Irrationality and Transcendence

An Introduction to Modern Transcendental Number Theory

Jasmine Burgess

April 2023

Declaration

This piece of work is a result of my own work except where it forms an assessment based on group project work. In the case of a group project, the work has been prepared in collaboration with other members of the group. Material from the work of others not involved in the project has been acknowledged and quotations and paraphrases suitably indicated.

Acknowledgments

I would like to thank my supervisor, Dan Evans, for his invaluable guidance in illuminating a path through the forest of number theory.

Contents

1 Introduction **1**
1.1 A Brief History . . . 1
1.2 Preliminaries . . . 1
1.3 Impossible Integers: The Fundamental Contradiction . . . 2

2 Beukers' Method for the Irrationality of $\zeta(3)$ **5**
2.1 Method and Motivation . . . 5
2.1.1 The Zeta Function . . . 5
2.1.2 Beukers' Method: An Overview . . . 6
2.1.3 An Arithmetic Aside . . . 7
2.1.4 Irrationality of $\log 2$ . . . 8
2.2 Irrationality of $\zeta(2)$ . . . 9
2.3 Irrationality of $\zeta(3)$ . . . 12
2.4 Further Study . . . 15

3 Approximation of Real Numbers with Rationals **17**
3.1 The rational case . . . 17
3.2 Dirichlet's Idea . . . 19
3.2.1 Hurwitz's Improvement . . . 20
3.3 Liouville's Theorem . . . 20
3.4 Thue, Siegel and Roth . . . 21
3.5 Approximation Exponent of Transcendental Numbers . . . 24
3.5.1 Liouville Numbers . . . 24
3.5.2 Other Transcendental Numbers . . . 26

4 Mahler's Number **29**
4.1 Some Naive Approximations . . . 29
4.2 Amou's Approximations . . . 30
4.2.1 Asymptotic Behaviour . . . 32
4.2.2 Approximation Exponent . . . 34
4.3 Generalisations . . . 35

5 Hermite's Method for Transcendence **36**
5.1 Transcendence of e^r . . . 36
5.1.1 Evaluating M and M_t . . . 37
5.1.2 Bounding the Integer . . . 39
5.2 The Transcendence Of π . . . 40
5.2.1 Symmetric Polynomials . . . 40
5.2.2 Lindemann's Proof . . . 41
5.2.3 The Original Direction of Proof . . . 45
5.3 Squaring The Circle . . . 46

6 Discussion **48**
6.1 Linear Independence of Transcendental Numbers . . . 48
6.2 Beukers-Lite Integrals for Irrationality and Periods . . . 49

Chapter 1

Introduction

1.1 A Brief History

The discovery of irrational numbers is often attributed to the Ancients Greeks, although it is impossible to be certain of many details. Pythagorus allegedly led a cult who believed all numbers could be expressed as arithmetic ratios (fractions) and he is fabled to have killed Hippasus of Metapontum, a disciple who proved the irrationality of $\sqrt{2}$ [7].

However, irrationality may actually have been first discovered in the Vedic civilisation of India and recorded in texts called the Shulba Sutras, written up to 300 years before Pythagorus. Like the Pythagoreans, the Vedic civilisation did not separate ideas of religiosity and mathematics, and it is possible that geometry was developed to meet the needs of ritual. Nevertheless, the Sutras were the world's first treatises on pure maths [21]. This report will be slightly more up to date: the methods for proving irrationality that I will look at were primarily developed by Hermite, Fourier, Niven and Beukers in the nineteenth and twentieth century.

The discovery of transcendental numbers was comparatively recent. Euler was the first to conjecture that certain class of numbers were transcendental in 1798, but it wasn't until the 1850s that Joseph Liouville found the first class of transcendental numbers. Liouville's work kick-started transcendental number theory and set the stage for a Field's medal winning discovery by Klaus Roth in 1955. I will be exploring this in detail in Chapter 3.

1.2 Preliminaries

This content in this section should be familiar to most readers.

Definition 1.1 (Irrationality)**.** A number is irrational if it cannot be written as a fraction p/q, where p and q are coprime integers and $q \neq 0$.

Throughout the rest of this report, unless otherwise stated, I will assume that all fractions are written sensibly: that is, p, q are coprime, and $q \neq 0$.

Definition 1.2 (Algebraic)**.** An algebraic number (over the rationals) is one which is the solution to a polynomial of finite degree with integer coefficients.

If α is algebraic, we can find a polynomial $F(x)$ with integer coefficients such that $F(\alpha) = 0$. Equivalently, we can find a monic polynomial $G(x)$ with rational coefficients such that $G(\alpha) = 0$.

Definition 1.3 (Minimal Polynomial)**.** The minimal polynomial of α over a field K is a monic polynomial $G(x)$ in $K[x]$ such that $G(x)$ is irreducible over K and $G(\alpha) = 0$.

Definition 1.4 (Degree)**.** The degree of α over K is the degree of its minimal polynomial over K.

If the field K is not specified, we will assume it it the rationals, $\mathbb{Q}$.

Definition 1.5 (Transcendence)**.** A transcendental number is one which is not algebraic.

We say that the degree of a transcendental number is infinite.

All rational numbers are algebraic, but an irrational number could be either algebraic or transcendental. For example, $\sqrt{2}$ is an algebraic number since it's the solution to $f(x) = x^2 - 2$, and thus $\sqrt{2}$ has degree 2 over $\mathbb{Q}$. The most famous transcendental numbers are e and π, and I will prove their transcendence in the final chapter.

1.3 Impossible Integers: The Fundamental Contradiction

Throughout this report, we show that numbers are irrational (and later, transcendental) by constructing integers which violate the "Fundamental Principle of Number Theory", i.e. that there is no integer between 0 and 1. This is done as follows:

- Assume our number, α, is rational in the form a/b .
- Define the numbers c_n dependent on a, b, α, and natural numbers n. For example, c_n may come from the difference between α and a rational approximation to α.
- Show c_n is an integer.
- Bound c_n between 0 and an upper bound which depends on n, so that as n tends to infinity, the upper bound tends to 0. This means that $0 < c_n < 1$ for large enough n.

Technically, we only need the upper bound on n to tend to some $0 \leq c < 1$, however every example in this report has the upper bound tending to 0.

Example 1.6. This is the most well-known proof for the irrationality of e and is attributed to Joseph Fourier in [10]. As expected, we assume that e is a rational number a/b, and consider the rational approximations to e found by truncating its Taylor series expansion:

$$\frac{p_n}{q_n} = \sum_{i=0}^{n} \frac{1}{i!} = 1 + 1 + \frac{1}{2!} + ... + \frac{1}{n!}.$$

Now define

$$\begin{aligned} c_n = n! \left| \frac{a}{b} - \frac{p_n}{q_n} \right| &= \frac{1}{n+1} + \frac{1}{(n+1)(n+2)} + ... \\ &< \frac{1}{n+1} + \frac{1}{(n+1)^2} + ... \\ &= \frac{1}{1 - \frac{1}{n+1}} - 1 \\ &= \frac{1}{n}. \end{aligned}$$

If we choose $n \geq b$, c_n is an integer, as clearly $q_n = n!$, so $n!$ will cancel out the denominators in c_n. However, c_n is non-zero since $p_n/q_n \neq e$, and we showed it was bounded above by $1/n < 1$. So we have created an impossible integer and a contradiction! Thus, e must be irrational.

Example 1.7 (Irrationality of π)**.** Assume $\pi = a/b$, and for n a natural number, define the functions:

$$\begin{aligned} f_n(x) &= \frac{x^n(a-bx)^n}{n!}, \\ F_n(x) &= f(x) - f^{(2)}(x) + f^{(4)}(x) - ... + (-1)^n f^{(2n)}(x). \end{aligned}$$

These functions have some useful properties.

1. $F_n(0)$ and $F_n(\pi)$ are integers.

Proof. $f_n(x)$ has a zero of order n at $x = 0$, thus for $i < n$, $f_n^{(i)}(0) = 0$. Then for $i \geq n$, differentiating n times will produce a factor of $n!$, which cancels out the denominator of $f^{(i)}(x)$, ensuring $f^{(i)}(0)$ is an integer for all i. Thus $F(0)$ is an integer.

It is also clear from substituting that $f_n(x) = f_n(a/b - x) = f_n(\pi - x)$, from which is follows that $F(\pi)$ is also an integer. □

2. $f_n(x) = F_n(x) + F_n^{(2)}(x)$.

This is evident after noting that as $f_n(x)$ is a polynomial of order $2n$, any derivatives of higher order than $2n$ must be 0.

Next, we use basic calculus to find an integer.

$$\begin{aligned}\frac{d}{dx}[F_n'(x)\sin x - F_n(x)\cos(x)] &= F_n''(x)\sin x + F_n(x)\sin x\\ &= f_n(x)\sin x,\end{aligned}$$

by the second property. Thus,

$$\begin{aligned}\int_0^\pi f_n(x)\sin x dx &= [F_n'(x)\sin x - F_n(x)\cos(x)]_0^\pi\\ &= F_n(\pi) + F_n(0).\end{aligned}$$

is an integer, by the first property. Now, as before, we aim to bound this integer. The integrand, $f_n(x)\sin(x)$ is non-negative on the whole interval $[0, \pi]$, and zero only at the endpoints, so the integer can be bounded below by zero. We therefore have

$$0 < \int_0^\pi f(x)\sin x dx = \int_0^\pi \frac{x^n(a-bx)^n}{n!}\sin x dx < \frac{\pi^n a^n}{n!},$$

by considering the maximum possible value for the integrand. However, for large n, $n!$ grows faster than any power of x^n, and so for large enough n, we have that our integer is bounded:

$$0 < F_n(0) + F_n(\pi) < \frac{\pi^n a^n}{n!} < 1,$$

which is the contradiction.

That nifty proof for the irrationality of π was found by Ivan Niven in 1946, who published it in a single page paper [20].

Auxiliary polynomials are another useful tool used in irrationality and transcendence proofs. Auxiliary polynomial are functions that are constructed during proofs to have many zeroes at different arguments, or to have zeroes of high order at certain arguments ([24], page 223).

I presented Niven's proof for the irrationality for π in the introduction as it is a simple irrationality proof using a family of integrals and auxiliary polynomials. In the next chapter, we will look at how more complicated families of integrals can evaluate to integers and prove irrationality. Niven's method of proof also has many similarities to Hermite's method for proving transcendence, which we will see in Chapter 5.

Chapter 2

Beukers' Method for the Irrationality of $\zeta(3)$

2.1 Method and Motivation

2.1.1 The Zeta Function

The Riemann-zeta function is defined as

$$\begin{aligned}\zeta(s) &= 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \frac{1}{5^s} + ... \\ &= \sum_{n=1}^{\infty} \frac{1}{n^s}.\end{aligned}$$

For real values of s greater than 1, the zeta function converges. It is known that the zeta function is irrational for even values of s, and also for $s = 3$. However, there is still active work on the irrationality of the zeta function for odd $s > 3$.

For even values of s, $\zeta(s)$ is a rational multiple of π^s, and since π is transcendental, $\zeta(s)$ must be both irrational and transcendental. For example, $\zeta(2)$ can be calculated to be $\pi^2/6$, however we will prove its irrationality without using this fact, because it will help us to prove $\zeta(3)$ is irrational.

Roger Apéry first proved the irrationality of $\zeta(3)$ in 1978 but as his proof was extremely complicated, it was met with suspicion until Frits Beukers simplified it the next year. We do not have a closed form expression for $\zeta(3)$ - it is sometimes referred to as Apéry's constant. It is also an open question as to whether $\zeta(3)$ is transcendental [19].

We are currently unable to use Beukers' method to prove $\zeta(4)$ is irrational, which is necessary to prove $\zeta(5)$ is irrational by the same method as $\zeta(3)$. In this chapter, I will bring together work in [15], [13] and [5] to explain Beuker's method for proving the irrationality of $\zeta(3)$. First, we will see the core ideas of the method, and how we can apply it to two simpler examples.

2.1.2 Beukers' Method: An Overview

This chapter extends the ideas of impossible integers. Again, we want to find an integer which is squeezed between 0 and 1 as its upper bound tends to 0. To do so, we use a series of integrals.

- Define a series of integrals I_n with n a natural number so that:

$$I_n = \int_0^1 x^n f(x)dx = A_n + B_n\alpha$$

  where A_n and B_n are rational numbers. If $f(x)$ is non-negative between 0 and 1 and not the zero function, clearly the entire integrand is non-negative in the interval, so $|I_n| > 0$.

- The Legendre polynomial of degree n satisfies Legendre's differential equation:

$$(1-x^2)\frac{d^2 f(x)}{dx^2} - 2x\frac{df(x)}{dx} - n(n+1)f(x) = 0,$$
$$\text{for } -1 < x < 1,$$

  and can be written

$$\begin{aligned} P_n(x) &= \frac{1}{n!}\frac{d^n}{dx^n}x^n(1-x^n) \\ &= \sum_{k=0}^{n}(-1)^k\binom{n}{k}\binom{n+k}{n}x^k. \end{aligned}$$

  The second line makes it clear that $P_n(x)$ is a polynomial with rational (in fact integer) coefficients, so since integrals are linear, we have that

$$J_n = \int_0^1 P_n(x)f(x)dx = A'_n + B'_n\alpha.$$

  Thus J_n can also be written as a rational sum of 1 and α.

