On the Arithmetic Nature of the Zeta Function and Its Generalizations

P. Y. Cruz

August, 2026

Abstract

In this paper, we present an exposition of the arithmetic nature of the Riemann zeta function, motivating our analysis through the study of Fourier series and the Basel problem. We develop the analytic properties of the zeta function, leading to an explicit formula for its values at even positive integers. We discuss irrationality results such as Apéry's theorem and Beukers' re-interpretation via multiple integrals. We consider results and directions following Apéry's theorem, specifically, we cover Zudilin's result that at least one of $\zeta(5), \zeta(7), \ldots, \zeta(25)$ is irrational, then propose refinements to the method. Finally, we discuss generalizations of the Riemann zeta function, including the Hurwitz zeta function and Dirichlet L-functions, with a particular emphasis on the Dirichlet beta function. The exposition is intended to be largely self-contained and accessible to an advanced undergraduate or graduate audience.

Contents

1 Introduction

The study of infinite sums dates back to the Greeks, around 300 BCE. A most well known account of this is sometimes called "Zeno's Paradox," named after Greek philosopher Zeno of Elea (c. 490 BCE – c. 430 BCE). The idea is as follows:

Remark 1.1: Zeno's Paradox

In order to traverse any distance, one must travel half-way there, then they must travel half the remaining distance, then another half and so on. That is, in order to get from point A to point B, one must travel half-way from A to B, then within this half, they must again travel half or $3/4$ in total, and this continues on forever. A diagram helps to visualize this.

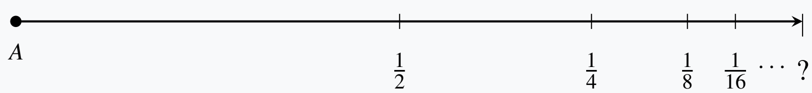

Figure 1

This is not immediately an infinite series, and at the time, such an object was unknown. However, if we consider the total amount of distance traveled as the sum of each step, this problem clearly becomes the evaluation of

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots .$$

The *paradox* that lies in Zeno's Paradox is that it is therefore impossible to get from point A to point B since there is an infinite number of steps required to be taken and we have only finite time. That is to say, the idea of an infinite sum evaluating to a finite value was paradoxical at the time.

This is, of course, however, not true, so what's going on here? This is where the study of infinite series comes in and the idea of *convergence*. It can be quickly shown that Zeno's Paradox is a specific case of what is regarded as a *geometric series*, which in particular converges for this case. One particularly famous instance of an infinite sum other than the geometric series being written in closed form was the so-called Basel problem.

Theorem 1.1: The Basel Problem

The sum of the squared reciprocals of positive integers is equal to the number π squared divided by six. That is,

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots = \frac{\pi^2}{6}. \tag{1.1}$$

This problem was posed by the Italian mathematician Pietro Mengoli (1626 – 1686) sometime in the mid 1600's and it was not until nearly a century later in 1734 that Swiss mathematician Leonhard Euler (1707 – 1783) solved it. As Zeno did, and many other mathematicians have, we can of course consider the structure of different infinite series, however we will restrict ourselves to the kind seen in Equation (1.1) until § 7 where we will consider its natural generalizations. A defining characteristic of infinite sums is the difficulty in finding their closed form. For example, the problem of evaluating Equation (1.1) with the exponent of n to be 3 rather than 2 was even elusive to Euler. In fact, almost three centuries later, current day, a closed form continues to be elusive. Unlike other well known Calculus concepts such as differentiation and (elementary) integration, there is no specific "method" to finding the closed form of a summation and there are typically numerous ways to find such a sum. For example, Euler solved the Basel problem in a rather intricate way, utilizing the so-called product formula for the sine function, which at the time, was actually yet to be proven[1], and we will use a completely different more elementary method.

Due to the difficulty in finding a closed form for these infinite series, mathematicians have turned to investigating the arithmetic nature of them, which is still rather difficult, but perhaps more accessible than a closed form. The Basel problem, was just one instance of a particular series, as it turns out Euler was able to generalize this result to any even positive integer. All of these evaluations happen to be rational multiples of powers of π. The first advancement in the arithmetic nature of infinite series of this form was due to French mathematician Adrien Legendre (1752 – 1833), who proved that π^2 was irrational in 1794. Using the closed form for our summation coupled with the fact that the product of a rational and irrational number is irrational proves that the sum in Equation (1.1) is irrational. However, since, in general, the product of two irrational numbers need not be irrational, this irrationality result does not imply anything further. Then, nearly a century later, in 1882, German mathematician Carl Lindemann (1852 – 1939) proved that π was *transcendental*. This result implied the transcendence and subsequent irrationality of all sums of the form (1.1) with exponent $2k$ for $k \in \mathbb{Z}_{>0}$.

It was not until another near century later that we received our most recent piece of information on the behavior of infinite series of this form due to the French mathematician Roger Apéry (1916 – 1994) in 1979, who proved the irrationality of both Equation (1.1) and the case of n cubed. His first result was already known and in fact superseded by the result of Lindemann, however, the main surprise lied in the novelty of his method.

As of current day, there has not been any success in further generalizing the results of Apéry to exponents of higher degree. Although the study of series of this form has not remained entirely stagnant. However, due to its evasive behavior, we have instead begun to consider non-constructive results. That is, instead of showing a specific power of our series is irrational, the most prominent results have been proving that infinitely many are or that at least one of a small set is.

1.1 Organization

The paper is organized as follows: In the proceeding subsection, we discuss some preliminaries to the method in which we solve the Basel problem and the reader comfortable with Fourier series may skim this. We then solve the Basel problem in § 1.3. Then, in § 2, we discuss the Riemann zeta function, drawing from ideas which may be found in [Apo76, § 12] and [Gam01, § 15.3]. In particular, we discuss the analytic properties of the Riemann zeta function, analytically continuing it and deriving its functional equation, which is then used to evaluate it at any even positive integer. In § 3 we consider the irrationality of $\zeta(3)$, otherwise known as Apéry's Theorem, following a detailed exposition [Ste05, § 4]. This is then followed up by Beukers' reinterpretation via multiple integrals in § 3.2. A detailed account of the story for $\zeta(2)$ is given in the following section. We then consider the concept of measuring *how* irrational a number may be, discussing

1. This identity was assumed by Euler in his solution to the Basel problem. It was, however, not proven until 1876 by German mathematician Karl Weierstraß (1815 – 1897).

Dirichlet's Approximation Theorem, giving us a definitive lower bound of the irrationality exponent for any irrational number. We proceed to determine an upper bound of the irrationality exponent for $\zeta(3)$ and $\zeta(2)$. We close the section by determining the exact irrationality exponent of $\sqrt{2}$ and then study the natural logarithm evaluated at two. We proceed to discuss Wadim Zudilin's result [Zud18] on the irrationality of at least one of the numbers $\{\zeta(5), \zeta(7), \ldots, \zeta(25)\}$ in § 6. Then in § 6.1 we give a brief and informal discussion on possible ways to use Zudilin's result to remove some items from the above set. In our last official section, § 7, we discuss natural generalization to the Riemann zeta function, and provide a similar, but less detailed story of these generalizations. In particular, we give special interest to the arithmetic of the so-called Dirichlet beta function. Finally, in § 8 we provide closing remarks and outline potential directions for further study.

Acknowledgments

The author would like to express their deepest gratitude to their advisor, Yunqing Tang, for their consistent support, patience, and inspiring discussions. A special thank you is also due to Jose J. Martinez, whose passion and mentorship convinced the author to pursue the study of mathematics, and to Avi Zeff, who provided the initial inspiration and encouragement to write this paper. Last but most certainly not least, the author would like to thank his wife, Tala, for her constant support, encouragement, and reading this paper front to back. Without the inspiration and encouragement of these individuals, this paper would not have been possible.

1.2 Fourier Series

In this section, we develop the necessary machinery to solve the Basel problem in one of many possible ways. We will utilize the ideas of the French mathematician Joseph Fourier (1768 – 1830).

First, recall that we say a sequence of complex-valued functions (f_n) on a set S **converges uniformly** to f on S if for all $\varepsilon > 0$ there exists some N such that if $n \geq N$, then for all $z \in S$,

$$|f_n(z) - f(z)| < \varepsilon.$$

Furthermore, if $\sum g_n(z)$ is a series of complex-valued functions on a set S, we say the **series converges uniformly** on S if the sequence of partial sums $(S_n(z))$ converges uniformly on S. Moreover, we have the following standard theorem on uniformly convergent sequences.

Theorem 1.2: Integrating term by term

Let (g_n) be a sequence of integrable functions on $[a, b]$. If the series $\sum g_n(z)$ converges uniformly on $[a, b]$, then

$$\int_a^b \left(\sum g_n(z)\right)\, dz = \sum \left(\int_a^b g_n(z)\, dz\right).$$

Fourier series are quite similar to the study of Laurent series. We define a **complex Fourier series** to be the two-tailed series

$$\sum_{n\in\mathbb{Z}} c_n e^{in\theta} = \cdots + c_{-2}e^{-2i\theta} + c_{-1}e^{-i\theta} + c_0 + c_1 e^{i\theta} + c_2 e^{2i\theta} + \cdots .$$

Example 1.1: Laurent Series Are Related To Fourier Series

If the Laurent series
$$f(z) = \sum_{n\in\mathbb{Z}} a_n z^n$$
converges uniformly on the unit circle $\{z \in \mathbb{C} : |z| = 1\}$, then the Fourier series expansion of $f(e^{i\theta})$ is
$$f(e^{i\theta}) = \sum_{n\in\mathbb{Z}} a_n e^{in\theta},$$
treated as a function of θ. Hence, the Fourier coefficients of this expansion are $c_n = a_n$.

Now, to solve for the coefficients of the series, suppose our Fourier series from above converges uniformly to $f(e^{i\theta})$, writing
$$f(e^{i\theta}) = \sum_{n\in\mathbb{Z}} c_n e^{in\theta}$$
to denote this. Then notice that if $n \neq k$,
$$\int_0^{2\pi} e^{in\theta}e^{-ik\theta}\, d\theta = \int_0^{2\pi} e^{(i(n-k))\theta}\, d\theta = \left.\frac{e^{i(n-k)\theta}}{i(n-k)}\right|_0^{2\pi} = \frac{e^{i(n-k)2\pi}-1}{i(n-k)} = 0$$
Then in the case where $n = k$, the integral becomes quite trivial and is simply
$$\int_0^{2\pi} e^{in\theta}e^{-ik\theta}\, d\theta = \int_0^{2\pi} d\theta = 2\pi.$$
Since we are looking for the coefficients of our series, this motivates introducing a factor of $1/2\pi$ to make computations cleaner, whence it follows that
$$\int_0^{2\pi} e^{in\theta}e^{-ik\theta}\, \frac{d\theta}{2\pi} = \begin{cases} 0, & n \neq k \\ 1, & n = k \end{cases}.$$
Using this, we have
$$\int_0^{2\pi} f(e^{i\theta})e^{-ik\theta}\, \frac{d\theta}{2\pi} = \int_0^{2\pi} \sum_{n\in\mathbb{Z}} c_n e^{in\theta}e^{-ik\theta}\, \frac{d\theta}{2\pi} = \sum_{n\in\mathbb{Z}} c_n \int_0^{2\pi} e^{in\theta}e^{-ik\theta}\, \frac{d\theta}{2\pi} = c_k,$$
where switching the order of summation and integration is justifiable, as we have assumed uniform convergence. We gather these ideas in a definition.

Definition 1.1: Fourier Coefficients

The Fourier coefficients of any integrable function $f(e^{i\theta})$ are
$$c_n = \int_0^{2\pi} f(e^{i\theta})e^{-in\theta}\, \frac{d\theta}{2\pi}.$$

We then use the notation
$$f(e^{i\theta}) \sim \sum_{n\in\mathbb{Z}} c_n e^{in\theta}$$

to say that $f(e^{i\theta})$ is **associated** to the given Fourier series, being careful to use a tilde as opposed to an equal sign to acknowledge that the series doesn't necessarily converge to $f(e^{i\theta})$ for all θ.

With all of these definitions, we are ready to tackle our first result, which is referred to as Parseval's identity, named after French mathematician Marc Parseval (1755 – 1836).

Theorem 1.3: Parseval's Identity

If $f(e^{i\theta}) \sim \sum c_n e^{in\theta}$ and the series converges uniformly to $f(e^{i\theta})$, then

$$\int_0^{2\pi} |f(e^{i\theta})|^2 \frac{d\theta}{2\pi} = \sum_{n\in\mathbb{Z}} |c_n|^2.$$

Proof Since $f(e^{i\theta}) = \sum_{n\in\mathbb{Z}} c_n e^{in\theta}$ converges uniformly, the conjugate $\overline{f(e^{i\theta})} = \sum_{k\in\mathbb{Z}} \bar{c}_k e^{-ik\theta}$ shall also converge uniformly. Therefore we may write

$$\int_0^{2\pi} |f(e^{i\theta})|^2 \frac{d\theta}{2\pi} = \int_0^{2\pi} f(e^{i\theta})\overline{f(e^{i\theta})} \frac{d\theta}{2\pi} = \int_0^{2\pi} f(e^{i\theta}) \sum_{k\in\mathbb{Z}} \bar{c}_k e^{-ik\theta} \frac{d\theta}{2\pi}.$$

We may then use the uniform convergence to justify writing

$$\int_0^{2\pi} f(e^{i\theta}) \sum_{k\in\mathbb{Z}} \bar{c}_k e^{-ik\theta} \frac{d\theta}{2\pi} = \sum_{k\in\mathbb{Z}} \bar{c}_k \int_0^{2\pi} f(e^{i\theta}) e^{-ik\theta} \frac{d\theta}{2\pi}.$$

Lastly by Definition (1.1), it follows that

$$\int_0^{2\pi} |f(e^{i\theta})|^2 \frac{d\theta}{2\pi} = \sum_{k\in\mathbb{Z}} \bar{c}_k c_k = \sum_{k\in\mathbb{Z}} |c_k|^2,$$

as desired. ■

However, note that although we stated and assumed uniform convergence in Theorem (1.3), it is not strictly necessary. In fact, as long as $f(e^{i\theta})$ satisfies $\int |f(e^{i\theta})|^2 \, d\theta < \infty$ (or for those comfortable, f is L^2), Parseval's identity holds.

1.3 The Basel Problem

After building up our definition of the so-called Fourier series, we are now equipped to solve the Basel Problem (1.1). Recall this asks for the evaluation of the sum of the reciprocals of squared positive integers.

We will now consider the following function:

$$f(e^{i\theta}) = \begin{cases} 1, & 0 < \theta < \pi \\ -1, & \pi < \theta < 2\pi \end{cases}.$$

Let $\mathbb{O} = \{2k+1 : k \in \mathbb{Z}\}$ be the set of odd integers and $\mathbb{E} = \{2k : k \in \mathbb{Z}\}$ the set of even integers. Then the Fourier coefficients of $f(e^{i\theta})$ are given by

$$c_n = -\int_\pi^{2\pi} e^{-in\theta} \frac{d\theta}{2\pi} + \int_0^\pi e^{-in\theta} \frac{d\theta}{2\pi} = \frac{1-(-1)^n}{\pi i n}$$

for $n \neq 0$. If $n = 0$, then we would find that

$$c_0 = -\int_{\pi}^{2\pi} \frac{d\theta}{2\pi} + \int_0^{\pi} \frac{d\theta}{2\pi} = \frac{-2\pi + \pi + \pi}{2\pi} = 0,$$

and if n were even, then $c_n = \frac{1-1}{\pi i n} = 0$. Lastly, if n were odd, then $c_n = \frac{1+1}{\pi i n} = 2/\pi i n$. Therefore the Fourier coefficients are

$$c_n = \begin{cases} 0, & n \in \mathbb{E} \\ 2/\pi i n, & n \in \mathbb{O} \end{cases}.$$

Thus one can see

$$|c_n|^2 = \begin{cases} 0, & n \in \mathbb{E} \\ 4/\pi^2 n^2, & n \in \mathbb{O} \end{cases},$$

and from the definition of $f(e^{i\theta})$, we have that $|f(e^{i\theta})|^2 = 1$ for $0 < \theta < 2\pi$. Now equipped with $|f(e^{i\theta})|^2$ and $|c_n|^2$, this suggests we should look into Parseval's identity. First we have

$$\int_0^{2\pi} |f(e^{i\theta})|^2 \frac{d\theta}{2\pi} = \int_0^{2\pi} \frac{d\theta}{2\pi} = 1 < \infty,$$

so we are good to use Parseval's identity[1]. Then for the coefficients,

$$\sum_{n\in\mathbb{Z}} |c_n|^2 = \sum_{n\in\mathbb{O}} \frac{4}{\pi^2 n^2}.$$

Due to the factor of n^2, we can instead sum over the odd positive integers $\mathbb{O}_{>0}$ and multiply by a factor of 2. Hence

$$\sum_{n\in\mathbb{Z}} |c_n|^2 = 2 \sum_{n\in\mathbb{O}_{>0}} \frac{4}{\pi^2 n^2}.$$

Thus applying Theorem (1.3), it follows that

$$2 \sum_{n\in\mathbb{O}_{>0}} \frac{4}{\pi^2 n^2} = 1.$$

So moving our terms around yields

$$\sum_{n\in\mathbb{O}_{>0}} \frac{1}{n^2} = \frac{\pi^2}{8}.$$

This is quite an interesting result in its own right, though not what we were looking for. However, just a little bit of algebra will reveal the result to us. Notice we can write[2]

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{1}{n^2} &= 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \frac{1}{6^2} + \cdots \\
&= \left(1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots\right) + \left(\frac{1}{2^2} + \frac{1}{4^2} + \frac{1}{6^2} + \cdots\right) \\
&= \sum_{n\in\mathbb{O}_{>0}} \frac{1}{n^2} + \sum_{n\in\mathbb{O}_{>0}} \frac{1}{(2n)^2} \\
&= \frac{\pi^2}{8} + \frac{1}{4} \sum_{n=1}^{\infty} \frac{1}{n^2}
\end{aligned}$$

1. This comes from the fact that $f \in L^2[0, 2\pi]$ and the comment which followed the proof of Parseval's Identity.
2. We are implicitly assuming reordering our terms in this way is allowed. Such a reordering is however valid since our sum is absolutely convergent. See Lemma (4.1).

