$

Now, let $\phi_6(z) = z^2 - z + 1$ be the sixth cyclotomic polynomial. The following identity holds

$$\frac{(z^2-z+1)^3}{(z^2-z)^2} = \phi_6(z)\cdot\phi_6(\sigma(z))\cdot\phi_6(\sigma^2(z)).$$

Hence,

$$\begin{aligned}\left|(abc)^2\prod_{i=1}^{n}\frac{(\xi_i^2-\xi_i+1)^3}{(\xi_i^2-\xi_i)^2}\right| &= \left|(abc)^2\prod_{i=1}^{n}\phi_6(\xi_i)\cdot\phi_6(\sigma(\xi_i))\cdot\phi_6(\sigma^2(\xi_i))\right| \\ &= \left|f(\omega)f(\overline{\omega})\cdot f_\sigma(\omega)f_\sigma(\overline{\omega})\cdot f_{\sigma^2}(\omega)f_{\sigma^2}(\overline{\omega})\right|.\end{aligned}$$

Because this product is symmetric in the roots ξ_i, it lies in the fixed field of the Galois group of f. As ω and $\overline{\omega}$ are algebraic integers, these products are non-negative integers. Equivalently,

the fact that these product are integers, can be shown by using Lemma 3.2.9. Since ω and $\overline{\omega}$ are not roots of f, these products are positive integers. So, the logarithm of this product is greater than zero. Hence, we can estimate our inequality by

$$-m(f) - m(f_\sigma) - m(f_{\sigma^2}) \leq -n \log |\alpha|$$

from which we conclude the theorem. Equality holds precisely if for all the roots of f equality holds in Lemma 5.4.2 and f is monic. Hence, equality holds if and only if $f, f \circ \sigma$ or f_{σ^2} is a power of $g(z)$. $\square$

Chapter 6

Mahler's Measure and Möbius Transformations

6.1 Möbius Transformations Acting on Polynomials

In this chapter, we will expand on the work done by Zagier and Dresden by giving a general procedure to find bounds of the sum of the Mahler measure of a polynomial under a Möbius transformation. To the best of my knowledge, this procedure cannot be found elsewhere.

First, we will define the action of a Möbius transformation on a polynomial. We already did this before stating Dresden's theorem, but now we will do it more rigorously.

Definition 6.1.1. For $\sigma \in \mathrm{PGL}(2,\mathbb{Q})$ and an irreducible polynomial $f \in \mathbb{Z}[x]$ with α as one of its roots and $\sigma(\alpha) \neq \infty$, define $f_{\sigma,\alpha}$ as the minimal polynomial of $\sigma(\alpha)$.

The following lemma implies that α is not needed in the notation $f_{\sigma,\alpha}$, that is, $f_{\sigma,\alpha}$ does not depend on the particular root α chosen.

Lemma 6.1.2. *Let $f \in \mathbb{Z}[x]$ be irreducible with roots $\xi_1, \xi_2, \ldots, \xi_n$. Let $\sigma \in \mathrm{PGL}(2,\mathbb{Q})$ and assume $\sigma(\xi_1) \neq \infty$. Then the roots of f_{σ,ξ_1} are $\sigma(\xi_1), \sigma(\xi_2), \ldots, \sigma(\xi_n)$.*

Proof. Let $g \in \mathbb{C}[x]$ be the monic polynomial with $\sigma(\xi_1), \sigma(\xi_2), \ldots, \sigma(\xi_n)$ as its roots, that is

$$g(z) = \prod_{i=1}^{n} (x - \sigma(\xi_i)).$$

We will show that beside a multiplicative factor g equals f_{σ,ξ_1}. First, we show that g is well-defined, that is that all the roots of g are non-infinite. If $\xi_1 \in \mathbb{Q}$, then it has no algebraic conjugates and because $\sigma(\xi_1) \neq \infty$, g is well-defined. If $\xi_1 \notin \mathbb{Q}$ all its conjugates are irrational numbers as well. As by definition a Möbius transformation can only be infinite for rational numbers, also in this case $\sigma(\xi_i) \neq \infty$ for all i, so g is well-defined. Note that the coefficients of g are symmetric fractions in the roots ξ_i, so these coefficients lie in the fixed field of $\mathrm{Gal}(f)$, which is $\mathbb{Q}$. Hence for some positive integer k, $\hat{g}(x) = k \cdot g(x)$ is a primitive polynomial in $\mathbb{Z}[x]$. Now, suppose $\hat{g}$ is reducible in $\mathbb{Z}[x]$, that is there exist integer polynomials h and j such that

$\hat{g} = h \cdot j$. Because f is irreducible, it is separable over $\mathbb{Q}$. As σ is a bijection on the Riemann sphere, it follows that $\hat{g}$ is separable as well. So, for every i we have that $\sigma(\xi_i)$ is a root of exactly one of the polynomials h and j. Take $S \subset \{1, 2, \dots n\}$ such that $\sigma(\xi_i)$ is a root of h for all $i \in S$ and $\sigma(\xi_i)$ is a root of j for all $i \in \{1, 2, \dots, n\} \backslash S$. Similarly to the construction of $\hat{g}$, we can show that there are primitive polynomials in $\mathbb{Z}[x]$ with $\sigma^{-1}\sigma(\xi_i) = \xi_i$ as their roots for $i \in S$ respectively for $i \in \{1, 2, \dots, n\} \backslash S$. Because f is the product of these polynomials, f should be reducible. This is a contradiction, so $\hat{g}$ is irreducible. Hence $\hat{g}$ is the minimal polynomial of $\sigma(\alpha)$, so $\hat{g} = f_{\sigma,\alpha}$. □

With this lemma, we can now define f_σ for all $f \in \mathbb{Z}[x]$.

Definition 6.1.3. For $\sigma \in \mathrm{PGL}(2, \mathbb{Q})$ and an irreducible polynomial $f \in \mathbb{Z}[x]$ with α as one of its roots and $\sigma(\alpha) \neq \infty$, let $f_\sigma = f_{\sigma,\alpha}$ be the minimal polynomial of $\sigma(\alpha)$. If $f \in \mathbb{Z}[x]$ is reducible, write it as a product of irreducible factors g_i with i in some finite index set I. Define f_σ as the product of $(g_i)_\sigma$ over all $i \in I$.

Example 6.1.4. Let

$$f(x) = x^3 - 3x^2 + x + 2 = (x - 2)(x^2 - x - 1)$$

and let

$$\sigma(z) = \frac{2z + 3}{2z - 1}.$$

We note that

$$\sigma(2) = \frac{7}{3} \quad \text{and} \quad \sigma\left(\frac{1 + \sqrt{5}}{2}\right) = \frac{5 + 4\sqrt{5}}{5}.$$

The corresponding minimal polynomials are $3x - 7$ and $5x^2 - 10x - 11$, so

$$f_\sigma(x) = (3x - 7)(5x^2 - 10x - 11) = 15x^3 - 65x^2 + 37x + 77.$$

Note that

$$\sigma\left(\frac{1 - \sqrt{5}}{2}\right) = \frac{5 - 4\sqrt{5}}{2}$$

is the other root of the polynomial $5x^2 - 10x - 11$.