- The Legendre polynomials were chosen because one can easily perform integration by parts with them. After performing integration by parts n times on J_n, we get the equation

$$J_n = \frac{(-1)^n}{n!}\int_0^1 x^n(1-x)^n\frac{d^n f}{dx^n}dx. \quad (2.1)$$

  We use this to find an upper bound for J_n in the form CM^n, such that M is less than 1.

- Now we follow the method explained in the introduction. Assume $\alpha = a/b$, so α is rational. Then

$$\begin{aligned} &0 < |J_n| \leq CM^n \\ \implies &0 < \left| A'_n + B'_n \frac{a}{b} \right| \leq CM^n. \end{aligned}$$

  Multiplying through by b and any denominators of A'_n and B'_n, we get an integer between 0 and some upper bound. We then show that the upper bound approaches 0 as n tends to infinity, creating an impossible integer.

2.1.3 An Arithmetic Aside

One complication we may find is that the denominator of B'_n grows with n. In fact in the examples below, it is the lowest common multiple of the first n natural numbers, which we write as d_n. Therefore, we need a preliminary lemma which relies on the well-known prime number theorem.

Theorem 2.1 (Prime Number Theorem)**.** Let n be a positive integer and define $\pi(n) := \#\{p \leq n : p \text{ is prime}\}$. Then

$$\pi(n) \sim \frac{n}{\log n}.$$

The prime number theorem is a key result of analytical number theory which tells us that for large n, $\pi(n)$ and $n/\log n$ have the same asymptotic behaviour.

Lemma 2.2. For n a positive integer, $d_n \leq n^{\pi(n)} \sim e^n$.

Proof. We have that

$$d_n = \text{lcm}(1, 2, \ldots, n) = \prod_{p \leq n} p^m,$$

where p are primes and m is the maximal integer such that:

$$\begin{aligned} &p^m \leq n \\ \implies &m \leq \log_p(n). \end{aligned}$$

Then

$$\begin{aligned} d_n &\leq \prod_{p \leq n} p^{\log_p n} = \prod_{p \leq n} n \\ &= n^{\pi(n)} \sim n^{n/\log(n)} = e^n. \end{aligned}$$

□

Lemma 2.3. For a sum of fractions, where the denominators b_i are all positive integers bounded above by n, we can write

$$S = \sum_i \frac{a_i}{b_i} = \frac{z}{d_n},$$

where z is an integer, but z and d_n are not necessarily coprime.

Proof. For any real integer $0 < b_i \leq n$, we have that $k \times b_i = d_n$ for some integer k, and so we can rewrite a_i/b_i as ka_i/d_n. Then a sum of fractions, all with denominator d_n, can be written as z/d_n. □

2.1.4 Irrationality of $\log 2$

It is possible to prove log(2) is irrational very quickly. If

$$\begin{aligned} \log(2) &= \frac{a}{b} \\ \implies 2 &= e^{a/b}. \end{aligned}$$

But in Chapter 5, we will show that e^r is in fact transcendental for all non-zero rational numbers r, so this is a contradiction. However, we will show log(2) is irrational without using this fact, in order to illustrate Beukers' method. We let

$$f(x) = \frac{1}{1+x}.$$

We then evaluate

$$I_0 = \int_0^1 f(x)dx = [\log(1+x)]_0^1 = \log 2.$$

For $n > 0$, we perform the substitution $u = 1 + x$:

$$\begin{aligned} I_n &:= \int_0^1 x^n f(x)dx \\ &= \int_1^2 \frac{(u-1)^n}{u} du \\ &= \int_1^2 u^{n-1} - nu^{n-2} + \binom{n}{2} u^{n-3} - \cdots + (-1)^{n-1} n + \frac{(-1)^n}{u} du \\ &= \left[\frac{1}{n} u^n - \frac{n}{n-1} u^{n-1} + \frac{\binom{n}{2}}{n-2} u^{n-2} - \cdots \pm nu \pm \log(u) \right]_1^2 \\ &= \frac{z_n}{d_n} \pm \log(2), \end{aligned}$$

where z_n is an integer, by Lemma (2.3), as all but the last term will integrate to a sum of fractions with denominators less than n. The final term integrates to $\pm\log(2)$. Thus

$$J_n = \int_0^1 \frac{P_n(x)}{1+x}dx = \frac{A_n}{d_n} + B_n \log(2),$$

with A_n and B_n integers.

Now we aim to bound J_n from above. By (2.1),

$$\begin{aligned}
|J_n| &= \frac{1}{n!}\left|\int_0^1 x^n(1-x)^n \frac{d^n}{dx^n}\frac{1}{1+x}\right| \\
&= \left|\int_0^1 x^n(1-x)^n \frac{1}{(1+x)^{n+1}}dx\right| \\
&\le M^n \left|\int_0^1 \frac{1}{1+x}dx\right| \\
&= M^n \log(2),
\end{aligned}$$

where M is defined as the maximum of $x(1-x)/(1+x)$ for x between 0 and 1, which can be calculated to be $3-2\sqrt{2}$, or roughly 0.17. Additionally it is clear that the integrand is non-negative over all $0 < x < 1$, so $|J_n| > 0$.

Assuming $\log(2) = a/b$ is rational, and multiplying through by the denominators, we get the inequality:

$$\begin{aligned}
& 0 < |J_n| < M^n \log(2) \\
\implies & 0 < |\frac{A_n}{d_n} + B_n \log(2)| < M^n \log(2) \\
\implies & 0 < |bA_n + aB_n d_n| < M^n d_n a.
\end{aligned}$$

But from Lemma (2.2), we know that for large n, $d_n < e^n$, and so

$$0 < |bA_n + aB_n d_n| < (eM)^n a < (0.5)^n a.$$

This upper bound will be less than 1 for some n, and as $bA_n + aB_n d_n$ is the sum and product of integers, this is a contradiction. Thus, $\log(2)$ is irrational.

2.2 Irrationality of $\zeta(2)$

The following two proofs follow much of the structure of [15]. To prove the irrationality of $\zeta(2)$, we choose

$$f(x) = \int_0^1 \frac{(1-y)^n}{1-xy}dy,$$

for some natural number n. We want to show that we have rational A_n and B_n so that

$$\begin{aligned} I_n &:= \int_0^1 x^n \int_0^1 \frac{(1-y)^n}{1-xy} dydx \\ &= \int_0^1 \int_0^1 \frac{x^n(1-y)^n}{1-xy} dydx \\ &= A_n + B_n \zeta(2). \end{aligned}$$

By expanding $(1-y)^n$ and splitting up this integral, we can see it is sufficient to show that every

$$I_{r,s} := \int_0^1 \int_0^1 x^r y^s \frac{1}{1-xy} dydx$$

has the required form, where $0 \leq r, s \leq n$. We do this by considering three cases.

Note that, for $x = y = 1$, the integrand diverges. This means $I_{r,s}$ is an improper integral, and we should consider limits. I will show how this works for the first case.

Case 1: $\mathbf{r = s = 0}$

For all three cases, we use the formula for an infinite geometric series to write

$$\frac{1}{1-xy} = \sum_{i=0}^{\infty} (xy)^i = \sum_{i=0}^{\infty} x^i y^i.$$

In our integral, x and y run between 0 and 1, and so the formula is valid for all values of xy except $(x, y) = (1, 1)$.

$$\begin{aligned} I_{0,0} &= \lim_{\epsilon \to 0^+} \int_0^{1-\epsilon} \int_0^{1-\epsilon} \frac{1}{1-xy} dydx \\ &= \lim_{\epsilon \to 0^+} \int_0^{1-\epsilon} \int_0^{1-\epsilon} \sum_{i=0}^{\infty} x^i y^i dydx \\ &= \sum_{i=0}^{\infty} \lim_{\epsilon \to 0^+} \int_0^{1-\epsilon} \int_0^{1-\epsilon} x^i y^i dydx \\ &= \sum_{i=0}^{\infty} \lim_{\epsilon \to 0^+} \left[\frac{1}{(i+1)^2} (xy)^{i+1} \right]_{0,0}^{1-\epsilon, 1-\epsilon} \\ &= \sum_{i=0}^{\infty} \frac{1}{(i+1)^2} \\ &= \zeta(2). \end{aligned} \tag{2.2}$$

From now on, I will leave out the analytic details of the limits, which do not affect the results.

Case 2: r=s $\neq$ 0

Similar to above, we get:

$$\begin{aligned} I_{r,r} &= \int_0^1 \int_0^1 x^r y^r \sum_{i=0}^{\infty} x^i y^i dy dx \\ &= \int_0^1 \int_0^1 \sum_{i=r}^{\infty} x^i y^i dy dx \\ &= \sum_{i=r}^{\infty} \frac{1}{(i+1)^2} \\ &= \zeta(2) - \sum_{i=1}^{r} \frac{1}{i^2}. \end{aligned}$$

In fact, as every $1/i$ can be written as a fraction with denominator d_r, we have that

$$\sum_{i=1}^{r} \frac{1}{i^2} = \frac{z_r}{(d_r)^2}$$

for some integer z_r, and thus

$$I_{r,r} = \zeta(2) - \frac{z_r}{(d_r)^2}.$$

Case 3: r $\neq$ s

Finally, following the same method as above, we achieve that

$$I_{r,s} = \sum_{i=1}^{\infty} \frac{1}{(i+r)(i+s)}.$$

Using partial fractions, we can rewrite

$$\frac{1}{(i+r)(i+s)} = \frac{1}{r-s} \left(\frac{1}{i+s} - \frac{1}{i+r} \right)$$

and so, assuming without loss of generality that $r > s$,

$$\begin{aligned} I_{r,s} &= \frac{1}{r-s} \sum_{i=1}^{\infty} \left(\frac{1}{i+s} - \frac{1}{i+r} \right) \\ &= \frac{1}{r-s} \left(\frac{1}{s+1} + \frac{1}{s+2} + ... + \frac{1}{r} \right) \\ &= \frac{z_{r,s}}{(d_r)^2}, \end{aligned}$$

for some integer $z_{r,s}$, by Lemma (2.3) as all the denominators are clearly less than or equal to $r \leq n$.

From this, we can conclude that

$$J_n = \int_0^1 \int_0^1 P_n(x) \frac{(1-y)^n}{1-xy} dydx = \frac{a'_n}{d_n^2} + B'_n \zeta(2)$$

where a'_n and B'_n are both integers (not just rational)! In addition, we use (2.1) to find

$$\begin{aligned} |J_n| &= \frac{1}{n!} \left| \int_0^1 x^n (1-x)^n \frac{d^n}{dx^n} \left(\int_0^1 \frac{(1-y)^n}{1-xy} dy \right) dx \right| \\ &= \frac{1}{n!} \left| \int_0^1 x^n (1-x)^n (1-y)^n \int_0^1 \frac{\partial^n}{\partial x^n} \left(\frac{1}{1-xy} \right) dxdy \right| \\ &= \left| \int_0^1 \int_0^1 (x(1-x)y(1-y))^n \frac{1}{(1-xy)^{n+1}} dxdy \right| . \end{aligned}$$

This is in the required form. Let M be the maximum of $x(1-x)y(1-y)/(1-xy)$ for x and y between 0 and 1, which can be calculated with differentiation to be around 0.09. Thus by Equation 2.2

$$\begin{aligned} |J_n| &\leq M^n \left| \int_0^1 \int_0^1 \frac{1}{1-xy} dxdy \right| \\ &= M^n \zeta(2). \end{aligned}$$

In addition, the integrand is positive over the entire range, and zero at the boundaries, and thus $J_n > 0$. As I explained at the start of this chapter, we assume $\zeta(2)$ is a rational a/b, and multiplying through the denominators b and d_n^2, we obtain an inequality bounding an integer:

$$0 < |a'_n b + B'_n d_n^2 a| < d_n^2 a M^n.$$

Using Lemma 2.2, we see that,

$$d_n^2 a M^n < a(Me^2)^n < a \times 0.7^n < 1$$

for large enough n. Therefore $a'_n b + B'_n d_n^2 a$ is an impossible integer, and $\zeta(2)$ is irrational.