Therefore, subtracting $\frac{1}{4}\sum \frac{1}{n^2}$ from both sides and multiplying by 4/3 yields

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{4}{3} \cdot \frac{\pi^2}{8} = \frac{\pi^2}{6},$$

as desired.

2 The Riemann Zeta Function

Now, the Basel Problem was only the evaluation of the reciprocals of squared positive integers. If we wanted to study series similar to that of Equation (1.1), but with an arbitrary exponent, it would be rather difficult to convey what series we were studying without consistently writing it in summation notation. Thus, it would be a significant service to give it a special name or label, that way, instead of saying "the evaluation of the Basel Problem" or "the reciprocals of squared positive integers," we could instead say "the evaluation of $f(2)$." German mathematician Georg Riemann (1826 – 1866) did just that, where he in fact extended this study to complex exponents. We follow the typical notation, $s = \sigma + it$ to denote this complex number.

Definition 2.1: Riemann Zeta Function

We define

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s} = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \cdots$$

for $\mathrm{Re}(s) > 1$.

Therefore, we could re-frame the Basel Problem to instead the evaluation of $\zeta(2)$. As it turns out, this mysterious function goes far beyond the evaluation of interesting series and is a fundamental concept in numerous famous problems, perhaps most notably the Riemann Hypothesis and the Prime Number Theorem. Now, in order to better understand this function, we consider another fundamental function.

Definition 2.2: Gamma Function

We define

$$\Gamma(s) := \int_0^{\infty} t^{s-1} e^{-t} \, dt$$

for $\mathrm{Re}(s) > 0$.

This function can be thought of as an extension of the factorial for non-integer input.

Example 2.1: The Gamma Function Acts Like A Factorial

We will evaluate $\Gamma(1)$. By the definition, we have that

$$\Gamma(1) = \int_0^{\infty} t^{1-1} e^{-t} \, dt = -e^{-t}\Big|_0^{\infty} = 1.$$

Therefore, we have $\Gamma(1) = 0! = 1$. We could also equally say $\Gamma(1) = 1!$, however, the reason we say zero factorial will be clear soon.

We can also find a functional equation for the gamma function.

Theorem 2.1: A functional equation for $\Gamma(s)$

If $x \in \mathbb{R}$, then we have

$$\Gamma(x+1) = x\Gamma(x). \tag{2.1}$$

Proof By definition, we have

$$\Gamma(x+1) = \int_0^\infty t^x e^{-t}\, dt. \tag{2.2}$$

Similarly, we have that

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, dt.$$

If we then use integration by parts on Equation(2.2), then we find

$$\Gamma(x+1) = -e^{-t}t^x\Big|_0^\infty + x\int_0^\infty e^{-t}t^{x-1}\, dt = x\Gamma(x),$$

as desired. ∎

The idea of Theorem (2.1) can further be extended to the complex plane in a similar manner. Moreover, Equation (2.1) can be applied inductively to see that

$$\Gamma(n+1) = n!, \tag{2.3}$$

for $n \in \mathbb{Z}_{>0}$ which explains why we wrote $\Gamma(1) = 0!$ in Example (2.1).

Example 2.2: The Gamma Function Extends The Factorial

Now we will evaluate $\Gamma(1/2)$. By definition, we write

$$\Gamma(1/2) = \int_0^\infty t^{-\frac{1}{2}} e^{-t}\, dt.$$

If we then take $t = r^2$ so $dt = 2r\ dr$, this becomes

$$\Gamma(1/2) = \int_{r=0}^\infty r^{-1} e^{-r^2}(2r)\, dr = 2\int_0^\infty e^{-r^2}\, dr = \sqrt{\pi},$$

where we used the standard result (which can be seen by a change of coordinates)

$$\int_{-\infty}^\infty e^{-x^2}\, dx = \sqrt{\pi}$$

and the fact that e^{-r^2} is an even function.

Like the Riemann zeta function, the Gamma function satisfies numerous properties and a great deal can be said about it. We, however, will only collect what we need.

Remark 2.1: Properties of The Gamma Function

- For all s,

$$\Gamma(1-s)\Gamma(s) = \frac{\pi}{\sin(\pi s)}$$

- The gamma function extends to be a meromorphic function over $\mathbb{C}$
- The gamma function has simple poles at $s = 0, -1, -2, \ldots$

2.1 Analytic Continuation

To start our study, we first consider a lemma.

Lemma 2.1: The ML-Estimate

Suppose γ is a piecewise smooth curve, $f(z)$ is continuous on γ, $|f(z)| \leq M$ on γ, and γ has length L. Then,

$$\left| \int_\gamma f(z)\, dz \right| \leq ML.$$

We understand how to describe the Riemann zeta function for $\mathrm{Re}(s) > 1$, but in order to better understand its properties we would like to extend it to be a *meromorphic* function for $\mathrm{Re}(s) \leq 1$. We do so by analytically continuing the Riemann zeta function. We thus consider the function

$$(-z)^{s-1} = e^{(s-1)(\log|-z| + i\,\mathrm{Arg}(-z))},$$

for $z \in \mathbb{C} \setminus [0, \infty)$. We can then make a branch cut along the positive real axis, and express $\zeta(s)$ as a contour integral. Now let $\mathcal{H}$ be the contour as seen in Figure (2) made up of three parts: The top edge of our branch cut extending from $+\infty$ to the second part, a circle of radius $\varepsilon < 2\pi$ centered at zero, and the last part extending from ε back to $+\infty$.

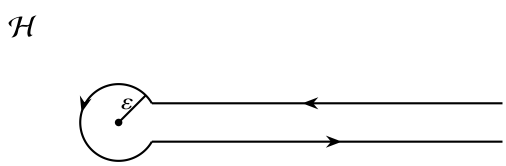

Figure 2

Then let us consider

$$\mathcal{I}(s) = \frac{1}{2\pi i} \int_{\mathcal{H}} \frac{(-z)^{s-1}}{e^z - 1}\, dz,$$

an entire function of s which converges. This then leads us to our first theorem.

Theorem 2.2

For $\mathrm{Re}(s) > 1$, *we have*

$$\zeta(s) = -\Gamma(1-s)\mathcal{I}(s).$$

Proof In order to compute $\mathcal{I}(s)$, we can write

$$\mathcal{I}(s) = \frac{1}{2\pi i} \int_\infty^\varepsilon \frac{e^{(s-1)(\log x - i\pi)}}{e^x - 1}\, dx + \frac{1}{2\pi i} \int_{|z|=\varepsilon} \frac{(-z)^{s-1}}{e^z - 1}\, dz + \frac{1}{2\pi i} \int_\varepsilon^\infty \frac{e^{(s-1)(\log x + i\pi)}}{e^x - 1}\, dx.$$

To clean things up a bit, switch the bounds of integration and sign of the first integral to get

$$\mathcal{I}(s) = \frac{1}{2\pi i} \int_\varepsilon^\infty \frac{e^{(s-1)i\pi} - e^{-(s-1)i\pi}}{e^x - 1} \cdot x^{s-1}\, dx + \frac{1}{2\pi i} \int_{|z|=\varepsilon} \frac{(-z)^{s-1}}{e^z - 1}\, dz.$$

Now we claim that this second integral tends to zero as $\varepsilon \to 0$. In order to see this, first notice that

$$|(-z)^{s-1}| = |e^{(s-1)(\log|-z|+i\,\mathrm{Arg}(-z))}| = e^{\mathrm{Re}(s-1)\log|z|} = |z|^{\mathrm{Re}(s)-1}.$$

Moreover, with $|e^z - 1| \approx |z|$ for z small, there exists some constant $c > 0$ such that for sufficiently small $|z|$, $|e^z - 1| \geq c|z|$. Putting these together yields

$$\left|\frac{(-z)^{s-1}}{e^z - 1}\right| \leq \frac{|z|^{\mathrm{Re}(s)-1}}{c|z|} = \frac{1}{c}|z|^{\mathrm{Re}(s)-2}.$$

Then since our circle of radius ε has length at most $2\pi\varepsilon$, along this path, Lemma (2.1) suggests

$$\left|\frac{1}{2\pi i}\int_{|z|=\varepsilon}\frac{(-z)^{s-1}}{e^z-1}\,dz\right| \leq \frac{\varepsilon^{\mathrm{Re}(s)-1}}{c}$$

which will tend to zero as $\varepsilon \to 0$ since $\mathrm{Re}(s) > 1$. Therefore, letting $\varepsilon \to 0$ in the limit yields

$$\mathcal{I}(s) = \frac{1}{2\pi i}\left(e^{(s-1)i\pi} - e^{-(s-1)i\pi}\right)\int_0^\infty \frac{x^{s-1}}{e^x-1}\,dx = -\frac{\sin(\pi s)}{\pi}\int_0^\infty \frac{x^{s-1}}{e^x-1}\,dx$$

using the fact that $\sin(x+\pi) = -\sin(x)$. Now to evaluate this integral, express part of the denominator of the integrand as a geometric series by writing $1/(e^x-1)$ as $e^{-x}/(1-e^{-x})$. Doing so gives us

$$\mathcal{I}(s) = -\frac{\sin(\pi s)}{\pi}\int_0^\infty\left(\sum_{n=1}^\infty e^{-nx}\right)x^{s-1}\,dx.$$

It can then be seen that series $\sum_{n=1}^\infty e^{-nx}x^{s-1}$ converges uniformly on any compact sub-interval of $(0,\infty)$ and the total integral converges absolutely for $\mathrm{Re}(s) > 1$, so we are justified in switching the order of integration and summation. Then using the change of variables $nx \mapsto t$ and the definition of the gamma function, we find that

$$\mathcal{I}(s) = -\frac{\sin(\pi s)}{\pi}\sum_{n=1}^\infty\left(\int_0^\infty e^{-nx}x^{s-1}\,dx\right) = -\frac{\sin(\pi s)}{\pi}\sum_{n=1}^\infty\left(\frac{1}{n^s}\int_0^\infty e^{-t}t^{s-1}\,dt\right) = -\frac{\sin(\pi s)}{\pi}\Gamma(s)\zeta(s).$$

Then using what we wrote down in Remark(2.1), we find that

$$\mathcal{I}(s) = -\frac{\zeta(s)}{\Gamma(1-s)},$$

as desired. ■

Since $\mathcal{I}(s)$ is an entire function of s and $\Gamma(1-s)$ is a meromorphic function of s on all of $\mathbb{C}$, we may use this equation to extend the definition of $\zeta(s)$ beyond the line $\mathrm{Re}(s) = 1$. We collect this idea in a corollary.

Corollary 2.1: The Riemann Zeta Function Is Meromorphic Over $\mathbb{C}$

We use the definition

$$\zeta(s) = -\Gamma(1-s)\mathcal{I}(s)$$

for $\mathrm{Re}(s) \leq 1$ to extend $\zeta(s)$ to the entire complex plane.

2.2 Functional Equation

The previous theorem allowed us to continue the zeta function analytically throughout the complex plane, but it lacked meaningful information on the behavior of the function for particular inputs.

Lemma 2.2: Uniqueness Principal

If $f(z)$ and $g(z)$ are analytic on a domain D, and if there exists a point $z_0 \in D$ such that there are infinitely many points $z \neq z_0$ with $z \to z_0$ for which $f(z) = g(z)$, then $f(z) = g(z)$ for all $z \in D$.

The next theorem aims to provide a more concrete equation for the zeta function.

Theorem 2.3

For all s, we have

$$\zeta(s) = 2^s \pi^{s-1} \sin\left(\frac{\pi s}{2}\right) \Gamma(1-s)\zeta(1-s). \tag{2.4}$$

Proof First, let us fix $s < 0$ to be a real number. Then consider R_n to be the contour as seen in Figure (3) where the rectangle has vertices $(\pm n, \pm(2n + \frac{1}{2})\pi)$.

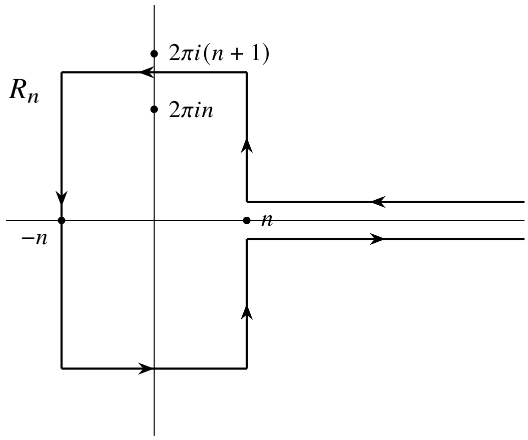

Figure 3

Then we define

$$\mathcal{I}_n(s) = \frac{1}{2\pi i} \int_{R_n} \frac{(-z)^{s-1}}{e^z - 1}\, dz.$$

Now notice that on the right vertical edge when $x = n$,

$$|e^z - 1| \geq ||e^z| - 1| = e^n - 1 > 1.$$

On the left vertical edge, when $x = -n$,

$$|e^z - 1| \geq |1 - |e^z|| = 1 - e^{-n} > \frac{1}{2}.$$

Then on the horizontal edges when $y = \pm(2n + \frac{1}{2})\pi$,

$$|e^z - 1|^2 = 2 - e^z - e^{-z} = 2 > 1.$$

Therefore, on all edges of the rectangle, $|e^z - 1| > 1/2$, implying that the integrand is bounded by $2n^{s-1}$ on the edges. Moreover, since the length of this rectangle is $2(2n + 4n\pi + \pi)$, using the ML-estimates suggests that these integrals are bounded by

$$(2n^{s-1})(4n + 8n\pi + 2\pi) = 8n^s + 16\pi n^s + 4\pi n^{s-1} \leq n^s(8 + 20\pi).$$

Then since we took $s < 0$ earlier, this bound tends to zero as $n \to \infty$, causing $\mathcal{I}_n(s) \to 0$ as $n \to \infty$. However, consider the difference $\mathcal{I}_n(s) - \mathcal{I}(s)$ which is the integral along the closed contour C as seen in Figure (4).