Remark. A more direct approach to define f_σ would be to use the biholomorphism between the Riemann sphere $\hat{\mathbb{C}}$ and the complex projective line $\mathbb{C}P^1$, that is a bijective holomorphic function $\hat{\mathbb{C}} \to \mathbb{C}P^1$ such that the inverse is also holomorphic (see page 12 of [2]). A Möbius transformation on $\mathbb{C}P^1$ can be written in homogeneous coordinates as

$$\left[\frac{az + b}{cz + d} : 1\right] = [az + b : cz + d].$$

A polynomial $f(z) = \sum_{i=0}^{n} a_i z^i$ corresponds to the homogeneous polynomial

$$g(X, Y) = \sum_{i=0}^{n} a_i X^i Y^{n-i}$$

such that $f(z) = g(z, 1)$. In these coordinates it would be natural to define $\hat{f}_\sigma$ as corresponding to

$$g(aX + bY, cX + dY) = \sum_{i=0}^{n} a_i (aX + bY)^i (cX + dY)^{n-i}.$$

In our former coordinates it then follows that

$$\hat{f}_\sigma(z) = (cz + d)^n f\left(\frac{az + b}{cz + d}\right) = (cz + d)^n f(\sigma(z)).$$

Note that the roots of $\hat{f}_\sigma$ are given by $\sigma^{-1}(\xi_i)$, where ξ_i are the roots of f. However, $\hat{f}_\sigma(z)$ is not necessarily primitive. For example, for $\sigma(z) = \frac{z+1}{z-1}$ and $f(z) = z + 1$ we find that

$$\hat{f}_\sigma(z) = (z - 1)\left(\frac{z + 1}{z - 1} + 1\right) = z + 1 + z - 1 = 2z.$$

However, if we divide $\hat{f}_\sigma$ by its content $c(\hat{f}_\sigma)$, we have that $\hat{f}_\sigma / c(\hat{f}_\sigma)$ is a primitive polynomial with the same roots as $f_{\sigma^{-1}}$. So,

$$\hat{f}_\sigma / c(\hat{f}_\sigma) = f_{\sigma^{-1}}.$$

In what follows, we will always be interested in sets like $\{|f_\sigma(z)| : \sigma \in G\}$ for some group $G \subset \mathrm{PGL}(2, \mathbb{Q})$. Hence, in what follows it does not matter which of the two definitions is used.

6.2 A Lehmer-like Problem

Definition 6.2.1. Let $f \in \mathbb{Z}[x]$ and let G be a finite subgroup of $\mathrm{PGL}(2, \mathbb{Q})$. If f_σ is well-defined for all $\sigma \in G$, define *the logarithmic Mahler measure of f under G* as

$$m_G(f) = \sum_{\sigma \in G} m(f_\sigma)$$

Extend the definition to include $m_G(f) = \infty$ if f_σ is not defined for some $\sigma \in G$, that is $\sigma(\xi) = \infty$ for some root of f. In inequalities we use that ∞ is greater than all real numbers. By this convention we can omit the condition that f_σ is well-defined.

Example 6.2.2. In Zhang's theorem we had $G = \langle -z + 1 \rangle \cong \mathbb{Z}_2$ and found that

$$m_G(f) \geq \frac{n}{2} \log\left(\frac{1 + \sqrt{5}}{2}\right),$$

except when f equals $z, -z + 1$ or $z^2 - z + 1$, where $m_G(f)$ vanished. In Dresden's theorem we had $G = \langle 1 - \frac{1}{z} \rangle \cong \mathbb{Z}_3$ and found that in this case for $\log|\alpha| \approx 0.42180$ we have

$$m_G(f) \geq n \log|\alpha|,$$

with a vanishing logarithmic Mahler measure under G when f equals $z^2 - z + 1$.

We will now generalize these two examples. The problem we will consider is the following.

Problem 6.2.3. *Let a finite subgroup G of $\mathrm{PGL}(2,\mathbb{Q})$ be given. Does there exist a lower bound $D > 0$ such that for all polynomials f in $\mathbb{Z}[x]$ of degree n we have that*

$$m_G(f) = 0 \quad or \quad m_G(f) \geq nD \ ?$$

If so, can we explicitely find an optimal value of D, that is, a $D > 0$ such that there exist a polynomial f for which $m_G(f) = nD$?

The answer to this question depends on which of the following categories G belongs to:

Lemma 6.2.4. *For a finite subgroup G of $\mathrm{PGL}(2,\mathbb{Q})$ exactly one of the following three holds:*

(i) *$m_G(f) = 0$ for all cyclotomic polynomials f.*

(ii) *$m_G(f) = 0$ for at least one and at most three irreducible polynomials. These irreducible polynomials are of the form z, $z+1$, $z-1$, z^2+1, z^2+z+1 or z^2-z+1.*

(iii) *There is no polynomial f such that $m_G(f) = 0$.*

Proof. Assume first that all $\sigma \in G$ map the unit circle onto itself. For a cyclotomic polynomial f and $\sigma \in G$ we then have that f_σ has all its roots on the unit circle. Moreover, because f_σ is a minimal polynomial, it is primitive. Therefore, by Theorem 3.3.3 we find that $m(f_\sigma) = 0$, from which it follows that $m_G(f) = 0$. Hence, if all $\sigma \in G$ map the unit circle to itself, we are in case (i).

Now, note that a circle in the complex plane is uniquely determined by three points lying on it. So, by Theorem 4.1.4, which states that a Möbius transformation maps circles on the Riemann sphere to circles on the Riemann sphere, if three roots of unity are mapped to other roots of unity the whole circle is mapped into itself. So, if there is a $\sigma \in G$ which does not map the unit circle into itself, it sends at most two roots of unity to other roots of unity. Therefore, the roots of a cyclotomic polynomial f of degree three or higher are not all mapped to other roots of unity by σ. For such a polynomial, f_σ is not cyclotomic. Assuming that $m_G(f) = 0$ and $f(0) \neq 0$ and using Theorem 3.3.3 again implies that f is a cyclotomic polynomial of degree at most two.

Because monomials also have a vanishing Mahler measure and $\sigma(0) \in \mathbb{Q}$, σ might fix 0, in which case $\{0\}$ is an orbit of σ acting on $\hat{\mathbb{C}}$. Other possible orbits containing 0 are $\{0,-1\}, \{0,1\}$ and $\{0,-1,1\}$. So, at most four polynomials satisfy $m_G(f) = 0$, namely the zero polynomial, the minimal polynomial of -1 and 1 if these roots are fixed on the unit circle or mapped to 0 and a cyclotomic polynomial of degree two whose roots are fixed on the unit circle. Observe that if both the minimal polynomial of 1 and -1 satisfies $m_G(f) = 0$, then at least one of these roots is fixed, and because all other roots of unity are of degree two or higher, there could not be an additional cyclotomic polynomial. We conclude that if there is a $\sigma \in G$ which does not map the unit circle into itself and there is a polynomial f with $m_G(f) = 0$ that at most three of the polynomials z, $z+1$, $z-1$, z^2+1, z^2+z+1 and z^2-z+1 satisfy $m_G(f) = 0$. This is case (ii).

Finally, it is possible that there is a $\sigma \in G$ which does not map the unit circle into itself such that for no polynomial f the measure $m_G(f)$ vanishes. That is case (iii). □

We will use the short-hand notation $\omega_{\pm,\pm'} = \pm\frac{1}{2} \pm' \frac{1}{2}i\sqrt{3}$, which are the roots of z^2+z+1 and z^2-z+1. In this notation, the set of roots of the six polynomials of case (ii) is given by $S = \{0, \pm 1, \pm i, \omega_{\pm,\pm}\}$.

Notice that the two examples considered in chapter 5 satisfy condition (ii). For each of the conditions we will now examine for which groups it holds. In fact, there are only a few possibilities for finite subgroups G which satisfy condition (i) or (ii). Namely, we will use that if for $\sigma \in \mathrm{PGL}(2,\mathbb{C})$ there is a polynomial $f \in \mathbb{Z}[x]$ such that $m(f_\sigma) = 0$, we already have some information about the orbits the roots of f under σ.

6.3 Case (i): Unit Circle Preserving Groups

We will call a subgroup of $\mathrm{PGL}(2,\mathbb{Q})$ unit circle preserving if it satisfies condition (i) in Lemma 6.2.4, which is equivalent to saying that for all $\sigma \in \mathrm{PGL}(2,\mathbb{Q})$ the unit circle is mapped to itself. Moreover, $\sigma \in \mathrm{PGL}(2,\mathbb{Q})$ is called unit circle preserving if $\langle\sigma\rangle$ is unit circle preserving. For unit circle preserving groups we have that $m_G(f) = 0$ for all cyclotomic polynomials.

$G = \{\mathrm{e}\}$ is a unit circle preserving group. Problem 6.2.3 with $G = \{\mathrm{e}\}$ is like Lehmer's problem, the only difference is the factor n on the right side of the proposed inequality. For other finite unit circle preserving groups G this question seems to be slightly easier, because we have more information: if $m_G(f) = 0$, we do not only have that $m(f) = 0$, but also that $m(f_\sigma) = 0$ for all $\sigma \in G$. Nevertheless, like Lehmer's problem, it is unknown what to answer is to Problem 6.2.3 is for unit circle preserving groups. In the rest of this section we will characterize all unit circle preserving groups, although we are not able to solve Problem 6.2.3 for these groups.

