On Mahler Measures and Digraphs

Joshua Coyston

Submitted in partial fulfillment of the requirements of the Degree of Doctor of Philosophy

Department of Mathematics

Royal Holloway, University of London

Date of Submission: December 9, 2021

Declaration of Authorship

I, Joshua Coyston, hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signature:

Date: December 9, 2021

Acknowledgements

Firstly, I would like to thank my supervisor, James McKee, for his continuous support and guidance. From his initial flexibility to taking me on as a student, to his patience whilst I reacquainted myself to mathematics, to the depth and breadth of his knowledge and to the amount of time he dedicated to the work contained here and around it, he has been a wonderful supervisor. And, of course, his help in proof reading and editing has been invaluable. This PhD would not have been as successful as it is without him.

I would also like to thank my advisor, Martin Widmer, for his useful insights throughout the process and always being on hand if needed. Moreover, thank you to my examiners – Jozef Skokan and Rainer Dietmann – for their useful insights and corrections about this thesis, and to them and Simon Blackburn for making my viva an enjoyable experience. I extend further thanks to all the academics I had contact with within the department during my time here, as well as the admin, IT and cleaning teams associated with the department.

I am grateful to fellow PhD students Nick, Reynold, Maya and Tom particularly for many useful mathematical conversations during my time. Certainly some of the ideas I have had have come from those conversations. I am also grateful to the help of Jeroen Sijsling – though he only supervised me for a short summer project over seven years ago, he has always been on hand to give words of wisdom and encouragement, as well as mathematical help, since.

Further thanks to the core team members of the Lockdown Quiz team I was a part of – Erin, Ben, Simon and Lenka – for many useful distractions during the pandemic for which much of my PhD was undertaken under (and also for being the best team, obviously!). Moreover, I extend gratitude to everyone in both the original 253 office, and the Bedford Building office, for their warmth, support and many useful distractions...a group of people sadly too long to exhaustively list.

My thanks go to my travel companions (and now former flatmates) Adam and Jack. Though our theme park trip plans were abruptly paused by several lockdowns, even the planning of many blue sky trip ideas kept me motivated and excited for the future when the going got tough. Highlight trips during my PhD include our 2019 Halloween trip to Germany and The Netherlands (I give a nervous, knowing nod to any reader who is familiar with *The Clinic* at Walibi Holland), our 2019 Helsinki trip and the strange 2020 Soquet-chasing French trip. Also, I am appreciative of “Team XScream” – in particular, Jay, Oli and Lee. Their flexibility and support has been nothing short of wonderful, and of course we have had many successful trips to a variety of escape rooms (including the record-destroying 18 minute 11 second run through a room which should have taken an hour!).

I also take this opportunity to express gratitude and thanks to James Batney, Giles Carey and Alex Green. They likely do not realise it, but conversations I have had with them in the past have either truly stuck with me, or given me some sort of inspiration, which helped me complete this journey.

Finally, I would like to thank my Mum and my little brother Zachary. My Mum’s constant support (even if she did refer to this PhD in Pure Mathematics as a “Masters in Statistics” for the best part of two years) and help is something which cannot be overlooked or understated. I probably would not begin to understand the level of help she has given me. And I cannot express how much Zachary’s love and admiration for me has meant, and what a drive it is for me on a daily basis.

To all of you,

Thank you.

Abstract

This thesis is based in the intersection of number theory and graph theory, introducing a link between the Mahler measure of a polynomial and specific types of graphs, which we refer to as digraphs. One of the main aims of this thesis is to find "small" Mahler measure values "from" digraphs. Our other aims are to further extend our knowledge of the digraphs presented, the theory linking Mahler measures and graphs more broadly and, finally, understanding when polynomials may share the same Mahler measure.

As such, this thesis has two distinctive themes which are intertwined together. In parts of the thesis, primarily Chapters 1, 3 and 6, our focus is more theoretical, and we aim to introduce and survey existing ideas, both from the topics of Mahler measures and digraph theory, as well as build upon these and present new concepts. Other parts, particularly Chapters 2 and 5, are focused on calculating Mahler measure values from a practical viewpoint, including introducing a new method for calculating the Mahler measures of two-variable polynomials, as well as detailing how exactly we performed our experiments.

Meanwhile, Chapter 4 is a blend of these themes: whilst it focuses on developing our theoretical knowledge, it is done in a way to prepare us for our experiments. We round off the thesis in Chapter 7 by summarizing potential future work.

In most circumstances, we present Mahler measure values to 8 decimal places, though this convention is broken on occasions where appropriate.

Contents

List of Notation

A list of standard notation used throughout this thesis:

χ_G Characteristic polynomial of the graph G

$\deg(\mathcal{P})$ The total degree of the polynomial $\mathcal{P}$

$\deg_{z_i}(\mathcal{P})$ The degree of the largest z_i term in the polynomial $\mathcal{P}$

$\ell(P)$ The term-length of P

$\mathrm{Gal}(f)$ The Galois group of the polynomial f

λ $= 1.1762808\cdots$; Lehmer's number

$\mathbb{P}(X = n)$ The probability that $X = n$

$\mathcal{L}$ A lattice

$\mathcal{P}$ A polynomial in n variables

$\mu_i(A)$ The i-th eigenvalue of the matrix A

$\mu_\alpha(x)$ The minimal polynomial of α

Φ_n The n-th cyclotomic polynomial

Π A parallelogram

$\mathrm{prim}(\mathcal{P})$ The polynomial $\mathcal{P}$ strongly primitized

θ $= 1.3247179\cdots$; the smallest Pisot number

A_G Adjacency matrix of the graph G

D_n The dihedral group of order $2n$

f A single variable polynomial

$G \cdot H$ The coalescence between G and H

G A graph

$G - [uv]$ The graph obtained by deleting the (multi)edge between vertices u and v

$G - u$ The graph obtained by deleting the vertex u

G^* The graph obtained by contracting a specified (multi)edge

$G_{(m)}$ The graph obtained by adding an edge between two specified vertices, and subdividing it by m vertices

$G_{[t_1,\cdots,t_r]}$ A graph with r subdivided edges, where the i-th edge is subdivided by t_i vertices

$G_{m_1,\cdots,m_k}$ A graph attached with k pendant paths, where the i-th path has m_i vertices

G_m A graph attached with a pendant path with m vertices

$G^n_{uv}H$ The n-bridge of G and H

$G_{uv}H$ The bridge of G and H

$M(\cdot)$ The Mahler measure of $\cdot$

$m(\cdot)$ The logarithmic Mahler measure of $\cdot$

P A two-variable polynomial

$P_K(k_1, k_2, \cdots, k_n)$ A pretzel knot with n tangles

P_n The path graph on n vertices

R_G The reciprocal polynomial of a graph G

$V(G)$ The set of vertices of a graph G

V_4 The Klein-4 group

List of Figures

List of Tables

Chapter 1

Introduction and Motivation

In this chapter we introduce the core definitions and ideas surrounding Mahler measures, both in one variable and in several variables. We then move on to explain our motivation for introducing combinatorial objects, and outline the rest of thesis.

1.1 Mahler Measures and Single Variable Polynomials

Mahler measures are closely related to height functions, in that they are a way of measuring the complexity of a polynomial. This thesis focuses on the study of them in their own right, but there exist applications to other areas of maths, notably dynamical systems (an overview of which can be seen in, for example, Smyth [41]).

Definition 1.1.1. *The* ***Mahler measure*** *of a non-zero polynomial* $f \in \mathbb{C}[x]$ *is defined as:*

$$M(f) := \exp\left(\int_0^1 \log|f(e^{2\pi i t})| \,\mathrm{d}t\right). \tag{1.1}$$

That is to say that the Mahler measure of a polynomial is the "geometric mean" of $|f(z)|$, for z on the unit circle.

In some situations, we may make reference to the *logarithmic Mahler measure*, which is written as $m(f)$. This is defined as one would perhaps expect: take logarithms of both sides of the above definition, giving us $m(f) = \log M(f)$. This can sometimes be more convenient to work with. Equally, some may define the Mahler measure as:

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

We note that this is the same as (1.1) in Definition 1.1.1, with the substitution $\theta = 2\pi t$. Throughout this thesis, we may make use of either of these formulations when necessary. Furthermore, any reference to the Mahler measure will specifically be referring to $M(f)$. If we make use of the logarithmic Mahler measure, we will make it explicit that we are doing so.

There is, however, an alternative way to view the Mahler measure, which is perhaps easier to visualise in practice.

Proposition 1.1.2. *The Mahler measure of a single variable polynomial $f(x)$ can be written as:*

$$M(f) = |a_n| \prod_{i=1}^{n} \max\{1, |\alpha_i|\}, \tag{1.2}$$

where a_n is the coefficient of the leading term and α_i are the zeros of f (where a zero of multiplicity k is counted k times).

This was, in fact, the original definition from Lehmer [17]. This means that we can view the Mahler measure of a polynomial as the product of the absolute values of the zeros of that polynomial which lie outside of the unit circle (multiplied by the modulus of the lead coefficient).

To prove that this is indeed equivalent to (1.1), we need to make use of the following:

Proposition 1.1.3 (Jensen's Formula)**.** *Let $F \in \mathbb{C}[z]$ be an analytic function for $|z| \le r$, for some $r \in \mathbb{R}_{>0}$, such that $F(0) \neq 0$. Let $\zeta_1, \zeta_2, \cdots, \zeta_N$ be the zeros of $F(z)$ such that $|\zeta_i| \le r$ for each i (where a zero of multiplicity b is listed b times).*

Then we have:

$$\frac{1}{2\pi}\int_0^{2\pi} \log|F(re^{i\theta})|\,\mathrm{d}\theta = \log|F(0)| + \sum_{i=1}^{N} \log\frac{r}{|\zeta_i|}.$$

Remark. The statement provided here is from Titchmarsh [45, p. 125], though essentially equivalent phrasings of the statement exist elsewhere with some superficial differences. For a particularly clear proof of Jensen's Formula, see Spiegel [43].

Proof of Proposition 1.1.2. Let $F(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$ in Jensen's Formula, and assume that $a_n \neq 0$ and $a_0 \neq 0$, meaning $\deg(F) = n$ and $F(0) \neq 0$. We note that these assumptions are reasonable to make, since any polynomial can be factored to this form, and we can then consider the other term (which would be a power of z) trivially by itself. Let $\zeta_1, \cdots, \zeta_n$ be the zeros of F.

Numbering these zeros appropriately, we can write them in the following way:

$$|\zeta_1| \leq |\zeta_2| \leq \cdots \leq |\zeta_N| \leq 1 < |\zeta_{N+1}| \leq |\zeta_{N+2}| \leq \cdots \leq |\zeta_n|,$$

for some $1 \leq N \leq n$.

Now set $r = 1$ in Jensen's Formula. This means that the integral side of the formula is now equal to the logarithmic Mahler measure of f, $m(f)$. As such, we now have:

$$\begin{aligned} m(f) &= \log|a_0| + \sum_{i=1}^{N} \log \frac{1}{|\zeta_i|} \\ &= \log \left| \frac{a_0}{\zeta_1 \zeta_2 \cdots \zeta_N} \right|. \end{aligned}$$

We note that $|a_0| = |a_n \zeta_1 \cdots \zeta_n|$, meaning that:

$$\begin{aligned} m(f) &= \log \left| \frac{a_n \zeta_1 \zeta_2 \cdots \zeta_N \zeta_{N+1} \cdots \zeta_n}{\zeta_1 \zeta_2 \ldots \zeta_N} \right| \\ &= \log|a_n \zeta_{N+1} \zeta_{N+2} \cdots \zeta_n| \\ &= \log|a_n| + \log|\zeta_{N+1} \zeta_{N+2} \cdots \zeta_n| \\ &= \log|a_n| + \sum_{i=1}^{n} \log^+|\zeta_i| \ , \end{aligned} \tag{1.3}$$

where

$$\log^+ a = \begin{cases} \max\{\log a, 0\}, & \text{if } a > 0\,, \\ 0, & \text{if } a = 0\,. \end{cases}$$

We note here that since $|\zeta_1| \leq \cdots \leq |\zeta_N| \leq 1$, then $\log^+|\zeta_j| = 0$ for $1 \leq j \leq N$, meaning they do not contribute to the value of $m(f)$ in (1.3). Equally, we see that for

$a \leq 1$, $\exp(\log^+ a) = 1$, since $\log^+ a = 0$ in this case.

We now take exponentials on both side of (1.3) to get:

$$\exp(m(f)) = \exp\left(\log|a_n| + \sum_{i=1}^{n} \log^+|\zeta_i|\right),$$
$$M(F) = |a_n| \prod_{i=1}^{n} \max\{1, |\zeta_i|\}.$$

By identifying F and f as the same polynomial, and rewriting the zeros as $\zeta_i = \alpha_i$, this completes the proof. □

Remark. This relationship was noted by Mahler, and the proof given here was based on work of Mahler [19] and [20].

A simple, but useful, relation regarding the Mahler measure is that it is multiplicative:

Lemma 1.1.4. *For polynomials $f, g \in \mathbb{C}[x]$, $M(fg) = M(f)M(g)$.*

Proof. To make use of (1.2), we let $f(x) = a_n \prod_{i=1}^{n}(x - \alpha_i)$ and $g(x) = b_m \prod_{j=1}^{m}(x - \beta_j)$. Then, we can write
$fg(x) = c \prod_{k=1}^{n+m}(x - \gamma_k)$, where $c = a_n b_m$ and:

$$\gamma_k = \begin{cases} \alpha_k, & \text{for } k = 1, \cdots, n, \\ \beta_{k-n}, & \text{for } k = n+1, \cdots, n+m. \end{cases}$$

Then, we have:

$$\begin{aligned} M(fg) &= |c| \prod_{k=1}^{n+m} \max\{1, |\gamma_k|\} \\ &= |a_n b_m| \left(\prod_{k=1}^{n} \max\{1, |\gamma_k|\}\right) \left(\prod_{k=n+1}^{n+m} \max\{1, |\gamma_k|\}\right) \\ &= |a_n| \left(\prod_{i=1}^{n} \max\{1, |\alpha_i|\}\right) |b_m| \left(\prod_{j=1}^{m} \max\{1, |\beta_j|\}\right) \\ &= M(f)M(g). \end{aligned}$$

□

1.1.1 Integer and Cyclotomic Polynomials

The focus on this thesis relates mainly to integer polynomials, especially when it comes to explicit calculations, although some of our results may be more general. As such, unless explicitly stated, it should be assumed that any results given exist for all complex coefficient polynomials. However, much of the study of Mahler measures has focused specifically on integer coefficient polynomials, $f \in \mathbb{Z}[x]$, and here we outline some key points from this specific section of study.

Definition 1.1.5. *For $n \geq 1$, the* ***n-th cyclotomic polynomial****, denoted by $\Phi_n(z)$, is the minimal polynomial of* $\exp(\frac{2\pi i}{n})$.

Definition 1.1.6. *A polynomial $f \in \mathbb{Z}[z]$ is said to be* ***Kronecker-cyclotomic*** *if it is the product of some number of $\Phi_n(z)$ and/or a power of z.*

Remark. This choice of nomenclature is due to a classification from Kronecker, seen below in Lemma 1.1.8.

Some authors may simply refer to Kronecker-cyclotomic polynomials as *cyclotomic polynomials*. To avoid any confusion between 'the' (n-th) cyclotomic polynomial and 'a' cyclotomic polynomial, we instead use Kronecker-cyclotomic to create a more explicit difference.

Lemma 1.1.7. $M(\Phi_n(z)) = 1$ *for* $n \geq 1$.

Proof. The n-th cyclotomic polynomial can be written as

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

For any such k and n, it is clear that $|e^{\frac{2\pi ik}{n}}| = 1$, and so $M(f) = 1$. □

Lemma 1.1.8. $M(f) = 1$ *if and only if f is Kronecker-cyclotomic.*

Proof. We know that the Mahler measure is multiplicative. Equally, it is easy to see that $M(z) = 1$. So it is easy to see if f in Kronecker-cyclotomic, then $M(f) = 1$.

The converse is a consequence of Kronecker's Theorem, a proof of which comes from Kronecker [15], with more convenient proofs from, for example, McKee [21] or McKee and Smyth [25]. □

Definition 1.1.9. *We say that a Mahler measure value is* ***non-trivial*** *if it is greater than* 1*.*

It is currently unknown if there is a smallest, non-trivial Mahler measure. This is what is known as *Lehmer's Problem*, and the following conjecture is associated with Lehmer [17]:

Conjecture 1.1.10 (Lehmer's Problem)**.** *For all polynomials $f \in \mathbb{Z}[x]$ such that $M(f)$ is non-trivial, we have that $M(f) \geq 1.17628081\cdots = \lambda$.*

The value λ comes from the polynomial $\Lambda(x) = x^{10}+x^9-x^7-x^6-x^5-x^4-x^3+x+1$; usually referred to as ***Lehmer's number*** *and* ***Lehmer's polynomial*** *respectively.*

This is also sometimes known as the *Strong Lehmer Conjecture*. In turn, there is a weaker version of this conjecture:

Conjecture 1.1.11. *There exists an absolute constant $c > 1$ such that for all polynomials $f \in \mathbb{Z}[x]$, where $M(f)$ is non-trivial, $M(f) \geq c$.*

Work in trying to answer Lehmer's Problem can hence go in a couple of different directions. One of these directions would involve trying to find a polynomial with Mahler measure $\lambda_1 < \lambda$. This would then change the question with regards to the strong version. Another direction would be to find conditions on c and see if an absolute constant can be found as a result.

Partial progress has been made on the latter. Due to work from Dobrowolski [12], and later from Voutier [47], we know that for polynomials f of sufficiently large enough degree d, we have that:

$$M(f) > 1 + \frac{1}{4}\Big(\frac{\log\log d}{\log d}\Big)^3.$$

1.1.2 Small Mahler Measures

Lehmer's Problem has been one of the biggest motivators for studying Mahler measures for many, and as such, much work has spawned from this. One particular object that has arisen as a result is the study of polynomials which admit Mahler measures which are comparatively *small.* From a crude viewpoint, understanding when polynomials have Mahler measures which are "small" (and, as such, close to λ, the smallest known Mahler measure) could well help us prove, or even possibly disprove, Lehmer's Problem.

Before we formalise the definition of a small Mahler measure, we must first look at some other number theoretic objects: Salem and Pisot numbers.

Salem and Pisot Numbers

Definition 1.1.12. *For a polynomial $f(z)$, we define its* ***reverse polynomial*** *as $f^*(z) := z^{\deg(f)} f(1/z)$.*

We say that f is ***reciprocal*** *if $f = f^*$.*

Remark. The definitions presented in Definition 1.1.12 are known under different names by other authors. Some may refer to the reverse polynomial as a reciprocal polynomial, and in turn a reciprocal polynomial as a *self-reciprocal* polynomial. Others may call the here-presented reciprocal polynomial a *palindromic* polynomial.

The use of reverse and reciprocal here is most convenient for the work presented throughout this thesis, hence the reason for their names.

Definition 1.1.13. *A* ***Salem number*** *is a real, algebraic integer $\alpha > 1$, such that:*

- *its Galois conjugates (other than itself) all have absolute value at most 1,*
- *at least one of its Galois conjugates has absolute value exactly 1.*

Lemma 1.1.14. *Let α be a Salem number, and $\mu_\alpha(x)$ be its minimal polynomial. Then, μ_α is reciprocal.*

Proof. We know that there is some zero of μ_α, say β, which has absolute value 1. As $|\beta| = 1$, then we know we can write $\overline{\beta} = 1/\beta$. Furthermore, we note that $\mu_\alpha(\overline{\beta}) = \overline{\mu_\alpha(\beta)} = 0$. Thus, $\mu_\alpha(1/\beta) = 0$.

Now consider the reverse polynomial of μ_α, which we know we can write as $\mu_\alpha^*(x) = x^d\mu_\alpha(1/x)$, where $d = \deg(\mu_\alpha)$. We can see that gcd $\big(\mu_\alpha(x), x^d\mu_\alpha(1/x)\big)$ is not trivial, hence we can write $\mu_\alpha(x) = \nu x^d \mu_\alpha(1/x)$, for some ν.

Finally, if we set $x = 1$, we see that $\nu = 1$. So, this means that $\mu_\alpha(x) = x^d\mu_\alpha(1/x)$, and hence μ_α is reciprocal. □

Lemma 1.1.15. *If α is a Salem number, then all of its Galois conjugates, except itself and $1/\alpha$, have absolute value* 1.

Proof. Consider some conjugate of α, say γ, which we assume not to be $1/\alpha$. As α is Salem, we know that $|\gamma| \leq 1$.

Now assume $|\gamma| < 1$, and let $\mu_\alpha(x)$ be the minimal polynomial of α. As γ is a Galois conjugate of α, it is a zero of μ_α. By Lemma 1.1.14, we know that μ_α is reciprocal, and so $1/\gamma$ is also a zero of μ_α. However, this means that $|1/\gamma| > 1$, contradicting the fact that all conjugates of α (except itself) have absolute value at most 1.

So, any Galois conjugate of α, except itself and $1/\alpha$, has absolute value exactly 1. □

Example. Consider λ, whose minimal polynomial is Lehmer's polynomial: $\Lambda(x) = x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1$. Let us look at the roots of $\Lambda(x) = 0$ plotted in relation to the unit circle:

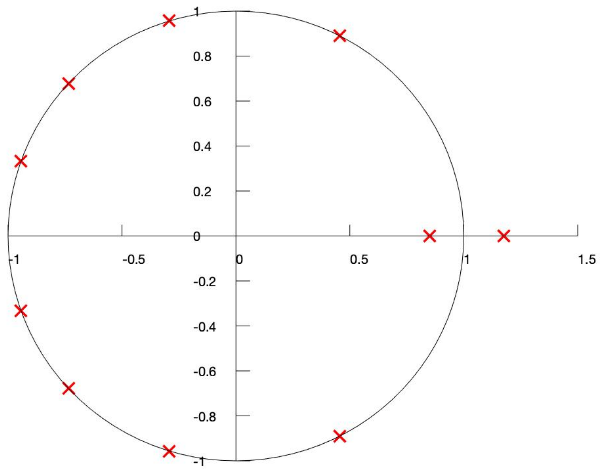


Figure 1.1: The roots of $\Lambda(x) = 0$.

We see that the only roots to not lie on the unit circle are λ and $1/\lambda$. So, this means in particular that the smallest known Mahler measure value is also a Salem number (and, in fact, is the smallest known Salem number).

Definition 1.1.16. *A **Pisot number** (sometimes called a **Pisot-Vijayaraghavan number**) is a real, algebraic integer $\alpha > 1$ such that all of its Galois conjugates (other than itself) have absolute value less than* 1.

Salem [36] showed that there is a smallest Pisot number, which was soon found by Siegel [37]:

Example. The smallest Pisot number is the real root of $f(x) = x^3 - x - 1$. This is $\theta = 1.32471795\cdots$.

Unlike with Salem numbers, the corresponding minimal polynomials of Pisot numbers are not reciprocal, except in some special quadratic cases, where they are of the shape $x^2 + ax + 1$, for $a \in \mathbb{Z}$ and $a \leq -3$.

It is perhaps not surprising that Salem and Pisot numbers are closely related, and that their study is closely related. Smyth [42], for example, offers detailed insight related to this topic.

Small Mahler Measures

As has been noted, the smallest known Mahler measure and smallest known Salem number are the same. Furthermore, the study of Mahler measures and Salem numbers can be linked very closely, due to the fact that the Mahler measure of the minimal polynomial of any Salem number is itself the Salem number (and similarly for the minimal polynomial of any Pisot number).

When studying Salem numbers, it has been of interest to some to find *small Salem numbers*, which are defined as Salem numbers which are less than 1.3. One reason, albeit not particularly rigorous, for choosing 1.3 as the bound is because it is slightly smaller than $\theta \approx 1.32$, the smallest Pisot number. A more concrete justification for choosing this bound is because it is conjectured that the smallest limit point of the set of Salem numbers is also θ (see Boyd [1]).

Definition 1.1.17. *The Mahler measure of an integer polynomial is* ***small*** *if it is non-trivial and less than* 1.3*.*

There are infinitely many small Mahler measures; this is discussed by Boyd and Mossinghoff [5], and we give a level of justification for this in Heuristic Proposition 1.2.7. However, we can look at the number of Mahler measures that occur when we restrict the degree of the polynomial, and only look at irreducible polynomials when doing so. For example, many authors have looked at polynomials of degree up to 180. Currently, there are 8574 small Mahler measures known up to degree 180; with online lists split across Mossinghoff [26] and Sac-Épée [34], and a newly found one by Coyston and McKee [9]. Additionally, all small Mahler measures are known up to degree 44, and the strong version of Lehmer's conjecture has been verified for all polynomials up to degree 56, by Mossinghoff et al. [27].

Whilst the strong version of Lehmer's conjecture has only been proven for a small bounded degree, the weak version is actually much easier to show for any bounded degree.

Proposition 1.1.18. *For all possible polynomial degrees D, there exists some constant $c_D > 1$ such that for all $f \in \mathbb{Z}[x]$ with $\deg(f) \leq D$ and where $M(f)$ is non-trivial, we have that $M(f) \geq c_D$.*

Proof. It is enough to show that there are only finitely many polynomials that have degree at most D and have Mahler measure in the interval $(1, \lambda)$. Any such polynomial must be monic. So, fix some target Mahler measure value for our polynomial f which is less than λ. A consequence of this is that f must therefore be monic.

All zeros of f must then have moduli less than λ. In turn, all moduli of all of the coefficients of f are bounded, because they are all sums of products of bounded numbers of zeros; this bound depends on D.

Hence, the number of possible polynomials f such that $1 < M(f) < \lambda$ is bounded. One can exhaustively check all of these possibilities to see if such a polynomial exists. □

One sensible question to ask is what types of polynomials may potentially yield a small Mahler measure. Fortunately, the following result gives us an answer to this question:

Theorem 1.1.19 (Smyth [39]). *If $M(f) < \theta$, then f is reciprocal.*

A weaker bound of this result was originally stated by Breusch [6], but the result of Smyth here is the best possible bound. Though this does not mean that all reciprocal polynomials admit small Mahler measures, as seen below, it gives a narrower collection of polynomials to search over when trying to find small Mahler measure values.

Example. Let $f(x) = x^6 + x^5 - x^4 + x^3 - x^2 + x + 1$. Then, $M(f) = 1.94685626\cdots$.

Another collection of polynomials we must restrict ourselves to when trying to find small Mahler measures is monic polynomials. As a result of (1.2), the Mahler measure of f is at least the value of the absolute value of its leading coefficient, meaning any non-monic (integer) polynomial will have a Mahler measure at least 2.

1.1.3 Mahler Measures from Matrices

Mahler measures are traditionally only associated to polynomials. However, we can in fact associate Mahler measures to other objects; matrices, for example.

Definition 1.1.20. *Let A be an $n \times n$ matrix with complex coefficients, with characteristic polynomial $\chi = \chi_A(x)$. Define the "reciprocalisation" of χ as $R_\chi(z) = z^{\deg(\chi)}\chi(z + 1/z)$. Then, we define the* ***Mahler measure of the matrix*** *A as $M(R_\chi)$.*

Remark. We refer to $R_\chi = R$ as the *reciprocal polynomial* of the matrix A.

Taking the reciprocalisation of the characteristic polynomial gives us a reciprocal polynomial, twice the degree of the characteristic polynomial. When first defined by McKee and Smyth [22], this was done by design, as it allows for a better chance to find small Mahler measures. It is once again worth stressing, though, that a polynomial being reciprocal does not guarantee that its Mahler measure will be small.

Example. Let:

$$A = \begin{pmatrix} 2 & 3 & 1 & 0 \\ 1 & -2 & 0 & 1 \\ 0 & -1 & -3 & 5 \\ 3 & 0 & 1 & 0 \end{pmatrix}.$$

Then:

$$\chi_A(x) = x^4 + 3x^3 - 12x^2 - 43x - 21,$$
$$R_\chi(z) = z^8 + 3z^7 - 8z^6 - 34z^5 - 39z^4 - 34z^3 - 8z^2 + 3z + 1,$$
$$M(A) = M(R_\chi) = 24.83245050\cdots.$$

As such, it is worth making careful choices for what type of matrices we consider, if seeking small Mahler measures. There are many ways to do this, but the methods explored in this thesis will look at finding the adjacency matrices of special types of graphs: we explore this is greater detail in chapter 3.

1.2 Mahler Measures and n-Variable Polynomials

We can also define the Mahler measure of an n-variable polynomial.

Definition 1.2.1. *Let $\mathcal{P}(z_1, \cdots, z_n) \in \mathbb{C}[z_1, \cdots, z_n]$. The* ***Mahler measure*** *of the polynomial $\mathcal{P}$ is defined as:*

$$M(\mathcal{P}) := \exp\left(\int_{t_1=0}^{1} \cdots \int_{t_n=0}^{1} \log|\mathcal{P}(e^{2\pi i t_1}, \cdots, e^{2\pi i t_n})| \, \mathrm{d}t_n \cdots \mathrm{d}t_1\right). \tag{1.4}$$

This is a generalisation of (1.1), and it is easy to see if we set $n = 1$ here, we get back Definition 1.1.1. A way of viewing this is as the geometric mean of $|\mathcal{P}|$ over the k-Torus $\mathbb{T}^k$.

A natural question to ask is whether there is an alternative, perhaps simpler, form to view the Mahler measure of an n-variable polynomial, as in the single variable case. Unfortunately, the short answer here is that there is not. The main reason for this is because (1.2), the "simpler" form of viewing Mahler measures for single variable polynomials, is found due to Jensen's Formula, which relates to single variable polynomials, and there is no analogue of this in multiple variables.

That is not to say there is nothing we can do. In chapter 2, we will explore ways in which we can simplify the viewpoint, particularly when it comes to calculations of Mahler measures of certain two-variable polynomials.

1.2.1 Integer and Cyclotomic Polynomials (Again)

As was the case with single variable polynomials, a lot of focus will remain on integer polynomials, though some results and methods may actually be generalised further. Here, we outline some analogous statements to those of single variable polynomials.

Definition 1.2.2. *A polynomial $\mathcal{P} \in \mathbb{Z}[z_1, \cdots, z_m]$ is said to be* ***m-Kronecker-cyclotomic*** *if its Mahler measure is* 1*.*

Remark. Following from this, a refinement to Definition 1.1.6 would see us call such a single variable polynomial that is the product of some number of $\Phi_n(z)$ and/or a power of z as **1-Kronecker-cyclotomic**. However, for simplicity, we will still refer to these simply as **Kronecker-cyclotomic**.

The following result from Boyd [3] gives us another way of looking at the Mahler measure of a particular, special, two-variable polynomial:

Proposition 1.2.3. $\log M(1+x+x^n) = \log M(1+x+y) + \frac{c(n)}{n^2} + O(\frac{1}{n^3})$*, where:*

$$c(n) = \begin{cases} \frac{-\sqrt{3}\pi}{6}, & \text{if } n \equiv 2 \mod 3\,, \\ \frac{\sqrt{3}\pi}{18}, & \text{if } n \equiv 0,1 \mod 3\,. \end{cases}$$

So, we can view the value of $\log M(1+x+y)$ as a true limit point of the values $\log M(1+x+x^n)$ as $n \to \infty$.

This, however, is a very special situation, and in general, the true Mahler measure values of n-variable polynomials are not known exactly. However, Boyd [3] also provides a broader statement which is very useful to us in practice:

Proposition 1.2.4. $\log M(P(x,y)) = \lim_{n\to\infty} \log M(P(x,x^n))$.

And, in fact, Lawton [16] generalised this result for polynomials of n-variables.

This gives us a significantly simpler way of viewing the values of the Mahler measure of n-variable polynomials with reasonably good accuracy. As noted, it is only in a very small selection of cases that we do know the true limit point, as in Proposition 1.2.3. In this thesis, we do not delve into the discussions of the exact values of Mahler measures in n-variables (though further discussion of this can be found by, for example, Boyd and Mossinghoff [5]).

Fortunately, m-Kronecker-cyclotomic polynomials are relatively easy to spot. They look like Kronecker-cyclotomic polynomials, but with the single variable replaced with a monomial, as formalised by Boyd [2]:

Proposition 1.2.5. *For each m, let K_m be the set of polynomials which are products of $\pm {z_1}^{c_1} \cdots {z_m}^{c_m}$ and polynomials of the form $\Phi_a({z_1}^{d_1} \cdots {z_m}^{d_m})$, for $a \geq 1$ and where $c_i, d_i \in \mathbb{Z}_{\geq 0}$ for each i. Then, for $\mathcal{P} \in \mathbb{Z}[z_1, \cdots, z_m]$, $M(\mathcal{P}) = 1$ if and only if $\mathcal{P} \in K_m$.*

For example, we know $X^2 - 1$ is Kronecker-cyclotomic (it is the product of $\Phi_1(X) = X - 1$ and $\Phi_2(X) = X + 1$), so it is easy to see that $x^2y^8 - 1$ is 2-Kronecker-cyclotomic, by the monomial replacement $X = xy^4$.

It should also come as no surprise that we have an extension to Lemma 1.1.4, and this would also be another useful tool for finding 2-Kronecker-cyclotomic polynomials.

Lemma 1.2.6. *For polynomials $P, Q \in \mathbb{C}[x, y]$, $M(PQ) = M(P)M(Q)$.*

We omit details of the proof, though again, this result is not surprising.

Currently, there are only two known Mahler measures less than 1.3 coming from two-variable polynomials: $1.25543386 \cdots$ and $1.28573486 \cdots$. However, finding two-variable polynomials which give Mahler measure less than 1.37 is of interest. Choosing 1.37 as a bound may seem arbitrary at first. However, there is some reasoning behind this.

It is conjectured that the smallest Mahler measure from a three-variable polynomial is $1.38135 \cdots$, from the polynomial $\mathcal{P}(x, y, z) = (x^2 + 1)yz + (y^2 + 1)xz + (z^2 + 1)yz$. Noting that Mahler measures from three-variable polynomials will be limit points of limit points of Mahler measures, we would expect to find a large number Mahler measure values coming from two-variable polynomials around this value. As such, taking a value comfortably below this ensures we will find a restricted number of values, keeping them of interest.

Incidentally, this logic could suggest a refinement to our definition of small Mahler measures coming from single variable polynomials. As mentioned, the smallest known Mahler measure coming from two-variable polynomials is $1.25543386 \cdots$. As this can be seen as a limit point of Mahler measure values, and in particular, one which is not obtained, we can expect to find a infinite number of small Mahler measures from single variable polynomials which will be around this value. As such, taking a value below this will ensure we find a restricted number of values. For example, McKee and Smyth [25] note that there are only 236 known single variable polynomials with Mahler measure below 1.25 (values which they call *tiny Mahler measures*).

Currently, there are 61 known Mahler measures from two-variable polynomials less than 1.37 (referred to as "small limit points of Mahler measures"), which can be found between Mossinghoff [26], Otmani et al. [30] and Sac-Épée et. al. [35].

Before moving on, we use our knowledge of Mahler measures coming particularly from two-variable polynomials to give some level of justification for a claim made earlier:

Heuristic Proposition 1.2.7. *There are infinitely many small Mahler measure values.*

Non-Proof. From Proposition 1.2.4, we can say that the Mahler measure coming from a two-variable polynomial is the limit of a sequence of Mahler measure values coming from single variable polynomials. We also have that the smallest known Mahler measure value from two-variable polynomials is $L = 1.25543386\cdots$. So, for some polynomial $P(x, y)$, we have that for all $\epsilon > 0$, there is some natural number N such that for all $n > N$, we have:

$$|\, M(P(x, x^n)) - L \,| < \epsilon\,.$$

□

Remark. We only list this as a "Heuristic Proposition" and give a "non-proof" because we are skipping some key details. Namely, we do not know if in this case, $L = 1.25543386\cdots$ is a true limit point of the sequence of values $M(P(x, x^n))$. Indeed, this was an issue we raised when introducing Propositions 1.2.3 and 1.2.4.

As such, we do not know if, in our proof above, $M(P(x, x^n))$ becomes constant for all values greater than some number $k \in \mathbb{N}$. Such discussion is beyond the scope of this thesis, but can be seen in greater detail in, for example, Boyd and Mossinghoff [5].

1.3 Motivation

It is striking that the smallest known Mahler measure, λ, comes from a relatively small and straightforward-looking polynomial. As has already been noted, λ is also the smallest known Salem number, and we can see that these two phenomena are closely linked, in the sense that if a smaller Salem number is found, so too is a smaller Mahler measure. And indeed, some of the smallest known Mahler measure values are also Salem numbers.

However, whilst there is a clear link between Mahler measures and the likes of Salem and Pisot numbers, they are perhaps too closely related. By this, we mean that the study of one will likely do little to further deepen our understanding of the other at this point. As such, this leads us to an interesting question: are there links between Mahler measures to other, less-closely related, mathematical areas, that can help deepen our understanding?

1.3.1 An Interlude Featuring Knots

As previously mentioned, the Mahler measure has links and uses to other fields of study, but these do not particularly enrich our understanding of Mahler measures. However, there is an interesting connection between Mahler measures and knot theory.

For the context of this thesis, we do not need to delve too deep into knot theory: see Appendix A for all necessary details. In short, we note that the Alexander polynomials of the pretzel knots $P_K(-2,3,3)$ and $P_K(-2,3,5)$ are cyclotomic polynomials, whilst the Alexander polynomial of the knot $P_K(-2,3,7)$ (which is shown in Figure A.1) is in fact $\Lambda(-t)$. This occurrence can be found in Reidemeister [32], which was published a year before Lehmer published the polynomial in the context of Mahler measures.

What this phenomenon exhibits is a family of topological objects which have an associated polynomial with trivial Mahler measure in the very first, smallest, examples of the family. Then, as we make small changes to the object, we get an associated polynomial with small Mahler measure. In some sense, this means that the small topological changes made to the object are leading us from cyclotomic polynomials to 'almost-cyclotomic' polynomials. These almost-cyclotomic polynomials are the ones which we would expect to have the best chance of having a small Mahler measure.

This connection has been studied, and one key result confirms that λ is the smallest Mahler measure to come from a pretzel knot (see, for example, Hironaka [14]). However, this connection might not be the most fruitful for helping further our understanding of Mahler measures, due to the fact there is no closed formula for calculating the Alexander polynomial of a knot, as discussed briefly in Appendix A. Nevertheless, what this does show is that there is worth in investigating other mathematical objects which have polynomials attached to them.

1.3.2 Combinatorial Inspiration

As noted in Section 1.1.3, McKee and Smyth introduced the notion of the Mahler measure of a matrix, by taking the reciprocalisation of its characteristic polynomial. In particular, McKee and Smyth focused on integer symmetric matrices for attempting to find small Mahler measures. And indeed, much of their focus was on integer symmetric matrices whose entries only took on the values of $-1, 0, 1$.

From here, and with the understanding that Mahler measures could be attached to other mathematical objects, a leap could be made to connect Mahler measures with graphs. We can get integer symmetric matrices from graphs by, for example, taking the adjacency matrix of the graph. This allows us to define the Mahler measure of a graph, as first defined by McKee and Smyth [22]:

Definition 1.3.1. *The **Mahler measure of a graph** G is the Mahler measure of the reciprocal polynomial of its adjacency matrix. That is, if we let $A = A_G$ be the adjacency matrix of G, $\chi = \chi_A(x)$ be the characteristic polynomial of A and $R_G = R_\chi(z)$ be the reciprocal polynomial of χ, then we define the Mahler measure of G as $M(G) := M(R_G)$.*

As had been done with families of knots, McKee and Smyth categorized graphs which have Mahler measure 1: so-called *cyclotomic graphs*. One can then consider ‘nudging’ or ‘growing’ these graphs, by means of adding edges and vertices, creating a graph which is combinatorially close to a cyclotomic graph. The hope from here is that a graph which is combinatorially close to being cyclotomic may be also have a Mahler measure that is small too. However, McKee and Smyth [24] only found five small Mahler measures using this method.

Large parts of this thesis aim to extend the work of McKee and Smyth, looking at more general graphs; these will be introduced in chapter 3. As we will see, looking at a wider variety of graphs means that we will keep, and be able to utilise, many graph theoretic properties to aid us in finding small Mahler measures. However, we also see that we lose some other powerful properties that only can be used with the more regularly used graphs. As such, we introduce suitable ways to work with these types of graphs in chapter 4.

In chapter 5, we then look at experiments which were run utilising the theory of these special types of graphs. We also utilise methods discussed in chapter 2 to explicitly calculate Mahler measures of two-variable polynomials that we can also associate to our combinatorial objects.

1.3.3 Polynomials with the Same Mahler Measure

Owing to the nature of the Mahler measure, work related to this thesis has involved a lot of experimentation and explicit computations. These experiments have primarily made use of PARI/GP [44].

One particular question this raises is the reliability of results we receive. Say we check the Mahler measures of two different polynomials using PARI/GP (or, indeed, any computational software), and they return the "same" values. We cannot conclude with absolute certainty that these polynomials share the same Mahler measure, but rather, that they agree up to the precision that PARI/GP has been set to. Though this is perhaps less of an issue in the case of single variable polynomials, where there are many methods for checking roots, this is a bigger problem with two-variable polynomials (and, indeed, n-variable polynomials more generally).

In chapter 6, we attempt to quash these potential concerns. We state a potential result related to single variable polynomials, including experimental and heuristic justification for why we believe this result to be true, as well as our attempts to prove this so far. We then move to two-variable polynomials, where we introduce a notion of 'equivalence', and how this helps us understand when two two-variable polynomial can, and cannot, share the same Mahler measure value.

1.3.4 Road Map

So far, we have introduced the Mahler measure and some important known results. We have also given motivation for why we wish to explore the connection between Mahler measures and digraphs. To complete this chapter, we now give a road map for the contents of the thesis: this includes the main contributions we will make, as well as the approaches of the thesis.

The key contribution of this thesis is a new method for finding small Mahler measures for both one and two-variable polynomials. This is a versatile method which can be easily adapted to switch between finding Mahler measures of one and two-variable polynomials; this is a unique feature of the method. A highlight is Result 5.3.1, which in fact gives us a new small Mahler measure coming from a single variable polynomial. In finding this new value, we also found over 97% of known small Mahler measure values from single variable polynomials; making this one of the most fruitful methods currently known. We also find a large proportion of known small Mahler measures from two-variable polynomials. These findings make chapter 5 the key chapter of the thesis.

Chapters 2–4 serve two purposes. Their primary purpose is to serve as building blocks towards chapter 5, allowing us to outline our new method for finding small Mahler measures in both one and two variables. However, these building blocks are also interesting in their own right, and some are noteworthy within the field. As such, we also add peripheral details where necessary, creating further contributions.

The highlight of chapter 2 is the introduction of a new method for calculating the Mahler measures of two-variable polynomials. We see that exact evaluations of such polynomials are incredible rare. Whilst methods exist for calculating these values to a certain number of significant figures exist, they each have their own flaws. Through Lemma 2.2.1 and Proposition 2.2.5, we are able to introduce a refined method for calculating such Mahler measures, which is practical to implement and is particularly relevant for searching for small Mahler measures of two-variable polynomials. Moreover, Section 2.5, discusses the distinction between the terms “variable” and “dimension” in terms of Mahler measures. This is a somewhat technical aside, although an important one, especially as this is something that is usually omitted, or only loosely touched upon, when studying Mahler measures.

Whilst the motivation of chapter 2 is to build upon theoretical knowledge, we take a very practical approach to the work presented. In particular, Example 2.3.1 highlights our new method in great detail, showcasing how all the theoretical knowledge is applied, and why our conditions are necessary.

Chapter 3 starts by providing background and context about what is already known at the interface of the study of Mahler measures and graphs. In particular, and as already noted, simple graphs are not fruitful for helping us find small Mahler measures, which motivates our introduction of digraphs. As such, our main focus in this chapter is to explain the intricate nature of digraphs, and why working with them is much more difficult compared to simple graphs. We also give some heuristic ideas which will help us later. Section 3.5 highlights specific theoretical results related to digraphs to highlight which tools we have at our disposal. Results presented here are regularly complemented with examples of specific digraphs, and Lemma 3.3.2 is particularly useful moving forward.

Chapter 4 is a direct continuation of chapter 3. We first use our knowledge of digraphs to find various ways to extend and "grow" digraphs, by attaching what we call pendant paths to them, which coincides with our heuristic ideas from before, which climaxes with Theorem 4.2.14. This result is then followed by Example 4.2.15, to demonstrate explicitly how it is applied in practice. This chapter also includes another key result, Theorem 4.3.1, which allows us to "internally grow" digraphs.

This then puts us in a position to look at our experiments which found small Mahler measure values, which is the area of concern for chapter 5. In Section 5.1, we explain the specific shape of digraphs we wish to grow when using our key theorems from chapter 4, and then outline how our experiments are run in Section 5.2. This then leads us onto the highlight of the thesis – Section 5.3 – which includes our newly found small Mahler measure and details the other small Mahler measures we found from one and two-variable polynomials, the latter of which are shown in Table D.1. We also include Example 5.3.3, which shows in great detail how we find the Mahler measures of two-variable polynomials from digraphs.

The thesis then steps away from finding small Mahler measures to a different, albeit related, topic. As our methods for calculating Mahler measures involve using computer software, we always have to be wary of our results and calculations, and ensure we are confident that our values are correct to our mentioned accuracy. Another way of viewing this concern is that we could consider different polynomials which share the same Mahler measure, and by computing the Mahler measure of these different polynomials, we are effectively wasting time in our experiments. As such, we investigate conditions for when two polynomials can share the same Mahler measure.

When we consider single variable polynomials, our main idea is presented as Open Problem 6.1.1. We justify the formulation of this problem in Section 6.1, and outline some theory which may help prove this in Section 6.2. We then move to considering the problem for two-variable polynomials; this is significantly more interesting, but equally much more intricate. The key concrete results here are Theorem 6.4.5 and Proposition 6.6.4. These give us two wide-reaching results which capture and explain many of the times when two two-variable polynomials share the same Mahler measure. Unfortunately, they do not capture all cases when this occurs, as described in the ending of the chapter.

Owing to the fact that chapter 6 does not fully answer the key questions raised, much of the what is presented here is speculative ideas. These speculations are justified by means of auxiliary results along the way, as well as several detailed examples. The aim of this final main chapter is to initiate the discussions about when polynomials share the same Mahler measure, and present the building blocks of the theory, in a hope that the key questions can be answered in the future.

The thesis concludes with chapter 7, which summarises the unanswered questions and open problems which have come about within the thesis. As such, it gives potential leads for future work. These include long-standing questions, such as Question 7.1.2, or questions first raised within the main body of the thesis, such as Questions 7.2.3 and 7.2.4, as well as the main problems from chapter 6.

Chapter 2

Calculating Mahler Measures of Two-Variable Polynomials

As mentioned in chapter 1, calculating explicit values of Mahler measures of polynomials becomes increasingly difficult as the number of variables increase. This should come as no surprise, as its definition involves several integrals of a non-standard multi-variable function. However, we have results, such as Proposition 1.2.4, which show us that Mahler measures of two-variable polynomials are limits of Mahler measures of single variable polynomials. Whilst this still does not give us a way for calculating an exact value for Mahler measures of polynomials in two-variables, it is a useful tool for knowing roughly what value the Mahler measure should take.

In this chapter, we look at ways in which Mahler measures of two-variable polynomials can be calculated, as well as ways in which we can make our calculations as simple as possible. We also introduce an evolution of one of these methods, which is particularly relevant to experiments which will be reported later in this thesis. We finally discuss a subtlety that occurs when calculating Mahler measures of several variables. This will show that, in the context of this thesis, we are in fact more concerned with *dimension* when calculating small Mahler measures of two-variable polynomials. More broadly, this highlights that the dimension of a polynomial is something which holds greater weight than the number of variables, especially when we are focusing on small Mahler measures.

2.1 Current Methods

Some work has gone into giving explicit forms of Mahler measures of certain two-variable polynomials. As already seen, Proposition 1.2.3 shows that $m(1+x+y)$ is a true limit point of $m(1+x+x^n)$ as $n \to \infty$, and gives us the error term as well. However, the following result of Smyth gives us an exact formulation:

Example 2.1.1 (Smyth [40])**.**

$$m(1+x+y) = \frac{3\sqrt{3}}{4\pi} L(2, \chi_3)\,, \tag{2.1}$$

where $L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$ is the Dirichlet L-function, and $\chi_3(n)$ is the Legendre symbol $(\frac{n}{3})$. In particular, we have:

$$L(2,\chi_3) = 1 - \frac{1}{4} + \frac{1}{16} - \frac{1}{25} + \cdots .$$

An alternative formulation for this specific polynomial can see us rewrite the integral in a different form, and use dynamical systems to evaluate this integral.

Example 2.1.2 (Pollicott and Felton [31])**.** By use of Jensen's Formula, we can write:

$$m(1+x+y) = \int_{1/6}^{5/6} \log 2\sin(\pi s)\ \mathrm{d}s\,. \tag{2.2}$$

By using the change of variables $t = \frac{3}{2}(s - \frac{1}{6})$, this gives us:

$$m(1+x+y) = \frac{2}{3}\int_0^1 \log\left(2\sin\left(\frac{2\pi}{3}t + \frac{\pi}{6}\right)\right)\ \mathrm{d}t\,. \tag{2.3}$$

Pollicott and Felton then describe how to estimate this to a high level of accuracy. Their method revolves around finding an explicit solution to the integral in terms of series, and then, as they describe it, "use periodic points of a simple dynamical system (the doubling map on the interval)" to in turn describe this series.

Remark. Here, we have provided a (trivial) correction to the original example provided by Pollicott and Felton, who state that the integrand is $\log\left(2\sin\left(\frac{2\pi}{3}t + \frac{1}{6}\right)\right)$.

It is perhaps striking (and, if not, at least important) in this instance that (2.2) uses Jensen's Formula, a tool for single variable polynomials, to simplify the integral of a two-variable polynomial to a single variable polynomial. So, whilst there is no analogue of Jensen's Formula for two-variable polynomials, it can still play a role in helping us understand Mahler measures of two-variable polynomials. Further still, Pollicott and Felton also note Jensen's Formula can be applied to polynomials in more than two variables, giving an example with a three-variable polynomial.

Of course, there are known ways of estimating double integrals to a high degree of accuracy. For example, Otmani et al. [30] utilised an algorithmic method for finding two-variable polynomials with small Mahler measures. To calculate the integrals involved for finding these values, they used the Monte Carlo method to approximate double integrals. Discussion on this statistical method can be found in various textbooks – such as Newman and Barkema [28] – but we give a basic heuristic overview of the method for single integrals here for reference.

Example. Consider a function $f(x)$ which we wish to integrate between two real values a and b, with $b > a$. Say:

$$I = \int_a^b f(x)\,\mathrm{d}x\,.$$

Choose N values in the range of the integral at random; say $\{x_1, \cdots, x_N\} \in [a, b]$. For each x_i, calculate $f(x_i)(b-a)$: this can be viewed as the area of a rectangle which has the width of the range we are integrating over. These values, in a crude sense, estimate the value of I (albeit rather poorly in some instances). However, we can average these estimates and, for large N, this can be thought of as a good approximation. This can be summarised by (2.4), which is usually referred to as the *basic Monte Carlo estimator*:

$$\langle I^N \rangle = \frac{b-a}{N} \sum_{i=1}^{N} f(X_i)\,, \tag{2.4}$$

where $\langle I^N \rangle$ represents an approximation of I over N samples.

Finally, the law of large numbers allows us to note that:

$$\mathbb{P}\left(\lim_{N\to\infty}\langle I^N\rangle = I\right) = 1\,.$$

That is, the probability that the basic Monte Carlo estimator gives us the value of our original integral for very large N is 1.

It is not too difficult to see how this method could be extended to double integrals, and how it is a powerful tool for giving good estimates.

Another method is outlined by Boyd and Mossinghoff [5]. Their approach, like that of Pollicott and Felton, uses Jensen's Formula to reduce the double integral to a more manageable single integral (in fact, more generally, n integrals to $(n-1)$ integrals). Their method, predating that of Pollicott and Felton, can also be used more generally.

Example 2.1.3. Let $P(x, y)$ have degree d in y, and $y_1(x), \cdots, y_d(x)$ be continuous, piecewise analytic functions in x that are the d solutions to $P(x, y) = 0$. This means we can write:

$$P(x,y) = a_0(x)\prod_{k=1}^{d}(y - y_k(x))\,,$$

where $a_0(x)$ is the coefficient of the y^d term. A similar argument used in the Proof of Proposition 1.1.2 shows that:

$$\int_0^1 \log|P(x, e^s)|\ \mathrm{d}s = |a_0(x)| + \sum_{k=1}^{d}\log^+|y_k(x)|\,. \tag{2.5}$$

Finally, letting $x = e^t$ in (2.5) and integrating over t gives:

$$m(P) = m(a_0(x)) + \sum_{k=1}^{d}\int_0^1 \log^+|y_k(e^t)|\ \mathrm{d}t\,. \tag{2.6}$$

We note here that $m(P)$ is the logarithmic Mahler measure of a two-variable polynomial, whilst $m(a_0(x))$ is the logarithmic Mahler measure of a single variable polynomial.

From here, Boyd and Mossinghoff [5] detail how to explicitly calculate the integrals of each of these branches. Most polynomials they considered are of small enough degree in y that these branches are sufficiently "nice" to integrate, but further details of calculating more complicated examples are also covered by Boyd [4].

These examples give us different ways of calculating two-variable Mahler measures. Statements of exact formulations, such as (2.1), are extremely rare, and only known in very special instances. The method from Otmani et al. did expand the known list of small Mahler measure values coming from two-variable polynomials, but they do note that their experiments took a long time and were very CPU expensive. However, the examples from Pollicott and Felton, in special cases, and Boyd and Mossinghoff, more generally, give us ways of calculating two-variable Mahler measures to a high degree of accuracy, and both do this by reducing the calculation to a single integral, which can be quicker and more efficient. This is useful, and sufficient, when trying to find small Mahler measures coming from two-variable polynomials, and we now turn to refining this method for our own benefit.

2.2 Evolution: A Refined Method

We now outline our refined method for calculating the Mahler measures of two-variable polynomials. This method is based on the method of Boyd and Mossinghoff, as seen in Example 2.1.3. It was born out of a need to create a routine we could use to calculate Mahler measures of two-variable polynomials. Whilst trying to create something which would follow the Boyd and Mossinghoff method, this similar method came to light. In fact, we will see that this method applies more specifically to the polynomials we discover by means of experiments which are reported upon later in this thesis.

A useful observation, which will be one of the main tools in constructing our method, is as follows:

Lemma 2.2.1. *For any two-variable polynomial $P(x, y)$, we have:*

$$\log M(P) = \frac{1}{\pi} \int_0^{\pi} \log M(P(x, e^{it}))\ \mathrm{d}t\,. \tag{2.7}$$

Before proving this, we first need a simple result related to the Mahler measure of a polynomial and its complex conjugate.

Definition 2.2.2. *Let $f(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$, with $a_i \in \mathbb{C}$. We define the* ***complex conjugate*** *of f, denoted $\overline{f}$, to be:*

$$\overline{f}(z) := \overline{a_n} z^n + \overline{a_{n-1}} z^{n-1} + \cdots + \overline{a_1} z + \overline{a_0}\,.$$

Lemma 2.2.3. *For a polynomial $f \in \mathbb{C}[z]$, $M(f) = M(\overline{f})$.*

Proof. We know that $f(z) = 0$ if and only if $\overline{f}(\overline{z}) = 0$. Furthermore, we also know that $|z| = |\overline{z}|$. As such, the moduli of the zeros of f and $\overline{f}$ are the same.

From here, we can use (1.2) to see that $M(f) = M(\overline{f})$. □

Proof of Lemma 2.2.1. By suitable substitutions in the definition of $M(P(x, y))$, we get:

$$M(P) = \exp\left(\frac{1}{(2\pi)^2} \int_{s=0}^{2\pi} \int_{t=0}^{2\pi} \log|P(e^{is}, e^{it})| \, \mathrm{d}t \, \mathrm{d}s\right).$$

We can then change the order of integration to get:

$$M(P) = \exp\left(\frac{1}{(2\pi)^2} \int_{t=0}^{2\pi} \int_{s=0}^{2\pi} \log|P(e^{is}, e^{it})| \, \mathrm{d}s \, \mathrm{d}t\right). \tag{2.8}$$

We can now evaluate the inner integral of (2.8) to get:

$$M(P) = \exp\left(\frac{1}{2\pi} \int_{t=0}^{2\pi} \log M(P(x, e^{it})) \, \mathrm{d}t\right).$$

We now split the integral into two halves:

$$M(P) = \exp\left(\frac{1}{2\pi} \int_{t=0}^{\pi} \log M(P(x, e^{it})) \, \mathrm{d}t + \frac{1}{2\pi} \int_{t=\pi}^{2\pi} \log M(P(x, e^{it})) \, \mathrm{d}t\right). \tag{2.9}$$

By use of the substitution $\theta = 2\pi - t$ on the second integral of (2.9) only, we now get:

$$M(P) = \exp\left(\frac{1}{2\pi} \int_{t=0}^{\pi} \log M(P(x, e^{it})) \, \mathrm{d}t + \frac{1}{2\pi} \int_{\theta=\pi}^{0} (-1) \log M(P(x, e^{-i\theta})) \, \mathrm{d}\theta\right).$$

And a suitable change to the second integral gives us:

$$M(P) = \exp\left(\frac{1}{2\pi} \int_{t=0}^{\pi} \log M(P(x, e^{it})) \, \mathrm{d}t + \frac{1}{2\pi} \int_{\theta=0}^{\pi} \log M(P(x, e^{-i\theta})) \, \mathrm{d}\theta\right). \tag{2.10}$$

If we now consider $Q_t(x) = P(x, e^{it})$, then $\overline{Q_t}(x) = P(x, e^{-it})$. As such, we have that $M(P(x, e^{it})) = M(P(x, e^{-it}))$, by Lemma 2.2.3. This means that both integrals in (2.10) are in fact equal.

So, by summing together and subsequently taking logs, we then achieve (2.7), our desired result. □

Lemma 2.2.1 reduces the problem of calculating the Mahler measure of a two-variable polynomial to two much simpler steps. The first step gives us an expression, in $\mathbb{C}[t]$, for the Mahler measure of some single variable polynomial, which is related to the two-variable polynomial. The second step then allows us to integrate this expression, which in turn evaluates the Mahler measure of the given two-variable polynomial. This is significantly easier to do than attempting to do the double integral directly.

We can evaluate integrals like in (2.7) with reasonable ease. If we treat $P(x, e^{it})$ as a single variable polynomial in x, we can then use Proposition 1.1.2 to get an expression for the Mahler measure in terms of t. Our integrand then becomes $\log h(t)$, for some function $h(t)$ and we can integrate this as we would any other function.

This does cause a problem, however. Generally speaking, the function $H(t) = \log M(P(x, e^{it}))$ is not smooth. Indeed, in most cases, H has many "spikes" (points where we cannot define the derivative of H), meaning that it is very hard to get an accurate value when integrating directly. On the other hand, if we are able to locate these spikes, we can integrate between them, and sum the resulting values together, to get the value of any integral of the form (2.7).

It is at this point we make an extension to Definition 1.1.12:

Definition 2.2.4. *We say a polynomial $\mathcal{P}(z_1, \cdots, z_n)$ is **reciprocal** if $\mathcal{P}(z_1, \cdots, z_n) = {z_1}^{d_1} \cdots {z_n}^{d_n} \mathcal{P}(1/z_1, \cdots, 1/z_n)$, such that $\deg_{z_i}(\mathcal{P}) = d_i$ for each i.*

Remark. We note that if we have $n = 1$, this indeed recovers our original Definition.

The polynomials we consider for our experiments later in this thesis will always be reciprocal. When this happens, we are able to locate the aforementioned spikes (by methods we will outline shortly). This does mean that our refined method will ultimately only apply for reciprocal polynomials, but this is not a problem for us.

Proposition 2.2.5. *Let $P \in \mathbb{Z}[x, y]$ be non-zero and reciprocal. Then, the spikes of the function $\log M(P(x, e^{it}))$ are in correspondence with a subset of the solutions of $disc_y P(x, y) = 0$.*

Proof. Let $\deg_x(P) = n$ and $\deg_y(P) = m$.

