VILNIUS UNIVERSITY

JONAS JANKAUSKAS

HEIGHTS OF POLYNOMIALS

Doctoral dissertation
Physical sciences, mathematics (01P)

Vilnius, 2012

The scientific work was carried out in 2008–2012 at Vilnius University

Scientific supervisor:

prof. dr. habil. Artūras Dubickas (Vilnius University, Physical sciences, Mathematics - 01P)

Scientific adviser:

doc. dr. Paulius Drungilas (Vilnius University, Physical sciences, Mathematics - 01P)

VILNIAUS UNIVERSITETAS

JONAS JANKAUSKAS

POLINOMŲ AUKŠČIAI

Daktaro disertacija
Fiziniai mokslai, matematika (01P)

Vilnius, 2012

Disertacija rengta 2008–2012 metais Vilniaus universitete.

Mokslinis vadovas:

prof. habil. dr. Artūras Dubickas (Vilniaus universitetas, fiziniai mokslai, matematika - 01P)

Konsultantas:

doc. dr. Paulius Drungilas (Vilniaus universitetas, fiziniai mokslai, matematika - 01P)

He turned the handle and the door opened.
Beyond it was another door.
He turned the handle again and the other door stood wide.
In this way, he opened one hundred twenty four doors.
Then he grew tired and he collapsed.
”Behind the one hundred twenty fifth door, there lies a garden,
roses have just begun to bloom there“ - he thought,
drowsily dying.
Behind the door, there was another door.

A. Škėma

Notation

$\mathbb{N}$ – the set of all positive integers

$\mathbb{Z}$ – the set of all integers

$\mathbb{Q}$ – the set of all rational numbers

$\mathbb{R}$ – the set of all real numbers

$\mathbb{C}$ – the set of all complex numbers

$\mathbb{Z}[x]$ – the set of polynomials with integer coefficients in one variable x

$\mathbb{Q}[x]$ – the set of polynomials with rational coefficients

$\mathbb{R}[x]$ – the set of polynomials with real coefficients

$\mathbb{C}[x]$ – the set of polynomials with complex coefficients

i – a complex number, satisfying $i^2 = -1$, unless used as an index in subscript or superscript.

$H(P)$ – the height of a polynomial P

$L(P)$ – the length of a polynomial P

$M(P)$ – the Mahler measure of a polynomial P

$||P||$ – the Euclidean norm of a polynomial P

$||P||_s$ – the L^s norm of a polynomial P, taken on the unit circle $|z| = 1$ in the complex plane

P^* – a reciprocal polynomial of a polynomial $P(x)$, defined by the equation $P^*(x) := x^{\deg(P)} P(1/x)$

$|z|$ – the absolute value of a complex number z

$[x]$, $\lfloor x \rfloor$ – the integer part of a real number x

Brackets $\{\dots\}$ are used to denote a set. Brackets $\{x\}$ containing a single number $x \in \mathbb{R}$ mean the fractional part of x. It shall be clear from the context in which they will be used

$\mathfrak{U}_n$ – unimodular polynomials of degree n (polynomials $P \in \mathbb{C}[x]$ whose coefficients are of modulus 1)

$\mathcal{L}_n$ – Littlewood polynomials of degree n (polynomials $P \in \mathbb{R}[x]$ whose coefficients $a_j \in \{-1, 1\}$)

$\mathcal{N}_n$ –Newman polynomials of degree n (polynomials $P \in \mathbb{R}[x]$ with coefficients $a_j \in \{0, 1\}$ and $P(0) \neq 0$)

Remarks

A perfectionist reader might be surprised by some small changes in notation through the thirteen chapters of the thesis. To avoid any confusion, we warn that the notation will be slightly dependent on the context of a mathematical problem considered in the respective chapters. The polynomials will be denoted by uppercase letters P, Q, R, G, H, ... or lowercase p, q, f, g, h, ... We use x for the variable of a real polynomial in $\mathbb{R}[x]$, or for a polynomial with coefficients taken from any (abstract) field K. The variable z is used when we want to emphasize the polynomial as a function of a complex variable $z \in \mathbb{C}$. The degree of a polynomial shall be denoted by d or n. The former notation of degree is preferred by number theorists an algebraists, the latter – by analysts and engineers.

Contents

Chapter 1

Introduction

1.1 The notion of height

Most of the mathematical research presented in this doctoral dissertation is centered around the notion of the *height* of a polynomial. We consider polynomials

$$P(x) = a_n x^n + \cdots + a_1 x + a_0$$

in one variable x with real or complex coefficients. If the leading coefficient $a_n \neq 0$, then the integer $n \geqslant 0$ is called *the degree* of the polynomial P. Polynomials with all coefficients in $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$ or $\mathbb{Z}$ are called *complex*, *real*, *rational* or *integer* polynomials, respectively.

The height of a polynomial, in the most general sense, is a quantity by which we measure the *complexity* of the polynomial P. There are several different types of heights. The most simple heights take into account the size of the coefficients of a polynomial. Such are the *naive height*,

$$H(P) = \max\{|a_0|, |a_1|, \ldots, |a_n|\},$$

the *length*:

$$L(P) = |a_0| + |a_1| + \ldots |a_n|,$$

and the *Euclidean norm* of a polynomial

$$||P|| = \left(|a_0|^2 + |a_1|^2 + \cdots + |a_n|^2\right)^{1/2}.$$

Another way to define the height is to consider the size of the roots of the polynomial P. The Fundamental Theorem of Algebra asserts that P splits in $\mathbb{C}[x]$ into

the product of linear factors

$$P(x) = a_n(x - \alpha_1)(x - \alpha_2)\dots(x - \alpha_n),$$

where α_1, α_2, $\dots$, α_n are the complex roots of P, not necessarily distinct (see Theorem 5.1 in Lang's book [129]). For such a factorization, the *Mahler measure* $M(P)$ is defined by the formula

$$M(P) = |a_n| \prod_{j=1}^{n} \max\{1, |\alpha_j|\}$$

as the product of of absolute values of roots α_j of modulus greater than one and the leading coefficient of P.

One more type of height measures the analytic behavior of the polynomial P in some compact subset of the complex plane $\mathbb{C}$. In the present thesis we consider the mean values of the polynomial P on the circle of radius 1, centered at the point $z = 0$. For arbitrary real number $s > 0$, the integral mean value is defined by the formula

$$||P||_s = \left(\frac{1}{2\pi} \int_0^{2\pi} \left| P(e^{it}) \right|^s \, dt \right)^{1/s}.$$

For the values $s \geqslant 1$, the mean value $||P||_s$ coincides with the norm of a polynomial P in the space L^s of functions which are absolutely s – integrable on the unit circle in the sense of Lebesgue. For $s = 0$, one defines the geometric mean of a polynomial $||P||_0$ by the formula

$$||P||_0 = \exp\left(\frac{1}{2\pi} \int_0^{2\pi} \log \left| P(e^{it}) \right| \, dt \right).$$

For historical reasons we will refer to the quantity $||P||_s$ as a L^s norm for all the values $s < 1$. However, one should keep in mind that $||P||_s$ is not a true norm of the vector space L^s for $s < 1$. One also defines the *infinity norm* $||P||_\infty$ of a polynomial by setting

$$||P||_\infty = \max_{t \in [0, 2\pi)} |P(e^{it})|.$$

It is known (see [102]) that the norm $||P||_s$ (for a fixed polynomial P) considered as a function of s, is non-decreasing and continuous in the interval $s \in (0, +\infty)$, and that

$$\lim_{s \to 0^+} ||P||_s = ||P||_0, \qquad \lim_{s \to +\infty} ||P||_s = ||P||_\infty.$$

In general, the norms $||P||_s$ are hard to calculate explicitly. For $s = 2$, the *Parseval's identity* yields

$$||P||_2 = ||P||.$$

This is the only easily computable case. For $s = 0$, it is known that

$$||P||_0 = M(P).$$

The equality of $||P||_0$ norm and the Mahler measure of P was established by Mahler himself [139] using *Jensen's formula*. There are various ways to prove it – see, for instance, Chapter 1 in [83], or [197].

There exist more heights for the polynomials of several variables, such as the *Bombieri norm* or *multivariate Mahler measures*. In the present thesis we do not delve in to these and keep focused on the univariate setting.

1.2 Applications

Heights of polynomials play a substantial role in the modern number theory. In Diophantine analysis heights are used to bound the rate of convergence of rational approximations to algebraic numbers. Heights are important technical tools in proving explicit lower bounds for linear forms in logarithms of algebraic numbers. Heights are useful in bounding the size of integer solutions to Diophantine equations or giving upper bounds on the total number of such solutions. For number theorists, the most useful heights are the naive height $H(P)$, the length $L(P)$, the Mahler measure $M(P)$ and the logarithmic variants of Mahler measure.

Heights which measure analytic properties of polynomials (mainly, the norms $||P||_s$) are important in various branches of mathematical analysis, such as the approximation theory in L^s spaces, theory of Fourier series, functional analysis – and this list is very far for being complete. Outside the realm of pure mathematics, heights of polynomials have applications in electric engineering and signal processing theory, where they are used to measure signal strength and energy.

1.3 Aims and problems

In this section we summarize the primary directions of the research presented in this thesis. Mathematical problems which received most attention in the course of the doctoral research are formulated here.

- **Height reduction in $\mathbb{R}[x]$.** In Chapter 3 we will study the following problem: given a polynomial $P(x) \in \mathbb{R}[x]$, find another real polynomial $Q(x)$, whose coefficients are as small as possible, such that $Q(x)$ is divisible by $P(x)$. This problem originates in Diophantine analysis. We will introduce

the reduced height of a polynomial and give results on the properties and the explicit computation of this quantity.

- **Maximal values of polynomials.** In Chapter 4, the maximal values of polynomials with real coefficients restricted to the interval $[-1, 1]$ on the complex unit circle $|z| = 1$ will be studied. We shall calculate exact and asymptotic formulas for the maximal values and outline how the main result may be generalized for polynomials with coefficients in arbitrary intervals $[a, b]$.

- **Newman and Littlewood numbers.** We will investigate the sets of complex numbers which are the zeros of polynomials with small integer coefficients in Chaper 5. The sets of roots of polynomials with coefficients $\{-1, 1\}$ will be called *Littlewood numbers* and denoted $V_{\mathcal{L}}$. Roots of polynomials with coefficients $\{0, 1\}$ will be called *Newman numbers*. The set of Newman numbers is denoted $V_{\mathcal{N}}$. We will prove non-trivial inclusion of Newman numbers of low degrees into the set $V_{\mathcal{L}}$ and find examples of Newman numbers which are not Littlewood numbers. Most of our results there rely on specially tailored fast computer algorithms. The divisibility of Littlewood polynomials by trinomials and quadrinomials of height 1 will be also studied.

- **Mahler measures of a polynomial and its derivative.** We shall study the inequality $M(f') > d/2M(f)$ for self-inversive polynomials $f \in \mathbb{C}[z]$. This inequality is a counterpart to the classical inequality $M(f') < dM(f)$, established by Mahler [140] himself. Following the remark of anonymous Referee in Chapter 6 we will give a slightly refined version of this classical inequality investigated in [70]. We also calculate the exact minimum of the ratio of Mahler measures $M(f')/M(f)$ for self-inversive cubic polynomials.

- **Numbers in geometric and arithmetic progressions.** Given an arbitrary geometric progression $\mathcal{G}$ of real numbers, how many of them lie in some arithmetic progression $\mathcal{A}$? This problem is connected to the question of equal values in the sequence of fractional parts of powers $\{\xi\alpha^n\}$, $n = 1, 2, \ldots$ of an algebraic number α. In Chapter 7, we will extend classical results of Ehlich [81], Supnick, Keston, Cohn [205], Posner and Rumsey [165]. It turns out that the answer to this problem depends on the geometrical and arithmetical properties of complex numbers which are the roots of polynomials having only three or four non-zero terms.

- **Factorization of quadrinomials.** In Chapter 8, we investigate the reducibility of quadrinomials of the form $x^i + x^j + x^k + 4$. We give an answer

to the question of Walsh which irreducible quadrinomials of this form split in $\mathbb{Z}[x]$ after the variable x is replaced by some power x^l.

- **Number systems.** We consider number systems in the ring $\mathbb{Z}[\alpha]$ for an arbitrary expanding algebraic integer α. An algebraic number α is called *expanding*, if all its conjugates are of modulus > 1. Our aim is an effective construction of the number system $(\mathcal{B}, \alpha)$ with the smallest possible digit set $\mathcal{B}$. In Chapter 9, we will show that this can be done efficiently in the case where all conjugates of α are of sufficiently large modulus. We will also give an effective construction for all expanding quadratic integers and also for the expanding integers of degree 3 and trace 0.

- **Metric Mahler measures.** In Chapter 10 we will study *metric Mahler measures*. This concept was introduced by Dubickas and Smyth in [74] in order to give a notion of distance in the multiplicative group of the field $\overline{\mathbb{Q}}$. We shall study a variant of these metric Mahler measures, called *t-metric Mahler measures*. We will determine when the infimum of t-metric Mahler measure $M_t(\alpha)$ can be achieved in the field $\mathbb{Q}(\alpha)$, provided that α is a rational number or a quadratic surd of some square-free positive integer. Our work complements earlier results of Samuels [175], [176], [177].

- **Barker sequences and polynomials.** Chapter 11 is devoted to research on the class of Laurent polynomials denoted $\mathcal{L}P_n$. This class consists of polynomials with small coefficients c_j, for $-n \leqslant j \leqslant n$, $j \neq 0$, and a large central coefficient c_0. Such polynomials arise in Signal Processing theory in connection to *aperiodic autocorrelations of binary sequences*. We have studied extremal extremal Mahler measures and L^s norms of such polynomials in the hope of finding an alternative approach to a long standing *Barker conjecture* in [31], [32]. We have determined extremal polynomials in the class $\mathcal{L}P_n$ and established non-trivial lower bounds for their Mahler measures $M(P)$ and L^s norms.

- **Composition equation in polynomials.** Chapter 12 is devoted to the solution of the composition equation $f(g(x)) = f(x)h^m(x)$ in unknown polynomials $f(x)$, $g(x)$, $h(x)$ with coefficients in an arbitrary field K. We will provide a complete solution to this polynomial equation in the case when the polynomial $f(x)$ is separable (has no multiple roots) and the integer $m \geqslant 2$ is not divisible by the characteristic of the field K. The composition equation we solve is of some interest in the context of the conjecture of Cassaigne et al. [54] on the sign changes in the sequence $\lambda(f(n))$, $n \in \mathbb{N}$ for

Louiville's lambda function λ evaluated at integer values n of polynomials $f(x) \in \mathbb{Z}[x]$.

1.4 Methods

One needs a variety of different methods in order to solve problems related to polynomials.

In computing reduced heights, we will need standard linear algebra, determinant theory [121], [167] and theorems from the linear optimization [18]. To determine maximal values of polynomials, we use vector representations of complex numbers, cosine sum formulas and the theorem of Weil [123] on the uniform distribution of fractional parts of complex arguments $\{m\phi/2\pi\}$ in the interval $[0, 1)$.

The investigations of the sets of Newman and Littlewood numbers require fast computer algorithms to check which numbers are zeros of Newman and Littlewood polynomials, and which Newman polynomials divide some Littlewood polynomial. We will develop two different efficient recursive algorithms for these purposes. The algorithms were implemented in C++ using specialized libraries for multi-precision computing [26], [100] and interval arithmetic [104]. In order to establish the results on the divisibility properties of trinomials and certain quadrinomials with coefficients $\{-1, 0, 1\}$, we will use the factorization of polynomials modulo 2 in the ring $\mathbb{F}_2[x]$. In our constructions we use theorems of Filaseta and other authors on the irreducibility of certain Newman polynomials.

Inequalities for the Mahler measure of derivative of self-inversive polynomials are based on the formula for the real and imaginary parts of the rational function $zf'(z)/f(z)$ on the unit circle, combined with Jensen's formula [83] and an integral version of Mahler's inequality [102]. In cubic case, we compute the precise minimal value of Mahler measures (a non-trivial optimization problem) by calculating critical points of partial derivatives, aided by the MAPLE [215] computer algebra package.

To explore real numbers in arithmetic and geometric progressions, or, alternatively, equal values of fractional parts of powers, one uses the techniques on the geometry of complex zeros of trinomial and quadrinomial equations. Results stated in Chapter 7 are obtained as a combination of geometric methods [81], [165], [205], Galois theory and estimates on the maximal zero multiplicity of non-degenerate linear recurrence sequences proven by Beukers [25].

In the problem of Walsh on irreducibility of quadrinomials $x^i + x^j + x^k + 4$, we shall use Ljunggren's [137] method to study the reducibility of non-reciprocal

part of polynomials with small Euclidean norm (in particular, trinomials and quadrinomials of height one).

For the construction of number systems $(\mathcal{B}, \alpha)$ in rings $\mathbb{Z}[\alpha]$, we use various criteria on the coefficients of polynomials with all roots outside the unit circle. We prove inequalities for the size of the digits in the set $\mathcal{B}$ in terms of the modulus of the smallest algebraic conjugate of α and the discriminant of polynomial. Finite automata theory [178] is used in the construction for the cubic case.

Number theoretical methods, such as the divisibility properties of integers and the formula for the norm $N(\alpha)$ of algebraic numbers in quadratic fields [156], are used to investigate t-metric Mahler measures.

In order to study extremal Mahler measures and L^s norms of polynomials related to Barker sequences, we go back to the realms of classical analysis. We use logarithmic and binomial Taylor series, the Dirichlet kernel $D_N(t)$ and sine phasor formulas in addition to the Weierstrass uniform convergence criteria in Chapter 11. Rudin's book [173] is considered to be a classical reference on the subject. Jensen's formula will be one of the main tools here.

1.5 Originality

Most of the results presented in this doctoral thesis are completely new and have not appeared before in the scientific literature. Some of the results are extensions or applications of previously known results in new areas, so they are also a valuable addition to existing theory.

Despite considerable progress in the theory of polynomials, many problems and conjectures are still wide open. We hope that the present thesis will be a step towards a more general mathematical theory.

1.6 Dissemination of results

The results of this thesis were presented in the following conferences:

- *27th Journées Arithmétiques*, Vilnius, Lithuania, June 27 – July 1, 2011.
- *Heights 2011*, Tossa de Mar, Spain, April 25 – 30, 2011.
- *PIMS/SFU/UBC Number Theory Seminar*, University of British Columbia, Vancouver, Canada,October 7, 2010.
- *19–th Czech and Slovak International Conference on Number Theory*, Hradec nad Moravicí, Czech Republic, August 31 – September 4, 2009.

- *Explicit Methods in Number Theory*, Institute of Mathematics, University of Debrecen, Debrecen, Hungary, January 26 – 30, 2009.
- *Number Theory seminar*, The School of Mathematics, University of Edinburgh, Edinburgh, Scotland, November 5, 2008.
- *Journeé Diophantienne*, L'Institut de Recherche Mathématique Avancée, L'Université de Strasbourg, Strasbourg, France, October 23, 2008.
- *International Conference on Number Theory, dedicated to the 60th birthday of A.Laurinčikas*, August 11 – 15, 2008, Šiauliai, Lithuania.

The results of the thesis were also presented and approved in the mathematical seminar of Department of Probability Theory and Number Theory on May 14, 2012 and joint seminar of Lithuanian Mathematical Society, held on May 28, 2012 at the Faculty of Mathematics and Informatics of Vilnius University in honor of the late J. Kubilius.

1.7 Publications

1.7.1 Principal publications

The results of the doctoral research will appear in 11 research papers. Eight of them have already been published, three are accepted for publishing in general or specialized periodical peer reviewed foreign mathematical journals. 10 out of 11 papers will appear in journals indexed by *ISI Web of Science*.

Published papers:

1. A. Dubickas, J. Jankauskas, *On the reduced height of a polynomial*, Publ. Math. Debrecen, **71** (6) (2007), 325–348.
2. A. Dubickas, J. Jankauskas, *The maximal value of polynomials with restricted coefficients*, Journal of the Korean Mathematical Society, **46** (1) (2009), 41–49.
3. A.Dubickas, J.Jankauskas, *On the intersection of infinite geometric and arithmetic progressions*, Bull. of the Brazilian Math. Soc., **41** (4) (2010), 551–566.
4. A.Dubickas, J.Jankauskas, *On Newman polynomials which divide no Littlewood polynomial*, Mathematics of Computation, **78** (265) (2009), 327–344.

5. A.Dubickas, J.Jankauskas,
   On Mahler measures of a self-inversive polynomial and its derivative,
   Bull. London Math. Soc., **42** (2) (2010), 195–209.

6. J. Jankauskas,*On the reducibility of certain quadrinomials*,
   Glasnik Matematički, **45** (65) (2010), 31–41.

7. J. Jankauskas, C. Samuels,
   The t-metric Mahler measures of surds of rational numbers,
   Acta Math. Hungar., **134** (4) (2012), 481–498.

8. P. Borwein, S. K. K. Choi and J.Jankauskas,
   On a class of polynomials related to Barker sequences,
   Proceedings of Amer. Math. Soc., **140** (8) (2012), 2613–2625.

Accepted for publication:

1. P. Borwein, S. K. K. Choi and J.Jankauskas,
   Extremal Mahler measures and L_s norms in the class of polynomials related to Barker sequences,
   Proceedings of Amer. Math. Soc.

2. S. Akiyama, P. Drungilas and J. Jankauskas,
   Height reducing problem on algebraic integers,
   Funct. Approx. Comment. Math.

3. H. Ganguli, J. Jankauskas,
   On the equation $f(g(x)) = f(x)h^m(x)$ for composite polynomials,
   J. Aust. Math. Soc. (special issue dedicated to Alfred van der Poorten).

1.7.2 Conference abstracts:

1. A. Dubickas, J. Jankauskas,
   On the intersection of infinite geometric and arithmetic progressions,
   27–th Journées Arithmétiques, June 27 – July 1, 2011, Vilnius, Lithuania: programme and abstract book. Vilnius, Vilniaus universitetas, 2011. Available online at
   `http://atlas-conferences.com/cgi-bin/abstract/cbbv-46`.

2. S. Akiyama, P. Drungilas, J. Jankauskas,
   Height reducing in number systems,
   27–th Journées Arithmétiques, June 27 – July 1, 2011, Vilnius, Lithuania:

programme and abstract book. Vilnius, Vilniaus universitetas, 2011. Available online at
http://atlas-conferences.com/cgi-bin/abstract/cbbv-68.

3. S. Akiyama, P. Drungilas, J. Jankauskas,
Aukščio mažinimas skaičiavimo sistemose,
Fizinių ir technologijos mokslų tarpdalykiniai tyrimai: pirmosios LMA Jaunųjų mokslininkų konferencijos pranešimų santraukos, Vilnius, 2011 02 08. Vilnius, LMA leidykla, 2011. Available online at
http://lma.lt/media/k2/attachments/SANTRAUKOS.pdf.

4. P. Borwein, S. Choi, J. Jankauskas,
On a class of polynomials related to Barker sequences,
Abstracts of talks given by doctoral students at *Heights 2011* conference, Tossa de Mar, Spain, April 29, 2011. Available online at
http://www.imub.ub.es/heights2011/Abstracts_young.pdf

1.8 Acknowledgments

I would like to thank my advisor, Professor Artūras Dubickas. I have benefited from his knowledge and insight, especially at the early stages of my mathematical work. Professor Dubickas will always be my personal example of a professional mathematician.

I am grateful to Dr. Paulius Drungilas, who was my mentor and a good friend during my studies at Vilnius University. With his help and support, difficult things could be handled more easily.

None of my scientific research would be possible without the help of my collaborators: Professor Peter Borwein and Professor Stephen Kwok-Kwong Choi at Simon Fraser University, Professor Shigeki Akiyama at Niigata University, my colleagues Dr. Charles Samuels, and Himadri Ganguli.

I thank Professor Christopher J. Smyth for a careful review of my thesis and a large number of corrections. I thank Professor Tamás Erdélyi for his comment which led to the shortened proof avoiding the monotone convergence argument in Theorem 11.13 of Chapter 11. I thank Professor Michael Mossinghoff for his remark on the case $s \in (2, 3)$ of Theorem 11.15. I am also grateful to Professor A. Schinzel for his note on the total number of irreducible factors of quadrinomials $x^i + x^j + x^k + n$ in Chapter 8. I am grateful to Sonata Juškevičienė for helping me improve English in my thesis. My coauthors and I are also grateful to the anonymous referees for their useful comments and corrections of our research papers.

In the course of my doctoral studies I benefited from several research visits:

- The 5 month research visit to Professor Peter Borwein at IRMACS, Simon Fraser University, Burnaby, British Columbia, Canada, from September 1, 2010 to January 30, 2011. My stay in Canada was funded by Lithuanian Science Council.

- The winter school *Explicit Methods in Number Theory*, Institute of Mathematics, University of Debrecen, Debrecen, Hungary, January 26–30, 2009. My visit was funded by Vilnius University.

- One week visit to Professor Y. Bugeaud, L'Institut de Recherche Mathématique Avancée, L'Université de Strasbourg, Strasbourg, France, October 20–25, 2008. The visit was supported by GILIBERT programme.

- One week visit to Professor C. J. Smyth, School of Mathematics, University of Edinburgh, Edinburgh, Scotland, November 2–8, 2008. The visit was funded by Vilnius University.

I would like to thank Dr. Antanas Apynis for making me interested in mathematics at the time I was studying at Pasvalys Petras Vileišis Gymnasium and keeping me involved in mathematical activities in Pasvalys. Without him, I would not be studying mathematics in Vilnius University.

I am grateful to Dr. Hamletas Markšaitis and Professor R. Garunkštis for their wonderful mathematical seminars at the Department of Probability Theory and Number Theory. During these seminars, I was exposed to the beauty of abstract algebra and elliptic curves.

I would like to thank Dr. Romualdas Kašuba for his optimism and unprecedented sense of humor. I have learned from him how to make my teaching work interesting (both for me and my students).

I thank my colleagues at Vilnius University, Faculty of Mathematics and Informatics, who made my studies and life in Vilnius interesting (in a good way): Arūnas Vareika, Aivaras Novikas, Mindaugas Skujus, Jonas Šiurys, Paulius Šarka, Albertas Zinevičius, Matas Šileikis, Justas Kalpokas, Donata Puplinskaitė, Jurgita Markevičiūtė, Marius Grigelionis. I am grateful to Žymantas Darbėnas and Tomas Juškevičius who kept in touch with me during these years. I thank Ron Ferguson for his hospitality in Vancouver.

And, last but not least, I must note that I am heavily in debt to my aunt Jūratė and uncle Romas Olšauskai who supported me while living in Vilnius for 9 years. Thanks to them, I could dedicate most of my time to mathematics.

Chapter 2

Literature review

Heights of polynomials are not lonely and isolated objects of mathematical curiosity. On the contrary, there exist important mathematical applications for them. In this section, we give a short survey of some of these applications. We also review several famous conjectures which are central for the theory of heights of polynomials. Most of the information presented in this Chapter is a condensed version of several surveys and books, including: surveys on Mahler measure and Lehmer's problem by Smyth [201] and the book of Everest and Ward [83], surveys on the Merit factor problem by Borwein, Fergusson, Knauer [36] and Jedwab [107]. The book [29] by Borwein is the ultimate reference on all problems concerning polynomials with small coefficients $\{-1, 0, 1\}$. For in-depth study of the subject, the reader should refer to these sources. For general references on polynomials, we can recommend [34] and [171].

2.1 Algebraic number theory

Recall that a number $\alpha \in \mathbb{C}$ is called *an algebraic number* if α is a root of some polynomial $P(x)$ with integer coefficients. Among all the integer polynomials $P \in \mathbb{Z}[x]$ satisfying the equation $P(\alpha) = 0$, there exists a polynomial of minimal degree which is irreducible in $\mathbb{Z}[x]$ and whose leading coefficient is a positive number. Such polynomial $P(x)$ is called *the minimal polynomial* of α in $\mathbb{Z}[x]$. The degree of the minimal polynomial $P(x)$ is called the degree of an algebraic number α. Usually, it is denoted $d = \deg(\alpha)$. In particular, if $P(x)$ has leading coefficient 1, then α is an *algebraic integer*. By the Galois theory, the roots of the minimal polynomial $P(x)$ are *algebraic conjugates* of α over the field of rational numbers $\mathbb{Q}$. The minimal polynomial $P(x)$ carries all the information on the number α and its conjugates, such as the size of the conjugates, their distribution in the complex plane and their arithmetical properties. The Mahler measure of α

is defined as the Mahler measure of its minimal polynomial:

$$M(\alpha) = M(P).$$

The Mahler measure $M(\alpha)$ is useful in proving Kronecker's theorem [122], which states that all algebraic integers α having all conjugates α' of modulus 1 are roots of unity; this result can be also derived as a consequence of the inequality

$$|a_k| \leqslant \binom{d}{k} M(\alpha), \quad d = \deg \alpha. \tag{2.1}$$

for the size of coefficients a_k of the minimal polynomial $P(x)$ (see [197]). This inequality was proved by Mahler himself [139]. Another inequality, also due to Mahler [142], bounds the size of the discriminant of $disc(P) = a_d^{2d-2} \prod_{1 \leqslant i < j \leqslant d} (\alpha_i - \alpha_j)^2$ by

$$|disc(P)| \leqslant d^d M(P)^{2d-2}.$$

In 1933, Lehmer in his (now famous) paper [132] asked whether there exists an absolute constant $C > 1$, such that, for any non-zero algebraic number α which is not a root of unity (that is, $M(\alpha) \neq 1$) one has $M(\alpha) > C$. The smallest Mahler measure that Lehmer could find was $M(\alpha) = 1.17628\ldots$. This measure is attained by the number α which is the root of the polynomial

$$P(x) = x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1. \tag{2.2}$$

It is rather surprising that, to date, no numbers $\alpha \neq 0$ other than the roots of unity, are known having Mahler measure smaller than that one found by Lehmer. Algebraic numbers like the one found by Lehmer are called *reciprocal*. They are characterized by the property that α and $1/\alpha$ are algebraic conjugates. Algebraic numbers which do not satisfy this property are called *non-reciprocal*. The celebrated theorem of Smyth [199] states that for non-zero non-reciprocal algebraic numbers one has

$$M(\alpha) \geqslant M(x^3 - x - 1) = 1.32471\ldots. \tag{2.3}$$

A weaker inequality $M(\alpha) > 1.1796\ldots$ was proved by Breusch [48].

In the case where all the conjugates of α are real, Schinzel [180] proved that

$$M(\alpha) \geqslant \left(\frac{1+\sqrt{5}}{2}\right)^{d/2},$$

provided that $\alpha \neq 0, -1, 1$.

The best general result towards Lehmer conjecture is a theorem of Dobrowolski [60], subsequently improved by Cantor and Strauss [53], Rausch [168], Louboutin [138] and Voutier [211]. Namely, for all α with $M(\alpha) \neq 1$, the inequality

$$M(\alpha) > 1 + \frac{1}{4}\left(\frac{\log\log d}{\log d}\right)^3$$

holds. The Lehmer conjecture, however, still remains unsolved, since the above bound depends on $d = \deg(\alpha)$ and tends to 1 as $d \to \infty$. A huge amount of computations, performed by Boyd [43], [44] and later by Mossinghoff, Rhin and Wu [153] produced a lot of supporting evidence towards the Lehmer's conjecture. See a nice survey of Smyth [201] on the current state of Lehmer problem and more applications of Mahler measure. The first and third chapters in the book of Everest and Ward [83] can be also recommended.

2.2 Diophantine analysis

Heights of polynomials turn out to be useful to measure the rate of convergence of rational approximations to algebraic numbers. Historically, the earliest example is Liouiville's Approximation Theorem, which states that, roughly, an irrational algebraic number cannot be approximated by any rational number too well. A precise statement of the theorem is as follows: for an algebraic number α of degree $d \geqslant 2$ and coprime integers $p \in \mathbb{Z}$, $q \in \mathbb{N}$, the inequality

$$\left|\alpha - \frac{p}{q}\right| > \frac{C(\alpha)}{q^d}$$

holds for some positive constant $C(\alpha)$ which depends on α only. It is possible to give an explicit formula for $C(\alpha)$ in terms of various heights. For instance, one can take $C(\alpha) = 1/(d^2(1+|\alpha|)^{d-1}H(P))$, where P is the minimal polynomial of α – see [202].

A more modern and more general result is stated in the Waldschmidt book [212]: for any polynomial $Q \in \mathbb{Z}[x]$, which does not vanish on α, one has

$$|Q(\alpha)| > \frac{1}{M(\alpha)^N L(Q)^{d-1}},$$

where N is the degree of Q and $L(Q)$ is the length of Q.

The Mahler measure is used in the theory of logarithms of algebraic numbers, most often in the form of the *Weil height* (see [197] or [212], for instance). The

Weil height (or *the absolute height*) of α is defined by the formula

$$h(\alpha) = \frac{\log M(\alpha)}{d}.$$

In particular, the Weil height provides upper bounds for the modulus of integers $t_1, t_2, \ldots, t_k$ which satisfy the equation

$$t_1 \log \beta_1 + \ldots t_k \log \beta_k = 0.$$

Here $\beta_1, \beta_2, \ldots, \beta_k$ are multiplicatively dependent algebraic integers, see [212]. Another recent example is related to the length $L(P)$ of a polynomial P. Dubickas [65] proved the following inequality for the distance between upper and lower limit points in the sequence of fractional parts of powers $\{\xi\alpha^n\}$, $n = 1, 2, \ldots$ of an algebraic number:

$$\Delta(\xi, \alpha) = \limsup_{n\to\infty}\{\xi\alpha^n\} - \liminf_{n\to\infty}\{\xi\alpha^n\} \geqslant \frac{1}{l(\alpha)}$$

The quantity $l(\alpha)$ is called *the reduced length* of α. It is defined by the formula

$$l(\alpha) := \inf_{\substack{Q\in\mathbb{R}[x],\\ Q\text{– monic or } Q(0)=1}} L(PQ).$$

The reduced length was studied in detail by Schinzel [187], [188], [189].

2.3 Number systems and tilings

Let α be an algebraic integer with minimal polynomial $P(x) \in \mathbb{Z}[x]$. A non-empty, finite integer set $\mathcal{B}$,

$$\mathcal{B} \subset \mathbb{Z}, \qquad \mathcal{B} \neq \varnothing, \qquad |\mathcal{B}| < \infty$$

is called *a digit set* in the ring $\mathbb{Z}[\alpha]$, if each element $\beta \in \mathbb{Z}[\alpha]$ has a finite expression of the form

$$\beta = b_0 + b_1\alpha + \cdots + b_m\alpha^m$$

with all coefficients (digits) b_j in set $\mathcal{B}$, $0 \leqslant j \leqslant m$. In other words, $\mathbb{Z}[\alpha] = \mathcal{B}[\alpha]$. In such a case, the pair $(\mathcal{B}, \alpha)$ is called *a number system*. If all digits in $\mathcal{B}$ are positive, then $(\mathcal{B}, \alpha)$ is called *a canonical number system* (CNS for short). Such number systems $(\mathcal{B}, \alpha)$ are the generalizations of usual numbers systems in $\mathbb{N}$ with positive integer bases. D. Knuth [117] considered canonical number systems in $\mathbb{Z}[-1 + i]$ which are related to fractals. The CNS in the ring of Gaussian integers

were investigated systematically for the first time by Kátai, Szabó [113]; these results were generalized to the rings of quadratic integers by Kátai, Kovács,[111], [112] and Gilbert [99]. The algorithms to decide if $(\mathcal{B}, \alpha)$ is a CNS and to compute expansions of elements in $\mathbb{Z}[\alpha]$ were devised by Kovács and Pethő [120], Akiyama and Pethő [6], Scheicher and Thuswaldner [179].

It is easy to prove that if $\mathbb{Z}[\alpha]$ possesses a number system $(\mathcal{B}, \alpha)$ (not necessarily a canonical system), then all algebraic conjugates α' of the number α are of modulus $|\alpha'| > 1$. Algebraic integers which have this property are called *expanding* algebraic integers. The converse is also true. Lagarias and Wang in [125], [124] and [126] proved that for any expanding algebraic integer α, the ring $\mathbb{Z}[\alpha]$ possesses a number system $(\mathcal{B}, \alpha)$ with the digits

$$\mathcal{B} = \{-(q-1),\ \ldots,\ -1,\ 0,\ 1,\ \ldots,\ q-1\}, \qquad \text{with } q = |P(0)|.$$

Such sets $\mathcal{B}$ are essentially the smallest possible digit sets. The proof of Lagarias and Wang is non-constructive. It relies on the existence of self–affine tilings of the space $\mathbb{R}^d$, associated to a certain integer matrix whose characteristic polynomial is $P(x)$. It is a very interesting and challenging problem to find an effective construction of the number systems for expanding algebraic integers. This problem is still unsolved.

2.4 Signal Processing

Let

$$a_0, \quad a_1, \quad a_2, \quad \ldots, \quad a_n$$

be a finite sequence of complex numbers. This sequence is called *unimodular* if all the numbers a_j are of modulus 1. Unimodular sequences are of considerable interest in signal processing theory, [85]. In particular, such sequences are used to convert discrete signals into continuous signal by phase modulation. Each number a_j in the sequence, written in the form $a_j = e^{i\phi}, \phi \in [0, 2\pi)$ corresponds to the pulse signal with the phase ϕ. The discrete signal is converted into a group of continuous pulses according to the pattern of the coefficients a_0, a_1, $\ldots$, a_n. The initial discrete signal is then restored back (demodulated) by cross correlating the received continuous signal with itself. For this, the autocorrelation coefficients are defined by the formula

$$c_k = \sum_{j=0}^{n-k} a_j \overline{a}_{j+k}, \qquad \text{and} \qquad c_{-k} = \overline{c}_k,$$

for $k = 0, 1, \ldots, n$. By the unimodularity property, the central autocorrelation coefficient c_0 is equal to

$$c_0 = |a_0|^2 + |a_1|^2 + \cdots + |a_n|^2 = n + 1.$$

In the signal processing theory, the central coefficient c_0 can be interpreted as the total signal energy, while the squared moduli of the non-central coefficients $|c_k|^2$, $k \neq 0$ can be thought as the energy in the k – th sidelobe of the signal. Energy in higher sidelobes represents the level of signal inefficiency – the energy which is lost due to self-interference [36]. Sequences which possess small autocorrelations have many practical applications in direct sequence spread spectrum telecommunications and radar pulse compression. A convenient mathematical framework to investigate this kind of problem is to associate a polynomial

$$P(z) = \sum_{j=0}^{n} a_j z^j$$

to each unimodular sequence. Such polynomials are called *unimodular*. The class of unimodular polynomials of degree n is denoted $\mathfrak{U}_n$. The signal energy then can be interpreted in terms of L^2 norm of unimodular polynomial:

$$c_0 = ||P||^2 = ||P||_2^2 = \frac{1}{2\pi} \int_0^{2\pi} \left| P(e^{it}) \right|^2 \, dt.$$

Another integral

$$\frac{1}{2\pi} \int_0^{2\pi} \left(|P(e^{it})|^2 - ||P||_2^2 \right)^2 dt = ||P||_4^4 - ||P||_2^4 =$$

$$= ||P||_4^4 - (n+1)^2 = 2 \sum_{k=1}^{n} |c_k|^2.$$

measures the deviation of the amplitude spectrum of the continuous time signal from the mean value.

Among all unimodular polynomials, polynomials with the coefficients $a_j \in \{-1, 1\}$, $0 \leqslant j \leqslant n$, are called *Littlewood* polynomials after [134], [135], [136]. The set of Littlewood polynomials of degree n is denoted $\mathcal{L}_n$. They correspond to the binary sequences, consisting of the numbers -1 and 1. A challenging mathematical problem, the determination of Littlewood polynomials which minimize the L^4 norm (and, consequently, minimize the sidelobe energy) is called the *merit factor problem*. As of present, the merit factor problem is still largely unsolved – see surveys [36], [107].

Another challenging problem of interest both in signal processing and ma-

thematical analysis, is the *Littlewood's problem on flat polynomials*. An infinite sequence of polynomials $P(z) \in \mathfrak{U}_n$ of increasing degree n is called a *flat sequence*, if the moduli $|P(z)|$ of polynomials $P(z)$ on the unit circle are uniformly bounded by their L^2 norms. More precisely, there exist absolute constants $c_2 > c_1 > 0$ (independent of n), such that the inequalities

$$c_1 < \frac{|P(z)|}{\sqrt{n+1}} < c_2$$

hold for all polynomials $P(z)$ in the sequence. This problem was considered by Littlewood [135], [136] and Erdős [82]. It is possible to construct flat unimodular polynomials by using the complex roots of unity (see the constructions of Littlewood [134], Kahane [110], Beck [21]). However, no infinite sequences of flat polynomials with all coefficients $a_j \in \{-1, 1\}$ are known to date, despite some supporting computational evidence given by Robinson [169].

Closely related to the above mentioned merit factor and flat polynomial problems are questions on the extremal Mahler measures and L^s norms of polynomials $P(z)$. It must be noted that the Lehmer conjecture for polynomials with all coefficients $a_j \in \{-1, 1\}$ was solved by Borwein, Dobrowolski and Mossinghoff [33]; the lower bounds were later improved in [72] and [98].

Another family of polynomials which appear often in the context of mathematical analysis are *Newman polynomials*. These are polynomials P with coefficients $a_j \in \{0, 1\}$, and a non-zero constant term $P(0) \neq 0$. The set of Newman polynomials of degree n is denoted $\mathcal{N}_n$. Newman polynomials, together with Littlewood polynomials, and general integer polynomials of height $H(P) = 1$ were studied in [39], [45], [52], [161], [200].

The minimal possible (asymptotic) logarithmic growth rate of $||P||_1$ norms of polynomials with restricted coefficients was proved by McGee, Pigno, Smith [144] and independently by Konyagin [118]. The sharp version of this result for minimal L^s norms for $s \in [0, 2]$ and maximal norms for $s \in [2, 4]$ of Littlewood polynomials was proved by Klemeš [116]. Very little is known about the upper bounds (largest possible Mahler measures and L^s norms). It is still not clear if Mahler measures or L^1 norms of Littlewood and Newman polynomials can be arbitrarily close to their L^2 norms – the question which dates back to Littlewood [135], [136], Newman [158], [159] and Mahler themselves, despite some progress made in [40], [81]. The book by P. Borwein [29] is a wonderful source of references on the subject.

2.5 Factorization of polynomials

It is not too surprising that heights of polynomial do appear in practical algorithms for factoring integer polynomials in $\mathbb{Z}[x]$. Let us suppose that an integer polynomial P is a product of two polynomials $Q, R \in \mathbb{Z}[x]$, that is $P = QR$. The Mahler measure of any integer polynomial is always greater or equal to 1. Also, it is multiplicative: $M(P) = M(QR) = M(Q)M(R)$. Thus $M(Q) \leqslant M(P)$ for any integer polynomial Q which divides P in $\mathbb{Z}[x]$. In a combination with a theorem of Smyth (2.3), this allows us to estimate the number of non-reciprocal polynomial factors of P in $\mathbb{Z}[x]$, counted with multiplicities: namely, it does not exceed

$$\left\lfloor \frac{\log M(P)}{\log \theta} \right\rfloor,$$

where $\theta = 1.32471...$ the Mahler measure of $x^3 - x - 1$ from (2.3) and $\lfloor \ldots \rfloor$ stands for the integer part of a real number. If Lehmer's conjecture is true, a similar bound should hold for the number of reciprocal factors of P – possibly, with θ replaced by the measure $1.17628\ldots$ of Lehmer's polynomial (2.2). One should refer to the papers of Schinzel [183], [184] for more estimates of this type.

A more sophisticated application of the Mahler measure in the form of the Mahler inequality (2.1) can be found in the factorization algorithm of Zassenhaus [216], [217]. Here the bound on the size of the coefficients of the polynomial factor Q is necessary in order to calculate the number of times Hensel's lift is performed.

The L^2 norm is also useful in the factorization of lacunary polynomials – polynomials which have only a few non-zero terms with small coefficients (such polynomials are also known as *sparse* polynomials). In particular, the identity $||P|| = ||QR|| = ||Q^*R||$ is very important in identifying the non-reciprocal factors of P. This identity is a key idea in the method of Ljunggren [137] who established that reducible trinomials and quadrinomials $P \in \mathbb{Z}[x]$ of height $H(P) = 1$ should vanish at some root of unity. Ljunggren's method was extended by Schinzel [181], [182]. See the papers of Filaseta [87], [92] for further applications of Ljunggren's method.

Chapter 3

Reduced Height

3.1 Statement of the problem

There are many diophantine applications when, for a given integer polynomial P, one needs either to find or to prove the existence of a nonzero integer polynomial G which is divisible by P and has the smallest possible height. In other words, one needs to evaluate $\min H(PQ)$ for a given $P(x) \in \mathbb{Z}[x]$, where the minimum is taken over every nonzero $Q(x) \in \mathbb{Z}[x]$. This problem is known as a special case of Siegel's lemma (see, for instance, [27], [146]). It is known that $\min H(PQ) \leqslant [M(P)]$, where $[\dots]$ stands for the integral part of a number. In particular, $\min H(PQ) = 1$ if $M(P) < 2$.

A similar quantity $\min ||PQ||$, where $P(x) \in \mathbb{Z}[x]$ and where the minimum is taken over every nonzero $Q(x) \in \mathbb{Z}[x]$, was introduced and studied by Filaseta, Robinson and Wheeler [91].

In principle, one can study similar problems for polynomials with coefficients in an arbitrary subring of $\mathbb{C}$ (not just $\mathbb{Z}$), for example, for polynomials with real or complex coefficients. Of course, for $P(x) \in \mathbb{R}[x]$, one should allow Q to have real coefficients too, whereas, for $P(x) \in \mathbb{C}[x]$, it is natural to take Q in $\mathbb{C}[x]$. On the other hand, the normalization of the problem should be different. The requirement that Q is in $\mathbb{Z}[x]$ can be replaced by the requirement that Q is a monic polynomial in $\mathbb{R}[x]$ or in $\mathbb{C}[x]$. So, for any given $P(x) \in \mathbb{R}[x]$ (or in $\mathbb{C}[x]$), we can study the quantities like

$$\inf ||PQ||, \ \inf L(PQ), \ \inf H(PQ),$$

where the infimum is taken over every monic $Q(x) \in \mathbb{R}[x]$ (or in $\mathbb{C}[x]$, respectively).

The first of these three quantities can be calculated using the result of Szegö.

For any $P(x) \in \mathbb{R}[x]$, Szegö's theorem (see, e.g., [206]) implies that

$$\inf_{Q \in \mathbb{R}[x] - \text{monic}} ||PQ|| = M(P).$$

This result was generalized to other L^p norms by Durand [76]. Lawton [130] noticed that it can be used for the practical calculation of $M(P)$ by introducing $M_n(P) := \min ||PQ||$, where the minimum is taken over monic $Q(x) \in \mathbb{R}[x]$ of degree at most n. These minima $M_n(P)$ can be calculated without computing the roots of P. They tend to $M(P)$ as $n \to \infty$. See the papers of Amoroso [10] and Dégot [58] for some further work on this problem.

Recently, in connection with the distribution of fractional parts of powers of an algebraic number, Dubickas [65] introduced and started to study the second quantity $l(P) := \inf_{Q \in \mathbb{R}[x] - \text{monic}} L(PQ)$. We called $l(P)$ the *reduced length* of a polynomial. The reduced length was then investigated in detail by Schinzel [187]. In particular, he proved that, in principle, the reduced length of polynomials having no roots of modulus 1 can be calculated. Schinzel's results show that there is no hope that a simple formula for $l(P)$ can be found. For example, the value of $l(2x^3 + 3x^2 + 4)$ is not known and is left as an open problem in [187].

In this chapter, we shall study the third quantity, namely, the *reduced height* of $P(x) \in \mathbb{R}[x]$ defined by the formula

$$\mathbb{H}(P) := \inf_{Q \in \mathbb{R}[x] - \text{monic}} H(PQ). \tag{3.1}$$

The main problem we will consider through Chapter 3 is:

Problem 3.1. *For a given polynomial $P \in \mathbb{R}[x]$, find an effective way to compute the reduced height $\mathbb{H}(P)$ and investigate its properties.*

We begin with the following basic properties of $\mathbb{H}(P)$:

Theorem 3.2. *Suppose that $P(x) \in \mathbb{R}[x]$, $c \in \mathbb{R}$, $w \in \mathbb{C}$, $k \in \mathbb{N}$. Then*

(i) $\mathbb{H}(cP) = |c|\mathbb{H}(P)$,

(ii) $\mathbb{H}((x-c)P(x)) = \mathbb{H}(P)$ *if* $|c| \leqslant 1$,

(iii) $\mathbb{H}((x-w)(x-\overline{w})P(x)) = \mathbb{H}(P)$ *if* $|w| \leqslant 1$,

(iv) $\mathbb{H}(x-c) = \max\{1, |c| - 1\}$,

(v) $\mathbb{H}(\pm P(\pm x^k)) = \mathbb{H}(P(x))$.

Theorem 3.2 (i) shows that in the study of $\mathbb{H}(P)$ we can restrict ourselves to monic polynomials $P(x) \in \mathbb{R}[x]$. For monic $P(x) \in \mathbb{R}[x]$, we clearly have

$$1 \leqslant \mathbb{H}(P) \leqslant H(P).$$

It is evident that each monic polynomial $P(x) \in \mathbb{R}[x]$ is completely determined by the list (multiset) of its roots counted with multiplicities. Suppose $\mathcal{S}$ is such a list. Obviously, $\mathcal{S}$ must be closed under the map $z \to \overline{z}$. We shall call any such list a *symmetric set of order* d if $\mathcal{S} \subset \mathbb{C}$ satisfies $\mathcal{S} = \overline{\mathcal{S}}$ and contains d elements counted with multiplicities. For instance, $1, 2, 1+i, 1+i, 1-i, 1-i$ is a symmetric set of order 6. If $\mathcal{S}$ contains d distinct elements (so $\mathcal{S}$ itself is a set), then a corresponding polynomial P in $\mathbb{R}[x]$ is separable, i.e. P has no multiple roots.

Note that, by Theorem 3.2 (ii), (iii), we can restrict ourselves to the study of polynomials which have all their roots in $|z| > 1$. The next theorem shows that it is sufficient to consider separable polynomials.

Theorem 3.3. *Suppose* $P, P_1, P_2, \cdots \in \mathbb{R}[x]$ *are monic polynomials such that* $||P_N - P|| \to 0$ *as* $N \to \infty$. *Then* $\lim_{N\to\infty} \mathbb{H}(P_N) = \mathbb{H}(P)$.

Indeed, each root α of P of multiplicity $m(\alpha) \geqslant 2$ can be replaced by $m(\alpha)$ distinct roots $\alpha, \alpha + 1/N, \ldots, \alpha + (m(\alpha)-1)/N$. For each N sufficiently large, say $N \geqslant N_0$, the polynomial P_N obtained in this way from P will be separable. The coefficients of a polynomial depend continuously on its roots. So $||P_N - P|| \to 0$ as $N \to \infty$. If we know $\mathbb{H}(P_N)$ for $N \geqslant N_0$, then using Theorem 3.3 we can calculate $\mathbb{H}(P) = \lim_{N\to\infty} \mathbb{H}(P_N)$, where P is a polynomial having multiple roots.

Summarizing, we see that in evaluation of $\mathbb{H}(P)$ it is sufficient to consider monic separable polynomials $P(x) \in \mathbb{R}[x]$ with all roots of modulus $|z| > 1$. Also, if $\mathcal{S}$ is a symmetric set, we can define

$$\mathbb{H}(\mathcal{S}) := \inf H(G - x^{\deg G}), \tag{3.2}$$

where the infimum is taken over every nonzero monic polynomial $G(x) \in \mathbb{R}[x]$ vanishing at each $\alpha \in \mathcal{S}$ with multiplicity $\geqslant m(\alpha)$ if $\mathcal{S}$ contains $m(\alpha)$ copies of α. Of course, for $G(x) = x^n + g_{n-1}x^{n-1} + \cdots + g_0$, we have $H(G - x^{\deg G}) = \max_{0\leqslant j\leqslant n-1} |g_j|$. The problem of finding $\mathbb{H}(\mathcal{S})$ is thus the problem of finding the infimum over the heights of monic polynomials vanishing at $\mathcal{S}$ with prescribed multiplicities. If P is a monic polynomial corresponding to $\mathcal{S}$ then

$$\mathbb{H}(P) = \max\{1, \mathbb{H}(\mathcal{S})\}. \tag{3.3}$$

Note that (3.2) implies that $\mathbb{H}(\mathcal{S}) \leqslant \mathbb{H}(\mathcal{S}')$ if $\mathcal{S} \subset \mathcal{S}'$ are two symmetric sets. Combined with (3.3), this yields that

$$\mathbb{H}(P) \geqslant \mathbb{H}(Q) \tag{3.4}$$

if P, Q are two monic polynomials in $\mathbb{R}[x]$ such that $Q|P$.

If $\mathcal{S} \subset \{z \in \mathbb{C} : \ 0 < |z| < 1\}$ is a finite symmetric set then one can consider a power series of the form $1+\sum_{j=1}^{\infty} h_j x^j$ vanishing at the points of $\mathcal{S}$ with respective multiplicities. Let $\mathbb{H}_{\rm ser}(\mathcal{S})$ be the infimum over all $h > 0$ for which there exists a power series $1 + \sum_{j=1}^{\infty} h_j x^j$, where $h_j \in \mathbb{R}$, $|h_j| \leqslant h$, vanishing at each $\alpha \in \mathcal{S}$ with multiplicity $\geqslant m(\alpha)$. (Here, $m(\alpha)$ is the number of copies of α in $\mathcal{S}$.) Using a standard compactness argument, we shall derive the following lemma showing that the value $\mathbb{H}_{\rm ser}(\mathcal{S})$ is attained, the proof of which appears at the end of Section 3.5.

Lemma 3.4. *For any finite symmetric set $\mathcal{S} \subset \{z \in \mathbb{C} : \ 0 < |z| < 1\}$, there exists a series $1 + \sum_{j=1}^{\infty} h_j x^j$, where $h_j \in \mathbb{R}$, $|h_j| \leqslant \mathbb{H}_{ser}(\mathcal{S})$, vanishing at each $\alpha \in \mathcal{S}$ with multiplicity $\geqslant m(\alpha)$.*

Evidently, $\mathcal{S}^{-1}$ (which contains elements reciprocal to those in $\mathcal{S}$) is a symmetric set if $\mathcal{S}$ is a symmetric set.

Theorem 3.5. *For any finite symmetric set $\mathcal{S}$, where $0 \notin \mathcal{S}$, we have $\mathbb{H}(\mathcal{S}) = \mathbb{H}_{ser}(\mathcal{S}^*)$, where $\mathcal{S}^* := \mathcal{S}^{-1} \cap \{z \in \mathbb{C} : \ |z| < 1\}$.*

For example, if $\mathcal{S} = \{3, 4\}$ then $\mathcal{S}^* = \{1/4, 1/3\}$. We will show below (see (3.16)) that $\mathbb{H}((x-3)(x-4)) = 6$. Combined with (3.3) and Theorem 3.5, this yields $\mathbb{H}_{\rm ser}(\mathcal{S}^*) = 6$. The value 6 is attained for the power series $1 - 6x + 6(x^2 + x^3 + \dots)$ vanishing at $1/4$ and $1/3$.

Although the value $\mathbb{H}_{\rm ser}(\mathcal{S})$ is attained by some coefficients of power series, this is not necessarily the case for $\mathbb{H}(\mathcal{S})$ and $\mathbb{H}(P)$. For example, taking $P(x) = x - 2$ and $\mathcal{S} = \{2\}$, by Theorem 3.2 (iv), we have $\mathbb{H}(x-2) = \mathbb{H}(\mathcal{S}) = 1$. However, there is no monic polynomial G divisible by $x - 2$ for which $H(G) = 1$. Indeed, for any $G(x) = x^n + g_{n-1}x^{n-1} + \dots + g_0$ satisfying $G(2) = 2^n + g_{n-1}2^{n-1} + \dots + g_0 = 0$, we have

$$H(G) \geqslant H(G - x^{\deg G}) = \max_{0 \leqslant j \leqslant n-1} |g_j| \geqslant 2^n/(1+2+\dots+2^{n-1}) = 1 + 1/(2^n - 1) > 1.$$

It follows that neither in (3.1) nor in (3.2) one can replace the infimum by the minimum.

Some problems for power series in terms of $\mathbb{H}_{\rm ser}(\mathcal{S})$, where $\mathcal{S} \subset \{z \in \mathbb{C} : \ 0 < |z| < 1\}$, were considered by Beaucoup, Borwein, Boyd and Pinner in [19] and [20]. For instance, using (3.3) and Theorem 3.5 we can restate the main problem considered in [19] as follows: for any $d \in \mathbb{N}$, find the maximal $\kappa = \kappa(d) > 1$ for which $\mathbb{H}((x-\kappa)^d) = 1$. Another problem which can be interpreted in terms of the reduced height of a polynomial was considered in [20]: given $\varphi \in (0, \pi)$, find the

largest $\varrho = \varrho(\varphi) > 1$ for which $\mathbb{H}((x - \varrho e^{i\varphi})(x - \varrho e^{-i\varphi})) = 1$. The results obtained in [19] and [20] combined with Theorem 3.5 serve as an additional motivation for the study of $\mathbb{H}(P)$.

We shall give the proofs of Theorems 3.2, 3.3, 3.5 in Section 3.6. In the next section we shall prove our main result, which is based on Theorem 3.5. Its application to quadratic polynomials is given in Section 3.3 (see Section 3.6 for the proofs). Some practical computations and examples will be given in Section 3.4. Section 3.5 contains the proof of Lemma 3 and some auxiliary results from linear algebra which will be used in Section 3.6.

3.2 The main result

Let $P(x) = (x - \alpha_1) \dots (x - \alpha_d) \in \mathbb{R}[x]$ be a separable polynomial with all roots of modulus $|z| > 1$. Put $\beta_1 := 1/\alpha_1, \dots, \beta_d := 1/\alpha_d$. Then $\mathcal{S} = \{\beta_1, \dots, \beta_d\}$ is a symmetric set with d distinct elements in $0 < |z| < 1$. We shall estimate $\mathbb{H}_{\text{ser}}(\mathcal{S})$ from below.

Suppose that $J = \{k_1, \dots, k_{d-1}\}$ is a subset of $\mathbb{N}$ such that

$$D(J) := \begin{vmatrix} 1 & \beta_1^{k_1} & \dots & \beta_1^{k_{d-1}} \\ \vdots & & & \\ 1 & \beta_d^{k_1} & \dots & \beta_d^{k_{d-1}} \end{vmatrix} \neq 0. \tag{3.5}$$

We define $S_n(J)$ by the formula

$$S_n(J) := \frac{1}{D(J)} \begin{vmatrix} \beta_1^n & \beta_1^{k_1} & \dots & \beta_1^{k_{d-2}} & \beta_1^{k_{d-1}} \\ \vdots & & & & \\ \beta_d^n & \beta_d^{k_1} & \dots & \beta_d^{k_{d-2}} & \beta_d^{k_{d-1}} \end{vmatrix}. \tag{3.6}$$

Clearly, from (3.5) and (3.6) we have

$$S_0(J) = 1, \quad S_{k_1}(J) = \dots = S_{k_{d-1}}(J) = 0. \tag{3.7}$$

We claim that

$$\mathbb{H}_{\text{ser}}(\mathcal{S}) \geqslant 1 / \sum_{j \in \mathbb{N} \setminus J} |S_j(J)|. \tag{3.8}$$

Indeed, suppose that $1 + \sum_{n=1}^{\infty} h_n x^n$ are arbitrary power series vanishing at $\mathcal{S}$.

Put

$$\ell_j := \frac{(-1)^{j-1}}{D(J)} \begin{vmatrix} \beta_1^{k_1} & \beta_1^{k_2} & \dots & \beta_1^{k_{d-2}} & \beta_1^{k_{d-1}} \\ \vdots & & & & \\ \beta_{j-1}^{k_1} & \beta_{j-1}^{k_2} & \dots & \beta_{j-1}^{k_{d-2}} & \beta_{j-1}^{k_{d-1}} \\ \beta_{j+1}^{k_1} & \beta_{j+1}^{k_2} & \dots & \beta_{j+1}^{k_{d-2}} & \beta_{j+1}^{k_{d-1}} \\ \vdots & & & & \\ \beta_d^{k_1} & \beta_d^{k_2} & \dots & \beta_d^{k_{d-2}} & \beta_d^{k_{d-1}} \end{vmatrix}.$$

By (3.6), we have $\ell_1\beta_1^n + \cdots + \ell_d\beta_d^n = S_n(J)$. Hence, multiplying each equality $1 + \sum_{n=1}^{\infty} h_n\beta_j^n = 0$, where $j = 1, \dots, d$, by ℓ_j and adding all d obtained equalities, we find that $S_0(J) + \sum_{n=1}^{\infty} h_n S_n(J) = 0$. Using (3.7), we deduce that $1 + \sum_{j\in\mathbb{N}\setminus J} h_j S_j(J) = 0$. Hence

$$\sup_{n\in\mathbb{N}} |h_n| \geqslant \sup_{j\in\mathbb{N}\setminus J} |h_j| \geqslant 1/\sum_{j\in\mathbb{N}\setminus J} |S_j(J)|.$$

This proves (3.8).

Is there any chance that the inequality (3.8) by an appropriate choice of J becomes an equality? In order to describe some cases when this can happen we shall introduce the following notation. For each $n \in \mathbb{N}$, put

$$\delta_n = \delta_n(J) := S_n(J)/|S_n(J)| \in \{-1, 1\} \quad \text{and} \quad \phi(x) = \phi(J, x) := \sum_{j\in\mathbb{N}\setminus J}^{\infty} \delta_j x^j. \quad (3.9)$$

Here, $\delta_n = 0$ in case $S_n(J) = 0$. Also, for each $j \in \{1, \dots, d, d+1\}$, let $D_j(J)$ denote the determinant of the matrix

$$\begin{pmatrix} \phi(\beta_1) & 1 & \beta_1^{k_1} & \dots & \beta_1^{k_{d-1}} \\ \vdots & & & & \\ \phi(\beta_d) & 1 & \beta_d^{k_1} & \dots & \beta_d^{k_{d-1}} \end{pmatrix} \quad (3.10)$$

with j-th column omitted, so that $D_1(J) = D(J)$.

Using (3.6), (3.9) and (3.10) we have

$$D_2(J) = \begin{vmatrix} \phi(\beta_1) & \beta_1^{k_1} & \dots & \beta_1^{k_{d-1}} \\ \vdots & & & \\ \phi(\beta_d) & \beta_d^{k_1} & \dots & \beta_d^{k_{d-1}} \end{vmatrix} = \sum_{j=1}^{\infty} \delta_j S_j(J) D(J) = D(J) \sum_{j\in\mathbb{N}\setminus J}^{\infty} |S_j(J)|.$$

So if $D(J) \neq 0$ then $D_2(J) \neq 0$ and

$$D(J)/D_2(J) = 1/\sum_{j\in\mathbb{N}\setminus J}^{\infty} |S_j(J)|. \quad (3.11)$$

In order to show that $\mathbb{H}_{\mathrm{ser}}(\mathcal{S}) \leqslant 1/\sum_{j\in\mathbb{N}\setminus J}|S_j(J)|$ for certain

$$J = \{k_1, \ldots, k_{d-1}\} \subset \mathbb{N}$$

(which combined with (3.8) would imply the equality), we shall look into the series

$$1 + \sum_{j\in J} h_{k_j} x^{k_j} + \sum_{j\in\mathbb{N}\setminus J} \delta_j h_0 x^j = 1 + \sum_{j\in J} h_{k_j} x^{k_j} + h_0\phi(x)$$

as a 'potential' candidate. By the definition of $D_j(J)$ (see (3.9) and (3.10)), we derive that the linear system

$$\begin{cases} h_0\phi(\beta_1) + h_{k_1}\beta_1^{k_1} + \cdots + h_{k_{d-1}}\beta_1^{k_{d-1}} = -1, \\ h_0\phi(\beta_2) + h_{k_1}\beta_2^{k_1} + \cdots + h_{k_{d-1}}\beta_2^{k_{d-1}} = -1, \\ \vdots \\ h_0\phi(\beta_d) + h_{k_1}\beta_d^{k_1} + \cdots + h_{k_{d-1}}\beta_d^{k_{d-1}} = -1, \end{cases}$$

has a unique solution

$$h_0 = -D_1(J)/D_2(J) = -D(J)/D_2(J),$$

$$h_{k_j} = (-1)^j D_{j+2}(J)/D_2(J),$$

where $j = 1, \ldots, d-1$. In particular, $|h_{k_j}| \leqslant |h_0|$ precisely when $|D_j(J)| \leqslant |D(J)|$ for each $j = 3, \ldots, d+1$. By the definition of $\mathbb{H}_{\mathrm{ser}}(\mathcal{S})$ and (3.11), for $\mathcal{S} = \{\beta_1, \ldots, \beta_d\} \subset \{z \in \mathbb{C} : \ 0 < |z| < 1\}$, we thus obtain that

$$\mathbb{H}_{\mathrm{ser}}(\mathcal{S}) \leqslant |h_0| = 1/\sum_{j\in\mathbb{N}\setminus J}|S_j(J)| = |D(J)/D_2(J)|$$

in case there is a $J \subset \mathbb{N}$ such that $D(J) \neq 0$ and $|D_j(J)| \leqslant |D(J)|$ for each $j = 3, \ldots, d+1$. Combined with (3.8) this implies that

$$\mathbb{H}_{\mathrm{ser}}(\mathcal{S}) = 1/\sum_{j\in\mathbb{N}\setminus J}|S_j(J)| = |D(J)/D_2(J)| \tag{3.12}$$

when there is a $J \subset \mathbb{N}$ such that $|D_j(J)| \leqslant |D(J)|$ for each $j = 3, \ldots, d+1$.

By Theorem 3.5, we have that $\mathbb{H}_{\mathrm{ser}}(\mathcal{S}) = \mathbb{H}(\{\alpha_1, \ldots, \alpha_d\})$, so (3.3) and (3.12) yield the following theorem:

Theorem 3.6. *Suppose that $P(x) \in \mathbb{R}[x]$ is a monic separable polynomial with*

all roots of modulus $|z| > 1$. *If* $J \subset \mathbb{N}$ *is such that* $D(J) \neq 0$ *then*

$$\mathbb{H}(P) \geqslant \frac{1}{\sum_{j\in\mathbb{N}\setminus J} |S_j(J)|} = \frac{|D(J)|}{|D_2(J)|}.$$

Furthermore, we have equality $\mathbb{H}(P) = \max\{1, |D(J)/D_2(J)|\}$ *in case* J *is a subset of* $\mathbb{N}$ *such that* $|D_j(J)| \leqslant |D(J)|$ *for each* $j = 3, \dots, d+1$.

In particular, taking $J_0 = \{1, \dots, d-1\}$, we have $D(J_0) = \prod_{1\leqslant i<j\leqslant d}(\beta_j - \beta_i) \neq 0$. If

$$S_n = S_n(J_0) = \frac{1}{D(J_0)} \begin{vmatrix} \beta_1^n & \beta_1 & \dots & \beta_1^{d-2} & \beta_1^{d-1} \\ \vdots & & & & \\ \beta_d^n & \beta_d & \dots & \beta_d^{d-2} & \beta_d^{d-1} \end{vmatrix}, \tag{3.13}$$

where $n = d, d+1, \dots$, all have the same sign then $\phi(x) = \pm x^d/(1-x)$. Thus

$$D_2(J_0) = \pm \begin{vmatrix} \beta_1^d/(1-\beta_1) & \beta_1 & \dots & \beta_1^{d-1} \\ \vdots & & & \\ \beta_d^d/(1-\beta_d) & \beta_d & \dots & \beta_d^{d-1} \end{vmatrix} = \frac{\pm\beta_1 \dots \beta_d D(J_0)}{(1-\beta_1)\dots(1-\beta_d)}.$$

Hence

$$|D(J_0)/D_2(J_0)| = |(\beta_1^{-1} - 1)\dots(\beta_d^{-1} - 1)| = |(\alpha_1 - 1)\dots(\alpha_d - 1)| = |P(1)|.$$

It follows that

$$\mathbb{H}(P) \geqslant |P(1)|$$

provided that all $S_n = S_n(J_0)$, $n = d, d+1, \dots$, have the same sign.