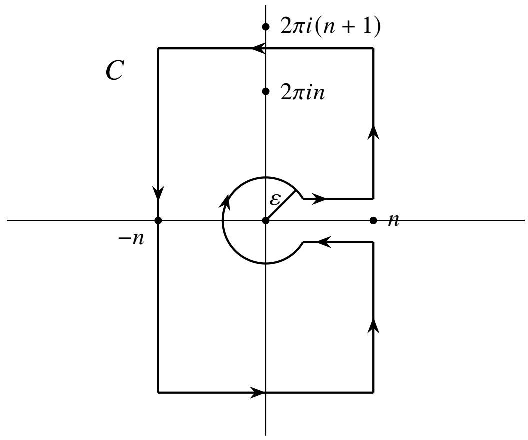

Figure 4

Within this contour, only the poles at the points $\pm 2\pi ik$ for $0 < k \leq n$ of our integrand are contained. Each pole of the integrand is simple with residue

$$\operatorname{Res}\left[\frac{(-z)^{s-1}}{e^z - 1}, 2\pi ik\right] = (-2\pi ik)^{s-1}$$

for positive z (the analogous case holds for negative z). Then, upon combining the results for positive and negative k, using the residue theorem, we have

$$\begin{aligned} \mathcal{I}_n(s) - \mathcal{I}(s) &= \left(e^{(s-1)i\pi/2} - e^{-(s-1)i\pi/2}\right)(2\pi)^{s-1}\sum_{k=1}^{n} k^{s-1} \\ &= 2\sin\left(\frac{\pi s}{2}\right)(2\pi)^{s-1}\sum_{k=1}^{n} k^{s-1} \end{aligned}$$

by the reflection property of cosine. Then letting n tend to ∞, since as we established, $\mathcal{I}_n(s) \to 0$, we get

$$-\mathcal{I}(s) = 2\sin\left(\frac{\pi s}{2}\right)(2\pi)^{s-1}\zeta(1-s). \tag{2.5}$$

Lastly, if we utilize the result from Theorem (2.2), we obtain Equation (2.4) for $s < 0$. If we then let P be the set of poles for both sides of Equation (2.5), then both sides will be analytic on the connected domain $\mathbb{C} \setminus P$. Then, since they agree on the set $\{s \in \mathbb{C} : s < 0\} \setminus P$, by the Uniqueness Principle (2.2), the equation holds for all s. ■

Corollary 2.2

We also have

$$\zeta(1-s) = 2^s \pi^{s-1}\cos\left(\frac{\pi s}{2}\right)\Gamma(s)\zeta(s). \tag{2.6}$$

Proof This follows by plugging $1-s$ into Equation (2.4) and using the cosine reflection property. ■

2.3 Zeta At Even Positive Integers

If we then make the change of variables $z \mapsto -z$ in the definition of $\mathcal{I}(s)$, we have

$$\mathcal{I}(s) = -\frac{1}{2\pi i}\int_{-\mathcal{H}} \frac{z^{s-1}e^z}{1-e^z}\, dz.$$

Then we can make the definition

$$\mathcal{J}(s) := \frac{1}{2\pi i} \int_{-\mathcal{H}} \frac{z^{s-1} e^z}{1 - e^z}\, dz$$

and, as a result of this, we have

$$\zeta(s) = \Gamma(1 - s)\mathcal{J}(s). \tag{2.7}$$

By Equation (2.3), when n is a positive integer, we can explicitly calculate the value of $\zeta(-n)$. Taking $s = -n$ in Equation (2.7) gives us

$$\zeta(-n) = \Gamma(1 + n)\mathcal{J}(-n) = n!\mathcal{J}(-n).$$

Moreover, it follows that

$$\mathcal{J}(-n) = \operatorname{Res}\left[\frac{z^{-n-1} e^z}{1 - e^z}, 0\right], \tag{2.8}$$

and the calculation of this residue is an example of an important classification of functions.

Definition 2.3: Bernoulli Polynomials

For any $\alpha \in \mathbb{C}$ we define the functions $B_n(\alpha)$ by

$$\frac{z e^{\alpha z}}{e^z - 1} = \sum_{n \in \mathbb{Z}_{\geq 0}} \frac{B_n(\alpha)}{n!} z^n, \quad |z| < 2\pi.$$

Then the evaluation of these functions at $\alpha = 0$ gives rise to an interesting class of numbers.

Definition 2.4: Bernoulli Numbers

The numbers $B_n(0)$ are called the Bernoulli numbers and are denoted by B_n. They are defined by

$$\frac{z}{e^z - 1} = \sum_{n \in \mathbb{Z}_{\geq 0}} \frac{B_n}{n!} z^n, \quad |z| < 2\pi.$$

However, as it turns out, we can also describe the Bernoulli numbers when $\alpha = 1$ for particular n. The next theorem covers this.

Theorem 2.4

The Bernoulli polynomials satisfy

$$B_n(\alpha + 1) - B_n(\alpha) = n\alpha^{n-1}$$

for $n \in \mathbb{Z}_{>0}$. Moreover, if $n \geq 2$ we have

$$B_n(0) = B_n(1). \tag{2.9}$$

Proof We can start by considering

$$\sum_{n=0}^{\infty} \frac{B_n(\alpha + 1) - B_n(\alpha)}{n!} \cdot z^n. \tag{2.10}$$

Then applying our definition of the Bernoulli polynomials, this becomes

$$\frac{ze^{(\alpha+1)z}}{e^z-1} - \frac{ze^{\alpha z}}{e^z-1} = ze^{\alpha z} = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} \cdot z^{n+1}.$$

Now comparing the coefficients of z^n in this series with those of Equation (2.10) establishes that $B_n(\alpha+1) - B_n(\alpha) = n\alpha^{n-1}$. Lastly, taking $\alpha = 0$ gives us

$$B_n(1) - B_n(0) = 0,$$

completing the proof. The reason for $n \geq 2$ should be clear by checking what happens when $n = 1$. ■

Now that we've shown $B_n(1) = B_n(0) = B_n$ for $n \geq 2$, we are ready to consider the next result.

Theorem 2.5

For integers $n \geq 2$, we have

$$\zeta(1-n) = -\frac{B_n}{n}.$$

Proof If we take $s = 1 - n$ in Equation (2.7), we obtain

$$\zeta(1-n) = \Gamma(n)\mathcal{J}(1-n) = (n-1)!\mathcal{J}(1-n).$$

Then utilizing our comment from Equation (2.8), we see

$$\mathcal{J}(1-n) = \operatorname{Res}\left[\frac{z^{-n}e^z}{1-e^z}, 0\right] = \operatorname{Res}\left[z^{-n-1}\sum_{m=0}^{\infty}\frac{B_m(1)}{m!}z^m, 0\right] = -\frac{B_n(1)}{n!}.$$

Therefore it follows that

$$(n-1)!\mathcal{J}(1-n) = -\frac{B_n(1)}{n}.$$

Lastly, using Equation (2.9) gives us the desired result. ■

Now equipped with a way to explicitly write down $\zeta(1-n)$ when $n \geq 2$, this gives us hope for writing down $\zeta(n)$. More specifically, since the formula holds for $n \geq 2$, it gives rise to a different approach in solving the Basel problem from § 1.3. Before we do so, we will consider a result which follows directly from Theorem (2.5) and then take a brief detour on the Bernoulli numbers/polynomials.

Corollary 2.3

For $n \in \mathbb{Z}_{>0}$,

$$\zeta(1-2n) = -\frac{B_{2n}}{2n}.$$

As we previously said, there are many useful facts about the Bernoulli numbers/polynomials in application to the zeta function. For example, values of the Bernoulli polynomials can be found recursively.

Theorem 2.6

The functions $B_n(\alpha)$ are given by

$$B_n(\alpha) = \sum_{m=0}^{n} \binom{n}{m} B_m \alpha^{n-m}.$$

Proof We can first start with Definition (2.3). Expanding the left side, we have

$$\frac{ze^{\alpha z}}{e^z - 1} = \left(\sum_{n=0}^{\infty} \frac{B_n}{n!} z^n \right) e^{\alpha z} = \left(\sum_{n=0}^{\infty} \frac{B_n}{n!} z^n \right) \left(\sum_{n=0}^{\infty} \frac{\alpha^n}{n!} z^n \right).$$

Upon opening the brackets and collecting like terms, we see that the coefficient of z^n is

$$\sum_{n=0}^{n} \frac{B_m}{m!} \cdot \frac{\alpha^{n-m}}{(n-m)!}.$$

Therefore, equating the coefficients of z^n, we have that

$$\frac{B_n(\alpha)}{n!} = \sum_{n=0}^{n} \frac{B_m}{m!} \cdot \frac{\alpha^{n-m}}{(n-m)!}$$

Finally , multiplying over $n!$ and using the definition of the binomial coefficient will finish the proof. ■

As it turns out, an analogous result holds for the Bernoulli numbers.

Corollary 2.4

For all integers $n \geq 2$,

$$B_n = \sum_{m=0}^{n} \binom{n}{m} B_m.$$

Proof Take $\alpha = 1$ in Theorem (2.6) and use Equation (2.9). ■

Now that we have a recursive way to determine the values of the Bernoulli numbers, it would be great if we could actually use it. However, with our definition only being defined for $n \geq 2$, we have to do just a little more work to find B_0 and B_1. To do so, we expand

$$\frac{z}{e^z - 1} = 1 - \frac{z}{2} + \ldots$$

and then equate corresponding coefficients for z^0 and z^1. Doing so yields $B_0 = 1$ and $B_1 = -\frac{1}{2}$, so we can finally use Corollary (2.4) to determine the first few Bernoulli numbers. They are

$$B_0 = 1, \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_3 = 0, \quad B_4 = -\frac{1}{30}.$$

Theorem 2.7: The Bernoulli Numbers Are Rational

For all $n \in \mathbb{Z}_{\geq 0}$, $B_n \in \mathbb{Q}$.

Proof We will use *strong* induction on n. As a base case, we have $B_0 = 1 \in \mathbb{Q}$. Then, suppose for some integer $n \geq 1$ *all* previous Bernoulli numbers $B_0, B_1, \ldots, B_{n-1} \in \mathbb{Q}$. We would then like to show that $B_n \in \mathbb{Q}$. First, we write

$$B_{n+1} = \sum_{m=0}^{n+1} \binom{n+1}{m} B_m$$

which then gives

$$\sum_{m=0}^{n} \binom{n+1}{m} B_m = 0$$

after expanding out the very last term of the sum and subtracting it over to the left side. Then, if we again expand out the very last term in the sum, we have

$$(n+1)B_n + \sum_{m=0}^{n-1} \binom{n+1}{m} B_m = 0,$$

or

$$B_n = -\frac{1}{n+1} \sum_{m=0}^{n-1} \binom{n+1}{m} B_m.$$

Then since $n+1 \in \mathbb{Z}_{>0}$ we have $\frac{1}{n+1} \in \mathbb{Q}$. Furthermore, we assume all $B_m \in \mathbb{Q}$ for $m \in \{0, 1, \ldots, n-1\}$ and the binomial coefficients are integers[1], so it follows that $B_n \in \mathbb{Q}$. Finally, putting all of this together proves our claim. ■

There is one last result that will be important to us later, concluding our Bernoulli detour.

Theorem 2.8

All Bernoulli numbers of the form B_{2k+1} for $k \in \mathbb{Z}_{>0}$ are zero.

Proof To see this, we consider the function

$$g(z) := \frac{z}{e^z - 1} + \frac{z}{2} = \frac{z + ze^z}{2(e^z - 1)}.$$

It can then easily be shown that $g(-z) = g(z)$, which in particular implies g is an even function. If we then recognize the definition of the Bernoulli numbers, we have

$$g(z) = \sum_{n \in \mathbb{Z}_{>0}} \frac{B_n}{n!} z^n + \frac{z}{2} = \left(B_0 + B_1 z + \frac{B_2}{2!} z^2 + \ldots \right) + \frac{z}{2}.$$

However, since as we previously saw, since g is even, its Taylor series shall include only even terms. In particular, this implies $B_1 + \frac{1}{2} = 0$, $\frac{B_3}{3!} = 0$, and generally $B_{2k+1}/(2k+1)! = 0$. Therefore $B_n = 0$ for all odd $n > 1$ and we have again confirmed $B_1 = -\frac{1}{2}$. ■

Now that we have a closed form for $\zeta(1-2n)$ and understand the Bernoulli numbers better, we can circle back to evaluations of the zeta function. We'll start by verifying our solution to the Basel problem in § 1.3.

1. See Theorem (3.6) for a discussion of this.

Theorem 2.9: Basel Redux

We have
$$\zeta(2) = \frac{\pi^2}{6}.$$

Proof We can take $s = 2$ in Equation (2.6) to get
$$-\frac{B_2}{2} = 2(2\pi)^{-2}\Gamma(2)\cos\left(\frac{2\pi}{2}\right)\zeta(2).$$
Then moving everything over to isolate $\zeta(2)$ gives us
$$\zeta(2) = \frac{B_2(2\pi)^2}{2^2} = B_2\pi^2.$$
Lastly, using the fact that $B_2 = 1/6$ from earlier, we have that
$$\zeta(2) = \frac{\pi^2}{6},$$
as desired. ■

To finish up this subsection, we'd like to see what the zeta function evaluated at any even positive integer is. The next theorem does just that.

Theorem 2.10

For all $n \in \mathbb{Z}_{>0}$, we have
$$\zeta(2n) = \frac{(-1)^{n+1}2^{2n-1}\pi^{2n}B_{2n}}{(2n)!}. \tag{2.11}$$

Proof Like Theorem (2.9), we can start with Equation (2.6) and take $s = 2n$. Thus,
$$-\frac{B_{2n}}{2n} = 2(2\pi)^{-2n}\Gamma(2n)\cos\left(\frac{2n\pi}{2}\right)\zeta(2n).$$
After isolating $\zeta(2n)$, we have
$$\zeta(2n) = \frac{(-1)^{n+1}B_{2n}(2\pi)^{2n}}{2^2n(2n-1)!},$$
which proves Equation (2.11). ■

2.4 Difficulty With Zeta At Odd Positive Integers

We have now seen a way to write down the zeta function evaluated at any even positive integer of our choice. This naturally motivates us to further develop our understanding of the zeta function evaluated at odd positive integers. Doing so will provide full information on the behavior of this function for positive integer inputs. As a motivating example, let's consider when $s = 1$. We have that
$$\zeta(1) = \sum_{n=1}^{\infty}\frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots.$$
Most will see this series and immediately conclude that it fails to converge. Such a conclusion is indeed true, but let us remind ourselves of *why* it is true. There are numerous ways to go about such a problem, but what is likely the most clear is the following.

Definition 2.5: Partial Sum And Convergence Of An Infinite Sum

Given a series of the form

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + \cdots ,$$

we can write S_n to denote the nth partial sum so that

$$S_n = \sum_{k=1}^{n} a_k = a_1 + a_2 + \cdots + a_n.$$

If the sequence (S_n) is convergent and $\lim S_n = S_\infty$, then we say that the series is convergent and write

$$\sum_{k=1}^{\infty} a_k = S_\infty.$$

If instead the sequence (S_n) is divergent, then we say the series is divergent.

Given this refresher, a reasonable approach would be to show that a partial sum of $\sum \frac{1}{n}$ diverges. As we will see, the choice of powers of two is particularly useful. We have

$$\begin{aligned}
S_2 &= 1 + \frac{1}{2} \\
S_{2^2} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) > 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) \\
S_{2^3} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6}\right) + \left(\frac{1}{7} + \frac{1}{8}\right) > 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8}\right) + \left(\frac{1}{8} + \frac{1}{8}\right) \\
&\vdots \\
S_{2^n} &> 1 + \frac{n}{2}.
\end{aligned}$$

If we then take $n \longrightarrow \infty$, we see the sequence (S_{2^n}) diverges and thus $\zeta(1)$ is divergent[1]. We can in fact take this one step further with the following theorem which we will omit the proof of.

Theorem 2.11

The function $\zeta(s)$ is a meromorphic function on the complex plane with only one pole at $s = 1$ which has residue 1.

What we take from this is that at $s = 1$, not only can $\zeta(s)$ not be written in a nice way compared to $\zeta(2n)$, but it also fails to converge. However, all hope is not lost. If we return to Definition (2.2), we see that the zeta function was originally defined for $\mathrm{Re}(s) > 1$, a property which ensures convergence; one that our last example failed. Thus, the start of our analysis was a rather poor one; but we should be better suited for the next one, where we take $s = 3$. We spent a decent amount of time developing our functional equation, so we

1. We will however see in § 5.1 that a slight modification to this infinite series will converge and is related to the natural logarithm.

would of course like to use it. Doing so, we find that

$$\begin{aligned}\zeta(3) &= 2(2\pi)^{3-1}\sin\left(\frac{3\pi}{2}\right)\Gamma(1-3)\zeta(1-3)\\ &= 8\pi^2(-1)\Gamma(-2)\left(-\frac{B_3}{3}\right).\end{aligned}$$

Recall that $\Gamma(s)$ has a pole at $s=-2$ and $B_3=0$, so the right hand side formally becomes $0\cdot\infty$. That is to say, we cannot naïvely plug in 3 and expect this to directly yield a value for $\zeta(3)$. Perhaps we used the wrong formula. Using our other formula gives

$$\zeta(3) = \frac{\zeta(1-3)(2\pi)^{3-1}}{2\Gamma(3)\cos\left(\frac{3\pi}{2}\right)},$$

which is undefined. Odd, but as it turns out, we already encountered this problem. Back in Theorem (2.8), we showed that for all $k\in\mathbb{Z}_{>0}$, $B_{2k+1}=0$, which shows that the argument used to derive the explicit formula for $\zeta(2n)$ does not extend to odd positive integers. A method failing for a particular case does not mean the zeta function is ill behaved for odd positive integer input, however it may suggest that these inputs are fundamentally different than their even relatives. In the even case, the functional equation related $\zeta(2n)$ to $\zeta(1-2n)$, and the latter could be expressed in terms of Bernoulli numbers. However, for odd positive integers, $\zeta(-2n)$ vanishes which causes this approach to collapse. Although there is difficulty in writing down a closed form for zeta at odd inputs we currently do bear more information on $\zeta(3)$ than the rest of the odd integers; the next section covers this in detail.

3 The Irrationality Of Zeta-3

We will discuss the irrationality of $\zeta(3)$. This result was first introduced by Roger Apéry [Apé79] in 1978. Then in 1979 Dutch mathematician Frits Beukers (1953–) presented his own interpretation of the problem [Beu79], providing a proof using multiple integrals. Many reference a very well-regarded exposition of Apéry's theorem by Dutch mathematician Alfred van der Poorten (1942 – 2010) [Poo79]. We however reference [Ste05, § 4], which in turn cites Poorten's work.