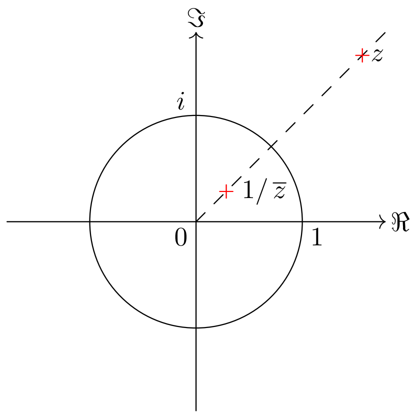
Since P is reciprocal and has real coefficients, we know that:

$$P(x, y) = x^n y^m P(1/x, 1/y), \; x, y \in \mathbb{C}^*, \tag{2.11}$$

$$\overline{P(x, y)} = P(\overline{x}, \overline{y}), \; x, y \in \mathbb{C}. \tag{2.12}$$

For $t \in [0, 2\pi]$, let $\gamma_1(t), \cdots, \gamma_n(t)$ be the branches of the solutions of $P(x, e^{it}) = 0$. If $P(x, e^{it}) = 0$, then (2.11) gives us $P(1/x, e^{-it}) = 0$ (since $x \neq 0$ here). Then, (2.12) gives us $\overline{P(1/x, e^{-it})} = P(1/\overline{x}, e^{it}) = 0$.

For reference, the sketch below demonstrates visually what happens when we map $z \mapsto 1/\overline{z}$:



.

In words, it means that any value outside the unit circle is mapped into the unit circle (and vice versa), and the newly mapped value remains on the straight line formed by the original value and the origin.

Hence, if $x = \gamma_j(t)$, for some j and t, then $1/\overline{x} = \gamma_k(t)$, for the same t. In view of (2.6), spikes occur when some $|\gamma_j(t)|$ crosses from being on the unit disc (i.e.: of value at most 1) to being outside the unit disc (i.e.: of value larger than 1), or vice versa.

We now say that t_1 is a "critical value" if $|\gamma_j(t_1)| = 1$ and $|\gamma_j(t)| > 1$ for $t \in (t_1, t_2]$, for some j and t_2, or for $t \in [t_2, t_1)$, for some j and t_2 (the former occurring when $|\gamma_j(t)|$ crosses from being on the unit disc to being outside it, the latter in the opposing case).

These critical values include all the spikes.

Then, the mapping $z \mapsto 1/\overline{z}$ maps γ_j to γ_k, for some $k \neq j$, since $\gamma_k(t_2) \neq \gamma_j(t_2)$. However, we have that $\gamma_k(t_1) = \gamma_j(t_1)$. We will call this value e^{is}.

Remark. A consequence of what we have shown here is that $k = j$ if and only if $\gamma_j(t)$ is strictly confined to the unit circle.

We note that the value $x = e^{is}$ is in fact a double root of $P(x, e^{it_1}) = 0$. So, the polynomial $\text{disc}_x P(x, e^{it})$ (a single variable polynomial in x) vanishes precisely at the point $t = t_1$. Therefore, $\text{disc}_y P(x, y)$ vanishes precisely at the point $y = e^{it}$. □

Remark. We see that the "critical values" mentioned here include all spikes, but this does not mean that all critical values are spikes. We see later, in Example 2.3.1, for instance, a situation where this is the case.

As a result of Proposition 2.2.5, we now have no more worries about using Lemma 2.2.1 for calculating the Mahler measure of a two-variable polynomial. Furthermore, it should be noted that this work applies analogously to the polynomial $P(e^{is}, y)$, giving us the following:

Lemma 2.2.6. *For any two-variable polynomial $P(x, y)$, we have:*

$$\log M(P) = \frac{1}{\pi} \int_0^{\pi} \log M(P(e^{is}, y)) \, \mathrm{d}s \,. \tag{2.13}$$

We present two proofs, which are essentially the same:

Proof (Version 1). We can follow an analogous argument as presented in the proof of Lemma 2.2.1. Equally, a very similar proof to that of Proposition 2.2.5 shows us that the spikes of the function $\log M(P(e^{is}, y))$ are in correspondence with the solutions of $\text{disc}_x P(x, y)$. □

Proof (Version 2). Consider $Q(x, y) = P(y, x)$ and apply our previous results to Q. The result then follows. □

In practice, this means that we take our two-variable polynomial P, treat it as a single variable polynomial in either x or y (replacing the other term with e^{it} or e^{is}

respectively). We then find the Mahler measure of this single variable polynomial (in terms of t or s respectively), and then integrate the resulting function (which will be in the form of $\log h(t)$ or $\log g(s)$ respectively). Broadly speaking, there is no preferred choice here: either option results in the same answer. However, in some cases, our substitutions may result in one of the functions $\log h(t)$ and $\log g(s)$ being "easier" to integrate, in which case, it is perhaps more sensible to work with that particular function.

We will soon look at examples of using this method in practice. Before doing so, we briefly establish how exactly we find spikes of functions like $\log M(P(x, e^{is}))$, and how we then go on to use them.

Doing this makes use of Proposition 2.2.5, and one can easily create a short code to do this. The procedure we follow is:

1. Find $\Delta(y) = \text{disc}_y P(x, y)$,
2. Establish which roots of $\Delta(y) = 0$ have absolute value 1; discard all others,
3. Label each conjugate pair of roots by $r_1, \cdots, r_n$ ($n \geq 1$, unless there are no spikes, in which case $n = 0$),
4. Calculate $t_i = \arccos(\Re(r_i))$, for each r_i. These are the potential locations of the spikes,
5. Calculate (2.14), as seen below.

If we are in the rare situation where there are no spikes, then we can calculate the integral $\int_0^\pi \log M(P(x, e^{it}))\ \mathrm{d}t$ directly.

As described earlier, the actual calculation made in practice is:

$$\log M(P) = \frac{1}{\pi}\left(\int_{t=0}^{t_1} \log M(P(x, e^{it}))\ \mathrm{d}t + \cdots \right.$$
$$\left. + \int_{t=t_{n-1}}^{t_n} \log M(P(x, e^{it}))\ \mathrm{d}t + \int_{t=t_n}^{\pi} \log M(P(x, e^{it}))\ \mathrm{d}t \right), \quad (2.14)$$

and since the function is smooth in each specified interval, this is fine.

When using PARI/GP, this routine is quick. However, perhaps unsurprisingly, it can struggle with very large, complicated polynomials which have many spikes (as do other methods). Since we are searching for small Mahler measures, we can circumvent this issue by choosing good candidates of polynomials which may potentially have small Mahler measure. We will outline how we do this later, in chapters 4 and 5.

2.3 Examples: The Good and The Bad

This Section is dedicated to looking at a couple of examples in detail, showing what is happening at each step. Of course, in practice, none of this is seen: it is all encapsulated within one command. Nevertheless, it is helpful to see this method executed, step-by-step. We will also show a more complicated example which this method struggles with, to highlight the importance of making sure we consider the "right" polynomials for attempting to find small Mahler measures.

2.3.1 Good Examples

Example 2.3.1. Let $P(x,y) = (y^4 - y^3)x^6 - y^4x^4 + x^2 + y - 1$.

Though in practice our code does not do this, here we look at a plot of the function $\log M(P(x, e^{it}))$ between 0 and π:

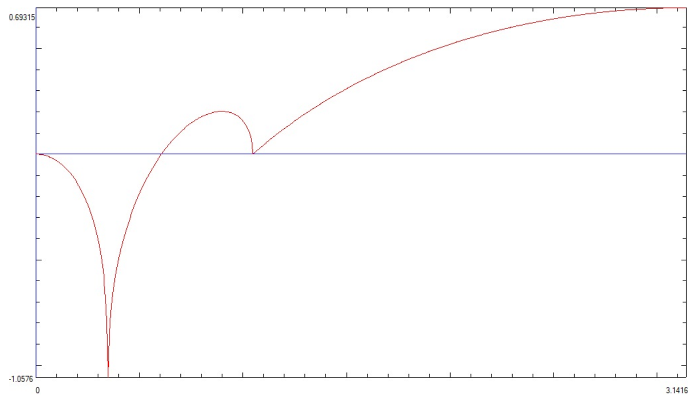


Figure 2.1: A plot of $\log M(P(x, e^{it}))$.

This plot was done using the PARI/GP command `ploth`, which gives a good overview of the function, although certainly not as in depth as other software options. Nonetheless, this suggests that the function is not smooth, and will likely have at least two spikes within this range.

We now follow our procedure. The first thing to do is establish what $\Delta(y) = \mathrm{disc}_y P(x, y)$ is. The PARI/GP routine `poldisc` gives us:

$$\begin{aligned}\Delta(y) =&(-64y^9)(16y^{22} - 64y^{21} + 96y^{20} - 280y^{19} + 1384y^{18} - 3808y^{17} + 6825y^{16}\\&- 12414y^{15} + 29107y^{14} - 63320y^{13} + 103197y^{12} - 121478y^{11} + 103197y^{10}\\&- 63320y^9 + 29107y^8 - 12414y^7 + 6825y^6 - 3808y^5 + 1384y^4 - 280y^3\\&+ 96y^2 - 64y + 16)\,.\end{aligned}$$

We now find the roots of $\Delta(y) = 0$ (we use `polroots` in PARI/GP), and discard those with modulus not equal to 1. This leaves us with four pairs of roots:

- $r_1 = 1 \pm 0i$,
- $r_2 = 0.94023 \cdots \pm 0.34053 \cdots i$,
- $r_3 = 0.93888 \cdots \pm 0.34423 \cdots i$,
- $r_4 = \frac{1}{2} \pm \frac{\sqrt{3}}{2} i$.

We now take the arccos of the real parts of each of these values, giving us:

- $t_0 = 0$,
- $t_1 = 0.34749 \cdots$,
- $t_2 = 0.35142 \cdots$,
- $t_3 = 1.04719 \cdots$.

Before moving on, we note a couple of things here. Firstly, we see that one of our identified points is 0, however, Figure 2.1 shows that it does not necessarily appear to be a spike. Indeed, Figure 2.2 suggests that the function is smooth here. However, we recall that Proposition 2.2.5 states that the spikes of our function are only in correspondence with a *subset* of the solutions of our $\Delta(y)$. As such, any identified points are just potential locations of spikes.

In practice, we still integrate between these identified points. Moreover, in this particular instance, we would be integrating from the value $t = 0$ anyway. So we can proceed as normal.

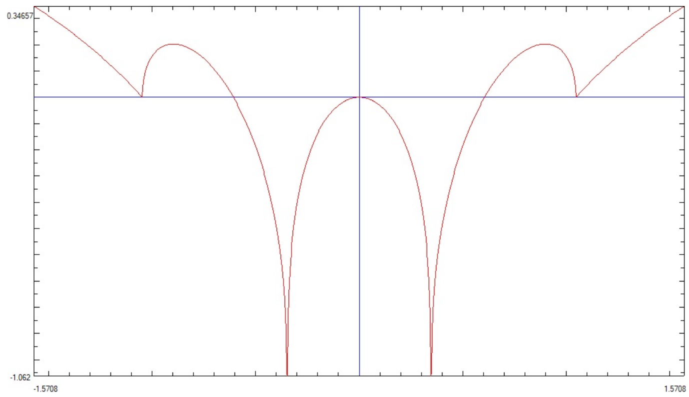


Figure 2.2: An alternate plot of $\log M(P(x, e^{it}))$, suggesting smoothness at the origin.

Secondly, our procedure as outlined earlier sees us calculate values t_i such that $t_i < t_{i+1}$ for each necessary i (so that we may calculate (2.14)), and that we have corresponding labels for our pairs of roots r_i. However, there is no easy systematic way to find and enumerate the pairs of roots r_i such that the corresponding t_i are such that $t_i < t_{i+1}$. In practice, we compute the t_i in whatever order they emerge, and then sort them afterwards. For the sake of this example, we have labelled the roots r_i so that the t_i values appear in the appropriate order.

Finally, upon a first look at Figure 2.1, one may think there are exactly two spikes. However, we note that t_1 and t_2 are close together, so this is not clear on a rough plot such as the one seen in Figure 2.1. A plot using a smaller range, such as Figure 2.3, or different software with the same range, makes these two spikes more distinguishable.

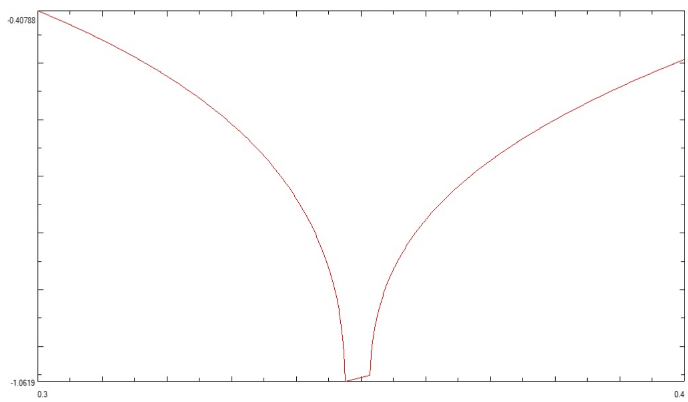


Figure 2.3: A plot of $\log M(P(x, e^{it}))$ in the range $x \in [0.3, 0.4]$, highlighting the spikes at t_1 and t_2.

We now calculate $\log M(P)$ as in (2.14), which gives us:

$$\log M(P) = \frac{1}{\pi}\int_{t=0}^{0.34749\cdots} M(P(x,e^{it}))\ \mathrm{d}t + \frac{1}{\pi}\int_{t=0.34749\cdots}^{0.35142\cdots} M(P(x,e^{it}))\ \mathrm{d}t$$
$$+ \frac{1}{\pi}\int_{t=0.35142\cdots}^{1.04719\cdots} M(P(x,e^{it}))\ \mathrm{d}t + \frac{1}{\pi}\int_{t=1.04719\cdots}^{\pi} M(P(x,e^{it}))\ \mathrm{d}t\,.$$

Using the standard precision provided by PARI/GP (38 significant digits), this gives us that $M(P) = 1.3486519908902990407430414102458380026$. This matches the value given by Otmani et al. [30] for this polynomial, up to their provided accuracy (which was 15 significant figures).

We now show an example where we treat our two-variable polynomial P as a single variable in y, hence replacing x with e^{is}. In fact, to show the versatility of our method, we will look at the same polynomial as in Example 2.3.1. In the following, we will not be going into as much detail as before, noting the key concepts are the same, but still outline how the procedure works.

Example 2.3.2. Let $P(x,y) = (y^4 - y^3)x^6 - y^4x^4 + x^2 + y - 1$. We plot the function $\log M(P(e^{is}, y))$ between 0 and π (though, again, there is no need to do this in practice):

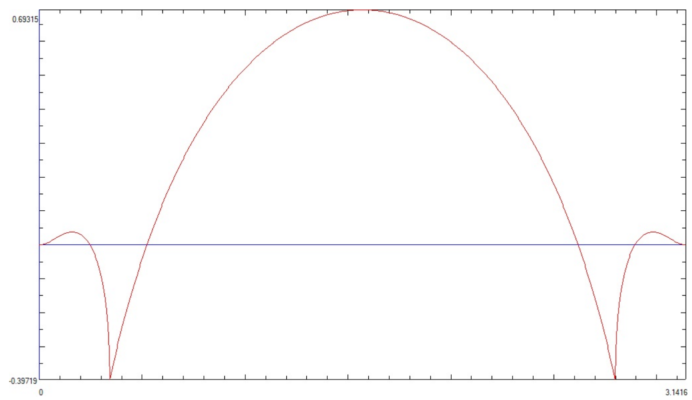


Figure 2.4: A plot of $\log M(P(e^{is}, y))$.

Though this perhaps should not be surprising, we note how strikingly different this is to Figure 2.1.

Following our explained routine, we find that there are indeed two spikes; these are located at $s_1 = 0.34148\cdots$ and $s_2 = 2.80010\cdots$. Calculating $\log M(P)$ as in (2.14) gives us:

$$\log M(P) = \frac{1}{\pi}\int_{s=0}^{0.34148\cdots} M(P(e^{is},y))\ \mathrm{d}s + \frac{1}{\pi}\int_{s=0.34148\cdots}^{2.80010\cdots} M(P(e^{is,y}))\ \mathrm{d}s$$
$$+ \frac{1}{\pi}\int_{s=2.80010\cdots}^{\pi} M(P(e^{is},y))\ \mathrm{d}s\,.$$

And again, using the standard precision provided by PARI/GP, this gives us that $M(P) = 1.3486519908902990407430414102458380026$. This is exactly as in Example 2.3.1, demonstrating that these methods do indeed give us the exact same Mahler measure value.

2.3.2 A More Difficult Example

We now turn our attention to a case where our refined method is not as efficient or practical to use.

Example 2.3.3. Let $P(x, y) = y^6x^{27} - (y^6 + y^5)x^{21} + y^3x^{14} + y^3x^{13} - (y + 1)x^6 + 1$. Again, let us look at the plot of the function $\log M(P(x, e^{it}))$ between 0 and π:

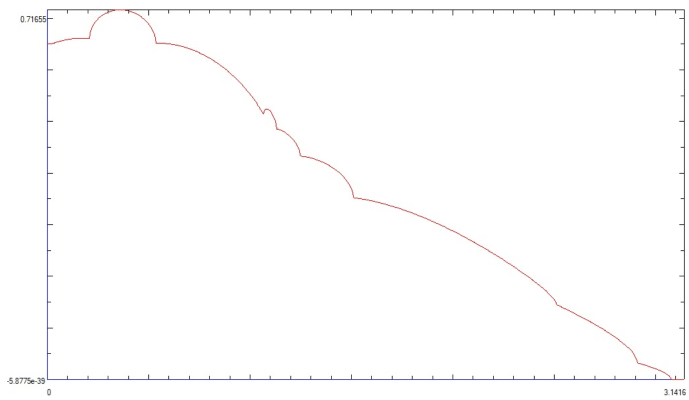


Figure 2.5: A plot of $\log M(P(x, e^{is}))$.

We see that this plot certainly has many more spikes. It perhaps should not be too surprising that there are more spikes: polynomials of larger degree lend themselves to having more spikes. Identifying all these spikes, and then integrating between each of them, is thus more time consuming.

Finally, we find that $M(P) = 1.46829299\cdots$.

Example 2.3.3 may not seem significantly more complex compared to Examples 2.3.1 and 2.3.2. Whilst it is true there are more spikes, the smooth components of the function look "nice" to integrate between. However, in practice, the computations still take significantly longer.

For example, using our routine, the calculation involved in Example 2.3.3 lasted more than 270 times longer than the calculation involved in Example 2.3.1. This represents a significant increase in time, and it is not inconceivable that there are polynomials which would take even longer than this. Our routine and method presented here could perhaps be optimised further, or be more efficient on more powerful machines or when run on software other than PARI/GP, but these explorations go beyond the scope of the work presented here.

This creates a difficult crossroad for us when trying to find small Mahler measures coming from two-variable polynomials. On the one hand, the known list of small values from Mossinghoff [26], Otmani et al. [30] and Sac-Épée et al. [35] all come from polynomials of reasonably small degree. As it appears likely that polynomials of "larger degree" (a concept we define formally shortly) take a longer time to calculate the Mahler measure of, one possible avenue for us to focus on is to disregard polynomials above a certain degree. On the other hand, one may wonder if it is perhaps fruitful to use this method to explore polynomials of larger degrees, in the hope of finding new small Mahler measure values.

In any case, we have demonstrated that we need to take caution with our evolved variant for calculating Mahler measures of two-variable polynomials. Not only does this only work for reciprocal polynomials – a caveat which we will see in chapters 4 and 5 is not an issue in the scope of this thesis – we have to balance between polynomials not being too complicated, whilst still being of interest to us.

2.4 Simplifying Calculations

As has been noted, calculating the Mahler measure of more complicated two-variable polynomials (those which have a high degree, or those which have a large number of terms) can take a long time, regardless of the method we use. As our primary interest is in searching for small Mahler measures, it is as such important to look at polynomials which may have a better chance of having small Mahler measure. We explain this in greater detail later.

However, there are still concerns that arise here. Firstly, there is no guarantee that good potential candidate polynomials will be sufficiently nice (of a small degree and of a short enough length). Secondly, this does not help if there are instances where one may not necessarily be interested solely in polynomials admitting small Mahler measure values. Though the latter concern here is more hypothetical and outside the scope of this thesis, the former is a more troubling issue.

As such, we now look at possible ways to simplify polynomials to make calculations easier. In particular, this means we want to look at ways in which we can take a polynomial and make it of a smaller degree.

Definition 2.4.1. *Let $\mathcal{P} \in \mathbb{C}[z_1, \cdots, z_n]$. We say that the* ***total degree*** *of a polynomial, denoted* $\deg(\mathcal{P})$*, is the maximal sum of the powers of all variables amongst all monomials.*

Example. Let $P(x, y) = x^4y^3 + x^3y + xy^4 + y^2 + 1$. Then, $\deg(P) = 7$.

The total degree is a basic, somewhat crude, way of deciding how simple a polynomial may be. Of course, it is not the only tool we should use. However, in a broad sense, it is reasonable to assume that if we can take a polynomial and reduce its total degree (without affecting its Mahler measure), then it should be easier to calculate the Mahler measure of the resulting polynomial.

Another, somewhat obvious, simplification one can do is to only ever calculate the Mahler measure of irreducible polynomials. Knowing that the Mahler measure is multiplicative, for a polynomial $P = P_1 \cdots P_n$, with each P_i irreducible, we can calculate each $M(P_i)$ and multiply the resulting values together. Moreover, we can ignore any 2-Kronecker-cyclotomic factors as well (and, in general, n-Kronecker-cyclotomic factors when considering polynomials in n variables).

Of course, in the practice of finding polynomials with small Mahler measures, we will only ever be interested in irreducible polynomials. This is because we expect any reducible polynomial with no 2-Kronecker-cyclotomic factors to have a non-small Mahler measure, as demonstrated below:

Example. Let $P(x, y) \in \mathbb{Z}[x, y]$ be a reducible polynomial with n factors, say $P = P_1 \cdots P_n$, and assume that none of these factors are 2-Kronecker-cyclotomic. If $M(P)$ is small, then there is at least one P_i such that $M(P_i) < (1.37)^{1/n}$.

In the case $n = 2$, this would mean that there is at least one factor, say P_1, such that $M(P_1) < \sqrt{1.37} = 1.17046999\cdots$. As the smallest known Mahler measure from two-variable polynomials is $1.25543386\cdots$, we do not expect this to happen.

2.4.1 Primitive Polynomials

One important idea is to look at polynomials which are "primitive"; that is, polynomials which are, in some sense, as simple as possible. We want our concept of primitive to be something that allows us to "primitize" a polynomial, whilst fixing the Mahler measure. In other words, we want to have some way of taking a polynomial, and seeing if we can make it simpler, and so easier to determine its Mahler measure. This tool would be something to apply after having checked if our polynomial is irreducible (although, in theory, there is no need to necessarily do this).

We will soon see that there is more than one notion of primitive that can be defined. Before turning to these, we first fix some notation and results. These will help give some further intuition to one definition of a primitive polynomial.

Definition 2.4.2. *Let $\boldsymbol{z} = (z_1, \cdots, z_n)$, and $A = (a_{ij})$ be some $n \times k$ matrix. We then define $\boldsymbol{z}^A$ to be the following k-tuple:*

$$\boldsymbol{z}^A := (z_1^{a_{11}} \cdots z_n^{a_{n1}}, \cdots, z_1^{a_{1k}} \cdots z_n^{a_{nk}}).$$

Proposition 2.4.3 (McKee and Smyth [25], Proposition 2.9)**.** *Let $\mathcal{P} \in \mathbb{R}[z_1, \cdots, z_n]$ and $A \in GL_n(\mathbb{Z})$. Then:*

$$M(\mathcal{P}(\boldsymbol{z})) = M(\mathcal{P}(\boldsymbol{z}^A)).$$

Proof. We provide the heart of the concept of the proof here; full details are in McKee and Smyth [25].

If we think of the integrals involved when defining the Mahler measure as defining the 'volume' of a given region, then replacing $\mathbf{z}$ by $\mathbf{z}^A$ will multiply the region by a factor of $|\det(A)|$. However, we find that the Jacobian of the necessary change of variables is also $|\det(A)|$, and so cancels the factor of $|\det(A)|$ which we had found. □

Definition 2.4.4. *Let $\mathcal{P} \in \mathbb{Z}[z_1, \cdots, z_n]$. We say that $\mathcal{P}$ is* ***strongly primitive*** *if there is no $A \in GL_n(\mathbb{Z})$, with $|\det(A)| > 1$, and $\mathcal{Q} \in \mathbb{Z}[z_1, \cdots, z_n]$ for which $\mathcal{P}(\boldsymbol{z}) = \mathcal{Q}(\boldsymbol{z}^A)$.*

This definition comes from Boyd and Mossinghoff [5]. Here, they simply refer to such polynomials *primitive*.

Strongly primitive polynomials are of interest because of their total degree. Say we have a polynomial $\mathcal{P}$ which is not strongly primitive. Then, the corresponding $\mathcal{Q}$ which reveals this will have a lower total degree and, by Proposition 2.4.3, will still have the same Mahler measure. That is to say, checking if a polynomial is strongly primitive is one quick way of determining if the polynomial we are looking at can be simplified.

At this point, we are saying that a quick and easy way of simplifying the calculation of the Mahler measure of a polynomial is to ensure it is irreducible and strongly primitive. We are *not* claiming this is all the simplification we can do, nor that this is the best we can do.

Unfortunately, it is not obvious to the eye whether a polynomial is strongly primitive or not. However, there is a quick and effective way to check. As our focus here is on two-variable polynomials, we look specifically at how to do this for two-variable polynomials P. The following discussion makes references to lattices; the unfamiliar reader is referred to Appendix B for an overview of the necessary concepts.

Let $P(x, y) = \sum_i c_i x^{a_i} y^{b_i}$, with $a_i, b_i, c_i \in \mathbb{Z}$. We first construct a matrix $M = \begin{pmatrix} a_1 & a_2 & \cdots \\ b_1 & b_2 & \cdots \end{pmatrix}$. We can then construct a lattice $\mathcal{L}$, which is the $\mathbb{Z}$-span on the columns of M. We have that $\operatorname{rank}(\mathcal{L}) \leq 2$, but the only case we are interested in is when $\operatorname{rank}(\mathcal{L}) = 2$. The case $\operatorname{rank}(\mathcal{L}) = 0$ is trivial, and $\operatorname{rank}(\mathcal{L}) = 1$ occurs when the columns of M are not linearly independent, effectively giving us a single variable polynomial, hence why the $\operatorname{rank}(\mathcal{L}) = 2$ case is the only one of interest.

We now briefly describe how, for two-variable polynomials, we approach this in practice, using PARI/GP. It is perhaps better understood with Example 2.4.5, which follows after, however.

If we have $P(\mathbf{z}) = Q(\mathbf{z}^A)$, for some $|\det(A)| > 1$, then the map $\mathbf{z} \mapsto \mathbf{z}^A$ sends $\mathbb{Z}^2$ to a proper sublattice of $\mathbb{Z}^2$. Now, assuming that P has t non-zero terms, then M will be a $2 \times t$ matrix. We can use the inbuilt LLL algorithm feature in PARI/GP, `qflll`, to find an LLL reduction of the lattice $\mathcal{L}$. Specifically, if `qflll( M ) = U`, then the columns of $MU = A$, with $A \in GL_2(\mathbb{Z})$, give us an LLL-reduced basis for the lattice $\mathcal{L}$.

Moreover, we have $|\det(A)| = 1$ if and only if the lattice is the whole of $\mathbb{Z}^2$. This occurs if and only if P is strongly primitive. Equally, $|\det(A)| > 1$ allows us to recover a strongly primitive Q, with the same Mahler measure as P. This will be demonstrated in the following examples.

Example 2.4.5. Let $P(x, y) = 1 + xy^2 + x^2y$.

We have $M = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix}$. Using PARI/GP, we have $U = \begin{pmatrix} 0 & 0 \\ -1 & 1 \\ 1 & 0 \end{pmatrix}$. This gives us $A = MU = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}$. We note that $|\det(A)| = 3$, so P is not strongly primitive.

As noted, we are interested in strongly primitive polynomials because they make computing the Mahler measure of a polynomial quicker. However, there will likely be polynomials we are interested in which are not strongly primitive. Hence, we want to be able to *strongly primitize* polynomials: take a polynomial P, and turn it into a strongly primitive polynomial, without affecting the Mahler measure.

Fortunately, it is reasonably straightforward to expand what we know to create a procedure for strongly primitizing a polynomial. When doing this, we also need to remember the coefficients of each term of our original polynomial P. This is easy to do in isolated examples, and can also do included with any code routines by creating a *companion vector*: $\mathbf{c} = \begin{pmatrix} c_1 & c_2 & \cdots \end{pmatrix}$.

We know that our matrix $A \in GL_2(\mathbb{Z})$ maps $\mathbb{Z}^2$ onto $\mathcal{L}$ by left multiplication. As such, A^{-1} 'reverses' this: it maps $\mathcal{L}$ onto $\mathbb{Z}^2$ by left multiplication. So, for $M = \begin{pmatrix} a_1 & a_2 & \cdots \\ b_1 & b_2 & \cdots \end{pmatrix}$, we have that $A^{-1}M \in \mathrm{Mat}_{2\times t}(\mathbb{Z})$ will correspond to a strongly primitive polynomial, which is the strongly primitized version of P.

More explicitly, for $A^{-1}M = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots \\ \beta_1 & \beta_2 & \cdots \end{pmatrix}$, the strongly primitized version of P, which we will denote as $Q = \mathrm{prim}(P)$, can be read as:

$$Q(x, y) = \sum_i c_i x^{\alpha_i} y^{\beta_i} . \tag{2.15}$$

Example 2.4.5 (Continued)**.** We note that our companion vector in this instance is $\mathbf{c} = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$.

We have that $A^{-1} = \begin{pmatrix} 2/3 & -1/3 \\ 1/3 & 1/3 \end{pmatrix}$. This gives us $A^{-1}M = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. From this, we follow (2.15) to read that $\mathrm{prim}(P) = Q(x, y) = 1 + y + xy$.

Remark. A^{-1} does not need to have integer entries (in fact, it probably will not in many instances). However, $A^{-1}M$ always will have integer entries.

Remark. By Proposition 2.4.3, we know that this procedure does not change the Mahler measure of the polynomial, since $P(\mathbf{z}) = Q(\mathbf{z}^A)$.

We can also look at a the notion of primitive from a different perspective, which is somewhat weaker.

Definition 2.4.6. *Let $\mathcal{P} \in \mathbb{Z}[z_1, \cdots, z_n]$. We say $\mathcal{P}$ is* ***weakly primitive*** *if, for each i, it can be viewed as a single variable polynomial in z_i, but not in ${z_i}^{d_i}$ for any $d_i \in \mathbb{Z}_{>1}$.*

A benefit this has over our definition of strongly primitive is that it is much easier to see by eye if a polynomial is weakly primitive or not. Furthermore, we are able to take any polynomial which is not weakly primitive and make it weakly primitive: for each i in which $\mathcal{P}$ can be viewed as a single variable polynomial in ${z_i}^{d_i}$, replace ${z_i}^{d_i}$ with z_i for the highest possible d_i. So, we are able to *weakly primitize* a polynomial.

As with strongly primitizing a polynomial, weakly primitizing a polynomial does not affect its Mahler measure. This gives us a way of reducing the degree of a polynomial, in a hope of simplifying the calculation of its Mahler measure.

There are two downsides to this, however. Firstly, this does not simplify calculations much. Routines which calculate the Mahler measure of a two-variable polynomial will not be noticeably slowed down when dealing with a polynomial which is not weakly primitive compared to one which is. More importantly – as our naming might suggest – a weakly primitive polynomial may not be strongly primitive.

Example. $P(x, y) = 1 + x^2y + xy^2$ is weakly primitive, but (as seen in Example 2.4.5) not strongly primitive.

On the other hand, we have the following Proposition which can describe this:

Proposition 2.4.7. *If $\mathcal{P}$ is strongly primitive, then $\mathcal{P}$ is weakly primitive.*

Proof. We show the contrapositive: if $\mathcal{P}$ is not weakly primitive, then $\mathcal{P}$ is not strongly primitive. So, assume $\mathcal{P}$ is not weakly primitive.

This means there is at least one variable in $\mathcal{P}$, say z_k, such that $\mathcal{P}$ can be viewed as a polynomial in ${z_k}^{d_k}$, for some $d_k \in \mathbb{Z}_{>1}$. Let us write $\mathcal{P}$ in the following form:

$$\mathcal{P}(z_1, \cdots, z_n) = \sum_i c_i {z_1}^{f_{1_i}} \cdots {z_n}^{f_{n_i}},$$

where $c_i, f_{j_i} \in \mathbb{Z}$, and there are exactly t non-zero c_i.

We can then construct a matrix of the powers of these coefficients:

$$M = \begin{pmatrix} f_{1_1} & f_{1_2} & \cdots \\ f_{2_1} & f_{2_2} & \cdots \\ & \vdots & \\ f_{k_1} & f_{k_2} & \cdots \\ & \vdots & \\ f_{n_1} & f_{n_2} & \cdots \end{pmatrix}.$$

As $\mathcal{P}$ can be viewed as a polynomial in ${z_k}^{d_k}$, that means there exists some $\delta \in \mathbb{Z}_{>1}$

such that $\delta | f_{ki}$ for each i. Letting $f_{ki} = \delta F_{ki}$, we can rewrite M as:

$$M = \begin{pmatrix} f_{11} & f_{12} & \cdots \\ f_{21} & f_{22} & \cdots \\ & \vdots & \\ \delta F_{k1} & \delta F_{k2} & \cdots \\ & \vdots & \\ f_{n1} & f_{n2} & \cdots \end{pmatrix} = D \begin{pmatrix} f_{11} & f_{12} & \cdots \\ f_{21} & f_{22} & \cdots \\ & \vdots & \\ F_{k1} & F_{k2} & \cdots \\ & \vdots & \\ f_{n1} & f_{n2} & \cdots \end{pmatrix}, \tag{2.16}$$

where D is a diagonal $n \times n$ matrix with 1 values in all diagonal entries except in the (k, k)-entry, which has value δ.

We can now construct a lattice $\mathcal{L}$, which is the $\mathbb{Z}$-span on the columns of M, and find a matrix U such that $MU = A$, with $A \in GL_n(\mathbb{Z})$, which gives us an LLL-reduced basis for the lattice. Following (2.16), we can write $M = DM_1$, and so we have $A = DM_1U$.

Taking determinants and absolute values, we have that:

$$\begin{aligned} |\det(A)| &= |\det(DM_1U)|\,, \\ &= |\det(D)||\det(M_1U)|\,, \\ &= \delta|\det(M_1U)|\,. \end{aligned}$$

Our matrices M_1 and U have entries in $\mathbb{Z}$, and $\det(M_1U) \neq 0$, meaning $\det(M_1U) \geq 1$. Since we have that $\delta > 1$, $|\det(A)| > 1$. Therefore, $\mathcal{P}$ is not strongly primitive. □

As such, it is more important to focus on ensuring we work with strongly primitive polynomials. However, we will look again at weakly primitive polynomials, more specifically, single variable polynomials, later, in chapter 6.

2.4.2 Further Reductions

Our aim remains to make polynomials which we are interested in knowing the Mahler measure of as simple as possible. Looking at strongly primitive polynomials is a good start, but it is natural to ask if there is anything else we can do to further simplify our polynomials.

Before we move on, let us look at an example to motivate why we might ask such a question.

Example 2.4.8. Let $P(x, y) = 1 + x^2y + xy^2$.

We know that $\text{prim}(P) = 1 + y + xy$, and this in particular means that $M(P) = M(\text{prim}(P))$. However, consider the polynomial $Q(x, y) = 1 + x + y$. We have that $M(Q) = M(P)$.

We further note that if we take $\text{prim}(P) = 1 + y + xy$, and do a substitution taking $x \mapsto x/y$ and $y \mapsto y$, we then get $Q(x, y) = 1 + x + y$.

Here, we see that the total degree of Q is less than that of P; making it in some sense "more simple". We do note, however, that Example 2.4.8 is a rather basic one. By this, we mean that any method for calculating the Mahler measure of a two-variable polynomial will return an answer for $M(Q)$ and $M(\text{prim}(P))$ in effectively the same time in this instance. At the very least, it makes no difference in practice which is more effective, or quicker, in this instance.

However, this does raise a potential problem, in theory. When we consider polynomials of higher total degree, or with more terms, it could be the case that a strongly primitive, irreducible polynomial is not as simple as we would perhaps like. It could even be a case that polynomials of interest are already strongly primitive and irreducible themselves. As these more complex polynomials are ones which we expect to take a longer time to compute the Mahler measure for, this means we could spend significantly more time on calculations than necessary.

As such, it would be beneficial to know if there is a way to further simplify our polynomials, or if there is another explanation for why these polynomials share the same Mahler measure. If it is the former, we can then create and implement a routine to further simplify any given polynomial before calculating its Mahler measure. If it is the latter, we have to accept the likelihood of increased time spent on calculations.

In the instance of Example 2.4.8, there is a way to further simplify $\text{prim}(P)$ to Q, which we outline in chapter 6, when we formalise the concept presented as a notion of "equivalence". The reasoning for not detailing this idea of further reduction is as follows:

- The theory is still in its infancy. Whilst we have ways to, in some sense, simplify some polynomials further, we do not have a complete understanding of the theory.
- There is little need to introduce such reduction now. For the polynomials we will look at for our experiments in chapter 5, this concept of simplification is not particularly necessary in most instances.
- This notion of "equivalence" we introduce opens discussion to a broader topic of study. Indeed, as we will see in chapter 6, we focus on a more general question of 'When do polynomials share the same Mahler measure?' rather than 'Can we simplify a polynomial without changing its Mahler measure?'.

In short, what we have firmly introduced so far is enough for the calculations presented in this thesis, and any further techniques for making our calculations simpler are still being improved upon.

2.5 Dimension vs Variable

In this Section, we address an interesting subtlety which arises particularly when considering small Mahler measures of polynomials in more than one variable. As mentioned in Section 1.2, the smallest known Mahler measure coming from two-variable polynomials is $1.25543386\cdots$. However, this is not exactly true.

Example. Let $P(x,y) = y(x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1)$. More precisely, $P(x,y) = y\Lambda(x)$.

By Lemma 1.2.6, we know that the Mahler measure of polynomials in several variables is multiplicative. Thus:

$$M(P) = M(y)M(\Lambda(x)) = \lambda\,.$$

This example certainly contradicts our earlier claim, and it is also easy to see why. A tempting quick fix would be to be more precise with our claim, and say that the smallest known Mahler measure coming from *irreducible* two-variable polynomials is $1.25543386\cdots$. However, once again, this is not exactly true.

Example. Let $P(x,y) = \Lambda(xy)$. More explicitly, we have that:

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

In this case, P is irreducible, but calculating the Mahler measure shows $M(P) = \lambda$.

This does mean that we can find plenty of examples of two-variable polynomials, and, indeed, irreducible polynomials, with Mahler measure smaller than $1.25543386\cdots$. For example, there are 236 examples of single variable polynomials with Mahler measure less than 1.25 (tiny Mahler measures). If we make any monomial substitution – that is to say, if we replace the single variable with any monomial in two variables – to any of these polynomials with tiny Mahler measure, we will find a two-variable polynomial with Mahler measure less than $1.25543386\cdots$.

This minor issue arises in some texts and situations. Ultimately, this is down to a somewhat unspoken convention about how precise we are with our restrictions and definitions of "smallness" of Mahler measures coming from polynomials in several variables. We address this situation explicitly here, to add clarity.

Firstly, we make the following claim:

Lemma 2.5.1. *We can calculate the Mahler measure of a Laurent polynomial in any number of variables.*

It is not immediately obvious that this should be true. As such, we give the following argument to help justify this:

Proof. Similarly to Definition 2.4.2, we write ${\mathbf{z}_n}^{\mathbf{j}} = {z_1}^{j_1} \cdots {z_n}^{j_n}$. Let $\mathcal{P} \in \mathbb{Z}[z_1, {z_1}^{-1}, \cdots, z_n, {z_n}^{-1}]$ be a Laurent polynomial. A somewhat crude way of writing $\mathcal{P}$ is:

$$\mathcal{P}(\mathbf{z}_n) = \sum_{\mathbf{j} \in J} a_{\mathbf{j}} {\mathbf{z}_n}^{\mathbf{j}}\,,$$

where $J = J_1 \oplus J_2 \oplus \cdots \oplus J_n$, with each $a_{\mathbf{j}} \in \mathbb{Z}$ and each $J_i \subset \mathbb{Z}$.

Now, for each i, define $t_i = |\min_{j \in J_i} j|$. Then, we can find the linear polynomial ${z_1}^{t_1} \cdots {z_n}^{t_n} \mathcal{P}(\mathbf{z}_n)$, and we note that $M({z_1}^{t_1} \cdots {z_n}^{t_n} \mathcal{P}(\mathbf{z}_n)) = M(\mathcal{P}(\mathbf{z}_n))$. □

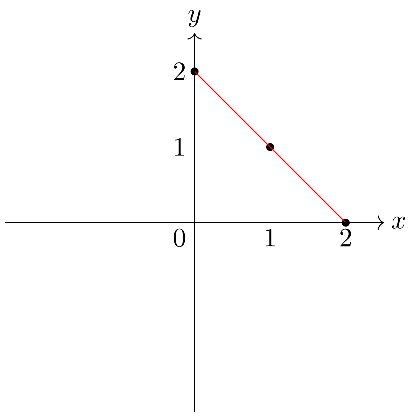
We now work towards defining the *dimension* of a polynomial. However, it is perhaps best understood by first looking at an example, where we introduce necessary concepts as we go along. We will make a formal definition afterwards.

Example 2.5.2. Let $P(x, y) = x^2 + xy + y^2$.

Whilst the powers of each monomial are positive, we can still adopt the notation introduced earlier. In other words, we can consider writing our polynomial in the form $P(x, y) = \sum_{\mathbf{j} \in J} a_{\mathbf{j}} \mathbf{z}_n{}^{\mathbf{j}}$, where $J = J_1 \oplus J_2$.

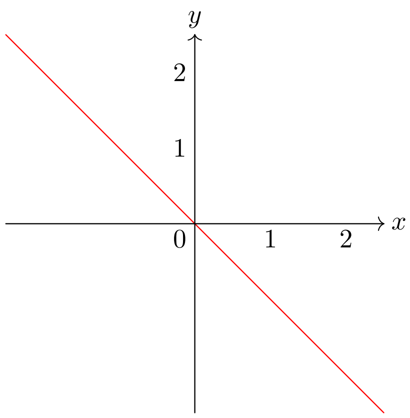
We first find the *convex hull* of the vectors $\mathbf{j}$. This is the intersection of all convex sets containing all the $\mathbf{j}$, where these $\mathbf{j}$ are elements in $\mathbb{Z}^2$, such that $a_{\mathbf{j}} \neq 0$. We call the resulting object the *exponent polytope* of P.

In this case, these vectors are $\mathcal{J} = \{(2, 0), (1, 1), (0, 2)\}$. So, the exponent polytope of P is the straight line, of gradient -1, joining $(0, 2)$ to $(2, 0)$, as seen in the following:



.

Next, we shift this exponent polytope so that it passes through the origin, and then look at the subspace it spans after suitable scaling:



.

We then define the *dimension* of the polynomial P, which we denote as $\dim(P)$, to be the dimension of this subspace. In this case, it is clear that the dimension of this subspace, a straight line contained in the plane, is 1, so we have that $\dim(P) = 1$.

Though this example is rather basic, it does demonstrate the heart of the concept. We now formalise the concepts we have just introduced:

Definition 2.5.3. *Let* $\mathcal{P}(\boldsymbol{z_n}) = \sum_{\boldsymbol{j}\in J} a_{\boldsymbol{j}}\boldsymbol{z_n}^{\boldsymbol{j}}$*, where* $J = J_1 \oplus J_2 \oplus \cdots \oplus J_n$*, with each* $a_{\boldsymbol{j}} \in \mathbb{Z}$ *and each* $J_i \subset \mathbb{Z}$.

The ***exponent polytope*** *of* $\mathcal{P}$ *is the convex hull of the* $\boldsymbol{j} \in \mathbb{Z}^n \subset \mathbb{R}^n$ *such that* $a_{\boldsymbol{j}} \neq 0$.

Definition 2.5.4. *The* ***dimension*** *of* $\mathcal{P}$ *is the dimension of the subspace spanned by the exponent polytope of* $\mathcal{P}$*, after it has been shifted to pass through the origin.*

The dimension of the polynomial has a further benefit. If we have a polynomial $\mathcal{P}$ in n variables such that its dimension is $k < n$, then we can transform our polynomial into a Laurent polynomial in k variables. This can, in theory, simplify calculations for finding the value of $M(\mathcal{P})$.

Example 2.5.2 (Continued)**.** For $P(x, y) = x^2 + xy + y^2$, we know that $\dim(P) = 1$. We now demonstrate how we can transform P into a single variable polynomial. To do this, we take any given point on the exponent polytope, (j_1, j_2), and consider $P(x, y)/x^{j_1}y^{j_2}$.

For example, taking $(j_1, j_2) = (1, 1)$ gives us:

$$\frac{P(x,y)}{xy} = \frac{x}{y} + 1 + \frac{y}{x}\,.$$

We then find a suitable substitution to transform this into a single variable Laurent polynomial. We note that the substitution $\frac{x}{y} = z$ gives us the polynomial $z + 1 + 1/z$.

When can then calculate the Mahler measure of this fairly straightforwardly. We have that:

$$\begin{aligned} M(z + 1 + 1/z) &= M\left(z(z + 1 + 1/z)\right) \\ &= M(z^2 + z + 1)\,. \end{aligned}$$

We know that $\Phi_4(z) = z^2 + z + 1$, so we have that $M(z + 1 + 1/z) = 1$, and thus $M(P) = 1$.

Though Example 2.5.2 is somewhat trivial, it again demonstrates the heart of the concepts being discussed here. In particular, we can see how we can simplify calculations by considering the dimension of the polynomial, but also we can now address our earlier concern.

The value $1.25543386\cdots$ is the smallest known Mahler measure that can be found from irreducible polynomials of dimension two. Polynomials of dimension two are in at least two variables, and any that are in more than two variables can be transformed into a polynomial of two variables. The adopted convention is that when we discuss the smallest Mahler measure that can be obtained from a polynomial in n variables, we do mean the smallest Mahler measure that can be obtained from an irreducible polynomial of dimension n.

We can also look at this from a slightly different perspective.

Definition 2.5.5. *Let α be an algebraic number which is the Mahler measure of some irreducible, non-zero integer polynomial $\mathcal{P}$. The* ***dimension*** *of α, denoted* $\dim(\alpha)$*, is the smallest k such that $\alpha = M(\mathcal{P})$ for such a $\mathcal{P}$ in k variables.*

From this, we can refocus some of the claims we have previously made. For example, instead of saying that the value $1.25543386\cdots$ is the smallest known Mahler measure that can be found from irreducible polynomials of dimension two, we can say that is the smallest known Mahler measure value of dimension two.

Example 2.5.6.

Claim: We have that $\dim(1) = 0$.

Justification: Let $\mathcal{P} = 1$. This is an irreducible, non-zero and integer polynomial of zero variables. As such, 0 is the smallest k such that $1 = M(\mathcal{P})$ and $\mathcal{P}$ has k variables.

Remark. This can be extended for all positive integers.

We can encapsulate some more of our knowledge about dimension, both of a value and a polynomial, in the following result:

Proposition 2.5.7. *Let α the Mahler measure of some n-variable polynomial $\mathcal{P}$. Then:*

$$\dim(\alpha) \leq \dim(\mathcal{P}) \leq n\,. \tag{2.17}$$

Proof. We first show that $\dim(\mathcal{P}) \leq n$. In fact, this follows fairly trivially from Definition 2.5.4.

Write:

$$\mathcal{P}(z_1, \cdots, z_n) = \sum_{\underline{i} \in \mathbb{Z}^n} c_{\underline{i}} \underline{z}^{\underline{i}}\,,$$

where $\underline{z}^{\underline{i}} = {z_1}^{i_1} \cdots {z_n}^{i_n}$, $\underline{i} = (i_1, \cdots, i_n)$ and only finitely many of the $c_{\underline{i}} \neq 0$. Furthermore, multiplying $\mathcal{P}$ by a suitable Laurent monomial means we can suppose that $c_{\underline{0}} \neq 0$.

We know that $\dim(\mathcal{P})$ is equal to the dimension of the space spanned by the $\underline{i}$ for which $c_{\underline{i}} \neq 0$. Since this space is a subspace of $\mathbb{R}^n$, this means that $\dim(\mathcal{P}) \leq n$.

We now show that $\dim(\alpha) \leq \dim(\mathcal{P})$. We let $\dim(\mathcal{P}) = d$.

Let $\mathcal{L}$ be the lattice formed by the $\mathbb{Z}$-span of the $\underline{i}$ for which we have non-zero $c_{\underline{i}}$. These are integer linear combinations of the $\underline{i}$. We can now choose a $\mathbb{Z}$-basis of $\mathcal{L}$:

$$\begin{aligned} \underline{e_1} &= (e_{1,1}, \cdots, e_{1,n})\,, \\ \underline{e_2} &= (e_{2,1}, \cdots, e_{2,n})\,, \\ &\vdots \\ \underline{e_d} &= (e_{d,1}, \cdots, e_{d,n})\,. \end{aligned} \tag{2.18}$$

Additionally, for all $\underline{i}$ such that $c_{\underline{i}} \neq 0$, we have that $\underline{i}$ can be written as a $\mathbb{Z}$-linear combination of these $\underline{e_k}$.

We can also consider extending to a $\mathbb{Z}$-basis for $\mathbb{Z}^n$, by introducing further basis vectors $\underline{e_{d+1}}, \cdots, \underline{e_n}$, defined as in (2.18). We can then consider:

$$\begin{aligned} y_1 &= {z_1}^{e_{1,1}} {z_2}^{e_{1,2}} \cdots {z_n}^{e_{1,n}}\,, \\ y_2 &= {z_1}^{e_{2,1}} {z_2}^{e_{2,2}} \cdots {z_n}^{e_{2,n}}\,, \\ &\vdots \\ y_d &= {z_1}^{e_{d,1}} {z_2}^{e_{d,2}} \cdots {z_n}^{e_{d,n}}\,. \end{aligned} \tag{2.19}$$

These are Laurent monomials in $z_1, \cdots, z_n$.

We can now write each $\underline{i}$, for which $c_{\underline{i}} \neq 0$, as a $\mathbb{Z}$-linear combination of $\underline{e_1}, \cdots, \underline{e_d}$. This gives us:

$$\mathcal{P}(z_1, \cdots, z_n) = \sum_{\substack{\underline{i} \in \mathbb{Z}^n \\ c_{\underline{i}} \neq 0}} c_{\underline{i}} \, \underline{y}^{\underline{u}(\underline{i})} , \tag{2.20}$$

where $\underline{u}(\underline{i})$ is the $\mathbb{Z}$-linear combination of $\underline{e_1}, \cdots, \underline{e_d}$ which depends on $\underline{i}$, and $\underline{y}^{(u_1, \cdots, u_d)} := {y_1}^{u_1} \cdots {y_d}^{u_d}$.

We have presented $\mathcal{P}$ as a Laurent polynomial in d variables in (2.20). However, we can extend the system presented in (2.19), and consider n monomials in $z_1, \cdots, z_n$:

$$\begin{aligned} y_1 &= {z_1}^{e_{1,1}} {z_2}^{e_{1,2}} \cdots {z_n}^{e_{1,n}} , \\ y_2 &= {z_1}^{e_{2,1}} {z_2}^{e_{2,2}} \cdots {z_n}^{e_{2,n}} , \\ &\vdots \\ y_d &= {z_1}^{e_{d,1}} {z_2}^{e_{d,2}} \cdots {z_n}^{e_{d,n}} , \\ y_{d+1} &= {z_1}^{e_{d+1,1}} {z_2}^{e_{d+1,2}} \cdots {z_n}^{e_{d+1,n}} , \\ &\vdots \\ y_n &= {z_1}^{e_{n,1}} {z_2}^{e_{n,2}} \cdots {z_n}^{e_{n,n}} . \end{aligned} \tag{2.21}$$

The system presented here in (2.21) is in fact invertible. Thus, we can apply Proposition 2.4.3, to give:

$$M(\mathcal{P}) = M \left(\sum_{\substack{\underline{i} \in \mathbb{Z}^n \\ c_{\underline{i}} \neq 0}} c_{\underline{i}} \, \underline{y}^{\underline{u}(\underline{i})} \right) .$$

Hence, $\dim(M(\mathcal{P})) \leq d = \dim(\mathcal{P})$, and so we are done. □

A consequence here is that if we are given some value α, and know that $\alpha = M(\mathcal{P})$ for some $\mathcal{P}$, we are able to determine $\dim(\alpha)$. From Proposition 2.5.7, we know that $\dim(\alpha) \leq \dim(\mathcal{P})$, and that this inequality is only strict when α is not a positive integer. So, we have that $\dim(\alpha) = 0$ if α is positive integer, and $\dim(\alpha) = \dim(\mathcal{P})$ otherwise.

The main relevance of this for this thesis moving forward is when we turn to extensive computations of finding Mahler measures of two-variable polynomials in chapter 5. We need to be alert to the possibility that some the Mahler measure values we are considering do, in fact, have dimension less than two. In general, we will continue to refer to "two-variable polynomials", under the assumption that we mean "two-variable polynomials whose dimension is also two". Whilst our focus later will be on Mahler measure values of dimension two, our methods for finding these values apply only to polynomials of two variables, and not any higher, hence the need to make this distinction.

Chapter 3

Digraph Theory

In this chapter, we expand upon Section 1.3.2, which introduced the concept of the Mahler measure of a graph. We first look at the known results related to simple graphs, and how these particular combinatorial objects are not overly helpful in our quest to find small Mahler measures. From here, we move to looking at more general graph theoretic objects, which will be much more fruitful with helping us find small Mahler measures, but come at a cost: we lose some powerful graph theoretic tools. This means we have to develop different techniques for working with digraphs specifically.

Small parts of Section 3.2 have previously appeared in Coyston and McKee [9].

3.1 Simple Graphs and Mahler Measures

For avoidance of doubt, we first recall and fix some straightforward definitions:

Definition 3.1.1. *A* ***graph*** *G is a pair (V, E), where V is a set of elements called* ***vertices****, and E is a set of elements of paired vertices, called* ***edges****.*

There are many algebraic notions which are of interest to those studying graphs, though the one of main interest to us here is the *adjacency matrix*, and the resulting characteristic polynomial. For a graph G on n vertices, one can label the vertices $1, 2, \cdots, n$. We then have that the adjacency matrix of the graph is an $n \times n$ matrix, say

A_G, with entries a_{ij} such that:

$$a_{ij} := \begin{cases} 1, & \text{if } ij \text{ is an edge}\,, \\ 0, & \text{otherwise}\,. \end{cases}$$

The characteristic polynomial of the graph, χ_G, is then defined to be the characteristic polynomial of A_G. It is also a well known fact that the characteristic polynomial of the graph is invariant of how we label the vertices of our graph.

A graph is a very general object. However, there are several different, more particular, variants of a graph:

- We can *direct* our edges; specifying the order in which the paired vertices are viewed.
- We can attach *weights* to our edges and vertices; associating a (usually integral) number to each edge and vertex.
- We can have *multiple edges* between a pair of vertices; that is to say, E can be a multiset.
- We can have *loops*; an edge where the pair of vertices are the same.

Depending on the situation and need, some of these additions can make working with graphs more complicated. However, there is much that can still be gained from studying the "simplest" type of graphs.

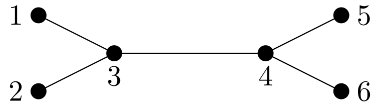
Definition 3.1.2. *A* ***simple graph*** *is a graph which is undirected, unweighted and has neither multiple edges nor loops.*

Though many authors will refer to these plainly as *graphs*, we will be more precise and always refer to these as simple graphs from now. If we refer to a "graph", it should be assumed that this is the most general notion of a graph.

We now recall Definition 1.3.1, in simpler terms:

Definition 3.1.3. *The* ***Mahler measure of a graph*** G *is the Mahler measure of its adjacency matrix.*

Example. Let G be the following simple graph, with its vertices labelled 1 to 6:



.

We have that the adjacency matrix of G is:

$$A_G = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix},$$

which gives us characteristic polynomial: $\chi_G(x) = x^6 - 5x^4 + 4x^2$.

We then calculate the reciprocal polynomial of this, as defined in Definition 1.1.12, giving us: $R_G(z) = z^{12} + z^{10} - z^8 - 2z^6 - z^4 + z^2 + 1$. We finally note that $M(R_G) = 1$, and so $M(G) = 1$.

Example. Consider K_4, the complete graph on 4 vertices. The characteristic polynomial is $\chi_{K_4} = x^4 - 6x^2 - 8x - 3$, and $R_{K_4} = z^8 - 2z^6 - 8z^5 - 9z^4 - 8z^3 - 2z^2 + 1$. Thus, $M(K_4) = 2.61803398\cdots$.

The latter example here shows that Mahler measures of simple graphs are not always trivial or small, and that even common simple graphs do not give us small Mahler measures.

A natural question is when do simple graphs have trivial Mahler measure, and when do they have small Mahler measure. Fortunately, both of these questions have been answered.

3.1.1 Interlacing

Interlacing is a powerful tool which relates specifically to matrices. By considering adjacency matrices of simple graphs, it can also be a powerful tool in graph theory too.

Definition 3.1.4. *For any real, symmetric $n \times n$ matrix M, we can label the eigenvalues of M in non-increasing order: $\mu_1(M) \geq \mu_2(M) \geq \cdots \geq \mu_n(M)$.*

*Let A and B be real, symmetric matrices of size $n \times n$ and $m \times m$ respectively, with $n \geq m$. We say that the eigenvalues of B **interlace** the eigenvalues of A if, for $i = 1, 2, \cdots, m$, we have:*

$$\mu_i(A) \geq \mu_i(B) \geq \mu_{n-m+i}(A) .$$

Proposition 3.1.5 (Interlacing Theorem)**.** *Let A be a real, symmetric $n \times n$ matrix, with eigenvalues $\mu_1(A) \geq \mu_2(A) \geq \cdots \geq \mu_n(A)$. Let B be a submatrix of A obtained by deleting the i-th row and i-th column of A, with eigenvalues $\mu_1(B) \geq \mu_2(B) \geq \cdots \geq \mu_{n-1}(B)$. Then, the eigenvalues of B interlace with those of A; that is to say:*

$$\mu_1(A) \geq \mu_1(B) \geq \mu_2(A) \geq \cdots \geq \mu_{n-1}(B) \geq \mu_n(A) .$$

A proof of this can be found in Godsil and Royle [13, p.193], who give a more general statement.

It is easy to see the relation and possible uses related to simple graphs here. Indeed, there is a common graph theoretic object which exists:

Definition 3.1.6. *Let $G = (V, E)$ be a graph. An **induced subgraph** of G is the graph formed by taking a subset of the vertices V, say V_I, and all edges of E that connect any vertices in V_I together.*

So, for a graph G, if we know its eigenvalues (that is to say, if we know the eigenvalues of its adjacency matrix), and we then remove a vertex, as well as all edges incident with it, then the eigenvalues of this induced subgraph interlace with the eigenvalues of G.

We will use interlacing to show that, for an induced subgraph H of a simple graph G, $M(H) \leq M(G)$. However, our current definition for the Mahler measure of a graph does not make use of the eigenvalues of its adjacency matrix. As such, we need the following statement to help us:

Lemma 3.1.7. *Let A be an $n \times n$ matrix and $\mu_1(A), \cdots, \mu_n(A)$ be its eigenvalues. Then:*

$$M(A) = \prod_{i=1}^{n} \max \left\{ \max\left(1, \frac{|\mu_i(A) + \sqrt{\mu_i(A)^2 - 4}\,|}{2}\right), \max\left(1, \frac{|\mu_i(A) - \sqrt{\mu_i(A)^2 - 4}\,|}{2}\right) \right\}. \tag{3.1}$$

Prior to proving this, it is helpful to further understand the behaviour of the two functions included here.

Lemma 3.1.8. *Let:*

$$f(x) = \frac{|x + \sqrt{x^2 - 4}\,|}{2} \quad \textit{and} \quad g(x) = \frac{|x - \sqrt{x^2 - 4}\,|}{2}\,.$$

Then:

I. We have $f(x)g(x) = 1$ for all $x \in \mathbb{C}$,

II. On the interval $[-2, 2]$, $f(x) = g(x) = 1$,

III. Outside the interval $[-2, 2]$, we have that $f(x)$ is strictly increasing, and $g(x)$ is strictly decreasing.

We also provide a visualisation in Figure 3.1, showcasing $f(x)$ and $g(x)$, to further demonstrate the behaviour of these functions.

