Arithmetic Applications
of Hankel Determinants

Daniel James Sutherland

BMath(Hons)/BSc

Dissertation Submitted to
The School of Mathematical and Physical Sciences
at The University of Newcastle
for the Degree of
Doctor of Philosophy

Submitted: December 2014       Revised: February 2015
Statement of Originality
The thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. I give consent to the final version of my thesis being made available worldwide when deposited in the University’s Digital Repository, subject to the provisions of the Copyright Act 1968.

Candidate name: _______________    Candidate signature: _______________

Declaration by the candidate
I declare that the details provided above are correct.

Candidate Signature: _______________

Endorsement by the supervisor
I, as the supervisor of the candidate, certify that the details provided above are correct.

Supervisor name: _______________

Supervisor signature: _______________    Date: _______________
Acknowledgements

I would like to thank my supervisors Michael Coons, Wadim Zudilin and Jonathan Borwein for their guidance at various stages throughout my research. Their support has been invaluable and their enthusiasm has always given me confidence.

I thank Timothy Trudgian and Jean-Paul Allouche for examining my thesis and providing me with helpful feedback.

I thank my parents, Brian and Heather, who have always been eager to hear about my mathematical successes and setbacks, and my brother Josh for helping me to improve my explanatory skills.

I also thank my wonderful fiancée Skye, who has helped to bear the burdens of my research and who has always encouraged me to press on. I am grateful for her belief in my ability and her willingness to become familiar with my work.

Lastly, I would also like to acknowledge the One from whom all things derive their order and beauty, my Lord of whom it is written:

I wisdom dwell with prudence, and find out knowledge of witty inventions.

— Proverbs 8:12 (KJV)
## Contents

Acknowledgements v  
Abstract ix  
Notation xi  

Chapter 0. Introduction 1  
0.1. Matrix determinants 1  
0.2. The Riemann zeta function 2  
0.3. Overview and original contributions 5  

Chapter 1. Hankel Determinants with Formal Parameters 9  
1.1. Observations and experimental mathematics 9  
1.2. Proof via row reduction 17  
1.3. Bounding denominators 24  

Chapter 2. Hankel Determinants and Rational Approximations 27  
2.1. The Hilbert matrix 27  
2.2. Padé approximation 32  
2.3. Applications toward irrationality 41  

Chapter 3. Hankel Determinants of Convergent Sequences 51  
3.1. Hankel in Wonderland 51  
3.2. Positive Hankel determinants 57  
3.3. The pursuit of positivity 61  

Chapter 4. Hankel Determinants of Dirichlet Series 69  
4.1. Hankel determinants of some specialised sequences 69  
4.2. Hankel determinants of ordinary Dirichlet series 74  
4.3. Hankel determinants of Dirichlet $L$-series 83  

Appendix A. Historical development of determinants 93  
Appendix B. Determinant calculations 97  
Bibliography 99
Abstract

This thesis focuses on the application of matrix determinants as a means of producing number-theoretic results. Motivated by an investigation of properties of the Riemann zeta function, we examine the growth rate of certain determinants of zeta values. We begin with a generalisation of determinants based on the Hurwitz zeta function, where we describe the arithmetic properties of its denominator and establish an asymptotic bound. We later employ a determinant identity to bound the growth of positive Hankel determinants. Noting the positivity of determinants of Dirichlet series allows us to prove specific bounds on determinants of zeta values in particular, and of Dirichlet series in general. Our results are shown to be the best that can be obtained from our method of bounding, and we conjecture a slight improvement could be obtained from an adjustment to our specific approach.

Within the course of this investigation we also consider possible geometric properties which are necessary for the positivity of Hankel determinants, and we examine the role of Hankel determinants in irrationality proofs via their connection with Padé approximation.
Notation

\[ \mathbb{N} \] The set of natural numbers \( \{1, 2, 3, \ldots\} \).

\[ \mathbb{Q} \] The set of rational numbers.

\[ \mathbb{R} \] The set of real numbers.

\[ \mathbb{R}^2 \] The set of ordered pairs \((x, y)\) of real numbers \(x, y \in \mathbb{R}\).

\[ \mathfrak{S}_n \] The symmetric group of order \( n \).

\[ \mathbb{Z} \] The set of integers.

\[ \text{adj}(M) \] The adjugate of the square matrix \( M \).

\[ \text{denom}(P(x)) \] The integer denominator of the coefficients of \( P(x) \in \mathbb{Q}[x] \).

\[ \text{lcm}\{n_1, n_2, \ldots\} \] The least common multiple of a set of natural numbers.

\[ \phi(n) \] The number of positive integers less than \( n \) that are coprime to \( n \).

\[ \Re(s) \] The real part \( \sigma \) of the complex number \( s = \sigma + it \), with \( \sigma, t \in \mathbb{R} \).

\[ \text{sgn}(\pi) \] The sign of a permutation \( \pi \in \mathfrak{S}_n \).

\[ \text{sgn}(x) \] The sign of \( x \in \mathbb{R} \), equal to \( x/|x| \), \( x \neq 0 \).

\[ f(n) = \mathcal{O}(g(n)) \] The ratio \( |f(n)/g(n)| \) is bounded as \( n \) tends to \( \infty \), \( g(n) \neq 0 \).

\[ f(n) = \mathcal{o}(g(n)) \] The ratio \( |f(n)/g(n)| \) tends to 0 as \( n \) tends to \( \infty \), \( g(n) \neq 0 \).

\[ f(n) \asymp g(n) \] Equivalent to \( f(n) = \mathcal{O}(g(n)) \) and \( g(n) = \mathcal{O}(f(n)) \).

\[ C, c \] A constant (different subscripts denote different constants).

\[ p \] A prime \( p \in \mathbb{N} \).
CHAPTER 0

Introduction

0.1. Matrix determinants

In this thesis we consider the evaluation of certain matrix determinants as well as their notable arithmetic properties. Throughout the development of linear algebra, the name “determinant” was not used in the modern sense until the early 19th century when Augustin-Louis Cauchy [12] began to establish concrete notation (for a more detailed description of the history of determinants see Appendix A). Determinants have broadened from their original algebraic use of investigating the nature of solutions to a system of linear equations, and are now present in many other mathematical fields (for instance, in calculus as Jacobians or Wronskians). In recent times, determinants have been useful in several scientific areas, such as lattice theory and general relativity (see Vein and Dale [40] for a number of applications in mathematical physics).

The determinants present in many of these cases often exist in special forms with additional properties. In this thesis we consider some consequences that follow from the properties of Hankel determinants, defined below.

Definition 0.1. An $n \times n$ matrix $M = (m_{i,j})_{i,j=1,\ldots,n}$ is a Hankel matrix if $m_{i,j} = m_{i'+j'}$ whenever $i + j = i' + j'$, and $M$’s determinant is a Hankel determinant.

This characterisation of a Hankel matrix reveals that it is simply a symmetric matrix with constant positive-sloping diagonals. From the definition we see that the entries $m_{i,j}$ of a Hankel matrix therefore only depend on $i + j$, which allows us to deduce an alternative construction. Suppose we have a sequence $\{h(k)\}_{k \geq 2}$ of real numbers. Then the $n \times n$ matrix $M = (m_{i,j})_{i,j=1,\ldots,n}$ where $m_{i,j} = h(i + j)$ is a Hankel matrix.
0. INTRODUCTION

Here, and throughout this thesis we write $H_n^{(r)}[h]$ to denote an $n \times n$ Hankel determinant, built on the underlying sequence $\{h(k)\}_{k \geq 2}$ such that

$$H_n^{(r)}[h] = \det_{1 \leq i, j \leq n} \left( h(i + j + r) \right),$$

where the integer $r \geq 0$ determines which term of the sequence $\{h(k)\}_{k \geq 2}$ first appears in $H_n^{(r)}[h]$.

While Hankel determinants may appear to be algebraic objects, their connection to orthogonal polynomials (see Krattenthaler [23]) often leads to useful combinatorial applications. Viennot [41] showed how certain Hankel determinants can be used to enumerate families of weighted paths and Aigner [2] demonstrated Hankel determinants of Catalan-like numbers were valuable in combinatorial enumeration of certain lattice paths and rooted binary trees. For a recent application, Hankel determinants were an important tool in the proof of the Refined Alternating Sign Matrix Conjecture (see Bressoud and Propp [10] for a commentary on the resolution of this result).

0.2. The Riemann zeta function

Among the most enticing aspects of mathematics is that a simply stated problem may require a sophisticated solution — or worse, is intractable. Several challenges of this form exist in relation to the Riemann zeta function:

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \quad (0.1)$$

This series was first studied by Leonhard Euler in the early 18th century for integers $s > 1$. Pafnuty Chebyshev later extended its definition to include all real $s > 1$ (see Devlin [15] for further historical remarks). Subsequent to the emergence of complex analysis, Bernhard Riemann [33] generalised the function that now bears his name to take complex values for $\Re(s) > 1$. Riemann went on to show that his function could be continued analytically to all complex $s \neq 1$ (where $\zeta(1)$ corresponds to the divergent harmonic series).

Since this time, there have been two significant problems related to the Riemann zeta function: one analytic and one arithmetic. The existence of trivial zeroes of the analytic continuation of the function were well understood by Riemann ($\zeta(s) = 0$ for even integers $s < 0$), and any nontrivial zeroes had to be located in the critical
strip $0 \leq \Re(s) \leq 1$. The Riemann Hypothesis, proposed in 1859, conjectured that all nontrivial zeroes satisfy $\Re(s) = \frac{1}{2}$, and many attempts have been made to settle this problem as interest in it grew. During a number theory conference held at Caltech in 1955, Tom Apostol performed an ode lamenting the “many good men, with vim and with vigor” that have left their mark on the Riemann Hypothesis (see Krantz [22, p.75] for an excerpt). Still this conjecture remains unsolved and is now one of the Millennium Prize Problems, which has a million dollar reward.

The second problem is related to the first in that it concerns the nature of $\zeta(n)$ for integers $n$ (known as zeta values). Just as the Riemann zeta function is easily evaluated at negative even integers, it can similarly be evaluated at positive even integers as

$$\zeta(2n) = (-1)^{n+1} \frac{B_{2n}}{2(2n)!} (2\pi)^{2n},$$

for integers $n > 0$, where $B_{2n}$ is a Bernoulli number, defined as

$$B_m = \sum_{k=0}^{m} \sum_{v=0}^{k} (-1)^v \binom{k}{v} \frac{v^m}{k+1},$$

for integers $m > 0$. Dunham [17, pp.212–222] outlines Euler’s discovery of this result explicitly up to $\zeta(26)$. This is not the only appearance of the Bernoulli numbers in connection with the Riemann zeta function, as the zeta values at negative integers can be given as

$$\zeta(-n) = -\frac{B_{n+1}}{n+1},$$

for integers $n > 0$. This explains the occurrence of the trivial zeroes at the negative even integers as $B_m = 0$ for odd integers $m > 1$.

The behaviour of the Riemann zeta function is entirely understood for negative integers (it takes rational values that can be stated explicitly), but its nature for positive integers is yet to be understood. The even zeta values were shown to be rational multiples of a power of $\pi$, and at the time of Euler’s evaluation it was known that $\pi$ is irrational (and hence $\zeta(2n)$ is irrational for $n \in \mathbb{N}$); this was shown by Johann Lambert [25] in 1761. It was not until 1882 when Ferdinand von Lindemann [26] proved the transcendence of $\pi$ that the even zeta values were finally shown to be transcendental. However, much less is known for odd zeta values. We noted earlier that $\zeta(1)$ diverges as a sum because it simply reduces to the harmonic series, but what about $\zeta(3)$, $\zeta(5)$ or $\zeta(7)$? The expected answer is that each odd
zeta value is likewise transcendental, but a proof of this remains elusive. Moreover, even demonstrating the irrationality of individual odd zeta values is a challenging endeavour.

In June 1978, Roger Apéry gave a talk in which he outlined a proof that $\zeta(3)$ is irrational. He did not receive recognition for this discovery until August that same year as many mathematicians simply dismissed his proof on the basis of his sketchy approach. His original proof [3] was quite elementary but very technical in detail. Shortly after the acceptance of Apéry’s result, alternate proofs began emerging for the irrationality of $\zeta(3)$. For example, the proof by Frits Beukers [5] replaces Apéry’s series with integral representations. Others involving Padé approximants (for an introduction to Padé approximation see Section 2.2) or hypergeometric series followed, and similar arguments were attempted for $\zeta(5)$ and additional odd zeta values without success.

The story ends here for irrationality of individual zeta values, although recently it has been possible to say more with less specificity. Tanguy Rivoal [34] proved in 2000 that infinitely many of the numbers $\zeta(3), \zeta(5), \zeta(7), \ldots$ are irrational by constructing linear forms of odd zeta values and examining the dimension of the spaces spanned over $\mathbb{Q}$. The following year, Wadim Zudilin [44] generalised Rivoal’s construction to demonstrate that each set $\{\zeta(s + 2), \zeta(s + 4), \ldots, \zeta(8s - 3), \zeta(8s - 1)\}$ with odd $s > 1$ must contain at least one irrational number. He later added (see [45]) that at least one of the four numbers $\zeta(5), \zeta(7), \zeta(9), \zeta(11)$ is irrational. Although there are still ongoing investigations into this mysterious function, there has been little significant progress towards establishing the irrationality (let alone the transcendence) of specific zeta values.

The Riemann zeta function is a versatile object with many applications outside number theory. It appears within various branches of statistical mechanics, for instance thermodynamic formulae (Schrödinger [36]) and spin correlation functions (Boos [6]). In quantum field theory, some Feynman diagrams can be expressed in terms of $\zeta(n)$ (Kreimer [24]), and zeta-function regularisation is a key component in quantitatively describing the Casimir effect (a review of new developments can be found in Bordag [7]). In light of this, any additional insight about the ubiquitous zeta function would have widespread impact on mathematics and the broader sciences.
0.3. Overview and original contributions

We are concerned in this thesis, as the title suggests, with the investigation of the role Hankel determinants play in number theory. Motivated by the desire to discover more about the Riemann zeta function, we study determinants of zeta values and other similar sequences. To this end, this thesis is organised as follows.

In Chapter 1 we begin our investigation by replacing the Riemann zeta function with the Hurwitz zeta function, where this is understood as removing a finite number of initial terms from (0.1). Here we seek to determine the effect on the determinant of removing these large terms, so we describe the asymptotic growth of a sequence of Hankel determinants formed from corresponding Hurwitz zeta values. Our general formulation of the Hankel determinants produces a polynomial with rational coefficients, whose denominator is described in Theorem 1.3.

**Theorem.** Let \( n \geq 2, \ r \geq 2, \ a \geq 1 \) and \( d \geq 1 \) be integers, and fix a prime \( p \leq r \). Set \( m = \left\lfloor \frac{\log r}{\log p} \right\rfloor \) and \( \lambda = \left\lfloor \frac{r^{-1}}{p^m} \right\rfloor \). If \( W_n^{(a,d)}(r)[x] \neq 0 \), then the power of \( p \) in \( \text{denom}(W_n^{(a,d)}(r)[x]) \) is bounded above by

\[
\begin{cases} 
  d(m - 1)n^2 + (a(m - 1) + 2d\lambda)n - \lambda(d\lambda - a) & \text{if } \lambda \leq n \\
  dmn^2 + amn & \text{if } \lambda \geq n.
\end{cases}
\]

Bounding the primes in this denominator also allows us to produce a bound in Proposition 1.8 for the asymptotic growth of the denominators of this sequence of Hankel determinants.

**Proposition.** Let \( n \geq 2, \ r \geq 2, \ a \geq 1 \) and \( d \geq 1 \) be integers. Write \( W_n^{(a,d)}(r)[x] = \frac{P(x)}{q} \) where \( P(x) \in \mathbb{Z}[x_{a+d}, \ldots, x_{a+(2n-1)d}] \) and where \( q \in \mathbb{N} \). If \( P(x) \neq 0 \) then

\[
\log q < 1.01624(dn^2 + an)(1 + \log 2 + \log \log r) + 2 \log r.
\]

In Chapter 2 we move from directly studying the number theoretic properties of the Riemann zeta function and instead turn our attention to an application of Hankel determinants, namely proving irrationality results. We first consider some classical results about the Hilbert matrix, before extending them in Proposition 2.6 to our generalisation.

---

1As we are stating our results out of their immediate context, the notation may not yet make sense. For a precise description of each result we direct the reader to the chapter in which it appears.
0. INTRODUCTION

Proposition. The determinant of the matrix obtained by removing the $k$th row and $m$th column of the general $n \times n$ Hilbert matrix is given by

$$(H_n^{(r)})^m = H_n^{(r)} \cdot (k + m + r - 1) \binom{k + n + r - 1}{n - m} \binom{n + m + r - 1}{n - k} \cdot \binom{k + m + r - 2}{k - 1} \binom{k + m + r - 2}{m - 1}.$$ 

This allows us to give a closed-form expression for most of the Padé approximation table of the natural logarithm (up to a shift) in Proposition 2.16.

Proposition. Let $f(x) = -\frac{1}{x} \log(1 - x)$ and fix positive integers $L$ and $M$ with $L \geq M - 1$. The Padé approximant to $f(x)$ of order $\left\lfloor \frac{L}{M} \right\rfloor$ exists and is given by

$$[L/M] f(x) = \sum_{j=0}^{M} (-1)^j \binom{M}{j} \binom{L + j + 1}{M} x^{M-j} \sum_{i=0}^{L-M+j} \frac{x^i}{i+1}.$$ 

We then use this rational function to produce an explicit sequence of rational numbers which closely approximate certain values of the natural logarithm, and so in Theorem 2.11 we give an original proof of the irrationality of these values of the natural logarithm.

Theorem. The real number $\log \left( 1 + \frac{1}{b} \right)$ is irrational for integers $b \geq 10$.

We attempt to use these techniques to resolve the irrationality of Catalan’s constant, and discuss the limitations of our approach.

In Chapter 3 we introduce the method of computing determinants known as Dodgson condensation. Expressed in terms of Hankel determinants we obtain the recurrence relation:

$$H_{n+1}^{(r)}[h] \cdot H_{n-1}^{(r+2)}[h] = H_n^{(r)}[h] \cdot H_n^{(r+2)}[h] - \left( H_n^{(r+1)}[h] \right)^2.$$ 

We use this important result to produce Proposition 3.6, an upper bound on positive Hankel determinants.

Proposition. Let $R \geq 0$ be a fixed integer. If $H_n^{(r)}[h] > 0$ for all $n \geq 1$ and for $r \geq R$, then

$$H_n^{(r)}[h] < h(2 + r) \prod_{k=0}^{n-2} \frac{H_{2k}^{(r+2k)}[h]}{H_k^{(r+2k)}[h]}.$$ 

for all $n \geq 3$ and for $r \geq R$. 
From this bound we deduce in Theorem 3.5 that a certain sequence of positive Hankel determinants formed from the terms of any convergent sequence strictly decreases to zero.

**Theorem.** Let \(\{h(k)\}_{k \geq 2}\) be a convergent sequence of positive real numbers. If \(H_n^{(r)}[h] > 0\) for all \(n \geq 2\) and \(r \geq R\), for some integer \(R \geq 0\), then

\[
\lim_{r \to \infty} H_n^{(r)}[h] = 0,
\]

for all integers \(n \geq 2\), where the limit is achieved strictly monotonically from above.

Noting that these results are based on the need for positive determinants, we begin an investigation into whether geometric considerations of Hankel determinants can provide any insight into their positivity.

Chapter 4 brings our investigation of Hankel determinants to a close by returning to values of the Riemann zeta function. We consider the sequence of determinants formed from increasingly larger Hankel matrices, and produce a recast upper bound on positive Hankel determinants in Proposition 4.4.

**Proposition.** Let \(K \geq 2\) be a fixed integer and \(h(k) > 0\) for \(k \geq K\). Suppose there are positive functions \(A(k), B(k)\) and \(\lambda(k)\), with \(A(k) < B(k + 1)\) for all \(k \geq K\), such that

\[
\begin{align*}
(i) \quad & A(k) \leq \frac{h(k + 1)}{h(k)} \leq B(k) \quad \text{for all } k \geq K, \\
(ii) \quad & h(k + 2) \leq \lambda(k + 2) \cdot \frac{B(k + 1)}{B(k + 1) - A(k)} \quad \text{for all } k \geq K.
\end{align*}
\]

If \(H_n^{(r)}[h] > 0\) for all \(n \geq 1\) and for \(r \geq K - 2\), then

\[
H_n^{(r)}[h] < h(2 + r) \prod_{k=2}^{n} \lambda(2k + r),
\]

for all \(n \geq 3\) and for \(r \geq K - 2\).

With this new formulation we turn our attention to Hankel determinants of Dirichlet series, proving in Theorem 4.10 a bound on their asymptotic growth.

**Theorem.** Let \(F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}\) be a nondegenerate ordinary Dirichlet series with \(f(n) \geq 0\) for all \(n\), which is convergent for \(\Re(s) \geq s_0\). Then \(H_n^{(r)}[F] > 0\) for all \(n \geq 1\) and \(r \geq s_0 - 2\), and for sufficiently large \(r\) there is a \(c = c(r) > 0\) such that

\[
\log H_n^{(r)}[F] < -cn^2,
\]
for all sufficiently large integers $n$.

Experimental observations suggest that our bound could be improved, so we return to Dodgson condensation in full to demonstrate that our result is the best that can be obtained from Proposition 4.4. We conclude by opening an investigation into Hankel determinants of Dirichlet $L$-series, which are not positive (as was the case for ordinary Dirichlet series) but whose magnitude appears to grow similarly.
CHAPTER 1

Hankel Determinants with Formal Parameters

We begin our work on Hankel determinants with a specific example, investigated in a recent study of a collection of Hankel matrices whose entries are zeta values. This particular collection of Hankel determinants ultimately acts as the main example for our work.

In Section 1.1 we introduce these determinants and propose a generalisation by inserting formal parameters into the entries of the matrices, making the determinants polynomials in these parameters with rational coefficients. We then describe the experimental approach taken to study the denominator of these determinants, conjecturing their form based on our observations. In Section 1.2 we prove the general case of our experimental observations and comment in Section 1.3 on what these results mean, closing with a conjecture based on our work.

1.1. Observations and experimental mathematics

A 2009 preprint by Hartmut Monien [27] investigated the sequences of determinants

\[
H_1^{(0)}[\zeta] = \zeta(2), \quad H_2^{(0)}[\zeta] = \begin{vmatrix} \zeta(2) & \zeta(3) \\ \zeta(3) & \zeta(4) \end{vmatrix} , \quad H_3^{(0)}[\zeta] = \begin{vmatrix} \zeta(2) & \zeta(3) & \zeta(4) \\ \zeta(3) & \zeta(4) & \zeta(5) \\ \zeta(4) & \zeta(5) & \zeta(6) \end{vmatrix} , \ldots
\]

and

\[
H_1^{(1)}[\zeta] = \zeta(3), \quad H_2^{(1)}[\zeta] = \begin{vmatrix} \zeta(3) & \zeta(4) \\ \zeta(4) & \zeta(5) \end{vmatrix} , \quad H_3^{(1)}[\zeta] = \begin{vmatrix} \zeta(3) & \zeta(4) & \zeta(5) \\ \zeta(4) & \zeta(5) & \zeta(6) \\ \zeta(5) & \zeta(6) & \zeta(7) \end{vmatrix} , \ldots
\]

Monien opens his article with the surprising observation that, in addition to all being positive, these determinants trend very rapidly to zero as the dimension of the matrix becomes large (see Table 1.1).
Recall the Hurwitz zeta function $\zeta(s, \alpha)$, a generalisation of the Riemann zeta function $\zeta(s)$, classically defined by the formula

$$\zeta(s, \alpha) = \sum_{k=0}^{\infty} \frac{1}{(k + \alpha)^s},$$

for $\Re(s) > 1$ and $\Re(\alpha) > 0$, although we are only interested in the case where $s, \alpha \in \mathbb{Z}$ with $s \geq 2$ and $\alpha \geq 1$. It seems reasonable to expect the same rapidly decreasing trend that was described for the determinants above could also be exhibited in the case that the Riemann zeta values are replaced with Hurwitz zeta values (note that $\zeta(s, 1) = \zeta(s)$). To that end, we define the determinant $W_n(r)[x]$ as

$$W_n(r)[x] = \det_{1 \leq i, j \leq n} \left( x_{i+j} - \sum_{k=1}^{r} \frac{1}{k^{i+j}} \right)$$

$$= \begin{vmatrix} x_2 - 1 - \cdots - \frac{1}{r^2} & \cdots & x_{n+1} - 1 - \cdots - \frac{1}{r^{n+1}} \\ \vdots & \ddots & \vdots \\ x_{n+1} - 1 - \cdots - \frac{1}{r^{n+1}} & \cdots & x_{2n} - 1 - \cdots - \frac{1}{r^{2n}} \end{vmatrix},$$

for all integers $n \geq 1$ and $r \geq 0$ (where we understand the $r = 0$ case as corresponding to an empty sum), and where the terms $x_2, \ldots, x_{2n}$ are formal parameters. This construction is such that $W_n(r)[x]$ is a polynomial in the parameters $x_2, \ldots, x_{2n}$ with rational coefficients. If we were to make the specification $x_k = \zeta(k)$ for
$k \in \{2, 3, \ldots, 2n\}$, then we would have

$$W_n(r)[\zeta] = \det_{1 \leq i,j \leq n} \left( \zeta(i+j) - \sum_{k=1}^{r} \frac{1}{k^{i+j}} \right) = \det_{1 \leq i,j \leq n} \left( \zeta(i+j,r+1) \right),$$

and so our formulation is indeed a generalisation of Monien’s determinants as we have $W_n(0)[\zeta] = H_n^{(0)}[\zeta]$ for all $n \geq 1$.

It is important to note that if we specify $x_k = \frac{p_k}{q_k} \in \mathbb{Q}$ for $k \in \{2, 3, \ldots, 2n\}$, then $W_n(r)[x] \in \mathbb{Q}$ because we are simply evaluating a rational polynomial with a rational argument. Thereby, it is advantageous for us to study the denominator of $W_n(r)[x]$ with the aim of providing a good bound.

The first step in our investigation is to explicitly calculate a selection of these determinants $W_n(r)[x]$. After making the appropriate definitions in Maple we are able to examine the denominator of the determinants directly:

\[
denom(W(3,2));
\]

\[
denom(W(5,3));
\]

\[
denom(W(4,7));
\]

\[
denom(W(6,50));
\]

\[
denom(W(6,50));
\]

\[
denom(W(6,50));
\]

Seemingly, there appears to be no particular trend here (aside from the obvious fact that the denominator becomes larger if $n$ or $r$ is made larger). However, a deeper structure is revealed by factorising these denominators.

\[
ifactor(denom(W(3,2)));
\]

\[
(2)^6
\]
A number of trends are now evident that are worth pursuing:

- every prime, up to \( r \), is present in the denominator of \( W_n(r)[x] \),
- the power of any prime never exceeds the power of any smaller prime,
- the power of each prime is even, and
- primes that lie within a certain interval all share a common power.

It is this last trend which seems most appealing to investigate and make precise as it could allow us to create a general expression for the power of any prime.