3.1 Apéry's Theorem

Before we begin to discuss the proof that $\zeta(3)$ is irrational, we will first review some core concepts and provide some necessary definitions.

Definition 3.1: Coprime

We say that two integers a and b are coprime if their only common positive integer divisor is 1. That is, $\gcd(a,b)=1$.

The next theorem will allow us to describe integers greater than one.

Theorem 3.1: Fundamental Theorem Of Arithmetic

Every integer greater than 1 is either prime or can (up to permuting factors) be uniquely expressed as the product of primes. We refer to this unique expression as the unique prime factorization of the number.

Then given an integer, we would like to be able to discuss its properties by considering the exponent of each prime found in its unique prime factorization.

Definition 3.2: p-adic Valuation

We will write $\nu_p(n)$ to denote the exponent of the prime p in the unique prime factorization of n.

So, for example, since $9 = 3^2$, we have $\nu_3(9) = 2$ and $\nu_5(9) = 0$. Since, loosely speaking[1], it is just the exponent of a given prime p in the unique prime factorization of n, it should behave like an exponent, which it does. We have

$$\nu_p\left(\frac{n}{m}\right) = \nu_p(n) - \nu_p(m) \tag{3.1}$$

and

$$\nu_p(n \cdot m) = \nu_p(n) + \nu_p(m). \tag{3.2}$$

As we will see later, the factorial of an integer ties this function together with the floor function nicely.

Definition 3.3: Floor Function

We define the floor function of a real number x to be

$$\lfloor x \rfloor = \max\{m \in \mathbb{Z} : m \leq x\}.$$

In words, the floor of x is the greatest integer less than or equal to x.

Although we cannot generally "open the brackets" of the floor function to add elements, when $n \in \mathbb{Z}$, it does follow that

$$\lfloor x + n \rfloor = \lfloor x \rfloor + n. \tag{3.3}$$

In the case that $n \notin \mathbb{Z}$, we instead have the chain of inequalities

$$\lfloor x \rfloor + \lfloor y \rfloor \leq \lfloor x + y \rfloor \leq \lfloor x \rfloor + \lfloor y \rfloor + 1 \tag{3.4}$$

which holds for all $x, y \in \mathbb{R}$.

Remark 3.1

We can also say that $\lfloor x \rfloor = y$ if and only if $y \leq x < y + 1$.

In order to describe the growth of functions, we make use of the following.

Definition 3.4: Big-Oh

We write $f(x) = O(g(x))$ as $x \to \infty$ if there exists a constant $C > 0$ and K such that whenever $x \geq K$, we have $|f(x)| \leq C \cdot |g(x)|$.
Moreover, we will use the notation $f(x) \ll g(x)$ at times which is read "$f(x)$ is much less than $g(x)$," but these may be thought of as interchangeable.

There is also a slightly stronger version of this definition.

1. The formal definition of this is indeed more concrete, but we believe this heuristic is enough to understand this function. For a more in-depth coverage of this function and its property, see any text on *p-adic analysis*.

Definition 3.5: Little-Oh

We write $f(x) = o(g(x))$ as $x \to \infty$ if there exists some nonnegative function h such that $\lim_{x\to\infty} h(x) = 0$ and $|f(x)| \le h(x) \cdot |g(x)|$.

This statement allows us to understand a very important theorem in number theory.

Theorem 3.2: Prime Number Theorem

Let $\pi(x)$ be the number of primes p satisfying $p \le x$. Then we have that

$$\pi(x) = \frac{x}{\log x} + o\left(\frac{x}{\log x}\right).$$

Example 3.1

To understand what $\pi(x)$ is more clearly, note that we have $\pi(2) = 1$, $\pi(3) = 2$, $\pi(\pi) = 2$, and $\pi(17) = 7$.

We are now ready to define the special sequences which will allow us to prove the irrationality of $\zeta(3)$. For $n, k \in \mathbb{Z}$ which satisfy $0 \le k \le n$ and $m \in \mathbb{Z}_{>0}$, consider

$$c_{n,k} := \sum_{m=1}^{n} \frac{1}{m^3} + \sum_{m=1}^{k} \frac{(-1)^{m+1}}{2m^3\binom{n}{m}\binom{n+m}{m}},$$

$$a_n := \sum_{k=0}^{n} c_{n,k} \binom{n}{k}^2 \binom{n+k}{k}^2, \qquad b_n := \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2. \tag{3.5}$$

Of course, with the restrictions we placed on n, m, and k, we have

$$c_{0,k} = \sum_{m=1}^{k} \frac{(-1)^m + 1}{2m^3\binom{n}{m}\binom{n+m}{m}}, \qquad c_{n,0} = \sum_{m=1}^{n} \frac{1}{m^3},$$

and $c_{0,0} = 0$.

To get started, we need a way to go about showing $\zeta(3)$ is irrational. Our first result will give a solid criterion to rely on in order to show a given number is irrational.

Lemma 3.1: Irrationality Criterion

If there are infinitely many coprime integers p, q such that

$$\left|\alpha - \frac{p}{q}\right| < q^{-(1+\delta)} \tag{3.6}$$

for fixed $\delta > 0$, then α is irrational.

Proof Suppose there exists infinitely many coprime p and q for which Equation (3.6) holds for fixed $\delta > 0$, $p \in \mathbb{Z}$, and $q \in \mathbb{Z}_{>0}$. Moreover, for the sake of contradiction, let us suppose α is rational. That is, suppose we can write $\alpha = \frac{a}{b}$, where $a \in \mathbb{Z}$, $b \in \mathbb{Z}_{>0}$, and $aq \ne pb$. Then we have

$$\left|\alpha - \frac{p}{q}\right| = \left|\frac{a}{b} - \frac{p}{q}\right| = \left|\frac{aq - pb}{bq}\right|.$$

Now, with the numerator being a nonzero whole number, it follows that it is at least one, so

$$\frac{1}{bq} \leq \left|\frac{aq - pb}{bq}\right| < q^{-(1+\delta)}.$$

This implies that $q^{\delta} < b$, or that $q < b^{1/\delta}$ and so there are only finitely many q which satisfy this condition. Then for any particular one such q, let us write

$$|p - \alpha q| < q^{-\delta}.$$

This then implies that $p \in (\alpha q - q^{-\delta}, \alpha q + q^{-\delta})$, which is an interval of length $2q^{-\delta} \leq 2$. With p being an integer in an interval of length at most two, this suggests that p can assume finitely many values for this particular q; more importantly, p is finite for each chosen q. Therefore, it must not be the case that there exist infinitely many coprime p and q, contradicting our supposition. ■

Now given a criterion to show a number is irrational, the method is clear: we will aim to show that there are infinitely many coprime integers p and q for which

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

for some fixed $\delta > 0$. To start, we show the ratio of (a_n) and (b_n) converges to $\zeta(3)$.

Lemma 3.2

For all $1 \leq m \leq n$*,*

$$\binom{n}{m}\binom{n+m}{m} \geq n^2.$$

Theorem 3.3

We have

$$\lim \frac{a_n}{b_n} = \zeta(3).$$

Proof If we expand the definition of (a_n), we find

$$a_n = b_n \sum_{m=1}^{n} \frac{1}{m^3} + \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 \sum_{m=1}^{k} \frac{(-1)^{m+1}}{2m^3\binom{n}{m}\binom{n+m}{m}}.$$

Now, notice the first sum is the partial sum of what we are interested in and so it would be nice to estimate the second in some way. Utilizing Lemma (3.2) allows us to make the estimate

$$\sum_{m=1}^{k} \frac{(-1)^{m+1}}{2m^3\binom{n}{m}\binom{n+m}{m}} \ll \frac{1}{n^2} \sum_{m=1}^{k} \frac{1}{m^3} \ll \frac{1}{n^2},$$

where we used the fact that an alternating sum with positive coefficients is less than its standard sum and a partial sum is less than its infinite sum. If we then apply this to (a_n), it follows that

$$a_n = b_n \sum_{m=1}^{n} \frac{1}{m^3} + O\left(\frac{b_n}{n^2}\right).$$

Now in order to incorporate $\zeta(3)$, write

$$\zeta(3) = \sum_{m=1}^{n} \frac{1}{m^3} + \sum_{m=n+1}^{\infty} \frac{1}{m^3}. \tag{3.7}$$

Notice that the first sum is already contained in (a_n), so we would like to estimate the infinite sum by something which tends to zero. We have that

$$\frac{1}{m^3} \le \int_{m-1}^{m} \frac{1}{u^3}\, du,$$

which can quickly be verified by evaluating the integral. In particular, this implies that

$$\frac{1}{(n+1)^3} \le \int_{n}^{n+1} \frac{1}{u^3}\, du.$$

If we inductively carry out the process of adding an additional fraction with the denominator shifted by one and a subsequent correction of the upper bound of integration, we see that

$$\sum_{m=n+1}^{\infty} \frac{1}{m^3} \le \int_{n}^{\infty} \frac{1}{u^3}\, du.$$

Then, by the evaluation of a simple integral, one finds

$$\int_{n}^{\infty} \frac{1}{u^3}\, du = \frac{1}{2n^2} \ll \frac{1}{n^2}.$$

Now, dividing through by b_n and subtracting by $\zeta(3)$ in Equation (3.7),

$$\left| \frac{a_n}{b_n} - \zeta(3) \right| = \left| \sum_{m=1}^{n} \frac{1}{m^3} + O\left(\frac{1}{n^2}\right) - \sum_{m=1}^{n} \frac{1}{m^3} - \sum_{m=n+1}^{\infty} \frac{1}{m^3} \right| \ll \frac{1}{n^2}.$$

Lastly, this implies that there exists some $c > 0$ such that

$$\left| \frac{a_n}{b_n} - \zeta(3) \right| \le c \left| \frac{1}{n^2} \right|.$$

Then, since $\lim \frac{1}{n^2} = 0$, it follows that

$$\lim \left| \frac{a_n}{b_n} - \zeta(3) \right| = 0,$$

as desired. ■

This result builds confidence that (a_n) and (b_n) provide a good rational approximation to $\zeta(3)$, however with the terms of the sequence (a_n) not being integers[2], we cannot yet consider applying Lemma (3.1). For now, we show these sequences have more properties than just approximating $\zeta(3)$ well.

2. We never explicitly showed this. However a simple application of the definition of (a_n) in Equation (3.5) will reveal $a_0 = 0$, $a_1 = 6$, and $a_2 = \frac{351}{4}$.

Theorem 3.4

For all $n \in \mathbb{Z}_{>0}$, the sequence (b_n) satisfies

$$(n+1)^3 b_{n+1} - B(n) b_n + n^3 b_{n-1} = 0$$

where $B(x) = 34x^3 + 51x^2 + 27x + 5$.

Proof We will start by defining

$$\lambda_{n,k} = \binom{n}{k}^2 \binom{n+k}{k}^2 = \frac{(n+k)!^2}{k!^4 (n-k)!^2}$$

and

$$B_{n,k} = 4(2n+1)(2k^2 + k - (2n+1)^2)\lambda_{n,k}.$$

We would then like to first show that for $1 \le k \le n$,

$$B_{n,k} - B_{n,k-1} = (n+1)^3 \lambda_{n+1,k} - B(n)\lambda_{n,k} + n^3 \lambda_{n-1,k}. \tag{3.8}$$

In order to motivate what may look like a random selection of items, notice that $\lambda_{n,k}$ is precisely the summand found in the series definition of (b_n). Moreover, with (b_n) being a series, this motivates summing the above claimed equality at some point, in which one may notice that we will get some nice telescoping behavior. Expanding the definition of $B_{n,k}$ will reveal

$$B_{n,k} - B_{n,k-1} = 4(2n+1)(2k^2 + k - (2n+1)^2)\lambda_{n,k} - 4(2n+1)(2(k-1)^2 + k - 1 - (2n+1)^2)\lambda_{n,k-1}.$$

Thus if we divide out a factor of $\lambda_{n,k}$, we want to instead show the equality.

$$\begin{aligned} 4(2n+1)(2k^2 + k - (2n+1)^2) - 4(2n+1)(2(k-1)^2 + k - 1 - (2n+1)^2)\frac{\lambda_{n,k-1}}{\lambda_{n,k}} \\ = (n+1)^3 \frac{\lambda_{n,k-1}}{\lambda_{n,k}} - B(n) + n^3 \frac{\lambda_{n-1,k}}{\lambda_{n,k}} \end{aligned}$$

is indeed true. If we then expand the definition of $\lambda_{n,k}$, it follows that

$$\frac{\lambda_{n,k-1}}{\lambda_{n,k}} = \frac{k^4}{(n-k+1)^2(n+k)^2}, \qquad \frac{\lambda_{n+1,k}}{\lambda_{n,k}} = \frac{(n+k+1)^2}{(n-k+1)^2}, \qquad \frac{\lambda_{n-1,k}}{\lambda_{n,k}} = \frac{(n-k)^2}{(n+k)^2}.$$

We can then plug these into our last equality. Finally, we can multiply both sides by a factor of $(n-k+1)^2(n+k)^2$ so we only have linear factors. The problem then reduces to showing the following equation holds:

$$\begin{aligned} 4(2n+1)(2k^2 + k - (2n+1)^2) - 4(2n+1)(2(k-1)^2 + k - 1 - (2n+1)^2)k^4 \\ = (n+1)^3(n+k)^2(n-k+1)^2 - B(n)(n+k)^2(n-k+1)^2 + n^3(n-k)^2(n-k+1)^2. \end{aligned} \tag{3.9}$$

Finally, we will delegate this task to a machine[3] which will confirm we do indeed have equality. If we then use this for $k = 0, 1, \ldots, n+1$, since $\binom{n}{k} = 0$ whenever $n < k$ or $k < 0$, it follows that $\lambda_{n,n+1} = \lambda_{n,-1} = \lambda_{n-1,n+1} = \lambda_{n-1,n} = 0$. This implies that $B_{n,n+1} = B_{n,-1} = 0$ and thus notice the beautiful telescoping of

$$B_{n,0} - B_{n,-1} + B_{n,1} - B_{n,0} + \cdots + B_{n,n+1} - B_{n,n} = 0.$$

3. In particular, if one moves all terms to one side of Equation (3.9), using your machine's equivalent of *expand* or *simplify* to test for equality with zero should return true.

Therefore,

$$\begin{aligned} 0 &= (n+1)^3 \sum_{k=0}^{n+1} \lambda_{n+1,k} - B(n) \sum_{k=0}^{n+1} \lambda_{n,k} + n^3 \sum_{k=0}^{n+1} \lambda_{n-1,k} \\ &= (n+1)^3 \sum_{k=0}^{n} \lambda_{n+1,k} - B(n) \sum_{k=0}^{n} \lambda_{n,k} + n^3 \sum_{k=0}^{n} \lambda_{n-1,k}. \end{aligned}$$

Lastly, recall we said that $\lambda_{n,k}$ are precisely the summands in the definition of (b_n), so we're done. ■

The same relation holds for (a_n).

Theorem 3.5

For all $n \in \mathbb{Z}_{>0}$, the sequence (a_n) satisfies

$$(n+1)^3 a_{n+1} - B(n) a_n + n^3 a_{n-1} = 0,$$

where $B(x) = 34x^3 + 51x^2 + 27x + 5$.