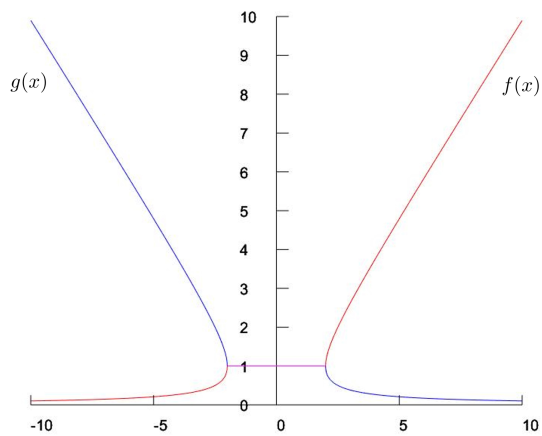


Figure 3.1: The functions $f(x)$ and $g(x)$.

Proof. We prove each claim individually.

Proof of Claim I: This is easy to see directly:

$$\begin{aligned} f(x)g(x) &= \frac{|x+\sqrt{x^2-4}\,|}{2}\frac{|x-\sqrt{x^2-4}\,|}{2} \\ &= \frac{|x^2-\big(x^2-4\big)\,|}{4} \\ &= 1\,. \end{aligned}$$

Proof of Claim II: Let $x \in [-2, 2]$ and consider $f(x)$. As $x \in [-2, 2]$, we can write:

$$\sqrt{x^2-4} = i\sqrt{4-x^2}\,.$$

Thus, we have:

$$\begin{aligned} f(x) &= \frac{|x+i\sqrt{4-x^2}\,|}{2} \\ &= \frac{\sqrt{x^2+(4-x^2)}}{2} \\ &= 1\,. \end{aligned}$$

As, by Claim I, $f(x)g(x) = 1$ for all $x \in \mathbb{R}$, we thus have $g(x) = 1$ in $[-2, 2]$ as well.

Proof of Claim III: We split this into two different cases. Firstly, consider $x > 2$. In this case, $f(x) = \frac{x+\sqrt{x^2-4}}{2}$. Next, we note that that:

$$f'(x) = \frac{1}{2} + \frac{x}{2\sqrt{x^2-4}} > 0\,,$$

and so $f(x)$ is strictly increasing. Equally, as $f(x)g(x) = 1$, this must mean $g(x)$ is strictly decreasing when $x > 2$.

Now consider $x < -2$. In this case, $f(x) = \frac{-x-\sqrt{x^2-4}}{2}$. Next, we note that that:

$$f'(x) = -\frac{1}{2} - \frac{x}{2\sqrt{x^2-4}} > 0\,,$$

and so again $f(x)$ is strictly increasing and $g(x)$ is strictly decreasing when $x < -2$. $\square$

Proof of Lemma 3.1.7. We recall that $M(A) = M(R_A(z))$, where we write $R_A(z) = z^n \chi_A(z + 1/z)$.

As $\chi_A(\mu_i(A)) = 0$ for each i, set $z + 1/z = \mu_i(A)$. Solving for z here gives us: $z = \frac{\mu_i(A) \pm \sqrt{\mu_i(A)^2 - 4}}{2}$. Due to our understanding of this function from Lemma 3.1.8, we can put this into (1.2), the result then follows. □

This now gives us enough to show that, for some simple graph G, the Mahler measure of any induced subgraph of is at most $M(G)$.

Proposition 3.1.9. *Let G be a simple graph, and H an induced subgraph of G. Then, $M(H) \leq M(G)$.*

Proof. We prove this by induction. We write H_k to represent an induced subgraph of G that has had k vertices removed.

Now, let A and B_1 be the adjacency matrices of G and H_1 respectively. As H_1 is an induced subgraph of G, the eigenvalues of B_1 will interlace with those of A:

$$\mu_1(A) \geq \mu_1(B_1) \geq \mu_2(A) \geq \cdots \geq \mu_{n-1}(B_1) \geq \mu_n(A)\,.$$

When combined with (3.1), it is easy to see that $M(B_1)$ certainly cannot exceed $M(A)$, with equality possible with appropriate equality of eigenvalues between A and B_1.

We can now continue inductively, since H_{j+1} is an induced subgraph of H_j. Setting $H_k = H$ for a desired k gives us our result.

□

3.1.2 Simple Cyclotomic Graphs

We now briefly look at all simple graphs with trivial Mahler measures.

Definition 3.1.10. *A graph with Mahler measure 1 is called a* ***cyclotomic graph****.*

When searching to see if a simple graph G is cyclotomic, we have two options. The first option would be as we have already seen: find the reciprocal polynomial, R_G, and calculate the Mahler measure of this polynomial (or, alternatively, check if it is Kronecker-cyclotomic). The second option involves looking at its eigenvalues.

Lemma 3.1.11. *A (simple) graph G is cyclotomic if and only if all of its eigenvalues lie in the interval* $[-2, 2]$.

Proof. Let A be the adjacency matrix of G. From Lemma 3.1.7, we see that if all our eigenvalues are such that $\frac{|\mu_i(A) \pm \sqrt{\mu_i(A)^2 - 4}|}{2}$ is at most 1 for each i, then $M(A) = 1$, and so G is cyclotomic.

We next recall Case II from the proof of Lemma 3.1.8. This shows that whenever $\mu_i(A) \in [-2, 2]$, we have that $\frac{|\mu_i(A) \pm \sqrt{\mu_i(A)^2 - 4}|}{2} = 1$. Thus, if all $\mu_i(A) \in [-2, 2]$, $M(G) = 1$.

Now assume there exists some eigenvalue $\mu_j(A)$ outside the interval $[-2, 2]$. Clearly, $|\mu_j(A) \pm \sqrt{\mu_j(A)^2 - 4}| > 2$ and so $\frac{|\mu_j(A) \pm \sqrt{\mu_j(A)^2 - 4}|}{2} > 1$. Thus, if there is such an eigenvalue $\mu_j(A) \notin [-2, 2]$, $M(G) \neq 1$. So, if $M(G) = 1$, then all $\mu_i(A) \in [-2, 2]$. □

Corollary 3.1.11.1. *Let G be a simple cyclotomic graph. Then, any induced subgraph of G is also cyclotomic.*

Proof. This follows immediately from the results of Lemma 3.1.7 and Lemma 3.1.11. □

As such, a categorisation of simple cyclotomic graphs will likely showcase families of simple graphs. Indeed, this is the case:

Proposition 3.1.12 (Smith [38])**.** *The connected cyclotomic simple graphs are precisely the induced subgraphs of the simple graphs: $\widetilde{E}_6$, $\widetilde{E}_7$, $\widetilde{E}_8$, $\widetilde{A}_n$ (for $n \geq 2$) and $\widetilde{D}_n$ (for $n \geq 4$), as seen in Figures 3.2 and 3.3.*

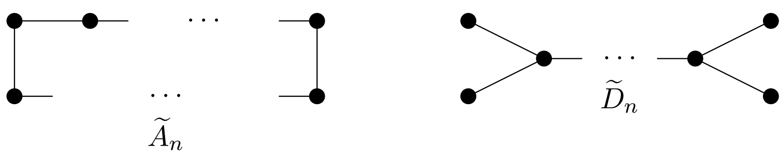


Figure 3.2: The family examples $\widetilde{A}_n$ and $\widetilde{D}_n$.

Remark. We have the convention here that the number of vertices of $\widetilde{A}_n$ and $\widetilde{D}_n$ is one more than the subscript. In other words, these are on $n + 1$ vertices.

Furthermore, it is worth noting that $\widetilde{A}_n = C_{n+1}$, the commonly used notion for the cycle graph on $n+1$ vertices.

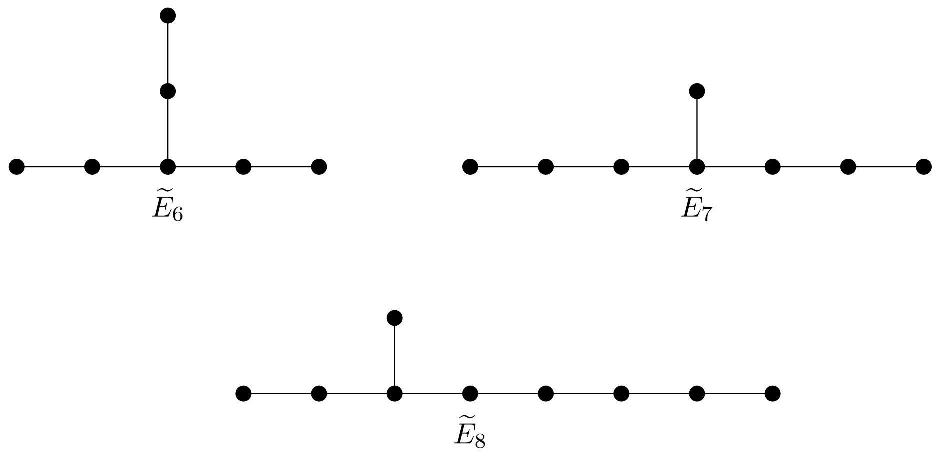


Figure 3.3: The sporadic examples $\widetilde{E}_6, \widetilde{E}_7$ and $\widetilde{E}_8$.

Remark. Again, we note that the number of vertices in each of these examples is one more than the subscript.

We note that in particular, $\widetilde{A}_n$ and $\widetilde{D}_n$ are true 'families', and the remaining three are 'sporadic' examples. An immediate consequence of this result is that we have the cycle graph on n vertices, $C_n = \widetilde{A}_{n-1}$, and the path graph, P_n, are cyclotomic.

3.1.3 Simple Graphs with Small Mahler Measure

We now turn our attention to simple graphs which have small Mahler measure. A complete list of simple graphs with small Mahler measure has been found, which we state in the following Proposition:

Proposition 3.1.13 (McKee and Smyth [24]). *Let $T_{a,b,c}$ be the tree graph represented in Figure 3.4.*

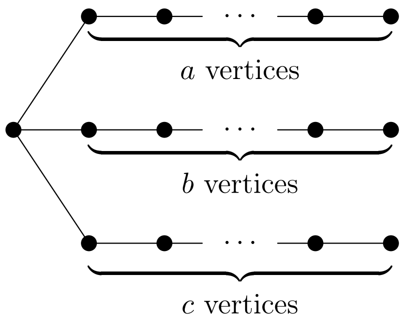


Figure 3.4: The tree $T_{a,b,c}$.

Then, the following are the only small Mahler measures that can be found from any simple graph:

- $M(T_{1,2,6}) = \lambda$,
- $M(T_{1,2,7}) = 1.23039143\cdots$,
- $M(T_{1,2,8}) = 1.26123096\cdots$,
- $M(T_{1,2,9}) = M(T_{1,3,4}) = 1.28063815\cdots$,
- $M(T_{1,2,10}) = 1.29349311\cdots$.

To have only five different Mahler measures be found from simple graphs is, obviously, a very disappointing outcome. If we were to further restrict ourselves to tiny Mahler measures (of which only a finite number are known), we only find two of the current known 236 values, which is again a disappointing return.

We have already mentioned that simple graphs are not as fruitful as we could have hoped they would be for finding small Mahler measures. However, as we have alluded to, there is still much to take away from this. Indeed, McKee and Smyth [24] find small Mahler measures coming from a more broader type of graph (we will explore these types of graphs shortly).

Perhaps a glimmer of hope, however, to take from Proposition 3.1.13 is that they can all be found from a specific family of graphs. Whilst at this stage it is too soon to say with certainty, this at least suggests that, when looking at broader types of graphs and trying to find small Mahler measures, we should aim to look for families of graphs. Indeed, this is a sensible idea to consider in the first place, but this result at least reaffirms that this is worthwhile.

3.2 Digraphs

We have seen that, whilst simple graphs allow us to use powerful graph theoretic tools to aid us, they are not versatile enough to help us find many small Mahler measures. As such, we should look further afield and study more general graphs. However, as previously alluded, this will come at a cost, which we will describe shortly. Therefore, we should not be looking at too general a graph.

We choose to expand the number of graphs we look at by including the following possibilities:

- We allow *arcs*, which roughly speaking can be thought of as a "directed edge", in that the order of the vertices matters. We explain this in more detail shortly.
- We attach weights to our arcs.
- We allow loops on our vertices.

In fact, we will look at a very specific type of graph:

Definition 3.2.1. *A* ***charged signed digraph*** *is a graph G such that:*

- *Each arc is signed; it is given a weight of* $+1$ *(positively signed) or* -1 *(negatively signed),*
- *Each vertex is charged; it is given a value of either* $+1$ *(positively charged),* -1 *(negatively charged) or* 0 *(neutrally charged, or uncharged).*

For sake of simplicity, we refer to these as **digraphs** throughout. Most other authors use "digraph" to mean a simple graph with the inclusion of directed arcs, but that is not the case here, and these should not be confused here.

Upon first viewing, the idea of a charged vertex may seem a little unusual. However, one alternative way of thinking about them, albeit in a somewhat clunky fashion, is as a signed loop attached to a vertex. So, a positively charged vertex is the same as the vertex having a loop of weight 1 attached to it, and a negatively charged vertex is the same as the vertex having a loop of weight -1 attached to it.

Furthermore, we need to take care when talking about edges (and arcs) now. With simple graphs, we define an edge as a pair of vertices, and the order of this pair does not matter. An arc, as previously mentioned, is like a directed edge, in the sense that the order of the pair of vertices matters. For our digraphs, we also have signs attached to them.

So, if we have vertices v_1, v_2, and a signed arc exists from v_1 to v_2, and a signed arc of the opposing sign exists from v_2 to v_1, then there does *not* exist any edge between v_1 and v_2. However, if there are arcs from v_1 to v_2 and v_2 to v_1 with the same sign, we say there is a *signed edge* between v_1 and v_2.

One reason for the particular nomenclature and descriptions of these digraphs is to make the drawing of them as straightforward as possible. There are two different ways for depicting arcs and edges in a digraph, which we see in Tables 3.1 and 3.2.

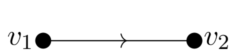
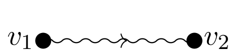
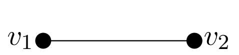
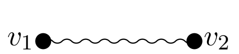
Object	**Meaning**
●	Uncharged vertex
⊕	Positive vertex
⊖	Negative vertex
v_1● ●v_2	No arc between v_1 and v_2
	Positive arc from v_1 to v_2
	Negative arc from v_1 to v_2
	Positive edge between v_1 and v_2
	Negative edge between v_1 and v_2

Table 3.1: All the possible objects that can be used in a digraph, and their literal meaning.

Object	Meaning
$v_1 \bullet\!\!\frac{\;1\;}{0}\!\!\bullet v_2$	Positive arc from v_1 to v_2
$v_1 \bullet\!\!\frac{\;0\;}{1}\!\!\bullet v_2$	Positive arc from v_2 to v_1
$v_1 \bullet\!\!\frac{\;-1\;}{0}\!\!\bullet v_2$	Negative arc from v_1 to v_2
$v_1 \bullet\!\!\frac{\;0\;}{-1}\!\!\bullet v_2$	Negative arc from v_2 to v_1
$v_1 \bullet\!\!\frac{\;1\;}{1}\!\!\bullet v_2$	Positive edge between v_1 and v_2
$v_1 \bullet\!\!\frac{\;-1\;}{-1}\!\!\bullet v_2$	Negative edge between v_1 and v_2

Table 3.2: An alternative way for representing arcs and edges digraphs.

In Table 3.2, we read the value of the sign of an arc between v_1 and v_2 by seeing what number is 'on the left' when travelling along the drawn line from v_1 to v_2.

Example 3.2.2. Figure 3.5 demonstrates a digraph using both of these forms of notation:

Figure 3.5: The same digraph under different notation.

Throughout this thesis, we will use a blended approach when drawing digraphs, utilising depictions from either Table where convenient and sensible to do so.

We fix some final pieces of notation which will be used throughout. If we wish to indicate the presence of a vertex whose charge is not specified, this will be represented as:

$$\circledast \; .$$

Additionally, if we wish to indicate the presence of a vertex whose charge is specified as

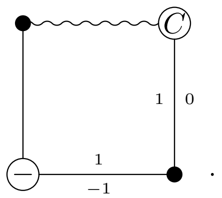
C, but not yet known, this will be represented as:

$$\textcircled{C}\ .$$

Remark. We adopt the follow convention with our notation: A charge C is represented in a matrix entry, and any corresponding equations, by c.

To further increase the clarity of this notation, we have that $C \in \{+, -, \bullet\}$ and $c \in \{+1, -1, 0\}$ respectively. So, C is a pictorial representation of a charge for a vertex, and c is the corresponding numerical value.

Example. The following is a digraph G, which is drawn using our "blended approach" of notation, and has a vertex whose charge is specified as C:



This has adjacency matrix:

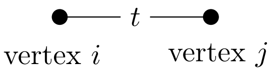
$$A_G = \begin{pmatrix} 0 & -1 & 0 & 1 \\ -1 & c & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & -1 \end{pmatrix}.$$

It may seem unclear why we need to have these two different pieces of notation. In short, there are situations where we may need to represent a vertex whose charge is not specified, but the charge itself is inconsequential to anything which follows in terms of calculations or results. In this case, we use the former piece of notation. However, there are times where the charge will have a consequence, which would be a situation when we use the latter notation.

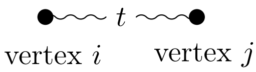
Definition 3.2.3. *We say a path on vertices $v_1, \cdots, v_n$ is **totally positive** if it consists only of positive edges between v_i and v_{i+1}, for $1 \leq i \leq n-1$.*

A parallel definition exists for ***totally negative****.*

In turn, we use

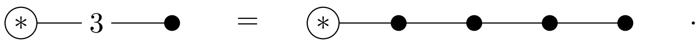


and



to represent totally positive and totally negative paths with t additional neutral vertices (and so $t+1$ positive or negative edges respectively) between vertices i and j.

For example, we have:



Remark. For avoidance of doubt, whenever we use these representations in visualisations of digraphs, it will only ever be to add neutral vertices, regardless of the charge of vertices at either end of the path.

In terms of the adjacency matrix, digraphs are much more broad than simple graphs. The adjacency matrices of simple graphs only have $\{0, 1\}$-entries, are symmetric and only have 0-entries on the diagonal. The adjacency matrices of digraphs have $\{-1, 0, 1\}$-entries, are not necessarily symmetric and can have $\{-1, 0, 1\}$-entries on the diagonal.

This means that the number of possible Mahler measure values increases dramatically too. It is easy to see that this gives us significantly more matrices with different collections of eigenvalues, and different characteristic polynomials and, as such, different reciprocal polynomials too. In turn, this gives us hope that studying these digraphs may give us significantly more small Mahler measure values.

3.3 Digraphs and Mahler Measures: Key Examples

3.3.1 Cyclotomic Digraphs: An Unfulfilled Quest

An interesting question to ask is whether there is an extension of Proposition 3.1.12; that is, if there is a complete classification of (connected) cyclotomic digraphs. The answer here is "no". However, we will soon see that there is still a reasonably strong related result, as shown by McKee and Smyth [23], in Theorems 3.3.3 and 3.3.4.

For now, though, it is useful to look at certain examples. Furthermore, we will verify that certain families of digraphs are indeed cyclotomic prior to introducing the related classification from McKee and Smyth, for reasons which will soon become clear.

Example. Let G be the digraph that appears in Figure 3.5.

Labelling the vertices clockwise starting from the top-left vertex, this gives adjacency matrix:

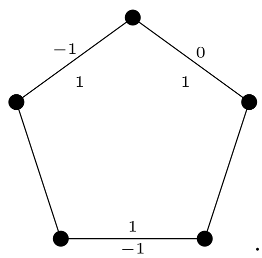
$$A_G = \begin{pmatrix} 0 & -1 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & -1 \end{pmatrix},$$

and characteristic polynomial $\chi_G(x) = x^4 - 2x^2 - x$.

The reciprocal polynomial of this is $R_G(z) = z^8 + 2z^6 - z^5 + 2z^4 - z^3 + 2z^2 + 1$, and this is in fact Kronecker-cyclotomic; we have $R_G(z) = \Phi_3(z)\Phi_4(z)\Phi_{10}(z)$. So G is actually cyclotomic.

This example illustrates also that there are digraphs which are cyclotomic. However, it may lead to a deceiving thought, in that we can just take a cyclotomic simple graph, include charges on vertices or replace edges with signed arcs or edges, and still have a resulting digraph which is cyclotomic. Unsurprisingly, this is not the case:

Example. Let G be the following:



.

We then have:

$$A_G = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ -1 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

This has reciprocal polynomial $R_G(z) = z^{10} + 5z^8 + 11z^6 - z^5 + 11z^4 + 5z^2 + 1$ and $M(G) = 5.24192758\cdots$.

Whilst we do not have a categorisation of all cyclotomic digraphs, we do know of a particular family of digraphs which is cyclotomic, a special type of *charged path.*

For convenience, when referring to any sort of path, we may talk about a specific vertex (the "first vertex", "5-th vertex", "last vertex" and so forth). If we do so, this will be accompanied by a visual representation of a path, and the numbering will be the canonical numbering that we would expect: vertices numbered left to right, 1 (first) to n (last).

Proposition 3.3.1. *Let $P_n^{C_1,C_2}$ be a path on n vertices, whereby the first and last vertices are given some charge, C_1 and C_2 respectively, and all other vertices in the path are neutral.*

Then, $M(P_n^{C_1,C_2}) = 1$.

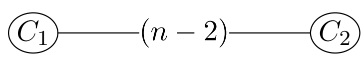


Figure 3.6: $P_n^{C_1,C_2}$.

There are different ways to prove this, but the method we use is particularly relevant for further methods that will be used throughout the thesis. We first need to make use of the following, more general, result:

Lemma 3.3.2. *Let G be a graph on $n \geq 1$ vertices with a distinguished vertex v. For each $m > 0$, let G_m^C be the graph obtained by attaching one endvertex of an m-vertex path to the vertex v, where C represents a non-neutral charge on the other endvertex (so G_m^C has m more vertices than G).*

Let $R_m^C(z)$ be the reciprocal polynomial of G_m^C. Then, for $m \geq 2$, we have:

$$(z + c)R_m^C(z) = z^{2m-1}B(z) + cB^*(z)\,, \tag{3.2}$$

for some monic integer polynomial $B(z)$ that depends on G and v, but not m.

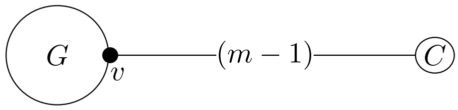


Figure 3.7: A visualisation of G_m^C.

Proof. We begin by noting that the uncharged case is covered by McKee and Smyth [22]. More explicitly, if we let $R_m(z)$ be the related reciprocal polynomial of $G_m = G_m^{\bullet}$, we have that:

$$(z^2 - 1)R_m(z) = z^{2m}B(z) - B^*(z)\,, \tag{3.3}$$

again, for some monic polynomial $B(z)$ depending on G and v, but not m. It is further noted that $B(z) = R_1(z) - R_0(z)$.

For convenience, we refer to the adjacency matrices of G_m^C and G as A and A_G respectively, and the corresponding characteristic polynomials as χ_m^C and χ_m respectively

It is noted by McKee and Smyth [22] that we have a recursive relation for the characteristic polynomial of G_m (valid for $m \geq 2$):

$$\chi_m(x) = x\chi_{m-1}(x) - \chi_{m-2}(x)\,. \tag{3.4}$$

Taking inspiration from this, we can label the graph G_m^C appropriately so that the adjacency matrix is:

$$A = \left(\begin{array}{cccccc|c} c & 1 & & & & & \\ 1 & 0 & 1 & & & & \\ & 1 & 0 & \ddots & & & \\ & & \ddots & \ddots & 1 & & \\ & & & 1 & 0 & 1 & \\ & & & & 1 & 0 & 1 \\ \hline & & & & & 1 & \\ & & & & & & A_G \end{array}\right).$$

Similarly to (3.4), one can find a relation for the characteristic polynomial of G_m^C: $\chi_m^C(x) = (x-c)\chi_{m-1}(x) - \chi_{m-2}(x)$. This is achieved by expanding $\det(xI - A)$ along the first row, at which point it is straightforward to see. We can use (3.4) to simplify this to:

$$\chi_m^C(x) = \chi_m(x) - c\chi_{m-1}(x)\,. \tag{3.5}$$

We then take the reciprocalisation of (3.5), giving us:

$$R_m^C(z) = R_m(z) - czR_{m-1}(z)\,. \tag{3.6}$$

We then substitute in (3.3), which gives us:

$$\begin{aligned} R_m^C(z) &= \left(\frac{z^{2m}}{z^2-1}B(z) - \frac{B^*(z)}{z^2-1}\right) - cz\left(\frac{z^{2(m-1)}}{z^2-1}B(z) - \frac{B^*(z)}{z^2-1}\right) \\ &= \frac{z^{2m-1}(z-c)}{z^2-1}B(z) - \frac{1-cz}{z^2-1}B^*(z)\,, \end{aligned}$$

where $B(z) = R_1(z) - R_0(z)$.

To proceed from here, we first consider explicitly our $1 - cz$ term. Noting that

$c \in \{\pm 1\}$, we have that:

$$
\begin{aligned}
1 - cz &= c(\frac{1}{c} - z) \\
&= c(c - z) \\
&= -c(z - c)\,.
\end{aligned}
$$

Hence, we have that:

$$
\begin{aligned}
R_m^C(z) &= \frac{z^{2m-1}(z-c)}{z^2-1}B(z) + c\frac{z-c}{z^2-1}B^*(z) \\
\Rightarrow \frac{z^2-1}{z-c}R_m^C(z) &= z^{2m-1}B(z) + cB^*(z)\,.
\end{aligned}
$$

Finally, we note that $\frac{z^2-1}{z-c} = z + c$, again since $c \in \{\pm 1\}$, which gives us (3.2), as required. □

This now gives us all we need to prove Proposition 3.3.1:

Proof of Proposition 3.3.1. Let A be the adjacency matrix of $P_n^{C_1,C_2}$. Then:

$$
A = \left(\begin{array}{cccccc|c}
c_1 & 1 & & & & & \\
1 & 0 & 1 & & & & \\
 & 1 & 0 & \ddots & & & \\
 & & \ddots & \ddots & 1 & & \\
 & & & 1 & 0 & 1 & \\
 & & & & 1 & 0 & 1 \\
\hline
 & & & & & 1 & c_2
\end{array}\right),
$$

where $c_1, c_2 \in \{\pm 1\}$ are the values representing the charges of C_1, C_2 respectively. We can then apply Lemma 3.3.2 (by treating the single vertex with charge C_2 as our graph G that we are growing), meaning that the reciprocal polynomial of $P_n^{C_1,C_2}$, say $R_n(z)$, is:

$$
(z + c_1)R_n(z) = z^{2n-1}B(z) + c_1B^*(z)\,. \tag{3.7}
$$

We recall that $B(z) = R_1(z) - R_0(z)$. By construction, we have that $R_0(z)$ and $R_1(z)$

are the reciprocal polynomials of $\chi_0(x) = |x - c_2|$ and $\chi_1(x) = \begin{vmatrix} x & -1 \\ -1 & x - c_2 \end{vmatrix}$ respectively. This gives us that $B(z) = z^4 - c_2 z^3$ and $B^*(z) = -c_2 z + 1$. We can then substitute this into (3.7), which gives us:

$$(z + c_1)R_n(z) = z^{2n-1}(z^4 - c_2 z^3) + c_1(-c_2 z + 1)\,.$$

We now expand this as appropriate and consider each case for c_2 individually:

$$\begin{aligned}(z + c_1)R_n(z) &= z^{2n+2}(z - c_2) + c_1(-c_2 z + 1)\ , \\ &= \begin{cases} (z^{2n+2} - c_1)(z - 1), & \text{for } c_2 = +1\,, \\ (z^{2n+2} + c_1)(z + 1), & \text{for } c_2 = -1\,. \end{cases}\end{aligned} \tag{3.8}$$

In each case represented in (3.8), we can then further consider individual cases for when $c_1 = +1$ and $c_1 = -1$. In each of these four cases, we have that the resulting right hand side of the formula in (3.8) are Kronecker-cyclotomic, meaning that $M\left(R_n^C(z)\right) = 1$. As this is the reciprocal polynomial of $P_n^{C_1,C_2}$, this means $M(P_n^{C_1,C_2}) = 1$, as required. □

We have exhibited a phenomenon here which we refer to as *graph growing*. Graph growing is the process of taking a graph, and adding several vertices in succession, resulting in a new, larger graph. Lemma 3.3.2 is a simple, yet key, tool for this: we attach a path to a graph, ending with a possibly non-neutrally charged vertex.

We have seen how this is useful for proving results related explicitly to paths. However, this will come to be an even more critical tool to use when trying to find small Mahler measures. We will formalise this concept further in chapter 4.

We now turn our attention to the aforementioned results from McKee and Smyth. They were able to give a classification for all cyclotomic *charged signed graphs*. That is to say, objects which can have charged vertices, signed edges, but no signed arcs. So these results do not relate to as broad a combinatorial object as we have studied here, but they are still of use to us.

These results are best described in two separate theorems; one specifically related to signed graphs (i.e. graphs with signed edges, but no charged vertices), and one more broadly related to charged signed graphs:

Theorem 3.3.3 (McKee and Smyth, [23]). *Every maximal connected cyclotomic signed graph is equivalent to one of the following:*

i. For some $k \geq 3$*, the* $2k$*-vertex toral tessellation* T_{2k} *(as seen in Figure C.1),*

ii. The 14*-vertex signed graph* S_{14} *(as seen in Figure C.2),*

iii. The 16*-vertex signed hypercube* S_{16} *(as seen in Figure C.3).*

Further, every connected cyclotomic signed graph is contained in a maximal one, up to equivalence.

Theorem 3.3.4 (McKee and Smyth, [23]). *Every maximal connected cyclotomic charged signed graph not included in Theorem 3.3.3 is equivalent to one of the following:*

i. For some $k \geq 2$*, one of the* $2k$*-vertex cylindrical tessellations* C_{2k}^{++} *or* C_{2k}^{+-} *(as seen in Figure 3.8),*

ii. One of the three sporadic digraphs S_7*,* S_8*,* $S_8{}'$ *(as seen in Figure C.4).*

Further, every connected cyclotomic charged signed graph is contained in a maximal one, up to equivalence.

Remark. The term *equivalence* can have many meanings. We explain the meaning of equivalence here in more detail in Section 3.5, but for now, it suffices to know that there is some notion of equivalence that allows us to capture all digraphs in this set up.

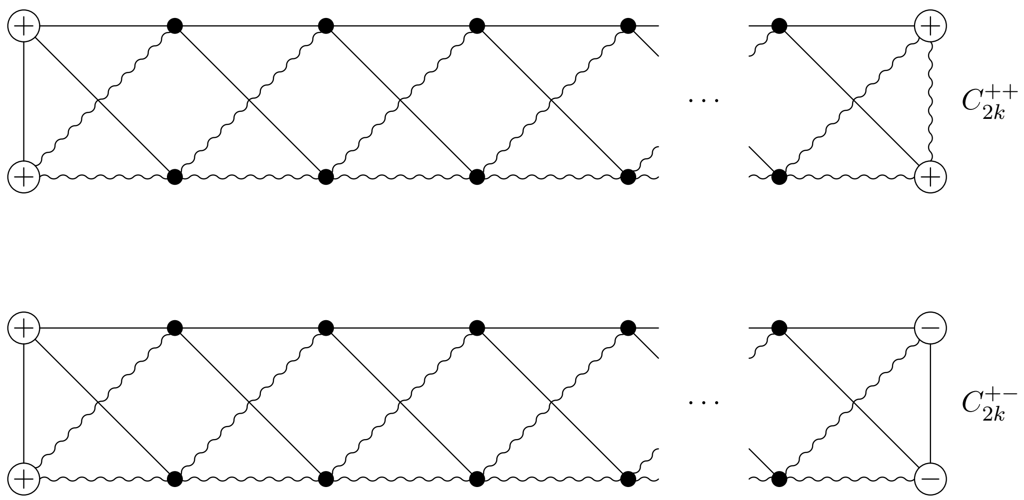


Figure 3.8: The charged signed graphs C_{2k}^{++} and C_{2k}^{+-}, with $2k$ vertices, for $k \geq 2$.

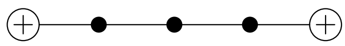
So, if we wish to see if a charged signed graph is cyclotomic, it suffices to check if it is a connected subgraph of one of the maximal examples listed in these theorems (or "equivalent" to one). We will make use of this on several occasions.

We recall that the digraph we saw in Figure 3.5 was cyclotomic. However, as it has signed arcs, it cannot be a subdigraph of any of the listed graphs here, and so this classification would not have been of use to us here. A classification of cyclotomic (charged signed) digraphs remains an open problem.

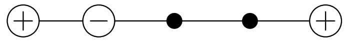
3.3.2 Digraphs with Small Mahler Measure

There are many approaches to searching for digraphs with small Mahler measures. Here, we show some examples of digraphs with small Mahler measure, before outlining a naïve approach that one may consider attempting for trying to find many digraphs with small Mahler measures.

Example. Let $G = P_5^{++}$:


,

and G_1 be the digraph that results in changing the second vertex from a neutral charge to a negative charge:

 .

Then, $M(G_1) = \lambda$.

This exhibits an interesting phenomenon. From Proposition 3.3.1, we know that P_5^{++} is cyclotomic. So, by making a small change to a cyclotomic digraph (changing the charge of a vertex), we have made a small change to the Mahler measure.

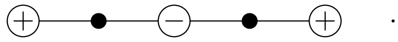
One may quickly ask an interesting question as a result of this: Do small changes to digraphs make small changes to its Mahler measure? This is, naturally, a vague and broad question. To try and increase precision, we could say that "small changes" to a digraph involve just changing the charge of a single vertex, or the sign of a single arc. Alternatively, we could allow a number of these small changes at the same time, so long as this number is significantly smaller than the number of vertices of our digraph.

Heuristically, there is justification for asking this question and believing that we may get some sort of positive result from this. Say we do just make a single change, either to a charge of a vertex or a sign of an arc. In terms of the matrix, this is the equivalent of mapping a single entry from $\{0, \pm 1\}$ to a new value in $\{0, \pm 1\}$. Though there is not much material to support this, we would wishfully expect that the resulting changes in eigenvalues would not be too drastic, especially when searching for Mahler measures.

We also have further justification for believing this is a good avenue to explore. In Section 1.3.1, we saw how we can attach Mahler measures to knots. Furthermore, we noted a situation where we took a knot with associated Mahler measure 1, made a small topological change to that knot, and the resulting knot had small Mahler measure. Though the study of knots proved to be of minimal interest for finding small Mahler measures, the fact that some parallels exist give us hope to at least explore this question further.

Owing to the broad nature of this question, we can focus on a more specific case. In particular, if we take some cyclotomic digraph and make small changes to it, will the new digraph have a small Mahler measure? Unfortunately, the answer here is "no":

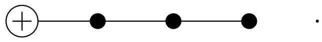
Example. Let $\widetilde{G}_1$ be the digraph that results in changing the middle (third) vertex of $G = P_5^{++}$ from a neutral charge to a negative charge:



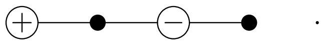
We have that $M(\widetilde{G}_1) = 1.40126836\cdots$.

It is, of course, disappointing that making small changes to cyclotomic digraphs does not always lead to a small enough change in the Mahler measure to give us a new digraph with a small Mahler measure. Yet, it is not wholly surprising either. But there is still hope yet that making small changes to digraphs may be a useful method for trying to find small Mahler measures.

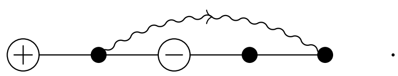
Example 3.3.5. Let $G = P_4^{+\bullet}$:



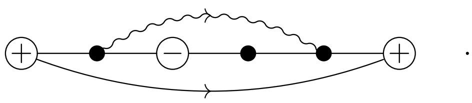
Let G_1 be the digraph that results in changing the third vertex of G from a neutral charge to a negative charge:



Next, let G_2 be the digraph that results in adding a fifth, neutral, vertex to G_1, attached to the fourth vertex of G_1 by a positive edge, as well as having a negative arc from the second and fifth vertex:



Finally, let G_3 be the digraph that results in joining a sixth vertex, which is positively charged, to the fifth vertex of G_2 with a positive edge, as well as having a positive arc from the first vertex to the sixth vertex:



Then, we have that: $M(G) = 1$, $M(G_1) = 1.28063815\cdots$, $M(G_2) = 1.23039143\cdots$ and $M(G_3) = 1.26123096\cdots$.

This again exhibits a way of taking some graph and growing it. However, this is slightly different to our previous concept of graph growing. Here, we start with a cyclotomic digraph, and make a small change, resulting in a new digraph with small Mahler measure. Indeed, we usually are growing it, in that we are adding new vertices in some instances. However, we can also see that we are, in some sense, increasing the complexity: at each instance, we are adding more non-zero terms to the resulting adjacency matrix.

This concept of taking a digraph and growing it requires a lot of care though. For a digraph on n vertices, say we wish to grow the digraph into a new digraph on $n+1$ vertices by attaching the vertex to just one existing vertex. Then, there are $24n$ possible ways to do this growing (since we have n vertices to choose from, 3 possible charges and 8 possible arc combinations to attach a new vertex). If we then include the possibility of adding new arcs between existing vertices too, changing the charge of individual vertices or attaching a new vertex to multiple new vertices, this increases the number of ways to increase the complexity of the digraph even further. Doing this repeatedly becomes increasingly complicated, and potentially time consuming, especially compared to our original concept of graph growing.

We explore the concept of graph growing further in chapter 4, and see explicit results in chapter 5.

3.4 The Problems with Digraphs

As has been mentioned, looking at the more general of these digraphs comes with costs: many results which are true for simple graphs do not extend. Here, we look at such results which do not extend from simple graphs to digraphs, and explain the subsequent consequences. We refer to these as “Non-Propositions”: statements which we write as if they were Propositions, but are false.

Firstly, we look at an attempted extension of Proposition 3.1.5:

Non-Proposition 3.4.1. *Let A be a real $n \times n$ matrix, with eigenvalues written as $\mu_1(A) \geq \mu_2(A) \geq \cdots \geq \mu_n(A)$. Let B be a submatrix of A obtained by deleting the i-th row and i-th column of A, with eigenvalues $\mu_1(B) \geq \mu_2(B) \geq \cdots \geq \mu_{n-1}(B)$. Then, the eigenvalues of B interlace with those of A; that is to say:*

$$\mu_1(A) \geq \mu_1(B) \geq \mu_2(A) \geq \cdots \geq \mu_{n-1}(B) \geq \mu_n(A)\,.$$

As the adjacency matrices of our digraphs are, in general, not symmetric, we need to lose that condition from Proposition 3.1.5. However, as this is the only condition we drop, and we make no subsequent changes to our statement, it should come as no surprise that such a result is not true. As a result, it is not the case that if we have a digraph G and an induced subdigraph H, then $M(G) \geq M(H)$.

A further consequence here relates to Example 3.3.5. There, we were taking a digraph and growing it by adding a single vertex. If we were to do this with a simple graph, we know that the Mahler measure of the new simple graph will be at least the Mahler measure of the original one. This gives our growing of simple graphs a natural end point: we either grow them to the point the resulting Mahler measures are too large to be of interest to us, or they possibly end up always being the same.

With digraphs, as we have no relations between that of the Mahler measure of a grown digraph and that of the original, we have no natural end point. Indeed, we see this occur in Example 3.3.5. We started with a cyclotomic digraph, and then as we grew the digraph by adding (or changing) vertices, as well as additional arcs, the Mahler measures of the resulting digraphs oscillated, whilst remaining small. Whilst this does mean that there is greater scope for finding more digraphs with small Mahler measures, we are inheriting a volatile environment of which we have little understanding.

With these in mind, we can now make another attempt to extend Proposition 3.1.5:

Non-Proposition 3.4.2. *Let A be a real $n \times n$ matrix, with eigenvalues written as $\mu_1(A) \geq \mu_2(A) \geq \cdots \geq \mu_n(A)$. Let B be a submatrix of A obtained by deleting the i-th row and i-th column of A, so that B is symmetric. Write the eigenvalues of B as $\mu_1(B) \geq \mu_2(B) \geq \cdots \geq \mu_{n-1}(B)$. Then, the eigenvalues of B interlace with those of A;*

that is to say:

$$\mu_1(A) \geq \mu_1(B) \geq \mu_2(A) \geq \cdots \geq \mu_{n-1}(B) \geq \mu_n(A)\,.$$

It is clear why it is desirable to have such a result. In terms of digraphs, it is claiming that if we have a digraph G which is "close" to being a charged signed graph (i.e. a charged graph which only has signed edges), then the resulting eigenvalues interlace. By "close", in the simplest sense, we can mean that only one single signed arc exists, and so we can remove one of the vertices it is incident with to arrive at a charged signed graph.

Compared to Non-Proposition 3.4.1, it is slightly less obvious whether this statement is true or false. However, we have already seen a counter-example to disprove this:

Example 3.4.3. Let A be the adjacency matrix of G_2 in Example 3.3.5, and B by the matrix obtained by deleting the fifth row and column of A. This means that B is in fact the adjacency matrix of G_1 from Example 3.3.5:

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

We have already seen that $M(G_2) < M(G_1)$. However, let us look at the eigenvalues explicitly to verify this beyond doubt:

Eigenvalue	A	B
$\mu_1(\,)$	$1.69649\ldots$	$1.76401\ldots$
$\mu_2(\,)$	$1.27683\ldots$	$0.69382\ldots$
$\mu_3(\,)$	$0.19993\ldots$	$-0.39633\ldots$
$\mu_4(\,)$	$-1.13012\ldots$	$-2.06149\ldots$
$\mu_5(\,)$	$-2.04314\ldots$	

We see that, in particular, $\mu_4(B) \not\geq \mu_5(A)$, and so interlacing fails.

And so this is our main problem with digraphs, especially when trying to find small Mahler measures. With no interlacing, and no plausible replacement, growing digraphs becomes more intricate. Let us make a final statement related to simple graphs:

Proposition 3.4.4. *Let G be a simple graph on n vertices, and G_1 be the simple graph obtained by attaching some new $(n+1)$-th vertex to at least one specified vertex in G. Then, $M(G_1) \geq M(G)$.*

Proof. The proof is effectively the same as that of Proposition 3.1.9. □

Indeed, this is just one of the reasons why searching over and growing simple graphs was so desirable in the first place. However, because – as we have seen – there is an oscillating nature of Mahler measures for digraphs which have been grown, there is nothing we can say in this setting. We capture this in one final Non-Proposition:

Non-Proposition 3.4.5. *Let G be a digraph on n vertices, and G_1 be the digraph obtained by attaching some new $(n+1)$-th vertex to at least one specified vertex in G. Then, $M(G_1) \geq M(G)$.*

We will see in chapters 4 and 5 how we tackle this issue. In fact, we see in chapter 5 that this oscillation of Mahler measures is a particularly interesting feature which we can use to expand our knowledge.

One may wonder if the problems we have seen with digraphs extend elsewhere. We have seen that interlacing fails for our digraphs and maybe it was too optimistic to think that such a tool would hold beyond symmetric, $\{0,1\}$-entry matrices (which translate to simple graphs). However, the class of matrices for which interlacing does hold is indeed broader. In fact, interlacing holds for all complex Hermitian matrices: matrices A of the form $A = \overline{A^T}$, and so in particular any real symmetric matrix.

Another known class of matrices for which interlacing does hold is symmetrizable matrices:

Definition 3.4.6. *A real $n \times n$ matrix A is* ***symmetrizable*** *if there is some real diagonal matrix D, with strictly positive entries, such that $D^{-1}AD$ is symmetric.*

Discussion on symmetrizable matrices, including a justification for the fact that interlacing does hold for these matrices, can be found in, for example, McKee and Smyth [25].

3.5 Equal and Opposite: Switching and Equivalence

So far, we have acknowledged that, whilst simple graphs have many powerful tools associated to them, they are particularly limited with regards to finding small Mahler measures. In turn, we overcome these limitations by introducing digraphs. However, whilst we have seen they certainly can cast a wider net for finding small Mahler measures, we have only focused on the difficulties we face when using them.

Here, we explore some straightforward, but useful, results which relate to digraphs more specifically. Some of these results are included out of a sense of interest and for completeness, whilst others will prove more useful later.

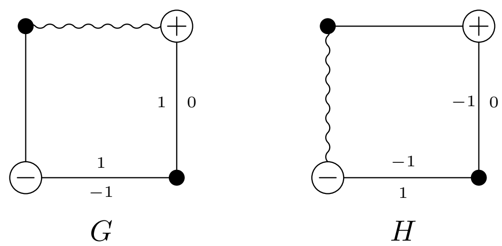
3.5.1 Switching, Degree and Negativity

Switching

Definition 3.5.1. *Let G be a digraph and i a distinguished vertex of G. A* ***sign switch*** *at the vertex i changes the signs of all arcs going in and out of the vertex i, whilst not changing the charge of the vertex i.*

Definition 3.5.2. *We say two digraphs G and H are* ***switch equivalent*** *if there exists a series of sign switches taking G to H.*

Example 3.5.3. The following digraphs G and H are switch equivalent. This can be achieved, for example, by performing sign switches at both of the neutral vertices.



.

Proposition 3.5.4. *A sign switch at any vertex preserves Mahler measure.*

Proof. Let G be a digraph and say we are performing a sign switch at a distinguished vertex labelled i. If we consider the adjacency matrix of G, a sign switch swaps the sign of all entries in the i-th row and column (meaning the (i, i)-entry is unchanged). This does not affect the characteristic polynomial, χ_G.

As such, the reciprocal polynomial R_G also remains the same, and so $M(G)$ will be preserved. □

Corollary 3.5.4.1. *Let G and H be switch equivalent digraphs. Then, $M(G) = M(H)$.*

If we consider the matrix theory behind sign switching, we will note that any series of sign switches has the effect of replacing the adjacency matrix A by $DAD = DAD^{-1}$, where D is a diagonal matrix with ± 1 entries. In fact, the location of the -1 entries of D flag the vertices at which the sign switches are performed. With this in mind, we now turn to the following Theorem which is important in the theory of sign switching.

Theorem 3.5.5. *Let G and H be connected digraphs on n vertices, with corresponding adjacency matrices $A = A_G$ and $B = A_H$ respectively, such that:*

- $a_{ii} = b_{ii}$ *for each* i,
- $a_{ij} = |b_{ij}|$ *for* $1 \leq i \leq n$, $1 \leq j \leq n$, $i \neq j$.

Then, $\chi_A = \chi_B$ if and only if there exists a diagonal matrix $D = \text{diag}(d_1, \cdots, d_n)$, with each $d_i \in \{\pm 1\}$, such that $A = DBD^{-1}$.

A proof of this is contained inside McKee and Smyth [25].

Theorem 3.5.5 gives us a way of testing if, given two digraphs, they are switch equivalent. For us, this is useful as it allows us to see if two digraphs, which may look rather different, share the same Mahler measure.

To explain the usefulness in short, our aim is to be taking different digraphs and growing them to see if the resulting new digraphs have small Mahler measures. If, given two digraphs, we find they are switch equivalent, then it is perhaps worth not growing both digraphs, as it is possible they will admit the same results if grown in the same way. This is particularly prevalent when we are making careful selections for digraphs to grow, and vertices to grow them from.

There is a subtlety around the notion of switch equivalence we have not yet addressed. We need to take into account our labelling of vertices, and how we can permute these. When we consider, for example, the adjacency matrices of digraphs, we have made a choice of the labelling of the vertices, and this choice can be changed and permuted in any way we see fit (since it does not affect other objects of interest, such as the characteristic polynomial). This leads to a concern, since when we consider switch equivalence, we specifically work with the adjacency matrices of digraphs.

As such, we introduce the following notion to help combat this concern:

Definition 3.5.6. *We say two digraphs G and H are* ***equal*** *if there is a bijection between $V(G)$ and $V(H)$ which preserves the charges of the vertices and the signs of arcs.*

A consequence of this means that if two digraphs G and H are equal, and have adjacency matrices A_G and A_H respectively, then $A_G = SA_HS = SA_HS^{-1}$, where S is some permutation matrix. This precisely addresses our concern, since this would reflect a re-labelling of our vertices. So we do not need to be concerned with the labelling of our vertices.

Before moving on from switching, we note an important result related to paths.

Proposition 3.5.7. *Let G be a path on n vertices, such that each vertex v_i has some charge C_i, and each edge (v_i, v_{i+1}) has a sign e_i. Let H be a path on n vertices, such that each vertex w_i has charge C_i (so it is the same charge as the corresponding vertex in G), and each edge (v_i, v_{i+1}) is positively signed.*

Then, G and H are switch equivalent.

Proof. Say that the first negatively signed edge is e_k. We can perform a sign switch at the vertex v_{k+1}. This means that e_k is now a positively signed edge. We then repeat this process, if any negatively signed edges remain, knowing that the next negatively signed edge will be some edge e_t, for some $t > k$. Of course, if no negatively signed edges remain, we are done.

We can repeatedly perform this process to achieve our result. We note, in particular, if we are in a situation where the 'first' negative edge is e_{n-1} (that is, the last edge), we perform a sign switch at the vertex v_n. As e_{n-1} is the only edge formed by v_n, this is the only edge affected by a sign switch. This would mean all edges are positively signed, as required. □

For a very simple visualisation of how a smart choice of vertices to switch at takes us to a case where all edges are positively signed, if we have a totally negative path on n neutral vertices, then switching at every second vertex will transform the path to a totally positive one. The following demonstrates this for a path on 6 and a path on 7 vertices, and it is easy to see how this extrapolates generally depending on when n is even and odd:

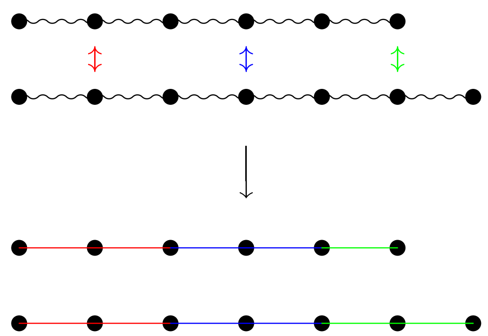


Figure 3.9: Switching along totally negative paths.

Degree

Now, we look to extend the notion of the *degree* of a vertex, which we are familiar with from simple graphs.

Definition 3.5.8. *Let G be a digraph, and v, w be vertices of G. We say v is* ***incident*** *with w if there exists an arc from v to w or w to v. Furthermore, if v has a non-neutral charge, we say that v is incident with itself.*

Remark. Previously, we have also used the word *incident* to mean any arc or edge which is connected to a vertex. Moving forward, when we use the word incident, it is clear from context whether we are referring to vertices being incident, or edges being incident with a vertex.

Definition 3.5.9. *The **degree** of a vertex v in a digraph is equal to the number of vertices incident with v.*

We note that this is a natural generalisation of the degree of a vertex in a simple graph. Furthermore, since a non-neutrally charged vertex is incident with itself, a non-neutral charge contributes 1 to the degree of the vertex.

In terms of the adjacency matrix, say A, the degree of the i-th vertex of a digraph is equal to the number of non-zero entries in the i-th row of A. This is also equal to the (i, i)-entry of A^2.

Proposition 3.5.10. *Let G be a cyclotomic charged signed graph. Then, the degree of any vertex of G is at most 4.*

Proof. Let $A = A_G$ be the adjacency matrix of G, and let d_i be the (i, i)-entry of A^2 (and so the degree of the i-th vertex).

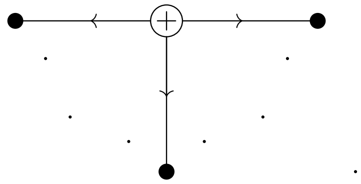
By interlacing, A^2 has an eigenvalue $\mu \geq d_i$, and so A has an eigenvalue μ_1 such that $|\mu_1| \geq \sqrt{d_i}$. As A is cyclotomic, all eigenvalues lie in $[-2, 2]$ (by Lemma 3.1.11) and so $2 \geq |\mu_1|$. Therefore, $d_i \leq 4$. □

A natural question to ask from here is if we can extend the result for cyclotomic (charged signed) digraphs:

Non-Proposition 3.5.11. *Let G be a cyclotomic digraph. Then, the degree of any vertex of G is at most some d, for some $d \in \mathbb{Z}$.*

However, for any given d, we can always find a cyclotomic digraph with a vertex that has degree larger than d, as can be seen by Example 3.5.12 below. That is to say, there is no upper bound for the maximum degree of any vertex of a cyclotomic digraph, at least, in the same way as Proposition 3.5.10. In fact, the only general claim we can make regarding cyclotomic digraphs and degree is that the degree of any vertex cannot exceed the number of vertices of the digraph. However, since the degree of a vertex cannot exceed the number vertices by definition, this is not a useful claim to make.

Example 3.5.12. Let G be following star-like digraph on $n+1$ vertices:



This has adjacency matrix:

$$A = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

From here, we note that $\chi_G(x) = x^n(x-1)$, and so $R_G(z) = (z^2 - z + 1)(z^2+1)^n$. Thus, $M(G) = 1$. We further note that the charged vertex here has degree $n+1$.

Of course, that is not to say there is nothing we could say to link vertex degree and cyclotomic digraphs, but there is little hope of a broad statement akin to Proposition 3.5.10.

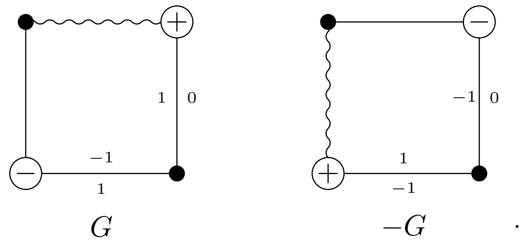
Negativity

We finally turn our attention to what it means to take the *negative* of a digraph:

Definition 3.5.13. *Let G be a digraph and $A = A_G$ be its adjacency matrix, with entries (a_{ij}).*

We say that the ***negative*** *of the digraph G, denoted $-G$, is the digraph whose corresponding adjacency matrix has entries $(-a_{ij})$.*

Example 3.5.14. Here we look at the negative of a cyclotomic digraph we saw Example 3.2.2:



We have that:

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

and in fact $A_G = -A_{-G}$.

We have characteristic polynomial $\chi_{-G}(x) = x^4 - 2x^2 + x$ and reciprocal polynomial $R_{-G}(z) = z^8 + 2z^6 + z^5 + 2z^4 + z^3 + 2z^2 + 1$. This reciprocal polynomial is also Kronecker-cyclotomic: $R_{-G}(z) = \Phi_4(z)\Phi_5(z)\Phi_6(z)$, but we note this is slightly different to $R_G(z)$. Another way of writing this, however, is that $R_{-G}(z) = R_G(-z)$.

Proposition 3.5.15. *Let G be any digraph on $n \geq 1$ vertices. We have that $M(G) = M(-G)$.*

Proof. Let A be the adjacency matrix of G and B be the adjacency matrix of $-G$. So, we have that $B = -A$.

The eigenvalues of B are precisely the negatives of those of A. Indeed, this means that each corresponding eigenvalue will have the same multiplicity as well, since if $A\mathbf{v} = \mu\mathbf{v}$ for some eigenvalue μ, then $B\mathbf{v} = -A\mathbf{v} = -\mu\mathbf{v}$.

Therefore, eigenvalues of A in $[-2, 2]$ correspond to eigenvalues of B in $[-2, 2]$, and thus do not contribute to the Mahler measure (as we have seen in Lemma 3.1.8). Any eigenvalue of A, say μ, greater than 2 contributes a factor of $\frac{|\mu+\sqrt{\mu^2-4}|}{2}$ to the Mahler measure of A (again, following from Lemma 3.1.8). This factor is equal to $\frac{|-\mu-\sqrt{(-\mu)^2-4}|}{2}$, which contributes to the Mahler measure of B, since $-\mu$ is an eigenvalue of B.

As such, the contributions to the Mahler measure are always equal, meaning $M(A) = M(B)$. Therefore, $M(G) = M(-G)$.

□

3.5.2 Equivalent Digraphs

Finally, we return to a remark made after Theorems 3.3.3 and 3.3.4, and formalise the notion of *equivalence* between two digraphs.

Definition 3.5.16. *Let G and H be digraphs on the same number of vertices. We say that G and H are* ***equivalent*** *if one of the following is true:*

- *G and H are switch equivalent,*
- *G and $-H$ are switch equivalent;*
- *G is switch equivalent to a digraph which is equal to H or $-H$.*

Returning to the classification theorems of cyclotomic charged signed graphs, this means that a cyclotomic charged signed graph may be contained inside one of the mentioned maximal examples, or it may be equivalent to one which is contained inside one of these maximal examples. This is the meaning of "up to equivalence" in these statements. As such, we do not see *all* cyclotomic charged signed graph inside the maximal examples, but we do see all up to equivalence.

We recall that prior to proving Proposition 3.3.1 (which states that the path with charges at either end, $P_n^{C_1,C_2}$, is cyclotomic), we mentioned that there were many ways to prove it. The method we used demonstrated the importance of graph growing, which is an important part of this thesis. With our knowledge of these classification theorems and equivalence, we can now give an alternative, simpler, proof:

Alternative Proof of Proposition 3.3.1. We look to use Theorem 3.3.4 and show that $P_n^{C_1,C_2}$ is either contained within one of the maximal examples, or equivalent to one contained within one of the maximal examples, for any combination of charges (C_1, C_2).

Firstly, we note that P_n^{++} is equivalent to P_n^{--}, and $P_n^{+\bullet}$ is equivalent to $P_n^{-\bullet}$. To see this, we can take the negative of one of the digraphs, to change the charges of the vertices, and then apply Proposition 3.5.7 to switch the signs of edges back. As such, we only need to verify that $P_n^{++}, P_n^{+-}, P_n^{+\bullet}$ and $P_n^{\bullet\bullet} = P_n$ are contained with the maximal examples.

This is straightforward to see: P_n^{++} and $P_n^{+\bullet}$ are contained with C_{2k}^{++}. Meanwhile, P_n^{+-} is contained within C_{2k}^{+-}. Finally, P_n is in fact contained within both of these cylindrical tessellations. Hence, $P_n^{C_1,C_2}$ is cyclotomic. □

We round off this Section by showing a certain digraph is cyclotomic, making use both of Theorem 3.3.4 (one of our classification theorems for charged signed graphs) and this notion of equivalence.

Proposition 3.5.17. *The "snake tongue" digraph, S, is cyclotomic.*

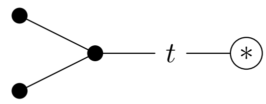
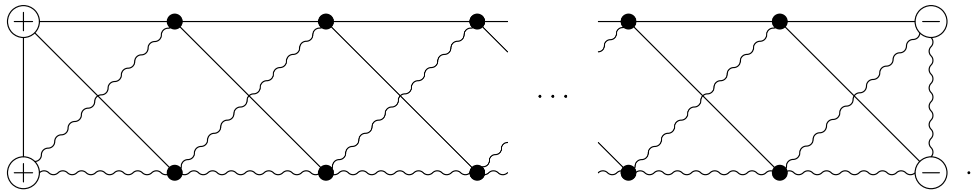


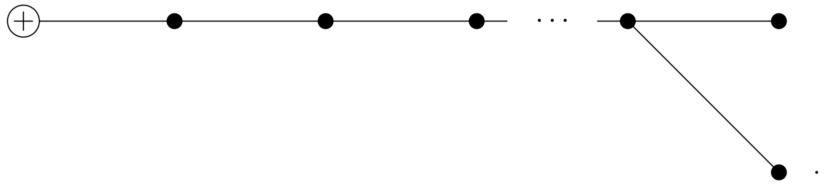
Figure 3.10: The digraph S.

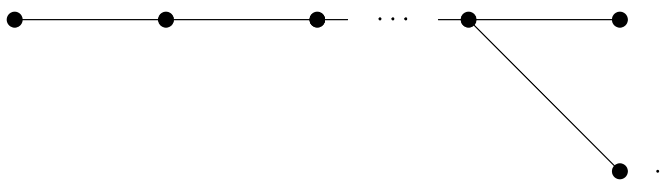
Remark. We note that we have referred to S as a (charged signed) *digraph.* However, more particularly, it is a charged signed graph, which means we will be allowed to use Theorem 3.3.4. We refer to it as a digraph since we will use it within that context later on. Of course, charged signed graphs are a special case of our digraphs, so this is just a technicality.