Lemma 6.3.1. *The only unit circle preserving $\sigma \in \mathrm{PGL}(2,\mathbb{Q})$ are the identity e and transformations of the form*

$$A_q = \begin{pmatrix} 1 & q \\ -q & -1 \end{pmatrix} \quad \textit{and} \quad B_\pm = \begin{pmatrix} 0 & 1 \\ \pm 1 & 0 \end{pmatrix}$$

for $q \in \mathbb{Q}$ with $q \neq \pm 1$.

Proof. Let $\sigma \in \mathrm{PGL}(2,\mathbb{Q})$ be of finite order and sending the unit circle to itself. Because $\sigma(\pm 1) \in \mathbb{Q}$ it follows that either $\sigma(1) = 1$ and $\sigma(-1) = -1$, or $\sigma(1) = -1$ and $\sigma(-1) = 1$. In the first case, by writing σ as

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

we find that $a = d$ and $b = c$ by solving two linear equations. If $a = d = 0$ we have that σ equals B_+. Otherwise, we can divide by a and for $q = \frac{b}{a}$ we find that σ equals

$$\begin{pmatrix} 1 & q \\ q & 1 \end{pmatrix}.$$

If σ is of finite order, by Lemma 4.3.3 the order of σ is a divisor of 4 or 6. Solving $\sigma^4 = 0$ and $\sigma^6 = 0$ for $q \in \mathbb{Q}$ yields $q = 0$ as the only solution. Hence, we find the identity transformation.

The second case can be done similarly. We find that $b = -c$ and $a = -d$. If $a = -d = 0$, we obtain B_-. Else, we find A_q, which is of order 2 for all $q \in \mathbb{Q}\backslash\{-1, 1\}$.

Finally, note that for $z = e^{i\theta}$ with $\theta \in [0, 2\pi)$ we have that

$$|A_q(z)| = \frac{q^2 + 1 - 2q\cos\theta}{q^2 + 1 - 2q\cos\theta} = 1 \quad (\text{as } q \neq \pm 1)$$

as well as $|e(z)| = 1$ and $|B_\pm(z)| = \frac{1}{1} = 1$. Hence, transformations of the form e, A_q or $B_\pm$ are indeed unit circle preserving. We conclude that all unit circle preserving transformations are of the form e, A_q or $B_\pm$. □

Corollary 6.3.2. *All non-trivial unit circle preserving subgroups of* $\mathrm{PGL}(2, \mathbb{Q})$ *are isomorphic to* $\mathbb{Z}_2$ *or the Klein four-group* D_2*.*

Proof. All transformations A_q and $B_\pm$ are of order 2. All the elements of the groups $\mathbb{Z}_n$ and D_n in the classification of all finite subgroups of $\mathrm{PGL}(2, \mathbb{Q})$ are of order 2 if and only if $n = 2$. □

Example 6.3.3. For all $q \in \mathbb{Q}\backslash\{-1, 0, 1\}$ the group

$$\left\{ \mathrm{Id}, \begin{pmatrix} 1 & q \\ q & -1 \end{pmatrix}, \begin{pmatrix} 1 & -1/q \\ -1/q & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\}$$

is a finite unit circle preserving subgroup of $\mathrm{PGL}(2, \mathbb{Q})$ isomorphic to the Klein four-group.

6.4 Case (ii)

The second case of Lemma 6.2.4 is the most interesting one. In this case the answer to Problem 6.2.3 is 'yes'. With a procedure similar to the proofs of Zhang's theorem by Zagier, and Dresden's theorem for cyclic groups, we can find the constant D Problem 6.2.3 is asking for. Recall that in case (ii) we have that $G \subset \mathrm{PGL}(2, \mathbb{Q})$ permutes at least one of the special roots in $S = \{0, \pm 1, \pm i, \omega_{\pm,\pm}\}$. First of all, we will characterize all finite transformations of case (ii).

To find these, we will consider all possible cases of orbits of the roots of the six polynomials in (ii) under σ, similarly to the proof of Lemma 6.3.1. For example, consider the polynomial $f(z) = z$. As f is not cyclotomic with $m(f) = 0$, this polynomial is by assumption excluded in Theorem 3.3.3 which characterizes all polynomials with a vanishing Mahler measure. We have the following result.

Lemma 6.4.1. *Let* $f(z) = z$ *and* $\sigma \in \mathrm{PGL}(2, \mathbb{Q})$ *of finite order. If* $m_{\langle\sigma\rangle}(f) = 0$*, then* σ *is either the identity* e *or of the form*

$$\begin{pmatrix} 1 & 0 \\ r & -1 \end{pmatrix}, \begin{pmatrix} 1 & \pm 1 \\ r & -1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -3 & 1 \end{pmatrix} \quad \textit{or} \quad \begin{pmatrix} 1 & -1 \\ 3 & 1 \end{pmatrix}$$

for some $r \in \mathbb{Q}$*.*

Proof. This proof is not clarifying, but only consists of simple number theoretic and linear algebra calculations. The details can be filled in by the reader.

Let $\sigma = \frac{az+b}{cz+d}$. Because $m(f_\sigma) = 0$ and $\sigma(0)$ is rational, σ should map 0 to $-1, 0$ or 1. Because σ is of order $1, 2, 3, 4$ or 6 we have that σ^4 or σ^6 is the identity.

If $\sigma(0) = 0$ we have that $b = 0$. Assuming σ^4 equals the identity, it follows that $c = 0$ or $a = -d$. The first case implies that $a = \pm d$, so $\sigma(z) = \pm z$ and σ is either the identity or of the form of the first matrix of this lemma with $r = 0$. In the second case it follows from $ad - bc \neq 0$ that $a \neq 0$. After division by $-a$ we find that σ is of the form of the the first matrix. If σ^6 equals the identity, we find the same transformations.

If $\sigma(0) = 1$ we have that $b = d$. Because $m(f_{\sigma^2}) = 0$ as well, we know that 1 is sent to $-1, 0, 1$. Because 0 is mapped to 1, it is not possible that 1 is mapped to 1. If 1 is sent to 0, we find that $a = -b$. Then, σ^2 is the identity and σ is of the form of the second matrix with a minus sign. If 1 is sent to -1, still assuming that $\sigma(0) = 1$, we have that $a+b+c+d = 0$. If σ^4 equals the identity, then $b = 0$, $a = -b$ or $a = 3b$. None of these cases result in a $\sigma \in \mathrm{PGL}(2, \mathbb{Q})$ such that $m_{\langle\sigma\rangle}(f) = 0$. By calculating σ^6 in this case, we find that $b = 0$, $a = b$, $a = -b$ or $a = 5b$. Only if $a = b$ we have that $m_{\langle\sigma\rangle}(f) = 0$, in which case σ is of the form of the third matrix.

If $\sigma(0) = -1$, then we similarly find that σ is of the form of the second matrix with a plus sign or of the form of the fourth matrix.

Finally, note that by construction all the transformations σ in this lemma have the orbit of 0 under $\langle\sigma\rangle$ contained in $\{0, -1, 1\}$. Hence, we indeed have that $m_{\langle\sigma\rangle}(f) = 0$. □

Actually, not all transformations of the previous lemma are transformations of case (ii). For example, taking $r = 0$ in the first transformation of the above lemma, we find

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which is a unit circle preserving transformation. However, neglecting the transformation we found in Lemma 6.3.1, we have now characterized all finite order transformations of case (ii) for which $m_G(z) = 0$. We can generalize this lemma to find all finite order case (ii) transformations.

Theorem 6.4.2. *Let $\sigma \in \mathrm{PGL}(2, \mathbb{Q})$. Then σ is a case* (ii) *transformation if and only if σ is in the cyclic group generated by some matrix in the first column of Table 6.1.*

Proof sketch. For all six polynomials in condition (ii) (or equivalently: for all elements of S) we do the same calculations as in the proof of Lemma 6.4.1. Just as 0 had to be sent to 1 or -1, we have that i should be sent to $\pm i$ and $\omega_{\pm,+}$ should be sent to $\omega_{\pm,+}$ or $\omega_{\pm,-}$. Hence, we only have a finite number of cases to consider. □

Some transformations can be found in more than one cell of the first column of Table 6.1. For example, the transformation

$$\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$$

corresponds to the first row for $p/q = 1$ and to the seventh row for $p/q = 0$. All these transformations can be found in Table 6.3.