Proof First note that from the definition of $(c_{n,k})$, we have

$$\begin{aligned} c_{n,k} - c_{n-1,k} &= \sum_{m=1}^{n} \frac{1}{m^3} + \sum_{m=1}^{k} \frac{(-1)^{m+1}}{2m^3 \binom{n}{m}\binom{n+m}{m}} - \sum_{m=1}^{n-1} \frac{1}{m^3} - \sum_{m=1}^{k} \frac{(-1)^{m+1}}{2m^3 \binom{n-1}{m}\binom{n-1+m}{m}} \\ &= \frac{1}{n^3} + \sum_{m=1}^{k} \left(\frac{(-1)^{m+1}}{2m^3 \binom{n}{m}\binom{n+m}{m}} - \frac{(-1)^{m+1}}{2m^3 \binom{n-1}{m}\binom{n-1+m}{m}} \right). \end{aligned} \tag{3.10}$$

After some algebraic manipulation, this reduces to

$$c_{n,k} - c_{n-1,k} = \frac{1}{n^3} + \sum_{m=1}^{k} \frac{(-1)^m m!^2 (n-m-1)!}{m^2 (n+m)!}.$$

If we then let

$$T_{n,k} = \frac{(-1)^k k!^2 (n-k-1)!}{n^2 (n+k)!},$$

a little more algebra will reveal

$$T_{n,m} - T_{n,m-1} = \frac{(-1)^m m!^2 (n-m-1)!}{m^2 (n+m)!},$$

which is the exact summand we are interested in. Therefore,

$$\begin{aligned} c_{n,k} - c_{n-1,k} = \frac{1}{n^3} + \sum_{m=1}^{k} (T_{n,m} - T_{n,m-1}) &= \frac{1}{n^3} + T_{n,1} - T_{n,0} + \cdots + T_{n,k} - T_{n,k-1} \\ &= \frac{1}{n^3} + T_{n,k} - T_{n,0} \\ &= \frac{1}{n^3} + \frac{(-1)^k k!^2 (n-k-1)!}{n^2 (n+k)!} - \frac{1}{n^3}. \end{aligned}$$

Then a slightly easier identity is the following. By pulling the kth term out and seeing the cancellation for all $k-1$ and below terms, we have

$$c_{n,k} - c_{n-1,k} = \frac{(-1)^{k+1}}{2k^3\binom{n}{k}\binom{n+k}{k}}. \tag{3.11}$$

We shall now define

$$P_{n,k} = (n+1)^3\lambda_{n+1,k}c_{n+1,k} - B(n)\lambda_{n,k}c_{n,k} + n^3\lambda_{n-1,k}c_{n-1,k}.$$

Again, to motivate that these are not completely random selections of objects put together, recognize that $\lambda_{n,k}c_{n,k}$ is precisely the summand which appears in the definition of (a_n). We then claim that

$$P_{n,k} = (B_{n,k} - B_{n,k-1})c_{n,k} + \frac{(-1)^k k!^2(n+1)\lambda_{n+1,k}(n-k)!}{(n+k+1)!} - \frac{(-1)^k k!^2 n\lambda_{n-1,k}(n-k-1)!}{(n+k)!}.$$

To see this, we can start with the definition of $P_{n,k}$ and try and transform it into the right hand side here. Recall Equations (3.8), (3.11), and (3.10), which give us

$$\begin{aligned} P_{n,k} &= (n+1)^3\lambda_{n+1,k}c_{n+1,k} - (-B_{n,k} + B_{n,k-1} + (n+1)^3\lambda_{n+1,k} + n^3\lambda_{n-1,k})c_{n,k} + n^3\lambda_{n-1,k}c_{n-1,k} \\ &= (n+1)^3\lambda_{n+1,k}c_{n+1,k} + (B_{n,k} - B_{n,k-1})c_{n,k} - (n+1)^3\lambda_{n+1,k}c_{n,k} - n^3\lambda_{n-1,k}(c_{n,k} - c_{n-1,k}) \\ &= (B_{n,k} - B_{n,k-1})c_{n,k} + \frac{(-1)^k k!^2(n+1)\lambda_{n+1,k}(n-k)!}{(n+k+1)!} - \frac{(-1)^k k!^2 n\lambda_{n-1,k}(n-k-1)!}{(n+k)!}, \end{aligned}$$

as desired. We will then make the definition

$$A_{n,k} = B_{n,k}c_{n,k} + \frac{5(-1)^{k+1}k(2n+1)}{n(n+1)}\binom{n}{k}\binom{n+k}{k}$$

and claim that

$$\begin{aligned} A_{n,k} - A_{n,k-1} = (B_{n,k} - B_{n,k-1})c_{n,k} &+ \frac{B_{n,k-1}(-1)^{k+1}}{2k^3\binom{n}{k}\binom{n+k}{k}} \\ &+ \underbrace{\frac{5(2n+1)}{n(n+1)}\left[(-1)^{k+1}k\binom{n}{k}\binom{n+k}{k} - (-1)^k(k-1)\binom{n}{k-1}\binom{n+k-1}{k-1}\right]}_{\ell}. \end{aligned}$$

This can be seen by simply writing out the definition of $A_{n,k}$. We have

$$\begin{aligned} A_{n,k} - A_{n,k-1} = B_{n,k}c_{n,k} &+ \frac{5(-1)^{k+1}k(2n+1)}{n(n+1)}\binom{n}{k}\binom{n+k}{k} - B_{n,k-1}c_{n,k-1} \\ &- \frac{5(-1)^k(k-1)(2n+1)}{n(n+1)}\binom{n}{k-1}\binom{n+k-1}{k-1}. \end{aligned}$$

After some simplification, this becomes

$$\begin{aligned} A_{n,k} - A_{n,k-1} &= B_{n,k}c_{n,k} - B_{n,k-1}c_{n,k-1} + \ell \\ &= B_{n,k}c_{n,k} - B_{n,k-1}c_{n,k} + B_{n,k-1}c_{n,k} - B_{n,k-1}c_{n,k-1} + \ell \\ &= (B_{n,k} - B_{n,k-1})c_{n,k} + B_{n,k-1}(c_{n,k} - c_{n,k-1}) + \ell, \end{aligned}$$

which gives the desired result after applying the identity we found in Equation (3.11). Lastly, we claim that $A_{n,k} - A_{n,k-1} = P_{n,k}$. However, we shall again delegate this laborious algebraic check to a machine. Now, for a similar reason to that of the proof of (b_n), we have that $A_{n,n+1} - A_{n,-1} = 0$ where we define $c_{n,k} = 0$ for $k > n$ to avoid division by zero. Notice the telescoping of

$$\begin{aligned} P_{n,0} + P_{n,1} + \cdots + P_{n,n} + P_{n,n+1} &= A_{n,0} - A_{n,1} + \cdots + A_{n,n+1} - A_{n,n} \\ &= A_{n,n+1} - A_{n,-1}. \end{aligned}$$

Therefore we have $\sum_{k=0}^{n+1} P_{n,k} = 0$ which is to say

$$\begin{aligned} 0 &= (n+1)^3 \sum_{k=0}^{n+1} \lambda_{n+1,k} c_{n+1,k} - B(n) \sum_{k=0}^{n+1} \lambda_{n,k} c_{n,k} + n^3 \sum_{k=0}^{n+1} \lambda_{n-1,k} c_{n-1,k} \\ &= (n+1)^3 \sum_{k=0}^{n} \lambda_{n+1,k} c_{n+1,k} - B(n) \sum_{k=0}^{n} \lambda_{n,k} c_{n,k} + n^3 \sum_{k=0}^{n} \lambda_{n-1,k} c_{n-1,k}, \end{aligned}$$

which is precisely what we were looking for. ■

Not only was the sequence (a_n) failing to be a sequence of integers an issue, but the rate of convergence derived from Theorem (3.3) was not as strong as we would like. Thus our goal is to express the sequence (a_n) as an integer and to improve our estimate. We will start with the latter.

Lemma 3.3

For $n \in \mathbb{Z}_{\geq 0}$, the sequence (b_n) is strictly increasing.

Theorem 3.6

For all $n \in \mathbb{Z}_{>0}$,

$$\left| \zeta(3) - \frac{a_n}{b_n} \right| \ll b_n^{-2}.$$

Proof For $n = 1$, a direct computation will show that $\frac{a_n}{b_n} = 1.2$. Therefore, using the fact that $\zeta(3) = 1.202\ldots$, we have

$$\left| \zeta(3) - \frac{a_n}{b_n} \right| \approx 0.002 \ll b_n^{-2}.$$

Thus let us consider $n \geq 2$. For n within this range, we may write

$$n^3 \bigstar_n - B(n-1) \bigstar_{n-1} + (n-1)^3 \bigstar_{n-2} = 0, \tag{3.12}$$

where $(\bigstar_n)$ stands for either sequence (a_n) or (b_n). The idea is to then simultaneously consider Equation (3.12) for both sequences (a_n) and (b_n) and then get rid of the $B(\cdot)$ term by multiplying the equation in (b_n) by a factor of a_{n-1} and vice versa for the equation in (a_n). After shifting around some terms, we will find

$$n^3(a_n b_{n-1} - b_n a_{n-1}) = (n-1)^3(b_{n-2} a_{n-1} - a_{n-2} b_{n-1}). \tag{3.13}$$

Now, if we consider the substitution $n \mapsto k-1$, we will quickly find[4] that $k^3(a_k b_{k-1} - b_k a_{k-1}) = (k-2)^3(b_{k-3} a_{k-2} - a_{k-3} b_{k-2})$. Thus if we carry this out inductively, we will arrive at the result

$$n^3(a_n b_{n-1} - b_n a_{n-1}) = 1^3(b_0 a_1 - a_0 b_1).$$

4. This may not be immediately clear. The idea is that if we apply this substitution in Equation (3.13), then the left hand side becomes $(k-1)^3(b_{k-2} a_{k-1} - a_{k-2} b_{k-1}$. But then by Equation (3.13), this is equal to $k^3(a_k b_{k-1} - b_k a_{k-1})$.

This is rather pleasant, as we now have a way to relate the nth term of sequences (a_n) and (b_n) to their particular values. In this case, simply applying the definition of each sequence will reveal that $a_0 = 0, a_1 = 6, b_0 = 1$, and $b_1 = 5$. Therefore

$$a_n b_{n-1} - b_n a_{n-1} = \frac{6}{n^3}.$$

We then make the definition $r_n = \zeta(3) - a_n/b_n$ so that

$$r_{n-1} - r_n = \frac{a_n b_{n-1} - a_{n-1} b_n}{b_n b_{n-1}} = \frac{6}{n^3 b_n b_{n-1}}.$$

Moreover[5],

$$r_n = r_n - r_{n+1} + r_{n+1} = r_n - r_{n+1} + r_{n+1} - \cdots = \sum_{m=n+1}^{\infty} (r_{m-1} - r_m) \tag{3.14}$$

and therefore

$$\zeta(3) - \frac{a_n}{b_n} = \sum_{m=n+1}^{\infty} (r_{m-1} - r_m) = \sum_{m=n+1}^{\infty} \frac{6}{m^3 b_m b_{m-1}}.$$

As a result of Lemma (3.3), we know that (b_n) is a strictly increasing sequence, so $b_{m-1} \geq b_n$ and $b_m b_{m-1} > b_n^2$. Therefore,

$$\sum_{m=n+1}^{\infty} \frac{6}{m^3 b_m b_{m-1}} < \frac{1}{b_n^2} \sum_{m=n+1}^{\infty} \frac{6}{m^3} \ll b_n^{-2},$$

verifying the claim. ■

This estimate is much better than the linear convergence we had in Theorem (3.3), but with our estimate now depending on the sequence (b_n), this motivates us to analyze its behavior in hopes of further improving our estimate.

Theorem 3.7

The sequence (b_n) satisfies

$$\left(1 + \sqrt{2}\right)^{4n} \ll b_n \ll \left(1 + \sqrt{2}\right)^{4n},$$

which then implies

$$\left|\zeta(3) - \frac{a_n}{b_n}\right| \ll \left(1 + \sqrt{2}\right)^{-8n} \tag{3.15}$$

Proof Recall when $n \geq 2$ we had $n^3 b_n - B(n-1) b_{n-1} + (n-1)^3 b_{n-2} = 0$. That is,

$$n^3 b_n - (34(n-1)^3 + 51(n-1)^2 + 27(n-1) + 5) b_{n-1} + (n-1)^3 b_{n-2} = 0.$$

Now, if we divide through by n^3, we obtain

$$b_n - A_n b_{n-1} + B_n b_{n-2} = 0$$

where

$$A_n = \frac{34(n-1)^3 + 51(n-1)^2 + 27(n-1) + 5}{n^3}, \qquad B_n = \left(\frac{n-1}{n}\right)^3.$$

5. Specifically, recall Definition (2.5) and Theorem (3.3), which, in particular, informs us that the convergence of (r_n) justifies Equation (3.14).

We then have that $\lim A_n = 34$ and $\lim B_n = 1$ and so our limiting characteristic equation is then

$$r^2 - 34r + 1 = 0$$

which gives us roots $r_{1,2} = (1 \pm \sqrt{2})^4$. Now heuristically speaking, since we are interested in the asymptotic growth of (b_n), we expect this limiting characteristic equation to do just that. Furthermore, with $b_n > 0$ for all n, $r_1 = (1+\sqrt{2})^4$ and $0 < r_2 = (1-\sqrt{2})^4 < 1$, the dominant growth shall be governed by r_1. Therefore, there must exist positive constants C_1, C_2 and N such that

$$C_1 r_1^n \leq b_n \leq C_2 r_1^n$$

whenever $n \geq N$. Note that this is precisely $r_1^n \ll b_n \ll r_1^n$ and so we are done. Lastly, Equation (3.15) follows directly from this. ■

This is great as we were able to further improve on our estimate, but we have yet to address the issue of the sequence (a_n) not being a sequence of integers. To amend this issue, we will consider a modified version of (a_n). We require some intermediate steps to start.

Lemma 3.4

The number of multiples of t in the list $\{1, 2, \ldots, n\}$ is $\lfloor \frac{n}{t} \rfloor$, where $t \in \mathbb{Z}_{>0}$.

Proof The multiples of t in the list $\{1, 2, \ldots, n\}$ are of the form jt where $j \in \mathbb{Z}_{>0}$. Therefore, the multiples of t will satisfy the inequality

$$1 \leq jt \leq n$$

which is equivalently

$$\frac{1}{t} \leq j \leq \frac{n}{t}. \tag{3.16}$$

Since $j \in \mathbb{Z}_{>0}$, it has the smallest possible value $j = 1$. Then its largest possible value is the greatest integer less than or equal to n/t by Equation (3.16). That is to say, the largest value of j which will satisfy $j \leq n/t$ is $j = \lfloor n/t \rfloor$. Hence j may take on the values $1, 2, \ldots, \lfloor n/t \rfloor$, of which there are precisely $\lfloor n/t \rfloor$. ■

Lemma 3.5: Legendre's Formula

For a prime number p,

$$\nu_p(n!) = \sum_{k=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor.$$

Proof Recall that $\nu_p(n!)$ is the exponent of the prime p in the unique prime factorization of $n!$. Note that since $n!$ is the product of integers 1 to n, by Lemma (3.4) the number of integers within this range divisible by p is $\lfloor n/p \rfloor$. However, each multiple of p^2 contributes an additional factor of p which was not previously accounted for. The number of such multiples is $\lfloor n/p^2 \rfloor$ and so we should add on $\lfloor n/p^2 \rfloor$. Likewise the multiples of p^3 contributes an additional factor of p which we did not previously consider. The number of such multiples is $\lfloor n/p^3 \rfloor$ and so we should add on $\lfloor n/p^3 \rfloor$. Continuing this argument will give us the desired result. ■

Lemma 3.6

The binomial coefficients are integers.

Proof We will consider the binomial coefficient

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

We would like to show that for an arbitrary prime number p, we have

$$\nu_p(n!) \geq \nu_p((n-k)!) + \nu(k!).$$

Recall what we found in Lemma (3.5) to see that

$$\begin{aligned}\nu_p((n-k)!) + \nu_p(k!) &= \sum_{m=1}^{\infty}\left(\left\lfloor \frac{n-k}{p^m} \right\rfloor + \left\lfloor \frac{k}{p^m} \right\rfloor\right) \\ &\leq \sum_{m=1}^{\infty} \left\lfloor \frac{n-k+k}{p^m} \right\rfloor \\ &= \nu_p(n!).\end{aligned}$$

This shows that not only must every prime in the unique prime factorization of $(n-k)!$ and $k!$ show up in that of $n!$, but the exponent in each subsequent prime of $n!$ must also be greater than or equal to $k!(n-k)!$, implying $\binom{n}{k} \in \mathbb{Z}$. ■

Lemma 3.7

For all $n, m \in \mathbb{Z}$ satisfying $1 \leq m \leq n$,

$$\nu_p \binom{n}{m} \leq \nu_p(\operatorname{lcm}(1, \ldots, n)) - \nu_p(m).$$

Proof First recall the factorial form of the binomial coefficients and Legendre's formula. Thus[6] we have[7]

$$\nu_p \binom{n}{m} = \sum_{k=1}^{\infty}\left(\left\lfloor \frac{n}{p^k} \right\rfloor - \left\lfloor \frac{n-m}{p^k} \right\rfloor - \left\lfloor \frac{m}{p^k} \right\rfloor\right).$$

Now we denote $N := \nu_p(\operatorname{lcm}(1, \ldots, n))$ as the largest integer for which $p^N \leq n$, which implies $p^N \leq n < p^{N+1}$. Then, for $k > N$ we have $p^k > n \geq m$ and $p^k > n - m$, so our infinite sum reduces to the finite sum

$$\nu_p \binom{n}{m} = \sum_{k=1}^{N}\left(\left\lfloor \frac{n}{p^k} \right\rfloor - \left\lfloor \frac{n-m}{p^k} \right\rfloor - \left\lfloor \frac{m}{p^k} \right\rfloor\right).$$

Moreover, let $\nu_p(m) = \ell$. Now for $1 \leq k \leq \ell$, p^k divides m which implies $m/p^k \in \mathbb{Z}$ and so by Equation (3.3)

$$\left\lfloor \frac{n-m}{p^k} \right\rfloor = \left\lfloor \frac{n}{p^k} \right\rfloor - \frac{m}{p^k}.$$

Therefore,

$$\sum_{k=1}^{\ell}\left(\left\lfloor \frac{n}{p^k} \right\rfloor - \left\lfloor \frac{n-m}{p^k} \right\rfloor - \left\lfloor \frac{m}{p^k} \right\rfloor\right) = \sum_{k=1}^{\ell} 0 = 0.$$

6. Noticing the slight abuse of notation.
7. Where we used Equations (3.1), (3.2).

Of course, when $k \geq \ell + 1$, $p^k \nmid m$, however notice we can write[8]

$$\left\lfloor \frac{n}{p^k} \right\rfloor = \left\lfloor \frac{n-m+m}{p^k} \right\rfloor \leq \left\lfloor \frac{n-m}{p^k} \right\rfloor + \left\lfloor \frac{m}{p^k} \right\rfloor + 1.$$

Thus we find that

$$\sum_{k=\ell+1}^{N} \left(\left\lfloor \frac{n}{p^k} \right\rfloor - \left\lfloor \frac{n-m}{p^k} \right\rfloor - \left\lfloor \frac{m}{p^k} \right\rfloor \right) \leq \sum_{k=\ell+1}^{N} 1 = N - \ell.$$

This is precisely what we wanted to show. ■

Now we are ready to consider the result which will ensure we may convert (a_n) to a sequence of integers.