2.3 Irrationality of $\zeta(3)$

For $\zeta(3)$, we choose

$$f(x) = \int_0^1 \frac{P_n(y)}{1-xy} \log(xy) dy.$$

To show that J_n is the sum of rational multiples of 1 and $\zeta(3)$, we will establish that

$$I_{r,s} := -\int_0^1 \int_0^1 \frac{x^r y^s \log(xy)}{1-xy} dxdy$$

is in this form for all $0 \leq r, s \leq n$. As before, we consider three different cases:

Case 1: $r = s = 0$

We use the geometric sum formula again to find:

$$\begin{aligned} I_{0,0} &= -\int_0^1 \int_0^1 \frac{\log(xy)}{1-xy} dxdy \\ &= -\int_0^1 \int_0^1 (\log(x) + \log(y)) \sum_{i=0}^{\infty} x^i y^i dxdy \\ &= -2 \int_0^1 \int_0^1 \log(x) \left(\sum_{i=0}^{\infty} x^i y^i \right) dxdy \\ &= -2 \sum_{i=0}^{\infty} \left(\int_0^1 \log(x) x^i dx \int_0^1 y^i dy \right). \end{aligned}$$

Integrating $-\log(x)x^i$ by parts, and considering the limit as $x \to 0^+$, we achieve:

$$\begin{aligned} -\int_0^1 \log(x) x^i dx &= -\left[\frac{\log(x) x^{i+1}}{i+1} \right]_0^1 + \frac{1}{i+1} \int_0^1 x^i dx \\ &= \frac{1}{(i+1)^2}, \end{aligned}$$

and thus

$$I_{0,0} = 2 \sum_{i=0}^{\infty} \frac{1}{(i+1)^3} = 2\zeta(3).$$

Case 2: $r = s \neq 0$

We showed when proving the irrationality of $\zeta(2)$ that

$$\int_0^1 \int_0^1 \frac{x^r y^r}{1-xy} dxdy = \sum_{i=1}^{\infty} \frac{1}{(i+r)^2}.$$

Now we perform the simple substitution $r \longrightarrow r + t$, for t some non-negative real number, and differentiate with respect to t to get:

$$\begin{aligned} \int_0^1 \int_0^1 \frac{1}{1-xy} \frac{\partial}{\partial t} (xy)^{t+r} dxdy &= \frac{\partial}{\partial t} \int_0^1 \int_0^1 \frac{(xy)^{t+r}}{1-xy} dxdy \\ &= \sum_{i=1}^{\infty} \frac{d}{dt} \frac{1}{(i+r+t)^2}. \end{aligned}$$

Using that

$$\frac{d}{dt}(a^t) = a^t \log(a),$$

we obtain

$$\int_0^1\int_0^1 \frac{(xy)^{r+t}}{1-xy}\log(xy)dxdy = -2\sum_{i=1}^{\infty}\frac{1}{(i+r+t)^3}.$$

Now set t=0:

$$\begin{aligned} I_{r,r} &= -\int_0^1\int_0^1 \frac{(xy)^r}{1-xy}\log(xy)dxdy = 2\sum_{i=1}^{\infty}\frac{1}{(i+r)^3} \\ &= 2\left(\zeta(3) - \sum_{i=1}^{r}\frac{1}{i^3}\right) \\ &= 2\zeta(3) - \frac{z_r}{d_r^3}, \end{aligned}$$

for an integer z_r.

Case 3: $r \neq s$

We apply the same method as above. Now using

$$\int_0^1\int_0^1 \frac{x^r y^s}{1-xy}dxdy = \frac{1}{r-s}\left(\frac{1}{s+1} + \frac{1}{s+1} + \cdots + \frac{1}{r}\right),$$

we substitute $r \longrightarrow r+t$ and $s \longrightarrow s+t$, and differentiate with respect to t to achieve:

$$\begin{aligned} \int_0^1\int_0^1 \frac{x^{r+t}y^{s+t}\log(xy)}{1-xy}dxdy &= \frac{1}{r-s}\frac{d}{dt}\left(\frac{1}{s+t+1} + \frac{1}{s+t+2} + \cdots + \frac{1}{r+t}\right) \\ &= \frac{-1}{r-s}\left(\frac{1}{(s+t+1)^2} + \frac{1}{(s+t+2)^2} + \cdots + \frac{1}{(r+t)^2}\right). \end{aligned}$$

Finally, setting $t = 0$,

$$\begin{aligned} I_{r,s} &= -\int_0^1\int_0^1 \frac{x^r y^s\log(xy)}{1-xy}dxdy = \frac{1}{r-s}\left(\frac{1}{(s+1)^2} + \frac{1}{(s+2)^2} + \cdots + \frac{1}{r^2}\right) \\ &= \frac{z_{r,s}}{(d_r)^3} \end{aligned}$$

for some integer $z_{r,s}$.

This has showed that

$$\begin{aligned} J_n &:= \int_0^1 P_n(x)f(x)dx \\ &= \int_0^1\int_0^1 \frac{P_n(x)P_n(y)\log(xy)}{1-xy}dxdy \\ &= \frac{A_n}{d_n^3} + B_n\zeta(3), \end{aligned}$$

where A_n, B_n are integers. Now we wish to bound $|J_n|$ using Equation (2.1). After some tedious but elementary algebraic manipulation and calculus, we can achieve that

$$\begin{aligned}|J_n| &= \int_0^1\int_0^1\int_0^1 \frac{(x-x^2)^n(y-y^2)^n(z-z^2)}{[(1-(1-z)x)(1-yz)]^{n+1}}\\ &\leq M^n\int_0^1 \frac{1}{((1-(1-z)x)(1-yz)},\end{aligned}$$

where here M is the maximum of

$$\frac{(x-x^2)(y-y^2)(z-z^2)}{(1-(1-z)x}$$

for x, y, z all in the range [0,1], and can be evaluated to be $17-12\sqrt{2} < 0.03$. Then as before, we can assume $\zeta(3) = a/b$ and find

$$0 < |A_n d_n^3 b + B_n a| < d_n^3 a M^n \sim a(Me^3)^n < a \times 0.6^n.$$

This shows that $\zeta(3)$ is irrational.

2.4 Further Study

In [23], the authors describe $\zeta(5)$ as "surely irrational, in the everyday sense of the word *sure* (like death and taxes)." Thus, we might hope to extend the ideas in this proof to show $\zeta(s)$ is irrational for larger odd values of s. Following the same strategy as for $\zeta(3)$, to show $\zeta(5)$ is irrational, we would first want to show it for $\zeta(4)$. To do this, we search for a family of integrals such that $I_0 = \zeta(4)$, and I_n is sum of rational multiples of 1 and $\zeta(4)$. We can find:

$$\zeta(4) = \int_0^1\int_0^1\int_0^1\int_0^1 \frac{1-xy}{(1-(1-xy)w)(1-(1-xy)v)} dxdydwdv$$

However, optimistically defining

$$f(x) = \int_0^1\int_0^1\int_0^1 \frac{1-xy}{(1-(1-xy)w)(1-(1-xy)v)} dydwdv$$

does not work, because

$$J_n = \int_0^1 P_n(x)f(x)dx$$

cannot be bounded sufficiently well [13]. Huylebrock suggests in [14] that a good family of integrals to consider is

$$I_n = \int_0^1\int_0^1\int_0^1\int_0^1 \frac{(1-xy)^{2n+1}((1-x)x(1-y)y(1-w)w(1-v)v)^n}{((1-(1-xy)w)(1-(1-xy)v))^{n+1}} dxdydwdv,$$

because it is easy to bound

$$I_n \leq M^n \zeta(4)$$

where M is the maximum in $[0,1]^4$ of

$$M = \frac{(1-xy)^2(1-x)x(1-y)y(1-w)w(1-v)v}{(1-(1-xy)w)(1-(1-xy)v)}$$

If we had that for all natural numbers n,

$$I_n = \frac{A_n}{d_n^4} + B_n\zeta(4),$$

where A_n, B_n were integers for all n, this would be sufficient to show $\zeta(4)$ was irrational. However, unfortunately, this has only been shown to be true for $n = 0, 1, 2$, and so a proof of the irrationality of $\zeta(4)$ and then $\zeta(5)$ is still unavailable.

However, progress has been made on the odd zeta values. It has been shown that infinitely many $\zeta(2n+1)$ are irrational, and we won't need to go far to find the next irrational value, as Wadim Zudilin has shown that at least one of $\zeta(5), \zeta(7), \zeta(9)$ and $\zeta(11)$ is irrational [28]. This number of odd zeta values required for one to be irrational has been being lowered over the years, and it is conjectured that $\zeta(2n+1)$ is irrational for all n [3].

Chapter 3

Approximation of Real Numbers with Rationals

When we first defined irrational numbers, they were framed using negatives: they *cannot* be written as a fraction. This chapter will allow us to define irrational numbers positively, and then extend these ideas to discover transcendental numbers.

The problem we are considering is the approximation of a real number, called α, by rational numbers, p/q. Our problem is initially uninteresting because the rationals become infinitely close together as we "zoom in" (that is, they are dense in the reals). Thus for any $\epsilon > 0$, we can find infinite rational numbers such that

$$\left|\alpha - \frac{p}{q}\right| < \epsilon,$$

by simply choosing any $q > 1/\epsilon$ and $p = \lceil q\alpha \rceil - 1$. However, this method requires q to get very large as ϵ approaches 0, and so it is more interesting to consider approximations where we limit the size of q. For example, consider the inequality

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}. \tag{3.1}$$

3.1 The rational case

The next two sections follow Hardy and Wright [12], Chapter 11, however I have filled in some details and added a proof of Proposition 3.2.

Proposition 3.1. If α is a rational number, a/b, there are only finitely many solutions to Inequality 3.1.

Proof. Let $\alpha = a/b$, and assume that $p/q \neq \alpha$. Then,

$$\left|\alpha - \frac{p}{q}\right| = \left|\frac{a}{b} - \frac{p}{q}\right| = \left|\frac{aq - pb}{bq}\right| \geq \left|\frac{1}{bq}\right|, \tag{3.2}$$

since we have that $aq - bp$ is an integer, and cannot be equal to zero. This means that we require $q < b$, so there are finite solutions for q. Then, for any fixed q, we need

$$|\alpha q^2 - pq| < 1,$$

so the only potential integer solutions for p are $\lceil q^2\alpha \rceil / q$ and $\lfloor q^2\alpha \rfloor / q$. Thus, there can only be a finite number of solutions to Inequality (3.1). □

We can naturally generalise this.

Proposition 3.2. If $\alpha = a/b$ is a rational number, then 1 is the largest possible value of μ so that there are infinite solutions for p/q to

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{q^\mu}.$$

Proof. By Inequality (3.2), for q to be a solution we require

$$b \geq q^{\mu - 1}.$$

If $\mu = 1$, this simply equivalent to saying $b \geq 1$, so there is no restriction on the number of solutions. Otherwise we find

$$q < b^{1/(\mu-1)} \text{ if } \mu > 1,$$
$$q > b^{1/(\mu-1)} \text{ if } \mu < 1,$$

and, by the same argument as the previous proposition, for $\mu > 1$ there can only be finite solutions, whereas for $\mu < 1$, the number of possible values of q is not restricted. □

Remark 3.3 (Angell [2], Section 3.4)**.** Apéry's original proof that $\zeta(3)$ is irrational involved showing that there are infinitely many solutions to

$$\left| \zeta(3) - \frac{p}{q} \right| < \frac{1}{q^{1.03}}.$$

The idea of finding the maximum possible value of μ so that Inequality (3.2) has infinite solutions will reoccur later in this chapter to define concretely how well different numbers can be approximated by rationals.

We now move onto the more interesting case: what happens when α is irrational?

3.2 Dirichlet's Idea

Theorem 3.4 (Dirichlet)**.** If α is irrational, there are infinite solutions p/q to (3.1).

Proof. Let us assume α is irrational and suppose we are given a large integer, N. We define the fractional part of α, (α), as

$$(\alpha) = \lceil \alpha \rceil - \alpha.$$

Now consider the $N + 1$ numbers:

$$0, (\alpha), (2\alpha), ..., (N\alpha),$$

and the intervals:

$$\left[0, \frac{1}{N}\right), \left[\frac{1}{N}, \frac{2}{N}\right), ..., \left[\frac{N-1}{N}, 1\right).$$

By the pigeonhole principle, there must be at least one interval which contains two points. Note that we cannot have that $(i\alpha) = (j\alpha)$ for $i \neq j$, as this implies $i\alpha$ and $j\alpha$ differ by an integer. Then,

$$i\alpha - j\alpha = m,$$
$$\alpha = \frac{i-j}{m},$$

so α would be rational. Thus, for some integers i, j between 0 and N, with $i \neq j$, we have that

$$0 < (i\alpha) - (j\alpha) < \frac{1}{N}$$

Then we can choose $q = (i - j) \leq N$, and there is an integer p such that

$$|q\alpha - p| < \frac{1}{N},$$

which implies

$$\left|\frac{p}{q} - \alpha\right| < \frac{1}{qN} \leq \frac{1}{q^2}.$$

This almost proves that there are infinite solutions to Inequality (3.1). If there were a finite number of solutions, assume that $p_1/q_1, p_2/q_2, \ldots, p_k/q_k$ was a full list of solutions. However, because α is irrational, and thus none of the solutions p_s/q_s are equal to α, there exists a N such that

$$\left|\frac{p_s}{q_s} - \alpha\right| > \frac{1}{N} \quad \text{for } s = 1, ..., k.$$

However, from above we know we can find another solution p, q such that

$$\left|\frac{p}{q} - \alpha\right| < \frac{1}{N},$$

and so p/q is a solution not on the original list of solutions. □

Interestingly, although the idea of the pigeonhole principle has been around for millennia, Dirichlet was the first to give it a name when he used it in this proof [11].

3.2.1 Hurwitz's Improvement

Dirichlet's theorem was later improved by Hurwitz.

Theorem 3.5 (Hurwitz)**.** For any irrational α, if $c \leq \sqrt{5}$, there are infinite solutions to

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{cq^2} \tag{3.3}$$

However for any $c > \sqrt{5}$, there are only finitely many solutions.

This theorem can be proved by using continued fractions or Farey sequences. If $c < \sqrt{5}$, we can demonstrate that at least one of any three consecutive convergents to the continued fraction of α must satisfy Inequality 3.3. Then, since there are infinite convergents to α, this shows that there are infinite solutions to Inequality 3.3. See [2], pages 88-90, for a full proof with continued fractions.

The bound on c for infinite solutions is closely linked to the golden ratio, $\phi = (\sqrt{5}+1)/2$, which is significant for having a continued fraction form consisting only of 1s.