First observe that the largest primes in each case have power \( 2n \). Outside this interval, a closer examination of \( W_6(50)[x] \) reveals that the primes 17, 19, and 23 also share a common power, and we observe that they all lie in \( \left( \frac{r}{3}, \frac{r}{2} \right) \). So why then a power of 22 in this case? We have already observed the powers are even, so perhaps it is more useful to ask why the power is \( 2 \cdot 11 \) instead. In this case we have \( n = 6 \) and so perhaps the power is given by \( 2(n + (n - 1)) = 4n - 2 \). We can confirm this by checking \( W_4(7)[x] \), for which we find the prime 3 located in the interval \( \left( \frac{r}{3}, \frac{r}{2} \right) \) with power \( 2(4 + 3) = 14 \).

Let us subdivide the interval \( (1, r] \) into nonoverlapping intervals \( \left( \frac{r}{\lambda + 1}, \frac{r}{\lambda} \right) \) for some integers \( \lambda \geq 1 \). Perhaps, under this subdivision, the power of each prime in an interval indexed by \( \lambda \) is

\[
e(\lambda) = 2(n + (n - 1) + \ldots + (n - \lambda + 1)) = 2\lambda n - \lambda(\lambda - 1).
\]

Testing this proposal for the remaining primes in our sample reveals that Equation (1.1) is not satisfied for \( p = 2 \) in \( W_4(7)[x] \) and for \( p \in \{2, 3, 5, 7\} \) in \( W_6(50)[x] \).

We observe that these exceptions are all primes less than \( \sqrt{r} \), and so to eliminate these exceptions we refine Equation (1.1) to hold only for \( 1 \leq \lambda < \sqrt{r} \). Additionally, it is worth noting that for fixed \( n \), the function \( e(\lambda) = 2\lambda n - \lambda(\lambda - 1) \) is increasing.
for \(1 \leq \lambda \leq n\) and decreasing for \(\lambda > n\), and so to avoid contradicting our earlier observation that the power of any prime never exceeds the power of any smaller prime we further restrict Equation (1.1) to be valid only for integers \(1 \leq \lambda \leq n\), which seems sensible given our earlier purpose for introducing \(\lambda\). All together this enables us to make the following partial conjecture.

**Conjecture 1.1.** Let \(n \geq 1\), \(r \geq 2\), and \(\lambda\) be integers and \(p\) a prime in \((\sqrt{r}, r]\). If

\[
\lambda < \min\{\sqrt{r}, n + 1\} \quad \text{and} \quad \frac{r}{\lambda + 1} < p \leq \frac{r}{\lambda}
\]


then the power of \(p\) in \(\text{denom}(W_n(r)|x|)\) is

\[
e(\lambda) = 2\lambda n - \lambda(\lambda - 1).
\]

Before we attempt to establish the validity of Conjecture 1.1 we first turn our attention to primes \(p\) with \(p \leq \sqrt{r}\) so that we might additionally conjecture the power of these smaller primes. To that end, let us examine the following determinant:

\[
\text{ifactor}\left(\text{denom}(W(3, 600))\right);
\]

\[
\begin{align*}
(2)^{99}(3)^{58}(5)^{36}(7)^{29}(11)^{24}(13)^{24}(17)^{22}(19)^{18}(23)^{18}(29)^{12}(31)^{12}(37)^{11}(41)^{12}(43)^{12}
\end{align*}
\]

\[
\begin{align*}
(47)^{12}(53)^{12}(59)^{12}(61)^{12}(67)^{12}(71)^{12}(73)^{12}(79)^{12}(83)^{12}(89)^{12}(97)^{12}(101)^{12}(103)^{12}
\end{align*}
\]

\[
\begin{align*}
(107)^{12}(109)^{12}(113)^{12}(127)^{12}(131)^{12}(139)^{12}(149)^{12}(151)^{12}(157)^{12}(163)^{12}
\end{align*}
\]

\[
\begin{align*}
(167)^{12}(173)^{12}(179)^{12}(181)^{12}(191)^{12}(193)^{12}(197)^{12}(199)^{12}(211)^{10}(223)^{10}(227)^{10}
\end{align*}
\]

\[
\begin{align*}
(229)^{10}(233)^{10}(239)^{10}(241)^{10}(251)^{10}(257)^{10}(263)^{10}(269)^{10}(271)^{10}(277)^{10}
\end{align*}
\]

\[
\begin{align*}
(281)^{10}(283)^{10}(293)^{6}(307)^{6}(311)^{6}(313)^{6}(317)^{6}(331)^{6}(337)^{6}(347)^{6}
\end{align*}
\]

\[
\begin{align*}
(349)^{6}(353)^{6}(359)^{6}(367)^{6}(373)^{6}(379)^{6}(383)^{6}(389)^{6}(397)^{6}(401)^{6}
\end{align*}
\]

\[
\begin{align*}
(409)^{6}(419)^{6}(421)^{6}(431)^{6}(433)^{6}(439)^{6}(443)^{6}(449)^{6}(457)^{6}(461)^{6}
\end{align*}
\]

\[
\begin{align*}
(463)^{6}(467)^{6}(479)^{6}(487)^{6}(491)^{6}(499)^{6}(503)^{6}(509)^{6}(521)^{6}(523)^{6}
\end{align*}
\]

\[
\begin{align*}
(541)^{6}(547)^{6}(557)^{6}(563)^{6}(569)^{6}(571)^{6}(577)^{6}(587)^{6}(593)^{6}(599)^{6}
\end{align*}
\]

First note the primes in the intervals \((300, 600]\), \((200, 300]\) and \((\sqrt{600}, 200]\) have power 6, 10 and 12 respectively. There are two exceptions (boxed above) where we propose an unexpected cancellation with the numerator has given these primes a slightly decreased power. As \(\sqrt{600} \approx 24.5\), we consider the power of primes less than 24. Since some of the larger primes in this interval share a common power it seems reasonable to further subdivide this interval. Note that

\[
\sqrt{\frac{600}{2}} \approx 17.3 \quad \text{and} \quad \sqrt{\frac{600}{\lambda}} \approx 14.1,
\]

and so we can group the primes by their inclusion in the intervals \(\left(\sqrt{\frac{600}{\lambda + 1}}, \sqrt{\frac{600}{\lambda}}\right)\) for \(\lambda \in \{1, 2, 3\}\). The primes 19 and 23 both lie in the interval \(\left(\sqrt{\frac{600}{2}}, \sqrt{600}\right]\) and both have a power of 18, giving support to this interval.
subdivision. We can see the primes in the intervals \( \left( \sqrt{\frac{600}{\lambda + 1}}, \sqrt{\frac{600}{\lambda}} \right) \) for \( \lambda = 1, 2, 3 \) have powers 18, 22 and 24 respectively. This is an identical progression to the one observed for primes \( p > \sqrt{r} \), as shown by the vector sum
\[
\begin{pmatrix}
18 \\
22 \\
24
\end{pmatrix}
= \begin{pmatrix}
6 \\
10 \\
12
\end{pmatrix} + \begin{pmatrix}
12 \\
12 \\
12
\end{pmatrix}.
\]

Since we are dealing with a 3 \( \times \) 3 determinant here we might argue that the primes in the interval \( \left( \sqrt{\frac{600}{\lambda + 1}}, \sqrt{\frac{600}{\lambda}} \right) \) have powers which exceed those of the larger primes in the interval \( \left( \frac{600}{\lambda + 1}, \frac{600}{\lambda} \right) \), for \( \lambda \in \{1, 2, 3\} \), by \( n(n + 1) \) because \( 12 = 3 \times 4 \). As a final observation, we see that the primes 11 and 13 both have a power of 24, but only 13 lies in the interval \( \left( \frac{600}{4}, \sqrt{\frac{600}{3}} \right) \). Keeping in line with our previous approach we extend this interval to \( \left( \sqrt[3]{600}, \sqrt{\frac{600}{3}} \right) \) where \( \sqrt[3]{600} \approx 8.4 \).

It is worth considering now whether this same routine of subdividing intervals could continue until we have partitioned all remaining primes. In other words, we wish to determine whether it is appropriate to partition the interval \((1, r]\) into the intervals
\[
\begin{cases}
\left( \frac{m}{\lambda + 1} \sqrt{\frac{r}{\lambda}}, \frac{m}{\lambda} \right) & \text{for all } 1 \leq \lambda < \min\{\sqrt{r} - 1, n\} \\
\left( \frac{m}{\lambda + 1} \sqrt[3]{r}, \frac{m}{\lambda} \right) & \text{when } \lambda = \min\{\lfloor \sqrt{r} - 1 \rfloor, n\},
\end{cases}
\]
for sufficiently many integers \( m \geq 1 \). When we increase \( m \) from 1 to 2 for a given \( \lambda \), the power of the primes also increase by \( n(n + 1) \), and so it might be for arbitrary \( m \) that
\[
e(m, \lambda) = (m - 1)(n(n + 1)) + (2\lambda n - \lambda(\lambda - 1))
= (m - 1)n^2 + (m - 1 + 2\lambda)n - \lambda(\lambda - 1).
\]

This allows us to make the following refinement of Conjecture 1.1.

**Conjecture 1.2.** Let \( n \geq 1 \) and \( r \geq 2 \) be integers and fix a prime \( p \leq r \). Set \( m = \left\lfloor \frac{\log r}{\log p} \right\rfloor \) and \( \lambda = \min\left\{ \left\lfloor \frac{r}{p^m} \right\rfloor, n \right\} \). The power of \( p \) in \( \text{denom}(W_n(r)\left[x\right]) \) is
\[
e(m, \lambda) = (m - 1)n^2 + (m - 1 + 2\lambda)n - \lambda(\lambda - 1).
\]

From Table 1.2 we see that the conjectured formula \( e(m, \lambda) \) agrees exactly with the power of each prime in the denominator of \( W_3(600)[x] \), with only a few exceptions.

---

**1. HANKEL DETERMINANTS WITH FORMAL PARAMETERS**
If these exceptions exist as unexpected cancellation of primes in the numerator with the denominator, then we can accept $e(m, \lambda)$ as providing a near-optimal upper bound on the power of each prime.

<table>
<thead>
<tr>
<th>Primes</th>
<th>$m$</th>
<th>$\lambda$</th>
<th>Interval</th>
<th>Power</th>
<th>$e(m, \lambda)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>307, ..., 599</td>
<td>1</td>
<td>1</td>
<td>$(600/2, 600/1] = (300, 600]$</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>211, ..., 293</td>
<td>1</td>
<td>2</td>
<td>$(600/3, 600/2] = (200, 300]$</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>29, ..., 199</td>
<td>1</td>
<td>3</td>
<td>$(\sqrt{600, 600/3]} \approx (24.5, 200]$</td>
<td>11 or 12</td>
<td>12</td>
</tr>
<tr>
<td>19, 23</td>
<td>2</td>
<td>1</td>
<td>$\left(\sqrt{600/2, \sqrt{600/1}}\right] \approx (17.3, 24.5]$</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>2</td>
<td>$\left(\sqrt{600/3, \sqrt{600/2}}\right] \approx (14.1, 17.3]$</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>11, 13</td>
<td>2</td>
<td>3</td>
<td>$\left(\sqrt{600, \sqrt{600/3}}\right] \approx (8.4, 14.1]$</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>1</td>
<td>$\left(\sqrt{600/2, \sqrt{600/1}}\right] \approx (6.7, 8.4]$</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>3</td>
<td>$\left(\sqrt{600, \sqrt{600/3}}\right] \approx (4.9, 5.8]$</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>2</td>
<td>$\left(\sqrt{600/3, \sqrt{600/2}}\right] \approx (2.9, 3.1]$</td>
<td>58</td>
<td>58</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>1</td>
<td>$\left(\sqrt{600/2, \sqrt{600/1}}\right] \approx (1.88, 2.04]$</td>
<td>99</td>
<td>102</td>
</tr>
</tbody>
</table>

Table 1.2. The primes less than 600 partitioned according to the parameters $m$ and $\lambda$, along with their actual and conjectured powers in the denominator of $W_3(600)[x]$.

The difficulty now lies in proving this conjectured bound. By way of investigation, let us attempt to simplify the determinant $W_n(r)[x]$ using elementary row operations. For our example to be manageable, we choose to focus on determining the power of 5 in the denominator of $W_3(55)[x]$. In much of the calculation we are not concerned with any expression without powers of 5 in its denominator, and so to that end we let the symbol $\rho$ denote such an expression (with possible subscripts and superscripts to differentiate between different expressions all represented as $\rho$). This convention allows for the following notational simplification:

$$x^2 - \sum_{k=1}^{55} \frac{1}{k^2} = \rho_1 - \left(\frac{1}{5^2} + \frac{1}{10^2} + \frac{1}{15^2} + \frac{1}{20^2} + \frac{1}{30^2} + \frac{1}{35^2} + \frac{1}{40^2} + \frac{1}{45^2} + \frac{1}{55^2}\right)$$

$$- \left(\frac{1}{25^2} + \frac{1}{50^2}\right)$$
This allows us to express the determinant \( W_3(55)[x] \) as

\[
\begin{vmatrix}
\frac{1}{5^2} & \frac{1}{12 \cdot 5^4} & \frac{1}{2 \cdot 5^4} \\
\frac{1}{5^3} & \frac{1}{13 \cdot 5^6} & \frac{1}{23 \cdot 5^6} \\
\frac{1}{5^4} & \frac{1}{14 \cdot 5^8} & \frac{1}{24 \cdot 5^8}
\end{vmatrix}
\]

Expanding this determinant yields \( W_3(55)[x] = \frac{\rho_8}{5^2} \). However, evaluating \( W_3(55)[x] \) in Maple gives us the exact form \( W_3(55)[x] = \frac{\rho_8}{5^2} \), and so we seek to find some row operations that would eliminate terms in the above representation of \( W_3(55)[x] \).

Denote the three rows of the above determinant by \( R_1, R_2 \) and \( R_3 \) respectively, and recall the determinant is invariant under certain row operations. Performing the row operations \( R_1 \leftarrow R_1 - (2 \cdot 5^2)R_2 \) and \( R_2 \leftarrow R_2 - (2 \cdot 5^2)R_3 \) gives

\[
\begin{vmatrix}
\frac{1}{5^2} & \frac{1}{12 \cdot 5^4} & \frac{1}{2 \cdot 5^4} \\
\frac{1}{5^3} & \frac{1}{13 \cdot 5^6} & \frac{1}{23 \cdot 5^6} \\
\frac{1}{5^4} & \frac{1}{14 \cdot 5^8} & \frac{1}{24 \cdot 5^8}
\end{vmatrix}
\]

Denote the three rows of this determinant by \( R'_1, R'_2 \) and \( R'_3 \). The common numerators in the first two rows suggests that we should perform the row operation \( R'_1 \leftarrow R'_1 - (1 \cdot 5^2)R'_2 \), giving

\[
\begin{vmatrix}
\frac{1}{5^2} & \frac{1}{5^3} & \frac{1}{5^4} \\
\frac{2}{5^3} & \frac{2}{5^4} & \frac{2}{5^5} \\
\frac{3}{5^4} & \frac{3}{5^5} & \frac{3}{5^6}
\end{vmatrix}
\]

Expanding this determinant gives \( W_3(55)[x] = \frac{\rho_{12}}{5^2} \) as required. Moreover, note in this example that since \( 2 \cdot 5^2 \leq 55 < 3 \cdot 5^2 \), the prime 5 has parameters \( m = 2 \) and \( \lambda = 2 \) and so Conjecture 1.2 gives \( e(2, 2) = 22 \).
We conclude our experimentation with an observation of the row operations we performed. The second row was resolved after only a single row operation, \( R_2 \leftarrow R_2 - (2 \cdot 5^2)R_3 \), whereas the first row required two successive row operations. Composing these two row operations into a single action gives

\[
R_1' - (1 \cdot 5^2)R_2' = (R_1 - (2 \cdot 5^2)R_2) - (1 \cdot 5^2) \cdot (R_2 - (2 \cdot 5^2)R_3)
\]

\[
= R_1 - ((1 + 2) \cdot 5^2)R_2 + ((1 \times 2) \cdot 5^4)R_3.
\]

This is superficially similar to the quartic

\[
(1 - x^2)(1 - 2x^2) = 1 - (1 + 2) \cdot x^2 + (1 \times 2) \cdot x^4,
\]

and so we end by remarking that a proof of Conjecture 1.2 with arbitrary choice of parameters would likely depend on the properties of certain polynomials constructed from the parameters \( m \) and \( \lambda \).

### 1.2. Proof via row reduction

Before we formalise any of the observations from Section 1.1 we further generalise the determinant \( W_n(r)[x] \). Recall the original determinants studied by Monien (see the beginning of Section 1.1) were based on evaluating the Riemann zeta function at consecutive positive integers. We propose the same observations made thus far ought to hold for the Riemann zeta function when evaluated along any positive arithmetic progression. To show this, we extend the previous determinant by including two new parameters \( a \) and \( d \) to control the arithmetic progression. So, fix integers \( a \geq 1 \) and \( d \geq 1 \) and let \( W_{n(a,d)}(r)[x] \) be

\[
W_{n(a,d)}(r)[x] = \det_{1 \leq i,j \leq n} \left( x_{a+(i+j-1)d} - \sum_{k=1}^{r} \frac{1}{k^{a+(i+j-1)d}} \right)
\]

\[
\begin{vmatrix}
    x_{a+d} - 1 - \cdots - \frac{1}{r^{a+d}} & \cdots & x_{a+nd} - 1 - \cdots - \frac{1}{r^{a+nd}} \\
    \vdots & \ddots & \vdots \\
    x_{a+nd} - 1 - \cdots - \frac{1}{r^{a+nd}} & \cdots & x_{a+(2n-1)d} - 1 - \cdots - \frac{1}{r^{a+(2n-1)d}}
\end{vmatrix},
\]
for all integers $n \geq 1$ and $r \geq 0$, and where the terms $x_{a+d}, \ldots, x_{a+(2n-1)d}$ are formal parameters. Therefore, specifying $x_{a+kd} = \zeta(a + kd)$ for $k \in \{1, 2, \ldots, 2n\}$ yields

$$W_n^{(a,d)}(r)[\zeta] = \det_{1 \leq i,j \leq n} \left( \zeta(a + (i + j - 1)d) - \sum_{k=1}^{r} \frac{1}{k^{a+(i+j-1)d}} \right)$$

and so our formulation is indeed a further generalisation of Monien’s determinants as we have $W_n^{(r+1,1)}(0)[\zeta] = H_n^{(r)}[\zeta]$ for all $n \geq 1$ and $r \geq 0$.

With this new formulation we can tidy up the clumsy conjectures from the previous section and prove a more general result.

**Theorem 1.3.** Let $n \geq 2$, $r \geq 2$, $a \geq 1$ and $d \geq 1$ be integers, and fix a prime $p \leq r$. Set $m = \left\lfloor \frac{\log r}{\log p} \right\rfloor$ and $\lambda = \left\lfloor \frac{r}{p^{m}} \right\rfloor$. If $W_n^{(a,d)}(r)[x] \neq 0$, then the power of $p$ in $\text{denom}(W_n^{(a,d)}(r)[x])$ is bounded above by

$$d(m - 1)n^2 + (a(m - 1) + 2d\lambda)n - \lambda(d\lambda - a) \quad \text{if } \lambda \leq n$$

and

$$dmn^2 + amn \quad \text{if } \lambda > n.$$ 

We approach the proof of this theorem by specifying a collection of row operations and observing how they simplify the determinant. We therefore need to refer to the underlying matrix that gives rise to the determinant $W_n^{(a,d)}(r)[x]$, which we denote by $W_n^{(a,d)}(r)[x]$.

**Proof.** Fix integers $n \geq 2$, $r \geq 2$, $a \geq 1$ and $d \geq 1$. Choose a prime $p \leq r$ and set the parameters $m$ and $\lambda$ as in the statement of the theorem (so that we have $\lambda p^m \leq r < (\lambda + 1)p^m$). Again let the symbol $\rho$ denote any expression that does not contain any powers of $p$ in its denominator. (This placeholder may be equal to zero in some cases, but we never divide by $\rho$ and it never results in the matrix having zero determinant as this was a hypothesis for the theorem.)

Using this notation we write the entries of the matrix $W_n^{(a,d)}(r)[x]$ as

$$\left( W_n^{(a,d)}(r)[x] \right)_{i,j} = x_{a+(i+j-1)d} - 1 - \frac{1}{2a+(i+j-1)d} - \cdots - \frac{1}{r(a+(i+j-1)d)}$$

and

$$= \frac{\rho_1^{ij}}{(p^{m-1})^{a+(i+j-1)d}} - \sum_{k=1}^{\lambda} \frac{1}{(kp^m)^{a+(i+j-1)d}}$$

$$= \frac{\rho_1^{ij}}{p(a+(i+j-1)d)(m-1)} - \sum_{k=1}^{\lambda} k^{(n-i)d} \frac{k_{(n-i)d}}{k^{a+(n+j-1)d}p(a+(i+j-1)d)m}.$$
Case $\lambda \geq n$. Here we do not perform any row operations but simply evaluate the determinant directly. Recall $\mathfrak{S}_n$ is the symmetric group of order $n$, and write $\text{sgn}(\pi)$ to denote the sign of the permutation $\pi \in \mathfrak{S}_n$. Then expanding the determinant gives

$$W_n^{(a,d)}(r)[x] = \left| \begin{array}{ccc} \frac{1,1}{p(a+d)m} & \cdots & \frac{1,n}{p(a+nd)m} \\ \vdots & \ddots & \vdots \\ \frac{n,1}{p(a+nd)m} & \cdots & \frac{n,n}{p(a+(2n-1)d)m} \end{array} \right| = \sum_{\pi \in \mathfrak{S}_n} \text{sgn}(\pi) \frac{1,\pi(1)}{p(a+(\pi(1))d)m} \frac{2,\pi(2)}{p(a+(\pi(2))d)m} \cdots \frac{n,\pi(n)}{p(a+(n+\pi(n)-1)d)m}.$$  

Now for any permutation $\pi \in \mathfrak{S}_n$ we have

$$\sum_{i=1}^{n} (a + (i + \pi(i) - 1)d)m = (a - d)mn + dm \sum_{i=1}^{n} i + dm \sum_{i=1}^{n} \pi(i) = (a - d)mn + dm(n+1) = dm^2 + amn,$$

and so the power of $p$ in the denominator of $W_n^{(a,d)}(r)[x]$ is at most $dm^2 + amn$.

Case $\lambda \leq n$. For each $\ell \in \{1, \ldots, \lambda\}$ define the function

$$f_\ell(x) = \prod_{k=\ell}^{\lambda} (x^d - k^d) = x^{(\lambda-\ell+1)d} + c_{\ell,\ell}x^{(\lambda-\ell)d} + \ldots + c_{\ell,\lambda},$$

where $c_{\ell,k} \in \mathbb{Z}$ for all $k \in \{\ell, \ldots, \lambda\}$. In particular, note that by construction we have $f_\ell(t) = 0$ for all integers $\ell \leq t \leq \lambda$. Define the $n \times n$ matrix $R_n$ as follows:

If $1 \leq i \leq n - \lambda$, then $(R_n)_{i,j} =

\begin{cases}
  c_{1,j-i}d^{(j-i)d} & \text{for } i < j \text{ and } j \leq \lambda + i \\
  0 & \text{for } i < j \text{ and } j > \lambda + i \\
  1 & \text{for } i = j \\
  0 & \text{for } i > j,
\end{cases}

If $n - \lambda < i \leq n$, then $(R_n)_{i,j} =

\begin{cases}
  c_{i-(n-\lambda-1),j-(n-\lambda)}d^{(j-i)d} & \text{for } i < j \\
  1 & \text{for } i = j \\
  0 & \text{for } i > j.
\end{cases}

The matrix $R_n$ is upper triangular and so its determinant is simply the product of its diagonal entries, i.e. $\det(R_n) = 1$. Write $V_n^{(a,d)}(r)[x]$ for the matrix obtained
from the product $R_n \times \mathcal{W}^{(a,d)}_{n}(r)[x]$, and observe

\[ W_{n}^{(a,d)}(r)[x] = \det \left( W_{n}^{(a,d)}(r)[x] \right) = \frac{\det \left( V_{n}^{(a,d)}(r)[x] \right)}{\det (R_n)} = \det \left( V_{n}^{(a,d)}(r)[x] \right). \]

All that remains is then to calculate the determinant of the matrix $V_{n}^{(a,d)}(r)[x]$. For ease of expression we abbreviate the matrices by omitting some of their parameters. Thus we obtain

\[
V_{i,j} = \begin{cases} 
W_{i,j} + \sum_{k=1}^{\lambda} c_{1,k} p_{a,d}^{k} m W_{i+k,j} & \text{for } 1 \leq i \leq n - \lambda \\
W_{i,j} + \sum_{k=1}^{n-i} c_{i-(n-\lambda-1),i-(n-\lambda-k)} p_{a,d}^{k} m W_{i+k,j} & \text{for } n - \lambda < i < n \\
W_{n,j} & \text{for } i = n.
\end{cases}
\]

We evaluate this multiplication in two stages. First suppose $1 \leq i \leq n - \lambda$, then by definition of $W_{n}^{(a,d)}(r)[x]$ we have

\[
V_{i,j} = W_{i,j} + \sum_{k=1}^{\lambda} c_{1,k} p_{a,d}^{k} m W_{i+k,j}
\]

\[
= \frac{\rho^{i,j}_{1}}{p(a+(i+j-1)d)(m-1)} - \frac{\rho^{(i-j),d}_{a+(a+(i+j-1)d)m}}{p(a+(a+(i+j-1)d)m-1)} + \sum_{k=1}^{\lambda} \left( \frac{\rho^{i+k,j}_{1} \cdot c_{1,k} p_{a,d}^{k} m}{p(a+(i+j+k-1)d)(m-1)} - \frac{c_{1,k} \rho^{(n-i-k),d}_{a+(a+(i+j+k-1)d)m}}{p(a+(a+(i+j+k-1)d)m-1)} \right)
\]

\[
= \frac{\rho^{i,j}_{1}}{p(a+(i+j-1)d)(m-1)} - \frac{\rho^{(n-i),d}_{a+(a+(i+j-1)d)m}}{p(a+(a+(i+j-1)d)m-1)} + \sum_{k=1}^{\lambda} \left( \frac{\rho^{i+k,j}_{1} \cdot c_{1,k} p_{a,d}^{k}}{p(a+(i+j+k-1)d)(m-1)} - \frac{c_{1,k} \rho^{(n-i-k)}_{a+(a+(i+j+k-1)d)m}}{p(a+(a+(i+j+k-1)d)m-1)} \right)
\]

\[
= \frac{\rho^{i,j}_{1}}{p(a+(i+j-1)d)(m-1)} - \frac{\rho^{(n-i),d}_{a+(a+(i+j-1)d)m}}{p(a+(a+(i+j-1)d)m-1)} - \sum_{k=1}^{\lambda} \sum_{t=1}^{\lambda} \frac{c_{1,k} \rho^{(n-i-k)}_{a+(a+(i+j-1)d)m}}{p(a+(a+(i+j-1)d)m-1)}.
\]
Grouping the summations and using the definition of \( f_i(t) \) gives

\[
V_{i,j} = \frac{\rho_{3}^{i,j}}{p(a+(i+j-1)d)(m-1)} - \sum_{t=1}^{\lambda} \frac{t^{(n-i)d} + \sum_{k=1}^{\lambda} c_{i,k} t^{(n-i-k)d}}{t a+(n+j-1)d p(a+(i+j-1)d)m} \\
= \frac{\rho_{3}^{i,j}}{p(a+(i+j-1)d)(m-1)} - \sum_{t=1}^{\lambda} \frac{t^{(n-i-\lambda)d} + t^{\lambda d} + \sum_{k=1}^{\lambda} c_{i,k} t^{(\lambda-k)d}}{t a+(n+j-1)d p(a+(i+j-1)d)m} \\
= \frac{\rho_{3}^{i,j}}{p(a+(i+j-1)d)(m-1)} - \sum_{t=1}^{\lambda} \frac{f_1(t)}{t a+(i+j+\lambda-1)d p(a+(i+j-1)d)m} \\
= \frac{\rho_{3}^{i,j}}{p(a+(i+j-1)d)(m-1)} .
\]