Theorem 3.8

For all $n, k \in \mathbb{Z}$ satisfying $0 \leq k \leq n$,

$$2\,\mathrm{lcm}(1,\ldots,n)^3 \binom{n+k}{k} c_{n,k} \in \mathbb{Z}.$$

Proof We can start by expanding the definition of $c_{n,k}$ to get

$$2\,\mathrm{lcm}(1,\ldots,n)^3 \binom{n+k}{k} c_{n,k} = 2\,\mathrm{lcm}(1,\ldots,n)^3 \binom{n+k}{k} \sum_{k=1}^{n} \frac{1}{m^3} + \mathrm{lcm}(1,\ldots,n)^3 \binom{n+k}{k} \sum_{m=1}^{k} \frac{(-1)^{m+1}}{m^3 \binom{n}{m} \binom{n+m}{m}}.$$

Then since the integers are closed under operations[9], all we really need to show is that

$$\frac{2\,\mathrm{lcm}(1,\ldots,n)^3 \binom{n+k}{k}}{m^3} \in \mathbb{Z}, \qquad \frac{\mathrm{lcm}(1,\ldots,n)^3 \binom{n+k}{k}}{m^3 \binom{n}{m} \binom{n+m}{m}} \in \mathbb{Z}.$$

For the first fraction, since $m \leq n$ this implies that m^3 divides $\mathrm{lcm}(1,\ldots,n)^3$, and since the binomial coefficients are integers by Lemma (3.6), this must be an integer. The second fraction requires a little more work. If we move the $\binom{n+k}{k}$ term to the denominator, then what we need to show is that the exponent of every prime in the numerator is greater than or equal to that of every prime in the denominator. To do so, we will first notice that[10]

$$\binom{n+m}{m} \binom{n+k}{k}^{-1} = \binom{k}{m} \binom{n+k}{k-m}^{-1}$$

and so our denominator becomes

$$m^3 \binom{n}{m} \binom{k}{m} \binom{n+k}{k-m}^{-1},$$

8. Recall Equation (3.4).
9. That is, the sum and product of any integer with another integer shall remain an integer.
10. Which can readily seen by writing both the left hand side and the right in factorial form.

Then using what we found in Lemma (3.7), we have that

$$\begin{aligned}
\nu_p\left(m^3\binom{n}{m}\binom{k}{m}\binom{n+k}{k-m}^{-1}\right) &= \nu_p(m^3)+\nu_p\binom{n}{m}+\nu_p\binom{k}{m}-\nu_p\binom{n+k}{k-m} \\
&\leq 3\nu_p(m)+\nu_p(\operatorname{lcm}(1,\dots,n))-\nu_p(m)+\nu_p(\operatorname{lcm}(1,\dots,n))-\nu_p(m)+0 \\
&\leq \nu_p(m)+2\nu_p(\operatorname{lcm}(1,\dots,n)) \\
&\leq 3\nu_p(\operatorname{lcm}(1,\dots,n)),
\end{aligned}$$

and since the binomial coefficients are integers $\nu_p\binom{n+k}{k-m} \geq 0$ and so we have the upper bound $-\nu_p\binom{n+k}{k-m} \leq 0$. This is exactly what we wanted to show and so we're done. ■

Now it may not be immediately clear how this amended our issue, however, notice

$$2\operatorname{lcm}(1,\dots,n)^3 a_n = 2\operatorname{lcm}(1,\dots,n)^3\left(\sum_{k=0}^{n}\binom{n}{k}^2\binom{n+k}{k}^2 c_{n,k}\right).$$

With the binomial coefficients being integers and the integers being closed under operations, this whole expression must be an integer. This discussion motivates some of our later definitions. For now, we have another claim which at first glance seems rather unrelated, but as we will see, is essential.

Theorem 3.9

For all $\varepsilon > 0$, there exists some $K(\varepsilon)$ such that if $n \geq K(\varepsilon)$, then $\operatorname{lcm}(1,\dots,n) < e^{n(1+\varepsilon)}$.

Proof First note that we can equivalently write the Prime Number Theorem as $\pi(x) = (1+o(1))\frac{x}{\log x}$. Moreover, if $f(n) = o(1)$, then by definition, $\lim f(n) = 0$. That is, for every $\varepsilon > 0$, there exists some constant $K(\varepsilon)$ for which $n \geq K(\varepsilon)$ implies $o(1) < \varepsilon$. Now, with e^x being strictly increasing, greater than one, and $\varepsilon - o(1) > 0$, we have $e^{\varepsilon - o(1)} > 1$. Therefore it follows that

$$e^{\varepsilon} = e^{o(1)}e^{\varepsilon - o(1)} > e^{o(1)}.$$

Then, the primes which divide $\operatorname{lcm}(1,\dots,n)$ are precisely those which satisfy $p \leq n$. If we then let j be the largest integer such that $p^j \leq n$, the inequality $p^j \leq n < p^{j+1}$ immediately follows. Upon taking the base p logarithm of this inequality, we have $j \leq \log_p n < j+1$, which holds if and only if $\lfloor \log_p n \rfloor = j$ by Remark (3.1). Therefore, the exponent of any prime p in $\operatorname{lcm}(1,\dots,n)$ is $\lfloor \log_p n \rfloor$. From this we can write

$$\operatorname{lcm}(1,\dots,n) = \prod_{p\leq n} p^{\lfloor \log_p n \rfloor} \leq \prod_{p \leq n} n.$$

Noticing the last product is the primes $p \leq n$ copies of n, which is the π function from Theorem (3.2), we get $\operatorname{lcm}(1,\dots,n) \leq n^{\pi(n)}$. Lastly, using Theorem (3.2) and our above discussion,

$$\operatorname{lcm}(1,\dots,n) \leq e^{\pi(n)\log n} < e^{n(1+\varepsilon)},$$

completing the proof. ■

We are finally ready to discuss the irrationality of $\zeta(3)$.

Theorem 3.10: Apéry's Theorem

The number $\zeta(3)$ is not rational.

Proof To set the stage, recall our Irrationality Criterion (3.1), which in short, told us that if we can find two integer sequences such that Equation (3.6) holds, then our constant of interest must be irrational. Therefore, to satisfy this integer sequence requirement, we are motivated to make the definitions

$$q_n := 2b_n \operatorname{lcm}(1,\dots,n)^3, \qquad p_n := 2a_n \operatorname{lcm}(1,\dots,n)^3$$

so that $p_n, q_n \in \mathbb{Z}$ and $a_n/b_n = p_n/q_n$. Therefore, the strategy is clear; show

$$\left|\zeta(3) - \frac{p_n}{q_n}\right| \ll q_n^{-(1+\delta)}$$

for some fixed $\delta > 0$. If we then use Theorems (3.7, 3.9), and simplifying the exponential term asymptotically, we have

$$q_n \ll r_1^n e^{n(3+\varepsilon)} = r_1^{n\left(1+\frac{3+\varepsilon}{\log r_1}\right)}.$$

This then tells us that

$$r_1^{-n\left(1+\frac{3+\varepsilon}{\log r_1}\right)} \ll q_n^{-1}.$$

Then considering what we found in Equation (3.15), this motivates us to find the values of $\delta > 0$ for which

$$r_1^{-2n} \ll r_1^{-n\left(1+\frac{3+\varepsilon}{\log r_1}\right)(1+\delta)} \ll q_n^{-(1+\delta)}.$$

That is, we are looking for the δ where

$$2n \geq n\left(1 + \frac{3+\varepsilon}{\log r_1}\right)(1+\delta).$$

After moving some terms around to isolate the δ, we find

$$\delta \leq \frac{\log r_1 - 3 - \varepsilon}{\log r_1 + 3 + \varepsilon}.$$

Then since $\varepsilon > 0$ is arbitrarily small, the acceptable values of δ can be arbitrarily close to

$$\delta = \frac{\log r_1 - 3}{\log r_1 + 3} = 0.080\dots. \tag{3.17}$$

Therefore we can choose $\delta = (\log r_1 - 3)/(\log r_1 + 3) > 0$ in order to satisfy $r_1^{-2n} \ll q_n^{-(1+\delta)}$, proving that $\zeta(3)$ fails to be rational by our Irrationality Criterion. ■

3.2 Beukers' Idea

The idea of this proof is distinct from that of Apéry although the central method remains unchanged. We will consider a sequence which is related to $\zeta(3)$ in some "nice" way. We then study the behavior of said sequence with hopes of controlling the denominators, suppose $\zeta(3)$ is irrational, and derive a necessary contradiction. We will start with some definitions and theorems.

Definition 3.6: Legendre Polynomials

We define

$$P_n(\alpha) := \frac{1}{n!}\frac{d^n}{d\alpha^n}\left(\alpha^n(1-\alpha)^n\right).$$

We will later make use of the simplest case of the so-called arithmetic and geometric mean inequality.

Theorem 3.11: AM-GM inequality

If $x, y \in \mathbb{R}_{\geq 0}$, then

$$x + y \geq 2\sqrt{xy}.$$

Theorem 3.12: Binomial Theorem

We have

$$(x+y)^n = \sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k = \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}.$$

To illustrate Beukers' idea, we will frequently make use of integrals over the unit square $[0,1]^2$ and the unit cube $[0,1]^3$. One may quickly notice that many of the integrals we consider are indeterminate. Thus, to rigorously justify the manipulations that follow, first consider the integrals over the truncated interval $[0, 1-\varepsilon]$ and then let $\varepsilon \to 0^+$ which will recover the original integral.

We first let $k, \ell \in \mathbb{Z}_{\geq 0}$ and $w \in \mathbb{R}_{\geq 0}$. We then have the following claim.

Theorem 3.13

For k, ℓ, w as above, and $k > \ell$,

$$\iint_{[0,1]^2} \frac{x^{k+w}y^{\ell+w}}{1-xy}\,dxdy = \frac{1}{k-\ell}\sum_{j=\ell}^{k-1}\frac{1}{j+w+1}. \tag{3.18}$$

Proof We start by recalling

$$\frac{1}{1-xy} = \sum_{m=0}^{\infty}(xy)^m$$

and so

$$\iint_{[0,1]^2} \frac{x^{k+w}y^{\ell+w}}{1-xy}\,dxdy = \iint_{[0,1]^2}\sum_{m=0}^{\infty}(xy)^m x^{k+w}y^{\ell+w}\,dxdy.$$

Then if we interchange the summation and integral symbols and compute the double integral, it will then become

$$\sum_{m=0}^{\infty}\iint_{[0,1]^2} x^{k+w+m}y^{\ell+w+m}\,dxdy = \sum_{m=0}^{\infty}\frac{1}{(m+k+w+1)(m+\ell+w+1)}. \tag{3.19}$$

If we then let $m + w + 1 = a$ we can use partial fraction decomposition to express our summand as

$$\frac{1}{(k+a)(\ell+a)} = \frac{A}{k+a} + \frac{B}{\ell+a}$$

Upon solving for A and B and pulling out a factor of $1/(k-\ell)$, we arrive at

$$\sum_{m=0}^{\infty} \frac{1}{(m+k+w+1)(m+\ell+w+1)} = \frac{1}{k-\ell}\left(\sum_{m=0}^{\infty} \frac{1}{m+\ell+w+1} - \sum_{m=0}^{\infty} \frac{1}{m+k+w+1}\right).$$

Lastly, if we make the substitutions $j \mapsto m+\ell$ and $j \mapsto m+k$ in the first and second sum respectively, we arrive at the claimed equality. ■

Now notice that if we put $w = 0$, we will obtain

$$\iint_{[0,1]^2} \frac{x^k y^\ell}{1-xy}\, dxdy = \frac{1}{k-\ell}\sum_{j=\ell}^{k-1} \frac{1}{j+1}. \tag{3.20}$$

We then have the following theorem.

Theorem 3.14

The equality

$$\iint_{[0,1]^2} \frac{x^{k+w} y^{\ell+w}}{1-xy} \log(xy)\, dxdy = -\frac{1}{k-\ell}\sum_{j=\ell}^{k-1} \frac{1}{(j+w+1)^2}.$$

holds.

Proof If we differentiate Equation (3.18) with respect to w, on the left hand side, we have

$$\frac{d}{dw}\iint_{[0,1]^2} \frac{x^{k+w} y^{\ell+w}}{1-xy}\, dxdy = \iint_{[0,1]^2} \frac{d}{dw}(xy)^w \frac{x^k y^\ell}{1-xy}\, dxdy = \iint_{[0,1]^2} \frac{x^{k+w} y^{\ell+w}}{1-xy} \log(xy)\, dxdy.$$

Similarly, on the right hand side,

$$\frac{d}{dw}\frac{1}{k-\ell}\sum_{j=\ell}^{k-1} \frac{1}{j+w+1} = \frac{1}{k-\ell}\sum_{j=\ell}^{k-1} \frac{d}{dw}\frac{1}{j+w+1} = -\frac{1}{k-\ell}\sum_{j=\ell}^{k-1} \frac{1}{(j+w+1)^2}.$$

giving us the desired equality. ■

As before, if we let $w = 0$, then we instead have

$$\iint_{[0,1]^2} \frac{x^k y^\ell}{1-xy} \log(xy)\, dxdy = -\frac{1}{k-\ell}\sum_{j=\ell}^{k-1} \frac{1}{(j+1)^2}. \tag{3.21}$$

Now we can analyze the behavior of the denominators of our results. If we assume that $1 \le \ell < k \le n$, then the integer $k-\ell$ will divide $\operatorname{lcm}(1,\dots,n)$. Moreover, each of the integers $(j+1)^2$ for $j \in \{\ell, \ell+1, \dots, k-1\}$ will divide $\operatorname{lcm}(1,\dots,k)^2$. Therefore we have that the product $(k-\ell)(j+1)^2$ will divide $\operatorname{lcm}(1,\dots,k)^3$ for each j. Thus each summand $1/(k-\ell)(j+1)^2$ has denominator dividing $\operatorname{lcm}(1,\dots,k)^3$, and thus the finite sums found in Equations (3.20), (3.21) are rational numbers with denominators dividing $\operatorname{lcm}(1,\dots,k)^3$. Then with $k \le n$, we have $\operatorname{lcm}(1,\dots,k)$ divides $\operatorname{lcm}(1,\dots,n)$.

If we then consider when $k = \ell$ in Equation (3.18) and differentiate with respect to w, on the right hand side, we have

$$\frac{d}{dw}\iint_{[0,1]^2} \frac{(xy)^{k+w}}{1-xy}\, dxdy = \iint_{[0,1]^2} \frac{d}{dw}\frac{(xy)^{k+w}}{1-xy}\, dxdy = \iint_{[0,1]^2} \frac{(xy)^{k+w}}{1-xy} \log(xy)\, dxdy.$$

Then when we consider the series representation in Equation (3.19), differentiate term by term, and let $w = 0$, we have the equality

$$\iint_{[0,1]^2} \frac{(xy)^k}{1-xy} \log(xy)\ dxdy = -2\sum_{m=0}^{\infty} \frac{1}{(m+k+1)^3}.$$

Upon making the substitution $j \mapsto m+k+1$, our series then becomes

$$-2\sum_{m=0}^{\infty} \frac{1}{(m+k+1)^3} = -2\sum_{j=k+1}^{\infty} \frac{1}{j^3} = -2\left(\zeta(3) - \sum_{j=1}^{k} \frac{1}{j^3}\right).$$

We may collect these observations in the following theorem.

Theorem 3.15

The equality

$$\iint_{[0,1]^2} \frac{(xy)^k}{1-xy} \log(xy)\ dxdy = -2\left(\zeta(3) - \sum_{j=1}^{k} \frac{1}{j^3}\right) \tag{3.22}$$

holds.