Put

$$R(x) := (x - 1/\alpha_1)\dots(x - 1/\alpha_d) = (x - \beta_1)\dots(x - \beta_d)$$
$$= x^d + r_1 x^{d-1} + \dots + r_d = P(1/x)x^d(-1)^d\beta_1 \dots \beta_d.$$

Then $S_n = S_n(J_0)$ satisfy the linear recurrence relation

$$S_{n+d} + S_{n+d-1}r_1 + \dots + S_n r_d = 0, \tag{3.14}$$

where $S_0 = 1$, $S_1 = \dots = S_{d-1} = 0$. In Section 3.5, we shall prove the following lemma:

Lemma 3.7. *For each* $j \in \{2, 3, \dots, d+1\}$ *we have*

$$\begin{vmatrix} \beta_1^d/(1-\beta_1) & 1 & \dots & \beta_1^{j-3} & \beta_1^{j-1} & \dots & \beta_1^{d-1} \\ \vdots & & & & & & \\ \beta_d^d/(1-\beta_d) & 1 & \dots & \beta_d^{j-3} & \beta_d^{j-1} & \dots & \beta_d^{d-1} \end{vmatrix} =$$

$$= \pm \frac{|D(J_0)||r_d + r_{d-1} + \cdots + r_{d-j+2}|}{|R(1)|}. \tag{3.15}$$

If S_n, $n = d, d+1, \ldots,$ all have the same sign then, by (3.10) and Lemma 3.7, we obtain that $|D_j(J_0)|/|D(J_0)| = |r_d + r_{d-1} + \cdots + r_{d-j+2}|/|R(1)|$ for $j \in \{2, 3, \ldots, d+1\}$. Combining this with Theorem 3.6 and using $|D(J_0)/D_2(J_0)| = |R(1)/r_d| = |P(1)|$ we derive the following corollary:

Corollary 3.8. *Let $P(x)$ be a separable polynomial with all roots in $|z| > 1$. Suppose $|r_d + \cdots + r_{d-j+1}| \leqslant |R(1)| = |r_d + \cdots + r_1 + 1|$ for each $j = 1, \ldots, d$, and suppose S_n, where $n = d, d+1, \ldots$, defined by (3.13) or (3.14) all have the same sign. Then $\mathbb{H}(P) = |P(1)|$.*

3.3 Quadratic polynomials

In this section, we give some corollaries of Theorem 3.6 to quadratic polynomials P. In particular, in the next two statements we compute explicitly the reduced length of a quadratic polynomial with two positive real roots. The proofs will be given in Section 3.6.

Corollary 3.9. *Let $u > v > 1$ be two real numbers, and let k be the largest positive integer for which $(1 - 2u^{1-k})/(u-1) \geqslant (1 - 2v^{1-k})/(v-1)$. Then*

$$\mathbb{H}((x-u)(x-v)) = \max\left\{1, \frac{u^k - v^k}{(u^k-2)/(u-1) - (v^k-2)/(v-1)}\right\}.$$

In particular, selecting $k = 1$ and combining this corollary with (3.12) (see also (3.3)), we obtain that

$$\mathbb{H}_{\text{ser}}(\{1/v, 1/u\}) = (u-1)(v-1) \tag{3.16}$$

if $(1 - 1/v)(1 - 1/u) \geqslant 1/2$.

For $P(x) = (x-u)^2 \in \mathbb{R}[x]$ the result is as follows:

Corollary 3.10. *Let $u \geqslant 0$ be a real number. Then*

$$\mathbb{H}((x-u)^2) = \begin{cases} (u-1)^2 \text{ if } u \geqslant 2 + \sqrt{2}, \\ 2u(u-1)^2/(u^2 - 2u + 2) \text{ if } u \in [2, 2+\sqrt{2}], \\ 3u^2(u-1)^2/(2u^3 - 3u^2 + 2) \text{ if } u \in [\kappa_1, 2], \\ 4u^3(u-1)^2/(3u^4 - 4u^3 + 2) \text{ if } u \in [\kappa_2, \kappa_1], \\ 1 \text{ if } u \in [0, \kappa_2], \end{cases}$$

where $\kappa_1 := 1.6279\ldots$ *and* $\kappa_2 := 1.5405\ldots$ *satisfy* $\kappa_1^4 - 8\kappa_1 + 6 = 0$ *and* $4\kappa_2^5 - 11\kappa_2^4 + 8\kappa_2^3 - 2 = 0$, *respectively.*

The minimal polynomial of $1/\kappa_2$ is $2x^5 - 8x^2 + 11x - 4$. This polynomial was found in [19] with respect to the above mentioned problem: find the maximal $\kappa = \kappa(d)$ for which $\mathbb{H}((x-\kappa)^d) = 1$. We have $\kappa(2) = \kappa_2 = 1.5405\ldots$. See [19] for the minimal polynomials of $1/\kappa(3)$ and $1/\kappa(4)$. We remark that the fourth line of Corollary 3.10 applied to $u = 8/5 \in (\kappa_2, \kappa_1)$ yields the equality $\mathbb{H}((x-8/5)^2) = 9216/8245$.

Our final statement deals with quadratic polynomials having two complex conjugate roots.

Corollary 3.11. *Let* $w = |w|e^{i\varphi}$ *be a complex number. If* $|w| \geqslant 2 + \sqrt{2}$ *then*

$$\mathbb{H}((x-w)(x-\overline{w})) = \frac{|w|^2}{1 + \sum_{j=1}^{\infty} |w|^{-j} |\sin((j+1)\varphi)/\sin(\varphi)|}. \tag{3.17}$$

This corollary is of interest in connection with the result of Schinzel, who proved in [187] that the reduced length $l(P)$ belongs to the field generated by the coefficients of P in the case when $P(x) \in \mathbb{R}[x]$ has all zeros outside the unit circle. It seems very likely that the value obtained at the right hand side of (3.17) can be transcendental for an algebraic integer w of degree 2 having a complex conjugate $\overline{w}$. In the next section we shall consider the example of $w = 9 + i$ with minimal polynomial $P(x) = x^2 - 18x + 82$.

3.4 Practical computations

Let $P(x) \in \mathbb{R}[x]$ be a monic polynomial. Let us define

$$\mathbb{H}_n(P) := \min H(PQ), \tag{3.18}$$

where the minimum is taken through all monic polynomials $Q(x) \in \mathbb{R}[x]$ of degree at most n. Clearly, $\mathbb{H}_0(P) \geqslant \mathbb{H}_1(P) \geqslant \mathbb{H}_2(P) \geqslant \ldots \geqslant \mathbb{H}(P)$ and $\lim_{n\to\infty} \mathbb{H}_n(P) = \mathbb{H}(P)$.

Using the simplex method of linear programming we can calculate $\mathbb{H}_n(P)$ explicitly for small values of n, e.g., for $n = 20$.

For example, with the input $P(x) = x^2 - 18x - 82$ the output for $\mathbb{H}_{20}(P)$ is the polynomial $x^{22} - g_1 x^{21} - g_2(x^{20} - x^{19} + \cdots - x + 1)$ with

$$g_1 = \frac{501374916424516058314281143579486 56}{263881534960299064058796703169185 1} =$$

$$= 18.999999999998023226693\ldots,$$

$$g_2 = \frac{151283283992847526085104101143692103 68}{240132196813872148293504999883958441} =$$

$$= 63.000000000045005485157\ldots.$$

This suggests that

$$\mathbb{H}(x^2 - 18x - 82) = 63. \tag{3.19}$$

Indeed, using Theorem 3.2 (v) we have $\mathbb{H}(x^2-18x-82) = \mathbb{H}(x^2+18x-82)$. Note that $x^2+18x-82 = (x-\alpha_1)(x-\alpha_2)$, where $\alpha_1 = -9-\sqrt{163}$, $\alpha_2 = -9+\sqrt{163}$. Setting $\beta_1 = 1/\alpha_1$ and $\beta_2 = 1/\alpha_2$, we get $R(x) = (x-\beta_1)(x-\beta_2) = x^2 - 9x/41 - 1/82$. The inequalities $1/82 \leqslant |R(1)| = 63/82$ and $1/82 + 9/41 \leqslant |R(1)| = 63/82$ of Corollary 3.8 hold. Moreover, by (3.13),

$$S_n = (\beta_1^n\beta_2 - \beta_2^n\beta_1)/(\beta_2 - \beta_1) = (\beta_2^{n-1} - \beta_1^{n-1})/2\sqrt{163}$$

are positive for every $n \geqslant 2$. Hence, by Corollary 3.8, we find that $\mathbb{H}(x^2 + 18x - 82) = |1 + 18 - 82| = 63$. This proves (3.19).

Another example is more interesting. For $P(x) = x^2 - 18x + 82$, the output for $\mathbb{H}_{20}(P)$ is the polynomial $x^{22} - g_3x^{21} + g_4(x^{20} + \cdots + x + 1)$, where

$$g_3 = \frac{39566397351975506821506664017372395 52}{23274351383515004012129299735108 4209} =$$

$$= 17.000000000000000000381\ldots,$$

$$g_4 = \frac{151283283992847526085104101143692103 68}{23274351383515004012129299735108 4209} =$$

$$= 65.000000000000000002691\ldots.$$

This suggests that the limit value is 65. However, this is not true! In fact, we have

$$\mathbb{H}(x^2 - 18x + 82) = 1/\sum_{j=2}^{\infty} |S_j| = 64.999999999999999999999999863\ldots,$$

where $S_0 = 1$, $S_1 = 0$ and $S_n = 9S_{n-1}/41 - S_{n-2}/82$ for $n = 2, 3, \ldots$. Indeed, since $|w| = |9 + i| \geqslant 2 + \sqrt{2}$, the condition of Corollary 3.11 is satisfied. By (3.25) (see the proof of Corollary 3.11 below, where we took $J = \{1\}$), the right hand side of (3.17) is equal to $1/\sum_{j=2}^{\infty} |S_j|$. It seems likely that the constant $1/\sum_{j=2}^{\infty} |S_j|$ is transcendental.

Finally, suppose that $P(x) = (x+3)(x+2)(x+1)x(x-1)(x-2)(x-3)$. Then, by Theorem 3.2 (ii), (v), (3.3) and (3.16), we find that

$$\mathbb{H}(P) = \mathbb{H}((x^2 - 9)(x^2 - 4)) = \mathbb{H}((x - 9)(x - 4)) = 8 \cdot 3 = 24.$$

In other words, the minimal height of a monic polynomial with real coefficients vanishing at $-3, -2, -1, 0, 1, 2$ and 3 is equal to 24.

3.5 Auxiliary lemmas from linear algebra

Throughout, we shall write the linear system

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,d}x_d = b_1, \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,d}x_d = b_2, \\ \vdots \\ a_{d,1}x_1 + a_{d,2}x_2 + \cdots + a_{d,d}x_d = b_d \end{cases}$$

in the matrix form $A\mathbf{x} = \mathbf{b}$, where $A = ||a_{i,j}||_{1 \leqslant i,j \leqslant d}$ is a $d \times d$ matrix, $\mathbf{x} := (x_1, \ldots, x_d)^T$, $\mathbf{b} := (b_1, \ldots, b_d)^T$. Here and below, T stands for the transpose.

Lemma 3.12. *Let $A = ||a_{i,j}||_{1 \leqslant i,j \leqslant d}$ be a $d \times d$ matrix with complex entries, $b_1, \ldots, b_d \in \mathbb{C}$, and $\varepsilon > 0$. Suppose that the linear system $A\mathbf{x} = \mathbf{b}$ has at least one real solution, and that there exist $y_1, \ldots, y_d \in \mathbb{R}$ such that $|a_{i,1}y_1 + a_{i,2}y_2 + \cdots + a_{i,d}y_d - b_i| < \varepsilon$ for each $i = 1, \ldots, d$. Then there is a constant $c = c(A) > 0$ and a real vector $\mathbf{x} = (x_1, \ldots, x_d)^T$, where $x_j \in (y_j - \varepsilon c, y_j + \varepsilon c)$ for each $j = 1, \ldots, d$, such that $A\mathbf{x} = \mathbf{b}$.*

Proof: By the condition of the lemma, there exist $\varepsilon_1, \ldots, \varepsilon_d \in \mathbb{C}$, all of moduli less than ε such that $a_{i,1}y_1 + a_{i,2}y_2 + \cdots + a_{i,d}y_d = b_i + \varepsilon_i$ for $i = 1, \ldots, d$. Take any real vector $\mathbf{x}$ satisfying $A\mathbf{x} = \mathbf{b}$. Then $\mathbf{z} := (y_1 - x_1, \ldots, y_d - x_d)^T$ is a real solution of $A\mathbf{z} = (\varepsilon_1, \ldots, \varepsilon_d)^T$. If the matrix A is non-singular, namely, $\det A \neq 0$, then $A\mathbf{z} = (\varepsilon_1, \ldots, \varepsilon_d)^T$ has a unique solution $z_j = \det A_j / \det A$ $(j = 1, 2, \ldots, d)$, where A_j is the matrix A with j-th column replaced by $(\varepsilon_1, \ldots, \varepsilon_d)^T$. This yields that for each $j \in \{1, \ldots, d\}$ we have $|y_j - x_j| < \varepsilon c$, where c depends on A only (and not on $b_1, \ldots, b_d$). Since $y_j - x_j \in \mathbb{R}$, the proof of the lemma in this (non-singular) case is completed.

In the alternative case, when $\det A = 0$, the equation $A\mathbf{z} = (\varepsilon_1, \ldots, \varepsilon_d)^T$ has infinitely many solutions. We may suppose without loss of generality that the largest nonzero minor corresponds to the matrix $A' = ||a_{i,j}||_{1 \leqslant i,j \leqslant r}$. Selecting $x_j = y_j$ for $j = r + 1, \ldots, d$, by the above argument applied to the non-singular matrix A' and to the vector $(x_1, \ldots, x_r)$ (instead of $(x_1, \ldots, x_d)$), we derive that $x_j \in (y_j - \varepsilon c, y_j + \varepsilon c)$ for $i = 1, \ldots, r$, where c depends on A' only. This completes the proof of the lemma. □

For $z \in \mathbb{C}$ we set

$$V(z) := (z, z^2, \ldots, z^d)$$

Likewise, let $V^{(m)}(z)$ be the vector obtained from $V(z)$ by replacing each entry with the m-th derivative. Given a symmetric set $\mathcal{S}$ of order $d = m_1 + \cdots + m_s$ (which is, say, a list of m_1 copies of β_1, ..., m_s copies of β_s), we define the matrix $A(\mathcal{S})$ by its d consecutive rows

$$V(\beta_1), \ldots, V^{(m_1-1)}(\beta_1), V(\beta_2), \ldots, V^{(m_2-1)}(\beta_2), \ldots, V(\beta_s), \ldots, V^{(m_s-1)}(\beta_s).$$

Its determinant is known as a version of confluent Vandermonde determinant. It is nonzero if the numbers $\beta_1, \ldots, \beta_s$ are distinct and $\beta_j \neq 0$ for $j = 1, \ldots, s$ (see, e.g., [121]).

Lemma 3.13. *Let $\mathcal{S} \subset \{z \in \mathbb{C} : \ 0 < |z| < 1\}$ be a finite symmetric set. Then $\mathbb{H}(\mathcal{S}^{-1}) \leqslant \mathbb{H}_{ser}(\mathcal{S})$.*

Proof: Suppose $f(x) = 1 + \sum_{j=1}^{\infty} h_j x^j$, where

$$h_1, h_2, \cdots \in \mathbb{R}, \quad \text{and} \quad h := \sup_{j \geqslant 1} |h_j|,$$

satisfies $f(\beta_1) = \cdots = f^{(m_1-1)}(\beta_1) = \cdots = f(\beta_s) = \cdots = f^{(m_s-1)}(\beta_s) = 0$. Set $d = m_1 + \cdots + m_s$. By (3.2), we see that it suffices to show that, for each $\varepsilon > 0$, there is a monic polynomial $G(x) = x^n + g_{n-1}x^{n-1} + \cdots + g_0 \in \mathbb{R}[x]$ satisfying $H(G - x^{\deg G}) = \max_{0 \leqslant j \leqslant n-1} |g_j| \leqslant h + \varepsilon$ which vanishes at $1/\beta_j$ with multiplicity $\geqslant m_j$ $(j = 1, \ldots, d)$. Here, $\beta_j \neq 0$.

Put $\beta := \max_{1 \leqslant j \leqslant s} |\beta_j| < 1$ and $m := \max\{m_1, \ldots, m_s\}$. Take n so large that

$$\sum_{j=n+1}^{\infty} |h_j| j^{m-1} \beta^{j-m+1} < \varepsilon.$$

Set $f_n(x) := 1 + h_{d+1}x^{d+1} + \cdots + h_n x^n$. Note that

$$h_1 x + \cdots + h_d x^d = f(x) - f_n(x) - \sum_{j=n+1}^{\infty} h_j x^j.$$

On applying Lemma 3.12 to the matrix $A = A(\mathcal{S})$, the real vector $(x_1, \ldots, x_d) = (h_1, \ldots, h_d)$ and

$$(b_1, \ldots, b_d) = (-f_n(\beta_1), \ldots, -f_n^{(m_1-1)}(\beta_1), \ldots, -f_n(\beta_s), \ldots, -f_n^{(m_s-1)}(\beta_s)),$$

we find that there is a constant c depending on $\mathcal{S}$ only and $g_j \in (h_j - \varepsilon c, h_j + \varepsilon c)$ $(j = 1, \ldots, d)$ such that

$$A(\mathcal{S})(g_1, \ldots, g_d)^T = (b_1, \ldots, b_d)^T.$$

This means that the polynomial $1 + g_1x + \cdots + g_dx^d + h_{d+1}x^{d+1} + \cdots + h_nx^n$ vanishes at β_j with multiplicity $\geqslant m_j$ $(j = 1, \ldots, s)$. Its reciprocal polynomial $G(x) = x^n + g_1x^{n-1} + \cdots + g_dx^{n-d} + h_{d+1}x^{n-d-1} + \cdots + h_n$ satisfies $H(G - x^{\deg G}) \leqslant h + \varepsilon$ and, for each $j \in \{1, \ldots, s\}$, vanishes at $1/\beta_j$ with multiplicity $\geqslant m_j$. It follows that $\mathbb{H}(\mathcal{S}^{-1}) \leqslant \mathbb{H}_{\mathrm{ser}}(\mathcal{S}) + \varepsilon$. Since ε can be taken arbitrarily small, this completes the proof. □

The next lemma shows that the minimum in (3.18) is attained.

Lemma 3.14. *Let $P(x) \in \mathbb{R}[x]$ be a monic polynomial of degree d, and let n be a nonnegative integer. Then there is a monic polynomial $Q_n(x) = x^n + b_{n-1}x^{n-1} + \cdots + b_0$ such that $\mathbb{H}_n(P) = H(PQ_n)$.*

Proof: Let $P(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_0$. Each monic polynomial divisible by $P(x)$ has the form $(x^d + a_{d-1}x^{d-1} + \cdots + a_0)(x^n + b_{n-1}x^{n-1} + \cdots + b_0)$. We need to minimize the maximum of the following $n + d + 1$ numbers $|a_0b_0|$, $|a_0b_1 + a_1b_0|$, $\ldots$, $|a_{d-1} + b_{n-1}|$, 1. Writing each inequality $|a| \leqslant H$ as two inequalities $H - a \geqslant 0$ and $H + a \geqslant 0$, we get a system of $2(n + d)$ inequalities in $n + 1$ unknowns $b_0, \ldots, b_{n-1}, H$. It is clear that the minimum H for which this system has a solution exists, so the quantity $\mathbb{H}_n(P)$ is equal to $\max\{1, H\}$. This completes the proof of the lemma. □

In fact, by the fundamental theorem of linear programming, if a linear programming problem has a solution, then at least one of the solutions always occurs at a corner point. We can thus find the polynomial $G(x) = P(x)Q_n(x) = x^{n+d} + g_{n+d-1}x^{n+d-1} + \cdots + g_0$, where Q_n is the polynomial of Lemma 3.14, such that at most $d - 1$ of the numbers $|g_{n+d-1}|$, $\ldots$, $|g_0|$ are smaller than the largest of the numbers $|g_{n+d-1}|$, $\ldots$, $|g_0|$. In other words, $|g_j| = H(G - x^{n+d})$ for all but at most $d - 1$ indices $j \in \{0, 1, \ldots, n + d - 1\}$. This explains our strategy used in the proof of Theorem 3.6. We just need to find the 'correct' set of $d - 1$ 'small' coefficients, because, for any symmetric set $\mathcal{S}$ of order d in $0 < |z| < 1$, there exists a series $h(x) = 1 + \sum_{j=1}^{\infty} h_jx^j$ vanishing at $\mathcal{S}$ with each (except for at most $d - 1$ coefficients) h_j equal to $\pm\mathbb{H}_{\mathrm{ser}}(\mathcal{S})$.

Proof of Lemma 3.7: Let us multiply both sides of (3.15) by

$$R(1) = (1 - \beta_1) \ldots (1 - \beta_d).$$

We need to show that

$$\begin{vmatrix} 1-\beta_1 & \beta_1-\beta_1^2 & \dots & \beta_1^{j-3}-\beta_1^{j-2} & \beta_1^{j-1} & \dots & \beta_1^d \\ \vdots & & & & & & \\ 1-\beta_d & \beta_d-\beta_d^2 & \dots & \beta_d^{j-3}-\beta_d^{j-2} & \beta_d^{j-1} & \dots & \beta_d^d \end{vmatrix} = \pm|D(J_0)\sum_{k=d-j+2}^{d} r_k|. \tag{3.20}$$

Notice that the left hand side of (3.20) is equal to

$$\begin{vmatrix} 1-\beta_1^{j-2} & \beta_1-\beta_1^{j-2} & \dots & \beta_1^{j-3}-\beta_1^{j-2} & \beta_1^{j-1} & \dots & \beta_1^d \\ \vdots & & & & & & \\ 1-\beta_d^{j-2} & \beta_d-\beta_d^{j-2} & \dots & \beta_d^{j-3}-\beta_d^{j-2} & \beta_d^{j-1} & \dots & \beta_d^d \end{vmatrix}$$

for each $j \in \{3, \dots, d+1\}$. Using the next well-known formula (see, e.g., Problem 346 in [167])

$$\begin{vmatrix} 1 & \beta_1 & \dots & \beta_1^{k-1} & \beta_1^{k+1} & \dots & \beta_1^d \\ \vdots & & & & & & \\ 1 & \beta_d & \dots & \beta_d^{k-1} & \beta_d^{k+1} & \dots & \beta_d^d \end{vmatrix} = (\beta_1 \dots \beta_{d-k} + \dots + \beta_{k+1} \dots \beta_d) \prod_{1 \leqslant i < j \leqslant d} (\beta_j - \beta_i)$$

$$= (-1)^{d-k} r_{d-k} D(J_0),$$

where $k = 0, \dots, d-1$, we can expand the left hand side of the next determinant by its first $j-2$ columns:

$$\begin{vmatrix} 1-\beta_1^{j-2} & \beta_1-\beta_1^{j-2} & \dots & \beta_1^{j-3}-\beta_1^{j-2} & \beta_1^{j-1} & \dots & \beta_1^d \\ \vdots & & & & & & \\ 1-\beta_d^{j-2} & \beta_d-\beta_d^{j-2} & \dots & \beta_d^{j-3}-\beta_d^{j-2} & \beta_d^{j-1} & \dots & \beta_d^d \end{vmatrix} =$$

$$= (-1)^{d-j+2} D(J_0) \sum_{k=d-j+2}^{d} r_k.$$

This implies (3.20) and completes the proof of the lemma. □

We conclude this section with the proof of Lemma 3.4:

By the definition of $\mathbb{H}_{\text{ser}}(\mathcal{S})$, for any $N \in \mathbb{N}$, there exists a power series

$$1 + \sum_{j=1}^{\infty} h_{j,N} x^j$$

vanishing at $\mathcal{S}$ such that $|h_{j,N}| \leqslant \mathbb{H}_{\text{ser}}(\mathcal{S}) + 1/N$. Put $h := \mathbb{H}_{\text{ser}}(\mathcal{S})$. We can choose a sequence $N_1 < N_2 < N_3 < \dots$ of positive integers such that $h_{1,N_k} \to h_1 \in [-h, h]$ as $k \to \infty$. Then, we choose its subsequence (denoted by $N_1 < N_2 < N_3 < \dots$ again) such that $h_{2,N_k} \to h_2 \in [-h, h]$ as $k \to \infty$ and so on. We claim that the

series $1+\sum_{j=1}^{\infty} h_j x^j$ vanishes at $\mathcal{S}$ with required multiplicities. Indeed, suppose that $\beta \in \mathcal{S}$ (so $|\beta|<1$), but $B := 1+\sum_{j=1}^{\infty} h_j \beta^j \neq 0$. Take $M \in \mathbb{N}$ so large that

$$3\max\{1,h\}|\beta|^{M+1}/(1-|\beta|) < |B|/2.$$

Next, take $N \in \mathbb{N}$ which is, e.g., an element of the above sequence $N_1 < N_2 < \dots$ after M steps are taken and is so large that

$$|h_1 - h_{1,N}| < |B|/2M, \quad |h_2 - h_{2,N}| < |B|/2M, \quad \dots, \quad |h_M - h_{M,N}| < |B|/2M.$$

Since $1+\sum_{j=1}^{\infty} h_{j,N_k}\beta^j = 0$ for any $N \in \mathbb{N}$, using $|h_j - h_{j,N}| \leqslant 3\max\{1,h\}$, we obtain that

$$|B| = |\sum_{j=1}^{\infty}(h_j - h_{j,N})\beta^j| < \sum_{j=1}^{M}|h_j - h_{j,N}| + \sum_{j=M+1}^{\infty}|(h_j - h_{j,N})\beta^j| <$$

$$< |B|/2 + |B|/2 = |B|,$$

a contradiction. It follows that $B=0$. The case when β is a multiple root can be treated in the same manner. This completes the proof of the lemma. □

3.6 Proofs

Proof of Theorem 3.2: Clearly, $H(cP) = |c|H(P)$ implies (i). Using (i) we can assume that P in $(ii), (iii), (v)$ is monic. Hence (3.4) implies that $\mathbb{H}((x-c)P(x)) \geqslant \mathbb{H}(P)$. On the other hand, suppose that $Q(x) \in \mathbb{R}[x]$ is such that $H(PQ) < \mathbb{H}(P)+\varepsilon$. Take $n > \deg(PQ)$. Since $(x^n - c^n)P(x)Q(x)$ is a polynomial of height $H(PQ)$ divisible by $(x-c)P(x)$, we obtain that $\mathbb{H}((x-c)P(x)) \leqslant \mathbb{H}(P)$. This proves (ii).

Similarly, by (3.4), we have $\mathbb{H}((x-w)(x-\overline{w})P(x)) \geqslant \mathbb{H}(P)$ for any $w \in \mathbb{C}$. Suppose that $Q(x) \in \mathbb{R}[x]$ is a polynomial for which $H(PQ) < \mathbb{H}(P)+\varepsilon$. We claim that, for $|w| \leqslant 1$, there is $n > \deg(PQ)$ such that the polynomial $(x^n - w^n)(x^n - \overline{w}^n) = x^{2n} - 2\Re(w^n)x^n + |w|^2$ is of height 1. This would imply (iii) as above. Writing $w = |w|e^{i\varphi}$, where $\varphi \in (0,\pi)$, we have $2\Re(w^n) = 2|w|^n\cos(n\varphi)$. Its modulus is at most 1 if $|\cos(n\varphi)| \leqslant 1/2$ which is equivalent to $\cos(2n\varphi) \leqslant -1/2$ and to $||n\varphi/\pi|| \geqslant 1/3$. (Here, $||x||$ denotes the distance from $x \in \mathbb{R}$ to the nearest integer, whereas in the proof of Theorem 2 the same notation is used for the Euclidean norm of a polynomial.) The inequality $||n\varphi/\pi|| \geqslant 1/3$ clearly holds for infinitely many $n \in \mathbb{N}$ if φ/π is irrational. For $\varphi/\pi \in \mathbb{Q}$, namely, $\varphi/\pi = u/v$ with integer $u < v$, where $v \geqslant 2$, by taking $n = vk + r$, where $ur \equiv [v/2] \pmod{v}$, we have $||(vk+r)u/v|| = ||ru/v|| = [v/2]/v \geqslant 1/3$. This completes the proof of (iii).

(The example $\varphi = \pi/3$ shows that the constant $1/3$ in $||n\varphi/\pi|| \geqslant 1/3$ cannot be improved.)

Obviously, $\mathbb{H}(P(x)) = \mathbb{H}(P(-x))$, so in the proof of (iv) we can assume that $c > 0$. Since $x^n - c^n$ is divisible by $x - c$, we have $\mathbb{H}(x-c) = 1$ for $c \in (0, 1]$. Suppose $c > 1$. Note that $x - c$ divides the polynomial $x^n - ((c-1)/(1-c^{-n}))(x^{n-1} + \cdots + x + 1)$ of height $\max\{1, (c-1)/(1-c^{-n})\}$. Since c^{-n} tends to zero as $n \to \infty$, we have $\mathbb{H}(x-c) \leqslant \max\{1, c-1\}$. On the other hand, each polynomial $x^n + c_{n-1}x^{n-1} + \cdots + c_0$ vanishing at c has at least one coefficient c_i of modulus greater than or equal to $(c-1)/(1-c^{-n})$, since otherwise $c^n \leqslant |c_0| + \cdots + |c_{n-1}|c^{n-1} < (c^n - 1)(c-1)/(c-1)(1-c^{-n}) = c^n$, which is a contradiction. Thus $\mathbb{H}_n(x-c) \geqslant \max\{1, (c-1)/(1-c^{-n})\}$. This implies that $\mathbb{H}(x-c) \geqslant \max\{1, c-1\}$ and proves (iv).

The proof of (v) is exactly the same as the proof of Proposition (iv) in [187]. The upper bound $\mathbb{H}(P(x^k)) \leqslant \mathbb{H}(P(x))$ is trivial by (3.1). For the lower bound $\mathbb{H}(P(x^k)) \geqslant \mathbb{H}(P(x))$, we write $P(x^k)Q(x)$ in the form

$$P(x^k)Q_0(x^k) + xP(x^k)Q_1(x^k) + \cdots + x^{k-1}P(x^k)Q_{k-1}(x^k).$$

Here, $Q_j(x^k) = x^{-j}\sum_{i \equiv j \pmod k} q_i x^i$, where $Q(x) = \sum_{i=0}^n q_i x^i$, is a polynomial in x^k. Observing that $H(P(x^k)Q(x)) \geqslant H(P(x^k)Q_0(x^k))$ we obtain that $\mathbb{H}(P(x^k)) \geqslant \mathbb{H}(P(x))$. So $\mathbb{H}(P(x^k)) = \mathbb{H}(P(x))$. Combined with $\mathbb{H}(\pm P(\pm x)) = \mathbb{H}(P(x))$ this completes the proof of (v). □

Proof of Theorem 3.3: Since $||P_N - P|| \to 0$ as $N \to \infty$ and $P, P_1, P_2, \ldots$ are all monic, we can assume without loss of generality that $P_N(x) = x^d + a_{N,d-1}x^{d-1} + \cdots + a_{N,0}$ for $N = 1, 2, \ldots$ and $P(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_0$. It follows that $\lim_{N\to\infty} a_{N,j} = a_j$ as $N \to \infty$ for each $j \in \{0, 1, \ldots, d-1\}$.

Let ε be a fixed positive number. Since $\mathbb{H}(P) = \lim_{n\to\infty} \mathbb{H}_n(P)$, there is a positive integer n so large that $\mathbb{H}_n(P) \geqslant \mathbb{H}(P) \geqslant \mathbb{H}_n(P) - \varepsilon$. Lemma 3.14 implies that $\mathbb{H}_n(P) = H(PQ_n)$ for some monic polynomial Q_n of degree n, thus $\mathbb{H}(P) \geqslant H(PQ_n) - \varepsilon$. Next, take N so large that $H(P_NQ_n) - H(PQ_n) \leqslant \varepsilon$. These inequalities show that $H(P_NQ_n) \leqslant \mathbb{H}(P) + 2\varepsilon$. But $\mathbb{H}(P_N) \leqslant H(P_NQ_n)$, so $\mathbb{H}(P_N) \leqslant \mathbb{H}(P) + 2\varepsilon$. Hence

$$\limsup_{N\to\infty} \mathbb{H}(P_N) \leqslant \mathbb{H}(P).$$

It remains to show that, for any $\varepsilon > 0$, there is a positive integer N_0 such that $\mathbb{H}(P) \leqslant \mathbb{H}(P_N) + \varepsilon$ for each $N \geqslant N_0$. This would imply that $\liminf_{N\to\infty} \mathbb{H}(P_N) \geqslant \mathbb{H}(P)$. By combining this inequality with the upper bound for the largest limit point, we will be able to conclude that $\lim_{N\to\infty} \mathbb{H}(P_N) = \mathbb{H}(P)$.

By Theorem 3.2 (ii), (iii), we can assume that the roots of P are all in $|z| \geqslant r > 1$. Since the roots of a monic polynomial depend continuously from its coefficients, there is no loss of generality to assume that the roots of P_N are all in $|z| \geqslant r_1 > 1$.

Let $\mathcal{S} = \{\beta_1, \dots, \beta_d\}$ and $\mathcal{S}_N = \{\beta_{N,1}, \dots, \beta_{N,d}\}$ be the multisets which are reciprocal to the multisets $\{\beta_1^{-1}, \dots, \beta_d^{-1}\}$ and $\{\beta_{N,1}^{-1}, \dots, \beta_{N,d}^{-1}\}$ of roots of P and P_N, respectively. It is clear that the multisets $\mathcal{S}_N$ tend pointwise to $\mathcal{S}$ as $N \to \infty$. By (3.3) and Theorem 3.5 (the proof of this theorem will be given below), it is sufficient to prove the inequality

$$\mathbb{H}_{\mathrm{ser}}(\mathcal{S}) \leqslant \mathbb{H}_{\mathrm{ser}}(\mathcal{S}_N) + \varepsilon$$

for $N \geqslant N_0(\varepsilon)$.

The argument is quite close to the one used in the proof of Lemma 3.13. Fix some constants $r_2 > 0$ and $r_3 < 1$ such that the elements of $\mathcal{S}$ and $\mathcal{S}_N$ all lie in the annulus $r_2 \leqslant |z| \leqslant r_3$.

Let $f_N(x) = 1 + \sum_{j=1}^{\infty} g_{N,j} x^j$ be the series vanishing at $\mathcal{S}_N$ for which

$$\sup_{j \geqslant 1} |g_{N,j}| = \mathbb{H}_{\mathrm{ser}}(\mathcal{S}_N).$$

(Lemma 3.4 implies their existence for any N.) We shall change the first d coefficients of f_N into $g_1, \dots, g_d$ and consider the series $\bar{f}_N(x) = 1 + \sum_{j=1}^{d} g_j x^j + \sum_{j=d+1}^{\infty} g_{N,j} x^j$. These d coefficients will be chosen in a way which guarantees that $\bar{f}_N$ vanishes at $\mathcal{S}$ with required multiplicities, namely, $\bar{f}_N(\beta_s) = \dots = \bar{f}_N^{(m_s-1)}(\beta_s) = 0$ if β_s occurs in $\mathcal{S}$ with multiplicity m_s. A corresponding linear system of d equations in d unknowns $g_1, \dots, g_d$ gives

$$g_j = F_{j,0}(\mathcal{S}) + F_{j,d+1}(\mathcal{S}) g_{N,d+1} + F_{j,d+2}(\mathcal{S}) g_{N,d+2} + \dots$$

for $j = 1, \dots, d$. Here, setting

$$D(\mathcal{S}) := \begin{vmatrix} \beta_1 & \beta_1^2 & \dots & \beta_1^d \\ \vdots & & & \\ \beta_d & \beta_d^2 & \dots & \beta_d^d \end{vmatrix}$$

and $D_{j,m}(\mathcal{S})$ for $D(\mathcal{S})$ with j-th column replaced by the column $(\beta_1^m, \dots, \beta_d^m)$, we have

$$F_{j,m}(\mathcal{S}) = -D_{j,m}(\mathcal{S}) / D(\mathcal{S}) \tag{3.21}$$

for each $m \in \{0, d+1, d+2, \dots\}$. Note that, although $D(\mathcal{S}) = 0$ in the case when at least one β_s belongs to $\mathcal{S}$ is with multiplicity $\geqslant 2$, the functions $F_{j,m}(\mathcal{S})$ are well

defined. Moreover, from (3.21) it is easily seen that, for each $m \geqslant d+1$, $F_{j,m}(\mathcal{S})$ is a symmetric polynomial in $\beta_1, \ldots, \beta_d$ of degree $\leqslant m$ with at most m^{r_4} terms, where r_4 is a positive constant depending on d only. For example, for $d = 2$, we have

$$g_1 = -(\beta_1+\beta_2)(\beta_1\beta_2)^{-1} + \beta_1\beta_2 \sum_{j=3}^{\infty} g_{N,j}(\beta_1^{j-3} + \beta_1^{j-2}\beta_2 + \cdots + \beta_2^{j-3})$$

and

$$g_2 = (\beta_1\beta_2)^{-1} - \sum_{j=3}^{\infty} g_{N,j}(\beta_1^{j-2} + \beta_1^{j-1}\beta_2 + \cdots + \beta_2^{j-2}).$$

Of course, the fact that f_N vanishes at $\mathcal{S}_N$ implies that

$$g_{N,j} = F_{j,0}(\mathcal{S}_N) + F_{j,d+1}(\mathcal{S}_N)g_{N,d+1} + F_{j,d+2}(\mathcal{S}_N)g_{N,d+2} + \ldots$$

for each $j = 1, \ldots, d$. Here, $F_{j,m}(\mathcal{S}_N)$ are defined as in (3.21), where each β_j in the above determinants is replaced by $\beta_{N,j}$. Subtracting $g_{N,j}$ from g_j, we deduce that

$$g_j - g_{N,j} = F_{j,0}(\mathcal{S}) - F_{j,0}(\mathcal{S}_N) + \sum_{t=d+1}^{\infty} g_{N,t}(F_{j,t}(\mathcal{S}) - F_{j,t}(\mathcal{S}_N)). \tag{3.22}$$

Now, fix $\varepsilon > 0$. We will show that, for $N \geqslant N_0(\varepsilon)$, $|g_j - g_{N,j}| < \varepsilon$ for each $j = 1, \ldots, d$. For this, we split the sum in (3.22) into two sums corresponding to $t \leqslant M$ and $t \geqslant M+1$, where M will be chosen later (M will be the same for all N).

We will first bound the sum over $t \geqslant M+1$. Note that, since β_j and $\beta_{N,j}$ all lie in the annulus $r_2 \leqslant |z| \leqslant r_3$ and since $F_{j,t}(\mathcal{S})$ and $F_{j,t}(\mathcal{S}_N)$ are symmetric polynomials in $\beta_1, \ldots, \beta_d$ and $\beta_{N,1}, \ldots, \beta_{N.d}$, respectively, of degree $\leqslant t$ with $\leqslant t^{r_4}$ terms, $|F_{j,t}(\mathcal{S})|$ and $|F_{j,t}(\mathcal{S}_N)|$ do not exceed $t^{r_4}r_3^t$. Thus

$$\sum_{t=M+1}^{\infty} |g_{N,t}||F_{j,t}(\mathcal{S}) - F_{j,t}(\mathcal{S}_N)| \leqslant 2\mathbb{H}_{\mathrm{ser}}(\mathcal{S}_N) \sum_{t=M+1}^{\infty} t^{r_4}r_3^t < \varepsilon/2$$

if M is large enough, say, $M \geqslant M(\varepsilon)$. Here, we bound each $\mathbb{H}_{\mathrm{ser}}(\mathcal{S}_N)$ from above by an absolute constant, so that $M(\varepsilon)$ is independent of N. Fix one of such large M, say, $M = M(\varepsilon)$.

Let us order the points of $\mathcal{S}_N$ so that, for each $j \in \{1, \ldots, d\}$, $\beta_{N,j}$ is 'close' to β_j, and put

$$\delta_N := \max_{1 \leqslant j \leqslant d} |\beta_{N,j} - \beta_j|.$$

Clearly, $\delta_N \to 0$ as $N \to \infty$. From the formula (3.21), we obtain that there exists

some positive constant r_5 which depends on r_2, r_3, M and $\mathcal{S}$ only such that

$$|F_{j,t}(\mathcal{S}) - F_{j,t}(\mathcal{S}_N)| \leqslant r_5\delta_N$$

for every $t \in \{0, d+1, d+2, \dots, M\}$ and $j \in \{1, \dots, d\}$. Then (3.22) implies that

$$|g_j - g_{N,j}| \leqslant r_5\delta_N + (M-d)\mathbb{H}_{\mathrm{ser}}(\mathcal{S}_N)r_5\delta_N + \varepsilon/2$$

for each $j \in \{1, \dots, d\}$, where r_5 is a positive constant. Taking $N_0(\varepsilon)$ so large that $r_5\delta_N + (M-d)\mathbb{H}_{\mathrm{ser}}(\mathcal{S}_N)r_5\delta_N < \varepsilon/2$ for $N \geqslant N_0(\varepsilon)$, we obtain that $|g_j - g_{N,j}| < \varepsilon$ for all $N \geqslant N_0(\varepsilon)$.

It follows that, for $N \geqslant N_0(\varepsilon)$, the moduli of the coefficients of the series $\bar{f}_N$, namely, $|g_1|, \dots, |g_d|, |g_{N,d+1}|, |g_{N,d+2}|, \dots$ are all smaller than $\sup_{j\geqslant 1}|g_{N,j}| + \varepsilon = \mathbb{H}_{\mathrm{ser}}(\mathcal{S}_N) + \varepsilon$. But $\bar{f}_N$ vanishes at $\mathcal{S}$ with required multiplicities, so

$$\mathbb{H}_{\mathrm{ser}}(\mathcal{S}) \leqslant \sup\{|g_1|, \dots, |g_d|, |g_{N,d+1}|, |g_{N,d+2}|, \dots\} \leqslant \mathbb{H}_{\mathrm{ser}}(\mathcal{S}_N) + \varepsilon,$$

as claimed. □

Proof of Theorem 3.5: Note that, as in Theorem 3.2 $(ii), (iii)$, we have $\mathbb{H}(\mathcal{S}) = \mathbb{H}(\mathcal{S}')$, where $\mathcal{S}' = \mathcal{S} \cap \{z \in \mathbb{C} : \ |z| > 1\}$. Evidently, $\mathcal{S}^* = \mathcal{S}'^{-1}$, so for the proof of $\mathbb{H}(\mathcal{S}) = \mathbb{H}_{\mathrm{ser}}(\mathcal{S}^*)$ it suffices to show that $\mathbb{H}(\mathcal{S}') = \mathbb{H}_{\mathrm{ser}}(\mathcal{S}'^{-1})$. The bound $\mathbb{H}(\mathcal{S}') \leqslant \mathbb{H}_{\mathrm{ser}}(\mathcal{S}'^{-1})$ follows immediately from Lemma 3.13.

As for the inequality $\mathbb{H}(\mathcal{S}') \geqslant \mathbb{H}_{\mathrm{ser}}(\mathcal{S}'^{-1})$, observe first that for each $\varepsilon > 0$ there is a polynomial $G(x)$ vanishing at the points of $\mathcal{S}'$ with prescribed multiplicities which satisfies $H(G - x^{\deg G}) < \mathbb{H}(\mathcal{S}') + \varepsilon$. On replacing $G(x)$ by its reciprocal and adding zero terms, we obtain a series that vanish at $\mathcal{S}'^{-1}$ with prescribed multiplicities. It follows that $\mathbb{H}_{\mathrm{ser}}(\mathcal{S}'^{-1}) \leqslant H(G - x^{\deg G}) + \varepsilon$. This, by (3.2) and (3.3), implies that $\mathbb{H}(\mathcal{S}') \geqslant \mathbb{H}_{\mathrm{ser}}(\mathcal{S}'^{-1})$ and completes the proof of $\mathbb{H}(\mathcal{S}) = \mathbb{H}_{\mathrm{ser}}(\mathcal{S}^*)$. □

Proof of Corollary 3.9: Let k be the largest positive integer for which $(1 - 2u^{1-k})/(u-1) \geqslant (1 - 2v^{1-k})/(v-1)$. Take $J = \{k\}$ and apply Theorem 3.6 to $\beta_1 = 1/u$, $\beta_2 = 1/v$. By (3.5) and (3.6), we have $D(J) = v^{-k} - u^{-k} > 0$ and $S_n(J) = (uv)^{-k}(u^{k-n} - v^{k-n})/D(J)$ which is positive for $n < k$ and negative for $n > k$. Hence, by (3.9),

$$\phi(x) = x + \cdots + x^{k-1} - x^{k+1} - x^{k+2} - \cdots = (x - x^k - x^{k+1})/(1-x).$$

Next, by (3.10), we obtain that

$$D_2(J) = \frac{\phi(u^{-1})}{v^k} - \frac{\phi(v^{-1})}{u^k} = \frac{u^{-1} - u^{-k} - u^{-k-1}}{(1-u^{-1})v^k} - \frac{v^{-1} - v^{-k} - v^{-k-1}}{(1-v^{-1})u^k}$$

$$= (uv)^{-k}\Big(\frac{u^k - u - 1}{u-1} - \frac{v^k - v - 1}{v-1}\Big) = (uv)^{-k}\Big(\frac{u^k - 2}{u-1} - \frac{v^k - 2}{v-1}\Big).$$

Thus

$$\frac{D(J)}{D_2(J)} = \frac{u^k - v^k}{(u^k-2)/(u-1) - (v^k-2)/(v-1)}. \tag{3.23}$$

Similarly, by (3.10), we have

$$D_3(J) = \phi(u^{-1}) - \phi(v^{-1}) = \frac{1 - u^{-k+1} - u^{-k}}{u-1} - \frac{1 - v^{-k+1} - v^{-k}}{v-1}. \tag{3.24}$$

By (3.23), (3.24) and Theorem 3.6, we see that the proof of the corollary will be completed if we show that $|D_3(J)| \leqslant D(J) = v^{-k} - u^{-k}$, namely,

$$\Big|\frac{1 - u^{-k+1} - u^{-k}}{u-1} - \frac{1 - v^{-k+1} - v^{-k}}{v-1}\Big| \leqslant v^{-k} - u^{-k}.$$

This inequality is equivalent to the system of two inequalities

$$\begin{cases} (1-2u^{-k})/(u-1) \leqslant (1-v^{-k})/(v-1), \\ (1-2u^{1-k})/(u-1) \geqslant (1-v^{1-k})/(v-1). \end{cases}$$

Clearly, both these inequalities hold if k is defined as above. This completes the proof of Corollary 3.9. □

Proof of Corollary 3.10: Let $\mathcal{S}_{u,v} = \{u, v\}$ and let $\mathcal{S}_{u,u}$ be the symmetric set u, u. Clearly,

$$\lim_{v\to u} \mathbb{H}(\mathcal{S}_{u,v}) = \mathbb{H}(\mathcal{S}_{u,u}).$$

(This follows from Theorem 3.3, where we proved that $\mathbb{H}(\mathcal{S}_N) \to \mathbb{H}(\mathcal{S})$ if $\mathcal{S}_N$ as a vector tends to $\mathcal{S}$ as $N \to \infty$.) By a standard computation, we find that

$$\lim_{v\to u} \frac{u^k - v^k}{(u^k-2)/(u-1) - (v^k-2)/(v-1)} = \frac{ku^{k-1}(u-1)^2}{(k-1)u^k - ku^{k-1} + 2}$$

and

$$\lim_{v\to u} \frac{(1-2u^{1-k})/(u-1) - (1-2v^{1-k})/(v-1)}{u-v} = -\frac{u^k - 2ku + 2(k-1)}{u^k(u-1)^2}.$$

Set $u_1 := \infty$, and suppose $u_k > 1$, where $k = 2, 3, \ldots$, is the largest real root of the equation $x^k - 2kx + 2(k-1) = 0$. Clearly, $u_1 > u_2 > u_3 > \ldots$, $\lim_{k\to\infty} u_k = 1$,

and the condition 'the largest k for which $(1-2u^{1-k})/(u-1) \geqslant (1-2v^{1-k})/(v-1)$ holds' becomes '$u \in [u_{k+1}, u_k)$'. So Corollary 3.9 implies that

$$\mathbb{H}(\mathcal{S}_{u,u}) = \frac{ku^{k-1}(u-1)^2}{(k-1)u^k - ku^{k-1} + 2}$$

for each $u \in [u_{k+1}, u_k)$. Observing that $u_2 = 2+\sqrt{2}$, $u_3 = 2$, $u_4 = \kappa_1$ and that at $u = \kappa_2$ the equality $4u^3(u-1)^2 = 3u^4 - 4u^3 + 2$ holds, we conclude the proof of Corollary 3.10 via (3.3). □

Proof of Corollary 3.11: Take $J = J_0 = \{1\}$ in Theorem 3.6. Suppose that $\mathcal{S} = \{w, \overline{w}\} = \{|w|e^{i\varphi}, |w|e^{-i\varphi}\}$. Then $\beta_1 = \beta = |w|^{-1}e^{i\varphi}, \beta_2 = \overline{\beta} = |w|^{-1}e^{-i\varphi}$. We now find that

$$S_n = S_n(J_0) = S_n = (\beta^n\overline{\beta} - \overline{\beta}^n\beta)/(\overline{\beta} - \beta) = -|w|^{-n}\sin((n-1)\varphi)/\sin(\varphi).$$

Since

$$\sum_{n=2}^{\infty} |S_n| = |w|^{-2}(1 + \sum_{j=1}^{\infty} |w|^{-j}|\sin((j+1)\varphi)/\sin(\varphi)|), \tag{3.25}$$

we find from Theorem 3.6 that

$$\mathbb{H}(\mathcal{S}) \geqslant |w|^2/(1 + \sum_{j=1}^{\infty} |w|^{-j}|\sin((j+1)\varphi)/\sin(\varphi)|).$$

In order to show that the equality holds, it suffices to prove that $|\phi(\beta)\overline{\beta} - \phi(\overline{\beta})\beta|$ and $|\phi(\beta) - \phi(\overline{\beta})|$ are both smaller than or equal to $|\beta - \overline{\beta}|$. Here, $\phi(x) = \sum_{j=2}^{\infty} \delta_j x^j$, $\delta_j \in \{-1, 1\}$.

Since $|\beta^j\overline{\beta} - \overline{\beta}^j\beta|/|\beta - \overline{\beta}| \leqslant (j-1)|\beta|^j = (j-1)|w|^{-j}$ and $|\beta^j - \overline{\beta}^j|/|\beta - \overline{\beta}| \leqslant j|\beta|^{j-1} = j|w|^{-j+1}$, we obtain that

$$\frac{|\phi(\beta)\overline{\beta} - \phi(\overline{\beta})\beta|}{|\beta - \overline{\beta}|} \leqslant \sum_{j=2}^{\infty}(j-1)|w|^{-j} = \frac{1}{(|w|-1)^2},$$

and

$$\frac{|\phi(\beta) - \phi(\overline{\beta})|}{|\beta - \overline{\beta}|} \leqslant \sum_{j=2}^{\infty} j|w|^{-j+1} = \frac{2|w|-1}{(|w|-1)^2}.$$

Clearly, both right hand sides $1/(|w|-1)^2$ and $(2|w|-1)/(|w|-1)^2$ are smaller than or equal to 1, because $|w| \geqslant 2+\sqrt{2}$. This completes the proof of the corollary. □

Finally, we remark that, by Theorem 3.6, for any $d \in \mathbb{N}$ there is a constant $\eta(d)$ such that the reduced height of $P(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_0 \in \mathbb{R}[x]$ which

has all its roots in $|z| \geqslant \eta(d)$ can be evaluated by the formula

$$\mathbb{H}(P) = \Big(\sum_{j=d}^{\infty} |S_j| \Big)^{-1},$$

where S_n, $n = 0, 1, 2, \ldots$, satisfy the linear recurrence relation (3.14). Indeed, then the roots $\beta_i = 1/\alpha_i$ of the polynomial

$$R(x) = (x - \beta_1) \ldots (x - \beta_d) = x^d + r_1 x^{d-1} + \cdots + r_d = P(1/x) x^d (-1)^d \beta_1 \ldots \beta_d$$

are so small that the conditions of Theorem 3.6 are satisfied for the set $J = J_0 = \{1, \ldots, d-1\}$.

The value $\eta(2)$ is equal to $2 + \sqrt{2}$. (See Corollaries 3.9-3.11.) It seems that, for each $d \geqslant 2$, the formula $\eta(d) = 1/(1 - (1 + 1/(d-1)!)^{-1/d})$ is true. Some evidence towards this formula can be given as follows. For $|z| \leqslant 1 - (1 + 1/(d-1)!)^{-1/d}$, the inequality $|\phi^{(d-1)}(z)| \leqslant 1$, where $\phi(z) = \sum_{j=d}^{\infty} \pm z^j$, holds for any distribution of signs $\pm$, giving $|D_{d+1}(J_0)| \leqslant |D(J_0)|$ (see (3.10) and Theorem 3.6).

Chapter 4

Maximal values of polynomials

4.1 Statement of the problem

A nonzero polynomial with $0, 1$ coefficients is called a *Newman polynomial* after [159]. There is a variety of different problems in number theory and analysis related to Newman polynomials. See, for instance, [39], [45], [52], [161], [200].

Our research presented in this chapter is motivated by the work of Akiyama, Brunotte, Pethö and Steiner [2] which, at the first glance, has nothing to do with Newman polynomials. They investigate the sequence of integers satisfying $a_{n+1} = -[\lambda a_n] - a_{n-1}$, $n = 1, 2, \ldots$. It is conjectured in [2] that, for any $a_0, a_1 \in \mathbb{Z}$ and $\lambda \in [-2, 2]$, the sequence a_n, $n = 0, 1, 2, \ldots$ is periodic. The nontrivial case is when $\lambda \in (-2, 2) \setminus \{-1, 0, 1\}$. This problem seems to be very difficult, especially, when the number ζ, defined by the equality $\zeta + \zeta^{-1} = -\lambda$ (so that $|\zeta| = 1$), is not a root of unity. In fact, the only case when the periodicity of the sequence a_n, $n = 0, 1, 2, \ldots$, is proved and published [2] is when $\lambda = (1 + \sqrt{5})/2 = 2\cos(\pi/5)$, so that ζ corresponding to λ is a root of unity. It seems that similar methods can be applied to some other λ of the form $2\cos(\pi r)$ with $r \in \mathbb{Q}$. However, for $\lambda \neq 2\cos(\pi r)$, i.e., when ζ is not a root of unity, the periodicity problem seems to be completely out of reach.

We now explain how this periodicity problem is related to polynomials with coefficients in $[0, 1]$ and, in particular, with Newman polynomials. Rewrite the recurrence equation as $a_{j+1} + \lambda a_j + a_{j-1} = \{\lambda a_j\}$. Multiplying each equality by ζ^j and adding all obtained equalities for $j = 1, \ldots, n$, using $\zeta + \zeta^{-1} = -\lambda$, we get

$$(a_{n+1} - \zeta a_n)\zeta^n = \sum_{j=1}^{n} \{\lambda a_j\}\zeta^j + (a_1 - \zeta a_0).$$

Put $r_n := |a_{n+1} - \zeta a_n|$. Then

$$|r_n| \leqslant |\sum_{j=1}^{n} \{\lambda a_j\}\zeta^j| + |r_0| = |\sum_{j=1}^{n} \{\lambda a_j\}\zeta^{j-1}| + |r_0|.$$

One can show easily (see Proposition 2.4 in [2]) that the periodicity of the sequence a_n, $n = 0, 1, 2, \ldots$, would follow from the inequality

$$\limsup_{n\to\infty} \frac{r_n^2}{n} < \frac{\sqrt{4-\lambda^2}}{\pi}.$$

The sum $\sum_{j=1}^{n} \{\lambda a_j\}\zeta^{j-1}$ is equal to the value at ζ of a certain polynomial of degree $\leqslant n-1$ with all coefficients in the interval $[0, 1)$. This suggests the problem to which Chapter 4 is devoted:

Problem 4.1. *Let $\zeta \in \mathbb{C}$ be a fixed complex number of modulus $|\zeta| = 1$. Denote by $S(\zeta, n)$ the maximal modulus of the real polynomial of degree at most $n-1$, evaluated at the point $z = \zeta$, assuming that the coefficients of the polynomial are restricted to the interval $[0, 1]$. Find the value $S(\zeta, n)$.*

We shall prove below that $\lim_{n\to\infty} S(\zeta, n)/n = 1/\pi$ for every ζ of modulus 1 which is not a root of unity, so one gets $\limsup_{n\to\infty} |r_n|/n \leqslant 1/\pi$. Moreover, we will show that the polynomials $f(z)$ which have the maximal modulus at the point $z = \zeta$ are Newman polynomials. Unfortunately, the result is still too weak to solve the above problem of periodicity.

Finally, let us consider the case $\lambda = 1/2$. Then $\zeta = (-1 + i\sqrt{15})/4$ satisfying $\zeta + \zeta^{-1} = -1/2$ is not a root of unity. We claim that the sequence a_n, $n = 0, 1, 2, \ldots$, defined by $a_{n+1} = -[a_n/2] - a_{n-1}$, $n = 1, 2, \ldots$, contains at least four equal elements. Indeed, without loss of generality suppose that the sequence $|a_n|$, $n = 0, 1, 2, \ldots$, is unbounded. Then, for any $N \in \mathbb{N}$, there is an index $n > N$ such that $|a_n| \geqslant |a_j|$ for $j = 0, 1, \ldots, n-1$. The corresponding polynomial $f(z) := \sum_{j=1}^{n} \{a_j/2\} z^{j-1}$ is a Newman polynomial multiplied by $1/2$. The inequality

$$|r_n| = |a_{n+1} - \zeta a_n| \leqslant |f(\zeta)|/2 + |a_1 - \zeta a_0|$$

combined with the inequality $|a_{n+1} - \zeta a_n| \geqslant |\Im(\zeta a_n)| = |a_n|\sqrt{15}/4$ implies that $|a_n| \leqslant 2|f(\zeta)|/\sqrt{15} + 4|a_1 - \zeta a_0|/\sqrt{15}$. Hence, by Theorem 4.5 below, for any $\varepsilon > 0$ and any sufficiently large $n > n(\varepsilon)$, we have $|a_n| < (2/(\pi\sqrt{15}) + \varepsilon)n < 0.165n$. The interval $[-0.165n, 0.165n]$ contains at most $0.33n + 1 < 0.333n < n/3$ distinct integers. Since $|a_n| \geqslant |a_j|$, $j = 0, 1, \ldots, n-1$, it includes all integers $a_0, a_1, \ldots, a_n$. If none of them is repeated more than three times, then the set $\{a_0, a_1, \ldots, a_n\}$ is of cardinality $\geqslant (n+1)/3 > n/3$. This leads to the contradiction.

4.2 Main results

Let Λ_n be the set of real polynomials of degree $\leqslant n-1$ with all coefficients in the interval $[0,1]$. Set

$$S(\zeta, n) := \max_{f \in \Lambda_n} |f(\zeta)|$$

for any $\zeta \in \mathbb{C}$. It is clear that

$$S(\zeta, n) = 1 + \zeta + \cdots + \zeta^{n-1}$$

for each nonnegative real number ζ.

We remark first that, for any fixed $\zeta \in \mathbb{C}$, the maximum $S(\zeta, n)$ is attained for some polynomial $f(z) = c_0 + c_1 z + \cdots + c_{n-1} z^{n-1} \in \Lambda_n$ Indeed, treating $f(\zeta)$ as a complex continuous function in n real variables $c_0, \ldots, c_{n-1} \in [0,1]$, by a standard argument of compactness, we see that its modulus $|f(\zeta)|$ attains its maximum for some fixed values of the coefficients $c_0, \ldots, c_{n-1} \in [0,1]$. It follows that, for any $\zeta \in \mathbb{C}$, there exists a (not necessarily unique) polynomial $f \in \Lambda_n$ such that $S(\zeta, n) = |f(\zeta)|$.

Below, we sometimes use the vector representation of complex numbers. Let us denote the value $f(\zeta)$ with the largest modulus $|f(\zeta)|$ among all $f \in \Lambda_n$ by the vector $\mathbf{s}$. As we already said above, the vector $\mathbf{s}$ satisfying $|\mathbf{s}| = S(\zeta, n)$ is not necessarily unique. We begin with the following simple but important observation:

Theorem 4.2. *Let $\zeta \neq 0$, and let $\mathbf{s} = f(\zeta) = \sum_{j=0}^{n-1} c_j \zeta^j$ be one of the vectors of maximal length, where $f \in \Lambda_n$. Then f is a Newman polynomial. Moreover, for each $j = 0, 1, \ldots, n-1$, we have $c_j = 1$ if the projection of the vector ζ^j to the vector $\mathbf{s}$ is positive, and $c_j = 0$ otherwise.*

In particular, if $\mathbf{s}$ is one of the extremal vectors, then the line passing through the origin and orthogonal to $\mathbf{s}$ contains none of the points $1, \zeta, \ldots, \zeta^{n-1}$. Therefore, Theorem 4.2 suggests the following practical method for the computation of $S(\zeta, n)$. Suppose that $\zeta \neq 0$. Let ℓ be any line passing through the origin but through none of the n points $D_n := \{1, \zeta, \ldots, \zeta^{n-1}\}$. Let us rotate the line ℓ, say, counterclockwise until it reaches at least one of the points of D_n. Then rotate ℓ again by an angle so small that no point of D_n lies on ℓ and stop. At this, first, stop we calculate the sums r_1 and l_1 of the numbers from D_n that lie on both sides, say, 'right hand side' and 'left hand side' of ℓ. (Note that $r_1 + s_1 = 1 + \zeta + \cdots + \zeta^{n-1}$.) Then rotate ℓ until it reaches at least one point of D_n again, slightly pass this point, stop for the second time, and calculate r_2, l_2, where $r_2 + l_2 = 1 + \zeta + \cdots + \zeta^{n-1}$, and so on. The last, say, kth stop will be when ℓ is rotated by the angle π, so that it reaches its original position (but changes its direction). It is easy to see

that $k \leqslant n$, where the value n for k is attained when no two points of D_n lie on a line passing through the origin. Theorem 4.2 implies that

$$S(\zeta, n) = \max\Big(|r_1|, |l_1|, |r_2|, |l_2|, \dots, |r_k|, |l_k|\Big).$$

In particular, if ζ is a negative real number, then all of its powers are positive and negative real numbers. Let us start with a line, say, orthogonal to the real axis and begin the process described above. Then there is only one stop, giving $r_1 = 1 + \zeta^2 + \cdots + \zeta^u$, where $u \leqslant n-1$ is the largest even integer, and $l_1 = -\zeta - \zeta^3 - \cdots - \zeta^v$, where $v \leqslant n-1$ is the largest odd integer. The formula $S(\zeta, n) = \max\Big(|r_1|, |l_1|\Big)$ yields the following corollary:

Corollary 4.3. *Let u and v be the largest even and odd numbers, respectively, satisfying $u, v \leqslant n-1$. If ζ is a negative real number then*

$$S(\zeta, n) = \max\Big(1 + \zeta^2 + \cdots + \zeta^u, -\zeta(1 + \zeta^2 + \cdots + \zeta^{v-1})\Big).$$

Suppose that ζ is a complex number of modulus 1. In the evaluation of $S(\zeta, n)$ there are two different cases depending on whether ζ is or is not a root of unity. Let throughout $\zeta_d := \exp(2\pi i/d)$ be a primitive dth root of unity. Let also U_d be the set of its conjugates over $\mathbb{Q}$, so that $|U_d| = \varphi(d)$, where $\varphi(d)$ stands for the Euler totient function. In the next theorem, we calculate the value $S(\zeta, md)$ for every $\zeta \in U_d$ and $m \in \mathbb{N}$.

Theorem 4.4. *Suppose that $m \in \mathbb{N}$ and $\zeta \in U_d$, where $d \geqslant 2$. Then $S(\zeta, md) = m/\sin(\pi/d)$ if d is even and $S(\zeta, md) = m/(2\sin(\pi/2d))$ if d is odd.*

The main theorem of this chapter can be stated as follows:

Theorem 4.5. *Let $\zeta \in \mathbb{C}$ be a complex number of modulus 1. If $\zeta \in U_d$, where $d \in \mathbb{N}$, then*

$$\lim_{n\to\infty} S(\zeta, n)/n = \begin{cases} 1, & \text{if } d = 1, \\ 1/(d\sin(\pi/d)) & \text{if } d \text{ is even,} \\ 1/(2d\sin(\pi/2d)) & \text{if } d > 1 \text{ is odd.} \end{cases}$$

If ζ is not a root of unity then $\lim_{n\to\infty} S(\zeta, n)/n = 1/\pi$.

In the next section, we shall prove Theorems 4.2, 4.4 and 4.5. Some numerical examples will be given in Section 4.4.

4.3 Proofs

Proof of Theorem 4.2: The vector $\mathbf{s}$ is the sum of the vectors ζ^j, where $j = 0, \ldots, n-1$, scaled by $c_j \in [0,1]$. Clearly, $|\mathbf{s}| > 0$. Put $\mathbf{s}_j := \zeta^j$. If there is an index $j \in \{0, \ldots, n-1\}$ such that the projection of $\mathbf{s}_j = \zeta^j$ to $\mathbf{s}$ is positive (i.e., the scalar product $(\mathbf{s}_j, \mathbf{s})$ is positive) and $c_j < 1$ then, by replacing c_j by 1, we obtain that the length of the vector $\mathbf{s} - c_j\mathbf{s}_j + \mathbf{s}_j = \mathbf{s} + (1-c_j)\mathbf{s}_j$ is greater than $|\mathbf{s}|$, a contradiction. Similarly, suppose that there is an index $j \in \{0, \ldots, n-1\}$ such that the projection of $\mathbf{s}_j = \zeta^j$ to $\mathbf{s}$ is negative or zero (i.e., $(\mathbf{s}_j, \mathbf{s}) \leqslant 0$) and $c_j > 0$. Then, by replacing c_j by 0, we obtain that the vector $\mathbf{s} - c_j\mathbf{s}_j$ is longer than $|\mathbf{s}|$, because $|\mathbf{s} - c_j\mathbf{s}_j|^2 - |\mathbf{s}|^2 = c_j^2|\mathbf{s}_j|^2 - 2c_j(\mathbf{s}_j, \mathbf{s}) \geqslant c_j^2|\mathbf{s}_j|^2 > 0$, a contradiction again. □

The following simple lemma will be used in the proof of Theorem 4.4 and in numerical examples of Section 4.4:

Lemma 4.6. *Let Γ_d be the set of complex roots of $z^d - 1 = 0$, where $d \geqslant 2$, and let ℓ be a line passing through the origin but through none of the points of Γ_d. Then the sum of all numbers from Γ_d that lie on one side of ℓ belongs to some axis of symmetry of a regular d-gon with vertices in Γ_d, and the modulus of this sum is equal to $1/\sin(\pi/d)$ for d even, and to $1/(2\sin(\pi/2d))$ for d odd.*

Proof of Lemma 4.6: Consider a half plane in that side of ℓ where exactly $k = [d/2]$ points of Γ_d are lying. Take $\zeta_d = \exp(2\pi i/d)$. Let r be the smallest positive integer such that ζ_d^r is the first vertex of Γ_d in that half plane counterclockwise. Then the points of Γ_d in this half plane are the powers ζ_d^j, where $j = r, \ldots, r+k-1$. Note that all sums $\zeta_d^{r+j} + \zeta_d^{r+k-1-j}$, where $j = 0, \ldots, [(k-1)/2]$, lie on the same axis of symmetry of a regular d-gon, hence so does their sum $\sum_{j=r}^{r+k-1} \zeta_d^j = \frac{1}{2}\sum_{j=0}^{k-1}(\zeta_d^{r+j} + \zeta_d^{r+k-1-j})$ on the same side of ℓ.