Next, suppose \( n - \lambda < i < n \). Again we have

\[
V_{i,j} = W_{i,j} + \sum_{k=1}^{n-i} c_{i-(n-\lambda-1),i-(n-\lambda-k)} p^{kd,m} W_{i+k,j} \\
= \left( \frac{\rho_{4}^{i,j}}{p(a+(i+j-1)d)(m-1)} - \sum_{t=1}^{\lambda} \frac{t^{(n-i)d}}{t a+(n+j-1)d p(a+(i+j-1)d)m} \right) \\
+ \sum_{k=1}^{n-i} \left( \frac{\rho_{4}^{i+k,j}}{p(a+(i+j+k-1)d)(m-1)} - \sum_{t=1}^{\lambda} \frac{c_{i-(n-\lambda-1),i-(n-\lambda-k)} t^{(n-i-k)d} p^{kd,m}}{t a+(n+j+k-1)d p(a+(i+j+k-1)d)m} \right) \\
= \left( \frac{\rho_{4}^{i,j}}{p(a+(i+j-1)d)(m-1)} - \sum_{t=1}^{\lambda} \frac{t^{(n-i)d}}{t a+(n+j-1)d p(a+(i+j-1)d)m} \right) \\
+ \sum_{k=1}^{n-i} \left( \frac{\rho_{4}^{i+k,j}}{p(a+(i+j+k-1)d)(m-1)} - \sum_{t=1}^{\lambda} \frac{c_{i-(n-\lambda-1),i-(n-\lambda-k)} t^{(n-i-k)d}}{t a+(n+j+k-1)d p(a+(i+j+k-1)d)m} \right) \\
- \sum_{k=1}^{n-i} \sum_{t=1}^{\lambda} \frac{c_{i-(n-\lambda-1),i-(n-\lambda-k)} t^{(n-i-k)d}}{t a+(n+j-1)d p(a+(i+j-1)d)m} .
\]

Grouping the summations and using the definition of \( f_i(t) \) gives

\[
V_{i,j} = \frac{\rho_{4}^{i,j}}{p(a+(i+j-1)d)(m-1)} - \sum_{t=1}^{\lambda} \frac{t^{(n-i)d} + \sum_{k=1}^{n-i} c_{i-(n-\lambda-1),i-(n-\lambda-k)} t^{(n-i-k)d}}{t a+(n+j-1)d p(a+(i+j-1)d)m} \\
= \frac{\rho_{4}^{i,j}}{p(a+(i+j-1)d)(m-1)} - \sum_{t=1}^{\lambda} \frac{f_{i-(n-\lambda-1)}(t)}{t a+(n+j-1)d p(a+(i+j-1)d)m} \\
= \frac{\rho_{4}^{i,j}}{p(a+(i+j-1)d)(m-1)} - \sum_{t=1}^{\lambda} \frac{f_{i-(n-\lambda-1)}(t)}{t a+(n+j-1)d p(a+(i+j-1)d)m} .
\]
Therefore

\[
V_{i,j} = \begin{cases} 
\frac{\rho^{i,j}_3}{p^{(a+(i+j-1)d)(m-1)}} & \text{for } 1 \leq i \leq n - \lambda \\
\frac{\rho^{i,j}_5}{p^{(a+(i+j-1)d)m}} & \text{for } n - \lambda < i \leq n.
\end{cases}
\]

Finally, we are able to expand the determinant \( V_n^{(a,d)}(r)[x] \) to obtain

\[
V_n^{(a,d)}(r)[x] = \sum_{\pi \in S_n} \text{sgn}(\pi) \left( \prod_{i=1}^{n-\lambda} \frac{\rho^{i,\pi(i)}_3}{p^{(a+(i+\pi(i)-1)d)(m-1)}} \right) \left( \prod_{i=n-\lambda+1}^{n} \frac{\rho^{i,\pi(i)}_5}{p^{(a+(i+\pi(i)-1)d)m}} \right).
\]

As before, for any permutation \( \pi \in S_n \) we have

\[
\sum_{i=1}^{n-\lambda} (a + (i + \pi(i) - 1)d)(m - 1) + \sum_{i=n-\lambda+1}^{n} (a + (i + \pi(i) - 1)d)m = \sum_{i=1}^{n} (a + (i + \pi(i) - 1)d)m - \sum_{i=1}^{n-\lambda} (a + (i + \pi(i) - 1)d) = dm n^2 + am n - (a - d)(n - \lambda) - d \sum_{i=1}^{n-\lambda} i - d \sum_{i=1}^{n-\lambda} \pi(i)
\]

\[
\leq dm n^2 + am n - (a - d)(n - \lambda) - \sum_{i=1}^{n-\lambda} i - \sum_{i=1}^{n-\lambda} i = dm n^2 + am n - (a - d)(n - \lambda) - \sum_{i=1}^{n-\lambda} i - \sum_{i=1}^{n-\lambda} i = dm n^2 + am n - (a - d)(n - \lambda) = dm n^2 - (a(m - 1) + 2d\lambda)n - \lambda(d\lambda - a),
\]

and hence the power of \( p \) in the denominator of \( W_n^{(a,d)}(r)[x] \) is no greater than

\[
d(m - 1)n^2 - (a(m - 1) + 2d\lambda)n - \lambda(d\lambda - a).
\]

The dense algebraic manipulations in the proof of Theorem 1.3 may obscure the details at work. In order to shed some light on this process, we revisit the determinant \( W_3^{(1,1)}(55)[x] \) from the end of Section 1.1 for a formal treatment.
Example 1.2.1. Consider the determinant $W^{(1,1)}_3(55)[x]$, with the intention of determining the power of 5 in its denominator. We wish to reduce the matrix

$$W^{(1,1)}_3(55)[x] = \begin{pmatrix}
x_2 - \sum_{k=1}^{55} \frac{1}{k^2} x_3 - \sum_{k=1}^{55} \frac{1}{k^3} x_4 - \sum_{k=1}^{55} \frac{1}{k^4} \\
x_3 - \sum_{k=1}^{55} \frac{1}{k^3} x_4 - \sum_{k=1}^{55} \frac{1}{k^4} x_5 - \sum_{k=1}^{55} \frac{1}{k^5} \\
x_4 - \sum_{k=1}^{55} \frac{1}{k^4} x_5 - \sum_{k=1}^{55} \frac{1}{k^5} x_6 - \sum_{k=1}^{55} \frac{1}{k^6}
\end{pmatrix}$$

using elementary row operations. Define the functions

$$f_1(x) = (x - 1)(x - 2) = x^2 - 3x + 2,$$

$$f_2(x) = x - 2,$$

and use these to construct the matrix

$$R_3 = \begin{pmatrix}
1 & -3 \cdot 25 & 2 \cdot 25^2 \\
0 & 1 & -2 \cdot 25 \\
0 & 0 & 1
\end{pmatrix}.$$
The matrix multiplication $V_3^{(1,1)}(55)[x] = R_3 \times W_3^{(1,1)}(55)[x]$ gives us

$$
\begin{pmatrix}
\frac{\rho_2^{1,1}}{5^4} - \frac{2}{2} \sum_{k=1}^{2} \frac{25^2 f_1(k)}{(25k)^4} & \frac{\rho_2^{1,2}}{5^4} - \frac{2}{2} \sum_{k=1}^{2} \frac{25^2 f_1(k)}{(25k)^4} & \frac{\rho_2^{1,3}}{5^4} - \frac{2}{2} \sum_{k=1}^{2} \frac{25^2 f_1(k)}{(25k)^4} \\
\frac{\rho_2^{2,1}}{5^3} - \frac{2}{2} \sum_{k=1}^{2} \frac{25 f_2(k)}{(25k)^3} & \frac{\rho_2^{2,2}}{5^3} - \frac{2}{2} \sum_{k=1}^{2} \frac{25 f_2(k)}{(25k)^3} & \frac{\rho_2^{2,3}}{5^3} - \frac{2}{2} \sum_{k=1}^{2} \frac{25 f_2(k)}{(25k)^3} \\
\frac{\rho_2^{3,1}}{5^2} - \frac{2}{2} \sum_{k=1}^{2} \frac{1}{(25k)^2} & \frac{\rho_2^{3,2}}{5^2} - \frac{2}{2} \sum_{k=1}^{2} \frac{1}{(25k)^2} & \frac{\rho_2^{3,3}}{5^2} - \frac{2}{2} \sum_{k=1}^{2} \frac{1}{(25k)^2}
\end{pmatrix}
$$

and we therefore conclude that the power of 5 in the denominator of $W_3^{(1,1)}(55)[x]$ does not exceed $2 + 2 \cdot 4 + 2 \cdot 6 = 22$. Indeed, we saw from Section 1.1 that it was equal to 22.

\[\diamondsuit\]

### 1.3. Bounding denominators

We conclude this chapter by estimating the magnitude of the denominator of $W_n^{(a,d)}(r)[x]$ as given in Theorem 1.3. Allow us first to prove some lemmas.

**Definition 1.4.** The **first Chebyshev function** $\vartheta(x)$ is the sum of the natural logarithms of all primes not exceeding $x$:

$$
\vartheta(x) = \sum_{p \leq x} \log p.
$$

Rosser and Schoenfield [35] gave a number of useful bounds for the Chebyshev function and other related functions over varying domains.

**Lemma 1.5.** For all $x > 0$ we have

$$
\vartheta(x) < 1.01624 x.
$$

In particular, $\vartheta(x) < x$ for $0 < x \leq 10^8$.

**Proof.** See Theorems 9 and 18 of [35]. \[\square\]
Lemma 1.6. For all integers \( n > 1 \) we have
\[
\sum_{k=1}^{n} \frac{1}{k} < 1 + \log n.
\]

**Proof.** Consider the function \( f(x) = \frac{1}{x} \) with the intention of estimating the integral \( \int_{1}^{n} f(x) \, dx \) for some integer \( n > 1 \). Since \( f(x) \) is convex, the right Riemann sum of \( f(x) \) over the interval \( [1, n] \) with \( n - 1 \) rectangles is an underestimate of the integral. Therefore we have
\[
\sum_{k=2}^{n} \frac{1}{k} < \int_{1}^{n} \frac{1}{x} \, dx = \log n,
\]
and so adding 1 to both sides proves the lemma. \( \Box \)

Lemma 1.7. For all integers \( n > 1 \) and \( x > 0 \) we have
\[
\sum_{k=1}^{n} \sqrt[k]{x} < (1 + \log n)x + n.
\]

**Proof.** Fix a positive integer \( k \) and consider the function \( f(x) = \sqrt[k]{x}, x > 0 \). Now \( f'(1) = \frac{1}{k} \) and so the tangent to \( y = f(x) \) at \( x = 1 \) has equation \( y = g(x) \) where
\[
g(x) = f(1) + f'(1)(x - 1) = 1 + \frac{1}{k}(x - 1) = \frac{1}{k}x + \frac{k - 1}{k}.
\]
Since \( f(x) \) is concave, the tangent satisfies \( g(x) \geq f(x) \) for all \( x > 0 \), and so for all integers \( n > 1 \) we have
\[
\sum_{k=1}^{n} \sqrt[k]{x} < \sum_{k=1}^{n} \left( \frac{1}{k}x + \frac{k - 1}{k} \right) < x \sum_{k=1}^{n} \frac{1}{k} + n.
\]
Combining this with Lemma 1.6 gives the desired bound. \( \Box \)

We can now estimate the denominator of \( W_{n}^{(a,d)}(r)[x] \).

**Proposition 1.8.** Let \( n \geq 2, r \geq 2, a \geq 1 \) and \( d \geq 1 \) be integers. Write \( W_{n}^{(a,d)}(r)[x] = \frac{P(x)}{q} \) where \( P(x) \in \mathbb{Z}[x_{a+d}, \ldots, x_{a+(2n-1)d}] \) and where \( q \in \mathbb{N} \). If \( P(x) \neq 0 \) then
\[
\log q < 1.01624 (dn^2 + an)((1 + \log 2 + \log \log r)r + 2 \log r).
\]

**Proof.** For each prime \( p \leq r \) define \( m_{p} = \left\lfloor \frac{\log r}{\log p} \right\rfloor \), and for ease of notation denote \( M = m_{2} \). Then by Theorem 1.3 we have for prime \( p \)
\[
q \leq \prod_{1 \leq p \leq r} p^{dn^2 + amn} = \prod_{m=1}^{M} \left( \prod_{1 \leq p \leq \sqrt[r]{M}} p^{dn^2 + an} \right).
\]
Taking the logarithm of both sides and using Lemma 1.5 we obtain
\[
\log q \leq (dn^2 + an) \sum_{m=1}^{M} \theta \left( \frac{m}{\sqrt{r}} \right) < 1.01624 (dn^2 + an) \sum_{m=1}^{M} \frac{m}{\sqrt{r}},
\]
and simplifying using Lemma 1.7 yields
\[
\log q < 1.01624 (dn^2 + an) ((1 + \log M)r + M).
\]
Finally observe that \( M \leq \frac{\log r}{\log 2} < 2 \log r \), and so
\[
\log q < 1.01624 (dn^2 + an) (1 + \log 2 + \log \log r)r + 2 \log r),
\]
as required. \(\square\)

We note, in passing, that the constant 1.01624 appearing in Proposition 1.8 can be replaced by \( 1 + \epsilon \) provided other parameters are taken to be sufficiently large.

In Chapter 4 we revisit \( W_n^{(a,d)}(r)[\zeta] \) in an attempt to describe its asymptotics as \( n \to \infty \). To bring this chapter to a close we draw attention to the following conjecture, which is based on other experimental observations.

**Conjecture 1.9.** Let \( n \geq 2, r \geq 2, a \geq 1 \) and \( d \geq 1 \) be integers. Then
\[
\log W_n^{(a,d)}(r)[\zeta] \asymp -n^2 \log n,
\]
for all sufficiently large \( n \) and \( r \).

This expected asymptotic behaviour allows us to make some mild statements about the nature of the Riemann zeta function. If we were to assume that \( \zeta(s) \) took rational values at appropriate arguments then we would have \( W_n^{(a,d)}(r)[\zeta] \in \mathbb{Q} \). We have already seen in Proposition 1.8 that the contribution to the denominator of \( W_n^{(a,d)}(r)[\zeta] \) from omitting the \( r \) terms in each entry does not produce the growth expected from the asymptotic in Conjecture 1.9. This means that zeta values which were assumed rational must possess denominators of a certain magnitude to make the dominant contribution required to produce the observed asymptotics. We do not give exact details here but defer a complete treatment of the bounds on these determinants and its consequences for Chapter 4.
Hankel Determinants and Rational Approximations

In this chapter we introduce Cauchy matrices for the purpose of studying their intersection with Hankel matrices; the Hilbert matrix and some generalisations of it. This enables us to evaluate various Hankel determinants using their additional structure as Cauchy determinants.

In Section 2.1 we reproduce several classical results on the Hilbert matrix and extend these results to our particular generalisation. Section 2.2 introduces the concept of Padé approximations as a generalisation of Taylor approximations, where we see that existence of these approximations is linked to the nonvanishing of certain Hankel determinants. By combining the results of Sections 2.1 and 2.2, we turn our attention in Section 2.3 towards producing an original proof of the irrationality of certain values of the natural logarithm. This is followed by a failed attempt to replicate the process for Catalan’s constant, and an explanation of the limitations of this approach.

2.1. The Hilbert matrix

In this section we investigate the determinant and minors of Cauchy matrices.

**Definition 2.1.** Let \( n \) be a positive integer, and let \( \{x_i\}_{i=1}^{n} \) and \( \{y_j\}_{j=1}^{n} \) be sequences such that \( x_i + y_j \neq 0 \) for all integers \( 1 \leq i, j \leq n \). A **Cauchy matrix** is an \( n \times n \) matrix of the form

\[
C = \begin{pmatrix}
\frac{1}{x_1+y_1} & \frac{1}{x_1+y_2} & \cdots & \frac{1}{x_1+y_n} \\
\frac{1}{x_2+y_1} & \frac{1}{x_2+y_2} & \cdots & \frac{1}{x_2+y_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{x_n+y_1} & \frac{1}{x_n+y_2} & \cdots & \frac{1}{x_n+y_n}
\end{pmatrix}
\tag{2.1}
\]

It is clear that if the sequence \( \{x_i\}_{i=1}^{n} \) contains any repeated elements, then its determinant is zero as at least two rows of the matrix are identical. Similarly if \( \{y_j\}_{j=1}^{n} \) contains any repeated elements then at least two columns are identical and
so the determinant is again zero. Lastly, one sequence cannot contain the additive inverse of any elements from the other sequence as this would cause matrix entries to be undefined. All together, this information about the zeroes and poles of the determinant produces the following closed form.

**Proposition 2.2** (Pólya and Szegő [29], 1971). Let $C$ be an $n \times n$ Cauchy matrix of the form (2.1). The determinant of $C$ is

$$\det(C) = \prod_{1 \leq i < j \leq n} (x_j - x_i) \prod_{1 \leq i < j \leq n} (y_j - y_i) \prod_{1 \leq i, j \leq n} (x_i + y_j).$$

(2.2)

An important observation about Cauchy matrices is that any submatrix of a Cauchy matrix is itself a Cauchy matrix. Specifically, if we remove the $k$th row and $m$th column of a Cauchy matrix then we have effectively removed the $k$th term and $m$th term of the sequences $\{x_i\}_{i=1}^n$ and $\{y_j\}_{j=1}^n$ respectively. This immediately gives the following corollary.

**Corollary 2.3.** Let $C$ be an $n \times n$ Cauchy matrix, and write $C_{km}^n$ for the matrix obtained by removing the $k$th row and $m$th column of $C$. Then

$$\det(C_{km}^n) = \prod_{1 \leq i < j \leq n, i \neq k, j \neq k} (x_j - x_i) \prod_{1 \leq i < j \leq n, i \neq m, j \neq m} (y_j - y_i) \prod_{1 \leq i, j \leq n, i \neq k, j \neq m} (x_i + y_j).$$

(2.3)

So what does this have to do with Hankel determinants? The classic example of a Hankel matrix that is also a Cauchy matrix is the Hilbert matrix, first introduced by Hilbert [20] in 1894. As an $n \times n$ Hankel matrix, the sequence $\{\frac{1}{k+1}\}_{k=2}^{2n}$ produces the Hilbert matrix, while as a Cauchy matrix it is expressed in terms of the sequences $\{x_i\}_{i=1}^n$ and $\{y_j\}_{j=1}^n$ where $x_i = i$ and $y_j = j - 1$, both producing the following matrix of unit fractions

$$\mathcal{H}_n = \begin{pmatrix}
1 & \frac{1}{2} & \cdots & \frac{1}{n} \\
\frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1}
\end{pmatrix}.$$
Much is known about this matrix: see Choi [13] for a number of results. As we are interested in determinants, we give the following closed form for the determinant of the \( n \times n \) Hilbert matrix.

**Proposition 2.4** (Choi [13], 1983). The determinant of the \( n \times n \) Hilbert matrix is given by

\[
\mathcal{H}_n = \begin{vmatrix}
1 & \frac{1}{2} & \ldots & \frac{1}{n} \\
\frac{1}{2} & \frac{1}{3} & \ldots & \frac{1}{n+1} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{n} & \frac{1}{n+1} & \ldots & \frac{1}{2n-1}
\end{vmatrix} = \frac{\left( \prod_{k=1}^{n} k! \right)^3}{2^{n-1} \prod_{k=n}^{n-1} k!} = \frac{1}{n! \prod_{k=0}^{n-1} \binom{n}{k} \binom{n+k}{n}},
\]

and hence is the reciprocal of a positive integer.

We wish to consider more general Hilbert matrices whose first entries are smaller unit fractions. In particular, for a given integer \( r \geq 0 \), we wish to evaluate

\[
\mathcal{H}_n^{(r)} = \begin{vmatrix}
\frac{1}{1+r} & \frac{1}{2+r} & \ldots & \frac{1}{n+r} \\
\frac{1}{2+r} & \frac{1}{3+r} & \ldots & \frac{1}{n+r+1} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{n+r} & \frac{1}{n+r+1} & \ldots & \frac{1}{2n+r-1}
\end{vmatrix}.
\]

To that end, we offer the following generalisation of Proposition 2.4.

**Proposition 2.5.** The determinant of the general \( n \times n \) Hilbert matrix is given by

\[
\mathcal{H}_n^{(r)} = \frac{\left( \prod_{k=1}^{n-1} k! \right)^2}{2^{n-1} \prod_{k=n}^{n-1} (r+k)!} = \frac{1}{n! \prod_{k=0}^{n-1} \binom{n}{k} \binom{n+r+k}{n}},
\]

and hence is the reciprocal of a positive integer.

**Proof.** This matrix is a Cauchy matrix whose sequences \( \{x_i\}_{i=1}^{n} \) and \( \{y_j\}_{j=1}^{n} \) are given by \( x_i = i + r \) and \( y_j = j - 1 \). Let us consider the three products in
Equation (2.2):

\[
\prod_{1 \leq i < j \leq n} (x_j - x_i) = \prod_{1 \leq i \leq n} \left( \prod_{i < j \leq n} (j - i) \right) = \prod_{k=1}^{n-1} k!,
\]

\[
\prod_{1 \leq i < j \leq n} (y_j - y_i) = \prod_{1 \leq i \leq n} \left( \prod_{i < j \leq n} (j - i) \right) = \prod_{k=1}^{n-1} k!,
\]

and

\[
\prod_{1 \leq i, j \leq n} (x_i + y_j) = \prod_{1 \leq i, j \leq n} \left( \prod_{1 \leq j \leq n} (i + j + r - 1) \right) = \prod_{k=0}^{n-1} \frac{(n + r + k)!}{(r + k)!}.
\]

Therefore by Proposition 2.2 we have

\[
\mathcal{H}_n^{(r)} = \frac{\left( \prod_{k=1}^{n-1} k! \right)^2}{\prod_{k=0}^{n-1} \left( \frac{(n + r + k)!}{(r + k)!} \right) \prod_{k=n}^{2n-1} (r + k)!},
\]

which is the desired result. \(\Box\)

As noted, removing a single row and column of a Cauchy matrix produces another Cauchy matrix. This allows for a closed form for some minors of the generalised Hilbert matrix.

**Proposition 2.6.** The determinant of the matrix obtained by removing the \(k\)th row and \(m\)th column of the general \(n \times n\) Hilbert matrix is given by

\[
(\mathcal{H}_n^{(r)})_{k, m} = \mathcal{H}_n^{(r)} \cdot (k + m + r - 1) \binom{k + n + r - 1}{n - m} \binom{n + m + r - 1}{n - k} \cdot \binom{k + m + r - 2}{k - 1} \binom{k + m + r - 2}{m - 1}.
\]

**Proof.** Consider the three products from Equation (2.3). We have

\[
\prod_{1 \leq i < j \leq n} (x_j - x_i) = \frac{\prod_{1 \leq i < j \leq n} (x_j - x_i)}{\prod_{1 \leq i < k} (x_k - x_i) \cdot \prod_{k < j \leq n} (x_j - x_k)},
\]

\[
\prod_{1 \leq i < j \leq n} (y_j - y_i) = \frac{\prod_{1 \leq i < j \leq n} (y_j - y_i)}{\prod_{1 \leq i < m} (y_m - y_i) \cdot \prod_{m < j \leq n} (y_j - y_m)}.
\]
OE 2.3 we have

\[
\prod_{\substack{1 \leq i, j \leq n \\
i \neq k, j \neq m}} (x_i + y_j) = \frac{\left( \prod_{1 \leq i, j \leq n} (x_i + y_j) \right) \cdot \left( \prod_{1 \leq i \leq n} (x_k + y_m) \right)}{\left( \prod_{1 \leq i \leq n} (x_i + y_m) \right) \cdot \left( \prod_{1 \leq j \leq n} (x_k + y_j) \right)}.
\]

Recalling the sequences \( \{x_i\}_{i=1}^n \) and \( \{y_j\}_{j=1}^n \) are given by \( x_i = i + r \) and \( y_j = j - 1 \), we can evaluate the above products to obtain:

\[
\prod_{1 \leq i < j \leq n} (x_j - x_i) = \prod_{1 \leq i \leq n} \left( \prod_{i < j \leq n} (j - i) \right) = \prod_{\ell=1}^{n-1} \ell!,
\]

\[
\prod_{1 \leq i < k} (x_k - x_i) = \prod_{i=1}^{k-1} (k - i) = (k - 1)!,
\]

\[
\prod_{k < j \leq n} (x_j - x_k) = \prod_{j=k+1}^{n} (j - k) = (n - k)!,
\]

\[
\prod_{1 \leq i < j \leq n} (y_j - y_i) = \prod_{1 \leq i \leq n} \left( \prod_{i < j \leq n} (j - i) \right) = \prod_{\ell=1}^{n-1} \ell!,
\]

\[
\prod_{1 \leq i < m} (y_m - y_i) = \prod_{i=1}^{m-1} (m - i) = (m - 1)!,
\]

\[
\prod_{m < j \leq n} (y_j - y_m) = \prod_{j=m+1}^{n} (j - m) = (n - m)!,
\]

\[
\prod_{1 \leq i, j \leq n} (x_i + y_j) = \prod_{1 \leq i \leq n} \left( \prod_{1 \leq j \leq n} (i + j + r - 1) \right) = \prod_{\ell=0}^{n-1} \frac{(n + r + \ell)!}{(r + \ell)!},
\]

\[
\prod_{1 \leq i \leq n} (x_i + y_m) = \prod_{i=1}^{n} (i + m + r - 1) = \frac{(n + m + r - 1)!}{(m + r - 1)!},
\]

and

\[
\prod_{1 \leq j \leq n} (x_k + y_j) = \prod_{j=1}^{n} (k + j + r - 1) = \frac{(k + n + r - 1)!}{(k + r - 1)!}.
\]

Therefore by Corollary 2.3 we have

\[
\begin{pmatrix} H_n^{(r)} \end{pmatrix}_k^n = \left( \frac{\prod_{\ell=1}^{n-1} \ell!}{(k-1)!(n-k)!} \right) \cdot \left( \frac{\prod_{\ell=1}^{n-1} \ell!}{(m-1)!(n-m)!} \right) \cdot \left( \prod_{\ell=0}^{n-1} \frac{(n + r + \ell)!}{(r + \ell)!} \cdot \frac{(m + r + 1)!}{(k + r - 1)!} \right) \cdot \left( \frac{(n + m + r - 1)!}{(m + r - 1)!} \cdot \frac{(k + n + r - 1)!}{(k + r - 1)!} \right).
\]
which simplifies to

\[
(H_n^{(r)})_k^m = \frac{\left(\prod_{\ell=1}^{n-1} \ell! \right)^2}{\prod_{\ell=0}^{n-1} (n+r+\ell)! (r+\ell)!} \cdot \frac{(n+m+r-1)! (k+n+r-1)!}{(m+r-1)! (k+r-1)!} \cdot \frac{(n-1)! (n-m)! (m-1)!}{(k-1)! (m-1)!} \cdot (k+m+r-1)
\]

\[
= H_n^{(r)} \cdot (k+m+r-1) \left( \frac{k+n+r-1}{n-m} \right) \left( \frac{n+m+r-1}{n-k} \right) \cdot \left( \frac{k+m+r-2}{k-1} \right) \left( \frac{k+m+r-2}{m-1} \right),
\]

as required. □

### 2.2. Padé approximation

Suppose that we have a real-valued function \( f \) we wish to approximate. One of the simplest ways to do this is to evaluate \( f \)’s Taylor series expansion and truncate appropriately to obtain a Taylor polynomial of the required accuracy. This polynomial agrees with \( f \) at a specified point, and agrees with the slope and some higher derivatives of \( f \) at that point depending on the degree of the Taylor polynomial and the differentiability of \( f \). For a smooth continuous function this is often a straightforward task (see Example 2.2.1).