3.3 Liouville's Theorem

Liouville investigated numbers that can be approximated by rationals extremely well, and showed that all irrational algebraic have a limit on how well they can be approximated. Thus, by finding numbers that can be approximated by rationals infinitely closely, Liouville constructed the first numbers which were proven to be transcendental.

Theorem 3.6 (Liouville)**.** If α is a real irrational number which is algebraic of degree d, then for any rational p/q,

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{c(\alpha)}{q^d} \tag{3.4}$$

where $c(\alpha)$ is some real constant that depends only on α.

When α is algebraic of degree 2, this follows from Hurwitz's Theorem.

Proof (Liouville's Theorem). This proof is based off the one given in Burger and Tubbs [8], pages 15-17. We choose some p, q that satisfy Inequality (3.4) and let the minimal polynomial for α over $\mathbb{Q}$ be $F(x) = a_d x^d + ... + a_1 x + a_0$. Consider $F(p/q)$. Since $F(x)$

is irreducible over $\mathbb{Q}$, it has no rational roots, so $F(p/q) \neq 0$.

$$\begin{aligned}\left|F\left(\frac{p}{q}\right)\right| &= \left|a_d(\frac{p}{q})^d + ... + a_1(\frac{p}{q}) + a_0\right| \\ &= \left|\frac{a_d p^d + ... + a_1 p q^{d-1} + a_0 q^d}{q^d}\right| = \frac{N}{q^d} \geq \frac{1}{q^d},\end{aligned}$$

where N is some positive integer. This is clear as $a_d p^d + \cdots + a_0 q^d$ is a non-zero integer.

We now find the exact value of $|F(p/q)|$ by Taylor expanding about $x = \alpha$). The higher order terms vanish because $F^{(k)}(x) = 0$ for $k > d$.

$$\begin{aligned}\left|F\left(\frac{p}{q}\right)\right| &= \left|\sum_{i=0}^{d} \frac{1}{i!} F^{(i)}(\alpha)\left(\alpha - \frac{p}{q}\right)^i\right| \\ &= \left|(\alpha - \frac{p}{q}) \sum_{i=1}^{d} \frac{1}{i!} F^{(i)}(\alpha)\left(\alpha - \frac{p}{q}\right)^{i-1}\right|,\end{aligned}$$

using that $F(\alpha) = 0$. Assume $|\alpha - p/q| \leq 1$. If not, then putting $c(\alpha) = 1$ is enough to satisfy the inequality in Liouville's Theorem. Then, by the triangle inequality,

$$\begin{aligned}\left|\sum_{i=1}^{d} \frac{1}{i!} F^{(i)}(\alpha)\left(\alpha - \frac{p}{q}\right)^{i-1}\right| &\leq \sum_{i=1}^{d} \frac{1}{i!}\left|F^{(i)}(\alpha)\left(\alpha - \frac{p}{q}\right)^{i-1}\right| \\ &\leq d \times M(\alpha),\end{aligned}$$

where we set $M(\alpha)$ to be the maximum value of $\left|F^{(i)}(\alpha)/i!\right|$ for i between 1 and d. $M(\alpha)$ cannot be zero, as $F(x)$ is not a constant. This now implies that, for all $|\alpha - p/q| \leq 1$,

$$\begin{aligned}\left|F\left(\frac{p}{q}\right)\right| = \left|\alpha - \frac{p}{q}\right| &\left|\sum_{i=1}^{d} \frac{1}{i!} F^{(i)}(\alpha)\left(\alpha - \frac{p}{q}\right)^{i-1}\right| = \frac{N}{q^d} \\ &\implies \frac{N}{q^d} \leq d \times M(\alpha)\left|\alpha - \frac{p}{q}\right| \\ &\implies \left|\alpha - \frac{p}{q}\right| \geq \frac{N}{d \times M(\alpha) q^d}.\end{aligned}$$

By setting $c(\alpha)$ to be the minimum of 1 and $N/(d \times M(\alpha))$, we get the required result. □

3.4 Thue, Siegel and Roth

The following two sections use results and definitions from Sally and Sally [24], and I will summarise the explanation of Thue's theorem given there. We can rewrite Liouville's Theorem to more closely resemble the structure of Dirichlet's Theorem, and to illuminate the following section.

Corollary 3.7. If α is a real algebraic number with degree d and $\delta > 0$, then

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{d+\delta}}$$

has only finitely many rationals solutions for p/q.

Proof. Assume there are infinite solutions p/q. As before, for any fixed q, p will be bounded, so there must be infinite solutions for q, and so q is unbounded. Thus for any $\delta > 0$, we can find q such that

$$q^\delta > \frac{1}{c(\alpha)}.$$

Then for this q,

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^d q^\delta} < \frac{c(\alpha)}{q^\delta},$$

contradicting Liouville's Theorem. □

Definition 3.8. For any $\alpha \in \mathbb{R}$, define the approximation exponent $\mu(\alpha) = \mu$ as the smallest possible real number so that for any $\delta > 0$

$$\left|\frac{p}{q} - \alpha\right| \leq \frac{1}{q^{\mu+\delta}}$$

has only finitely many solutions.

The approximation exponent measures how well a number can be approximated by rationals. Using our new notation, we have

- If α is a rational number, $\mu(\alpha) = 1$, by Proposition 3.2.
- If α is an irrational number, by Dirichlet's theorem, $\mu(\alpha) \geq 2$.
- If α is an algebraic number of degree d, by Liouville's theorem, $\mu(\alpha) \leq d$.
- Thus if α is a quadratic irrational number, $\mu(\alpha) = 2$.

The obvious question to ask is if we can find the approximation exponent for non-quadratic irrational algebraic numbers. This question challenged mathematicians throughout the first half of the twentieth century, as various people worked to reduce the upper bound on the approximation exponent for algebraic numbers. I have converted the stages of their progress into the table on the next page.

Mathematician	Year	Upper Bound on $\mu(\alpha)$
Liouville	1844	d
Thue	1909	$d/2+1$
Siegel	1921	$2\sqrt{d}$
Dyson	1947	$\sqrt{2d}$
Roth	1955	2

Table 3.1: Upper Bound of the Approximation Exponent of Irrational Algebraic Numbers of Degree d

The final discovery, Roth's Theorem, resulted in Roth winning the Field's Medal. Combined with Dirichlet's Theorem, this tells us that the approximation exponent of irrational algebraic numbers is completely independent of their degree - it is always 2!

Roth's Theorem is sometimes also called the Thue-Seigel-Roth Theorem to acknowledge the work of Thue and Seigel. Thue first improved Liouville's bound using auxiliary polynomials.

In the proof of Liouville's Theorem, we used a polynomial with a zero at α of order one, so we might attempt to improve on Liouville's result by using an auxiliary polynomial with a zero at α of order $h > 1$. Suppose $G(x)$ is such a polynomial with degree $r \geq h$, and follow the same strategy as in Liouville's Theorem. We have that

$$\left|G\left(\frac{p}{q}\right)\right| = \frac{N}{q^r}$$

where again p, q are approximations to α and N is a positive integer. Then, using that $G^{(i)}(\alpha) = 0$ for all $i < h$ and Taylor expanding $G(p/q)$ about $x = \alpha$, we can get that

$$\begin{aligned}\left|G\left(\frac{p}{q}\right)\right| &= \left|\left(\alpha - \frac{p}{q}\right)^h \sum_{i=h}^{r} \frac{1}{i!} G^{(i)}(\alpha) \left(\alpha - \frac{p}{q}\right)^{i-h}\right| \\ &\leq M(\alpha) \left|\alpha - \frac{p}{q}\right|^h\end{aligned}$$

so long as $|\alpha - p/q| < 1$. Putting this together, we can achieve the bound

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{c(\alpha)}{q^{r/h}},$$
$$\text{where } c(\alpha) = \left(\frac{N}{M(\alpha)}\right)^{1/h}.$$

Although this initially looks like an improvement on Liouville's bound, in fact $r/h \geq d$, where d is still the degree of α. This is because, as α is a zero of $G(x)$ of order h, $F(x)^h$ must divide $G(x)$, where $F(x)$ is the minimal polynomial of α. Therefore $r \geq hd$ and it is not possible to improve on Liouville's Theorem by taking an auxiliary polynomial in one variable with α as a root of order greater than 1. We need a new idea for auxiliary polynomials.

Thue's brilliant idea was to consider an auxiliary polynomials in 2 variables, $F(x, y)$, where $F(x, y)$ had integer coefficients and was in the form

$$F(x, y) = P(x) + yQ(x).$$

We also demand that $F(x, \alpha)$ has a zero of high order h at $x = \alpha$. It is not clear that it is possible to find such a polynomial, but we aim to find a condition on d so that it is.

Suppose that the overall degree of $F(x, y)$ is n, which implies that $\deg(P(x)) \leq n$ and $\deg(Q(x)) \leq n - 1$, so there are at most $2n + 1$ non-zero coefficients, which we label as $\{c_1, c_2, \dots, c_{2n+1}\}$.

If $F(x, \alpha)$ has a zero of order h at $x = \alpha$, then we have the system of h linear equations, where the variables are the c_i, given by:

$$F(\alpha, \alpha) = 0, \frac{\partial F}{\partial x}(\alpha, \alpha) = 0, \dots, \frac{\partial^{(h-1)} F}{\partial x^{h-1}}(\alpha, \alpha) = 0.$$

Each linear equation has coefficients in $\mathbb{Q}[\alpha]$, so, because α is algebraic of degree d, we can rewrite the system as hd linear equations with rational coefficients.

If there are fewer linear equations than variables, then we will have a non-zero solution. That is, if $hd < 2n + 1$, then we can find a non-zero polynomial $F(x, y)$ with coefficients $c_1, c_2, \dots, c_{2n+1}$ so $F(x, y)$ has the required properties. Rearranging $hd < 2n + 1$, we can see that $d/2 < r/h + 1/2h$ is already reminiscent of Thue's bound $\mu(\alpha) < d/2 + 1$.

Once we have the auxiliary polynomial $F(x, y)$, the rest of Thue's proof uses similar ideas to Liouville, although the actual details are much more complex. It requires taking approximations p_1/q_1 and p_2/q_2 to α and showing that

$$\frac{\partial^j F}{\partial x^j}\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) \neq 0,$$

for some j which depends only on α. Then by bounding this object from above in terms of $|\alpha - p_1/q_1|$ and $|\alpha - p_2/q_2|$, it is possible to create a contradiction.

The proof of Roth's Theorem is still more complicated. It involves constructing polynomials in m variables, where m is dependent on δ, as in Definition 3.8.

3.5 Approximation Exponent of Transcendental Numbers

3.5.1 Liouville Numbers

Definition 3.9 (Liouville Numbers)**.** A Liouville number is a (necessarily transcendental) number α, such that $\mu(\alpha) = \infty$.

If $\mu(\alpha) = \infty$, then α is clearly transcendental by Liouville theorem, as α cannot be the root of any polynomial of degree $d < \infty$. To our disappointment, not all transcendental numbers are Liouville numbers; in fact, most transcendental numbers, including our favourites like e and π are not Liouville numbers.

Nevertheless, this gives us a way to construct our first transcendental numbers! We require numbers that can be approximated by rationals extremely well. I found the following Liouville number by tweaking the canonical example and proof given in [8], Section 1.3.

Example 3.10. Consider:

$$\alpha = \sum_{n=1}^{\infty} 10^{-n^n}.$$

We can approximate α by truncating the series after N terms. This is a rational number where the denominator, $q(N)$, is 10^{N^N}. Then,

$$\begin{aligned}\left|\alpha - \sum_{n=1}^{N} 10^{-n^n}\right| = \sum_{n=N+1}^{\infty} 10^{-n^n} &= \frac{1}{10^{(N+1)^{(N+1)}}} + \frac{1}{10^{(N+2)^{(N+2)}}} + ... \\ &< \frac{1}{10^{(N+1)^{(N+1)}}} + \frac{1}{10^{(N+1)^{(N+2)}}} + ... \\ &< \frac{1}{10^{(N+1)^{(N+1)}}} \times \frac{10^{N+1}}{10^{N+1}-1}\end{aligned}$$

We have that the constant $10^{N+1}/(10^{N+1} - 1)$ is largest when N takes its smallest possible value, i.e $N = 1$. In addition, we use that $(N + 1)^{N+1} > N^{N+1}$ for positive real N. Thus we obtain

$$\begin{aligned}\left|\alpha - \sum_{n=1}^{N} 10^{-n^n}\right| &< \frac{100}{99} \times \frac{1}{10^{N^{N+1}}} \\ &= \frac{100}{99} \times \frac{1}{q(N)^N}.\end{aligned}$$

This is enough to show that α is transcendental and a Liouville number. Considering Liouville's theorem, for any $d, c(\alpha)) > 0$, let us take N large enough that $q(N)^{N-d} > 100/(99 \times c(\alpha))$. Then:

$$\left|\alpha - \sum_{n=1}^{N} 10^{-n^n}\right| < \frac{c(\alpha)}{q(N)^d}.$$

Thus $\mu(\alpha) = \infty$.

From the example, it is clear that there are infinite Liouville numbers. This follows from writing any real number in its (base 10) decimal form as $a_0 + 0.a_1a_2a_3\ldots$, where

a_0 is an integer and for $i \geq 1$, a_i is an integer between 0 and 9. Then

$$\sum_{n=1}^{\infty} a_n 10^{-n^n}$$

must also be a Liouville number.