σ	Orbits in S	$\operatorname{ord}\sigma$	ϕ	B	$\exp(D)$	$D\approx$ (for $p/q=5$)	Equality
$\begin{pmatrix}1 & 0\\ p/q & -1\end{pmatrix}$	$\{0\}$	2	$\frac{z^2}{-pz+q}$	1	$\max(\lvert\lvert p\rvert - q\rvert, \lvert q\rvert)$	1.38629	$\begin{cases}1 & \text{if } p/q>0\\ -1 & \text{if } p/q<0\end{cases}$
$\begin{pmatrix}1 & \pm 1\\ p/q & -1\end{pmatrix}$	$\{0,\mp 1\}$	2	$\frac{z(z\pm 1)}{pz-q}$	1	$\max(\lvert p\mp q\rvert/2, \lvert q\rvert)$	0.69315	± 1 if $2 \mid p\pm q$
$\begin{pmatrix}1 & p/q\\ p/q\pm 2 & -1\end{pmatrix}$	$\{\pm 1\}$	2	$\frac{(z\mp 1)^2}{(p\pm 2q)z-q}$	1	$\max(\lvert p\pm 3q\rvert/4, \lvert p\mp q\rvert/4)$	0.69315	∓ 1 if $4 \mid p\mp q$
$\begin{pmatrix}1 & -1\\ 3 & 1\end{pmatrix}$	$\{0,-1,1\}$	3	$\frac{z(z-1)(z+1)}{(3z-1)(3z+1)}$	1	5	1.60944	i
$\begin{pmatrix}1 & p/q\\ p/q & -1\end{pmatrix}$	$\{i,-i\}$	2	$\frac{(z^2+1)^2}{(pz-q)^2}$	$\frac{1}{2}$	$\lvert\lvert p\rvert + q\rvert/2$	1.09861	$1,-1$ if $2 \mid p+q$
$\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}$	$\{i\},\{-i\}$	4	$\frac{(z^2+1)^4}{z^2(z-1)^2(z+1)^2}$	0.19409	$\lvert\alpha\rvert$	0.73286	α, maximal root[1] of $(z^2+1)^4+z^2(z^2-1)^2$
$\begin{pmatrix}1 & p/q\\ p/q\pm 1 & -1\end{pmatrix}$	$\{\omega_{\pm,+},\omega_{\pm,-}\}$	2	$\frac{(z^2\mp z+1)^2}{((p\pm q)z-q)^2}$	$\frac{1}{2}$	$\max(\lvert p\pm 2q\rvert/3, \lvert p\mp q\rvert/3)$	0.84730	∓ 1 if $3 \mid p\mp q$
$\begin{pmatrix}0 & 1\\ -1 & \pm 1\end{pmatrix}$	$\{\omega_{\pm,+}\},\{\omega_{\pm,-}\}$	3	$\frac{(z^2\mp z+1)^3}{z^2(z\mp 1)^2}$	0.11724	$\lvert\alpha\rvert$	0.42180	α, maximal root of $(z^2\mp z+1)^3-(z^2-z)^2$
$\begin{pmatrix}2 & \mp 1\\ \pm 1 & 1\end{pmatrix}$	$\{\omega_{\pm,+}\}$	6	$\frac{\left(z^2\mp z+1\right)^6}{(z(z\mp 2)(2z\mp 1)(z-1)(z+1))^2}$	0.30503	$\left\lvert\frac{\alpha(2\alpha-1)}{\alpha+1}\right\rvert$	1.75737	α, maximal root of $(z^2\mp z+1)^6+$ $z^2(2z^2\mp 5z+2)^2(z^2-1)^2$

Table 6.1: The first column of this table gives a list of all finite order $\sigma\in\mathrm{PGL}(2,\mathbb{Q})$ of case (ii) in Theorem 6.4.2, in the other columns information needed to solve Problem 6.2.3 is displayed. In this table, $p/q\in\mathbb{Q}$ such that $\det(\sigma)\neq 0$. In the first row we let $p/q\neq 0$, because for $p/q=0$ this is a unit circle preserving transformation. The column 'Orbits in S' shows all orbits of σ which are contained in $S=\{0,\pm 1,\pm i,\omega_{\pm,\pm}\}$. When B is displayed with a floating point, its exact value (in terms of α) can be found in Table 6.2. The inequality $B\log|\phi(z)|-\sum_{i=0}^{\operatorname{ord}\sigma-1}\log^+|\sigma^i(z)|\leq -D$ holds where the values of B,ϕ and D can be found in the corresponding row of the columns 'B', 'ϕ' and 'D'. From this, it follows that $m_{\langle\sigma\rangle}(f)=0$ or $m_{\langle\sigma\rangle}(f)\geq nD$ for all $f\in\mathbb{Z}[x]$. For powers of the minimal polynomial of the root in the last column equality holds in $m_{\langle\sigma\rangle}(f)\geq nD$.

[1]Maximal in absolute value and with non-negative imaginary part.

σ	$B \approx$	B	f	Gal(f)
$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$	0.19409	$\frac{\alpha^4-1}{2\alpha^4-12\alpha^2+4}$	$(z^2+1)^4+z^2(z^2-1)^2$	D_4
$\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$	0.11724	$\frac{\alpha^2-\alpha+1}{2(2\alpha^2+\alpha-1)}$	$(z^2-z+1)^3-(z^2-z)^2$	D_3
$\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$	0.30503	$\frac{2\alpha^6-6\alpha^5+10\alpha^3-15\alpha^2+9\alpha-1}{4\alpha^6-12\alpha^5-30\alpha^4+80\alpha^3-30\alpha^2-12\alpha+4}$	$(z^2-z+1)^6+z^2(2z^2-5z+2)^2(z^2-1)^2$	D_6

Table 6.2: The corresponding value of B expressed in terms of α, the maximal root of f (see the footnote on the previous page) for three $\sigma \in \mathrm{PGL}(2,\mathbb{Q})$. In all three cases the Galois group of f is dihedral, generated by σ and an element of order 2.

σ	Orbits in S	ord σ	ϕ_1	ϕ_2	B_1	B_2	exp(D)	$D \approx$	Equality
$\begin{pmatrix} 1 & 0 \\ \pm 1 & -1 \end{pmatrix}$	$\{0\}, \{\omega_{\pm+}, \omega_{\pm-}\}$	2	$\frac{z^2}{z-1}$	$\frac{z^2 \mp z+1}{z \mp 1}$	1/2	1/2	$\sqrt{\frac{1+\sqrt{5}}{2}}$	0.24061	$\mp e^{\frac{2\pi i}{5}}$
$\begin{pmatrix} 1 & 0 \\ \pm 2 & -1 \end{pmatrix}$	$\{0\}, \{\pm 1\}$	2	$\frac{z^2}{2z-1}$	$\frac{(z \mp 1)^2}{2z-1}$	2/3	1/3	$\sqrt{3}$	0.54931	$\omega_{\pm+}$
$\begin{pmatrix} 1 & \mp 1 \\ 0 & -1 \end{pmatrix}$	$\{0, \pm 1\}, \{\omega_{\pm+}, \omega_{\pm-}\}$	2	$z(z \pm 1)$	$(z^2 \mp z+1)^2$	$\frac{\sqrt{5}-1}{2\sqrt{5}}$	$\frac{1}{4\sqrt{5}}$	$\sqrt{\frac{1+\sqrt{5}}{2}}$	0.24061	$\pm e^{\frac{\pi i}{5}}$
$\begin{pmatrix} 1 & \mp 1 \\ \mp 1 & -1 \end{pmatrix}$	$\{0, \pm 1\}, \{i, -i\}$	2	$\frac{z(z \mp 1)}{z \pm 1}$	$\frac{(z^2+1)^2}{(z \pm 1)^2}$					
$\begin{pmatrix} 1 & \mp 1 \\ \mp 2 & -1 \end{pmatrix}$	$\{0, \pm 1\}, \{\omega_{\mp+}, \omega_{\mp-}\}$	2	$\frac{z(z \mp 1)}{2z \pm 1}$	$\frac{(z^2 \pm z+1)^2}{(2z \pm 1)^2}$					
$\begin{pmatrix} 1 & \mp 1 \\ \mp 3 & -1 \end{pmatrix}$	$\{0, \pm 1\}, \{\mp 1\}$	2	$\frac{z(z \mp 1)}{3z \pm 1}$	$\frac{(z \pm 1)^2}{3z \pm 1}$					