Proof. We recall the digraph C_{2k}^{+-}, one of the maximal connected cyclotomic charged signed graphs from Theorem 3.3.4, seen in Figure 3.8:

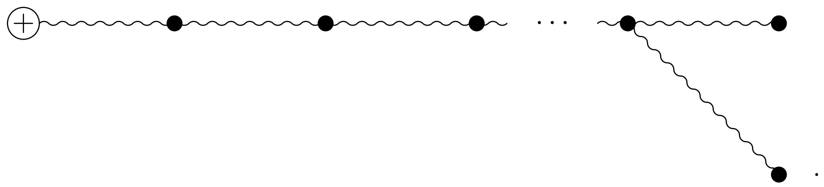


When the charged vertex of S is neutrally or positively charged, we see that it is an induced subgraph of C_{2k}^{+-}, and so by Theorem 3.3.4, S is cyclotomic. We see that the removal of all vertices on the lower layer, except the "last" uncharged vertex, as well as the negatively charged vertices and, in the case when S has only neutrally charged vertices, the upper positively charged vertex, gives us the induced subgraphs in these instances. This can be seen below:





Now consider the case where the charged vertex of S is negatively charged. In this case, S cannot be an induced subgraph of C_{2k}^{+-}: we would always be left with a negative edge regardless of which vertices we keep and which we remove. So in this case, we need to find an induced subgraph of C_{2k}^{+-} which is (switch) equivalent to the negative of this digraph. We see that the negative of S with a negatively charged vertex at the end is:



Now, by switching at the vertex which is joined to three vertices, and then switching in a similar fashion to that in the proof of Proposition 3.5.7, we arrive at the digraph S with a positively charged vertex at the end. We already know this is an induced subgraph of C_{2k}^{+-}.

So, up to equivalence, S is contained in a maximal connected cyclotomic charged signed graph, regardless of the charge of the vertex. Thus, S is also cyclotomic. □

The snake tongue S will be a useful addition to growing digraphs, as we will see in chapter 4.

Chapter 4

Growing Digraphs and More

In this chapter, we explore in greater detail fruitful ways to grow digraphs. In chapter 3, we saw two different ways to grow digraphs: attaching a path (with a possible non-neutral charge at the end) to a specified vertex, or by attaching smart choices of graph objects to a digraph. As noted, the latter of these choices is, on the surface, more complex. As such, we look in more detail at the former of these options, as well as related methods. We will see that there is much depth to this idea, and this will be of significant use to us when finding small Mahler measures from digraphs.

We have two key goals in this chapter. The first is to find explicit formulas related to certain grown digraphs. This will help us to find small Mahler measures in chapter 5, as well as being of use to our second goal, which is to investigate the general shape of the reciprocal polynomial of a digraph which has been grown using our outlined method.

Parts of Sections 4.1 and 4.3 have appeared in Coyston and McKee [9].

4.1 Formalising Definitions and Concepts

We begin by formalizing some concepts seen in chapter 3, and expanding upon these notions.

Definition 4.1.1. *A **pendant path** in a graph G is a sequence of neutral vertices $v_1, \cdots, v_t$ whereby:*

- *There are positive edges between v_i and v_{i+1} $(1 \leq i < t)$,*

- *If w is any vertex of G that is not one of the v_i, then there are no arcs in either direction between w and any of $v_1, \cdots, v_{t-1}$, although arcs between w and v_t are allowed.*

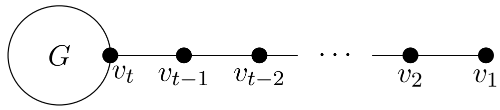


Figure 4.1: A visualisation of a graph G with a highlighted pendant path.

We note that pendant paths can exist in general graphs, but our focus will be for digraphs in particular.

We can extend this definition to include versions of charged paths:

Definition 4.1.2. *A **charged pendant path** in a digraph G is a sequence of vertices $v_1, \cdots, v_t$ whereby:*

- *The vertex v_1 has a non-neutral charge,*
- *There are positive edges between v_i and v_{i+1} $(1 \leq i < t)$,*
- *If w is any vertex of G that is not one of the v_i, then there are no arcs in either direction between w and any of $v_1, \cdots, v_{t-1}$, although arcs between w and v_t are allowed.*

*If the charge is positive, we refer to this as a **positively charged pendant path**, whilst if it is negative, we refer to it as a **negatively charged pendant path**.*

For convenience, from this point, we say *pendant path* to mean either a charged or uncharged pendant path. If we know exactly what type of pendant path we are working with, we will mention this explicitly. If we have an unspecified charge, say represented by C, at the end of our pendant path, we will refer to this as a C-charged pendant path.

Our definitions here make reference to pendant paths existing *within* a specified digraph. However, we can also use pendant paths to grow digraphs. In fact, we have already seen, in Lemma 3.3.2, an example of us attaching a C-charged pendant path to a specified digraph.

A further way of growing pendant paths is to attach more than one pendant path. Indeed, much like how McKee and Smyth [22] have a result for attaching an uncharged pendant path to a graph G, they also have a more general result for attaching multiple uncharged pendant paths:

Lemma 4.1.3. *Let G be a graph on $n \geq 1$ vertices with a list of k (not necessarily distinct) distinguished vertices $v_1, \cdots, v_k$. Let $G_{m_1,\cdots,m_k}$ be the graph obtained by attaching one endvertex of a new m_i-vertex uncharged pendant path to vertex v_i (so $G_{m_1,\cdots,m_k}$ has $m_1 + \cdots + m_k$ more vertices than G).*

Let $R_{m_1,\cdots,m_k}$ be the reciprocal polynomial of $G_{m_1,\cdots,m_k}$. Then, if $m_i \geq 2$ for each i, we have:

$$(z^2 - 1)^k R_{m_1,\cdots,m_k}(z) = \sum_{(\epsilon_1,\cdots,\epsilon_k)\in\{0,1\}} z^{2\sum \epsilon_i m_i} B_{\epsilon_1,\cdots,\epsilon_k}(z)\,, \tag{4.1}$$

for some monic integer polynomials $B_{\epsilon_1,\cdots,\epsilon_k}(z)$ that depends on G and $(v_1, \cdots, v_k)$, but not $m_1, \cdots, m_k$.

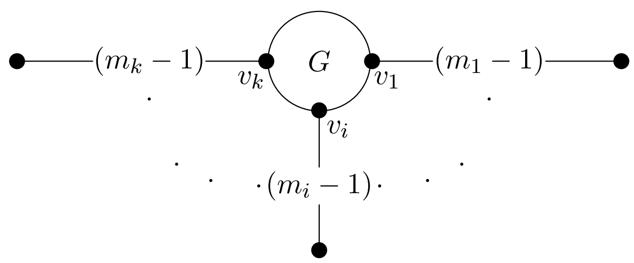


Figure 4.2: A visualisation of $G_{m_1,\cdots,m_k}$.

The concept of attaching pendant paths is a natural way to think of growing a graph; it is easy to visualise how the graph is growing. However, the heart of growing a graph is that we are adding vertices to the graph. As such, another way to achieve this is to *subdivide* an edge between two vertices:

Definition 4.1.4. *Let G be a digraph and v, w be distinct neutral vertices of G joined by a (signed) edge e. A **subdivision** (of size $t \geq 1$) of the edge e introduces t new neutral* *vertices between* v *and* w*, say* $u_1, \cdots, u_t$*, such that the edges* $(v, u_1), (u_t, w)$ *and* (u_i, u_{i+1}) *(for* $1 \leq i \leq t-1$*) are of the same sign as* e*.*

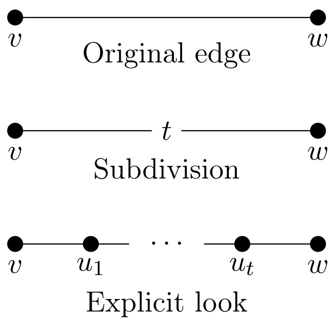


Figure 4.3: A subdivision of a positive edge.

Subdividing an edge can be thought of as growing an "internal path" within our digraph. Furthermore, we notice that if we subdivide an edge where one of the vertices is a *leaf* (a vertex that only has one adjacent vertex), this is effectively the same as attaching a neutral pendant path. So, in some sense, a neutral pendant path is a special case of a subdivided edge.

In the context of growing digraphs for finding small Mahler measures, pendant paths and subdivisions are of particular interest because they have trivial Mahler measure in their own right: they are just (possibly charged) paths. If we go back to our heuristic thought process from chapter 3, we have the hopeful idea that if we take a cyclotomic digraph and make a small change to it, we may make a small change to the Mahler measure as a result. In essence, we have a digraph that is "almost cyclotomic", and so we hope that the Mahler measure is "almost 1", or rather, small.

Attaching pendant paths and subdividing edges almost reverses this thought process. If we take a suitable digraph, and attach a large enough cyclotomic digraph, we are left with a new digraph which itself is almost cyclotomic. Again, this means that we can hope that the resulting Mahler measure is small. This, of course, has a certain level of dependence on the original digraph we choose. If we choose a digraph which has a large Mahler measure, and itself is very far away from being cyclotomic, even attaching large cyclotomic digraphs may not be enough to create a new digraph which is almost cyclotomic.

We will explore which types of digraphs are suitable for this idea later, in Section 5.1. Our focus for now remains on confirming that the idea of growing digraphs for finding small Mahler measures remains possible (i.e. if these digraphs have reciprocal polynomials which could have small Mahler measure) and feasible (i.e. if we can find ways to explicitly compute Mahler measures in a reasonable time).

4.2 Twice as Nice: Two Pendant Paths

We used Lemma 3.3.2 to find an explicit formula – namely (3.2) – for the reciprocal polynomial of a digraph which had a C-charged pendant path attached to a specified vertex. This was an extension of a result covered by McKee and Smyth [22].

Furthermore, Lemma 4.1.3 is also a result from McKee and Smyth. So, it is natural to ask if we can extend this result to also cover charged pendant paths. In short, the answer is yes, and it is not hard to conceive how. The problem that arises, however, is when we want to find *explicit* formulas in each case. For k pendant paths, we have 3^k possible combinations of pendant paths we could attach to specified vertices $v_1, \cdots, v_k$. As such, finding and stating all of these formulas becomes increasingly intricate and somewhat convoluted, albeit not extremely difficult, as k increases.

As such, we choose to find all explicit formulas for the situations where we add two pendant paths to a digraph. Not only does this mean our calculations will be manageable, we will see in chapter 5 that this is all that is necessary for the experiments we wish to run.

Before moving to the cases of charged pendant paths, let us re-state Lemma 4.1.3 for the case $k = 2$, with a slight extension and a refocused result:

Lemma 4.2.1. *Let G be a graph on $n \geq 1$ vertices and (v_1, v_2) be a list of two (not necessarily distinct) vertices of G. Let G_{m_1,m_2} be the graph obtained by attaching one endvertex of a new m_i-vertex pendant path to the vertex v_i, for $i = 1, 2$ (so, G_{m_1,m_2} will have $m_1 + m_2$ more vertices than G).*

Let R_{m_1,m_2} be the reciprocal polynomial of G_{m_1,m_2}. Then, if both $m_i \geq 0$, we have:

$$\begin{aligned}(z^2-1)^2 R_{m_1,m_2} &= (z^{2m_1}-1)(z^{2m_2}-1)R_{1,1} - (z^{2m_1}-1)(z^{2m_2}-z^2)R_{1,0} \\ &\quad - (z^{2m_1}-z^2)(z^{2m_2}-1)R_{0,1} + (z^{2m_1}-z^2)(z^{2m_2}-z^2)R_{0,0}\,.\end{aligned} \tag{4.2}$$

Firstly, we note the difference in the presentation of the formulas between (4.1) and (4.2). This is because the former equation is given in a more succinct format, and in the case $k = 2$, we have that:

$$\begin{aligned}B_{1,1}(z) &= R_{1,1}(z) - R_{1,0}(z) - R_{0,1}(z) + R_{0,0}(z)\,,\\ B_{1,0}(z) &= R_{1,1}(z) - z^2R_{1,0}(z) - R_{0,1}(z) + z^2R_{0,0}(z)\,,\\ B_{0,1}(z) &= R_{1,1}(z) - R_{1,0}(z) - z^2R_{0,1}(z) + z^2R_{0,0}(z)\,,\\ B_{0,0}(z) &= R_{1,1}(z) - z^2R_{1,0}(z) - z^2R_{0,1}(z) + z^4R_{0,0}(z)\,.\end{aligned} \tag{4.3}$$

Secondly, we note that we are extending this result for $m_i \geq 0$, whilst the version from McKee and Smyth is stated only for $m_i \geq 2$.

Proof of Lemma 4.2.1. As stated, this has already been shown when $m_i \geq 2$. In fact, the method presented by McKee and Smyth proves this for $m_i \geq 1$.

To show that (4.2) is valid for $m_i \geq 0$, we simply need to verify for two cases: when $m_1 = m_2 = 0$, and when $m_i > 0$ and $m_j = 0$, for $i, j \in \{1, 2\}, i \neq j$ (the case $m_i = 0$ and $m_j > 0$ would follow by a symmetric argument).

If we let $m_1 = m_2 = 0$ in (4.2), it is easy to see the result holds.

If we let $m_1 = 0$, and write $R_{i,j}$ for $R_{i,j}(z)$, then we get:

$$\begin{aligned}(z^2-1)^2 R_{0,m_2} &= (z^0-1)(z^{2m_2}-1)R_{1,1} - (z^0-1)(z^{2m_2}-z^2)R_{1,0}\\ &\quad - (z^0-z^2)(z^{2m_2}-1)R_{0,1} + (z^0-z^2)(z^{2m_2}-z^2)R_{0,0}\,,\\ &= -(1-z^2)(z^{2m_2}-1)R_{0,1} + (1-z^2)(z^{2m_2}-z^2)R_{0,0}\,.\end{aligned}$$

By dividing through by (z^2-1) and suitably expanding and rearranging, we get:

$$\begin{aligned}(z^2-1)R_{0,m_2} &= z^{2m_2}R_{0,1} - R_{0,1} - z^{2m_2}R_{0,0} + z^2R_{0,0} \\ &= z^{2m_2}(R_{0,1}-R_{0,0}) - (R_{0,1} - z^2R_{0,0})\,.\end{aligned}$$

We note that this situation is precisely the same as adding a single uncharged pendant path to some vertex. This is described in McKee and Smyth; we have already seen this in (3.3). So, following the above notation, the result can be written as:

$$(z^2-1)R_{0,m_2} = z^{2m_2}(R_{0,1}-R_{0,0}) - (R_{0,1}-R_{0,0})^*\,.$$

As such, to verify (4.2) is valid for when one $m_i = 0$, it is sufficient for us to verify that:

$$(R_{0,1} - z^2R_{0,0}) = (R_{0,1}-R_{0,0})^*\,. \tag{4.4}$$

For convenience, we will now let $R_{0,1} = R_1$ and $R_{0,0} = R_0$. We also note that, by definition, $\deg(R_1) = \deg(R_0) + 2$. This means that, for $a_i, b_i \in \mathbb{Z}$, we can write:

$$\begin{aligned} R_1(z) &= a_nz^n + a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \cdots + a_1z + a_0 \\ -R_0(z) &= b_{n-2}z^{n-2} + \cdots + b_1z + b_0\,, \\ \hline R(z) &= a_nz^n + a_{n-1}z^{n-1} + (a_{n-2}-b_{n-2})z^{n-2} + \cdots + (a_1-b_1)z + (a_0-b_0)\,, \\ \Rightarrow R^*(z) &= (a_0-b_0)z^n + (a_1+b_1)z^{n-1} + \cdots + (a_{n-2}-b_{n-2})z^2 + a_{n-1}z + a_n\,. \end{aligned}$$

This gives us the right hand side of (4.4) (where we have written $R(z) = R_1(z) - R_0(z)$ for convenience).

For the left hand side, we set $S(z) = (R_1 - z^2R_0)(z)$, giving us:

$$S(z) = (a_n - b_{n-2})z^n + (a_{n-1}-b_{n-3})z^{n-1} + \cdots + (a_2-b_0)z^2 + a_1z + a_0\,.$$

This gives us the left hand side of (4.4).

Since $R_1(z)$ and $R_0(z)$ are reciprocal, we know that $a_i = a_{n-i}$ and $b_j = b_{(n-2)-j}$. If we write $S(z) = \sum_{i=1}^{n} s_i z^i$ and $R^*(z) = \sum_{i=1}^{n} r_i^* z^i$, we get that:

$$\begin{aligned} s_i &= \begin{cases} a_i - b_{i-2}, & \text{for } i \geq 2\,, \\ a_i, & \text{for } i < 2\,. \end{cases} \\ &= \begin{cases} a_{n-i} - b_{n-i}, & \text{for } i \geq 2\,, \\ a_{n-i}, & \text{for } i < 2\,. \end{cases} \\ &= r_i^*\,, \end{aligned}$$

which verifies the equality of (4.4), and hence completes the proof. □

This will now serve as a strong foundation for us to find explicit formulas for the shape of the reciprocal polynomial of a digraph grown with combinations of (charged or uncharged) pendant paths. We will look at these in a case by case basis, as this helps add clarity to the results, before stating a final result in larger generality.

4.2.1 Charging One Pendant

We first look at the shape of the reciprocal polynomial of a digraph that has been grown by attaching a negatively charged pendant path, and an uncharged pendant path. We will show the proof and required manipulation of equations in great detail in this instance. What should be noted, however, is that the proof itself is not particularly difficult: it simply revolves around manipulating equations at the correct time. As such, we will note that the proof attaching one positively charged pendant path and one uncharged pendant path is effectively the same, and as such we will not look at that in as much detail.

Lemma 4.2.2. *Let G be a digraph on $n \geq 1$ vertices and (v_1, v_2) be a list of two (not necessarily distinct) vertices of G. Let $G^{\bullet-}_{m_1,m_2}$ be the digraph obtained by attaching one endvertex of a new neutrally charged pendant path with m_1 vertices to the vertex v_1, and attaching one endvertex of a new negatively charged pendant path with m_2 vertices to the vertex v_2 (so, $G^{\bullet-}_{m_1,m_2}$ will have $m_1 + m_2$ more vertices than G).*

Let $R^{\bullet-}_{m_1,m_2}$ be the reciprocal polynomial of $G^{\bullet-}_{m_1,m_2}$. Then, if both $m_i \geq 1$, we have:

$$(z-1)^2(z+1)R^{\bullet-}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon_2\in\{0,1\}} c_{\bullet-}(\epsilon_1,\epsilon_2) z^{2(\sum \epsilon_i m_i)-\epsilon_2} B_{\epsilon_1,\epsilon_2}(z)\,, \tag{4.5}$$

for some integer polynomials $B_{\epsilon_1,\epsilon_2}(z)$, as defined in (4.3), that depend on G and (v_1, v_2), but not on m_1, m_2, and where $c_{\bullet-}$ is the function:

$$c_{\bullet-}(\epsilon_1,\epsilon_2) = \begin{cases} 1\,, & \text{if } \epsilon_1 = \epsilon_2 = 1\,, \\ -1\,, & \text{if } \epsilon_1 = 1, \epsilon_2 = 0\,, \\ -1\,, & \text{if } \epsilon_1 = 0, \epsilon_2 = 1\,, \\ 1\,, & \text{if } \epsilon_1 = \epsilon_2 = 0\,. \end{cases}$$

Proof. Let $H = G^{\bullet}_{m_1} = G_{m_1}$; the digraph G with a neutrally charged pendant path on length m_1 attached to the specified vertex v_1. Then, we can think of the digraph $G^{\bullet-}_{m_1,m_2}$ as $H^{-}_{m_2}$; the digraph H with a negatively charged pendant path on length m_2 attached to the specified vertex v_2.

We can then apply (3.6), as found in Lemma 3.3.2, to give us:

$$R^{\bullet-}_{m_1,m_2}(z) = R_{m_1,m_2}(z) + zR_{m_1,m_2-1}(z)\,.$$

We can now multiply through by $(z^2-1)^2$ to get:

$$(z^2-1)^2 R^{\bullet-}_{m_1,m_2}(z) = (z^2-1)^2 R_{m_1,m_2}(z) + z(z^2-1)^2 R_{m_1,m_2-1}(z)\,,$$

and then note the we can use (4.2) to expand the right hand side.

Writing $R_{i,j}$ for $R_{i,j}(z)$ and $R^{\bullet-}_{i,j}$ for $R^{\bullet-}_{i,j}(z)$, this leads to:

$$\begin{aligned}(z^2-1)^2 R^{\bullet-}_{m_1,m_2} = \Big(&(z^{2m_1}-1)(z^{2m_2}-1)R_{1,1} - (z^{2m_1}-1)(z^{2m_2}-z^2)R_{1,0} \\ &- (z^{2m_1}-z^2)(z^{2m_2}-1)R_{0,1} + (z^{2m_1}-z^2)(z^{2m_2}-z^2)R_{0,0}\Big) \\ &+ z\Bigg((z^{2m_1}-1)(z^{2(m_2-1)}-1)R_{1,1} - (z^{2m_1}-1)(z^{2(m_2-1)}-z^2)R_{1,0} \\ &- (z^{2m_1}-z^2)(z^{2(m_2-1)}-1)R_{0,1} + (z^{2m_1}-z^2)(z^{2(m_2-1)}-z^2)R_{0,0}\Bigg)\,.\end{aligned}$$

Appropriate expansions then give us:

$$\begin{aligned}(z^2-1)^2 R^{\bullet -}_{m_1,m_2} &= (z^{2m_1}-1)(z^{2m_2}-1)R_{1,1} - (z^{2m_1}-1)(z^{2m_2}-z^2)R_{1,0}\\ &\quad - (z^{2m_1}-z^2)(z^{2m_2}-1)R_{0,1} + (z^{2m_1}-z^2)(z^{2m_2}-z^2)R_{0,0}\\ &\quad + (z^{2m_1}-1)(z^{2m_2-1}-z)R_{1,1} - (z^{2m_1}-1)(z^{2m_2-1}-z^3)R_{1,0}\\ &\quad - (z^{2m_1}-z^2)(z^{2m_2-1}-z)R_{0,1} + (z^{2m_1}-z^2)(z^{2m_2-1}-z^3)R_{0,0}\,.\end{aligned}$$

If we now collect the $R_{i,j}$ terms for $i,j \in \{0,1\}$, we now get:

$$\begin{aligned}(z^2-1)^2 R^{\bullet -}_{m_1,m_2} &= (z^{2m_1}-1)(z^{2m_2}+z^{2m_2-1}-z-1)R_{1,1}\\ &\quad - (z^{2m_1}-1)(z^{2m_2}+z^{2m_2-1}-z^3-z^2)R_{1,0}\\ &\quad - (z^{2m_1}-z^2)(z^{2m_2}+z^{2m_2-1}-z-1)R_{0,1}\\ &\quad + (z^{2m_1}-z^2)(z^{2m_2}+z^{2m_2-1}-z^3-z)R_{0,0}\,,\end{aligned}$$

$$\begin{aligned}&= (z^{2m_1}-1)(z^{2m_2-1}-1)(z+1)R_{1,1}\\ &\quad - (z^{2m_1}-1)(z^{2m_2-1}-z^2)(z+1)R_{1,0}\\ &\quad - (z^{2m_1}-z^2)(z^{2m_2-1}-1)(z+1)R_{0,1}\\ &\quad + (z^{2m_1}-z^2)(z^{2m_2-1}-z^2)(z+1)R_{0,0}\,.\end{aligned}$$

Next, we divide through by $(z+1)$ and expand the remaining terms on the left hand side:

$$\begin{aligned}(z-1)^2(z+1) R^{\bullet -}_{m_1,m_2} &= (z^{2m_1+2m_2-1}-z^{2m_1}-z^{2m_2-1}+1)R_{1,1}\\ &\quad - (z^{2m_1+2m_2-1}-z^{2m_1+2}-z^{2m_2-1}+z^2)R_{1,0}\\ &\quad - (z^{2m_1+2m_2-1}-z^{2m_1}-z^{2m_2+1}+z^2)R_{0,1}\\ &\quad + (z^{2m_1+2m_2-1}-z^{2m_1+2}-z^{2m_2+1}+z^4)R_{0,0}\,.\end{aligned}$$

Finally, using (4.3), we get:

$$(z-1)^2(z+1)R^{\bullet -}_{m_1,m_2}(z) = z^{2m_1+2m_2-1}B_{1,1}(z) - z^{2m_1}B_{1,0}(z) - z^{2m_2-1}B_{0,1}(z) + B_{0,0}(z)\,. \tag{4.6}$$

This is precisely (4.5), in its expanded form. □

As previously mentioned, finding this formula is not particularly difficult. The complications arise when trying to decide what exactly is the 'best' formula (that is, what is the most useful for our future experiments), and how this can best be presented. Once these are decided, the actual proof is relatively straightforward.

Moreover, once we have done this once, it is even more straightforward to find explicit formulas for all other combinations. Let us next look at a mirror of Lemma 4.2.2:

Lemma 4.2.3. *Let G be a digraph on $n \geq 1$ vertices and (v_1, v_2) be a list of two (not necessarily distinct) vertices of G. Let $G^{-\bullet}_{m_1,m_2}$ be the digraph obtained by attaching one endvertex of a new negatively charged pendant path with m_1 vertices to the vertex v_1, and attaching one endvertex of a new neutrally charged pendant path with m_2 vertices to the vertex v_2 (so, $G^{-\bullet}_{m_1,m_2}$ will have $m_1 + m_2$ more vertices than G).*

Let $R^{-\bullet}_{m_1,m_2}$ be the reciprocal polynomial of $G^{-\bullet}_{m_1,m_2}$. Then, if both $m_i \geq 1$, we have:

$$(z-1)^2(z+1)R^{-\bullet}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon_2\in\{0,1\}} c_{-\bullet}(\epsilon_1,\epsilon_2)z^{2(\sum \epsilon_i m_i)-\epsilon_1}B_{\epsilon_1,\epsilon_2}(z)\,, \tag{4.7}$$

for some integer polynomials $B_{\epsilon_1,\epsilon_2}(z)$, as defined in (4.3)*, that depend on G and (v_1, v_2), but not on m_1, m_2, and where $c_{-\bullet} = c_{\bullet-}$.*

Remark. We note that (4.7) is *almost* the same formula as (4.5); the sole exception is in the former, we have a "$-\epsilon_1$" term in the exponent of z, as opposed to a "$-\epsilon_2$" term.

The proof, again, is almost the same as the proof of Lemma 4.2.2. We highlight the main difference, which occurs at the start, and then briefly state the effect that this difference results in.

Proof. Let $H = G^{\bullet}_{m_2} = G_{m_2}$; the digraph G with a neutrally charged pendant path of length m_2 attached to the specified vertex v_2. Then, we can think of the digraph $G^{-\bullet}_{m_1,m_2}$ as $H^{-}_{m_1}$; the digraph H with a negatively charged pendant path of length m_1 attached to the specified vertex v_1.

We can then apply (3.6), as found in Lemma 3.3.2, to give us:

$$R^{\bullet-}_{m_1,m_2}(z) = R_{m_1,m_2}(z) + zR_{m_1-1,m_2}(z)\,.$$

We see that the second term on the left hand side here is slightly different. This creates changes to the values obtained throughout the calculations and manipulations of the proof (compared to that of Lemma 4.2.2), but does not affect the method. This method will give us the final result, in a form similar to that of (4.6). □

We now state the corresponding results for the addition of one positively charged pendant path, and one neutral pendant path.

Lemma 4.2.4. *Let G be a digraph on $n \geq 1$ vertices and (v_1, v_2) be a list of two (not necessarily distinct) vertices of G. Let $G^{+\bullet}_{m_1,m_2}$ be the digraph obtained by attaching one endvertex of a new positively charged pendant path to the vertex v_1 with m_1 vertices, and attaching one endvertex of a new neutrally charged pendant path with m_2 vertices to the vertex v_2 (so, $G^{+\bullet}_{m_1,m_2}$ will have $m_1 + m_2$ more vertices than G).*

Let $R^{+\bullet}_{m_1,m_2}$ be the reciprocal polynomial of $G^{+\bullet}_{m_1,m_2}$. Then, if both $m_i \geq 1$, we have:

$$(z+1)^2(z-1)R^{+\bullet}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon_2\in\{0,1\}} c_{+\bullet}(\epsilon_1,\epsilon_2) z^{2(\sum \epsilon_i m_i)-\epsilon_1} B_{\epsilon_1,\epsilon_2}(z)\,,$$

for some integer polynomials $B_{\epsilon_1,\epsilon_2}(z)$, as defined in (4.3), that depend on G and (v_1, v_2), but not on m_1, m_2. Here,

$$c_{+\bullet}(\epsilon_1,\epsilon_2) = \begin{cases} 1\,, & \text{if } \epsilon_1 = \epsilon_2 = 1\,,\\ -1\,, & \text{if } \epsilon_1 = 1, \epsilon_2 = 0\,,\\ 1\,, & \text{if } \epsilon_1 = 0, \epsilon_2 = 1\,,\\ -1\,, & \text{if } \epsilon_1 = \epsilon_2 = 0\,. \end{cases}$$

Lemma 4.2.5. *Let G be a digraph on $n \geq 1$ vertices and (v_1, v_2) be a list of two (not necessarily distinct) vertices of G. Let $G^{\bullet+}_{m_1,m_2}$ be the digraph obtained by attaching one endvertex of a new neutrally charged pendant path with m_1 vertices to the vertex v_1, and attaching one endvertex of a new positively charged pendant path with m_2 vertices to the* *vertex v_2 (so, $G^{\bullet+}_{m_1,m_2}$ will have m_1+m_2 more vertices than G).*

Let $R^{\bullet+}_{m_1,m_2}$ be the reciprocal polynomial of $G^{\bullet+}_{m_1,m_2}$. Then, if both $m_i \geq 1$, we have:

$$(z+1)^2(z-1)R^{\bullet+}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon_2\in\{0,1\}} c_{\bullet+}(\epsilon_1,\epsilon_2) z^{2(\sum \epsilon_i m_i)-\epsilon_2} B_{\epsilon_1,\epsilon_2}(z)\,,$$

for some integer polynomials $B_{\epsilon_1,\epsilon_2}(z)$, as defined in (4.3), that depend on G and (v_1,v_2), but not on m_1, m_2, and where $c_{\bullet+}$ is the function:

$$c_{\bullet+}(\epsilon_1,\epsilon_2) = \begin{cases} 1\,, & \text{if } \epsilon_1=\epsilon_2=1\,,\\ 1\,, & \text{if } \epsilon_1=1, \epsilon_2=0\,,\\ -1\,, & \text{if } \epsilon_1=0, \epsilon_2=1\,,\\ -1\,, & \text{if } \epsilon_1=\epsilon_2=0\,.\end{cases}$$

We do not give details of the proofs for these Lemmas, beyond noting that it is effectively the same as the proof of Lemma 4.2.2 and stating the following, which follow after applying (3.6) from the proof of Lemma 3.3.2:

$$R^{+\bullet}_{m_1,m_2}(z) = R_{m_1,m_2}(z) - zR_{m_1-1,m_2}(z)\,,$$
$$R^{\bullet+}_{m_1,m_2}(z) = R_{m_1,m_2}(z) - zR_{m_1,m_2-1}(z)\,.$$

4.2.2 Charging Two Pendants

We now look at the shape of the reciprocal polynomial of a digraph that has been grown by attaching two charged pendant paths. We first look at the situation where we attach two pendant paths of the same charge (either both positive or both negative), and then we move to the case where we attach opposing charged pendant paths.

It should come as no surprise that the statements, results and proofs will all be extremely similar to those that we have already seen.

Lemma 4.2.6. *Let G be a digraph on $n \geq 1$ vertices and (v_1, v_2) be a list of two (not necessarily distinct) vertices of G. Let $G^{--}_{m_1,m_2}$ be the digraph obtained by attaching one endvertex of a new m_i-vertex negatively charged pendant path to the vertex v_i, for $i = 1, 2$ (so, $G^{--}_{m_1,m_2}$ will have m_1+m_2 more vertices than G).*

Let $R^{--}_{m_1,m_2}$ be the reciprocal polynomial of $G^{--}_{m_1,m_2}$. Then, if both $m_i \geq 1$, we have:

$$(z-1)^2 R^{--}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon_2\in\{0,1\}} (-1)^{\sum \epsilon_i} z^{2(\sum \epsilon_i m_i) - \sum \epsilon_i} B_{\epsilon_1,\epsilon_2}(z)\,, \tag{4.8}$$

for integer polynomials $B_{\epsilon_1,\epsilon_2}(z)$, as defined in (4.3)*, that depend on G and (v_1, v_2), but not on m_1, m_2.*

Proof. Let $H = G^-_{m_1}$; the digraph G with a negatively charged pendant path of length m_1 attached to the specified vertex v_1. Then, we can think of the digraph $G^{--}_{m_1,m_2}$ as $H^-_{m_2}$; the digraph H with a negatively charged pendant path with m_2 vertices attached to the specified vertex v_2.

Writing $R^{**}_{m_1,m_2}$ for $R^{**}_{m_1,m_2}(z)$, we can then apply (3.6) from Lemma 3.3.2 twice, to give us:

$$\begin{aligned} R^{--}_{m_1,m_2} &= R^{-\bullet}_{m_1,m_2} + zR^{-\bullet}_{m_1,m_2-1} \\ &= (R_{m_1,m_2} + zR_{m_1-1,m_2}) + z(R_{m_1,m_2-1} + zR_{m_1-1,m_2-1})\,. \end{aligned}$$

Whilst further manipulation returns:

$$R^{--}_{m_1,m_2} = R_{m_1,m_2} + z(R_{m_1-1,m_2} + R_{m_1,m_2-1}) + z^2 R_{m_1-1,m_2-1}\,.$$

We then proceed in the same fashion as shown in the proof of Lemma 4.2.2 to arrive at the result. □

It is worth noting that because we have to apply (3.6), as found in Lemma 3.3.2, twice, we have to work with more terms when simplifying our formula. This makes the calculations and required manipulations slightly longer, but the heart of the proof remains the same.

We now state the equivalent Lemma for attaching two positively charged pendant paths. We shall omit any details of the proof, owing to the fact that all details should be clear at this point.

Lemma 4.2.7. *Let G be a digraph on $n \geq 1$ vertices and (v_1, v_2) be a list of two (not necessarily distinct) vertices of G. Let $G^{++}_{m_1,m_2}$ be the digraph obtained by attaching one* *endvertex of a new positively charged pendant path with m_i vertices to the vertex v_i, for $i = 1, 2$ (so, $G^{++}_{m_1,m_2}$ will have $m_1 + m_2$ more vertices than G).*

Let $R^{++}_{m_1,m_2}$ be the reciprocal polynomial of $G^{++}_{m_1,m_2}$. Then, if both $m_i \geq 1$, we have:

$$(z+1)^2 R^{++}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon_2 \in \{0,1\}} z^{2(\sum \epsilon_i m_i) - \sum \epsilon_i} B_{\epsilon_1,\epsilon_2}(z)\,, \tag{4.9}$$

for integer polynomials $B_{\epsilon_1,\epsilon_2}(z)$, as defined in (4.3)*, that depend on G and (v_1, v_2), but not on m_1, m_2.*

It again should not be too much of a surprise that the formulas in these cases are very similar to each other, as well as being incredibly similar to those that we have seen already.

Finally, let us turn to the cases where we attach two pendant paths with different charges. Again, we do not provide details of the corresponding proofs.

Lemma 4.2.8. *Let G be a digraph on $n \geq 1$ vertices and (v_1, v_2) be a list of two (not necessarily distinct) vertices of G. Let $G^{+-}_{m_1,m_2}$ be the digraph obtained by attaching one endvertex of a new positively charged pendant path with m_1 vertices to the vertex v_1, and attaching one endvertex of a new negatively charged pendant path with m_2 vertices to the vertex v_2 (so, $G^{+-}_{m_1,m_2}$ will have $m_1 + m_2$ more vertices than G).*

Let $R^{+-}_{m_1,m_2}$ be the reciprocal polynomial of $G^{+-}_{m_1,m_2}$. Then, if both $m_i \geq 1$, we have:

$$(z^2-1) R^{+-}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon_2 \in \{0,1\}} c_{+-}(\epsilon_1,\epsilon_2) z^{2(\sum \epsilon_i m_i) - \sum \epsilon_i} B_{\epsilon_1,\epsilon_2}(z)\,, \tag{4.10}$$

for integer polynomials $B_{\epsilon_1,\epsilon_2}(z)$, as defined in (4.3)*, that depend on G and (v_1, v_2), but not on m_1, m_2, and where c_{+-} is the function:*

$$c_{+-}(\epsilon_1, \epsilon_2) = \begin{cases} 1\,, & \textit{if } \epsilon_1 = \epsilon_2 = 1\,, \\ -1\,, & \textit{if } \epsilon_1 = 1, \epsilon_2 = 0\,, \\ 1\,, & \textit{if } \epsilon_1 = 0, \epsilon_2 = 1\,, \\ -1\,, & \textit{if } \epsilon_1 = \epsilon_2 = 0\,. \end{cases}$$

Lemma 4.2.9. *Let G be a digraph on $n \geq 1$ vertices and (v_1, v_2) be a list of two (not necessarily distinct) vertices of G. Let $G^{-+}_{m_1,m_2}$ be the digraph obtained by attaching one endvertex of a new negatively charged pendant path with m_1 to the vertex v_1, and attaching one endvertex of a new positively charged pendant path with m_2 vertices to the vertex v_2 (so, $G^{-+}_{m_1,m_2}$ will have $m_1 + m_2$ more vertices than G).*

Let $R^{-+}_{m_1,m_2}$ be the reciprocal polynomial of $G^{-+}_{m_1,m_2}$. Then, if both $m_i \geq 1$, we have:

$$(z^2 - 1)R^{-+}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon_2 \in \{0,1\}} c_{-+}(\epsilon_1, \epsilon_2) z^{2(\sum \epsilon_i m_i) - \sum \epsilon_i} B_{\epsilon_1,\epsilon_2}(z)\,, \qquad (4.11)$$

for integer polynomials $B_{\epsilon_1,\epsilon_2}(z)$, as defined in (4.3)*, that depend on G and (v_1, v_2), but not on m_1, m_2, and where c_{-+} is the function:*

$$c_{-+}(\epsilon_1, \epsilon_2) = \begin{cases} 1\,, & \textit{if } \epsilon_1 = \epsilon_2 = 1\,, \\ 1\,, & \textit{if } \epsilon_1 = 1, \epsilon_2 = 0\,, \\ -1\,, & \textit{if } \epsilon_1 = 0, \epsilon_2 = 1\,, \\ -1\,, & \textit{if } \epsilon_1 = \epsilon_2 = 0\,. \end{cases}$$

Again, we see that these formulas are also very similar.

4.2.3 Overview of Charged Pendants

We can encapsulate all of these results in one Proposition:

Proposition 4.2.10. *Let G be a digraph on $n \geq 1$ vertices and (v_1, v_2) be a list of two (not necessarily distinct) vertices of G. Let $G^{**}_{m_1,m_2}$ be the digraph obtained by attaching one endvertex of a new pendant path with m_1 to the vertex v_1, and attaching one endvertex of a new pendant path with m_2 vertices to the vertex v_2 (so, $G^{**}_{m_1,m_2}$ will* *have $m_1 + m_2$ more vertices than G).*

Let $R^{**}_{m_1,m_2}$ be the reciprocal polynomial of $G^{**}_{m_1,m_2}$. Then, if both $m_i \geq 1$, we have:

$$k(z)R^{**}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon_2\in\{0,1\}} c_{**}(\epsilon_1,\epsilon_2) z^{2(\sum \epsilon_i m_i)-d_{**}(\epsilon_1,\epsilon_2)} B_{\epsilon_1,\epsilon_2}(z)\,, \tag{4.12}$$

where:

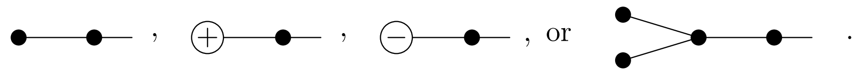
- *the $B_{\epsilon_1,\epsilon_2}(z)$ are integer polynomials, as defined in (4.3), that depend on G and (v_1, v_2), but not on m_1, m_2,*
- *the $c_{**}(\epsilon_1, \epsilon_2)$ take values ± 1 and depend only on the charges of the pendant paths,*
- *the $d_{**}(\epsilon_1, \epsilon_2)$ take values in $\{0, 1, 2\}$ and depend only on the charges of the pendant paths,*
- *$k(z) \neq 1$ divides $\left(z^2 - 1\right)^2$.*

Proposition 4.2.10 does give us an insight of the shape of the reciprocal polynomial. It also further helps us in calculating the Mahler measure of a digraph with two pendant paths attached to it. However, seeing each situation individually and more explicitly, as shown through Lemmas 4.2.1–4.2.9, is arguably the clearest way to see exactly what is happening in each individual scenario, even if it is perhaps a verbose way of doing so.

4.2.4 Attaching Snake Tongues – A Slithering Extra

So far, we have looked at attaching pendant paths to a digraph where the final vertex has some possible charge. Our reason for attaching these sorts of paths is because they are cyclotomic. However, we also know from Proposition 3.5.17 that the *snake tongue* (the digraph S seen in Figure 3.10) is cyclotomic. As such, it is worth asking what happens if we attach these to a digraph, as well as our charged pendant paths.

In practice, these snake tongue pendant paths are very similar to our charged pendant paths. We simply add a "fork" of two vertices to the end of the pendant path, rather than a charged vertex. So, by considering these as well, it means that there are four possible ways that our pendant paths can stop:



For avoidance of doubt, if we attach a snake tongue to a digraph, we are attaching an uncharged pendant path of, say, m vertices, and then attaching the two additional vertices (which make up the fork of the snake tongue) to the end of the pendant path. This means if we attach a snake tongue to a digraph, we are attaching a total of $m+2$ vertices to our digraph. This is slightly different to attaching charged pendant paths, where we are attaching a total of m vertices. Figure 4.4 may help make this slightly clearer.

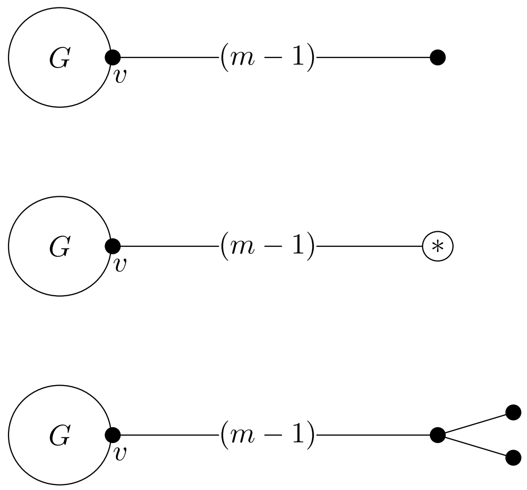


Figure 4.4: Visualisations of attaching pendant paths and snake tongues to a digraph.

Of course, there is nothing stopping us from changing this. We could, for example, fix it so that the snake tongue adds a total of m vertices (by having a path of $m-2$ neutral vertices, then attaching the fork to the leaf of that path).

However, we have chosen this convention for two reasons. Firstly, it streamlines our future calculations. Secondly, it fixes a convention that with our attached path, the final vertex – and only the final vertex – forms part of a cyclotomic digraph attached to the end of our path. This is a convention we have implicitly, and up until now, unnecessarily, used (since a single vertex, regardless of charge, is a cyclotomic digraph). We see why we have adopted this convention in Section 4.4.

Our aim here is as it was with attaching charged pendant paths: we want to find the shape of the reciprocal polynomial after attaching these pendant paths. The ideas for each individual case are effectively the same as those seen in Lemmas 4.2.1–4.2.9, with the obvious change that we are working with forks rather than charges. As such, we give an analogue of Proposition 4.2.10, featuring snake tongues:

Proposition 4.2.11. *Let G be a digraph on $n \geq 1$ vertices and (v_1, v_2) be a list of two (not necessarily distinct) vertices of G. Let $G^{\dagger\dagger}_{m_1,m_2}$ be the digraph obtained by attaching one endvertex of a new pendant path with m_1 vertices to the vertex v_1, and attaching one endvertex of a new pendant path with m_2 vertices to the vertex v_2, such that at least one of the pendant paths has a fork attached, and so is in fact a snake tongue (so, $G^{\dagger\dagger}_{m_1,m_2}$ will have $m_1 + m_2 + 2s$ more vertices than G, where s is the number of forks attached).*

Let $R^{\dagger\dagger}_{m_1,m_2}$ be the reciprocal polynomial of $G^{\dagger\dagger}_{m_1,m_2}$. Then, if both $m_i \geq 1$,we have:

$$\kappa(z)R^{\dagger\dagger}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon_2\in\{0,1\}} \gamma_{\dagger\dagger}(\epsilon_1, \epsilon_2) z^{2(\sum \epsilon_i m_i)-\delta_{\dagger\dagger}(\epsilon_1,\epsilon_2)} B_{\epsilon_1,\epsilon_2}(z)\,, \tag{4.13}$$

where:

- *the $B_{\epsilon_1,\epsilon_2}(z)$ are integer polynomials, as defined in (4.3), that depend on G and (v_1, v_2), but not on m_1, m_2,*
- *the $\gamma_{\dagger\dagger}(\epsilon_1, \epsilon_2)$ take values ± 1 and depend only on the ends of the pendant paths,*
- *the $\delta_{\dagger\dagger}(\epsilon_1, \epsilon_2)$ take values in $\{0, 1\}$ and depend only on the ends of the pendant paths,*
- *$\kappa(z)$ is a Kronecker-cyclotomic polynomial .*

We omit details of the full proof here, as we have acknowledged the fact that the method and thought process behind it remains the same as what we have seen with Proposition 4.2.10. However, it is worth seeing what the reciprocal polynomial of a digraph is when we attach a solitary snake tongue, and no other pendants. This helps create an analogy with Lemma 3.3.2, and gives us a parallel for (3.6); important features needed to understand how (4.13) can be derived.

Lemma 4.2.12. *Let G be a digraph on $n \geq 1$ vertices with a distinguished vertex v. For each $m \geq 0$, let $G_m^<$ be the graph obtained by attaching the end of a snake tongue, with a total of $m+2$ vertices, to the vertex v (so $G_m^<$ has $m+2$ more vertices than G).*

Let $R_m^<(z)$ be the reciprocal polynomial of $G_m^<$. Then, for $m \geq 2$, we have:

$$\frac{1}{z^2+1} R_m^<(z) = z^{2m} B(z) + B^*(z)\,, \tag{4.14}$$

for some monic integer polynomial $B(z)$ that depends on G and v, but not m.

Proof. We can label the digraph $G_m^<$ appropriately so that the adjacency matrix is:

$$A = \left(\begin{array}{cccccccc|c}
0 & 0 & 1 & & & & & & \\
0 & 0 & 1 & & & & & & \\
1 & 1 & 0 & 1 & & & & & \\
 & & 1 & 0 & \ddots & & & & \\
 & & & \ddots & \ddots & 1 & & & \\
 & & & & 1 & 0 & 1 & & \\
 & & & & & 1 & 0 & & 1 \\
\hline
 & & & & & & 1 & & \\
 & & & & & & & & A_G
\end{array}\right).$$

Labelling the characteristic polynomial of $G_m^<$ as $\chi_m^<$, we find a relation similar to (3.4) and (3.5) by expanding $\det(xI - A)$ along the first three rows appropriately:

$$\chi_m^<(x) = x\chi_{m+1}(x) - x\chi_{m-1}(x)\,. \tag{4.15}$$

We now take the reciprocalisation of this, which after rearranging appropriately gives us:

$$R_m^<(z) = (z^2+1)\left(R_{m+1}(z) - z^2 R_{m-1}(z)\right)\,. \tag{4.16}$$

Finally, we use (3.3) to arrive at our result. □

For additional exposure around Proposition 4.2.11, we include the explicit formulas for each case covered therein. We use a clear form of the notation just introduced: when a pendant path has a fork at the end of it, we use the symbols $>$ and $<$ to represent this (with the former representing a fork attached to the end of the "left" pendant path, with m_1 vertices, and the latter to the "right" pendant path, with m_2 vertices). Each of these formulas that follow is, of course, only valid under the conditions of Proposition 4.2.11.

Two forks:

$$R^{><}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon\in\{0,1\}} z^{2(\sum \epsilon_i m_i)} B_{\epsilon_1,\epsilon_2}(z)\,. \tag{4.17}$$

One fork, one neutral vertex:

Here, we have two possible ways to attach these 'pendant paths', giving us two different results. The first is:

$$(z^2-1)R^{>\bullet}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon\in\{0,1\}} c_{>\bullet}(\epsilon_1,\epsilon_2) z^{2(\sum \epsilon_i m_i)} B_{\epsilon_1,\epsilon_2}(z)\,, \tag{4.18}$$

where:

$$c_{>\bullet}(\epsilon_1,\epsilon_2) = \begin{cases} 1\,, & \text{if } \epsilon_1=\epsilon_2=1\,, \\ -1\,, & \text{if } \epsilon_1=1, \epsilon_2=0\,, \\ 1\,, & \text{if } \epsilon_1=0, \epsilon_2=1\,, \\ -1\,, & \text{if } \epsilon_1=\epsilon_2=0\,. \end{cases}$$

The second is:

$$(z^2-1)R^{\bullet<}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon\in\{0,1\}} c_{\bullet<}(\epsilon_1,\epsilon_2) z^{2(\sum \epsilon_i m_i)} B_{\epsilon_1,\epsilon_2}(z)\,, \tag{4.19}$$

where:

$$c_{\bullet<}(\epsilon_1,\epsilon_2) = \begin{cases} 1\,, & \text{if } \epsilon_1=\epsilon_2=1\,, \\ 1\,, & \text{if } \epsilon_1=1, \epsilon_2=0\,, \\ -1\,, & \text{if } \epsilon_1=0, \epsilon_2=1\,, \\ -1\,, & \text{if } \epsilon_1=\epsilon_2=0\,. \end{cases}$$

One fork, one positive vertex:

Here, we again have two possible ways to attach these 'pendant paths', giving us two different results. The first is:

$$(z+1)R^{>+}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon\in\{0,1\}} z^{2(\sum \epsilon_i m_i)-\epsilon_2} B_{\epsilon_1,\epsilon_2}(z)\,. \tag{4.20}$$

The second is:

$$(z+1)R^{+<}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon\in\{0,1\}} z^{2(\sum \epsilon_i m_i)-\epsilon_1} B_{\epsilon_1,\epsilon_2}(z)\,, \tag{4.21}$$

One fork, one negative vertex:

Here, we once again have two possible ways to attach these 'pendant paths', giving us two different results. The first is:

$$(z-1)R^{>-}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon\in\{0,1\}} c_{>-}(\epsilon_1,\epsilon_2) z^{2(\sum \epsilon_i m_i)-\epsilon_2} B_{\epsilon_1,\epsilon_2}(z)\,, \tag{4.22}$$

where:

$$c_{>-}(\epsilon_1,\epsilon_2) = \begin{cases} 1\,, & \text{if } \epsilon_1 = \epsilon_2 = 1\,, \\ -1\,, & \text{if } \epsilon_1 = 1, \epsilon_2 = 0\,, \\ 1\,, & \text{if } \epsilon_1 = 0, \epsilon_2 = 1\,, \\ -1\,, & \text{if } \epsilon_1 = \epsilon_2 = 0\,. \end{cases}$$

The second is:

$$(z-1)R^{-<}_{m_1,m_2}(z) = \sum_{\epsilon_1,\epsilon\in\{0,1\}} c_{-<}(\epsilon_1,\epsilon_2) z^{2(\sum \epsilon_i m_i)-\epsilon_1} B_{\epsilon_1,\epsilon_2}(z)\,, \tag{4.23}$$

where:

$$c_{-<}(\epsilon_1,\epsilon_2) = \begin{cases} 1\,, & \text{if } \epsilon_1 = \epsilon_2 = 1\,, \\ 1\,, & \text{if } \epsilon_1 = 1, \epsilon_2 = 0\,, \\ -1\,, & \text{if } \epsilon_1 = 0, \epsilon_2 = 1\,, \\ -1\,, & \text{if } \epsilon_1 = \epsilon_2 = 0\,. \end{cases}$$

4.2.5 A Final Result

Between the results encapsulated in Propositions 4.2.10 and 4.2.11, we have sixteen ways to attach two pendant paths (including pendant paths with forks attached to them) to a given digraph. In each case, the pendant paths we add are cyclotomic. Therefore, these results give us a way of taking a digraph – which we can choose to not be cyclotomic – and, in some combinatorial sense, make it close to being a cyclotomic digraph.

In chapter 5, we show how we implement these results in practice for finding digraphs with small Mahler measures. Before we can use these results, however, it would be useful to simplify them as much as possible. If nothing else, this just makes the corresponding formulas more user-friendly to input when running experiments.

What is striking about the results from Propositions 4.2.10 and 4.2.11 is they can be written in the following general form:

$$\kappa_1(z)R(z) = \kappa_2(z)L(z)\,,$$

where κ_1 and κ_2 are Kronecker-cyclotomic polynomials, R is the reciprocal of the resulting digraph and L is linear combination of some monic integer polynomials. Since the Mahler measure is multiplicative, we then have that $M(R) = M(L)$. Therefore, it is worthwhile trying to see if we can make the function $L(z)$ as simple as possible.

Indeed, it is possible to make such a simplification. Before stating the result, we introduce a quick definition for convenience:

Definition 4.2.13. *A **decorated pendant path** is a either a charged pendant path, or a pendant path which ends in a fork.*

Theorem 4.2.14. *Let G be a digraph on $n \geq 1$ vertices and (v_1, v_2) be a list of two (not necessarily distinct) vertices of G. Let $G^{\ddagger\ddagger}_{m_1,m_2}$ be the digraph obtained by attaching one endvertex of a new decorated pendant path with m_1 to the vertex v_1, and attaching one endvertex of a new decorated pendant path with m_2 vertices to the vertex v_2 (so, $G^{\ddagger\ddagger}_{m_1,m_2}$ will have $m_1 + m_2 + 2s$ more vertices than G, where s is the number of pendant paths decorated with forks at the end).*

Let $R^{\ddagger\ddagger}_{m_1,m_2}$ be the reciprocal polynomial of $G^{\ddagger\ddagger}_{m_1,m_2}$. Then, if both $m_i \geq 1$,we have:

$$\frac{\kappa_1(z)R^{\ddagger\ddagger}_{m_1m_2}(z)}{\kappa_2(z)} = z^{2(m_1+m_2)}B_{1,1}(z) + c_{10}(z)z^{2m_1}B_{1,0}(z) + c_{01}(z)z^{2m_2}B_{0,1}(z) + c_{00}(z)B_{0,0}(z)\,, \tag{4.24}$$

where:

$$c_{10}(z),\ c_{01}(z) \in \{1, -1, z, -z\}\,, \tag{4.25}$$

$$c_{00}(z) = c_{10}(z)c_{01}(z)\,, \tag{4.26}$$

and $\kappa_1(z)$ *and* $\kappa_2(z)$ *are Kronecker-cyclotomic.*

We again omit a proof here, but note that this result can be derived from Propositions 4.2.10 and 4.2.11 and Lemmas 4.2.1–4.2.9, albeit with slightly different manipulation of the equations therein.

It is worth noting, however, that Theorem 4.2.14 gives the right hand side of (4.24) in a slightly different form to what we have seen so far. In particular, we have that the powers of z which accompany $B_{1,0}$ and $B_{0,1}$ are always at least $2m_1$ and $2m_2$ respectively, whereas in our previous formulations there were cases where they could have been $2m_i - k$, for some k at most 2. As such, the Kronecker-cyclotomic factor on the left hand side of (4.24) is usually slightly different to the previous equations as well. This also explains why we have to divide by another Kronecker-cyclotomic factor, κ_2, as well.

Each decoration has a corresponding c_{10} and c_{01}, which give a result for c_{00}. There are precisely sixteen pairwise combinations for the values c_{10} and c_{01}, so this is a complete correspondence. These are listed in Table 4.1, alongside the corresponding κ_1 and κ_2 factors.

We note that Table 4.1 shows us that the κ_1 in particular are in fact very similar to the Kronecker-cyclotomic terms we had previously seen. Indeed, we take the original factor for the corresponding decoration, and then multiply by z^γ, where γ is the number of non-neutrally charged pendant paths attached to the original digraph. The κ_2 factors are in fact $(z^2+1)^s$, where s is the number of snake tongues attached to the original digraph.

D_L	D_R	$c_{10}(z)$	$c_{01}(z)$	κ_1	κ_2
$\bullet$	$\bullet$	-1	-1	$(z^2-1)^2$	1
$\bullet$	$+$	z	-1	$(z+1)^2(z-1)z$	1
$\bullet$	$-$	$-z$	-1	$(z-1)^2(z+1)z$	1
$\bullet$	$<$	1	-1	(z^2-1)	(z^2+1)
$+$	$\bullet$	-1	z	$(z+1)^2(z-1)z$	1
$+$	$+$	z	z	$(z+1)^2z^2$	1
$+$	$-$	$-z$	z	$(z+1)(z-1)z^2$	1
$+$	$<$	1	z	$z(z+1)$	(z^2+1)
$-$	$\bullet$	-1	$-z$	$(z-1)^2(z+1)z$	1
$-$	$+$	z	$-z$	$(z+1)(z-1)z^2$	1
$-$	$-$	$-z$	$-z$	$(z-1)^2z^2$	1
$-$	$<$	1	$-z$	$z(z-1)$	(z^2+1)
$>$	$\bullet$	-1	1	(z^2-1)	(z^2+1)
$>$	$+$	z	1	$z(z+1)$	(z^2+1)
$>$	$-$	$-z$	1	(z^2-1)	(z^2+1)
$>$	$<$	1	1	1	$(z^2+1)^2$

Table 4.1: The possible decorations D_L and D_R and the corresponding values of c_{10}, c_{01}, κ_1 and κ_2, as seen in (4.24).

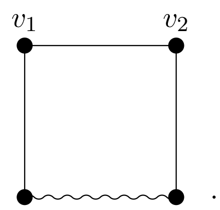
This result gives us the sixteen different formulas that arise from the sixteen pairs of decorated pendant paths we can attach to a given digraph. Since we are interested in finding the Mahler measure of $R^{\ddagger\ddagger}_{m_1m_2}$, we can omit $\kappa_1(z)$ and $\kappa_2(z)$ from any calculations (because they are Kronecker-cyclotomic and so have Mahler measure 1). This gives us much simpler formulas to work, which primarily rely on the monic integer polynomial $B_{i,j}$, which themselves rely on our original digraph G, and choice of vertices to grow from.

4.2.6 Some Straightforward Examples

As the concepts and results we have introduced here are particularly important, we give some straightforward examples to demonstrate all the results we have seen. This will be inclusive of everything we have seen so far in this Section.

We note, however, that because of their simplicity, the aim here is not to showcase digraphs with small Mahler measure, nor to exhibit that our results are better than a "brute force" method of calculation. Instead, the aim here is to give a further sense of validity to the results we have seen so far by seeing examples in which they hold.

Example 4.2.15. Let G be the following digraph, with vertices v_1 and v_2 labelled:



We intend to attach two decorated pendant paths to G; one at v_1 and one at v_2. In this case, we will fix pendant paths of certain lengths, and then see all sixteen reciprocal polynomials of the resulting digraphs arising from all possible combinations of pendant paths.

Before even fixing the lengths of our pendant paths, though, we can make some useful preliminary calculations. As seen in Theorem 4.2.14, specifically (4.24), the monic polynomials $B_{1,1}(z), B_{1,0}(z), B_{0,1}(z)$ and $B_{0,0}(z)$ are always present when we calculate the reciprocal polynomial of our new digraph, regardless of the length of our pendant paths, or their decorations. These $B_{i,j}$ are exactly as defined in (4.3), which themselves depend on the polynomials $R_{i,j}(z) = R_{G_{i,j}}(z)$.

As such, our first step would be to calculate these polynomials. To calculate $R_{i,j}(z)$, we take G, and then attach i vertices to v_1 and j vertices to v_2. We then calculate the reciprocal polynomial of the new resulting digraph. For example, to find $R_{1,0}(z)$, we attach one vertex to v_1, no vertices to v_2, and then calculate the reciprocal polynomial of the resulting digraph directly.

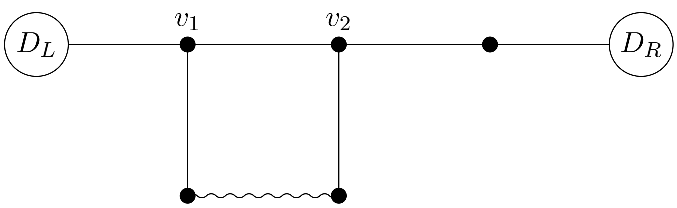
These are:

$$
\begin{aligned}
R_{0,0} &= z^8 + 2z^4 + 1\,,\\
R_{1,0} &= z^{10} + z^6 + z^4 + 1\,,\\
R_{0,1} &= z^{10} + z^6 + z^4 + 1\,,\\
R_{1,1} &= z^{12} + z^6 + 1\,.
\end{aligned} \tag{4.27}
$$

Following (4.3), we can now use these results to directly calculate our $B_{i,j}$:

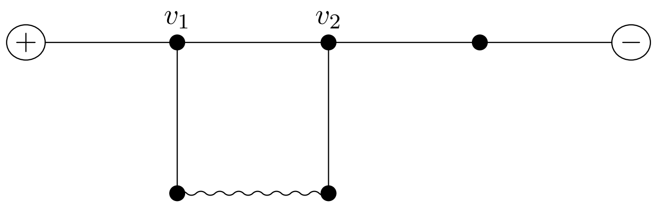
$$
\begin{aligned}
B_{1,1} &= z^{12} - 2z^{10} + z^8 - z^6\,,\\
B_{1,0} &= -z^8 + z^6 - z^4\,,\\
B_{0,1} &= -z^8 + z^6 - z^4\,,\\
B_{0,0} &= -z^6 + z^4 - 2z^2 + 1\,.
\end{aligned} \tag{4.28}
$$

Now, create a new digraph by attaching a path with $m_1 = 1$ vertices to the vertex v_1, with a decoration D_L at the end, and a path with $m_2 = 2$ vertices to the vertex v_2, with a decoration D_R at the end. So, our new digraph, which we refer to as $G^{\ddagger\ddagger}_{1,2}$, looks like:



.