3.5.2 Other Transcendental Numbers

It has been shown that, in terms of measure theory, almost all real numbers have approximation exponent equal to 2 ([24], section 10). Most specific transcendental numbers we know of have approximation exponent 2 or their approximation exponent is not yet known; Liouville numbers are an exception, as is Mahler's number, which we look at next chapter. This section is based on Borwein and Borwein [6].

Approximation Exponent for e

We know $\mu(e) = 2$ due to e having a very nice continued fraction representation.

Continued fractions are closely related to rational approximation, because the convergents - that is, the rational numbers formed by terminating the continued fraction before a plus sign - form a series of rational approximations. For example, the first few convergents to e are $2, 3, 8/3, 11/4$ and $19/7$.

Bounding the Approximation Exponent with Beukers Integrals

Unfortunately, the continued fraction of π is less easy to work with, and we are unable to use it to find $\mu(\pi)$ exactly. However we can find a bound on $\mu(\zeta(2))$ and then $\mu(\pi)$ by using Beukers' integrals from the previous chapter. I will briefly sketch out the ideas now.

Theorem 3.11. Assume we have an infinite sequence of approximations $\{p_n/q_n\}$, which satisfy the following conditions.

1. For some $\delta > 0$,
$$\left| \alpha - \frac{p_n}{q_n} \right| < \frac{1}{q_n^{1+\delta}}. \tag{3.5}$$

2. The denominators q_n form a monotonically increasing sequence such that
$$q_{n+1} = q_n^{1+\gamma_n}$$
where γ_n is always positive and tends to 0 as n tends to infinity.

Then for sufficiently large q, either

$$\frac{p}{q} = \frac{p_n}{q_n}$$

for some n, or there exists an $\epsilon > 0$ such that

$$\left|\alpha - \frac{p}{q}\right| > \frac{1}{q^{1+1/\delta+\epsilon}}.$$

I will not prove this here, but the proof is relatively simple: it involves choosing a q_n so that

$$\frac{1}{2}q_n^\delta \leq q < \frac{1}{2}q_{n+1}^\delta,$$

which is possible as q_n must be unbounded, and using the triangle inequality (see [6]).

Our aim is use the Beukers' Integrals to find a sequence of approximations to $\zeta(2)$ which satisfy the conditions in Theorem 3.11. Recall that we have

$$0 < |J_n| = \left|\int_0^1 \int_0^1 P_n(x)\frac{(1-y)^n}{1-xy}dydx\right| = \left|B_n\zeta(2) - \frac{a_n}{d_n^2}\right| < M^n\zeta(2),$$

where B_n and a_n are integers and I have simplified the notation from the previous chapter to remove dashes. We now define $c_n := {d_n}^2 B_n$ and divide through B_n to achieve

$$\left|\zeta(2) - \frac{a_n}{c_n}\right| < \frac{M^n\zeta(2){d_n}^2}{c_n} < \frac{1}{{c_n}^{1+\delta}},$$

where

$$\delta = \frac{5\log(1+\sqrt{5}) - \log M}{2 + 5\log(1+\sqrt{5}) - 5\log(2)} - 1 \approx 0.092.$$

While the value of δ looks very arbitrary, it can be derived by finding the exact value of B_n. Then one can use Stirling's formula and Lemma 2.2 to bound $M^n\zeta(2)d_n^2$ in terms of c_n and show that c_n satisfies both conditions of Lemma 3.11.

Therefore, for sufficiently large q, if

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{1+1/\delta+\epsilon}}, \tag{3.6}$$

we must have that $p = a_n$ and $q = c_n$.

However, for large n, it is possible to achieve that

$$\left|\zeta(2) - \frac{a_n}{\gamma_n}\right| \geq \frac{1}{c_n^3} > \frac{1}{c_n^{1+1/\delta}},$$

and so for some maximum q^*, there are no values $q > q^*$ satisfying Inequality 3.6. Thus must only be a finite number of solutions and we can bound

$$\mu(\zeta(2)) \leq 1 + 1/\delta < 11.86.$$

From this, we can immediately find a bound for $\mu(\pi)$. Because $\zeta(2) = \pi^2/6$, we have, for sufficiently large q,

$$\begin{aligned} \left|\pi - \frac{p}{q}\right| &= \frac{1}{|\pi^2 + p/q|}\left|\pi - \frac{p^2}{q^2}\right| \\ &> \frac{1}{|\pi^2 + p/q|} \times \frac{1}{q^{2\times 11.86...}} \\ &= \frac{c(\pi)}{q^{23.72}}. \end{aligned}$$

Using the argument from the proof of Corollary 3.7, this must imply that

$$\mu(\pi) < 23.72.$$

In fact, mathematicians have been steadily decreasing the upper bound on $\mu(\pi)$ over the past century. Karl Mahler was the first to bound $\mu(\pi)$ from above, showing that $\mu(\pi) < 42$ in 1953 (see [16]). He was also the first to construct a non-Liouville transcendental number, which we will look at in the next chapter.

The most recent bound is $\mu(\pi) \leq 7.10320534$, due to Zudilin and Zeilberger in 2020 (see [29]) It is naturally conjectured that $\mu(\pi) = 2$, and there is numerical evidence to support this [9].

Chapter 4

Mahler's Number

We now investigate a specific constant with approximation exponent between 2 and infinity.

Definition 4.1. Mahler's number in base 10 is $\mathcal{M} = 0.12345678910111213...$, the number formed by writing all the natural numbers in base 10 in the standard order, after the decimal point.

This number is more often referred to as Champernowne's constant after David Champernowne, who introduced $\mathcal{M}$ in 1933 and showed it was normal in base 10; that is, each digit appears with equal frequency in the decimal expansion of $\mathcal{M}$ (see [22]). However, I choose to call it Mahler's number as Mahler showed that $\mathcal{M}$ was transcendental and investigated its approximation exponent [17].

We will explore two approaches to find rational approximations to $\mathcal{M}$, and most of the ideas will generalise to a class of numbers, $\mathcal{M}(g)$.

4.1 Some Naive Approximations

This section follows Burger and Tubbs [8], Section 1.6. We aim to find an infinite series of rational approximations p_n/q_n to $\mathcal{M}$. Our first approach is to consider the rational numbers $\mathcal{M}_k$, formed from writing the numbers with k digits in base 10 in order after the decimal point. For example,

$$\begin{aligned}\mathcal{M}_1 &= 0.123456789,\\ \mathcal{M}_2 &= 0.1011121314....9899,\\ \mathcal{M}_3 &= 0.100101102...999.\end{aligned}$$

These are good approximations to the fractional parts of $\mathcal{M}, 10^9\mathcal{M}, 10^{189}\mathcal{M}$ respectively; for instance, we see that

$$|\mathcal{M} - \mathcal{M}_1| < \frac{1}{10^9}.$$

The most obvious idea is to put $p_n/q_n = \mathcal{M}_n$. Unfortunately, the denominator of $\mathcal{M}_n$ grows too large; because there are $9n \times 10^{n-1}$ digits in $\mathcal{M}_n$, q_n is on the order of 10^{10^n}. If we let

$$\mathcal{M}_1 = \frac{p_1}{q_1} = \frac{123456789}{10^9},$$

we only get

$$\left|\mathcal{M} - \frac{p_1}{q_1}\right| < \frac{1}{q_1},$$

and these approximations are not even good enough to establish that $\mathcal{M}$ is irrational.

Instead, we will first approximate $\mathcal{M}_k$ by rationals with much smaller denominators. We achieve this by noticing that $\mathcal{M}_k$ is formed of k character long strings of digit, each formed by adding 1 to the previous one. Therefore $(10^k - 1)\mathcal{M}_k$ is very closely approximated by a decimal of period k, and thus by a fraction with denominator $10^k - 1$, which is a lot smaller than our previous q_n!

This approach will give us an infinite series of approximations that satisfy

$$\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q^{4.5}}. \tag{4.1}$$

Example 4.2. Lets see how this works for $\mathcal{M}_1$. We have that

$$10\mathcal{M}_1 - \mathcal{M}_1 = 1.11111101 \approx \frac{10}{9}$$
$$\left|9\mathcal{M}_1 - \frac{10}{9}\right| < 0.000000101$$
$$\left|\mathcal{M}_1 - \frac{10}{81}\right| < \frac{1}{81^{4.5}}$$

Since $\mathcal{M}$ and $\mathcal{M}_1$ agree to the first 9 decimal places, we also have that

$$\left|\mathcal{M} - \frac{10}{81}\right| < \frac{1}{81^{4.5}}$$

and we can take $p_1 = 10, q_1 = 81$.

We are able to repeat the same trick for any $\mathcal{M}_n$ to get an approximation p_n/q_n which satisfies inequality 4.1. Thus this proves $\mu(\mathcal{M}) \geq 4.5$, and by Roth's Theorem, this means $\mathcal{M}$ must be transcendental. However it is clear that 4.5 is a fairly arbitrary constant, and we are able to find better approximations.

4.2 Amou's Approximations

In 1976, Mahler found a series expansion for $\mathcal{M}$ in [17] and used this to show that $\mu(\mathcal{M}) \leq 200/9$, and then in 1989, Masaaki Amou built on Mahler's work to find the exact value for $\mu(\mathcal{M})$ [1]. This section follows Amou's proof, however I have restructured it for clarity and proved several statements which he does not. The series expansion for $\mathcal{M}$, which is also used in Amou's proof, is

$$\mathcal{M} = \frac{10}{9^2} - \sum_{n=1}^{\infty} u_n 10^{-E_n},$$

where:

$$u_n = \frac{10^{2n} - 10^n + 1}{(10^n - 1)^2} - \frac{10^{2n+1} - 10^n + 1}{(10^{n+1} - 1)^2},$$
$$E_n = 9\sum_{k=1}^{n} k10^{k-1} = 9(1 + 20 + 300 + ... + n10^{n-1}).$$

This series expansion has links to the numbers $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3, \ldots$ defined in the previous section: we saw that 10/81 closely approximated $\mathcal{M}_1$ and $\mathcal{M}$. The full derivation is tedious but elementary; it is explained step-by-step in [22].

As in the proof of the irrationality of e in Example 1.6, we find rational approximations to $\mathcal{M}$ by truncating its series expansion. We define:

$$D_n = \prod_{k=1}^{n}(10^k - 1)^2 = 9^2 \times 99^2 \times (10^n - 1)^2,$$
$$B_n = D_{n+1}10^{E_n},$$
$$A_n = B_n\left(\frac{10}{9^2} - \sum_{k=1}^{n} u_k 10^{-E_k}\right),$$
$$R_n = \sum_{k=n+1}^{\infty} u_k 10^{-E_k}.$$

By these definitions, we have that

$$\left|\mathcal{M} - \frac{A_n}{B_n}\right| = R_n.$$

Note that A_n, B_n are integers. For B_n, this is clear. For A_n, when $k \leq n$, the factors of $(10^k - 1)^2$ and $(10^{k+1} - 1)^2$ in the denominator of u_k are cancelled out by D_{n+1}, and the powers 10^{-E_k} become non-negative powers when multiplied by 10^{E_n}.

However, A_n and B_n are **not** necessarily coprime. We choose rational approximations to $\mathcal{M}$ by putting

$$\frac{A_n}{B_n} = \frac{p_n}{q_n}$$

where now p_n, q_n are coprime.

4.2.1 Asymptotic Behaviour

We first study the asymptotic behaviour of our various sequences as n tends to infinity.

Theorem 4.3. We have the bound

$$\gcd(A_n, B_n) \leq D_{n+1} 10^n.$$

Proof. From the definitions of A_n and B_n, we have that

$$\begin{aligned} \frac{A_n}{B_n} &= \frac{10}{9^2} - \sum_{k=1}^{n} u_k 10^{-E_k} \\ &= \frac{A_{n-1}}{B_{n-1}} - \frac{u_n}{10^{E_n}}. \end{aligned}$$

Multiplying through by B_n and substituting in the definition of B_n, we achieve

$$A_n = A_{n-1}(10^{n+1} - 1)^2 10^{E_n - E_{n-1}} - u_n D_{n+1}.$$

The first term is a multiple of 10^{n+1}, since

$$E_n - E_{n-1} = 9n10^{n-1} > n + 1.$$

We write the second term as

$$\begin{aligned} u_n D_{n+1} &= \left(\frac{10^{2n} - 10^n + 1}{(10^n - 1)^2} - \frac{10^{2n+1} - 10^n + 1}{(10^{n+1} - 1)^2} \right) \prod_{k=1}^{n} (10^k - 1)^2 \\ &= \left((10^{2n} - 10^n + 1)(10^{n+1} - 1)^2 - (10^n - 1)^2 (10^{2n+1} - 10^n + 1) \right) D_{n-1}. \end{aligned}$$

We can see that if we expanded the brackets, the unit terms would cancel out, so the lowest power of 10 left would be 10^n. Since D_n is a product containing no factors of 10, this means that the second term is a multiple of 10^n but not 10^{n+1}.

Thus A_n is likewise divisible by 10^n but not 10^{n+1}. Finally, we compute

$$\gcd(A_n, B_n) = \gcd(A_n, D_{n+1} 10^{E_n}) \leq D_{n+1} 10^n.$$

□

Corollary 4.4. We can write

$$q_n = 10^{E_n(1+\alpha_n)},$$

where $\alpha_n \to 0$ for $n \to \infty$.