Table 6.3: All finite order $\sigma \in \mathrm{PGL}(2,\mathbb{Q})$ of case (ii) which belong to more than one line in Table 6.1 together with extra information needed to solve Problem 6.2.3. The column 'Orbits in S' shows all orbits of σ which are contained in $S = \{0, \pm 1, \pm i, \omega_{\pm,\pm}\}$. By Theorem 6.4.3 we have that $B_1 \log|\phi_1(z)| + B_2 \log|\phi_2(z)| - \sum_{i=0}^{\mathrm{ord}\,\sigma-1} \log^+|\sigma^i(z)| \leq -D$ for corresponding values in the columns 'B_1', 'B_2' and 'D'. From this it follows that $m_{\langle\sigma\rangle}(f) = 0$ or $m_{\langle\sigma\rangle}(f) \geq nD$ for all $f \in \mathbb{Z}[x]$. For powers of the minimal polynomial of the root in the last column equality holds in $m_{\langle\sigma\rangle}(f) \geq nD$. For the last three cases I was unable to find B_1, B_2 and D, such that the inequality $m_{\langle\sigma\rangle}(f) \geq D$ is strict.

Theorem 6.4.3. *For all case* (ii) *finite cyclic subgroups* $\langle\sigma\rangle$ *of* $\mathrm{PGL}(2,\mathbb{Q})$ *there exists a* $D \in \mathbb{R}$ *with* $D > 0$*, such that*

$$m_{\langle\sigma\rangle}(f) = 0 \quad \text{or} \quad m_{\langle\sigma\rangle}(f) \geq nD$$

for all $f \in \mathbb{Z}[x]$*. Moreover, the value for* D *found in Table 6.1 or 6.3 is optimal in the sense that equality holds for powers of the minimal polynomial of the root in the corresponding cell in the column 'Equality' in these tables.*

Proof for one case. For all groups $\langle\sigma\rangle$ the proofs are similar. Note that the eighth row in Table 6.1 and the third row in Table 6.3 correspond to Dresden's respectively Zhang's theorem, which we already proved in the previous chapter. Therefore, we will prove only one more instance, namely the inequality corresponding to the first row of Table 6.1.

For $\sigma = \frac{z}{p/q \cdot z - 1}$ we have to prove the inequality

$$\log\left|\frac{z^2}{pz-q}\right| - \log^+|z| - \log^+\left|\frac{z}{p/q\cdot z - 1}\right| \leq -\max(||p|-q|, |q|).$$

We will assume that p/q does not equal $0, \pm 1$ or ± 2, because these cases are considered in Table 6.3. Consider

$$L(z) = B\log\left|\frac{z^2}{pz-q}\right| - \log^+|z| - \log^+\left|\frac{z}{p/q\cdot z - 1}\right|,$$

for $0 < B < 1$. The reason for this factor B is that we now have that $L(z)$ tends to $-\infty$ if z tends to ∞, 0 or q/p. Next, we use the maximum modulus principle. As $L(z)$ is harmonic except for the circles $|z| = 1$ and $|\sigma(z)| = 1$, it attains its maximum on these circles. Because $L(z)$ is symmetric under σ, its maximum can be found on the unit circle. Assume first that $|p/q \cdot e^{i\theta} - 1|^{-1} < 1$. Then we have that

$$L(e^{i\theta}) = -\tfrac{1}{2}B\log(p^2 + q^2 - 2pq\cos\theta).$$

This is maximized for $\cos\theta = 1$ if $p/q > 0$ and for $\cos\theta = -1$ if $p/q < 0$. For $p/q > 0$ we find as maximum $-B\log|p-q|$, while for $p/q < 0$ we get $-B\log|p+q|$. In both cases, this equals $-B\log||p|-q|$.

Next, assume that $|p/q \cdot e^{i\theta} - 1|^{-1} \geq 1$. In this case we have

$$L(e^{i\theta}) = \tfrac{1}{2}(1-B)\log(p^2 + q^2 - 2pq\cos\theta) - \log|q|$$

which is maximized by $(1-B)\log||p|+q| - \log|q|$. So, in any case, we have that

$$L(z) \leq \max(B\log||p|-q|, (B-1)\log||p|+q| - \log|q|)$$

with equality for $z = 1$ or $z = -1$. By calculating $L(1)$ and $L(-1)$ we get the stronger inequality

$$L(z) \leq \min(B\log||p|-q|, (B-1)\log||p|+q| - \log|q|).$$

As $L(z)$ is continuous for all $z \in \mathbb{C}$ except for $z = q/p$ and (non-strict) inequalities are preserved under limits, it follows that for $z \neq q/p$

$$\begin{aligned}\lim_{B\to 1} L(z) &= \log\left|\frac{z^2}{pz-q}\right| - \log^+|z| - \log^+\left|\frac{z}{p/q\cdot z-1}\right| \\ &\leq -\max(\log||p|-q|, \log|q|).\end{aligned}$$

Because

$$\lim_{\substack{B\to 1\\ z\to q/p}} L(z) = \begin{cases} -\log|p| & \text{if } |p/q|>1 \\ -\log|q| & \text{if } |p/q|<1 \end{cases}$$

we deduce that for all $z\in\mathbb{C}$

$$\left|\frac{z^2}{pz-q}\right| - \log^+|z| - \log^+\left|\frac{z}{p/q\cdot z-1}\right| \le -\max(\log||p|-q|, \log|q|).$$

Let $f\in\mathbb{Z}[x]$ be a polynomial with roots ξ_i and leading coefficient a. By adding $2\log|a| - \log|a| - \log|a|$ and summing the inequality over all i for $z=\xi_i$ we obtain

$$\log\left|a^2\prod_{i=1}^n \xi_i\sigma(\xi_i)\right| - m(f) - m(f_\sigma) \le -n\max(\log||p|-q|, \log|q|).$$

As $\log\left|a^2\prod_{i=1}^n \xi_i\sigma(\xi_i)\right| = \log|f(0)f(\sigma(0))| \ge 0$ we conclude that

$$m(f)+m(f_\sigma) \ge n\max(\log||p|-q|, \log|q|).$$

Equality holds for $f=z\pm 1$ (depending on the sign of p/q). Namely, we have that

$$f_\sigma(z) = z - \frac{1}{p/q\mp 1} = z - \frac{q}{p\mp q}.$$

with $m(f_\sigma)=\log\max(|p\mp q|,|q|)$. □

Remark. Although it is not clear from the proof, there is a general way to find ϕ for a finite group $G\subset \mathrm{PGL}(2,\mathbb{Q})$. Recall that for every case (ii) transformation there is an orbit contained in S. Let g be the minimal polynomial of some element of this orbit with degree m. Then we define

$$\phi(z) = \frac{1}{E}\prod_{\sigma\in G} g(\sigma(z)),$$

where E is the content of the numerator of $\phi(z)$ when $\phi(z)$ is written in the form $p(z)/q(z)$ for relatively prime integer polynomials p and q. Because ϕ is symmetric under σ, it does not matter which element we have chosen to define g. Now, write

$$\sigma(z) = \frac{a_\sigma z + b_\sigma}{c_\sigma z + d_\sigma}$$

for rational numbers $a_\sigma, b_\sigma, c_\sigma$ and d_σ. Rewrite ϕ as

$$\phi(z) = \frac{1}{E}\prod_{\sigma\in G}\frac{(c_\sigma z+d_\sigma)^m g(\sigma(z))}{(c_\sigma z+d_\sigma)^m}.$$