**Definition 2.7.** Let \( n \geq 0 \) be an integer and the function \( f : \mathbb{R} \to \mathbb{R} \) be \( n \) times differentiable at the point \( a \in \mathbb{R} \). The **Taylor polynomial of order** \( n \) is

\[
T_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n
\]

\[
= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.
\]

**Example 2.2.1.** The prototypical example when studying Taylor series is the exponential function \( f(x) = e^x \) at the point \( x = 0 \). Here, we simply have

\[
\frac{d^k}{dx^k} (e^x) = e^x \quad \implies \quad f^{(k)}(0) = 1,
\]

for all integers \( k \geq 0 \). This means the function \( f(x) = e^x \) has a Taylor approximation \( T_2(x) = 1 + x + \frac{1}{2}x^2 \) of order 2, which agrees with the function value and its first two derivatives at \( x = 0 \). Improved accuracy can be obtained from approximating using the Taylor polynomial \( T_3(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 \) of order 3, which agrees with the function value and its first three derivatives at \( x = 0 \). Figure 2.1 shows the
improvement in approximation gained by taking Taylor polynomials of higher order.

Figure 2.1. The first four Taylor approximations to $f(x) = e^x$ centred at $x = 0$, with increasing dash spacing for increasing order of approximation. The bold curve shows the original function.

Example 2.2.2. Consider the function $f(x) = \cos(x)$, locally at $x = \frac{\pi}{2}$. The periodic nature of the derivatives gives us

$$\frac{d^k}{dx^k}(\cos(x)) = \begin{cases} 
\cos(x) & \text{if } k \equiv 0 \pmod{4} \\
-\sin(x) & \text{if } k \equiv 1 \pmod{4} \\
-\cos(x) & \text{if } k \equiv 2 \pmod{4} \\
\sin(x) & \text{if } k \equiv 3 \pmod{4}
\end{cases}$$

$$\implies f^{(k)}(0) = \begin{cases} 
0 & \text{if } k \equiv 0 \pmod{4} \\
-1 & \text{if } k \equiv 1 \pmod{4} \\
0 & \text{if } k \equiv 2 \pmod{4} \\
1 & \text{if } k \equiv 3 \pmod{4},
\end{cases}$$

for all integers $k \geq 0$. Thus the function $f(x) = \cos(x)$ is approximated by the Taylor polynomial $T_3(x) = -1 \left(x - \frac{\pi}{2}\right) + \frac{1}{6} \left(x - \frac{\pi}{2}\right)^3$ of order 3, which agrees with
the function value and its first three derivatives at \( x = \frac{\pi}{2} \). However, it is worth noting in this case that \( T_{2n}(x) = T_{2n-1}(x) \) for all positive integers \( n \), and so the Taylor polynomial \( T_{2n}(x) \) has only degree \( 2n - 1 \). Figure 2.2 shows the improvement in approximation by taking Taylor polynomials of higher order.

\[ \text{Figure 2.2. The first four (odd) Taylor approximations to } f(x) = \cos(x) \text{ centred at } x = \frac{\pi}{2}, \text{ with increasing dash spacing for increasing order of approximation. The bold curve shows the original function.} \]

Examples 2.2.1 and 2.2.2 demonstrate that increasing the order of these approximating polynomials increases their fit to the original function, with an exact match being achieved in the limit. However, there are some obvious limitations. Polynomials are continuous and everywhere differentiable, and therefore only serve as good approximations to discontinuous functions when applied over restricted domains. Figure 2.3 illustrates the first few approximations to \( f(x) = x \tan(x) \), where we see an increasingly better approximation over the domain \(-\frac{\pi}{2} < x < \frac{\pi}{2}\).

To capture the discontinuities of the function we require an approximation that also possesses poles, and so the sensible choice would be to take the ratio of two polynomials (a rational function) as an improved approximation. The best rational function approximation to a given function is the Padé approximant.
Figure 2.3. The first four (even) Taylor approximations to $f(x) = x \tan(x)$ centred at $x = 0$, with increasing dash spacing for increasing order of approximation. The bold curve shows the original function.

**Definition 2.8.** For a given function $f(x)$, a Padé approximant of $f$ of order $[L/M]$, written $[L/M]f(x)$, is a rational function $\frac{P(x)}{Q(x)}$ satisfying

$$f(x) - \frac{P(x)}{Q(x)} = O\left(x^{L+M+1}\right),$$

where the polynomials $P(x)$ and $Q(x)$ have (exact) degrees $L$ and $M$ respectively.

**Example 2.2.3.** Suppose we wish to find the Padé approximant of order $[1/1]$ for $f(x) = e^x$. Then we seek two polynomials $P(x) = a_0 + a_1 x$ and $Q(x) = b_0 + b_1 x$, such that

$$e^x - \frac{a_0 + a_1 x}{b_0 + b_1 x} = O\left(x^3\right).$$

Using the Taylor polynomial approximation for $e^x$ and multiplying through by $Q(x)$, we obtain

$$\left(1 + x + \frac{1}{2} x^2 + \cdots\right) \cdot (b_0 + b_1 x) - (a_0 + a_1 x) = O\left(x^3\right).$$
Equating the coefficients of \(x^0, x^1\) and \(x^2\) gives us the following three equations in four unknowns

\[
\begin{align*}
  b_0 - a_0 &= 0 \\
  b_0 + b_1 - a_1 &= 0 \\
  \frac{1}{2}b_0 + b_1 &= 0
\end{align*}
\]

\[
\Rightarrow \quad \begin{align*}
  a_0 &= b_0 \\
  a_1 &= \frac{1}{2}b_0 \\
  b_1 &= -\frac{1}{2}b_0.
\end{align*}
\]

The additional parameter here corresponds to scaling all coefficients in the rational function, thereby changing its appearance but not its value. Choosing the smallest integer solution allows us to write

\[
\frac{[1/1]}{f(x)} = \frac{P(x)}{Q(x)} = \frac{2 + x}{2 - x}.
\]

It can be a useful convention to choose the scaling such that the constant term in the denominator is 1 (i.e. set \(b_0 = 1\)), which leaves us with the same number of equations as unknowns.

To find the Padé approximant of order \([2/1]\), equate the coefficients of \(x^0, x^1, x^2\) and \(x^3\):

\[
\left(1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \ldots\right) \cdot (1 + b_1x) - (a_0 + a_1x + a_2x^2) = O\left(x^4\right),
\]

giving a similar four equations in four unknowns. Therefore we have

\[
\frac{[2/1]}{f(x)} = \frac{1 + \frac{2}{3}x + \frac{1}{3}x^2}{1 - \frac{2}{3}x} = \frac{6 + 4x + x^2}{6 - 2x},
\]

as the Padé approximant to \(f(x) = e^x\) of order \([2/1]\). Figure 2.4 shows several approximants of varying order, which when compared to Figure 2.1 illustrates the improved rate of convergence of Padé approximants over Taylor polynomials, even in the case where the denominator is only a linear approximation.

\[\diamondsuit\]

**Example 2.2.4.** Consider again approximating \(f(x) = \cos(x)\) at \(x = \frac{\pi}{2}\) and suppose we wish to find the Padé approximant of order \([1/1]\). Then we seek two polynomials \(P(x) = a_0 + a_1(x - \frac{\pi}{2})\) and \(Q(x) = b_0 + b_1(x - \frac{\pi}{2})\), such that

\[
\cos(x) - \frac{a_0 + a_1\left(x - \frac{\pi}{2}\right)}{b_0 + b_1\left(x - \frac{\pi}{2}\right)} = O\left((x - \frac{\pi}{2})^3\right).
\]
2.2. PADÉ APPROXIMATION

Figure 2.4. The Padé approximants of \( f(x) = e^x \) at \( x = 0 \) of order \([L/1]\) for \( L \in \{0, 1, 2, 3\} \), with increasing dash spacing for increasing order of approximation. The bold curve shows the original function.

Using the Taylor polynomial approximation for \( \cos(x) \), and multiplying through by \( Q(x) \), we obtain

\[
\left( -1 \left( x - \frac{\pi}{2} \right) + \frac{1}{6} \left( x - \frac{\pi}{2} \right)^3 - \cdots \right) \cdot \left( b_0 + b_1 \left( x - \frac{\pi}{2} \right) \right) \\
- \left( a_0 + a_1 \left( x - \frac{\pi}{2} \right) \right) = O \left( \left( x - \frac{\pi}{2} \right)^3 \right).
\]

Setting \( b_0 = 1 \) and equating coefficients gives three equations in three unknowns. Therefore we have

\[
[1/1] f(x) = \frac{0 - 1 \left( x - \frac{\pi}{2} \right)}{1 + 0 \left( x - \frac{\pi}{2} \right)} = -1 \left( x - \frac{\pi}{2} \right).
\]

Finding the Padé approximant of order \([2/1]\) requires equating coefficients such that

\[
\left( -1 \left( x - \frac{\pi}{2} \right) + \frac{1}{6} \left( x - \frac{\pi}{2} \right)^3 - \cdots \right) \cdot \left( b_0 + b_1 \left( x - \frac{\pi}{2} \right) \right) \\
- \left( a_0 + a_1 \left( x - \frac{\pi}{2} \right) + a_2 \left( x - \frac{\pi}{2} \right)^2 \right) = O \left( \left( x - \frac{\pi}{2} \right)^4 \right),
\]
resulting in a similar four equations in five unknowns. Choosing the smallest integer solution means we are forced to conclude
\[ [2/1]f(x) = \frac{0 + 0(x - \frac{\pi}{2}) - 1(x - \frac{\pi}{2})^2}{0 + 1(x - \frac{\pi}{2})} = -1(x - \frac{\pi}{2}) = [1/1]f(x). \]

The difficulty here arises from the requirement that Padé approximants of order \([L/M]\) should agree with the first \(L + M + 1\) terms in the Taylor series of \(f(x)\). In this case we have
\[
\begin{align*}
\frac{\cos(x) - [2/1]f(x)}{0 + 0(x - \frac{\pi}{2}) - 1(x - \frac{\pi}{2})^2} &= \left( -1(x - \frac{\pi}{2}) + \frac{1}{6}(x - \frac{\pi}{2})^3 - \cdots \right) - 1(x - \frac{\pi}{2}) \\
&= \frac{1}{6}(x - \frac{\pi}{2})^3 - \frac{1}{120}(x - \frac{\pi}{2})^5 + \cdots \\
&\neq O \left( \left( x - \frac{\pi}{2} \right)^4 \right),
\end{align*}
\]
and so Definition 2.8 is not satisfied. Thus we must conclude there is no rational function with the required order of approximation for \(f\). For the sake of comparison with Figure 2.2, Figure 2.5 shows the existence of several approximants of varying order and again demonstrates the more rapid rate of convergence.

To better understand how the previous example failed to produce a Padé approximant we need to refine the above procedure. In general, seeking to find a Padé approximant of order \([L/M]\) requires solving \(L + M + 1\) equations in \(L + M + 2\) unknowns. Writing \(c_j\) for the coefficient of \(x^j\) in the Taylor series approximation of \(f(x)\) (and for consistency defining \(c_j = 0\) for \(j < 0\), and equating coefficients of \(x^{L+1}, x^{L+2}, \ldots, x^{L+M}\) gives the following \(M\) equations in the \(M + 1\) unknowns \(b_0, \ldots, b_M\)
\[
\begin{align*}
b_M c_{L-M+1} + b_M c_{L-M+2} + \cdots + b_0 c_{L+1} &= 0, \\
b_M c_{L-M+2} + b_M c_{L-M+3} + \cdots + b_0 c_{L+2} &= 0, \\
&\vdots \\
b_M c_L + b_M c_{L+1} + \cdots + b_0 c_{L+M} &= 0.
\end{align*}
\]
Figure 2.5. The Padé approximants of \( f(x) = \cos(x) \) at \( x = \frac{\pi}{2} \) of order \( \lfloor L/2 \rfloor \) for \( L \in \{1, 3, 5, 7\} \), with increasing dash spacing for increasing order of approximation. The bold curve shows the original function.

Moving the last term to the right-hand side produces

\[
\begin{pmatrix}
c_{L-M+1} & c_{L-M+2} & \cdots & c_L \\
c_{L-M+2} & c_{L-M+3} & \cdots & c_{L+1} \\
\vdots & \vdots & \ddots & \vdots \\
c_L & c_{L+1} & \cdots & c_{L+M-1}
\end{pmatrix}
\begin{pmatrix}
b_M \\
b_{M-1} \\
\vdots \\
b_1
\end{pmatrix}
\begin{pmatrix}
c_{L+1} \\
\vdots \\
c_{L+M}
\end{pmatrix}
= -b_0
\begin{pmatrix}
c_{L+2} \\
\vdots \\
c_{L+M-1}
\end{pmatrix}.
\] (2.4)

For a given \( b_0 \) we expect Equation (2.4) to have a unique solution, and so we require that the \( M \times M \) (Hankel) matrix be invertible. Moreover, we require \( b_0 \neq 0 \) because otherwise \( b_i = 0 \) for all \( 1 \leq i \leq M \), which makes \( Q(x) \) identically zero. Without loss of generality, set \( b_0 \) to the (non-zero) determinant of the \( M \times M \) matrix in Equation (2.4). We then use Cramer’s rule to solve Equation (2.4) and hence express \( Q(x) \) in terms of the coefficients of the Taylor series approximation of \( f(x) \). We can express \( P(x) \) in a similar way by solving a similar system of equations. This procedure is carried out by Baker [4, pp.2–3] and allows for Padé approximantion to arbitrary order (if one exists).
Lemma 2.9 (Baker [4], 1996). Let \( f(x) \) be a function with power series \( \sum_{n=0}^{\infty} c_n x^n \), and define

\[
P(x) = \begin{vmatrix} c_{L-M+1} & \cdots & c_L & c_{L+1} \\ \vdots & \ddots & \vdots & \vdots \\ c_{L-M} & \cdots & c_{L+M-1} & c_{L+M} \\ \sum_{i=0}^{L-M} c_{i} x^{M+i} & \cdots & \sum_{i=0}^{L-1} c_{i} x^{i+1} & \sum_{i=0}^{L} c_{i} x^{i} \end{vmatrix}
\]

and

\[
Q(x) = \begin{vmatrix} c_{L-M+1} & \cdots & c_L & c_{L+1} \\ \vdots & \ddots & \vdots & \vdots \\ c_{L-M} & \cdots & c_{L+M-1} & c_{L+M} \\ x^M & \cdots & x & 1 \end{vmatrix}
\]

for some fixed non-negative integers \( L \) and \( M \). If \( Q(0) \neq 0 \) then there exists a Padé approximant to \( f(x) \) of order \( \lfloor L/M \rfloor \) given by the rational function \( \frac{P(x)}{Q(x)} \).

Accuracy is the final point to consider in our introduction to Padé approximants. Padé approximants are designed to agree with the first \( L + M + 1 \) terms of a Taylor series approximation, so we might ask how closely they agree with the remaining terms to which they are not equal. Fortunately, the approximants have been widely studied for their connection and use within: orthogonality and moment problems [39], rational interpolation for partial differential or integral equations [14], and a variety of other applications [42]. The relationship between Padé approximants and formal orthogonal polynomials was studied in detail by Brezinski [11] who demonstrated the following, which gives the accuracy of an approximant by showing how it compares with the remaining terms to which it is not equal.

Proposition 2.10 (Brezinski [11], 1980). Let \( f(x) \) be a function with power series \( \sum_{n=0}^{\infty} c_n x^n \), and define

\[
H_n^{(r)}[c] = \begin{vmatrix} c_r & \cdots & c_{n+r-1} \\ \vdots & \ddots & \vdots \\ c_{n+r-1} & \cdots & c_{2n+r-2} \end{vmatrix}
\]
for integers \( n \geq 1, r \geq 0 \). If \( H_M^{[L-M+1]}[c] \neq 0 \), then the Padé approximant \([L/M]_f(x)\) exists and satisfies

\[
f(x) - [L/M]_f(x) = \frac{H_M^{[L-M+1]}[c]}{H_M^{[L-M+1]}[c]} x^{L+M+1} + \mathcal{O}(x^{L+M+2}).
\]

2.3. Applications toward irrationality

In addition to the applications given at the end of Section 2.2, Padé approximants are also useful in irrationality proofs. As irrational numbers can be approximated to arbitrary precision by rational numbers, any irrationality proof requires sufficiently “good” rational approximations to a given real number. As we have just seen that Padé approximants are useful for approximating real-valued functions, evaluating a Padé approximant at a rational number might just give the “good” rational approximation we require. To that end, we use the previous results of this chapter to give an original and roundabout proof of the following (for the usual elementary proof of the irrationality of integer values of the natural logarithm, see Feldvoss [18]).

**Theorem 2.11.** The real number \( \log \left(1 + \frac{1}{b}\right) \) is irrational for integers \( b \geq 10 \).

This result is known for all integers \( b \geq 1 \), but we require \( b \geq 10 \) for the bound in Equation (2.8). Removing this limitation would require improvements to be made to Lemma 2.15. We will make use of a simple irrationality criterion to prove Theorem 2.11.

**Lemma 2.12** (Van Assche [39], 2006). Let \( x \in \mathbb{R} \). Suppose there are two integer sequences \( \{p_k\}_{k \geq 1} \) and \( \{q_k\}_{k \geq 1} \) such that

(i) \( q_kx - p_k \neq 0 \) for all integers \( k \geq 1 \), and

(ii) \( \lim_{k \to \infty} |q_kx - p_k| = 0 \).

Then \( x \) is irrational.

**Proof.** Suppose \( x \) is rational and write \( x = \frac{p}{q} \) where \( p \) and \( q > 0 \) are integers satisfying \( \gcd(p, q) = 1 \), and suppose that there exist integer sequences as in the statement of the lemma. By condition (i) we have \( q_k \frac{p}{q} - p_k \neq 0 \), and so \( 0 \neq q_k p - p_k q \in \mathbb{Z} \) also. Thereby \( |q_k p - p_k q| \geq 1 \) for all integers \( k \geq 1 \). However, this yields

\[
\lim_{k \to \infty} |q_kx - p_k| = \lim_{k \to \infty} \left| q_k \frac{p}{q} - p_k \right| = \lim_{k \to \infty} \frac{|q_k p - p_k q|}{q} \geq \frac{1}{q} > 0,
\]

which contradicts condition (ii). Hence \( x \) cannot be rational. \( \square \)
The evaluation of appropriate Padé approximants ultimately leads to expressions involving Legendre polynomials, and so we provide a definition with a few crucial identities these polynomials satisfy. A good reference for the definition and properties of Legendre polynomials is the monograph by Borwein and Erdélyi [8].

**Definition 2.13.** For integers \( n \geq 0 \), the *Legendre polynomials* \( P_n(x) \) are the coefficients in the Taylor series expansion
\[
\frac{1}{\sqrt{1 - 2xt + t^2}} = \sum_{n=0}^{\infty} P_n(x)t^n, \quad |t| < 1.
\]

The *shifted Legendre polynomials* \( \tilde{P}_n(x) \) can be written explicitly as
\[
\tilde{P}_n(x) = (-1)^n \sum_{k=0}^{n} \binom{n}{k} \binom{n+k}{k} (-x)^k,
\]
and satisfy the shifted relation \( \tilde{P}_n(x) = P_n(2x - 1) \).

**Lemma 2.14.** For all integers \( n \geq 0 \), we have
\[
\sum_{k=0}^{n} (-1)^{n+k} \binom{n}{k} \binom{n+k}{k} = 1.
\]

**Proof.** Using Definition 2.13, we obtain
\[
(-1)^n \sum_{k=0}^{n} (-1)^k \binom{n}{k} \binom{n+k}{k} = \tilde{P}_n(1) = P_n(1).
\]

Moreover, for \( |t| < 1 \) we have
\[
\sum_{n=0}^{\infty} P_n(1)t^n = \frac{1}{\sqrt{1 - 2t + t^2}} = \frac{1}{1 - t} = \sum_{n=0}^{\infty} t^n.
\]
So \( P_n(1) = 1 \) for all integers \( n \geq 0 \), which proves the lemma.

As we seek to approximate real numbers arbitrarily well by rational numbers, we require an asymptotic estimate for the growth of the Legendre polynomials. To that end we record the following formula.

**Lemma 2.15** (Szegő [38], 1939). Let \( x > 1 \) be a real number. For any constant \( C > 1 \),
\[
P_n(x) < C(2\pi n)^{-\frac{1}{2}}(x^2 - 1)^{-\frac{1}{4}} \left[ x + (x^2 - 1)^{\frac{1}{2}} \right]^{n+\frac{1}{2}},
\]
for all sufficiently large \( n \).
The final ingredient necessary for the proof of Theorem 2.11 is a collection of Padé approximants from which we construct the required sequence of rational numbers.

**Proposition 2.16.** Let \( f(x) = -\frac{1}{x} \log(1 - x) \) and fix positive integers \( L \) and \( M \) with \( L \geq M - 1 \). The Padé approximant to \( f(x) \) of order \( \lfloor L/M \rfloor \) exists and is given by

\[
[L/M]f(x) = \frac{\sum_{j=0}^{M} (-1)^j \binom{M}{j} \left( \frac{L + j + 1}{M} \right) x^{M-j} \sum_{i=0}^{L-M+j} \frac{x^i}{i+1}}{\sum_{j=0}^{M} (-1)^j \binom{M}{j} \left( \frac{L + j + 1}{M} \right) x^{M-j}}.
\]

**Proof.** Fix positive integers \( L \geq M - 1 \) and consider the function

\[
f(x) = -\frac{1}{x} \log(1 - x) = \sum_{n=0}^{\infty} \frac{x^n}{n+1}.
\]

Define the polynomial \( Q(x) \) as follows

\[
Q(x) = \begin{vmatrix}
  c_{L-M+1} & \cdots & c_L & c_{L+1} \\
  \vdots & \ddots & \vdots & \vdots \\
  c_L & \cdots & c_{L+M-1} & c_{L+M} \\
  x^M & \cdots & x & 1
\end{vmatrix} = \begin{vmatrix}
  \frac{1}{L-M+2} & \cdots & \frac{1}{L+1} & \frac{1}{L+2} \\
  \vdots & \ddots & \vdots & \vdots \\
  \frac{1}{L+1} & \cdots & \frac{1}{L+M} & \frac{1}{L+M+1} \\
  x^M & \cdots & x & 1
\end{vmatrix},
\]

and observe that \( Q(0) = \mathcal{H}_M^{(L-M+1)} \neq 0 \) by Proposition 2.5. Thus by Lemma 2.9 the Padé approximant to \( f(x) \) of order \( [L/M] \) exists and is given as in the statement of Lemma 2.9. Expanding along the last row of \( Q(x) \) gives

\[
Q(x) = \begin{vmatrix}
  \frac{1}{L-M+2} & \cdots & \frac{1}{L+1} & \frac{1}{L+2} \\
  \vdots & \ddots & \vdots & \vdots \\
  \frac{1}{L+1} & \cdots & \frac{1}{L+M} & \frac{1}{L+M+1} \\
  x^M & \cdots & x & 1
\end{vmatrix} = \sum_{j=0}^{M} (-1)^{M-j} \left( \mathcal{H}_M^{(L-M+1)} \right)_M^{j+1} x^{M-j}.
\]
Using Proposition 2.6 we simplify \( Q(x) \) to obtain
\[
Q(x) = \sum_{j=0}^{M} (-1)^{M-j} \cdot \mathcal{H}_{M+1}^{(L-M+1)} \cdot (L + j + 2)
\]
\[
\cdot \left( \frac{L + M + 2}{M-j} \right) \left( \begin{array}{c}
M + 2 \\
0
\end{array} \right) \left( \begin{array}{c}
L + j + 2 \\
M
\end{array} \right) \left( \begin{array}{c}
L + j + 1 \\
j
\end{array} \right) x^{M-j}
\]
\[
= \sum_{j=0}^{M} (-1)^{M+j} \cdot \mathcal{H}_{M+1}^{(L-M+1)} \cdot \frac{(L + M + 2)! (L + j + 1)!}{(M-j)! M! (L - M + j + 1)! j! (L+1)!} \cdot x^{M-j}.
\]
Note that
\[
\frac{(L + j + 1)!}{j! (M-j)! (L - M + j + 1)!} = \binom{M}{j} \binom{L + j + 1}{M} \in \mathbb{N},
\]
and so
\[
Q(x) = (-1)^{M} \cdot \frac{(L + M + 2)!}{M! (L + 1)!} \cdot \mathcal{H}_{M+1}^{(L-M+1)} \cdot \sum_{j=0}^{M} (-1)^{j} \binom{M}{j} \frac{(L + j + 1)!}{M} x^{M-j}.
\]
Next we define the polynomial \( P(x) \) according to Lemma 2.9 as follows
\[
P(x) = \left| \begin{array}{cccc}
\frac{1}{L-M+2} & \cdots & \frac{1}{L+1} & \frac{1}{L+2} \\
\vdots & \ddots & \vdots & \vdots \\
\frac{1}{L+1} & \cdots & \frac{1}{L+M} & \frac{1}{L+M+1} \\
\sum_{i=0}^{L-M} x^{M+i} & \cdots & \sum_{i=0}^{L-1} x^{i+1} & \sum_{i=0}^{L} x^{i}
\end{array} \right|
\]
\[
= \sum_{j=0}^{M} \left( (-1)^{M-j} \mathcal{H}_{M+1}^{(L-M+1)} \frac{L-M+j}{(M+1)} \sum_{i=0}^{L-M+j} \frac{x^{M+i-j}}{i+1} \right).
\]
Again using Proposition 2.6 we simplify \( P(x) \) to obtain
\[
P(x) = \sum_{j=0}^{M} (-1)^{M-j} \cdot \mathcal{H}_{M+1}^{(L-M+1)} \cdot (L + j + 2) \cdot \left( \frac{L + M + 2}{M-j} \right) \left( \begin{array}{c}
L + j + 2 \\
0
\end{array} \right) \left( \begin{array}{c}
L + j + 1 \\
M
\end{array} \right) \left( \begin{array}{c}
L + j + 1 \\
j
\end{array} \right) \cdot \sum_{i=0}^{L-M+j} \frac{x^{M+i-j}}{i+1}
\]
\[
= \sum_{j=0}^{M} (-1)^{M+j} \cdot \mathcal{H}_{M+1}^{(L-M+1)} \cdot \frac{(L + M + 2)! (L + j + 1)!}{(M-j)! M! (L - M + j + 1)! j! (L+1)!} \cdot \sum_{i=0}^{L-M+j} \frac{x^{i}}{i+1}
\]
\[
= (-1)^{M} \cdot \frac{(L + M + 2)!}{M! (L + 1)!} \cdot \mathcal{H}_{M+1}^{(L-M+1)}
\]
2.3. APPLICATIONS TOWARD IRRATIONALITY

\[
\sum_{j=0}^{M} \left( (-1)^j \binom{M}{j} \left( \frac{L + j + 1}{M} \right) x^{M-j} \cdot \sum_{i=0}^{L-M+j} \frac{x^i}{i+1} \right).
\]