We are referring to this digraph as $G^{\ddagger\ddagger}_{1,2}$, to mirror the notation used in Theorem 4.2.14. In this notation, "‡‡" represents our decorations (one "‡" for each decoration). As such, an alternative way to label this digraph could be $G^{D_L D_R}_{1,2}$. We choose to not label the digraph in this way, owing to how untidy it looks in comparison. In either case, we use the same notation we have introduced so far. For example, the digraph $G^{+-}_{1,2}$ would be:



This corresponds with our notation as presented in Lemma 4.2.8, for example, as the resulting reciprocal polynomial of this digraph would be $R^{+-}_{1,2}(z)$.

Now let D_L and D_R be neutral vertices, giving us the digraph $G^{\bullet\bullet}_{1,2}$. It is not difficult to calculate directly that the reciprocal polynomial here is $R^{\bullet\bullet}_{1,2}(z) = z^{14} + 1$.

By our knowledge of Theorem 4.2.14, we can use (4.28) and Table 4.1 to calculate the right hand side of (4.24) explicitly in this case. In this case, as both of our decorations are neutral vertices, we refer to the first line of Table 4.1 and see that $c_{10}(z) = -1$ and $c_{01}(z) = -1$. So, letting $m_1 = 1$, $m_2 = 2$ and $c_{10}(z) = c_{01}(z) = -1$ in (4.24), we get:

$$
\begin{aligned}
z^{2(m_1+m_2)}&B_{1,1}(z) + c_{10}(z)z^{2m_1}B_{1,0}(z) + c_{01}(z)z^{2m_2}B_{0,1}(z) + c_{00}(z)B_{0,0}(z) \\
&= z^6B_{1,1}(z) + (-1)z^2B_{1,0}(z) + (-1)z^4B_{0,1}(z) + (-1)(-1)B_{0,0}(z) \\
&= z^6\left(z^{12} - 2z^{10} + z^8 - z^6\right) - z^2\left(-z^8 + z^6 - z^4\right) \\
&\qquad - z^4\left(-z^8 + z^6 - z^4\right) + \left(-z^6 + z^4 - 2z^2 + 1\right) \\
&= z^{18} - 2z^{16} + z^{14} + z^4 - 2z^2 + 1\,.
\end{aligned}
$$

Again from the first line of Table 4.1 (and also via Lemma 4.2.1), we have that $\kappa_1(z) = \left(z^2 - 1\right)^2$ and $\kappa_2(z) = 1$ in (4.24). This gives us that the left hand side of (4.24) is $\left(z^2 - 1\right)^2 R^{\bullet\bullet}_{1,2}(z)$. Combining this with the result above and using (4.24), we get:

$$
\left(z^2 - 1\right)^2 R^{\bullet\bullet}_{1,2}(z) = z^{18} - 2z^{16} + z^{14} + z^4 - 2z^2 + 1\,.
$$

This gives us that $R^{\bullet\bullet}_{1,2}(z) = z^{14} + 1$ which agrees with our direct calculation.

Before moving on, we again stress that our aim here is to provide further validity to Theorem 4.2.14, and demonstrate the calculations in practice, and not to show a case of where this Theorem is simpler than direct calculations.

We now look at a different pair of decorations. Consider the case where D_L is a positive vertex and D_R is a fork. So, we are considering the digraph $G_{1,2}^{+<}$, and want to find $R_{1,2}^{+<}(z)$. We are familiar with this object and notation from (4.21), and again intend to verify that the result presented in Theorem 4.2.14 holds in this case. We can calculate (again, easily by hand) that the reciprocal polynomial in this case is: $R_{1,2}^{+<}(z) = z^{18} - z^{17} - z^{14} - 2z^{10} + 2z^9 - 2z^8 - z^4 - z + 1$.

By our knowledge of Theorem 4.2.14, we can use (4.28) and Table 4.1 to calculate the right hand side of (4.24) explicitly in this case. In this case, our left decoration, D_L, is a positive vertex, and our right decoration, D_R, is a fork. So, we refer to the eighth line of Table 4.1 and see that $c_{10}(z) = 1$ and $c_{01}(z) = z$. Thus, letting $m_1 = 1$, $m_2 = 2$, $c_{10}(z) = 1$ and $c_{01}(z) = z$ in (4.24), we get:

$$
\begin{aligned}
z^{2(m_1+m_2)}B_{1,1}(z) + c_{10}(z)z^{2m_1}B_{1,0}(z) + c_{01}(z)z^{2m_2}B_{0,1}(z) + c_{00}(z)B_{0,0}(z)& \\
=& z^6 B_{1,1}(z) + (1)z^2 B_{1,0}(z) + (z)z^4 B_{0,1}(z) + (1)(z)B_{0,0}(z) \\
=& z^6\left(z^{12} - 2z^{10} + z^8 - z^6\right) + z^2\left(-z^8 + z^6 - z^4\right) \\
&+ z^5\left(-z^8 + z^6 - z^4\right) + z\left(-z^6 + z^4 - 2z^2 + 1\right) \\
=& z^{18} - 2z^{16} + z^{14} - z^{13} - z^{12} + z^{11} - z^{10} - z^9 + z^8 - z^7 - z^6 + z^5 - 2z^3 + z\,.
\end{aligned}
$$

Again from the eighth line of Table 4.1 (and via (4.21)), we have that $\kappa_1(z) = z(z+1)$ and $\kappa_2(z) = (z^2+1)$ in (4.24). This gives us that the left hand side of (4.24) is $\frac{z(z+1)}{z^2+1}R_{1,2}^{+<}(z)$. Combining this with the result above and using (4.24), we get:

$$
\begin{aligned}
&\frac{z(z+1)}{z^2+1}R_{1,2}^{+<}(z) = \\
&\qquad z^{18} - 2z^{16} + z^{14} - z^{13} - z^{12} + z^{11} - z^{10} - z^9 + z^8 - z^7 - z^6 + z^5 - 2z^3 + z\,.
\end{aligned}
$$

This gives us that $R_{1,2}^{+<}(z) = z^{18} - z^{17} - z^{14} - 2z^{10} + 2z^9 - 2z^8 - z^4 - z + 1$, which again agrees with our direction calculation.

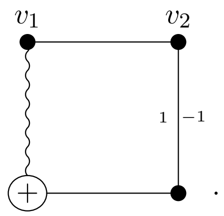
Finally, we give the reciprocal polynomials for all sixteen possible decorations, in Table 4.2. Using Table 4.1 and (4.28), one can verify that Theorem 4.2.14 holds and is indeed valid for this particular digraph. The results presented in Table 4.2 list each pair of decorations, D_L and D_R, and the resulting reciprocal polynomial $R_{1,2}^{\ddagger\ddagger}(z) = R_{1,2}^{D_L D_R}(z)$.

D_L	D_R	$R^{\ddagger\ddagger}_{1,2}(z)$
•	•	$z^{14}+1$
•	$+$	$z^{14}-z^{13}-z^{7}-z+1$
•	$-$	$z^{14}+z^{13}+z^{7}+z+1$
•	$<$	$z^{18}-z^{14}-2z^{10}-2z^{8}-z^{4}+1$
$+$	•	$z^{14}-z^{13}-z^{9}-z^{5}+z+1$
$+$	$+$	$z^{14}-2z^{13}+z^{12}-z^{9}+z^{8}-z^{7}+z^{6}-z^{5}+z^{2}-2z+1$
$+$	$-$	$z^{14}-z^{12}-z^{9}-z^{8}+z^{7}-z^{6}-z^{5}-z^{2}+1$
$+$	$<$	$z^{18}-z^{17}-z^{14}-2z^{10}+2z^{9}-2z^{8}-z^{4}-z+1$
$-$	•	$z^{14}+z^{13}+z^{9}+z^{5}+z+1$
$-$	$+$	$z^{14}-z^{12}+z^{9}-z^{8}-z^{7}-z^{6}+z^{5}-z^{2}+1$
$-$	$-$	$z^{12}+2z^{13}+z^{12}+z^{9}+z^{8}+z^{7}+z^{6}+z^{5}+z^{2}+2z+1$
$-$	$<$	$z^{18}+z^{17}-z^{14}-2z^{10}-2z^{9}-2z^{8}-z^{4}+z+1$
$>$	•	$z^{18}-z^{14}-2z^{12}-2z^{10}-2z^{8}-2z^{6}-z^{4}+1$
$>$	$+$	$z^{18}-z^{17}-z^{14}+z^{13}-2z^{12}+z^{11}-2z^{10}$ $+2z^{9}-2z^{8}+z^{7}-2z^{6}+z^{5}-z^{4}-z+1$
$>$	$-$	$z^{18}+z^{17}-z^{14}-z^{13}-2z^{12}-z^{11}-2z^{10}$ $-2z^{9}-2z^{8}-z^{7}-2z^{6}-z^{5}-z^{4}+z+1$
$>$	$<$	$z^{22}-2z^{18}-2z^{16}-3z^{14}-2z^{12}-2z^{10}-3z^{8}-2z^{6}-2z^{4}+1$

Table 4.2: Reciprocal polynomials for each pair of decorations in Example 4.2.15.

We now look at a second example, which exhibits an actual digraph (as opposed to a signed graph). This example is much the same as the previous example, with its main aim being to give further verification that Theorem 4.2.14 is indeed valid. We follow the same methods as presented in Example 4.2.15, but omit some of the details of calculations.

Example 4.2.16. Let G be the following digraph, with vertices v_1 and v_2 labelled:



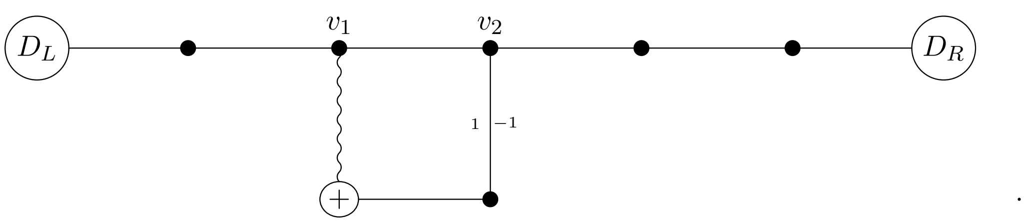
Under this set up, we can now calculate the $R_{i,j}(z) = R_{G_{i,j}}(z)$ for $i, j \in \{0, 1\}$, meaning we can then calculate the corresponding $B_{i,j}(z)$. These are:

$$\begin{aligned} R_{0,0} &= z^8 - z^7 + 2z^6 - 3z^5 + 2z^4 - 3z^3 + 2z^2 - z + 1\,, \\ R_{1,0} &= z^{10} - z^9 + 2z^8 - 3z^7 + z^6 - 3z^5 + z^4 - 3z^3 + 2z^2 - z + 1\,, \\ R_{0,1} &= z^{10} - z^9 + 2z^8 - 3z^7 + 3z^6 - 4z^5 + 3z^4 - 3z^3 + 2z^2 - z + 1\,, \\ R_{1,1} &= z^{12} - z^{11} + 2z^{10} - 3z^9 + 2z^8 - 4z^7 + z^6 - 4z^5 + 2z^4 - 3z^3 + 2z^2 - z + 1\,, \end{aligned} \tag{4.29}$$

and

$$\begin{aligned} B_{1,1} &= z^{12} - z^{11} - z^9 - z^8 + z^7 - z^6\,, \\ B_{1,0} &= z^8 - z^7 - z^6 - z^4\,, \\ B_{0,1} &= -z^8 - z^6 - z^5 + z^4\,, \\ B_{0,0} &= -z^6 + z^5 - z^4 - z^3 - z + 1\,. \end{aligned} \tag{4.30}$$

Now, create a new digraph by attaching a path with $m_1 = 2$ vertices to the vertex v_1, with a decoration D_L at the end, and a path with $m_2 = 3$ vertices to the vertex v_2, with a decoration D_R at the end. So, our new digraph, which we refer to as $G^{\ddagger\ddagger}_{2,3}$, looks like:



Now let D_L be a positively charged vertex and D_R be the fork of a snake tongue, giving us the digraph $G_{2,3}^{+<}$. It is not difficult to calculate directly that the reciprocal polynomial here is:

$$\begin{aligned} R_{2,3}^{+<}(z) = & z^{22} - 2z^{21} + 3z^{20} - 5z^{19} + 4z^{18} - 4z^{17} + 2z^{16} - 2z^{15} + z^{14} - 3z^{13} + 3z^{12} \\ & - 4z^{11} + 3z^{10} - 3z^{9} + z^{8} - 2z^{7} + 2z^{6} - 4z^{5} + 4z^{4} - 5z^{3} + 3z^{2} - 2z + 1 \,. \end{aligned}$$

By our knowledge of Theorem 4.2.14, we can use (4.30) and Table 4.1 to calculate the right hand side of (4.24) explicitly in this case:

$$\begin{aligned} & z^{2(2+3)}B_{1,1}(z) + (1)z^{2(2)}B_{1,0}(z) + (z)z^{2(3)}B_{0,1}(z) + (1)(z)B_{0,0}(z) = \\ & \quad z^{22} - z^{21} - z^{19} - z^{18} + z^{17} - z^{16} - z^{15} - z^{13} - z^{10} - z^{8} - z^{7} + z^{6} - z^{5} - z^{4} - z^{2} + z \,. \end{aligned}$$

From Table 4.1 (and via Lemma 4.2.1), we have that $\kappa_1(z) = z(z+1)$ and $\kappa_2(z) = z^2+1$ in (4.24). This gives us that the left hand side of (4.24) is:

$$\begin{aligned} & \frac{z(z+1)}{z^2+1} R_{2,3}^{+<}(z) = \\ & \quad z^{22} - z^{21} - z^{19} - z^{18} + z^{17} - z^{16} - z^{15} - z^{13} - z^{10} - z^{8} - z^{7} + z^{6} - z^{5} - z^{4} - z^{2} + z \,. \end{aligned}$$

So both sides agree and our results holds in this case.

Finally, for this example, we now give the reciprocal polynomials for all sixteen possible decorations. Using Table 4.1 and (4.30), one can verify that Theorem 4.2.14 holds in each case and is indeed valid. We present each polynomial in its factorised form for ease of presentation.

D_L	D_R	$R^{\ddagger\ddagger}_{1,2}(z)$
$\bullet$	$\bullet$	$\Phi_4(z)(z^{16}-z^{15}+z^{14}-2z^{13}+z^{12}-2z^{11}$ $-3z^9+z^8-3z^7-2z^5+z^4-2z^3+z^2-z+1)$
$\bullet$	$+$	$\Phi_4(z)\Phi_6(z)(z^{14}-z^{13}-2z^{11}+z^{10}+z^8-2z^7+z^6+z^4-2z^3-z+1)$
$\bullet$	$-$	$\Phi_4(z)(z^{16}-z^{13}-z^{12}-z^{11}-2z^{10}-3z^9-2z^8-3z^7-2z^6-z^5-z^4-z^3+1)$
$\bullet$	$<$	$\Phi_4(z)^2(z^{18}-z^{17}-z^{15}-z^{12}-z^{11}+z^{10}+z^8-z^7-z^6-z^3-z+1)$
$+$	$\bullet$	$z^{18}-2z^{17}+3z^{16}-5z^{15}+5z^{14}-6z^{13}+5z^{12}-7z^{11}$ $+6z^{10}-9z^9+6z^8-7z^7+5z^6-6z^5+5z^4-5z^3+3z^2-2z+1$
$+$	$+$	$\Phi_6(z)(z^{16}-2z^{15}+2z^{14}-4z^{13}+4z^{12}-3z^{11}+4z^{10}$ $-5z^9+4z^8-5z^7+4z^6-3z^5+4z^4-4z^3+2z^2-2z+1)$
$+$	$-$	$\Phi_4(z)(z^{16}-z^{15}-z^{13}-z^{10}-2z^9-2z^7-z^6-z^3-z+1)$
$+$	$<$	$\Phi_4(z)(z^{20}-2z^{19}+2z^{18}-3z^{17}+2z^{16}-z^{15}-z^{13}+z^{12}$ $-2z^{11}+2z^{10}-2z^9+z^8-z^7-z^5+2z^4-3z^3+2z^2-2z+1)$
$-$	$\bullet$	$\Phi_3(z)\Phi_6(z)(z^{14}-z^{11}-2z^{10}-z^9-z^8-z^7-z^6-z^5-2z^4-z^3+1)$
$-$	$+$	$\Phi_3(z)\Phi_6(z)(z^{14}-z^{13}-z^{11}-z^{10}+z^9+z^5-z^4-z^3-z+1)$
$-$	$-$	$\Phi_2(z)^2\Phi_3(z)\Phi_4(z)^2\Phi_6(z)(z^8-z^7-z^6+z^4-z^2-z+1)$
$-$	$<$	$\Phi_2(z)^2\Phi_3(z)\Phi_4(z)\Phi_6(z)(z^{14}-2z^{13}+2z^{12}-3z^{11}+2z^{10}$ $-z^9+z^8-z^7+z^6-z^5+2z^4-3z^3+2z^2-2z+1)$
$>$	$\bullet$	$\Phi_4(z)(z^{20}-z^{19}+z^{18}-2z^{17}-z^{15}-z^{14}-z^{13}-2z^{12}$ $-z^{11}-4z^{10}-z^9-2z^8-z^7-z^6-z^5-2z^3+z^2-z+1)$
$>$	$+$	$\Phi_4(z)\Phi_6(z)(z^{18}-z^{17}-2z^{15}+z^{13}+z^{12}-2z^{10}-2z^8+z^6+z^5-2z^3-z+1)$
$>$	$-$	$\Phi_2(z)\Phi_4(z)^2\Phi_{10}(z)z(z^{12}-z^{11}-z^9-z^7-z^5-z^3-z+1)$
$>$	$<$	$\Phi_2(z)^2\Phi_4(z)^2(z^{20}-3z^{19}+5z^{18}-8z^{17}+10z^{16}-11z^{15}$ $+11z^{14}-11z^{13}+10z^{12}-9z^{11}+8z^{10}-9z^9+10z^8$ $-11z^7+11z^6-11z^5+10z^4-8z^3+5z^2-3z+1)$

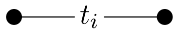
Table 4.3: Reciprocal polynomials for each pair of decorations in Example 4.2.16.

Before moving on, we again stress that Examples 4.2.15 and 4.2.16 do not exist to demonstrate the usefulness of our results here. They are provided solely to aid in the clarity of the results presented so far, and to give a sense of validity for some where we have not provided all of the details of the proof. However, this puts us in a good position to run our experiments and begin to actually find digraphs with small Mahler measures explicitly. This outlines the usefulness of these results, as we will see in chapter 5.

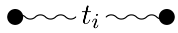
4.3 Growing Towards a Nice Shape

Our aim in this Section is to establish the following Theorem concerning the general shape of the reciprocal polynomial of a digraph that has one or more of its signed edges subdivided:

Theorem 4.3.1. *Let G be a digraph with r distinguished signed edges $e_1, \cdots, e_r$, each joining two neutral vertices. For integers $t_1, \cdots, t_r \geq 0$, let $G_{[t_1,\cdots,t_r]}$ be the digraph obtained from G by subdividing each edge e_i by adding t_i new vertices. Thus e_i is replaced by*



or



according as its sign is positive or negative. Then there is a polynomial

$$Q(z_0, z_1, \cdots, z_r) \in \mathbb{Z}[z_0, z_1, \cdots, z_r],$$

depending on G but independent of $t_1, \cdots, t_r$, such that $G_{[t_1,\cdots,t_r]}$ has reciprocal polynomial:

$$\frac{Q(z, z^{t_1}, z^{t_2}, \cdots, z^{t_r})}{z^{2r}(z^2-1)^r}.$$

Remark. This Theorem does not exclude the possibility of pendant paths; for example, if one of the edges e_i has a vertex which is a leaf.

Remark. If *all* growing paths are pendant (that is to say, if all the edges have a vertex that has a leaf), then this is precisely Lemma 4.1.3, just stated in a different way.

The latter remark here shows how Theorem 4.3.1 is an expansion of Lemma 4.1.3. As mentioned, the concept of subdividing edges within our digraph will be able to help us turn suitable digraph candidates into ones which are, combinatorially, almost cyclotomic. As such, this is a very important tool in our hunt for small Mahler measures.

4.3.1 Deletion and Contraction

A tool in our proof will be a tiny modification of a deletion-contraction theorem due to Rowlinson [33]. The theorem from Rowlinson was stated and proved for multigraphs

(graphs where multiple edges are allowed between vertices), but an examination of the proof shows that it easily extends to $\mathbb{Z}$-*weighted digraphs* (more general than the digraphs we need). So, prior to proving our Theorem, we need to state and prove this generalisation.

Firstly, we fix some notation and definitions, many of which are extensions to those that we are familiar with.

Definition 4.3.2. *A* $\mathbb{Z}$***-weighted digraph*** *is a type of graph such that:*

- *We have (directed) arcs between vertices,*
- *We allow directed, weighted loops,*
- *We can attach weights to our arcs, which are arbitrary elements of* $\mathbb{Z}$.

We see how this is more general than the digraphs which we use: we only allow signed arcs (values ± 1) and our loops are represented by charges (which again are values ± 1 only). As with our digraphs, the integer values for loops (the extension of charges) appear as diagonal entries of the adjacency matrix, and the integer arc weights appear as off-diagonal entries.

Definition 4.3.3. *Let G be a $\mathbb{Z}$-weighted digraph, and $A = (a_{ij})$ be its adjacency matrix. If $a_{ij} = a_{ji} \neq 0$ for some i and j, then we say that there is a* ***signed multiedge*** *between i and j.*

Perhaps unsurprisingly, the deletion-contraction theorem from Rowlinson revolves around 'deleting' and 'contracting' multiedges within multigraphs. We now formalize these definitions and fix some convenient notation which we will use:

Definition 4.3.4 (**Deletion of a multiedge**)**.** *Let G be a $\mathbb{Z}$-weighted digraph for which there is a signed multiedge between vertices u and v. We denote the deletion of the signed multiedge between u and v by $G - [uv]$. This deletes (i.e. removes) the signed multiedge, and nothing else.*

Remark. If G has adjacency matrix $A = (a_{ij})$, then the adjacency matrix of $G - [uv]$ is obtained from A by setting $a_{uv} = a_{vu} = 0$.

Definition 4.3.5 (Deletion of a vertex). *Let G be a $\mathbb{Z}$-weighted digraph. We denote the deletion of the vertex u by $G - u$. This deletes the vertex u, and all multiedges and arcs incident with u.*

Remark. If G has adjacency matrix $A = (a_{ij})$, then the adjacency matrix of $G - u$ is obtained from A by deleting row u and column u.

Remark. We can delete more than one vertex from a $\mathbb{Z}$-weighted digraph. For example, if we delete vertices u and v, we denote this by $G - u - v$, and to obtain the adjacency matrix, we delete rows u and v, and columns u and v from $A = A_G$.

We further note that $G - u - v = G - v - u$.

Definition 4.3.6 (Contraction of a multiedge). *Let G be a $\mathbb{Z}$-weighted digraph for which there is a signed multiedge between vertices u and v. We denote the contraction of the signed multiedge between u and v by G^*.*

This 'amalgamates' the vertices u and v, by summing the weights of their loops, whilst the weight of the multiedge does not contribute.

Remark. If G has adjacency matrix $A = (a_{ij})$, then the adjacency matrix of G^* is obtained from A by:

i. Adding row u to row v,

ii. Adding column u to column v,

iii. Subtracting $2a_{uv}$ from the new (v, v)-entry,

iv. Deleting row u and column u.

Remark. The notation G^* is somewhat vague, since the contraction of course depends on u and v. As such, notation such as G^*_{uv} may seem more suitable. However, the use of G^* is standard notation, in part because it is always clear what is being contracted, and in part because G^*_{uv} is rather clunky.

Example 4.3.7. Here we look at a straightforward $\mathbb{Z}$-weighted digraph G and the resulting $\mathbb{Z}$-weighted digraphs that occur from examples of deletion and contraction. For sake of simplicity, within the Figure, we do not label the vertices u and v, but these should be clear from context. Furthermore, our 'signed multiedge' in this case is in fact just "an edge" (that is, a positively signed multiedge with weight 1).

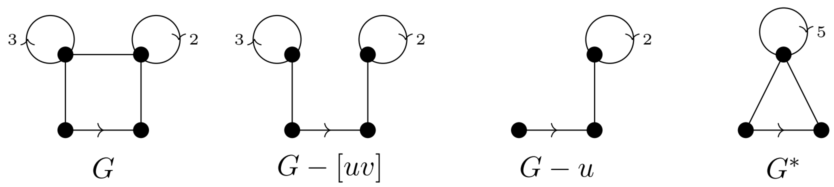


Figure 4.5: A look at $G-[uv], G-u$ and G^* for a straightforward G.

We can now turn to stating, and proving, the generalisation of the deletion-contraction theorem from Rowlinson:

Theorem 4.3.8 (Compare Rowlinson [33, Theorem 1.3])**.** *Let G be a $\mathbb{Z}$-weighted digraph with at least three vertices for which there is a signed multiedge between (distinct) vertices u and v, and let $c \in \mathbb{Z}$ be the weight of this multiedge. Writing $\chi_H = \chi_H(x)$ for the characteristic polynomial of any given digraph H, we have:*

$$\chi_G = \chi_{G-[uv]} + c\chi_{G^*} + c(x-c)\chi_{G-u-v} - c\chi_{G-u} - c\chi_{G-v}\,. \tag{4.31}$$

Proof. We follow the proof from Rowlinson, commenting on the tiny differences needed for the generalisation.

We may as well assume that $u = 1$ and $v = 2$. Then the adjacency matrix (a_{ij}) of G has the shape

$$\begin{pmatrix} a & c & \mathbf{r} \\ c & b & \mathbf{s} \\ \mathbf{p} & \mathbf{q} & A \end{pmatrix}.$$

(In the multigraph setting, $c > 0$, $\mathbf{p} = \mathbf{r}^T$, $\mathbf{q} = \mathbf{s}^T$, and A is symmetric. The only symmetry we require in this generalisation is that $a_{12} = a_{21} = c$.) The adjacency matrices of $G-[uv]$, $G-u$, $G-v$, $G-u-v$, and G^* are then:

$$\underbrace{\begin{pmatrix} a & 0 & \mathbf{r} \\ 0 & b & \mathbf{s} \\ \mathbf{p} & \mathbf{q} & A \end{pmatrix}}_{G-[uv]}, \quad \underbrace{\begin{pmatrix} b & \mathbf{s} \\ \mathbf{q} & A \end{pmatrix}}_{G-u}, \quad \underbrace{\begin{pmatrix} a & \mathbf{r} \\ \mathbf{p} & A \end{pmatrix}}_{G-v}, \quad \underbrace{\begin{pmatrix} A \end{pmatrix}}_{G-u-v}; \quad \underbrace{\begin{pmatrix} a+b & \mathbf{r}+\mathbf{s} \\ \mathbf{p}+\mathbf{q} & A \end{pmatrix}}_{G^*}.$$

Using the fact that the determinant is a linear function of any row or column, and that swapping rows or columns changes the sign, we have:

$$\begin{aligned}
\chi_G(x) &= \begin{vmatrix} x-a & -c & -\mathbf{r} \\ -c & x-b & -\mathbf{s} \\ -\mathbf{p} & -\mathbf{q} & xI-A \end{vmatrix} \\
&= \begin{vmatrix} x-a & 0 & -\mathbf{r} \\ -c & x-b & -\mathbf{s} \\ -\mathbf{p} & -\mathbf{q} & xI-A \end{vmatrix} + \begin{vmatrix} 0 & -c & \mathbf{0} \\ -c & x-b & -\mathbf{s} \\ -\mathbf{p} & -\mathbf{q} & xI-A \end{vmatrix} \\
&= \begin{vmatrix} x-a & 0 & -\mathbf{r} \\ 0 & x-b & -\mathbf{s} \\ -\mathbf{p} & -\mathbf{q} & xI-A \end{vmatrix} + \begin{vmatrix} x-a & 0 & -\mathbf{r} \\ -c & 0 & \mathbf{0} \\ -\mathbf{p} & -\mathbf{q} & xI-A \end{vmatrix} + \begin{vmatrix} 0 & -c & \mathbf{0} \\ -c & x-b & -\mathbf{s} \\ -\mathbf{p} & -\mathbf{q} & xI-A \end{vmatrix} \\
&= \chi_{G-[uv]} + c \begin{vmatrix} 1 & 0 & \mathbf{0} \\ x-a & 0 & -\mathbf{r} \\ -\mathbf{p} & -\mathbf{q} & xI-A \end{vmatrix} + c \begin{vmatrix} 1 & 0 & \mathbf{0} \\ x-b & -c & -\mathbf{s} \\ -\mathbf{q} & -\mathbf{p} & xI-A \end{vmatrix} \\
&= \chi_{G-[uv]} + c \begin{vmatrix} 0 & -\mathbf{r} \\ -\mathbf{q} & xI-A \end{vmatrix} + c \begin{vmatrix} 0 & -\mathbf{s} \\ -\mathbf{p} & xI-A \end{vmatrix} + c \begin{vmatrix} -c & \mathbf{0} \\ -\mathbf{p} & xI-A \end{vmatrix} \\
&= \chi_{G-[uv]} - c^2\chi_{G-u-v} + c \begin{vmatrix} 0 & -\mathbf{r} \\ -\mathbf{q} & xI-A \end{vmatrix} + c \begin{vmatrix} 0 & -\mathbf{s} \\ -\mathbf{p} & xI-A \end{vmatrix}. \qquad (4.32)
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\chi_{G^*}(x) &= \begin{vmatrix} x-(a+b) & -\mathbf{r}-\mathbf{s} \\ -\mathbf{p}-\mathbf{q} & xI-A \end{vmatrix} \\
&= \begin{vmatrix} x-(a+b) & \mathbf{0} \\ -\mathbf{p}-\mathbf{q} & xI-A \end{vmatrix} + \begin{vmatrix} 0 & -\mathbf{r} \\ -\mathbf{p}-\mathbf{q} & xI-A \end{vmatrix} + \begin{vmatrix} 0 & -\mathbf{s} \\ -\mathbf{p}-\mathbf{q} & xI-A \end{vmatrix} \\
&= (x-a-b)\chi_{G-u-v} \\
&\quad + \begin{vmatrix} x-a & -\mathbf{r} \\ -\mathbf{p} & xI-A \end{vmatrix} - \begin{vmatrix} x-a & \mathbf{0} \\ -\mathbf{p} & xI-A \end{vmatrix} + \begin{vmatrix} 0 & -\mathbf{s} \\ -\mathbf{p} & xI-A \end{vmatrix} \\
&\quad + \begin{vmatrix} x-b & -\mathbf{s} \\ -\mathbf{q} & xI-A \end{vmatrix} - \begin{vmatrix} x-b & \mathbf{0} \\ -\mathbf{q} & xI-A \end{vmatrix} + \begin{vmatrix} 0 & -\mathbf{r} \\ -\mathbf{q} & xI-A \end{vmatrix} \\
&= (x-a-b)\chi_{G-u-v} + \chi_{G-v} - (x-a)\chi_{G-u-v} + \chi_{G-u} - (x-b)\chi_{G-u-v} \\
&\quad + \begin{vmatrix} 0 & -\mathbf{r} \\ -\mathbf{q} & xI-A \end{vmatrix} + \begin{vmatrix} 0 & -\mathbf{s} \\ -\mathbf{p} & xI-A \end{vmatrix} \\
&= \chi_{G-u} + \chi_{G-v} - x\chi_{G-u-v} + \begin{vmatrix} 0 & -\mathbf{r} \\ -\mathbf{q} & xI-A \end{vmatrix} + \begin{vmatrix} 0 & -\mathbf{s} \\ -\mathbf{p} & xI-A \end{vmatrix}. \qquad (4.33)
\end{aligned}$$

Multiplying (4.33) by c and then subtracting from (4.32) gives the result. □

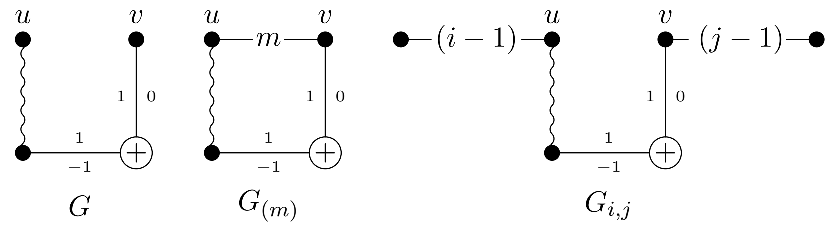
4.3.2 Adding to Divide

For the next ingredient we need to prove Theorem 4.3.1, we need to define a modification to a digraph G that has distinct vertices u and v for which there are no signed arcs in either direction between u and v:

Definition 4.3.9. *Let G be a digraph, and u and v be vertices of G for which there are no signed arcs in either direction between u and v. For $m \geq 2$, we denote by $G_{(m)}$ the digraph obtained by adding a positive edge between u and v, and then subdividing this edge with m vertices. Thus, $G_{(m)}$ has m more vertices than G.*

The next result describes the shape of the reciprocal polynomial of $G_{(m)}$. This result will make reference to digraphs $G_{i,j}$; those which have had two neutrally charged pendant paths attached to them. As such, it may help to see an example:

Example 4.3.10. *The following is a comparison between $G_{(m)}$ and $G_{i,j}$, for the same digraph G and distinguished vertices u and v:*



Lemma 4.3.11. *Let G be a digraph with two distinguished neutral vertices u and v that have no signed arc between them in either direction.*

Let $R_{(m)} = R_{(m)}(z)$ and $R_{i,j} = R_{i,j}(z)$ be the reciprocal polynomials of $G_{(m)}$ and $G_{i,j}$ respectively. Then, for $m \geq 2$, we have:

$$\begin{aligned}(z^2-1)R_{(m)} = z^{m-2}&(z^2-1)R_{(2)} \\ &+ (z^{m-2}-1)(z^m+1)R_{11} \\ &- z^2(z^{m-2}-1)(z^{m-2}+1)(R_{10}+R_{01}) \\ &+ z^4(z^{m-2}-1)(z^{m-4}+1)R_{00}\,. \end{aligned} \tag{4.34}$$

Proof. Let v_1 be the vertex on our subdivided edge that is adjacent to u, and let v_2 be that which is adjacent to v_1. We shall apply the deletion-contraction formula (4.31) for the digraph $G_{(m)}$ with the edge between v_1 and v_2:

$$\chi_{G_{(m)}}(x) = \chi_{G_{1,m-1}}(x) + \chi_{G_{(m-1)}}(x) + (x-1)\chi_{G_{0,m-2}}(x) - \chi_{G_{0,m-1}}(x) - \chi_{G_{1,m-2}}(x)\,. \tag{4.35}$$

To find the reciprocal polynomial, we replace x by $z + 1/z$ and multiply by z^m, noting that $G_{1,m-1}$ has the same number of vertices as $G_{(m)}$, $G_{0,m-2}$ has two vertices fewer, and all the other digraphs whose characteristic polynomial appears on the right in (4.35) have one fewer vertex:

$$R_{(m)} = R_{1,m-1} + zR_{(m-1)} + z(z^2 - z + 1)R_{0,m-2} - zR_{0,m-1} - zR_{1,m-2}\,. \tag{4.36}$$

Now we use (3.3) to write each $R_{0,j}$ in terms of $R_{0,0}$ and $R_{0,1}$ and each $R_{1,j}$ in terms of $R_{1,0}$ and $R_{1,1}$, simplifying by multiplying by (z^2-1):

$$\begin{aligned}(z^2-1)R_{(m)} = z(z^2-1)&R_{(m-1)} \\ &+ (z-1)(z^{2m-3}+1)R_{1,1} \\ &- z^2(z-1)(z^{2m-5}+1)(R_{1,0}+R_{0,1}) \\ &+ z^4(z-1)(z^{2m-7}+1)R_{0,0}\,. \end{aligned} \tag{4.37}$$

Now we claim that for $0 \le t \le m-2$ we have:

$$\begin{aligned}(z^2-1)R_{(m)} = z^t(z^2-1)&R_{(m-t)} \\ &+ (z^t-1)(z^{2m-2-t}+1)R_{1,1} \\ &- z^2(z^t-1)(z^{2m-4-t}+1)(R_{1,0}+R_{0,1}) \\ &+ z^4(z^t-1)(z^{2m-6-t}+1)R_{0,0}\,. \end{aligned} \tag{4.38}$$

This is trivial for $t=0$, and for $t=1$ it is precisely (4.37). If (4.38) holds for some $t<m-2$, then applying (4.37) (with m replaced by $m-t$) to the right hand side of (4.38) gives (after some manipulation) (4.38) for $t+1$. Inductively, our claim is established.

Putting $t=m-2$ in (4.38) gives (4.34), as is required. □

To further aid understanding of this proof, it may help to see explicit formulas for $R_{0,j}$ and $R_{1,j}$. We state these here, using (4.2), as to not disrupt the flow and clarity of the proof:

$$(z^2-1)R_{0,j} = (z^{2j}-1)R_{0,1} + (z^{2j}-z^2)R_{0,0}\,,$$
$$(z^2-1)R_{1,j} = (z^{2j}-1)R_{1,1} + (z^{2j}-z^2)R_{1,0}\,.$$

Upon a first viewing, it may seem like this only applies to subdivisions of positive edges. Indeed, $G_{(m)}$ has only been defined for positive edges, and we have not given any alternative. However, we are fortunate in that we do not need to make any alterations if we wish to consider subdivisions of negative edges.

For the analogue of Lemma 4.3.11 when we subdivide a negative edge, we note that Theorem 4.3.8 still applies to give (4.36) unchanged. Then, having noted that (3.3) still holds regardless of signs of edges on any pendant paths, we get all the subsequent formulas of Lemma 4.3.11 too. Of course the analogue of $R_{(2)}$ for the subdivision of a negative edge may be different from that for the subdivision of a positive edge.

4.3.3 Proof

We now have everything we need to prove the main Theorem of this chapter.

Proof of Theorem 4.3.1. We proceed by induction on the number, r, of subdivided signed edges: the case $r = 0$ being trivial.

Inductively, if we subdivide $r-1$ chosen signed edges of any digraph K, adding t_i new vertices to the i-th chosen signed edge to produce $K_{[t_1,\cdots,t_{r-1}]}$, we have

$$z^{2(r-1)}(z^2-1)R_{K_{[t_1,\cdots,t_{r-1}]}} = Q_{K_{[t_1,\cdots,t_{r-1}]}}(z, z^{t_1}, \cdots, z^{t_{r-1}}),$$

for some polynomial $Q_{K_{[t_1,\cdots,t_{r-1}]}}(z) \in \mathbb{Z}[z_0, z_1, \cdots, z_{r-1}]$, depending on K and the chosen signed edges. In particular, if we subdivide the first $r-1$ chosen signed edges of K, and then apply Lemma 4.3.11 to the final signed edge, each reciprocal polynomial on the right of (4.34) has this form for various choices of K. Multiplying (4.34) by $z^{2r}(z^2-1)^{r-1}$, we deduce that $z^{2r}(z^2-1)^r R_{G_{[t_1,\cdots,t_r]}}$ has the desired shape. □

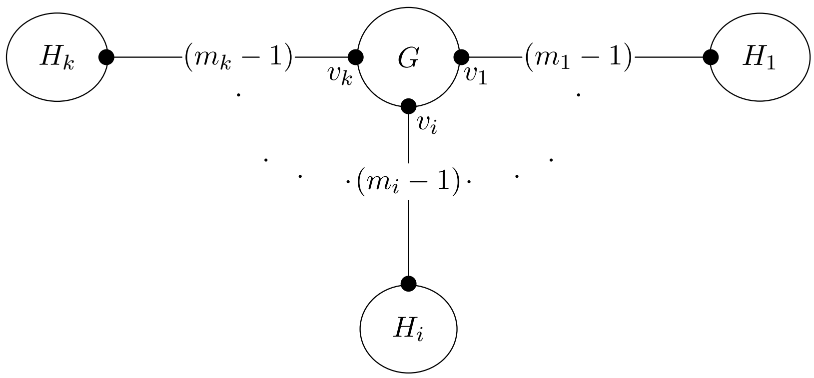
4.4 Bridges Between Digraphs

Our current results in this chapter have allowed us to grow digraphs in very simple settings. In effect, our results allow us to either attach pendant paths (with some sort of small decoration at the end), or to grow some form of internal path. Of course, in the context of finding digraphs with small Mahler measures, these are perhaps the most sensible things to consider when growing digraphs.

However, as noted in Section 4.2.4, when we were attaching pendant paths and snake tongues to digraphs, we were actually doing special situations of something more general. We were attaching a path of m vertices, and fixing it so that the final vertex of this path was part of some, very small, cyclotomic digraph. So, we could view this as taking two digraphs, and then connecting them together with a path.

4.4.1 Bridges Between Digraphs: A Justification

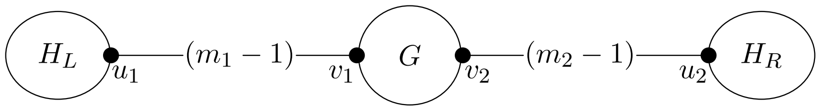
We currently have a good collection of ways to grow our digraphs. Before we move to running experiments with these ideas, however, it is worth asking if there are any other ideas we should be considering. One such idea, inspired by the revelation that we are connecting digraphs with long paths, is as follows: take a digraph G and attach k pendant paths, of length $m_1, \cdots, m_k$ respectively, to a list of k (not necessarily distinct) vertices $v_1, \cdots, v_k$. At the end of these pendant paths, attach further digraphs $H_1, \cdots, H_k$, and calculate the Mahler measure of the resulting new digraph. It may help to see a rough visualisation of this idea:



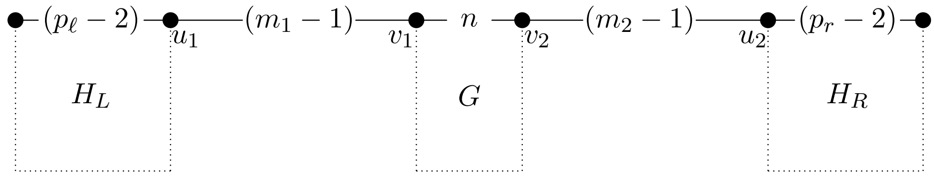
.

Upon first viewing, this idea may seem somewhat brash. Indeed, the idea of growing digraphs is to take a digraph G which is not cyclotomic, and obtain a new digraph which is combinatorially close to being cyclotomic, in hope of getting a small Mahler measure. This certainly seems difficult to achieve under this idea. However, if we look at this idea in a more precise situation, a potential use comes from this.

Instead of looking at the general case, we look at the case where $k = 2$ (and in turn, begin to mimic somewhat the ideas we have already seen). For convenience, we label the newly attached digraphs as H_L and H_R, and say that the vertices within these digraphs that are attached to the pendant paths are labelled u_1 and u_2 respectively. This gives us the following:



Next, consider a situation where the original digraph G has a path which exists between v_1 and v_2 (say of length n vertices). Since we choose our G, this is a valid consideration to make. Furthermore, assume that H_L and H_R each have paths of a "long" length which start at u_1 and u_2 respectively (say of lengths p_ℓ and p_r vertices respectively). In terms of a visualisation, this gives us:



Now, we can view this new digraph as a long path, on $p_\ell + m_1 + n + m_2 + p_r$ vertices, that has been nudged in some way with the additions of the remain content of G, H_L and H_R. If we have our original digraph G, and attached digraphs H_L and H_R, have a total number of vertices which is comparatively significantly smaller than $p_\ell + m_1 + n + m_2 + p_r$, and say choose H_L and H_R to themselves be cyclotomic, then the resulting digraph can be thought to be combinatorially close to being cyclotomic. From here, we can then employ our heuristic hope that a digraph which is combinatorially close to being cyclotomic will have a Mahler measure close to being 1, and in turn be small.

This idea is perhaps not as intuitive as attaching decorated pendant paths to our original digraph, but it does at least seem that there is some potential for the idea to give fruitful results. We now need to mirror our earlier results and see if we can find the general shape of the reciprocal polynomials of these digraphs.

4.4.2 Bridges and Simple Graphs

We approach this idea in a similar way to how we approached attaching decorated pendant paths. That is to say, we first look at similar situations involving simple graphs, and see if, and how, we can generalise those results.

Cvetković et al. [10] have some results for simple graphs which are relevant to us. We first fix some notation and definitions, and then move to giving these results.

Definition 4.4.1. *Let G and H be (simple) graphs, and u and v be vertices of G and H respectively.*

*The **coalescence** of G and H, denoted by $G \cdot H$, is the graph obtained by identifying the vertex u from G and the vertex v from H.*

Remark. $G \cdot H$ depends on the choice of u and v, however these are not included in the notation to avoid over-cluttering.

Definition 4.4.2. *Let G and H be (simple) graphs, and u and v be vertices of G and H respectively.*

*The **bridge** of G and H, denoted by $G_{uv}H$, is the graph obtained by adding an edge between the vertex u from G and the vertex v from H.*

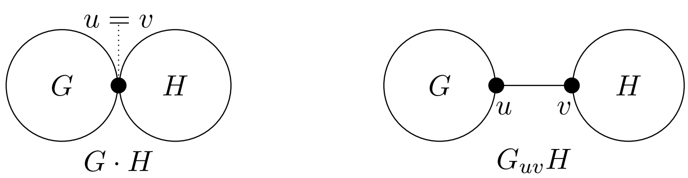


Figure 4.6: A visualisation of $G \cdot H$ and $G_{uv}H$.

We wish to understand the shape of the reciprocal polynomial of these new (simple) graphs. These are covered by Cvetković et al.; we state the results and give the corresponding proofs, so that we may modify them to get results for our digraphs.

Proposition 4.4.3 (Cvetković et al. [10], Theorem 2.2.3)**.** *For any simple graphs G*

and H, we have:

$$\chi_{G\cdot H}(x) = \chi_G(x)\chi_{H-v}(x) + \chi_{G-u}(x)\chi_H(x) - x\chi_{G-u}(x)\chi_{H-v}(x)\,. \qquad (4.39)$$

Proof. Let the adjacency matrices of G and H be

$$\begin{pmatrix} A' & \mathbf{r} \\ \mathbf{r}^\top & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & \mathbf{s}^\top \\ \mathbf{s} & B' \end{pmatrix}$$

respectively, for some matrices A', B' and vectors $\mathbf{r}, \mathbf{s}$. Then, the adjacency matrix of $G \cdot H$ is:

$$\begin{pmatrix} A' & \mathbf{r} & O \\ \mathbf{r}^\top & 0 & \mathbf{s}^\top \\ O & \mathbf{s} & B' \end{pmatrix},$$

where each O is a zero matrix of the necessary size.

This now gives us the following:

$$\begin{aligned}\chi_{G\cdot H}(x) &= \begin{vmatrix} xI - A' & -\mathbf{r} & O \\ -\mathbf{r}^\top & x & -\mathbf{s}^\top \\ O & -\mathbf{s} & xI - B' \end{vmatrix} \\ &= \begin{vmatrix} xI - A' & -\mathbf{r} & O \\ -\mathbf{r}^\top & x & -\mathbf{s}^\top \\ O & \mathbf{0} & xI - B' \end{vmatrix} + \begin{vmatrix} xI - A' & \mathbf{0} & O \\ -\mathbf{r}^\top & x & -\mathbf{s}^\top \\ O & -\mathbf{s} & xI - B' \end{vmatrix} - \begin{vmatrix} xI - A' & \mathbf{0} & O \\ -\mathbf{r}^\top & x & -\mathbf{s}^\top \\ O & \mathbf{0} & xI - B' \end{vmatrix}.\end{aligned}$$

Now consider $\begin{vmatrix} xI - A' & -\mathbf{r} & O \\ -\mathbf{r}^\top & x & -\mathbf{s}^\top \\ O & \mathbf{0} & xI - B' \end{vmatrix}$. In the sub-determinant $\begin{vmatrix} x & -\mathbf{s}^\top \\ \mathbf{0} & xI - B' \end{vmatrix}$, we require a non-zero contribution from every row. To achieve this, we clearly need a contribution from every from of the matrix $xI - B'$, the adjacency matrix of the graph H with the vertex v removed. This gives a contribution of $\chi_{H-v}(x)$ to the determinant.

Equally, we have the sub-determinant $\begin{vmatrix} xI - A' & -\mathbf{r} \\ -\mathbf{r}^\top & x \end{vmatrix}$, which simply contributes $\chi_G(x)$. Multiplying these two contributions together we get $\chi_G(x)\chi_{H-v}(x)$, which is the first part of the result.

The remaining two parts of the result follow similarly from the remaining two determinants. □

Before moving to our next result, we make a comment about notation that we will use going forward. If we wish to write down the characteristic polynomial of a digraph G which has had a pendant path with m vertices attached, we use the notation χ_{G_m}. This is a more formal and precise (and unfortunately more dense) version of the notation which was first introduced in Lemma 3.3.2 (where this was written as χ_m).

The reason for introducing this notation now is because we are now working with more than one original digraph. Previously, writing χ_m would cause no issue, as it would be clear which digraph we were attaching the pendant path to. Now, ambiguity could arise from such a notation, and so we will use this more formal notation here.

Proposition 4.4.4 (Cvetković et al. [10], Theorem 2.2.4)**.** *For simple graphs G and H, we have:*

$$\chi_{G_{uv}H}(x) = \chi_G(x)\chi_H(x) - \chi_{G-u}(x)\chi_{H-v}(x)\,. \tag{4.40}$$

Proof. We first note that $G_{uv}H = G_1 \cdot H$. That is to say, it is the coalescence of G with a pendant path of one vertex attached to the vertex u, and H, where we identify the vertex of the end of the pendant with the vertex v in H. Therefore, $\chi_{G_{uv}H}(x) = \chi_{G_1 \cdot H}(x)$. From here, we use (4.39) to obtain:

$$\chi_{G_1 \cdot H}(x) = \chi_{G_1}(x)\chi_{H-v}(x) + \chi_G(x)\chi_H(x) - x\chi_G(x)\chi_{H-v}(x)\,. \tag{4.41}$$

We note that we are already have a formula to simplify $\chi_{G_1}(x)$; we set $m = 1$ in (3.4), which was seen in the proof of Lemma 3.3.2. This gives us:

$$\begin{aligned}\chi_{G_1 \cdot H}(x) &= \Big(x\chi_G(x) - \chi_{G-u}(x)\Big)\chi_H(x) + \chi_G(x)\chi_H(x) - x\chi_G(x)\chi_{H-v}(x)\\ &= \chi_G(x)\chi_H(x) - \chi_{G-u}(x)\chi_{H-v}(x)\,,\end{aligned}$$

as required. □

This result only applies to situations where a single edge joins two graphs together.

However, it is easy to extend this result to a longer 'bridge' between the graphs. To state this result, we first have to introduce some further, unfortunately messy, notation:

Definition 4.4.5. *Let G and H be (simple) graphs, and u and v be vertices of G and H respectively.*

*The **n-bridge** of G and H, denoted by $G^n_{uv}H$, is the graph obtained by adding a path of n vertices (and so $n+1$ edges) between the vertex u from G and the vertex v from H.*

Remark. The bridge of G and H can be thought of as the 0-bridge of G and H.

Remark. We can view the n-bridge as a coalescence of two graphs: $G^n_{uv}H = G_n \cdot H$ (where, again, G_n is the graph G with a pendant path of n vertices attached to the vertex u).

Corollary 4.4.5.1. *For simple graphs G and H, we have:*

$$\chi_{G^n_{uv}H}(x) = \chi_{G_n}(x)\chi_{H-v}(x) + \chi_{G_{n-1}}(x)\chi_H(x) - x\chi_{G_{n-1}}(x)\chi_{H-v}(x)\,. \tag{4.42}$$

Proof. The result follows in an analogous fashion to Theorem 4.4.4. □

4.4.3 Bridges and Digraphs

Our aim in this Section is to extend the results of Cvetković et al. to digraphs. In particular, if we are able to extend the result of coalescence to digraphs, we are then immediately in a position to extend their result of a simple graph with a bridge to digraphs, since we can define the n-bridge of two digraphs in terms of coalescence.

Our next step would be to understand the shape of the reciprocal polynomial of these newly formed digraphs; especially that of the n-bridge of two digraphs.

From here, this puts us in a position to tackle our goal. We can think of the n-bridge of two digraphs as taking a digraph, and then attaching a pendant path to it, with a second digraph at the end of the pendant path. If we are able to extend the result to adding a second digraph, we then have the reciprocal polynomial of the digraph we want, and could then look to use this in our experiments.

Fortunately, it is quite straightforward to extend the result of coalescence to digraphs, provided our set-up is correct:

Lemma 4.4.6. *For any digraphs G and H which share at least one common charged vertex, we have that:*

$$\chi_{G \cdot H}(x) = \chi_G(x)\chi_{H-v}(x) + \chi_{G-u}(x)\chi_H(x) - x\chi_{G-u}(x)\chi_{H-v}(x)\,,$$

where the identified vertices u in G and v in H have the same charge.

Proof. In this case, the adjacency matrix of $G \cdot H$ is:

$$\begin{pmatrix} A' & \mathbf{r} & O \\ \hat{\mathbf{r}}^\top & c & \hat{\mathbf{s}}^\top \\ O & \mathbf{s} & B' \end{pmatrix},$$

for some vectors $\mathbf{r}, \mathbf{s}, \hat{\mathbf{r}}$ and $\hat{\mathbf{s}}$, and where c is the corresponding value for the charge of the identified vertex $u = v$.

The result follows from here, as in Theorem 4.4.3, albeit with minor differences to the adjacency matrices. □

It should not come as surprise that the coalescence of two digraphs can be found in this way as well. And indeed, it should not come as a surprise that we are not able to find the coalescence of two digraphs which do not share a common charged vertex in our case.

A consequence of this is that we are now able to extend the results of Theorem 4.4.4 and Corollary 4.4.5.1 to digraphs. We simply apply our knowledge from Lemma 4.4.6 and ensure we have suitable digraphs.

Lemma 4.4.7. *For digraphs G and H that each have at least one neutral vertex, we have:*

$$\chi_{G^n_{uv}H}(x) = \chi_{G_n}(x)\chi_{H-v}(x) + \chi_{G_{n-1}}(x)\chi_H(x) - x\chi_{G_{n-1}}(x)\chi_{H-v}(x)\,, \tag{4.43}$$

where the vertices u and v are neutrally charged.

Proof. The result also follows in an analogous fashion to Theorem 4.4.4. □

We are now in a position to find the general shape of the reciprocal polynomial of the n-bridge of two digraphs.

Proposition 4.4.8. *For digraphs G and H, neutral vertices u in G and v in H, and $n \geq 1$, the reciprocal polynomial of $G^n_{uv}H$ satisfies the following:*

$$(z^2-1)R_{G^n_{uv}H}(z) = L_1(z)B(z) + L_0(z)B^*(z)\,, \tag{4.44}$$

where L_1 and L_0 are Laurent polynomials and B is some monic integer polynomial that depends on G but not on u, v or n.

Proof. For convenience and ease of notation, let $\mathcal{G} = G^n_{uv}H$. So, we can rewrite (4.43) as:

$$\chi_{\mathcal{G}}(x) = \chi_{G_n}(x)\chi_{H-v}(x) + \chi_{G_{n-1}}(x)\chi_H(x) - x\chi_{G_{n-1}}(x)\chi_{H-v}(x)\,. \tag{4.45}$$

If we say that G has g vertices and H has h vertices, we see that $\mathcal{G}$ has $g+h+n$ vertices. To find $R_{\mathcal{G}}$, we take reciprocalisation of $\chi_{\mathcal{G}}$; that is, we calculate $z^{\deg \mathcal{G}}\chi_{\mathcal{G}}(z+1/z)$:

$$R_{\mathcal{G}}(z) = zR_{G_n}(z)R_{H-v}(z) + zR_{G_{n-1}}(z)R_H(z) - (z+1/z)z^2R_{G_{n-1}}(z)R_{H-v}(z)\,.$$

We rearrange this into a slightly more convenient form:

$$R_{\mathcal{G}}(z) = \Big(zR_{G_n}(z) - z(z^2+1)R_{G_{n-1}}(z)\Big)R_{H-v}(z) + zR_{G_{n-1}}(z)R_H(z)\,. \tag{4.46}$$

We now recall (3.3), which gives us the shape of the reciprocal polynomial of a digraph with a pendant path (with m vertices) attached:

$$(z^2-1)R_m(z) = z^{2m}B(z) - B^*(z)\,,$$

where B is some monic integer polynomial that depends on G, but not m or the vertex the pendant path is adjoined to.

So, by multiplying (4.46) by (z^2-1), we can use (3.3) to simplify the shape of the reciprocal polynomial further. In doing this, we get:

$$\begin{aligned}(z^2-1)R_{\mathcal{G}}(z) &= \Big(z(z^2-1)R_{G_n}(z) - z(z^2+1)(z^2-1)R_{G_{n-1}}(z)\Big)R_{H-v}(z) \\ &\qquad + z(z^2-1)R_{G_{n-1}}(z)R_H(z) \\ &= \bigg(z\Big(z^{2n}B(z)-B^*(z)\Big) - z(z^2+1)\Big(z^{2(n-1)}B(z)-B^*(z)\Big)\bigg)R_{H-v}(z) \\ &\qquad + z\Big(z^{2(n-1)}B(z)-B^*(z)\Big)R_H(z) \\ &= \underbrace{\Big(z^{2n-1}R_H(z) - z^{2n-1}R_{H-v}(z)\Big)}_{=L_1(z)}B(z) \\ &\qquad\qquad + \underbrace{\Big(z^3R_{H-v}(z) - zR_H(z)\Big)}_{=L_0(z)}B^*(z)\,.\end{aligned}$$

We note that the monic polynomial B which comes from simplifying R_{G_n} and $R_{G_{n-1}}$ is indeed the same. In fact, we know from McKee and Smyth [22] – as noted in the proof of Lemma 3.3.2 – that $B(z) = R_{G_1}(z) - R_{G_0}(z)$.

This gives us the desired shape of the reciprocal polynomial. □

4.4.4 Key Result for Bridges

This now puts us in a position to state and prove our main result related to these bridged digraphs.

Theorem 4.4.9. *Let G be a digraph with $g \geq 1$ vertices and (v_1, v_2) be a list of two (not necessarily distinct) vertices of G. Construct a new digraph from G, called $\mathcal{G}$, in the following fashion:*

- *Attach an uncharged pendant path, with m_1 vertices, such that the final vertex – denoted by u_1 – is part of a new digraph H_L, with $\ell \geq 2$ vertices,*
- *Attach a second uncharged pendant path, with m_2 vertices, such that the final vertex – denoted by u_2 – is part of a new digraph H_R, with $r \geq 2$ vertices,*

So, $\mathcal{G}$ will have $g + l + r + m_1 + m_2 - 2$ vertices. Writing $R_\#$ to represent the reciprocal polynomial $R_\#(z)$ for a given digraph, we then have that the reciprocal polynomial of $\mathcal{G}$ is:

$$
\begin{aligned}
&(z^2-1)^2 R_{\mathcal{G}} = \\
&\qquad z^{2m_1+2m_2-4}\left(R_{H_L}R_{H_R} - R_{H_L}R_{H_R-u_2} - R_{H_L-u_1}R_{H_R} + R_{H_L-u_1}R_{H_R-u_2}\right) B_{1,1} \\
&\qquad -z^{2m_1-2}(R_{H_L}R_{H_R} - z^2 R_{H_L}R_{H_R-u_2} - R_{H_L-u_1}R_{H_R} + z^2 R_{H_L-u_1}R_{H_R-u_2})B_{1,0} \\
&\qquad -z^{2m_2-2}(R_{H_L}R_{H_R} - R_{H_L}R_{H_R-u_2} - z^2 R_{H_L-u_1}R_{H_R} + z^2 R_{H_L-u_1}R_{H_R-u_2})B_{0,1} \\
&\qquad + (R_{H_L}R_{H_R} - z^2 R_{H_L}R_{H_R-u_2} - z^2 R_{H_L-u_1}R_{H_R} + z^4 R_{H_L-u_1}R_{H_R-u_2})B_{0,0}\,,
\end{aligned}
\tag{4.47}
$$

where:

$$
\begin{aligned}
B_{1,1}(z) &= R_{G_{1,1}}(z) - R_{G_{1,0}}(z) - R_{G_{0,1}}(z) + R_{G_{0,0}}(z)\,, \\
B_{1,0}(z) &= R_{G_{1,1}}(z) - z^2 R_{G_{1,0}}(z) - R_{G_{0,1}}(z) + z^2 R_{G_{0,0}}(z)\,, \\
B_{0,1}(z) &= R_{G_{1,1}}(z) - R_{G_{1,0}}(z) - z^2 R_{G_{0,1}}(z) + z^2 R_{G_{0,0}}(z)\,; \\
B_{0,0}(z) &= R_{G_{1,1}}(z) - z^2 R_{G_{1,0}}(z) - z^2 R_{G_{0,1}}(z) + z^4 R_{G_{0,0}}(z)\,.
\end{aligned}
\tag{4.48}
$$

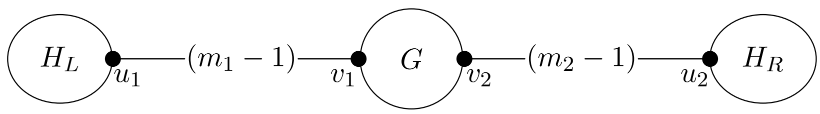


Figure 4.7: A visualisation of $\mathcal{G}$.