By the alternative definition of g_σ we have $g_{\sigma^{-1}}(z) = \frac{1}{E_\sigma}(c_\sigma z+d_\sigma)^m g(\sigma(z))$ where E_σ is a constant such that $g_{\sigma^{-1}}$ is primitive. Hence,

$$\phi(z) = \frac{1}{E}\prod_{\sigma\in G}\frac{E_\sigma\cdot g_{\sigma^{-1}}(z)}{(c_\sigma z+d_\sigma)^m}.$$

Assume

$$\prod_{\sigma\in G} E_\sigma\cdot g_{\sigma^{-1}}(z) \quad\text{and}\quad \prod_{\sigma\in G}(c_\sigma z+d_\sigma)^m$$

have a common root, then $-\frac{d_\sigma}{c_\sigma}$ is a root of g_τ for some $\sigma, \tau \in G$. As $\sigma(-\frac{d_\sigma}{c_\sigma}) = \infty$ it follows that $g_{\sigma\tau}(\infty) = 0$. This is a contradiction with the fact that $g_{\sigma\tau}$ is a polynomial. Hence,

$$\prod_{\sigma \in G} E_\sigma \cdot g_{\sigma^{-1}}(z) \quad \text{and} \quad \prod_{\sigma \in G} (c_\sigma z + d_\sigma)^m$$

are relatively prime polynomials over $\mathbb{Q}[z]$. By scaling the coefficients $a_\sigma, b_\sigma, c_\sigma$ and d_σ we can assume that the latter is primitive in $\mathbb{Z}[z]$. Then, $\prod_{\sigma \in G} E_\sigma = E$ and

$$\phi(z) = \prod_{\sigma \in G} \frac{g_\sigma(z)}{(c_\sigma z + d_\sigma)^m}.$$

When there is more than one orbit contained in S, as is the case in Table 6.3, one can define ϕ for every orbit with the above construction.

Remark. The above proof does not explain how the constant B is obtained as well. In fact, for a finite Möbius transformation σ there is no known general procedure for computing the constant B or B_1 and B_2.

One way to find B or B_1 and B_2 is by minimizing the maximum of

$$B \log|\phi(z)| - \sum_{i=0}^{k-1} \log^+ |\sigma^i(z)| \quad \text{or} \quad B_1 \log|\phi_1(z)| + B_2 \log|\phi_2(z)| - \sum_{i=0}^{k-1} \log^+ |\sigma^i(z)| \qquad (*)$$

for B respectively B_1 and B_2. Sometimes, as in the proof of Dresden's theorem, this can be done with the standard techniques of (complex) analysis for finding extrema. However, the equations involved are sometimes impossible or at least too difficult to solve algebraically.

A trick for solving these equations is by guessing the polynomial f for which $m_{\langle\sigma\rangle}(f)$ is minimal yet greater than 0. A good guess is a cyclotomic polynomial of low degree. Another good guess is the numerator of $\phi(z)$ plus or minus the denominator of $\phi(z)$. If the guess is correct, the roots of this polynomial might maximize (*), which can be used to simplify the equations found by minimizing. This way the constant in the last row of Table 6.1 is found. As the white space in Table 6.3 shows, I did not succeed in finding these constants in all cases.

Example 6.4.4. For non-cyclic groups the same procedure works, provided that one can find a suitable constant B. Table 6.4 shows two examples of dihedral groups for which the constant D of Problem 6.2.3 is found.

Generators of G	$G \simeq$	Orbit in S	ϕ	B	$\exp(D)$	Equality
$\begin{pmatrix}1 & -1\\ 0 & -1\end{pmatrix}, \begin{pmatrix}1 & 1\\ 2 & -1\end{pmatrix}$	D_2	$\{\omega_{++}, \omega_{+-}\}$	$\frac{(z^2-z+1)^4}{(2z-1)^4}$	$\frac{1}{4}$	2	-1
$\begin{pmatrix}1 & -1\\ 3 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$	D_3	$\{-1, 0, 1\}$	$\frac{(z(z-1)(z+1))^2}{(9z^2-1)^2}$	1	25	i

Table 6.4: Two finite dihedral subgroups of $\mathrm{PGL}(2, \mathbb{Q})$ for which property (ii) holds. The column 'Orbits in S' shows all orbits of G which are contained in $S = \{0, \pm 1, \pm i, \omega_{\pm,\pm}\}$. The inequality $B \log|\phi(z)| - \sum_{\sigma \in G} \log^+ |\sigma(z)| \leq -D$ implies that $m_G(f) = 0$ or $m_G(f) \geq nD$ for all $f \in \mathbb{Z}[x]$. For powers of the minimal polynomial of the root in the last column equality holds in $m_G(f) \geq nD$.

6.5 Case (iii)

For finite groups satisfying case (iii) the answer to Problem 6.2.3 is 'yes': there exists a $D > 0$ such that $m_G(f) \geq nD$ for all $f \in \mathbb{Z}[x]$. Notice that in this case this inequality holds for all integer polynomials. There are no exceptions.

Lemma 6.5.1. *For a finite cyclic subgroup $G = \langle \sigma \rangle$ of $PGL(2, \mathbb{Q})$ of case* (iii) *one of the following holds:*

(a) *Every orbit of the action of G on $\hat{\mathbb{C}}$ contains an element z with $|z| \neq 1$.*

(b) *There exist an $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ such that the only orbits of the action of G on $\hat{\mathbb{C}}$ for which $|z| = 1$ for all z in the orbit, are $\{\alpha\}$ and $\{\alpha^*\}$, or $\{\alpha, \alpha^*\}$.*

Proof. Assume first that there are at least three elements on the unit circle which are mapped by σ to other elements on the unit circle. As Möbius transformations send circles to circles, and circles are uniquely determined by three points, it follows that G is unit circle preserving. This contradicts the assumption that $\langle \sigma \rangle$ satisfies case (iii). Hence, there are at most two elements on the unit circle which are mapped to other element on the unit circle.

Now, assume that we are not in case (a). Let $\alpha \in \hat{\mathbb{C}}$ with $|\alpha| = 1$ be an element of an orbit contained in the unit circle. If $\sigma(\alpha) = \alpha$, it follows by complex conjugation that $\sigma(\alpha^*) = \alpha^*$. Note that if $\alpha = \pm 1$, $m_G(z \pm 1) = 0$ and σ is of case (i) or (ii). Hence, $\alpha \neq \alpha^*$, so $\{\alpha\}$ and $\{\alpha^*\}$ are two different orbits of σ. As there are at most two elements on the unit circle which are mapped to other elements on the unit circle, it follows that all other orbits contain an element z with $|z| \neq 1$.

Next, assume $\sigma(\alpha) = \beta$ with $\beta \neq \alpha$ and $|\beta| = |\alpha| = 1$. As there are at most two elements on the unit circle which are mapped to other element on the unit circle, it follows that $\sigma(\beta) = \alpha$. Moreover, $\sigma(\alpha^*) = \beta^*$, from which it follows that $\alpha = \beta^*$. So, $\{\alpha, \alpha^*\}$ is an orbit of the action of G on $\hat{\mathbb{C}}$ and all other orbits contain an element z with $|z| \neq 1$. □

Theorem 6.5.2. *Let G be a finite subgroup of $PGL(2, \mathbb{Q})$ and assume that every orbit of the action from G on $\mathbb{C}$ contains an element z with $|z| \neq 1$. Then,*

$$\sum_{\sigma \in G} \left(\log |\sigma(z)| - \log^+ |\sigma(z)| \right)$$

attains a maximum $-D < 0$. Let $f \in \mathbb{Z}[x]$ with degree n and $z \nmid f_\sigma$ for all $\sigma \in G$. Then, we have that

$$m_G(f) \geq nD.$$

Proof. Note that

$$\tfrac{1}{2} \log |z| - \log^+ |z| = \begin{cases} \tfrac{1}{2} \log |z| & \text{if } |z| \leq 1 \\ -\tfrac{1}{2} \log |z| & \text{if } |z| > 1. \end{cases}$$

Therefore, $\frac{1}{2} \log |z| - \log^+ |z| \leq 0$ with equality if and only if $|z| = 1$. Because every orbit of the action from G on $\mathbb{C}$ contains an element which does not lie on the unit circle, for every $z \in \hat{\mathbb{C}}$ there exist a $\sigma \in G$ such that $|\sigma(z)| \neq 1$. Hence,

$$\sum_{\sigma \in G} \left(\tfrac{1}{2} \log |\sigma(z)| - \log^+ |\sigma(z)| \right) < 0.$$

for all $z \in \mathbb{C}$. Let $-D$ be the maximum of the left-hand side, that is for all $z \in \mathbb{C}$ we have

$$\sum_{\sigma \in G} \left(\tfrac{1}{2} \log |\sigma(z)| - \log^+ |\sigma(z)| \right) \leq -D.$$

This maximum exists and is attained on one of the circles $|\sigma(z)| = 1$ for $\sigma \in G$, because the left-hand side is a bounded harmonic function. Moreover, $D > 0$.