Next, recall that $1 + \zeta_d + \cdots + \zeta_d^{d-1} = 0$. Hence, on both sides of ℓ we get the sums lying on the same axis of symmetry. The moduli of sums are equal to

$$|1 + \zeta_d + \cdots + \zeta_d^{[d/2]-1}| = |(\zeta_d^{[d/2]} - 1)/(\zeta_d - 1)| = \frac{\sin(\pi[d/2]/d)}{\sin(\pi/d)}.$$

This is equal to $1/\sin(\pi/d)$ for d even, and to

$$\cos(\pi/2d)/\sin(\pi/d) = 1/(2\sin(\pi/2d))$$

for d odd. □

Proof of Theorem 4.4: Suppose that $\zeta \in U_d$, where $d \geqslant 2$ is an integer. Since

$\zeta^d = 1$, we can write the value $f(\zeta)$ of the polynomial $f \in \Lambda_{md}$ at $z = \zeta$ as

$$f(\zeta) = f_1(\zeta) + \cdots + f_m(\zeta),$$

where $f_1, \ldots, f_m \in \Lambda_d$. Hence $S(\zeta, md) \leqslant mS(\zeta, d)$. Moreover, if $f_0 \in \Lambda_d$ is a polynomial for which $S(\zeta, d) = |f_0(\zeta)|$ then, by setting $f(z) := f_0(z)(1 + z^d + \cdots + z^{(m-1)d}) \in \Lambda_{md}$, we find that $f(\zeta) = mf_0(\zeta)$. Hence $S(\zeta, md) = mS(\zeta, d)$. It remains to show that $S(\zeta, d) = 1/\sin(\pi/d)$ if d is even and $S(\zeta, d) = 1/(2\sin(\pi/2d))$ if $d > 1$ is odd.

Let f be a Newman polynomial of degree $\leqslant d-1$ for which we have $S(\zeta, d) = |f(\zeta)|$. Put $\mathbf{s} = f(\zeta)$. By Theorem 4.2, $\mathbf{s}$ is the sum of all numbers ζ^j, where $j \in \{0, \ldots, d-1\}$, that lie on one side of a line ℓ orthogonal to $\mathbf{s}$ but not on ℓ itself. Moreover, none of the points ζ^j lies on ℓ. Since $\zeta \in U_d$, the set $\{\zeta^j : j = 0, \ldots, d-1\}$ is precisely the set of roots of $z^d - 1$, i.e., Γ_d. By Lemma 4.6, $|\mathbf{s}| = 1/\sin(\pi/d)$ for d even and $|\mathbf{s}| = 1/(2\sin(\pi/2d))$ for $d > 1$ odd. This completes the proof of the theorem. □

Proof of Theorem 4.5: The case $\zeta = 1$ is obvious. The maximal sum is $1 + \zeta + \cdots + \zeta^{n-1}$, so $S(1, n) = n$ for every positive integer n. Suppose that $\zeta \in U_d$ with $d \geqslant 2$. Choose an integer m such that $md \leqslant n < (m+1)d$. Since $S(\zeta, n)$ is a nondecreasing function in n, we have $S(\zeta, md) \leqslant S(\zeta, n) \leqslant S(\zeta, (m+1)d)$. Thus, by Theorem 4.4, for even $d \geqslant 2$, we have

$$\frac{1 - d/n}{d\sin(\pi/d)} = \frac{n/d - 1}{n\sin(\pi/d)} < \frac{m}{n\sin(\pi/d)} = \frac{S(\zeta, md)}{n} \leqslant \frac{S(\zeta, n)}{n}$$

$$\leqslant \frac{S(\zeta, (m+1)d)}{n} = \frac{m+1}{n\sin(\pi/d)} \leqslant \frac{n/d + 1}{n\sin(\pi/d)} = \frac{1 + d/n}{d\sin(\pi/d)}.$$

It follows that $\lim_{n\to\infty} S(\zeta, n)/n = 1/(d\sin(\pi/d))$ for each even $d \geqslant 2$. The proof of the case when $d > 1$ is odd is similar: one just uses the 'odd' part of Theorem 4.4 instead of its 'even' part.

Finally, suppose that $\zeta = e^{i\phi}$, where $0 < \phi < 2\pi$, is a complex number of modulus 1 which is not a root of unity. Then $\phi/\pi \notin \mathbb{Q}$. Suppose that $\mathbf{s} = f(\zeta) = \sum_{j=0}^{n-1} c_j\zeta^j$ is one of the vectors of maximal length. Then, by Theorem 4.2, $c_j \in \{0, 1\}$ with $c_j = 1$ if and only if the projection of ζ^j to $\mathbf{s}$ is positive. Let ℓ be the line passing through the origin and orthogonal to $\mathbf{s} = |\mathbf{s}|e^{i\tau}$. The line ℓ divides the complex plane into two half planes. Let us divide the open half plane with the point $e^{i\tau}$ into $2M$ equal sectors, where for each $k \in \{-M, \ldots, -1, 1, \ldots, M\}$ the k-th sector consists of complex numbers with arguments in the interval $[\tau + \pi(k-1)/2M, \tau + \pi k/2M)$ for $k > 0$ and in the interval $[\tau + \pi k/2M, \tau + \pi(k+1)/2M)$ for $k < 0$. (Since this half plane needs to be open, one exception is that the interval

corresponding to $k=-M$ is open $(\tau-\pi/2, \tau-\pi(M-1)/2M)$.)

For any $j \in \{0,1,\dots,n-1\}$ the vector ζ^j belongs to the sum $\mathbf{s}$ if and only if it lies in one of the above $2M$ sectors. The sum of the vectors $\zeta^j = \cos(j\phi)+i\sin(j\phi)$ is $f(\zeta) = \mathbf{s} = |\mathbf{s}|e^{i\tau}$, hence $f(\zeta)e^{-i\tau}$ is a real number. Using the fact that the number

$$f(\zeta)e^{-i\tau} = \sum_{j=0}^{n-1} c_j \zeta^j e^{-i\tau} = \sum_{j=0}^{n-1} c_j(\cos(j\phi-\tau) + i\sin(j\phi-\tau))$$

is real, we obtain that $\sum_{j=0}^{n-1} c_j \sin(j\phi-\tau) = 0$, so

$$|f(\zeta)| = f(\zeta)e^{-i\tau} = \sum_{j=0}^{n-1} c_j \cos(j\phi-\tau).$$

Suppose that the sector corresponding to the index k contains n_k vectors of the set $\{1,\dots,\zeta^{n-1}\}$, say, ζ^j with $j \in N_k$, where N_k is a subset of $\{0,1,\dots,n-1\}$ of cardinality n_k. Then $\sum_{j\in N_k}\cos(j\phi-\tau)$ is at least $n_k\cos(|k|\pi/2M)$ and at most $n_k\cos((|k|-1)\pi/2M)$. It follows that

$$\sum_{k=1}^{M}(n_k+n_{-k})\cos(k\pi/2M) \leqslant |f(\zeta)| \leqslant \sum_{k=1}^{M}(n_k+n_{-k})\cos((k-1)\pi/2M).$$

By the theorem of Weyl [214] (see, e.g., Example 2.1 in [123]), the sequence of fractional parts $\{m\phi/2\pi\}$, $m=0,1,2,\dots$, is uniformly distributed in the interval $[0,1)$, because $\phi/2\pi \notin \mathbb{Q}$. Fix $\varepsilon>0$. Then fix any $M=M(\varepsilon)\in\mathbb{N}$ satisfying

$$\frac{1}{4M}\left(1+\frac{1}{\tan(\pi/4M)}\right) < \frac{1+\varepsilon}{\pi} \quad \text{and} \quad \frac{1}{4M}\left(-1+\frac{1}{\tan(\pi/4M)}\right) > \frac{1-\varepsilon}{\pi}.$$

Such an M exists, because $\lim_{x\to\infty} x\tan(\pi/x) = \pi$. Given $k\in\{1,\dots,M\}$, ζ^j belongs to the k-th sector if and only if there is an $l\in\mathbb{Z}$ such that $\tau+\pi(k-1)/2M \leqslant j\phi - 2\pi l < \tau + \pi k/2M$, i.e., $(k-1)/4M \leqslant \{j\phi/2\pi - \tau/2\pi\} < k/4M$. Using uniform distribution of $\{j\phi/2\pi-\tau/2\pi\}$, $j=0,1,\dots$, in $[0,1)$, we deduce that $(1-\varepsilon)n/4M < n_k < (1+\varepsilon)n/4M$ for each sufficiently large $n\in\mathbb{N}$. The same bounds hold for $k\in\{-M,\dots,-1\}$. Hence

$$(1-\varepsilon)\frac{n}{2M}\sum_{k=1}^{M}\cos(k\pi/2M) \leqslant |f(\zeta)| \leqslant (1+\varepsilon)\frac{n}{2M}\sum_{k=1}^{M}\cos((k-1)\pi/2M).$$

Setting $x=\pi/2M$ into the identity

$$1/2+\cos(x)+\cdots+\cos((M-1)x) = \frac{\sin((M-1/2)x)}{2\sin(x/2)},$$

we derive that

$$\sum_{k=1}^{M} \cos((k-1)\pi/2M) = \frac{1}{2}\left(1 + \frac{1}{\tan(\pi/4M)}\right)$$

and

$$\sum_{k=1}^{M} \cos(k\pi/2M) = \frac{1}{2}\left(-1 + \frac{1}{\tan(\pi/4M)}\right).$$

Hence

$$(1-\varepsilon)\frac{n}{4M}\left(-1 + \frac{1}{\tan(\pi/4M)}\right) \leqslant |f(\zeta)| \leqslant (1+\varepsilon)\frac{n}{4M}\left(1 + \frac{1}{\tan(\pi/4M)}\right).$$

By the choice of M, this implies that $(1-\varepsilon)^2 n/\pi \leqslant |f(\zeta)| \leqslant (1+\varepsilon)^2 n/\pi$. Thus

$$(1-\varepsilon)^2/\pi \leqslant S(\zeta, n)/n = |f(\zeta)/n| \leqslant (1+\varepsilon)^2/\pi$$

for each $n \geqslant n(\varepsilon)$. However, ε can be arbitrarily small, so $\lim_{n\to\infty} S(\zeta, n)/n = 1/\pi$, as claimed. □

4.4 Practical computations

Take $\zeta = \exp(2\pi i/5)$ and $n = 5$. By Lemma 4.6, we can take any ℓ which goes through none of the roots of $z^5 - 1 = 0$. Take ℓ such that 1 and ζ are on one of its sides. Then, by Lemma 4.6, we find that $|1+\zeta| = 1/(2\sin(\pi/10)) = (1+\sqrt{5})/2 = 1.61803\ldots$.

Similarly, taking $\zeta = \exp(9\pi i/7)$ to be one of the roots of $z^{14} - 1 = 0$ and $n = 14$, one can choose ℓ to be the imaginary axis. Then one of the extremal Newman polynomials will be $f(z) = 1 + z^3 + z^5 + z^6 + z^8 + z^9 + z^{11}$, because $0, 3, \ldots, 11$ are the only powers of ζ that are on the right hand side of ℓ. Lemma 4.6 and Theorem 4.4 gives $f(\zeta) = 1/\sin(\pi/14) = 4.49395\ldots$.

Take $\zeta = i$ and $n = 5$. By Theorem 4.2, there are four possible quadrants for the location of $\mathbf{s}$. The maximum for $|f(i)|$ is attained by Newman polynomials $1 + z + z^4$ and $1 + z^3 + z^4$, giving $\mathbf{s} = 2 \pm i$. Hence $S(i, 5) = \sqrt{5}$. Note that the maximal vectors $2 \pm i$ do not lie on an axis of symmetry of the square with vertices $1, i, -1, -i$. So Lemma 4.6 does not hold, because there is one 'double' vector $1 = i^4$.

It seems likely that when ζ is not a root of unity, one cannot expect any simple formula for $S(\zeta, n)$. For example, for ζ satisfying $\zeta^2 - \zeta/2 + 1 = 0$, we calculated the value $S(\zeta, 100) = 31.8928\ldots$. It is easy to see that $S(\zeta, 100)/100 = 0.31892\ldots$ is quite close to the limit value $1/\pi = 0.31830\ldots$, given by Theorem 4.5. The value

$S(\zeta, 100)$ is attained by the polynomial $f(z) = z^{97} + z^{96} + z^{95} + z^{92} + z^{91} + z^{90} + z^{87} + z^{86} + z^{82} + z^{81} + z^{78} + z^{77} + z^{76} + z^{73} + z^{72} + z^{71} + z^{68} + z^{67} + z^{63} + z^{62} + z^{58} + z^{57} + z^{54} + z^{53} + z^{52} + z^{49} + z^{48} + z^{44} + z^{43} + z^{39} + z^{38} + z^{35} + z^{34} + z^{33} + z^{30} + z^{29} + z^{28} + z^{25} + z^{24} + z^{20} + z^{19} + z^{16} + z^{15} + z^{14} + z^{11} + z^{10} + z^9 + z^6 + z^5 + z + 1$.

Finally, we remark that our results may be applied to polynomials with coefficients in any real interval $[a, b]$. In this case, if $\zeta \neq 1$, the constant factor $b - a$ will appear on the right hand side of the formulas established by Theorems 4.4 and 4.5. Indeed, any polynomial $f(z) = \sum_{j=0}^{n-1} c_j z^j$ with coefficients $c_j \in [a, b]$ can be written as

$$f(z) = (b - a)g(z) + ah(z),$$

where $g(z) = \sum_{j=0}^{n-1}((c_j - a)/(b - a))z^j$ is a polynomial with coefficients in $[0, 1]$ and $h(z) = 1 + \cdots + z^{n-1} = (z^n - 1)/(z - 1)$. Now, $h(\zeta) = 0$ if $\zeta \neq 1$ is an nth root of unity. Furthermore, $|h(\zeta)|$ is bounded by an absolute constant depending on ζ only if $|\zeta| \leqslant 1$ and $\zeta \neq 1$, so that $|h(\zeta)|/n \to 0$ as $n \to \infty$. Taking $n = d$, Theorem 4.4 may be applied immediately to $g(z)$. To obtain a corresponding limit in Theorem 4.5, one can divide the equality by n, and then let $n \to \infty$.

Chapter 5

Newman and Littlewood polynomials

5.1 Statement of the problem

Let $V_{\mathcal{N}}$ and $V_{\mathcal{L}}$ be sets of roots of Newman and Littlewood polynomials, respectively. We refer to the numbers in the set $V_{\mathcal{N}}$ as *Newman numbers*, while the numbers in the set $V_{\mathcal{L}}$ will be called *Littlewood numbers*. This notation has originated in the doctoral Thesis of P. Drungilas, [61]. Let also V be the set of roots of polynomials P with coefficients in the set $\{-1, 0, 1\}$ and $P(0) \neq 0$.

The sets $V_{\mathcal{N}}, V_{\mathcal{L}}$ and V have been investigated by several authors, e.g., [42], [63], [161]. It is well known that the set $V_{\mathcal{N}}$ is contained in the intersection of the annulus $1/\phi < |z| < \phi$ with $\Re(z) < 3/2$, where $\phi = (1+\sqrt{5})/2$. A more precise bounding contour was given in [161], where it was also shown that the closure of this set $\overline{V_{\mathcal{N}}}$ is path-connected. The points of $\overline{V_{\mathcal{L}}}$ inside the unit circle are related to the points of vanishing of power series with ± 1 coefficients. Beacoup, Borwein, Boyd and Pinner studied the extremal zeros of such power series and their multiplicity in [19] and [20].

Clearly, every $\alpha \in V$ is an algebraic integer. Moreover, it is a unit, and it is not difficult to show that all such α are located in the annulus $1/2 < |z| < 2$. The converse of this statement does not hold, namely, there are many units α that lie with their conjugates in the annulus $1/2 < |z| < 2$ but $\alpha \notin V$. For instance, the minimal polynomial $P(x) = x^4 + x^3 + 2x^2 - x + 1$ of the number $\theta = (-1 + i\sqrt{3})(1+\sqrt{5})/4$ does not divide any polynomial with coefficients $\{-1, 0, 1\}$ (see [63]).

It is evident that $V_{\mathcal{N}} \subseteq V$ and $V_{\mathcal{L}} \subseteq V$. Moreover, $V_{\mathcal{N}}$ is a proper subset of V, because $V_{\mathcal{N}}$ contains no positive numbers, whereas, say, $1 \in V_{\mathcal{L}} \subseteq V$. In order to show that $V_{\mathcal{L}}$ is a proper subset of V it would be sufficient to prove that there is

an $\alpha \in V_{\mathcal{N}}$ which is not in $V_{\mathcal{L}}$. This would imply that $V_{\mathcal{N}}$ is not a subset of $V_{\mathcal{L}}$, so that all three sets $V_{\mathcal{N}}$, $V_{\mathcal{L}}$ and V are distinct and both sets $V_{\mathcal{N}} \setminus V_{\mathcal{L}}$ and $V_{\mathcal{L}} \setminus V_{\mathcal{N}}$ are not empty. We shall formulate now the main problem of Chapter 5.

Problem 5.1. *Does there exist an algebraic number which is a Newman number but is not a Littlewood number? In other words, is it true that* $V_{\mathcal{N}} \not\subset V_{\mathcal{L}}$*?*

For this, it suffices to show that there is an irreducible polynomial $P(x) \in \mathbb{Z}[x], P(0) \neq 0$, which divides some Newman polynomial but does not divide any Littlewood polynomial. However, it seems that the problem of finding such examples is non-trivial, so the first named author posed this question as an open problem 006:07 at the 2006 West Coast Number Theory conference. Below, we shall use a numerical algorithm (see Theorem 5.5) to determine whether a given polynomial $P(x) \in \mathbb{Z}[x]$ with at least one zero outside the unit circle divides some Littlewood polynomial or not. In particular, using this test, we will show that the irreducible Newman polynomial $x^9 + x^6 + x^2 + x + 1$ of degree 9 does not divide any Littlewood polynomial. In addition, it will be shown that there are no such Newman polynomials of degree at most 8.

Divisibility properties of polynomials with coefficients $\{-1, 0, 1\}$ have been studied on many occasions, since they have many applications to various diophantine problems. For example, the order of vanishing of such polynomials at 1 was studied in [9] and the multiplicity of cyclotomic and non-cyclotomic factors of such polynomials in [38] and [149]. The paper [149] was partly motivated by the hope to establish an absolute upper bound B for the multiplicity of a non-cyclotomic factor P in the factorization of polynomials with coefficients $\{-1, 0, 1\}$. Such a bound would lead to the proof of Lehmer's conjecture on Mahler's measure. If the bound B exists, then $B \geqslant 4$. Recently, several new results on Lehmer's conjecture have been obtained in [33] and [35]. In [33] Lehmer's conjecture was confirmed for polynomials with odd coefficients, so, in particular, for Littlewood polynomials. See also [72] for better numerical estimates.

The above mentioned papers contain some interesting examples which are in some sense 'special cases' of our problem. For instance, in [33] it was observed that if $P(x) \in \mathbb{Z}[x]$ is not a product of cyclotomic polynomials Φ_m modulo 2 and P divides some Littlewood polynomial L, then the quotient $Q = L/P$ has Mahler's measure greater than 1. In other words, Q cannot be a product of cyclotomic polynomials. Similarly, from the result on the Mahler's measure of Littlewood polynomials given in [35] it follows that if $1 < M(P) < (1 + \sqrt{5})/2$ and P divides a non-reciprocal Littlewood polynomial L, then L/P must have at least one non-cyclotomic factor. Otherwise, we have $M(L) = M(P) < (1 + \sqrt{5})/2$, where L is a non-reciprocal Littlewood polynomial, which is impossible by [35]. Mossinghoff

[149] found Littlewood polynomials divisible by ℓ and $\{-1,0,1\}$ polynomials divisible by ℓ^3, where $\ell(x) = x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1$ is Lehmer's polynomial.

Chapter 5 is organized as follows. The main results are given in Sections 5.2 and 5.3. In Section 5.4 we give an auxiliary result for polynomials over the finite field $\mathbb{F}_2$ with two elements. Section 5.5 contains the proofs of most of our theoretical results. We then develop some algorithms used in our computations (see Section 5.6). The details of computations are provided in Section 5.7.

5.2 Main results

Note that the sets $V_{\mathcal{N}}$ and $V_{\mathcal{L}}$ have infinitely many common elements. For example, the root of unity $\zeta = e^{2\pi i/u}$, where $u \geqslant 2$, is a root of $x^{u-1} + \cdots + x + 1$. This is a Newman polynomial and a Littlewood polynomial. So every root of unity except for 1 belongs to $V_{\mathcal{N}} \cap V_{\mathcal{L}}$. Our next result implies that the structure of the set $V_{\mathcal{N}} \cap V_{\mathcal{L}}$ is non-trivial: every Newman trinomial divides some Littlewood polynomial. (Trinomial is a polynomial with three non-zero coefficients.) In fact, our statement is more general:

Theorem 5.2. *For each trinomial P with $\{-1,0,1\}$ coefficients and $P(0) \neq 0$, there exists a polynomial Q with coefficients $\{-1,0,1\}$ such that the product PQ is a Littlewood polynomial.*

Similar results for certain special quadrinomials will be given in Section 5.3. Our computations show that the roots of Newman polynomials of degree at most 8 also belong to $V_{\mathcal{N}} \cap V_{\mathcal{L}}$:

Theorem 5.3. *Every Newman polynomial of degree at most 8 divides some Littlewood polynomial.*

We also have the following:

Theorem 5.4. *Every Newman polynomial divides some integer polynomial with odd coefficients.*

The main purpose of our work is to give some examples of algebraic numbers $\alpha \in V_{\mathcal{N}} \setminus V_{\mathcal{L}}$. We found several irreducible Newman polynomials not dividing any Littlewood polynomial. Some of them are given in Table 1. Polynomials given in rows $2, 4, 6, 8$ are reciprocal to polynomials given in rows $1, 3, 5, 7$, respectively.

All these polynomials were found using numerical tests based on the following statement:

Table 5.1: Some Newman polynomials of degree 9 not dividing any Littlewood polynomial.

	Polynomial $P(x)$
1.	$1 + x^4 + x^6 + x^7 + x^9$
2.	$1 + x^2 + x^3 + x^5 + x^9$
3.	$1 + x^3 + x^7 + x^8 + x^9$
4.	$1 + x + x^2 + x^6 + x^9$
5.	$1 + x + x^2 + x^4 + x^6 + x^9$
6.	$1 + x^3 + x^5 + x^7 + x^8 + x^9$
7.	$1 + x + x^4 + x^5 + x^6 + x^7 + x^9$
8.	$1 + x^2 + x^3 + x^4 + x^5 + x^8 + x^9$

Theorem 5.5. *Let $P(x) \in \mathbb{Z}[x]$ be a monic polynomial with roots $\alpha_1, \ldots, \alpha_k$, of modulus strictly greater than 1, where $k \geqslant 1$. Suppose that there exist a positive integer N and a real number $\delta \geqslant 0$ with the property that, for each of the 2^N vectors $\mathbf{b} = (b_1, \ldots, b_N)$, where $b_1, \ldots, b_N \in \{-1, 1\}$, there are two positive integers $n = n(\mathbf{b}) \leqslant N$ and $i = i(\mathbf{b}) \leqslant k$ such that*

$$(|\alpha_i| - 1)|\alpha_i^n + b_1\alpha_i^{n-1} + \cdots + b_n| \geqslant 1 + \delta.$$

Then P does not divide any Littlewood polynomial.

At the first glance, the statement of the theorem may look somewhat strange, because one obtains the weakest inequality when $\delta = 0$. However, on several occasions below, we shall use the statement of the theorem with strictly positive δ. This is why we prefer to state the theorem in the above form.

Using the examples from Table 1 it is possible to construct infinitely many irreducible Newman polynomials not dividing any Littlewood polynomial. This shows that set $V_{\mathcal{N}} \setminus V_{\mathcal{L}}$ is infinite. We will prove the following statement:

Theorem 5.6. *There exist infinitely many primitive irreducible Newman polynomials which do not divide any Littlewood polynomial.*

In this context, a polynomial $P(x) \in \mathbb{Z}[x]$ is called *primitive* if P cannot be written as $P(x) = G(x^k)$ with some integer $k \geqslant 2$ and some $G(x) \in \mathbb{Z}[x]$.

Take $P(x) = 1 + x^4 + x^6 + x^7 + x^9$. Since $P(x)$ does not divide any Littlewood polynomial, the polynomial $P(-x) = 1 + x^4 + x^6 - x^7 - x^9$ also does not divide any Littlewood polynomial. The polynomial $1 + x^4 + x^6 - x^7 - x^9$ has a positive root α, so it does not divide any Newman polynomial. This shows that $\alpha \in V$, but $\alpha \notin V_{\mathcal{L}} \cup V_{\mathcal{N}}$, so $V_{\mathcal{L}} \cup V_{\mathcal{N}}$ is strictly contained in V.

5.3 Other results

For each complex root α of the polynomial $P(x) \in \mathbb{R}[x]$, the complex conjugate $\overline{\alpha}$ is also a root of P. If P is a Newman (resp. Littlewood) polynomial, then its reciprocal P^* is also a Newman (resp. Littlewood) polynomial. Thus each of the sets $V_{\mathcal{L}}$, $V_{\mathcal{N}}$, $V_{\mathcal{N}} \cap V_{\mathcal{L}}$, $V_{\mathcal{N}} \setminus V_{\mathcal{L}}$, $V_{\mathcal{L}} \setminus V_{\mathcal{N}}$ map into itself by the complex conjugation $z \mapsto \overline{z}$ and the inversion $z \mapsto 1/z$. In the next statement we consider the map $z \mapsto z^{1/k}$.

Lemma 5.7. *Let k be a positive integer. Then $P(x) \in \mathbb{Z}[x]$ divides some Littlewood polynomial if and only if $P(x^k)$ divides some Littlewood polynomial.*

By the same method as that in the proof of Theorem 5.2, one can show that certain quadrinomials also divide polynomials with small odd coefficients and sometimes even Littlewood polynomials.

Theorem 5.8. *Let P be a quadrinomial with coefficients in $\{-1, 0, 1\}$ such that $P(0) = 1$. Then there is a Newman polynomial Q such that all coefficients of the product PQ belong to the set $\{-3, -1, 1, 3\}$ and, moreover, to the set $\{1, 3\}$ if P itself is a Newman polynomial. Furthermore, if $a < b < c$ are positive integers and P is of one of the forms*

i) $1 + x^a - x^b - x^c$,

ii) $1 - x^a - x^b - x^c$, *where exactly one of the numbers a, b, c is odd,*

iii) $1 + x^a + x^b + x^c$, *where exactly one of the numbers a, b, c is even,*

iv) $1 + x^a + x^b - x^c$, *where all the exponents a, b, c are odd or, alternatively, c is even and precisely one of the numbers a, b is odd,*

then the quadrinomial P divides some Littlewood polynomial L, and L/P is a polynomial with coefficients in $\{-1, 0, 1\}$.

It would be of interest to find out whether this result can be extended to the full analogue of Theorem 5.2, namely, to all quadrinomials of height 1. If so, then this would imply that our example $x^9+x^6+x^2+x+1$ is minimal not only in terms of its degree (nine), but also in terms of the number of its non-zero coefficients (five).

Suppose that $P(x) \in \mathbb{Z}[x]$ divides some Littlewood polynomial L. One may ask which values its degree $\deg L$ can take. The answer is given in terms of factorization of L modulo 2.

Lemma 5.9. *Suppose that a polynomial $P(x) \in \mathbb{Z}[x]$ divides a Littlewood polynomial L. Let $\widetilde{P}(x) \in \mathbb{F}_2[x]$ be the reduction of P modulo 2. Then $\deg L + 1$ is a multiple of $\deg_2 \widetilde{P}$. (This quantity will be defined in Section 5.4.)*

In fact, the value of $\deg L$ grows exponentially with the degree of P. If, for instance, a monic polynomial P of degree 10 is a prime divisor of the cyclotomic polynomial Φ_{1023} in $\mathbb{F}_2[x]$, then $\deg_2 \tilde{P} = 2^{10} - 1$. The degree of any Littlewood polynomial L divisible by P must be of the form $1023k - 1$, where $k \in \mathbb{N}$, so it is greater than or equal to 1022. One has thus to consider 2^{1022} different possibilities in trying to find a polynomial L of degree 1022 divisible by P. This simple example demonstrates the computational complexity of the problem.

One possible strategy is to search for a factor $Q(x) \in \mathbb{Z}[x]$ of small height, say, $H(Q) \leqslant 2$ such that the product PQ is a Littlewood polynomial. The following lemma implies that one can restrict himself with only finitely many choices for Q. This will be used in Algorithm 5.15 below.

Lemma 5.10. *Suppose that a polynomial $P(x) \in \mathbb{Z}[x]$ of degree $d \geqslant 1$ divides a Littlewood polynomial L. Let $h = H(L/P)$. Then there exists a polynomial $Q(x) = \sum_{j=0}^{n} b_j x^j \in \mathbb{Z}[x]$ of degree*

$$n \leqslant (2h+1)^d + d - 2$$

and height $H(Q) \leqslant h$ such that the product PQ is also a Littlewood polynomial. Moreover, the vector of coefficients $(b_0, b_1, b_2, \ldots, b_{n-1}, b_n)$ of Q does not contain any identical blocks $b_j, b_{j+1}, \ldots, b_{j+d-1}$ of length d.

Suppose that a Newman polynomial P does not divide any Littlewood polynomial. We shall construct infinitely many examples of such polynomials by perturbing the roots of P.

Theorem 5.11. *Suppose that P satisfies the conditions of Theorem 5.5 with some $\delta > 0$. Then there exists an $\varepsilon > 0$ which depends on P and δ only with the following property: if the polynomial $P_1(x) \in \mathbb{Z}[x]$ has some roots $\beta_1, \ldots, \beta_k$, each of modulus strictly greater than 1 such that $|\alpha_j - \beta_j| < \varepsilon$ for $j = 1 \ldots k$, where $\alpha_1, \ldots, \alpha_k$ are the roots of P of modulus strictly greater than 1, then P_1 does not divide any Littlewood polynomial.*

Such approximations may be obtained from the sequence of polynomials of the form $x^n P(x) + R(x)$, where R is a Newman polynomial relatively prime to P and $n > \deg R$.

Theorem 5.12. *Suppose that the polynomial P satisfies the conditions of Theorem 5.5 with some $\delta > 0$. Then, for any $R(x) \in \mathbb{Z}[x]$, there exists a positive integer n_0 such that, for each $n \geqslant n_0$, the polynomial $x^n P(x) + R(x)$ does not divide any Littlewood polynomial.*

All polynomials given in Table 1 have at least two zeros outside the unit circle. It would be of interest to find out whether there exists an irreducible polynomial $P(x) \in \mathbb{Z}[x]$ with exactly one root outside the unit circle such that P divides some Newman polynomial but no Littlewood polynomial. In other words, does there exist a Pisot or a Salem number α such that $-\alpha$ is a root of some Newman polynomial but not a root of any Littlewod polynomial?

5.4 Auxiliary facts about polynomials from $\mathbb{F}_2[x]$

Every polynomial $f(x) \in \mathbb{F}_2[x]$ with $f(0) \neq 0$ modulo 2 may be written uniquely as a product

$$f(x) = (x+1)^m \prod_{j=1}^{r} \phi_j(x)^{m_j},$$

where $m \geqslant 0$ and $\phi_j(x) \in \mathbb{F}_2[x]$ are irreducible polynomials of degree greater than or equal to 2 and multiplicity $m_j \geqslant 1, j = 1, \ldots, r$. The product is empty if $r = 0$. Every polynomial ϕ_j divides a unique cyclotomic polynomial Φ_{e_j} of odd index e_j. Let s be the least positive integer satisfying $2^s \geqslant \max\{m+1, m_1, \ldots, m_r\}$. Define the number

$$\deg_2 f = 2^s \mathrm{lcm}(e_1, \ldots, e_r).$$

Lemma 5.13. *If a polynomial $f(x) \in \mathbb{F}_2[x]$ divides the polynomial $h(x) = x^n + \ldots + x + 1$, then $n+1$ is divisible by the number $\deg_2 f$. Conversely, if $\deg_2 f$ divides $n+1$, then there exists a polynomial $g(x) \in \mathbb{F}_2[x]$ such that $f(x)g(x) = h(x)$.*

Proof: Write $h(x) = (x^{n+1}+1)/(x+1)$ in $\mathbb{F}_2[x]$. Let $n+1 = 2^l k$, where k is odd and $l \geqslant 0$. Then, in $\mathbb{F}_2[x]$, we have

$$h(x) = (x^k+1)^{2^l}/(x+1).$$

Let α be a root of the irreducible factor $\phi_j(x)$ of f. Note that the order of α in the multiplicative group of the field $\mathbb{F}_{2^{\deg \phi_j}}$ is e_j, so e_j divides k for every $j = 1, \ldots, r$. The polynomial $x^k + 1$ has no multiple roots, therefore the power 2^l is greater than or equal to the maximum of the numbers $m+1, m_1, \ldots, m_r$. Hence $n+1$ must be divisible by 2^s and the least common multiple of the integers $e_1, \ldots, e_r$. On the other hand, if we take $n = n_1 \deg_2 f - 1$, for some positive integer n_1, then $h(x)$ vanishes at all roots of $f(x)$ with required multiplicities. Thus $h(x)$ is divisible by $f(x)$, and $g(x) \in \mathbb{F}_2[x]$ is the quotient $h(x)/f(x)$. □

We shall give an example of the computation of $\deg_2 \widetilde{P}$. Consider the polynomial $P(x) = 1 + x^2 + x^5 + x^9 \in \mathbb{Z}[x]$. Reducing modulo 2, the polynomial splits

over $\mathbb{F}_2$ into the following irreducible factors

$$\widetilde{P}(x) = (x+1)^2(x^2+x+1)(x^5+x^4+x^3+x+1).$$

In this example, we have $m = 2$, $r = 2$, $m_1 = m_2 = 1$, $\phi_1(x) = x^2+x+1$ and $\phi_2(x) = x^5+x^4+x^3+x+1$. The polynomial ϕ_1 is the cyclotomic polynomial Φ_3. The polynomial ϕ_2 divides the cyclotomic polynomial Φ_{31}. Hence $e_1 = 3, e_2 = 2^5-1 = 31$ and $s = 2$. Thus $\deg_2 \widetilde{P} = 2^2\mathrm{lcm}(3,31) = 372$. Therefore, by Lemma 5.9, any Littlewood polynomial L divisible by $1+x^2+x^5+x^9$, must be of degree $\deg L = 372k-1$, where $k = 1, 2, \ldots$.

Two of our statements are very simple corollaries of Lemma 5.13.

Proof of Lemma 5.9. The reduction of any Littlewood polynomial L modulo 2 is $\widetilde{L}(x) = x^{\deg L} + \ldots + x + 1$. Since $\widetilde{L}$ is divisible by $\widetilde{P}$, the result follows from Lemma 5.13. □

Proof of Theorem 5.4. Let P be a Newman polynomial of degree b. By Lemma 5.13, there exists a polynomial $\widetilde{Q}(x) = \sum_{j=0}^{n-b} \widetilde{q}_j x^j \in \mathbb{F}_2[x]$, satisfying $\widetilde{P}(x)\widetilde{Q}(x) = x^n+\ldots+x+1$ in $\mathbb{F}_2[x]$. Set $Q(x) = \widetilde{Q}(x)$, where 0 and 1 are understood as positive integers rather than elements of $\mathbb{F}_2$. It follows that PQ has all odd coefficients. □

5.5 Proofs

Proof of Lemma 5.7. The proof of the lemma is similar to the proof of Proposition *(iv)* in [187]. If the polynomial $P(x) \in \mathbb{Z}[x]$ divides some Littlewood polynomial L, and Q is a Littlewood polynomial of degree $k-1$, then the product $L(x^k)Q(x)$ is a Littlewood polynomial divisible by $P(x^k)$. One can take, for instance, $Q(x) = 1+x+\cdots+x^{k-1}$.

For the converse, suppose that $P(x^k)$ divides a Littlewood polynomial L. Rewrite $L(x)$ putting the powers x^i, x^j satisfying $i \equiv j \pmod k$ together:

$$L(x) = L_0(x^k) + xL_1(x^k) + \ldots + x^{k-1}L_{k-1}(x^k).$$

Note that each $L_j(x), j = 0\ldots, k-1$, is either a Littlewood polynomial or zero. For each $0 \leqslant j \leqslant k-1$ there exist $Q_j, R_j \in \mathbb{Z}[x]$, such that $L_j = PQ_j + R_j$, where $\deg R_j < \deg P$. Since $P(x^k)|L(x)$, it follows that $P(x^k)$ divides $R(x) = R_0(x^k)+xR_1(x^k)+\ldots+x^{k-1}R_{k-1}(x^k)$. The degree of R is $\leqslant k(\deg P-1)+k-1$, so $\deg R < k \deg P$. Hence all the polynomials R_j must be zeros identically. This implies that all non-zero polynomials L_j are Littlewood polynomials divisible by P. (There must be at least one non-zero L_j, because L is non-zero.) □

Proof of Lemma 5.10. Let $P(x) = \sum_{j=0}^{d} a_j x^j$. Among all polynomials Q of height $H(Q) \leqslant h$ such that product PQ is a Littlewood polynomial, there is a polynomial of minimal degree, say, $Q(x) = \sum_{j=0}^{n} b_j x^j$. Write $P(x)Q(x) = L(x)$, where all coefficients of L are ± 1.

We begin from the second part of the statement. Suppose that the vector of coefficients of the polynomial Q, $(b_0, b_1, \ldots, b_{n-1}, b_n)$, contains two identical blocks $b_r, b_{r+1}, \ldots, b_{r+d-1}$ and $b_s, b_{s+1}, \ldots, b_{s+d-1}$ of length d, where $r < s$. After removing $s-r$ coefficients $b_r, b_{r+1}, \ldots, b_{s-1}$ from this vector, we obtain the vector $(b_0, \ldots, b_{r-1}, b_s, \ldots, b_n)$. Define the polynomial $T(x) = \sum_{j=0}^{n-(s-r)} t_j x^j$ by

$$t_j = \begin{cases} b_j & \text{if } j < r, \\ b_{j+(s-r)} & \text{if } j \geqslant r. \end{cases}$$

Since $b_{r+j} = b_{s+j}$ for $j = 0, \ldots, d-1$, the first $r+d$ coefficients of T and Q coincide, $t_j = b_j, 0 \leqslant j \leqslant r+d-1$. Hence $Q(x) \equiv T(x) \pmod{x^{r+d}}$. Similarly, the last $n-s+1$ coefficients of Q and T are equal. So, for their reciprocal polynomials, we have $Q^*(x) \equiv T^*(x) \pmod{x^{n-s+1}}$. It follows that $L(x) = P(x)Q(x) \equiv P(x)T(x) \pmod{x^{r+d}}$ and $L^*(x) = P^*(x)Q^*(x) \equiv P^*(x)T^*(x) \pmod{x^{n-s+1}}$. Hence the first $r+d$ and the last $n-s+1$ coefficients of L and PT are the same. But PT has precisely $n-s+r+d+1$ coefficients, so each of those coefficients must be ± 1. Hence PT is a Littlewood polynomial. It follows that $\deg T < \deg Q$, which is a contradiction with the minimality of $\deg Q$.

Now, we turn to the first part of the statement. Since the right hand side of the inequality, $(2h+1)^d + d - 2$, is greater than d for every $d \geqslant 1$, we may assume $n > d$. The number of blocks of length d in the vector of coefficients of Q is $n-d+2$. On the other hand, this number must be less than or equal to the total number of different possible blocks, otherwise two of them will be identical, which is already proved to be impossible. By choosing any element of the block from the set of $2h+1$ integers $\{-h, \ldots, 0, \ldots, h\}$, one obtains exactly $(2h+1)^d$ different blocks. This implies that $n-d+2 \leqslant (2h+1)^d$, as claimed. □

Proof of Theorem 5.2. Without loss of generality, we may assume that the constant coefficient of P is 1 (otherwise multiply P by -1). The trinomial P has one of the four forms

$$\textit{(i)}\ 1 - x^a + x^b, \qquad \textit{(ii)}\ 1 + x^a - x^b, \qquad \textit{(iii)}\ 1 - x^a - x^b, \qquad \textit{(iv)}\ 1 + x^a + x^b,$$

where $a < b$ are two positive integers.

Write $P(x) = 1 + \varepsilon_a x^a + \varepsilon_b x^b$, where the coefficients $\varepsilon_a, \varepsilon_b \in \{-1, 1\}$. We first consider the cases *(i)-(iii)*, when at least one of the coefficients $\varepsilon_a, \varepsilon_b$ is negative.

The reduction of the polynomial P mod 2 is $\widetilde{P}(x) = 1 + x^a + x^b$. By Lemma 5.13, there exists a polynomial $\widetilde{Q}(x) = \sum_{j=0}^{n-b} \widetilde{q}_j x^j \in \mathbb{F}_2[x]$, satisfying $\widetilde{P}(x)\widetilde{Q}(x) = x^n + \ldots + x + 1$ in $\mathbb{F}_2[x]$ provided that $n+1$ is divisible by $\deg_2 \widetilde{P}$. Take the number $n = \deg_2 \widetilde{P} - 1$ to obtain the polynomial of the least possible degree. Define the polynomial $Q(x) = \sum_{j=0}^{n-b} q_j x^j \in \mathbb{Z}[x]$ by

$$q_j = \begin{cases} 0 & \text{if } \widetilde{q}_j = 0, \\ 1 & \text{if } \widetilde{q}_j = 1, \end{cases}$$

so that $\widetilde{Q}(x)$ is a reduction mod 2 of the polynomial $Q(x)$.

Writing $P(x)Q(x) = (1 + \varepsilon_a x^a + \varepsilon_b x^b) \sum_{j=0}^{n-b} q_j x^j = L(x) = \sum_{j=0}^{n} l_j x^j$, we see that the coefficients $l_j \in \mathbb{Z}, j = 0, \ldots, n$, are given by the formula

$$l_j = \begin{cases} q_j & \text{for } 0 \leqslant j < a, \\ q_j + \varepsilon_a q_{j-a} & \text{for } a \leqslant j < b, \\ q_j + \varepsilon_a q_{j-a} + \varepsilon_b q_{j-b} & \text{for } b \leqslant j \leqslant n-b, \\ \varepsilon_a q_{j-a} + \varepsilon_b q_{j-b} & \text{for } n-b < j \leqslant n-b+a, \\ \varepsilon_b q_{j-b} & \text{for } n-b+a < j \leqslant n. \end{cases}$$

The third line is excluded in case $n < 2b$. Since $L(x) \equiv \widetilde{P}(x)\widetilde{Q}(x) \equiv x^n + \ldots + x + 1 \pmod 2$, all the coefficients l_j are odd. There are at most three non-zero terms in the formula for l_j, so $l_j \in \{-3, -1, 1, 3\}$. Note that $l_j = \pm 3$ may appear only in the third line when all three terms q_j, $\varepsilon_a q_{j-a}$ and $\varepsilon_b q_{j-b}$ are 1 or all three -1. This is impossible, because $q_j, q_{j-a}, q_{j-b} \in \{0, 1\}$ and at least one of $\varepsilon_a, \varepsilon_b$ is negative. Thus PQ is a Littlewood polynomial, where Q is a Newman polynomial.

Now consider the remaining case *(iv)*, where $P(x) = 1 + x^a + x^b$. Write $k = \gcd(a, b)$. Then $a = ka_1, b = kb_1$. At least one of the integers a_1, b_1 is odd. Note that $P(x) = P_1(x^k)$, where $P_1(x) = 1 + x^{a_1} + x^{b_1}$. The polynomial $P_1(-x)$ has one of the forms *(i)*, *(ii)* or *(iii)*. It follows from the earlier part of the proof that there exists a polynomial $Q_1(x)$ with coefficients 0 or 1, such that $P_1(-x)Q_1(x)$ is a Littlewood polynomial. Thus $P_1(x)Q_1(-x)$ is a Littlewood polynomial, so that

$$P_1(x^k)Q_1(-x^k)(1 + x + \ldots + x^{k-1}) = P(x)Q_1(-x^k)(1 + x + \ldots + x^{k-1})$$

is also a Littlewood polynomial. Clearly, in this case, the factor $Q_1(-x^k)(1 + \cdots + x^{k-1})$ is a polynomial with $\{-1, 0, 1\}$ coefficients. □

Proof of Theorem 5.8. The proof is very similar to the proof of Theorem 5.2 for trinomials, thus we will omit the details. For a given quadrinomial P, there exists

a $Q(x) \in \mathbb{Z}[x]$ with $0, 1$ coefficients, such that $P(x)Q(x) \equiv x^n + \ldots + x + 1 \pmod 2$. In the formula for the coefficients l_j of the polynomial $L = PQ$, there are at most four non-zero terms and all of the l_j must be odd, by the choice of Q. Hence $l_j \in \{-3, -1, 1, 3\}$. Moreover, $l_j \in \{1, 3\}$ if all the coefficients of P are non-negative. This proves the first part of the theorem.

Suppose that exactly two coefficients of the quadrinomial P are 1 and other two are -1. (For any quadrinomial P listed in *(i)-(iv)* either $P(x)$ or $P(-x)$ has this property.) The number $l_j = \pm 3$ may appear only in equations with three or four non-zero terms (see the formula for the coefficients l_j). This is impossible, because two of all non-zero terms have opposite signs. Therefore, for any polynomial P as in *(i)-(iv)*, $P(x)Q(x)$ or $P(x)Q(-x)$ must be a Littlewood polynomial. □

Proof of Theorem 5.5. Suppose that there is a Littlewood polynomial L which is divisible by P. Since, for any positive integer m, $L(x)(1+x^{\deg L+1}+\cdots+x^{m(\deg L+1)})$ is a Littlewood polynomial too, we can assume without loss of generality that $\deg L \geqslant N$. Write $L(x) = x^M + b_1 x^{M-1} + \cdots + b_M$, where $b_j \in \{-1, 1\}$ and $M \geqslant N$. By the assumption of the theorem, there exist positive integers $n \leqslant N$ and $i \leqslant k$ such that $(|\alpha_i| - 1)|\alpha_i^n + b_1\alpha_i^{n-1} + \cdots + b_n| \geqslant 1 + \delta$.

On the other hand, using the fact that L is divisible by P, we have $L(\alpha_i) = 0$. Hence $\alpha_i^n + b_1\alpha_i^{n-1} + \cdots + b_n = -(b_{n+1}\alpha_i^{-1} + \cdots + b_M\alpha_i^{n-M})$. Thus

$$|\alpha_i^n + b_1\alpha_i^{n-1} + \cdots + b_n| = |b_{n+1}\alpha_i^{-1} + \cdots + b_M\alpha_i^{n-M}| \leqslant$$

$$\leqslant \sum_{j=1}^{n-M} |\alpha_i|^{-j} < \sum_{j=1}^{\infty} |\alpha_i|^{-j} = 1/(|\alpha_i| - 1),$$

giving $(|\alpha_i| - 1)|\alpha_i^n + b_1\alpha_i^{n-1} + \cdots + b_n| < 1$, a contradiction. □

Proof of Theorem 5.11. By Theorem 5.5, for each of the 2^N vectors

$$\mathbf{b} = (b_1, \ldots, b_N) \in \{-1, 1\}^N,$$

there exist positive integers $n = n(\mathbf{b}) \leqslant N$ and $i = i(\mathbf{b}) \leqslant k$ such that the function

$$f_{\mathbf{b}}(z) = (|z| - 1)|z^n + b_1 z^{n-1} + \cdots + b_n|$$

is greater than $1 + \delta$ at α_i, namely, $f_{\mathbf{b}}(\alpha_i) \geqslant 1 + \delta$. By continuity of $f_{\mathbf{b}}(z)$, the inequality $f_{\mathbf{b}}(z) > 1$ holds for all z in the circle $|z - \alpha_i| < \varepsilon_{\mathbf{b}}$. Here, $\varepsilon_{\mathbf{b}} > 0$ depends on $\mathbf{b}$, α_i and δ only. Set

$$\varepsilon = \min_{\mathbf{b} \in \{-1,1\}^N} \varepsilon_{\mathbf{b}}.$$

Now if $|\alpha_i - \beta_i| < \varepsilon$, then $f_{\mathbf{b}}(\beta_i) > 1$. Hence the roots $\beta_j, |\beta_j| > 1, j = 1 \ldots k$,

of P_1 satisfy the same conditions of Theorem 5.5 as the roots of P with the same numbers $i = i(\mathbf{b})$, $n = n(\mathbf{b})$ and the number $\delta = 0$. □

Proof of Theorem 5.12. There exists a real number $\rho > 1$ such that all the roots $\alpha_j, j = 1 \dots k$, of the polynomial P outside the unit circle are of moduli strictly greater than ρ, i.e., $|\alpha_j| > \rho$. For any $\varepsilon > 0$, choose a sufficiently small positive number $r < \varepsilon$ such that, firstly, all the points of the set $S = \bigcup_{j=1}^{k} \{z : |z - \alpha_j| = r\}$ are of modulus $|z| > \rho$ and, secondly, the polynomial P does not vanish for any $z \in S$. Let $m = \inf_{z \in S} |P(z)|$, $M = \sup_{z \in S} |R(z)|$. Since S is a compact set, by the continuity of P, we have $m > 0$. Hence there exists a positive integer n_0 such that, for every $n > n_0$ and each $z \in S$, the inequality $|z^n P(z)| > \rho^n m > M \geqslant |R(z)|$ is satisfied. By Rouché's theorem, the polynomial $z^n P(z) + R(z)$ has the same number of zeros inside each circle $|z - \alpha_j| = r$ as the polynomial P. Now, choose $\varepsilon = \varepsilon(P, \delta)$ given by Theorem 5.11. For $n > n_0$, the polynomial $z^n P(z) + R(z)$ has k roots, say, $\beta_{n,1}, \dots, \beta_{n,k}$, satisfying $|\beta_{n,j}| > 1$ and $|\alpha_j - \beta_{n,j}| < \varepsilon$. By Theorem 5.11, $z^n P(z) + R(z)$ does not divide any Littlewood polynomial. □

Proof of Theorem 5.6. We first show that there exist infinitely many irreducible Newman polynomials which do not divide any Littlewood polynomial. Take some irreducible non-reciprocal polynomial P from Table 1. It does not divide any Littlewood polynomial. By Lemma 5.7, for any positive integer k, $P(x^k)$ also does not divide any Littlewood polynomial. Theorem 3 in the paper of Filaseta [87] asserts that the Newman polynomial $P(x^k)$ is irreducible, because $P(x)$ is irreducible and non-reciprocal.

In order to prove the stronger version asserting that there are infinitely many such primitive irreducible polynomials, let us consider the polynomial $P(x) = 1 + x^4 + x^6 + x^7 + x^9$ from Table 1. Numerical computations show that it satisfies the conditions of Theorem 5.12 with $\delta = 2$ (see Table 6 below). By Theorem 5.12, there exists an integer n_0 such that for every $n > n_0$ the polynomial $x^n P(x) + 1$ is a Newman polynomial not dividing any Littlewood polynomial.

We shall prove the existence of irreducible polynomials among $x^n P(x) + 1$ using standard techniques from the paper [90]. In particular, the direct consequence of Theorem 2 in [90] is that if $P_n(x) = x^n P(x) + 1$, where $n > 2 \deg P = 18$, is reducible then must have a common non-constant reciprocal factor with the reciprocal polynomial $P_n^*(x)$. If $\alpha \neq 0$ is a root of this common factor, then $\alpha^n P(\alpha) = -1$ and $\alpha^{-n} P(1/\alpha) = -1$. Multiplying the corresponding sides of these equalities, we obtain $P(\alpha)P(1/\alpha) - 1 = 0$. This implies that the minimal polynomial of α over $\mathbb{Q}$ must divide the polynomial $G(x) = x^9 P(x) P(1/x) - x^9$. The polynomial G splits in $\mathbb{Z}[x]$ as $G = G_1^2 G_2 G_3 G_4$, where the irreducible factors

G_1, G_2, G_3 and G_4 are given by

$$G_1(x) = x + 1, \qquad G_2(x) = x^2 + 1, \qquad G_3(x) = x^2 - x + 1,$$

$$G_4(x) = x^{12} - x^{11} + x^{10} - x^9 + 2x^8 - 2x^7 + 3x^6 - 2x^5 + 2x^4 - x^3 + x^2 - x + 1.$$

If an irreducible polynomial divides two polynomials $x^m P(x)+1$ and $x^n P(x)+1$ for some positive integers $n > m$, then it must divide their difference $x^m P(x)(x^{n-m} - 1)$. Hence the above non-cyclotomic factor of degree 12 divides at most one polynomial of the sequence $P_n(x) = x^n P(x) + 1$, $n = 1, 2, \ldots$. The roots of cyclotomic factors are $\alpha = -1, \pm i, e^{\pm \pi i/3}$. It is easy to check that $P_n(-1) = (-1)^n + 1$, $P_n(\pm i) = (\pm i)^n + 1$, $P(e^{\pm \pi i/3}) = e^{\pm \pi n i/3} + 1$. For any positive integer $n = 4k$, we have $P_{4k}(\alpha) \neq 0$, therefore $P_{4k}(x)$ is not divisible by any of the cyclotomic factors of $G(x)$. Hence the subsequence $x^{4k} P(x) + 1$, $k = 1, 2, \ldots$, contains infinitely many irreducible polynomials. □

5.6 Algorithms and implementation

5.6.1 Polynomials not dividing any Littlewood polynomial

Using Theorem 5.5 one can test whether a given polynomial $P(x) \in \mathbb{Z}[x]$ divides some Littlewood polynomial. Of course, there is no guarantee at all that this will give the required result if such an N exists but is very large, or if such an N does not exist. In both cases, it can still happen that P does not divide any Littlewood polynomial.

In practice, we choose a positive integer N (typical values are $N = 50, 70, 100$), a real number $\delta \geqslant 0$ ($\delta = 0, 1, 2$) and check if there exists at least one vector $\mathbf{b} \in \{-1, 1\}^N$ of length N for which the left hand side of the inequality of Theorem 5.5, namely, the quantity $(|\alpha_j| - 1)|\alpha_j^n + b_1 \alpha_j^{n-1} + \cdots + b_n|$ is less than $1 + \delta$ for every n, where $1 \leqslant n \leqslant N$, and each root α_j, $j = 1 \ldots k$, of P of modulus strictly greater than 1. This is accomplished by the recursive search through the binary tree of all vectors $\mathbf{b}$. If the program reports that the reached depth is less than $N + 1$, then such vector $\mathbf{b}$ does not exist. The conditions of Theorem 5.5 are satisfied, so that P does not divide any Littlewood polynomial. A similar method was already used in computer graphics (see [161]).

Algorithm 5.14. *Numerical test checking whether $P(x) \in \mathbb{Z}[x]$ can divide some Littlewood polynomial*

Input:	*an integer $N \geqslant 1$, a real number $\delta \geqslant 0$,*
	the set S of roots α of P lying in $\lvert z\rvert > 1$.
Output:	*integer depth – reached depth. Initially it is* 0.
Other variables:	*vector $\boldsymbol{b}$.*
Method:	*check the conditions of Theorem 5.5.*

Step 0: Set $n = 0$.

Step 1: Given n.

if $n >$ depth then set depth $= n$.

if depth $\leqslant N$ then do

for each $\alpha \in S$ test whether
$(|\alpha| - 1)|\alpha^n + b_1\alpha^{n-1} + \cdots + b_n| < 1 + \delta$.

if all inequalities hold then do
i) set $b_{n+1} = 1$ and invoke Step 1 on $n + 1$.
ii) set $b_{n+1} = -1$ and invoke Step 1 on $n + 1$.
end if.

end if.

We used the following method in order to reduce the amount of numerical calculations. Let $T_n(\alpha) = \alpha^n + b_1\alpha^{n-1} + \cdots + b_n$. The value $T_{n+1}(\alpha)$ can be found by the formula $T_{n+1}(\alpha) = \alpha T_n(\alpha) + b_{n+1}$. Moreover, if $\alpha, \overline{\alpha}$ are two complex conjugate roots of P, then it suffices to evaluate the inequality only at one of them, say, at α with $\Im(\alpha) \geqslant 0$.

The most critical part of this program is the numerical evaluation of the inequality of Theorem 5.5. Any rounding error may cause an incorrect termination of the recursion. To overcome this difficulty we coded the algorithm in C++ using C–XSC library for validated real and complex bounding interval arithmetics (see [104]). The details concerning the initial approximation and the computation of

enclosing rectangles of roots α are given in Section 5.7. We store the enclosures of the numbers $\alpha, T_n(\alpha)$ using C–XSC data types `cinterval` or `l_cinterval`. In order to evaluate the left hand side of the inequality, we compute its interval enclosure (data types `interval` and `l_interval`). Then the lower bound of this enclosure, given by `interval.Inf()`, is compared to the right hand side of our inequality. It must be strictly smaller than $1 + \delta$ for the inequality to hold. In addition, setting the variable δ to be a quite large non-zero number, say, $\delta = 2$ helps to prevent any accidental rounding errors.

5.6.2 Littlewood polynomials divisible by a given polynomial

Given $P(x) \in \mathbb{Z}[x]$, we search for a polynomial $Q(x) \in \mathbb{Z}[x]$ of height $H(Q) \leqslant h$ such that the product PQ is a Littlewood polynomial. Let

$$P(x) = a_0 + a_1 x + \ldots + a_d x^d, \qquad Q(x) = b_0 + b_1 x + \ldots + b_n x^n.$$

Clearly, $P(0)Q(0) = \pm 1$, hence we may assume that $a_0 = b_0 = 1$ (otherwise, replace P, Q by $-P, -Q$, respectively). If such Q exists, then, by Lemma 5.10, it is possible to find Q of degree at most $(2h+1)^d + d - 2$. We used the following approach.

Suppose that the first j coefficients $b_0, b_1, \ldots, b_{j-1}$, where $0 \leqslant j \leqslant n$, of Q are already known. The coefficient l_j of the product $P(x)Q(x) = \sum_{j=0}^{n+d} l_j x^j$ is given by the equality

$$l_j = b_j a_0 + b_{j-1} a_1 + \ldots + b_{\max\{0,j-d\}} a_{\min\{j,d\}}.$$

Since $a_0 = 1$, from this equation we find that

$$b_j = l_j - b_{j-1} a_1 - \ldots - b_{\max\{0,j-d\}} a_{\min\{j,d\}}.$$

The coefficient b_j depends on the previous coefficients $b_i, 0 \leqslant i \leqslant j-1$, and the value of the coefficient $l_j \in \{-1, 1\}$. We must consider only those choices for l_j which give $|b_j| \leqslant h$. Suppose that we have determined the correct value of the coefficient l_j and computed the number b_j. If $j = n$, we are done. If no, we proceed to compute the next coefficient b_{j+1}.

This approach leads to the algorithm which recursively iterates through all candidates for the polynomial Q by trying all possible values $l_j = -1$ and $l_j = 1$

and finding the coefficients b_j. The recursion terminates when the factor Q is found or when two identical blocks of coefficients of length d are detected in the vector of coefficients b_j computed in the current branch of the recursion. Two identical blocks will necessarily occur if $j > (2h+1)^d + d - 2$ (see the proof of Lemma 5.10), hence the depth of recursion is finite. This also prevents the algorithm from searching through non-optimal candidates for the factor Q with repeated blocks of coefficients. Search ends immediately when the polynomial Q is found. Otherwise, it continues until all possible candidates for Q are rejected.

Algorithm 5.15. *Determines whether $P(x) \in \mathbb{Z}[x]$ divides a Littlewood polynomial L with $H(L/P) \leqslant h$.*

Input: *A Newman polynomial $P(x) = a_0 + \ldots + a_d x^d \in \mathbb{Z}[x]$ of degree d.*
A positive integer h.

Output: *A polynomial $Q(x) = b_0 + \ldots + b_n x^n \in \mathbb{Z}[x]$ of height $\leqslant h$ such that PQ is a Littlewood polynomial.*
Prints '$H(Q) > h$' if such Q does not exist.

Level 0: 1. Set $b_0 = 1$, FOUND = false.
2. Iterate to the level 1.
3. If (FOUND is false), then print "$H(Q) > h$".
4. Exit.

Level j: 1. If (FOUND is true) or two identical blocks of length d are detected in the vector of coefficients $(b_0, b_1, \ldots, b_{j-1})$ computed so far, return.

2. Check if $P(x)(b_0 + b_1 x + \ldots + b_{j-1}x^{j-1})$ is already a Littlewood polynomial:

a) if it is, set FOUND = true, print the coefficients of Q and return;

b) if it is not:
i) set $l_j = 1$, compute $b_j = l_j - b_{j-1}a_1 - \ldots - b_{\max\{0,j-d\}}a_{\min\{j,d\}}$;
if $|b_j| \leqslant h$, iterate to the next level $(j+1)$;

ii) set $l_j = -1$, compute $b_j = l_j - b_{j-1}a_1 - \ldots - b_{\max\{0,j-d\}}a_{\min\{j,d\}}$;
if $|b_j| \leqslant h$, iterate to the next level $(j+1)$.

3. Return to the previous level $(j-1)$.

We coded Algorithm 5.15 in C++. The detection of repeating blocks and checking if the product $P(x)(b_0 + b_1x + \ldots + b_{j-1}x^{j-1})$ is already a Littlewood polynomial are very important to the program performance. We shall describe our implementation.

Before the search is started, we create an array of integers `A[ ]` of size $(2h+1)^d$ and initially fill this array with zeros. To every block $B_s = b_s, b_{s+1}, \ldots, b_{s+d-1}$ of length d we assign a non-negative integer

$$c(B_s) = (b_s + h)(2h + 1)^{d-1} + (b_{s+1} + h)(2h + 1)^{d-2} + \ldots + (b_{s+d-1} + h)$$

which is the representation of block B_s in base $2h + 1$. If B_s and B_{s+1} are two adjacent blocks in the vector of coefficients, then the number $c(B_{s+1})$ can be quickly computed from the identity

$$c(B_{s+1}) = c(B_s)(2h + 1) + b_{s+d} + h \pmod{(2h + 1)^d}.$$

For each new block B_{j-d} found at the recursion level $j \geqslant d$, we check whether `A[`$c(B_{j-d})$`]` is equal to 0. If so, we store the value 1 at `A[`$c(B_{j-d})$`]`. If `A[`$c(B_{j-d})$`]` is already equal to 1, the new block is identical to the one of the blocks computed before.

In order to check if the product of the polynomials P and $Q(x) = b_0 + b_1x + \ldots + b_{j-1}x^{j-1}$ is already a Littlewood polynomial, it suffices to check the last d coefficients $l_j, l_{j+1}, \ldots l_{j+d-1}$ of the product PQ. Indeed, all the coefficients $b_0, b_1, \ldots b_{j-1}$ in the course of the recursive search are computed in such a way that the first j coefficients $l_0, l_1, \ldots, l_{j-1}$ are -1 or 1. The values $l_j, l_{j+1}, \ldots l_{j+d-1}$ depend only on the last d values $b_{j-d}, b_{j-d+1}, \ldots, b_{j-1}$. We call block

$$B = b_{j-d}, b_{j-d+1}, \ldots, b_{j-1}$$

the *endblock,* if the last d coefficients of the product

$$P(x)(b_{j-d} + b_{j-d+1}x + \ldots + b_{j-1}x^{d-1})$$

belong to the set $\{-1, 1\}$. After initialization of the array `A[ ]`, before the recursive search is started, we precompute all possible endblocks B and store the values -1 at `A[`$c(B)$`]`. When the block B_s with `A[`$c(B_s)$`]`$= -1$ is found, the polynomial $Q(x)$ is printed and the search algorithm is stopped.

We remark that another algorithm for computing Littlewood polynomials with prescribed factors is given by Mossinghoff in the paper [149]. His approach is quite different from ours.

5.7 Computations

All the computations described bellow were performed on the Linux desktop computer with the Intel Pentium 4 class 2.4 Ghz processor and 1 GB of RAM. We used the GNU C++ compiler v.4.1.2.

5.7.1 Newman polynomials dividing Littlewood polynomials

We ran the implementation of Algorithm 5.15 on the list of all Newman polynomials P of degree $\deg P \leqslant 8$. For each of the 255 polynomials P in the list, the program computed a polynomial $Q(x) \in \mathbb{Z}[x]$, $H(Q) \leqslant 2$, such that the product PQ is a Littlewood polynomial L. The total program running time was less than one second. For most polynomials P, there exists a factor Q of height 1. There are only four exceptional Newman polynomials P of degree eight with the property that $H(L/P) \geqslant 2$ for any Littlewood polynomial L divisible by P. They are given in Table 2. The degree and height of the factor Q are also given here, together with the recursion depth reached until all candidates for Q of height 1 were rejected.

Table 5.2: Four exceptional Newman polynomials of degree 8 with $H(L/P) \geqslant 2$.

	Polynomial $P(x)$	$H(Q)$	$\deg Q$	Recursion depth ($h = 1$)
1.	$1 + x + x^4 + x^6 + x^8$	2	208	135
2.	$1 + x^2 + x^4 + x^7 + x^8$	2	208	78
3.	$1 + x + x^2 + x^5 + x^6 + x^8$	2	47	20
4.	$1 + x^2 + x^3 + x^6 + x^7 + x^8$	2	47	48

Then we ran the program on the list of Newman polynomials of degree 9. As a result, we found that 220 of the 256 polynomials divide some Littlewood polynomial L with the height of the quotient $H(L/P) = 1$. The program running time was also less than one second. For instance, for the Newman quadrinomial $P(x) = 1+x^2+x^5+x^9$, which was used as an example in Section 5.4, the program found a Littlewood polynomial L divisible by P. The coefficients of L are given below. Note that the degree 371 of the polynomial L is exactly as predicted by Lemma 5.9.

Then we used our program once again for $h = 2$. As a result, we found that 18 polynomials of degree nine of the remaining 36 divide Littlewood polynomials L with $H(L/P) = 2$. The computations were completed in 1.6 sec. We launched the

Table 5.3: The signs of the coefficients $l_0, l_1, \ldots, l_{371} \in \{-1, 1\}$ of the Littlewood polynomial $L(x) = \sum_{j=0}^{371} l_j x^j$ divisible by the polynomial $P(x) = 1 + x^2 + x^5 + x^9$.

+ + + + + + + − − + + − − + + − + + + + − − + + + + + + + + +
+ + + + + + − + + − + + + + + + − + + + + + − − + − − + + − −
+ + + + + + + − − + + − − + − − − + + + − + − − + + + − + + −
+ + + − + + − + + − + + − + + + + − − + + − + + − + + − − + +
− + − + + − + + − + + − + + + − + − − + + + + + + − − + + − +
− − + + + − + + + + + + + + + − + + + − + + − − + − − + + − −
+ − + − + + + − + + + + + + − − + − + + + + − − + + + + + + − +
+ − + + + − − + + − + + + + + + + − + + − + + − + − − + + − + +
+ − − + + − − + − + + + − + − − + + + + + + − + + − + + − + −
+ + + + + − + + + − + + − + + − − + + − + + − − + − − + + + −
+ − − + + − + + + − + + + − + + + − − + + − + + + + + + − + +
− − + + − + + + − + + + + + + − + + + − + + − − + − − + + − −

program once more to check if any of the remaining 18 polynomials divide some Littlewood polynomial with $H(L/P) = 3$. The time required for the program to complete the computations increased to 15.3 sec. The program gave a negative answer to all 18 polynomials. For the polynomials $1 + x^2 + x^6 + x^7 + x^9$ and $1 + x + x^2 + x^5 + x^7 + x^8 + x^9$ the algorithm reached the depths of the recursion 4640 and 4648, respectively, before rejecting all possible candidates for Q of height at most 3. For other 16 polynomials the maximal depth of the iterations required was not greater than 373.