Proof. From the definition of q_n, we have that

$$q_n = \frac{B_n}{gcd(A_n, B_n)} \geq \frac{D_{n+1} 10^{E_n}}{D_{n+1} 10^n} = 10^{E_n - n}.$$

We also have the obvious bound that $q_n \leq B_n$. Putting the bounds together, we have that

$$10^{E_n - n} \leq q_n \leq D_{n+1} 10^{E_n} < 10^{E_n + n(n+1)},$$
$$q_n = 10^{E_n(1+\alpha_n)},$$

where α_n is some number that between $-n/E_n$ and $n(n+1)/E_n$. By the squeezing theorem, $\alpha_n \to 0$ as $n \to \infty$, because E_n grows much faster than n. □

Theorem 4.5. For large n, we can write

$$R_n = 10^{-E_{n+1}(1+\beta_n)},$$

where $\beta_n \to 0$ from above.

Proof. As n tends to infinity, a power of 10^{2n} will dominate a power of 10^n, therefore

$$\lim_{n\to\infty} u_n = \frac{10^{2n}}{10^{2n}} - \frac{10^{2n+1}}{10^{2n+2}} = 1 - \frac{1}{10} = \frac{9}{10}.$$

Using this,

$$\lim_{n\to\infty} R_n = \sum_{k=n+1}^{\infty} \frac{9}{10} 10^{-E_k}$$
$$= \frac{9}{10} 10^{-E_{n+1}}$$
$$= 10^{-E_{n+1}(1+\beta_n)},$$

since the higher terms in R_n will exponentially damped. We need define $\beta(n)$ by:

$$10^{-E_{n+1}\beta_n} = \frac{9}{10}.$$

The exponent $-E_{n+1}\beta_n$ must be almost zero (to be precise, $\log(9/10) = 0.0045...$). Since E_{n+1} is large, we require β_n to be positive but small. Then as E_n clearly tends to infinity as $n \to \infty$, we must have that $\beta_n \to 0$. □

Theorem 4.6. As n goes to infinity, E_{n+1}/E_n tends to 10.

Proof. We first evaluate E_n. We have

$$E_n = \sum_{k=0}^{n} k 10^{k-1} = \sum_{k=0}^{n-1} (k+1) 10^k.$$

Thus, using the finite geometric sum formula, we achieve

$$\begin{aligned}
10E_n - E_n &= \sum_{k=0}^{n} k10^k - \sum_{k=0}^{n-1}(k+1)10^k \\
&= n10^n - \sum_{k=0}^{n-1} 10^k \\
&= n10^n - \frac{10^n - 1}{10 - 1} \\
\implies E_n &= \frac{9n10^n - 10^n + 1}{81}.
\end{aligned}$$

Therefore, the fraction E_{n+1}/E_n can be rewritten as

$$\begin{aligned}
\frac{E_{n+1}}{E_n} = \frac{\sum_{k=1}^{n+1} k10^{k-1}}{\sum_{k=1}^{n} k10^{k-1}} &= 1 + \frac{(n+1)10^n}{\sum_{k=1}^{n} k10^{k-1}} \\
&= 1 + \frac{81(n+1)10^n}{9n10^n - 10^n + 1}.
\end{aligned}$$

In the limit:

$$\begin{aligned}
\lim_{n\to\infty} \frac{E_{n+1}}{E_n} &= 1 + \lim_{n\to\infty} \frac{81(n+1)10^n}{9n10^n} \\
&= 1 + 9 \lim_{n\to\infty} \frac{n+1}{n} = 10.
\end{aligned}$$

When taking limits, we reduced the size of the denominator, so E_{n+1}/E_n approaches 10 from above: that is for any finite value of n, $E_{n+1}/E_n > 10$. □

4.2.2 Approximation Exponent

Now we can combine everything to get a bound on the approximation exponent!

$$|q_n\mathcal{M} - p_n| = R_n q_n$$

For large n, we substitute in our limits for R_n, q_n, and the ratio of adjacent terms E_n to get that

$$\begin{aligned}
|q_n\mathcal{M} - p_n| &= 10^{E_n(1+\alpha_n)} \times 10^{-E_{n+1}(1+\beta_n)} \\
&\leq 10^{E_n(1+\alpha_n)} \times 10^{-10E_n} \\
&= 10^{-9E_n} \times 10^{E_n\alpha_n} \\
&\leq q_n^{-9}.
\end{aligned}$$

Now we simply divide the inequality by q_n to get for large n, we have that:

$$\left|\mathcal{M} - \frac{p_n}{q_n}\right| \leq q_n^{-10},$$

or equivalently that there are infinite solutions to

$$\left|\mathcal{M} - \frac{p}{q}\right| < \frac{1}{q^{10}},$$

which immediately tells us that $\mu(\mathcal{M}) \geq 10$.

Compared to the previous section, this bound is clearly less arbitrary.

Theorem 4.7. For any $\delta > 0$,

$$\left|\mathcal{M} - \frac{p}{q}\right| < \frac{1}{q^{10+\delta}} \tag{4.2}$$

has only finite solutions, and thus $\mu(\mathcal{M}) = 10$.

Proof. This is clear from the asymptotic behaviour of the functions we defined. Because α_n and β_n tend to zero and E_{n+1}/E_n tends to 10 as n tends to infinity, we can modify the proof of the previous theorem to see

$$\left|\mathcal{M} - \frac{p_n}{q_n}\right| \sim \frac{1}{q_n^{10}}.$$

Thus for any $\delta > 0$, only finitely many p_n/q_n can satisfy 4.2. □

4.3 Generalisations

Mahler's constant can be generalised to different bases [1].

Definition 4.8. For $g \geq 2$ a positive integer, $M(g)$ is the number in base g formed by writing all the natural numbers in base g in the standard order after the decimal point.

Example 4.9.

- Clearly, $\mathcal{M} = \mathcal{M}(10)$.
- For $g = 3$, we have $\mathcal{M}(3) = 0.12101112202122..._3$.
  We can also write $\mathcal{M}(3)$ in base 10, although it loses the clear pattern:
  $\mathcal{M}(3) = 0.5989581675384...$

The series expansion and subsequent approximations and analysis can easily to generalised to $\mathcal{M}(g)$ for $g \geq 3$ by simply replacing any incidents of 10 with g, and 9 with $g - 1$. This leads to

$$\mu(\mathcal{M}(g)) = g,$$

for all $g \geq 3$, which implies the following theorem.

Theorem 4.10. For any positive integer n, we have a real number α so that $\mu(\alpha) = n$.

For $g = 2$, slightly more work has to be done, however it is possible to show that $\mu(\mathcal{M}(2)) = 2$, and additionally that $\mathcal{M}(2)$ is transcendental.

Chapter 5

Hermite's Method for Transcendence

Although $\mathcal{M}$ is a very interesting number, it is not naturally occurring; we do not find it in other formulae. Similarly, while Liouville showed that a class of numbers are transcendental, he had to construct his examples to have the required properties. The first naturally occurring number to be proved to be transcendental was e, by Hermite in 1873, and in 1882, Lindemann used Hermite's method to show π was also transcendental.

These proofs both use auxiliary polynomials to create a contradiction, and the following ingenious method to demonstrate an integer is non-zero.

- Assume we have a number written in the form
$$C_p = c_{0,p} + \sum_{i=1}^{n} c_{i,p},$$
where p is a large prime which we define later, and $c_{i,p}$ are all integers.
- We show that the sum of $c_{i,p}$ for i between 1 and n is always divisible by p.
- We next show that $c_{0,p}$ can be written as a product of factors, where none of the factors are dependent on p. Thus by choosing p to be larger than all of the factors of $c_{0,p}$, we ensure that $c_{0,p}$ cannot be divisible by p.
- Therefore C_p is the sum of two numbers where only one is divisible by p and therefore p cannot divide C_p. Then, since 0 is divisible by p, this means C cannot be zero!

5.1 Transcendence of e^r

Theorem 5.1 (Hermite)**.** Euler's constant, e, is transcendental.

Because $\mu(e) = 2$, we cannot use the approximation exponent to show that e is transcendental.

Proof. This proof is based on the one given by [26], although I restructured it so we get the standard contradiction. We start by assuming e is algebraic, that is, it is the root of a polynomial with integer coefficients, a_t:

$$\sum_{t=0}^{n} a_t e^t = 0, \quad \text{with} \quad n \geq 1, \;\; a_0, a_n \neq 0. \tag{5.1}$$

We approximate e^t for $t \geq 1$ by rationals in the following way:

$$e^t = \frac{M_t + \epsilon_t}{M},$$

where $M, M_1, M_2, ..., M_n$ are integers, chosen so that the correction term, ϵ_t/M, is small. We can then substitute into Equation (5.1) to get

$$a_0 + \sum_{t=1}^{n} a_t \left(\frac{M_t + \epsilon_t}{M} \right) = 0$$

$$\implies M a_0 + \sum_{t=1}^{n} a_t M_t = - \sum_{t=1}^{n} a_t \epsilon_t. \tag{5.2}$$

We define M, M_t, and ϵ_t as follows:

$$M = \int_0^\infty x^{p-1} e^{-x} \frac{((x-1)(x-2)...(x-n))^p}{(p-1)!} dx,$$

$$M_t = e^t \int_t^\infty x^{p-1} e^{-x} \frac{((x-1)(x-2)...(x-n))^p}{(p-1)!} dx,$$

$$\epsilon_t = e^t \int_0^t x^{p-1} e^{-x} \frac{((x-1)(x-2)...(x-n))^p}{(p-1)!} dx,$$

where p is a large prime. It is clear that $M_t + \epsilon_t = e^t M$ as required by the definition of M, M_t and ϵ_t, however we need to show that M, M_t are integers.

5.1.1 Evaluating M and M_t

We aim to show that M and M_t are integers and then that the left hand side of Equation (5.2) is a non-zero integer. We start with a simple preliminary proposition.

Proposition 5.2. $\int_0^\infty x^m e^{-x} dx = m!$.

Proof. Integrating by parts, we achieve that

$$\int_0^\infty x^m e^{-x} dx = [-x^m e^{-m}]_0^\infty + m \int_0^\infty x^{m-1} e^{-x} dx$$
$$= m \int_0^\infty x^{m-1} e^{-x} dx = m! \int_0^\infty e^{-x} dx = m!.$$

□

Next we rewrite the polynomial in the integrand of M and M_t:

$$[(x-1)...(x-n)]^p = x^{pn} + ... + (-1)^n (n!)^p$$
$$= (-1)^n (n!)^p + \sum_{i=1}^{n} b_i x^{ip},$$

where the coefficients of the polynomial, b_i, are integers. M is thus an integer since it can be written as

$$M = (-1)^n (n!)^p \int_0^\infty \frac{e^{-x} x^{p-1}}{(p-1)!} dx + \int_0^\infty \frac{e^{-x} x^{p-1}}{(p-1)!} \sum_{i=1}^{n} b_i x^{ip} dx$$
$$= (-1)^n (n!)^p + \sum_{i=1}^{n} b_i \frac{(ip+p-1)!}{(p-1)!},$$

using the previous proposition. The second term is clearly divisible by p, because, expanding the factorials, we have that

$$\frac{(ip+p-1)!}{(p-1)!} = p(p+1)\ldots(ip+p+1).$$

We now choose p so that $p > n$, and thus the first term, $(-1)^n (n!)^p$ cannot be divisible by p. This is because $n!$ is formed from multiplying numbers $1 \leq n < p$, and since p is prime, it has no factors less than itself.

Therefore M is not divisible by p, and by also choosing $p > a_0$, we further ensure that $a_0 M$ is not divisible by p.

We evaluate M_t similarly. Making the substitution $y = x - t$, M_t becomes

$$\int_0^\infty (y+t)^{p-1} e^{-y} \frac{((y+t-1)(y+t-2)...y...(y+t-n))^p}{(p-1)!} dy.$$

Remembering that n is the degree of the minimal polynomial for e, and that t runs between 1 and n, we know that $t - n \leq 0$ so there must be a factor of y in the polynomial. The powers of y in the polynomial run between 1 and $(np + p - 1)$, with coefficients, $c_i/(p-1)!$, where c_i is an integer for all i. That is:

$$\begin{aligned} M_t &= \int_0^\infty \frac{e^{-y}}{(p-1)!} \sum_{i=1}^{np+p-1} c_i y^i dy \\ &= \sum_{i=1}^{np+p-1} c_i \frac{(ip+p+1)!}{(p-1)!} \\ &= \sum_{i=1}^{np+p-1} c_i p(p+1)...(p+1+ip). \end{aligned}$$

Thus, M_t is an integer divisible by p, and, using the trick explained at the start of this chapter,

$$Ma_0 + \sum_{t=1}^{n} a_t M_t$$

is a non-zero integer.

5.1.2 Bounding the Integer

Using Equation (5.2), if we had that, for large enough p,

$$\left| \sum_{t=1}^{n} a_t \epsilon_t \right| < 1, \tag{5.3}$$

then e must be transcendental by our standard contradiction, as a non-zero integer cannot have absolute value less than 1.