Next, let $f \in \mathbb{Z}[x]$ be a polynomial of degree n and with roots ξ_i for $i \in \{1, \ldots, n\}$. Let a_σ be the leading coefficient of f_σ. By summing the inequality for $z = \xi_i$ over all $i \in \{1, \ldots, n\}$ and adding

$$\sum_{\sigma \in G} \left(\tfrac{1}{2} \log |a_\sigma| - \log |a_\sigma| \right) \leq 0,$$

we obtain

$$\sum_{\sigma \in G} \tfrac{1}{2} \log \left| a_\sigma \prod_{i=1}^{n} \sigma(\xi_i) \right| - m_G(f) \leq -nD.$$

Hence,

$$\sum_{\sigma \in G} \tfrac{1}{2} \log |f_\sigma(0)| - m_G(f) \leq -nD.$$

Since $z \nmid f_\sigma$ for all $\sigma \in G$ by assumption, we have that $\sum_{\sigma \in G} \frac{1}{2} \log |f_\sigma(0)| \geq 0$. We conclude that $m_G(f) \geq nD$. □

Example 6.5.3. Let

$$\sigma = \begin{pmatrix} 1 & 3 \\ 9 & -1 \end{pmatrix}$$

be a Möbius transformation of order 2. For $z = e^{i\theta}$ with $\theta \in \mathbb{R}$ we have that

$$|\sigma(z)|^2 = \frac{10 + 6\cos\theta}{82 - 18\cos\theta}.$$

As $\cos\theta \leq 1$ we find that $|\sigma(z)| \leq \frac{1}{2}$. Hence, if O is some orbit of σ acting on $\mathbb{C}$ and O contains an element on the unit circle, say of the form $e^{i\theta}$, it also contains an element which does not lie on the unit circle. So, motivated by Theorem 6.5.2 we want to maximize

$$\tfrac{1}{2} \log |z| + \tfrac{1}{2} \log |\sigma(z)| - \log^+ |z| - \log^+ |\sigma(z)|.$$

over $\mathbb{C}$. As this is a harmonic function except on the circles $|z| = 1$ and $|\sigma(z)| = 1$, symmetric under $z \to \sigma(z)$, the maximum is attained for z on the unit circle. Setting $z = e^{i\theta}$ we have to maximize the following function

$$\tfrac{1}{4} \log \left(\frac{10 + 6\cos\theta}{82 - 18\cos\theta} \right).$$

The maximum is found for $\theta = 0$ and equals $-\frac{1}{2}\log(2)$. Hence,

$$m_{\langle \sigma \rangle} \geq \tfrac{n}{2} \log(2)$$

for all $f \in \mathbb{Z}[x]$ with $f(0)$ and $f_\sigma(0)$ different from 0 and of degree n. As $\sigma(0) = -3$ we have that $m_{\langle \sigma \rangle}(z) = \log(3) > \frac{1}{2}\log(2)$. Hence, we can conclude that

$$m_{\langle \sigma \rangle} \geq \tfrac{n}{2} \log(2)$$

for all $f \in \mathbb{Z}[x]$. Notice that for $-z+1$, the minimal polynomial of $1 = e^{i\cdot 0}$, we have that $m_{\langle\sigma\rangle} = \log(2)$ because $\sigma(1) = \frac{1}{2}$. So, we do not have a strict inequality. However, if we would have instead chosen to maximize

$$\log|z| + \log|\sigma(z)| - \log^+|z| - \log^+|\sigma(z)|,$$

we would have found that

$$m_{\langle\sigma\rangle}(f) \geq n\log(2)$$

for all $f \in \mathbb{Z}[x]$. Equality then holds if and only if f or f_σ is a power of $-z+1$.

We conclude this chapter with the theorem below. This theorem applies not only to groups of case (iii), but also to groups of case (ii). We only have to assume that the action of G on $\hat{\mathbb{C}}$ has a finite number of orbits X with the property that all non-zero elements of X lie on the unit circle. For transformations of case (i) this is not the case as all orbits are finite and there is an infinite number of elements on the unit circle. Hence, there is an infinite number of orbits X for which all the non-zero elements lie on the unit circle. Unfortunately, the theorem does not give the optimal value of D, that is, a value such that equality holds in $m_G(f) \geq nD$ for some polynomial f. It only asserts that such a D exists.

Theorem 6.5.4. *Let G be a finite subgroup of $PGL(2,\mathbb{Q})$. Let $\mathcal{O} = \{O_i \mid i \in I\}$ be the set of all the orbits of G for which every non-zero element lies on the unit circle. Assume $\mathcal{O}$ is finite and non-empty. Let p_i be the minimal polynomial of an element in O_i. Let*

$$\phi_i(z) = \prod_{\sigma\in G} p_i(\sigma(z))$$

Then, there exists a constant $A > 0$ such that if $0 < B_i \leq A$ for all $i \in I$, we have that

$$\sum_{i\in I} B_i \log|\phi_i(z)| + \sum_{\sigma\in G}\left(\tfrac{1}{2}\log|\sigma(z)| - \log^+|\sigma(z)|\right)$$

attains a maximum $-D < 0$, where D depends on the values of B_i. Let $f \in \mathbb{Z}[x]$ be of degree n such that $p_i \nmid f_\sigma$ and $z \nmid f_\sigma(z)$ for all $i \in I$ and $\sigma \in G$. Then

$$m_G(f) \geq nD.$$

Proof. Let $B = \{B_i | i \in I\}$. Write ϕ_i as a fraction of relative prime polynomials. Let k_i be the degree of the numerator of ϕ_i minus the degree of the denominator of ϕ_i. Let k be the maximum of the k_i, which exists because $\mathcal{O}$ is finite. Assuming $A < \frac{1}{2k|I|}$ and $0 < B \leq A$ we have for $|z|$ approaching ∞ that the quantity

$$L_B(z) = \sum_{i\in I} B_i \log|\phi_i(z)| + \sum_{\sigma\in G}\left(\tfrac{1}{2}\log|\sigma(z)| - \log^+|\sigma(z)|\right)$$

goes to $-\infty$ by similar arguments as in Zagier's en Dresden's proof. Also, by taking the limit to y for all $y \in \mathbb{C}$ with $\sigma(y) = \infty$ we find that L_B approaches $-\infty$. Hence L_B is bounded and because L_B is a harmonic function, it attains a maximum.

By the proof of the previous theorem we know that

$$\sum_{\sigma\in G}\left(\tfrac{1}{2}\log|\sigma(z)| - \log^+|\sigma(z)|\right) \leq 0.$$

for all $z \in \mathbb{C}$. Equality holds if and only if z is an element α whose whole orbit is contained in the unit circle. We have that $\phi_i(\alpha) = 0$ for some $i \in I$, as there is a $\sigma \in G$ such that p_i is the minimal polynomial of $\sigma(\alpha)$. Hence $L_B(z) \to -\infty$ for $z \to \alpha$ and $B_i > 0$. As $\lim_{B \to 0} L_B(z) \leq 0$ and L is continuous in B, we can find an $0 < A < \frac{1}{2k|I|}$ such that for all sets B with $0 < B \leq A$ we have $L_B(z) < 0$. By maximizing L_B we can find a $D > 0$ such that $L_B(z) \leq -D$ for all $z \in \mathbb{C}$.