Naturally, the polynomials from the last list are very good candidates for Newman polynomials not dividing any Littlewood polynomial. At least this gave us a realistic hope that such polynomials do exist. So we tested them using Algorithm 5.14 (see the next subsection).

In addition, we experimented with some special polynomials of higher degrees. For instance, we computed the following factor Q of height 1 for the Lehmer polynomial ℓ of degree 10 given above:

$$Q(x) = 1 + x^2 - x^3 + x^4 + x^7 - x^8 + x^{10} + x^{12} - x^{13} + x^{16} - x^{17} + x^{18} - x^{20}.$$

The product ℓQ is a Littlewood polynomial of degree 30. See also [149] for other examples.

5.7.2 Irreducible Newman polynomials not dividing any Littlewood polynomial

We used the numerical solver program MPSolve [26] which is based on the GMP library [100] for the computations with extended precision. For every root $\alpha = \Re(\alpha) + i\Im(\alpha)$ with modulus $|\alpha| > 1$ and imaginary part $\Im(\alpha) \geqslant 0$ of a given Newman polynomial P, we calculated the approximations a and b of real and imaginary parts of α to the 100 digits. We then chose a real number $\varepsilon > 0$ which is sufficiently large to compute correct open bounding intervals $R = (a - \varepsilon, a + \varepsilon)$ and $I = (b - \varepsilon, b + \varepsilon)$ for $\Re(\alpha), \Im(\alpha)$, so that $\Re(\alpha) \in R, \Im(\alpha) \in I$. The values of ε are provided bellow. This procedure was applied to every polynomial P tested by Algorithm 5.14.

We launched the initial test to check the conditions of Theorem 5.5 for the values $N = 100$, $\delta = 0$ on 18 polynomials of degree 9. We set the variable ε which controls the accuracy of bounding intervals to a relatively large value, namely, $\varepsilon = 10^{-14}$. This accuracy was consistent with the capacity of the data types `interval` and `cinterval` used by the `C-XSC` library. The recursion depths reached by the program and corresponding times are summarized in Table 4.

Table 5.4: $N = 100, \delta = 0, \varepsilon = 10^{-14}$

$P(x)$	The depth of recursion	Time, sec.
1.	67	8.9
2.	50	2.3
3.	71	0.7
4.	35	0.4
5.	59	1.7
6.	101	0.4
7.	101	0.1
8.	49	1.9

The row number in the first column of the table corresponds to the number of the polynomial in Table 1. If the recursion depth reached is less than $N + 1$, then the corresponding polynomial is confirmed to be the polynomial which does not divide any Littlewood polynomial (see Section 5.6).

The initial test gave no information about the polynomials numbered 6 and 7. In contrast, their reciprocals, numbers 5 and 8, were identified as those which do not divide any Littlewood polynomial. To deal with these two examples (which obviously must be the polynomials which do not divide any Littlewood polynomial too), we rewrote the code of the program replacing data types `interval` and `cinterval` with multiprecision data types `l_interval` and `l_cinterval`. We

then increased the precision of the bounding intervals to $\varepsilon = 10^{-30}$. The numerical test confirmed both polynomials as not dividing any Littlewood polynomial:

Table 5.5: $N = 100, \delta = 0, \varepsilon = 10^{-30}$

$P(x)$	The depth of recursion	Time, sec.
6.	81	24.2
7.	58	7.6

Then we tested whether some polynomials from Table 1 satisfy the conditions of Theorem 5.5 with the strictly positive δ. We used the first version of the code due to a considerable increase in the time required by the program to complete the tests. The results for $\delta = 2$ which terminated up to $N = 100$ are given in Table 6.

Table 5.6: $N = 100, \delta = 2, \varepsilon = 10^{-14}$

$P(x)$	The depth of recursion	Time, sec.
1.	83	1662.0
2.	59	346.7
4.	49	140.3
5.	75	685.6
8.	67	629.5

It is important to note that Table 1 contains only those Newman polynomials of degree 9 for which numerical tests confirmed that both polynomial P and its reciprocal P^* do not divide any Littlewood polynomial. Moreover, in order to be absolutely sure, in each case, using the test based on Algorithm 5.14, we found that at least one of the polynomials P and P^* has the required property with quite large value of δ. See Table 6, where $\delta = 2$.

In each of the remaining 10 cases, the classification problem was not completely solved. All 10 remaining polynomials are listed in the following table.

Naturally, one can hope that further searches performed using Algorithm 14 expecting quotients L/P of height $H(L/P) \geqslant 4$ or additional tests based on Algorithm 13 with increased recursion depths and more accurate bounding intervals for the roots of P will complete the classification of polynomials given in Table 7. However, as we said above, in principle, it is possible that P does not divide any Littlewood polynomial, but this cannot be established by Algorithm 13 applied to P or P^*.

Table 5.7: Newman polynomials which are not confirmed to divide a Littlewood polynomial.

1.	$1 + x^2 + x^3 + x^7 + x^9$
2.	$1 + x^2 + x^6 + x^7 + x^9$
3.	$1 + x + x^3 + x^4 + x^7 + x^9$
4.	$1 + x^2 + x^5 + x^6 + x^8 + x^9$
5.	$1 + x + x^2 + x^3 + x^6 + x^7 + x^9$
6.	$1 + x^2 + x^3 + x^6 + x^7 + x^8 + x^9$
7.	$1 + x + x^2 + x^3 + x^5 + x^8 + x^9$
8.	$1 + x + x^4 + x^6 + x^7 + x^8 + x^9$
9.	$1 + x + x^2 + x^4 + x^7 + x^8 + x^9$
10.	$1 + x + x^2 + x^5 + x^7 + x^8 + x^9$

Chapter 6

Mahler measure of derivative

6.1 Statement of the problem

There exists a classical inequality between the Mahler measure of a polynomial $f \in \mathbb{C}[z]$ of degree d and its derivative

$$M(f') \leqslant dM(f). \tag{6.1}$$

It was proved by Mahler himself [140]. See also Section D in [83] for another proof of (6.1) and two recent papers of Pereira [163] and Pritsker [166], where they showed that (6.1) can be derived from the earlier result of de Bruijn and Springer [49]. The example $f(z) = z^d + 1$ shows that the constant d in (6.1) is best possible.

By the definition of Mahler's measure, we have

$$M(f') \geqslant d|a_d| \tag{6.2}$$

for every polynomial $f(z) = a_d z^d + \cdots + a_0 \in \mathbb{C}[z]$ of degree d. The example $f(z) = a_d z^d + a_0$ with $a_d \neq 0$ and $a_0 \in \mathbb{C}$ shows that we can have $M(f') = d|a_d|$, so inequality (6.2) is sharp. Furthermore, by adding a "large" positive integer n to any polynomial f we see that $\lim_{n\to\infty} M(f+n) = \infty$, whereas the derivative of $f + n$ is equal to f'. So one cannot replace the right-hand side of (6.2), $d|a_d|$, by $\varepsilon dM(f)$ with some 'small' positive ε, unless there is some restriction on the constant term of a polynomial. The most natural restriction is $f(0) = 0$. Then the result of Storozhenko [203] asserts that

$$M(f') \geqslant s(d)M(f)$$

for each $f \in \mathbb{C}[z]$ of degree d satisfying $f(0) = 0$, where

$$s(d) = \frac{d}{\prod_{d/6<k<5d/6}(2\sin(\pi k/d))} = \frac{d}{M((1+z)^d - 1)}.$$

The example $f(z) = (1+z)^d - 1$ shows that the bound of Storozhenko is best possible. However, $s(d)$ is approximately 1.4^{-d} for large d, so the constant $s(d)$ is very small. Hence, it is natural to raise the following question:

Problem 6.1. *For which class of polynomials $f \in \mathbb{C}[z]$ there is a lower bound for $M(f')$ given in terms of $M(f)$ and d such that the dependence on d is similar to the Mahler's result* (6.1)*?*

An important class of such polynomials is given in Theorem 6.2.

Let f^* denote the *reciprocal polynomial* of $f \in \mathbb{C}[z]$, defined by $f^*(z) = z^{\deg f}\overline{f}(1/z)$, where the bar denotes complex conjugation. More precisely, for $f(z) = a_d z^d + \cdots + a_0$, where $a_d \neq 0$, we have $f^*(z) = \overline{a_0}z^d + \cdots + \overline{a_d}$. Generally speaking, a polynomial $f \in \mathbb{C}[z]$ is called *self-inversive* if $f^*(z) = \theta f(z)$ for some $\theta \in \mathbb{C}$. Evidently, $\theta^2 = \overline{a_d a_0}/(a_d a_0)$, so such a number θ must be of modulus 1. In particular, a polynomial $f \in \mathbb{R}[z]$ is called *reciprocal* if its reciprocal polynomial f^* is $\pm f$, i.e., $f^*(z) = \pm f(z)$. In other words, f is reciprocal if it is a real self-inversive polynomial. The set of roots of a self-inversive polynomial is invariant under the map $z \mapsto 1/\overline{z}$, i.e., the multiset $\{\alpha_1, \ldots, \alpha_d\}$ coincides with the multiset $\{1/\overline{\alpha_1}, \ldots, 1/\overline{\alpha_d}\}$.

Reciprocal polynomials play a very important role in the theory of Mahler measure of algebraic numbers. In 1933, Lehmer [132] posed a question whether for every $\varepsilon > 0$ there is a polynomial $f \in \mathbb{Z}[z]$ satisfying $1 < M(f) < 1 + \varepsilon$. In 1951, Breusch [48] proved a result which implies that such a polynomial f, if exists, must be reciprocal. One should also mention the papers of Schinzel [180], Amoroso and Dvornicich [11], Borwein, Dobrowolski and Mossinghoff [33], where Lehmer's problem was resolved for some other classes of polynomials with integer coefficients. The theorem of Smyth [199] remains one of the most important results in this area, because it reduces Lehmer's problem to a 'small' class of reciprocal polynomials. Clearly, the next theorem, which gives a kind of reverse inequality to (6.1), is applicable to reciprocal polynomials in $\mathbb{Z}[z]$ and more generally to reciprocal polynomials in $\mathbb{R}[z]$:

Theorem 6.2. *Suppose that $f \in \mathbb{C}[z]$ is a self-inversive polynomial of degree $d \geqslant 2$. Then*

$$M^2(f') \geqslant \frac{d^2}{4}M^2(f) + M^2\left(zf' - \frac{d}{2}f\right).$$

In particular, one has

$$M(f') \geqslant \frac{d}{2}\left(M(f)^2 + |f(0)|^2\right)^{1/2}.$$

We remark that two trivial cases $d = 0$ and $d = 1$ are excluded from Theorem 6.2. If $d = 1$ then $f(z) = a_1 z + a_0 \in \mathbb{C}[z]$ is self-inversive if and only if $|a_0| = |a_1| \neq 0$. Then $M(f) = |a_1|$ and $M(f') = |a_1|$, giving $M(f') = M(f)$ for each linear self-inversive polynomial f. The proof of Theorem 6.2 comes as a combination of our original Lemma 6.6 and a remark of anonymous Referee, who suggested to use Mahler's inequality to strengthen the result stated in the paper [70]. Mahler's inequality will be stated and proved in Lemma 6.8 in Section 6.2. A weaker version of Lemma 6.6, namely, Corollary 6.7 has already been known to various authors. Together with Corollary 6.10, they have already been used in earlier papers (e. g., [14], [15]) in the context on the L^s norms of polynomials. See also books [171], [34]. It seems that neither our Lemma 6.6 nor theorem 6.2 were known in this (stronger) form and, moreover, were never applied in the context of Mahler measures of polynomials. (One can find many references on Mahler's measure in a survey of Smyth [201].)

Note that if $M(f) < \sqrt{3}|a_d|$ then the trivial inequality (6.2), which holds for every $f \in \mathbb{C}[z]$ of degree d, is stronger than the inequality of Theorem 6.2. In particular, Theorem 6.2 is of no use if $f \in \mathbb{Z}[z]$ is a monic reciprocal polynomial whose Mahler measure $M(f)$ satisfies $1 \leqslant M(f) \leqslant \sqrt{3}$. In this case, the trivial inequality $M(f') \geqslant d$ is stronger. Moreover, the inequality $M(f') \geqslant d$ is strict in case a self-inversive monic polynomial f of degree $d \geqslant 2$ has at least one root of modulus $\neq 1$. Indeed, by the result of Marden (see Theorem (45,2) in [143]), a self-inversive polynomial f and its derivative f' have the same number of zeros in $|z| > 1$, so f' has at least one root in $|z| > 1$ provided that f has a root of modulus $\neq 1$. Combining Marden's inequality with Theorem 6.2 we thus obtain

$$M(f') \geqslant d \max\left\{1, \quad \frac{1}{2}\sqrt{M^2(f) + 1}\right\}$$

for each monic self-inversive polynomial f. Self-inversive polynomials with all zeros on the unit circle have been studied by Bonsall and Marden [28], Lakatos and Losonczy [127], Schinzel [186], Sinclair and Vaaler [196]. For those polynomials of degree d we obviously have $M(f') = dM(f)$.

Combining Theorem 6.2 with Mahler's inequality (6.1), we have

$$\frac{d}{2} < \frac{M(f')}{M(f)} \leqslant d$$

for each self-inversive polynomial $f \in \mathbb{C}[z]$ of degree d. How close is the number $\inf \frac{M(f')}{M(f)}$, where the infimum is taken over every self-inversive polynomial $f \in \mathbb{C}[z]$ of degree d, to the value $d/2$? Our next theorem shows that for d even the lower bound $M(f')/M(f) > d/2$ is best possible even we are restricted to the class of monic integer reciprocal polynomials.

Theorem 6.3. *Let $d \geqslant 2$ be an even integer. Then there is a sequence of monic integer reciprocal polynomials f_n, $n = 1, 2, 3, \ldots$, of degree d such that $M(f_n')/M(f_n) \to d/2$ as $n \to \infty$. Furthermore, for every $w \in (d/2, d]$ there is a monic reciprocal polynomial $f \in \mathbb{R}[z]$ of degree d such that $M(f')/M(f) = w$.*

For $d \geqslant 3$ odd we prove the following:

Theorem 6.4. *Let $d \geqslant 3$ be an odd integer. Then there is a monic integer reciprocal polynomial f of degree d such that $M(f') < \frac{d+1}{2} M(f)$.*

For $d = 3$ we shall find the exact value of $\inf \frac{M(f')}{M(f)}$, where the infimum is taken over every cubic self-inversive polynomial $f \in \mathbb{C}[z]$:

Theorem 6.5. *Let $f \in \mathbb{C}[z]$ be a cubic self-inversive polynomial. Then*

$$M(f')/M(f) \geqslant \frac{3(827 + 384\sqrt{2})^{1/3}}{32} + \frac{219}{32(827 + 384\sqrt{2})^{1/3}} + \frac{9}{32} = 1.93867997\ldots. \tag{6.3}$$

The equality in (6.3) *is attained for the polynomial $f(z) = z^3 - B_0 z^2 - B_0 z + 1 \in \mathbb{R}[z]$, where*

$$B_0 = \frac{(827 + 384\sqrt{2})^{1/3}}{8} + \frac{73}{8(827 + 384\sqrt{2})^{1/3}} + \frac{11}{8} = 3.58490663\ldots. \tag{6.4}$$

One can also express B_0 as $B_0 = t + 1/t - 1$ with

$$t = (3 + 2\sqrt{2})^{1/3} + (3 - 2\sqrt{2})^{1/3} + 2 = 4.355301398\ldots.$$

Then

$$f(z) = z^3 - B_0 z^2 - B_0 z + 1 = (z + 1)(z - t)(z - 1/t),$$

where t satisfies

$$t^3 - 6t^2 + 9t - 8 = 0.$$

The minimal polynomial of the number B_0 over $\mathbb{Q}$ is

$$8z^3 - 33z^2 + 18z - 9. \tag{6.5}$$

By (6.4), the right hand side of (6.3) is equal to $3(B_0 - 1)/4$. The minimal polynomial of the cubic number $3(B_0 - 1)/4$ is

$$32z^3 - 27z^2 - 54z - 27. \tag{6.6}$$

Let $\mathcal{S}_d$ be the set of complex self-inversive polynomials of degree d, and let $\mathcal{R}_d \subset \mathcal{S}_d$ be the subset of (real) reciprocal polynomials. Theorems 6.2, 6.3 and 6.4 imply that $\inf_{f\in\mathcal{S}_d} \frac{M(f')}{M(f)} = (d + \tau_d)/2$, where $\tau_d = 0$ for d even and $0 \leqslant \tau_d < 1$ for $d \geqslant 3$ odd. Theorem 6.5 asserts that $\min_{f\in\mathcal{S}_3} \frac{M(f')}{M(f)} = 1.93867997\ldots$. Their proofs will be given in Sections 3 and 4. In Section 5 we shall study the quotients of L^s norms $||f'||_s/||f||_s$ for $f \in \mathcal{S}_d$ and $f \in \mathcal{R}_d$. We also give some numerical examples of polynomials of odd degree and low value of the quotient $M(f')/M(f)$ in Section 6.

6.2 Some lemmas

Lemma 6.6. *Let $f \in \mathbb{C}[z]$ be a self-inversive polynomial of degree $d \geqslant 1$, and let z_0 be a complex number of modulus 1. Then*

$$|f'(z_0)|^2 = \left(\frac{d}{2}\right)^2 |f(z_0)|^2 + |z_0 f'(z_0) - \frac{d}{2} f(z_0)|^2. \tag{6.7}$$

If, in addition, z_0 is not a root of f then $\Re(z_0 f'(z_0)/f(z_0)) = d/2$.

Proof: Since $f \in \mathbb{C}[z]$ is self-inversive, $f^*(z) = z^d\overline{f}(1/z) = \theta f(z)$ for some $\theta \in \mathbb{C}$ of modulus 1. Calculating the derivative of θf, we obtain

$$\theta f'(z) = dz^{d-1}\overline{f}(1/z) - z^{d-2}\overline{f}'(1/z) = \theta d z^{-1} f(z) - z^{d-2}\overline{f}'(1/z)$$

for every non-zero complex number z. Dividing both sides of the above equality by $\theta z^{-1} f(z)$, we deduce that

$$d = \frac{z f'(z)}{f(z)} + \frac{z^{d-1}\overline{f}'(1/z)}{\theta f(z)} = \frac{z f'(z)}{f(z)} + \frac{\overline{f}'(1/z)}{z\overline{f}(1/z)} \tag{6.8}$$

if z is non-zero and not a root of f.

Fix any z_0 of modulus 1 which is not a root of f. Then $1/z_0 = \overline{z_0}$, because $|z_0| = 1$. Setting $z = z_0$ into the right hand side of (6.8), we see that the second term is equal to the complex conjugate $\overline{z_0 f'(z_0)/f(z_0)}$ of the first term $z_0 f'(z_0)/f(z_0)$. It follows that $d = 2\Re(z_0 f'(z_0)/f(z_0))$, giving $\Re(z_0 f'(z_0)/f(z_0)) = d/2$. This proves the second part of the lemma.

It is clear that equality (6.7) holds for each z_0 of modulus 1 which is a root of f.

Suppose that $|z_0| = 1$, where z_0 is not a root of f. Then $\Re(z_0 f'(z_0)/f(z_0)) = d/2$ implies that

$$z_0 f'(z_0)/f(z_0) = d/2 + i\Im(z_0 f'(z_0)/f(z_0)).$$

Hence $|z_0 f'(z_0)/f(z_0) - d/2| = |\Im(z_0 f'(z_0)/f(z_0))|$. Since $|z_0| = 1$, we have

$$|f'(z_0)/f(z_0)|^2 = |z_0 f'(z_0)/f(z_0)|^2 = \Re(z_0 f'(z_0)/f(z_0))^2 + \Im(z_0 f'(z_0)/f(z_0))^2$$

$$= (d/2)^2 + |z_0 f'(z_0)/f(z_0) - d/2|^2.$$

Multiplying by $|f(z_0)|^2 \neq 0$ yields (6.7).

We remark that identity (6.7) of Lemma 6.6 implies that if $f \in \mathbb{C}[z]$ is a self-inversive polynomial, then its derivative f' has no zeros on the unit circle $|z| = 1$, except for the multiple zeros of f (if any). This is a result of Bonsall and Marden [28] (see Lemma (45,2) in [143]) which was proved by an entirely different argument. Note that Lemma 6.6 yields the following corollary:

Corollary 6.7. *Let $f \in \mathbb{C}[z]$ be a self-inversive polynomial of degree $d \geqslant 1$, and let z_0 be a complex number of modulus 1. Then $|f'(z_0)| \geqslant \frac{d}{2}|f(z_0)|$, where equality holds if and only if z_0 is one of at most d points of the circle $|z| = 1$ satisfying $z_0 f'(z_0) = \frac{d}{2} f(z_0)$.*

Proof: From (6.7) we see that $|f'(z_0)| \geqslant \frac{d}{2}|f(z_0)|$, where equality holds if and only if $z_0 f'(z_0) = \frac{d}{2} f(z_0)$. Note that $zf'(z) - \frac{d}{2} f(z)$ is a polynomial of degree d, so it has at most d complex roots. In particular, the equality $z_0 f'(z_0) = \frac{d}{2} f(z_0)$ holds for at most d complex numbers z_0 lying on the unit circle $|z| = 1$. □

The proof of our main theorem relies on the following inequality.

Lemma 6.8. *Let us suppose that $g(t)$ and $h(t)$ are continuous real valued functions, defined in the interval $t \in [0, 2\pi]$. If both functions $g(t)$ and $h(t)$ are strictly positive, then the inequality*

$$||g + h||_0 \geqslant ||g||_0 + ||h||_0$$

holds for the L^0 norms of the functions g, h and $g + h$.

Recall that L^0 norm in the interval $[0, 2\pi]$ is defined as a geometric mean

$$||g||_0 = \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \log|g(t)|dt\right).$$

Since g and h are positive and continuous, the logaritms of g, h and $g + h$ are integrable in the interval $[0, 2\pi)$. Hence, the geometric means $||g||_0$, $||h||_0$ and

$||g+h||_0$ are defined properly. Lemma 6.8 is known as Mahler's inequality [102] (or reversed triangle inequality). Usually, it is stated for the sequences of positive real numbers. Here we give a short proof for positive continuous functions.
Proof: By the monotonicity property of L^s norms, one has

$$\left|\left|\frac{g}{g+h}\right|\right|_0 \leqslant \left|\left|\frac{h}{g+h}\right|\right|_1 \quad \text{and} \quad \left|\left|\frac{g}{g+h}\right|\right|_0 \leqslant \left|\left|\frac{h}{g+h}\right|\right|_1 .$$

Since g and h are positive, L^1 norms are equal to:

$$\left|\left|\frac{g}{g+h}\right|\right|_1 = \frac{1}{2\pi}\int_0^{2\pi} \frac{g(t)}{g(t)+h(t)}\, dt,$$

and

$$\left|\left|\frac{h}{g+h}\right|\right|_1 = \frac{1}{2\pi}\int_0^{2\pi} \frac{h(t)}{g(t)+h(t)}\, dt.$$

Hence

$$\left|\left|\frac{g}{g+h}\right|\right|_0 + \left|\left|\frac{h}{g+h}\right|\right|_0 \leqslant \left|\left|\frac{g}{g+h}\right|\right|_1 + \left|\left|\frac{h}{g+h}\right|\right|_1 = 1.$$

Since L^0 norm is multiplicative,

$$\left|\left|\frac{g}{g+h}\right|\right|_0 = \frac{||g||_0}{||g+h||_0} \quad \text{and} \quad \left|\left|\frac{h}{g+h}\right|\right|_0 = \frac{||h||_0}{||g+h||_0}.$$

The inequality

$$\left|\left|\frac{g}{g+h}\right|\right|_0 + \left|\left|\frac{h}{g+h}\right|\right|_0 \leqslant 1.$$

leads us to the desired result

$$||g||_0 + ||h||_0 \leqslant ||g+h||_0 .$$

□

Here is another auxiliary result we will use.

Lemma 6.9. *Let D be a fixed positive integer, and let $g_1, g_2, g_3, \dots \in \mathbb{C}[z]$ be a sequence of polynomials satisfying $\deg g_n \leqslant D$ and $\lim_{n\to\infty} L(g_n) = 0$. Then $\lim_{n\to\infty} M(g+g_n) = M(g)$ for each $g \in \mathbb{C}[z]$.*

Here and below, for any polynomial $f(z) = \sum_{j=0}^{d} a_j z^j \in \mathbb{C}[z]$, its *length* $L(f)$ is defined by the formula $L(f) = \sum_{j=0}^{d} |a_j|$. The inequality for the difference of two Mahler measures

$$|M(f)^{1/d} - M(g)^{1/d}| \leqslant 2L(f-g)^{1/d}, \tag{6.9}$$

where $f, g \in \mathbb{C}[z]$ are any polynomials of degree $\leqslant d$, was obtained by Chern and Vaaler [55]. Setting $d = \max(D, \deg g)$ and $f = g + g_n$ into (6.9) yields Lemma 6.9. The fact that $\lim_{n\to\infty} M(f_n) = M(f)$ if $\lim_{n\to\infty} L(f - f_n) = 0$ and the polynomials $f, f_1, f_2, \ldots$ have the same degree was earlier proved by Boyd [47]. See also Corollary 14 on p. 251 in [185].

6.3 Proofs of Theorems 1, 2 and 3

Proof of Theorem 6.2: Suppose that $f \in \mathbb{C}[z]$ is a self inversive polynomial. Pick an arbitrary number $\varepsilon > 0$. For $t \in [0, 2\pi]$, define the functions

$$g(t) := \left|\frac{d}{2} f(e^{it})\right|^2 + \varepsilon, \qquad h(t) := \left|e^{it} f'(e^{it}) - \frac{d}{2} f(e^{it})\right|^2 + \varepsilon$$

By Lemma 6.6, one has

$$g(t) + h(t) = \left|f'(e^{it})\right|^2 + 2\varepsilon.$$

By Lemma 6.8, one has

$$\exp\left(\frac{1}{2\pi}\int_0^{2\pi} \log\left(\left|f'(e^{it})\right|^2 + 2\varepsilon\right) dt\right) = ||g + h||_0 \geqslant ||g||_0 + ||h||_0 =$$

$$\begin{aligned} &= \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \log\left(\left|\frac{d}{2} f(e^{it})\right|^2 + \varepsilon\right) dt\right) + \\ &+ \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \log\left(\left|e^{it} f'(e^{it}) - \frac{d}{2} f(e^{it})\right|^2 + \varepsilon\right) dt\right). \end{aligned}$$

For a complex number $z = e^{it}$, the function $g(t) + h(t)$ under the logarithm of the first integrand

$$\begin{aligned} \left|f'(e^{it})\right|^2 + 2\varepsilon = f'(e^{it})\bar{f}'(e^{-it}) + 2\varepsilon = \\ = \left|\frac{z^{d-1} f'(z)\bar{f}'(1/z) + 2\varepsilon z^{d-1}}{z^{d-1}}\right| = \\ = \left|z^{d-1} f'(z)\bar{f}'(1/z) + 2\varepsilon z^{d-1}\right|, \end{aligned}$$

is equal to the absolute value of the polynomial $f' f'^* + 2\varepsilon z^{d-1}$ on the unit circle. Hence, by Jensen's formula, the first term is the Mahler measure $M(f' f'^* + 2\varepsilon z^{d-1})$.

In the same way, last two terms are equal to the Mahler measures of polynomials

$$(d/2)^2 ff^* + \varepsilon z^d \qquad \text{and} \qquad (zf' - (d/2)f)(zf' - (d/2)f)^* + \varepsilon z^d,$$

respectively. Hence,

$$M(f'f'^* + 2\varepsilon z^{d-1}) \geqslant M((d/2)^2 ff^* + \varepsilon z^d) + M((zf' - (d/2)f)(zf' - (d/2)f)^* + \varepsilon z^d).$$

Now, suppose that $\varepsilon \to 0$. By the inequality of Chern and Vaaler (6.9), Mahler measures of the polynomials are continuous with respect to ε. Hence

$$\lim_{\varepsilon \to 0} M(f'f'^* + 2\varepsilon z^{d-1}) = M(f'f'^*) = M^2(f'),$$

$$\lim_{\varepsilon \to 0} M((d/2)^2 ff^* + \varepsilon z^d) = \lim_{\varepsilon \to 0} M((d/2)ff^*) = \frac{d^2}{4} M^2(f),$$

and

$$\begin{aligned} \lim_{\varepsilon \to 0} M((zf' - (d/2)f)(zf' - (d/2)f)^* + \varepsilon z^d) = \\ = M((zf' - (d/2)f)(zf' - (d/2)f)^*) = \\ = M^2\left(zf' - \frac{d}{2}f\right). \end{aligned}$$

Thus, by taking the limit as ε tends to zero we get

$$M^2(f') \geqslant \frac{d^2}{4} M^2(f) + M^2\left(zf' - \frac{d}{2}f\right),$$

as claimed. It remains to note that the Mahler measure $M(zf' - (d/2)f)$ is greater than or equal to the modulus of the constant coefficient of the polynomial $zf'(z) - (d/2)f(z)$, which is equal to $-(d/2)f(0)$. This yields

$$M^2(zf' - (d/2)f) \geqslant (d/2)^2 |f(0)|^2.$$

Consequently,

$$M(f') \geqslant \frac{d}{2}\left(M(f)^2 + |f(0)|^2\right)^{1/2}.$$

Proof of Theorem 6.3: Put

$$f_n(z) = z^d - nz^{d/2} + 1.$$

The Mahler measure of f_n is equal to the Mahler measure of the polynomial $z^2 - nz + 1$, so $M(f_n) = (n + \sqrt{n^2 - 4})/2$ if $n \geqslant 2$. From $f_n'(z) = dz^{d/2-1}(z^{d/2} - n/2)$

it follows that $M(f_n') = dn/2$ for $n \geqslant 2$. Thus

$$\frac{M(f_n')}{M(f_n)} = \frac{d}{1+\sqrt{1-4/n^2}} \to \frac{d}{2} \quad \text{as} \quad n \to \infty,$$

which proves the first part of the theorem.

For the second part, fix $w \in (d/2, d]$ and select a real number $u \geqslant 2$ such that $d/w - 1 = \sqrt{1-4/u^2}$. By the above, for the polynomial $f(z) = z^d - uz^{d/2} + 1 \in \mathbb{R}[z]$, we have $M(f')/M(f) = du/(u+\sqrt{u^2-4}) = w$. □

Proof of Theorem 6.4: For $d = 3$ we can take $f(z) = z^3 - 4(z^2+z) + 1$ and find that $M(f')/M(f) = 1.939249\ldots$ is smaller than 2.

Suppose that $d \geqslant 5$. Set

$$f_n(z) = z^d - ng(z) + 1 \quad \text{with} \quad g(z) = (z+1)(z^2-4z+1)z^{(d-3)/2}.$$

Using $L(f_n/n + g) = 2/n \to 0$ as $n \to \infty$, by Lemma 6.9, we obtain

$$\lim_{n\to\infty} \frac{M(f_n)}{n} = M(g) = M(z^2-4z+1) = 2+\sqrt{3}.$$

Note that

$$g'(z) = ((d+3)z^3 - 3(d+1)z^2 - 3(d-1)z + d - 3)z^{(d-5)/2}/2.$$

Hence, using $L(f_n'/n + g') = d/n \to \infty$ as $n \to \infty$, we find that

$$\lim_{n\to\infty} \frac{M(f_n')}{n} = M(g') = \frac{1}{2}M((d+3)z^3 - 3(d+1)z^2 - 3(d-1)z + d - 3) = \frac{1}{2}M(h),$$

where $h(z) = (d+3)z^3 - 3(d+1)z^2 - 3(d-1)z + d - 3$.

In case the limit $\lim_{n\to\infty} M(f_n')/M(f_n) = (2-\sqrt{3})M(h)/2$ is smaller than $(d+1)/2$, we can take $f = f_n$ with sufficiently large n. Indeed, if $(2-\sqrt{3})M(h) = d+1-\kappa(d)$ with some positive number $\kappa(d)$ then $M(f_n')/M(f_n) < (d+1-\kappa(d)/2)/2$ for each sufficiently large n, say, $n \geqslant n_0$. Taking $f = f_n$, where $n \geqslant n_0$, we would obtain

$$M(f')/M(f) = M(f_n')/M(f_n) < (d+1-\kappa(d)/2)/2 < (d+1)/2,$$

as claimed.

It remains to prove the inequality $M(h) < (2+\sqrt{3})(d+1)$. Note that $h(-1) = -12$, $h(0) = d-3 > 0$ and $h(1) = -4d < 0$. So h has a root in $(-1,0)$, a root in $(0,1)$, and a root $\beta > 1$. Hence $M(h) = (d+3)\beta$. The required inequality is

equivalent to the inequality $\beta < (2+\sqrt{3})(d+1)/(d+3)$. For this, it suffices to show that $h((2+\sqrt{3})(d+1)/(d+3)) > 0$. Indeed, by a simple computation, we have

$$h((2+\sqrt{3})(d+1)/(d+3)) = 4(3(\sqrt{3}+1)d + 3\sqrt{3} - 1)/(d+3)^2 > 0,$$

which completes the proof. □

6.4 The cubic case: Proof of Theorem 4

If all three roots of f lie on the unit circle, then, by the Gauss-Lucas theorem, all roots of f' lie in the unit circle $|z| \leqslant 1$. In such case, $M(f')/M(f) = 3$. This is more than we claim in (6.3).

Next, since $M(uf) = |u|M(f)$ and $M(uf') = |u|M(f')$, $u \in \mathbb{C}$, we can assume that f is monic. The Mahler measure of $f(z)$ and that of $f(ze^{i\phi})$, $\phi \in \mathbb{R}$, are equal. The same holds for their derivatives. Therefore, we may assume that f has a real root t greater than 1. Since the polynomial f is self-inversive, its two other roots are $1/t$ and $e^{i\phi}$, where $\phi \in [0, 2\pi)$. However, as $M(f) = M(\overline{f})$ and $M(f') = M(\overline{f}')$, without loss of generality, we can assume that $\phi \in [0, \pi]$.

It remains to minimize the quotient $M(f')/M(f)$, where f runs through the polynomials $f(z) = (z - e^{i\phi})(z - t)(z - 1/t)$ with $t > 1$ and $\phi \in [0, \pi]$. Evidently, $M(f) = t$. Set

$$w = e^{i\phi} \quad \text{and} \quad \lambda = t + 1/t.$$

Since

$$f(z) = (z-w)(z^2 - \lambda z + 1) = z^3 - (w+\lambda)z^2 + (w\lambda + 1)z - w,$$

we find that

$$f'(z) = 3z^2 - 2(w+\lambda)z + w\lambda + 1.$$

By the above mentioned theorem of Marden, f' has a unique root outside the unit circle, say, $\beta = \beta(\phi, t)$. Hence

$$M(f')/M(f) = 3|\beta|/t. \tag{6.10}$$

Clearly,

$$\beta = \beta(\phi, t) = \frac{\lambda + e^{i\phi} + \sqrt{\lambda^2 - \lambda e^{i\phi} + e^{2i\phi} - 3}}{3}, \tag{6.11}$$

where

$$3\beta(\phi,t)^2 - 2(t + 1/t + e^{i\phi})\beta(\phi,t) + (t + 1/t)e^{i\phi} + 1 = 0. \tag{6.12}$$

The sign of the square root in (6.11) is chosen so that its real part is positive. We claim that then $|\beta(\phi,t)|>1$. Indeed, this is the case if $\lambda=t+1/t$ is large enough. Suppose for some $t>1$ and $\phi\in[0,\pi]$ we have $|\beta(\phi,t)|<1$. Then, by continuity, there exist $t_0>1$ and $\phi_0\in[0,\pi]$ such that $|\beta(\phi_0,t_0)|=1$. However, this is impossible, because f has no multiple roots on $|z|=1$ (see the remark after Lemma 6.6).

Set

$$\gamma=\gamma(\phi,t)=\overline{\beta(\phi,t)}=\frac{\lambda+e^{-i\phi}+\sqrt{\lambda^2-\lambda e^{-i\phi}+e^{-2i\phi}-3}}{3}.$$

Then

$$3\gamma(\phi,t)^2-2(t+1/t+e^{-i\phi})\gamma(\phi,t)+(t+1/t)e^{-i\phi}+1=0. \tag{6.13}$$

Note that if $t\leqslant 3/2$ then $M(f')/M(f)\geqslant 3/t\geqslant 2$. Also, $|\beta(\phi,t)|/t\to 2/3$ as $t\to\infty$, hence $M(f')/M(f)$ tends to 2 as $t\to\infty$. The extremal polynomial $f(z)=z^3-B_0z^2-B_0z+1$ given in the statement of the theorem shows that $M(f')/M(f)$ attains a smaller value. We can thus restrict t to the interval $3/2\leqslant t\leqslant t_0$ with some absolute constant t_0. By (6.10), we have $M(f')^2/9M(f)^2=|\beta|^2/t^2=\beta\gamma/t^2$, so we need to find the minimum of the function $h(\phi,t)=\beta(\phi,t)\gamma(\phi,t)/t^2$ in the rectangle $\phi\in[0,\pi]$, $t\in[3/2,t_0]$.

Suppose that the minimum is attained at the point (ϕ,t), where $0<\phi<\pi$ and $3/2\leqslant t\leqslant t_0$. (It will be clear from the context when the same letters ϕ,t are used to denote the variables of the functions $\beta(\phi,t)$, $\gamma(\phi,t)$ and $h(\phi,t)$ and when the pair (ϕ,t) is used to denote the point, where the minimum of h is attained.) Then this point is a critical point of the function $h(\phi,t)$, so

$$\frac{\partial h}{\partial\phi}=\frac{\partial\beta}{\partial\phi}\gamma+\frac{\partial\gamma}{\partial\phi}\beta=0\quad\text{and}\quad\frac{\partial h}{\partial t}=\frac{1}{t^2}\Big(\frac{\partial\beta}{\partial t}\gamma+\frac{\partial\gamma}{\partial t}\beta\Big)-\frac{2\beta\gamma}{t^3}=0. \tag{6.14}$$

Our goal is to prove that this is not the case. Suppose for a moment that this is already established. Then, as we know, the minimum is not attained at $t=3/2$ and at $t=t_0$, so it must be attained at some point (ϕ,t), where $\phi\in\{0,\pi\}$ and $3/2<t<t_0$. Moreover, it is easy to check that $h(0,t)>h(\pi,t)$. Indeed,

$$t\sqrt{h(0,t)}=\beta(0,t)=\frac{\lambda+1+\sqrt{\lambda^2-\lambda-2}}{3}>$$

$$>\frac{\lambda-1+\sqrt{\lambda^2+\lambda-2}}{3}=\beta(\pi,t)=t\sqrt{h(\pi,t)},$$

because by squaring the inequality $2+\sqrt{\lambda^2-\lambda-2}>\sqrt{\lambda^2+\lambda-2}$ we have

$2\sqrt{\lambda^2-\lambda-2} > \lambda-2$. By squaring again, we see that this inequality holds, because $\lambda > 2$. Thus the minimum of h must be attained at some point (π, t) with $3/2 < t < t_0$. Then

$$\frac{M(f')}{M(f)} = 3\sqrt{h(\pi,t)} = \frac{3\beta(\pi,t)}{t} = \frac{2(\lambda-1+\sqrt{\lambda^2+\lambda-2})}{\lambda+\sqrt{\lambda^2-4}}.$$

Setting $x = \lambda - 1$, we thus need to find the minimum of the function

$$\Psi(x) = \frac{2(x+\sqrt{x^2+3x})}{x+1+\sqrt{(x+1)^2-4}}$$

for $x > \lambda - 1 = t + 1/t - 1 > 7/6$. A simple computation with Maple shows that the function $\Psi(x)$ has only one critical point in $[1,\infty)$. From

$$\Psi'(x) = \frac{2+(2x+3)/\sqrt{x^2+3x}}{x+1+\sqrt{x^2+2x-3}} - \frac{2(x+\sqrt{x^2+3x})(1+(x+1)/\sqrt{x^2+2x-3})}{(x+1+\sqrt{x^2+2x-3})^2}$$

we find that the minimum of $\Psi(x)$ in $[1,\infty)$ is attained at the point

$$x = B_0 = \frac{(827+384\sqrt{2})^{1/3}}{8} + \frac{73}{8(827+384\sqrt{2})^{1/3}} + \frac{11}{8} = 3.58490663\ldots,$$

where $\Psi'(B_0) = 0$. It is equal to $\Psi(B_0) = 1.93867997\ldots$. One of the extremal polynomials at which the equality $M(f')/M(f) = \Psi(B_0)$ is attained is the polynomial $f(z) = z^3 - B_0 z^2 - B_0 z + 1$. The minimal polynomials for B_0 and $\Psi(B_0)$, namely, (6.5) and (6.6) have been found with Maple. This proves the theorem.

In the remainder of this section we will show that for no point (ϕ, t), where $0 < \phi < \pi$ and $3/2 < t < t_0$ all four equalities given in (6.12), (6.13), (6.14) can hold.

Differentiating (6.12) with respect to the variable ϕ, we obtain

$$6\beta\frac{\partial\beta}{\partial\phi} - 2(\lambda+e^{i\phi})\frac{\partial\beta}{\partial\phi} - 2i\beta e^{i\phi} + ie^{i\phi}\lambda = 0.$$

Hence

$$(6\beta - 2\lambda - 2e^{i\phi})\frac{\partial\beta}{\partial\phi} = ie^{i\phi}(2\beta-\lambda). \tag{6.15}$$

Analogously, (6.13) yields

$$6\gamma\frac{\partial\gamma}{\partial\phi} - 2(\lambda+e^{-i\phi})\frac{\partial\gamma}{\partial\phi} + 2i\gamma e^{-i\phi} - ie^{-i\phi}\lambda = 0,$$

thus

$$(6\gamma - 2\lambda - 2e^{-i\phi})\frac{\partial \gamma}{\partial \phi} = ie^{-i\phi}(-2\gamma + \lambda). \tag{6.16}$$

Since, by (6.14), $\frac{\partial \beta}{\partial \phi}\gamma + \frac{\partial \gamma}{\partial \phi}\beta = 0$, multiplying (6.15) by $(3\gamma - \lambda - e^{-i\phi})\gamma$ and (6.16) by $(3\beta - \lambda - e^{i\phi})\beta$ and then adding, we obtain

$$0 = (3\gamma - \lambda - e^{-i\phi})\gamma ie^{i\phi}(2\beta - \lambda) + (3\beta - \lambda - e^{i\phi})\beta ie^{-i\phi}(-2\gamma + \lambda).$$

Hence

$$(2\beta - \lambda)(3\gamma^2 - (\lambda + e^{-i\phi})\gamma)e^{i\phi} = (2\gamma - \lambda)(3\beta^2 - (\lambda + e^{i\phi})\beta)e^{-i\phi}.$$

By (6.12) and (6.13), we know that $3\beta^2 - (\lambda + e^{i\phi})\beta = (\lambda + e^{i\phi})\beta - \lambda e^{i\phi} - 1$ and $3\gamma^2 - (\lambda + e^{-i\phi})\gamma = (\lambda + e^{-i\phi})\gamma - \lambda e^{-i\phi} - 1$. So

$$(2\beta - \lambda)((\lambda + e^{-i\phi})\gamma - \lambda e^{-i\phi} - 1)e^{i\phi} = (2\gamma - \lambda)((\lambda + e^{i\phi})\beta - \lambda e^{i\phi} - 1)e^{-i\phi}.$$

Here, the left hand side is equal to

$$2\beta\gamma(1 + \lambda e^{i\phi}) - \lambda(1 + \lambda e^{i\phi})\gamma - 2\beta(\lambda + e^{i\phi}) + \lambda(\lambda + e^{i\phi}),$$

and the right hand side is equal to

$$2\beta\gamma(1 + \lambda e^{-i\phi}) - \lambda(1 + \lambda e^{-i\phi})\beta - 2\gamma(\lambda + e^{-i\phi}) + \lambda(\lambda + e^{-i\phi}).$$

Subtracting one side from another and multiplying by $w = e^{i\phi}$ yields

$$2\lambda\beta\gamma(w^2 - 1) + \beta(-\lambda w - 2w^2 + \lambda^2) + \gamma(\lambda w + 2 - \lambda^2 w^2) + \lambda(w^2 - 1) = 0. \tag{6.17}$$

Differentiating (6.12) with respect to the variable t and using $\frac{\partial \lambda}{\partial t} = 1 - 1/t^2$, we obtain

$$6\beta\frac{\partial \beta}{\partial t} - 2(\lambda + w)\frac{\partial \beta}{\partial t} - 2(1 - 1/t^2)\beta + w(1 - 1/t^2) = 0.$$

Since, by (6.12), $\beta(3\beta - \lambda - w) = (\lambda + w)\beta - (\lambda w + 1)$, this yields

$$\frac{1}{\beta}\frac{\partial \beta}{\partial t} = \frac{(2\beta - w)(1 - 1/t^2)}{2((\lambda + w)\beta - \lambda w - 1)}.$$

Taking the complex conjugates of both sides, we obtain

$$\frac{1}{\gamma}\frac{\partial \gamma}{\partial t} = \frac{(2\gamma - \overline{w})(1 - 1/t^2)}{2((\lambda + \overline{w})\gamma - \lambda\overline{w} - 1)}.$$

By (6.14), their sum $\frac{1}{\beta}\frac{\partial\beta}{\partial t} + \frac{1}{\gamma}\frac{\partial\gamma}{\partial t}$ must be equal to $2/t$, hence

$$\frac{(2\beta - w)(1 - 1/t^2)}{2((\lambda + w)\beta - \lambda w - 1)} + \frac{(2\gamma - \overline{w})(1 - 1/t^2)}{2((\lambda + \overline{w})\gamma - \lambda\overline{w} - 1)} = \frac{2}{t} \tag{6.18}$$

Expressing $\overline{w} = 1/w$, $\lambda = t + 1/t$ and multiplying by common denominators, we can rewrite (6.18) in the form

$$p_3(w,t)\beta\gamma + p_2(w,t)\beta + p_1(w,t)\gamma + p_0(w,t) = 0, \tag{6.19}$$

where $p_j(w,t) \in \mathbb{Z}[w,t]$ for $j = 0,1,2,3$.

Clearly, (6.12) can be written as

$$3\beta^2 - 2(\lambda + w)\beta + \lambda w + 1 = 0 \tag{6.20}$$

and (6.13) can be written as

$$3w\gamma^2 - 2(\lambda w + 1)\gamma + \lambda + w = 0. \tag{6.21}$$

Observe that (6.17), (6.20), (6.21) are three equations in four variables β, γ, λ, w. Here, β and γ are the values of the corresponding functions at the critical point (ϕ, t) of h. By computing the resultant of (6.17) and (6.20) with respect to β, we shall obtain a polynomial F_0 in γ, λ, w. Then, by computing the resultant of this polynomial F_0 and the polynomial on the right hand side of (6.21) with respect to γ, we shall obtain the polynomial

$$\begin{aligned} F_1(w,\lambda) = &-3w^2(w-1)^2(w+1)^2(\lambda-2)^2(\lambda+2)^2(w^2 - \lambda w + 1) \\ &((\lambda^4 - 2\lambda^2 - 3)w^2 - (\lambda^5 - 6\lambda^3 + 13\lambda)w + \lambda^4 - 2\lambda^2 - 3). \end{aligned}$$

Similarly, we can substitute $\lambda = t + 1/t$ in (6.20) and (6.21), multiply them by t and then obtain a polynomial $G_1(w,t)$ as a result of first computing the resultant G_0 of (6.19) (obtained from (6.18)) and (6.20) with respect to β and then the resultant of this polynomial G_0 in γ, w, t and the polynomial (6.21) with respect to γ. The polynomial

$$G_1(w,t) = 8(w^7 + w^5)t^{23} + (21w^8 + 54w^6 + 21w^4)t^{22} + \cdots - 888(w^7 + w^5)t + 1152w^6$$

is too 'large' to be reproduced here. It has 120 nonzero terms, its total degree is 30, its degree in w is 12, and its degree in t is 23. (Of course, the polynomial $G_1(w,t)$ can be quickly computed with Maple exactly as it is described above.)

Since $w \neq \pm 1$ and $\lambda > 2$, only factor of F_1 which may vanish at points w with

modulus 1 is

$$F_2(w,\lambda) = (\lambda^4 - 2\lambda^2 - 3)w^2 - (\lambda^5 - 6\lambda^3 + 13\lambda)w + \lambda^4 - 2\lambda^2 - 3.$$

We substitute $\lambda = t + 1/t$ into F_2 and eliminate t by computing the resultant of $t^5 F_2(w, t + 1/t)$ and $G_1(w,t)$ with respect to t. This resultant $F_3(w)$ factors in $\mathbb{Z}[w]$ into irreducible factors as follows:

$$F_3(w) = -4398046511104(w-1)^{32}(w+1)^{32}F_4(w)F_5(w)F_6(w)F_6^*(w),$$

where

$$\begin{aligned}
F_4(w) =& 121w^{16} - 1184w^{14} + 3086w^{12} - 4544w^{10} \\
&+ 19379w^8 - 4544w^6 + 3086w^4 - 1184w^2 + 121, \\
F_5(w) =& 7623w^{24} - 51426w^{22} + 150903w^{20} - 294044w^{18} \\
&+ 469180w^{16} - 620138w^{14} + 694272w^{12} - 620138w^{10} \\
&+ 469180w^8 - 294044w^6 + 150903w^4 - 51426w^2 + 7623, \\
F_6(w) =& 5887w^{20} + 140216w^{18} + 48780w^{16} + 358498w^{14} \\
&- 1562910w^{12} + 1867560w^{10} - 914417w^8 \\
&+ 127414w^6 - 15237w^4 + 21096w^2 + 4356.
\end{aligned}$$

The polynomials F_6, F_6^* are irreducible in $\mathbb{Z}[w]$ and non-reciprocal, so they have no zeros of modulus 1. A simple numerical computation with Maple with extended precision shows that neither F_4 nor F_5 have such zeros. It follows that the function $h(\phi, t)$ has no critical points if $0 < \phi < \pi$. This proves our claim and completes the proof of Theorem 6.5. □

Another way to compute the extremal value t is to use (6.18). Since it is already established that the minimum is attained at $w = -1$ (and so $\beta = \gamma$), by (6.18), we have

$$\frac{2\beta + 1}{(\lambda - 1)(\beta + 1)} = \frac{2t}{t^2 - 1}.$$

Thus $\beta = (t^2 - 2t + 3)/2(t - 2)$. Substituting this expression into $3\beta^2 - (t + 1/t - 1)(2\beta + 1) = 0$ (which is obtained from (6.20) with $w = -1$), we find that $3t(t^2 - 2t + 3)^2 = 4(t^2 - t + 1)^2(t - 2)$. Hence $(t^3 - 6t^2 + 9t - 8)(t + 1)^2 = 0$, giving $t^3 - 6t^2 + 9t - 8 = 0$. Thus

$$t = (3 + 2\sqrt{2})^{1/3} + (3 - 2\sqrt{2})^{1/3} + 2 = 4.355301398\ldots$$

is the critical point of $h(\pi, t)$ and $f(z) = (z+1)(z-t)(z-1/t)$ is a corresponding

extremal polynomial.

6.5 L^s norms of a polynomial and its derivative

A direct analogue of (6.1) is the upper bound

$$||f'||_s \leqslant d||f||_s, \tag{6.22}$$

which holds for every $s > 0$ and every $f \in \mathbb{C}[z]$ of degree d. For $s = \infty$ inequality (6.22) is a classical inequality of Bernstein. Later, Zygmund [218] proved (6.22) for $1 \leqslant s < \infty$, while Arestov [13] established the inequality for $0 < s < 1$. Fix any $s > 0$. Which values does the quotient $||f'||_s/||f||_s$ take as f runs through self-inversive polynomials of degree d?

For $s = \infty$, we have $||f||_\infty = \max_{|z|=1} |f(z)|$. Theorem 14.3.1 of [171] asserts that if $f \in \mathbb{C}[z]$ is a self-inversive polynomial of degree $d \geqslant 1$, then

$$||f'||_\infty = \frac{d}{2}||f||_\infty.$$

So the quotient $||f'||_\infty/||f||_\infty$, where $f \in \mathcal{S}_d$, takes only one value $d/2$.

The upper bound for $||f'||_s/||f||_s$, where $f \in \mathbb{C}[z]$ is a self-inversive polynomial of degree d, was established by Rahman and Schmeisser [170]. Combining Theorem 14.6.5 and Remark 14.6.6 of [171], we have

$$\frac{||f'||_s}{||f||_s} \leqslant \frac{d}{||z^d + 1||_s} = \frac{d}{||z + 1||_s} = \frac{d}{2}\Big(\frac{\sqrt{\pi}\Gamma(s/2 + 1)}{\Gamma(s/2 + 1/2)}\Big)^{1/s} \tag{6.23}$$

for every $f \in \mathcal{S}_d$. Equality in (6.23) is attained for the polynomial $f(z) = z^d + 1 \in \mathcal{R}_d$.

Since the integrands $|f(e^{it})|^s$ and $|f'(e^{it})|^s$, where f is a polynomial, are defined for every point $t \in [0, 2\pi]$, Corollary 6.7 immediately implies the following lower bound for L^s-norms of a self-inversive polynomial and its derivative:

Corollary 6.10. *Suppose that $0 < s < \infty$. If $f \in \mathbb{C}[z]$ is a self-inversive polynomial of degree $d \geqslant 2$, then $||f'||_s > \frac{d}{2}||f||_s$.*

The bound of Corollary 6.10 is best possible for d even. Indeed, one can take the same example $f_n(z) = z^d - nz^{d/2} + 1$ with n large as in the proof of Theorem 6.3. Then, for every fixed $s > 0$, we have $||f_n||_s \sim n$ and $||f_n'||_s \sim dn/2$ as $n \to \infty$. Hence $||f_n'||_s/||f_n||_s \to d/2$ as $n \to \infty$. Combining with (6.23), by continuity, we deduce that for every $0 < s < \infty$ the quotient $||f'||_s/||f||_s$ takes every value in the interval $(d/2, \big(\sqrt{\pi}\Gamma(s/2 + 1)/\Gamma(s/2 + 1/2)\big)^{1/s} d/2]$ as f runs through $\mathcal{S}_d$ (or

$\mathcal{R}_d$), where $d \geqslant 2$ is even. For example, one can take $f(z) = z^d - uz^{d/2} + 1 \in \mathcal{R}_d$ with $u \in [0, +\infty)$. Corollary 6.10 for $p \geqslant 1$ was already established in [15]: see Theorem 2 and Remark 2 in [15], and also the earlier paper [14].

The question of finding $\inf_{f\in\mathcal{S}_d} ||f'||_s/||f||_s$ and $\inf_{f\in\mathcal{R}_d} ||f'||_s/||f||_s$ remains open for $d \geqslant 3$ odd. One case when both infimums can be established is $s = 2$. For $f(z) = a_d z^d + a_{d-1} z^{d-1} + \cdots + a_0$ we have $||f||_2 = \sqrt{|a_d|^2 + |a_{d-1}|^2 + \cdots + |a_0|^2}$. Thus if $f \in \mathbb{C}[z]$ is a self-inversive polynomial of degree $d = 1$ then $f(z) = a(z+\theta)$ with $a \neq 0$ and $|\theta| = 1$. Hence $||f'||_2 = \frac{1}{\sqrt{2}}||f||_2$.

For $d \geqslant 3$ we prove the following:

Theorem 6.11. *If $f \in \mathbb{C}[z]$ is a self-inversive polynomial of odd degree $d \geqslant 3$, then $||f'||_2 > \frac{\sqrt{d^2+1}}{2}||f||_2$. This inequality is best possible.*

Proof: Without loss of generality we may assume that f is monic. Then

$$f(z) = z^d + a_1 z^{d-1} + \cdots + a_{(d-1)/2} z^{(d+1)/2} + \theta\overline{a_{(d-1)/2}} z^{(d-1)/2} + \cdots + \theta\overline{a_1} z + \theta$$

with some $\theta \in \mathbb{C}$ of modulus 1. Hence

$$||f||_2^2 = 2(1 + |a_1|^2 + |a_2|^2 + \cdots + |a_{(d-1)/2}|^2)$$

and

$$||f'||_2^2 = d^2 + ((d-1)^2 + 1^2)|a_1|^2 + ((d-2)^2 + 2^2)|a_2|^2 + \cdots + \frac{d^2+1}{2}|a_{(d-1)/2}|^2.$$

The coefficients $(d-j)^2 + j^2$ are greater than $(d^2+1)/2$ for each j in the range $0 \leqslant j < (d-1)/2$. Therefore,

$$||f'||_2^2 > \frac{d^2+1}{2}(1 + |a_1|^2 + |a_2|^2 + \cdots + |a_{(d-1)/2}|^2) = \frac{d^2+1}{4}||f||_2^2.$$

This yields the required inequality $||f'||_2 > \frac{\sqrt{d^2+1}}{2}||f||_2$ for $d \geqslant 3$ odd.

In order to show that this inequality is tight, let us take the reciprocal polynomial $f_n(z) = z^d + n(z^{(d+1)/2} + z^{(d-1)/2}) + 1 \in \mathbb{Z}[z]$. We have $||f_n||_2^2 = 2(1+n^2)$ and $||f_n'||_2^2 = d^2 + (d^2+1)n^2/2$. So $\lim_{n\to\infty} ||f_n'||/||f_n|| = \sqrt{d^2+1}/2$. □

Theorem 6.11 implies that

$$\inf_{f\in\mathcal{S}_d} ||f'||_2/||f||_2 = \inf_{f\in\mathcal{R}_d} ||f'||_2/||f||_2 = \sqrt{d^2+1}/2$$

for $d \geqslant 3$ odd. Note that the right hand side of (6.23) for $s = 2$ is equal to $d/\sqrt{2}$.

By continuity, i.e., taking

$$f(z) = z^d + u(z^{(d+1)/2} + z^{(d-1)/2}) + 1 \in \mathcal{R}_d,$$

where $u \in [0, +\infty)$, we deduce that if f runs through every reciprocal polynomial f of odd degree $d \geqslant 3$ then the quotient $||f'||_2/||f||_2$ takes every value in the interval $(\sqrt{d^2+1}/2, d/\sqrt{2}]$.

6.6 Numerical examples

By the above results, it seems that the constant $d/2$ in the L^s norm inequality $||f'||_s > \frac{d}{2}||f||_s$, where $0 \leqslant s < \infty$, is not optimal for self-inversive polynomials $f(z)$ of odd degree d. In general, the computation of the infimum $||f'||_s/||f||_s$ over self-inversive polynomials f of odd degree d, is non-trivial, except in two cases $s = 2$ (Theorem 6.11) and $s = \infty$.

We give several examples of integer monic self-reciprocal polynomials with low value of $M(f')/M(f)$ (corresponding to the case $s = 0$), computed with Maple.

Table 6.1: Monic reciprocal polynomials $f \in \mathbb{Z}[z]$ of small odd degree d and low value $M(f')/M(f)$.

d	$f(z)$	$M(f)$	$M(f')$	$M(f')/M(f)$
3	$1, -4, -4, 1$	4.79128	9.29150	1.93924
5	$1, -2, 7, 7, -2, 1$	9.08609	26.58617	2.92602
7	$1, -1, 2, -8, -8, 2, -1, 1$	10.60148	41.56526	3.92070
9	$1, -1, 2, -3, 11, 11, -3, 2, -1, 1$	15.15276	74.51235	4.91740

Theorem 6.5 gives a precise answer for cubic self-reciprocal polynomials $f(z) \in \mathbb{C}[z]$. The cubic polynomial in Table 1 is the monic polynomial with integer coefficients closest to the optimal cubic polynomial in Theorem 6.5. The value $1.93924\ldots$ is the minimal value of $M(f')/M(f)$ among all monic cubic reciprocal polynomials with integer coefficients. For $d = 5$, we searched for optimal monic reciprocal polynomial with integer coefficients in the interval $[-100, 100]$. For $d = 7$ and $d = 9$, the search interval was reduced to $[-20, 20]$ and to $[-15, 15]$, respectively. We must note that the values of $M(f')/M(f)$ in Table 1 are not far from $(d+1)/2$ (see also Theorem 6.4), but the distance from those values to $(d+1)/2$ is increasing with d.

In the case $d = 5$, we have also computed the polynomial

$$g(z) = z^5 - 1.732z^4 + 6.165z^3 + 6.165z^2 - 1.732z + 1$$

with $M(g')/M(g) = 2.92557564\ldots$. It seems that this value is close to the infimum of $M(f')/M(f)$, where $f \in \mathcal{R}_5$.

Chapter 7

Intersections of arithmetic and geometric progressions

7.1 Statement of the problem

Let

$$\mathcal{G} = \mathcal{G}(u, q) := u, uq, uq^2, uq^3, \dots \tag{7.1}$$

be an infinite geometric progression with the first term $u > 0$ and the ratio $q > 1$. Let also

$$\mathcal{A} = \mathcal{A}(v, D) := v, v + D, v + 2D, v + 3D, \dots \tag{7.2}$$

be an infinite arithmetic progression with the first term $v \geqslant 0$ and the difference $D > 0$. Arithmetic and geometric progressions appear everywhere from elementary mathematics to the celebrated Green-Tao theorem on prime numbers that are consecutive terms of arithmetic progressions of arbitrary finite length. Obviously, for every $\mathcal{G}$ given in (7.1), there is an $\mathcal{A}$, where, e.g., $v := u$ and $D := u(q-1)$, as in (7.2) such that the intersection $\mathcal{G} \cap \mathcal{A}$ contains at least 2 elements. In this Chapter we are interested in the following natural question:

Problem 7.1. *How large can the intersection $\mathcal{G} \cap \mathcal{A}$ be?*

It is easily seen that $\mathcal{G} \cap \mathcal{A}$ contains at most two elements if the ratio q of $\mathcal{G}$ is a transcendental number. Indeed, suppose that uq^a, uq^b, uq^c, where $0 \leqslant a < b < c$ are integers, are three elements of the arithmetic progression $v + Dm$, $m = 0, 1, 2, \dots$, with indices $m_1 < m_2 < m_3$. Then

$$\frac{q^{c-a} - 1}{q^{b-a} - 1} = \frac{uq^c - uq^a}{uq^b - uq^a} = \frac{m_3 - m_1}{m_2 - m_1} \in \mathbb{Q}.$$

Thus q is a root of a polynomial in $\mathbb{Z}[x]$ which means that q must be an algebraic number over $\mathbb{Q}$.

The following result is a consequence of [66] and formula (7.5). It is stated here only for the sake of completeness.

Theorem 7.2. *For any geometric progression $\mathcal{G}$ with ratio $q > 1$, we have $|\mathcal{G} \cap \mathcal{A}| = \infty$ for some arithmetic progression $\mathcal{A}$ if and only if $q = m^{1/d}$ for some integers $m \geqslant 2$ and $d \geqslant 1$.*

The next result covers the ratios $q = r^{1/d}$, where $r > 1$ is a rational number and $d \in \mathbb{N}$. It shows that $|\mathcal{G} \cap \mathcal{A}|$ can take any nonnegative integer value:

Theorem 7.3. *Suppose that $r > 1$ is a rational non-integer number, $d \in \mathbb{N}$ and $q = r^{1/d}$. Then, for every nonnegative integer s, there is a geometric progression $\mathcal{G}$ with ratio q which contains exactly s positive integers, so that $|\mathcal{G} \cap \mathbb{N}| = s$.*

The main theorems of Chapter 7 are the following:

Theorem 7.4. *Suppose that the ratio $q > 1$ is not of the form $\beta^{1/d}$, where $d \in \mathbb{N}$ and where $\beta > 1$ is either a rational number or a cubic algebraic number with two nonreal conjugates over $\mathbb{Q}$ of moduli distinct from β. Then $|\mathcal{G} \cap \mathcal{A}| \leqslant 3$ for each $\mathcal{G} = \mathcal{G}(u, q)$ and each $\mathcal{A}$.*

Moreover, for every integer $s \geqslant 2$, there exist an algebraic number $q > 1$ of degree s (satisfying $q \neq \beta^{1/d}$ for $s \neq 3$) and a positive real number $u \in \mathbb{Q}(q)$ such that $|\mathcal{G} \cap \mathcal{A}| = 3$ for the geometric progression $\mathcal{G} = \mathcal{G}(u, q)$ and some arithmetic progression $\mathcal{A}$.

Theorem 7.5. *For $q = \beta^{1/d}$, where $d \in \mathbb{N}$ and $\beta > 1$ is a cubic algebraic number with two nonreal conjugates of moduli distinct from β, we have $|\mathcal{G} \cap \mathcal{A}| \leqslant 6$ for each $\mathcal{G} = \mathcal{G}(u, q)$ and each $\mathcal{A}$.*

In the proof of Theorem 7.5 we use a deep (and sharp!) result of Beukers [25] on the zero multiplicity of ternary recurrence sequences. Unlike in all other cases, we do not have any examples with $|\mathcal{G} \cap \mathcal{A}|$ equal to $4, 5$ or 6 under conditions of Theorem 7.5. So we conjecture that the sharp upper bound on $|\mathcal{G} \cap \mathcal{A}|$, where $q = \beta^{1/d}$ with a cubic β and d as in Theorem 7.5, should be 3, the same as in Theorem 7.4.

The problems on the intersection of $\mathcal{G}$ and $\mathcal{A}$ can be easily transformed into the language of multiplicities of fractional parts. This will be explained in Section 2, where we give some additional motivation for the study of those problems and also remind some earlier relevant results. In particular, the reduction of the problem on the size of $|\mathcal{G} \cap \mathcal{A}|$ to the corresponding problem on multiplicities of the fractional parts of the geometric sequence $\mathcal{G}$ shows that the result given in [66] implies Theorem 7.2. The proofs of Theorems 7.3, 7.4, 7.5 will be given in Section 4 after stating all necessary auxiliary lemmas and other results in Section 3.

7.2 Fractional parts of geometric progressions

Throughout, let $\{y\}$ be the fractional part of a real number y. Let also $\xi > 0$, $\alpha > 1$ and $t \in [0, 1)$ be arbitrary real numbers. We shall consider the fractional parts of powers $\{\xi\alpha^n\}$, $n = 1, 2, 3, \ldots$. Let $M_{\xi,\alpha}(t)$ (or simply $M(t)$) be the number of times the value t occurs in the sequence $\{\xi\alpha^n\}$, $n = 1, 2, 3, \ldots$. We call the number $M_{\xi,\alpha}(t)$ the *multiplicity* of t in the sequence $\{\xi\alpha^n\}$, $n = 1, 2, 3, \ldots$. In other words, $M_{\xi,\alpha}(t)$ is the number of positive integers n for which $\{\xi\alpha^n\} = t$. For example, $M_{1,3/2}(1/3) = 0$ and $M_{1,\sqrt{2}}(0) = \infty$.

Let us consider two progressions $\mathcal{G}$ and $\mathcal{A}$ as in (7.1), (7.2). We are interested in the upper bound for $|\mathcal{G} \cap \mathcal{A}|$, so assume, without loss of generality, that $0 \leqslant v < D$. Evidently, $uq^j = v + iD$ for some integers $i, j \geqslant 0$ if and only if

$$\frac{u}{qD} q^{j+1} = \frac{v}{D} + i. \tag{7.3}$$

Setting

$$\xi := u/qD, \quad \alpha := q, \quad n := j + 1, \quad t := v/D \tag{7.4}$$

in (7.3), we see that $\{\xi\alpha^n\} = t$. So each element of $\mathcal{G} \cap \mathcal{A}$ corresponds to the solution of the equation $\{\xi\alpha^n\} = t$ in positive integers n. In particular,

$$|\mathcal{G} \cap \mathcal{A}| = M_{\xi,\alpha}(t) \tag{7.5}$$

for $\mathcal{G} = \mathcal{G}(u, q)$, $\mathcal{A} = \mathcal{A}(v, D)$ with $u > 0, q > 1, D > 0$, $v \in [0, D)$ and ξ, α, t given in (7.4).