Hence, by Lemma 2.9

\[
\frac{[L/M]f(x)}{Q(x)} = \frac{\sum_{j=0}^{M} \left( (-1)^j \binom{M}{j} \left( \frac{L + j + 1}{M} \right) x^{M-j} \cdot \sum_{i=0}^{L-M+j} \frac{x^i}{i+1} \right)}{\sum_{j=0}^{M} (-1)^j \binom{M}{j} \left( \frac{L + j + 1}{M} \right) x^{M-j}},
\]

proving the proposition.

\[\square\]

**Proof of Theorem 2.11.** Fix an integer \( b \geq 11 \) and consider the function

\[f(x) = -\frac{1}{x} \log(1-x) = \sum_{n=0}^{\infty} c_n x^n, \quad c_n = \frac{1}{n+1}.\]

Observe that

\[H_n^{(0)}[c] = \begin{vmatrix} c_0 & \cdots & c_{n-1} \\ \vdots & \ddots & \vdots \\ c_{n-1} & \cdots & c_{2n-2} \end{vmatrix} = \begin{vmatrix} 1 & \cdots & \frac{1}{n} \\ \vdots & \ddots & \vdots \\ \frac{1}{n} & \cdots & \frac{1}{2n-1} \end{vmatrix} = \mathcal{H}_n \neq 0,
\]

by Proposition 2.4. Then, by Proposition 2.10 we have that the Padé approximant to \( f(x) \) of order \( \left[ \frac{(k-1)}{k} \right] \) exists and satisfies

\[f(x) - \left[ \frac{(k-1)}{k} \right] f(x) = \frac{\mathcal{H}_{k+1} x^{2k}}{\mathcal{H}_k} + \mathcal{O}(x^{2k+1}), \tag{2.5}\]

for all integers \( k \geq 1 \). We need to bound the coefficients on the right of Equation (2.5) to control the convergence of the Padé approximants. By Proposition 2.16 we have \( \left[ \frac{(k-1)}{k} \right] f(x) \) is given by

\[
\sum_{j=0}^{k} \left( (-1)^j \binom{k}{j} \left( \frac{k+j}{k} \right) x^{k-j} \sum_{i=0}^{j-1} \frac{x^i}{i+1} \right).
\]

Now, \( \left[ \frac{(k-1)}{k} \right] f(x) \) is a rational function and so we write \( \hat{f}(x) = \sum_{n=0}^{\infty} \hat{c}_n x^n \) for its power series expansion. Lemma 2.14 ensures that \( \left[ \frac{(k-1)}{k} \right] f(1) \) is finite, and so \( \hat{f}(1) = \sum_{n=0}^{\infty} \hat{c}_n \) is finite. In particular, this means \( \{\hat{c}_n\}_{n \geq 0} \) is a bounded sequence, and hence we have

\[0 < \left| f(x) - \left[ \frac{(k-1)}{k} \right] f(x) \right| < C_1 x^{2k} (1 + x^2 + \cdots) = \frac{C_1 x^{2k}}{1-x}, \]

for all integers \( k \geq 1 \).
for some constant $C_1 > 0$. Set $x = \frac{1}{b}$ to obtain

$$0 < \left| f\left(\frac{1}{b}\right) - \left[(k - 1)/k\right] f\left(\frac{1}{b}\right) \right| < \frac{C_1}{b^{2k-1}(b - 1)} ,$$

(2.6)

with

$$[(k - 1)/k] f\left(\frac{1}{b}\right) = \sum_{j=0}^{k} \binom{k}{j} \binom{k + j}{k} \left(\frac{1}{b^{k-j}} \sum_{i=0}^{j-1} \frac{1}{i+1} b^{i} \right)$$

Let us define $D_k = \text{lcm}\{1, 2, \ldots, k\}$ and construct two integer sequences $\{p_k\}_{k \geq 1}$ and $\{q_k\}_{k \geq 1}$ given by

$$p_k = (-1)^k D_k \sum_{j=0}^{k} \binom{k}{j} \binom{k + j}{k} \sum_{i=0}^{j-1} \frac{b^{j-i}}{i+1} \in \mathbb{Z},$$

and

$$q_k = (-1)^k D_k \sum_{j=0}^{k} \binom{k}{j} \binom{k + j}{k} b^j \in \mathbb{Z},$$

where $[(k - 1)/k] f\left(\frac{1}{b}\right) = \frac{p_k}{q_k}$. By Definition 2.13 we write

$$q_k = (-1)^k D_k \sum_{j=0}^{k} \binom{k}{j} \binom{k + j}{k} (-b)^j = D_k \tilde{P}_k(b) = D_k P_k(2b - 1) > 0,$$

and therefore by Lemma 2.15 for sufficiently large $k$ we have

$$q_k < C_2 e^{k} \frac{\left( (2b - 1) + \sqrt{(2b - 1)^2 - 1} \right)^{k+\frac{1}{2}}}{\sqrt{2\pi k} \sqrt{(2b - 1)^2 - 1}} < C_2 e^{k} (2(2b - 1))^{k+1} < 4^{k+1} e^{k} b^{k+1} C_2 ,$$

(2.7)

for some constant $C_2 > 1$. Returning to Equation (2.6), it follows that

$$0 < \left| f\left(\frac{1}{b}\right) - \frac{p_k}{q_k} \right| < \frac{C_1}{b^{2k-1}(b - 1)}.$$

Multiplying through by $q_k$ and using the bound from Equation (2.7) gives

$$0 < \left| q_k f\left(\frac{1}{b}\right) - p_k \right| < \frac{q_k C_1}{b^{2k-1}(b - 1)} < \frac{4^{k+1} e^{k} b^{k+1} C_3}{b^{2k-1}(b - 1)} = \frac{4b^2 C_3}{b - 1} \left(\frac{4e}{b}\right)^k.$$
Hence \( q_k f \left( \frac{1}{b} \right) - p_k \neq 0 \) for all integers \( k \geq 1 \), and

\[
\lim_{k \to \infty} \left| q_k f \left( \frac{1}{b} \right) - p_k \right| \leq \frac{4b^2 C_3}{b-1} \lim_{k \to \infty} \left( \frac{4e}{b} \right)^k \leq \frac{4b^2 C_3}{b-1} \lim_{k \to \infty} \left( \frac{4e}{11} \right)^k = 0. \tag{2.8}
\]

Thus by Lemma 2.12, \( f \left( \frac{1}{b} \right) = b \log \left( \frac{b}{b-1} \right) \) is irrational for integers \( b \geq 11 \), and so \( \log \left( \frac{b+1}{b} \right) \) is irrational for integers \( b \geq 10 \). \( \square \)

The success of this method lies in the ability to generate good rational approximations to the natural logarithm while maintaining moderately large denominators.

The usefulness of the Hankel determinants in this case arises from the fact that these determinants decrease to zero sufficiently fast and avoid being overwhelmed by the denominators of the Padé approximants. Following on from the moderate success of this result for the natural logarithm, we may well ask what other functions could be well approximated in this way? Rather than again attempting to demonstrate the irrationality of a known irrational (even transcendental) number, let us now investigate a constant whose (ir)rationality remains unknown.

Consider instead the function \( f(x) = \sum_{n=0}^{\infty} \frac{x^n}{(2n+1)^2} \) for the purpose of investigating the real number

\[
G = \frac{1}{1^2} - \frac{1}{3^2} + \frac{1}{5^2} - \frac{1}{7^2} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^2} = f(-1),
\]

known as Catalan’s constant. We proceed in the same way as before but with less rigour (as our investigation is ultimately fruitless). We first need to investigate the determinants

\[
H_n^{(0)}[c] = \begin{vmatrix}
\frac{1}{1^2} & \cdots & \frac{1}{(2n-1)^2} \\
\vdots & \ddots & \vdots \\
\frac{1}{(2n-1)^2} & \cdots & \frac{1}{(4n-3)^2}
\end{vmatrix}.
\]

Figure 2.6 shows these determinants for increasing \( n \), and so it seems reasonable to proceed as though \( H_n^{(0)}[c] \neq 0 \) (in fact greater than zero). This allows us to invoke Proposition 2.10 to conclude the Padé approximant to \( f(x) \) of order \( \lfloor (k-1)/k \rfloor \) exists and satisfies

\[
f(x) - \lfloor (k-1)/k \rfloor f(x) = \frac{\mathcal{H}_{k+1}}{\mathcal{H}_k} x^{2k} + \mathcal{O} \left( x^{2k+1} \right),
\]

for all integers \( k \geq 1 \).
Our next point of interest is the coefficients of the terms on the right-hand side, and so for this we again proceed experimentally to compute some coefficients for small values of $k$. The results as displayed in Figure 2.7 suggest that the coefficients can be bounded above by the equivalent coefficients (starting from the $x^{2k}$ term) obtained from taking the Padé approximant of order $[0/1]$. That is, because $[0/1]f(x) = \frac{1}{1 - \frac{1}{9}x}$ we write

$$\left| f(x) - \frac{(k-1)/k}{f(k/x)} \right| \leq \sum_{n=2k}^{\infty} \left( \frac{1}{(2n+1)^2} - \frac{1}{9^n} \right) x^n,$$

and therefore obtain

$$\left| f(-1) - \frac{(k-1)/k}{f(-1)} \right| \leq \sum_{n=2k}^{\infty} \left( \frac{1}{(2n+1)^2} - \frac{1}{9^n} \right). \tag{2.9}$$

Using Maple we plot the right-hand side of Equation (2.9), illustrated in Figure 2.8, from which we conclude

$$\left| f(-1) - \frac{(k-1)/k}{f(-1)} \right| \leq \frac{1}{8k},$$

for all integers $k \geq 1$. 
Figure 2.7. The logarithm of the coefficients of $x^n$ (for $0 \leq n \leq 200$) in the power series obtained by subtracting from $f(x)$ the Padé approximant of order $[(k-1)/k]$ for $1 \leq k \leq 16$. The bold curve shows the approximation of order $[0/1]$.

Figure 2.8. The close approximation of the reciprocal of the right-hand side of Equation (2.9) for $1 \leq k \leq 1000$ (solid), by $8k$ (dashed and slightly offset for ease of display).
The difficulty here is that \[ ((k - 1)/k) f(-1) \] is not a very good rational approximation to \( f(-1) \), as the term on the right-hand side would be easily overwhelmed once denominators are cleared. Figure 2.9 suggests the denominator of \[ ((k - 1)/k) f(-1) \] is bounded by \( e^{O(k^2)} \), eliminating any hope of even coming close to proving the irrationality of Catalan’s constant using this method.

**Figure 2.9.** The bounding above of the logarithm of the denominator of \[ ((k - 1)/k) f(-1) \] for \( 1 \leq k \leq 30 \) (solid), by \( f(k) = 3k^2 \) (dashed).
CHAPTER 3

Hankel Determinants of Convergent Sequences

We studied determinants of specific sequences in the first two chapters, zeta values in Chapter 1 and unit fractions in Chapter 2. We now make a more general investigation into Hankel determinants by considering arbitrary sequences.

Section 3.1 begins with an introduction to Dodgson condensation, another method of determinant computation. This allows us to produce Lemma 3.3, a recurrence relation on Hankel determinants that proves valuable for the remainder of our work. In Section 3.2 we use this recurrence relation to bound positive Hankel determinants, from which we can demonstrate that the determinants $H^{(r)}_n[h]$ strictly decrease to zero as $r \to \infty$. In an attempt to better understand the scope of this result, we conclude this chapter with some comments on the positivity of Hankel determinants in Section 3.3.

3.1. Hankel in Wonderland

In Appendix A we describe a number of methods which can be used to calculate the determinant of a matrix, however there remains a lesser known method that is central to our remaining work. To motivate this approach we require a matrix identity that was proven by Jacobi [21] but which existed in a number of special cases that were previously derived. The case we require was originally due to Desnanot in 1819 who proved it for some small values of $n$, (see [28]), and is now known as the Desnanot-Jacobi identity. We present here a proof based on the treatment given by Bressoud [9]; for an interesting combinatorial proof see Zeilberger [43].

Proposition 3.1 (Jacobi [21], 1833). Let $M$ be an $n \times n$ matrix and $A$ an $m \times m$ minor of $M$, where $m < n$. Write $A'$ for the corresponding minor of $\text{adj}(M)$, and $A^*$ for the complementary $(n - m) \times (n - m)$ minor of $M$. Then

$$\det A' = (\det M)^{m-1} \cdot \det A^*.$$
Proposition 3.2 (Desnanot-Jacobi, 1833). Let $M$ be an $n \times n$ matrix and write $M_i^j$ for the matrix obtained from $M$ by removing the $i$th row and $j$th column (with multiple indices if multiple rows/columns are removed). Then

$$\det(M_1^1) \cdot \det(M_n^n) - \det(M_n^1) \cdot \det(M_1^n) = \det(M) \cdot \det(M_1^{1,n}) \cdot \det(M_n^{n,1}).$$

Proof. Write $M = (m_{i,j})_{1 \leq i,j \leq n}$ and recall that the adjugate matrix of $M$ is given by $\text{adj}(M) = (a_{i,j})_{1 \leq i,j \leq n}$ where $a_{i,j} = (-1)^{i+j} \det(M_i^j)$. Construct another matrix, $N$, given by

$$N = \begin{pmatrix}
a_{1,1} & 0 & \cdots & 0 & a_{1,n} \\
a_{2,1} & 1 & \cdots & 0 & a_{2,n} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
a_{n-1,1} & 0 & \cdots & 1 & a_{n-1,n} \\
a_{n,1} & 0 & \cdots & 0 & a_{n,n}
\end{pmatrix}.$$

That is, the first and last column of $N$ are those of $\text{adj}(M)$, and $N_1^{1,n} = I_{n-2}$ (the $(n-2) \times (n-2)$ identity matrix). Now consider the matrix product $MN$, and recall that the definition of the adjugate matrix is such that $M \text{adj}(M) = \det(M) I_n$. Thus, we have

$$MN = \begin{pmatrix}
\det(M) & m_{1,2} & \cdots & m_{1,n-1} & 0 \\
0 & m_{2,2} & \cdots & m_{2,n-1} & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & m_{n-1,2} & \cdots & m_{n-1,n-1} & 0 \\
0 & m_{n,2} & \cdots & m_{n,n-1} & \det(M)
\end{pmatrix},$$

and so $\det(MN) = \det(M)^2 \det(M_1^{1,n})$. However, we also have

$$\det(MN) = \det(M) \det(N)$$

$$= \det(M) \cdot (a_{1,1} a_{n,n} - a_{n,1} a_{1,n})$$

$$= \det(M) \cdot (\det(M_1^1) \det(M_n^n) - \det(M_n^1) \det(M_1^n)).$$

The identity follows by equating the two expressions and cancelling the common factor $\det(M)$ (which is permissible when the identity is viewed as a correspondence of two polynomials in the $n^2$ variables $m_{i,j}$). \qed
Example 3.1.1. Using Proposition 3.2 we reduce the determinant of an \( n \times n \) matrix down to four determinants of size \((n-1) \times (n-1)\) and one of size \((n-2) \times (n-2)\). This is a substantial computational saving for large matrices when compared to the method of Laplace expansion, which reduces an \( n \times n \) determinant down to \( n \) separate \((n-1) \times (n-1)\) determinants. To demonstrate, let us find the determinant of a 3 \( \times \) 3 matrix:

\[
\begin{vmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9 \\
\end{vmatrix}
\]

\[
\begin{vmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9 \\
\end{vmatrix}
\]

\[
\Rightarrow
\begin{vmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9 \\
\end{vmatrix}
\]

\[
= (-3) \cdot (-3) - (-3) \cdot (-3) = 0. \quad \Box
\]

In 1866, Charles Lutwidge Dodgson\textsuperscript{1} used Proposition 3.2 recursively to create the method known as Dodgson condensation \cite{16}. His method involves “condensing” a given matrix into a smaller matrix whose entries are \( 2 \times 2 \) minors of the previous matrix, with some additional division. The procedure ultimately terminates when a 1 \( \times \) 1 matrix is produced, whose single entry is the determinant of the original matrix.

Let us write \( M_n \) for an \( n \times n \) matrix \((n \geq 3)\) whose entries are given by \( m_{i,j}^{(n)} \) for \( 1 \leq i, j \leq n \). In more precise terms, Dodgson condensation is the process of calculating the determinant of \( M_n \) as follows:

1. Construct the \((n-1) \times (n-1)\) matrix \( M_{n-1} \) with entries \( m_{i,j}^{(n-1)} \) given by

\[
m_{i,j}^{(n-1)} = \begin{vmatrix}
m_{i,j}^{(n)} & m_{i,j+1}^{(n)} \\
m_{i+1,j}^{(n)} & m_{i+1,j+1}^{(n)} \\
\end{vmatrix},
\]

and set the index \( k = n - 2 \).

2. Construct the \( k \times k \) matrix \( \sim M_k \) with entries \( \sim m_{i,j}^{(k)} \) given by

\[
\sim m_{i,j}^{(k)} = \begin{vmatrix}
m_{i,j}^{(k+1)} & m_{i,j+1}^{(k+1)} \\
m_{i+1,j}^{(k+1)} & m_{i+1,j+1}^{(k+1)} \\
\end{vmatrix}.
\]

\textsuperscript{1}Who is perhaps better known by his pseudonym Lewis Carroll.
(3) Construct the $k \times k$ matrix $M_k$ with entries $m_{i,j}^{(k)}$ given by

$$m_{i,j}^{(k)} = \frac{\sim m_{i,j}^{(k)}}{m_{i+1,j+1}^{(k+2)}}.$$

(4) If $k > 1$ then decrease $k$ by one and repeat steps 2 and 3. If $k = 1$, then

$$\det(M_n) = \det(M_1) = m_{1,1}^{(1)}.$$ 

This complicated method is (perhaps) best illustrated through example.

**Example 3.1.2.** Let us return to the $3 \times 3$ matrix from Example 3.1.1 which we denote $M_3$. By “condensing” $M_3$ we obtain the $2 \times 2$ matrix $M_2$

$$M_2 = \begin{pmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
4 & 5 & 6 \\
7 & 8 & 9
\end{pmatrix} = \begin{pmatrix}
-3 & -3 \\
-3 & -3
\end{pmatrix}.$$

Condensing further from $M_2$ produces $\sim M_1$:

$$\sim M_1 = \begin{pmatrix}
-3 & -3 \\
-3 & -3
\end{pmatrix} = (0),$$

from which we finally obtain $M_1$ by dividing the entries of $\sim M_1$ by the corresponding “interior” entries of $M_3$ (those that are obtained by deleting the first and last row and column of $M_3$)

$$M_1 = \begin{pmatrix}
0 \\
0
\end{pmatrix} = (0).$$

Therefore, the determinant of $M_3$ is 0. A close comparison of this method and the previous example makes clear that this condensation method is indeed simply a reworking of Proposition 3.2.

**Remark 3.1.1.** A keen observer may notice that the method of Dodgson condensation is doomed to fail whenever a 0 entry appears within the interior of any of the matrices $M_i$ for $i \geq 3$, due to the division required in step 3. For example, consider
the following condensation.

\[
M_4 = \begin{pmatrix}
1 & 2 & 3 & 4 \\
1 & 4 & 2 & 3 \\
4 & 2 & 1 & 3 \\
4 & 3 & 2 & 1
\end{pmatrix} \quad \Rightarrow \quad M_3 = \begin{pmatrix}
2 & -8 & 1 \\
-14 & 0 & 3 \\
4 & 1 & -5
\end{pmatrix}
\]

\[
\Rightarrow \tilde{M}_2 = \begin{pmatrix}
-112 & -24 \\
-14 & -3
\end{pmatrix} \quad \Rightarrow \quad M_2 = \begin{pmatrix}
-28 & -12 \\
-7 & -3
\end{pmatrix}
\]

\[
\Rightarrow \tilde{M}_1 = (0) \quad \Rightarrow \quad M_1 = \begin{pmatrix}
0 \\
0
\end{pmatrix}
\]

One way to avoid this problem would be to make an adjustment to the starting matrix \(M_4\) in such a way that its determinant remains unchanged (or at least changed in a known way). This could be done by exchanging rows or columns to ensure that the offending zero appears in either the first or the last row or column. Alternatively, adding a multiple of one row or column to another can prevent the zero from appearing in the first place. The drawback to these approaches is that we need to restart the condensation process with a new matrix. Fortunately, if we make only minor changes then we can reuse a number of the previous condensation calculations.

Further reading into the history and applications of Dodgson’s condensation method for evaluating determinants can be found in \([1]\) and \([32]\). Our interest in Dodgson condensation is its application toward evaluating Hankel determinants. Specifically, if we rewrite Proposition 3.2 (which, as the backbone of Dodgson condensation, is sometimes referred to as the “Lewis Carroll identity”) in the notation of Hankel determinants, then we have access to the following recurrence relation.

**Lemma 3.3.** Let \(n \geq 2\) and \(r \geq 0\) be integers and \(\{h(k)\}_{k \geq 2}\) be a sequence of real numbers. Then

\[
H_{n+1}^{(r)}[h] \cdot H_{n-1}^{(r+2)}[h] = H_n^{(r)}[h] \cdot H_n^{(r+2)}[h] - \left( H_n^{(r+1)}[h] \right)^2.
\]

As it was for Dodgson, Lemma 3.3 is the backbone of our remaining work on determinant evaluations. If the determinants we consider are non-zero (and they
3. HANKEL DETERMINANTS OF CONVERGENT SEQUENCES

are) then rearranging allows us to calculate any determinant from determinants of smaller size \( n \) or of lesser offset \( r \). We note that Pólya [30, p.97] provides a useful arrangement in which to organise these determinants, as demonstrated in Figure 3.1, thereby giving a certain visual appeal to the application of Lemma 3.3 that allows it to be used almost effortlessly.

\[
\begin{array}{cccccc}
H_1^{(0)}[h] & H_1^{(1)}[h] & H_1^{(2)}[h] & H_1^{(3)}[h] & \ldots & * & * & * \\
H_2^{(0)}[h] & H_2^{(1)}[h] & H_2^{(2)}[h] & \ldots & * & * & * \\
H_3^{(0)}[h] & * & * & H_3^{(r+2)}[h] & * & \cdots & \cdots & H_n^{(r)}[h] & H_n^{(r+1)}[h] & H_n^{(r+2)}[h] & * & \cdots & H_{n+1}^{(r)}[h] & * & * & * \\
\end{array}
\]

Figure 3.1. A triangular array displaying the five “cross-shaped” determinants that featured in Lemma 3.3.

Many consequences of Lemma 3.3 are evident from Figure 3.1 and are given in [30, pp.97–98]. For instance, any right angle opening upwards and bisected by the vertical contains only and all of the minors of the determinant standing at the vertex of the angle. We offer an additional result regarding the positivity of these determinants when viewed in the above arrangement.

**Lemma 3.4.** If \( H_n^{(R)}[h] > 0 \) and \( H_n^{(R+1)}[h] > 0 \) for all integers \( n \geq 1 \) and for some integer \( R \geq 0 \), then \( H_n^{(r)}[h] > 0 \) for all integers \( n \geq 1 \) and \( r \geq R \).

**Proof.** Define an ordering on the determinants in Figure 3.1 beginning with the upper-left and progressing down each column from left to right. This produces the sequence

\[
\{H_1^{(0)}[h], H_1^{(1)}[h], H_2^{(0)}[h], H_1^{(2)}[h], H_2^{(1)}[h], H_3^{(0)}[h], H_1^{(3)}[h], \ldots\}
\]

which contains the subsequence

\[
\{H_1^{(R)}[h], H_1^{(R+1)}[h], H_2^{(R)}[h], H_1^{(R+2)}[h], H_2^{(R+1)}[h], H_3^{(R)}[h], \ldots\}.
\]

Set \( H_n^{(R)}[h] > 0 \) and \( H_n^{(R+1)}[h] > 0 \) for all integers \( n \geq 1 \), and suppose that there exists some integers \( \bar{n} \geq 1 \) and \( \bar{r} \geq R + 2 \) such that \( H_{\bar{n}}^{(\bar{r})}[h] \leq 0 \) and \( H_n^{(r)}[h] > 0 \)
for all \( n \geq 1 \) and \( r \geq R \) where either \( n + r < \hat{n} + \hat{r} \) or both \( n + r = \hat{n} + \hat{r} \) and \( n < \hat{n} \) (that is, \( H_\hat{r}^{(\hat{r})}[h] \) is the first term in the above subsequence that is not strictly positive). In particular, this means that

\[
H_{\hat{n}+1}^{(\hat{r}−2)}[h] > 0, \quad H_{\hat{n}}^{(\hat{r}−1)}[h] > 0, \quad H_{\hat{n}}^{(\hat{r})}[h] > 0, \quad H_{\hat{n}−1}^{(\hat{r})}[h] > 0.
\]

Now by Lemma 3.3 we have

\[
0 < H_{\hat{n}+1}^{(\hat{r}−2)}[h] \cdot H_{\hat{n}−1}^{(\hat{r})}[h] = H_{\hat{n}}^{(\hat{r}−2)}[h] \cdot H_{\hat{n}}^{(\hat{r})}[h] − \left( H_{\hat{n}}^{(\hat{r}−1)}[h] \right)^2 < 0,
\]

and so there cannot exist such a matrix with non-positive determinant. Note that if we follow the convention of a \( 0 \times 0 \) matrix having determinant 1 then the above calculation still holds when \( \hat{n} = 1 \). \( \square \)

### 3.2. Positive Hankel determinants

The recurrence relation in Lemma 3.3 forces Hankel determinants to behave in a very predictable way. In this section we seek to investigate the effects of increasing the offset \( r \) of a Hankel determinant while keeping its size \( n \) fixed. This can be visualised as progressing to the right along any given row of the arrangement in Figure 3.1. To this end, we prove the following theorem.

**Theorem 3.5.** Let \( \{ h(k) \} \) be a convergent sequence of positive real numbers. If \( H_n^{(r)}[h] > 0 \) for all \( n \geq 2 \) and \( r \geq R \), for some integer \( R \geq 0 \), then

\[
\lim_{r \to \infty} H_n^{(r)}[h] = 0,
\]
for all integers \( n \geq 2 \), where the limit is achieved strictly monotonically from above.

The assumption that the determinants are strictly positive ultimately does much of the work in proving Theorem 3.5. Once we know that the determinants are non-zero we can rearrange Lemma 3.3 so as to cancel terms, and once we know that the determinants are not negative we can omit some positive terms to obtain useful inequalities. The culmination of this is the following proposition, which provides a way to bound Hankel determinants when their positivity is known, and which is the main ingredient in the proof of Theorem 3.5.