Now that we were able to explicitly relate our study of these seemingly random integrals to $\zeta(3)$, this motivates us to further investigate their properties in hopes of deriving a result which could eventually lead to a contradiction. Recall we made the effort to define the so-called Legendre polynomials before our analysis; perhaps now would be a good time to make use of our definition.

Theorem 3.16

The equality

$$\iiint_{[0,1]^3} \frac{P_n(x)P_n(y)}{1-(1-xy)z}\ dxdydz = \frac{A_n + B_n\zeta(3)}{\operatorname{lcm}(1,\dots,n)^3}$$

holds where $P_n(\alpha)$ is the Legendre polynomial of α and (A_n), (B_n) are sequences of integers.

Proof Recall that from Lemma (3.12), we may write

$$(1-\alpha)^n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \alpha^k.$$

Thus we can express the Legendre polynomial of α by

$$P_n(\alpha) = \frac{1}{n!}\frac{d^n}{dx^n}\sum_{k=0}^{n} (-1)^k \binom{n}{k} \alpha^{k+n} = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{(n+k)!}{n!\ k!} \alpha^k. \tag{3.23}$$

Then, if we set

$$a_j = (-1)^j \binom{n}{j} \frac{(n+j)!}{n!\ j!},$$

since $\nu_p((n+j)!) \geq \nu_p(n!) + \nu_p(j!)$, we have $a_j \in \mathbb{Z}$. Then, we may write

$$P_n(x)P_n(y) = \sum_{k=0}^{n} a_k x^k \sum_{\ell=0}^{n} a_\ell y^\ell = \sum_{k=0}^{n}\sum_{\ell=0}^{n} a_k a_\ell x^k y^\ell.$$

It then follows that

$$\iint_{[0,1]^2} \frac{-\log(xy)P_n(x)P_n(y)}{1-xy}\,dxdy = \iint_{[0,1]^2} \frac{-\log(xy)}{1-xy}\sum_{k=0}^{n}\sum_{\ell=0}^{n} a_k a_l x^k y^\ell\,dxdy$$
$$= \sum_{k=0}^{n}\sum_{\ell=0}^{n} a_k a_\ell \iint_{[0,1]^2} \frac{-\log(xy)}{1-xy} x^k y^\ell\,dxdy.$$

Then, as we saw before, when $k \neq \ell$, Equation (3.20) results in rational numbers with denominators dividing $\mathrm{lcm}(1,\ldots,n)^3$. Then in the case that $k = \ell$, we get some term involving $\zeta(3)$. We previously saw that $a_k, a_\ell \in \mathbb{Z}$, and so they will not introduce any new denominators. That is to say, there must exist some rational sequences (S_n) and (T_n) for which

$$\iint_{[0,1]^2} \frac{-\log(xy)P_n(x)P_n(y)}{1-xy}\,dxdy = S_n + T_n\zeta(3),$$

where the denominators of both (S_n) and (T_n) divide $\mathrm{lcm}(1,\ldots,n)^3$. Therefore, it follows that

$$\iint_{[0,1]^2} \frac{-\log(xy)P_n(x)P_n(y)}{1-xy}\,dxdy = \frac{A_n + B_n\zeta(3)}{\mathrm{lcm}(1,\ldots,n)^3} \tag{3.24}$$

for some integer sequences (A_n) and (B_n). Lastly, we notice that a quick u-substitution[1] will reveal

$$\int_0^1 \frac{1}{1-(1-xy)z}\,dz = -\frac{\log(xy)}{1-xy}. \tag{3.25}$$

Applying this to our previous integral gets the desired result. ■

Now we have a result which relates our integrals to $\zeta(3)$ and two integer sequences, just as Apéry did in his proof. So we are in good shape, however, it appears we simultaneously made our integral more complicated. Though, as we will see in the next two theorems, a neat trick will tie this all together.

Theorem 3.17

If we apply partial integration n-times with respect to x, then we have

$$\iiint_{[0,1]^3} \frac{P_n(x)P_n(y)}{1-(1-xy)z}\,dxdydz = \iiint_{[0,1]^3} \frac{(xyz)^n(1-x)^n P_n(y)}{(1-(1-xy)z)^{n+1}}\,dxdydz.$$

Proof First, let $g(x) = x^n(1-x)^n$ and

$$f(x) = \frac{1}{1-(1-xy)z},$$

so we want to consider

$$\frac{1}{n!}\int_0^1 g^{(n)}(x)f(x)\,dx.$$

Then a first application of partial integration will yield

$$g^{(n-1)}(x)f(x)\Big|_0^1 - \int_0^1 g^{(n-1)}(x)f(x)\,dx.$$

1. Let $u = 1-(1-xy)z$ so that $du = -(1-xy)\,dz$.

A second application of integration by parts results in

$$g^{(n-1)}(x)f(x)\Big|_0^1 - g^{(n-2)}(x)f'(x)\Big|_0^1 + \int_0^1 g^{(n-2)}(x)f''(x)\,dx.$$

This leads us to make the guess

$$\frac{1}{n!}\int_0^1 g^{(n)}(x)f(x)\,dx = \frac{1}{n!}\left(\sum_{k=0}^{n-1}(-1)^k g^{(n-k-1)}(x)f^{(k)}(x)\Big|_0^1 + (-1)^n\int_0^1 g(x)f^{(n)}(x)\,dx\right), \tag{3.26}$$

which may be proven by induction. We now consider the behavior of the derivatives of f and g in hope of further simplifying this equation. First, notice that

$$f'(x) = -\frac{yz}{(1-(1-xy)z)^2},$$

and that

$$f''(x) = \frac{2(yz)^2}{(1-(1-xy)z)^3}.$$

Therefore we guess that

$$f^{(n)}(x) = \frac{(-1)^n n!(yz)^n}{(1-(1-xy)z)^{n+1}},$$

which can again be shown with induction on n. Then, we can see that $g(x)$ is a polynomial of degree at least n. Therefore, it is readily seen that all derivatives of order less than n will vanish at zero. If we then observe the symmetry $g(x) = x^n(1-x)^n = g(1-x)$ and the derivative

$$\frac{d^s}{dx^s}\,g(1-x) = (-1)^s g^{(s)}(1-x) = g^{(s)}(x),$$

we find that when $s < n$, $g^{(s)}(1) = (-1)^s g^{(s)}(1-1) = 0$. Therefore, we have shown that the term

$$\sum_{k=0}^{n-1}(-1)^k g^{(n-k-1)}(x)f^{(k)}(x)\Big|_0^1 = 0.$$

Lastly, substituting what we found for $f^{(n)}$ into Equation (3.26) will complete the proof. ■

Given this, we can then consider the substitution $w = (1-z)/(1-(1-xy)z)$ where an application of the quotient rule will give

$$dw = -\frac{xy}{(1-(1-xy)z)^2}\,dz.$$

Then when $z = 0$, one can see that $w = 1$, and when $z = 1$, $w = 0$. Therefore, this substitution would, of course, flip the bounds of our outermost integral, however, with this substitution, we also introduced a minus sign, in which they effectively cancel each other out, so we may continue to consider our integrand over the unit cube $[0,1]^3$. So we have the following.

Lemma 3.8

Making the substitution w as discussed above results in

$$\iiint_{[0,1]^3}\frac{(xyz)^n(1-x)^n P_n(y)}{(1-(1-xy)z)^{n+1}}\,dxdydz = \iiint_{[0,1]^3}\frac{(1-x)^n(1-w)^n P_n(y)}{1-(1-xy)w}\,dxdydw.$$

Proof This is easily seen by showing that both integrands are the exact same. Plugging w into

$$\frac{(1-x)^n(1-w)^n}{1-(1-xy)w}$$

will do the trick. Then one will find we are missing a factor of xy in the numerator and two factors of $1-(1-xy)z$ in the denominator, however these are precisely what the differential dw introduces. ■

Now we perform a similar trick to the last theorem in order to better describe the behavior of these integrals.

Theorem 3.18

If we apply partial integration n-times with respect to y, then we have

$$\iiint_{[0,1]^3} \frac{P_n(x)P_n(y)}{1-(1-xy)z}\, dxdydz = \iiint_{[0,1]^3} \frac{(xyw)^n(1-x)^n(1-y)^n(1-w)^n}{(1-(1-xy)w)^{n+1}}\, dxdydw.$$

Proof We will first let $g(y) = y^n(1-y)^n$ and

$$f(y) = \frac{1}{1-(1-xy)w}.$$

Then the exact same analysis as in Theorem (3.17) will reveal

$$\frac{1}{n!}\int_0^1 g^{(n)}(x)f(x)\, dx = \int_0^1 \frac{(xyw)^n(1-y)^n}{(1-(1-xy)w)^{n+1}}\, dx.$$

A more important justification, however, is the interchanging of our variables of integration. For this second application of n integration by parts with respect to y to be of any use to our integral in its current state, we must implicitly switch the variables of integration from x to y temporarily, and then back to x. ■

Of course, this equality then implies that we can write

$$\iiint_{[0,1]^3} \frac{(xyw)^n(1-x)^n(1-y)^n(1-w)^n}{(1-(1-xy)w)^{n+1}}\, dxdydw = \frac{A_n + B_n\zeta(3)}{\text{lcm}(1,\ldots,n)^3},$$

as we said before, the analysis of this proof. The method for showing $\zeta(3)$ is irrational typically involves a contradiction at some point, however, our current integral representations will not directly allow us to derive any specific contradiction. The following theorem will get us one step closer.

Theorem 3.19

For x, y, w as defined throughout this analysis,

$$\frac{xyw(1-x)(1-y)(1-w)}{1-(1-xy)w} \le \frac{1}{27}.$$

Proof First, let us note that $w(1-w) = w - w^2$ is a parabola opening downward and thus obtains its maximum at its vertex. In this case, that vertex is at $\frac{-1}{-2}$ in which $w(1-w) = \frac{1}{4}$, so $w(1-w) \le \frac{1}{4}$. Now, by Theorem (3.11), if we write $1-(1-xy)w = (1-w) + wxy$, then we have

$$1-(1-xy)w \ge 2\sqrt{w(1-w)xy}.$$

Therefore, we have

$$\frac{xyw(1-x)(1-y)(1-w)}{1-(1-xy)w} \leq \frac{xyw(1-x)(1-y)(1-w)}{2\sqrt{w(1-w)xy}} \leq \frac{\sqrt{x}(1-x)\sqrt{y}(1-y)}{4}.$$

Then for some u,

$$\frac{d}{du}\sqrt{u}(1-u) = \frac{1}{2\sqrt{u}} - \frac{3}{2}\sqrt{u}.$$

Upon setting this equal to zero and solving for u, one finds that $u = \frac{1}{3}$. We then have

$$\frac{d^2}{du^2}\sqrt{u}(1-u) = -\left(\frac{1}{4u^{3/2}} + \frac{3}{4\sqrt{u}}\right) < 0$$

for all u. Thus, the maximum of $\sqrt{u}(1-u)$ will be at $u = \frac{1}{3}$ in which

$$\sqrt{x}(1-x)\sqrt{y}(1-y) \leq \frac{1}{3}\left(1-\frac{1}{3}\right)^2 = \frac{4}{27}.$$

The claimed inequality directly follows after this. ■

Now that we have a numerical bound on our integrand, we have a chance to derive some sort of contradiction.

Theorem 3.20: Beukers' Theorem

The number $\zeta(3)$ is irrational.

Proof As a result of Theorems (3.19, 3.18), we have the inequality

$$\frac{A_n + B_n\zeta(3)}{\operatorname{lcm}(1,\ldots,n)^3} \leq \frac{1}{27^n}\iiint_{[0,1]^3}\frac{dxdydw}{1-(1-xy)w}.$$

Moreover, if we set $k = 0$ in Equation (3.22), then we get

$$\iint_{[0,1]^2}\frac{-\log(xy)}{1-xy}\,dxdy = 2\zeta(3).$$

Then, applying Equation (3.25) yields

$$\frac{A_n + B_n\zeta(3)}{\operatorname{lcm}(1,\ldots,n)^3} \leq \frac{1}{27^n}\iint_{[0,1]^2}\frac{-\log(xy)}{1-xy}\,dxdy = \frac{2\zeta(3)}{27^n}.$$

Then, since Equation (3.24) is nonzero, we can write

$$0 < |A_n + B_n\zeta(3)| < \frac{2\zeta(3)\operatorname{lcm}(1,\ldots,n)^3}{27^n}.$$

We then recall that as a result of Theorem (3.9), for $\varepsilon = 0.1$, there exists some $K(\varepsilon)$ such that whenever $n \geq K(\varepsilon)$, $\operatorname{lcm}(1,\ldots,n)^3 < e^{n(3+0.1)} < 23^n$. Moreover using the fact that $\zeta(3) = 1.202\cdots < 2$, we have the bound

$$0 < |A_n + B_n\zeta(3)| < 4\left(\frac{23}{27}\right)^n.$$

Now, for the sake of contradiction, suppose that $\zeta(3)$ were rational. More specifically, suppose $\zeta(3) = \frac{a}{b}$ where $a \in \mathbb{Z}$ and $b \in \mathbb{Z}_{>0}$. This would then imply

$$|A_n + B_n\zeta(3)| = \frac{|bA_n + aB_n|}{b}$$

where $|bA_n + aB_n|$ is a nonzero integer. As a result, we have

$$\frac{1}{b} \leq |A_n + B_n\zeta(3)| < 4\left(\frac{23}{27}\right)^n.$$

Now, we can let $n \to \infty$. Notice the rightmost side must tend to zero, while the leftmost side will remain unchanged. This suggests as $n \to \infty$, $|A_n + B_n\zeta(3)|$ is simultaneously bounded above by zero, strictly positive, and not zero, a contradiction. ∎

4 The Irrationality Of Zeta-2

Of course as we have said many times throughout, we have known $\zeta(2)$ is irrational since 1794 due to the Basel Problem and Legendre's proof that π^2 is irrational. We will discuss the irrationality of $\zeta(2)$ as well, however the interesting part of the proceeding proofs is that we will make no assumption of $\zeta(2)$'s closed form. That is, the novelty lies in that our argument will be independent of π.

4.1 Apéry's Proof

Using Apery's method to prove that $\zeta(2)$ is irrational follows the same structure as proving $\zeta(3)$ is irrational. The biggest change we'll notice is the use of different defining sequences, the use of a different factor to clear the denominators of our numerator sequence, and a different recurrence relation for our sequences. Otherwise, the proof follows the same structure as that of $\zeta(3)$. We start by defining

$$\tilde{c}_{n,k} := \sum_{m=1}^{n} \frac{2(-1)^{m+1}}{m^2} + \sum_{m=1}^{k} \frac{(-1)^{m+n+1}}{m^2\binom{n}{m}\binom{n+m}{m}},$$
$$\tilde{a}_n := \sum_{k=0}^{n} \tilde{c}_{n,k}\binom{n}{k}^2\binom{n+k}{k}, \qquad \tilde{b}_n := \sum_{k=0}^{n}\binom{n}{k}^2\binom{n+k}{k}. \tag{4.1}$$

Now, like in the proof of $\zeta(3)$, we claim the ratio of $(\tilde{a}_n)$ and $(\tilde{b}_n)$ converges to $\zeta(2)$. We however require a lemma first.

Lemma 4.1: Series Rearrangements

If the series $\sum x_n$ is absolutely convergent, then any reordering of its terms will result in the same value.