In [66], using the result of Boyd [46], the first named author proved that $M_{\xi,\alpha}(t)$ can be infinite for $\xi > 0$, $\alpha > 1$ and $t \in [0, 1)$ if and only if $\alpha = m^{1/d}$ and $\xi = \frac{a}{b} m^{\ell/d}$ for some integers $m \geqslant 2$, $d, a, b \in \mathbb{N}$, $\ell \in \{0, 1, \ldots, d-1\}$. By (7.5), this yields Theorem 7.2.

The problem of determining the multiplicity $M_{1,\alpha}(t)$, where $\alpha > 1$, has been studied by Supnick, Cohen and Keston [205] and also by Ehlich [80] and Posner [164]. They proved (independently) the following result which was conjectured by Vijayaraghavan (see p. 153 in [204]). If $\{\alpha^a\} = \{\alpha^b\} = \{\alpha^c\}$, where $a < b < c$ are positive integers, then $\alpha^a, \alpha^b, \alpha^c$ are integers. In other words, the multiplicity of the sequence $\{\alpha^n\}$, $n = 1, 2, 3, \ldots$, satisfies $M(t) \leqslant 2$, unless $\alpha > 1$ is a root of an integer in which case $M(0) = \infty$. The above bound 2 is attained. Indeed, the identity $\{\alpha^a\} = \{\alpha^b\}$ holds for any number $\alpha > 1$ which is a root of the trinomial

$$x^b - x^a - k,$$

where $k, a, b \in \mathbb{N}$, $a < b$. Thus, if this root α is not of the form $m^{1/d}$ with some integers $m \geqslant 2$ and $d \geqslant 1$, we have $M_{1,\alpha}(\{\alpha^a\}) = 2$. Clearly, $M_{1,\alpha}(0) = 0$ and $M_{1,\alpha}(\{\alpha\}) = 1$ for each transcendental number α. This shows that, for every $\alpha > 1$ and every $t \in [0, 1)$, we may have only the following multiplicities:

$$M_{1,\alpha}(t) \in \{0, 1, 2, \infty\}. \tag{7.6}$$

There are many more cases for general ξ. In contrast to (7.6), by Theorems 7.2, 7.3 and (7.5), we obtain

$$M_{\xi,\alpha}(t) \in \mathbb{N} \cup \{0, \infty\},$$

where each value in $\mathbb{N} \cup \{0, \infty\}$ occurs for some ξ, α, t.

There exists a direct correspondence between equal values of the fractional parts $\{\xi\alpha^n\}$ for different n's and equal values which appear in certain linear recurrent sequences. Indeed, let α be an algebraic number of degree d and $\xi \in \mathbb{Q}(\alpha)$. Then one can write $\xi\alpha^n$ as a linear combination

$$\xi\alpha^n = c_{d-1,n}\alpha^{d-1} + c_{d-2,n}\alpha^{d-2} + \cdots + c_{1,n}\alpha + c_{0,n}$$

in the basis $1, \alpha, , \ldots, \alpha^{d-1}$ with rational coefficients $c_{j,n}$, $j = 0, 1, \ldots, d-1$. It is easy to see that the coefficients $c_{j,n}$ (for a fixed index j) satisfy a homogenous linear recurrence of order d. The minimal polynomial of α is the characteristic polynomial of the linear recurrence. Then $\{\xi\alpha^a\} = \{\xi\alpha^b\}$ holds precisely if

$$\xi\alpha^a - \xi\alpha^b = \sum_{j=0}^{d-1}(c_{j,a} - c_{j,b})\alpha^j \in \mathbb{Z},$$

which implies the simultaneous $d-1$ equalities $c_{j,a} = c_{j,b}$ for $j = 1, 2, \ldots, d-1$.

Integer, rational, algebraic and complex linear recurrences with repeating terms received considerable attention for a long time. See, for instance, [57], [198]. Effective upper bounds for the multiplicity of general linear recurrent sequences were obtained by Schlickewei [190], [191], Schmidt [192], [193] and then improved in some cases in [8], [12], [84]; see also a survey [194]. The complete classification of rational binary and ternary recurrences of highest multiplicity was accomplished by Beukers [24], [25]. These results provide additional tools and new motivation for the study of the multiplicity function $M_{\xi,\alpha}(t)$.

7.3 Lemmata

We shall derive Theorems 7.4 and 7.5 from the next theorem combined with Theorem 7.8 below and (7.5).

Theorem 7.6. *Let $\xi > 0$ and $\alpha > 1$ be arbitrary real numbers. Then the fractional parts $\{\xi\alpha^n\}$, $n = 1, 2, 3, \ldots$, take a fixed value $t \in [0, 1)$ at most 3 times, with two possible exceptions. The first exception occurs for $\alpha = r^{1/d}$, where $r \in \mathbb{Q}$, $r > 1$ and $d \in \mathbb{N}$. The other (possible) exception may occur for $\alpha = \beta^{1/d}$, where $\beta > 1$ is a real cubic algebraic number with two nonreal conjugates of moduli distinct from β. In the cubic case, the sequence of fractional parts can take the same value at most 6 times.*

We first give a simple lemma stating necessary and sufficient conditions for the sequence of fractional parts of powers to take some values two or three times. Firstly, the inequality $M(t) \geqslant 2$ implies that there exist two distinct positive integers, say, $a < b$ such that $t = \{\xi\alpha^b\} = \{\xi\alpha^a\}$. This is equivalent to the fact that the difference $k = \xi\alpha^b - \xi\alpha^a$ is a positive integer, thus $\xi = k/(\alpha^b - \alpha^a)$.

Secondly, assume that $M(t) \geqslant 3$ for some $t \in [0, 1)$. Now, there exist three positive integers $a < b < c$ such that the differences $k = \xi\alpha^b - \xi\alpha^a$ and $l = \xi\alpha^c - \xi\alpha^a$ belong to $\mathbb{N}$. This implies

$$\xi = \frac{k}{\alpha^b - \alpha^a} = \frac{l}{\alpha^c - \alpha^a}, \qquad k < l, \quad k, l \in \mathbb{N}. \tag{7.7}$$

These equations are *necessary and sufficient* for the sequence $\{\xi\alpha^n\}$, $n = 1,2,3,\ldots$, to take values of multiplicity at least 3. The second equality in (7.7) can be rewritten as $k\alpha^c - l\alpha^b + (l-k)\alpha^a = 0$. Hence α must be a root of the trinomial (a polynomial with three nonzero coefficients) $kx^{c-a} - lx^{b-a} + l - k$. In particular, α is an algebraic number over $\mathbb{Q}$. Summarizing, we have the following lemma:

Lemma 7.7. *The sequence $\{\xi\alpha^n\}$, $n = 1, 2, 3, \ldots$, takes two equal values at $n = a, b$ if and only if $\xi = k/(\alpha^b - \alpha^a)$, where $a < b$, $a, b, k \in \mathbb{N}$. It takes three equal values at $n = a, b, c$ if and only if α is a root of the trinomial $T(x) = kx^{c-a} - lx^{b-a} + l - k \in \mathbb{Z}[x]$, $a, b, c, k, l \in \mathbb{N}$, $a < b < c$, $k < l$ and ξ is the form given by* (7.7).

Lemma 7.7 allows us to construct infinitely many examples of algebraic numbers α and ξ, such that the multiplicity of the sequence $\{\xi\alpha^n\}$, $n = 1, 2, 3, \ldots$, is greater than or equal to 3. By Lemma 7.12 bellow, $T(x)$ has a root $\alpha > 1$ if and only if $l/k > (c-a)/(b-a)$. For instance, selecting in Lemma 7.7 $a = 1$, $b = 2$,

$c = 4$, $k = 1$, $l \geqslant 4$ one has

$$T(x) = x^3 - lx + l - 1 = (x-1)(x^2 + x - l + 1).$$

Then $\alpha = (-1 + \sqrt{4l-3})/2 > 1$ is quadratic over $\mathbb{Q}$ provided that $4l - 3$ is not a perfect square. Set $\xi := 1/(\alpha^2 - \alpha)$. Then $t = \{\xi\alpha\} = \{\xi\alpha^2\} = \{\xi\alpha^4\}$. Hence $M(t) \geqslant 3$. In fact, we have $M(t) = 3$, by Theorem 7.6, because α is not of the form $\beta^{1/d}$ (see the proof of Theorem 7.8 below).

More generally, given any integer $s \geqslant 2$, let us take in Lemma 7.7 $k = 1$, $a = 1$, $b = s + 1$, $c = s + 2$, $l \geqslant 2$, $l \in \mathbb{N}$. Then $\alpha > 1$ is a root of the trinomial

$$T(x) = x^{s+1} - lx^s + l - 1 = (x-1)(x^s - (l-1)(x^{s-1} + \cdots + x + 1)).$$

The polynomial

$$P(x) := x^s - (l-1)(x^{s-1} + \cdots + x + 1), \tag{7.8}$$

where $l \geqslant 2$, divides T. Hence, by Rouché's theorem, it has a unique root outside the unit circle, and so is irreducible. This unique root must be α. By Lemma 7.7, for this α and for $\xi := 1/(\alpha^{s+1} - \alpha)$, $t := \{\xi\alpha\}$, we have

$$\{\xi\alpha\} = \{\xi\alpha^{s+1}\} = \{\xi\alpha^{s+2}\} = t. \tag{7.9}$$

Below, using these α and ξ, we shall prove the following theorem:

Theorem 7.8. *For every integer $s \geqslant 2$, there is an algebraic number $\alpha > 1$ (which, for $s \neq 3$, is not of the form $\beta^{1/d}$ with $d \in \mathbb{N}$ and β either a rational or a cubic number with two nonreal conjugates) of degree s over $\mathbb{Q}$, a positive number $\xi \in \mathbb{Q}(\alpha)$ and $t \in [0, 1)$ such that $M_{\xi,\alpha}(t) = 3$.*

In order to prove Theorem 7.6, we shall use the following lemma.

Lemma 7.9. *Suppose that a real number $\beta > 1$ satisfies the equations*

$$\frac{k}{\beta^u - 1} = \frac{l}{\beta^v - 1} = \frac{m}{\beta^w - 1}, \tag{7.10}$$

where $k < l < m$ and $u < v < w$ are positive integers satisfying $\gcd(u, v, w) = 1$. Then β is a rational number, a real quadratic number or a cubic number with two nonreal conjugates. Moreover, if β' is conjugate to β over $\mathbb{Q}$, $\beta' \neq \beta$, then $|\beta'| \neq \beta$.

We shall also need two simple lemmas on the roots of trinomials. Lemmas

7.10 and 7.11 will be used in the proof of Lemma 7.9. The first one is essentially given on p. 248 in [205].

Lemma 7.10. *Suppose that the complex number $z = \rho e^{i\phi}$, where $\rho = |z| > 0$ and $\phi = \arg z \in [0, 2\pi)$, is a root of the trinomial $f(x) = Ax^r + Bx^s + C$, where A, B, C are nonzero real numbers and r, s are distinct positive integer exponents. Then*

$$\cos r\phi = -(A^2\rho^{2r} - B^2\rho^{2s} + C^2)/(2AC\rho^r).$$

Proof. The equation

$$f(\rho e^{i\phi}) = A\rho^r(\cos r\phi + i \sin r\phi) + B\rho^s(\cos s\phi + i \sin s\phi) + C = 0$$

implies

$$A\rho^r \cos r\phi + C = -B\rho^s \cos s\phi, \tag{7.11}$$

$$A\rho^r \sin r\phi = -B\rho^s \sin s\phi. \tag{7.12}$$

Squaring both sides of (7.11) and (7.12) and adding them we obtain

$$A^2\rho^{2r} + 2AC\rho^r \cos r\phi + C^2 = B^2\rho^{2s}.$$

The result now follows. □

The second Lemma was proved by Posner and Rumsey [165]. (We also give a short proof for the sake of completeness.)

Lemma 7.11. *Suppose that the complex numbers z_1 and z_2 are roots of the trinomial $f(x) = Ax^r + Bx^s + C \in \mathbb{R}[x]$. If $|z_1| = |z_2|$, then one has either $z_1^r = z_2^r$ or $z_1^r = \overline{z_2^r}$, where $\overline{z}$ denotes the complex conjugate of $z \in \mathbb{C}$.*

Proof. Let $\phi_1 = \arg z_1$, $\phi_2 = \arg z_2$. Set $|z_1| = |z_2| = \rho > 0$ in Lemma 7.10. The formula of Lemma 7.10 then gives $\cos r\phi_1 = \cos r\phi_2$. This yields $\sin r\phi_1 = \pm \sin r\phi_2$. Hence $z_1^r = z_2^r$ or $z_1^r = \overline{z_2^r}$. □

Here is another useful observation.

Lemma 7.12. *The trinomial $f(x) = Ax^r - Bx^s + B - A \in \mathbb{R}[x]$, where $r > s > 0$, $A > 0$, $B > 0$ are integers, has a real root $y > 1$ if and only if $B/A > r/s$.*

Proof. Note that $f(1)=0$. The derivative $f'(x) = rAx^{r-1} - sBx^{s-1}$ has one positive real root $x_0 = (sB/rA)^{1/(r-s)} > 1$ if and only if $B/A > r/s > 1$. The necessity now follows from the theorem of Rolle. Conversely, $B/A > r/s > 1$ implies

$f'(x) < 0$ for any positive $x < x_0$. Thus $f(x) < 0$ in $(1, x_0)$ and $f(+\infty) = +\infty$. Hence $f(y) = 0$ for some $y \geqslant x_0 > 1$. □

In addition, we make use of two classical results from the theory of linear recurrent sequences. Recall that a nonzero (which means that $u_n \neq 0$ for at least one $n \geqslant 0$) linear recurrence sequence $u_n, n = 0, 1, 2, \ldots$, satisfying

$$u_{n+d} = c_{d-1}u_{n+d-1} + \cdots + c_0 u_n$$

for $n = 0, 1, 2, \ldots$ is called *non-degenerate*, if the quotient α'/α'' is not a root of unity for any two distinct roots α', α'' of its characteristic polynomial $x^d - c_{d-1}x^{d-1} - \cdots - c_0$. The sequences corresponding to $d = 2$ and $d = 3$ are called *binary* and *ternary*, respectively.

Theorem 7.13. *Suppose that the non-degenerate rational binary linear recurrent sequence u_n, $n = 0, 1, 2, \ldots$, has a characteristic polynomial with two distinct real roots. Then the multiplicity of any rational value in the sequence u_n is at most* 3.

Theorem 7.13 was first proved by Chowla, Dunton and Lewis [57]. Their proof for integer sequences works for the rational (and real) sequences as well. An alternative proof of this result using an earlier result of Smiley [198] is also given on p. 837 in [57].

To settle the cubic case in Theorem 7.6, we use the following result of Beukers [25] on the multiplicity of rational ternary linear recurrent sequences.

Theorem 7.14. *Let u_n, $n = 0, 1, 2, \ldots$, be a non-degenerate ternary linear recurrent sequence of rational numbers. Then the zero multiplicity of u_n is at most* 6.

The proofs of Lemma 7.9 and of all our theorems will be given in the next section.

7.4 Proofs

Proof of Lemma 7.9. The equation $k/(\beta^u - 1) = l/(\beta^v - 1)$ yields $k\beta^v - l\beta^u + l - k = 0$. Similarly, using $l/(\beta^v - 1) = m/(\beta^w - 1)$ we obtain $l\beta^w - m\beta^v + m - l = 0$. Thus β is a common root of the trinomials

$$P(x) := kx^v - lx^u + l - k \quad \text{and} \quad Q(x) := -mx^v + lx^w + m - l.$$

Let $z \in \mathbb{C}$ be any common root of $P(x)$ and $Q(x)$. Set $\rho := |z|$, $\phi := \arg z$.

Using Lemma 7.10 with $f(x) = P(x)$, $r = v$, one has

$$\cos v\phi = -\frac{k^2\rho^{2v} - l^2\rho^{2u} + (l-k)^2}{2k(l-k)\rho^v}.$$

Similarly, using Lemma 7.10 with $f(x) = Q(x)$, $r = v$, one obtains

$$\cos v\phi = -\frac{m^2\rho^{2v} - l^2\rho^{2w} + (m-l)^2}{2m(l-m)\rho^v}.$$

Hence

$$\frac{k^2\rho^{2v} - l^2\rho^{2u} + (l-k)^2}{k(l-k)} = \frac{m^2\rho^{2v} - l^2\rho^{2w} + (m-l)^2}{m(l-m)}.$$

This yields

$$kl(l-k)\rho^{2w} - mk(m-k)\rho^{2v} + ml(m-l)\rho^{2u} - (m-l)(m-k)(l-k) = 0.$$

By Descartes rule of signs, the quadrinomial

$$H(x) := kl(l-k)x^{2w} - mk(m-k)x^{2v} + ml(m-l)x^{2u} - (m-l)(m-k)(l-k)$$

has 1 or 3 positive real roots (counting with multiplicities), because $0 < u < v < w$ and $0 < k < l < m$. Observe that $H(1) = 0$ and, by (7.10), $H(\beta) = 0$. Hence $H(x)$ has three positive roots.

Let $\gamma > 0$ be a real number which is the third root of $H(x)$ (including the possible case of a multiple root $\gamma = 1$ or $\gamma = \beta$). Suppose that $\beta' \neq \beta$ is a conjugate of β over $\mathbb{Q}$. Since β' is a common root of P and Q, $|\beta'|$ is a root of H. Thus $H(|\beta'|) = 0$. Hence the number β and all its algebraic conjugates over $\mathbb{Q}$ must lie on three circles with the common center at $z = 0$ and the radii 1, β and γ.

If β has no other conjugates over $\mathbb{Q}$ then $\beta \in \mathbb{Q}$. Suppose β' is a conjugate of β over $\mathbb{Q}$ satisfying $\beta' \neq \beta$. We claim that β' cannot lie on either of the circles $|z| = 1$ or $|z| = \beta$. Suppose first that $|\beta'| = \beta$. Then, as β' and β lie on the same circle, by Lemma 7.11, one has either $\beta^r = \beta'^r$ or $\beta^r = \overline{\beta'^r}$ for every positive integer exponent r in $P(x)$ or $Q(x)$, namely, for $r = u, v, w$. Since β^r is a real number, we must have $\beta^r = \beta'^r$. Thus $\zeta^u = \zeta^v = \zeta^w = 1$, where $\zeta := \beta/\beta'$. This implies $\zeta = \zeta^{\gcd(u,v,w)} = 1$. So $\beta = \beta'$, a contradiction. Similarly, if β' lies on the circle $|z| = 1$, then $|\beta'| = 1$. Since $\beta' \neq \pm 1$, it must be a complex (nonreal) number. The number 1 is also the common root of $P(x)$ and $Q(x)$. Hence, as above, by Lemma 7.11, we obtain $\beta'^r = 1^r = 1$ for every positive integer exponent r in $P(x)$ or $Q(x)$, i.e., for $r = u, v, w$. So β' is a root of unity. However, its conjugate $\beta > 1$ over $\mathbb{Q}$ is not a root of unity, a contradiction.

Therefore, if $\beta \notin \mathbb{Q}$, then all its conjugates over $\mathbb{Q}$ (except for β itself but including β') must lie on the circle $|z| = \gamma$. In particular, $\gamma \neq 1$ and $\gamma \neq \beta$. Algebraic numbers lying with their conjugates on two circles were studied in [73]. However, as in [73] it is assumed that their norms are 1 (which is not the case here), we shall give an independent argument.

We first claim that β may have at most two such conjugates, so β is of degree 2 or 3 over $\mathbb{Q}$. Indeed, assume that β has at least three distinct conjugates on $|z| = \gamma$, namely, $\beta', \overline{\beta'}$ and, say, $\beta'' \notin \{\beta', \overline{\beta'}\}$. Then

$$\beta' \overline{\beta'} = \beta'' \overline{\beta''},$$

where $\overline{\beta''}$ is a conjugate of β which is equal to β'' if $\beta'' \in \mathbb{R}$. Taking an automorphism σ of the Galois group of $\mathbb{Q}(\beta)/\mathbb{Q}$ which maps β' to β we obtain

$$\beta \sigma(\overline{\beta'}) = \sigma(\beta'') \sigma(\overline{\beta''}). \tag{7.13}$$

However, the modulus of the left hand side of (7.13) is equal to $\beta\gamma$, whereas the modulus of its right hand side is equal to γ^2, because all conjugates of β except for β itself lie on the circle $|z| = \gamma$. This is a contradiction, because $\beta\gamma \neq \gamma^2$. Hence β is of degree at most 3 over $\mathbb{Q}$ and, in case $\deg \beta = 3$, β has two nonreal conjugates on $|z| = \gamma$, where $\gamma \neq \beta$. □

Proof of Theorem 7.6. Suppose that the multiplicity $M(t)$ of some sequence $\{\xi\alpha^n\}$, $n = 1, 2, 3, \ldots$, where $\xi > 0$, $\alpha > 1$, is greater than or equal to 4. Then there exist positive integers $a < b < c < e$ such that $\{\xi\alpha^a\} = \{\xi\alpha^b\} = \{\xi\alpha^c\} = \{\xi\alpha^e\}$. This occurs precisely if and only if the differences $k = \xi\alpha^b - \xi\alpha^a$, $l = \xi\alpha^c - \xi\alpha^a$, $m = \xi\alpha^e - \xi\alpha^a$ are positive integers $k < l < m$. Then

$$\xi\alpha^a = \frac{k}{\alpha^{b-a} - 1} = \frac{l}{\alpha^{c-a} - 1} = \frac{m}{\alpha^{e-a} - 1}. \tag{7.14}$$

Let $d := \gcd(b - a, c - a, e - a)$. Then $b - a = du$, $c - a = dv$, $e - a = dw$, where $u < v < w$, $d, u, v, w \in \mathbb{N}$ and $\gcd(u, v, w) = 1$. Set $\beta := \alpha^d$. Then, by (7.14), we find that

$$\frac{k}{\beta^u - 1} = \frac{l}{\beta^v - 1} = \frac{m}{\beta^w - 1}.$$

Hence, Lemma 7.9 implies that $\beta > 1$ is a rational number, a real quadratic number or a cubic number with two nonreal conjugates. In the first case, $\alpha = r^{1/d}, r \in \mathbb{Q}$, which is an exceptional case in the statement of the theorem. The third, cubic, case is also exceptional and will be considered below. Before this, let us examine the quadratic case.

Quadratic case. Set $\nu := k/(\beta^u - 1)$. Obviously, $\nu \in \mathbb{Q}(\beta)$. Observe that the fractional parts $\{\nu\}$, $\{\nu\beta^u\}$, $\{\nu\beta^v\}$ and $\{\nu\beta^w\}$ are equal, by (7.7). Write

$$\nu\beta^n = a_n\beta + b_n, \quad n = 0, 1, 2, \ldots,$$

where $a_n, b_n \in \mathbb{Q}$. The sequences a_n, $n = 0, 1, 2, \ldots$, and b_n, $n = 0, 1, 2, \ldots$, satisfy a linear homogeneous recurrence of order two with characteristic polynomial which is the minimal polynomial of β over $\mathbb{Q}$. Then $\nu\beta^u - \nu = k$, $\nu\beta^v - \nu = l$, $\nu\beta^w - \nu = m$ leads to $a_0 = a_u = a_v = a_w$, because all differences are positive integers. So the multiplicity of the value a_0 in the sequence a_n, $n = 0, 1, 2, \ldots$ at a_0 is at least 4.

By Theorem 7.13, the sequence a_n, $n = 0, 1, 2, \ldots$, must be degenerate. This means that either a_n, $n = 0, 1, 2, \ldots$, is the zero sequence or β/β' is a root of unity. In the first case, $a_1 = a_0 = 0$ implies $\beta = \nu\beta/\nu = b_1/b_0 \in \mathbb{Q}$, a contradiction. Also, by Lemma 7.9, β has one real conjugate $\beta' \neq \pm\beta$. Thus β/β' is not a root of unity. This is also a contradiction which shows that β (which is a power of α) cannot be a quadratic number with two real conjugates β, β' satisfying $|\beta'| \neq \beta$ in case $M_{\xi,\alpha}(t) \geqslant 4$.

Cubic case (with two nonreal conjugates). Now, we will show that $M(t) \leqslant 6$. Suppose that $M_{\xi,\alpha}(t) > 6$ for some $\xi > 0, \alpha > 1, t \in [0, 1)$, where $\beta = \alpha^d > 1$ is a cubic number with two nonreal conjugates and $d \in \mathbb{N}$. Assume that d is the smallest positive number for which α^d is a cubic number.

We claim that the number α^m is cubic if and only if $d|m$. There is nothing to prove for $d = 1$. Assume that $d > 1$. Clearly, $\alpha = \beta^{1/d}$ has conjugates on two circles $|z| = \beta^{1/d}$ and $|z| = \gamma^{1/d}$, where β' and $\beta'' = \overline{\beta'}$ are two conjugates of β lying on $|z| = \gamma$, $\gamma \neq \beta$. Moreover, the conjugates of $\alpha = \alpha_1$ lying on the circle $|z| = \beta^{1/d}$ must be of the form $\alpha_j = \zeta_j\beta^{1/d}$, $j = 2, \ldots, s$, where ζ_j is a root of unity satisfying $\zeta_j^d = 1$. Since α is not a root of a rational number or a quadratic number, $\deg(\alpha^m) = 3$ if and only if

$$\alpha_1^m = \alpha_2^m = \cdots = \alpha_s^m. \tag{7.15}$$

Write $m \in \mathbb{N}$ in the form $m = dk + l$, where $0 \leqslant l < d$. Assume that $l > 0$. Then $\alpha_1^{dk} = \alpha_2^{dk} = \cdots = \alpha_s^{dk}$ combined with (7.15) yields $\alpha_1^l = \alpha_2^l = \cdots = \alpha_s^l$ Thus $\deg(\alpha^l) = 3$, a contradiction with the minimality of d.

Since $M(t) > 6$, there exist some positive integers $e_0 < e_1 < \cdots < e_6$ such that $\{\xi\alpha^{e_0}\} = \cdots = \{\xi\alpha^{e_6}\}$. In particular, the differences $k_i = \xi\alpha^{e_i} - \xi\alpha^{e_0}$, $i = 1, \ldots, 6$, must be positive integers. Hence

$$\xi\alpha^{e_0} = \frac{k_i}{\alpha^{e_i - e_0} - 1}$$

for $i = 1, \dots, 6$. Set $g := \gcd(e_1 - e_0, \dots, e_6 - e_0)$ and $w_i := (e_i - e_0)/g \in \mathbb{N}$ for $i = 1, \dots, 6$. Then $\gcd(w_1, \dots, w_6) = 1$ and

$$\frac{k_1}{\alpha^{gw_1} - 1} = \cdots = \frac{k_6}{\alpha^{gw_6} - 1}. \tag{7.16}$$

Put $g_1 := \gcd(w_1, w_2, w_3)$ and $v_i := w_i/g_1 \in \mathbb{N}$ for $i = 1, 2, 3$. By Lemma 7.9 applied to the first two equalities of (7.16), we deduce that α^{gg_1} is a rational, a quadratic or a cubic number. Since α^d is cubic and no positive integer power of α is of degree smaller than 3, the number α^{gg_1} must be cubic. Hence d divides gg_1, by the above claim. It follows that d divides $gg_1v_1 = gw_1$, and gw_2, gw_3. By the same argument applied to the last two equalities of (7.16), d divides gw_4, gw_5, gw_6. Since $\alpha^d = \beta$, equalities (7.16) can be written in the form

$$\frac{k_1}{\beta^{u_1} - 1} = \cdots = \frac{k_6}{\beta^{u_6} - 1} \tag{7.17}$$

with positive integers $u_1 < \cdots < u_6$.

Set $\nu := k_1/(\beta^{u_1} - 1) \in \mathbb{Q}(\beta)$. Note that the fractional parts $\{\nu\}$, $\{\nu\beta^{u_1}\}$, $\dots$, $\{\nu\beta^{u_6}\}$ are equal, by (7.17). Thus the sequence $\{\nu\beta^n\}$, $n = 0, 1, 2, \dots$, takes the value $t = \{\nu\}$ at least seven times. Let

$$H(x) = x^3 - Ax^2 - Bx - C, \qquad A, B, C \in \mathbb{Q}, \tag{7.18}$$

be the minimal polynomial of β over $\mathbb{Q}$. Write

$$\nu\beta^n = a_n\beta^2 + b_n\beta + c_n, \qquad n = 0, 1, 2, \dots, \qquad a_n, b_n, c_n \in \mathbb{Q}. \tag{7.19}$$

Note that the sequences a_n, $n = 0, 1, 2, \dots$, and b_n, $n = 0, 1, 2, \dots$, satisfy a homogeneous linear recurrence of order 3 with characteristic polynomial $H(x)$. The relation $\nu\beta^m - \nu\beta^s \in \mathbb{Z}$ implies $a_m = a_s$, $b_m = b_s$. Hence, there exist $a, b \in \mathbb{Q}$ such that $(a_n, b_n) = (a, b)$ for seven distinct nonnegative integers n.

Consider the sequence $d_n := ba_n - ab_n$, $n = 0, 1, 2, \dots$. Then

$$d_{n+3} = Ad_{n+2} + Bd_{n+1} + Cd_n$$

for $n = 0, 1, 2, \dots$. Note that $d_n = 0$ for seven distinct non-negative indices n. We shall prove that $d_n = 0$ for each $n \geqslant 0$. By Theorem 7.14, it suffices to show that the quotient of two distinct conjugates of β is not a root of unity. Indeed, by Lemma 7.9, the conjugates β' and $\beta'' = \overline{\beta'}$ are of modulus $|\beta'| \neq \beta$. So only β'/β'' can be a root of unity. But then $\beta'^m = \beta''^m$ for some $m \in \mathbb{N}$. Mapping β' to β, we get $\beta^m = \sigma(\beta'')^m$, where $\sigma(\beta'') \in \{\beta', \beta''\}$, which is a contradiction, by

modulus consideration. This proves that $d_n = 0$ for each $n \geqslant 0$.

Therefore, $ba_n = ab_n$ for each $n \geqslant 0$. Obviously, by (7.19), $a \neq 0$ or $b \neq 0$, since otherwise $a_n = b_n = 0$ for each $n \geqslant 0$, by Theorem 7.14. Assume, without loss of generality, that $a \neq 0$. By (7.18), $\beta^3 = A\beta^2 + B\beta + C$. So (7.19) yields

$$\nu\beta^{n+1} = a_n\beta^3 + b_n\beta^2 + c_n\beta = (b_n + Aa_n)\beta^2 + (c_n + Ba_n)\beta + Ca_n. \tag{7.20}$$

It follows that $a_{n+1} = (b_n + Aa_n) = (b/a + A)a_n$. Hence $a_n = (b/a + A)^n a_0$ for each integer $n \geqslant 0$. Since $a_0 \neq 0$, two terms of the sequence a_n, $n = 0, 1, 2, \ldots$, are equal only if $b/a + A = \pm 1$. Then $a_n = a_{n+2}$ for each $n \geqslant 0$. This implies $c_n = c_{n+2}$ for each $n \geqslant 1$, since $c_{n+1} = Ca_n$, by (7.20). In the same way, $b_{n+2} = b_n$, because $b_{n+1} = c_n + Ba_n$, by (7.20). In view of (7.19), we obtain $\nu\beta^{n+2} = \nu\beta^n$. This leads to $\beta^2 = 1$ and gives the desired contradiction. □

Proof of Theorem 7.8. Suppose first that $s \neq 3$. We will show that α defined in (7.8) is not of the form $\beta^{1/d}$ with rational or cubic β. Then, applying Theorem 7.6 and (7.9), we will immediately obtain $M_{\xi,\alpha}(t) = 3$.

Assume that α^d is a rational or a cubic number for some $d \in \mathbb{N}$. Since the minimal polynomial (7.8) of α is irreducible, α is of degree s. Since $s \geqslant 2$, every positive integral power of α is of degree s too, because α is an algebraic integer having a unique conjugate outside the unit circle α itself, hence no two powers of distinct conjugates of α can be equal. This proves the theorem for $s \neq 3$.

For $s = 3$, we select in (7.8) $l = 2$, so that $\alpha = 1.839286\ldots$ is the root of $x^3 - x^2 - x - 1 = 0$. This time, we take

$$\xi := \frac{1}{\alpha^4 - \alpha} = \frac{3}{2} + 2\alpha - \frac{3}{2}\alpha^2.$$

By (7.9),

$$\{\xi\alpha\} = \{\xi\alpha^4\} = \{\xi\alpha^5\} = \xi\alpha = \frac{1}{\alpha^3 - 1} = \frac{1}{2}\alpha^2 - \frac{3}{2}.$$

Hence $M_{\xi,\alpha}(t_0) \geqslant 3$ for $t_0 := \alpha^2/2 - 3/2 = 0.191487\ldots$.

We will show that

$$M_{\xi,\alpha}(t) \leqslant 3 \tag{7.21}$$

for each $t \in [0, 1)$. In particular, combining (7.21) with the reverse inequality for $t = t_0$, we find that $M_{\xi,\alpha}(t_0) = 3$.

Indeed, for α and ξ as above, let us write $\xi\alpha^n = a_n\alpha^2 + b_n\alpha + c_n$ with rational a_n, b_n, c_n. Then the sequences a_n, b_n, c_n, where $n = 1, 2, 3, \ldots$, satisfy the same linear recurrence

$$x_{n+3} = x_{n+2} + x_{n+1} + x_n$$

for $n = 1, 2, 3, \ldots$. The first terms of those recurrence sequences are given in the following table:

n	a_n	b_n	c_n
1	1/2	0	−3/2
2	1/2	−1	1/2
3	−1/2	1	1/2
4	1/2	0	−1/2
5	1/2	0	1/2
6	1/2	1	1/2
7	3/2	1	1/2

It is easy to see that $a_{n+1} > \max(a_1, \ldots, a_n)$ for each $n \geqslant 6$, because the numbers a_n are positive for $n \geqslant 4$. So the equality $a_u = a_v$ can only hold for integers $u, v \in \mathbb{N}$ in the range $1 \leqslant u, v \leqslant 6$. As above, the sequence $\{\xi\alpha^n\} = \{a_n\alpha^2 + b_n\alpha + c_n\}$, $n = 1, 2, 3, \ldots$, may take equal values at $n = u$ and $n = v$ only if $(a_u, b_u) = (a_v, b_v)$. (In fact, they are equal if, in addition, $c_v - c_u \in \mathbb{Z}$.) Note that there are only three equal vectors among (a_n, b_n), $n = 1, \ldots, 6$, namely, $(a_1, b_1) = (a_4, b_4) = (a_5, b_5) = (1/2, 0)$. So at most 3 values of the sequence $\{\xi\alpha^n\}$, $n = 1, 2, 3, \ldots$, can be equal. This proves inequality (7.21). □

Proof of Theorem 7.3. Write $r^{1/d}$ in the form r_1^{1/d_1} with $r_1 \in \mathbb{Q}$ and the smallest possible $d_1 \in \mathbb{N}$. By abuse of notation, we shall keep the same notation r, d for r_1, d_1. Write $r = a/b$, where $a > b > 1$ are integers satisfying $\gcd(a, b) = 1$. Take $u := b^{s-1}$. We claim that the geometric progression

$$\mathcal{G}(u, q) = b^{s-1}, b^{s-1}(a/b)^{1/d}, b^{s-1}(a/b)^{2/d}, b^{s-1}(a/b)^{3/d}, \ldots \tag{7.22}$$

contains exactly s positive integers.

Indeed, the number $(a/b)^{n/d}$ is rational if and only if $n = 0, d, 2d, 3d, \ldots$. Hence $b^{s-1}(a/b)^{n/d} \in \mathbb{N}$ implies $n = md$, where $m \geqslant 0$ is an integer. It is clear that $b^{s-1}(a/b)^{md/d} = b^{s-1-m}a^m$ is an integer for $m = 0, \ldots, s-1$. Thus the sequence $b^{s-1-m}a^m$, $m = 0, 1, 2, \ldots$, (and so $\mathcal{G}$ defined in (7.22)) contains exactly s positive integers. □

Proof of Theorem 7.4. By Theorem 7.6 and (7.5), we have $|\mathcal{G} \cap \mathcal{A}| \leqslant 3$. On the other hand, Theorem 7.8 combined with (7.5) implies the existence of $\mathcal{A}$ with parameters given in (7.4) for which $|\mathcal{G} \cap \mathcal{A}| = 3$.

Finally, it is easy to see that Theorem 7.5 follows from the last statement of Theorem 7.6 combined with (7.5).

Chapter 8

Reducibility of quadrinomials

8.1 Statement of the problem

Let $P(x)$ be a polynomial with integer coefficients. The polynomial $P(x)$ is called *primitive*, if one cannot write $P(x) = P_1(x^l)$ for some polynomial $P_1 \in \mathbb{Z}[x]$ and integer $l > 1$.

Through the Chapter, *reducibility* shall always mean reducibility in $\mathbb{Z}[x]$.

The following problem was posed by Walsh [155] at the West Coast Number Theory Conference in 2007 .

Problem 8.1. *Let $i > j > k$ be positive integers. Does there exist an irreducible polynomial $P(x) = x^i + x^j + x^k + 4$ of degree $\deg P > 17$, such that for some integer $l > 1$, the polynomial $P(x^l)$ factors in $\mathbb{Z}[x]$?*

In addition, Walsh asked for the examples of reducible primitive quadrinomials of the form $x^i + x^j + x^k + n$ with integer constant coefficient $n > 4$, which have no linear or quadratic factors. He gave one such example $x^7 + x^5 + x^3 + 8 = (x^3 - x^2 - x + 2)(x^4 + x^3 + 3x^2 + 2x + 4)$.

The choice of the constant coefficient 4 in the polynomial $x^i + x^j + x^k + 4$ is not accidental. A similar example is the binomial $x^2 + 4$ which factors after the change of variable x to x^2. The polynomials $x^{4m} + 4b^4$ are the exceptional case in the theorem of Capelli [185] on the reducibility of binomials. In the trinomial case, all reducible polynomials of the form $x^i \pm x^j \pm 4$ were completely determined by Jonassen [109]. By the theorem given in his paper [109], there are no irreducible trinomials $P(x) = x^i \pm x^j \pm 4$, such that for some positive integer l, the polynomial $P(x^l)$ is reducible. In contrast, there exist quadrinomials $P(x) = x^i + x^j + x^k + 4$ which have this property. In Section 8.2, we shall give a complete description of such quadrinomials.

The questions on the reducibility of trinomials and quadrinomials have received a lot of interest. The reducible trinomials and quadrinomials with all non-zero

coefficients equal to 1 or -1 were investigated by Selmer [195] and Ljunggren [137]. The missing cases in Ljunggren's work were settled by Mills [147]. Many important results and generalizations were established by Schinzel in the long series of papers starting with [181], [182]. In 1972 Fried and Schinzel [94] proved a deep result on the reducibility of quadrinomials. Theorems 2 and 3 in [94] state that for the fixed integers a, b, c, d, any reducible quadrinomial $P(x) = ax^i+bx^j+cx^k+d$ either factors into the product of certain polynomials of standard shape, or such polynomial has the form $P(x) = P_1(x^l)$, $l \in \mathbb{Z}$, $l > 0$, were $P_1 \in \mathbb{Z}[x]$ is primitive reducible quadrinomial of degree less than or equal to the effectively computable constant $C(a, b, c, d)$. Unfortunately, this constant is too large for almost any practical applications: in our case, $C(1, 1, 1, 4) > 2^{8045222}$. In the present Chapter, we shall use Ljunggren's method [137] to determine all such exceptional quadrinomials $x^i + x^j + x^k + 4$ which appear in the question of Walsh. See [87], [92] for a good exposition on the Ljunggren's method. More recent results on reducibility of trinomials can be found in [89]. For efficient factoring algorithms, we refer to [88].

8.2 Main results

The following theorem answers the first question of Walsh. We note that in this Chapter we do not use the same terminology as in [155].

Theorem 8.2. *The only primitive irreducible polynomial $P \in \mathbb{Z}[x]$ of the form $P(x) = x^i + x^j + x^k + 4$, $i > j > k > 0$, such that the polynomial $P(x^l)$ for some positive integer l factors in $\mathbb{Z}[x]$, is the polynomial $P(x) = x^4 + x^3 + x^2 + 4$. More precisely, for $l = 2$,*

$$P(x^2) = x^8 + x^6 + x^4 + 4 = (x^4 - x^3 + x^2 - 2x + 2)(x^4 + x^3 + x^2 + 2x + 2).$$

Indeed, if the polynomial $P(x) = x^i + x^j + x^k + 4$ has the property asked in Problem 007 : 14, then $P(x) = P_1(x^m)$ for some primitive polynomial $P_1(x)$ and some positive integer m. By Theorem 8.2, $P_1(x) = x^4+x^3+x^2+4$. In Lemma 8.3 bellow, we prove that $P(x^l)$ is reducible only for even values $l = 2g$; in this case the polynomial $P(x^l)$ splits into two irreducible factors $x^{4g} - x^{3g} + x^{2g} - 2x^g + 2$ and $x^{4g} + x^{3g} + x^{2g} + 2x^g + 2$.

Lemma 8.3. *Let $P(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1x + a_0$ be a monic integer polynomial, such that $|a_0| = p^2 > |a_1| + |a_2| + \cdots + |a_{d-1}| + 1$, where p is a prime. Let $l > 1$ be a positive integer, such that $P(x^l)$ is reducible in $\mathbb{Z}[x]$ but $P(x^m)$ is irreducible for any positive integer $m < l$ which divides l. Then l is even and*

$P(x^l) = \pm Q(x)Q(-x)$, $Q \in \mathbb{Z}[x]$. *Moreover, for any integer* $r \geqslant 1$, *both factors* $Q(\pm x^r) \in \mathbb{Z}[x]$ *are irreducible.*

In Lemma 8.4 we shall determine all the possible forms the product polynomial PP^* can take. This will be used in the proof of Theorem 8.2.

Lemma 8.4. *Let* P *be a quadrinomial* $P(x) = x^i + x^j + x^k + 4$ *with the integer exponents* $i > j > k > 0$. *Then the polynomial* PP^* *takes one of the following forms:*

1) $4x^{2i} + x^{2i-k} + x^{2i-j} + x^{i+j-k} + 4x^{i+j} + 4x^{i+k} + 19x^i + 4x^{i-k} + 4x^{i-j} + x^{i-j+k} + x^j + x^k + 4$,

$$\text{if} \quad i \neq 2j, \quad i \neq 2k, \quad j \neq 2k, \quad i + k \neq 2j, \quad i \neq j + k;$$

2) $4x^{2i} + x^{2i-k} + 5x^{2i-j} + x^{i+j-k} + 4x^{i+k} + 19x^i + 4x^{i-k} + x^{i-j+k} + 5x^j + x^k + 4$,

$$\text{if} \quad i = 2j, \quad i \neq 2k, \quad j \neq 2k, \quad i + k \neq 2j, \quad i \neq j + k;$$

3) $4x^{2i} + 5x^{2i-k} + x^{2i-j} + x^{i+j-k} + 4x^{i+j} + 19x^i + 4x^{i-j} + x^{i-j+k} + x^j + 5x^k + 4$,

$$\text{if} \quad i \neq 2j, \quad i = 2k, \quad j \neq 2k, \quad i + k \neq 2j, \quad i \neq j + k;$$

4) $4x^{2i} + x^{2i-k} + x^{2i-j} + 5x^{i+j-k} + 4x^{i+j} + 19x^i + 4x^{i-j} + 5x^{i-j+k} + x^j + x^k + 4$,

$$\text{if} \quad i \neq 2j, \quad i \neq 2k, \quad j = 2k, \quad i + k \neq 2j, \quad i \neq j + k;$$

5) $4x^{2i} + x^{2i-k} + 2x^{2i-j} + 4x^{i+j} + 4x^{i+k} + 19x^i + 4x^{i-k} + 4x^{i-j} + 2x^j + x^k + 4$,

$$\text{if} \quad i \neq 2j, \quad i \neq 2k, \quad j \neq 2k, \quad i + k = 2j, \quad i \neq j + k;$$

6) $4x^{2i} + 5x^{2i-k} + 5x^{2i-j} + x^{i+j-k} + 19x^i + x^{i-j+k} + 5x^j + 5x^k + 4$,

$$\text{if} \quad i \neq 2j, \quad i \neq 2k, \quad j \neq 2k, \quad i + k \neq 2j, \quad i = j + k;$$

7) $4x^{2i} + x^{2i-k} + 5x^{2i-j} + 5x^{i+j-k} + 19x^i + 5x^{i-j+k} + 5x^j + x^k + 4$,

$$\text{if} \quad i = 2j, \quad i \neq 2k, \quad j = 2k, \quad i + k \neq 2j, \quad i \neq j + k;$$

8) $4x^{2i} + 5x^{2i-k} + 2x^{2i-j} + 4x^{i+j} + 19x^i + 4x^{i-j} + 2x^j + 5x^k + 4$,

$$\text{if} \quad i \neq 2j, \quad i = 2k, \quad j \neq 2k, \quad i + k = 2j, \quad i \neq j + k;$$

9) $4x^{2i} + 5x^{2i-k} + 6x^{2i-j} + 19x^i + 6x^j + 5x^k + 4$,

$$\text{if} \quad i \neq 2j, \quad i \neq 2k, \quad j = 2k, \quad i + k = 2j, \quad i = j + k;$$

8.3 Computations

In order to answer the second part of Problem 007 : 14, we used the computer to search for the examples of reducible polynomials of the form $P(x) = x^i + x^j + x^k + n$. Since the polynomial $P(x)$ has no roots of modulus less than or equal

to 1 if $n \geqslant 4$, the polynomial $P(x)$ is irreducible provided the coefficient n is equal to the prime integer $p \geqslant 5$. With MAPLE computer algebra package we factored all primitive quadrinomials $P(x)$ with composite constant coefficient n and exponents i, j, k in the range $5 < n \leqslant 120$, $i - j \leqslant 20$, $j - k \leqslant 20$, $k \leqslant 20$. In addition, we factored all primitive quadrinomials $P(x)$ of this form satisfying inequalities $120 < n \leqslant 1000$, $i - j \leqslant 15$, $j - k \leqslant 15$, $k \leqslant 15$. We also searched for the irreducible polynomials $P(x)$, such that $P(x^l)$ is reducible for some integer l in the range $5 < n \leqslant 120$, $(i - j)l \leqslant 12$, $(j - k)l \leqslant 12$, $kl \leqslant 12$. In all the cases reducible polynomials $P(x)$ had a factor of the form $Q(x^l)$, where $Q(x)$ was a linear polynomial or a quadratic polynomial. The example $x^7 + x^5 + x^3 + 8$ of Walsh was the only notable exception. However, it does not seem easy to prove this.

All found examples of reducible quadrinomials $P(x)$ had two or three irreducible factors. In the recent exchanges of emails, A. Schinzel sent a short remark that any irreducible polynomial dividing the quadrinomial $P(x)$ has constant coefficient greater than 1, hence the number of irreducible factors cannot exceed $\Omega(n)$, the total number of prime factors of n. The sharpness of this estimate may be shown by the example

$$x^{12} + x^8 + x^4 + 52 = (x^2 - 2x + 2)(x^2 + 2x + 2)(x^8 - 3x^4 + 13).$$

Since $\Omega(n) \leqslant \log n / \log 2$, the number $\log n / \log 2$ is the best known bound for the total number of prime factors of the quadrinomials $P(x)$ in question.

Finally, we note that there exist the examples of reducible quadrinomials $P(x)$ with no linear or quadratic factors and negative coefficient $n < -5$, namely, the polynomials

$$\begin{aligned}
x^6 + x^4 + x^2 - 16 &= (x^3 - 3x^2 + 5x - 4)(x^3 + 3x^2 + 5x + 4), \\
x^{12} + x^8 + x^4 - 16 &= (x^3 - x^2 - x + 2)(x^3 + x^2 - x - 2)(x^6 + 3x^4 + 5x^2 + 4), \\
x^7 + x^3 + x^2 - 98 &= (x^3 - x^2 + 2x - 7)(x^4 + x^3 - x^2 + 4x + 14),
\end{aligned}$$

and

$$x^{17}+x^{14}+x^8-16 = (x^5+x^3-x^2-2)(x^{12}-x^{10}+2x^9+x^8-x^7+x^6+2x^4+4x^3-4x^2+8).$$

8.4 Proofs

Proof of Lemma 8.3. Let $P(x^l) = Q(x)R(x)$ for some integer $l > 0$ and polynomials $Q, R \in \mathbb{Z}[x]$. The inequality $|a_0| > |a_1| + |a_2| + \cdots + |a_{d_1}| + 1$ implies

that P, Q and R have no roots of modulus $|z| \leqslant 1$. Thus the constant terms of Q and R are equal to $\pm p$ and they are irreducible in $\mathbb{Z}[x]$. Otherwise one of the polynomials Q or R would be divisible by the monic non-constant polynomial $S \in \mathbb{Z}[x]$ with the constant term $S(0) = \pm 1$. This is impossible, since such a polynomial S has at least one root of modulus less or equal to 1. The same argument also implies the irreducibility of polynomials $Q(x^r)$ and $R(x^r)$ which divide $P(x^{rl})$.

Now, assume that the exponent l has the property that for any positive integer $m < l$ which divides l, the polynomial $P(x^m)$ is irreducible. If $m = 1$, this means that $P(x)$ is irreducible. Let α be the root of the irreducible factor $Q(x)$. Since $P(x^l) = Q(x)R(x)$, the power of this root $\beta = \alpha^l$ is the root of the irreducible polynomial P. Let $K = \mathbb{Q}(\beta), L = \mathbb{Q}(\alpha)$, $K \subset L$. Let g be the degree $[L : K]$. The absolute norm of an algebraic integer β over $\mathbb{Q}$ $N_{L/\mathbb{Q}}(\beta) = N_{L/\mathbb{Q}}(\alpha^l) = N_{L/\mathbb{Q}}(\alpha)^l = \pm p^l$. In the other hand, by the relative norm property, $N_{L/\mathbb{Q}}(\beta) = N_{K/\mathbb{Q}}(\beta)^{[L:K]} = \pm p^{2g}$. Thus $l = 2g$.

Since l is even, the number $-\alpha$ is the root of $P(x^l)$. The irreducible polynomial $Q(-x)$ divides $P(x^l)$. Then $R(x) = \pm Q(-x)$. Indeed, otherwise $Q(-x) = -Q(x)$ or $Q(-x) = Q(x)$. The first case is impossible: the identity $Q(-x) = -Q(x)$ with $x = 0$ implies $Q(0) = P(0) = 0$, contradicting the inequality $|a_0| > 1$. The second identity $Q(-x) = Q(x)$ implies $Q(x) = T(x^2)$ for the polynomial $T \in \mathbb{Z}[x]$, thus $R(x) = P(x^l)/Q(x) = P(x^{2g})/T(x^2) = S(x^2), S \in \mathbb{Z}[x]$. This leads to the expression $P(x^{2g}) = T(x^2)S(x^2)$, hence $P(x^g) = T(x)S(x)$. This implies that $P(x^g)$ is reducible for the integer g wich is a proper divisor of l, contradicting the condition of Lemma 8.3. □

Proof of Lemma 8.4. Assume that integers i, j, k satisfy the inequality $i > j > k > 0$. The reciprocal of the polynomial $P(x) = x^i + x^j + x^k + 4$ is $P^*(x) = 4x^i + x^{i-k} + x^{i-j} + 1$. The product $F = PP^*$ takes the form

$$4x^{2i}+x^{2i-k}+x^{2i-j}+4x^{i+j}+x^{i+j-k}+4x^{i+k}+19x^i+4x^{i-k}+x^{i-j+k}+4x^{i-j}+x^j+x^k+4.$$

Let

$$v_1 = 2i, \quad v_2 = 2i-k, \quad v_3 = 2i-j, \quad v_4 = i+j, \quad v_5 = i+j-k,$$

$$v_6 = i+k, \quad v_7 = i, \quad v_8 = i-k, \quad v_9 = i-j+k, \quad v_{10} = i-j,$$

$$v_{11} = j, \quad v_{12} = k, \quad v_{13} = 0.$$

The multi-set $V = \{v_r, r = 0 \dots 13\}$ contains all possible exponents which appear in the polynomial F. If none of them are equal, F takes the form in *Case 1.* We

shall classify all other cases where some exponents v_r and $v_s, r \neq s$ are equal, and the terms of F with equal exponents add together. Set

$$e_1 = \{i = 2j\}, \quad e_2 = \{i = 2k\}, \quad e_3 = \{j = 2k\},$$

$$e_4 = \{i + k = 2j\}, \quad e_5 = \{i = j + k\}.$$

The elements of the set $E = \{e_r, r = 1 \dots 5\}$ denote the linear relations among the integers i, j, k. Note that v_1 and v_{13} are the largest and smallest integers in V. Since $i > j > k > 0$, all the exponents $v_1, v_2, v_3, v_4, v_5, v_6$ are strictly greater than the exponent of the middle term $v_7 = i$, exponents $v_8, v_9, v_{10}, v_{11}, v_{12}, v_{13}$ are strictly less than $v_7 = i$. F is self-reciprocal, since $F = PP^* = F^*$. Hence $v_s = 2i - v_{13-s+1}, s = 8 \dots 13$. Thus it suffices to check all possible cases when some of the integers v_2, v_3, v_4, v_5, v_6 are equal. Observe that

$$v_2 > v_3, \quad v_2 > v_5, \quad v_4 > v_5, \quad v_4 > v_6.$$

Hence, all possible pairs of equal exponents $v_s, s = 2 \dots 6$ are: $v_2 = v_4$ (this is equivalent to the linear relation e_5), $v_2 = v_6$ (equivalent to e_2), $v_3 = v_4$ (e_1), $v_3 = v_5$ (e_4), $v_3 = v_6$ (e_5), $v_5 = v_6$ (e_3). The remaining pairs of equal exponents $v_s, s = 8 \dots 13$ are determined uniquely by the symmetry $v_s = 2i - v_{13-s+1}$.

The forms of the polynomial F where the integer exponents i, j, k satisfy precisely one linear relation in the set E are listed in Cases (2)-(6). In order to check Cases (2)-(6), use the expression in Case (1) and add terms with equal powers x^{v_r} and x^{v_s} together if $v_r = v_s$. The next step is to determine all possible forms of F where the integers i, j, k satisfy two linear relations $\{e_s, e_t\} \subset E$. Observe that no one pair of the linear relations $\{e_1, e_2\}$, $\{e_1, e_4\}$, $\{e_1, e_5\}$, $\{e_2, e_3\}$, $\{e_2, e_5\}$ is possible if $i > j > k > 0$. The possible pairs are $\{e_1, e_3\}$, $\{e_2, e_4\}$, $\{e_3, e_4\}$, $\{e_3, e_5\}$, $\{e_4, e_5\}$. Consider the pair $\{e_1, e_3\}$ as a system of two linear equations in three integer variables i, j, k. All positive integer solutions of this system are vectors $(i, j, k) = (4u, 2u, u), u \in \mathbb{Z}, u > 0$. In this case, the form of the polynomial F with equal exponents $v_3 = v_4$, $v_5 = v_6$, $v_8 = v_9$, $v_{10} = v_{11}$ is described in Case (7) of Lemma 8.4. Similarly, the integer solutions to the equations $\{e_2, e_4\}$ are $(i, j, k) = (4u, 3u, 2u), u \in \mathbb{Z}, u > 0$ and F takes the form given in *Case 8*. Observe that all the pairs of equations from the system $\{e_3, e_4, e_5\}$ are equivalent and have the solution $(i, j, k) = (3u, 2u, u), u \in \mathbb{Z}, u > 0$. Hence, any two of the three relations $\{e_3, e_4, e_5\}$ imply the third one. This situation is depicted in *Case 9*. It remains to show that there are no other cases where three or more linear relations $e_s \in E$ hold. Indeed, in such case three different pairs of linear relations, other than all 3 possible pairs from the set $\{e_3, e_4, e_5\}$ must be satisfied. There would be

at least one of pairs $\{e_1, e_3\}$, $\{e_2, e_4\}$ and one pair from the set $\{e_3, e_4, e_5\}$. This is impossible, since all the intersections of the sets of integer triples (i, j, k) which satisfy such linear relations

$$\{(4u, 2u, u), u \in \mathbb{Z}, u > 0\}, \quad \{(4u, 3u, 2u), u \in \mathbb{Z}, u > 0\},$$

$$\{(3u, 2u, u), u \in \mathbb{Z}, u > 0\}$$

are empty. □

Sketch of the proof of Theorem 8.2. Before proceeding to prove Theorem 8.2, we give the sketch of the proof. First, we show that any primitive quadrinomial $P(x) = x^i + x^j + x^k + 4$ in question with exponents (i, j, k) which satisfy a certain linear relation is precisely the quadrinomial $x^4 + x^3 + x^2 + 4$. Secondly, we consider the polynomial $G(x) = Q(-x)Q^*(x)$, where Q is the polynomial from the factorization $P(x^l) = \pm Q(x)Q(-x)$. This factorization is a consequence of Lemma 8.3. We determine the form of the polynomial G using reduction modulo 2 and the identity $||P|| = ||G||$ for the Euclidean norms of P and G. Following [137], we refer to the equality $||P|| = ||G||$ as the *identity of Ljunggren*. Thirdly, we use the expression $GG^*(x) = PP^*(x^l)$ and compare $PP^*(x^l)$ from Lemma 8.4 to $GG^*(x)$. We establish that the case $x^4 + x^3 + x^2 + 4$ is the only possible.

Proof of Theorem 8.2. First suppose that the exponents i, j, k of $P(x) = x^i + x^j + x^k + 4$ satisfy linear relations $i = 2k$ and $i + k = 2j$. Then $(i, j, k) = (4u, 3u, 2u)$ for some positive integer u so $P(x) = x^{4u} + x^{3u} + x^{2u} + 4$, which is primitive for $u = 1$. A simple computation shows that the polynomial $P_1(x) = x^4 + x^3 + x^2 + 4$ is irreducible, and $P_1(x^2) = (x^4 - x^3 + x^2 - 2x + 2)(x^4 + x^3 + x^2 + 2x + 2)$. For even integers $l = 2g > 0$ the polynomial $P_1(x^l)$ splits in $\mathbb{Z}[x]$ into $P_1(x^l) = (x^{4g} - x^{3g} + x^{2g} - 2x^g + 2)(x^{4g} + x^{3g} + x^{2g} + 2x^g + 2)$. By Lemma 8.3, both factors are irreducible. By Lemma 8.3, $P_1(x^l)$ is irreducible for odd exponents l. Below we shall show that this case is the only possible. Note that the linear relations $i = 2k, i + k = 2j$ appear in Case (8) of Lemma 8.4. Hence we have to prove that every quadrinomial $P(x)$ in the question of Walsh satisfies $P(x^l)P^*(x^l) = F(x^l)$, where F is a polynomial in Case (8) of Lemma (8.4).

Let $P(x) = x^i + x^j + x^k + 4$ be an irreducible polynomial and $l > 0$ be an integer such that $P(x^l)$ splits in $\mathbb{Z}[x]$, while $P(x^m)$ is irreducible for any integer m, $1 \leqslant m < l$ dividing the exponent l. By Lemma 8.3, $l = 2g, g \in \mathbb{Z}$, and $P(x^l) = \pm Q(x)Q(-x)$. Without loss of generality, we may assume that $Q(x)$ is monic. Then $P(x^l) = (-1)^{ig} Q(x)Q(-x)$. The polynomials $Q(x)$ and $Q(-x)$ have equal constant terms ± 2, hence $P(x) = Q(x)Q(-x)$. This implies that the degree ig is even. Also, $Q(0) = 2$. Otherwise $Q(x)$ has a positive real root which is

impossible, since $P(x) > 0$ if $x > 0$. Consider the reduction of $P(x^{2g})$ modulo 2:

$$P(x^g)^2 \equiv P(x^{2g}) = Q(x)Q(-x) \equiv Q(x)^2 \pmod 2.$$

Hence $Q(x) \equiv P(x^g) \equiv x^{ig} + x^{jg} + x^{kg} \pmod 2$. Let G be the product $G(x) = Q(-x)Q^*(x)$ of degree $2ig$. Note that the polynomial G satisfies the identity $x^{2ig}G(1/x) = (-1)^{ig}G(-x)$. Since ig is even, $G^*(x) = G(-x)$. Hence the integer coefficients of $G(x) = \sum_{s=0}^{2ig} b_s x^s$ which are symmetric with the respect of middle term are equal in modulus, more precisely,

$$b_s = (-1)^s b_{2ig-s}, 0 \leqslant s \leqslant 2ig. \tag{8.1}$$

Reduce $G(x)$ modulo 2:

$$G(x) = Q^*(x)Q(-x) \equiv Q^*(x)Q(x) \equiv (x^{(i-k)g} + x^{(i-j)g} + 1)(x^{ig} + x^{jg} + x^{kg}) \equiv$$

$$\equiv x^{(2i-k)g} + x^{(2i-j)g} + x^{(i+j-k)g} + x^{ig} + x^{(i-j+k)g} + x^{jg} + x^{kg} \pmod 2. \tag{8.2}$$

Note that in (8.2) the commuting of the operation $*$ and reduction $\pmod 2$ is essentially used, which makes sense since $f^*(x) \pmod 2 = (f(x) \pmod 2)^*$ if and only if the leading coefficient of $f \in \mathbb{Z}[x]$ is odd.

Since Q is monic, $Q(0) = 2$, the leading and constant coefficients of G are equal to 2. Since $i > j > k > 0$, the exponents in G modulo 2 satisfy the inequalities

$$(2i-k)g > (2i-j)g \geqslant (i+j-k)g > ig > (i-j+k)g \geqslant jg > kg,$$

provided $i + k \geqslant 2j$ or

$$(2i-k)g > (i+j-k)g \geqslant (2i-j)g > ig > jg \geqslant (i-j+k)g > kg,$$

provided $i + k \leqslant 2j$. Hence the polynomial $G(x)$ in (8.2) has 7 odd coefficients if $i + k \neq 2j$. If $i + k = 2j$, $G(x)$ modulo 2 takes the form

$$G(x) \equiv x^{(2i-k)g} + x^{ig} + x^{kg} \pmod 2, \tag{8.3}$$

with 3 odd coefficients. Observe that $P(x^l)P^*(x^l) = G(x)G^*(x)$. The equality holds since $(Q^*)^* = Q$ which is true since $Q(0) \neq 0$. By the identity of Ljunggren, $||G||^2 = ||P||^2 = 1^2 + 1^2 + 1^2 + 4^2 = 19$. The leading and constant coefficients of G are equal to 2, thus the sum of squares of coefficients $b_s, 1 \leqslant s \leqslant 2ig - 1$ is

equal to 11. In addition, there must be precisely 3 or 7 odd coefficients by (8.2) and (8.3). All such possible sums of squares are

$$11 = 3^2+1^2+1^2 = 2^2+2^2+1^2+1^2+1^2 = 2^2+1^2+1^2+1^2+1^2+1^2+1^2+1^2 \quad (8.4)$$

The absolute values of the coefficients of G are symmetric with respect to the middle term $b_{ig} \equiv 1 \pmod 2$. Thus an integer which appears in the sum of squares above the odd number of times must be the absolute value of the middle coefficient b_{ig}. Hence the summands in the third sum in (8.4) cannot be the squares of coefficients of the polynomial G. This implies that G has 3 odd coefficients and the exponents i, j, k satisfy the relation $i + k = 2j$. Using the identities (8.1), (8.3), and (8.4) we deduce that G takes one of the forms:

$$G(x) = 2x^{2ig} + \varepsilon x^{(2i-k)g} + 3\delta x^{ig} + (-1)^{kg}\varepsilon x^{kg} + 2, \quad (8.5)$$

$$G(x) = 2x^{2ig}+\varepsilon_1 x^{(2i-k)g}+2\varepsilon_2 x^{t+ig}+\delta x^{ig}+(-1)^t 2\varepsilon_2 x^{ig-t}+(-1)^{kg}\varepsilon_1 x^{kg}+2, \quad (8.6)$$

the coefficients $\varepsilon, \varepsilon_1, \varepsilon_2, \delta$ are all equal to -1 or 1 and t is an integer $0 < t < ig$. Also, note that the terms x^{t+ig}, x^{ig-t} do not coincide with any other term of the polynomial G in (8.6) by (8.4). We shall determine the coefficients $\varepsilon, \varepsilon_1, \varepsilon_2, \delta$. Observe that $G(1) = Q(1)Q(-1) = P(1) = 7$. In (8.5), this is possible if and only if $\delta = 1$ and $\varepsilon + (-1)^{kg}\varepsilon = 0$, hence the exponent kg is odd in (8.5). Moreover, we may assume that $\varepsilon = 1$. Otherwise, replace x by $-x$. Thus G in (8.5) takes the form

$$i) \qquad 2x^{2ig} + x^{(2i-k)g} + 3x^{ig} - x^{kg} + 2, 2 \nmid kg.$$

Consider G in (8.6). There are three cases where $G(x)$ takes the value $G(1) = 7$:

a) $2 \mid kg$, $2 \nmid t$, $\varepsilon_1 = 1$, $\delta = 1$. Assume that $\varepsilon_2 = 1$, otherwise change x to $-x$;

b) $2 \mid t$, $2 \nmid kg$, $\varepsilon_2 = 1$, $\delta = -1$. Assume that $\varepsilon_1 = 1$, otherwise change x to $-x$;

c) $2 \mid kg, 2|t$, $\varepsilon_1 = -1$, $\varepsilon_2 = 1$, $\delta = 1$.

The polynomial G in Cases *(a),(b),(c)* takes the forms *(ii)*, *(iii)*, *(iv)* below, respectively.

$$ii) \qquad 2x^{2ig} + x^{(2i-k)g} + 2x^{t+ig} + x^{ig} - 2x^{ig-t} + x^{kg} + 2,$$

$$iii) \qquad 2x^{2ig} + x^{(2i-k)g} + 2x^{t+ig} - x^{ig} + 2x^{ig-t} - x^{kg} + 2,$$

$$iv) \qquad 2x^{2ig} - x^{(2i-k)g} + 2x^{t+ig} + x^{ig} + 2x^{ig-t} - x^{kg} + 2.$$

In each case *(i), (ii), (iii), (iv)* we check if $G(x)G^*(x) = P(x^{2g})P^*(x^{2g}) = F(x^{2g})$, where $F(x)$ is one of the polynomials in Lemma 8.4. Since $i+k=2j$, it suffices to check Cases (5), (8), (9) in Lemma 8.4. First, assume that G takes the form *(i)*. Then

$$G(x)G^*(x) = 4x^{4ig} + 12x^{3ig} - x^{(4i-2k)g} + 19x^{2ig} - x^{2kg} + 12x^{ig} + 4$$

has 7 non-zero coefficients, hence it must coincide with $F(x^{2g})$ in Case (9) of Lemma 8.4. This is impossible, since the coefficients of F are different from the coefficients of GG^*. Next, assume that G takes the form (iii). Compute the product GG^* modulo 4:

$$G(x)G^*(x) \equiv -x^{(4i-2k)g} - x^{2ig} - x^{2kg} \pmod 4.$$

None of the polynomials F in Lemma 8.4 satisfy $F(x^{2g}) \equiv G(x)G^*(x) \pmod 4$. Thus the form *(iii)* is impossible.

Assume that G takes the form *(iv)*. The integer $2ig$ is the largest exponent in G. Let v be the second largest exponent in G. Clearly, $v = (2i-k)g$ or $v = t+ig$. Observe that $2ig + v > s + r$ if at least one inequality $r \leqslant 2ig, s \leqslant v$ is strict. Hence the second largest exponent in GG^* is $2ig+v$. Thus the first two terms of GG^* are $x^{4ig} - 4x^{(4i-k)g}$ if $(2i-k)g > t+ig$ or $x^{4ig} + 8x^{3ig+t}$ if $(2i-k)g < t+ig$. Such terms do not occur in any polynomial in Lemma 8.4. Hence we reject the form (iv).