**Proposition 3.6.** Let \( R \geq 0 \) be a fixed integer. If \( H_n^{(r)}[h] > 0 \) for all \( n \geq 1 \) and for \( r \geq R \), then

\[
H_n^{(r)}[h] < h(2 + r) \cdot \prod_{k=0}^{n−2} \frac{H_2^{(r+2k)}[h]}{H_1^{(r+2k)}[h]}.
\]
for all \( n \geq 3 \) and for \( r \geq R \).

**Proof.** Let \( r \geq R \) and suppose that \( n \geq 2 \). By Lemma 3.3 and the positivity of \( H_n^{(r)}[h] \), we have

\[
H_{n+1}^{(r)}[h] = \frac{H_n^{(r)}[h] \cdot H_{n+2}^{(r)}[h] - \left( H_{n+1}^{(r+1)}[h] \right)^2}{H_{n+1}^{(r+2)}[h]} < \frac{H_n^{(r)}[h] \cdot H_{n+2}^{(r)}[h]}{H_{n+1}^{(r+2)}[h]}.
\]

Combining this with the analogous result for \( H_n^{(r+2)}[h] \) gives

\[
H_{n+1}^{(r)}[h] < H_n^{(r)}[h] \cdot \frac{H_{n-1}^{(r+2)}[h]}{H_{n+1}^{(r+2)}[h]} = H_n^{(r)}[h] \cdot \frac{H_{n-1}^{(r+4)}[h]}{H_{n-2}^{(r+4)}[h]}.
\]

We continue this process to keep reducing \( n \) (while increasing \( r \)) to give

\[
H_{n+1}^{(r)}[h] < H_n^{(r)}[h] \cdot \frac{H_2^{(r+2(n-1))}[h]}{H_1^{(r+2(n-1))}[h]}.
\]

Now let \( n \geq 3 \). The previous inequality gives \( H_n^{(r)}[h] < H_{n-1}^{(r)}[h] \cdot \frac{H_2^{(r+2(n-2))}[h]}{H_1^{(r+2(n-2))}[h]} \).

Repeated application of this inequality gives

\[
H_n^{(r)}[h] < H_1^{(r)}[h] \cdot \prod_{k=2}^{n} \frac{H_2^{(r+2(k-2))}[h]}{H_1^{(r+2(k-2))}[h]} = h(2 + r) \cdot \prod_{k=0}^{n-2} \frac{H_2^{(r+2k)}[h]}{H_1^{(r+2k)}[h]},
\]

which is the desired result. \(\square\)

From the positivity of these determinants alone we have been able to bound them above in terms of the positive sequence \( h(k) \) and its \( 2 \times 2 \) determinants. An immediate consequence of the proposition is that if the ratio \( \frac{H_2^{(r)}[h]}{H_1^{(r)}} \) remains small then the determinants must decrease towards zero. To see this, let us revisit the Hankel determinants formed from a sequence of unit fractions (i.e. the Hilbert matrices from Chapter 2).

**Example 3.2.1.** Let us consider the sequence \( \{ h(k) \}_{k \geq 2} \) where \( h(k) = \frac{1}{k-1} \). Then \( H_n^{(0)}[h] \) is the determinant of the \( n \times n \) Hilbert matrix

\[
H_n^{(0)}[h] = \begin{vmatrix}
1 & \frac{1}{2} & \cdots & \frac{1}{n} \\
\frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1}
\end{vmatrix}.
\]
Now the entries $h_{i,j}$ of the determinant $H_n^{(r)}[h]$ can be expressed as

$$h_{i,j} = \frac{1}{i + j + r - 1} = \int_0^1 x^{i+j+r-2} \, dx = \int_0^1 x^{i+\frac{r}{2}-1}x^{j+\frac{r}{2}-1} \, dx,$$

and so $H_n^{(r)}[h]$ can be viewed as a Gram determinant. That is, as the determinant of a matrix whose entries are the collection of inner products of the vectors $\{x^\frac{r}{2}, x^{1+\frac{r}{2}}, \ldots, x^n+\frac{r}{2}-1\}$ on the $L^2$ space of continuous functions over $[0, 1]$. Since these vectors are linearly independent within the vector space, the Gram matrix is non-singular and hence $H_n^{(r)}[h] \neq 0$. Moreover, the Gram determinant of a set of vectors equals the square of the volume of the parallelootope formed from the vectors, and therefore $H_n^{(r)}[h] > 0$ for all integers $n \geq 1$ and $r \geq 0$. Therefore we can apply Proposition 3.6 to obtain the bound

$$H_n^{(r)}[h] < \frac{1}{r+1} \cdot \prod_{k=0}^{n-2} \frac{1}{\frac{r+2k+1}{r+2k+3} - \frac{1}{(r+2k+2)^2}}$$

$$= \frac{1}{r+1} \cdot \prod_{k=0}^{n-2} \frac{1}{(r+2k+2)^2(r+2k+3)}$$

$$= \frac{1}{2^{n-1} \Gamma(r+2) \Gamma\left(\frac{r+2}{2}\right) \Gamma\left(\frac{r+4n}{2}\right)}.$$ 

From this we see that $H_n^{(r)}[h]$ decreases towards zero as $r$ increases (and indeed as $n$ increases, but that does not concern us until Chapter 4), but we would need more to determine that it does so monotonically. In this case the sequence $\{\frac{1}{r+1}, \frac{1}{r+2}, \ldots\}$ converges, so monotonicity is guaranteed by Theorem 3.5. ♦

It is worth noting that we gave a closed form for $H_n^{(r)}[h]$ in Proposition 2.5, whereas here we are only able to give bounds. However, the result in Proposition 2.5 was obtained because the Hilbert matrices had additional structure as Cauchy matrices, whereas the bounds achieved using Proposition 3.6 required no special classes of matrices but only the positivity of the determinants. Ideally a closed form for any Hankel determinant could be calculated from a more precise understanding of how the arithmetic properties of $\{h(k)\}_{k \geq 2}$ affect the determinants, but it is often much more accessible to simply confirm the positivity of such determinants and then assert the appropriate upper bound.

Proposition 3.6 is most useful for the case that $\{h(k)\}_{k \geq 2}$ is a convergent sequence, but it is not confined to this. In Section 4 we are interested in sequences whose terms
are values of positive Dirichlet series, and so \( \{h(k)\}_{k \geq 2} \) is a decreasing sequence. However, we can still apply the proposition to an increasing (and divergent) sequence, as seen in the next example.

**Example 3.2.2.** Let us consider the sequence \( \{h(k)\}_{k \geq 2} \) where \( h(k) = (k - 2)! \). Strehl has shown that \( H_n^{(r)}[h] > 0 \) for all \( n \geq 1 \) and \( r \geq 0 \); in fact, exact values for these determinants are known—see Radoux [31] for details. To gain an upper bound, we apply Proposition 3.6 to obtain

\[
H_n^{(r)}[h] < r! \cdot \prod_{k=0}^{n-2} \frac{(r + 2k)!((r + 2k + 2)!) - ((r + 2k + 1)!)^2}{(r + 2k)!} = r! \cdot \prod_{k=0}^{n-2} (r + 2k + 1)!
\]

Here the determinants increase without bound as \( r \) increases, but we can still bound how quickly they increase using nothing more than their positivity.

**Proof of Theorem 3.5.** Let \( n \geq 2 \) and \( r \geq R \). By Proposition 3.6 we have that

\[
0 < H_n^{(r)}[h] \leq h(2 + r) \cdot \prod_{k=0}^{n-2} \frac{H_2^{(r+2k)}[h]}{H_1^{(r+2k)}[h]},
\]

for all \( n \geq 2 \) and for \( r \geq R \). Now the sequence \( \{h(k)\}_{k \geq 2} \) converges to some limit, \( \ell \), say, where \( \ell \geq 0 \). If \( \ell > 0 \) then we have

\[
\lim_{r \to \infty} \frac{H_2^{(r)}[h]}{H_1^{(r)}[h]} = \lim_{r \to \infty} \left( h(4 + r) \cdot \frac{h(3 + r)}{h(2 + r)} \right) = \ell - \ell \cdot \frac{\ell}{\ell} = 0,
\]

otherwise \( \ell = 0 \) and thus

\[
\lim_{r \to \infty} \frac{H_2^{(r)}[h]}{H_1^{(r)}[h]} \leq \lim_{r \to \infty} h(4 + r) = \ell = 0.
\]

Therefore we have

\[
0 \leq \lim_{r \to \infty} H_n^{(r)}[h] \leq \lim_{r \to \infty} h(2 + r) \cdot \prod_{k=0}^{n-2} \left( \lim_{r \to \infty} \frac{H_2^{(r+2k)}[h]}{H_1^{(r+2k)}[h]} \right) = 0.
\]

Hence \( H_n^{(r)}[h] \to 0 \) as \( r \to \infty \).

We now prove \( \{H_n^{(r)}[h]\}_{r \geq R} \) is a strictly decreasing sequence. By Lemma 3.3 and the positivity of \( H_n^{(r)}[h] \), we have

\[
(H_n^{(r+1)}[h])^2 = H_n^{(r)}[h] \cdot H_n^{(r+2)}[h] - H_{n+1}^{(r)}[h] \cdot H_{n-1}^{(r+2)}[h] < H_n^{(r)}[h] \cdot H_n^{(r+2)}[h].
\]
Since the geometric mean of two positive numbers cannot exceed their arithmetic mean, we have

\[ H_n^{(r+1)}[h] < \sqrt{H_n^{(r)}[h] \cdot H_n^{(r+2)}[h]} \leq \frac{H_n^{(r)}[h] + H_n^{(r+2)}[h]}{2} \]

which can be rearranged to give

\[ H_n^{(r+1)}[h] - H_n^{(r+2)}[h] < H_n^{(r)}[h] - H_n^{(r+1)}[h]. \] (3.1)

We assume, towards a contradiction, that there exist some integers \( \hat{n} \geq 2 \) and \( \hat{r} \geq R \) such that \( H_{\hat{n}}^{(\hat{r}+1)}[h] \geq H_{\hat{n}}^{(\hat{r})}[h] \). Then using (3.1), we have

\[ 0 \geq H_{\hat{n}}^{(\hat{r})}[h] - H_{\hat{n}}^{(\hat{r}+1)}[h] > H_{\hat{n}}^{(\hat{r}+1)}[h] - H_{\hat{n}}^{(\hat{r}+2)}[h], \]

and so \( H_{\hat{n}}^{(\hat{r}+2)}[h] > H_{\hat{n}}^{(\hat{r}+1)}[h] \). Thus we have that \( H_{\hat{n}}^{(r+1)}[h] > H_{\hat{n}}^{(r)}[h] \) for all \( r > \hat{r} \), which contradicts that \( H_{\hat{n}}^{(r)}[h] \to 0 \) as \( r \to \infty \). \( \square \)

**Remark 3.2.1.** We mentioned previously that the assumption of the positivity of the determinants in Theorem 3.5 is essential to its proof. In light of the above proof we see that the consequences of this assumption are far-reaching. For convenience take \( R = 0 \), so that Equation (3.1) implies for fixed integer \( n \geq 1 \) that the sequence \( \{H_n^{(r)}[h]\}_{r \geq 0} \) is strictly convex. This provides the intuition for the contradiction from the above proof: if the determinants were increasing, then their convexity would imply the determinants grew without bound, and so this cannot be the case. In particular, notice \( H_1^{(r)}[h] \) is simply \( h(2 + r) \), and so the convexity of \( \{H_1^{(r)}[h]\}_{r \geq 0} \) is just the convexity of \( \{h(k)\}_{k \geq 2} \). All together we have the positive sequence \( \{h(k)\}_{k \geq 2} \) is strictly convex but also convergent, and so must be strictly decreasing. There is of course an analogous conclusion about \( \{h(k)\}_{k \geq 2} \) in terms of the convexity of \( \{H_2^{(r)}[h]\}_{r \geq 0} \) and other higher-order determinants, and these together supply the hidden restrictions that the sequence \( \{h(k)\}_{k \geq 2} \) is satisfying. \( \Diamond \)

### 3.3. The pursuit of positivity

The observations in Remark 3.2.1 focus on how the positivity of the determinants \( H_n^{(r)}[h] \) impose certain conditions on the behaviour of the underlying sequence \( \{h(k)\}_{k \geq 2} \). This may have been useful in understanding the hypotheses of Theorem 3.5, but the converse may be even more useful. Specifically, it is often hard to
tell whether or not a Hankel determinant is positive in general (as seen in Example 3.2.1 it can be done for specific cases by appealing to other structures that the determinant may possess), and so we wish to draw some conclusions about what possible conditions could be imposed on the sequence \( \{h(k)\}_{k \geq 2} \) to guarantee the positivity of the associated determinants.

Let us start with small matrices and work our way up. Clearly the determinant of a 1 \( \times \) 1 Hankel matrix is positive whenever the sequence \( \{h(k)\}_{k \geq 2} \) is positive.

For the case of a 2 \( \times \) 2 Hankel determinant we obtain

\[
0 < H_2^{(r)}[h] = \begin{vmatrix} h(2 + r) & h(3 + r) \\ h(3 + r) & h(4 + r) \end{vmatrix} = h(2 + r)h(4 + r) - h(3 + r)^2,
\]

and so we have \( \{H_2^{(r)}[h]\}_{r \geq 0} \) is a positive sequence whenever \( \left\{ \frac{h(k+1)}{h(k)} \right\}_{k \geq 2} \) is a strictly increasing sequence, which is slightly stronger than requiring that \( \{h(k)\}_{k \geq 2} \) be a strictly convex sequence. Unfortunately it is not very difficult to calculate any given 2 \( \times \) 2 determinant and see directly whether or not it is positive, so we ought to proceed to higher order matrices. It is clear that this algebraic approach is not going to be of much assistance, for if we consider a 3 \( \times \) 3 Hankel determinant we obtain

\[
0 < H_3^{(r)}[h] = \begin{vmatrix} h(2 + r) & h(3 + r) & h(4 + r) \\ h(3 + r) & h(4 + r) & h(5 + r) \\ h(4 + r) & h(5 + r) & h(6 + r) \end{vmatrix} = h(2 + r)h(4 + r)h(6 + r) + 2h(3 + r)h(4 + r)h(5 + r) - h(2 + r)h(5 + r)^2 - h(3 + r)^2h(6 + r) - h(4 + r)^3,
\]

from which we draw no meaningful conclusions about conditions to impose on the sequence \( \{h(k)\}_{k \geq 2} \). A more robust approach to higher dimensional thinking could be to consider the determinants from a geometric viewpoint rather than an algebraic one. Recall that the determinant of a 2 \( \times \) 2 matrix is the signed area of the parallelogram formed from the column vectors of the matrix. Given that we require \( \{h(k)\}_{k \geq 2} \) to be a positive sequence, we need only work in the positive quadrant of \( \mathbb{R}^2 \). Figure 3.2 demonstrates the two possible cases for the orientation of the two vectors, leading to either a positive or a negative determinant (clearly the determinant is zero if the vectors are parallel).
When viewed from this perspective, it is valid to say that the determinant is positive if a line parallel to the second vector is steeper than a line parallel to the first vector. In the case of $H_2^{(r)}[h]$ the vectors are $\langle h(k), h(k+1) \rangle$ and $\langle h(k+1), h(k+2) \rangle$, and so the conclusion is that $\frac{h(k+1)}{h(k)} > \frac{h(k+1)}{h(k)}$ for all integers $k \geq 2$. That is, $\left\{ \frac{h(k+1)}{h(k)} \right\}_{k \geq 2}$ must be a strictly increasing sequence as before. The idea of “steepness” does not translate very well into higher dimensions, so perhaps a better way to think about these vectors is with regards to the angle they make with the positive $x$-axis. If we write $\theta_2(k)$ for the angle corresponding to $\langle h(k), h(k+1) \rangle$, then we require $\left\{ \theta_2(k) \right\}_{k \geq 2}$ to be an increasing sequence. Now we have

$$\cos \theta_2(k) = \frac{h(k)}{\sqrt{h(k)^2 + h(k+1)^2}} = \frac{1}{\sqrt{1 + \left(\frac{h(k+1)}{h(k)}\right)^2}},$$

and so again we confirm the conclusion that $\left\{ \frac{h(k+1)}{h(k)} \right\}_{k \geq 2}$ must be a strictly increasing sequence. Figure 3.3 shows the sequence of vectors $\left\{ \langle h(k), h(k+1) \rangle \right\}_{k \geq 2}$ associated with the $2 \times 2$ Hilbert matrices, which demonstrates the trend that we are expecting.

This formulation has greater potential to be lifted to higher dimensions since we are always able to examine the angle between two vectors. The sequence of vectors $\left\{ \langle h(k), h(k+1), h(k+2) \rangle \right\}_{k \geq 2}$ associated with the $3 \times 3$ Hilbert matrices can be seen in Figure 3.4. Denote by $\theta_3(k)$ the angle between the vector $\langle h(k), h(k+1), h(k+2) \rangle$ and the positive $x$-axis, and $\phi_3(k)$ for the angle with the positive $y$-axis. In the same
Figure 3.3. The vectors $\langle h(k), h(k+1) \rangle$ for $2 \leq k \leq 10$ where $h(k) = \frac{1}{k-1}$, normalised for appearance. The bold vector is parallel to $\langle 1, 1 \rangle$.

manner as before, we might ask whether the sequences $\{\theta_3(k)\}_{k \geq 2}$ and $\{\phi_3(k)\}_{k \geq 2}$ are increasing or decreasing, and whether this is related to the positivity of the determinants of the $3 \times 3$ Hilbert matrices.

Figure 3.5 suggests that $\{\theta_3(k)\}_{k \geq 2}$ is an increasing sequence, while $\{\phi_3(k)\}_{k \geq 2}$ is a decreasing sequence, both converging to $\arccos \left(\frac{1}{\sqrt{3}}\right)$. Now we have

$$\cos \theta_3(k) = \frac{h(k)}{\sqrt{h(k)^2 + h(k+1)^2 + h(k+2)^2}} = \frac{1}{\sqrt{1 + \left(\frac{h(k+1)}{h(k)}\right)^2 + \left(\frac{h(k+2)}{h(k)}\right)^2}},$$

and

$$\cos \phi_3(k) = \frac{h(k+1)}{\sqrt{h(k)^2 + h(k+1)^2 + h(k+2)^2}} = \frac{1}{\sqrt{\left(\frac{h(k)}{\pi(k+1)}\right)^2 + 1 + \left(\frac{h(k+2)}{\pi(k+1)}\right)^2}}.$$

While examining the $2 \times 2$ determinants we concluded that $\left\{\frac{h(k+1)}{h(k)}\right\}_{k \geq 2}$ was a strictly increasing sequence. Given that $\frac{h(k+2)}{h(k)} = \frac{h(k+2)}{h(k+1)} \frac{h(k+1)}{h(k)}$, is the product of two strictly increasing sequences, we can conclude that $\left\{\frac{h(k+2)}{h(k)}\right\}_{k \geq 2}$ is also strictly increasing, and so $\{\theta_3(k)\}_{k \geq 2}$ is a strictly increasing sequence. For $\{\phi_3(k)\}_{k \geq 2}$ to be
Figure 3.4. The vectors $\langle h(k), h(k + 1), h(k + 2) \rangle$ for $2 \leq k \leq 20$ where $h(k) = \frac{1}{k-1}$, normalised for appearance and shown in two orientations for perspective. The bold vector is parallel to $\langle 1, 1, 1 \rangle$. 
a decreasing sequence we require that
\[
\left\{ \left( \frac{h(k)}{h(k+1)} \right)^2 + \left( \frac{h(k+2)}{h(k+1)} \right)^2 \right\}_{k \geq 2}
\]
be decreasing also.

\[
\begin{array}{c}
\text{Figure 3.5. The angles } \theta_3(k) \text{ (solid) and } \phi_3(k) \text{ (dashed) for } 2 \leq k \leq 50.
\end{array}
\]

We are yet to demonstrate whether any of these conditions resulting from the behaviour of the sequences \( \{\theta_3(k)\}_{k \geq 2} \) and \( \{\phi_3(k)\}_{k \geq 2} \) are connected to the positivity of general Hankel determinants, but are simply observations arising from the determinant of the \( 3 \times 3 \) Hilbert matrices. We shall suspend attempting to establish this connection in light of our next observation. Denote by \( \theta_4(k) \) the angle between the vector \( \langle h(k), h(k+1), h(k+2), h(k+3) \rangle \) and the positive \( x \)-axis, \( \phi_4(k) \) for the angle with the positive \( y \)-axis, and \( \psi_4(k) \) for the angle with the positive \( z \)-axis. Figure 3.6 illustrates that these sequences all converge to \( \frac{\pi}{3} \) as expected, but that \( \{\phi_4(k)\}_{k \geq 2} \) is not a monotonic sequence. Thus our attempts to describe the general behaviour of the angles between these vectors and the co-ordinate axes might not yield any useful information regarding the positivity of Hankel determinants.

It may be possible to make some more concrete observations about the positivity of Hankel determinants from a geometric interpretation of the determinant, but we do not pursue this here; we hope to address this in future work.
Figure 3.6. The angles $\theta_4(k)$ (solid), $\phi_4(k)$ (dashed) and $\psi_4(k)$ (short dashed) for $2 \leq k \leq 50$. 
CHAPTER 4

Hankel Determinants of Dirichlet Series

The previous chapter focussed on developing a recurrence relation which allowed us to give an upper bound on positive Hankel determinants \( H_n^{(r)}[h] \) as \( r \to \infty \). To conclude our work we now turn our attention to investigating the growth of these determinants as \( n \to \infty \), with a particular focus on Hankel determinants of zeta values as studied by Monien [27].

In Section 4.1 we motivate our investigation by presenting the anticipated decay of these Hankel determinants. In the general case, we demonstrate how the growth of a sequence \( \{h(k)\}_{k \geq 2} \) and the ratio of its successive terms can be used to arrive at a useful bound on \( H_n^{(r)}[h] \). We find that Hankel determinants of certain Dirichlet series are always positive, and so in Section 4.2 we make use of our results to bound these determinants and examine some consequences of this bound. We make a short investigation into the elements of Dodgson condensation that were central to our results, and we explain why our bound is the best possible from this method. Lastly we examine in Section 4.3 whether these results can be applied to Dirichlet \( L \)-series and conclude our observations with some conjectures.

4.1. Hankel determinants of some specialised sequences

Recall the opening of Chapter 1 where we introduced the following determinants:

\[
H_1^{(0)}[\zeta] = \zeta(2), \quad H_2^{(0)}[\zeta] = \begin{vmatrix} \zeta(2) & \zeta(3) \\ \zeta(3) & \zeta(4) \end{vmatrix}, \quad H_3^{(0)}[\zeta] = \begin{vmatrix} \zeta(2) & \zeta(3) & \zeta(4) \\ \zeta(3) & \zeta(4) & \zeta(5) \\ \zeta(4) & \zeta(5) & \zeta(6) \end{vmatrix}, \ldots
\]

and

\[
H_1^{(1)}[\zeta] = \zeta(3), \quad H_2^{(1)}[\zeta] = \begin{vmatrix} \zeta(3) & \zeta(4) \\ \zeta(4) & \zeta(5) \end{vmatrix}, \quad H_3^{(1)}[\zeta] = \begin{vmatrix} \zeta(3) & \zeta(4) & \zeta(5) \\ \zeta(4) & \zeta(5) & \zeta(6) \\ \zeta(5) & \zeta(6) & \zeta(7) \end{vmatrix}, \ldots
\]
In an effort to understand the asymptotic decay of these determinants as \( n \to \infty \), Monien produced a heuristic which suggests that
\[
\log H_n^{(0)}[\zeta] \sim \log H_n^{(1)}[\zeta] \sim -n^2 \left( \log(2n) - \frac{3}{2} \right).
\]
Figure 4.1 illustrates the extent to which these three expressions agree as a function of \( n \).

Monien also found, experimentally, that
\[
-H_n^{(0)}[\zeta] H_n^{(1)}[\zeta] = -\frac{1}{2n+1} + \frac{2}{(2n+1)^2} - \frac{7}{3} \frac{1}{(2n+1)^3} + \cdots,
\]
and
\[
-H_{n+1}^{(0)}[\zeta] H_n^{(1)}[\zeta] = -\frac{1}{2n} - \frac{1}{(2n)^2} + \frac{2}{3} \frac{1}{(2n)^3} - \frac{6}{5} \frac{1}{(2n)^4} + \frac{56}{45} \frac{1}{(2n)^5} + \cdots.
\]
Additionally, according to Monien, detailed numerical experiments by Don Zagier suggest that
\[
H_n^{(0)}[\zeta] = A^{(0)} \left( \frac{2n+1}{e\sqrt{e}} \right)^{-(n+\frac{1}{2})^2} \left( 1 + \frac{1}{24} \frac{1}{(2n+1)^2} - \frac{12319}{259200} \frac{1}{(2n+1)^4} + \cdots \right),
\]
and

\[ H_{n-1}^{(1)}[\zeta] = A^{(1)} \left( \frac{2n}{e\sqrt{e}} \right)^{-n^2+\frac{3}{4}} \left( 1 - \frac{17}{240} \frac{1}{(2n)^2} - \frac{199873}{7257600} \frac{1}{(2n)^4} - \cdots \right), \]

where \( A^{(0)} \approx 0.351466738331 \ldots \) and \( A^{(1)} = \frac{e^9/8}{\sqrt{6}} A^{(0)} \).

Our goal is twofold: bound \( \{ H_n^{(r)}[\zeta]\}_{n \geq 1} \) in such a way as to produce this expected asymptotic decay in particular, and bound \( \{ H_n^{(r)}[h]\}_{n \geq 1} \) in general. In fact, we can make substantial progress towards confirming this particular decay with the following result (a special case of Proposition 4.5).

Lemma 4.1 (Monien [27], 2009). The Hankel determinant \( H_n^{(r)}[\zeta] \) can be expressed as

\[ H_n^{(r)}[\zeta] = \frac{1}{n!} \sum_{m_1, \ldots, m_n=1}^{\infty} \frac{1}{(m_1 m_2 \cdots m_n)^{2n+r}} \prod_{i<j} (m_i - m_j)^2, \]

for all integers \( n \geq 1 \) and \( r \geq 0 \).

Since the summations are over nonnegative terms only, we can trivially bound the expression in Lemma 4.1 by a single term from the summations. However, we first need the following lemma, which resembles Stirling’s formula for \( \log n! \).

Lemma 4.2. For all integers \( n \geq 1 \) we have

\[ \log n! < (n + 1) \log(n + 1) - n. \]

Proof. Let us write \( \log n! = \sum_{k=1}^{n} \log k \), and treat this as a left Riemann sum approximation (with rectangles of width 1) for the area underneath \( f(x) = \log x \).