Remark. We note that the formulas in (4.48) are the same as those in (4.3). These have been restated under our more formal notation for absolute clarity.

Proof. Let $\mathcal{H} = (H_L){}_u^{m_1}{}_v G$. We can also think of this as $\mathcal{H} = H_L \cdot G_{m_1}$.

Then, $\mathcal{G} = \mathcal{H}_{v_2}^{m_2}{}_{u_2} H_R$. Again, we can view this as $\mathcal{G} = \mathcal{H}_{m_2} \cdot H_R$. Applying Lemma 4.4.6 to this, we get:

$$\chi_{\mathcal{G}} = \underbrace{\chi_{\mathcal{H}_{m_2}}\chi_{H_R-u_2}}_{(*)} + \underbrace{\chi_{\mathcal{H}_{m_2-1}}\chi_{H_R}}_{(**)} - x\,\underbrace{\chi_{\mathcal{H}_{m_2-1}}\chi_{H_R-u_2}}_{(***)} . \tag{4.49}$$

We expand each part of (4.49) separately. Firstly, we expand $(*)$, by applying Lemma 4.4.6 to $\mathcal{H}_{m_2} = H_L \cdot G_{m_1,m_2}$:

$$\chi_{\mathcal{H}_{m_2}}\chi_{H_R-u_2} = \Big(\chi_{G_{m_1,m_2}}\chi_{H_L-u_1} + \chi_{G_{m_1-1,m_2}}\chi_{H_L} - x\chi_{G_{m_1-1,m_2}}\chi_{H_L-u_1}\Big)\,\chi_{H_R-u_2} . \tag{4.50}$$

Next, we expand $(**)$ by applying Lemma 4.4.6 to $\mathcal{H}_{m_2-1}$:

$$\chi_{\mathcal{H}_{m_2-1}}\chi_{H_R} = \Big(\chi_{G_{m_1,m_2-1}}\chi_{H_L-u_1} + \chi_{G_{m_1-1,m_2-1}}\chi_{H_L} - x\chi_{G_{m_1-1,m_2-1}}\chi_{H_L-u_1}\Big)\,\chi_{H_R} . \tag{4.51}$$

Finally, we expand $(***)$ by applying Lemma 4.4.6 to $\mathcal{H}_{m_2-1}$:

$$\chi_{\mathcal{H}_{m_2-1}}\chi_{H_R-u_2} = \Big(\chi_{G_{m_1,m_2-1}}\chi_{H_L-u_1} + \chi_{G_{m_1-1,m_2-1}}\chi_{H_L} - x\chi_{G_{m_1-1,m_2-1}}\chi_{H_L-u_1}\Big)\,\chi_{H_R-u_2} . \tag{4.52}$$

Combining (4.50), (4.51) and (4.52) back together into (4.49) and rearranging gives us:

$$\begin{aligned}
\chi_{\mathcal{G}} = {} & \chi_{G_{m_1,m_2}}\chi_{H_L-u_1}\chi_{H_R-u_2} \\
& + \chi_{G_{m_1-1,m_2}}\left(\chi_{H_L}\chi_{H_R-u_2} - x\chi_{H_L-u_1}\chi_{H_R-u_2}\right) \\
& + \chi_{G_{m_1,m_2-1}}\left(\chi_{H_L-u_1}\chi_{H_R} - x\chi_{H_L-u_1}\chi_{H_R-u_2}\right) \\
& + \chi_{G_{m_1-1,m_2-1}}\left(\chi_{H_L}\chi_{H_R} - x\chi_{H_L-u_1}\chi_{H_R} - x\chi_{H_L}\chi_{H_R-u_2} + x^2\chi_{H_L-u_1}\chi_{H_R-u_2}\right) .
\end{aligned} \tag{4.53}$$

We now find the reciprocal polynomial, which we will write as $R_{\mathcal{G}}(z) = R_{\mathcal{G}}$.

$$
\begin{aligned}
R_{\mathcal{G}} &= R_{G_{m_1,m_2}} R_{H_L-u_1} R_{H_R-u_2} \\
&\quad + R_{G_{m_1-1,m_2}} \left(R_{H_L} R_{H_R-u_2} - z(z+1/z) R_{H_L-u_1} R_{H_R-u_2} \right) \\
&\quad + R_{G_{m_1,m_2-1}} \left(R_{H_L-u_1} R_{H_R} - z(z+1/z) R_{H_L-u_1} R_{H_R-u_2} \right) \\
&\quad + R_{G_{m_1-1,m_2-1}} (R_{H_L} R_{H_R} - z(z+1/z) R_{H_L-u_1} R_{H_R} \\
&\qquad\qquad - z(z+1/z) R_{H_L} R_{H_R-u_2} + z(z+1/z)^2 R_{H_L-u_1} R_{H_R-u_2}) \,.
\end{aligned} \tag{4.54}
$$

At this point, we now wish to exploit (4.2) (as presented in Lemma 4.2.1). Thus, we need to multiply (4.54) by $(z^2-1)^2$. From here, we can then expand each of the $(z^2-1)^2 R_{G_{i,j}}$ terms into a linear combination of $R_{1,1}$, $R_{1,0}$, $R_{0,1}$ and $R_{0,0}$.

Once we have done this, we can collect each of the $R_{1,1}$, $R_{1,0}$, $R_{0,1}$ and $R_{0,0}$ terms together. This gives us an equation of the form:

$$
(z^2-1)^2 R_{\mathcal{G}}(z) = \alpha_1(z) R_{1,1}(z) + \alpha_2(z) R_{1,0}(z) + \alpha_3(z) R_{0,1}(z) + \alpha_4(z) R_{0,0}(z) \,.
$$

We look at each of these α_i individually, for clarity:

$$
\begin{aligned}
\alpha_1(z) &= (z^{2m_1-2}-1)(z^{2m_2-2}-1) R_{H_L} R_{H_R} \\
&\quad - (z^{2m_1-2}-1)(z^{2m_2-2}-z^2) R_{H_L} R_{H_R-u_2} \\
&\quad - (z^{2m_1-2}-z^2)(z^{2m_2-2}-1) R_{H_L-u_1} R_{H_R} \\
&\quad + (z^{2m_1-2}-z^2)(z^{2m_2-2}-z^2) R_{H_L-u_1} R_{H_R-u_2} \,,
\end{aligned} \tag{4.55}
$$

$$
\begin{aligned}
\alpha_2(z) &= -(z^{2m_1-2}-1)(z^{2m_2-2}-z^2) R_{H_L} R_{H_R} \\
&\quad + (z^{2m_1-2}-1)(z^{2m_2-2}-z^4) R_{H_L} R_{H_R-u_2} \\
&\quad + (z^{2m_1-2}-z^2)(z^{2m_2-2}-z^2) R_{H_L-u_1} R_{H_R} \\
&\quad - (z^{2m_1-2}-z^2)(z^{2m_2-2}-z^4) R_{H_L-u_1} R_{H_R-u_2} \,,
\end{aligned} \tag{4.56}
$$

$$\begin{aligned}\alpha_3(z) &= -(z^{2m_1-2} - z^2)(z^{2m_2-2} - 1)R_{H_L}R_{H_R} \\ &+ (z^{2m_1-2} - z^2)(z^{2m_2-2} - z^2)R_{H_L}R_{H_R-u_2} \\ &+ (z^{2m_1-2} - z^4)(z^{2m_2-2} - 1)R_{H_L-u_1}R_{H_R} \\ &- (z^{2m_1-2} - z^4)(z^{2m_2-2} - z^2)R_{H_L-u_1}R_{H_R-u_2}\,,\end{aligned} \tag{4.57}$$

$$\begin{aligned}\alpha_4(z) &= (z^{2m_1-2} - z^2)(z^{2m_2-2} - z^2)R_{H_L}R_{H_R} \\ &- (z^{2m_1-2} - z^2)(z^{2m_2-2} - z^4)R_{H_L}R_{H_R-u_2} \\ &- (z^{2m_1-2} - z^4)(z^{2m_2-2} - z^2)R_{H_L-u_1}R_{H_R} \\ &+ (z^{2m_1-2} - z^4)(z^{2m_2-2} - z^4)R_{H_L-u_1}R_{H_R-u_2}\,.\end{aligned} \tag{4.58}$$

We can now combine (4.55)–(4.58) and apply (4.3) to get:

$$\begin{aligned}(z^2-1)^2 R_{\mathcal{G}} &= \left(z^{2m_1+2m_2-4}B_{1,1} - z^{2m_1-2}B_{1,0} - z^{2m_2-2}B_{0,1} + B_{0,0}\right) R_{H_L}R_{H_R} \\ &- \left(z^{2m_1+2m_2-4}B_{1,1} - z^{2m_1}B_{1,0} - z^{2m_2-2}B_{0,1} + z^2B_{0,0}\right) R_{H_L}R_{H_R-u_2} \\ &- \left(z^{2m_1+2m_2-4}B_{1,1} - z^{2m_1-2}B_{1,0} - z^{2m_2}B_{0,1} + z^2B_{0,0}\right) R_{H_L-u_1}R_{H_R} \\ &- \left(z^{2m_1+2m_2-4}B_{1,1} - z^{2m_1}B_{1,0} - z^{2m_2}B_{0,1} + z^4B_{0,0}\right) R_{H_L-u_1}R_{H_R-u_2}\,.\end{aligned} \tag{4.59}$$

We can rearrange this to get (4.47), our desired result. □

Remark. An alternative way to write (4.47) would be as:

$$(z^2-1)^2 R_{\mathcal{G}} = z^{2m_1+2m_2-4}B_{1,1}C_{1,1} - z^{2m_1-2}B_{1,0}C_{0,1} - z^{2m_2-2}B_{0,1}C_{1,0} + B_{0,0}C_{0,0}\,, \tag{4.60}$$

where the $B_{i,j}$ are as in (4.48) and

$$\begin{aligned}C_{1,1} &= R_{H_L}R_{H_R} - R_{H_L}R_{H_R-y} - R_{H_L-u}R_{H_R} + R_{H_L-u}R_{H_R-y}\,, \\ C_{0,1} &= R_{H_L}R_{H_R} - z^2R_{H_L}R_{H_R-y} - R_{H_L-u}R_{H_R} + z^2R_{H_L-u}R_{H_R-y}\,, \\ C_{1,0} &= R_{H_L}R_{H_R} - R_{H_L}R_{H_R-y} - z^2R_{H_L-u}R_{H_R} + z^2R_{H_L-u}R_{H_R-y}\,; \\ C_{0,0} &= R_{H_L}R_{H_R} - z^2R_{H_L}R_{H_R-y} - z^2R_{H_L-u}R_{H_R} + z^4R_{H_L-u}R_{H_R-y}\,.\end{aligned} \tag{4.61}$$

4.4.5 A Bridge Too Far?

Now that we have the shape of the reciprocal polynomial, (4.47), of these bridged digraphs, we can now include digraphs of these shape in our experiments. At least, in theory we can.

There is an unfortunate issue with the shape of the reciprocal polynomial in (4.47): it is very complicated. In computing this, we need to create a routine which handles a lot of data. To summarize, the routine has to have the following features:

i. Allows us to input details of the starting digraph G, as well as choose suitable vertices v and w,

ii. Can vary over different suitable digraphs H_L and H_R, and choose appropriate vertices u and y,

iii. Calculate the reciprocal polynomials of eight different digraphs simultaneously,

iv. Is not too time consuming.

Parts ii and iv are large hurdles for us. We are limited with our ability to create a routine which can run over suitable digraphs and choose appropriate vertices. To overcome this, we could go for a watered-down version, where we as the user choose and input details of specific H_R and H_L, but this is perhaps less practical.

Furthermore, if we were to do this, we still have the issue of time. We will address this more in chapter 5, once we have outlined how our experiments work in detail. However, in short, whilst creating a routine that can calculate these reciprocal polynomials in a reasonable amount of time is achievable, we have to consider two further factors. The first factor is the time taken to complete the rest of our experiment. The second factor is the comparative time taken to calculate these polynomials compared to reciprocal polynomials where we just attach pendant paths with some decoration at the end.

4.5 Consequences and Next Steps

This chapter has demonstrated an effective way of growing graphs, which can indeed be applied to our digraphs. We have explored ways to increase the size (number of vertices) of our digraphs by growing a path of uncharged vertices, either internally (by means of subdividing an edge) or externally (by means of attaching a pendant path). In the latter scenario, we also have the option to attach a small decoration at the end of the path. In each scenario, we know what the shape of the reciprocal polynomial will be.

We have also extended the idea even further, by means of attaching new digraphs at the ends of pendant paths. Heuristically, this is a step removed from our original idea. Furthermore, from a practical perspective, this idea is not one which seems worthwhile to experiment with, However, we have at least been able to build a more comprehensive theory around growing digraphs.

Now, we are in a position to experiment and attempt to find digraphs with small Mahler measure. Again, we go back to our key idea: if a digraph is combinatorially close to being cyclotomic, we hope that its Mahler measure will be close to being 1, and hence small. We can obtain digraphs which are combinatorially close to being cyclotomic in one of two ways. The first way is to take a cyclotomic digraph, and then 'nudge' it (by making a small number of changes or additions), breaking its cyclotomic nature. The second way is to take a non-cyclotomic digraph, and grow it – using the methods outlined here – so the resulting digraph resembles, in a combinatorial sense, a digraph which is close to being cyclotomic. In other words, it could be a cyclotomic digraph which had been nudged.

In the next chapter, we formalise these ideas, including making decisions on which cyclotomic digraphs we should nudge, and which non-cyclotomic digraphs make good candidates to grow.

Chapter 5

Experiments and Results

In this chapter, we explain in detail the experiments we ran to attempt to find small Mahler measures from digraphs, as well as the results thereafter. In fact, we have two key focuses here: the first is digraphs with small Mahler measures, the second is associating two-variable polynomials to certain digraphs, and then attempting to find small Mahler measures of two-variable polynomials from families of digraphs. We will give details of how this was achieved after looking at the single variable case.

As alluded to in Section 4.5, our experiments will require us to find good candidates of digraphs to either 'nudge' or 'grow'. So, before detailing our experiments, we first find such good candidates.

The experiments described here have appeared in a less detailed format in Coyston and McKee [9], with the final results also being included there.

5.1 Good Candidates

5.1.1 Nudged Paths

We know that the path P_n is cyclotomic and, in some sense, is perhaps the most straightforward family of cyclotomic (di)graphs to consider. As such, these are ideal candidates to first consider for our experiments.

We have mentioned previously our desire to 'nudge' digraphs, by which we mean, making a small number of changes or additions to it. Before continuing, we formalise this concept a bit more:

Definition 5.1.1. *A **nudge** to a digraph G is the following process:*

- *Choose a value n, much less than the number of vertices of G,*
- *Identify vertices $v_1, \cdots, v_n$ in G,*
- *Attach a new, possibly charged, vertex, say u, to these vertices, with some signed arcs or edges.*

The reasoning for calling this a nudge is somewhat intuitive: we are attaching one new vertex to a digraph with a small number of arcs, compared to the size of the digraph. By choosing n to be much less than the number of vertices, this gives us a certain level of confidence that we are not changing the digraph too much, and in a combinatorial sense, just nudging it to become something different.

Example. Figure 5.1 shows P_{30} nudged by attaching an uncharged vertex to three vertices: the fifth, thirteenth and last. The new vertex u is attached by signed arcs $s_{1,1}, \cdots, s_{3,2}$, some of which may be 0, but not both $s_{i,1}$ and $s_{i,2}$ for any given i.

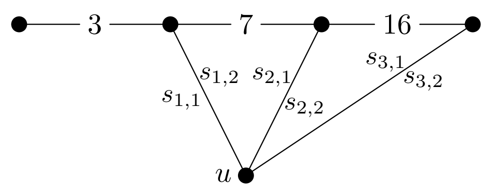
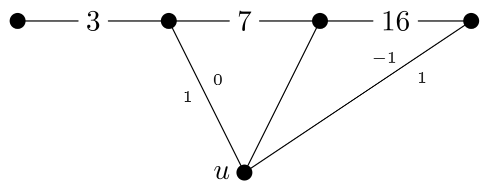


Figure 5.1: P_{30} after it has been nudged.

It is now worth refocusing: our aim with nudging digraphs is to help us find digraphs with small Mahler measures. We know that we *need* to be looking at digraphs which are not simple graphs too. That is to say, we need to be looking for digraphs which have charged vertices and signed arcs or edges. As such, when we nudge P_n, we should consider adding a charged vertex and, more importantly, signed arcs. Indeed, it would be best if at least one one of the vertices v_i was not joined to u by a signed edge, but only with a signed arc in one direction (so that the resulting adjacency matrix is non-symmetric). For example, the following would be a sensible choice of nudge to P_{30} consider:

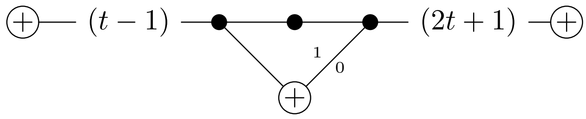


.

Now again, we reconsider our aim of nudging digraphs is to help us find digraphs with small Mahler measures. When we nudge P_n, we want – in a combinatorial sense – to keep it close to being cyclotomic, whilst ensuring it is no longer cyclotomic. Therefore, it is sensible to attach the new vertex to as few of the existing vertices as possible. As such, we restrict ourselves to nudging by attaching the new vertex to at most $n = 2$ vertices of the path.

Remark. The process of nudging can be applied broadly. Though restricting ourselves to attaching a new vertex to at most two existing vertices when nudging a path is the sensible choice here, this is not necessarily the optimal value of n for a general digraph.

Example 5.1.2. Consider the following family of digraphs, which we label as $G^{++}_{t,2t+2}$:



We note that we are calling this digraph $G^{++}_{t,2t+2}$, but our pendant paths are listed as a having $t-1$ and $2t+1$ vertices in our visual. This is because the labelling in the visual shows the number of neutral vertices in our pendant path, and does not include the endvertex. We choose to name our digraph in this way to ensure we are including the decoration of our pendant path.

Then, for various values of t, we get digraphs which have small Mahler measures:

t	$M(G^{++}_{t,2t+2})$
2	$1.24072642\cdots$
3	$1.20261674\cdots$
4	$1.17628081\cdots$
5	$1.26911783\cdots$
6	$1.27829231\cdots$
7	$1.27410221\cdots$
8	$1.26552733\cdots$
9	$1.25555243\cdots$
10	$1.24537290\cdots$

Table 5.1: Various values of t and the resulting Mahler measure of $G^{++}_{t,2t+2}$.

Remark. In fact, as $t \to \infty$, the value of $M(G^{++}_{t,2t+2})$ will tend to a limit point. We discuss this fact later. This also highlights the oscillating nature of Mahler measures of digraphs we have grown, which we noted in Section 3.4.

This example exhibits just why the path is a good candidate to nudge. Figure 5.2 gives us a look at a more general version of the suitably nudged path:

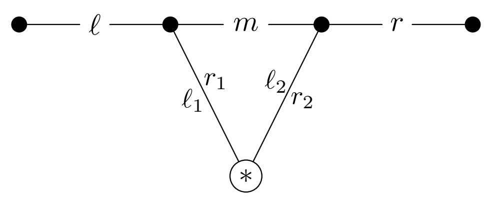


Figure 5.2: A good general example of a nudged path.

We see that we have a path on $\ell + m + r + 4$ vertices, which is being nudged by attaching the new vertex to the $(\ell+2)$-th vertex and the $(\ell+m+3)$-th vertex, with some signed arcs ℓ_1, r_1, ℓ_2 and r_2. This is a particularly fruitful-looking nudge to consider, and forms the heart of our experiments.

5.1.2 Core Digraphs and Decorations

Nudging is primarily useful for us when we have a cyclotomic digraph. We are taking a digraph with Mahler measure 1 and breaking its combinatorial cyclotomic nature slightly, in the hope of getting a new digraph which has small Mahler measure. Whilst this looks fruitful (and, as we will shortly see, is indeed fruitful!), it does restrict us and the number of experiments we could perform.

Therefore, we look to 'reverse-engineer' the process. So, instead of taking a cyclotomic digraph and nudging it, find a good choice of digraph and attach a suitable cyclotomic digraph (or digraphs) to it. If the attached cyclotomic digraphs are comparatively significantly larger than our original, then we effectively end up with a new digraph which is close to be cyclotomic.

Of course, the main problem which arises here is what is "good choice" of digraph, and what are "suitable" cyclotomic digraphs to attach. We first tackle the former of these problems before moving onto the latter, but fortunately, we already have a good basis to build on to answer these.

Recall Figure 5.2, where we attached one vertex to two vertices of a path. If we set $\ell = r = -1$, we have the specific case of a nudged path, whereby we have attached the vertex to the first and last vertices of a path, as seen in Figure 5.3.

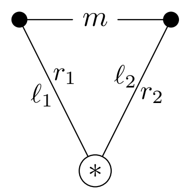


Figure 5.3: A nudged path, with a new vertex attached to both ends of the path.

Now, if instead of nudging the path by attaching one vertex, consider the situation where we add two vertices, u_1 and u_2, joined by a negative edge, and attach u_1 to the first vertex of the path, and u_2 to the last vertex of the path. This results in the following new digraph, seen in Figure 5.4:

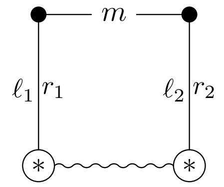


Figure 5.4: A nudged path with two new vertices.

This is a step away from our idea of nudging paths, in that we are adding more than one new vertex. But at the same time, it is very much within the spirit of the idea. Nudging makes a small number of changes to a digraph to give us a new digraph. So, assuming that the path we are nudging has a large number of vertices, adding two new vertices, and a maximum of six new signed arcs (treating an edge as two arcs) is still making a small number of changes.

More aptly, this new digraph is still very close to being cyclotomic. To see this, we need the following Proposition:

Proposition 5.1.3. *Let G be the disjoint union of graphs $G_1, \cdots, G_k$. Then, the Mahler measure of G is:*

$$M(G) = \prod_{i=1}^{k} M(G_i)\,.$$

Proof. Let the adjacency matrix of G be A. Labelling the vertices of G appropriately, we can write A in the following block diagonal form:

$$A = \begin{pmatrix} A_1 & 0 & & & \\ 0 & A_2 & 0 & & \\ & 0 & \ddots & \ddots & \\ & & \ddots & A_{k-1} & 0 \\ & & & 0 & A_k \end{pmatrix},$$

where each A_i is the adjacency matrix of G_i. As A has this block diagonal form, our result follows straightforwardly from here. □

An easy consequence of Proposition 5.1.3 is that the disjoint union of cyclotomic graphs, and indeed disjoint union of cyclotomic digraphs, is itself cyclotomic. If we return to Figure 5.4, we see that if we remove the arcs joining the new vertices to the original path, we are left with (digraphs which are equivalent to) charged paths. Hence, this would be the disjoint union of cyclotomic digraphs, meaning the digraph considered in Figure 5.4 is combinatorially close to being cyclotomic.

This now gives us a good choice of digraph to reverse-engineer our nudging process. If we take the digraphs shown in Figure 5.4 and attach pendant paths to the ends of the original path, using the methods as described in Section 4.2, we get a resulting digraph which is close to being cyclotomic (as seen in Figure 5.5). Furthermore, we have an additional bonus in that we have a method to calculate its reciprocal polynomial (and hence Mahler measure) already our disposal!

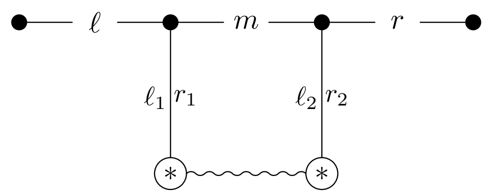


Figure 5.5: A new digraph worth experimenting with.

In fact, we can use similar methods to find many good choices of digraphs to experiment with. Before stating a result which captures all of these good choices, we attempt to firm up what exactly we mean by a “good choice”:

(Informal) Definition 5.1.4. *We say a digraph is a **good choice** of digraph for our reverse-engineered process if it is a digraph which is combinatorially close to being cyclotomic. This itself is not well-defined, so we go with visual evidence here to decide is a digraph is combinatorially close to being cyclotomic.*

Remark. As the concept of “good choice” is not well-defined, we cannot properly formalise our result about which digraphs are indeed good choices. As such, the following statement is referred to as a “Heuristic Proposition”, as opposed to a Proposition.

Heuristic Proposition 5.1.5. *The digraphs CD_1, CD_2 and CD_3, as showcased in Figures 5.6 and 5.7, give us particularly good choices of digraph for our reverse-engineered nudging process:*

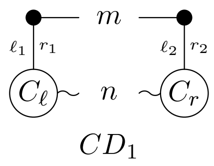


Figure 5.6: The digraph CD_1.

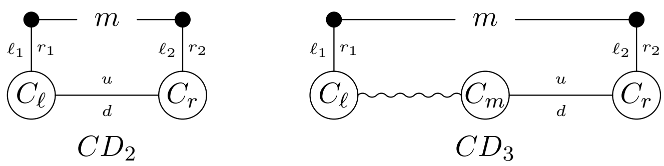


Figure 5.7: The digraphs CD_2 and CD_3.

Owing to the lack of precision in the statement of Heuristic Proposition 5.1.5, we cannot formally prove it. Furthermore, we note that not all combinations are good choices, but most are. As such, we give a justification for why it is true, and make this justification as formal as possible.

Justification of Heuristic Proposition 5.1.5. We look at each digraph on a case-by-case basis after removing the signed arcs ℓ_1, r_1, ℓ_2 and r_2. Our aim is to show that, after removing these arcs, each digraph in the resulting disjoint is cyclotomic in most (or all) cases.

In each case, we have two paths, of which the first is an uncharged path on $(m+2)$ vertices, which we know is cyclotomic. So we look exclusively at the second path.

The case of CD_1 is not difficult to tackle. The second path is a path on $(n+2)$ vertices, where the ends are charged, and $(n+1)$ negative edges. By Proposition 3.5.7, we know that this second path is switch equivalent to $P_n^{C_\ell C_r}$, which we know is cyclotomic.

The remaining two cases are not so trivial. The second path of CD_2 has adjacency matrix $\begin{pmatrix} c_\ell & u \\ d & c_r \end{pmatrix}$. We find that, in most cases that the values of these take, the eigenvalues of this matrix always lie in $[-2, 2]$. So, by Lemma 3.1.11, CD_2 is cyclotomic in most cases.

We can use a similar exhaustive approach to show the same for CD_3, again noting that some choice of values give us a non-cyclotomic digraph, but most cases do. □

Remark. We can choose to omit the values which give us digraphs which are not "good choices" in our experiments.

This now gives us a variety of good candidates to experiment with growing and attaching cyclotomic digraphs to. We call these candidates **core digraphs**.

We have now answered the first of our two main problems we encountered prior to experimenting; namely what is a "good choice" of digraph to start with. We now have to tackle the second problem: what are "suitable" cyclotomic digraphs to attach to our core digraphs?

This is, in fact, much easier to answer. In Section 4.2, we attached pendant paths to some digraph, and the ends of these pendant paths had some form of *decoration*. In particular, they had some (possibly neutral) charge, or had a "fork" from a snake tongue at the end. All of these are cyclotomic digraphs, and so are suitable choices for cyclotomic digraphs to attach.

Furthermore, as we saw in Section 4.4, attaching more general digraphs at the end of our pendant paths creates significantly more complex results for the shape of the reciprocal polynomial. As such, we cannot work with these reciprocal polynomials as quickly as we would perhaps want to. We discuss the issues related to this in slightly more detail in Section 5.4.

As such, whilst other cyclotomic digraphs may prove useful in theory, they are harder to implement in practice. This gives us a short, but still useful, list of suitable cyclotomic digraphs which are worth attaching to our digraphs.

5.1.3 The Heart of the Experiments

We now give an outline of what we wish to achieve with our experiments of using digraphs to find small Mahler measures. We do not detail how our experiments will work – this is done in Section 5.2 – but rather overview what we have covered in a simple, closed fashion.

The specific digraphs we will be looking at will be in the form of Figure 5.8, where:

- D_L, D_R are decorations in the form of one vertex of some possibly neutral charge, or the snake tongue,
- CD is a core digraph (which includes the internal path with m vertices) as seen in Heuristic Proposition 5.1.5.

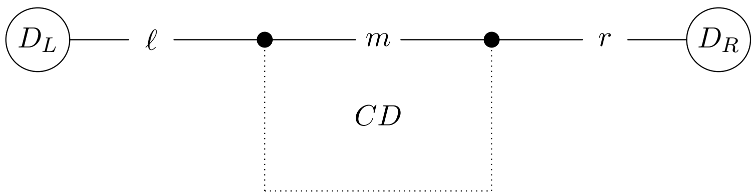


Figure 5.8: The general digraph experimented on.

Of course, within this general digraph are several different parameters. Some of these parameters do only arise in certain choice of core digraphs, but we have:

- The number of vertices in our pendant paths, ℓ and r, which must be non-negative,
- The number of vertices in our main path of the core digraph, m, which must be non-negative,
- The signs of our arcs within the core digraph, ℓ_1, r_1, ℓ_2 and r_2, which can take values in $\{-1, 0, 1\}$, and cannot all be zero,
- The charges of our main vertices in our core digraph, C_ℓ and C_r, which have corresponding values c_ℓ and c_r in $\{-1, 0, 1\}$,

- If our core digraph is CD_1 (as seen in Figure 5.6), the number of vertices in our negative path of the core digraph, n, which must be at least -1,
- If our core digraph is CD_2 or CD_3 (as seen in Figure 5.7), the signs of our arcs within these core digraphs, u and d, which can take values in $\{-1, 0, 1\}$,
- If our core digraph is CD_3 (as seen in Figure 5.7), the charge of our additional vertex, C_m, whose corresponding value is c_m is in $\{-1, 0, 1\}$.

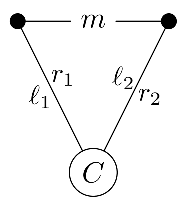
Remark. We note that in the case of CD_1, we allow choices of $n \geq -1$. In the case $n = -1$, we identify the charged vertices C_ℓ and C_r to be the same vertex.

Unsurprisingly, this means that our experiments for finding digraphs with small Mahler measures are extremely involved. This is despite the fact we have already made very careful choices to ensure we are looking at a smaller selection of digraphs which we are hopeful will be fruitful. Our next step is to determine how to make our experiments as optimal as possible. With so many parameters (some of which have no bound), we need to ensure that our experiments run efficiently, know when to continue and when to stop.

5.1.4 Reducing Choices

One way to make our experiments run as efficiently as possible is to ensure we are not experimenting on digraphs which are equivalent. In particular, in Corollary 3.5.4.1 we saw that digraphs which are switch equivalent share the same Mahler measure. As such, if we have two core digraphs which are equivalent, it stands to reason that we do not need to experiment on both of them.

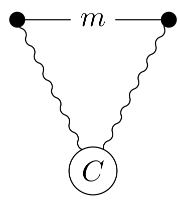
We will not detail in great depth how to find all non-equivalent core digraphs, nor will we give any sort of comprehensive lists. Instead, we will look at a specific example of a certain family of core digraphs, and outline how we find all non-equivalent digraphs in this case. The specific case we look at is the same digraph we saw in Figure 5.3, where we nudged a path by attaching a charged vertex to the ends of the path. This is one of the simplest core digraphs, and we show it again here for convenience:



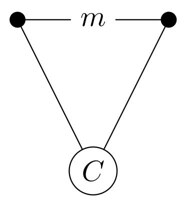
.

This is CD_1, with $n = -1$, meaning that the charged vertices C_ℓ and C_r are identified as the same vertex.

If we fix m, we have 3^5 possible digraphs from our parameters (ℓ_1, r_1, ℓ_2, r_2 and c all have three possible values). We can reduce the number of digraphs we consider immediately by ignoring situations where $\ell_1 = r_1 = 0$ and $\ell_2 = r_2 = 0$, since we want to attach the charged vertex to two vertices. It is straightforward to find digraphs which are obviously equivalent. For example,



and



are equivalent, by switching at the charged vertex.

Table 5.2 showcases the values ℓ_1, r_1, ℓ_2, r_2 for all non-equivalent digraphs, with a fixed charge C in this case. There are 13 different combinations of parameters here. This means that there are a total of $13 \times 3 = 36$ non-equivalent relevant digraphs in the shape of Figure 5.3, which represents a not-insignificant decrease from our original 81.

In this case, there was a reasonably manageable number of unique digraphs. However, to demonstrate the usefulness of this, it is better to look at one of our other core digraphs.

The core digraph CD_2, again for a fixed m, has eight different parameters, and so 3^8 unique digraphs to consider (although this number can again be immediately reduced, after we ignore combinations of cases when $\ell_1 = r_1 = 0$, $\ell_2 = r_2 = 0$ and $u = d = 0$, and combinations which are not good choices). However, there are only 645 non-equivalent relevant digraphs, which represents a significant decrease from our original number.

ℓ_1	r_1	ℓ_2	r_2
-1	-1	-1	-1
-1	-1	-1	0
-1	-1	-1	1
-1	-1	0	1
-1	-1	1	1
-1	0	-1	0
-1	0	-1	1
-1	0	0	-1
-1	0	0	1
-1	0	1	-1
-1	0	1	0
-1	1	-1	1
-1	1	1	-1

Table 5.2: Values of ℓ_1, r_1, ℓ_2, r_2 for non-equivalent digraphs in Figure 5.3.

Using our knowledge of equivalence of digraphs, we have been able to reduce the number of good candidates that we have to experiment with. However, we still have a significant number of candidates, on top of the adding complexity of growing our digraphs. In any case, we are now in the best possible position to move onto our experiments, which we now outline.

5.2 Outline of Experiments

We now turn our attention to the exact nature of how our experiments were run. To reiterate, our overall goal is to find digraphs with small Mahler measures. Due to Theorem 4.2.14, we know what the general shape of the reciprocal polynomial is for the general digraph we are experimenting on. As such, the key idea is to vary over different values for each of our parameters, as outlined in Section 5.1.3.

However, before we approach the question of how far we should stretch our parameters, we need to know what exactly to do with good candidate polynomials once we have them. This is because if we have a good candidate polynomial (that is to say, we have a good choice of parameters in our digraph which give us a good reciprocal polynomial), there is a good chance that digraphs with similar parameters will also be fruitful to consider. Therefore, we need to determine a way to not only just find digraphs, but also push our searches further for potentially good families of digraphs.

5.2.1 An Aside on Hyperplanes

Here we take a short, and somewhat technical, diversion and look a little into hyperplanes. At first, this may seem like an obtuse diversion to take, especially in the context of what we are trying to achieve with our experiments, but this will allow us to create and run experiments which are as fruitful and optimal as possible.

Definition 5.2.1. *A* ***hyperplane*** *is an affine subspace of* $\mathbb{R}^n$ *consisting of all vectors* $\boldsymbol{x} = (x_1, \cdots, x_n)$ *such that* $a_1x_1 + \cdots + a_nx_n = c$*, for some constant* c *and for scalars* $a_1, \cdots, a_n$*, not all zero.*

One way to develop a link between hyperplanes and Mahler measures comes from the following result of Lawton:

Theorem 5.2.2 (Lawton, [16])**.** *For* $\mathbf{n} = (n_1, \cdots, n_r) \in \mathbb{Z}^r$*, define*

$$\mu(\mathbf{n}) = \min_{\substack{\mathbf{c}=(c_1,\cdots,c_r)\in\mathbb{Z}^r\setminus\{\mathbf{0}\},\\ \mathbf{c}\cdot\mathbf{n}=0}} \max_{1\le i\le r} |c_i| \, . \tag{5.1}$$

Then, for sequences $\mathbf{n}_i = (n_{i,1}, \cdots, n_{i,r}) \in \mathbb{Z}^r$ *for which* $\mu(\mathbf{n}_i) \to \infty$ *as* $i \to \infty$*,*

$$M(Q(z^{n_{i,1}}, \cdots, z^{n_{i,r}})) \to M(Q(z_1, \cdots, z_r)) \, ,$$

as $i \to \infty$*.*

The link to hyperplanes might not be immediately obvious. Firstly, we take some $\mathbf{c} = (c_1 : c_2 : \cdots : c_r) \in \mathbb{P}^{r-1}(\mathbb{Q})$. Scaling $\mathbf{c}$, we can suppose that all the c_i are in $\mathbb{Z}$, and their greatest common divisor is 1. Then the **height** of $\mathbf{c}$ is defined to be the maximum of the $|c_i|$. Such points $\mathbf{c}$ correspond precisely to hyperplanes through the origin in $\mathbb{Q}^r$, with the hyperplane corresponding to $\mathbf{c}$ being the set of points $\mathbf{x}$ satisfying $\mathbf{c} \cdot \mathbf{x} = 0$. We define the **height** of the hyperplane to be the height of the corresponding point $\mathbf{c}$.

One way to interpret $\mu(\mathbf{n})$ in (5.1) is as the minimum height of a hyperplane through the origin that contains the point $\mathbf{n}$ (viewing $\mathbb{Z}^r$ as a subset of the vector space $\mathbb{Q}^r$). The condition $\mu(\mathbf{n}_i) \to \infty$ in Theorem 5.2.2 then immediately translates as the following Lemma:

Lemma 5.2.3. *Let* $(\mathbf{n}_i)$ *be a sequence of elements of* $\mathbb{Z}^r$*. Then* $\mu(\mathbf{n}_i) \to \infty$ *as* $i \to \infty$ *if and only if there is no hyperplane in* $\mathbb{Q}^r$ *through the origin (i.e., no* $(r-1)$*-dimensional subspace of* $\mathbb{Q}^r$*) containing infinitely many terms of the sequence.*

Proof. Suppose that for some constant $\mathbf{c} = (c_1, \cdots, c_r) \neq \mathbf{0}$ the hyperplane $\mathbf{c} \cdot \mathbf{x} = 0$ contains $\mathbf{n}_i$ for infinitely many i, say for $i = i_1 < i_2 < \cdots$. Clearing denominators, we can suppose that $\mathbf{c} \in \mathbb{Z}^r$. Then $\mu(\mathbf{n}_{i_j}) \leq \max_{1 \leq k \leq r} |c_k|$ for all j, and so $\mu(\mathbf{n}_{i_j}) \not\to \infty$ as $j \to \infty$, and $\mu(\mathbf{n}_i) \not\to \infty$ as $i \to \infty$.

Conversely, if $\mu(\mathbf{n}_i)$ has a bounded subsequence, then infinitely many of the $\mathbf{n}_i$ lie on a finite set of hyperplanes, and so there is some hyperplane containing infinitely many of the $\mathbf{n}_i$. $\square$

Proposition 1.2.4 tells us that a special case of Theorem 5.2.2 exists because of Boyd [3]:

$$\lim_{n \to \infty} M(P(z, z^n)) = M(P(x, y)) \,. \tag{5.2}$$

Here the points $(1, n)$ lie on a line, but not on a line through the origin.

We are now able to, loosely, explain how our experiments work now. And, in fact, the following approach is suitable for searching not just for small Mahler measures of single variable polynomials, but also for limits of sequences of such. In other words, we can also find Mahler measures of two-variable polynomials as well. We give the full details shortly afterwards.

<u>Method:</u>

- Find a good $(r+1)$-variable polynomial $Q(z_0, z_1, \cdots, z_r) \in \mathbb{Z}[z_0, \cdots, z_r]$.

- For each of a selection of good hyperplanes through the origin in $\mathbb{Q}^{r+1}$ (or for the intersection of several), choose a sequence of $\mathbf{n}_i = (1, n_{i,1}, \cdots, n_{i,r})$ on the hyperplane (or the intersection) and consider:
  (i) $M(Q(z, z^{n_{i,1}}, \cdots, z^{n_{i,r}}))$ for small Mahler measures;
  (ii) $\lim_{i\to\infty} M(Q(z, z^{n_{i,1}}, \cdots, z^{n_{i,r}}))$ for small limits.

In practice, because of the way we are finding our polynomials, we are choosing the $\mathbf{n}_i$ in such a way that the limit point can be written as a two-variable Mahler measure, as in (5.2), which can then be computed numerically reasonably quickly.

We are now at a point where we know how to get our good candidate polynomials (which we have labelled Q here), and we know what to do once we have a "good sequence", $\mathbf{n}_i$, from a good hyperplane. What we do not know, however, is *how* to find these good sequences once we have our good polynomials. We will outline the strategy to do so next, which will then finally put us in a position to carry out our experiments.

5.2.2 The Search for Good Lines on Planes

In this Section, we are assuming we have found a digraph which admits a good candidate polynomial. Our aim is understand how we can find a sequence $\mathbf{n}_i$ from a hyperplane associated to this polynomial to potentially find more good polynomials and digraphs.

Let $Q(z_0, \cdots, z_r) \in \mathbb{Z}[z_0, \cdots, z_r]$ be a candidate polynomial. For $n_1 \cdots, n_r$ positive integers, and putting $\mathbf{n} = (n_1, \cdots, n_r)$, we define:

$$P_{\mathbf{n}}(z) = Q(z, z^{n_1}, \cdots, z^{n_r}) \,.$$

Take a sequence of $\mathbf{n}_i = (n_{1,i}, \cdots, n_{r,i})$ with all the $n_{j,i}$ positive integers. Provided there is no hyperplane through the origin containing infinitely many of the points $(1, n_{1,i}, \cdots, n_{r,i})$, Theorem 5.2.2 tells us that:

$$\lim_{i\to\infty} M(P_{\mathbf{n}_i}) = M(Q(z_0, \cdots, z_r)) \,.$$

We shall deliberately break the constraints of this Theorem by choosing points that lie on hyperplanes through the origin. As we are seeking small Mahler measures, it is reasonable for us to seek Q for which the $(r+1)$-variable Mahler measure is not large: although we avoid the generic limit, it would seem more surprising to find small Mahler measures if the generic limit were large.

For fixed integers a_1, b_1, a_2, $b_2, \cdots, a_r$, b_r, the Laurent polynomial

$$Q(z, z^{a_1 t+b_1}, z^{a_2 t+b_2}, \cdots, z^{a_r t+b_r})$$

can be written as a Laurent polynomial in z and z^t, say:

$$Q(z, z^{a_1 t+b_1}, z^{a_2 t+b_2}, \cdots, z^{a_r t+b_r}) = P_{\mathbf{a},\mathbf{b}}(z, z^t)\,,$$

where $\mathbf{a} = (a_1, \cdots, a_r)$ and $\mathbf{b} = (b_1, \cdots, b_r)$. In practice we will take $a_i \geq 0$ for all i, and if $a_i = 0$ then $b_i \geq 0$, but there is no need in theory for such restrictions. Note that for $r \geq 2$ and $t \in \mathbb{Z}$ the line of points

$$(n_0, \cdots, n_r)_t = (1, a_1 t + b_1, \cdots, a_r t + b_r)$$

lie on a plane through the origin in $\mathbb{Q}^{r+1}$, namely the intersection of the $r-1$ hyperplanes

$$(-a_i b_1 + a_1 b_i)n_0 + a_i n_1 - a_1 n_i = 0\,, \tag{5.3}$$

for $2 \leq i \leq r$.

Thus we are not constrained by the generic limit in Theorem 5.2.2, and, by using (5.2), the limits

$$\lim_{t\to\infty} M(P_{\mathbf{a},\mathbf{b}}(z, z^t)) = M(P_{\mathbf{a},\mathbf{b}}(x, y)) \tag{5.4}$$

might be different for different choices of $\mathbf{a}$ and $\mathbf{b}$ (and indeed generally are different). From a single polynomial Q we have the (somewhat surprising!) prospect of finding more than one small limit.

It is now reasonable to ask how we are going to arrange this search. One option would be to search over pairs $(\mathbf{a}, \mathbf{b})$ of bounded height. However, this can take a long time as r grows. Our alternative, which is less painful as r grows, is the following three phase strategy:

Phase 1: Searching for Points

1.1 Choose a bound H for the heights of points, and a bound B for the largest Mahler measure to be tolerated.

  1.1.i When searching for small Mahler measures, we might chose $B = 1.3$,

  1.1.ii When searching for small limits, we might choose $B = 1.38$.

1.2 Loop over the r-dimensional box $0 \leq n_i \leq H$, $1 \leq i \leq r$, and store points $(n_1, \cdots, n_r)$ for which $M(Q(z, z^{n_1}, \cdots, z^{n_r})) \leq B$.

Phase 2: Searching for Lines

2.1 For each triple of points found in Phase 1, see if they lie on a line in $\mathbb{Q}^r$.

2.2 If so, write this line in canonical form $\mathbf{a}t + \mathbf{b}$, where $\mathbf{a}$, $\mathbf{b} \in \mathbb{Z}^r$, the content of $\mathbf{a}$ is 1 and the first nonzero entry in $\mathbf{a}$ is positive.

2.3 Store distinct lines found.

Phase 3: Check the Lines

3.1 For each line found in Phase 2, perform one of two tests, depending on whether one is seeking small Mahler measures or small limit points.

  3.1.i If seeking small limit points, compute $M(P_{\mathbf{a},\mathbf{b}}(x, y))$.

  3.1.ii If seeking small single-variable Mahler measures, then we are interested in the precise values obtained as we approach the limit, whilst constraining the degree, and so we compute $M(P_{\mathbf{a},\mathbf{b}}(z, z^t))$ for a range of t, constrained by a target bound for the degree (we chose a degree bound of 180, in order to compare our results with those with the online tables of Mossinghoff [26] and Sac-Épée [34]).

For part 3.1.ii of this process, we performed two reductions before assessing the degree of $P_{\mathbf{a},\mathbf{b}}(z, z^t)$:

- We removed any cyclotomic factors,
- If the polynomial was not weakly primitive, we weakly primitized the polynomial.

5.2.3 Experiments in Practice

We have now detailed how our experiments work in theory. Before outlining our results, we explain in a bit more detail how these experiments and computations were run in practice.

For each choice of our parameters (ℓ_1, r_1, etc.), we organised our search for small Mahler measures as follows:

1. Loop over m and n up to some bound, determined by our patience.
2. Given m and n (and a decision as to whether one of the $n+1$ edges is to be positive), compute the $B_{i,j}$ in (4.3) (with $\ell, r \in \{-1, 0\}$).
3. Loop over ℓ and r, and for each pair (ℓ, r), compute the polynomial on the right hand side of (4.24) with the sixteen possibilities given by (4.25) to give sixteen reciprocal polynomials (multiplied and/or divided by polynomials of Mahler measure 1) corresponding to the different decorations for the ends of the paths.
4. Compute the Mahler measure, and keep if it is small.

That is Phase 1 of the general strategy. Note that Theorem 4.3.1 tells us that there is a polynomial $Q(v, w, x, y, z)$ such that for each ℓ, m, r, n our reciprocal polynomial has the shape $Q(z, z^\ell, z^m, z^r, z^n)$. But we do not in fact compute Q explicitly. We merely compute various instances of $Q(z, z^\ell, z^m, z^r, z^n)$.

Phase 2 is done precisely as previously described.

For Phase 3, if we are after small limits, we have to ask the following: how do we compute P, given that we have not computed Q explicitly? We could – in principle – work through the detail to get explicit formulas for all sixteen Q. In practice, however, for a line given by $\ell = a_1 t + b_1$, $m = a_2 t + b_2$, $r = a_3 t + b_3$, $n = a_4 t + b_4$, we simply computed $Q(z, z^\ell, z^m, z^r, z^n)$ with $t = 50$ to obtain $P(z, z^{50})$ (though other suitably large enough values of t would also work). Then, writing the monomials in $P(z, z^{50})$ in the shape cz^{50a+b} and replacing this monomial by $cy^a z^b$, we expect to have recovered $P(z, y)$. Indeed, for the relatively small digraphs used we can be certain that we have done so.

5.3 Results

We are now finally in a position to run our experiments and outline our results.

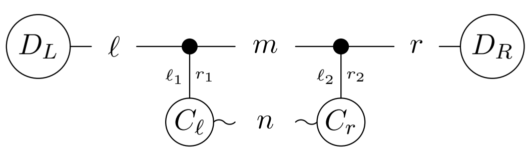
As with all computations in this thesis, we used PARI/GP for our experiments. By the fact we know what the shapes of the reciprocal polynomials will be in each possible digraph, as outlined in Figure 5.8, we were able to create a command which could give us the reciprocal polynomial in each case, depending on core digraph and decorations. From here, we create a routine which can run over the number of vertices (up to a set bound), and all possibilities of signed arcs and charges within the core digraphs.

We outline the experiments and corresponding results of small Mahler measures (of single variable polynomials) and small limit points (that is, small Mahler measures of two-variable polynomials) individually.

5.3.1 Small Mahler Measures

As with all other authors, we restrict our results to small Mahler measures coming from polynomials of degree at most 180.

For small Mahler measures, we exclusively used the core digraph CD_1, as shown in Figure 5.6. Our experiments, of course, attached pendant paths with decorations at the end, as detailed in Section 5.1.3. For reference, the digraphs we were therefore looking at were of the following shape:



.

In the first instance, we put a bound of 10 on ℓ, m, r, n, for the number of vertices in each section of the path, along with trying all possibilities for ℓ_1, ℓ_2, r_1, r_2, C_ℓ, C_r. For those that were especially fruitful, we pushed the bound on ℓ, m, r, n up to 20. In each case we looped over all sixteen patterns of decorations, and employed our three-phase campaign: search for points (with a threshold of 1.3 for their Mahler measure), then search for lines passing through triples of points and then finally test the lines.

In the online list of Mossinghoff [26], there are 8458 small measures, and in the online list of Sac Épée [34] (along with two additional examples in the paper Otmani et al. [29]) there are a further 115. We found 8334 of these, and also 1 new one. Including our newly found value, this thus means we found just over 97% of known small Mahler measures.

Additionally, there are 236 known tiny Mahler measures (those below 1.25): we found 235 of these. Achieving this required us to push the bound on ℓ, m, n, r up to 30 for a particularly good family, as seen in Figure 5.9, and reducing the Mahler measure bound to 1.25.

We state our newly found small Mahler measure here:

Result 5.3.1. *The value* $m_{\text{new}} = 1.252826882865\cdots$ *is the Mahler measure of a degree* 180 *irreducible polynomial,* f_{new}.

Remark. The polynomial f_{new} has 81 terms, making it somewhat impractical to state explicitly. To state it more simply, it is the only non-cyclotomic irreducible factor of the short polynomial $z^{221} - z^{160} - z^{155} + z^{66} + z^{61} - 1$. More details on "short polynomials" and *shortness* can be found in McKee and Smyth [25].

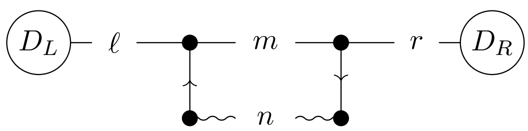


Figure 5.9: A particularly good family of digraphs for finding small Mahler measures.

5.3.2 Small Limit Points

When searching for limit points, the bound on Mahler measures used in Phase 1 of our strategy was increased to 1.38 (at least for promising values of ℓ_1, ℓ_2, r_1, r_2, c_ℓ, c_r), but the maximum bound for ℓ, m, r, n was necessarily lower: 15. We found 57 of the 61 known small limit points (so just over 93% of all known limit points).

In Table D.1, seen in Appendix D, we describe the limit points found and the corresponding details of the core digraph, decorations and size of pendant paths for each digraph with that limit point. This Table is dense with information; each row consists of the following data:

- The first two columns consist of the parameters ℓ_1, r_1, ℓ_2 and r_2; the signed arcs within the core digraph.
- The third column consists of the parameters c_ℓ, c_r, D_L and D_R; c_ℓ and c_r being the corresponding values of charged vertices which appear in every core digraph, and D_L and D_R representing the decorations at the ends of the pendant paths.
- The fourth column consists of the parameters u, d and c_m; u and d being the signed arcs which features only in the core digraphs CD_2 and CD_3, and c_m being the corresponding value of the charged vertex which appears exclusively in CD_3.
- The fifth column features the line of (ℓ, m, r, n), the number of vertices in each path of the overall digraph, parametrized by the variable t.
- The final column gives the value of the limit point.

To encapsulate all of this information into one Table, some conventions have been adopted. We detail these, as well as some quirks of the Table, briefly to aid clarity.

In many examples, we have $n = -1$ in the digraph CD_1. This therefore means that C_ℓ and C_r are in fact the same vertex. We repeat the corresponding value for c_ℓ and c_r in Table D.1 to avoid confusion.

If a limit points comes from the digraph CD_3, the value of n is listed as 1. This reflects the existence of the negative edge between the vertices C_ℓ and C_m.

If a limit point comes from the digraph CD_2, the value of n is listed as 0. This reflects the lack-of existence of a guaranteed negative edge, whilst also making it clear that we are not associating two vertices together. Unfortunately, this causes the caveat that if the column featuring n is read locally, it may suggest that a negative edge does in fact exist. However, if read in context of the entire Table, one can see that if values for u and d are present, and a value for c_m is not present, then we are indeed looking at the digraph CD_2, and so there is no such guaranteed negative edge.

Furthermore, the decorations D_L and D_R are indicated as follows:

- no decoration: $\bullet$,
- a charged vertex: $+$ or $-$ to indicate the necessary charge;
- a fork of two neutral leaves: $>$ if on the left (i.e.: to represent D_L) and $<$ if on the right (i.e.: to represent D_R).

For example, the entry

ℓ_1, r_1	ℓ_2, r_2	c_ℓ, c_r D_L, D_R	u, d c_m	ℓ, m r, n	Limit point
-1, -1	1, -1	0, 0 $+$, $+$		t, $0t+2$ $2t+1$, -1	$1.36443581\ldots$

represents the family of digraphs as seen in Figure 5.10:

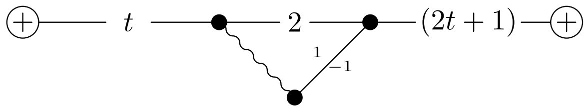


Figure 5.10: A family of digraphs with Mahler measure $1.36443581\cdots$.

We note that when we are talking about small Mahler measures of two-variable polynomials (that is to say, small limit points) coming from digraphs, we actually are referring to a family of digraphs, as opposed to a specific digraph. This should not come as a surprise; the fact that we can view these values as limits of Mahler measure values from single variable polynomials at least alludes to this fact.

Furthermore, as we saw in Example 5.1.2, when we were taking specific values of t, the values of the Mahler measure for each digraph seemed to be tending to a value, as demonstrated in Table 5.1. Indeed, if we did take larger values of t, we would see that the values were tending to $1.25543386\cdots$, which is the smallest known limit point of Mahler measures.

As has been previously mentioned, this means that if we start with a specific core digraph, and fix certain decorations for our pendant paths, we can still get a variety of different limit point values, depending on what rate we grow the pendant paths.

Explicit Calculation

Up to this point, we have adopted a theoretical approach, and subsequent understanding, of how to find Mahler measure values from families of digraphs. We will soon look at an explicit example of how to find both the associated two-variable polynomial and Mahler measure of a family of digraphs. Before doing this, we clearly define what we are finding:

Definition 5.3.2. *Let $G = G_t$ be a family of digraphs, whose number of vertices is defined by a function of t. We define the* ***Mahler measure of the family of digraphs,*** *$M(G)$ as:*

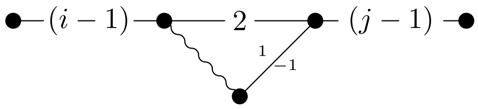
$$M(G) = \lim_{t\to\infty} M(G_t)\,,$$

To make this absolutely clear, for a fixed t, G_t is some digraph, and so we can find its reciprocal polynomial and $M(G_t)$ is a Mahler measure value of dimension one. On the other hand, G can be viewed as an "infinite digraph", and so $M(G)$ is the limit point of the values $M(G_t)$ (a value which we assume to exist). Thus, $M(G)$ is a Mahler measure value of dimension two, and has an associated two-variable polynomial. We explicitly show how to find these two-variable polynomials shortly.

Drawing upon both our knowledge of growing digraphs from chapter 4, and how to calculate the Mahler measure of two-variable polynomials from chapter 2, we now explicitly find both the associated two-variable polynomial and Mahler measure of a family of digraphs. In particular, we will look at the family of digraphs represented in Figure 5.10.

Example 5.3.3. We begin by noting that we will label the reciprocal polynomial as $R^{++}_{t+1,2t+2}(z)$, since $\ell = t$ and $r = 2t + 1$, and we have an additional vertex from our decoration at each end.

We recall that (4.24) gives us a general shape for the reciprocal polynomial of the digraph. So, our first step is to calculate this polynomial; that is, we calculate the $B_{i,j}$ and the c_{ij}. The c_{ij} are determined by the decorations at the ends of the pendant paths. The $B_{i,j}$ are linear combinations of the $R_{i,j}$, which come from calculating the reciprocal polynomial of the basic core digraph, with one vertex attached where appropriate. The below shows exactly the look of the digraphs where our $R_{i,j}$ originate from:



.

Computing these $R_{i,j}$ is straightforward. Letting G represent our core digraph, we can find the characteristic polynomial of $G_{i,j}$, and then the corresponding reciprocal polynomial in each instance.

This gives us the following:

$$\begin{aligned} R_{1,1}(z) &= z^{10} + 2z^8 + 2z^6 + 2z^4 + 2z^2 + 1\,, \\ R_{1,0}(z) &= z^8 + 2z^6 + z^4 + 2z^2 + 1\,, \\ R_{0,1}(z) &= z^8 + 2z^6 + 3z^4 + 2z^2 + 1\,, \\ R_{0,0}(z) &= z^6 + 2z^4 + 2z^2 + 1\,. \end{aligned} \tag{5.5}$$

We now find the $B_{i,j}$, using (4.3):

$$\begin{aligned} B_{1,1}(z) &= z^{10} - z^6\,, \\ B_{1,0}(z) &= z^6 - z^4\,, \\ B_{0,1}(z) &= -z^6 + z^4\,, \\ B_{0,0}(z) &= -z^4 + 1\,. \end{aligned} \tag{5.6}$$

Furthermore, for the given decoration, we have that $c_{10} = z$, $c_{01} = z$ and so $c_{00} = z^2$ (this can be found from Table 4.1). So, (4.24) gives us:

$$\frac{\kappa_1(z) R^{++}_{t+1,2t+2}(z)}{\kappa_2(z)} = z^2(z^2-1)(z^{6t+10}(z^2+1) - z^{4t+7} + z^{2t+5} - (z^2+1))\,. \tag{5.7}$$

Since $\kappa_1(z)$ and $\kappa_2(z)$ are both Kronecker-cyclotomic polynomials (they are $(z+1)^2 z^2$ and 1 respectively), we do not need actually to include them in our calculations, and work solely with the right hand side of (5.7) rather than computing the reciprocal polynomial itself. Furthermore, we can ignore the $z^2(z^2-1)$ term, since this also has Mahler measure 1.

We now need to calculate the two-variable polynomial whose Mahler measure gives the limit of the sequence of Mahler measures obtained as t varies, as explained towards the end of Section 5.2.3. For this particular example, however, we can perform the manipulation directly. Noting that we only have even coefficients of t in our powers of z, we put $y = -z^{2t+3}$, where the '3' is chosen to produce a polynomial of smallest degree in z. This gives us:

$$P(z,y) = -\left(y^3z^3 + (y+1)z^2 + (y^3+y^2)z + 1\right) . \tag{5.8}$$

We can now calculate the Mahler measure of (5.8). We find that the most efficient way to calculate this using PARI/GP is to make use of Lemma 2.2.1 and the refined method we introduced. We note our P here is indeed reciprocal so we can use our method. Then, calculating $M(P(z,y))$ gives us the limit point $1.36443581\cdots$, as stated.