This implies that G takes the form *(ii)*. The product GG^*

$$G(x)G^*(x) = 4x^{4ig} + 4x^{(4i-k)g} - 4x^{2t+2ig} + x^{(4i-2k)g} + 4x^{3ig} + 2x^{(3i-k)g} + 4x^{(2i+k)g} + 19x^{2ig} + 4x^{(2i-k)g} + 2x^{(i+k)g} + 4x^{ig} + x^{2kg} - 4x^{2ig-2t} + 4x^{kg} + 4.$$

Thus GG^* coincides with the polynomial $F(x^{2g})$ in Case (5), (8) or (9) of Lemma 8.4, since $i+k=2j$ in *(ii)*. Since $i > k > 0$, the integer $4i-k$ is strictly greater than $3i, 4i-2k, 3i-k, 2i+k$. If $(4i-k)g > 2ig+2t$, then the second leading term of GG^* is $4x^{(4i-k)g}$. This leads directly to Case (8) of the Lemma 8.4. Indeed, only polynomials $F(x)$ in Case (5) or Case (8) have terms with coefficients equal to 4 which are not leading nor constant terms. The term with the second highest exponent $4x^{(4i-k)g}$ in GG^* must coincide with the term $4x^{2(i+j)g}$ or $4x^{2(i+k)g}$ in $F(x^{2g})$ in Case (5). Since $j > k$, the exponent $2(i+j)g$ is greater than $2(i+k)g$. Thus $(4i-k)g = 2(i+j)g$, so $4i-k = 2i+2j$. Together with the identity $i+k=2j$ in Case (5) the linear relation $4i-k = 2i+2j$ implies $i = 2k$, which implies Case (8) of Lemma 8.4.

Hence we may assume that $(4i-k)g \leqslant 2ig+2t$. If the inequality is strict, then the second leading term of GG^* is $-4x^{2ig+2t}$. However, the polynomials in Lemma 8.4 have no negative terms. Hence $(4i-k)g = 2ig+2t$. Thus

$$\begin{aligned} GG^* = 4x^{4ig} + x^{(4i-2k)g} + 4x^{3ig} + 2x^{(3i-k)g} + 4x^{(2i+k)g} + \\ + 19x^{2ig} + 4x^{(2i-k)g} + 2x^{(i+k)g} + 4x^{ig} + x^{2kg} + 4. \end{aligned}$$

Let $F(x)$ be the polynomial in Case (5) of Lemma 8.4. Replace x by x^{2g} and use the identity $2j = i+k$. The resulting polynomial

$$\begin{aligned} F(x^{2g}) &= 4x^{4ig} + x^{(4i-2k)g} + 2x^{(4i-2j)g} + 4x^{(2i+2j)g} + 4x^{(2i+2k)g} + \\ &+ 19x^{2ig} + 4x^{(2i-2k)g} + 4x^{(2i-2j)g} + 2x^{2jg} + x^{2kg} + 4 \\ &= 4x^{4ig} + x^{(4i-2k)g} + 2x^{(3i-k)g} + 4x^{(3i+k)g} + 4x^{(2i+2k)g} + \\ &+ 19x^{2ig} + 4x^{(2i-2k)g} + 4x^{(i-k)g} + 2x^{(i+k)g} + x^{2kg} + 4. \end{aligned}$$

The difference $F(x^{2g}) - GG^* = 4x^{(3i+k)g} + 4x^{(2i+2k)g} + 4x^{(2i-2k)g} + 4x^{(i-k)g} - 4x^{3ig} - 4x^{(2i+k)g} - 4x^{(2i-k)g} - 4x^{ig} \neq 0$, since the exponent $(3i+k)g$ is larger than the other exponents in $F(x^{2g}) - GG^*$. Thus GG^* does not coincide with a polynomial given in Case (5) of Lemma 8.4.

Let $F(x)$ be the polynomial in Case (9) of Lemma 8.4. The equations $j = 2k$, $2j = i+k$, $i = j+k$ imply $(i,j,k) = (3u, 2u, u), u \in \mathbb{Z}, u > 0$. Hence

$$GG^* = 4x^{12ug} + x^{10ug} + 4x^{9ug} + 2x^{8ug} + 4x^{7g} + 19x^{6ug} + 4x^{5ug} + 2x^{4ug} + 4x^{3ug} + x^{2ug} + 4.$$

Also, $F(x^{2g}) = 4x^{12ug} + 5x^{10ug} + 6x^{8ug} + 19x^{6ug} + 6x^{4ug} + 5x^{2ug} + 4$ and $F(x^{2g}) \neq GG^*$, so Case (9) is impossible. Hence we conclude that Case (8) is the only possible. This completes the proof. □

Chapter 9

Height reduction and Number systems

9.1 Statement of the problem

Let α be an algebraic integer with conjugates $\alpha_1 = \alpha$, α_2, $\ldots$, α_d lying outside the unit circle (including α itself). Such numbers are called *expanding* algebraic numbers. We are interested in the *height reducing property* of α, that is

$$\mathbb{Z}[\alpha] = \mathcal{B}[\alpha]$$

for a certain finite set $\mathcal{B} \subset \mathbb{Z}$. We note that

Lemma 9.1. *If an algebraic integer α, $|\alpha| > 1$, has height reducing property, then α is expanding.*

Proof. Suppose α has height reducing property with a finite set $\mathcal{B} \subset \mathbb{Z}$. First assume it has a conjugate β with $|\beta| < 1$. Set $B = \max_{b\in\mathcal{B}} |b|$ and take an integer $K > \frac{B}{1-|\beta|}$. Then K has an expression $K = \sum_{i=0}^{n} b_i\alpha^i$ for some integer n. Taking conjugate, we have

$$K < \sum_{i=0}^{\infty} B|\beta|^i$$

which gives a contradiction. Therefore all the conjugates of α must be not less than one in modulus. Assume that there is a conjugate β with $|\beta| = 1$. Then β must be a complex number and $\beta\beta' = 1$ where β' is a complex conjugate of β. By taking conjugate map which sends β to α, we get a contradiction. □

Note that roots of unity (with all their conjugates on the unit circle) also have height reducing property with a set $\mathcal{B} = \{-1, 0, 1\}$.

When α is expanding, it is of interest whether it has height reducing property, and how small the set $\mathcal{B}$ we can we. Denote by $N(\alpha)$ the absolute norm of α over $\mathbb{Q}$, i.e., $N(\alpha) = \alpha_1 \cdot \alpha_2 \dots \alpha_d$. If we can choose $\mathcal{B} = \{0, 1, \dots, |N(\alpha)| - 1\}$, we say $(\alpha, \mathcal{B})$ forms a *canonical number system* (CNS for short). The question of finding all α which gives CNS is studied by many authors. The early studies are found in [111, 112, 99]. Readers may consult [7, 3] for recent developments to solve the problem in a general frame work called a *shift radix system*.

However not every expanding algebraic integer α generates CNS. Indeed, if there is a positive conjugate β of α, one sees that -1 can not be in $\mathcal{B}[\alpha]$ which is shown by taking conjugate.

For the rest of the Chapter let $\mathcal{B} = \{0, \pm 1, \dots, \pm (|N(\alpha)| - 1)\}$.

Kirat and Lau [114] introduced a slightly different height reducing property for expanding polynomials (all roots in $|z| > 1$, not necessarily irreducible) to consider the connectedness of a class of self-affine tiles. In our notation, they are interested in $N(\alpha) \in \mathcal{B}[\alpha]$ (see [115] for details).

In this Chapter we are mainly concerned with the following type of height reducing problem:

Problem 9.2. *Does the equality $\mathbb{Z}[\alpha] = \mathcal{B}[\alpha]$ hold for any expanding algebraic integer?*

In the study of self-affine tilings, Lagarias and Wang [126] answered this question in affirmative manner using the wavelet analysis by extending the result of [101]. To read this result out of their consecutive works, see Corollary 6.2 in [126] and Theorem 1.2 (ii) of [124]. However their proof is rather indirect and intricate, although the statement itself looks simple in nature. The first author [1] asked for a direct proof of $\mathbb{Z}[\alpha] = \mathcal{B}[\alpha]$ (see problem 12). In this Chapter we shall give several attempts to solve this question. For the moment, it is far from satisfactory but we hope our work gives a starting point for other trials. First we show

Theorem 9.3. *For any expanding quadratic algebraic integer α the equality*

$$\mathbb{Z}[\alpha] = \mathcal{B}[\alpha]$$

holds.

Theorem 9.3 is derived from Theorem 9.5. We obtain a similar result for expanding cubic trinomials.

Theorem 9.4. *Let α be an expanding algebraic integer which is a root of a cubic monic integer trinomial (i.e., polynomial of the form $x^3 + ax^2 + c$ or $x^3 + bx + c$). Then $\mathbb{Z}[\alpha] = \mathcal{B}[\alpha]$.*

The set of expanding cubic trinomials splits into two disjoint subsets, say, A and B. For the trinomials from A we apply Theorem 9.5. The subset B consists of trinomials of the form $x^3 - cx \pm c$, $c \geqslant 2$, $c \neq 8$. Theorem 9.11 (see Section9.2.1) shows that in case of a trinomial from B it is impossible to derive Theorem 9.4 from Theorem 9.5. Theorem 9.4 for trinomials from B is proved by constructing certain finite automaton, the so called counting automaton (see Section 9.2.4).

In general, we have the following result.

Theorem 9.5. *Suppose that an expanding algebraic integer α is a root of a polynomial*

$$P(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_0 \in \mathbb{Z}[x]$$

with

$$|a_0| \geqslant |a_1| + |a_2| + \ldots + |a_{d-1}| + 1.$$

Then $\mathbb{Z}[\alpha] = \tilde{\mathcal{B}}[\alpha]$ with $\tilde{\mathcal{B}} = \{0, \pm 1, \ldots, \pm(|a_0| - 1)\}$.

Theorem 9.5 follows from Proposition 3.1 of [95]. Nevertheless, we present an alternative proof of Theorem 9.5 in Section 9.2.1.

Note that the strict inequality $|a_0| > |a_1| + |a_2| + \ldots + |a_{d-1}| + 1$ would imply that all the roots of $P(x)$ are expanding algebraic integers.

Unfortunately, not every expanding algebraic integer α possesses a polynomial $P(x)$ satisfying the conditions of the theorem with $P(0) = \pm N(\alpha)$. In the Note at the end of Subsection 9.2.1, we provide an infinite family of such algebraic numbers which are roots of certain cubic integer trinomials. Such examples are minimal in terms of degree and the number of non-zero coefficients.

The best result we could obtain using Theorem 9.5 for a general expanding algebraic integer is the following:

Theorem 9.6. *Let α be an expanding algebraic integer of degree d (over $\mathbb{Q}$). Suppose that α_1 is a conjugate of α of least modulus. Then for any integer $n \geqslant -\log(2^{1/d} - 1)/\log|\alpha_1|$ we have*

$$\mathbb{Z}[\alpha] = \mathcal{B}_n[\alpha]$$

with $\mathcal{B}_n = \{0, \pm 1, \ldots, \pm(|N(\alpha)|^n - 1)\}$.

The upper bound $|N(\alpha)|^n - 1$ for the size of digits in $\mathcal{B}_n$ is large. By using more sophisticated division procedure, we were able to prove the next result.

Theorem 9.7. *Let α be an expanding algebraic integer of degree d with conjugates $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_d$. For any $\beta \in \mathbb{Z}[\alpha]$ there exists a nonzero polynomial $P(x) \in$*

$\mathbb{Z}[x]$ *of height at most*

$$\max\left\{\frac{|N(\alpha)|}{2\sqrt{D(\alpha)}}\sum_{i=1}^{d}\frac{\sqrt{|\alpha_i|^2-1}}{(|\alpha_i|-1)\sqrt{|\alpha_i|^{2d}-1}}\prod_{j=1}^{d}\sqrt{\frac{|\alpha_j|^{2d}-1}{|\alpha_j|^2-1}},\ |N(\alpha)|/2\right\}$$

such that $\beta = P(\alpha)$. *Here* $D(\alpha)$ *stands for the discriminant of* α.

The bound in our Theorem 9.7 seems to be much smaller than that of Theorem 9.6, however, there is no way of direct comparison. Nevertheless, in the division algorithm used in Theorem 9.7 we prove that in order to find the representations of elements of $\mathbb{Z}[\alpha]$ with the smallest possible digits, it suffices to find the expansions of finitely many elements of $\mathbb{Z}[\alpha]$ with conjugates in $\mathbb{Z}[\alpha_i]$ of absolute value less than or equal to $N(\alpha)/2(|\alpha_i|-1)$.

9.2 Proofs

9.2.1 Proofs of Theorems 9.5 and 9.6

Theorem 9.5 follows from the next lemma.

Lemma 9.8. *Suppose that an expanding algebraic integer* α *is a root of a polynomial* $P(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_0 \in \mathbb{Z}[x]$ *with*

$$|a_0| \geqslant |a_1| + |a_2| + \ldots + |a_{d-1}| + 1,$$

and $\tilde{\mathcal{B}} = \{0, \pm 1, \ldots, \pm(|a_0|-1)\}$. *Let* $A_0, A_1, \ldots, A_{d-1}$ *be integers with* $A_0 \notin \tilde{\mathcal{B}}$. *Then there exist integers* $A_0', A_1', \ldots, A_{d-1}'$ *and* $c_0, c_1, \ldots, c_k \in \tilde{\mathcal{B}}$ *such that*

$$A_0 + A_1\alpha + \ldots + A_{d-1}\alpha^{d-1} = c_0 + c_1\alpha + \ldots + c_k\alpha^k +$$

$$\left(A_0' + A_1'\alpha + \ldots + A_{d-1}'\alpha^{d-1}\right)\alpha^{k+1}$$

and $|A_0'| + |A_1'| + \ldots + |A_{d-1}'| < |A_0| + |A_1| + \ldots + |A_{d-1}|$.

Proof of Lemma 9.8. If $A_0 + A_1\alpha + \ldots + A_{d-1}\alpha^{d-1} = 0$ then we can take $k = 0$, $c_0 = 0$ and $A_i' = 0$ for all $i = 0, 1, \ldots, d-1$.

Further, assume that $A_0 + A_1\alpha + \ldots + A_{d-1}\alpha^{d-1} \neq 0$.

Assume without loss of generality that $A_0 > 0$. Then $A_0 \notin \tilde{\mathcal{B}}$ implies $A_0 \geqslant |a_0|$. Divide A_0 by a_0:

$$A_0 = c_0 + qa_0, \quad 0 \leqslant c_0 < |a_0|, \quad q \neq 0.$$

(Note that $qa_0 > 0$.) Then $P(\alpha) = 0$ implies

$$a_0 = -a_1\alpha - a_2\alpha^2 - \ldots - a_{d-1}\alpha^{d-1} - \alpha^d$$

and

$$A_0 = c_0 + qa_0 = c_0 - qa_1\alpha - qa_2\alpha^2 - \ldots - qa_{d-1}\alpha^{d-1} - q\alpha^d.$$

Hence

$$A_0 + A_1\alpha + \ldots + A_{d-1}\alpha^{d-1} = c_0 + (A_1 - qa_1)\alpha + \ldots +$$

$$(A_{d-1} - qa_{d-1})\alpha^{d-1} - q\alpha^d = c_0 + \left(B_0 + B_1\alpha + \ldots + B_{d-1}\alpha^{d-1}\right)\alpha$$

where $B_{d-1} = -q$ and $B_i = A_{i+1} - qa_{i+1}$, $i = 0, 1, \ldots, d-2$.

Further, $|a_0| \geqslant |a_1| + |a_2| + \ldots + |a_{d-1}| + 1$ implies

$$\sum_{i=0}^{d-1} |B_i| = \sum_{i=1}^{d-1} |A_i - qa_i| + |q| \leqslant \sum_{i=1}^{d-1} |A_i| + |q| \left(\sum_{i=1}^{d-1} |a_i| + 1\right) \leqslant$$

$$\sum_{i=1}^{d-1} |A_i| + |q||a_0| \leqslant \sum_{i=0}^{d-1} |A_i|.$$

If $c_0 \neq 0$, then the last inequality is strict, since $A_0 = |c_0 + qa_0| > |q||a_0|$. On the other hand, if $\sum_{i=0}^{d-1} |B_i| < \sum_{i=0}^{d-1} |A_i|$, then we can take $k = 0$, $A'_i = B_i$, $i = 0, 1, \ldots, d-1$, and we are done.

Further, assume that $\sum_{i=0}^{d-1} |B_i| = \sum_{i=0}^{d-1} |A_i|$. (Then $c_0 = 0$.)

If $B_i \in \tilde{\mathcal{B}}$ for all $i = 0, 1, \ldots, d-1$, then we can take $k = d$, $c_j = B_{j-1}$, $j = 1, 2, \ldots, d$, $A'_i = 0$ for all $i = 0, 1, \ldots, d-1$, and we are done in this case.

Now suppose that $B_t \notin \tilde{\mathcal{B}}$ for some $t \in \{0, 1, \ldots, d-1\}$. Let $s \in \{0, 1, \ldots, d-1\}$ be the smallest integer for which $B_s \neq 0$. If $B_s \in \tilde{\mathcal{B}}$ (in that case $s < d-1$), then we can take $k = s+1$, $c_1 = \ldots = c_s = 0$, $c_{s+1} = B_s$ and $A'_i = B_{s+i+1}$, $i = 0, 1, \ldots, d-s-2$ and $A'_i = 0$ for $i > d-s-2$. Indeed,

$$\sum_{i=0}^{d-1} |A'_i| = \sum_{i=s+1}^{d-1} |B_i| < \sum_{i=s}^{d-1} |B_i| = \sum_{i=0}^{d-1} |A_i|.$$

Finally, if $B_s \notin \tilde{\mathcal{B}}$ then we can repeat the above procedure with the number $B_s + B_{s+1}\alpha + \ldots$. After a finite number of iterations, we will receive the inequality $\sum_{i=0}^{d-1} |A'_i| < \sum_{i=0}^{d-1} |A_i|$. Otherwise the number

$$A_0 + A_1\alpha + \ldots + A_{d-1}\alpha^{d-1} \neq 0$$

would be divisible by α^n for every positive integer n, which is impossible, since α is expanding. $\square$

We will derive Theorem 9.6 from Theorem 9.5 using the following lemma.

Lemma 9.9. *Let $P(x) \in \mathbb{Z}[x]$ be a monic polynomial such that all roots of $P(x)$ are of modulus strictly greater than one. Then there exists a monic polynomial*

$$Q(x) = x^m + b_{m-1}x^{m-1} + \ldots + b_1 x + b_0 \in \mathbb{Z}[x]$$

which is a multiple of $P(x)$ and

$$|b_0| \geqslant |b_1| + |b_2| + \ldots + |b_{m-1}| + 1.$$

Moreover, for any integer $n \geqslant -\log(2^{1/d} - 1)/\log|\alpha_1|$ one can choose $Q(x)$ with $Q(0) = P(0)^n$, where d is the degree of $P(x)$ and α_1 is the root of $P(x)$ of least modulus.

Proof of Lemma 9.9. Let d be the degree of $P(x)$. Suppose that $\alpha_1, \alpha_2, \ldots, \alpha_d$ are all complex roots of $P(x)$ (not necessarily distinct). Assume without loss of generality that

$$1 < |\alpha_1| \leqslant |\alpha_2| \leqslant \ldots \leqslant |\alpha_d|.$$

Let n be a positive integer. Set

$$G(x) = \prod_{i=1}^{d}(x - \alpha_i^n) = x^d + g_{d-1}x^{d-1} + \ldots + g_1 x + g_0.$$

Clearly, all the coefficients g_i are integers. Now the inequality $1+|g_{d-1}|+\ldots+|g_1| \leqslant |g_0|$ is equivalent to

$$1 + |g_{d-1}| + \cdots + |g_1| + |g_0| \leqslant 2|g_0|.$$

Dividing both sides by $|g_0|$ we obtain

$$\frac{1}{|g_0|} + \frac{|g_{d-1}|}{|g_0|} + \ldots + \frac{|g_1|}{|g_0|} + 1 \leqslant 2.$$

Here the left hand side is

$$1 + \left|\sum_{i=1}^{d}\alpha_i^{-n}\right| + \left|\sum_{i<j}\alpha_i^{-n}\alpha_j^{-n}\right| + \ldots + \left|\prod_{i=1}^{d}\alpha_i^{-n}\right| \leqslant \prod_{i=1}^{d}(1 + |\alpha_i^{-n}|) \leqslant (1 + |\alpha_1^{-n}|)^d.$$

Hence the inequality $1 + |g_{d-1}| + \ldots + |g_1| \leqslant |g_0|$ holds provided $(1 + |\alpha_1^{-n}|)^d \leqslant 2$ which is equivalent to $n \geqslant -\log(2^{1/d} - 1)/\log|\alpha_1|$. Finally, note that the polynomial $Q(x) = G(x^n) = \prod_{i=1}^{d}(x^n - \alpha_i^n)$ is the required one. □

Remark 9.10. In Lemma 9.9 we get $g_0 = \pm P(0)$ provided the conjugates of α of

degree d all lie in $|z| > (2^{1/d} - 1)^{-1}$.

Proof of Theorem 9.6. Let α be an expanding algebraic integer with minimal polynomial $P(x)$. By Lemma 9.9 for any integer $n \geqslant -\log(2^{1/d}-1)/\log|\alpha_1|$ there is a monic polynomial $Q(x)$ with $Q(0) = P(0)^n$ which satisfies the condition of Theorem 9.5. Finally, note that $P(0) = \pm N(\alpha)$. □

Note. Suppose that α is an expanding algebraic integer. In order to prove the equality $\mathbb{Z}[\alpha] = \mathcal{B}[\alpha]$ using Theorem 9.5, one needs a polynomial $P(x)$ satisfying the conditions of Theorem 9.5 and $P(0) = \pm N(\alpha)$. Unfortunately, this is false in general. Consider an algebraic integer α which is the root of cubic trinomial $p(x) = x^3 - cx + c$, $c \geqslant 2, c \neq 8, c \in \mathbb{Z}$. If $p(x)$ is reducible in $\mathbb{Z}[x]$, then it has an integer root, say, m. The equation $m^3 = c(m-1)$ implies that $m-1$ divides m^3. Since $\gcd(m^3, m-1) = 1$ and $c > 0$, this implies $m - 1 = 1$. Thus $m = 2, c = 8$. Hence, the polynomial $p(x)$ is irreducible in $\mathbb{Z}[x]$ if $c \geqslant 2, c \neq 8$. By direct substitution one easily checks that $p(x)$ has three real roots in intervals $(-\sqrt{c}, -\sqrt{c}+1)$, $(1+1/c, 3/2)$ and $(\sqrt{c}-1, \sqrt{c})$ if $c \geqslant 7$, all of modulus strictly greater than one. For $c = 2, 3, 4, 5, 6$, the polynomial $p(x)$ has one real and two complex roots outside the unit circle, which can be verified by direct computation. Alternatively, use the Shur-Cohn criterion [97], [143]. Thus α is a cubic expanding algebraic integer. In Theorem 9.11 below, we prove that $\mathbb{Z}[\alpha] = \mathcal{B}[\alpha]$ in principle cannot be established by Theorem 9.5.

Theorem 9.11. *The polynomial $p(x) = x^3 - cx + c, c \in \mathbb{Z}, c \geqslant 2, c \neq 8$ does not divide any polynomial $P(x) = a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$ with $|a_0| \geqslant |a_1| + |a_2| + \cdots + |a_n|$ and $a_0 = \pm c$.*

Proof of Theorem 9.11. Assume that there exists a polynomial $P(x) = a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$ which is a multiple of $p(x)$ and satisfies $|a_0| \geqslant |a_1| + |a_2| + \cdots + |a_n|$ with $a_0 = \pm c$. Then $P(x) = p(x)q(x)$ for some non constant polynomial $q \in \mathbb{Z}[x]$. Since $a_0 = \pm c$, $q(0) = \pm 1$. Hence, any irreducible factor of $q(x)$ has a root of modulus less or equal to 1. Let ζ be one of such roots. Then $P(\zeta) = 0$ implies

$$-a_0 = a_1\zeta + a_2\zeta^2 + \ldots a_n\zeta^n. \tag{9.1}$$

This implies $\zeta^k = \pm 1$ for any coefficient $a_k \neq 0, k = 1 \ldots n$. Otherwise, by comparing the real parts of the complex numbers in both sides of (9.1), one has

$$|a_1| + |a_2| + \cdots + |a_n| > |\Re(a_1\zeta + a_2\zeta^2 + \ldots a_n\zeta^n)| = |a_0|,$$

which contradicts the assumption. This shows that ζ is a root of unity. Thus $q(x)$

is a product of cyclotomic polynomials and a constant $a \in \mathbb{Z}$. Since $q(0) = \pm 1$, $a = \pm 1$. We claim that

$$q(x) = \pm(x-1)^r(x+1)^s(x^2+1)^t(x^2+x+1)^u(x^2-x+1)^v, \qquad (9.2)$$

with integer exponents $r, s, t, u, v \geqslant 0$. To prove this, it suffices to show that at least one coefficient a_1, a_2 or a_3 is not equal to 0, so we have $\zeta = \pm 1$, $\zeta^2 = \pm 1$ or $\zeta^3 = \pm 1$ in (9.1).

Assume that $a_1 = a_2 = a_3 = 0$. Let α be the root of polynomial $p(x) = x^3 - cx + c$. Then (9.1) with ζ replaced by α implies that α^4 divides $a_0 = \pm c$ in the ring R of algebraic integers of $\mathbb{Q}(\alpha)$. Note that $p(\alpha) = 0$ gives $\alpha^3 = c(\alpha - 1)$. Thus $\alpha^4 | c$ in R implies $\alpha^4 | \alpha^3$, so α is a unit in R. This is impossible, since $c \geqslant 2$ and $p(x)$ is irreducible if $c \neq 8$, so the claim (9.2) is proved.

Observe that $t \geqslant 1$ in (9.2) implies $2|k$ for every non zero coefficient a_k, $k = 1 \ldots n$ in (9.1), since $i^k = \pm 1$ if and only if $2|k$ (here, as usual, $i^2 = -1$). In this case, $P(x) = P(-x) = P_1(x^2)$ for some polynomial $P_1 \in \mathbb{Z}[x]$. This is impossible, since such a polynomial $P(x)$ would be divisible by $p(x)$ and $p(-x)$ so $p(0)^2 = c^2$ divides $a_0 = P(0) = \pm c$ contradicting $c \geqslant 2$.

Similarly, $3|k$ for any non-zero a_k in (9.1) provided $u \geqslant 1$ or $v \geqslant 1$, since $(\pm e^{\pm 2\pi i/3})^k = \pm 1$ if and only if $3|k$. In this case, $P(x) = P_1(x^3)$ for some $P_1 \in \mathbb{Z}[x]$. Set $\zeta = e^{2\pi i/3}$. Then $P(\alpha) = P(\zeta\alpha) = P_1(\alpha^3) = 0$ for any root α of $p(x)$. The polynomials $p(x)$ and $p(\zeta x)$ have no roots in common, since

$$p(\zeta\alpha) - p(\alpha) = (\zeta^3\alpha^3 - c\zeta\alpha + c) - (\alpha^3 - c\alpha + c) = c(1-\zeta)\alpha \neq 0.$$

This implies that $P(x)$ is a multiple of $p(x)p(\zeta x)$. Since all roots of P are of modulus greater or equal to one, one has $|P(0)| \geqslant |p(0)p(\zeta 0)| = |p(0)|^2 = c^2 > c = |a_0| = |P(0)|$, which again leads to the contradiction.

From the arguments given above, it follows that $t = u = v = 0$, thus $q(x) = (x-1)^r(x+1)^s$ is the only remaining possibility. Then

$$|P(i)|^2 = |p(i)q(i)|^2 = |(i^3 - ci + c)^2(i-1)^r(i+1)^s|^2 = ((1+c)^2 + c^2)2^{r+s}.$$

The inequality

$$|P(i)| \leqslant |a_n| + \cdots + |a_1| + |a_0| \leqslant 2|a_0| = 2c,$$

implies

$$((1+c)^2 + c^2)2^{r+s} \leqslant 4c^2$$

which is impossible unless $r = s = 0$. This contradicts the assumption that $q(x)$

is a non constant polynomial and concludes the proof of Theorem 9.11.

9.2.2 Proof of Theorem 9.3

The following lemma provides a necessary condition for a quadratic algebraic integer to be expanding which will be used in the proof of Theorem 9.3.

Lemma 9.12. *Let α be an expanding quadratic algebraic integer with the minimal polynomial $x^2 + ax + b$. Then $|a| \leqslant |b|$. The equality $|a| = |b|$ holds if and only if $b = |a| \geqslant 2$ and $|a| \neq 4$.*

One could employ the necessary and sufficient conditions (see Corollary 2.1 of [5]) developed using the Schur-Cohn criterion [97], [143]. Nevertheless, we provide the proof of Lemma 9.12.

Proof of Lemma 9.12. We might assume that $a \geqslant 0$, since $a = -(\alpha + \alpha')$ and α is expanding if and only if $-\alpha$ is expanding. Here α' stands for the conjugate of α.

Suppose, contrary to our claim, that $a > |b|$. This implies the inequalities

$$(a-2)^2 \leqslant a^2 - 4b < (a+2)^2,$$

$$a - 2 \leqslant \sqrt{a^2 - 4b} < a + 2$$

and

$$\left| \frac{-a + \sqrt{a^2 - 4b}}{2} \right| \leqslant 1$$

which is a contradiction, since

$$\{\alpha,\, \alpha'\} = \left\{ \frac{-a \pm \sqrt{a^2 - 4b}}{2} \right\}.$$

Now, suppose that $|b| = a > 0$ and α is expanding. We claim that $b = a$. Indeed, $b = -a$ implies

$$0 < \frac{-a + \sqrt{a^2 + 4a}}{2} = \frac{2a}{a + \sqrt{a^2 + 4a}} < \frac{2a}{a + a} = 1$$

which again leads to the contradiction.

Thus $b = a > 0$. Assume that $b = a \geqslant 5$. Then

$$\min\{|\alpha|, |\alpha'|\} = \min\left\{ \left| \frac{-a \pm \sqrt{a^2 - 4a}}{2} \right| \right\} = \frac{a - \sqrt{a^2 - 4a}}{2} =$$

$$= \frac{2a}{a + \sqrt{a^2 - 4a}} > \frac{2a}{a + a} = 1$$

which implies that α is expanding.

Finally, one easily checks that $b = a = 2$ or 3 implies that α is expanding, whereas $b = a = 1$ or 4 implies that α is not expanding quadratic algebraic integer. □

Proof of Theorem 9.3. Let α be an expanding quadratic algebraic integer with the minimal polynomial x^2+ax+b. Assume without loss of generality that $a \geqslant 0$. (Indeed, Theorem 9.3 holds for α if and only if it holds for $-\alpha$.) By Lemma 9.12, $0 \leqslant a \leqslant |b|$. If $a+1 \leqslant |b|$ then the result follows from Theorem 9.5 with $P(x) = x^2 + ax + b$. Suppose that $a = |b|$. By Lemma 9.12 $b = a \geqslant 2$ and $a \neq 4$. Now the minimal polynomial of α is $x^2 + ax + a$ and we can apply Theorem 9.5 with $P(x) = (x-1)(x^2 + ax + a) = x^3 + (a-1)x^2 - a$. □

9.2.3 Proof of Theorem 9.4

In the proof of Theorem 9.4 we will construct a finite automaton which is called *transducer* (cf. [23], [79]). We follow the notations of [178].

Definition 9.13. *The 6-tuple $A = (Q, \Sigma, \Delta, q, q_0, \delta)$ is called a finite transducer automaton if*

- *Q, Σ and Δ are non empty, finite sets, and*
- *$q : Q \times \Sigma \to Q$ and $\delta : Q \times \Sigma \to \Delta$ are unique mappings.*

The sets Σ and Δ are called input and output alphabet, respectively. Q is the set of states and q_0 is the starting state. The mappings q and δ are called transformation and result function, respectively.

We will use the following characterization of expanding cubic polynomials.

Lemma 9.14. *The polynomial $p(x) = x^3 + ax^2 + bx + c$ with integer coefficients is expanding if and only if*

$$\begin{cases} |b - ac| < c^2 - 1, \\ |b + 1| < |a + c|. \end{cases} \tag{9.3}$$

Proof. This is Lemma 1 from Akiyama and Gjini [5]. □

Proof of Theorem 9.4. Suppose that α is an expanding cubic algebraic integer which is a root of the trinomial $p(x) = x^3 + ax^2 + bx + c$. Then either $a = 0$ or $b = 0$. If $b = 0$ then the first inequality of (9.3) implies $|a||c| < c^2 - 1$ and $|a| < |c|$. Hence, each expanding cubic trinomial $x^3 + ax^2 + c$ satisfies $1 + |a| \leqslant |c|$, and we can apply Theorem 9.5. Now suppose that $a = 0$. Then the second inequality of

(9.3) implies $|b+1| < |c|$. If $b \geqslant 0$, then $1+|b| < |c|$, and again we can apply Theorem 9.5. Let $b < 0$. Then the inequality $|b+1| < |c|$ implies $b \geqslant -|c|$. If $b \geqslant -|c|+1$, then $1+|b| \leqslant |c|$, and once again we can apply Theorem 9.5. Finally, we are left with the trinomials $p_1(x) = x^3 - cx + c$, and $p_2(x) = x^3 - cx - c$, $c \geqslant 2$. Note that $p_2(-x) = -p_1(x)$. Hence, it is enough to consider the trinomial $x^3 - cx + c$, $c \geqslant 2$. This trinomial is irreducible provided $c \neq 8$ (see the note before Theorem 9.11). However, Theorem 9.11 shows that in this case it is impossible to apply Theorem 9.5. Instead, we will construct a finite automaton for this trinomial.

Now we briefly discuss how to construct the counting automaton $A_0(1)$ which performs the addition of 1 in $\mathcal{B}[\alpha]$. We will follow the explanation presented in [178]. Denote $(\sigma_N, \ldots, \sigma_0) = \sum_{j=0}^{N} \sigma_j \alpha^j$. We say that $(\sigma_N, \ldots, \sigma_0)$ is an α-adic representation of $v \in \mathbb{Z}[\alpha]$ if $v = (\sigma_N, \ldots, \sigma_0)$ and $\sigma_0, \ldots, \sigma_N \in \mathcal{B}$. Suppose $v \in \mathbb{Z}[\alpha]$ has α-adic representation $v = (d_N(v), d_{N-1}(v), \ldots, d_0(v))$. We want to add 1 to the α-adic representation of v, i.e., we want to construct the α-adic representation of $v+1 = (d_{N'}(v+1), d_{N'-1}(v+1), \ldots, d_0(v+1))$, $d_j(v+1) \in \mathcal{B}$. We perform the addition digit wise, from right to left. First, we add 1 to the first digit $d_0(v)$. The addition produces a carry $q_1 \in \mathbb{Z}[\alpha]$ obeying the scheme $d_0(v) + 1 = d_0(v+1) + \alpha q_1$. Note that in contrast to [178], our $d_0(v+1)$ and q_1 are not unique unless $d_0(v+1) = 0$. This reduces the problem of adding 1 to v to the problem of adding q_1 to $(d_N(v), d_{N-1}(v), \ldots, d_1(v))$. Iterating this procedure yields the general scheme

$$d_j(v) + q_j = d_j(v+1) + \alpha q_{j+1}, \quad j \geqslant 0. \tag{9.4}$$

Since the division procedure (9.4) is not unique, we restrict our iteration procedure to the following: for each pair $(q_j, d_j(v))$ we fix the pair $(q_{j+1}, d_j(v+1))$ satisfying (9.4), and each time the iteration starts with $(q_j, d_j(v))$ we will use the same pair $(q_{j+1}, d_j(v+1))$. Adopting the notation of Definition 9.13, we define the counting automaton $A_0(1)$ by setting

$Q =$ the set of possible carries,

$\Sigma = \Delta = \mathcal{B}$,

$q_0 = 1$,

$q : Q \times \Sigma \to Q : (q_j, d_j(v)) \mapsto q_{j+1}$ according to (9.4),

$\delta : Q \times \Sigma \to \Delta : (q_j, d_j(v)) \mapsto d_j(v+1)$ according to (9.4).

Now we explicitly construct the counting automaton $A_0(1)$ for α – a root of $x^3 - cx + c$, $c \geqslant 2$, $c \neq 8$. Consider the following table.

number of carry	carry	:	input/output	next carry
0	0	:	$k\|k$	0
1	1	:	$k \leqslant c-2,\ k\|k+1$	0
		:	$c-1\|0$	2
2	$\overline{1}\,0\,c$	:	$k\|k$	3
3	$\overline{1}\,\overline{1}\,c$	:	$k\|k$	4
4	$\overline{1}\,\overline{1}\,c-1$	:	$k \leqslant 0,\ k\|k+c-1$	5
		:	$k \geqslant 1,\ k\|k-1$	4
5	$\overline{1}\,\overline{1}$	:	$k \geqslant \overline{c}+2,\ k\|k-1$	6
		:	$\overline{c}+1\|0$	7
6	$\overline{1}$	:	$k \geqslant \overline{c}+2,\ k\|k-1$	0
		:	$\overline{c}+1\|0$	8
7	$1\,0\,\overline{c}-1$	:	$k \geqslant \overline{c}+2,\ k\|k-1$	9
		:	$\overline{c}+1\|0$	10
8	$1\,0\,\overline{c}$	:	$k\|k$	9
9	$1\,1\,\overline{c}$	:	$k\|k$	11
10	$2\,1\,\overline{c}+\overline{c}$	:	$k\|k$	12
11	$1\,1\,\overline{c}+1$	:	$k \leqslant \overline{1},\ k\|k+1$	11
		:	$k \geqslant 0,\ k\|k-c+1$	13
12	$2\,2\,1+\overline{c}+\overline{c}$	:	$k \leqslant \overline{1},\ k\|k+1$	14
		:	$k \geqslant 0,\ k\|k-c+1$	15
13	$1\,1$	:	$k \leqslant c-2,\ k\|k+1$	1
		:	$c-1\|0$	16
14	$2\,2\,2+\overline{c}+\overline{c}$	:	$k \leqslant \overline{2},\ k\|k+2$	14
		:	$k \geqslant \overline{1},\ k\|k-c+2$	15

number of carry	carry	:	input/output	next carry
15	$1\,2\,\overline{c}+2$	:	$k \leqslant \overline{2},\ k\vert k+2$	17
		:	$k \geqslant \overline{1},\ k\vert k-c+2$	18
16	$\overline{1}\,0\,c+1$	:	$k \leqslant c-2,\ k\vert k+1$	3
		:	$c-1\vert 0$	19
17	$1\,1\,\overline{c}+2$	:	$k \leqslant \overline{2},\ k\vert k+2$	11
		:	$k \geqslant \overline{1},\ k\vert k-c+2$	13
18	$1\,2$	:	$k \leqslant c-3,\ k\vert k+2$	1
		:	$k \geqslant c-2,\ k\vert k-c+2$	16
19	$\overline{2}\,\overline{1}\,c+c$	:	$k\vert k$	20
20	$\overline{2}\,\overline{2}\,c+c+\overline{1}$	:	$k \leqslant 0,\ k\vert k+c-1$	21
		:	$k \geqslant 1,\ k\vert k-1$	22
21	$\overline{1}\,\overline{2}\,c-2$	:	$k \leqslant 1,\ k\vert k+c-2$	23
		:	$k \geqslant 2,\ k\vert k-2$	24
22	$\overline{2}\,\overline{2}\,\overline{2}+c+c$	:	$k \leqslant 1,\ k\vert k+c-2$	21
		:	$k \geqslant 2,\ k\vert k-2$	22
23	$\overline{1}\,\overline{2}$	:	$k \leqslant \overline{c}+2,\ k\vert k+c-2$	7
		:	$k \geqslant \overline{c}+3,\ k\vert k-2$	6
24	$\overline{1}\,\overline{1}\,c-2$	:	$k \leqslant 1,\ k\vert k+c-2$	5
		:	$k \geqslant 2,\ k\vert k-2$	4

Here $\overline{a}$ denotes $-a$. The second column *carry* indicates the carry. Carries are numbered in the first column *number of carry*. The third column *input/output* defines the result function δ: $k \in \mathcal{B}$ denotes the input digit and $k|u(k)$ means that the corresponding output is $u(k) \in \mathcal{B}$. The fourth column *next carry* defines the transformation function q indicating the number of the next carry.

One can check that this counting automaton $A_0(1)$ has no *zero cycles*, i.e., if we begin with any carry from the second column and start walking the zero path (each time taking input 0), eventually we will reach the sync point – carry 0. This means that we can add 1 to any α-adic representation $v \in \mathbb{Z}[\alpha]$ and obtain an α-adic representation of $v+1$.

If we run the counting automaton $A_0(1)$ starting with the carry no. 6 (i.e. $q_0 := \overline{1}$), this would produce the subtraction of 1. Now if we run $A_0(1)$ starting with the carry no. 13, this would produce addition of $1\,1 = \alpha + 1$. Then we take the resulting representation and subtract 1. This gives the addition of $1\,0 = \alpha$. Similarly running $A_0(1)$ with the starting carry no. 5 and then adding 1, we obtain the subtraction of α. If we run $A_0(1)$ starting with the carry no. 11, then subtract

$1\,0 = \alpha$ and then for $c-1$ times add 1, we would get the addition of $1\,0\,0 = \alpha^2$. Finally, running $A_0(1)$ with starting carry no. 4, then adding $1\,0 = \alpha$ and then for $c-1$ times subtracting 1, we obtain the subtraction of $1\,0\,0 = \alpha^2$. Hence starting with 0 and applying ± 1 or $\pm\alpha$ or $\pm\alpha^2$, we can find α-adic representation of any number lying in $\mathbb{Z}[\alpha] = \mathbb{Z} + \mathbb{Z}\alpha + \mathbb{Z}\alpha^2$. □

Note. The polynomial $x^3 - cx + c$, $c \geqslant 2$, $c \neq 8$ is not a CNS polynomial (see Theorem 3 of [50]).

9.2.4 Proof of Theorem 9.7

Proof. Let $p(x) = x^d + a_{d-1}x^{d-1} + \ldots + a_1 x + a_0$ be the minimal polynomial of α . (Then $N(\alpha) = \alpha_1\alpha_2 \cdot \ldots \cdot \alpha_d = \pm a_0$.) Let $\gamma \in \mathbb{Z}[\alpha]$, $\gamma = C_0 + C_1\alpha + \ldots + C_{d-1}\alpha^{d-1}$, $C_j \in \mathbb{Z}$. Then the conjugates of γ are $\gamma_i = C_0 + C_1\alpha_i + \ldots + C_{d-1}\alpha_i^{d-1}$, $i = 1, 2, \ldots, d$. Consider the following division procedure. There are integers r and q such that $C_0 = r + a_0 q$ and $|r| \leqslant |a_0|/2$. The equality $p(\alpha_i) = 0$ implies

$$a_0 = -a_1\alpha_i - \ldots - a_{d-1}\alpha_i^{d-1} - \alpha_i^d.$$

Thus

$$C_0 = r - \alpha_i\left(a_1 q + a_2 q\alpha_i + \ldots + a_{d-1}q\alpha_i^{d-2} + q\alpha_i^{d-1}\right).$$

Denote

$$\gamma_i = r + \alpha_i\gamma_i'$$

where $\gamma_i' = C_0' + C_1'\alpha_i + \ldots + C_{d-1}'\alpha_i^{d-1}$ with integers $C_j' = C_{j+1} - a_{j+1}q$, $0 \leqslant j \leqslant d-2$, $C_{d-1}' = -q$. (Note that the numbers C_j' do not depend on the choice of conjugate γ_i.)

Now fix $i \in \{1, 2, \ldots, d\}$ and define the sequence $x_n^{(i)}$ as follows.

$$x_0^{(i)} = \beta_i = B_0 + B_1\alpha_i + \ldots + B_{d-1}\alpha_i^{d-1},$$

$B_j \in \mathbb{Z}$, $j = 0, 1, \ldots, d-1$, and $x_{n+1}^{(i)}$ is obtained from $x_n^{(i)}$ via the division procedure described above, i. e.,

$$x_n^{(i)} = r_n + \alpha_i x_{n+1}^{(i)}, \quad |r_n| \leqslant |a_0|/2, \ \ n \geqslant 0. \tag{9.5}$$

Then

$$\beta_i = r_0 + r_1\alpha_i + \ldots + r_{n-1}\alpha_i^{n-1} + \alpha_i^n x_n^{(i)} \tag{9.6}$$

and

$$|x_n^{(i)}| = \left|\frac{\beta_i}{\alpha_i^n} - \frac{r_0}{\alpha_i^n} - \ldots - \frac{r_{n-1}}{\alpha_i}\right| \leqslant \frac{|\beta_i|}{|\alpha_i|^n} + \frac{|r_0|}{|\alpha_i|^n} + \ldots + \frac{|r_{n-1}|}{|\alpha_i|} \leqslant$$

$$\frac{|\beta_i|}{|\alpha_i|^n} + \frac{|a_0|}{2}\left(\frac{1}{|\alpha_i|} + \frac{1}{|\alpha_i|^2} + \dots\right) = \frac{|\beta_i|}{|\alpha_i|^n} + \frac{|a_0|}{2(|\alpha_i| - 1)}.$$

Let $m = \min_{1\leqslant i\leqslant d} |\alpha_i|$ and $M = \max_{1\leqslant i\leqslant d} |\beta_i|$. Then the last inequality yields

$$|x_n^{(i)}| \leqslant \frac{M}{m^n} + \frac{|a_0|}{2(|\alpha_i| - 1)} \leqslant \frac{M}{m^n} + \frac{|a_0|}{2(m-1)}. \tag{9.7}$$

Thus the set $\{x_n^{(i)} : 1 \leqslant i \leqslant d,\ n \geqslant 0\}$ is finite, since it consists of algebraic integers of degree at most d with bounded conjugates. Now (9.5) implies that the sequence $x_n^{(i)}$ is periodic starting from certain $n \geqslant n_0$. (Note that n_0 does not depend on the choice of conjugate $x_n^{(i)}$.)

Further, take any $\delta_i \in \{x_n^{(i)} : n \geqslant n_0\}$. Since $\delta_i = x_n^{(i)}$ for infinitely many positive integers n, (9.7) shows that

$$|\delta_i| \leqslant \frac{|a_0|}{2(|\alpha_i| - 1)} = \frac{|N(\alpha)|}{2(|\alpha_i| - 1)} \tag{9.8}$$

for all $i = 1, 2, \dots, d$. Since $\delta_i \in \mathbb{Z}[\alpha_i]$, there exist integers $A_0, A_1, \dots, A_{d-1}$ such that

$$A_0 + A_1\alpha_i + \dots + A_{d-1}\alpha_i^{d-1} = \delta_i, \quad i = 1, 2, \dots, d.$$

By Cramer's rule,

$$A_j = \frac{1}{\det(\alpha_i^r)} \begin{vmatrix} 1 & \alpha_1 & \cdots & \alpha_1^{j-1} & \delta_1 & \alpha_1^{j+1} & \cdots & \alpha_1^{d-1} \\ 1 & \alpha_2 & \cdots & \alpha_2^{j-1} & \delta_2 & \alpha_2^{j+1} & \cdots & \alpha_2^{d-1} \\ & & & \cdots & \cdots & & & \\ 1 & \alpha_d & \cdots & \alpha_d^{j-1} & \delta_d & \alpha_d^{j+1} & \cdots & \alpha_d^{d-1} \end{vmatrix} \tag{9.9}$$

for $j = 0, 1, \dots, d-1$. Denote by U_k, $1 \leqslant k \leqslant d$, the determinant obtained from the last determinant by omitting the $k-$th row and the $j+1-$th column. On applying Hadamard's inequality, one obtains

$$|U_k| \leqslant \prod_{r\neq k} \sqrt{\frac{|\alpha_r|^{2d} - 1}{|\alpha_r|^2 - 1}}. \tag{9.10}$$

It's well-known that $\det^2(\alpha_i^r) = D(\alpha)$, where $D(\alpha)$ stands for the discriminant of α (see, e. g., Chapter 2 of [156]). Then in view of (9.9), (9.8) and (9.10), we have

$$|A_j| = \frac{1}{\sqrt{D(\alpha)}}\left|\sum_{k=1}^{d} \delta_k U_k\right| \leqslant \frac{1}{\sqrt{D(\alpha)}} \sum_{k=1}^{d} \frac{|N(\alpha)|}{2(|\alpha_k| - 1)} \prod_{r\neq k} \sqrt{\frac{|\alpha_r|^{2d} - 1}{|\alpha_r|^2 - 1}} =$$

$$\frac{|N(\alpha)|}{2\sqrt{D(\alpha)}}\prod_{r=1}^{d}\sqrt{\frac{|\alpha_r|^{2d}-1}{|\alpha_r|^2-1}}\sum_{k=1}^{d}\frac{\sqrt{|\alpha_k|^2-1}}{(|\alpha_k|-1)\sqrt{|\alpha_k|^{2d}-1}}. \tag{9.11}$$

Now, $\delta_i = x_n^{(i)}$ for certain n. Then in view of (9.6), we obtain

$$\beta = \beta_1 = r_0 + r_1\alpha + \ldots + r_{n-1}\alpha^{n-1} + \alpha^n\delta_1 =$$

$$r_0 + r_1\alpha + \ldots + r_{n-1}\alpha^{n-1} + A_0\alpha^n + A_1\alpha^{n+1} + \ldots + A_{d-1}\alpha^{n+d-1}.$$

Finally, in view of (9.11), the polynomial

$$P(x) = r_0 + r_1x + \ldots + r_{n-1}x^{n-1} + A_0x^n + A_1x^{n+1} + \ldots + A_{d-1}x^{n+d-1}$$

is the required one. □

Chapter 10

Metric Mahler measure

10.1 Statement of the problem

Through the Chapter 10, the *(logarithmic) Mahler measure* of the polynomial $f \in \mathbb{C}[z]$ is denoted by

$$m(f) = \log|a| + \sum_{n=1}^{N} \log^{+}|\alpha_n|.$$

If α is a non-zero algebraic number, the *(logarithmic) Mahler measure* $m(\alpha)$ *of* α is defined as the Mahler measure of the minimal polynomial of α over $\mathbb{Z}$.

For an algebraic number α, Dubickas and Smyth [74] introduced the *metric Mahler measure* $m_1(\alpha)$ by

$$m_1(\alpha) = \inf\left\{\sum_{n=1}^{N} m(\alpha_n) : N \in \mathbb{N},\ \alpha_n \in \overline{\mathbb{Q}}^{\times},\ \alpha = \prod_{n=1}^{N} \alpha_n\right\}. \tag{10.1}$$

Here, the infimum is taken over all ways to write α as a product of algebraic numbers. The advantage of m_1 over m is that it satisfies the triangle inequality

$$m_1(\alpha\beta) \leqslant m_1(\alpha) + m_1(\beta)$$

for all algebraic numbers α and β. In view of this observation, m_1 is well-defined on the quotient group $G = \overline{\mathbb{Q}}^{\times}/\mathrm{Tor}(\overline{\mathbb{Q}}^{\times})$, and the map $(\alpha, \beta) \mapsto m_1(\alpha\beta^{-1})$ defines a metric on G. This metric induces the discrete topology if and only if Lehmer's conjecture is true.

The metric Mahler measure m_1 is only a special case of the *t-metric Mahler*

measures, which are defined for $t \geqslant 1$ by

$$m_t(\alpha) = \inf\left\{\left(\sum_{n=1}^{N} m(\alpha_n)^t\right)^{1/t} : N \in \mathbb{N},\ \alpha_n \in \overline{\mathbb{Q}}^{\times},\ \alpha = \prod_{n=1}^{N} \alpha_n\right\}.$$

In addition, the ∞-*metric Mahler measure* of α is defined by

$$m_\infty(\alpha) = \inf\left\{\max_{1\leqslant n\leqslant N}\{m(\alpha_n)\} : N \in \mathbb{N},\ \alpha_n \in \overline{\mathbb{Q}}^{\times},\ \alpha = \prod_{n=1}^{N} \alpha_n\right\}.$$

The t-metric Mahler measures were introduced and studied in [176, 177]. It follows from the results of [176] that these functions have analogues of the triangle inequality

$$m_t(\alpha\beta)^t \leqslant m_t(\alpha)^t + m_t(\beta)^t \quad \text{and} \quad m_\infty(\alpha\beta) \leqslant \max\{m_\infty(\alpha), m_\infty(\beta)\}$$

Hence, the map $(\alpha, \beta) \mapsto m_t(\alpha\beta^{-1})$ defines a metric on G that induces the discrete topology if and only if Lehmer's conjecture is true.

If $t \in [1, \infty]$ and $\alpha \in \overline{\mathbb{Q}}$, we say that *the infimum in* $m_t(\alpha)$ *is attained by* $\alpha_1, \ldots, \alpha_N$ if we have that

$$\alpha = \alpha_1 \cdots \alpha_N \quad \text{and} \quad m_t(\alpha) = \begin{cases} \left(\sum_{n=1}^{N} m(\alpha_n)^t\right)^{1/t} & \text{if } t < \infty \\ \max_{1\leqslant n\leqslant N}\{m(\alpha_n)\} & \text{if } t = \infty. \end{cases}$$

If S is any subset of $\overline{\mathbb{Q}}$, we say *the infimum in* $m_t(\alpha)$ *is attained in* S if there exist points $\alpha_1, \ldots, \alpha_N \in S$ that attain the infimum in $m_t(\alpha)$.

It is not immediately obvious that $m_t(\alpha)$ is attained for all values of α and t. Dubickas and Smyth [74] conjectured that the infimum in $m_1(\alpha)$ is always attained a fact later proved by Samuels [175]. More specifically, if K_α is the Galois closure of $\mathbb{Q}(\alpha)$ over $\mathbb{Q}$ and

$$\mathrm{Rad}(K_\alpha) = \{\gamma \in \overline{\mathbb{Q}} : \gamma^n \in K_\alpha \text{ for some } n \in \mathbb{N}\},$$

then the infimum in $m_1(\alpha)$ is attained in $\mathrm{Rad}(K_\alpha)$. Using the same method, this result was generalized for all t-metric Mahler measures in [176]. That is, for every $t \geqslant 1$, the infimum in $m_t(\alpha)$ is attained in $\mathrm{Rad}(K_\alpha)$.

It is natural to ask if these results can be improved, having a smaller set S in place of $\mathrm{Rad}(K_\alpha)$. In particular, for each $\alpha \in \overline{\mathbb{Q}}$, we would like to identify a set S_α whose points generate a finite extension of $\mathbb{Q}$ and the infimum in $m_t(\alpha)$ is attained in S_α for all t. This problem is of considerable importance if we hope to compute exact values of $m_t(\alpha)$. For example, Conjecture 2.1 of [177] predicts that

if α is rational, then the infimum in $m_t(\alpha)$ is attained in $\mathbb{Q}$. With this assumption, it is possible to graph some examples of the function $t \mapsto m_t(\alpha)$, where $\alpha \in \mathbb{Q}$. In particular, the main question we will investigate through Chapter 10 is:

Problem 10.1. *Is it true that the infimum in $m_t(\alpha)$ is always attained in $\mathbb{Q}(\alpha)$?*

It follows from [74] and [93] that Conjecture 2.1 of [176] holds for $t = 1$ and $t = \infty$. Unfortunately, these methods seem genuinely distinct and cannot be easily generalized to handle all values of t and α. As our first result, we prove this conjecture for all $t \geqslant 1$.

Theorem 10.2. *If α is a non-zero rational number and $t \in [1, \infty]$, then the infimum in $m_t(\alpha)$ is attained in $\mathbb{Q}$.*

Our next question is whether Theorem 10.2 can be extended to arbitrary algebraic numbers α. In view of Theorem 10.2, one might suspect that the infimum in $m_t(\alpha)$ is always attained in K_α. This turns out to be false, however, as we are able to produce an infinite family of quadratic counterexamples. More specifically, if D is a square-free positive integer, we show precisely when $m_t(\sqrt{D})$ is attained in $K_{\sqrt{D}} = \mathbb{Q}(\sqrt{D})$.

Theorem 10.3. *Suppose that $p_1, \dots, p_L$ are distinct primes written in decreasing order, $D = p_1 \cdots p_L$, and $t \in (1, \infty]$. The infimum in $m_t(\sqrt{D})$ is attained in $\mathbb{Q}(\sqrt{D})$ if and only if $D < p_1^2$. In this situation, the infimum is attained by points*

$$\sqrt{\frac{p_1}{p_2 \cdots p_L}}, p_2, \dots, p_L \in \mathbb{Q}(\sqrt{D}),$$

and we have that

$$m_t(\sqrt{D}) = \begin{cases} \left(\sum_{\ell=1}^{L} (\log p_\ell)^t\right)^{1/t} & \text{if } t \in (1, \infty) \\ \log p_1 & \text{if } t = \infty. \end{cases}$$

Theorem 10.3 enables the construction of infinitely many integers D such that $m_t(\sqrt{D})$ is not attained in $K_{\sqrt{D}} = \mathbb{Q}(\sqrt{D})$ for any $t > 1$. Theorem 10.4 below gives a set of points that attain the infimum in $m_t(\alpha)$ for algebraic numbers $\alpha = D^{1/k}$, where $D > 0$ is a square-free integer.

Theorem 10.4. *If $p_1, \dots, p_L$ are distinct primes, $D = p_1 \cdots p_L$, and $t \in [1, \infty]$, then the infimum in $m_t(D^{1/k})$ is attained by $p_1^{1/k}, \dots, p_L^{1/k}$ and*

$$m_t(D^{1/k}) = \begin{cases} \left(\sum_{\ell=1}^{L} (\log p_\ell)^t\right)^{1/t} & \text{if } t \in [1, \infty) \\ \max_{1 \leqslant \ell \leqslant L} \{\log p_\ell\} & \text{if } t = \infty. \end{cases}$$

As an example, for $D = 30 = 2 \cdot 3 \cdot 5$, Theorem 10.4 asserts that $m_t(\sqrt{30})$ is attained by $\sqrt{2}, \sqrt{3}, \sqrt{5}$, and

$$m_t(\sqrt{30})^t = (\log 5)^t + (\log 3)^t + (\log 2)^t.$$

While it is obvious that $\sqrt{2}, \sqrt{3}, \sqrt{5} \notin \mathbb{Q}(\sqrt{30})$, the infimum in $m_t(\sqrt{30})$ might be attained by some distinct set of points in $\mathbb{Q}(\sqrt{30})$. Theorem 10.3 excludes any such possibilities.

If we take $D = 42 = 2 \cdot 3 \cdot 7$, Theorem 10.4 establishes that $m_t(\sqrt{42})$ is attained by $\sqrt{2}, \sqrt{3}, \sqrt{7}$, and

$$m_t(\sqrt{42})^t = (\log 7)^t + (\log 3)^t + (\log 2)^t.$$

Nonetheless, Theorem 10.3 identifies the slightly more subtle points $\sqrt{7/6}, 3, 2 \in \mathbb{Q}(\sqrt{42})$ that also attain the infimum in $m_t(\sqrt{42})$. In this example, we note that the infimum is not attained by a unique set.

At first glance, one might think that the infimum in $m_t(\sqrt{D})$ can be attained only by rational numbers and their square roots. This intuition is misleading, however, as we see in the following example. Let $t = \infty$ and take $D = 21 = 7 \cdot 3$. We know from Theorem 10.3 that the infimum in $m_t(\sqrt{21})$ is attained by the points $\sqrt{7/3}, 3 \in \mathbb{Q}(\sqrt{21})$ and

$$m_t(\sqrt{21})^t = (\log 7)^t + (\log 3)^t.$$

Now consider

$$\sqrt{21} = (-1) \cdot \left(\frac{7 + \sqrt{21}}{2}\right) \cdot \left(\frac{3 - \sqrt{21}}{2}\right), \tag{10.2}$$

and we verify easily that

$$m\left(\frac{7 + \sqrt{21}}{2}\right) = \log 7 \quad \text{and} \quad m\left(\frac{3 - \sqrt{21}}{2}\right) < \log 7.$$

In other words, $m_\infty(\sqrt{21})$ is attained by the points on the right hand side of (10.2), and these points belong to $\mathbb{Q}(\sqrt{21})$. It is important to note that $m\left((3 - \sqrt{21})/2\right)$ is greater than $\log 3$, so these points cannot be used to attain the infimum in $m_t(\sqrt{21})$ for other values of t. Nonetheless, this example illustrates that the infimum in $m_t(\sqrt{D})$ may be attained by using distinct non-trivial sets of points contained in $\mathbb{Q}(\sqrt{D})$.

We would like to conclude with the following question.

Question 10.5. *Is the infimum in $m_t(\alpha)$ always attained by points $\alpha_1, \ldots, \alpha_N$*

such that $[\mathbb{Q}(\alpha_n) : \mathbb{Q}] \leqslant [\mathbb{Q}(\alpha) : \mathbb{Q}]$ *for all* n*?*

According to Theorem 10.4, the answer is 'yes' when α is a surd, although we know of little other evidence.

10.2 The rational case

Recall that the *(logarithmic) Weil height* of an algebraic number α is given by

$$h(\alpha) = \frac{m(\alpha)}{\deg \alpha}.$$

It is well-known that if ζ is a root of unity, then $h(\alpha) = h(\zeta\alpha)$ so that h is well-defined on our quotient group G. Furthermore, if n is an integer, then we have that $h(\alpha^n) = |n| \cdot h(\alpha)$. Also recall that a surd is an algebraic number α such that $\alpha^n \in \mathbb{Q}$ for some positive integer n.

Suppose now that F is any number field containing the algebraic number α. Further assume that K is an extension of F which is Galois over $\mathbb{Q}$. We set

$$G = \mathrm{Gal}(K/\mathbb{Q}) \quad \text{and} \quad H = \mathrm{Gal}(K/F),$$

and let S be a set of left coset representatives of H in G. Recall that the *norm of* α *from* F *to* $\mathbb{Q}$ is given by

$$\mathrm{Norm}_{F/\mathbb{Q}}(\alpha) = \prod_{\sigma \in S} \sigma(\alpha).$$

It follows from standard Galois Theory that $\mathrm{Norm}_{F/\mathbb{Q}}$ is a homomorphism from F to $\mathbb{Q}$ which does not depend on the choice of K or S. In addition, if E is any extension of F, then it is easily verified that

$$\mathrm{Norm}_{E/\mathbb{Q}}(\alpha) = \mathrm{Norm}_{F/\mathbb{Q}}(\alpha)^{[E:F]}. \tag{10.3}$$

We begin our proof of Theorem 10.2 with a lemma that relates the Mahler measure of a surd to the Mahler measure of its norm.

Lemma 10.6. *If* γ *is a surd then* $m(\gamma) = m\left(\mathrm{Norm}_{\mathbb{Q}(\gamma)/\mathbb{Q}}(\gamma)\right)$.

Proof. Since γ is a surd, its conjugates over $\mathbb{Q}$ are given by

$$\{\zeta_1\gamma, \zeta_2\gamma, \ldots, \zeta_d\gamma\}$$

where $d = \deg \gamma$ and ζ_j are roots of unity. It now follows that

$$\gamma^d \prod_{j=1}^{d} \zeta_j = \mathrm{Norm}_{\mathbb{Q}(\gamma)/\mathbb{Q}}(\gamma) \in \mathbb{Q}.$$

Since $\mathrm{Norm}_{\mathbb{Q}(\gamma)/\mathbb{Q}}(\gamma)$ is clearly a rational number, we have that

$$m\left(\mathrm{Norm}_{\mathbb{Q}(\gamma)/\mathbb{Q}}(\gamma)\right) = h\left(\gamma^d \prod_{j=1}^{d} \zeta_j\right) = d \cdot h(\gamma) = \deg \gamma \cdot h(\gamma) = m(\gamma)$$

completing the proof. □

In our proof of Theorem 10.2, it will be necessary to replace an arbitrary representation $\alpha = \alpha_1 \cdots \alpha_N$ with another representation of $\alpha = \beta_1 \cdots \beta_N$ that uses only rational numbers and satisfies

$$\sum_{n=1}^{N} m(\beta_n)^t \leqslant \sum_{n=1}^{N} m(\alpha_n)^t.$$

Our next lemma provides us with the necessary elementary number theoretic tools to do this.

Lemma 10.7. *Suppose that $m, r_1, \ldots, r_N$ are positive integers such that*

$$m \mid \prod_{n=1}^{N} r_n.$$

For $1 \leqslant n \leqslant N$, recursively define the points m_n by

$$m_1 = \gcd(r_1, m) \quad \text{and} \quad m_n = \gcd\left(r_n, \frac{m}{\prod_{i=1}^{n-1} m_i}\right). \tag{10.4}$$

Then we have that

$$m = \prod_{n=1}^{N} m_n.$$

Before we provide the proof of Lemma 10.7, we make one clarification regarding the definition of m_n. Naively, it would appear that

$$\frac{m}{\prod_{i=1}^{n-1} m_i}$$

is not necessarily an integer, so that taking its greatest common divisor with another integer might not be well-defined. However, we note immediately that $m_1 \mid m$, which also implies that m_2 is well-defined. Then clearly we have that $m_2 \mid m/m_1$ implying that m_3 is also well-defined. As we can see, it follows

inductively that

$$m_n \mid \frac{m}{\prod_{i=1}^{n-1} m_i}$$

for all $1 \leqslant n \leqslant N$, meaning, in particular, that m_n is well-defined for all such n. Now we may proceed with the proof of Lemma 10.7.

Proof of Lemma 10.7. We will assume that $m \neq \prod_{n=1}^{N} m_n$ and find a contradiction. Since the product $\prod_{n=1}^{N} m_n$ divides m, there must exist a prime number p for which

$$\nu_p(m) > \sum_{j=1}^{N} \nu_p(m_j), \tag{10.5}$$

where $\nu_p(x)$ denotes the highest power of p dividing the integer x. It now follows that

$$\nu_p(m_n) < \nu_p(m) - \sum_{j=1}^{n-1} \nu_p(m_j) = \nu_p\left(\frac{m}{\prod_{j=1}^{n-1} m_j}\right)$$

for every $n \in \{1, \ldots, N\}$. Hence, the definition of m_n implies that

$$\nu_p(m_n) = \min\left\{\nu_p(r_n), \nu_p\left(\frac{m}{\prod_{j=1}^{n-1} m_j}\right)\right\} = \nu_p(r_n)$$

for every $n \in \{1, \ldots, N\}$. It now follows from (10.5) that $\nu_p(m) > \sum_{n=1}^{N} \nu_p(r_n)$, contradicting our assumption that m divides $\prod_{n=1}^{N} r_n$. □

Now that we have established our key lemmas, we may now proceed with the proof of Theorem 10.2.

Proof of Theorem 10.2. As we have noted in the introduction, the case $t = \infty$ is known [93], so we proceed immediately to the situation where $1 \leqslant t < \infty$.