Now, let $f \in \mathbb{Z}[x]$ be a polynomial with degree n and roots ξ_i for $i \in \{1, \ldots, n\}$. Assume that $p_i \nmid f_\sigma$ and $z \nmid f_\sigma(z)$ for all $i \in I$ and $\sigma \in G$. Let a_σ be the leading coefficient of f_σ. Note that $\sum_{i \in I} kB_i + \frac{1}{2} < 1$. By summing the inequality $L_B(z) \leq -D$ for $z = \xi_j$ over all $j \in \{1, \ldots, n\}$ and adding

$$\sum_{\sigma \in G} \left(\left(\sum_{i \in I} kB_i + \tfrac{1}{2} \right) \log |a_\sigma| - \log |a_\sigma| \right) \leq 0,$$

we obtain

$$\sum_{i \in I} B_i \log \left| \prod_{\sigma \in G} a_\sigma^k \cdot \prod_{j=1}^{n} \phi_i(\xi_j) \right| + \sum_{\sigma \in G} \left(\tfrac{1}{2} \log \left| a_\sigma \prod_{j=1}^{n} \sigma(\xi_j) \right| \right) - m_G(f) \leq -nD.$$

Per definition of ϕ we then get

$$\prod_{\sigma \in G} a_\sigma^k \cdot \prod_{j=1}^{n} \phi_i(\xi_j) = \prod_{\sigma \in G} \left(a_\sigma^k \cdot \prod_{j=1}^{n} p_i(\sigma(\xi_j)) \right).$$

By Lemma 3.2.10 we know that $a_\sigma^k \prod_{j=1}^{n} p_i(\sigma(\xi_j))$ and $a_\sigma \prod_{j=1}^{n} \sigma(\xi_j)$ are integers. They are non-zero, because $p \nmid f_\sigma$ and $z \nmid f_\sigma(z)$ for all $\sigma \in G$. Hence,

$$\sum_{i \in I} B_i \log \left| \prod_{\sigma \in G} a_\sigma^k \cdot \prod_{j=1}^{n} \phi_i(\xi_j) \right| + \sum_{\sigma \in G} \left(\tfrac{1}{2} \log \left| a_\sigma \prod_{j=1}^{n} \sigma(\xi_j) \right| \right) \geq 0.$$

We conclude that $m_G(f) \geq nD$. □

Theorem 6.5.5. *Let G be a finite subgroup of $\mathrm{PGL}(2, \mathbb{Q})$ of case (ii) or (iii). There exists a lower bound $D > 0$ such that for all polynomials f in $\mathbb{Z}[x]$ of degree n we have that*

$$m_G(f) = 0 \quad \text{or} \quad m_G(f) \geq nD.$$

Proof. As G is of case (ii) or (iii) we have that $\mathcal{O}$ has at most three elements, so $\mathcal{O}$ is finite. If $\mathcal{O}$ is empty, we use Theorem 6.5.2 and else Theorem 6.5.4. These theorems imply there exists a $D > 0$ such that

$$m_G(f) \geq nD$$

for all f of degree n except when f is divisible by z or p_i for all i in some finite set I. As $m_G(f \cdot g) = m_G(f) + m_G(g)$ for $f, g \in \mathbb{Z}[x]$, it is enough to prove inequalities on $m_G(f)$ for irreducible polynomials. Observe that the inequality

$$m_G(f) \geq nD$$

holds for all irreducible f except when f equals one of a finite number of polynomials. As for one of these polynomials f with degree n we have that $m_G(f)/n$ is minimal yet greater than 0, we conclude that there is a constant $D' > 0$, possibly smaller than D, such that

$$m_G(f) = 0 \quad \text{or} \quad m_G(f) \geq nD',$$

concluding this theorem. □

Chapter 7

Concluding Remarks

In this thesis we started with Mahler's measure on polynomials with complex coefficients. Next, we saw Lehmer's unsolved problem for the Mahler measure for polynomials with integer coefficients. As a consequence of Kronecker's lemma, it was possible to determine when the Mahler measure vanished. After that, we found all finite groups of Möbius transformations and we saw Zhang's and Dresden's theorem involving Mahler's measure and two such finite groups of Möbius transformations. In the sixth chapter we generalised Mahler's measure and Lehmer's problem. In Lemma 6.2.4 three cases for this generalised problem are described and for case (ii) and (iii) Problem 6.2.3 is solved in Theorem 6.5.5. Moreover, all transformations of case (i) and (ii) are classified.

Besides the fact that Problem 6.2.3 is still unsolved for G of case (i), there are a few other things to wonder about. Theorem 6.5.5 does not indicate how the optimal value of D, that is a value such that there is a polynomial $f \in \mathbb{Z}[x]$ of degree n such that $m_G(f) = nD$, can be found. In Section 6.4 we found optimal values of D for nearly all cyclic groups of case (ii). However, it would be interesting as well to find more values of D for non-cyclic groups and groups of case (iii). Furthermore it would be interesting to determine a general lower bound for D.

Another interesting thing to do is to replace $\mathbb{Q}$ in $\mathrm{PGL}(2, \mathbb{Q})$ by $\overline{\mathbb{Q}}$, the algebraic closure of $\mathbb{Q}$. The transformation

$$\sigma = \begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}$$

is then another transformation of case (ii). Does in this case Theorem 6.5.5 still hold?

Finally, as suggested by Merlijn Staps, the results for $m_G(f)$ might be useful to approach Lehmer's problem. Namely, if we combine lower bounds of $m_G(f)$ with other results on $m(f)$, for example upper bounds on $m(f)$, it might be possible to solve Lehmer's problem.

Bibliography

[1] Mark Anthony Armstrong. *Groups and symmetry*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1988.

[2] Werner Ballmann. *Lectures on Kähler manifolds*. ESI Lectures in Mathematics and Physics. European Mathematical Society (EMS), Zürich, 2006.

[3] Gregory Dresden. Orbits of algebraic numbers with low heights. *Math. Comp.*, 67(222):815–820, 1998.

[4] Gregory Dresden. There Are Only Nine Finite Groups of Fractional Linear Transformations with Integer Coefficients. *Math. Mag.*, 77(3):211–218, 2004.

[5] Gertrude Ehrlich. *Fundamental Concepts of Abstract Algebra*. Dover Publications, Inc., Mineola, N.Y., 2011.

[6] Graham Everest and Thomas Ward. *Heights of polynomials and entropy in algebraic dynamics*. Universitext. Springer-Verlag London, Ltd., London, 1999.

[7] Aleksandr Osipovič Gel′fond. *Transcendentnye i algebraičeskie čisla*. Gosudarstv. Izdat. Tehn.-Teor. Lit., Moscow, 1952.

[8] Johan Jensen. Sur un nouvel et important théorème de la théorie des fonctions. *Acta Math.*, 22(1):359–364, 1899.

[9] Felix Klein. *Lectures on the icosahedron and the solution of equations of the fifth degree*. Dover Publications, Inc., New York, N.Y., revised edition, 1956. Translated into English by George Gavin Morrice.

[10] Leopold Kronecker. Zwei Sätze über Gleichungen mit ganzzahligen Coefficienten. *J. Reine Angew. Math.*, 53:173–175, 1857.

[11] Serge Lang. *Complex analysis*, volume 103 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, fourth edition, 1999.

[12] Derrick Henry Lehmer. Factorization of certain cyclotomic functions. *Ann. of Math. (2)*, 34(3):461–479, 1933.

[13] Kurt Mahler. An application of Jensen's formula to polynomials. *Mathematika*, 7:98–100, 1960.

[14] Kurt Mahler. On some inequalities for polynomials in several variables. *J. London Math. Soc.*, 37:341–344, 1962.

[15] Peter M. Neumann. A lemma that is not Burnside's. *Math. Sci.*, 4(2):133–141, 1979.

[16] Joseph Rotman. *Galois theory*. Universitext. Springer-Verlag, New York, second edition, 1998.

[17] René L. Schilling. *Measures, integrals and martingales*. Cambridge University Press, New York, 2005.

[18] Jerry Shurman. *Geometry of the quintic*. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1997.

[19] Don Zagier. Algebraic numbers close to both 0 and 1. *Math. Comp.*, 61(203):485–491, 1993.

[20] Shouwu Zhang. Positive line bundles on arithmetic surfaces. *Ann. of Math. (2)*, 136(3):569–587, 1992.

[21] Antoni Zygmund. *Trigonometric series. Vol. I, II*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2002. With a foreword by Robert A. Fefferman.