Finally, we note that the polynomial $-P(z,y)$ agrees with that found by Boyd and Mossinghoff [5].

Remark. We have implicitly used the fact that $M(\pm P(\pm x, \pm y)) = M(P(x,y))$ here. We have yet to confirm this is true however; we discuss this in more detail in chapter 6.

We have previously noted that we are not constrained by the generic limit in Theorem 5.2.2 and, as seen in (5.4), we can get different limit points. In practice, this means that if we change how our pendant paths grow (whilst keeping all other variables the same), we do actually find a different limit point. This did sound like a surprising notion as first, but we can see from our method in Example 5.3.3 why this is the case. We demonstrate this phenomenon with the following example:

Example 5.3.4. Consider the following family of digraphs:

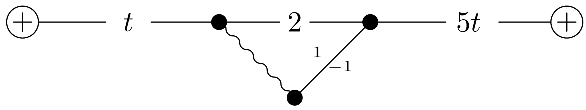


Figure 5.11: A different family of digraphs.

We note how the only difference here compared to Figure 5.10 is we are changing how our “right” pendant path is growing.

Following this through (noting in particular that the $B_{i,j}$ and c_{ij} are the same as in Example 5.3.3), we eventually get:

$$P(z,y) = (y^6z^2 - 1)(z^2 + 1) - (y^5 - yz^4)\,.$$

Finally, we find that $M(P) = 1.36272428\cdots$, which is the Mahler measure of our family of digraphs.

Remark. These explicit examples have been included for clarity of what actually happened in our experiments. In practice, everything outlined here is computed within the experiment; the only user required element is to input initial bounds and values.

5.4 A Remark on Bridges

In Section 4.4, we saw how we can take a digraph, and create a new, large, digraph by attaching pendant paths to our original digraph, and then attaching new digraphs at the ends of these pendant paths. However, towards the end of the Section, we stated that we would not be performing experiments with these types of digraphs. After seeing the results of the experiments we have run, we take the opportunity here to further justify why we have not experimented on these digraphs.

Strong Set of Results

From experimenting only on a select few cases of digraphs which have had two decorated pendant paths attached, we have found a large percentage of known small Mahler measure values and small limit points. It is not unreasonable to suggest that with stronger computational power, and more time, that these experiments could find more as well. With such a positive set of results, which grew as we expanded our search parameters, it made sense to work with a simpler setup which was proving fruitful.

Furthermore, the experiments we ran were on digraphs which supported our heuristic intuition. The concept of bridges between digraphs was one which veered away somewhat from this original idea. In turn, veering away from this idea may not yield as positive a set of results.

Time

As mentioned in Section 4.4.5, the calculation of the reciprocal polynomial of these bridged digraphs was more time-consuming than for our other digraphs. This in itself is not too problematic. However, we also need to combine the calculations of this reciprocal polynomial with our three phase process, as outlined in Section 5.2.2.

This causes two problems.

- We need to find the reciprocal polynomial of many different bridged digraphs (as our parameters vary), meaning we need to run the routine a large number of times,
- The three phase process is itself time-consuming.

So, with these problems coupled together, this increases the time taken to run the experiment and achieve the results by a noteworthy amount.

Hope for the Future?

Of course, there is a possibility that experimenting on these bridged digraphs may be fruitful. Performing calculations on a more powerful machine, the creation of a more efficient routine and even trying a different piece of software are just some immediate ways one can think of to make this idea less time-consuming.

Equally, this idea and method might be a way of finding an even greater yield of digraphs with small Mahler measures, and more families of digraphs with small limit points of Mahler measures. This may also be the only way to find some of the values we are missing. So there is hope that, at the very least, this idea can be considered for improvement and implemented in the future.

Chapter 6

Polynomials with the Same Mahler Measure

When calculating the Mahler measures of digraphs throughout this thesis, we have implicitly used some results which explain why different polynomials (and, in turn, different digraphs) have the same Mahler measure. For example, we have made use of the multiplicative nature of the Mahler measure repeatedly, to see that if an n-variable polynomial is an n-Kronecker-cyclotomic multiple of another, they will share the same Mahler measure. In Example 5.3.3, we made use of the fact that changing the sign of variables fixes the Mahler measure. Though this latter fact is something we have yet to verify, we will do so soon. These are fairly trivial facts. However, it does lead us to an interesting question: when can two polynomials share the same Mahler measure?

In this chapter, we look at attempts to answer this question for both single variable and two-variable polynomials. As we will see, these are different situations and so we look at them completely separately. Before moving to look at these attempts, we stress that these are indeed only *attempts* to answer this question: as of yet, we do not have a complete answer for when two polynomials do share the same Mahler measure. As such, this remains an open problem, but we conjecture ideas which may lead to a resolution.

6.1 Single Variable Polynomials – Motivating Examples

We begin by stating our over-arching claim related to single variable polynomials.

Open Problem 6.1.1. *Let $f, g \in \mathbb{Z}[x]$ be monic, weakly primitive, irreducible, reciprocal polynomials, with non-trivial Mahler measures.*

Is it the case that $M(f) \neq M(g)$?

This is phrased in a very specific way. In this case, we are not giving a list of conditions which necessarily guarantees that two polynomials share the same Mahler measure. Instead, we are giving conditions which – we believe – guarantee that two polynomials cannot share the same Mahler measure. Then, in turn, if one of these conditions is broken, we can find polynomials which share the same Mahler measure.

We phrase it in this way because of the wider context of this thesis and the work presented. Our focus has been on attempting to find polynomials with small Mahler measures by means of several different experiments. Whilst our experiments are versatile and largely efficient, it is fair to ask how much these could be improved if we were to ignore polynomials which share the same Mahler measure. So, it would be ideal to attempt to find conditions which we could constrict polynomials to, which guarantee they would admit different Mahler measure values, as presented in Open Problem 6.1.1.

As it stands though, Open Problem 6.1.1 is a very specific statement to make with very little justification. As such, before moving on, we give examples which show that if any one of these conditions are broken, then it is possible to have an equality of Mahler measures. This gives a certain level of justification that these conditions are necessary.

Example (Breaking the Non-Trivial Condition)**.**
We have seen from Lemma 1.1.8 that all Kronecker-cyclotomic polynomials have trivial Mahler measure.

Indeed, if we consider $\Phi_2(x) = x + 1$ and $\Phi_3(x) = x^2 + x + 1$, these are both monic, weakly primitive, irreducible and reciprocal, but both have trivial Mahler measure.

Even without knowledge of Lemma 1.1.8, it would not be a surprise that we would need to exclusively consider non-trivial Mahler measure values within the context of our Open Problem.

Example (Breaking the Monic Condition).
Let $f(x) = 2x^2 + x + 2$ and $g(x) = 2x^2 + 3x + 2$. These are both weakly primitive, irreducible and reciprocal, and have a non-trivial Mahler measure, but we have that $M(f) = M(g) = 2$.

Again, this is not too much of a surprise. More importantly, however, in the context of finding small Mahler measure values, we do indeed only wish to consider monic polynomials, since for polynomials $f \in \mathbb{Z}[x]$ with leading coefficient a_n, we know that $M(f) \geq |a_n|$.

Example (Breaking the Weakly Primitive Condition).
Let $f(x) = x^{16} + x^{10} - x^8 + x^6 + 1$ and $g(x) = x^{40} + x^{25} - x^{20} + x^{15} + 1$. These are both monic, irreducible and reciprocal, and have a non-trivial Mahler measure, but we have that $M(f) = M(g) = 1.28063815\cdots$.

If we were to weakly primitize both of these polynomials, by replacing x^2 with x in f, and replacing x^5 with x in g, we then get the same polynomial: $h(x) = x^8 + x^5 - x^4 + x^3 + 1$.

Once again, this condition should not come as a surprise, especially since we know that the procedure of weakly primitizing a polynomial does not affect its Mahler measure. As such, if two polynomials share the same weakly primitized polynomial, then they will of course share the same Mahler measure.

Example (Breaking the Irreducible Condition).
Let $g(x) = \kappa(x)f(x)$, where κ is some Kronecker-cyclotomic polynomial. Then:

$$\begin{aligned} M(g) &= M(\kappa f) \\ &= M(\kappa)M(f) \\ &= M(f)\,. \end{aligned}$$

This example is both unsatisfying, and also avoidable. For instance, we could specify that at least one of our polynomials simply does not have any Kronecker-cyclotomic factors, whilst still remaining reducible. It could then be possible that this is enough to always ensure that our polynomials share different Mahler measure values. However, as we now see, this is still possible.

Example (Breaking the Irreducible Condition Again)**.**

Let $f(x) = x^2-18x+1$ and $g(x) = x^4-10x^3+23x^2-10x+1 = (x^2-3x+1)(x^2-7x+1)$. These are both monic, weakly primitive and reciprocal, and have a non-trivial Mahler measure, but we have that $M(f) = M(g) = 9+4\sqrt{5}$.

This example is actually of a very specific nature and structure, which in turn would allow us to create many other examples of this form. We look at this structure now:

Example 6.1.2. Let α be a Salem number with minimal polynomial $\mu_\alpha(x)$. Let $f(x) = \mu_{\alpha^3}(x)$ and $g(x) = \mu_\alpha(x)\mu_{\alpha^2}(x)$.

It is straightforward to see that, for any positive integer n, $M(\mu_{\alpha^n}) = \alpha^n$. So, it follows that we have $M(f) = M(g) = \alpha^3$.

Remark. The structure here also applies if we were to take a quadratic Pisot number; that is, Pisot numbers whose minimal polynomial is of the form x^2+ax+1, for $a \in \mathbb{Z}$ and $a \leq -3$.

We note that when breaking the irreducible condition, especially in this instance, we are only taking one of our polynomials to be irreducible, whereas when we have demonstrated breaking our other conditions, we have chosen both polynomials to break the condition. Of course, we could create an example where both polynomials are reducible (by letting $f(x) = \mu_\alpha(x)\mu_{\alpha^4}(x)$ and $g(x) = \mu_{\alpha^2}(x)\mu_{\alpha^3}(x)$, for example). Our presented examples simply demonstrate the most straightforward example of this phenomenon.

Another point worth raising is that in the context of finding small Mahler measures, it is likely that we want to be considering irreducible polynomials in any case.

Example. Let $f \in \mathbb{Z}[x]$ be a reducible polynomial with n irreducible factors, say $f = f_1 \cdots f_n$, and assume that none of these factors are Kronecker-cyclotomic. If $M(f)$ is small, then there is at least one f_i such that $M(f_i) < (1.3)^{1/n}$.

In the case $n = 2$, this would mean that there is at least one factor, say f_1, such that $M(f_1) < \sqrt{1.3} = 1.14017542\cdots$. As this is smaller than the smallest known Mahler measure value, we do not expect this to happen.

In particular, with polynomials which follow the structure presented in Example 6.1.2, the smallest Mahler measure we can obtain would be $\lambda^3 = 1.62754514\cdots$, since λ is the smallest known Salem number.

Example 6.1.3 (Breaking the Reciprocal Condition)**.** (Dixon and Dubickas [11]).
Let $f(x) = x^6 - x^4 - x^3 - x^2 + 1$ and $g(x) = x^4 - x + 1$. These are both monic, weakly primitive and irreducible, and have a non-trivial Mahler measure, but we have that $M(f) = M(g) = 1.40126836\cdots$.

Once more, in context of finding small Mahler measures, we know from Theorem 1.1.19 that we have to restrict ourselves to reciprocal polynomials anyway. However, it is interesting to see that we also require this condition if we wish to only ever consider polynomials which admit different Mahler measure values. Furthermore, there are results from Dixon and Dubickas [11] which may prove useful in attempts to answer Open Problem 6.1.1, which we will now explore.

6.2 Attempts at a Proof

Now that we have justified the conditions set out in Open Problem 6.1.1, we now move to looking at attempts to answer and prove the presented idea. We stress that the results and discussion that follow may not necessarily form the basis of the way to answer our Open Problem, but they do present a possible route which seems fruitful.

Definition 6.2.1. *Let $f \in \mathbb{Z}[x]$ and α be a zero of f. We say that α is **large** if $|\alpha| > 1$.*

This is a fairly intuitive definition; in particular, a consequence of this definition is that we see only large zeros contribute to the Mahler measure of a polynomial. This now allows us to introduce the following, more interesting, concept:

Definition 6.2.2. *Let $f \in \mathbb{Z}[x]$ and $M(f) = \mathcal{M}$. We say that f is **basal** (for $\mathcal{M}$) if it is the polynomial of least degree such that:*

- *$M(f) = \mathcal{M}$;*
- *At most half the zeros of each irreducible factor of f are large.*

Remark. We can choose to omit the "for $\mathcal{M}$", since a polynomial can only ever be basal for one value.

Example. We look at simple examples of polynomials which are basal and which are not.

- Lehmer's polynomial, $\Lambda(x)$, is basal (for λ), since it is irreducible, and only has one large zero.
- $x^4 - 2x^3 - x^2 + x + 2 = (x^3 - x - 1)(x - 2)$ is not basal, since one of its irreducible factors only has large roots.

We note in the latter example that the polynomial has at most half of its zeros large, but since one its irreducible factors has more than half of its zeros large, the polynomial is not basal.

There is also a somewhat stronger notion we can introduce:

Definition 6.2.3. *Let $f \in \mathbb{Z}[x]$ be irreducible and $M(f) = \mathcal{M}$. We say that f is* ***basal irreducible*** *if it is the polynomial of least degree among irreducible polynomials such that:*

- $M(f) = \mathcal{M}$*;*
- *At most half of the zeros of f are large.*

Our focus will be on basal irreducible polynomials. Indeed, if the statement of Open Problem 6.1.1 is correct, it makes sense to do so, since we wish to be finding polynomials which will always have different Mahler measures. More importantly, there is a known result related to irreducible polynomials which are basal that we can take advantage of:

Proposition 6.2.4 (Dixon and Dubickas [11])**.** *Let $E \subset \mathbb{C}$ be a finite Galois extension of $\mathbb{Q}$, $f \in \mathbb{Z}[x]$ and $M(f) \in E$.*

If f is basal irreducible and monic, then all zeros of f lie in E.

We omit a proof; details are provided by Dixon and Dubickas.

It is easy to see the relevance of such a result, especially in relation to our problem. Indeed, a basal irreducible polynomial must also be weakly primitive by definition, and so a reciprocal, monic polynomial with non-trivial Mahler measure will have at most half its zeros large. So this result seems to fit nicely with our problem.

Indeed, we can use this result to show the following:

Lemma 6.2.5. *Let $f, g \in \mathbb{Z}[x]$ be monic and irreducible of degree at least 2 such that $M(f) = M(g)$. Suppose that f has exactly one large zero and is the polynomial of least degree with Mahler measure $M(f)$ (so $\deg(g) \geq \deg(f)$). Then, the splitting field of f over $\mathbb{Q}$ is contained inside the splitting field of g over $\mathbb{Q}$.*

Proof. Let α be the large zero of f and $\beta_1, \cdots, \beta_m$ be the large zeros of g. Additionally, let K and L be the splitting fields over $\mathbb{Q}$ of f and g respectively.
Claim: f is basal irreducible.
Justification: As f is the polynomial of least degree with Mahler measure $M(f)$, we only need to check that at most half of the zeros of f are large. However, we know that there is exactly one large zero, and $\deg(f) \geq 2$, so this is indeed the case. Hence, f is indeed basal irreducible.

As $M(f) = M(g)$, we have that $|\alpha| = |\beta_1 \cdots \beta_m|$. As such, it follows that $M(f) \in L$. Then, by Proposition 6.2.4, all zeros of f are inside L. Thus, $K \subset L$, as required. □

As we see in the proof here, the equality of the Mahler measures is particularly crucial for arriving at our final result. In particular, if we were to fix all other conditions, and use the notation from the proof of Lemma 6.2.5, we can arrive to a logical conclusion that:

$$\text{``}M(f) = M(g) \Longrightarrow K \subset L\text{''}.$$

In turn, taking the contrapositive of this would suggest that:

$$\text{``}K \not\subset L \Longrightarrow M(f) \neq M(g)\text{''}. \tag{6.1}$$

Following this argument, we can pose a slightly weaker question to spawn from Open Problem 6.1.1, which appears more approachable to solve.

Open Problem 6.2.6. *Let $f, g \in \mathbb{Z}[x]$ be monic, weakly primitive, irreducible, reciprocal polynomials, with non-trivial Mahler measures. Let f have exactly one large zero and be* *the polynomial of least degree with Mahler measure $M(f)$, and g have exactly two large zeros.*

Is it the case that $M(f) \neq M(g)$?

Of course, this is a weaker question than our original one, and if we were to able to answer Open Problem 6.1.1, we do indeed get an answer here too. However, we can also see the appeal in adding the additional constraints presented here. If we are able to answer this positively, one would hope that maybe we can learn enough to tackle the problem more generally; for example, by considering an increased number of large zeros for g, or an arbitrary number of large zeros for g.

Though we still do not have an answer to this question, we do note that, because of Lemma 6.2.5 and the logical conclusion reached by (6.1), to solve this problem, it would suffice to show that the splitting field of f over $\mathbb{Q}$ is not contained in the splitting field of g over $\mathbb{Q}$.

Depending on the success of this potential strategy, it could also be extended further. In theory, we could also extend Lemma 6.2.5 to consider the case where the polynomial f has more than one large zero as well, assuming that $\deg(f)$ is suitably large enough. However, we expect that the use of this, and the subsequent care that would have to be taken when using this, can only become apparent if we can understand how to solve Open Problem 6.2.6.

6.2.1 Survey of Further Ideas

We have now established a weaker problem, as well a direct avenue for solving this problem. Now, we give a brief overview of routes explored so far to attempt to answer Open Problem 6.2.6. Again, we stress that this is simply an overview exploring potential ideas and the justification for their exploration, and they may or may not be fruitful options.

The results of Proposition 6.2.4 and Lemma 6.2.5 certainly suggest that we should familiarise ourselves further with some useful Galois theoretic results. Indeed, there are some results related to the Galois theory of reciprocal polynomials and, more narrowly, Salem polynomials, which may be of relevance.

Theorem 6.2.7 (Viana and Veloso [46]). *Let $\mathbb{F}$ be a field, $a_1, \cdots, a_m \in \mathbb{F}$ be distinct values and $n = 2m$. Then, the polynomial:*

$$f(x) = x^n + 1 + \sum_{i=1}^{m} (-1)^i a_i (x^{n-i} + x^i)\,, \tag{6.2}$$

has Galois group $\mathrm{Gal}(f) = (\mathbb{Z}/2\mathbb{Z})^m \rtimes S_m$, where S_m is the symmetric group on m elements.

This result is related to a very specific shape of reciprocal polynomial. However, this does at least give us a level of insight into how we might expect the Galois group of any given reciprocal polynomial to look. A more specific result related to Salem polynomials (that is, polynomials that are minimal polynomials of a Salem number) can also be derived.

Definition 6.2.8. *Let $f \in \mathbb{Z}[x]$ be monic, irreducible and reciprocal, with $\deg(f) = 2n$ and zeros $\alpha_1, 1/\alpha_1, \cdots, \alpha_n, 1/\alpha_n$. The corresponding* ***trace polynomial*** *of f is the integer polynomial $t(x)$ with zeros $\gamma_1 = \alpha_1 + 1/\alpha_1, \cdots, \gamma_n = \alpha_n + 1/\alpha_n$.*

Theorem 6.2.9 (Christopoulos and McKee [8]). *Let f be a Salem polynomial of degree $2n$, t be its corresponding trace polynomial and $\mathrm{Gal}(f)$ and $\mathrm{Gal}(t)$ be the Galois groups of f and t respectively. Then:*

$$\mathrm{Gal}(f) \cong N \rtimes \mathrm{Gal}(t)\,,$$

where N is the kernel of the natural map $\psi : \mathrm{Gal}(f) \to \mathrm{Gal}(t)$. In particular, we have that $N \cong (\mathbb{Z}/2\mathbb{Z})^n$ or $N \cong (\mathbb{Z}/2\mathbb{Z})^{n-1}$, with the latter only possible when n is odd.

A detailed proof of this is provided by Christopoulos [7]. However, the key thing to note with this is the similarity of the results between this and Theorem 6.2.7. Indeed, the statement of Theorem 6.2.9 is perhaps more insightful, and we explain this in slightly more depth now.

Definition 6.2.10. *A group H is a* ***group extension*** *of a group K by N if there exists a homomorphism $\phi : H \to K$ with kernel N.*

Furthermore, we say a group extension is ***split*** *if $H \cong N \rtimes K$.*

In the case of Theorem 6.2.9, we can clearly see that $\mathrm{Gal}(f)$ is a group extension of $\mathrm{Gal}(t)$ by N, and that this does indeed split. We could also see this phenomenon in Theorem 6.2.7. Unfortunately, it is not the case that we can find a split group extension for every reciprocal polynomial. For example, Viana and Veloso [46] show that for the n-th cyclotomic polynomial Φ_n, which is reciprocal for $n > 1$, we have a split extension if and only if either $4 \mid n$, or $p \mid n$ for some prime p such that $p \equiv 3 \pmod 4$.

Now we turn to our set up more specifically. Consider some monic, weakly primitive, irreducible, reciprocal polynomial $f \in \mathbb{Z}[x]$, with zeros $\alpha_1, 1/\alpha_1, \cdots, \alpha_n, 1/\alpha_n$, $\deg(f) = 2n$ and with corresponding trace polynomial t. Labelling the Galois groups of f and t as $\mathrm{Gal}(f)$ and $\mathrm{Gal}(t)$ respectively, we again know that there exists a natural map $\psi : \mathrm{Gal}(f) \to \mathrm{Gal}(t)$ with kernel N. In particular, elements of N must map each zero α to either itself or $1/\alpha$. So, elements of N swap pairs of reciprocal zeros, meaning that N must be some subgroup of $(\mathbb{Z}/2\mathbb{Z})^n$. Furthermore, if we have a split extension, we then have that $\mathrm{Gal}(f) \cong N \rtimes \mathrm{Gal}(t)$. Knowing that $|N| \leq 2^n$, and $\mathrm{Gal}(t)$ cannot be larger than S_n, we have that $|\mathrm{Gal}(f)| \leq 2^n n!$.

We now return to the final conditions of Open Problem 6.2.6. Let $\deg(f) = 2n$ and $\deg(g) = 2m$, for $m \geq n$. We set up $\deg(g) \geq \deg(f)$ since our set up states that f is the polynomial of least degree with Mahler measure $M(f)$. So, if we are to answer the posed question, we need to consider a polynomial of weakly greater degree.

Since f has exactly one large zero, we know it is a Salem polynomial. Hence, by Theorem 6.2.9, we know we can find a group extension which splits, and by the argument above, we know that $|\mathrm{Gal}(f)| \leq 2^n n!$, where $\mathrm{Gal}(f)$ is the Galois group of f.

However, less can be said about g. We can consider its corresponding trace polynomial, but we do not know if this leads us to a split extension. If we are in the situation where we do have a split extension, we can say that $|\mathrm{Gal}(g)| \leq 2^m m!$, where $\mathrm{Gal}(g)$ is the Galois group of g. If we do not have a split extension, all we can say is that $|\mathrm{Gal}(g)| \leq (2m)!$.

Now say that K and L are the splitting fields of f and g respectively. We know that $|K| = |\mathrm{Gal}(f)|$ and $|L| = |\mathrm{Gal}(g)|$. Using (6.1), this would suggest that, to show $K \not\subset L$ and hence $M(f) \neq M(g)$, we have to show that $|\mathrm{Gal}(f)| \nmid |\mathrm{Gal}(g)|$. As it stands, we do not have enough information to verify this either way.

6.3 Two-Variable Polynomials: An Introduction

The rest of this chapter will now focus on two-variable polynomials, as we attempt to answer the question of whether two different two-variable polynomials can share the same Mahler measure. Unlike with the single variable case, and at the start of Section 6.1, we do not give an over-arching claim or statement. We will see that this is because we currently do not have an over-arching claim to make, but rather claims for specific circumstances. Indeed, there are many more subtleties to consider with two-variable polynomials, which make it more difficult to give a “one size fits all” style claim, similar to Open Problem 6.1.1.

We note that when we have studied limit points of Mahler measures, and two-variable polynomials, throughout this thesis so far, we have again exploited the multiplicative nature of the Mahler measure in many cases. However, as mentioned, in Example 5.3.3 we made use of the fact that changing the signs of variables also fixes the Mahler measure. This is something we have yet to verify, but will in Section 6.4.

A particular issue we have experienced with our calculations of the Mahler measures of two-variable polynomials is our lack of certainty about their exact value. We have already briefly discussed how this is a wider issue and not something we particularly address within this thesis. However, the consequence of this is when we find two two-variable polynomials which, according to PARI/GP, for example, share the same Mahler measure value, we do not have complete certainty that they actually are the same. This is significantly less of a problem with single variable polynomials, since we have a deep pool of knowledge at our disposal for finding roots (and consequently, zeros) of such polynomials. As such, we can feel *very* confident in saying when two single variable polynomials have equal Mahler measure, as in Example 6.1.3, for instance.

As such, we will look at different ways of determining if two two-variable polynomials share the same Mahler measure without having to actually compute their Mahler measures. We hope that these different ways fully explain when two-variable polynomials do share the same Mahler measure, but we as of yet have no justification to support such a claim.

6.4 A Notion of Equivalence

As stated, our aim is to find some way of determining if two two-variable polynomials share the same Mahler measure value, without having to rely on calculations. Before doing this, however, we first look at some basic results which will help us in determining such a concept, as well as giving wider context to what we already know.

6.4.1 Basic Results

The first of the basic results we state is one which we have already seen, albeit in a slightly different form. However, we restate it to note its importance here:

Lemma 6.4.1 (See also: Lemma 1.2.6)**.** *For polynomials* $P, Q \in \mathbb{C}[x, 1/x, y, 1/y]$*,* $M(PQ) = M(P)M(Q)$*.*

This is particularly useful for helping us with some other simple results related to two-variable polynomials. However, there is one crucial change we have made to our statement: we are now considering Laurent polynomials, as opposed to linear polynomials. We note that we can consider Laurent polynomials as a result of Lemma 2.5.1. This does not affect the associated proof in any way, but the results that follow will also apply to Laurent polynomials, hence why we have stressed this extension.

Corollary 6.4.1.1. *For* $P \in \mathbb{C}[x, 1/x, y, 1/y]$ *and* $r, s \in \mathbb{Z}$*, we have that:*

$$M(x^r y^s P(x, y)) = M(P(x, y)) \,. \tag{6.3}$$

Remark. This result follows trivially, but we state it owing to its importance later on.

Lemma 6.4.2. *For* $P \in \mathbb{C}[x, 1/x, y, 1/y]$*, we have that:*

$$M(P(-x, y)) = M(P(x, y)) \,; \tag{6.4a}$$

$$M(P(x, -y)) = M(P(x, y)) \,. \tag{6.4b}$$

Proof. We first prove (6.4a). Furthermore, this follows a similar argument to a method we have seen for calculating the Mahler measure of a two-variable polynomial.

Let $P(-x,y)$ have degree d in y, and $y_1(-x),\cdots,y_d(-x)$ be continuous, piecewise analytic functions in $-x$ that are the d solutions to $P(-x,y)=0$. This means we can write:

$$P(-x,y)=a_0(-x)\prod_{k=1}^{d}(y-y_k(-x)),$$

where $a_0(-x)$ is the coefficient of the y^d term. A similar argument used in the Proof of Proposition 1.1.2 shows that:

$$\int_0^1 \log|P(-x,e^{2\pi is})|\ \mathrm{d}s = |a_0(-x)| + \sum_{k=1}^{d}\log^+|y_k(-x)|\,. \tag{6.5}$$

Now consider substituting in $x=e^{2\pi it}$ in (6.5), and then integrating over t. This gives us:

$$\begin{aligned}\int_{t=0}^{1}\int_{s=0}^{1}\log|P(-e^{2\pi it},e^{2\pi is})|\ \mathrm{d}s\,\mathrm{d}t &= \int_{t=0}^{1}|a_0(-e^{2\pi it})|\ \mathrm{d}t + \sum_{k=1}^{d}\int_{t=0}^{1}\log^+|y_k(-e^{2\pi it})|\ \mathrm{d}t\\ \Rightarrow m(P(-x,y)) &= m(a_0(-x)) + \underbrace{\sum_{k=1}^{d}\int_{t=0}^{1}\log^+|y_k(-e^{2\pi it})|\ \mathrm{d}t}_{=S}\ .\end{aligned} \tag{6.6}$$

We now need to consider the summation S. Firstly, we note that $-e^{2\pi it}=e^{2\pi i(t+1/2)}$. So, we can rewrite S as:

$$S=\sum_{k=1}^{d}\int_{t=\frac{1}{2}}^{\frac{3}{2}}\log^+|y_k(e^{2\pi it})|\ \mathrm{d}t\,.$$

However, we note that if we were to split these into two integrals, such as:

$$S=\sum_{k=1}^{d}\int_{t=\frac{1}{2}}^{1}\log^+|y_k(e^{2\pi it})|\ \mathrm{d}t+\sum_{k=1}^{d}\int_{t=1}^{\frac{3}{2}}\log^+|y_k(e^{2\pi it})|\ \mathrm{d}t\,,$$

then the second integral is actually equal to

$$\sum_{k=1}^{d}\int_{t=0}^{\frac{1}{2}}\log^+|y_k(e^{2\pi it})|\ \mathrm{d}t\,.$$

This is by a similar argument to that seen in the Proof of Lemma 2.2.1, as demonstrated in equations (2.9) and (2.10).

In turn, this gives us that:

$$S = \sum_{k=1}^{d} \int_{t=0}^{1} \log^{+}|y_k(e^{2\pi i t})| \, \mathrm{d}t\,.$$

Substituting this back into (6.6) thus results in:

$$m(P(-x,y)) = m(a_0(-x)) + \sum_{k=1}^{d} \int_{0}^{1} \log^{+}|y_k(e^{t})| \, \mathrm{d}t\,. \tag{6.7}$$

As in (2.6), we note here that $m(P)$ is the logarithmic Mahler measure of a two-variable polynomial, whilst $m(a_0(-x))$ is the logarithmic Mahler measure of a single variable polynomial. However, the absolute values of the roots of a_0 are not changed with the map $x \mapsto -x$. Therefore, $m(a_0(-x)) = m(a_0(x))$.

In particular, this means that (6.7) is equal to (2.6), and so $m(P(-x,y)) = m(P(x,y))$, and our result follows.

We note that a similar argument holds for the proof for (6.4b). □

Remark. An alternate proof of Lemma 6.4.2 would see us make use of Jensen's Formula, from Proposition 1.1.3. In particular, we could take the double integral used for calculating the Mahler measure of P and rewrite it as single integral (in a similar fashion to what we saw in Example 2.1.2, for example). We know that the Mahler measure of a single variable polynomial is invariant under a change of sign, since the absolute values of the zeros do not change, which would give us our result.

Lemma 6.4.3. *A polynomial P which can be weakly primitized in one variable has the same Mahler measure as the resulting weakly primitized polynomial. That is to say, for $P \in \mathbb{C}[x, 1/x, y, 1/y]$ and $r, s \in \mathbb{Z}$, we have that:*

$$M(P(x^r, y)) = M(P(x,y))\,; \tag{6.8a}$$

$$M(P(x, y^s)) = M(P(x,y))\,. \tag{6.8b}$$

Proof. Again, we will only outline a proof of (6.8a), noting that (6.8b) follows by an analogous argument.

Firstly, we fix y. Then, for a given a root ρ of $P(x, y) = 0$, we can find r roots of $P(x^r, y) = 0$ which correspond to ρ; say $\rho_1, \cdots, \rho_r$. Each of these ρ_i have modulus $\rho^{1/r}$.

As such, the contribution to $M(P(x, y))$ from ρ equals the contribution to $M(P(x^r, y))$ from these ρ_i. Thus, we find that $M(P(x, y)) = M(P(x^r, y))$. □

Before moving on, we note that, for positive integers r and s, it is convenient to extend our defined construction of the Mahler measure of a Laurent polynomial (first introduced in Lemma 2.5.1) in x and y to define the Mahler measure of a Laurent polynomial in $x^{1/r}$ and $y^{1/s}$. To see how we define this, let

$$P(x, y) = \sum_{i,j} c_{i,j} x^{i/r} y^{j/s}$$

be a non-zero Laurent polynomial in $x^{1/r}$ and $y^{1/s}$, for finitely many non-zero terms $c_{i,j} \in \mathbb{C}$. For absolute clarity, we are here summing over pairs of integers, (i, j). Then, $P(x^r, y^s)$ is a Laurent polynomial in x and y, and we can define:

$$M(P(x, y)) := M(P(x^r, y^s)) .$$

For this construction, we do not require that r and s are chosen minimally such that $P(x^r, y^s)$ is a Laurent polynomial. This condition is not required due to Lemma 6.4.3. For example, if we had integers r_1, s_1, r_2 and s_2, then we note that we have $M(P(x^{r_1}, y^{s_1})) = M(P(x^{r_1 r_2}, y^{s_1 s_2}))$ and $M(P(x^{r_2}, y^{s_2})) = M(P(x^{r_1 r_2}, y^{s_1 s_2}))$, meaning that $M(P(x^{r_1}, y^{s_1})) = M(P(x^{r_2}, y^{s_2}))$.

6.4.2 Introducing and Justifying Equivalence

We are now in a position to introduce our notion of equivalence here.

Definition 6.4.4. *Let $P, Q \in \mathbb{Z}[x, y]$. We say that P and Q are **equivalent** if there exists an affine transformation $x^r y^s \mapsto x^{ar+bs+e} y^{cr+ds+f}$, for $a, b, c, d, e, f \in \mathbb{Q}$ and $ad - bc \neq 0$, taking $\pm P(\pm x, \pm y)$ to $Q(\pm x, \pm y)$.*

At this point, it is sensible to question our choice of nomenclature here, as well as the reasoning for the precise set up of the definition. In other words, we have yet to justify why such a concept can be referred to as "equivalence". This is encapsulated in the following result:

Theorem 6.4.5. *Let $P, Q \in \mathbb{Z}[x, y]$ be equivalent. Then, $M(P) = M(Q)$.*

So, our reasoning for introducing this definition of equivalence is to create a result which allows us to determine, with relative ease, if two polynomials share the same Mahler measure. Of course, the word "equivalent" is somewhat overused broadly across mathematics, and this definition should not be linked to, or confused with, any other concepts of 'polynomial equivalence' which may exist.

We are not yet in a position to see the proof of Theorem 6.4.5. Before we move to the necessary auxiliary results, we look at an example of two equivalent polynomials.

Example 6.4.6. Let $P(x, y) = y^8x^8 - y^8x^6 + y^5x^5 + y^3x^3 - x^2 + 1$ and $Q(x, y) = y^2x^8 + yx^8 + yx^5 + yx^3 + y + 1$.

Using PARI/GP, for example, we find that $M(P) = 1.36365149\cdots$ and $M(Q) = 1.36365149\cdots$. However, we remain unsure if these are indeed equal, or just agree to a large number of decimal places.

The affine transformation $x^ry^s \mapsto x^{-4r+3s+8}y^{\frac{-r}{2}+\frac{s}{2}+1}$ takes $-P(x, y)$ to $Q(x, -y)$. In this transformation, $a = -4$, $b = -\frac{1}{2}$, $c = 3$, $d = \frac{1}{2}$, $e = 8$ and $f = 1$. As such, $|ad - bc| = |\,(-4)\left(\frac{1}{2}\right) - \left(-\frac{1}{2}\right)(3)\,| = \frac{1}{2} \neq 0$.

Thus, P and Q are equivalent and Theorem 6.4.5 tells us that $M(P) = M(Q)$.

For some of the next results, we will consider the logarithmic Mahler measure of a two-variable polynomial in its integral form, as seen in Definition 1.2.1:

$$m(P) = \int_{t=0}^{1} \int_{s=0}^{1} \log|P(e^{2\pi is}, e^{2\pi it})| \,\mathrm{d}s \,\mathrm{d}t \,. \tag{6.9}$$

Lemma 6.4.7. *The integral seen in (6.9) is invariant under the transformation $(s, t) \mapsto (s + m, t + n)$, for $m, n \in \mathbb{Z}$.*

Proof. This follows from the periodicity of $e^{2\pi i}$. □

Proposition 6.4.8. *Let* Π *be a parallelogram in* $\mathbb{R}^2$ *with all four corners in* $\mathbb{Z}^2$ *and area* A*. Let* P *be a non-zero Laurent polynomial in* $\mathbb{C}[x, 1/x, y, 1/y]$*. Then:*

$$\iint_{\Pi} \log|P(e^{2\pi i s}, e^{2\pi i t})| \,\mathrm{d}s\,\mathrm{d}t = A \cdot m(P)\,. \tag{6.10}$$

Before proving this, we state and prove the following useful fact:

Lemma 6.4.9. *Let* Π *be a parallelogram in* $\mathbb{R}^2$ *with all four corners in* $\mathbb{Z}^2$*. Then, the area of* Π*, denoted* A*, is an integer.*

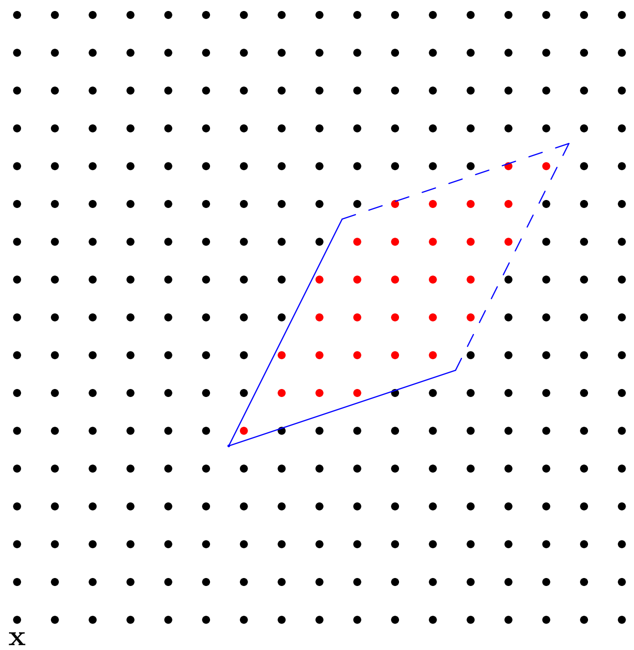
Proof. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be displacement column vectors of two adjacent sides of Π, and M be the 2×2 matrix whose columns are $\mathbf{v}_1$ and $\mathbf{v}_2$. Then, the area of Π is given by $|\det(M)|$.

Since the corners of Π are in $\mathbb{Z}^2$, it follows that the entries of M are integers, and hence $A \in \mathbb{Z}$. □

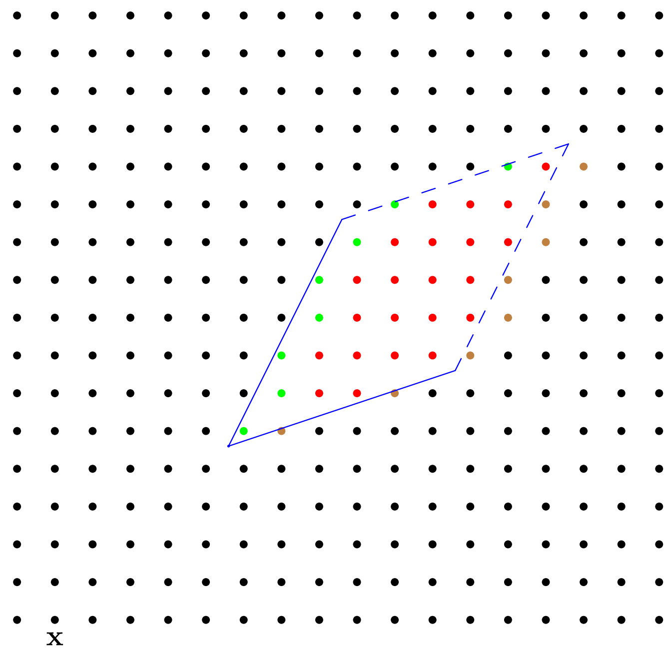
Proof of Proposition 6.4.8. Firstly, we note that if Π is a rectangle (that is, it has vertical and horizontal sides), then the result follows immediately from Lemma 6.4.7. However, the following argument holds whether our parallelogram is a rectangle or not.

For any point $\mathbf{x} \in \mathbb{R}^2$, consider the intersection of $\mathbf{x} + \mathbb{Z}^2$ and Π. In this instance, when we consider Π, we only include one corner and its two adjacent sides, and not the other two sides, nor the remaining three corners. We treat Π in this way as it will allow us to then perfectly tile the plane with the parallelogram (each point in the plane would be covered exactly once). This is visualised in top half of Figure 6.1: we have our point $\mathbf{x}$, and a resulting lattice-like structure formed by considering $\mathbf{x} + \mathbb{Z}^2$. The solid blue lines are our 'considered' sides of Π, with the dashed lines the not-included boundary of Π. The highlighted red points are precisely the intersection of $\mathbf{x} + \mathbb{Z}^2$ and Π.

<u>Claim:</u> The number of points inside the intersection of $\mathbf{x} + \mathbb{Z}^2$ and Π is constant.
<u>Justification:</u> As $\mathbf{x}$ moves, whenever a point enters Π, this is balanced by a point leaving Π (and vice versa). This is demonstrated visually in the bottom half of Figure 6.1: we have shifted $\mathbf{x}$ to the right, and see that the same number of points which have entered Π (as highlighted in green) have also left Π (as highlighted in brown).



The original intersection, before shifting $\mathbf{x}$.



The intersection after shifting $\mathbf{x}$.

Figure 6.1: A visualisation of the intersection of $\mathbf{x} + \mathbb{Z}^2$ and Π, before and after shifting $\mathbf{x}$.

A consequence of this is that, in $\mathbb{R}^2/\mathbb{Z}^2$, each point is covered the same number of times by the parallelogram, say B times. Then, we have that:

$$\begin{aligned} B &= B \iint\limits_{[0,1)^2} 1 \,\mathrm{d}x \,\mathrm{d}y \\ &= \iint\limits_{\Pi} 1 \,\mathrm{d}x \,\mathrm{d}y \\ &= A\,. \end{aligned} \tag{6.11}$$

Thus, each point is in fact covered A times by the parallelogram. Furthermore, by Lemma 6.4.9, we know that this is an integer.

Using this and Lemma 6.4.7, it follows that

$$\iint\limits_{\Pi} \log|P(e^{2\pi i s}, e^{2\pi i t})| \,\mathrm{d}s \,\mathrm{d}t = A \cdot m(P)\,,$$

as required. □

This puts us in a position to prove the following result.

Proposition 6.4.10. *For $a, b, c, d \in \mathbb{Z}$ with $ad - bc \neq 0$, we have that $m(P(x^a y^c, x^b y^d)) = m(P(x, y))$.*

Proof. We write

$$m(P(x^a y^c, x^b y^d)) = \int_{t=0}^{1} \int_{s=0}^{1} \log|P(e^{2\pi i(as+ct)}, e^{2\pi i(bs+dt)})| \,\mathrm{d}s \,\mathrm{d}t\,. \tag{6.12}$$

Now, putting $u = as + ct$ and $v = bs + dt$, we can calculate the Jacobian determinant of the above function:

$$\mathcal{J} = \begin{vmatrix} \frac{\partial u}{\partial s} & \frac{\partial u}{\partial t} \\ \frac{\partial v}{\partial s} & \frac{\partial v}{\partial t} \end{vmatrix} = ad - bc\,.$$

We next consider the square $0 \leq s \leq 1$, $0 \leq t \leq 1$, which has corners with coordinates $(s, t) = (0, 0)$, $(1, 0)$, $(0, 1)$ and $(1, 1)$. Under our transformation, this becomes a parallelogram, say Π, in (u, v)-space, which will have corners with coordinates $(u, v) = (0, 0)$, (a, b), (c, d) and $(a + b, c + d)$ respectively. Thus, the area of Π is precisely

$|ad - bc|$. So, using (6.12), we get:

$$\begin{aligned} m(P(x^a y^c, x^b y^d)) &= \iint\limits_{\Pi} \log|P(e^{2\pi i u}, e^{2\pi i v})| \frac{1}{|ad-bc|} \mathrm{d}u\, \mathrm{d}v \\ &= \frac{1}{|ad-bc|} \iint\limits_{\Pi} \log|P(e^{2\pi i u}, e^{2\pi i v})| \,\mathrm{d}u\, \mathrm{d}v \\ &= \frac{1}{|ad-bc|} |ad-bc| \cdot m(P)\,,\text{ by Proposition 6.4.8,} \\ &= m(P)\,. \end{aligned}$$

□

This pushes us a lot of the way to proving Theorem 6.4.5. However, we have so far only considered values in $\mathbb{Z}$, as opposed to $\mathbb{Q}$. We address this within the proof.

Proof of Theorem 6.4.5. As we know that weakly primitizing a polynomial does not affect its Mahler measure, we can suppose that P and Q are weakly primitive.

Next, say $P(x, y) = \sum_{r,s} a_{r,s} x^r y^s$, where $r, s \in \mathbb{Q}$ and $a_{r,s} \in \mathbb{Z}$, with only finitely many $a_{r,s}$ non-zero. As P and Q are equivalent, we can write:

$$Q(x, y) = \sum_{r,s} a_{r,s} x^{ar+bs+e} y^{cr+ds+f}\,,$$

where $a, b, c, d, e, f \in \mathbb{Q}$ are fixed, and $ad - bc \neq 0$.

Now let g and h be the lowest common denominators of a, b, e and c, d, f respectively. Then, we know in particular that $M(Q(x, y)) = M(Q(x^g, y^h))$, by Lemma 6.4.3. Thus, we get:

$$\begin{aligned} M(Q(x,y)) &= M(Q(x^g, y^h)) \\ &= M(\sum_{r,s} a_{r,s} x^{agr+bgs+eg} y^{chr+dhs+fh}) \\ &= M(\underbrace{\sum_{r,s} a_{r,s} x^{a'r+b's+e'} y^{c'r+d's+f'}}_{=Q_1(x,y)})\,, \end{aligned} \tag{6.13}$$

where $a', b', c', d', e', f' \in \mathbb{Z}$ and $a'd' - b'c' = gh(ad - bc) \neq 0$.

Next, we multiply $Q_1(x, y)$ by $x^{-e'}y^{-f'}$. This does not affect the Mahler, since $M(x^{-e'}y^{-f'}) = 1$. Hence, we have:

$$M(Q(x,y)) = M(\sum_{r,s} a_{r,s} x^{a'r+b's} y^{c'r+d's}) . \tag{6.14}$$

We can then apply Proposition 6.4.10 to (6.14) (it follows trivially that the Mahler measures are equal if the logarithmic Mahler measures are equal). Hence, we have that:

$$\begin{aligned} M(Q(x,y)) &= M(\sum_{r,s} a_{r,s} x^r y^s) \\ &= M(P(x,y)) , \end{aligned} \tag{6.15}$$

as required. □

So we now know that if two (two-variable) polynomials are equivalent, they share the same Mahler measure. This also gives us a level of comfort about our choice of nomenclature.

6.5 Equivalence in Practice

We now have a definition of equivalence and know that equivalent polynomials share the same Mahler measure. However, this is only of use to us if we have a method for checking if two polynomials are equivalent. That is to say, if we have two polynomials P and Q, we need a way to determine if an affine transformation $x^r y^s \mapsto x^{ar+bs+e} y^{cr+ds+f}$, for $a, b, c, d, e, f \in \mathbb{Q}$ and $ad - bc \neq 0$, taking $\pm P(\pm x, \pm y)$ to $Q(\pm x, \pm y)$, actually exists.

In practice, any method we use should allow us to check if two polynomials are equivalent, and then if they are, give us an appropriate affine transformation. We first outline the method which we will use, and then comment on how we can use PARI/GP to check for equivalences.

6.5.1 A Method for Checking Equivalence

Throughout, we will consider polynomials $P(x, y)$ and $Q(x, y)$, and aim to find if there is an affine transformation taking $\pm P(\pm x, \pm y)$ to $Q(\pm x, \pm y)$. As our focus has been on integer polynomials throughout this thesis, we will assume $P, Q \in \mathbb{Z}[x, y]$, although we do not need to make such an assumption in theory.

We will write $P(x, y) = \sum a_i x^{r_i} y^{s_i}$, with $a_i, r_i, s_i \in \mathbb{Z}$, and where all but finitely many of the a_i are non-zero. For each non-zero a_i, we write the powers of the corresponding monomial in the form of a column vector:

$$\begin{pmatrix} r_i \\ s_i \end{pmatrix} .$$

For convenience, we will refer to these vectors as $\mathbf{p}_i$. We similarly write $Q(x, y) = \sum b_j x^{r_j} y^{s_j}$ and write powers of the corresponding monomials of non-zero b_j as column vectors, referred to as $\mathbf{q}_j$.

We then fix three monomials from P, $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$, such that $\mathbf{p}_1 - \mathbf{p}_3$ and $\mathbf{p}_2 - \mathbf{p}_3$ are linearly independent. These may not be the 'first' three monomials as written (as, depending on how these are written, such monomials may not be suitable), but for convenience, these can be chosen in most cases. From here, we fix two matrices:

$$A = (\mathbf{p}_1 - \mathbf{p}_3 \,|\, \mathbf{p}_2 - \mathbf{p}_3) \text{ and } B = (\mathbf{p}_1 \,|\, \mathbf{p}_2 \,|\, \mathbf{p}_3) . \tag{6.16}$$

For avoidance of doubt, A is a 2×2 matrix, whilst B is a 2×3 matrix.

We now need to vary over all possible transformations. To do this, we have to look at all possible triples of monomials from Q, which we write as: $\mathbf{q}_1$, $\mathbf{q}_2$, $\mathbf{q}_3$. For each possible triple, we calculate:

$$N = (\mathbf{q}_1 - \mathbf{q}_3 \,|\, \mathbf{q}_2 - \mathbf{q}_3) . \tag{6.17}$$

We are now in a position to find some of our integer values. In particular, we have that:

$$N\, A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} . \tag{6.18}$$

We refer to the matrix found in (6.18) as M. Owing to the conditions of our affine transformation, we require that $M \in \mathrm{Mat}_2(\mathbb{Q})$ and $|\det(M)| \neq 0$.

Assuming these conditions are met, we next set:

$$n = (\mathbf{q}_1 \,|\, \mathbf{q}_2 \,|\, \mathbf{q}_3) \ . \tag{6.19}$$

We then find the final two values which form our affine transformation:

$$\begin{pmatrix} e & e & e \\ f & f & f \end{pmatrix} = n - M\,B\,. \tag{6.20}$$

We refer to the matrix found in (6.20) as C.

Combining these results together, this gives us our map: $X \longrightarrow M\,X + C$ which takes P to Q.

As described up to now, this method does not tell us the signs of the variables within P and Q for which any affine transformation is valid. However, these are easy to find once we have found the affine transformation.

Such a method can easily be executed within PARI/GP (and other software programs). It is straightforward to create a routine which, when given a polynomial, returns the powers of the monomials as vectors. From here, it is simple to choose suitable $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ such that $\mathbf{p}_1 - \mathbf{p}_3$ and $\mathbf{p}_2 - \mathbf{p}_3$ are linearly independent, and then run over all possible $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ and return suitable results.

6.5.2 Applying the Method

We now look at an example of how this method works in practice. In particular, we look at Example 6.4.6 in much more detail:

Example 6.5.1. Let $P(x,y) = y^8x^8 - y^8x^6 + y^5x^5 + y^3x^3 - x^2 + 1$ and $Q(x,y) = y^2x^8 + yx^8 + yx^5 + yx^3 + y + 1$. In our routine in PARI/GP, we consider these as a vector of column vector entries, which we write as $\mathbf{p}$ and $\mathbf{q}$ respectively:

$$\mathbf{p} = \left(\begin{pmatrix} 8 \\ 8 \end{pmatrix} \begin{pmatrix} 6 \\ 8 \end{pmatrix} \begin{pmatrix} 5 \\ 5 \end{pmatrix} \begin{pmatrix} 3 \\ 3 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right),$$
$$\mathbf{q} = \left(\begin{pmatrix} 8 \\ 2 \end{pmatrix} \begin{pmatrix} 8 \\ 1 \end{pmatrix} \begin{pmatrix} 5 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right).$$

We choose $\mathbf{p}_1 = \begin{pmatrix} 8 \\ 8 \end{pmatrix}$, $\mathbf{p}_2 = \begin{pmatrix} 6 \\ 8 \end{pmatrix}$ and $\mathbf{p}_3 = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$, noting that $\mathbf{p}_1 - \mathbf{p}_3 = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$ and $\mathbf{p}_2 - \mathbf{p}_3 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ are linearly independent, giving us $A = \begin{pmatrix} 1 & 3 \\ 3 & 3 \end{pmatrix}$.

By running over all possible combinations for $\mathbf{q}_1, \mathbf{q}_2$ and $\mathbf{q}_3$, we find that when we have:

$$\mathbf{q}_1 = \begin{pmatrix} 8 \\ 2 \end{pmatrix}, \ \mathbf{q}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ and } \mathbf{q}_3 = \begin{pmatrix} 3 \\ 1 \end{pmatrix},$$

we get an N such that:

$$N\,A^{-1} = \begin{pmatrix} -4 & 3 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Labelling this as M, we note that $|\det(M)| = \frac{1}{2}$. As such, we are now in a position to calculate C:

$$C = \begin{pmatrix} 0 & 8 & 3 \\ 1 & 2 & 1 \end{pmatrix} - \begin{pmatrix} -4 & 3 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 8 & 6 & 5 \\ 8 & 8 & 5 \end{pmatrix}$$
$$= \begin{pmatrix} 8 & 8 & 8 \\ 1 & 1 & 1 \end{pmatrix}.$$

So, we have that $a = -4$, $b = 3$, $c = -\frac{1}{2}$, $d = \frac{1}{2}$, $e = 8$ and $f = 1$. This gives us the affine transformation:

$$x^r y^s \mapsto x^{-4r+3s+8} y^{\frac{-r}{2}+\frac{s}{2}+1}, \tag{6.21}$$

as stated earlier, and a short routine can show us that this takes $-P(x, y)$ to $Q(x, -y)$.

As previously stated, this is not a particularly difficult method to apply, and is very quick. However, as noted, our routine runs over *all* possible combinations for $\mathbf{q}_1, \mathbf{q}_2$ and $\mathbf{q}_3$; in other words, all possible triples of monomials from Q. Therefore, it is possible that there is a different triple of monomials which also gives us an affine transformation.

In fact, since our routine is quick, and does not need to stop after finding a single affine transformation, we can actually find all possible triples which result in an affine transformation.

Example 6.5.1 (Continued)**.** With the same $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ as before, the values

$$\mathbf{q}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ \mathbf{q}_2 = \begin{pmatrix} 8 \\ 1 \end{pmatrix} \text{ and } \mathbf{q}_3 = \begin{pmatrix} 5 \\ 1 \end{pmatrix},$$

give us matrices N_1 and n_1 such that:

$$M_1 = N_1\, A^{-1} = \begin{pmatrix} 4 & -3 \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \text{ and } C = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

As a result, this gives as the strikingly similar affine transformation:

$$x^r y^s \mapsto x^{4r-3s} y^{\frac{r}{2} - \frac{s}{2} + 1},$$

which also takes $-P(x, y)$ to $Q(x, -y)$.

We note that the matrix M_1 in this second instance is in fact $-M$ from the first instance of Example 6.5.1. This may come as a surprise at first, however, there is an explanation for this.

Firstly, we take a closer look at the matrices associated to M and M_1. We recall that we have fixed A, and we have seen that $M = -M_1$. As such, we have that:

$$\begin{aligned} M &= -M_1 \\ NA^{-1} &= -N_1 A^{-1} \\ \Rightarrow N &= -N_1\,. \end{aligned}$$

That is to say, the matrices formed when we run over all possible triples from our target polynomial only differ by sign.

This has occurred because our target polynomial is reciprocal. As such, if we have a monomial represented by the column vector $\begin{pmatrix} r_i \\ s_i \end{pmatrix}$, then there is also a monomial in our polynomial Q which can be represented by the column vector $\begin{pmatrix} r_m - r_i \\ s_m - s_i \end{pmatrix}$, where $r_m = \deg_x(Q)$ and $s_m = \deg_y(Q)$ (when we consider Q as a linear polynomial with all powers of all monomial non-negative). This means that if we can find a suitable triple $\mathbf{q}_1$, $\mathbf{q}_2$, $\mathbf{q}_3$ to give us an N, we can find a second suitable triple to give us $-N$. As our focus throughout this thesis has been on reciprocal polynomials, this is a common occurrence in our examples.

This does not mean that these are necessarily the only other affine transformations though. Indeed, returning back to Example 6.5.1, we will see that there is indeed a somewhat different transformation too.

Example 6.5.1 (Continued)**.** Again with the same $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ as in the first instance, the values

$$\mathbf{q}_1 = \begin{pmatrix} 8 \\ 2 \end{pmatrix}, \ \mathbf{q}_2 = \begin{pmatrix} 8 \\ 1 \end{pmatrix} \text{ and } \mathbf{q}_3 = \begin{pmatrix} 5 \\ 1 \end{pmatrix},$$

give us matrices N and n such that:

$$N\,A^{-1} = \begin{pmatrix} 0 & 1 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \text{ and } C = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

As a result, this gives us the affine transformation:

$$x^r y^s \mapsto x^s y^{-\frac{r}{2}+\frac{s}{2}+1}.$$

which, again, also takes $-P(x, y)$ to $Q(x, -y)$.

This is somewhat more striking. Yet, the key takeaway is that regardless of the monomials used, and the resulting affine transformation, we are seeing the same adjusted version of P being taken to same adjusted version of Q. However, this is not the norm: indeed, there are situations where we have equivalent polynomials, but there do not exist these 'different' transformations.

Example 6.5.2. Consider Laurent polynomials $P(x,y) = y^3x^3+yx^2-x^2+y^3x-y^2x-1$ and $Q(x,y) = x + x^{-1} + y^3 + y^{-3} + yx^2 + y^{-1}x^{-2}$.

Choosing

$$\mathbf{p}_1 = \begin{pmatrix} 3 \\ 3 \end{pmatrix}, \ \mathbf{p}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ and } \mathbf{p}_3 = \begin{pmatrix} 2 \\ 0 \end{pmatrix},$$

we get $A = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}$.

The vectors:

$$\mathbf{q}_1 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \ \mathbf{q}_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \text{ and } \mathbf{q}_3 = \begin{pmatrix} -2 \\ -1 \end{pmatrix}$$

give us matrices N and n such that:

$$M = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \text{ and } C = \begin{pmatrix} 0 & 0 & 0 \\ 3 & 3 & 3 \end{pmatrix}.$$

This gives us an affine transformation mapping $x^r y^s \mapsto x^{-r+s}y^{r+s-3}$, taking $P(-x,y)$ to $Q(x,-y)$.

Meanwhile, the vectors:

$$\mathbf{q}_1 = \begin{pmatrix} 0 \\ -3 \end{pmatrix}, \ \mathbf{q}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \mathbf{q}_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

give us matrices N_1 and n_1 such that:

$$M = \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix} \text{ and } C = \begin{pmatrix} 0 & 0 & 0 \\ -3 & -3 & -3 \end{pmatrix}.$$

This gives us an affine transformation mapping $x^r y^s \mapsto x^{r-s}y^{-r-s+3}$, which also takes $P(-x,y)$ to $Q(x,-y)$.

Example 6.5.2 demonstrates much more intuitively how polynomials which are reciprocal will give rise to different affine transformations. Moreover, it shows that this method can also be used for Laurent polynomials. Indeed, it is exploiting this that helps clearly show why reciprocal polynomials give us these 'pairs' of affine transformations. Equally, it shows a case where we only have one truly different transformation.

6.5.3 Automorphism Groups of Polynomials

We have demonstrated a method which not only can find the associated transformations between two equivalent polynomials P and Q, but can find all possible transformations between these polynomials. These transformations will take the same adjusted version of P to the same adjusted version of Q in each instance.

However, a natural question to ask is what happens if we were to consider the *same* polynomial. Indeed, it would be intuitive to say that some polynomial P is equivalent to itself. From what we have seen so far, though, we cannot say how many different affine transformations exist which would map P to itself.