Since \( f(x) \) is concave, we write

\[ \sum_{k=1}^{n} \log k < \int_{1}^{n+1} \log x \, dx = (n + 1) \log(n + 1) - n, \]

which proves the result. \( \square \)

Proposition 4.3. The Hankel determinant \( H_n^{(r)}[\zeta] \) satisfies

\[ \log H_n^{(r)}[\zeta] > -(2n + r + 1)(n + 1) \log(n + 1) + 2n^2, \]

for all integers \( n \geq 1 \) and \( r \geq 0 \). In particular, \( \log H_n^{(r)}[\zeta] > -cn^2 \log n \) for some constant \( c > 0 \) and all sufficiently large \( n \).
Proof. Set the summation indices $m_i = i$ in Lemma 4.1. Then we have
\[
H_n^{(r)}[\varsigma] > \frac{1}{n! n^{2n + r}} \prod_{1 \leq i < j \leq n} (i-j)^2 \geq \frac{1}{n! n^{2n + r + 1}}.
\]
Taking the logarithm of both sides and using Lemma 4.2 gives
\[
\log H_n^{(r)}[\varsigma] > -(2n + r + 1) \log n! > -(2n + r + 1) \left( (n + 1) \log(n + 1) - n \right),
\]
and so the result follows. \(\square\)

To successfully produce an upper bound on the Hankel determinants $H_n^{(r)}[h]$ in the general case, we require bounds on the sequence $\{h(k)\}_{k \geq 2}$ and on the ratio of consecutive terms. To this end, we prove the following proposition.

**Proposition 4.4.** Let $K \geq 2$ be a fixed integer and $h(k) > 0$ for $k \geq K$. Suppose there are positive functions $A(k)$, $B(k)$ and $\lambda(k)$, with $A(k) < B(k + 1)$ for all $k \geq K$, such that

1. $A(k) \leq \frac{h(k+1)}{h(k)} \leq B(k)$ for all $k \geq K$, and
2. $h(k+2) \leq \lambda(k+2) \cdot \frac{B(k+1)}{B(k+1) - A(k)}$, for all $k \geq K$.

If $H_n^{(r)}[h] > 0$ for all $n \geq 1$ and for $r \geq K - 2$, then
\[
H_n^{(r)}[h] < h(2+r) \prod_{k=2}^{n} \lambda(2k+r),
\]
for all $n \geq 3$ and for $r \geq K - 2$.

Proof. Fix $r \geq K - 2$. By definition, $H_2^{(r)}[h] = h(2+r)h(4+r) - h(3+r)^2$ and $H_1^{(r)}[h] = h(2+r)$. Using the positivity of $h(k)$ along with assumptions (i) and (ii), we have
\[
H_2^{(r)}[h] = h(2+r)h(4+r) \left( 1 - \frac{h(3+r)}{h(4+r)} \frac{h(3+r)}{h(2+r)} \right)
\leq h(2+r)h(4+r) \left( 1 - \frac{A(2+r)}{B(3+r)} \right)
\leq h(2+r) \cdot \lambda(4+r) \cdot \frac{B(3+r)}{B(3+r) - A(2+r)} \left( 1 - \frac{A(2+r)}{B(3+r)} \right)
= H_1^{(r)}[h] \cdot \lambda(4+r).
\]
Suppose that $n \geq 2$. By Lemma 3.3 and the positivity of $H_n^{(r)}[h]$, we have
\[
H_n^{(r)}[h] = \frac{H_n^{(r)}[h] \cdot H_n^{(r+2)}[h] - \left( H_n^{(r+1)}[h] \right)^2}{H_n^{(r+2)}[h]} < H_n^{(r)}[h] \cdot \frac{H_n^{(r+2)}[h]}{H^{(r+2)}[h]}. 
\]
Combining this with the analogous result for $H_n^{(r+2)}[h]$ gives

\[ H_{n+1}^{(r)}[h] < H_n^{(r)}[h] \cdot \frac{H_n^{(r+2)}[h]}{H_n^{(r+2)}[h]} = H_n^{(r)}[h] \cdot \frac{H_n^{(r+4)}[h]}{H_n^{(r+4)}[h]} \cdot \frac{H_{n+1}^{(r+2)}[h]}{H_n^{(r+2)}[h]} \cdot \frac{H_{n+1}^{(r+4)}[h]}{H_{n+1}^{(r+4)}[h]} = H_n^{(r)}[h] \cdot \frac{H_n^{(r+2)}[h]}{H_n^{(r+2)}[h]} \cdot \frac{H_{n+1}^{(r+2)}[h]}{H_n^{(r+2)}[h]} \cdot \frac{H_{n+1}^{(r+4)}[h]}{H_{n+1}^{(r+4)}[h]} < H_n^{(r)}[h] \cdot H_n^{(r+2)}[h] \cdot H_{n+1}^{(r+2)}[h] \cdot H_{n+1}^{(r+4)}[h]. \]

We continue this process of reducing $n$ (while increasing $r$) until we can apply the $n = 1$ case to give

\[ H_{n+1}^{(r)}[h] < H_n^{(r)}[h] \cdot H_{n+1}^{(r+2)}[h] < H_n^{(r)}[h] \cdot \frac{H_n^{(r+2)}[h]}{H_n^{(r+2)}[h]} < H_n^{(r)}[h] \cdot H_{n+1}^{(r+2)}[h] \cdot H_{n+1}^{(r+4)}[h] \cdot \frac{H_{n+1}^{(r+4)}[h]}{H_{n+1}^{(r+4)}[h]} < H_n^{(r)}[h] \cdot H_{n+1}^{(r+2)}[h] \cdot H_{n+1}^{(r+4)}[h] \cdot H_{n+1}^{(r+4)}[h]. \]

Now let $n \geq 3$. The previous inequality gives $H_n^{(r)}[h] < H_{n+1}^{(r)}[h] \cdot \lambda(2n + r)$. Repeated application of this inequality gives

\[ H_n^{(r)}[h] < H_{n+1}^{(r)}[h] \cdot \prod_{k=2}^{n} \lambda(2k + r) = h(2 + r) \prod_{k=2}^{n} \lambda(2k + r), \]

which is the desired result. \(\square\)

We note that if additionally $h(k)$ is bounded above for all $k$, say by $M$, then condition (ii) of Proposition 4.4 is satisfied by defining $B(k) = B$ for some constant $B > 0$ and $\lambda(k) = (B - A(k - 2)) \frac{M}{B}$. That is, we have

\[ \lambda(k+2) \cdot \frac{B(k+1)}{B(k+1) - A(k)} = (B - A(k)) \frac{M}{B} \cdot \frac{B}{B - A(k)} = M \geq h(k+2), \]

for all $k \geq K$. To demonstrate this we revisit Example 3.2.1.

**Example 4.1.1.** Recall the sequence \{h(k)\}_{k \geq 2} where $h(k) = \frac{1}{k-1}$, which is bounded above by 1. Fix $K = 2$ and observe that $\frac{h(k+1)}{h(k)} = \frac{k-1}{k} \leq 1$. Hence choose $A(k) = \frac{k-1}{k}$ and set $B(k) = 1$, and note that $A(k) < B(k + 1)$ for all $k \geq K$. Moreover, using our above observation define

\[ \lambda(k) = (1 - A(k - 2)) \frac{1}{1} = 1 - \frac{k-3}{k-2} = \frac{1}{k-2}, \]

from which we confirm

\[ \lambda(k+2) \cdot \frac{B(k+1)}{B(k+1) - A(k)} = \frac{1}{k} \cdot \frac{1}{1 - \frac{k-1}{k}} = 1 \geq h(k+2), \]

for all $k \geq K$. Therefore by Proposition 4.4 we obtain

\[ H_n^{(r)}[h] < \frac{1}{r+1} \cdot \prod_{k=2}^{n} \frac{1}{r+2k-2}. \]

Once more we have shown these determinants to be bounded above and thus must decrease to zero (this time as $n \to \infty$), but here we have seen how the rate at which
the sequence decreases can be used to bound the rate at which the determinant decreases. In this case the bound is not as good as in Example 3.2.1, but often this may be sufficient (as is the case in Section 4.2).

In what follows we are only interested in decreasing sequences \( \{h(k)\}_{k \geq 2} \), as seen in Example 4.1.1 (though Proposition 4.4 holds for increasing functions as well).

**Example 4.1.2.** To show Proposition 4.4 in action for an increasing function, we return to the sequence \( \{h(k)\}_{k \geq 2} \) where \( h(k) = (k - 2)! \) from Example 3.2.2. Fix \( K = 2 \) and observe that \( \frac{h(k+1)}{h(k)} = \frac{(k-1)!}{(k-2)!} = k - 1 \). Since there is no constant upper bound for \( h(k) \) we instead choose \( A(k) = B(k) = k - 1 \), and note that \( A(k) < B(k+1) \) for all \( k \geq K \). Moreover, define \( \lambda(k) = (k - 3)! \) so that

\[
\lambda(k+2) \cdot \frac{B(k+1)}{B(k+1) - A(k)} = (k - 1)! \cdot \frac{k}{k - (k - 1)} = k \geq h(k+2),
\]

for all \( k \geq K \). Therefore by Proposition 4.4 we obtain

\[
H_n^{(r)}[h] < r! \cdot \prod_{k=2}^{n} (r + 2k - 3)!.
\]

Again these determinants increase without bound (this time as \( n \to \infty \)), but we have now demonstrated that understanding the rate at which the sequence increases allows us to bound the rate at which the determinant increases.

### 4.2. Hankel determinants of ordinary Dirichlet series

In this section, we ultimately make use of Proposition 4.4 to bound Monien’s determinants of zeta values. To begin with we consider ordinary Dirichlet series with nonnegative coefficients. There is a natural split between what can be considered a degenerate case, and a nondegenerate case. As the degenerate case, we consider a Dirichlet series having only finitely many nonzero coefficients, and as the nondegenerate case, we consider Dirichlet series having infinitely many nonzero coefficients.

We use the following result for the nonvanishing of values of Hankel determinants.

**Proposition 4.5** (Monien [27], 2009). Let \( F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \) be an ordinary Dirichlet series, which is convergent for \( \Re(s) \geq s_0 \). Then for any \( n \geq 1 \) and \( r \geq s_0 - 2 \), we have

\[
H_n^{(r)}[F] = \frac{1}{n!} \sum_{m_1, m_2, \ldots, m_n=1}^{\infty} \prod_{i=1}^{n} \frac{f(m_i)}{m_i^{2n+r}} \prod_{i<j} (m_i - m_j)^2.
\]
While Proposition 4.5 holds in great generality, a classical result due to Kronecker (see also Pólya and Szegö [30, Part 7]) provides a nice dichotomy.

**Proposition 4.6** (Kronecker, 1881). Let \( \{h(k)\}_{k \geq 0} \) be a sequence of real numbers. The function \( \sum_{k=0}^{\infty} h(k)x^k \) is rational if and only if finitely many of the determinants \( H_n^{(r)}[h] \) are nonzero.

This leads to the following result.

**Lemma 4.7.** Let \( F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \) be an ordinary Dirichlet series with \( f(n) \geq 0 \) for all \( n \), which is convergent for \( \Re(s) \geq s_0 \). Then

(i) if there are only finitely many \( f(n) \neq 0 \) (degenerate case), then \( H_n^{(r)}[F] = 0 \) for sufficiently large \( n \) and any \( r \geq s_0 - 2 \);

(ii) if there are infinitely many \( f(n) \neq 0 \) (nondegenerate case), then \( H_n^{(r)}[F] > 0 \) for any \( n \geq 1 \) and \( r \geq s_0 - 2 \).

**Proof.** Proposition 4.5 implies both (i) and (ii) immediately. Part (i) also follows from Kronecker’s theorem.

To further examine the Hankel determinants in the nondegenerate case and to prepare for an application of Proposition 4.4, we require a lower bound on the ratio of successive values of \( F(s) \).

**Lemma 4.8.** Let \( F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \) be a nondegenerate ordinary Dirichlet series with \( f(n) \geq 0 \) for all \( n \), which is convergent for \( \Re(s) \geq s_0 \). If \( N < M \) are the two minimal indices of the nonzero coefficients of \( F(s) \), then the following hold:

(i) if \( N = 1 \), then
\[
\lim_{s \to \infty} \frac{M^{s+1}}{(M-1)f(M)} \left( 1 - \frac{F(s+1)}{F(s)} \right) = 1;
\]

(ii) if \( N > 1 \), then
\[
\lim_{s \to \infty} \frac{f(N)N^{(\alpha-1)s}}{f(M)(1-N^{1-\alpha})} \left( 1 - N \cdot \frac{F(s+1)}{F(s)} \right) = 1,
\]
where \( \alpha = \log M / \log N \).

**Proof.** We treat the two cases, \( N = 1 \) and \( N > 1 \), separately.

**Case** \( N = 1 \). We have
\[
F(s) = 1 + f(M) \cdot M^{-s} + O(M^{-\delta s}),
\]
where $\delta = \log(M + 1)/\log M > 1$. Thus

$$
\frac{F(s + 1)}{F(s)} = \frac{1 + \frac{f(M)}{M} M^{-s} + O(M^{-\delta s})}{1 + f(M) M^{-s} + O(M^{-\delta s})} = 1 - (M - 1) \frac{f(M)}{M^{s+1}} + O\left(\frac{1}{M^{\delta s}}\right),
$$

and so

$$
\frac{M^{s+1}}{(M - 1)f(M)} \left(1 - \frac{F(s + 1)}{F(s)}\right) = 1 + O\left(\frac{M^{s+1}}{M^{\delta s}}\right).
$$

Since $\delta > 1$, taking $s$ to infinity gives the required result.

**Case** $N > 1$. We have

$$
F(s) = f(N) \cdot N^{-s} + f(M) \cdot N^{-\alpha s} + O\left(N^{-\beta s}\right)
$$

where $\alpha = \log M/\log N$ and $\beta = \log(M + 1)/\log N$. Note that $\beta > \alpha > 1$. Thus

$$
\frac{F(s + 1)}{F(s)} = \frac{f(N) \cdot N^{-(s+1)} \left(1 + \frac{f(M)}{f(N)} \cdot N^{(1-\alpha)(s+1)} + O\left(N^{(1-\beta)s}\right)\right)}{f(N) \cdot N^{-s} \left(1 + \frac{f(M)}{f(N)} \cdot N^{(1-\alpha)s} + O\left(N^{(1-\beta)s}\right)\right)}
$$

$$
= \frac{1}{N} \left(1 - \frac{f(M)}{f(N)} \left(1 - N^{1-\alpha}\right) N^{(1-\alpha)s} + O\left(N^{(1-\alpha)s}\right)\right),
$$

and so

$$
\frac{f(N) N^{(\alpha-1)s}}{f(M)(1 - N^{1-\alpha})} \left(1 - N \cdot \frac{F(s + 1)}{F(s)}\right) = 1 + O\left(1\right).
$$

This completes the proof of the lemma. \qed

Lemma 4.8 yields the following immediate corollary.

**Corollary 4.9.** Let $F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$ be a nondegenerate ordinary Dirichlet series with $f(n) \geq 0$ for all $n$, which is convergent for $\Re(s) \geq s_0$. If $N < M$ are the two minimal indices of the nonzero coefficients of $F(s)$, then the following hold:

(i) if $N = 1$, then there are positive constants $c_1 > 0$ and $K_1 \geq s_0$ such that for all $s \geq K_1 \geq s_0$, we have

$$
0 < 1 - \frac{c_1}{M^s} \leq \frac{F(s + 1)}{F(s)} < 1;
$$

(ii) if $N > 1$, then there are positive constants $c_2 > 0$ and $K_2 \geq s_0$ such that for all $s \geq K_2 \geq s_0$, we have

$$
0 < \frac{1}{N} - \frac{c_2}{N^{(\alpha-1)s}} \leq \frac{F(s + 1)}{F(s)} < \frac{1}{N},
$$

where $\alpha = \log M/\log N$. 


We are now in a position to prove a general upper bound on nondegenerate ordinary Dirichlet series.

**Theorem 4.10.** Let \( F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \) be a nondegenerate ordinary Dirichlet series with \( f(n) \geq 0 \) for all \( n \), which is convergent for \( \Re(s) \geq s_0 \). Then \( H_n^{(r)}[F] > 0 \) for all \( n \geq 1 \) and \( r \geq s_0 - 2 \), and for sufficiently large \( r \) there is a \( c = c(r) > 0 \) such that

\[
\log H_n^{(r)}[F] < -cn^2,
\]

for all sufficiently large integers \( n \).

**Proof.** As in the proof of Lemma 4.8, we split this proof into the two cases, \( N = 1 \) and \( N > 1 \), where \( N \) is the minimal index of nonzero coefficients of \( F(s) \).

**Case** \( N = 1 \). Let \( c_1, K_1 > 0 \) be such that the conclusion of Corollary 4.9(i) holds, and for all \( k \), set

\[
A(k) = 1 - \frac{c_1}{M^k}.
\]

Then we have \( A(k) \leq \frac{F(k+1)}{F(k)} < B(k) \), where \( A(k) \) and \( B(k) = 1 \) are positive functions for \( k \geq K_1 \). Also,

\[
\frac{B(k+1)}{B(k+1) - A(k)} = \frac{M^k}{c_1}.
\]

Since \( F(s) \) is a monotonically decreasing function of \( s \geq s_0 \), we may set

\[
\lambda(k) = F(s_0) \cdot (1 - A(k-2)) = F(s_0) \cdot \frac{c_1}{M^{k-2}},
\]

so that

\[
\lambda(k+2) \cdot \frac{B(k+1)}{B(k+1) - A(k)} = F(s_0) > F(k+2).
\]

We may now apply Proposition 4.4, with \( K = K_1 \), so that

\[
H_n^{(r)}[F] < F(2 + r) \prod_{k=2}^{n} \lambda(2k + r) \leq F(s_0) \prod_{k=2}^{n} \lambda(2k),
\]
for all integers \( n \geq 3 \) and \( r \geq K \). Thus
\[
\log H_n^{(r)}[F] < \log \left( F(s_0) \prod_{k=2}^{n} \lambda(2k) \right)
= \log F(s_0) + \sum_{k=2}^{n} \log \lambda(2k)
= \log F(s_0) + \sum_{k=2}^{n} \log \left( F(s_0) \cdot \frac{c_1}{M^{2k-2}} \right)
= \log F(s_0) + \sum_{k=2}^{n} \log(F(s_0) \cdot c_1 M^2) - \sum_{k=2}^{n} 2k \log M
= \log F(s_0) + (n - 1) \log(F(s_0) \cdot c_1 M^2) - 2 \log M \cdot \frac{n(n + 1) - 2}{2}
\sim -n^2 \log M.
\]
This completes the \( N = 1 \) case.

Case \( N > 1 \). Let \( c_2, K_2 > 0 \) be such that conclusion of Corollary 4.9(ii) holds, and for all \( k \), set
\[
A(k) = \frac{1}{N} - \frac{c_2}{N^{(\alpha-1)k}}.
\]
Then we have \( A(k) \leq F(k+1)/F(k) < B(k) \), where \( A(k) \) and \( B(k) = \frac{1}{N} \) are positive functions for \( k \geq K_2 \). Also,
\[
\frac{B(k+1)}{B(k+1) - A(k)} = \frac{1}{N} \cdot \frac{N^{(\alpha-1)k}}{c_2}.
\]
Since \( F(s) \) is a monotonically decreasing function of \( s \geq s_0 \), we may set
\[
\lambda(k) = N \cdot F(s_0) \cdot \left( \frac{1}{N} - A(k-2) \right) = N \cdot F(s_0) \cdot \frac{c_2}{N^{(\alpha-1)(k-2)}},
\]
so that
\[
\lambda(k+2) \cdot \frac{B(k+1)}{B(k+1) - A(k)} = F(s_0) > F(k + 2),
\]
We may now apply Proposition 4.4, with \( K = K_2 \), so that
\[
H_n^{(r)}[F] < F(2 + r) \prod_{k=2}^{n} \lambda(2k + r) \leq F(s_0) \prod_{k=2}^{n} \lambda(2k),
\]
for all integers \( n \geq 3 \) and \( r \geq K \). Thus

\[
\log H_n^{(r)}[F] < \log \left( F(s_0) \prod_{k=2}^{n} \lambda(2k) \right)
\]

\[
= \log F(s_0) + \sum_{k=2}^{n} \log \lambda(2k)
\]

\[
= \log F(s_0) + \sum_{k=2}^{n} \log \left( N \cdot F(s_0) \cdot \frac{c_2}{N^{(\alpha-1)(k-2)}} \right)
\]

\[
= \log F(s_0) + \sum_{k=2}^{n} \log \left( N \cdot F(s_0)c_2N^{2(\alpha-1)} \right) - \sum_{k=2}^{n} 2(\alpha - 1)k \log N
\]

\[
= \log F(s_0) + (n - 1) \log \left( F(s_0)c_2N^{2\alpha-1} \right)
\]

\[
- 2(\alpha - 1) \log N \cdot \frac{n(n+1) - 2}{2}
\]

\[
\sim -n^2(\alpha - 1) \log N.
\]

Recalling \( \alpha > 1 \) completes the \( N > 1 \) case, and with that, the proof of the theorem. \( \square \)

This bound does not allow us to deduce any results on the possible (ir)rationality of particular values of Dirichlet series, but it does allow us to give some description of their behaviour in the case that infinitely many values of a Dirichlet series are rational.

**Corollary 4.11.** Let \( F(s) = \sum_{m=1}^{\infty} \frac{f(m)}{m^s} \) be a nondegenerate ordinary Dirichlet series with \( f(n) \geq 0 \) for all \( n \), which is convergent for \( \Re(s) \geq s_0 \). Suppose that \( F(k) \in \mathbb{Q} \) for all integers \( k \geq K \geq s_0 \), and for each \( k \geq K \), define the positive pair of coprime integers \( p_k \) and \( q_k \) by \( \frac{p_k}{q_k} = F(k) \). For any fixed integer \( R > K \), set

\[
D_m[F] = \text{lcm}\{q_{R+2}, q_{R+3}, \ldots, q_{R+m}\}.
\]

Then the common denominator \( D_m[F] \) grows at least exponentially in \( m \).

**Proof.** Suppose \( F(k) \) is rational for all integers \( k \geq K \). Fix \( R > K \) and let \( N > 0 \) such that for \( n \geq N \) we have \( \log H_n^{(R)}[F] < -cn^2 \) as in Theorem 4.10. As in the statement of the theorem, write \( \frac{p_k}{q_k} = F(k) \) and set

\[
D_m[F] = \text{lcm}\{q_{R+2}, q_{R+3}, \ldots, q_{R+m}\}.
\]
Then, using the positivity of $H_n^{(R)}[F]$ given by Theorem 4.10, we have
\[(D_{2n}[F])^n \cdot H_n^{(R)}[F] \in \mathbb{N},\]
so that
\[
\log H_n^{(R)}[F] + n \log D_{2n}[F] \geq 0,
\]
for all integers $n \geq N$. Suppose, for a contradiction, that for large enough $m$ the sequence $D_m[F]$ satisfied $D_m \leq C^m$ for some $C \in (1, e^{c/2})$ where $c > 0$ is as given by Theorem 4.10. Then using Theorem 4.10, for large enough $n$ there is a $\delta > 0$ such that
\[
0 \leq \log H_n^{(R)}[F] + n \log D_{2n}[F] \leq \log H_n^{(R)}[F] + 2n^2 \log C = \log H_n^{(R)}[F] + (c - \delta)n^2 < -\delta n^2,
\]
which is a contradiction. Thus the sequence $D_m[F]$ cannot be bounded in this way. \[\square\]

We note in the case that $F(s) = \zeta(s)$, if one keeps track of constants, then Theorem 4.10 implies for each $\varepsilon \in (0, 1)$ that $D_m[\zeta] > (2 - \varepsilon)^m$ for large enough $m$ depending on $\varepsilon$. This in turn implies the $q_k$, as defined in Corollary 4.11, satisfy
\[
\max_{k \leq m} \{q_k\} > m \log(2 - \varepsilon)
\]
for any $\varepsilon \in (0, 1)$ and $m$ large enough. It is worth noting this bound reveals very little in practice. Since the sequence $\{\zeta(k)\}_{k \geq 2}$ converges to 1 (and is always strictly greater than 1), then any subsequence of possibly rational zeta values would need to have unboundedly increasing denominators. However, if the upper bound in Theorem 4.10 can be improved then we can improve our explicit lower bound on $D_m[F]$. We observe the bound in Theorem 4.10 is close to the asymptotic one suggested by Monien, but is still clearly lacking. To this end, we offer up the following conjecture.

**Conjecture 4.12.** The bound in Theorem 4.10 can be improved to give
\[
\log H_n^{(r)}[F] \asymp -n^2 \log n,
\]
for all sufficiently large integers $n$ and $r$. 

4.2. Hankel Determinants of Ordinary Dirichlet Series

Note in the case of \( \zeta(s) \) we already have the lower bound in Proposition 4.3. This unfortunately means we are unable to fully resolve Conjecture 1.9, although we can make some progress towards a solution. Fix integers \( a \geq 1 \) and \( d \geq 1 \) and recall the definition of \( W_n^{(a,d)}(r)[x] \) as

\[
W_n^{(a,d)}(r)[x] = \det_{1 \leq i,j \leq n} \left( x_{a+(i+j-1)d} - \sum_{k=1}^{r} \frac{1}{k^{a+(i+j-1)d}} \right)
\]

for all integers \( n \geq 1 \) and \( r \geq 0 \), and where the terms \( x_{a+d}, \ldots, x_{a+(2n-1)d} \) are formal parameters. If we again set \( x_{a+kd} = \zeta(a+kd) \) for \( k \in \{1, 2, \ldots, 2n-1\} \), then we have

\[
W_n^{(a,d)}(r)[\zeta] = \det_{1 \leq i,j \leq n} \left( \zeta(a+(i+j-1)d) - \sum_{k=1}^{r} \frac{1}{k^{a+(i+j-1)d}} \right)
\]

\[
= H_n^{(0)}[F],
\]

where \( F(k) = \sum_{m=r+1}^{\infty} \frac{1}{m^{a+(k-1)d}} \). Note \( F(k) \) is then a nondegenerate ordinary Dirichlet series with nonnegative coefficients, and that for some integer \( R > 0 \) we have

\[
H_n^{(R)}[F] = \det_{1 \leq i,j \leq n} \left( \zeta(a+(i+j+R-1)d, r+1) \right) = W_n^{(a+Rd,d)}(r)[\zeta].
\]

Applying Theorem 4.10 then allows us to conclude that \( W_n^{(a,d)}(r)[\zeta] > 0 \) for all \( n \geq 1 \) and \( a \geq 1 \), and that there is a \( c > 0 \) such that

\[
\log W_n^{(a,d)}(r)[\zeta] < -cn^2,
\]

for all sufficiently large integers \( n \) and \( a \).