To prove that ϵ_t can be made small enough that Equation (5.3) holds, we consider bounds:

$$\begin{aligned} \epsilon_t &= e^t \int_0^t x^{p-1} e^{-x} \frac{((x-1)(x-2)...(x-n))^p}{(p-1)!} dx \\ &< e^t t^p \int_0^t \frac{((x-1)(x-2)...(x-n))^p}{(p-1)!} dx. \end{aligned}$$

Clearly the polynomial $((x-1)(x-2)...(x-n))$ has finite values for x between 0 and t, so we can let U be the maximal absolute value of $((x-1)(x-2)...(x-n))$ for x in this range. Then,

$$|\epsilon_t| < \frac{e^t t^{p+1} U^p}{(p-1)!} = e^t t \times \frac{(tU)^p}{(p-1)!}.$$

Note t is some fixed number between 0 and n, but p can be any prime larger than n. As we make p larger, $(p-1)!$ grows faster than any power of p, so ϵ_t approaches 0. Thus we can choose p such that (5.3) holds and e is transcendental. □

Corollary 5.3. For any non-zero rational number r, e^r is a transcendental number.

Proof. I prove this using basic Galois theory with results from [25]. Let $r = a/b$, and assume that e^r is algebraic, and let d be the degree of its minimal polynomial over $\mathbb{Q}$. Then,

$$[\mathbb{Q}(e^r)] : \mathbb{Q}] = d < \infty.$$

Clearly, $e^a = (e^r)^b$ is in the field extension $\mathbb{Q}(e^r)$, so we define the polynomial

$$f(x) = x^a - e^a \in \mathbb{Q}(e^r). \tag{5.4}$$

Since e is a root of $f(x)$, we have that

$$[\mathbb{Q}(e) : \mathbb{Q}(e^r)] \leq a$$
$$[\mathbb{Q}(e)] : \mathbb{Q}] \leq ad < \infty$$

by the Tower Theorem. But this says that e has finite degree over $\mathbb{Q}$, or equivalently that e is algebraic, a contradiction to the previous section. Note this proof breaks down if $r = a = 0$ because Equation (5.4) becomes the zero polynomial. Thus e^r is transcendental for rational $r \neq 0$.

□

5.2 The Transcendence Of π

5.2.1 Symmetric Polynomials

The next section fuses the proofs of the transcendence of π in [25] and [2]. First, we need to develop some more Galois Theory.

Definition 5.4. A symmetric polynomial in n variables, $f(x_1, x_2, \ldots x_n)$, is a polynomial such that

$$f(\tau(x_1, \ldots x_n)) = f(x_1, x_2, \ldots x_n)$$

for all $\tau \in S_n$.

Example 5.5. Let

$$f(x_1, x_2, x_3) = x_1x_2x_3 + x_1 + x_2 + x_3$$

It is clear that any permutation of $\{x_1, x_2, x_3\}$ will not change $f(x_1, x_2, x_3)$.

We introduce the elementary symmetry polynomials in n variables, which can be thought of as the building blocks of symmetric polynomials:

$$\begin{aligned} e_1 &= x_1 + x_2 + ... + x_n \\ e_2 &= x_1x_2 + x_1x_3 + ... + x_{n-1}x_n \\ &\vdots \\ e_n &= x_1x_2 \ldots x_n. \end{aligned}$$

Definition 5.6. The elementary symmetric polynomials in n variables are, for $0 < k \leq n$,

$$e_k(x_1, \dots x_n) = \sum_{1 \leq i_1 < \dots < i_k \leq n} x_{i_1} x_{i_2} \dots x_{i_k}$$

with $e_k = 0$ for $k > n$.

Theorem 5.7. Any symmetric polynomial in n variables can be expressed as a polynomial in the elementary symmetric polynomials.

Corollary 5.8. If we have a polynomial $F(t)$

$$F(t) = t^n + c_{n-1}t^{n-1} + \dots c_0,$$

such that $F(t)$ has roots $\{x_1, \dots x_n\}$, then any symmetric polynomial in $\{x_1, \dots x_n\}$ can be expressed in terms of the coefficients $\{c_{n-1}, \dots, c_0\}$.

Proof. We write $F(t)$ as:

$$\begin{aligned} F(t) &= c_n[(t - x_1)(t - x_2) \dots (t - x_n)] \\ &= c_n[t^n - e_1 t^{n-1} + e_2 t^{n-2} \dots + (-1)^n e_n] \end{aligned}$$

Therefore, we can express the elementary symmetric polynomials in $\{x_1, \dots, x_n\}$ as

$$e_k = (-1)^k c_{n-k}$$

and any symmetric polynomial in $\{x_1, \dots, x_n\}$ can written in terms of these coefficients by Theorem 5.7. □

5.2.2 Lindemann's Proof

Theorem 5.9 (Lindemann)**.** π is a transcendental number.

Proof. In order to prove this, we must consider complex transcendental numbers as well as real ones. We assume π is an algebraic number. Then, since i is algebraic (as it is a root of $f(x) = x^2 + 1$), $i\pi$ must be algebraic as well.

Now consider the minimal polynomial for $i\pi$ over the rationals and label the roots of it $\alpha_1 = i\pi, \alpha_2, \dots, \alpha_n$, counted with multiplicity, so the degree of π over the rationals is n. Then:

$$(e^{i\alpha_1} + 1)(e^{\alpha_2} + 1) \dots (e^{\alpha_n} + 1) = 0, \tag{5.5}$$

by Euler's identity:

$$e^{i\pi} + 1 = 0.$$

Expanding (5.5), all the terms will be in the form e^{β_s} where s is a subset of $\{1, 2, \dots, n\}$,

$$\beta_s = \sum_{k \in S} \alpha_k.$$

and we put $\beta_\emptyset = 0$. For example, we have the term

$$e^{\alpha_1} \times e^{\alpha_2} \times 1 \times 1 \times \cdots \times 1 = e^{\alpha_1 + \alpha_2} = e^{\beta_{\{1,2\}}}.$$

There are 2^n different subsets of $\{1, 2, \dots, n\}$, and thus 2^n different β_s values. However, the β_s terms are not necessarily distinct, as different combinations of α_k might sum to the same number. We relabel the non-zero values of β_s as $\beta_1, \beta_2, ...\beta_r$ with multiplicity. Then (5.5) becomes

$$\begin{aligned} 0 &= e^{\beta_1} + ... + e^{\beta_r} + e^0 + ... + e^0 \\ &= e^{\beta_1} + ... + e^{\beta_r} + k, \end{aligned} \tag{5.6}$$

where k is an integer greater than 1, since $\beta_\emptyset = 0$. We create a polynomial with roots which are the non-zero sums β_s:

$$\tilde{\theta}(z) = \prod_{i=1}^{r} (z - \beta_i).$$

Now the coefficients of $\tilde{\theta}$ are elementary symmetry polynomials in the sums β_i. Because we can permute $\{\alpha_k\}$ without changing the set of non-zero sums, the coefficients of $\tilde{\theta}$ can also be written as elementary symmetric polynomials in $\alpha_1, ...\alpha_n$.

But the elementary symmetric polynomials of $\{\alpha_1, ...\alpha_n\}$ are nothing more than plus or minus the coefficients of the minimal polynomial of $i\pi$, which are rational!

Therefore $\tilde{\theta}$ has rational coefficients. By multiplying by a suitable constant, we create a polynomial $\theta(z)$ with integer coefficients with $\beta_1, ...\beta_r$ as roots:

$$\theta(z) = c_r z^r + c_{r-1} z^{r-1} + ... + c_0$$

where the degree of $\theta(z)$ is $r = 2^n - k$. We now define

$$\begin{aligned} f(z) &= \frac{{c_r}^{rp} z^{p-1} (\theta(z))^p}{(p-1)!}, \\ F(x) &= f(z) + f'(z) + ... + f^{(pr-p-1)}(z) \\ I_\beta &= \int_0^\beta f(z) e^{\beta - z} dz, \end{aligned}$$

where p is a large prime.

Similarly to in Example 1.7, where we proved the irrationality of π, we can see that

$$F'(z) - F(z) = -f(z),$$

and thus
$$\frac{d}{dz}[F(z)e^{\beta-z}] = -f(z)e^{\beta-z}.$$
Because $f(z)e^{\beta-z}$ is an entire function, the value of I_β is independent of the contour taken, and so
$$I_\beta = F(0)e^\beta - F(\beta).$$

Now we consider the object which will eventually be proved to be our impossible integer:

$$\begin{aligned} J := \sum_{i=1}^{r} I_{\beta_i} &= F(0)\sum_{i=1}^{r} e^{\beta_i} - \sum_{i=1}^{r} F(\beta_i) \\ &= -kF(0) - \sum_{i=1}^{r}\sum_{j=1}^{pr+p-1} f^{(j)}(\beta_i) \\ &= -kF(0) - \sum_{j=0}^{\infty}\sum_{i=1}^{r} f^{(j)}(\beta_i), \end{aligned} \tag{5.7}$$

using 5.6 and the fact that $f^{(n)}(z) = 0$ for $n \geq pr + p$. Note that the second term is actually a finite sum, since $f(z)$ is a polynomial.

Evaluating J

Consider
$$\frac{1}{{c_r}^{rp}} \sum_{i=1}^{r} f^{(j)}(\beta_i). \tag{5.8}$$

Since $f(z)$ has zeroes of order at least p at every $z = \beta_i$, if $j < p$, $f^{(j)}(\beta_i) = 0$.

Then for $j \geq p$, differentiating p times will create a factor of $p!$. After dividing by $(p-1)!$ in the denominator of $f(z)$, $f^{(j)}(z)$ will be a polynomials with integer coefficients which are all multiples of ${c_r}^{rp}p$ (as $f(z)$ also contains a factor of ${c_r}^{rp}$).

Thus 5.8 is a symmetric polynomial in $\beta_1, \ldots \beta_r$ with integer coefficients divisible by p. So, by Corollary 5.8, 5.8 can be written as a polynomial with coefficients divisible by p in the elementary symmetric polynomials, that is:

$$\begin{aligned} \frac{1}{{c_r}^{rp}} \sum_{i=1}^{r} f^{(j)}(\beta_i) &= g\left(e_1, e_2, ...e_r\right) \\ &= g\left(-\frac{c_{r-1}}{c_r}, \frac{c_{r-2}}{c_r}, ..., \pm\frac{c_0}{c_r}\right). \end{aligned}$$

Then, since
$$\deg(g) \leq \deg(f^{(j)}) \leq \deg(f) - j \leq pr,$$

multiplying by ${c_r}^{rp}$ will clear all the denominators in g, meaning that

$$\sum_{i=1}^{r} f^{(j)}(\beta_i)$$

is an integer divisible by p.

Evaluating $F(0)$ Using the argument about $f^{(j)}(z)$ from the previous section, and the fact that $z = 0$ is a root of $f(z)$ of order exactly $p - 1$, we can find that

$$f^{(j)}(0) = \begin{cases} 0, & \text{if } j < p - 1 \\ {c_0}^p {c_r}^{rp}, & \text{if } j = p - 1 \\ \text{a multiple of } p, & \text{if } j > p - 1. \end{cases}$$

This implies that

$$\begin{aligned} F(0) &= \sum_{i=0}^{\infty} f^{(j)}(0) \\ &= {c_0}^p {c_r}^{rp} + Np, \end{aligned}$$

where N is some integer. Thus, using 5.7 we can write

$$J = -k{c_0}^p {c_r}^{rp} + N'p,$$

where N' is another integer. Now we can choose p to be larger than the absolute values of c_0, c_r and k; this means $kc_0c_r{}^{rp}$ cannot contain a factor of p. Therefore, J is an integer not divisible by p, and is thus non-zero.

Bounding J

We now aim to bound J from above. Recall that

$$I_{\beta_i} = \int_0^{\beta_i} f(z)e^{\beta_i - z}dz = \frac{{c_r}^{rp}}{(p-1)!} \int_0^{\beta_i} z^{p-1}(\theta(z))^p e^{\beta_i - z}dz$$

and, since the value of I_{β_i} is independent of the contour taken, we assume that z takes the values $\beta_i t$ for t between 0 and 1.

Now let C be the maximum absolute value of the coefficients of $\theta(z)$, that is:

$$C = \max\{|c_1|, |c_2|, ..., |c_r|\}.$$

Then we bound

$$\begin{aligned} |\theta(z)| &\le \sum_{j=0}^{r} |c_j||z|^j \\ &\le C \sum_{j=0}^{r} |\beta_i|^j = C(1 + |\beta_i|)^r \end{aligned}$$

for $z = \beta_i t$ with t between 0 and 1. Additionally, we bound

$$\begin{aligned}\left|e^{\beta_i - z}\right| &= e^{\mathrm{Re}\,(\beta_i - z)} \\ &\leq e^{|\,\mathrm{Re}\,(\beta_i)|}.\end{aligned}$$

This all implies

$$\begin{aligned}|I_{\beta_i}| &\leq \frac{{c_r}^{rp}}{(p-1)!} \times |\beta_i|^p \times C^p(1+\beta_i)^{rp} \times e^{|\,\mathrm{Re}\,(\beta_i)|} \\ &= \frac{[{c_r}^r \times \beta_i \times C(1+|\beta_i|)^r]^p \times e^{|\,\mathrm{Re}\,(\beta_i)|}}{(p-1)!} \\ &= \frac{A_i \times {B_i}^p}{(p-1)!}\end{aligned}$$

where A_i and B_i are independent of p. Then

$$\begin{aligned}|J| = \left|\sum_{i=1}^{r} I_{\beta_i}\right| &\leq \sum_{i=1}^{r} |I_{\beta_i}| \\ &\leq \frac{1}{(p-1)!}\sum_{i=1}^{r}(A_i \times {B_i}^p) \\ &\leq \frac{A \times B^p}{(p-1)!}\end{aligned}$$

where

$$\begin{aligned}B &= \max\{B_1, B_2, \ldots B_r\} \\ A &= r\max\{A_1, A_2, \ldots A_r\}\end{aligned}$$

As we have seen before, for large p, $(p-1)!$ will grow faster than any power, and the upper bound on J will tend to 0. By our standard contradiction, π is transcendental.