Definition 6.5.3. *The* ***automorphism group of a polynomial*** P*, denoted* $\mathrm{Aut}(P)$*, is the set of affine transformations taking* P *to* P*, coupled with the standard composition operation.*

This may seem like a strange piece of nomenclature at first. However, we can see that the set of any such collection of affine transformations which take P and send it to itself will form a group, which is a subgroup of the infinite group of affine transformations over $\mathbb{Q}$. And because these transformations are mapping a polynomial to itself, it is sensible enough to refer to such a group as an automorphism group.

Example. Let $P(x, y) = x + x^{-1} + y + y^{-1} + y^4x^5 + y^{-4}x^{-5}$. We find that there are exactly two affine transformations which take $P(x, y)$ to $P(x, y)$:

$$x^r y^s \mapsto x^r y^s \,;$$

$$x^r y^s \mapsto x^{-r} y^{-s} \,.$$

As such, we have that $\mathrm{Aut}(P) \cong \mathbb{Z}/2\mathbb{Z}$.

Of course, in this specific Example, it should come as no surprise that these are indeed two such affine transformations.

Having now seen an example of the automorphism group of a polynomial, it is reasonable to ask what the purpose of such an object is. In short, the hope would be that by introducing this, it may give us some further insight for further problems, and we will refer back to this concept later.

For now, we will look at some more examples. In particular, as our focus is on polynomials with small Mahler measure, and seeing when polynomials share the same Mahler measure, we will look at polynomials $P \in \mathbb{Z}[x, y]$ such that $M(P)$ is small. We know that there are currently 61 known small Mahler measure values coming from polynomials of dimension two. Furthermore, McKee and Smyth, [25, Table D.2] give reciprocal Laurent polynomials for each of these values. We will make use of these polynomials in particular for our examples.

Example 6.5.4. Let $P(x, y) = x + x^{-1} + 1 + y + y^{-1}$, which has Mahler measure $M(P) = 1.28573486\cdots$. We have that there are eight affine transformations taking $P(x, y)$ to $P(x, y)$; these map $x^r y^s$ to one of $x^{\pm r} y^{\pm s}$ or $x^{\pm s} y^{\pm r}$.

From here, it is not difficult to see that $\mathrm{Aut}(P) \cong D_4$, the dihedral group of order 8.

Example. Let $P(x, y) = x^{10} + x^{-10} + y^{10} + y^{-10} + x^9 y^9 + x^{-9} y^{-9}$, which has Mahler measure $M(P) = 1.35246806\cdots$. We have that there are four affine transformations taking $P(x, y)$ to $P(x, y)$; these map $x^r y^s$ to one of $x^r y^s$, $x^{-r} y^{-s}$, $x^s y^r$ or $x^{-s} y^{-r}$.

From here, it is not difficult to see that $\mathrm{Aut}(P) \cong V_4$, the Klein-4 group.

We end by noting that all of the 61 polynomials listed by McKee and Smyth in this format have automorphism group isomorphic to one of the three groups exhibited here. More specifically, Example 6.5.4 is the only one of their polynomials with automorphism group isomorphic to D_4, forty-three have automorphism group isomorphic to V_4 and the remaining seventeen have automorphism group isomorphic to $\mathbb{Z}/2\mathbb{Z}$.

6.6 Equivalence II: Longing for More

We have been able to show that if two polynomials are equivalent, they share the same Mahler measure. This is a useful tool, but there is a perhaps glaring issue with our definition of equivalence: it relies on the two polynomials having the same number of monomials. Our affine transformation will map one monomial to another monomial, but cannot "create" new ones.

To simplify this idea, we introduce the following definition:

Definition 6.6.1. *The* ***term-length*** *of a two-variable polynomial P, denoted $\ell(P)$, is the number of (non-zero) monomial terms in P.*

So, to re-word the highlighted issue: if P and Q do not have the same term-length, they cannot be equivalent. So whilst we do have a useful, and powerful, tool for deciding if two polynomials share the same Mahler measure, we are still a long way off from giving a complete description of when two two-variable polynomials share the same Mahler measure (which, of course, is our main goal).

At this point, our concerns about polynomials with different term-length sharing the same Mahler measure are purely theoretical; we have not seen polynomials P and Q with $\ell(P) \neq \ell(Q)$, but $M(P) = M(Q)$. The following example shows that this is possible:

Example. Let $P(x,y) = y^2x^4 + (-y^2 - y)x^3 + yx^2 + (-y-1)x + 1$ and $Q(x,y) = y^2x^8 - yx^7 + (-y^2-1)x^4 - yx + 1$.

We have that $\ell(P) = 7$ and $\ell(Q) = 6$, and $M(P) = M(Q) = 1.25543386\cdots$.

So, our concern is very much a real one. As such, our main focus now should be trying to understanding how two polynomials of different term-length can share the same Mahler measure.

One may wonder if the automorphism groups of polynomials possibly give some explanation to what is happening here. For example, it may be the case that whenever $M(P) = M(Q)$, $\mathrm{Aut}(P) \cong \mathrm{Aut}(Q)$. However, this is not the case:

Example. If we let $P(x,y) = y^2x^4 + (-y^2 - y)x^3 + yx^2 + (-y-1)x + 1$ and $Q(x,y) = y^2x^8 - yx^7 + (-y^2-1)x^4 - yx + 1$. We have seen from above that it seems to be the case that $M(P) = M(Q)$ (and indeed, we confirm this in Example 6.6.2). However, we find that $\mathrm{Aut}(P) \cong D_6$ (the dihedral group of order 12) and $\mathrm{Aut}(Q) \cong V_4$.

6.6.1 A Look at Subsequences

We recall Proposition 1.2.4, which states that:

$$M(P(x,y)) = \lim_{n\to\infty} M(P(x,x^n)) \,. \tag{6.22}$$

Thinking in these terms, we can consider a sequence of values $M(P(x,x^n))$ for each n, treating the limit of this sequence as the Mahler measure of the two-variable polynomial.

Since we know our sequence does indeed converge to a limit, we can take any subsequence, and know that will also converge to the same limit. This somewhat trivial statement can give us some level of understanding as to why two polynomials of different term-lengths share the same Mahler measure.

Example 6.6.2. Let $P(x,y) = y^2x^4 + (-y^2 - y)x^3 + yx^2 + (-y-1)x + 1$ and $Q(x,y) = y^2x^8 - yx^7 + (-y^2 - 1)x^4 - yx + 1$, as before.

Now consider $M(P(x,x^n))$ and $M(Q(x,x^n))$ for each n. We look at a selection of explicit values for some small values of n in each case:

n	$M(P(x,x^n))$	n	$M(Q(x,x^n))$
1	$1.55603019\cdots$	3	$1.55603019\cdots$
3	$1.35098033\cdots$	7	$1.35098033\cdots$
7	$1.21972085\cdots$	15	$1.21972085\cdots$
10	$1.17628081\cdots$	21	$1.17628081\cdots$
11	$1.24949933\cdots$	23	$1.24949933\cdots$
12	$1.26911783\cdots$	25	$1.26911783\cdots$

These values, and in fact all values we experiment on, give us reason to suspect the following claim:

$$M(P(x,x^n)) = M(Q(x,x^{2n+1})) \,, \tag{6.23}$$

for each n.

Of course, at this stage, we are not claiming with absolute certainty that this is actually true for all n. Rather, we are simply saying that we have the empirical evidence to suggest that this could be the case.

The importance of (6.23), and the subsequent claim, is that the values $M(Q(x,x^{2n+1}))$ form a subsequence of $M(Q(x,x^n))$. As such, we can make the conclusion that $\lim_{n\to\infty} M(Q(x,x^{2n+1})) = M(Q(x,y))$. However, since this subsequence of values coincides with the sequence of values of $M(P(x,x^n))$, we know that this is the same as $M(P(x,y))$. In other words, we can conclude that $M(P) = M(Q)$.

We now return back to Example 6.6.2, with a view of verifying that (6.23) is true for all n.

Example 6.6.2 (Continued)**.** We have that:

$$\begin{aligned} P(x,x^n) &= x^{2n+4} - x^{2n+3} - x^{n+3} + x^{n+2} - x^{n+1} - x + 1\,, \\ Q(x,x^{2n+1}) &= x^{4n+10} - x^{4n+6} - x^{2n+8} - x^{2n+2} - x^4 + 1\,. \end{aligned} \tag{6.24}$$

We next note that $Q(x,x^{2n+1})$ can be factored:

$$\begin{aligned} Q(x,x^{2n+1}) &= (x^2+1)\left(x^{4n+8} - x^{4n+6} - x^{2n+6} + x^{2n+4} - x^{2n+2} - x^2 + 1\right) \\ &= (x^2+1)p(x^2)\,, \end{aligned} \tag{6.25}$$

where $p(x) = P(x,x^n)$.

We know that $M(p(x^2)) = M(p(x))$, and so not only does this verify $M(P(x,x^n)) = M(Q(x,x^{2n+1}))$ for each positive integer n, but also gives us a link between the polynomials themselves. Thus, by our explanation from earlier, we have shown that $M(P) = M(Q)$.

This highlights a particularly interesting situation. For the polynomials P and Q, we have infinitely many situations where $(x^2+1)P(x^2,x^{4n+2}) = Q(x,x^{2n+1})$. This phenomenon explains precisely why the values of $M(P)$ and $M(Q)$ are indeed exactly the same. The natural next step after seeing this is to ask if this is a one-off coincidence, or if this is perhaps just one example of something more powerful.

As a matter of fact, we will see that this appears to exhibit something more powerful.

Before stating the over-arching concept which explains how polynomials of different term-lengths share the same Mahler measure, we look at some more examples. This will help give further heuristic justification to our idea when we arrive at it.

Example. Let:

$$P(x,y) = (y^2+y)(x^{10}+x^9+x^8+x^7)+(y^2+y+1)(x^6+x^5+x^4)+(y+1)(x^3+x^2+x+1)\,,$$

$$Q(x,y) = y^2x^{28}+y^2x^{25}+(-y^2+1)x^{14}-yx^3-1\,.$$

For each n, we have that:

$$(x-1)(x+1)P(x^2,x^{2n}) = Q(x,x^{2n-3})\,,$$

and so $M(P) = M(Q)$.

Example 6.6.3. Let:

$$P(x,y) = x^8+x^7+x^6+x^5+(y^2+y+1)x^4+y^2x^3+y^2x^2+y^2x+y^2\,,$$
$$Q(x,y) = y^2x^{20}+yx^{11}+(y^2+1)x^{10}+yx^9+1\,.$$

For each n, we have that:

$$(x^2+1)P(-x^2,x^{4n}) = x^8Q(x,x^{4n-9})\,. \tag{6.26}$$

and so $M(P) = M(Q)$.

This case perhaps looks slightly more convoluted than our previous ones. With Q, we are simply replacing y with x^{4n-9}, and so find a subsequence of values. With P, we first replace y with x^{2n}, also finding a subsequence of values. Then, we replace x with $-x^2$ (a change which still does not affect the Mahler measure). So, we have that $P(x,x^{2n}) \mapsto P(-x^2,(-x^2)^{2n}) = P(-x^2,x^{4n})$.

The following is a good start to seeing explicitly that (6.26) does indeed hold, although we omit the full calculation:

$$P(-x^2, x^{4n}) = (x^{8n} + x^8)(x^8 - x^6 + x^4 - x^2 + 1) + x^{4n+8},$$
$$Q(x, x^{4n-9}) = x^{8n+2} + x^{8n-8} + x^{4n+2} + x^{4n} + x^{10} + 1.$$

6.6.2 Further Ideas and Examples

The preceding Examples have demonstrated the loose idea which could well describe when two polynomial share the same Mahler measure. What we have seen is that if we expect two two-variable polynomials, say P and Q, to share the same Mahler measure, then we can implement the following strategy:

1. Consider the sequences $M(P(x, x^n))$ and $M(Q(x, x^n))$; these are sequences of Mahler measure values of single variable polynomials.

2. Attempt to find subsequences of both of these sequences which agree.

3. Attempt to find an equality of polynomials, with suitable substitutions and Kronecker-cyclotomic factors included, related to these corresponding subsequences.

In practice, it is usually fairly quick and straightforward to find the Mahler measure of a specific polynomial to a good level of accuracy (as we have seen in chapter 2). As such, this strategy need only be implemented on polynomials where we expect the Mahler measure values to be the same. In turn, our attempts to find agreeable subsequences – if this strategy indeed always works – will be successful.

Moreover, our subsequences have only ever been defined by a linear relation of the form $s_1 n + s_2$, for integers s_1 and s_2. We have seen situations where we may only need to find the subsequence of one of our original sequences, and also both. We have also seen that s_2 need not be strictly positive.

When we attempt to find an equality of polynomials related to the corresponding subsequences, we note that we are restricted to "operations" which do not change the Mahler measure of a single variable polynomial. This allows us to multiply by Kronecker-cyclotomic polynomials and either change the sign of a polynomial or the sign of the variable of a polynomial. Furthermore, since we know that weakly primitizing a polynomial does not change the Mahler measure, we can also reverse this process, and replace x with x^t, where t is some positive integer.

We are now in a position to state and prove the following result:

Proposition 6.6.4. *Let P and Q be polynomials such that there exists Kronecker-cyclotomic polynomials κ_1 and κ_2, and integers t, a_1, a_2, b_1 and b_2 such that:*

$$\kappa_1(x)P(\pm x^t, \pm x^{a_1 n+a_2}) = \pm\kappa_2(x)Q(\pm x, \pm x^{b_1 n+b_2}), \tag{6.27}$$

for each $n \in \mathbb{N}$.

Then, $M(P) = M(Q)$.

Proof. Firstly, we know from Lemma 6.4.2 that for any choice of sign, $M(\pm P(\pm x, \pm y)) = M(P(x,y))$ (and similarly for Q). So it suffices to only consider the case where all signs are $+1$.

We then make use of Proposition 1.2.4, to give us:

$$\lim_{n\to\infty} M(P(x,x^n)) = M(P(x,y)).$$

We treat $P(x,x^n)$ as a single variable polynomial. We know that replacing x with x^t, for some positive integer t, does not affect the Mahler measure of a polynomial. So, $M(P(x,x^n)) = M(P(x^t,x^{nt}))$. In turn, we have:

$$\lim_{n\to\infty} M(P(x^t,x^{nt})) = M(P(x,y)). \tag{6.28}$$

Moreover, we also know that multiplying a polynomial by a Kronecker-cyclotomic polynomial, say κ_1, does not affect the Mahler measure. So, we have that $M(P(x,x^n)) = M(\kappa_1(x)P(x^t,x^{nt}))$. Therefore,

$$\lim_{n\to\infty} M(\kappa_1(x)P(x^t,x^{nt})) = M(P(x,y)).$$

We now take a subsequence of the natural numbers, say $A_1 n + A_2$. In turn, we can now consider a sequence of values

$$M\left(\kappa_1(x)P(x^t, x^{(A_1n+A_2)t})\right)\,, \tag{6.29}$$

which is a subsequence of the sequence of values $M\left(\kappa_1(x)P(x^t, x^{nt})\right)$. As we know that this is a convergent sequence, the sequence of values shown in (6.29) will converge to the same limit point. This hence gives us:

$$\lim_{n\to\infty} M\left(\kappa_1(x)P(x^t, x^{(A_1n+A_2)t})\right) = M(P(x,y))\,.$$

If we now let $(A_1n + A_2)t = a_1n + a_2$, we can then use (6.27) and (6.28) to find that:

$$\lim_{n\to\infty} M\left(\kappa_2(x)Q(x, x^{b_1n+b_2})\right) = M(P(x,y))\,, \tag{6.30}$$

where κ_2 is Kronecker-cyclotomic and $b_1, b_2 \in \mathbb{Z}$.

This now gives us the limit of the Mahler measures of polynomials related to Q in terms of $M(P)$. However, we can also determine this limit in a different way as well. So, now let

$$\lim_{n\to\infty} M\left(\kappa_2(x)Q(x, x^{b_1n+b_2})\right) = L\,,$$

where L is a limit point value we are attempting to find.

We know that multiplying a polynomial by a Kronecker-cyclotomic polynomial does not affect the Mahler measure, so we choose to ignore this factor. Thus, we have:

$$L = \lim_{n\to\infty} M\left(Q(x, x^{b_1n+b_2})\right)\,.$$

Furthermore, we can clearly see that the sequence with values $M\left(Q(x, x^{b_1n+b_2})\right)$ is a subsequence of $M\left(Q(x, x^n)\right)$. We know that this main sequence converges; in particular:

$$\lim_{n\to\infty} M\left(Q(x, x^n)\right) = M(Q(x,y))\,,$$

and so $\lim_{n\to\infty} M\left(Q(x, x^{b_1 n+b_2})\right) = M(Q(x,y))$. This therefore gives us that:

$$L = M(Q(x,y)) .$$

Combining this with (6.30) shows that $M(P) = M(Q)$. □

This is not the most powerful result, but this is at least a tool we can use to see if two polynomials do share the same Mahler measure. However, if this fails, we do not have enough certainty either way to say whether the Mahler measures are different or not. In turn, we would want to know if the reverse of this statement is true, which would indeed be a more powerful result too. Unfortunately, this is not the case:

Non-Proposition 6.6.5. *If $M(P) = M(Q)$, then there exists Kronecker-cyclotomic polynomials κ_1 and κ_2, and integers t, a_1, a_2, b_1 and b_2 such that:*

$$\kappa_1(x)P(\pm x^t, \pm x^{a_1 n+a_2}) = \pm\kappa_2(x)Q(\pm x, \pm x^{b_1 n+b_2}) , \tag{6.31}$$

for each $n \in \mathbb{N}$.

We now aim to exhibit a situation showing that this result does not hold.

Example 6.6.6. Let $P(x,y) = y^8x^8 - y^8x^6 + y^5x^5 + y^3x^3 - x^2 + 1$ and $Q(x,y) = y^2x^8 + yx^8 + yx^5 + yx^3 + y + 1$. By Example 6.4.6, we know that P and Q are equivalent, and so $M(P) = M(Q)$.

Trying to prove that there cannot exist Kronecker-cyclotomic polynomials κ_1 and κ_2 and integers t, a_1, a_2, b_1 and b_2 such that (6.31) does not hold is a difficult task. Instead, we take a different, albeit not rigorous, approach.

Consider $M(P(x,x^i))$ and $M(Q(x,x^j))$ for $i,j \in [1,150]$. In this range, we do not find any pair of i and j such that $M(P(x,x^i)) = M(Q(x,x^j))$. Our ideas related to when single variable polynomials share the same Mahler measure, as seen in Section 6.1, would suggest that if we did find two polynomials which shared the same Mahler measure, then we would be able to express them in a form such as in (6.31). However, as we cannot find two polynomials sharing a Mahler measure, then we cannot express them in the form of (6.31).

Exploring over $i, j \in [1, 150]$ gives us a level of empirical evidence that we will not find any polynomials sharing the same Mahler measure. Certainly from a practical standpoint, if we were to find subsequences of $M(P(x, x^n))$ and $M(Q(x, x^n))$ which were to coincide, we would expect to find have found one instance within this explored range. Indeed, for all examples we have seen where there are coinciding subsequences, we do see evidence of this by exploring over similar ranges.

As such, we have a situation where we have polynomials P and Q such that $M(P) = M(Q)$, but we do not expect (6.31) to hold.

There are some obvious issues present in Example 6.6.6. Firstly, it has not been rigorously shown that Non-Proposition 6.6.5 does indeed fail. Instead, it only presents evidence – although fairly convincing within context – that it will fail. Secondly, we are relying somewhat on Open Problem 6.1.1 holding true for the presented evidence to be convincing. This does mean that whilst, with all information available to us, we do not expect this to hold, there is an air of uncertainty around our claims.

If Example 6.6.6 was the only situation where we had equivalent polynomials and could not find suitable Kronecker-cyclotomic polynomials and integers for (6.31) to hold, this would certainly increase our uncertainty. However, we can find plenty of other examples where we have equivalent polynomials but are unable to show that they satisfy (6.31):

Example. The following pairs of polynomials P and Q are equivalent, but do not appear to satisfy our conditions:

i. $P(x, y) = y^5x^5 + y^4x^4 - y^5x^3 + x^2 - yx + 1\,,$
$Q(x, y) = y^2x^5 + yx^5 - yx^4 - yx + y + 1\,.$

ii. $P(x, y) = y^4x^2 - y^4x + y^3x + yx - x + 1\,,$
$Q(x, y) = y^2x^4 + yx^4 + yx^3 + yx + y + 1\,.$

This may lead us to consider the possibility that potential results depend entirely on if the term-lengths of our polynomials are the same. In other words, the reverse of Proposition 6.6.4 may only hold when $\ell(P) \neq \ell(Q)$. However, this is not the case:

Example 6.6.7. Let

$$P(x,y) = y^2x^6 + yx^6 + y^2x^5 + x + y + 1\,;$$
$$Q(x,y) = yx^6 + y + y^2x^3 + x^3 + y^2x^4 + x^2\,.$$

We find that P and Q are equivalent; for example, we have the affine transformation $x^r y^s \mapsto x^{-r+2s+4}y^{-s+2}$ taking $P(x,y)$ to $Q(x,y)$. Furthermore we have that:

$$x^2P(x,x^n) = Q(x,x^{n+2})\,.$$

So, Example 6.6.7 dispels the possibility that our introduced notion of equivalence is mutually exclusive with the concept we have introduced here. Unfortunately, this is of little solace to us.

6.6.3 To Equivalence and Beyond

We end this chapter by posing open problems which relate to the topic of when two two-variable polynomials share the same Mahler measure. These, in turn, give a summary of the work seen so far, and the issues encountered.

Open Problem 6.6.8. *Let $P, Q \in \mathbb{Z}[x,y]$. Is there a way to determine if $M(P) = M(Q)$?*

We have gone some way into answering this problem, by introducing our notion of equivalence, as well as going further by looking at subsequences of Mahler measure values. However, we do not know if these encapsulate all situations; there may be other ways to determine equality of Mahler measures.

Open Problem 6.6.9. *Let $P, Q \in \mathbb{Z}[x,y]$ and $M(P) = M(Q)$. Are P and Q, in some sense, the same?*

This is a vaguer statement, and in many ways, just posing Open Problem 6.6.8 in a different way. However, this statement is an attempt at mirroring the statement of Open Problem 6.1.1. Our main problem for single variable polynomials gives a list of conditions on polynomials f and g which we believe guarantee that $M(f) \neq M(g)$.

In other words, if $M(f) = M(g)$, then the polynomials are "the same", up to these conditions.

Open Problem 6.6.9 asks if there is a similar formulation for two-variable polynomials. That is, if we have polynomials P and Q, then the polynomials are "the same", up to a set of given conditions. Based on our work so far, a more precise formulation of this statement would be:

Open Problem 6.6.10. *Let $P, Q \in \mathbb{Z}[x, y]$ and $M(P) = M(Q)$. Then, at least one of the following holds:*

1. *P and Q are equivalent.*
2. *There exist Kronecker-cyclotomic polynomials κ_1 and κ_2, and integers t, a_1, a_2, b_1 and b_2 such that (6.27) holds for each $n \in \mathbb{N}$.*

The wording of this Open Problem is somewhat daring, as we are suggesting that there are only two possible scenarios which explain when two-variable polynomials share the same Mahler measure. However, our evidence for this is that in every case where we found $M(P) = M(Q)$ numerically, at least one of the two conditions in Open Problem 6.6.10 has been found to hold. This certainly gives us reason to believe that the statement as presented could indeed be true.

Chapter 7

An Epic Epilogue of Open Problems

To round off, we present a selection of Open Problems and unanswered questions which have been posed throughout. There is no new content introduced here; rather, we aim to summarize everything we have not been able to fully answer, and questions which were beyond the scope of this thesis. However, some of these questions / problems may not have been explicitly stated previously. We make reference back to appropriate Sections for each point raised.

In doing this, we make clear potential avenues for future work following from this thesis.

7.1 Calculating Mahler Measures

Broadly speaking, the central concern and motivation for studying Mahler measures has been Lehmer's Problem:

Conjecture 7.1.1 (See Conjecture 1.1.10, in Section 1.1.1)**.** *For all polynomials $f \in \mathbb{Z}[x]$ such that $M(f)$ is non-trivial, we have that $M(f) \geq 1.17628081 \cdots = \lambda$.*

This thesis has not attempted to prove or disprove this. However, we hope that some of the results and ideas introduced provide some additional context and support for believing this to be true.

Question 7.1.2 (See Sections 1.1.1 and 5.3.1)**.** *Can we find any new small Mahler measure values from single variable polynomials?*

We have been able to find a new small Mahler measure value: $1.252826882865\cdots$. Whilst our aim was not specifically to find new small Mahler measures, it is certainly a nice by-product of our work, and any future experimental work that may follow!

Question 7.1.3 (See Section 1.2.1)**.** *Is the list of tiny Mahler measures complete?*

Question 7.1.4 (See Section 1.2.1)**.** *Is the list of small limit points of Mahler measures complete?*

Again, this thesis has not attempted to answer these questions directly. However, our experiments have given us a different way of finding these values, and at the very least, help convince us that the lists are likely near-complete, if not already.

Question 7.1.5 (See Section 2.2)**.** *Can we improve upon methods for calculating the Mahler measure of two-variable polynomials?*

We presented a new method for calculating the Mahler measures of two-variable polynomials (albeit one which is only applicable for reciprocal polynomials), which is quick and efficient. It would be useful to know if we can improve this, or one of the other known methods, to have a more general, quicker and more efficient method.

7.2 Mahler Measures and Digraphs

Open Problem 7.2.1 (See Section 3.3.1)**.** *Can we give a classification of all cyclotomic (charged signed) digraphs?*

We have seen that we have a classification of cyclotomic charged signed graphs, but a classification remains unknown when we extend to digraphs.

Question 7.2.2 (See Sections 4.4 and 5.4)**.** *Can we create and implement an efficient routine which calculates the Mahler measures of "bridged digraphs", and families of bridged digraphs?*

We extended our construction of growing digraphs to, effectively, attach digraphs to the ends of pendant paths which themselves are already attached to a 'central' digraph. We chose not to experiment with this construction, partially because of the difficulty of implementing it. However, if this could be implemented successfully, it may help improve the results presented within this thesis.

Question 7.2.3 (See Section 5.3.1)**.** *Can we find all tiny Mahler measures from digraphs? Further, can we find all known small Mahler measures (whose associated polynomial is of degree at most* 180*) from digraphs?*

Question 7.2.4 (See Section 5.3.2)**.** *Can we find all small limit points of Mahler measures from digraphs?*

Our experiments have found all but one of the (known) tiny Mahler measures, approximately 97% of the known small Mahler measures whose associated polynomial is of degree at most 180 and all but four of the (known) small limit points of Mahler measures. This included one new small Mahler measure value. With access to more powerful computers, more efficient routines and more time, we may be able to find the remaining known values, and possibly even more new values.

7.3 Polynomials with the Same Mahler Measures

Open Problem 7.3.1 (See Open Problem 6.1.1, in Section 6.1)**.** *Let $f, g \in \mathbb{Z}[x]$ be monic, weakly primitive, irreducible, reciprocal polynomials, with non-trivial Mahler measures.*

Is it the case that $M(f) \neq M(g)$?

Open Problem 7.3.2 (See Open Problem 6.2.6, in Section 6.2)**.** *Let $f, g \in \mathbb{Z}[x]$ be monic, weakly primitive, irreducible, reciprocal polynomials, with non-trivial Mahler measures. Let f have exactly one large zero and be the polynomial of least degree with Mahler measure $M(f)$, and g have exactly two large zeros.*

Is it the case that $M(f) \neq M(g)$?

Question 7.3.3. *Can we answer Open Problem 6.2.6 for f having n large zeros and g having m large zeros, for other fixed values n and m?*

We have justified our reasoning for these stating these problems with these conditions, and have sufficient empirical evidence to believe that these will be true. However, we still have yet to be able to prove any of these statements. We expect that Question 7.3.3 can only be properly posed and answered if we get an answer to Open Problem 6.2.6.

Question 7.3.4 (See Section 6.5.3)**.** *Can the automorphism group of a (two-variable) polynomial be used to greater effect than already presented?*

We introduce the concept of an automorphism group of a polynomial, in hope that it may help give us further insight to other problems. Unfortunately, it does not give us much in the way of answering any questions and problems we raise. On the other hand, we hope that there may be some use to the concept, or a possible related concept.

Question 7.3.5 (See Open Problems 6.6.8, 6.6.9 and 6.6.10 in Section 6.6.3)**.** *Let P and Q be two-variable polynomials such that $M(P) = M(Q)$. Are there conditions on P and Q which can explain this?*

The question of when two two-variable polynomials share the same Mahler measure has proved to be significantly more involved than the single variable scenario. There are many ways of posing and attempt to answer this general question, and our ideas give a possible explanation, and are supported again by a plethora of empirical evidence.

Appendix A

Knot Theory: Tying Up Loose Ends

Here, we aim to give a brief overview of what the pretzel knot is, what it looks like and a definition of the Alexander polynomial, to help aid with the visualisation and understanding of Section 1.3.1. We do not go into the intricate nature of the field; this is here to just give a high level, almost crude, overview for the unfamiliar reader.

A **knot**, K, is a loop in 3-space without self-intersections. In practice, we can think of it like a piece of string. A **twist** is a counter-clockwise movement involving two separate sections of a curve, causing them to overlap. Very crudely speaking, if we were to hold two different pieces of a string, one piece in each hand, a twist would involve us overlapping the piece of string in our right hand over the piece of string in our left (this is a counter-clockwise movement). We can make multiple twists, with the convention that after one twist, the 'left' and 'right' pieces of string have now swapped, and as such have swapped hands. We can also make a negative number of twists, which involves a clockwise movement, rather than counter-clockwise (moving our left hand over our right hand).

The **pretzel knot** is a special example of a knot, which is split into n different sections, called **tangles**. This can be written as $P_K(k_1, k_2, \cdots, k_n)$, where the k_i represent the number of twists in i-th tangle. So, for example, the following is $P_K(-2, 3, 7)$:

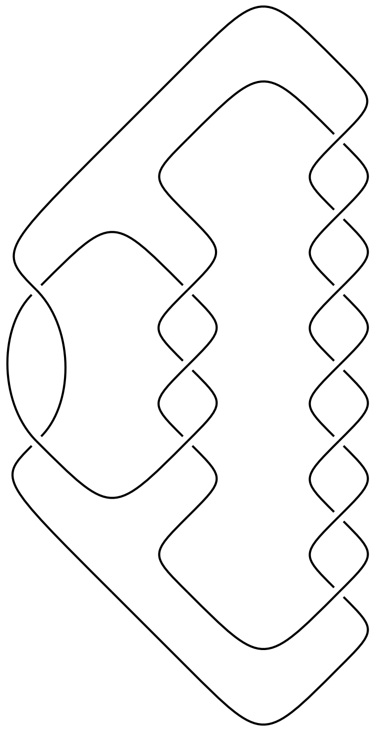


Figure A.1: The pretzel knot $P_K(-2, 3, 7)$.

When we are presented with a diagram of a knot, we make reference to a **crossing** within the knot: this is when a section of the knot goes above another section of the knot. In a sense, it is a section of the knot which has been twisted. The **Alexander polynomial** of a knot K, denoted by Δ_K, is a polynomial that is determined by the number of crossings made in the knot.

Due to the nature of the Alexander polynomial, there is no closed formula for calculating the Alexander polynomial of a pretzel knot. This means that there can be a sense of volatility with the Alexander polynomial of any knot, including those of interest to us, the pretzel knots $P_K(k_1, \cdots, k_n)$, This gives us no guarantees that the polynomial will follow any set pattern or shape.

As such, we give a worked example on how to calculate the Alexander polynomial for a small knot. In short, the method one follows for determining the Alexander polynomial of a knot K is as follows:

1. Choose an **orientation** for our knot (a way to follow the loop),

2. Label each crossing in our knot $1, \cdots, n$,

3. Construct an $n \times n$ matrix, labelled M_K, in the following manner:

    a. At crossing i, identify the first crossing to the left of i, which will be labelled j, where $j \neq i$,

    b. At crossing i, identify the first crossing to the right of i, which will be labelled k, where $k \neq i$ and $k \neq j$,

    c. The (i, i)-entry of M_K is given value $1 - t$,

    d. The (i, j)-entry of M_K is given value t,

    e. The (i, k)-entry of M_K is given value -1,

    f. All other entries in row i are given value 0.

4. Delete the last row and column of M_K, to give an $(n-1) \times (n-1)$ matrix, say N_K,

5. Set $\Delta_K(t) = \det(N_K)$; this is the Alexander polynomial of K.

We note that this method involves making various choices (such as orientation, the labelling of the crossings and so forth). The Alexander polynomial is invariant of these choices; a fact we do not prove, but can be found in any introductory text book about knot theory (see, for example, Reidemeister [32]).

Example A.0.1. Let K be the "figure eight knot", as shown in Figure A.2. We note that we have given this an orientation, marked by the arrowhead, and labelled the crossings. In particular, our crossings have been labelled in a sort of consecutive manner of how they appear when following the loop, though there is no need to actually do this in practice.

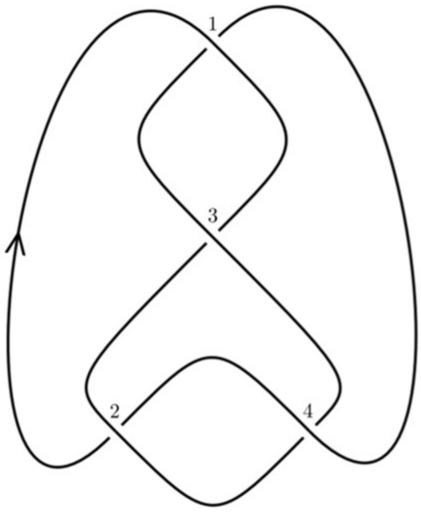


Figure A.2: The figure eight knot.

We now begin to construct our matrix M_K. We will explicitly look at how we construct the first row, and then just state the rest. Starting at crossing 1, we see that following our orientation means we would be 'facing down'. Thus, the first crossing to the left of crossing 1 we encounter is crossing 4, whilst the first crossing to the right of crossing 1 we encounter is crossing 3.

So, we have that the $(1,1)$-entry of M_K is $1-t$, the $(1,4)$-entry is t and finally the $(1,3)$-entry is -1. The remaining entries in the first row are zero, which in this case is just the $(1,2)$-entry.

Repeating the procedure at each crossing gives us:

$$M_K = \begin{pmatrix} 1-t & 0 & -1 & t \\ -1 & 1-t & 0 & t \\ -1 & t & 1-t & 0 \\ 0 & t & -1 & 1-t \end{pmatrix}.$$

Deleting the last row and column of M_K gives us:

$$N_K = \begin{pmatrix} 1-t & 0 & -1 \\ -1 & 1-t & 0 \\ -1 & t & 1-t \end{pmatrix}.$$

Finally, we have that:

$$\begin{aligned}\Delta_K(t) &= -t^3 + 3t^2 - t \\ &= t^2\left(-t + 3 - \frac{1}{t}\right).\end{aligned}$$

We note that when we calculated the Alexander polynomial in this case, we had a linear polynomial where all powers of t were non-negative. However, we also presented a version where we factored out a suitable power of t so that we have a Laurent polynomial whose powers are symmetric about the t^0 term. This follows the convention when working with Alexander polynomials, and some authors may refer to this specific Laurent polynomial as *the* Alexander polynomial of the knot.

As mentioned in Section 1.3.1, studying the Alexander polynomials of pretzel knots is of particular interest to us in the context of Mahler measures. As mentioned, the Alexander polynomials of $P_K(-2, 3, 3)$ and $P_K(-2, 3, 5)$ are Kronecker-cyclotomic, whilst the Alexander polynomial of $P_K(-2, 3, 7)$ is $\Lambda(-t)$ (when we write these as linear polynomials, as opposed to Laurent polynomials). However, due to the lack of closed formula, this is a less productive method for trying to find objects which we can associate polynomials to to find small Mahler measures compared to our work surrounding digraphs, as we see through this thesis.

Appendix B

Lattices: Connecting the Dots

Here, we give an overview of the key definitions and concepts of lattices needed from Section 2.4.1. We do not go into the intricate nature of the field; this is here to just give a high level, almost crude, overview for the unfamiliar reader

A **lattice**, $\mathcal{L}$, is a set of points in n-dimensional space with a periodic structure. Perhaps the simplest example to consider is $\mathbb{Z}^n$ (in $\mathbb{R}^n$); n-dimensional vectors with integer entries. More formally, given m linearly independent vectors $b_1, \cdots, b_m \in \mathbb{R}^n$, for $m \leq n$, we can generate a lattice with them, defined to be:

$$\mathcal{L} = \mathcal{L}(b_1, \cdots, b_m) = \left\{ \sum_m x_i b_i \,|\, x_i \in \mathbb{Z} \right\} ,$$

where we call these vectors the **basis** of the lattice $\mathcal{L}$, and denote this as B. Following the notation from this more formal set up, we say that the **dimension** of a lattice is the value n, and the **rank** is the value m.

We note that a consequence of our definition is that every lattice necessarily has a basis. However, the basis of a lattice usually is not unique. For example, $B_1 = \{(1,0),(0,1)\}$ and $B_2 = \{(1,1),(2,1)\}$ are both bases for $\mathbb{Z}^2$, viewed as a lattice over $\mathbb{R}^2$.

For a lattice $\mathcal{L}$ with a given basis B, we can look to "reduce" the basis B. The method of reduction of interest to use is the **LLL reduction** method, first introduced by Lenstra et al. [18], which we detail shortly.

Before detailing the LLL reduction of a basis, we state the known Gram-Schmidt process, which gives us an orthogonalization of a basis. For a basis $B = \{b_1, \cdots, b_m\}$, define vectors ${b_i}^*$ inductively by:

$$ {b_i}^* = b_i - \sum_{j=1}^{i-1} \nu_{i,j} {b_j}^* \, , $$

where:

$$ \nu_{i,j} = \frac{\langle b_i, {b_j}^* \rangle}{\langle {b_j}^*, {b_j}^* \rangle} \, , \tag{B.1} $$

with $\langle \cdot, \cdot \rangle$ representing the standard inner product on $\mathbb{R}^n$. We call the vectors $B^* = \{{b_1}^*, \cdots, {b_n}^*\}$ the corresponding Gram-Schmidt orthogonal basis.

We are now in a position to define what an LLL reduced basis of a lattice $\mathcal{L}$ is. Let $B = \{b_1, \cdots, b_m\}$ be a basis of $\mathcal{L}$ and $B^* = \{{b_1}^*, \cdots, {b_n}^*\}$ be its orthogonal basis, with $\nu_{i,j}$ as defined in (B.1). Then, B is an **LLL reduced basis** if:

1. $|\nu_{i,j}\,| \leq \frac{1}{2}$, for $1 \leq j < i \leq m$,

2. $|{b_i}^* + \nu_{i,i-1} {b_{i-1}}^*|^2 \geq \frac{3}{4} |{b_{i-1}}^*|^2$, for $1 < i \leq m$.

If our basis B is not LLL reduced, then we can use an algorithm proposed by Lenstra et al. [18] to find an LLL reduced basis – the LLL algorithm – to find one in polynomial time. This algorithm is well known and used. For example, in PARI/GP, we have the command `qflll` to return a marix we can be used to find an LLL reduced basis.

Appendix C

Cyclotomic Charged Signed Graphs

In Theorems 3.3.3 and 3.3.4, we stated every maximal connected cyclotomic signed graph and every maximal connected cyclotomic charged signed graph respectively. After stating these, we only included figures of the charged signed graphs which were necessary for work which followed within the thesis. Here, we include figures for the remaining ones.

As with the statements themselves, these figures appear in McKee and Smyth [23]. They are included here for easy reference and a sense of completeness.

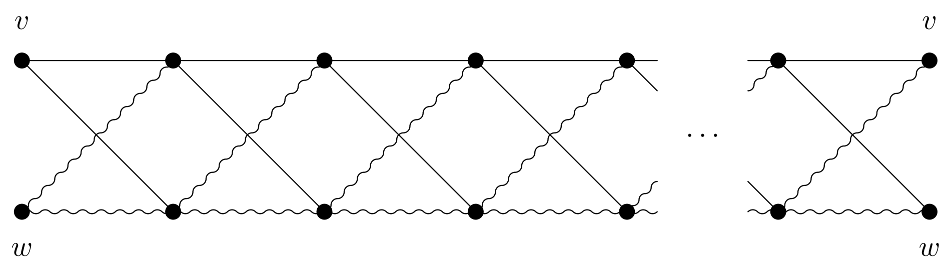


Figure C.1: The signed graph T_{2k}, with $2k$ vertices, for $k \geq 3$.

Remark. In Figure C.1, we have copies of the vertices v and w. These are identified and are the same vertex.

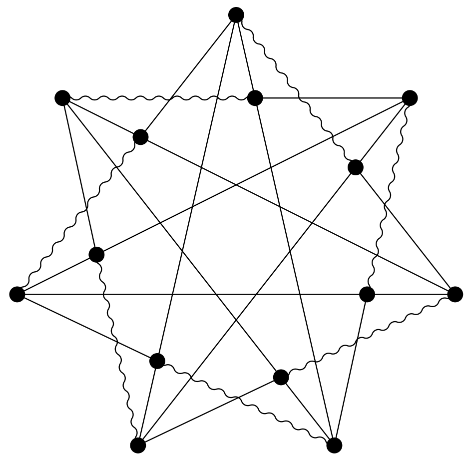


Figure C.2: The signed graph S_{14}, with 14 vertices.

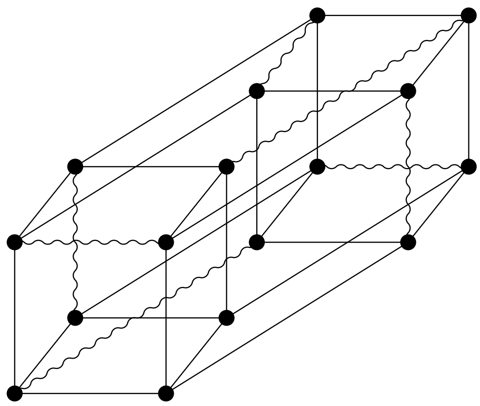
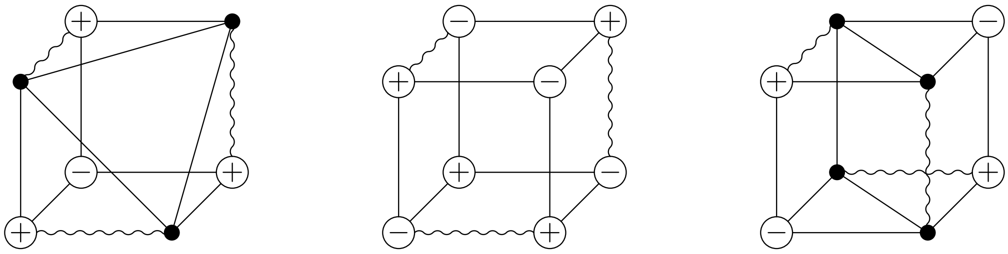


Figure C.3: The signed graph S_{16}, with 16 vertices.



S_7 S_8 S_8'

Figure C.4: The three sporadic maximal charged signed graphs S_7, S_8 and S_8'.

Appendix D

Table of Small Limits Points

ℓ_1, r_1	ℓ_2, r_2	c_ℓ, c_r D_L, D_R	u, d c_m	ℓ, m r, n	Limit point
-1, -1	1, -1	0, 0 $-$, $\bullet$		$t+1$, $0t+2$ $2t$, -1	$1.25543386\cdots$
-1, -1	1, -1	0, 0 $+$, $+$		t, $0t+2$ $3t+4$, -1	$1.28573486\cdots$
-1, 0	0, -1	-1, 0 $-$, $<$	0, -1 0	$0t+0$, $t+0$ $0t+1$, 1	$1.30909838\cdots$
-1, -1	1, -1	0, 0 $\bullet$, $-$		$2t+2$, $0t+2$ $3t+2$, -1	$1.31569270\cdots$
0, -1	0, -1	0, 0 $+$, $+$		$t+0$, $0t+2$ $t+0$, -1	$1.32471795\cdots$
-1, -1	1, -1	0, 0 $+$, $+$		$t+1$, $0t+2$ $3t+2$, -1	$1.32537249\cdots$
-1, -1	1, -1	0, 0 $\bullet$, $\bullet$		$t+1$, $0t+4$ $2t+1$, -1	$1.33205110\cdots$
-1, -1	1, 0	0, 0 $-$, $+$		$t+0$, $0t+5$ $3t+1$, -1	$1.33239612\cdots$

ℓ_1, r_1	ℓ_2, r_2	c_ℓ, c_r D_L, D_R	u, d c_m	ℓ, m r, n	Limit point
-1, -1	1, -1	0, 0 $+$, $+$		$t+1$, $0t+5$ $3t+2$, -1	$1.33813743\cdots$
-1, -1	-1, 0	0, 0 $\bullet$, $-$	1, 0	$0t+0$, $t+0$ $0t+1$, 0	$1.33999992\cdots$
-1, -1	1, -1	0, 0 $+$, $+$		$t+0$, $0t+2$ $2t+2$, -1	$1.34050688\cdots$
-1, -1	1, 0	0, 0 $-$, $-$		$t+1$, $0t+5$ $3t+1$, -1	$1.34865199\cdots$
-1, 0	0, -1	0, 0 $\bullet$, $-$	1, 1	$0t+2$, $0t+2$ $t-1$, 0	$1.34971610\cdots$
-1, 0	0, -1	-1, 1 $-$, $+$	0, 1	$0t+1$, $t+0$ $0t+1$, 0	$1.35001483\cdots$
-1, -1	1, 0	0, 0 $-$, $\bullet$		$t+0$, $0t+5$ $4t+5$, -1	$1.35031697\cdots$
-1, -1	1, 0	0, 0 $-$, $\bullet$		$t+1$, $0t+5$ $4t+1$, -1	$1.35114589\cdots$
-1, -1	1, 0	0, 0 $\bullet$, $\bullet$		$4t+3$, $0t+5$ $5t+0$, -1	$1.35246806\cdots$
-1, -1	1, 0	0, 0 $-$, $+$		$t+0$, $0t+5$ $t+4$, -1	$1.35369764\cdots$
-1, -1	-1, 0	-1, 0 $\bullet$, $+$	-1, 0	$0t+0$, $2t+0$ $t+1$, 0	$1.35674810\cdots$
-1, -1	1, -1	0, 0 $+$, $+$		$3t+1$, $0t+2$ $5t+0$, -1	$1.35678598\cdots$
-1, -1	1, -1	0, 0 $+$, $+$		$3t+1$, $0t+2$ $5t+3$, -1	$1.35854559\cdots$

ℓ_1, r_1	ℓ_2, r_2	c_ℓ, c_r D_L, D_R	u, d c_m	ℓ, m r, n	Limit point
-1, -1	-1, 0	0, 0 $\bullet$, $+$	1, 0	$0t+0$, $t+0$ $0t+2$, 0	$1.35920806\cdots$
-1, -1	1, 0	0, 0 $\bullet$, $-$		$t+1$, $0t+5$ $2t+0$, -1	$1.35937564\cdots$
-1, -1	1, 0	0, 0 $\bullet$, $\bullet$		$5t+4$, $0t+5$ $6t+1$, -1	$1.35981177\cdots$
-1, -1	1, 0	0, 0 $-$, $\bullet$		$t+0$, $0t+5$ $6t+5$, -1	$1.35981589\cdots$
-1, -1	-1, 0	0, 0 $\bullet$, $+$	-1, -1	$0t+0$, $0t+7$ $t-1$, 0	$1.35991414\cdots$
-1, 0	0, -1	-1, 1 $-$, $+$	0, -1 0	$0t+1$, $t+0$ $0t+0$, 1	$1.36022084\cdots$
-1, -1	-1, 0	0, 0 $\bullet$, $+$	1, 0	$0t+0$, $t+7$ $3t-2$, 0	$1.36195645\cdots$
-1, -1	1, -1	0, 0 $+$, $+$		$t+0$, $0t+2$ $5t+0$, -1	$1.36272428\cdots$
-1, -1	1, 0	0, 0 $-$, $+$		$5t+5$, $0t+5$ $3t+0$, -1	$1.36365149\cdots$
-1, -1	1, 0	0, 0 $\bullet$, $-$		$2t+1$, $0t+5$ $3t+0$, -1	$1.36419954\cdots$
-1, -1	1, -1	0, 0 $+$, $+$		t, $0t+2$ $2t+1$, -1	$1.36443581\cdots$
-1, 0	0, -1	-1, 1 $-$, $+$	0, -1 0	$0t+3$, $t+0$ $0t+0$, 1	$1.36465572\cdots$
-1, 0	0, -1	-1, -1 $+$, $+$		$0t+3$, $t-2$ $t+0$, $t+3$	$1.36506231\cdots$

ℓ_1, r_1	ℓ_2, r_2	c_ℓ, c_r D_L, D_R	u, d c_m	ℓ, m r, n	Limit point
-1, -1	1, -1	0, 0 $\bullet$, $-$		$4t+3$, $0t+2$ $5t+1$, -1	$1.36526954\cdots$
-1, 0	0, -1	-1, 1 $\bullet$, $+$	-1, -1	$2t-2$, $0t+5$ $t+0$, 0	$1.36546873\cdots$
-1, 0	0, -1	-1, 1 $-$, $+$	0, -1 0	$0t+0$, $t+0$ $0t+0$, 1	$1.36614596\cdots$
-1, -1	1, -1	0, 0 $+$, $+$		$t+0$, $0t+2$ $5t+3$, -1	$1.36629907\cdots$
-1, -1	1, 0	0, 0 $-$, $\bullet$		$3t+0$, $0t+5$ $2t+0$, -1	$1.36640199\cdots$
-1, -1	1, 0	0, 0 $\bullet$, $-$		$2t+0$, $0t+5$ $t+1$, -1	$1.36643553\cdots$
-1, 0	0, -1	-1, 1 $\bullet$, $+$	0, -1 0	$0t+0$, $t+0$ $0t+0$, 1	$1.36657097\cdots$
-1, -1	1, -1	0, 0 $+$, $+$		$2t+2$, $0t+2$ $3t+4$, -1	$1.36680788\cdots$
-1, -1	1, -1	0, 0 $+$, $+$		$t+1$, $0t+6$ $t+2$, -1	$1.36688307\cdots$
-1, 0	0, -1	-1, -1 $-$, $\bullet$		$0t+2$, $t+1$ $0t+5$, $0t+6$	$1.36699091\cdots$
-1, -1	1, 0	0, 0 $-$, $\bullet$		$t+1$, $0t+5$ $4t+2$, -1	$1.36751103\cdots$
-1, 0	0, -1	0, 0 $-$, $-$	-1, -1 0	$t+2$, $t-3$ $0t+5$, 1	$1.36779885\cdots$
-1, 0	0, -1	-1, 1 $-$, $\bullet$	-1, -1	$t-1$, $0t+8$ $0t+0$, 0	$1.36785463\cdots$

ℓ_1, r_1	ℓ_2, r_2	c_ℓ, c_r D_L, D_R	u, d c_m	ℓ, m r, n	Limit point
-1, 1	1, 0	1, 1 $-$, $\bullet$	1, 1	$0t+1$, $0t+1$ $t-1$, 0	$1.36813222\cdots$
-1, 0	0, -1	0, 0 $\bullet$, $+$	-1, -1	$t+2$, $2t-2$ $0t+7$, 0	$1.36819625\cdots$
-1, -1	-1, 0	0, 1 $\bullet$, $+$	-1, -1	$0t+0$, $t+0$ $0t+2$, 0	$1.36821400\cdots$
-1, 0	0, -1	1, 0 $>$, $<$	0, -1 -1	$0t+4$, $t+0$ $0t+1$, 1	$1.36834343\cdots$
-1, -1	1, -1	0, 0 $+$, $+$		$3t+0$, $0t+2$ $5t+2$, -1	$1.36839671\cdots$
-1, 0	0, -1	0, 0 $-$, $\bullet$	1, 1	$0t+4$, $2t+0$ $t+2$, 0	$1.36874744\cdots$
-1, -1	1, 0	0, 0 $\bullet$, $-$		$3t+1$, $0t+5$ $2t+0$, -1	$1.36892221\cdots$
-1, -1	1, 0	0, 0 $-$, $+$		$t+2$, $0t+5$ $3t+0$, -1	$1.36897873\cdots$
-1, 0	0, -1	-1, -1 $+$, $<$		$t-1$, $0t+2$ $3t+1$, $2t+1$	$1.36948937\cdots$
-1, 0	0, -1	-1, 1 $+$, $<$	0, -1 0	$0t+0$, $t+1$ $4t-2$, 1	$1.36978231\cdots$

Table D.1: Small limits of sequences of Mahler measures from families of digraphs.

This Table exhibits families of digraphs which have small limit points of Mahler measures, where small here is taken to mean "less than 1.37". For a full understanding of how to read the Table, see Section 5.3.2.

Bibliography

[1] D. Boyd. Small Salem numbers. *Duke Mathematical Journal*, 44(2):315 – 328, 1977. 23

[2] D. Boyd. Kronecker's Theorem and Lehmer's Problem for Polynomials in Several Variables. *Journal of Number Theory*, 13:116–121, 1981. 28

[3] D. Boyd. Speculations concerning the range of Mahler's measure. *Canad. Math. Bull.*, 24(4):453–469, 1981. 27, 178

[4] D. Boyd. Mahler's measure and invariants of hyperbolic manifolds. *Number Theory for the Millennium*, 01:127–143, 2002. 40

[5] D. Boyd and M. Mossinghoff. Small Limit Points of Mahler's Measure. *Exp. Math.*, 14(4):403–414, 2005. 24, 28, 29, 39, 54, 189

[6] R. Breusch. On the distribution of the roots of a polynomial with integral coefficients. *Proc. Amer. Math. Soc.*, 2:939–941, 1951. 25

[7] C. Christopoulos. *On Galois groups of Salem polynomials.* PhD thesis, Royal Holloway, University of London, 2008. 201

[8] C. Christopoulos and J. McKee. Galois theory of Salem polynomials. *Mathematical Proceedings of the Cambridge Philosophical Society*, 148(1):47–54, 2010. 201

[9] J. Coyston and J. McKee. Small Mahler Measures from Digraphs. *Experimental Mathematics*, (13 pages), 2021. 24, 68, 107, 164

[10] D. Cvetković, P. Rowlinson, and S. Simić. *An Introduction to the Theory of Graph Spectra*. London Mathematical Society Student Texts. Cambridge University Press, 2009. 151, 153

[11] J. D. Dixon and A. Dubickas. The values of Mahler measures. *Mathematika*, 51(1-2):131–148, 2004. 197, 198

[12] E. Dobrowolski. On a question of Lehmer and the number of irreducible factors of a polynomial. *Acta Arithmetica*, 34(4):391–401, 1979. 20

[13] C. Godsil and G. Royle. *Algebraic Graph Theory. No. 207 in Graduate Texts in Mathematics*. Springer-Verlag New York, 2001. 71

[14] E. Hironaka. The Lehmer Polynomial and Pretzel Links. *Canadian Mathematical Bulletin*, 44:440–451, 12 2001. 31

[15] L. Kronecker. Zwei Sätze über Gleichungen mit ganzzahligen Coefficienten. *J. Reine Angew. Math.*, 53:173–175, 1857. 20

[16] W. Lawton. A problem of Boyd concerning geometric means of polynomials. *J. Number Theory*, 16:356–362, 1983. 27, 177

[17] D. H. Lehmer. Factorization of certain cyclotomic functions. *Ann. Math.*, 34(2):461–479, 1933. 16, 20

[18] A. Lenstra, H. Lenstra, and L. László. Factoring Polynomials with Rational Coefficients. *Mathematische Annalen*, 261:515–534, 12 1982. 242, 243

[19] K. Mahler. An Application of Jensen's Formula to Polynomials. *Mathematika*, 7:98–100, 1960. 18

[20] K. Mahler. On Some Inequalities for Polynomials in Several Variables. *Journal London Math. Soc.*, 37:341–344, 1962. 18

[21] J. McKee. Conjugate Algebraic Numbers on Conics: A Survey. *London Mathematical Society Lecture Note Series*, 352:211–240, 2008. 20

[22] J. McKee and C. Smyth. Salem numbers, Pisot numbers, Mahler measure, and graphs. *Experimental Mathematics*, 14(2):211–229, 2005. 25, 31, 85, 109, 111, 157

[23] J. McKee and C. Smyth. Integer symmetric matrices having all their eigenvalues in the interval $[-2, 2]$. *Journal of Algebra*, 317:260–290, 2007. 83, 89, 244

[24] J. McKee and C. Smyth. Integer symmetric matrices of small spectral radius and small Mahler measure. *International Mathematics Research Notices*, 2012(1):102–136, 2012. 31, 76, 77

[25] J. McKee and C. Smyth. *Around the Unit Circle: Mahler Measure, Integer Matrices and Roots of Unity.* Universitext. Springer International Publishing, 2021. 20, 29, 53, 97, 98, 184, 221

[26] M. Mossinghoff. Mossinghoff: List of Mahler measures. `http://wayback.cecm.sfu.ca/~mjm/Lehmer/`, 2011. (Date accessed: May 6, 2022). 24, 29, 51, 181, 184

[27] M. Mossinghoff, G. Rhin, and Q Wu. Minimal Mahler Measures. *Experimental Mathematics*, 17(4):451–458, 2008. 24

[28] M. E. J. Newman and G. T. Barkema. *Monte Carlo methods in statistical physics.* Clarendon Press ; Oxford University Press, Oxford : New York, 1999. 38

[29] S. El Otmani, A. Maul, G. Rhin, and J-M. Sac Épée. Integer linear programming applied to determining monic hyperbolic irreducible polynomials with integer coefficients and span less than 4. *J. de Théorie des Nombres de Bordeaux*, 25(1):71–78, 2013. 184

[30] S. El Otmani, G. Rhin, and J.-M. Sac-Épée. Finding New Limit Points of Mahler's Measure by Genetic Algorithms. *Exp. Math.*, 28(2):129–131, 2019. 29, 38, 49, 51

[31] M. Pollicott and P. Felton. Estimating Mahler measures using periodic points for the doubling map. *Indagationes Mathematicae*, 25:619–631, 06 2014. 37

[32] K. Reidemeister. *Knotentheorie.* Reprint of 1932 original, Springers, 1974. 30, 239

[33] P. Rowlinson. A deletion-contraction algorithm for the characteristic polynomial of a multigraph. *Proc. Royal Soc. Edinburgh*, 105A:153–160, 1987. 140, 143

[34] J.-M. Sac-Épée. `http://www.iecl.univ-lorraine.fr/~Jean-Marc.Sac-Epee/SMM3.txt`. (Date accessed: May 6, 2022). 24, 181, 184

[35] J.-M. Sac-Epee, S. Elotmani, A. Maul, and G. Rhin. Finding new limit points of Mahler measure by methods of missing data restoration. *BAU Journal - Science and Technology*, 2(2):10, June 2021. 29, 51

[36] R. Salem. A remarkable class of algebraic integers. Proof of a conjecture of Vijayaraghavan. *Duke Mathematical Journal*, 11(1):103 – 108, 1944. 23

[37] C. L. Siegel. Algebraic integers whose conjugates lie in the unit circle. *Duke Mathematical Journal*, 11(3):597 – 602, 1944. 23

[38] J. Smith. Some properties of the spectrum of a graph. *Combinatorial Structures and Their Applications, Gordon and Breach, New York*, pages 403–406, 1970. 75

[39] C. J. Smyth. On the product of the conjugates outside the unit circle of an algebraic integer. *Bull. London. Math. Soc.*, 2:169–175, 1971. 25

[40] C. J. Smyth. On measures of polynomials in several variables. *Bulletin of the Australian Mathematical Society*, 23:49–63, 02 1981. 37

[41] C. J. Smyth. The Mahler measure of algebraic numbers: A survey. *Number Theory and Polynomials*, 352:322–349, 01 2007. 15

[42] C. J. Smyth. Seventy Years of Salem Numbers. *Bull. London. Math. Soc.*, 47:379–395, 2015. 23

[43] R. Spiegel. *Theory and Problems of Complex Variables - SI (Metric) Edition*. McGraw-Hill Publishing Co., 1974. 16

[44] The PARI-Group, Univ. Bordeaux. *PARI/GP version `2.11.2`*, 2019. available from `http://pari.math.u-bordeaux.fr/`. (Date accessed: May 6, 2022). 32

[45] E. C. Titchmarsh. *The Theory of Functions*. Oxford University Press, 1939. 16

[46] P. Viana and P. M. Veloso. Galois Theory of Reciprocal Polynomials. *The American Mathematical Monthly*, 109(5):466–471, 2002. 201, 202

[47] P. Voutier. An effective lower bound for the height of algebraic numbers. *Acta Arithmetica*, 74(1):81–95, 1996. 20

Index