To conclude our comments on Theorem 4.10 and its consequences, we once more return to Dodgson condensation. Recall the crucial step in producing the bound in Proposition 4.4 was to bound the recurrence relation from Lemma 3.3. We could, however, proceed by applying Dodgson condensation without bounding any of the
determinants. Specifically, if we have \( H_{n}^{(r)}[h] > 0 \) for all integers \( n \geq 1 \) and \( r \geq 0 \), then repeated use of Dodgson condensation gives

\[
H_{n+1}^{(r)}[h] = H_{n}^{(r)}[h] \cdot \frac{H_{n}^{(r+2)}[h]}{H_{n-1}^{(r+2)}[h]} \left( 1 - \frac{\left( H_{n}^{(r+1)}[h] \right)^2}{H_{n}^{(r)}[h] \cdot H_{n}^{(r+2)}[h]} \right)
\]

so that replacing \( n+1 \) with \( n \) we have

\[
H_{n}^{(r)}[h] = H_{n-1}^{(r)}[h] \cdot \frac{H_{2}^{(r+2(n-2))}[h]}{H_{1}^{(r+2(n-2))}[h]} \prod_{j=0}^{n-3} \left( 1 - \frac{\left( H_{n-j}^{(r+2j+1)}[h] \right)^2}{H_{n-j}^{(r+2j)}[h] \cdot H_{n-j-1}^{(r+2j+2)}[h]} \right)
\]

In the case of a nondegenerate ordinary Dirichlet series \( F(s) \) with nonnegative coefficients, the double product is bounded above by 1, and so we have

\[
H_{n}^{(r)}[F] < F(2+r) \prod_{i=2}^{n} \frac{H_{i}^{(r+2(i-2))}[F]}{F(r + 2(i - 1))} = F(2+r) \prod_{i=2}^{n} \left( F(r + 2i) - \frac{F(r + 2i - 1)^2}{F(r + 2i - 2)} \right).
\]

Now if we again use \( N \) and \( M \) for the minimal indices of positive coefficients of \( F(s) \), then we have

\[
F(s) = \frac{f(N)}{N^s} + \frac{f(M)}{M^s} + \mathcal{O} \left( \frac{1}{(M+1)^s} \right),
\]

and

\[
F(s)^2 = \frac{f(N)^2}{N^{2s}} + \frac{2f(N)f(M)}{(NM)^s} + \mathcal{O} \left( \frac{1}{M^{2s}} \right).
\]
Thus

\[
F(r + 2i) - \frac{F(r + 2i - 1)^2}{F(r + 2i - 2)} = \frac{f(N)}{N^{r + 2i}} + \frac{f(M)}{M^{r + 2i}} + O\left(\frac{1}{(M + 1)^{r + 2i}}\right)
\]

\[
- \frac{f(N)^2}{N^{2(r + 2i - 1)}} + \frac{2f(N)f(M)}{(NM)^{r + 2i - 1}} + O\left(\frac{1}{M^{2(r + 2i - 1)}}\right)
\]

\[
= \frac{f(N)}{N^{r + 2i - 2}} + \frac{f(M)}{M^{r + 2i - 2}} + O\left(\frac{1}{(M + 1)^{r + 2i - 2}}\right)
\]

\[
= \frac{f(N)}{N^{r + 2i}} + \frac{f(M)}{M^{r + 2i}} + O\left(\frac{1}{(M + 1)^{r + 2i}}\right)
\]

Putting this together gives

\[
H_{n}^{(r)}[F] < F(2 + r) \prod_{i=2}^{n} \left[ F(r + 2i) - \frac{F(r + 2i - 1)^2}{F(r + 2i - 2)} \right]
\]

\[
= F(2 + r) \prod_{i=2}^{n} \left[ \frac{f(M)}{M^{r + 2i}} \left( \frac{M - N}{N} \right)^2 \left( 1 + O\left(\frac{1}{(M + 1)^{2(r + 2i)}}\right) \right) \right].
\]

This immediately implies that there is a constant \( c > 0 \) such that

\[
\log H_{n}^{(r)}[F] < -c \cdot n^2,
\]

for \( n \) large enough. As our work is based on using this term, Theorem 4.10 cannot be improved further using our method. Getting upper bounds on the double product term involves obtaining good lower bounds on \( \log H_{n}^{(r)}[F] \); we hope to address this in future work.

### 4.3. Hankel determinants of Dirichlet L-series

In Section 4.2 we considered ordinary Dirichlet series with nonnegative coefficients. This allowed us to use Lemma 4.7 to show that the associated Hankel determinants were positive, and thus we took advantage of the inequalities that followed from bounding Dodgson’s condensation. This prompts the question whether
these conditions are necessary or could be relaxed, and so a natural place to look is to the related case of Dirichlet $L$-series.

**Definition 4.13.** A real *Dirichlet character* is a function $\chi : \mathbb{N} \to \mathbb{R}$ where:

(i) There is an integer $m > 0$ such that $\chi(n) = \chi(n + m)$ for all integers $n > 0$,

(ii) $\chi(n) \neq 0$ if and only if $\gcd(n, k) = 1$, and

(iii) $\chi(n_1 n_2) = \chi(n_1) \chi(n_2)$ for all integers $n_1$ and $n_2$.

The Dirichlet character which takes values of only 0 and 1 only is called principal.

Note that Definition 4.13 indicates that the function $\chi$ is periodic with period $m$. We refer to the integer $m$ as the modulus of the character $\chi$.

**Definition 4.14.** An *induced modulus* is a divisor $d$ of $m$ such that $\chi(n) = 1$ whenever $\gcd(n, m) = 1$ and $n \equiv 1 \pmod{d}$. The Dirichlet character $\chi$ is primitive if it has no induced modulus $d < m$, and imprimitive otherwise.

**Definition 4.15.** A *Dirichlet L-series* is a series of the form

$$L_m(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s},$$

for some Dirichlet character $\chi$ with modulus $m$.

Let us start by considering the $L$-series of a character $\chi$ with modulus 3. Now $\chi(n)$ must be either of:

$$\chi_0(n) = \begin{cases} 0 & \text{if } n \equiv 0 \pmod{3} \\ 1 & \text{if } n \equiv 1 \pmod{3} \\ 1 & \text{if } n \equiv 2 \pmod{3} \end{cases} \quad \text{or} \quad \chi_1(n) = \begin{cases} 0 & \text{if } n \equiv 0 \pmod{3} \\ 1 & \text{if } n \equiv 1 \pmod{3} \\ -1 & \text{if } n \equiv 2 \pmod{3} \end{cases}.$$

Note that $\chi_0$ is a principal character, which means that Hankel determinants of the Dirichlet series $L_3(s, \chi_0)$ fall within the hypotheses of Theorem 4.10. So we only consider the case of $L_3(s, \chi_1)$ and disregard principal Dirichlet characters of other moduli. Table 4.1 lists the values of $H_n^{(r)}[L_3(s, \chi_1)]$ for several choices of $n$ and $r$, and gives rise to two immediate observations. The first is that the sign of $H_n^{(r)}[L_3(s, \chi_1)]$ appears to depend only on $n$, and is only positive when $n \equiv 0, 1 \pmod{4}$. Further evidence for this can be seen in Figure 4.2. The second is that $H_n^{(r)}[L_3(s, \chi_1)]$ approaches zero in magnitude at a similar rate to $H_n^{(r)}[\zeta]$, as illustrated by Figure 4.3.
4.3. Hankel Determinants of Dirichlet $L$-Series

Table 4.1. Values of the Hankel determinant $H_n^{(r)}[L_3(s, \chi_1)]$ for different values of $n$ and $r$.

<table>
<thead>
<tr>
<th>$H_n^{(r)}$</th>
<th>$r = 0$</th>
<th>$r = 1$</th>
<th>$r = 10$</th>
<th>$r = 50$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 2$</td>
<td>$-4.7 \times 10^{-2}$</td>
<td>$-2.7 \times 10^{-2}$</td>
<td>$-6.1 \times 10^{-5}$</td>
<td>$-5.6 \times 10^{-17}$</td>
</tr>
<tr>
<td>$n = 3$</td>
<td>$-3.9 \times 10^{-5}$</td>
<td>$-6.4 \times 10^{-6}$</td>
<td>$-1.1 \times 10^{-13}$</td>
<td>$-9.6 \times 10^{-50}$</td>
</tr>
<tr>
<td>$n = 4$</td>
<td>$1.8 \times 10^{-10}$</td>
<td>$6.4 \times 10^{-12}$</td>
<td>$6.9 \times 10^{-26}$</td>
<td>$6.2 \times 10^{-90}$</td>
</tr>
<tr>
<td>$n = 5$</td>
<td>$2.8 \times 10^{-18}$</td>
<td>$1.5 \times 10^{-20}$</td>
<td>$1.0 \times 10^{-41}$</td>
<td>$2.5 \times 10^{-139}$</td>
</tr>
<tr>
<td>$n = 6$</td>
<td>$-8.1 \times 10^{-29}$</td>
<td>$-6.1 \times 10^{-32}$</td>
<td>$-4.4 \times 10^{-61}$</td>
<td>$-8.2 \times 10^{-195}$</td>
</tr>
<tr>
<td>$n = 7$</td>
<td>$-3.3 \times 10^{-42}$</td>
<td>$-2.8 \times 10^{-46}$</td>
<td>$-5.1 \times 10^{-84}$</td>
<td>$-2.7 \times 10^{-257}$</td>
</tr>
<tr>
<td>$n = 8$</td>
<td>$1.4 \times 10^{-58}$</td>
<td>$1.1 \times 10^{-63}$</td>
<td>$1.4 \times 10^{-110}$</td>
<td>$2.1 \times 10^{-325}$</td>
</tr>
<tr>
<td>$n = 9$</td>
<td>$4.3 \times 10^{-78}$</td>
<td>$3.0 \times 10^{-84}$</td>
<td>$9.3 \times 10^{-141}$</td>
<td>$1.6 \times 10^{-399}$</td>
</tr>
<tr>
<td>$n = 10$</td>
<td>$-8.1 \times 10^{-101}$</td>
<td>$-4.5 \times 10^{-108}$</td>
<td>$-1.2 \times 10^{-174}$</td>
<td>$-4.9 \times 10^{-479}$</td>
</tr>
<tr>
<td>$n = 20$</td>
<td>$-8.5 \times 10^{-522}$</td>
<td>$4.4 \times 10^{-542}$</td>
<td>$4.3 \times 10^{-726}$</td>
<td>$3.9 \times 10^{-1560}$</td>
</tr>
<tr>
<td>$n = 50$</td>
<td>$-1.5 \times 10^{-4255}$</td>
<td>$-8.7 \times 10^{-4326}$</td>
<td>$-2.0 \times 10^{-4959}$</td>
<td>$-2.1 \times 10^{-7801}$</td>
</tr>
</tbody>
</table>

Let us next consider the $L$-series of a character $\chi$ with modulus 4. If $\chi$ is not a principal character we must have

$$\chi(n) = \begin{cases} 
0 & \text{if } n \equiv 0 \pmod{4} \\
1 & \text{if } n \equiv 1 \pmod{4} \\
0 & \text{if } n \equiv 2 \pmod{4} \\
-1 & \text{if } n \equiv 3 \pmod{4}.
\end{cases}$$

Note that $\phi(3) = \phi(4) = 2$, and so both nonprincipal characters of moduli 3 and 4 take the same number of non-zero values per period. Figure 4.4 illustrates the sign alternation for $H_n^{(r)}[L_4(s, \chi)]$ is the same as for $H_n^{(r)}[L_3(s, \chi_1)]$, and Figure 4.5 also shows $H_n^{(r)}[L_4(s, \chi)]$ approaching zero in magnitude again at a similar rate to $H_n^{(r)}[\zeta]$. 

Figure 4.2. The sign of $H_n^{(r)}[L_3(s, \chi_1)]$ for different values of $n$ and $r$, where a black fill corresponds to a positive Hankel determinant.

Figure 4.3. Values of $\log(H_n^{(r)}[L(s, \chi_1)])$ for increasing $n$ where $r \in \{0, 10, 20, 30\}$, with increasing dash spacing for increasing $r$. The bold curve shows $\log(H_n^{(0)}[\zeta])$ for comparison.
Figure 4.4. The sign of $H_n^{(r)}[L_4(s, \chi)]$ for different values of $n$ and $r$, where a black fill corresponds to a positive Hankel determinant.

Figure 4.5. Values of $\log(H_n^{(r)}[L_4(s, \chi)])$ for increasing $n$ where $r \in \{0, 10, 20, 30\}$, with increasing dash spacing for increasing $r$. The bold curve shows $\log(H_n^{(0)}[\zeta])$ for comparison.
Since $\phi(5) = 4$, we next consider the $L$-series of the nonprincipal character $\chi$ with modulus 5, given by

$$\chi(n) = \begin{cases} 
0 & \text{if } n \equiv 0 \pmod{5} \\
1 & \text{if } n \equiv 1 \pmod{5} \\
-1 & \text{if } n \equiv 2 \pmod{5} \\
-1 & \text{if } n \equiv 3 \pmod{5} \\
1 & \text{if } n \equiv 4 \pmod{5}.
\end{cases}$$

This time we see a different effect taking place when we consider the sign of $H_n^{(r)}[L_5(s, \chi)]$ for different values of $n$ and $r$. Figure 4.6 shows the sign of the Hankel determinants varying as $r$ varies. In addition, there seems to be some self-similarity in the distribution of signs, suggesting there is some period with which the sign alternations of these determinants repeat. Figure 4.7 shows $H_n^{(r)}[L_5(s, \chi)]$ approaching zero in magnitude at a similar rate to $H_n^{(r)}[\zeta]$.

![Figure 4.6](image_url)

**Figure 4.6.** The sign of $H_n^{(r)}[L_5(s, \chi)]$ for different values of $n$ and $r$, where a black fill corresponds to a positive Hankel determinant.
4.3. HANKEL DETERMINANTS OF DIRICHLET $L$-SERIES

Figure 4.7. Values of $\log(H_n^{(r)}[L_5(s,\chi)])$ for increasing $n$ where $r \in \{0, 10, 20, 30\}$, with increasing dash spacing for increasing $r$. The thicker curve shows $\log(H_n^{(0)}[\zeta])$ for comparison.

Since $\phi(8) = \phi(5) = 4$, we also consider each of the $L$-series of characters $\chi$ with modulus 8, given by

$$
\chi_1(n) = \begin{cases} 
0 & \text{if } n \equiv 0 \pmod{8} \\
1 & \text{if } n \equiv 1 \pmod{8} \\
0 & \text{if } n \equiv 2 \pmod{8} \\
-1 & \text{if } n \equiv 3 \pmod{8} \\
0 & \text{if } n \equiv 4 \pmod{8} \\
0 & \text{if } n \equiv 5 \pmod{8} \\
0 & \text{if } n \equiv 6 \pmod{8} \\
-1 & \text{if } n \equiv 7 \pmod{8}
\end{cases}
$$

or

$$
\chi_2(n) = \begin{cases} 
0 & \text{if } n \equiv 0 \pmod{8} \\
1 & \text{if } n \equiv 1 \pmod{8} \\
0 & \text{if } n \equiv 2 \pmod{8} \\
-1 & \text{if } n \equiv 3 \pmod{8} \\
0 & \text{if } n \equiv 4 \pmod{8} \\
0 & \text{if } n \equiv 5 \pmod{8} \\
0 & \text{if } n \equiv 6 \pmod{8} \\
1 & \text{if } n \equiv 7 \pmod{8}
\end{cases}
$$

The patterns of sign alternations in $H_n^{(r)}[L_8(s,\chi_1)]$ and $H_n^{(r)}[L_8(s,\chi_2)]$ are very similar to those in $H_n^{(r)}[L_5(s,\chi)]$, producing a graph similar to Figure 4.6, and these determinants again approach zero in the expected way. It is worth noting for the
Dirichlet character given by

\[
\chi_3(n) = \begin{cases} 
0 & \text{if } n \equiv 0 \pmod{8} \\
1 & \text{if } n \equiv 1 \pmod{8} \\
0 & \text{if } n \equiv 2 \pmod{8} \\
-1 & \text{if } n \equiv 3 \pmod{8} \\
0 & \text{if } n \equiv 4 \pmod{8} \\
1 & \text{if } n \equiv 5 \pmod{8} \\
0 & \text{if } n \equiv 6 \pmod{8} \\
-1 & \text{if } n \equiv 7 \pmod{8},
\end{cases}
\]

that \(\chi_3(1) = \chi_3(5) = 1\), and so this character with modulus 8 also has 4 as an induced modulus. Thus \(\chi_3\) is imprimitive, and the Dirichlet \(L\)-series associated with this character is identical to the series associated with the nonprincipal character of modulus 4.

As \(\phi(7) = 6\), we see the Hankel determinants of the \(L\)-series whose character \(\chi\) of modulus 7, given by

\[
\chi(n) = \begin{cases} 
0 & \text{if } n \equiv 0 \pmod{7} \\
1 & \text{if } n \equiv 1 \pmod{7} \\
1 & \text{if } n \equiv 2 \pmod{7} \\
-1 & \text{if } n \equiv 3 \pmod{7} \\
1 & \text{if } n \equiv 4 \pmod{7} \\
-1 & \text{if } n \equiv 5 \pmod{7} \\
-1 & \text{if } n \equiv 6 \pmod{7},
\end{cases}
\]

produces another distinct pattern of sign alternations as displayed in Figure 4.8.

Our observations from these examples motivate the following conjectures.

**Conjecture 4.16.** Let \(L_m(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}\) be a Dirichlet \(L\)-series with Dirichlet character \(\chi\) of modulus \(m\). Then \(H_n^{(r)}[L_m(s, \chi)] \neq 0\) for all \(n \geq 1\) and \(r \geq 0\), and there is a \(c > 0\) such that

\[
\log |H_n^{(r)}[L_m(s, \chi)]| < -cn^2 \log n,
\]

for all sufficiently large integers \(n\) and \(r\).
Conjecture 4.17. Let \( L_m(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} \) be a Dirichlet \( L \)-series with primitive Dirichlet character \( \chi \) of modulus \( m \). For any integer \( r \geq 0 \), the sequence \( \left\{ \text{sgn} \left( H_n^{(r)}[L_m(s, \chi)] \right) \right\}_{n \geq 1} \) is eventually periodic, and its period depends on \( \phi(m) \).

It is not presently clear how these conjectures might be solved, but their resolution would reveal that the magnitude of Hankel determinants of ordinary Dirichlet series and of Dirichlet \( L \)-series appear to grow similarly. An understanding of the sign changes in these determinants might allow for the application of an altered form of Proposition 4.4, which would further broaden our earlier results.
Historical development of determinants

Historically, the theory of determinants predates the theory of matrices. Straf-fin [37] cites the ability of Chinese mathematicians to solve systems of linear equations in the 3rd century, where the determinant “determines” whether the system has a unique solution. It was not until the correspondence between Gottfried Wilhelm Leibniz and Guillaume de l’Hôpital in the late 17th century that the theory of determinants began to spread to the wider mathematical community. As recorded by Muir [28], Leibniz revealed to l’Hôpital that he had developed a rule for writing out what he called the “resultant” of a set of linear equations.

The following is a determinant evaluation, historically attributed to Leibniz, written in modern notation. Recall $S_n$ is the symmetric group of order $n$, and $\text{sgn}(\pi)$ is the signature of a permutation $\pi \in S_n$ (where $\text{sgn}(\pi)$ is 1 when $\pi$ is an even permutation and $-1$ when odd). Let $M = (m_{i,j})_{i,j=1,...,n}$ be an $n \times n$ matrix. Then the determinant of $M$ is

$$\text{det}(M) = \sum_{\pi \in S_n} \text{sgn}(\pi) \prod_{i=1}^{n} m_{i,\pi(i)}.$$  \hspace{1cm} (A.1)

As a demonstration, let us use this formula to calculate the determinant of the matrix

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}.$$  \hspace{1cm} (A.2)

From Leibniz’s definition in Equation (A.1) we obtain

$$\text{det}(M) = m_{1,1}m_{2,2}m_{3,3} - m_{1,1}m_{2,3}m_{3,2} + m_{1,2}m_{2,1}m_{3,3}$$

$$- m_{1,2}m_{2,1}m_{3,3} + m_{1,3}m_{2,1}m_{3,2} - m_{1,3}m_{2,2}m_{3,1}$$

$$= 1 \cdot 2 \cdot 3 - 1 \cdot 2 \cdot 2 + 1 \cdot 2 \cdot 1 - 1 \cdot 1 \cdot 3 + 1 \cdot 1 \cdot 2 - 1 \cdot 2 \cdot 1$$

$$= 1.$$
Note that this method of evaluating a determinant is often quite lengthy as $\mathfrak{S}_n$ consists of $n!$ permutations.

By 1772, Pierre-Simon Laplace had developed a way to express these determinants as an aggregate of other determinants, which extended the work of his contemporary Alexandre-Théophile Vandermonde. Let $M = (m_{i,j})_{i,j=1,...,n}$ be an $n \times n$ matrix, and write $M^i_j$ for the matrix obtained from $M$ by removing the $i$th row and $j$th column. Then for any fixed $i', j' \in \{1, 2, \ldots, n\}$, the determinant of $M$ is

$$\det(M) = \sum_{i=1}^{n} (-1)^{i+j'} m_{i,j'} \det \left( M^{i'}_j \right) = \sum_{j=1}^{n} (-1)^{i'+j} m_{i',j} \det \left( M^i_{j'} \right). \quad (A.3)$$

In its simplest form, the Laplace expansion gives the determinant of a $2 \times 2$ matrix as

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$  

To show the recursive nature of Equation (A.3), let us re-evaluate the determinant of the matrix in (A.2) via Laplace expansion. Fix $i' = 1$, which corresponds to expanding the determinant along its first row. Using Equation (A.3) we obtain

$$\det(M) = (-1)^2 m_{1,1} \det \left( M^1_1 \right) + (-1)^3 m_{1,2} \det \left( M^2_1 \right) + (-1)^4 m_{1,3} \det \left( M^3_1 \right)$$

$$= 1 \cdot \begin{vmatrix} 2 & 2 \\ 2 & 3 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} + 1 \cdot \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix}$$

$$= 1 \cdot 2 - 1 \cdot 1 + 1 \cdot 0$$

$$= 1.$$

As Equation (A.3) suggests, determinants are independent of the choice of row or column used in the expansion. For instance, if we instead fix $j' = 2$, corresponding to an expansion along the second column, we similarly obtain

$$\det(M) = (-1)^3 m_{1,2} \det \left( M^1_2 \right) + (-1)^4 m_{2,2} \det \left( M^2_2 \right) + (-1)^5 m_{3,2} \det \left( M^3_2 \right)$$

$$= (-1)^3 m_{1,2} \det \left( M^1_2 \right) + (-1)^4 m_{2,2} \det \left( M^2_2 \right) + (-1)^5 m_{3,2} \det \left( M^3_2 \right)$$

$$= (-1) \cdot \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} + 2 \cdot \begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} - 2 \cdot \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix}$$

$$= (-1) \cdot 1 + 2 \cdot 2 - 2 \cdot 1$$

$$= 1.$$
This method of determinant evaluation is very useful when a matrix has many zero entries. In particular, this allows for the efficient computation of the determinant of an upper triangular matrix (in which all of the entries under the main diagonal are zero). For example, by expanding along the first column of the determinant below, we calculate

\[
\begin{vmatrix}
1 & 2 & 3 \\
0 & 4 & 5 \\
0 & 0 & 6
\end{vmatrix}
= 1 \cdot \begin{vmatrix}
4 & 5 \\
0 & 6
\end{vmatrix} - 0 \cdot \begin{vmatrix}
2 & 3 \\
0 & 6
\end{vmatrix} + 0 \cdot \begin{vmatrix}
2 & 3 \\
4 & 6
\end{vmatrix}
= 1 \cdot 4 \cdot 6 = 24,
\]

and conclude that the determinant of an upper triangular matrix is equal to the product of the entries on the main diagonal.

Gaussian elimination is arguably the most widely recognised development in the theory of determinants. Historically, it was used as a method for solving simultaneous linear equations by many people, but now bears the namesake of Carl Friedrich Gauss (see Grcar [19] for an investigation of the wider history of Gaussian elimination). The method consists of a certain sequence of elementary row operations (equivalently column operations) which reduces a linear system to a simpler form. These three operations are:

(i) Swap the position of any two rows.
(ii) Multiply any row by a nonzero scalar.
(iii) Add a scalar multiple of one row to another.

The utility of these operations lie in the way they affect the determinant of the matrix being reduced: the determinant is invariant under (iii), while (ii) scales the determinant by the same scalar and (i) changes the sign of the determinant. This allows for the reduction of any matrix into upper triangular form, from which we can simply make use of our previous observations to calculate its determinant.

To illustrate this process, let us once more evaluate the determinant of the matrix in (A.2) instead using Gaussian elimination. Subtracting the second row from the third, and subtracting the first row from the second, yields

\[
\det(M) = \begin{vmatrix}
1 & 1 & 1 \\
1 & 2 & 2 \\
1 & 2 & 3
\end{vmatrix} = \begin{vmatrix}
1 & 1 & 1 \\
1 & 2 & 2 \\
0 & 0 & 1
\end{vmatrix} = \begin{vmatrix}
1 & 1 & 1 \\
0 & 1 & 1 \\
0 & 0 & 1
\end{vmatrix} = 1 \cdot 1 \cdot 1 = 1.
\]
Here we have reduced $M$ to upper triangular form and thus the determinant is equal to product of the diagonal entries.
APPENDIX B

Determinant calculations

Our main tool in evaluating determinants is Lemma 3.3, rearranged to give

$$H_n^{(r)}[h] = \frac{H_n^{(r)}[h] \cdot H_n^{(r+2)}[h] - \left( H_n^{(r+1)}[h] \right)^2}{H_n^{(r+2)}[h]}.$$ 

Using this recursively (and interpreting the determinant of a $0 \times 0$ matrix as 1) we express $H_n^{(r)}[h]$ in terms of some $1 \times 1$ determinants, simply given as $\{h(k)\}_{k=2+r}^{2n+r}$.

By way of example, we demonstrate this process in the creation of Figure 4.1.

All calculations in this thesis were performed in Maple 12.

```maple
with(plots):
# Set higher decimal precision for reliable evaluations
Digits := 300:
# Set the maximum value of n in the plot
n_max := 50:
# Define the Hankel determinant recursively via Dodgson condensation
H := proc (n, r) option remember;
    evalf((H(n-1, r)*H(n-1, r+2)-H(n-1, r+1)^2)/H(n-2, r+2))
end proc:
# Set the 0x0 determinants to 1
for r from 0 to 2*(n_max-1) do
    H(0, r) := 1
end do:
# Set the 1x1 determinants to zeta values
for r from 0 to 2*(n_max-1) do
    H(1, r) := Zeta(2+r)
end do:
# Generate the sequence of determinants and plot the result
plot([seq([n, log(H(n, 0))], n = 0 .. n_max)]);
```
Additionally, the following Maple code produces Figure 4.8.

```maple
with(plots):

# Set higher decimal precision for reliable evaluations
Digits := 700:

# Set the maximum values of n and r in the plot
n_max := 120:

r_max := 50:

# Define the Hankel determinant recursively via Dodgson condensation
H := proc(n, r) option remember;
  evalf((H(n-1, r)*H(n-1, r+2)-H(n-1, r+1)^2)/H(n-2, r+2))
end proc:

# Set the 0x0 determinants to 1
for r from 0 to 2*(n_max-1)+r_max do
  H(0, r) := 1
end do:

# Set the 1x1 determinants to L-series values
for r from 0 to 2*(n_max-1)+r_max do
  H(1, r) := evalf((-1)^r*(Psi(r+1, 1/7)+Psi(r+1, 2/7)
                  -Psi(r+1, 3/7)+Psi(r+1, 4/7)-Psi(r+1, 5/7)
                  -Psi(r+1, 6/7))/(7^(r+2)*factorial(r+1)))
end do:

# Record the indices of positive determinants and plot the result
S := []:
for n from 0 to n_max do
  for r from 0 to r_max do
    if H(n, r) > 0 then
      S := [op(S), [n, r]]
    end if
  end do
end do:

pointplot(S, symbol = solidbox, scaling = constrained);
```
Bibliography

33. G. F. B. Riemann, *On the number of prime numbers less than a given quantity (Über die Anzahl der Primzahlen unter einer gegebenen Größe)*, Monatsberichte der Berliner Akademie (1859).