We may assume without loss of generality that $\alpha > 0$. Since α is rational, there exist positive integers m and m' such that $\gcd(m, m') = 1$ and $\alpha = m/m'$. Furthermore, by the results of [176], there exist surds $\alpha_1, \ldots, \alpha_N$ such that

$$\alpha = \alpha_1 \cdots \alpha_N \quad \text{and} \quad m_t(\alpha)^t = \sum_{n=1}^{N} m(\alpha_n)^t. \tag{10.6}$$

Let K be a number field containing $\alpha_1, \ldots \alpha_N$. Now we may take the norm from K to $\mathbb{Q}$ of both sides of the first equation in (10.6). We apply (10.3) and the fact that the $\mathrm{Norm}_{K/\mathbb{Q}}$ is a homomorphism to establish that

$$\left(\frac{m}{m'}\right)^{[K:\mathbb{Q}]} = \prod_{n=1}^{N} \mathrm{Norm}_{K/\mathbb{Q}}(\alpha_n) = \prod_{n=1}^{N} \left(\mathrm{Norm}_{\mathbb{Q}(\alpha_n)/\mathbb{Q}}(\alpha_n)\right)^{[K:\mathbb{Q}(\alpha_n)]}.$$

Suppose further that, for each $1 \leqslant n \leqslant N$, r_n and s_n are relatively prime positive

integers such that

$$\frac{r_n}{s_n} = \pm \mathrm{Norm}_{\mathbb{Q}(\alpha_n)/\mathbb{Q}}(\alpha_n).$$

Therefore, we have that

$$\left(\frac{m}{m'}\right)^{[K:\mathbb{Q}]} = \pm \prod_{n=1}^{N} \left(\frac{r_n}{s_n}\right)^{[K:\mathbb{Q}(\alpha_n)]}.$$

It is obvious that $[K:\mathbb{Q}(\alpha_n)] \mid [K:\mathbb{Q}]$ so we obtain that

$$m^{[K:\mathbb{Q}]} \mid \left(\prod_{n=1}^{N} r_n\right)^{[K:\mathbb{Q}]} \quad \text{and} \quad m'^{[K:\mathbb{Q}]} \mid \left(\prod_{n=1}^{N} s_n\right)^{[K:\mathbb{Q}]}.$$

It follows from elementary number theory facts that

$$m \mid \prod_{n=1}^{N} r_n \quad \text{and} \quad m' \mid \prod_{n=1}^{N} s_n. \tag{10.7}$$

Setting up the hypotheses of Lemma 10.7, we define recursive sequences corresponding to m and m'. First set

$$m_1 = \gcd(r_1, m) \quad \text{and} \quad m_n = \gcd\left(r_n, \frac{m}{\prod_{i=1}^{n-1} m_i}\right)$$

and

$$m'_1 = \gcd(s_1, m') \quad \text{and} \quad m'_n = \gcd\left(s_n, \frac{m'}{\prod_{i=1}^{n-1} m'_i}\right)$$

so we clearly have that

$$|r_n| \geqslant |m_n| \quad \text{and} \quad |s_n| \geqslant |m'_n|. \tag{10.8}$$

Applying Lemma 10.7, we have that

$$m = \prod_{n=1}^{N} m_n \quad \text{and} \quad m' = \prod_{n=1}^{N} m'_n$$

so that

$$\alpha = \frac{m}{m'} = \prod_{n=1}^{N} \frac{m_n}{m'_n}. \tag{10.9}$$

Now it follows from the definition of $m_t(\alpha)$ that

$$m_t(\alpha)^t \leqslant \sum_{n=1}^{N} m \left(\frac{m_n}{m'_n}\right)^t, \tag{10.10}$$

so we must show that the right hand side of (10.10) is also a lower bound for

$m_t(\alpha)^t$.

To see this, note that by Lemma 10.6, we have that

$$m(\alpha_n) = m\left(\mathrm{Norm}_{\mathbb{Q}(\alpha_n)/\mathbb{Q}}(\alpha_n)\right) = m\left(\frac{r_n}{s_n}\right)$$

for all $1 \leqslant n \leqslant N$. We have assumed that r_n and s_n are relatively prime, so it follows from known facts about the Mahler measure that

$$m(\alpha_n) = \log \max\{|r_n|, |s_n|\}.$$

Then applying (10.8), we find that

$$m(\alpha_n) \geqslant \log \max\{|m_n|, |m_n'|\} \geqslant m\left(\frac{m_n}{m_n'}\right),$$

and consequently,

$$m_t(\alpha)^t = \sum_{n=1}^{N} m(\alpha_n)^t \geqslant \sum_{n=1}^{N} m\left(\frac{m_n}{m_n'}\right)^t.$$

Combining this with (10.9) and (10.10), the result follows.

□

10.3 The quadratic case

Our first lemma gives one particular set of points that attain the infimum in $m_t(\sqrt{D})$ for all $t \in [1, \infty]$. When $t > 1$, we can also identify the Mahler measures of any points $\alpha_1, \ldots, \alpha_N$ attaining the infimum in $m_t(D^{1/k})$.

Lemma 10.8. *Suppose that $p_1, \ldots, p_L$ are distinct primes written in decreasing order, $D = p_1 \cdots p_L$, $t \in [1, \infty)$, and $k \in \mathbb{N}$. The infimum in $m_t(D^{1/k})$ is attained by $p_1^{1/k}, \ldots, p_L^{1/k}$ and*

$$m_t(D^{1/k})^t = \sum_{\ell=1}^{L} (\log p_\ell)^t.$$

If $t > 1$ and $\alpha_1, \cdots, \alpha_N$ are algebraic numbers attaining the infimum in $m_t(D^{1/k})$ then $N \geqslant L$. Moreover, it is possible to relabel the elements $\alpha_1, \ldots, \alpha_N$ so that

(i) $m(\alpha_n) = \log p_n$ for all $n \leqslant L$, and

(ii) $m(\alpha_n) = 0$ for all $n > L$.

In particular, $m(\alpha_n) \leqslant \log p_1$ for all n.

Proof. We certainly have that $D^{1/k} = p_1^{1/k} \cdots p_\ell^{1/k}$, and by the definition of m_t, we know that

$$m_t(D^{1/k})^t \leqslant \sum_{\ell=1}^{L} m(p_\ell^{1/k})^t.$$

For each ℓ, we know that $x^k - p_\ell$ vanishes at $p_\ell^{1/k}$ and is irreducible by Eisenstein's criterion, so that $m(p_\ell^{1/k}) = m(x^k - p_\ell) = \log p_\ell$. Hence, we find that

$$m_t(D^{1/k})^t \leqslant \sum_{\ell=1}^{L} (\log p_\ell)^t. \tag{10.11}$$

To prove the first statement of the lemma, it is now sufficient to show that

$$m_t(D^{1/k})^t \geqslant \sum_{\ell=1}^{L} (\log p_\ell)^t. \tag{10.12}$$

Now suppose $\alpha_1, \ldots, \alpha_N \in \overline{\mathbb{Q}}$ attain the infimum in $m_t(D^{1/k})$ and select a number field K containing $D^{1/k}, \alpha_1, \ldots, \alpha_N$. By definition, we know that $D^{1/k} = \alpha_1 \cdots \alpha_N$. Using the fact that $\mathrm{Norm}_{K/\mathbb{Q}}$ is a multiplicative homomorphism, we obtain that

$$\mathrm{Norm}_{K/\mathbb{Q}}(D^{1/k}) = \prod_{n=1}^{N} \mathrm{Norm}_{K/\mathbb{Q}}(\alpha_N)$$

so that

$$\left(\mathrm{Norm}_{\mathbb{Q}(D^{1/k})/\mathbb{Q}}(D^{1/k})\right)^{[K:\mathbb{Q}(D^{1/k})]} = \prod_{n=1}^{N} \left(\mathrm{Norm}_{\mathbb{Q}(\alpha_n)/\mathbb{Q}}(\alpha_N)\right)^{[K:\mathbb{Q}(\alpha_n)]}. \tag{10.13}$$

Each of the above norms is a rational number. Hence, for each n, there exist positive relatively prime integers r_n and s_n such that

$$|\mathrm{Norm}_{\mathbb{Q}(\alpha_n)/\mathbb{Q}}(\alpha_n)| = \frac{r_n}{s_n}.$$

Again using Eisenstein's Criterion, we know that $x^k - D$ is the minimal polynomial of $D^{1/k}$ over $\mathbb{Q}$, implying that $|\mathrm{Norm}_{\mathbb{Q}(D^{1/k})/\mathbb{Q}}(D^{1/k})| = D$. Substituting these values into (10.13), we find that

$$D^{[K:\mathbb{Q}(D^{1/k})]} = \prod_{n=1}^{N} \left(\frac{r_n}{s_n}\right)^{[K:\mathbb{Q}(\alpha_n)]}. \tag{10.14}$$

For each n, α_n has minimal polynomial of the form

$$\hat{f}_n(x) = x^d + \frac{a_{d-1}}{b_{d-1}} x^{d-1} + \cdots + \frac{a_1}{b_1} x \pm \frac{r_n}{s_n}$$

over $\mathbb{Q}$ for integers $a_1, \dots, a_{d-1}, b_1, \dots, b_{d-1}$ with $b_i \neq 0$ and $(a_i, b_i) = 1$. Hence, its minimal polynomial over $\mathbb{Z}$ is given by

$$f_n(x) = \operatorname{lcm}(s_n, b_{d-1}, \dots, b_1) \cdot x^d + \cdots \pm r_n \cdot \operatorname{lcm}(s_n, b_{d-1}, \dots, b_1)$$

and its Mahler measure satisfies

$$m(\alpha_n) \geqslant \log\left(\frac{r_n}{s_n} \cdot \operatorname{lcm}(s_n, b_{d-1}, \dots, b_1)\right) \geqslant \log r_n.$$

For each n, let

$$P_n = \{p \in \{p_1, \dots, p_L\} : p \mid r_n\}.$$

We have assumed that $\alpha_1, \dots, \alpha_N$ attains the infimum in $m_t(D^{1/k})$, so we get that

$$m_t(D^{1/k})^t = \sum_{n=1}^{N} m(\alpha_n)^t \geqslant \sum_{n=1}^{N} (\log r_n)^t \geqslant \sum_{n=1}^{N} \left(\sum_{p \in P_n} \log p\right)^t. \tag{10.15}$$

Since $t \geqslant 1$, we always have that

$$\left(\sum_{p \in P_n} \log p\right)^t \geqslant \sum_{p \in P_n} (\log p)^t, \tag{10.16}$$

which implies that

$$m_t(D^{1/k})^t \geqslant \sum_{n=1}^{N} \sum_{p \in P_n} (\log p)^t.$$

However, applying (10.14), we know that for each $\ell \in \{1, \dots, L\}$, there exists $n \in \{1, \dots, N\}$ such that $p_\ell \in P_n$, establishing (10.12) and the first statement of the lemma.

Now assume that $t > 1$. If $|P_n| \geqslant 2$, then we must have strict inequality in (10.16). Therefore, if $|P_n| \geqslant 2$ for some n, then (10.15) implies that

$$m_t(D^{1/k})^t > \sum_{n=1}^{N} \sum_{p \in P_n} (\log p)^t \geqslant \sum_{\ell=1}^{L} (\log p_\ell)^t$$

contradicting (10.11). Therefore, $|P_n| \leqslant 1$ for every n and we have established that

(a) For every ℓ, there exists n such that $p_\ell \mid r_n$, and

(b) If $\ell_1 \neq \ell_2$ then we can never have that $p_{\ell_1} \mid r_n$ and $p_{\ell 2} \mid r_n$.

It follows from the box principle that $N \geqslant L$. Moreover, we may reorder the

numbers $\alpha_1, \ldots, \alpha_N$ so that $p_n \mid r_n$ for all $1 \leqslant n \leqslant L$, which shows that

$$m(\alpha_n) \geqslant \log r_n \geqslant \log p_n \quad \text{for } 1 \leqslant n \leqslant L. \tag{10.17}$$

If we have strict inequality in (10.17) for some n, then

$$m_t(D^{1/k})^t = \sum_{n=1}^{N} m(\alpha_n)^t > \sum_{\ell=1}^{L} (\log p_\ell)^t \tag{10.18}$$

contradicting (10.11) and establishing (i). Similarly, if $m(\alpha_n) > 0$ for some $n > L$, then (10.18) holds as well verifying (ii).

□

Now that we have proven Lemma 10.8, the proof of Theorem 10.4 is essentially complete. Indeed, when $t \in [1, \infty)$ Theorem 10.4 is simply the first statement of Lemma 10.8, and the case $t = \infty$ was given already in [93]. The only task remaining is to prove Theorem 10.3, in which the second statement of Lemma 10.8 plays a key role.

Before proceeding, we establish some conventions that will be used for the remainder of this article. For $d \in \mathbb{Z}$ and $r \in \mathbb{Q}$, we say that d *divides* $r = m/n$ with $m \in \mathbb{N}$, $n \in \mathbb{Z} \setminus \{0\}$ and $(m, n) = 1$, if $d \mid m$ or $d \mid n$. We say that d *divides the numerator or denominator of* r if d divides m or n, respectively.

We say that an algebraic number α is *stable* if all of its conjugates lie either inside the open unit disk, on the unit circle, or outside the closed unit disk. Otherwise, we say that α is *unstable*. It is clear that all rational numbers and all imaginary quadratic numbers are stable, while real quadratic numbers can be either stable or unstable. If α is any algebraic number having minimal polynomial

$$f(x) = a_N x^N + \cdots + a_1 x + a_0,$$

then it is simple to verify that

$$m(\alpha) \geqslant \log \max\{|a_N|, |a_0|\}$$

with equality if and only if α is stable. We now state a simple criterion which allows us to determine if a quadratic algebraic number is stable by considering the coefficients of the minimal polynomial.

Lemma 10.9. *Suppose that α is a quadratic algebraic number having minimal polynomial $f(x) = ax^2 + bx + c$ over $\mathbb{Z}$. We have that α is stable if and only if $|a + c| > |b|$. In this situation, the following hold.*

(i) If $|a| < |c|$ then both conjugates of α have modulus greater than one.

(ii) If $|a| = |c|$ then both conjugates of α have modulus one.

(iii) If $|a| > |c|$ then both conjugates of α have modulus less than one.

Proof. Suppose that $f(x) = a(x-\alpha)(x-\beta)$. If $f(1)$ and $f(-1)$ have opposite signs, then f has precisely one root in the interval $(-1,1)$. The other root must also be real and lie outside of $(-1,1)$, so α is unstable. If $f(1)$ and $f(-1)$ have the same sign, then f has either zero or two roots in $(-1,1)$. In the case of two roots in $(-1,1)$, α is clearly stable. If f has zero roots in $(-1,1)$, then it either has two complex roots, in which case α is certainly stable, or two real roots both lying outside of $[-1,1]$, also implying that α is stable.

We have now shown that α is stable if and only if $f(1) = a+b+c$ and $f(-1) = a-b+c$ have the same sign. Clearly, $f(1)$ and $f(-1)$ are both positive if and only if $a+c > |b|$ and both negative if and only if $-a-c > |b|$. Thus, α is stable if and only if $|a+c| > |b|$.

If, in addition, $|a| < |c|$, then $|\alpha\beta| = |c|/|a| > 1$, so both α and β have modulus greather than 1. Similarly, if $|a| > |c|$, then $|\alpha\beta| = |c|/|a| < 1$ implying that both α and β have modulus less than 1. Finally, if $|a| = |c|$, then $|\alpha\beta| = 1$. Since α is stable, α and β must be complex conjugate numbers both of modulus 1. □

The following lemma shows us that certain quadratic algebraic numbers, which we will encounter in the proof of Theorem 10.3, have relatively simple minimal polynomials.

Lemma 10.10. *Let D be a square-free integer, p be a prime divisor of D, and suppose that α a quadratic algebraic number in $\mathbb{Q}(\sqrt{D})$. If $m(\alpha) \leqslant \log p$ and p divides the numerator of $\mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha)$, then α is stable. Moreover, the minimal polynomial of α satisfies*

$$f(x) = ax^2 \pm p \quad \text{or} \quad f(x) = ax^2 \pm px + p,$$

where a is a positive integer with $a < p$.

Proof. Suppose that $f(x) = ax^2 + bx + c \in \mathbb{Z}[x]$ is the minimal polynomial of α over $\mathbb{Z}$, so we may assume that $a > 0$. Since α has degree 2, we have that

$$\mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha) = \alpha\bar{\alpha} = \frac{c}{a},$$

where $\bar{\alpha}$ is the conjugate of α over $\mathbb{Q}$. We have assumed that p divides the numerator of $\mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha)$, which itself must divide c, implying that $p \mid c$.

Since $m(\alpha) \geqslant \log \max\{|a|, |c|\}$, we have that

$$\log p \leqslant \log |c| \leqslant \log \max\{|a|, |c|\} \leqslant m(\alpha) \leqslant \log p,$$

and we conclude that

$$m(\alpha) = \log |c| = \log p. \tag{10.19}$$

It now follows that $|a| \leqslant |c|$ and, since $m(\alpha)$ is the log of an integer, we further obtain that α is stable. Hence, Lemma 10.9 implies that $|a + c| > |b|$.

We cannot have $|a| = |c|$, since

$$\mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha) = \frac{c}{a} = \pm 1$$

is not divisible by p, so it follows that $|a| < |c|$. In view of Lemma 10.9 (i), we have that $|\alpha|, |\bar{\alpha}| > 1$. Therefore, we find that

$$|b| < |a + c| \leqslant |a| + |c| < 2|c| = 2p. \tag{10.20}$$

Now let $\Delta = b^2 - 4ac$. Since $\mathbb{Q}(\sqrt{\Delta}) = \mathbb{Q}(\sqrt{D})$, and D is square-free, we have $\Delta = Dv^2$ for some $v \in \mathbb{Z}$. The quadratic formula gives

$$\mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha) = \alpha\bar{\alpha} = \frac{b^2 - Dv^2}{4a^2},$$

and since $p \mid D$ and the numerator of $\mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha)$, it follows that p divides b^2. Of course, this implies that $p \mid b$. Using (10.20), we now see that $b \in \{0, p, -p\}$.

If $b = 0$, then we have by (10.19) that $f(x) = ax^2 \pm p$, establishing the lemma in this case. If $b = \pm p$, then $|a + c| > |b|$ holds if and only if a and c have the same sign. So in this situation, (10.19) yields that $c = p$ which leads to $f(x) = ax^2 \pm px + p$. □

Proof of Theorem 10.3. By Theorem 10.4, we know that

$$m_t(\sqrt{D}) = \begin{cases} \left(\sum_{\ell=1}^{L} (\log p_\ell)^t\right)^{1/t} & \text{if } t \in (1, \infty) \\ \log p_1 & \text{if } t = \infty. \end{cases}$$

We also observe that

$$\sqrt{D} = \sqrt{\frac{p_1}{p_2 \cdots p_L}} \cdot p_2 \cdots p_L \tag{10.21}$$

and that each term in the product on the right hand side of (10.21) belongs to $\mathbb{Q}(\sqrt{D})$. We obviously have that $m(p_\ell) = \log p_\ell$ for all ℓ. Furthermore, our

assumption that $D < p_1^2$ ensures that $p_2 \cdots p_L < p_1$, so it follows that

$$m\left(\sqrt{\frac{p_1}{p_2 \cdots p_L}}\right) = \log p_1.$$

Combining these observations, we see that

$$m\left(\sqrt{\frac{p_1}{p_2 \cdots p_L}}\right)^t + \sum_{\ell=2}^{L} m(p_\ell)^t = \sum_{\ell=1}^{L} (\log p_\ell)^t = m_t(\sqrt{D})^t$$

when $1 < t < \infty$ and

$$\max\left\{m\left(\sqrt{\frac{p_1}{p_2 \cdots p_L}}\right), m(p_2), \ldots, m(p_L)\right\} = \log p_1 = m_\infty(\sqrt{D}).$$

establishing one direction of the theorem as well as the second statement.

To prove the other direction, we assume that there exist points $\alpha_1, \ldots, \alpha_N \in \mathbb{Q}(\sqrt{D})$ that attain the infimum in $m_t(\sqrt{D})$, and for simplicity, we set $p = p_1$. When $t \in (1, \infty)$, Lemma 10.8 establishes that that $m(\alpha_n) \leqslant \log p$ for all n. In the case $t = \infty$, we also have $m(\alpha_n) \leqslant \log p$ for all n as a consequence of Theorem 10.4. Since $\sqrt{D} = \alpha_1 \cdots \alpha_N$, we have that

$$-D = \mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\sqrt{D}) = \prod_{n=1}^{N} \mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha_n). \tag{10.22}$$

Defining the set

$$\Lambda = \left\{1 \leqslant n \leqslant N : p \mid \mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha_n)\right\}$$

we apply (10.22) to see that

$$\sum_{n \in \Lambda} \nu_p\left(\mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha_n)\right) = \nu_p(D) = 1, \tag{10.23}$$

where the last equality follows since D is square-free. If Λ contains no irrational points, then we have that

$$p \mid \mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha_n) = \alpha_n^2$$

for all $n \in \Lambda$. However, this implies that $\nu_p(\mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha_n))$ is even for all $n \in \Lambda$. It follows that the left hand side of (10.23) is also even, a contradiction.

We have shown that there must exist n such that α_n is quadratic, $m(\alpha_n) \leqslant \log p$, and p divides $\mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha_n)$. If p divides the numerator of the rational number $\mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha_n)$, then we may apply Lemma 10.10 to see that α_n is

stable and is a root of

$$f(x) = ax^2 \pm p \quad \text{or} \quad f(x) = ax^2 \pm px + p$$

for some positive integer $a < p$.

Suppose now that Δ is the discriminant of f. Since α_n is quadratic over $\mathbb{Q}$, we have $\mathbb{Q}(\sqrt{\Delta}) = \mathbb{Q}(\sqrt{D})$. Furthermore, since D is a square-free, we have that $\Delta = Dv^2$ for some $v \in \mathbb{N}$. If $f(x) = ax^2 \pm p$, we see that $\Delta = \pm 4ap$, so that

$$pp_2 \cdots p_L v^2 = Dv^2 = \pm 4ap.$$

Since $p_2, \ldots, p_L$ are distinct primes, we obtain that $p_2 \cdots p_L \mid a$, and hence,

$$p_2 \cdots p_L \leqslant a < p,$$

establishing that $D < p^2$ in this case.

If $f(x) = ax^2 \pm px + p$ then $\Delta = p^2 - 4ap = p(p-4a)$. We have assume that D is positive so that $p - 4a > 0$, and, trivially, $p - 4a < p$. Hence, $D \leqslant \Delta = p(p-4a) < p^2$ completing the proof when p divides the numerator of $\mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha_n)$.

If p divides the denominator of $\mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha_n)$ instead, the p must divide the numerator of $\mathrm{Norm}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}(\alpha_n^{-1})$. Of course, we also have that $m(\alpha_n^{-1}) \leqslant \log p$ and $\alpha_n^{-1} \in \mathbb{Q}(\sqrt{D})$ is quadratic, so we may apply the above argument to α_n^{-1} in place of α_n. $\square$

Chapter 11

Barker sequences and polynomials

11.1 Statement of the problem

Let us recall some definitions from Section 2.4 of Chapter 2. A finite sequence $a_0, a_1, \ldots, a_n$ of complex numbers is called a *binary sequence*, if all numbers a_j are equal to 1 or -1. For this sequence, the *aperiodic autocorrelation* coefficients $c_k, -n \leq k \leq n$ are defined by

$$c_k := \sum_{j=0}^{n-k} a_j a_{k+j} \quad \text{for} \quad 0 \leq k \leq n \quad \text{and} \quad c_{-k} := c_k. \tag{11.1}$$

Binary (and more generally, unimodular) sequences a_k having small autocorrelation values $|c_k|$ have a long history in signal processing, see [17], [85], [36], and [107]. In Chapter 11, we shall consider binary sequences with autocorrelation coefficients c_k, $k \neq 0$ equal either to 0, -1 or 1. These sequences were introduced by Barker [17] in 1953 and are called *Barker* sequences. Thus Barker sequences are the binary sequences which possess minimal possible (in absolute value) autocorrelations. Since we enumerate the terms in the sequence starting with 0, the *length* of the sequence is equal to $n + 1$ – we shall keep this convention in mind through the entire Chapter 11. All the currently known Barker sequences [210], [40], are summarized in Table 11.1 bellow (sequences which can be obtained from these in the table by negating or rewriting them backwards are not included).

Clearly, all the restrictions to which Barker sequences are subjected seem to be quite special and hard to satisfy. Because of this it is widely believed that only finitely many Barker sequences exist. Turyn and Storer [210] proved that no Barker sequences of *odd length* exist for $n \geq 13$. Thus, by the result of Turyn and Storer [210], all the Barker sequences which are longer than 12 should have *even*

Length	Sequence
2	$++$
3	$++-$
4	$++-+$ and $+++-$
5	$+++-+$
7	$+++--+-$
11	$+++---+--+-$
13	$+++++--++-+-+$

Table 11.1: Currently known Barker sequences

length. However, all the known Barker sequences of even length are very short:

$$n = 1:$$

$$++, \quad --, \quad +-, \quad -+,$$

$$n = 3:$$

$$++-+, \quad --+-, \quad +---, \quad -+++,$$

$$+++-, \quad ---+, \quad +-++, \quad -+--.$$

In view of this situation, one would expect to find a simple and concise proof of the following conjecture:

Conjecture 11.1. *There exist no Barker sequences of even length* > 4.

However, the proof of the non-existence of long Barker sequences still remains elusive despite a substantial amount of research in the last 45 years. This problem has been attacked by combinatorial and number theoretical methods. We remark that various restrictions on the possible values of n were derived by Eliahou, Kervaire, Saffari [77], [78], Jedwab and Lloyd [108], Leung and Schmidt [131] and Turyn [207], [208], [209]. These restrictions were used to check the nonexistence of Barker sequences on computers for very large values of n. The current computation record belongs to Mossinghoff [150], who showed that if a Barker sequence of even length exists, then either $n = 189260468001034441522766781604$ or $n > 2 \cdot 10^{30}$. All the mentioned results provide a strong evidence in support of Conjecture 11.1.

In recent literature [40] and [174], the question of the existence of long Barker sequences was tied to the existence of polynomials with small integer coefficients $\{-1, 1\}$ having remarkable analytic properties. In particular, the polynomials constructed by means of long Barker sequences are thought to have extremely large Mahler measures and L^s norms on the unit circle — which seems to be

unlikely. The questions about the existence of such extremal polynomials go back to Littlewood [135], [136], Mahler [141] and Erdős [82] themselves and have been open for half a century now.

In the present Chapter we will also focus on the polynomial setting. In Section 11.2, we recall the definition of Littlewood polynomials and define Barker polynomials. We will explain in detail the relation between the Barker conjecture and the conjectures on extremal Mahler measures and L^s norms of polynomials. We will state *two problems* which imply the Barker conjecture. Finally (and most importantly), we will solve *one* of these two problems in Section 11.3.

11.2 Polynomials on the unit circle

11.2.1 Barker polynomials

Recall that a polynomial $p(z) \in \mathbb{Z}[z]$ is called a Littlewood polynomial if all coefficients of $p(z)$ are equal to 1 or -1. The set of Littlewood polynomials is denoted by

$$\mathcal{L}_n := \{p(z) = a_n z^n + \cdots + a_0 : a_j = 1 \text{ or } a_j = -1,\ 0 \le j \le n\}.$$

Littlewood polynomials are very convenient in signal analysis to study properties of binary sequences and Barker sequences in particular.

The polynomial $p(z) \in \mathcal{L}_n$ is called a *Barker polynomial* if the coefficients $a_0, a_1, \cdots, a_n$ form a Barker sequence of length $n+1$. Since autocorrelations of the coefficients do not change in magnitude by replacing the polynomial $p(z)$ with $p(-z)$ or $-p(-z)$, one can normalize Barker polynomials using the conditions $a_n = a_{n-1} = 1$. Also, if $p(z)$ is a Barker polynomial, then the *reciprocal* polynomial $p^*(z) = z^n p(1/z)$ is also a Barker polynomial. With our convention in mind, the Barker sequences of *even* length $n+1$ correspond to the Barker polynomials of *odd* degree n. Similarly, the Barker sequences of *odd* length $n+1$ correspond to the Barker polynomials of *even* degree n.

By the results of [210], there exist precisely five normalized Barker polynomials of even degree n, namely: z^2+z-1, $z^4+z^3+z^2-z+1$, $z^6+z^5+z^4-z^3-z^2+z-1$, $z^{10}+z^9+z^8-z^7-z^6-z^5+z^4-z^3-z^2+z-1$, and $z^{12}+z^{11}+z^{10}+z^9+z^8-z^7-z^6+z^5+z^4-z^3+z^2-z+1$. According to Conjecture 11.1, the only two normalized Barker polynomials of odd degree are the polynomials $z+1$ and z^3+z^2+z-1.

11.2.2 L^s norms

Let us recall some definitions. Let $P(z) = a_n(z-\alpha_1)(z-\alpha_2)\cdots(z-\alpha_n) \in \mathbb{C}[z]$ be a complex polynomial. The *Mahler measure* of $P(z)$ is defined by

$$M(P) = |a| \prod_{j=1}^{n} \max\{1, |\alpha_j|\}.$$

In view of Jensen's formula [139], one has

$$\log M(P) = \frac{1}{2\pi} \int_0^{2\pi} \log |P(e^{it})| dt.$$

More generally, for a real number $s > 0$, the integral L^s norm of a complex polynomial is defined by the formula

$$\|P\|_s = \left(\frac{1}{2\pi} \int_0^{2\pi} |P(e^{it})|^s dt \right)^{1/s}.$$

We should make a note here that $\|\cdot\|_s$ is a *norm* in $\mathbb{C}[z]$ (as a vector space over $\mathbb{C}$) only if $s \geq 1$. L^s norms have the *continuity property*: if the polynomial $P(z)$ is fixed, then the norm $||P||_s$ is a monotonically increasing and continuous function of s. Moreover,

$$\lim_{s\to 0^+} ||P||_s = M(P), \qquad \lim_{s\to\infty} ||P||_s = ||P||_\infty = \sup_{|z|=1} |P(z)|.$$

We remark that, in general, the computation of Mahler measures or L^s norms of an arbitrary complex polynomial is a hard problem. For practical computations, the most useful is the L^2 norm. By Parseval's identity, the L^2 norm of the polynomial

$$P(z) = a_n z^n + \cdots + a_1 z + a_0,$$

is given explicitly by

$$||P||_2 = \left(|a_n|^2 + \cdots + |a_1|^2 + |a_0|^2 \right)^{1/2}.$$

As for the values of $M(P)$ and $||P||_s$, $s \neq 2$, some rough estimates in terms of coefficients of the polynomial are more useful, see [34] and [171]. Most often, these estimates use the easy-computable L^2 norm as a reference point. For instance, $M(P) < ||P||_2$ by the monotonicity property. Hence, it is natural to raise the following question:

Question 11.2. *Given a polynomial $P(z) \in \mathbb{C}[z]$, how large is the Mahler measure $M(P)$ of this polynomial in relation to its L^2 norm?*

Mahler [141] investigated the maximum of $M(P)$ for polynomials with bounded coefficients. Mahler proved that $M(P)$ is maximized if P has complex coefficients of equal modulus (that is, P is a constant multiple of some unimodular polynomial). Subsequently, Fielding [86], Beller and Newman [22] proved that for such polynomials, the maximum of $M(P)/||P||_2$ tends to 1, as the degree n increases to infinity.

Obviously, the L^2 norm of a Littlewood polynomial $p \in \mathcal{L}_n$ is equal to $\sqrt{n+1}$. In contrast with a general complex case $P(z) \in \mathbb{C}[z]$, it is not known whether Mahler measures of $p \in \mathcal{L}_n$ can be arbitrarily close to or they are bounded away from the L^2 norm, i.e. if there exists a constant $c > 0$, such that

$$\frac{M(p)}{\sqrt{n+1}} < c < 1$$

for all $p \in \mathcal{L}_n$. Here we formulate a weaker conjecture:

Conjecture 11.3. *For any polynomial $p(z) \in \mathcal{L}_n, n \geqslant 1$, there exists an absolute constant $c > 0$, such that*

$$M(p) < \sqrt{n+1} - c.$$

Even in this weak form, Conjecture 11.3 is still open. Newman [158], [159] and Littlewood [135], [136] posed similar questions for the L^1 norm instead of the Mahler measure. By the monotonicity property of L^s norms, proving Conjecture 11.3 for the L^1 norm also implies the same result for the Mahler measure. The best known result towards Conjecture 11.3 is

$$M(p) < ||p||_1 < \sqrt{n+0.91}.$$

This inequality has been obtained in [40] by optimizing previous estimate of Newman [158]. Recently, Erdélyi made a significant progress in this direction by proving L^1 version of Conjecture 11.3 for cyclotomic Littlewood polynomials (the paper [81] is still in preparation).

In [174], Saffari proved that Barker polynomials of large degree would possess the property of *flatness*, namely,

$$c_1\sqrt{n+1} < |p(z)| < c_2\sqrt{n+1},$$

on the complex unit circle $|z| = 1$ for some positive constants c_1 and c_2. The very existence of such flat polynomials is yet another old question which dates back to Littlewood [135], [136] and Erdős [82]. By using the estimate of Saffari, Borwein and Mossinghoff showed in [40] that if an infinite sequence of Barker polynomials

would exist, they should have extremely large Mahler measures:

$$M(p) > \sqrt{n+1} - 1$$

by Theorem 4.1 in [40]. In view of the last estimate, we formulate the second conjecture on Mahler measures of Barker polynomials.

Conjecture 11.4. *Suppose that $p_{n_k}(z)$ is an infinite sequence of Barker polynomials of increasing degree n_k. Then one has*

$$\lim_{k\to\infty} \left(M(p_{n_k}) - \sqrt{n_k+1} \right) = 0.$$

We note that if both Conjectures 11.3 and 11.4 are true, then there are only finitely many Barker sequences. In Section 11.3 we shall completely solve Conjecture 11.4. Although Conjecture 11.3 is still open, recent results [81] provide some hope that Conjecture 11.3 can be settled in the affirmative.

11.2.3 Class $\mathcal{L}P_n$

We start by giving some definitions. Recall that a function $P(z)$ is called a *Laurent polynomial* (centered at origin) if it is a polynomial in z and $1/z$ with complex coefficients, $P(z) \in \mathbb{C}[z, 1/z]$. If $P \neq 0$ is a Laurent polynomial, then $P(z) = z^{-m}Q(z)$ for some polynomial $Q \in \mathbb{C}[z]$, $Q(0) \neq 0$ and some $m \in \mathbb{N}$:

$$Q(z) := c_0 + c_1 z^2 + \cdots + c_n z^n = c_n \prod_{j=1}^{n} (z - \alpha_j) \in \mathbb{C}[z].$$

Since P and Q satisfy $|P(z)| = |Q(z)|$ on the unit circle $|z| = 1$, one can define the L^s norm of $P(z) \in \mathbb{C}[z, 1/z]$ on the unit circle $|z| = 1$ in a usual way, by setting

$$||P||_s := ||Q||_s = \left(\frac{1}{2\pi} \int_0^{2\pi} |P(e^{it})|^s dt \right)^{1/s}.$$

The Mahler measure of $P(z)$ and $Q(z)$ is defined by

$$M(P) := M(Q) = |c_n| \prod_{j=1}^{n} \max\{1, |\alpha_j|\}.$$

In view of Jensen's formula, one has

$$\log M(P) = \frac{1}{2\pi} \int_0^{2\pi} \log |P(e^{it})| dt. \tag{11.2}$$

Now we are ready to introduce the class $\mathcal{L}P_n$ of Laurent polynomials, which

arise in a natural way in the investigations of Barker polynomials of odd degree. The special form that the polynomials $P \in \mathcal{L}P_n$ take allows us to simplify the mathematical notation considerably.

Definition 11.5. *Let n be an odd integer. Define the class $\mathcal{L}P_n$ as the class of Laurent polynomials of the form*

$$P(z) = (n+1) + \sum_{k=-n,\ k\ odd}^{n} c_k z^k,$$

with coefficients $c_k = -1$ or 1 and $c_{-k} = c_k$.

If $p(z)$ is a Barker polynomial of degree n, then the product $p(z)p(1/z)$ is a Laurent polynomial, which belongs to the class $\mathcal{L}P_n$. In Proposition 11.6 below we identify a more precise subclass of the polynomials in $\mathcal{L}P_n$, related to the products of Barker polynomials. We note that this is just a restatement of the theorem of Turyn and Storer [207], [210]. The proof of Proposition 11.6 can be also found in [40].

Proposition 11.6. *Suppose that $p(z) \in \mathcal{L}_n$ is a Barker polynomial of odd degree n. Then $n = 4m^2 - 1$ for some $m \in \mathbb{Z}$ and the coefficients c_k of the polynomial $P(z) = p(z)p(1/z) \in \mathcal{L}P_n$ in the formula* (11.1) *satisfy $c_{n+1-k} = -c_k$ for $1 \leqslant k \leqslant n$, so that $P(z)$ can be written as*

$$P(z) = (n+1) + Q(z) + Q(1/z),$$

where the polynomial

$$Q(z) := c_n z^n + \ldots c_k z^{k-1} + \cdots + c_1 z$$

is a negative reciprocal: $z^{n+1}Q(1/z) = -Q(z)$.

One should note that, while Barker polynomials of large degree are only hypothetical, the class $\mathcal{L}P_n$ exists and has some very peculiar extremal properties. Among all polynomials in $\mathcal{L}P_n$, polynomials with all coefficients c_k equal to 1 or, alternatively, all coefficients $c_k = -1$ are of special interest.

Definition 11.7. *Let us denote*

$$R_n(z) := (n+1) + \sum_{\substack{k=-n \\ k\ -\ odd}}^{n} z^k.$$

We will show that $R_n(z)$ and $R_n(-z)$ has several interesting extremal properties. We begin with some elementary observations.

Proposition 11.8. *Let $P \in \mathcal{L}P_n$. Then $P(e^{it})$ takes real non-negative values for $t \in [0, 2\pi)$. $P(z) = 0$ holds for some z of modulus 1 if and only if $P(z) = R_n(z)$ and $z = -1$ or $P(z) = R_n(-z)$ and $z = 1$. For each $P \in \mathcal{L}P_n$, one has*

$$||P||_1 = n + 1,$$

and

$$||P||_2 = \Big((n+1)^2 + (n+1)\Big)^{1/2}.$$

In addition, we also have

$$||P||_4 \leqslant ||R_n||_4 = \Big((n+1)(3n^3 + 29n^2 + 49n + 24)/3\Big)^{1/4}.$$

After some computer experimentation, we have conjectured in [31] that the polynomials $R_n(z)$ and $R_n(-z)$ have smallest possible Mahler measures in $\mathcal{L}P_n$.

Conjecture 11.9. *Among all the polynomials $P \in \mathcal{L}P_n$ the polynomials $R_n(z)$ and $R_n(-z)$ have minimal Mahler measures:*

$$\min_{P \in \mathcal{L}P_n} M(P) = M(R_n).$$

We have verified Conjecture 11.9 computationally up to degree $n = 39$.

Conjecture 11.10. *The Mahler measures of polynomials $R_n(z)$ satisfy*

$$M(R_n) = n - 1 + \delta_n,$$

where the error term $\delta_n < 0.725$ for all n sufficiently large and the sequence δ_n is slowly decreasing. It seems that δ_n is bounded from below, probably $\delta_n > 0$. If this is true, then the monotone convergence theorem implies the existence of the limit $\lim_{n\to\infty}(M(R_n) - (n-1))$.

This second conjecture has been verified for polynomials R_n of degree $n < 250$.

We would like to conclude this section with another conjecture on the reducibility of polynomials in the class $\mathcal{L}P_n$ which, combined with the Proposition (11.6), implies the non-existence of Barker polynomials of odd degree $n \geqslant 5$. This conjecture has been verified computationally for $n < 75$. An analogous conjecture for Barker polynomials in the even degree case was introduced in [37] by Borwein, Kaltofen and Mossinghoff.

Conjecture 11.11. *For each $n \geqslant 5$, the subset of polynomials $P \in \mathcal{L}P_n$ with coefficients $c_{n+1-k} = -c_k$ for $1 \leqslant k \leqslant n$ in (11.1) contains only two elements which are reducible in $\mathbb{Z}[z, 1/z]$ (up to the factor of the form $\pm z^m$, $m \in \mathbb{Z}$). These*

elements are $P(z) = R_n(z)$ *and* $P(z) = R_n(-z)$ *defined by* (11.7). *The polynomial* $R_n(z)$ *factors into* $(z-1)(1-1/z)Q(z)$, *where* $Q \in \mathbb{Z}[z, 1/z]$ *is irreducible.*

11.3 Main results

We proceed to state our main results. We will use them to show that Conjecture 11.4 is true. In the first step, we prove that extremal polynomials R_n, defined in previous Section 11.2.3, actually have very large Mahler measures. This result can be seen as a step towards Conjecture 11.10.

Theorem 11.12. *For the polynomials* $R_n(z)$ *defined above:*

$$M(R_n) > n - \frac{2}{\pi}\log n + O(1),$$

as $n \to \infty$.

The second step is to solve Conjecture 11.9.

Theorem 11.13. *If a polynomial* $P \in \mathcal{L}P_n$, *then* $M(P) \geqslant M(R_n)$.

One should note that precise extremal results like Theorem 11.13 are quite rare. A considerably weaker estimate $M(R_n) > (n+1)/2$ was established already in the paper [31]. In general, it is hard to establish non–trivial lower bounds for Mahler measures of polynomials – see, for example, a nice survey of Smyth [201]. Surprisingly, Theorem 11.13 admits a simple (but not trivial) analytical proof. In the present thesis, we will give a short proof of Theorem 11.13 by replacing the more lengthy monotone convergence argument used in paper [32] with an argument based on the continuity of the Mahler measure of the shifted polynomial $P + \varepsilon$. This argument was suggested by Erdélyi. For the proof of the continuity of Mahler measure, see papers of Chern and Vaaler [55] or Boyd [47].

Let us see how our results imply Conjecture 11.4. Let $p(z)$ be a Barker polynomial. Then $P(z) = p(z)p(1/z) \in \mathcal{L}P_n$. Since Mahler measure is multiplicative, one has

$$M(p)^2 = M(p(z))M(p(1/z)) = M(p(z)p(1/z)) = M(P).$$

Hence, by theorems 11.13 and 11.12:

Corollary 11.14. *For any Barker polynomial* $p(z)$ *of degree* n, *we have*

$$M(p) = M(P)^{1/2} \geqslant M(R_n)^{1/2} \geq \left(n - \frac{2}{\pi}\log n + O(1)\right)^{1/2}.$$

This improves Theorem 4.1 in [40]. In addition to this,

$$\left|\sqrt{n+1}-\left(n-\frac{2}{\pi}\log n+O(1)\right)^{1/2}\right| \sim \frac{\log n}{\pi\sqrt{n+1}} \to 0$$

as $n \to \infty$. Hence Conjecture 11.4 is proved.

In fact, one can extend the argument of the proof of Theorem 11.13 to prove a more general result on the extremal L^s norms of the polynomials in the class $\mathcal{L}P_n$.

Theorem 11.15. *For $s \in [0,1] \cup [2,3]$, the polynomials $R_n(\pm z)$ have minimal L^s norms in the class $\mathcal{L}P_n$. On the other hand, $R_n(z)$ have maximal L^s norms in $\mathcal{L}P_n$ for $s \in [2j-1, 2j]$, $j \in \mathbb{N}$, and also for all s which are sufficiently large: $s > s_0(n)$.*

Theorem 11.15 is of less significance for studying Barker polynomials. Nevertheless, we include this result to demonstrate that the integral norms in the set $\mathcal{L}P_n$ are bounded in a predictable way, which is not the case for general complex polynomials in $\mathbb{C}[z]$. We remark that the case $s \in [2,3]$ added to Theorem 11.15 comes as an observation of Mossinghoff. This case does not appear in the original paper [32].

11.4 Proofs

11.4.1 Lemmas

Next Lemma 11.16 is a simple formula for the addition of two sine waves (sometimes referred as the sine phasor formula).

Lemma 11.16. *Let a, b, x and α be real numbers. Suppose $a + b\cos\alpha \geq 0$. Then we have*

$$a\sin x + b\sin(x+\alpha) = c\sin(x+\beta),$$

where

$$c = \sqrt{a^2+b^2+2ab\cos\alpha}$$

and

$$\beta = \tan^{-1}\left(\frac{b\sin\alpha}{a+b\cos\alpha}\right).$$

The proof is straightforward.

Theorem 11.12 will be proven by estimating the integral which appears in Jensen's formula. Here are two lemmas which will provide the necessary estimates.

Lemma 11.17. *Let $n \in \mathbb{N}$ be odd and $I(a) := [a - 1/n, a + 1/n] \subseteq \mathbb{R}$ for $a \in \mathbb{R}$. Then the function*

$$h(t) := (n+1)\sin t + \sin(n+1)t$$

satisfies the following

$$\log|h(t)| = \begin{cases} \log n + \log|t| + O(1), & \text{if } t \in I(0), \\ \log n + O(1/n), & \text{if } t \in I(\pi/2) \cup I(3\pi/2), \\ 3\log n + 3\log|t-\pi| + O(1), & \text{if } t \in I(\pi). \end{cases}$$

Proof. Since $h(t)$ belongs to $C^\infty(\mathbb{R})$, it is well approximated by its Taylor expansion. Indeed, if we let $N := n + 1$, then the derivatives of $h(t)$ are

$$h^{(k)}(t) = \begin{cases} (-1)^{\frac{k}{2}}(N\sin t + N^k \sin Nt) & \text{if } k \equiv 0 \pmod 2, \\ (-1)^{\frac{k-1}{2}}(N\cos t + N^k \cos Nt) & \text{if } k \equiv 1 \pmod 2. \end{cases}$$

We also observe that N is even and $\log N = \log n + O(1/n)$.

We now suppose $t \in I(a)$.

Suppose first that $a = \pi/2$ or $a = 3\pi/2$. Then we have $h(a) = \pm N$, $|h'(t)| \leqslant 2N$, $|t - a| \leqslant 1/n$ and the Taylor expansion of $h(t)$ about a is

$$h(t) = \pm N + h'(\theta)(t-a)$$

for some $\theta \in I(a)$. It follows that

$$\log|h(t)| = \log N + \log\left(1 + \frac{h'(\theta)(t-a)}{N}\right) = \log n + O(1/n).$$

For $a = 0$, we have $h(0) = h''(0) = 0$, $h'(0) = 2N$ and $|h'''(t)| < 2N^3$. Hence, we have

$$h(t) = 2Nt + \frac{h'''(\theta)t^3}{6} = 2Nt\left(1 + \frac{h'''(\theta)t^2}{12N}\right)$$

for some $\theta \in I(0)$. Since

$$\left|\frac{h'''(\theta)t^2}{12N}\right| < \frac{1}{6}\left(\frac{N}{n}\right)^2 = \frac{1}{6}\left(1 + \frac{1}{n}\right)^2 \leq \frac{2}{3},$$

so

$$\frac{1}{3} < 1 - \left|\frac{h'''(\theta)t^2}{12N}\right| \leq \left|1 + \frac{h'''(\theta)t^2}{12N}\right| \leq 1 + \left|\frac{h'''(\theta)t^2}{12N}\right| < \frac{5}{3}.$$

Hence

$$\log|h(t)| = \log N + \log|t| + O(1).$$

Finally, for $a = \pi$, we have $h(a) = h'(a) = h''(a) = h^{(4)}(a) = 0$, $h'''(t) = N - N^3$ and $|h^{(5)}(t)| < 2N^5$.We get

$$h(t) = \frac{N - N^3}{3!}(t-\pi)^3 + \frac{h^{(5)}(\theta)}{5!}(t-\pi)^5 = \frac{N - N^3}{6}(t-\pi)^3 \left(1 + \frac{h^{(5)}(\theta)(t-\pi)^2}{20(N - N^3)}\right)$$

for some $\theta \in I(\pi)$. Since

$$\left|\frac{h^{(5)}(\theta)(t-\pi)^2}{20(N - N^3)}\right| < \frac{N^5}{10(N^3 - N)n^2} = \frac{1}{10}\left(\frac{(n+1)^4}{n^3(n+2)}\right) < \frac{1}{10}\left(1 + \frac{1}{n}\right)^3 \leq \frac{4}{5},$$

so

$$\frac{1}{5} < \left|1 + \frac{h^{(5)}(\theta)(t-\pi)^2}{20(N - N^3)}\right| < \frac{9}{5}.$$

Hence

$$\log|h(t)| = \log(N^3 - N) + \log|t-\pi|^3 + O(1) = 3\log n + 3\log|t-\pi| + O(1).$$

This completes the proof. □

Lemma 11.18. *Let $n \geq 2$. If $a \in \{0, \pi\}$, then we have*

$$\int_{a-2/n}^{a+2/n} \log|\sin t|dt = -\frac{4}{n}\log n + O\left(\frac{1}{n}\right),$$

and

$$\int_{a-1/n}^{a+1/n} \log|\sin(t + \beta(n,t))|dt = -\frac{2}{n}\log n + O\left(\frac{1}{n}\right)$$

where

$$\beta(n,t) = \tan^{-1}\left(\frac{\sin(nt)}{n+1+\cos(nt)}\right).$$

If $a \in \{\pi/2, 3\pi/2\}$, then we have

$$\int_{a-2/n}^{a+2/n} \log|\sin t|dt \ll \frac{1}{n} \quad \textit{and} \quad \int_{a-1/n}^{a+1/n} \log|\sin(t + \beta(n,t))|dt \ll \frac{1}{n^2}.$$

Proof. If $a = \frac{\pi}{2}$ or $a = \frac{3\pi}{2}$, then $0 < \sin\left(\frac{\pi}{2} - 1\right) \leqslant |\sin t| \leqslant 1$ in the interval $[a - 2/n, a + 2/n]$. Hence

$$\left|\int_{a-2/n}^{a+2/n} \log|\sin t|dt\right| \leq \left(\frac{4}{n}\right)\left|\log\sin\left(\frac{\pi}{2} - 1\right)\right| \ll \frac{1}{n}.$$

If $a = 0$ or $a = \pi$, then we write

$$\log|\sin t| = \log\left|\frac{\sin(t-a)}{t-a}\right| + \log|t-a|.$$

Note that the function $|\sin(t-a)/(t-a)|$ is continuous and strictly positive in $[a-2/n, a+2/n]$. Hence

$$\int_{a-2/n}^{a+2/n} \log\left|\frac{\sin(t-a)}{t-a}\right| dt \ll \frac{1}{n}.$$

On the other hand,

$$\int_{a-2/n}^{a+2/n} \log|t-a|dt = \frac{4}{n}\left(\log\frac{2}{n} - 1\right) = -\frac{4}{n}\log n + O\left(\frac{1}{n}\right)$$

and hence

$$\int_{a-2/n}^{a+2/n} \log|\sin t|dt = -\frac{4}{n}\log n + O\left(\frac{1}{n}\right).$$

From Lemma 11.16 with $a = n+1, b = 1, x = t$ and $\alpha = nt$, we have

$$\sin(t+\beta(n,t)) = \frac{(n+1)\sin t + \sin(n+1)t}{c(n,t)}$$

where $c(n,t) = \sqrt{(n+1)^2 + 1 + 2(n+1)\cos(nt)}$. Let $h(t) = (n+1)\sin t + \sin(n+1)t$. By Lemma 11.17, if $a = \pi/2$ or $3\pi/2$, then $\log|h(t)| = \log n + O(1/n)$ for all $t \in [a-1/n, a+1/n]$. Hence

$$\int_{a-1/n}^{a+1/n} \log|h(t)|dt = 2\frac{\log n}{n} + O(1/n^2).$$

In addition, $\log c(n,t) = \log n + O(1/n)$. Hence

$$\int_{a-1/n}^{a+1/n} \log|\sin(t+\beta(n,t))|dt \ll \frac{1}{n^2}.$$

If $a = 0$, then $\log|h(t)| = \log n + \log|t| + O(1)$ for all $t \in [a-1/n, a+1/n]$ and hence

$$\int_{a-1/n}^{a+1/n} \log|h(t)|dt = \frac{2\log n}{n} + \int_{-1/n}^{1/n} \log|t|dt + O(1/n) = O(1/n).$$

If $a = \pi$, then $\log|h(t)| = 3\log n + 3\log|t-\pi| + O(1)$ for all $t \in [a-1/n, a+1/n]$ and hence

$$\int_{a-1/n}^{a+1/n} \log|h(t)|dt = \frac{6\log n}{n} + 3\int_{\pi-1/n}^{\pi+1/n} \log|t-\pi|dt + O(1/n) = O(1/n).$$

Hence

$$\int_{a-1/n}^{a+1/n} \log|\sin(t+\beta(n,t))|dt = -\frac{2}{n}\log n + O\left(\frac{1}{n}\right). \qquad \square$$

11.4.2 Propositions and Theorems

Proof of Proposition 11.8. Let $P \in \mathcal{L}P_n$. Then

$$P(z) = (n+1) + \sum_{k=-n, k\ -\ \text{odd}}^{n} c_k z^k = (n+1) + \sum_{k=1, k\ -\ \text{odd}}^{n} c_k (z^k + z^{-k}).$$

For $z = e^{it}$, $t \in [0, 2\pi)$,

$$P(e^{it}) = (n+1) + \sum_{\substack{k=1 \\ k-\ \text{odd}}}^{n} c_k (e^{ikt} + e^{-ikt}) =$$

$$= (n+1) + \sum_{\substack{k=1 \\ k-\ \text{odd}}}^{n} 2c_k \cos kt \geqslant (n+1) - (n+1) = 0,$$

since n is odd and $c_k = \pm 1$. The equality is possible if and only if all $c_k \cos(kt) = -1$. In particular, for $k = 1$, $\cos t = 1$ or $\cos t = -1$. In the first case, $t = 0$, hence all the coefficients $c_k = -1$. In the second case $t = \pi$ and all $c_k = 1$. This proves that $P(e^{it}) \geqslant 0$ for all $P \in \mathcal{L}P_n$ and $P(e^{it}) = 0$ only if $P(z) = R_n(\pm z)$. Thus for any $P \in \mathcal{L}P_n$

$$\|P\|_1 = \frac{1}{2\pi} \int_0^{2\pi} |P(e^{it})| dt = \frac{1}{2\pi} \int_0^{2\pi} P(e^{it}) dt =$$

$$= (n+1) + \sum_{\substack{k=1 \\ k-\ \text{odd}}}^{n} \frac{2c_k}{2\pi} \int_0^{2\pi} \cos kt\, dt = (n+1).$$

By Parserval's formula,

$$\|P\|_2^2 = (n+1)^2 + \sum_{\substack{k=1 \\ k\ -\ \text{odd}}}^{n} 2c_k^2 = (n+1)^2 + (n+1).$$

It remains to compute L_4 norm. By Parserval's formula, $||P||_4^4$ is equal to the sum of squares of coefficients of the polynomial P^2. Clearly, if all $c_k = 1$, then all the coefficients of P^2 are positive and achieve maximal values. Thus maximal L_4 norm is achieved by the polynomial R_n. By the direct calculation,

$$\begin{aligned} R_n(z)^2 &= \left((n+1) + \sum_{k=-n, k\ \text{odd}}^{n} z^k \right)^2 \\ &= (n+1)^2 + 2(n+1) \sum_{k=-n, k\ \text{odd}}^{n} z^k + \left(\sum_{k=-n, k\ \text{odd}}^{n} z^k \right)^2 \end{aligned}$$

$$
\begin{aligned}
&= (n+1)^2 + 2(n+1) \sum_{k=-n, k \text{ odd}}^{n} z^k + \sum_{\substack{m=-2n \\ m \text{ even}}}^{2n} \left(\sum_{\substack{k_1,k_2=-n,k_i \text{ odd} \\ k_1+k_2=m}}^{n} \right) z^m \\
&= \sum_{m=-2n}^{2n} d_m z^m
\end{aligned}
$$

where

$$
d_m := \begin{cases} 2(n+1) & \text{if } m \text{ is odd and } |m| \le n, \\ 0 & \text{if } m \text{ is odd and } |m| > n, \\ (n+1)(n+2) & \text{if } m = 0, \\ n+1-\frac{|m|}{2} & \text{if } m \text{ is even } 0 < |m| \le 2n. \end{cases}
$$

Therefore,

$$
\begin{aligned}
\|R_n(z)\|_4^4 &= \sum_{m=-2n}^{2n} d_m^2 \\
&= ((n+1)(n+2))^2 + (n+1)(2(n+1))^2 + 2\sum_{m=1}^{n} (n+1-m)^2 \\
&= \frac{1}{3}(n+1)\left(3n^3 + 29n^2 + 49n + 24\right).
\end{aligned}
$$

□

We are ready for the proof of Theorem 11.12.

Proof of Theorem 11.12. It suffices to consider the "+" case in R_n, since the "−" case is $R_n(-z)$. We may also assume $n \ge 3$. Let $z := e^{it}$. Then

$$
\begin{aligned}
R_n(z) &= (n+1) + \sum_{k=-n,\ k \text{ odd}}^{n} z^k \\
&= (n+1) + \frac{z^{n+1} - \bar{z}^{(n+1)}}{z - \bar{z}} \\
&= \frac{(n+1)\sin t + \sin(n+1)t}{\sin t}.
\end{aligned}
$$

Now using Lemma 11.16 with $a = n+1$, $b = 1$, $x = t$ and $\alpha = nt$, we have

$$
R_n(e^{it}) = c(n,t)\frac{\sin(t + \beta(n,t))}{\sin t}
$$

with

$$
c(n,t) = \sqrt{(n+1)^2 + 1 + 2(n+1)\cos(nt)}, \qquad \beta(t) = \tan^{-1}\left(\frac{\sin(nt)}{n+1+\cos(nt)}\right),
$$

since $a + b\cos\alpha = (n+1) + \cos nt \geqslant n > 0$. Observe that

$$n \leqslant c(n,t) \leqslant n+2, \tag{11.3}$$

and

$$|\beta(t)| \leq \tan^{-1}\left(\frac{1}{n}\right) < \frac{1}{n}. \tag{11.4}$$

The last inequality follows from $\tan^{-1}|t| \leq |t|$ for all $t \in \mathbb{R}$. By Jensen's formula, we have

$$\begin{aligned} 2\pi \log M(R_n) - \int_0^{2\pi} \log c(n,t)dt &= \int_0^{2\pi} \Big(\log|R_n(e^{it})| - \log c(n,t)\Big)\, dt \\ &= \int_0^{2\pi} \log|\sin(t+\beta(n,t)| - \log|\sin t|dt. \end{aligned}$$

We now estimate the integral of $\log|\sin(t+\beta(n,t))|$ by using Lemma 11.18. Let

$$I = \int_0^{2\pi} \log|\sin(t+\beta(n,t))|dt = \int_{-1/n}^{2\pi-1/n} \log|\sin(t+\beta(n,t))|dt$$

because the integral is periodic with period 2π. We now write

$$I = \int_{I(0)\cup I(\pi/2)\cup I(\pi)\cup I(3\pi/2)} \log|\sin(t+\beta(n,t))|dt + \int_J \log|\sin(t+\beta(n,t))|dt \tag{11.5}$$

where $I(a) = [a-1/n, a+1/n]$ as in Lemma 11.17 and J is the complement of the disjoint union $I(0)\cup I(\pi/2)\cup I(\pi)\cup I(3\pi/2)$ in $[-1/n, 2\pi-1/n)$.

In view of Lemma 11.18, we first have

$$\int_{I(0)\cup I(\pi/2)\cup I(\pi)\cup I(3\pi/2)} \log|\sin(t+\beta(n,t))|dt = -\frac{4}{n}\log n + O\left(\frac{1}{n}\right). \tag{11.6}$$

Since the function $|\sin t|$ is increasing for $t \in [0,\pi/2]$ and $t \in [\pi, 3\pi/2]$, we find

$$\begin{aligned} \int_{J\cap([0,\pi/2]\cup[\pi,3\pi/2])} \log|\sin(t+\beta(n,t))|dt &\geq \int_{J\cap([0,\pi/2]\cup[\pi,3\pi/2])} \log|\sin(t-1/n)|dt \\ &= \left(\int_0^{\pi/2-2/n} + \int_\pi^{3\pi/2-2/n}\right) \log|\sin t|dt. \end{aligned}$$

Similarly, since the function $|\sin t|$ is decreasing in intervals $[\pi/2,\pi]$ and $[3\pi/2, 2\pi]$, we find

$$\begin{aligned} \int_{J\cap([\pi/2,\pi]\cup[3\pi/2,2\pi])} \log|\sin(t+\beta(n,t))|dt &\geq \int_{J\cap([\pi/2,\pi]\cup[3\pi/2,2\pi])} \log|\sin(t+1/n)|dt \\ &= \left(\int_{\pi/2+2/n}^{\pi} + \int_{3\pi/2+2/n}^{2\pi}\right) \log|\sin t|dt. \end{aligned}$$

Therefore, we have

$$\begin{aligned}\int_J \log|\sin(t+\beta(n,t))|dt &\geq \int_0^{2\pi}\log|\sin t|dt - \left(\int_{\pi/2-2/n}^{\pi/2+2/n} + \int_{3\pi/2-2/n}^{3\pi/2+2/n}\right)\log|\sin t|dt \\ &= \int_0^{2\pi}\log|\sin t|dt + O\left(\frac{1}{n}\right) \qquad (11.7)\end{aligned}$$

by Lemma 11.18. In view of (11.5), (11.6) and (11.7), we have

$$\int_0^{2\pi}\log|\sin(t+\beta(n,t))|dt \geq \int_0^{2\pi}\log|\sin t|dt - \frac{4}{n}\log n + O\left(\frac{1}{n}\right).$$

Hence from (11.5),

$$2\pi\log M(R_n) > \int_0^{2\pi}\log c(n,t)dt - \frac{4}{n}\log n + O\left(\frac{1}{n}\right).$$

Finally, since

$$c(n,t) = \sqrt{(n+1)^2+1+(2n+1)\cos t} = n\sqrt{1+O(1/n)}$$

so $\log c(n,t) = \log n + O(1/n)$. Hence

$$\log M(R_n) > \log n - \frac{2}{n\pi}\log n + O\left(\frac{1}{n}\right).$$

Applying the exponent on both sides of the above inequality and using $e^{-t} > 1-t$ for $t > 0$ one obtains

$$M(R_n) > n\left(1 - \frac{2}{n\pi}\log n + O\left(\frac{1}{n}\right)\right) = n - \frac{2}{\pi}\log n + O(1),$$

as claimed. □

Finally, we give proofs of Theorem 11.13 and Theorem 11.15.

We need some notation. Let

$$T(z) = \sum_{\substack{k=-n \\ k\,-\,\text{odd}}}^{n} c_k z^k \qquad \text{and} \qquad U_n(z) = \sum_{\substack{k=-n \\ k\,-\,\text{odd}}}^{n} z^k.$$

Then $P(z) = (n+1) + T(z)$ and $R_n(z) = (n+1) + U_n(z)$. From the proof of Theorem 2.3 in [31], for any real number $t \in [0, 2\pi)$, one has

$$T(e^{it}) = 2\sum_{\substack{k=1 \\ k\,-\,\text{odd}}}^{n} c_k\cos kt \qquad \text{and} \qquad U_n(e^{it}) = 2\sum_{\substack{k=1 \\ k\,-\,\text{odd}}}^{n}\cos kt = \frac{\sin(n+1)t}{\sin t}.$$

We start with a simple observation.

Lemma 11.19. *For any $m \in \mathbb{Z}$, $m \geqslant 0$,*

$$\frac{1}{2\pi}\int_0^{2\pi} T^m(e^{it})dt \leqslant \frac{1}{2\pi}\int_0^{2\pi} U_n^m(e^{it})dt.$$

Moreover, the equality only holds for $T(z) = U_n(z)$. For odd n, we have

$$\frac{1}{2\pi}\int_0^{2\pi} T^m(e^{it})dt = \frac{1}{2\pi}\int_0^{2\pi} U_n^m(e^{it})dt = 0 \tag{11.8}$$

for any odd m.

Proof. Write

$$T^m(z) = \left(\sum_{\substack{k=-n,\\ n\,-\,\text{odd}}}^{n} z^k\right)^m = \sum_{k=-mn}^{mn} A_k z^k,$$

where the coefficients A_k are defined by

$$A_k := \sum_{\substack{k_1+\cdots+k_m=k,\\ -n\leqslant k_j\leqslant n,\\ k_j\,-\,\text{odd}}} c_{k_1}\cdots c_{k_m}.$$

Since

$$\frac{1}{2\pi}\int_0^{2\pi} e^{ikt}dt = \begin{cases} 0, \text{ if } k \neq 0, \\ 1, \text{ if } k = 0, \end{cases}$$

this yields

$$\frac{1}{2\pi}\int_0^{2\pi} T^m(e^{it})dt = \sum_{k=-mn}^{mn} \frac{A_k}{2\pi}\int_0^{2\pi} e^{ikt}dt = A_0.$$

Observe that A_0 achieves the maximal value if and only if all products $c_{k_1}\cdots c_{k_m}$ in the sum are equal to 1. This is possible if all the coefficients $c_k = 1$, i.e., $T(z) = U_n(z)$. This proves the result.

To prove the last assertion, since all m and k_j are odd, so

$$k_1 + \cdots + k_m \equiv 1 + \cdots + 1 \equiv m \equiv 1 \not\equiv 0 \pmod 2$$

and hence $A_0 = 0$. □

Proof of Theorem 11.13. From now on, we assume that n is fixed. Choose arbitrary $\varepsilon > 0$ and set $N_\varepsilon := n + 1 + \varepsilon$. Let $P(z) \in \mathcal{L}P_n$. Recall that

$$-\log(1-u) = u + \frac{u^2}{2} + \cdots + \frac{u^m}{m} + \ldots \tag{11.9}$$

holds for any real number $u \in [-1, 1)$. Moreover, the infinite series converges absolutely if $|u| < 1$. Note that the polynomial $T(z)$ of $P(z)$ satisfies

$$\max_{t\in[0,2\pi)} \frac{|T(e^{it})|}{N_\varepsilon} < 1.$$

To see this, note that $|T(e^{it})| \leqslant n + 1 = N_\varepsilon - \varepsilon$. Hence, for $u = T(e^{it})/N_\varepsilon$ in (11.9), the Weierstrass M–criterion implies that the series converges uniformly in the interval $t \in [0, 2\pi)$ for each polynomial $P(z) \in \mathcal{L}P_n$. Since the convergence is uniform with respect to t, we can integrate and exchange the integration and summation to obtain

$$-\int_0^{2\pi} \log\left(1 - \frac{T(e^{it})}{N_\varepsilon}\right) dt = \sum_{m=1}^{\infty} \int_0^{2\pi} \frac{T^m(e^{it})}{mN_\varepsilon^m} dt.$$

The application of Lemma 11.19 gives

$$\sum_{m=1}^{\infty} \int_0^{2\pi} \frac{T^m(e^{it})}{mN_\varepsilon^m} dt \leqslant \sum_{m=1}^{\infty} \int_0^{2\pi} \frac{U_n^m(e^{it})}{mN_\varepsilon^m} dt.$$

By the uniform convergence argument, the exchange of the summation and integration yields

$$\sum_{m=1}^{\infty} \int_0^{2\pi} \frac{U_n^m(e^{it})}{mN_\varepsilon^m} dt = -\int_0^{2\pi} \log\left(1 - \frac{U_n(e^{it})}{N_\varepsilon}\right) dt. \tag{11.10}$$

Thus we have proved that

$$-\int_0^{2\pi} \log\left(1 - \frac{T(e^{it})}{N_\varepsilon}\right) dt \leqslant -\int_0^{2\pi} \log\left(1 - \frac{U_n(e^{it})}{N_\varepsilon}\right) dt.$$

It remains to observe that the integral on the right hand side of the above identity is equal to $2\pi(\log M(P(-z) + \varepsilon) - \log N_\varepsilon)$, and the integrand on the right hand side is $2\pi(\log M(R_n(-z) + \varepsilon) - \log N_\varepsilon)$ by Jensen's formula. By multiplying the last inequality by -1 and exponentiating, one gets

$$M(P + \varepsilon)) \geqslant M(R_n + \varepsilon),$$

since $M(P(-z) + \varepsilon) = M(P(z) + \varepsilon)$, $M(R_n(-z) + \varepsilon) = M(R_n(z) + \varepsilon)$. Now we use the fact that Mahler measures are continuous functions with respect to ε, and obtain the inequality $M(P) \geqslant M(R_n)$ by letting $\varepsilon \to 0$. □

Proof of Theorem 11.15. We use the integral L^s norm formula instead of Jensen's

formula and the binomial formula instead of $-\log(1-u)$:

$$(1+u)^s = \sum_{m=0}^{\infty} \binom{s}{m} u^m, \qquad \text{for } |u| < 1.$$

As in the proof Theorem 11.13, for arbitrary $\varepsilon > 0$, let again $N_\varepsilon := n+1+\varepsilon$. We have

$$\begin{aligned} \|(P+\varepsilon)/N_\varepsilon\|_s^s &= \frac{1}{2\pi}\int_0^{2\pi} \left|\frac{P(e^{it})+\varepsilon}{N_\varepsilon}\right|^s dt \\ &= \frac{1}{2\pi}\int_0^{2\pi} \left(1+\frac{T(e^{it})}{N_\varepsilon}\right)^s dt \\ &= \frac{1}{2\pi}\int_0^{2\pi} \sum_{m=0}^{\infty} \binom{s}{m} \left(\frac{T(e^{it})}{N_\varepsilon}\right)^m dt. \end{aligned}$$

The exchange of the integration and summation in the binomial series is possible by the same uniform convergence argument. One has

$$\frac{1}{2\pi}\int_0^{2\pi} T^m(e^{it})dt = \frac{1}{2\pi}\int_0^{2\pi} U_n^m(e^{it})dt = 0, \qquad \text{if } m \text{ is odd.}$$

Hence

$$\|(P+\varepsilon)/N_\varepsilon\|_s^s = \frac{1}{2\pi}\sum_{m=0}^{\infty} \binom{s}{2m} \int_0^{2\pi} \left(\frac{T(e^{it})}{N_\varepsilon}\right)^{2m} dt.$$

The binomial coefficients are given by

$$\binom{s}{0} = 1, \quad \binom{s}{m} = \frac{s(s-1)\cdots(s-m+1)}{m!} \quad \text{for } m = 1, 2, 3, \ldots.$$

If $0 < s < 1$, the coefficients $\binom{s}{2m}$ are negative for $m \geqslant 1$, whereas $\binom{s}{2m+1}$ are positive for $m \geqslant 0$. By the first part of Lemma 11.19,

$$\begin{aligned} -\|(P+\varepsilon)/N_\varepsilon\|_s^s &= \frac{1}{2\pi}\sum_{m=0}^{\infty} -\binom{s}{2m} \int_0^{2\pi} \left(\frac{T(e^{it})}{N_\varepsilon}\right)^{2m} dt \\ &\leqslant \frac{1}{2\pi}\sum_{m=0}^{\infty} -\binom{s}{2m} \int_0^{2\pi} \left(\frac{U_n(e^{it})}{N_\varepsilon}\right)^{2m} dt \\ &= -\frac{1}{2\pi}\int_0^{2\pi} \left(1+\frac{U_n(e^{it})}{N_\varepsilon}\right)^s dt \\ &= -\frac{1}{2\pi}\int_0^{2\pi} \left|\frac{R_n(e^{it})+\varepsilon}{N_\varepsilon}\right|^s dt = -\|(R_n+\varepsilon)/N_\varepsilon\|_s^s, \end{aligned}$$

since the constant term $m = 0$ in the binomial power series is the same on both sides. Thus, we have found that $||P+\varepsilon||_s \geqslant ||R_n+\varepsilon||_s$. Taking the limits on both sides as $\varepsilon \to 0$, one obtains $||P||_s \geqslant ||R||_s$, since the L^s norms are continuous

with respect to the argument function.