□

5.2.3 The Original Direction of Proof

In fact, the direction of Lindemann's proof was reversed [2]. He first showed the more general theorem:

Theorem 5.10 (Lindemann)**.** If α is any non-zero algebraic number, then e^{α} is transcendental.

Proving this theorem is more involved than our proof for the transcendence of π and e, and requires more Galois Theory. However it then has the immediate corollary that $i\pi$ and thus π must be transcendental, since we have that

$$e^{i\pi} = -1.$$

This is the underlying reason for the similarities between the transcendence proof of e and π.

5.3 Squaring The Circle

We started this report by discussing irrational numbers in antiquity, so it is satisfying to end the report by using our previous results to answer a problem from Ancient Greece. This section follows [25], Chapter 7.

The Ancient Greeks were interested in ruler and compass construction. In particular, they wanted to construct a square with the same area as a given circle using only an unmarked straight edge and a pair of compasses. This problem was known as "squaring the circle", and we will show that it is impossible.

We can reformulate this problem. Given a set of distinct points P_0 in $\mathbb{R}^2$, we consider the operators:

- **Ruler:** Draw a straight line passing through any 2 points in P_0.
- **Compasses:** Draw a circle with a centre at any point in P_0 and radius ω, where ω is a distance between two points in P_0.

Definition 5.11 (Constructable Numbers)**.** We can construct a point from P_0 in one step if it is at the intersection of any two lines or circles drawn using the operators above. A point $r \in \mathbb{R}^2$ is constructable from P_0 if there is a finite sequence $r_1, \ldots, r_n = r$ in $\mathbb{R}$ such that the point r_i can be constructed in one step from the set $P_0 \cup \{r_1, \ldots, r_{i-1}\}$.

We now realise K_0 as the minimal field extension of $\mathbb{Q}$ containing all the x and y coordinates of the points in P_0.

Example 5.12. If $P_0 = \{(1, 2), (3, \sqrt{2}), (\sqrt{5}, 0)\}$, then

$$K_0 = \mathbb{Q}(\sqrt{2}, \sqrt{5}).$$

If a point r_i has coordinates (x_i, y_i), we define K_j as the minimal field extension of K_{j-1} containing (x_i, y_i), that is

$$K_i = K_{i-1}(x_1, y_1).$$

We thus have a tower of subfields:

$$\mathbb{Q} \subseteq K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n \subseteq \mathbb{R},$$

since all coordinates x and y are in $\mathbb{R}$.

Lemma 5.13. With the notation above, x_i and y_i each have at most quadratic degree over K_{j-1}.

The coordinates (x_i, y_i) could have been formed in 3 ways: at the intersection of 2 lines, 2 circles, or a circle and a line. Finding the intersection between them is equivalent to solving simultaneous equations, which leads to an equation in (x_i, y_i) of at most quadratic degree over K_{j-1}.

Theorem 5.14. For any constructable point $r = (x, y) \in \mathbb{R}^2$, both x and y are algebraic numbers with degree a power of 2 over K_0.

Proof. The previous lemma states that

$$[K_{j-1}(x_i) : K_{j-1}] = 1 \text{ or } 2,$$

for all $1 \leq i \leq n$, and similarity for y_i. Therefore

$$[K_i : K_{i-1}] = [K_{i-1}(x_i, y_i)) : K_{i-1}] = 1, 2 \text{ or } 2^2.$$

□

By the Tower Theorem, this implies that

$$[K_{n-1} : K_0] = 2^m$$

for some natural number m, and so both x and y have degrees of either 2^m or 2^{m+1} over K_0 and are thus algebraic over K_0.

Corollary 5.15. It is impossible to square the circle.

Proof. Constructing a square with the same area as a unit circle is equivalent to the point $(\sqrt{\pi}, 0)$ being constructable from $P_0 = \{(0, 0), (1, 0)\}$.

In this case, we have $K_0 = \mathbb{Q}$ and, if it is possible, by the previous lemma, $\sqrt{\pi}$ must be algebraic over $\mathbb{Q}$ with degree 2^m for some natural number m. This then implies π is algebraic, but by Theorem 5.9, it is not. □

Chapter 6

Discussion

We have now built several methods and criteria for proving numbers are irrational or transcendental, and applied them to many significant numbers. In particular, we have seen classes of transcendental numbers, such as Liouville numbers and $\mathcal{M}(g)$, which have different properties.

However transcendental number theory is a huge topic and, with more time and space, I could have explored many more areas. For example, a natural extension of this work would be continued fractions, as the continued fraction form of real numbers is closely linked to their approximation exponent. I also would have liked to examined more methods to proving the transcendence of numbers, such as Mahler's method, which uses functional equations and automata. In addition, I would like to briefly highlight some open questions in transcendental number theory.

6.1 Linear Independence of Transcendental Numbers

Considering I proved the irrationality of e and π in the introduction, using only school-level mathematics, it is perhaps surprising that there is no proof that $e + \pi$ and $e\pi$ are irrational. We can be sure that at least one of these number is irrational by considering the quadratic polynomial with roots e and π:

$$f(x) = (x - e)(x - \pi) = x^2 - (e + \pi)x + e\pi.$$

Because we have seen that e and π are transcendental, $f(x)$ cannot have rational coefficients, and thus at least one of $e + \pi$ and $e\pi$ is irrational. If we could prove that e and π are linearly independent over the rational numbers, then immediately $e + \pi$ would be irrational.

Although this has so far alluded us, in 1993, Masayoshi Hata showed that the numbers $\log 2$ and π are linearly independent over $\mathbb{Q}$. He did this using a type of rational approximation called Hermite-Pade rational approximation. Apéry and Beukers' proof of the irrationality of $\zeta(3)$ is closely related to Hermite-Pade rational approximation, and can in fact be reformulated in terms of it [3].

Additionally in 1885, Weierstrass generalised Lindermann's Theorem (Theorem 5.10). I state the version reformulated by Baker in [4], Chapter 1.

Theorem 6.1 (Weierstrass)**.** If $\alpha_1, \ldots, \alpha_n$ are distinct algebraic numbers, then $e^{\alpha_1}, \ldots, e^{\alpha_n}$ are linearly independent over the algebraic numbers.

6.2 Beukers-Lite Integrals for Irrationality and Periods

Proving the irrationality of specific naturally occurring, real constants is often highly difficult, and numbers like e and π are very much the exception. As we explored earlier, attempts to prove the irrationality of $\zeta(5)$ have so far not been successful.

Anther number which has resisted efforts to prove its irrationality is Catalan's constant C, which arises naturally in combinatorics:

$$C = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)^2}.$$

Wadim Zudilin discovered that C can be written as an integral:

$$C = \frac{1}{8} \int_0^1 \int_0^1 \frac{x^{-1/2}(1-y)^{-1/2}}{1-xy} dxdy.$$

Using this, he constructed Beuker's-like series of double integrals, which come close to proving the irrationality of C. Zudilin's ideas have inspired other mathematicians to tweak Beukers' integral to attempt to prove the irrationality of other constants such as Euler-Mascheroni's constant, γ (see [23]), which is a number closely linked to the zeta function, defined below.

$$\begin{aligned} \gamma &= \lim_{n \to \infty} \left(-\log n + \sum_{i=1}^{n} \frac{1}{i} \right) \\ &= \int_1^\infty \left(\frac{1}{\lfloor x \rfloor} - \frac{1}{x} \right) dx. \end{aligned} \tag{6.1}$$

We can generalise the ideas of the report to define periods (this final part draws on [27]).

Definition 6.2 (Periods)**.** A period is a number which can be expressed as an integral of an algebraic function over an algebraic domain.

Periods are a class of numbers between the algebraic numbers and the transcendental numbers: all algebraic numbers are periods.

We have seen that $\zeta(2), \zeta(3)$, and $\zeta(4)$ can be written as integrals of algebraic functions over a domain in the form $[0,1]^n$, and are therefore periods. In fact, $\zeta(s)$ is a period for all integer values of s where it is defined. Unsurprisingly, due to the links between the zeta function and π, we also have that π is a period, specifically:

$$\pi = \int_0^1 \frac{4}{1+x^2} dx.$$

In addition, Zudilin's discovery tells us that $\mathcal{C}$ is a period, however it is not known whether γ is: note that the floor function is not an algebraic function so 6.1 does not satisfy the definition for γ to be a period. It is also an open question whether e is a period!

Bibliography

[1] Masaaki Amou. Approximation to certain transcendental decimal fractions by algebraic numbers. *Journal of Number Theory*, 37(2):231–241, 1991.

[2] David Angell. *Irrationality and Transcendence in Number Theory.* Chapman and Hall/CRC, 2021.

[3] Walter Van Assche. Hermite-pade rational approximation to irrational numbers. *Computational Methods and Function Theory*, 10:585–602, 2010.

[4] Alan Baker. *Transcendental Number Theory.* Cambridge Mathematical Library. Cambridge University Press, 1975.

[5] Frits Beukers. A note on the irrationality of $\zeta(2)$ and $\zeta(3)$. *Bulletin of the London Mathematical Society*, 11:268–272, 1969.

[6] Jonathan Borwein and Peter Borwein. Irrationality measures. In *Pi and the AGM*, chapter 11. Wiley-Interscience, 1987.

[7] Brilliant.org. A history of irrational numbers. Electronic Document, retrieved April 2023, `https://brilliant.org/wiki/history-of-irrational-numbers/`.

[8] Edward B. Burger and Robert Tubbs. *Making Transcendence Transparent: An intuitive approach to classical transcendental number theory.* Springer Science+Business Media, July 2004.

[9] Sully F. Chen and Erin P. J. Pearse. The irrationality measure of π as seen through the eyes of $\cos(n)$. *Elemente der Mathematik*, 75(4):152–165, October 2020.

[10] Janot de Stainville. *Melanges d'Analyse Algebrique et de Geometrie (A Mixture of Algebraic Analysis and Geometry).* Veuve Courcier, 1815.

[11] Jeff Miller (ed.). Earliest known uses of some of the words of mathematics. *https://mathshistory.st-andrews.ac.uk/Miller/mathword/*, August 2018. Electronic document, retrieved March 2023.

[12] G.H. Hardy and E.M. Wright. *Introduction to the theory of numbers.* Oxford University Press, 1960.

[13] Dirk Huylebrouck. Similarities in irrationality proofs for pi, $\ln 2$, zeta(2), and zeta(3). *American Mathematical Monthly*, 108:222 – 231, March 2001.

[14] Dirk Huylebrouck. A way to prove the irrationality of zeta(4). *arxiv: Number Theory*, 2014. `https://doi.org/10.48550/arXiv.1406.4464`.

[15] F. M. S. Lima. Beukers-like proofs of irrationality for $\zeta(2)$ and $\zeta(3)$. *arXiv: Number Theory*, 2016. `https://doi.org/10.48550/arXiv.1308.2720`.

[16] Karl Mahler. On the approximation of pi. *Koninklijke Nederlandse Akademie van Wetenschappen*, 56(1):30–42, 1953.

[17] Karl Mahler. On a class of transcendental decimal fractions. *Communications on Pure and Applied Mathematics*, 29:717–725, 1976.

[18] David Masser. *Auxillary Polynomials in Number Theory.* Cambridge University Press, 2016.

[19] Paul J. Nahin. *In Pursuit of Zeta-3.* Princeton University Press, October 2021.

[20] Ivan Niven. A simple proof that π is irrational. *Bulletin of the American Mathematical Society*, 56(6):509–509, June 1946.

[21] Ian G Pearce. Mathematics in the service of religion: Sulba sutras. `https://mathshistory.st-andrews.ac.uk/Projects/Pearce/chapter-5/`, Electronic document, retrieved March 2023.

[22] Gertjan Pomstra. The constant of champernowne. July 2018. Bachelor's Thesis: `https://fse.studenttheses.ub.rug.nl/17907/1/The%20Constant%20of%20Champernowne.pdf`.

[23] Christoph Koutschan Robert Dougherty-Bliss and Doron Zeilberger. Tweaking the beukers integrals in search of more miraculous irrationality proofs a la apery. *The Ramanujan Journal*, 58(3):973–994, July 2021.

[24] Judith Sally and Paul Sally. Rational approximation. In *Roots to Research: A Vertical Development of Mathematical Problems.* Providence RI: American Mathematical Society, 2007.

[25] Ian Stewart. *Galois Theory.* Chapman and Hall/CRC, 1973.

[26] Marco Tavora. Proving the transcendence of e. *Towards Data Science*, March 2020. Electronic document, retrieved April 2023, `https://towardsdatascience.com/proving-the-transcendence-of-e-65aaa2de69c3`.

[27] Don Zagier and Maxim Kontsevich. Periods. In Björn Engquist and Wilfried Schmid, editors, *Mathematics unlimited—2001 and beyond*, pages 771–808. Springer Science+Business Media, 2001.

[28] Wadim Zaudlin. One of the number $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, $\zeta(11)$ is irrational. *Russian Math. Surveys*, 56(4):774–776, 2001.

[29] Doron Zeilberger and Wadim Zudilin. The irrationality measure of π is at most 7.103205334137... *Moscow Journal of Combinatorics and Number Theory*, 9(4):407–419, December 2020.