Suppose now that $s \geqslant 1$. The binomial coefficients $\binom{s}{m}$ are positive for $m < s+1$. If s is an integer, the binomial coefficient $\binom{s}{m} = 0$ for $m \geqslant s+1$. If $s \notin \mathbb{N}$, binomial coefficients alternate in sign for $m > s+1$, and the first negative coefficient occurs at $m = [s]+2$ (here $[s]$ denotes the integer part of a real number). Thus

$$\binom{s}{2m} \geqslant 0 \text{ for } m = 0, 1, 2, \ldots \qquad \text{if } s \in (2j-1, 2j), j \in \mathbb{N}.$$

Hence

$$\begin{aligned} \|(P+\varepsilon)/N_\varepsilon\|_s^s &= \frac{1}{2\pi} \sum_{m=0}^{\infty} \binom{s}{2m} \int_0^{2\pi} \left(\frac{T(e^{it})}{N_\varepsilon} \right)^{2m} dt \\ &\leqslant \frac{1}{2\pi} \sum_{m=0}^{\infty} \binom{s}{2m} \int_0^{2\pi} \left(\frac{U_n(e^{it})}{N_\varepsilon} \right)^{2m} dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} \left| \frac{R_n(e^{it})+\varepsilon}{N_\varepsilon} \right|^s dt = \|(R_n+\varepsilon)/N_\varepsilon\|_s^s. \end{aligned}$$

By the same continuity argument, it follows that $||P||_s \leqslant ||R_n||_s$.

Let us consider now the case $s \in [2,3]$. Observe that only the $m = 0$ term in the sum has a positive binomial coefficient, and the rest of non zero terms are all negative. For the negative ones, replacing T by U_n will decrease the sum. Replacing T by U_n does not change the constant term $m = 0$. Hence, $||R_n||_s$ is minimal in the class $\mathcal{L}P_n$ for $s \in [2,3]$.

Observe that the proof fails if $s \in (2j, 2j+1)$, since $\binom{s}{2m} > 0$ for $2m < s+1$ while $\binom{s}{2m} < 0$ for $2m > s+1$.

It remains to prove the last statement of Theorem 11.15. Recall that for a fixed polynomial $P(z)$, the norm $\|P\|_s$ is a continuous, monotonically increasing function of s and $\lim_{s\to\infty} \|P\|_s = \|P\|_\infty$. Since the polynomials $R_n(z)$ have maximal infinity norms $\|R_n\|_\infty = 2(n+1)$ in the class $\mathcal{L}P_n$, it follows that $R_n(z)$ have maximal norms for all s sufficiently large. □

Chapter 12

Composition equations

12.1 Statement of the problem

In Chapter 12, we shall investigate the following problem:

Problem 12.1. *Do there exist integer polynomials $f(x)$, $g(x)$ and $h(x)$ of degrees $\deg f \geqslant 3$, $\deg g \geqslant 2$, $f(x)$ separable (and possibly irreducible in $\mathbb{Z}[x]$), such that $f(g(x)) = f(x)h^2(x)$?*

This question has been posed in connection with a recent work of Borwein, Choi and Ganguli [30] on the sign changes of the *Liouville's lambda function* $\lambda(f(n))$ for the values of integer quadratic polynomials $f(x) \in \mathbb{Z}[x]$ at integer points $n \in \mathbb{Z}$. Recall that for $n \in \mathbb{Z}$, the lambda function $\lambda(n)$ is defined by $\lambda(n) = (-1)^{\Omega(n)}$, where $\Omega(n)$ is the total number of prime factors of n, counted with multiplicity. Alternatively, $\lambda(n)$ is the completely multiplicative function defined by $\lambda(p) = -1$ for each prime p dividing n. Chowla [56] conjectured that

$$\sum_{n \leqslant x} \lambda(f(n)) = o(x)$$

for any integer polynomial $f(x)$ which is not of the form $f(x) = bg(x)^2$, where $b \in \mathbb{Z}$ and $g(x) \in \mathbb{Z}[x]$. For $f(x) = x$, Chowla's conjecture is equivalent to the prime number theorem and has been proven for linear polynomials $f(x)$, but is open for polynomials of higher degrees. Even the much weaker conjecture of Cassaigne et al. [54] which states

Conjecture 12.2. *If $f(x) \in \mathbb{Z}[x]$ and is not of the form of $bg^2(x)$ for some $g(x) \in \mathbb{Z}[x]$, then $\lambda(f(n))$ changes sign infinitely often.*

has not been proved unconditionally for the polynomials of degree $\deg f \geqslant 2$.

In the paper [30], it was proved that the sequence $\lambda(f(n))$ cannot be eventually constant for quadratic integer polynomials $f(x) = ax^2 + bx + c$, provided that

at least one sign change occurs for $n > (|b| + (|D| + 1)/2)/2a$, where D is the discriminant of $f(x)$. The proof is based on the solutions of Pell-type equations. In practice, using this conditional result, one can prove the Cassaigne's conjecture for any particular integer quadratic $f(x)$, for instance, $f(x) = 3x^2 + 2x + 1$. In contrast, the only examples of degree $\deg f \geqslant 3$ for which the conjecture has been proven in [54] are $f(x) = \prod_{j=1}^{k}(ax + b_j)$, where $a, b_k \in \mathbb{N}$, b_k are all distinct, $b_1 \equiv \cdots \equiv b_k \pmod{a}$. No similar examples of irreducible integer polynomials of degree $d \geqslant 3$ are known. It appears that the problem of finding an irreducible example of degree $d = 3$ is interesting and probably difficult.

We now explain how the composition identity in Question 12.1 could be of use to prove that $\lambda(f(n))$ or $\lambda(f(-n))$ is not eventually constant for cubic polynomials $f(x)$. Assume that the leading coefficient of $g(x)$ is positive. Since $\deg g \geqslant 2$, there exists a positive integer n_0 such that $g(n) > n$ for integers $n > n_0$. Suppose that there exist two integers $k_0, l_0 > n_0$ such that $\lambda(f(k_0)) = -\lambda(f(l_0))$. Then $\lambda(f(k_j))$ and $\lambda(f(l_j))$ also differ in sign for infinite sequences of integers k_j and l_j, defined by $k_{j+1} = g(k_j)$ and $l_{j+1} = g(l_j)$, $j \geqslant 0$, since $\lambda(f(g(n))) = \lambda(f(n))$ follows by the composition identity.

Unfortunately, the answer to Question 12.1 is negative. In the next section, we prove a general result which holds for polynomials with coefficients in an arbitrary field K. Our result shows that one cannot prove the conjecture for cubic polynomials $f(x)$ by using the composition identity in Question 12.1. We also refer to [64], where a certain composition identity was used to investigate multiplicative dependence of integer values of quadratic integer polynomials and [62] for further results in this direction.

12.2 Main Result

The main result of Chapter 12 is the following theorem:

Theorem 12.3. *Let $m \geqslant 2$ be an integer not divisible by the characteristic of the field K. Suppose that $f(x) \in K[x]$ is non constant and separable, and the polynomial $g(x)$, $\deg g \geqslant 2$, has a non-zero derivative. Then the equation*

$$f(g(x)) = f(x)h^m(x)$$

holds if and only if:

$$I) \qquad f(x) = ax + b, \qquad a, b \in K, a \neq 0, \qquad g(x) = \left(x + \frac{b}{a}\right) h^m(x) - \frac{b}{a}$$

or

$$II) \qquad f(x) = ax^2 + bx + c, \qquad a, b, c \in K, \quad a \neq 0, \qquad m = 2,$$

with

$$g(x) = \frac{1}{2a}\left(\pm T_n\left(\frac{2ax+b}{\sqrt{D}}\right)\sqrt{D} - b\right), \qquad h(x) = \pm U_{n-1}\left(\frac{2ax+b}{\sqrt{D}}\right),$$

where $T_n(x)$, $U_n(x)$ are Chebyshev polynomials of the first and second kind, respectively, $D = b^2 - 4ac$ is the discriminant of $f(x)$.

We remark that the condition on the separability of $f(x)$ cannot be weakened in Theorem 12.3, which can be seen by taking $f(x) = g(x) = x(x-1)^m$ in $\mathbb{Q}[x]$. The requirement that $g(x)$ has a non-zero derivative for fields K of characteristic $p \neq 0$ also cannot be weakened. Indeed, consider the simple example given by $f(x) = x^d - 1$, $g(x) = x^{p^l}$ in $\mathbb{F}_p[x]$. Also, if the characteristic p divides the exponent $m \neq 0$ in the equation $f(g(x)) = f(x)h^m(x)$, then one can write $h^m(x) = h_1^{m/p}(x^p) = h_2^{m/p}(x)$, where $h_2(x)$ is a polynomial with coefficients in K.

Recall that for the field K of characteristic not equal to 2, the *Chebyshev polynomials* $T_n(x) \in K[x]$ of the *first kind* are defined by the linear recurrence of order two:

$$T_0(x) = 1, \quad T_1(x) = x, \qquad T_{n+2}(x) = 2xT_{n+1}(x) - T_n(x). \tag{12.1}$$

In the similar way, the Chebyshev polynomials of the *second kind* $U_n(x) \in K[x]$ are defined by the recurrence

$$U_0(x) = 1, \quad U_1(x) = 2x, \qquad U_{n+2}(x) = 2xU_{n+1}(x) - U_n(x). \tag{12.2}$$

Polynomials $T_n(x)$ and $U_n(x)$ contain only even powers of x for even n, odd powers of x for odd n. Thus, the coefficients of $g(x)$ and $h(x)$ in Theorem 12.3, (II) lie in K if n is odd and in $K(\sqrt{D})$ if n is even. Chebyshev polynomials have many other remarkable properties, see, for instance, [172]. They play a key role in the theorems of Ritt for decompositions of polynomials [185]. In addition, Chebyshev polynomials are related to permutation polynomials over finite fields called Dickson polynomials [133]. In our proof, the following property of Chebyshev polynomials will be useful:

Proposition 12.4. *Suppose that the characteristic of the field K is not equal to 2. Then all solutions of the Pell equation*

$$P^2(x) - (x^2 - 1)Q^2(x) = 1$$

in the ring $K[x]$ are given by

$$P(x) = \pm T_n(x), \qquad Q(x) = \pm U_{n-1}(x),$$

where $T_n(x)$ and $U_n(x)$ are Chebyshev polynomials of the first and second kind, respectively.

The equation which appears in Proposition 12.4 is a special case of a general polynomial Pell equation $P(x)^2 - D(x)Q^2(x) = 1$. Solutions to general Pell equations in polynomials over complex number field $K = \mathbb{C}$ were investigated by Pastor [162]. Dubickas and Steuding [75] gave an elementary algebraic proof for arbitrary field K. The proof of Proposition 12.4 can be found in [75]. Alternative proofs (in the case $K = \mathbb{C}$) are given in [16] and [162].

12.3 Proof of main theorem

Proof. Set $d = \deg f$. Let $a \in K$ and $b \in K$ be the leading coefficients of polynomials $f(x)$ and $g(x)$, respectively, $ab \neq 0$. Suppose that L is the field extension of K generated by the roots of the polynomials $f(x)$, $x^m - 1$ and $x^m - b$. Then

$$f(x) = a \prod_{\alpha \in V(f)} (x - \alpha). \tag{12.3}$$

Here $V(f) \subset L$ denotes the set of the roots of the polynomial $f(x)$. The composition equation $f(g(x)) = f(x)h^m(x)$ factors in $L[x]$ into

$$a \prod_{\alpha \in V(f)} (g(x) - \alpha) = a \prod_{\alpha \in V(f)} (x - \alpha)h^m(x), \tag{12.4}$$

and one can cancel a on both sides. Observe that distinct factors $g(x) - \alpha$ on the left hand side of (12.4) are relatively prime in $L[x]$, since their difference is a non-zero constant. We claim that at most one factor $g(x) - \alpha$ may be relatively prime with $f(x)$ if $m \geqslant 2$, and the characteristic of K does not divide m. Indeed, suppose that $g(x) - \beta$, $\beta \in V(f)$, $\beta \neq \alpha$ is another such factor. Then both $g(x) - \alpha$ and $g(x) - \beta$ divide $h^m(x)$, so $g(x) - \alpha$ and $g(x) - \beta$ must be the m-th powers of some polynomials $u(x)$ and $v(x)$ in $L[x]$ which divide $h(x)$, say, $g(x) - \alpha = u^m(x)$ and $g(x) - \beta = v(x)^m$. (Note that $u(x)$ and $v(x)$ belong to $L[x]$, since the field

L contains all roots of $f(x)$ and the m–th roots of the leading coefficient b of the polynomial $g(x)$). Then $u(x)^m - v(x)^m = \beta - \alpha$ is a non-zero constant polynomial. On the other hand,

$$u^m(x) - v^m(x) = \prod_{j=0}^{m-1} (u(x) - \zeta^j v(x)),$$

where ζ is a primitive m–th root of unity in L, and at least one of polynomials $u(x) - \zeta^j v(x)$ has degree greater than or equal to one, which is impossible.
Now, suppose that $V(f) = \{\alpha_1, \alpha_2, \ldots, \alpha_d\}$. Let V_j be the set containing all distinct common roots of the polynomial $g(x) - \alpha_j$ and the polynomial $f(x)$,

$$V_j := V(g(x) - \alpha_j) \cap V(f).$$

Then $g(x) - \alpha_j = f_j(x)u_j(x)$, where $u_j(x) \in L[x]$ and

$$f_j(x) := \prod_{\alpha \in V_j} (x - \alpha).$$

Note that $f_j(x)$ are all separable and coprime in $L[x]$. Since $f(x)$ is also separable, the equation (12.4) implies

$$a \prod_{j=1}^{d} f_j(x) = f(x) \qquad \text{and consequently,} \qquad \prod_{j=1}^{d} u_j(x) = h^m(x). \tag{12.5}$$

The polynomials $u_j(x)$ are relatively prime, thus $u_j(x) = h_j^m(x)$, $j = 1,\ldots,$ d, for some polynomials $h_j(x) \in L[x]$ whose product is equal to $h(x)$ in (12.5). Let $n_j := \deg f_j$, for $j = 1, \ldots, d$. Without loss of generality, assume that $n_1 \leqslant n_2 \leqslant \ldots \leqslant n_d$. Then $n_1 \geqslant 0$. Observe that $n_2 \geqslant 1$ if $n_1 = 0$, since no two factors $g(x) - \alpha_j$ can be coprime with $f(x)$, as noted above. The first identity in (12.5) gives

$$n_1 + n_2 + \cdots + n_d = \deg f = d. \tag{12.6}$$

Since $g(x) = f_j(x)h_j(x)^m + \alpha_j$, one also has $\deg g \equiv n_j \pmod m$. We now consider two cases for $\deg g$ modulo m.

Case 1). Assume that $\deg g \equiv 0 \pmod m$. Then $n_j \geqslant m$ for $j \geqslant 2$, hence

$$d \geqslant m(d-1) \tag{12.7}$$

by (12.6). Since $m \geqslant 2$, one has $d \geqslant 2d - 2$ which is possible for $d = 1$ or $d = 2$ only. Suppose that $d = 2$. Then one also has $m \leqslant 2$ by (12.7).

Case 2). Assume that $\deg g \not\equiv 0 \pmod m$. Then $n_1 = \cdots = n_d = 1$ by (12.6).

Let $\deg g = sm+1$, where $s := \deg h_j \geqslant 1$ for $1 \leqslant j \leqslant d$. Since $h_j^m(x) \mid g(x) - \alpha_j$, the polynomials $h_j^{m-1}(x)$ are (relatively prime) factors of the derivative $g'(x)$. By conditions of Theorem, $g'(x)$ is a non-zero polynomial, hence

$$ms \geqslant \deg g' \geqslant \deg h_1^{m-1} + \cdots + \deg h_d^{m-1} = d(m-1)s$$

and, consequently,

$$m \geqslant d(m-1). \tag{12.8}$$

Then $d \leqslant m/(m-1) \leqslant 2$. Suppose $d = 2$. Then, in addition, (12.8) gives $m \leqslant 2$.

Thus, it remains to consider the cases $d = 1$ and $d = 2$. In the first case, the polynomial $f(x)$ is linear, thus $f(x) = ax + b$ with $a, b \in K$, $a \neq 0$. The equation $f(g(x)) = f(x)h^m(x)$ is equivalent to

$$ag(x) + b = (ax + b)h^m(x),$$

so one simplification solves $g(x)$, and this completes the proof in the case $d = 1$. Suppose $d = 2$. Then $f(x) = ax^2+bx+c$ with $a, b, c \in K$, $a \neq 0$. Let $D = b^2-4ac$, $D \neq 0$, since $f(x)$ is separable. One also has $m = 2$ by the conditions of Theorem 12.3 and the degree inequalities in the two cases above. Hence, it suffices to find the polynomials $g(x)$ and $h(x)$ in the equation $f(g(x)) = f(x)h^2(x)$. Since the characteristic of the field K is not equal to 2 by the conditions of Theorem 12.3, the linear change of variables $x \to x(t)$ defined by

$$x = \frac{t\sqrt{D} - b}{2a}$$

transforms the polynomial $f(x)$ into

$$f(x) = \frac{D}{4a}F(t),$$

where $F(t) = t^2 - 1$. Set

$$G(t) := \frac{1}{\sqrt{D}}\left(2ag\left(\frac{t\sqrt{D} - b}{2a}\right) + b\right), \qquad H(t) := h\left(\frac{t\sqrt{D} - b}{2a}\right).$$

By straightforward substitution, one easily checks that the map $x \to x(t)$ transforms the composition equation $f(g(x)) = f(x)h^2(x)$ into

$$D/4aF(G(t)) = D/4aF(t)H^2(t).$$

Canceling the factor $D/4a$ on both sides, one obtains

$$F(G(t)) = F(t)H^2(t),$$

or, equivalently,

$$G^2(t) - (t^2 - 1)H^2(t) = 1.$$

By Proposition 12.4, all the solutions to this equation are given by the formulas $G(t) = \pm T_n(t)$, $H(t) = \pm U_{n-1}(t)$, where $T_n(t)$ and $U_n(t)$ are Chebyshev polynomials of the first and second kind, respectively. Application of the inverse map $t \to t(x)$ now yields the result. □

12.4 Rational and integer examples

Let $f(x) = ax^2 + bx + c$ be a quadratic polynomial with rational coefficients. For $n = 3$ in Theorem 12.3, one has $T_3(x) = 4x^3 - 3x$ and $U_2(x) = 4x^2 - 1$. Then $f(g(x)) = f(x)h^2(x)$ holds by Theorem 12.3 for

$$\begin{aligned} g(x) &= (16a^2x^3 + 24abx^2 + (9b^2 + 12ac)x + 8bc)/D, \\ h(x) &= (16a^2x^2 + 16abx + 3b^2 + 4ac)/D. \end{aligned} \tag{12.9}$$

Extend the definition of λ function to the whole set of rationals $\mathbb{Q}$ by the complete multiplicativity of λ. Then, using the method outlined in Section 12.1, one can prove easily the following analogue of Theorem 2 in [30] for the sign changes of λ function at rational points $f(r)$, $r \in \mathbb{Q}$, namely: either $\lambda(f(r))$ is constant for all rational numbers r greater than the largest real root of $g(x) - x$ or it changes sign infinitely many often.

The question of finding all solutions of the composition equation in integer polynomials $f(x)$, $g(x)$ and $h(x)$ is closely related to the solution of the polynomial Pell equations in $\mathbb{Z}[x]$, see [145], [157], [213]. This does not seem to be easy. The examples of such polynomials are $f(x) = x^2 \pm 1$, $f(x) = x^2 \pm 2$, $f(x) = x^2 \pm 4$. Respective polynomials $g(x)$ and $h(x)$ with integer coefficients can be found using (12.9). See Table 12.1 bellow.

Table 12.1: Examples of polynomials $f(x), g(x), h(x) \in \mathbb{Z}[x]$ in Theorem 12.3.

$f(x)$	$g(x)$	$h(x)$
x^2+1	$4x^3+3x$	$4x^2+1$
x^2-1	$4x^3-3x$	$4x^2-1$
x^2+2	$2x^3+3x$	$2x^2+1$
x^2-2	$2x^3-3x$	$2x^2-1$
x^2+4	x^3+3x	x^2+1
x^2-4	x^3-3x	x^2-1

Chapter 13

Conclusions

During the doctoral studies at Vilnius University, we have worked on mathematical problems focused on polynomials and their heights. These problems originate in diverse fields, from number theory and Diophantine analysis to signal processing. By the combination of considerable efforts and pure luck, several mathematical results were obtained. In this last Chapter, a brief summary of these results is given.

- For any real polynomial $P(x) \in \mathbb{R}[x]$, we introduced the reduced height by the formula $\mathbb{H}(P) = \inf_{Q \in \mathbb{R}[x] \text{ – monic}} H(PQ)$. We studied the properties of the reduced height. We restated the problem of explicit calculation of reduced heights as a problem for power series vanishing at points of a finite symmetric set in $|z| < 1$ with prescribed multiplicities. Some relations of the reduced height of a polynomial to other measures of height were also given. Most of the results obtained in our work show that the calculation of the reduced height of a polynomial is not straightforward. For example, we found that $\mathbb{H}(x^2 - 18x - 82) = 63$ and $\mathbb{H}((x - 8/5)^2) = 9216/8245$, but $\mathbb{H}(x^2 - 18x + 82) = 64.999999999999999999999863\ldots$ which is the sum of certain infinite series. It seems likely that this constant is a transcendental number. Our work on the reduced height was carried out in parallel with the Schinzel's work [187], [188] on the reduced length of real and complex [189] polynomials.

- We studied the quantity $S(\zeta, n) := \max_{f \in \Lambda_n} |f(\zeta)|$, where Λ_n is the set of all real polynomials of degree at most $n - 1$ with coefficients in the interval $[0, 1]$ and ζ is a fixed complex number of modulus 1. We first demonstrated that in principle, for any given $\zeta \in \mathbb{C}$ and $n \in \mathbb{N}$, the quantity $S(\zeta, n)$ can be calculated. Then we computed the limit $\lim_{n\to\infty} S(\zeta, n)/n$ for every $\zeta \in \mathbb{C}$ of modulus 1. It is equal to $1/\pi$ if ζ is not a root of unity. If

$\zeta = \exp(2\pi ik/d)$, where $d \in \mathbb{N}$ and $k \in [1, d-1]$ is an integer satisfying $\gcd(k,d) = 1$, then the answer depends on the parity of d. More precisely, the limit is 1, $1/(d\sin(\pi/d))$ and $1/(2d\sin(\pi/2d))$ for $d = 1$, d even and $d > 1$ odd, respectively.

- We investigated the sets of Newman and Littlewood numbers: these are algebraic numbers which are complex roots of Newman polynomials (polynomials with coefficients $0, 1$ and constant term 1) and Littlewood polynomials (polynomials with coefficients $-1, 1$). For each Newman polynomial $P(x)$ of degree at most 8, we found a Littlewood polynomial divisible by $P(x)$. Moreover, we showed that every trinomial $1 + ux^a + vx^b$, where $a < b$ are positive integers and $u, v \in \{-1, 1\}$, so, in particular, every Newman trinomial $1 + x^a + x^b$ divides some Littlewood polynomial. Nevertheless, we proved that there exist irreducible Newman polynomials which divide no Littlewood polynomial, e.g., $x^9 + x^6 + x^2 + x + 1$. These results show that the sets of Newman numbers $V_{\mathcal{N}}$, Littlewood numbers $V_{\mathcal{L}}$ and the set of all complex zeros polynomials with coefficients$\{-1, 0, 1\}$, which is denoted by V, are distinct in the sense that between them there are only trivial relations $V_{\mathcal{N}} \subset V$ and $V_{\mathcal{L}} \subset V$. Moreover, $V \neq V_{\mathcal{L}} \cup V_{\mathcal{N}}$. Our example of Newman numbers which are not Littlewood numbers settles the problem 006:07 posed by prof. A. Dubickas at the 2006 West Coast Number Theory conference. The proofs of the main results use both mathematical theory and computer algorithms.

- We investigated the Mahler measures of a derivative $f'(z)$ of a self-inversive polynomial $f(z) \in \mathbb{C}[z]$. Mahler proved that $M(f') \leqslant dM(f)$ for each $f \in \mathbb{C}[z]$ of degree d. In contrast, it is known that for self-inversive polynomials $f(z)$ is a self-inversive polynomial of degree $d \geqslant 2$ then $M(f') > \frac{d}{2}M(f)$. Following the remark of Smyth, we improved the later inequality to

$$M(f') \geqslant \frac{d}{2}(M(f)^2 + |f(0)|^2)^{1/2}.$$

  We showed that the constant $d/2$ is the best possible for d even, namely, that the quotient $M(f')/M(f)$ takes every value in the interval $(d/2, d]$ as f runs through reciprocal polynomials $f \in \mathbb{R}[z]$ of degree d. It seems likely that for d odd the constant $d/2$ is not optimal. For instance, for $d = 3$, the optimal value of the constant is $1.93867997\ldots$ instead of $3/2$. For each odd $d \geqslant 5$, we proved that there exists a monic reciprocal polynomial $f \in \mathbb{Z}[z]$ of degree d such that $M(f') < \frac{d+1}{2}M(f)$. A corresponding problem for L^s norms of a self-inversive polynomial and its derivative was also considered.

- We considered the geometric progressions $\mathcal{G}$ of real numbers which have common elements with some arithmetic progression $\mathcal{A}$. We proved that the intersection $\mathcal{G} \cap \mathcal{A}$ of an infinite geometric progression $\mathcal{G} = u, uq, uq^2, uq^3, \ldots$, where $u > 0$ and $q > 1$ are real numbers, and an infinite arithmetic progression $\mathcal{A}$ contains at most 3 elements, except for two kinds of ratios q. The first exception occurs for $q = r^{1/d}$, where $r > 1$ is a rational number and $d \in \mathbb{N}$. Then this intersection can be of any cardinality $s \in \mathbb{N}$ or infinite. The other (possible) exception may occur for $q = \beta^{1/d}$, where $\beta > 1$ is a real cubic algebraic number with two nonreal conjugates of moduli distinct from β and $d \in \mathbb{N}$. In this (cubic) case, we proved that the intersection $\mathcal{G} \cap \mathcal{A}$ contains at most 6 elements. We also formulated an equivalent result on the values of fractional parts of powers $\{\xi \alpha^n\}$, extending the previous results of Supnick, Cohen, Keston [205], Ehlich [81], Posner and Rumsey [164], [165].

- We investigated the problem of explicit construction of number systems $(\mathcal{B}, \alpha)$ in the rings $\mathbb{Z}[\alpha]$ for expanding algebraic integers α. We proved that such number systems with certain finite digit sets $\mathcal{B} \subset \mathbb{Z}$ can be constructed by elementary means. We proved inequalities for the size of digits in the set $\mathcal{B}$. We showed that if α is quadratic or cubic trinomial, then one can choose $\mathcal{B} = \{0, \pm 1, \ldots, \pm(|N(\alpha)| - 1)\}$, where $N(\alpha)$ stands for the absolute norm of α over $\mathbb{Q}$.

- In the 2007 West Coast Number Theory conference Problem 007 : 14 Walsh asked to determine all irreducible polynomials of the form $P(x) = x^i + x^j + x^k + 4$ with integer exponents $i > j > k > 0$, such that for some positive integer l the polynomial $P(x^l)$ is reducible in $\mathbb{Z}[x]$. We proved that such polynomials are quadrinomials $x^{4m} + x^{3m} + x^{2m} + 4$, where m is an odd positive integer. In addition, Walsh asked for the examples of reducible quadrinomials $x^i + x^j + x^k + n$, $n > 4$ with no linear or quadratic factors. We found some examples of reducible polynomials of such form above with a negative coefficient n.

- A. Dubickas and C. Smyth introduced the metric Mahler measure

$$m_1(\alpha) = \inf \left\{ \sum_{n=1}^{N} m(\alpha_n) : N \in \mathbb{N}, \alpha_1 \cdots \alpha_N = \alpha \right\},$$

  where $m(\alpha)$ denotes the usual (logarithmic) Mahler measure of $\alpha \in \overline{\mathbb{Q}}$. Samuels extended this definition in a natural way to the t-metric Mahler

measure by replacing the sum with the usual (discrete) ℓ^t norm of the vector

$$(m(\alpha_1), \ldots, m(\alpha_N))$$

for any $t \geqslant 1$. In a joint work with Samuels, for $\alpha \in \mathbb{Q}$, we proved that the infimum in $m_t(\alpha)$ may be attained using only rational points, establishing an earlier conjecture. However, we demonstrated that the natural analogue of this result fails for general $\alpha \in \overline{\mathbb{Q}}$ by giving an infinite family of quadratic counterexamples. As part of this construction, we proved an explicit formula for the computation of $m_t(D^{1/k})$, where $D \in \mathbb{N}$ is square-free.

- For odd integer $n > 0$, we introduced and studied Laurent polynomials

$$P(z) = (n+1) + \sum_{\substack{k=1 \\ k \text{ – odd}}}^{n} c_k(z^k + z^{-k}),$$

with all coefficients c_k equal to -1 or 1. We denoted the class of such polynomials by $\mathcal{L}P_n$. Such polynomials arise in the study of Barker sequences of even length – integer sequences having minimal possible autocorrelations. In particular, we studied extremal polynomials

$$R_n(z) = (n+1) + \sum_{\substack{k=-n \\ k \text{ – odd}}}^{n} z^k$$

with all the coefficients $c_k = 1$. We proved

$$M(R_n) > n - \frac{2}{\pi}\log n + O(1).$$

By using an elementary (but not trivial) analytic argument we established that polynomials $R_n(z)$ with all coefficients $c_k = 1$ have minimal Mahler measures in the class $\mathcal{L}P_n$. This allowed us to deduce that Barker polynomials of large degree would possess unlikely large Mahler measures. A generalization of this result to L^s norms was also given. The proofs given in the present thesis benefit from the remarks made by Erdélyi and Mossinghoff, cf. [32].

- We solved the equation $f(g(x)) = f(x)h^m(x)$ where $f(x)$, $g(x)$ and $h(x)$ are unknown polynomials with coefficients in an arbitrary field K, $f(x)$ is non-constant and separable, $\deg g \geqslant 2$, the polynomial $g(x)$ has non-zero derivative $g'(x) \neq 0$ in $K[x]$, and the integer $m \geqslant 2$ is not divisible by the characteristic of the field K. We proved that this equation has no solutions

if $\deg f \geqslant 3$. If $\deg f = 2$, we proved that $m = 2$, and we wrote down all solutions explicitly in terms of Chebyshev polynomials. Diophantine applications for such polynomials $f(x)$, $g(x)$, $h(x)$ with coefficients in $\mathbb{Q}$ or $\mathbb{Z}$ were considered in the context of the conjecture of Cassaigne et al. on the values of Liouiville's λ function at points $f(r)$, $r \in \mathbb{Q}$.

We hope that the results which were obtained in the course of the doctoral research will be of use to other researchers. One day, these results may become *just another brick in the wall* of some more general mathematical theory.

Bibliography

[1] S. Akiyama, *Finiteness and periodicity of beta expansions - number theoretical and dynamical open problems*, Actes des rencontres du CIRM, **1** no. 1: Numération: mathématiques et informatique (2009), p. 3–9.

[2] S. Akiyama, H. Brunotte, A. Pethö, W. Steiner, *Remarks on a conjecture on certain integer sequences,* Period. Math. Hungar., **52** (2006), 1–17.

[3] S. Akiyama, H. Brunotte, A. Pethő, J. Thuswaldner, *Generalized radix representations and dynamical systems II*, Acta Arith. **121** (2006) no. 1 , 21-61.

[4] S. Akiyama, P. Drungilas and J. Jankauskas, *Height reducing problem on algebraic integers*, Funct. Approx. Comment. Math., (to appear).

[5] S. Akiyama, N. Gjini, *Connectedness of number theoretic tilings,* Discret. Math. Theoret. Comput. Sci. **7** (2005), no. 1, 269–312.

[6] S. Akiyama and A. Pethő, *On canonical number systems*, Theoret. Comput. Sci., **270** (2002), 921–933.

[7] S. Akiyama and K. Scheicher, *From number systems to shift radix systems*, Nihonkai Math. J. **16** (2005) no. 2, 95–106.

[8] P. Allen, *On the multiplicity of linear recurrence sequences*, J. Number Theory, **126** (2007), 212–216.

[9] F. Amoroso, *Polynomials with prescribed vanishing at roots of unity*, Boll. Un. Mat. Ital. B(7), **9** (1995), 1021–1024.

[10] F. Amoroso, *A remark on a theorem of Szegö,* Ramanujan J., **1** (1997), 357–362.

[11] F. Amoroso and R. Dvornicich, *A lower bound for the height in abelian extensions,* J. Number Theory, **80** (2000), 260–272.

[12] F. Amoroso and E. Viada, *On the zeros of linear recurrent sequences*, (submitted).

[13] V.V. Arestov, *On integral inequalities for trigonometric polynomials and their derivatives,* Math. USSR, Izv., **18** (1982), 1–18.

[14] A. Aziz, *Inequalities for the derivative of a polynomial,* Proc. Amer. Math. Soc., **89** (1989), 259–268.

[15] A. Aziz and N.A. Rather, *L^p inequalities for polynomials,* Glasnik Mat., **32** (52) (1997), 39–43

[16] E.J. Barbeau, *Pell's equation*, Springer, 2003.

[17] R.H. Barker, *Group synchronizing of binary digital systems*, Communication theory, 273–287, Butterworths Sci. Pub., London, 1953.

[18] D. Bertsimas, J. N. Tsitsiklis *Introduction to Linear Optimization (Athena Scientific Series in Optimization and Neural Computation*, Athena Scientific, A Series in Optimization and Neural Computation, 1997.

[19] F. Beaucoup, P. Borwein, D .W. Boyd and C. Pinner, *Multiple roots of $[-1, 1]$ power series*, J. London Math. Soc. (2), **57** (1998), 135–147.

[20] F. Beaucoup, P. Borwein, D .W. Boyd and C. Pinner, *Power series with restricted coefficients and a root on a given ray*, Math. Comp., **67** (1998), 715–736.

[21] J. Beck, *Flat polynomials on the unit circle – note on a problem of Littlewood*, Bull. London Math. Soc. **23** (1991), 269–277.

[22] E. Beller, D.J. Newman, *An extremal problem for the geometric mean of polynomials*, Proc. Amer. Math. Soc. **39** (1973), 313–317.

[23] J. Berstel, *Transductions and context-free languages,* Taubner, 1979.

[24] F. Beukers, *The multiplicity of binary recurrences*, Compositio Math., **80** (1980), 251–267.

[25] F. Beukers, *The zero-multiplicity of ternary recurrences*, Compositio Math., **77** (1991), 165–177.

[26] D. A. Bini and G. Fiorentino, *Numerical computation of polynomial roots v. 2.0*, FRISCO report (1998) (available online at `http://www.dm.unipi.it/cluster-pages/mpsolve/index.htm`).

[27] E. Bombieri, J. D. Vaaler, *On Siegel's lemma,* Invent. Math., **73** (1983), 11–32.

[28] F. F. Bonsall and M. Marden, *Zeros of self-inversive polynomials,* Proc. Amer. Math. Soc., **3** (1952), 471–475.

[29] P. Borwein, *Computational Excursions in Analysis and Number Theory, CMS books in mathematics*, Springer, 2002.

[30] P. Borwein, S.K.K. Choi, H. Ganguli, *Sign Changes of the Liouville Function on Quadratics*, Canad. Math. Bull., (to appear).

[31] P. Borwein, S.K.K. Choi and J. Jankauskas, *On a class of polynomials related to Barker sequences*, Proceedings of Amer. Math. Soc., (to appear).

[32] P. Borwein, S.K.K. Choi and J. Jankauskas, *Extremal Mahler measures and L_s norms in the class of polynomials related to Barker sequences*, Proceedings of Amer. Math. Soc., (to appear).

[33] P. Borwein, E. Dobrowolski and M.J. Mossinghoff, *Lehmer's problem for polynomials with odd coefficients*, Ann. of Math. (2) **166** (2007), no. 2, 347–366.

[34] P. Borwein and T. Erdélyi, *Polynomials and Polynomial Inequalities*, Springer-Verlag, New York, 1995.

[35] P. Borwein, K.G. Hare and M.J. Mossinghoff, *The Mahler measure of polynomials with odd coefficients*, Bull. London Math. Soc., **36** (2004), 332–338.

[36] P. Borwein, R. Fergusson, J. Knauer, *The merit factor problem*, in: Number theory and polynomials, Cambridge University Press, 52 (2000).

[37] P. Borwein, E. Kaltofen, M.J. Mossinghoff *Irreducible polynomials and Barker sequences*, ACM Commun. Comput. Algebra **41** (2007), no. 3-4, 118–121.

[38] P. Borwein and M. J. Mossinghoff, *Polynomials with height 1 and prescribed vanishing at 1*, Exper. Math., **9** (2000), 425–433.

[39] P. Borwein, M.J. Mossinghoff, *Newman polynomials with prescribed vanishing and integer sets with distinct subset sums,* Math. Comput., **72** (2003), 787–800.

[40] P. Borwein, M.J. Mossinghoff, *Barker sequences and flat polynomials*, in: Number Theory and Polynomials (Bristol, U.K., 2006), J. McKee and C. Smyth, eds., 71–88, London Math. Soc. Lecture Note Ser. **352**, Cambridge Univ. Press, 2008.

[41] P. Borwein, M.J. Mossinghoff, J. Vaaler, *Generalizations of Gonçalves' inequality*, Proc. Amer. Math. Soc., **135**, (2007), no. 1, 253–261.

[42] P. Borwein and C. Pinner, *Polynomials with* $\{0, +1, -1\}$ *coefficients and a root close to a given point.* Canadian J. Math., **49** (1997), 887–915.

[43] D. W. Boyd, *Reciprocal polynomials having small measure*, Math. Comp. **35** (1980), 1361–1377.

[44] D. W. Boyd, *Reciprocal polynomials having small measure*, Math. Comp. **53** (1989), 453–469.

[45] D. W. Boyd, *Large Newman polynomials,* in: Diophantine Analysis (Kensigton, 1985), London Math. Soc. Lecture Note Ser. 109, Cambridge Univ. Press, Cambridge (1986), 159–170.

[46] D. W. Boyd, *Irreducible polynomials with many roots of maximal modulus*, Acta Arith., **68** (1994), 85–88.

[47] D. W. Boyd, *Uniform approximation to Mahler's measure is several variables,* Canad. Math. Bull., **41** (1998), 125–128.

[48] R. Breusch, *On the distribution of the roots of a polynomial with integral coefficients,* Proc. Amer. Math. Soc. **2** (1951), 939–941.

[49] N.G. de Bruijn and T.A. Springer, *On the zeros of composition polynomials,* Indag. Math., **9** (1947), 406–414.

[50] H. Brunotte, *Characterization of CNS trinomials,* Acta Sci. Math. (Szeged), **68** (2002), 673–679.

[51] H. Brunotte, A. Huszti, A. Pethő, *Bases of canonical number systems in quartic algebraic number fields,* J. Théor. des Nombres de Bordeaux **18** (2006), 537–557.

[52] D.M. Campbell, H.R.P. Ferguson, R.W. Forcade, *Newman polynomials on* $|z| = 1$*,* Indiana Univ. Math. J., **32** (1983), 517–525.

[53] D.C. Cantor, E.G. Straus, *On a conjecture of D.H. Lehmer*, Acta Arith. **42** (1982/83), 97–100. Correction: **42** (1983), 327.

[54] J. Cassaigne , S. Ferenczi, C. Mauduit, J. Rivat and A. Sarkozy, *On finite pseudorandom binary sequences IV: The Liouville function II*, Acta Arithmetica XCV. **4** (2000) 343–359.

[55] S.-J. Chern and J.D. Vaaler, *The distribution of values of Mahler's measure,* J. Reine Angew. Math., **540** (2001), 1–47.

[56] S. Chowla, *The Riemann Hypothesis and Hilbert's Tenth Problem*, Gordon and Breach, New York, 1965.

[57] S. Chowla, M. Dunton and D. J. Lewis, *Linear recurrences of order two*, Pacific J. Math., **11** (1961), 833–845.

[58] J. Dégot, Finite-dimensional Mahler measure of a polynomial and Szegö's theorem, *J. Number Theory,* **62** (1997), 422–427.

[59] J.D. Dixon, A. Dubickas, *The values of Mahler measures*, Mathematika, **51** (2004), 131–148.

[60] E. Dobrowolski, *On a question of Lehmer and the number of irreducible factors of a polynomial*, Acta Arith. **34** (1979), no. 4, 391–401.

[61] P.Drungilas, *Heights of algebraic numbers*, doctoral thesis. Vilniaus Universiteto leidykla, Vilnius, (2008).

[62] P. Drungilas, A. Dubickas, *Multiplicative dependence of shifted algebraic numbers*, Colloq. Math., **96** (1) (2003), 75–81.

[63] P. Drungilas and A. Dubickas, *Roots of polynomials of bounded height*, Rocky Mt. J. Math. (to appear).

[64] A. Dubickas, *Multiplicative dependence of quadratic polynomials*, Liet. Matem. Rink., **38** (3) (1998), 295–303.

[65] A. Dubickas, *Arithmetical properties of powers of algebraic numbers,* Bull. London Math. Soc., **38** (2006), 70–80.

[66] A. Dubickas, *On the fractional parts of natural powers of a fixed number*, Siberian Math. J., **47** (2006), 879–882.

[67] A. Dubickas, J. Jankauskas, *On the reduced height of a polynomial*, Publ. Math. Debrecen , **17** (3-4) (2007), 325–348.

[68] A. Dubickas, J. Jankauskas, *The maximal value of polynomials with restricted coefficients*, Journal of the Korean Mathematical Society, **46** (1) (2009), 41–49.

[69] A. Dubickas, J. Jankauskas, *On Newman polynomials which divide no Littlewood polynomial*, Mathematics of Computation, **78** (265) (2009), 327–344.

[70] A. Dubickas, J. Jankauskas, *On Mahler measures of a self-inversive polynomial and its derivative*, Bull. London Math. Soc., **42** (2) (2010), 195–209.

[71] A. Dubickas, J. Jankauskas, *On the intersection of infinite geometric and arithmetic progressions*, Bull. of the Brazilian Math. Soc., **41** (4) (2010), 551–566.

[72] A. Dubickas and M. J. Mossinghoff, *Auxiliary polynomials for some problems regarding Mahler's measure,* Acta Arith., **119** (2005), 65–79.

[73] A. Dubickas and C.J. Smyth, *On the Remak height, the Mahler measure and conjugate sets of algebraic numbers lying on two circles*, Proc. Edinburgh Math. Soc., **44** (2001), 1–17.

[74] A. Dubickas and C.J. Smyth, *On the metric Mahler measure*, J. Number Theory, **86** (2001), 368–387.

[75] A. Dubickas, J. Steuding, *The polynomial Pell equation*, Elemente der Math., **59** (2004), 133–143.

[76] A. Durand, *On Mahler's measure of a polynomial,* Proc. Amer. Math. Soc., **83** (1981), 75–76.

[77] S. Eliahou, M. Kervaire, *Barker sequences and difference sets*, Enseign. Math. (2) **38** (1992), no. 3–4, 345–382.

[78] S. Eliahou, M. Kervaire, B. Saffari, *A new restriction on the lengths of Golay complementary sequences,* J. Combin. Theory Ser. A **55** (1990), no. 1, 49–59.

[79] S. Eilenberg, *Automata, Languages and Machines. Vol A,* Academic Press, New York, 1974.

[80] H. Ehlich, *Die positiven Lösungen der Gleichung* $y^a - [y^a] = y^b - [y^b] = y^c - [y^c]$, Math. Zeitschr., **76** (1960), 1–4.

[81] T. Erdélyi, *On the* L_q *norm of cyclotomic Littlewood polynomials on the unit circle*, (in preparation).

[82] P. Erdős, *An inequality for the maximum of trigonometric polynomials*, Annales Polonica Math. **2** (1940), 310–313.

[83] G. Everest and T. Ward, *Heights of polynomials and entropy in algebraic dynamics,* Universitext, Springer-Verlag, London, 1999.

[84] J.-H. Evertse, H.-P. Schlickewei and W. Schmidt, *Linear equations in variables which lie in a finitely generated group*, Annals of Math., **155** (2002), 807–836.

[85] P. Fan, M. Darnell, *Sequence design for communications applications*, Research Studies Press, Somerset, England, 1996.

[86] G.T. Fielding, *The expected value of the integral around the unit circle of a certain class of polynomials*, Bull. London Math. Soc. **2** (1970), 301–306.

[87] M. Filaseta, *On the factorization of polynomials with small Euclidean norm*, In: Number theory in progress (Zakopane-Kościelisko, 1997), de Gruyter, Berlin, **1** (1999), 143–163.

[88] M. Filaseta, A. Granville, A. Schinzel, *Irreducibility and greatest common divisor algorithms for sparse polynomials*, In: Number Theory and Polynomials (ed. James McKee and Chris Smyth), LMS Lecture Note Series 352, Cambridge Univ. Press, (2008), pp. 155–176.

[89] M. Filaseta, F. Luca, P. Stănică, R. Underwood, *Two Diophantine approaches to the irreducibility of certain trinomials,* Acta Arithmetica 128 (2007), 149–156.

[90] M. Filaseta and M. Matthews, Jr., *On the irreducibility of* 0*,* 1 *- polynomials of the form* $f(x)x^n + g(x)$, Colloq. Math., **99** (2004), 1–5.

[91] M. Filaseta, M.L. Robinson, F.S. Wheeler, *The minimal Euclidean norm of an algebraic number is effectively computable,* J. Algorithms, **16** (1994), 309–333.

[92] M. Filaseta, I. Solan, *An extension of a theorem of Ljunggren*, Math. Scand. **84** (1) (1999), 5–10.

[93] P. Fili and C.L. Samuels, *On the non-Archimedean metric Mahler measure*, J. Number Theory, **129** (2009), no. 7, 1698–1708.

[94] M. Fried, A. Schinzel, *Reducibility of quadrinomials*, Acta Arith., **21** (1972), 153–171.

[95] C. Frougny, W. Steiner, *Minimal weight expansions in Pisot bases,* J. Math. Cryptol. **2** (2008), no. 4, 365–392.

[96] H. Ganguli, J. Jankauskas, *On the equation $f(g(x)) = f(x)h^m(x)$ for composite polynomials,* J. Aust. Math. Soc., (special issue dedicated to Alfred van der Poorten, to appear).

[97] I. Gargantini, *A generalization of the Schur-Cohn criterion,* Proceedings of the 9-th Manitoba Conference of Numerical Mathematics and Computing, 213–217, Congress. Numer. XXVII, Utilitas Math., Winnipeg, Man. 1980.

[98] J. Garza, M.I.M. Ishak, M.J. Mossinghoff, C.G. Pinner, and B. Wiles *Heights of roots of polynomials with odd coefficients*, J. Théor. Nombres Bordeaux **22** (2010), no. 2, 369–381.

[99] W.J. Gilbert, *Radix representations of quadratic fields,* J. Math. Anal. and Appl. **83** (1981), 264–274.

[100] GMP, *The GNU Multiple Precision Arithmetic Library* (available online at `http://swox.com/gmp/`).

[101] K. Gröchenig and A. Haas, *Self-similar lattice tilings*, J. Fourier Anal. Appl. 1 (1994), no. 2, pp. 131–170.

[102] G.H. Hardy, J.E. Littlewood and G. Pólya, *Inequalities,* Cambridge Univ. Press, London, 1952.

[103] X.-G. He, I. Kirat, K.-S. Lau, *Height reducing property of polynomials and self-affine tiles*, Geom. Dedicata, **152** 1 (2011), 153–164.

[104] W. Hofschuster and W. Krämer, *C–XSC 2.0: A C++ Class Library for Extended Scientific Computing*, Numerical Software with Result Verification, Leture Notes in Computer Science, **2991**, Springer-Verlag, Heidelberg, (2004) 15–35 (available online at `http://www.math.uni-wuppertal.de/~xsc/`).

[105] J. Jankauskas, *On the reducibility of certain quadrinomials*, Glasnik Matematički, **45** (65) (2010), 31–41.

[106] J. Jankauskas, C.L. Samuels, *The t-metric Mahler measures of surds of rational numbers*, Acta Math. Hungar., **134** (8) (2012), 481–498.

[107] J. Jedwab, *A survey of the merit factor problem for binary sequences*, Sequences and Their Applications, Proceedings of SETA 2004, Lecture Notes in Computer Science **3486**, 30–55, Springer Verlag, Berlin, 2005.

[108] J. Jedwab, S. Lloyd, *A note on the nonexistence of Barker sequences*, Des. Codes Cryptogr. **2** (1992), no. 1, 93–97.

[109] A. T. Jonassen, *On the irreducibility of the trinomials* $x^m \pm x^n \pm 4$, Math. Scand. **21** (1967), 177-189.

[110] J.-P. Kahane, *Sur les polynômes à coefficients unimodulaires*, Bull. London. Math. Soc. **12** (1980), 321–342.

[111] I. Kátai, B. Kovács, *Kanonische Zahlensysteme in der Theorie der Quadratischen Zahlen,* Acta Sci. Math. (Szeged) **42** (1980), 99–107.

[112] I. Kátai, B. Kovács, *Canonical number systems in Imaginary Quadratic Fields,* Acta Math. Acad. Sci. Hungar. **37** (1981), 159–164.

[113] I. Kátai, J. Szabó, *Canonical number systems for complex integers,* Acta Sci. Math. (Szeged) **37** (1975), 255–260.

[114] I. Kirat, K.-S. Lau, *On the connectedness of self-affine tiles,* J. London Math. Soc. **62** (2000), 291–304.

[115] I. Kirat, K.-S. Lau and H. Rao, *Expanding polynomials and Connectedness of self-affine tiles,* Discrete Comput. Geom. **31** (2004), 275–286.

[116] I. Klemeš, *Finite Toeplitz matrices and sharp Littlewood conjectures*, Algebra i Analiz **13** (2001), 39–59.

[117] D.E. Knuth, *The Art of Computer Programming, Vol. 2 Semi-numerical Algorithms*, Addison Wesley (1998) London 3rd-edition.

[118] S. Konyagin, *On a question of Pichorides*, C. R. Acad. Sci. Paris Sér. I Math. **324** (1997), 385–388.

[119] B. Kovács, *Canonical number systems in algebraic number fields*, Acta Math. Acad. Sci. Hungar. **37** (1981), 405–407.

[120] B. Kovács and A. Pethő, *Number systems in integral domains, especially in orders of algebraic number fields*, Acta Sci. Math. (Szeged), **55** (1991), 286–299.

[121] C. Krattenthaler, *Advanced determinant calculus,* Sémin. Lothar. Combin., **42** (1999), Art. B42q, 67 pp. (electronic).

[122] L. Kronecker, *Zwei Sätze über Gleichungen mit ganzzahligen Coefficienten*, J. für die Reine Angew. Math. **53** (1857), 173–175.

[123] L. Kuipers, H. Niederreiter, *Uniform Distribution of Sequences*, John Wiley & Sons, New York, 1974.

[124] J.C. Lagarias, Y. Wang, *Self-affine tiles in* $\mathbb{R}^n$*,* Adv. Math. **121** (1996), 21–49.

[125] J.C. Lagarias, Y. Wang, *Integral self-affine tiles in* $\mathbb{R}^n$ *I. Standard and nonstandard digit sets,* J. London Math. Soc. **54** (1996), no. 2, 161–179.

[126] J.C. Lagarias, Y. Wang, *Integral self-affine tiles in* $\mathbb{R}^n$ *II. Lattice tilings,* J. Fourier Anal. Appl. **3** (1997), no. 1, 83–102.

[127] P. Lakatos and L. Losonczi, *On zeros of reciprocal polynomials of odd degree,* JIPAM. J. Inequal. Pure Appl. Math., **4**, no. 3, article 60 (2003), 8 pp. (electronic).

[128] S. Lang, *Algebra. 3rd revised ed.,* Graduate Texts in Mathematics. 211. New York, NY: Springer. xv, 914 p., 2002.

[129] S. Lang, *Undergraduate Algebra. 3rd revised ed.,* Undergraduate Texts in Mathematics. New York, NY: Springer. xi, 385 p., 2005.

[130] W. Lawton, *Heights of algebraic numbers and Szegö's theorem,* Proc. Amer. Math. Soc., **49** (1975), 47–50.

[131] K.H. Leung, B. Schmidt, *The field descent method*, Des. Codes Cryptogr. **36** (2005), no. 2, 171–188.

[132] D.H. Lehmer, *Factorization of certain cyclotomic functions,* Ann. of Math., **34** (1933), 461–479.

[133] R. Lidl, H. Niederreiter, *Finite fields* , Encycl. Math. Appl. **20**, Cambridge Univ. Press (1997), 347–393.

[134] J.E. Littlewood, *On the mean values of certain trigonometric polynomials*, J. London Math. Soc. **36** (1961), 307–334.

[135] J.E. Littlewood, *On polynomials* $\sum^n \pm z^m$,$\sum^n e^{\alpha_m i} z^m$, $z = e^{\theta i}$, J. London Math. Soc. **41** (1966), 367–376.

[136] J.E. Littlewood, *Some Problems in real and complex analysis*, D.C. Heath and Co., Lexington, Mass., 1968.

[137] W. Ljunggren, *On the reducibility of certain trinomials and quadrinomials*, Math. Scand. **8** (1965), 65–70.

[138] R. Louboutin, *Sur la mesure de Mahler d'un nombre algébrique*, C. R. Acad. Sci. Paris Sér. I Math. **296** (1983), 707–708.

[139] K. Mahler, *An application of Jensen's formula to polynomials,* Mathematika, **7** (1960), 98–100.

[140] K. Mahler, *On the zeros of the derivative of a polynomial,* Proc. Royal Soc. London, Ser. A, **264** (1961), 145–154.

[141] K. Mahler, *On two extremum properties of polynomials*, Illinois J. Math. **7** (1963), 681–701.

[142] K. Mahler, *An inequality for the discriminant of a polynomial*, Michigan Math. J. **11** (1964), 257–262.

[143] M. Marden, *Geometry of polynomials,* Amer. Math. Soc., Providence, 1989.

[144] O.C. McGehee, L. Pigno and B. Smith, *Hardy's inequality and the L^1 norm of exponential sums*, Ann. of Math.**113** (2) (1981), 613–618.

[145] J. McLaughlin, *Polynomial solutions of Pell's equation and fundamental units in real quadratic fields*, J. London Math. Soc. **67** (2003), 16—28.

[146] M. Mignotte, *Sur les multiples des polynômes irréductibles,* Bull. Soc. Math. Belg., **27** (1975), 225–229.

[147] W.H. Mills, *The factorization of certain quadrinomials*, Math. Scand. **57** (1985), 44–50.

[148] M.J. Mossinghoff, *Algorithms for the determination of polynomials with small Mahler measure*, Ph.D. Thesis, University of Texas at Austin, 1995.

[149] M.J. Mossinghoff, *Polynomials with restricted coefficients and prescribed noncyclotomic factors*, LMS J. Comput. Math., **3** (2003), 314–325.

[150] M.J. Mossinghoff, *Wieferich Pairs and Barker Sequences*, Des. Codes Cryptogr., **53** (2009), no. 3, 149–163.

[151] M.J. Mossinghoff, website, *Lehmer's Problem*, `http://www.cecm.sfu.ca/~mjm/Lehmer`.

[152] M.J. Mossinghoff, C.G. Pinner and J.D. Vaaler, *Perturbing polynomials with all their roots on the unit circle*, Math. Comp. **67** (1998), 1707–1726.

[153] M.J. Mossinghoff, G. Rhin and Q. Wu, *Minimal Mahler measures*, Experiment. Math. **17** (2008), no. 4, 451–458.

[154] M. Meyer, *Le problème de Lehmer: méthode de Dobrowolski et lemme de Siegel à la Bombieri-Vaaler*, Publ. Math. Univ. P. et M. Curie (Paris VI), Problèmes Diophantiens , **90** (1988–89), no. 6, 15 pp.

[155] G. Myerson, *Western Number Theory Problems, 17–19 Dec 2007*, 6. Available online at `http://www.math.colostate.edu/~achter/wntc/problems/problems2007.pdf`

[156] W. Narkiewicz, *Elementary and analytic theory of algebraic numbers,* 3rd ed., Springer, Berlin, 2004.

[157] M.B. Nathanson, *Polynomial Pell equations*, Proc. Amer. Math. Soc. **56** (1976), 89-–92.

[158] D.J. Newman, *Norms of polynomials,* Amer. Math. Monthly **67** (1960), 778–779.

[159] D.J. Newman, *An L_1 extremal problem for polynomials,* Proc. Amer. Math. Soc. **16** (1965), 1287–1290.

[160] D.G. Northcott, *An inequality on the theory of arithmetic on algebraic varieties*, Proc. Cambridge Philos. Soc., **45** (1949), 502–509.

[161] A.M. Odlyzko, B. Poonen, *Zeros of polynomials with* $0, 1$ *coefficients,* Enseign. Math. (2), **39** (1993), 317–348.

[162] A.V. Pastor, *Generalized Chebyshev polynomials and the Pell-Abel equation*, Fundam. Prikl. Mat. **7** (2001), 1123—1145.

[163] R. Pereira, *Weak log-majorization, Mahler measure and polynomial inequalities,* Linear Alg. Appl., **421** (2007), 117–121.

[164] E.C. Posner, *Diophantine problems involving powers modulo one*, Illinois J. Math., **6** (1962), 251–263.

[165] E.C. Posner and H.R. Rumsey, *Polynomials that divide infinitely many trinomials*, Michigan Math. J., **12** (1965), 339–348.

[166] I.E. Pritsker, *Polynomial inequalities, Mahler's measure, and multipliers,* in: *Number theory and polynomials*, Proceedings of the conference held at the University of Bristol, 3-7 April, 2006, (eds. James McKee and Chris Smyth), LMS Lecture Note Series 352 (2008), pp. 255–276.

[167] I.V. Proskuriakov, *A collection of problems on linear algebra,* 6th ed., Nauka, Moscow, 1978 (in Russian).

[168] U. Rausch, *On a theorem of Dobrowolski about the product of conjugate numbers*, Colloq. Math.**50** (1985), 137–142.

[169] L. Robinson, *Polynomials with plus ir minus one coefficients: growth properties on the unit circle*, M. Sc. thesis, Simon Fraser university, 1997.

[170] Q.I. Rahman and G. Schmeisser, *L^p inequalities for polynomials,* J. Approx. Theory, **53** (1988), 26–32.

[171] Q.I. Rahman and G. Schmeisser, *Analytic theory of polynomials,* Oxford Univ. Press, Oxford, 2002.

[172] T.J. Rivlin, *Chebyshev Polynomials: From Approximation Theory to Algebra and Number Theory*, John Wiley & Sons, 1990.

[173] W. Rudin, *Real and complex analysis*, McGraw-Hill, Science, Engineering and Mathematics Series, New York, 3rd edition, 483pp.

[174] B. Saffari, *Barker sequences and Littlewood two-sided conjectures on polynomials with ± 1 coefficients,* Séminaire d'Analyse Harmonique, Année 1989/90, 139–151, Univ. Paris XI, Orsay, 1990.

[175] C.L. Samuels, *The infimum in the metric Mahler measure*, Canad. Math. Bull. **54** (2011), 739–747.

[176] C.L. Samuels, *A collection of metric Mahler measures*, J. Ramanujan Math. Soc. **25** (2010), no. 4, 433–456.

[177] C.L. Samuels, *The parametrized family of metric Mahler measures*, J. Number Theory **131** (2011), no. 6, 1070–1088.

[178] K. Scheicher, J. Thuswaldner, *Canonical number systems, counting automata and fractals,* Math. Proc. Cambridge Philos. Soc. **133** (2002), 163–182.

[179] K. Scheicher and J. Thuswaldner, *On the charecterization of canonical number systems*, J.M. Osaka J. Math. **41** (2004), 327—351.

[180] A. Schinzel, *On the product of the conjugates outside the unit circle of an algebraic number,* Acta Arith., **24** (1973), 385–399; Addendum, *ibid.* **26** (1975), 329–331.

[181] A. Schinzel, *On the reducibility of polynomials and in particular of trinomials*, Acta Arith., **11** (1965), 1–34.

[182] A. Schinzel, *On the reducibility of lacunary polynomials I*, Acta Arith., **16** (1969), 123–159.

[183] A. Schinzel, *On the number of irreducible factors of a polynomial*, Topics in number theory, (Proc. Colloq., Debrecen, 1974), 305–314, Colloq. Math. Soc. Janos Bolyai, **13**, North-Holland, Amsterdam, 1976.

[184] A. Schinzel, *On the number of irreducible factors of a polynomial. II*, Ann. Polon. Math. **42** (1983), 309–320.

[185] A. Schinzel, *Polynomials with special regard to reducibility*, Encycl. Math. Appl. **77**, Cambridge Univ. Press, 93pp, 2000.

[186] A. Schinzel, *Self-inversive polynomials with all zeros on the unit circle,* Ramanujan J., **9** (2005), 19–23.

[187] A. Schinzel, *On the reduced length of a polynomial with real coefficients,* Funct. Approximatio, Comment. Math., **35** (2006), 271–306.

[188] A. Schinzel,, *On the reduced length of a polynomial with real coefficients. II.* Funct. Approximatio, Comment. Math. **37** (2007), Part 2, 445–459.

[189] A. Schinzel, *The reduced length of a polynomial with complex coefficients*, Acta Arith. **133** (1) (2008), 73–81.

[190] H.-P. Schlickewei, *Multiplicities of recurrence sequences*, Acta Math., **176** (1996), 171–243.

[191] H.-P. Schlickewei, *The multiplicity of binary recurrences*, Invent. Math., **129** (1997), 11–36.

[192] W.M. Schmidt, *The zero multiplicity of linear recurrence sequences*, Acta Math., **182** (1999), 243–282.

[193] W. M. Schmidt, *Zeros of linear recurrence sequences*, Publ. Math. Debrecen, **56** (2000), 609–630.

[194] W.M. Schmidt, *Linear recurrence sequences*, in: Diophantine Approximation, Lectures given at the C.I.M.E. summer school, Cetraro, Italy, June 28–July 6, 2000, (Amoroso, F. (ed.) et al.), Berlin, Springer, Lect. Notes Math. 1819 (2003), pp. 171–247.

[195] E.S. Selmer, *On the irreducibility of certain trinomials*, Math. Scand. **4** (1956), 281–286.

[196] C.D. Sinclair and J.D. Vaaler, *Self-inversive polynomials with all zeros on the unit circle,* in: *Number theory and polynomials*, Proceedings of the conference held at the University of Bristol, 3-7 April, 2006, (eds. James McKee and Chris Smyth), LMS Lecture Note Series 352 (2008), pp. 312–321.

[197] H.-P. Skoruppa, *Heights*, Graduate Lecture course, Bordeaux, 1999.

[198] M.F. Smiley, *On the zeros of a cubic recurrence*, Amer. Math. Monthly, **63** (1956), 171–172.

[199] C.J. Smyth, *On the product of the conjugates outside the unit circle of an algebraic integer,* Bull. London Math. Soc., **3** (1971), 169–175.

[200] C.J. Smyth, *Some results on Newman polynomials,* Indiana Univ. Math. J., **34** (1985), 195–200.

[201] C.J. Smyth, *Mahler measure of algebraic numbers: a survey,* Number theory and polynomials, 322–349, London Math. Soc. Lecture Note Ser., **352**, Cambridge Univ. Press, Cambridge, 2008.

[202] J. Steuding, *Diophantine Analysis*, Chapman and Hall/CRC, 2005.

[203] E.A. Storozhenko, *A problem of Mahler on the zeros of a polynomial and its derivative,* Sb. Math., **187** (1996), 735–744.

[204] O. Strauch and Š. Porubský, *Distribution of sequences: A sampler,* Schriftenreihe der Slowakischen Akademie der Wissenschaften 1, Peter Lang, Frankfurt, 2005.

[205] F. Supnick, H.J. Cohen and J.F. Keston, *On the powers of a real number reduced modulo one*, Trans. Amer. Math. Soc., **94** (2) (1960), 244–257.

[206] G. Szegö, *Orthogonal polynomials,* AMS, Providence, 1975.

[207] R. Turyn, *On Barker codes of even length,* IEEE Trans. Inform. Theory **51** (1963), no. 9, 1256.

[208] R. Turyn, *Character sums and difference sets*, Pacific J. Math. **15** (1965), 319–346.

[209] R. Turyn, *Sequences with small correlation,* Error Correcting Codes (Proc. Sympos. Math. Res. Center, Madison, Wis.), 195–228, John Wiley, New York, 1968.

[210] R. Turyn, J. Storer, *On binary sequences*, Proc. Amer. Math. Soc. **12** (1961), 394–399.

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

[212] M. Waldschmidt, *Diophantine approximation on linear algebraic groups. Transcendence properties of the exponential function in several variables*, Grundlehren der Mathematischen Wissenschaften **326**, Springer-Verlag, Berlin, 2000.

[213] W.A. Webb, H. Yokota, *Polynomial Pell's equation*, Proc. Amer. Math. Soc. **131** (2002), 993–1006.

[214] H. Weyl, *Über die Gleichverteilung von Zahlen modulo Eins,* Math. Ann., **77** (1916), 313–352.

[215] F. Wright, *Computing with Maple*, CRC Mathematics Series, Chapman Hall, 2001.

[216] H. Zassenhaus, *On Hensel Factorization*, I. J. Number Theory, **1** (1969), 291–311.

[217] H. Zassenhaus, *A Remark on the Hensel Factorization Method*, Math. Comp., **32** (1978), 287–292.

[218] A. Zygmund, *A remark on conjugate series,* Proc. London Math. Soc., **34** (1932), 392–400.