Theorem 4.1

We have

$$\lim \frac{\tilde{a}_n}{\tilde{b}_n} = \zeta(2).$$

Proof If we expand the definition of $(\tilde{a}_n)$, we find

$$\tilde{a}_n = 2\tilde{b}_n \sum_{m=1}^{n} \frac{(-1)^{m+1}}{m^2} + \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k} \sum_{m=1}^{k} \frac{(-1)^{m+n+1}}{m^2 \binom{n}{m}\binom{n+m}{m}}.$$

Now, notice the first sum is the partial sum of what we're interested in and so it'd be nice to estimate the second in some way. Utilizing Lemma (3.2) allows us to make the estimate

$$\sum_{m=1}^{k} \frac{(-1)^{m+n+1}}{m^2 \binom{n}{m}\binom{n+m}{m}} \ll \frac{1}{n^2} \sum_{m=1}^{k} \frac{1}{m^2} \ll \frac{1}{n^2},$$

where we used the fact that an alternating sum with positive coefficients is less than its standard sum and a partial sum is less than its infinite sum. If we then apply this to $(\tilde{a}_n)$, it follows that

$$\tilde{a}_n = 2\tilde{b}_n \sum_{m=1}^{n} \frac{(-1)^{m+1}}{m^2} + O\left(\frac{\tilde{b}_n}{n^2}\right).$$

Then

$$\sum_{m=1}^{\infty} \left| \frac{(-1)^{m+1}}{m^2} \right| = \sum_{m=1}^{\infty} \frac{1}{m^2} = \zeta(2),$$

so this term is absolutely convergent and therefore by Lemma (4.1), any rearrangement of it will not change its value. Thus we write

$$\zeta(2) = \sum_{m=1}^{\infty} \frac{1}{m^2} = \sum_{m \in \mathbb{E}_{>0}} \frac{1}{m^2} + \sum_{m \in \mathbb{O}_{>0}} \frac{1}{m^2}. \tag{4.2}$$

We then have that

$$\sum_{m \in \mathbb{E}_{>0}} \frac{1}{m^2} = \frac{1}{4}\zeta(2).$$

If we then shift some terms around in Equation (4.2)), we can also find the sum over the odd values. We then have

$$\sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m^2} = 1 - \frac{1}{2^2} + \frac{1}{3^2} - \cdots = \sum_{m \in \mathbb{O}_{>0}} \frac{1}{m^2} - \sum_{m \in \mathbb{E}_{>0}} \frac{1}{m^2},$$

which implies that

$$\zeta(2) = \sum_{m=1}^{\infty} \frac{2(-1)^{m+1}}{m^2}.$$

Now in order to incorporate $\zeta(2)$, write

$$\zeta(2) = \sum_{m=1}^{n} \frac{2(-1)^{m+1}}{m^2} + \sum_{m=n+1}^{\infty} \frac{2(-1)^{m+1}}{m^2}.$$

Notice that the first sum is already contained in $(\tilde{a}_n)$, so we would like to estimate the infinite sum by something which tends to zero. We have that

$$\left| \sum_{m=n+1}^{\infty} \frac{2(-1)^{m+1}}{m^2} \right| \leq \frac{2}{(n+1)^2} \ll \frac{1}{n^2}.$$

Thus it follows that

$$\left|\frac{\tilde{a}_n}{\tilde{b}_n} - \zeta(2)\right| = \left|\sum_{m=1}^{n} \frac{2(-1)^{m+1}}{m^2} + O\left(\frac{1}{n^2}\right) - \sum_{m=1}^{n} \frac{2(-1)^{m+1}}{m^2} - \sum_{m=n+1}^{\infty} \frac{2(-1)^{m+1}}{m^2}\right| \ll \frac{1}{n^2}.$$

That is to say there exists some $c > 0$ and N such that whenever $n \geq N$,

$$\left|\frac{\tilde{a}_n}{\tilde{b}_n} - \zeta(3)\right| \leq c\left|\frac{1}{n^2}\right|.$$

Then since $\lim \frac{1}{n^2} = 0$, it follows that

$$\lim\left|\frac{\tilde{a}_n}{\tilde{b}_n} - \zeta(2)\right| = 0,$$

as desired. ∎

Then as one will notice, in these Apéry-like sequences, both the numerator and denominator sequences $(\tilde{a}_n)$ and $(\tilde{b}_n)$ respectively satisfy the same recurrence relation.

Theorem 4.2

If $(\tilde{\bigstar}_n)$ represents either sequence $(\tilde{a}_n)$ or $(\tilde{b}_n)$, then for all $n \in \mathbb{Z}_{>0}$,

$$(n+1)^2\tilde{\bigstar}_{n+1} - B(n+1)\tilde{\bigstar}_n - n^2\tilde{\bigstar}_{n-1} = 0$$

where $B(n) = 11n^2 - 11n + 3$.

We will omit this proof as it mostly relies on numerous messy algebraic manipulations. However, it is not too difficult to compare the recursion to the sequence definition and at least convince oneself of the formula.

We will then utilize the fact that both our numerator and denominator sequences satisfy the same recurrence relation to improve our previous estimate on the behavior of the absolute error between $\zeta(2)$ and its approximating sequence.

Lemma 4.2

For $n \geq 1$, the sequence $\tilde{b}_n$ is strictly increasing.

Theorem 4.3

For all $n \in \mathbb{Z}_{>0}$,

$$\left|\zeta(2) - \frac{\tilde{a}_n}{\tilde{b}_n}\right| \ll \tilde{b}_n^{-2}.$$

Proof For $n = 1$, a direct computation will show that $\frac{\tilde{a}_n}{\tilde{b}_n} < 1.7$. Therefore, using the fact that $\zeta(2) = 1.644\ldots$, we have

$$\left|\zeta(2) - \frac{\tilde{a}_n}{\tilde{b}_n}\right| \approx 0.02 \ll \tilde{b}_n^{-2}.$$

Thus let us consider when $n \geq 2$. For n within this range, we may write

$$n^2\tilde{\bigstar}_n - B(n)\tilde{\bigstar}_{n-1} - (n-1)^2\tilde{\bigstar}_{n-2} = 0,$$

The idea is to then simultaneously consider Equation (4.1)) for both sequences $(\tilde{a}_n)$ and $(\tilde{b}_n)$ and then get rid of the $B(\cdot)$ term by multiplying the equation in $(\tilde{b}_n)$ by a factor of $(\tilde{a}_{n-1})$ and vice versa for the equation in $(\tilde{a}_n)$. After shifting around some terms, we will find

$$n^2(\tilde{a}_n\tilde{b}_{n-1} - \tilde{b}_n\tilde{a}_{n-1}) = (n-1)^2(\tilde{b}_{n-2}\tilde{a}_{n-1} - \tilde{a}_{n-2}\tilde{b}_{n-1}).$$

Now, if we consider the substitution $n \mapsto k-1$, we will quickly find that $n^2(\tilde{a}_n\tilde{b}_{n-1} - \tilde{b}_n\tilde{a}_{n-1}) = (k-2)^2(\tilde{b}_{k-3}\tilde{a}_{k-2} - \tilde{a}_{k-3}\tilde{b}_{k-2})$. Thus if we carry this out inductively, we will arrive at the result

$$n^2(\tilde{a}_n\tilde{b}_{n-1} - \tilde{b}_n\tilde{a}_{n-1}) = 1^2(\tilde{b}_0\tilde{a}_1 - \tilde{a}_0\tilde{b}_1).$$

This is rather pleasant, as we now have a way to relate the nth terms of sequences $(\tilde{a}_n)$ and $(\tilde{b}_n)$ to their particular values. In this case, simply applying the definition of each sequence will reveal that $\tilde{a}_0 = 0, \tilde{a}_1 = 5, \tilde{b}_0 = 1$, and $\tilde{b}_1 = 3$. Therefore

$$\tilde{a}_n\tilde{b}_{n-1} - \tilde{b}_n\tilde{a}_{n-1} = \frac{5}{n^2}.$$

We then make the definition $\tilde{r}_n = \zeta(2) - \tilde{a}_n/\tilde{b}_n$ so that

$$\tilde{r}_{n-1} - \tilde{r}_n = \frac{\tilde{a}_n\tilde{b}_{n-1} - \tilde{a}_{n-1}\tilde{b}_n}{\tilde{b}_n\tilde{b}_{n-1}} = \frac{5}{n^2\tilde{b}_n\tilde{b}_{n-1}}.$$

Moreover[1],

$$\tilde{r}_n = \tilde{r}_n - \tilde{r}_{n+1} + \tilde{r}_{n+1} = \tilde{r}_n - \tilde{r}_{n+1} + \tilde{r}_{n+1} - \cdots = \sum_{m=n+1}^{\infty} (\tilde{r}_{m-1} - \tilde{r}_m) \tag{4.3}$$

and therefore

$$\left|\zeta(2) - \frac{\tilde{a}_n}{\tilde{b}_n}\right| = \left|\sum_{m=n+1}^{\infty} (\tilde{r}_{m-1} - \tilde{r}_m)\right| = \left|\sum_{m=n+1}^{\infty} \frac{5}{m^2\tilde{b}_m\tilde{b}_{m-1}}\right|.$$

As a result of Lemma (4.2), we know that $\tilde{b}_n$ is a sequence of strictly increasing integers, and since we have that $\tilde{b}_{m-1} \geq \tilde{b}_n$, this implies $\tilde{b}_m\tilde{b}_{m-1} > \tilde{b}_n^2$. Therefore,

$$\left|\sum_{m=n+1}^{\infty} \frac{5}{m^2\tilde{b}_m\tilde{b}_{m-1}}\right| < \frac{1}{\tilde{b}_n^2} \sum_{m=n+1}^{\infty} \frac{5}{m^2} \ll \tilde{b}_n^{-2},$$

verifying the claim. ■

Now that our estimate of absolute error depends on $(\tilde{b}_n)$, like in the proof of $\zeta(3)$, we are motivated to analyze the growth of $(\tilde{b}_n)$ in order to improve our estimate even more.

Theorem 4.4

The sequence $(\tilde{b}_n)$ satisfies

$$\left(\frac{11+5\sqrt{5}}{2}\right)^n \ll \tilde{b}_n \ll \left(\frac{11+5\sqrt{5}}{2}\right)^n,$$

1. Specifically, recall Definition (2.5) and Theorem (4.1), which, in particular, informs us that the convergence of $\tilde{r}_n$ justifies (4.3)).

which then implies

$$\left|\zeta(2) - \frac{\tilde{a}_n}{\tilde{b}_n}\right| \ll \left(\frac{11+5\sqrt{5}}{2}\right)^{-2n} \tag{4.4}$$

Proof Recall when $n \geq 2$ we had $n^2\tilde{b}_n - B(n)\tilde{b}_{n-1} - (n-1)^2\tilde{b}_{n-2} = 0$. That is,

$$n^2\tilde{b}_n - (11n^2 - 11n + 3)\tilde{b}_{n-1} + (n-1)^2\tilde{b}_{n-2} = 0.$$

Now, since we are interested in the asymptotic behavior of $\tilde{b}_n$, we will divide through by n^2 and then consider

$$\tilde{b}_n - A_n\tilde{b}_{n-1} + B_n\tilde{b}_{n-2} = 0$$

where

$$A_n = \frac{11n^2 - 11n + 3}{n^2}, \qquad B_n = \left(\frac{n-1}{n}\right)^2.$$

We then have that $\lim A_n = 11$ and $\lim B_n = 1$ and so our limiting characteristic equation is then

$$\tilde{r}^2 - 11\tilde{r} + 1 = 0.$$

which gives us roots $\tilde{r}_{1,2} = \frac{11\pm3\sqrt{13}}{2}$. Then every nontrivial solution of the recurrence will grow asymptotically like a linear combination of $\tilde{r}_1^n$ and $\tilde{r}_2^n$. Furthermore, with $\tilde{b}_n > 0$ for all n, $\tilde{r}_1 = \frac{11+3\sqrt{13}}{2}$ and $0 < \tilde{r}_2 = \frac{11-3\sqrt{13}}{2} < 1$, the dominant growth shall be governed by $\tilde{r}_1$. Therefore, there must exists positive constants C_1, C_2 and N such that

$$C_1\tilde{r}_1^n \leq b_n \leq C_2\tilde{r}_1^n$$

whenever $n \geq N$. Note that this is precisely $\tilde{r}_1^n \ll \tilde{b}_n \ll \tilde{r}_1^n$ and so we are done. Lastly, Equation (4.4)) follows directly from this. ■

Now, if we have any hopes of using The Irrationality Criterion (3.1), we require both sequences $(\tilde{a}_n)$, $(\tilde{b}_n)$ to be sequences of integers. By Lemma (3.6), it is readily seen that $(\tilde{b}_n)$ is a sequence of integers, however, $(\tilde{a}_n)$ is only a rational sequence. We thus introduce an additional factor to clear the denominators of $(\tilde{a}_n)$ so that this scaled version is a sequence of integers.

Theorem 4.5

For all $n, k \in \mathbb{Z}$ satisfying $0 \leq k \leq n$,

$$2\,\mathrm{lcm}(1,\ldots,n)^2\binom{n+k}{k}\tilde{c}_{n,k} \in \mathbb{Z}.$$

Proof We can start by expanding the definition of $\tilde{c}_{n,k}$ to get

$$\begin{aligned}2\,\mathrm{lcm}(1,\ldots,n)^2\binom{n+k}{k}\tilde{c}_{n,k} &= 4\,\mathrm{lcm}(1,\ldots,n)^2\binom{n+k}{k}\sum_{k=1}^{n}\frac{(-1)^{m+1}}{m^2}\\ &\quad + 2\,\mathrm{lcm}(1,\ldots,n)^2\binom{n+k}{k}\sum_{m=1}^{k}\frac{(-1)^{m+n+1}}{m^2\binom{n}{m}\binom{n+m}{m}}.\end{aligned}$$

Then since the integers are closed under addition, all we really need to show is that

$$\frac{4\operatorname{lcm}(1,\ldots,n)^2\binom{n+k}{k}}{m^2} \in \mathbb{Z}, \qquad \frac{2\operatorname{lcm}(1,\ldots,n)^2\binom{n+k}{k}}{m^2\binom{n}{m}\binom{n+m}{m}} \in \mathbb{Z}.$$

For the first fraction, since $m \leq n$ this implies that m^2 divides $\operatorname{lcm}(1,\ldots,n)^2$, and since the binomial coefficients are integers, this must be an integer. The second fraction requires a little more work. If we move the $\binom{n+k}{k}$ term to the denominator, then what we need to show is that the exponent of every prime in numerator is greater than or equal to that of every prime in the denominator. To do so, we will first notice that[2]

$$\binom{n+m}{m}\binom{n+k}{k}^{-1} = \binom{k}{m}\binom{n-k}{k-m}^{-1}$$

and so our denominator becomes

$$m^2\binom{n}{m}\binom{k}{m}\binom{n-k}{k-m}^{-1},$$

Then using what we found in Lemma (3.7), we have that

$$\nu_p\left(m^2\binom{n}{m}\binom{k}{m}\binom{n-k}{k-m}^{-1}\right) = 2\nu_p(m) + \nu_p\binom{n}{m} + \nu_p\binom{k}{m} - \nu_p\binom{n-k}{k-m}$$

Then we have that

$$2\nu_p(m) + \nu_p\binom{n}{m} + \nu_p\binom{k}{m} - \nu_p\binom{n-k}{k-m} \leq 2\nu_p(\operatorname{lcm}(1,\ldots,n)).$$

This is precisely what we wanted to show and so we're done. ■

We now have enough to prove that $\zeta(2)$ may not be written as the ratio of two integers.

Theorem 4.6: Apéry's Theorem

The number $\zeta(2)$ is irrational.

Proof To set the stage, recall our Irrationality Criterion (3.1), which in short, told us that if we can find two integer sequences such that Equation (3.6)) holds, then our constant of interest must be irrational. Therefore, to satisfy this integer sequence requirement, we are motivated to make the following definitions

$$\tilde{q}_n := 2\tilde{b}_n \operatorname{lcm}(1,\ldots,n)^2, \qquad \tilde{p}_n := 2\tilde{a}_n \operatorname{lcm}(1,\ldots,n)^2$$

so that $\tilde{p}_n, \tilde{q}_n \in \mathbb{Z}$ and $\tilde{a}_n/\tilde{b}_n = \tilde{p}_n/\tilde{q}_n$. Therefore, the strategy is clear; show

$$\left|\zeta(2) - \frac{\tilde{p}_n}{\tilde{q}_n}\right| \ll \tilde{q}_n^{-(1+\delta)}$$

for some fixed $\delta > 0$. If we then use Theorems (4.4, 3.9), and choose $\varepsilon/2$, we have

$$\tilde{q}_n \ll \tilde{r}_1^n e^{n(2+\varepsilon)} = \tilde{r}_1^{n\left(1+\frac{2+\varepsilon}{\log \tilde{r}_1}\right)}.$$

2. Which can readily be seen by writing both the left and right hand sides in factorial form.

This then tells us that

$$\tilde{r}_1^{-n\left(1+\frac{2+\varepsilon}{\log \tilde{r}_1}\right)} \ll \tilde{q}_n^{-1},$$

and so for every $\delta > 0$,

$$\tilde{r}_1^{-n\left(1+\frac{2+\varepsilon}{\log \tilde{r}_1}\right)(1+\delta)} \ll \tilde{q}_n^{-(1+\delta)}.$$

Then considering what we found in Equation (4.4)), we will have

$$\left|\zeta(2) - \frac{\tilde{p}_n}{\tilde{q}_n}\right| \ll \tilde{q}_n^{-(1+\delta)}$$

provides that

$$2 \geq \left(\frac{2 + \log \tilde{r}_1 + \varepsilon}{\log \tilde{r}_1}\right)(1+\delta).$$

After moving some terms around to isolate the δ, we find

$$\delta \leq \frac{\log \tilde{r}_1 - 2 - \varepsilon}{\log \tilde{r}_1 + 2 + \varepsilon}.$$

Then since $\varepsilon > 0$ is arbitrarily small, the acceptable values of δ can be arbitrarily close to

$$\delta = \frac{\log \tilde{r}_1 - 2}{\log \tilde{r}_1 + 2} = 0.09\ldots. \tag{4.5}$$

Therefore we can choose $\delta = \frac{\log \tilde{r}_1 - 2}{\log \tilde{r}_1 + 2} > 0$ in order to satisfy $r_1^{-2n} \ll q_n^{-(1+\delta)}$, proving that $\zeta(2)$ fails to be rational by our Irrationality Criterion. ■

4.2 Beukers' Idea

Like that of Apéry's proof, Beukers' idea for the proof of $\zeta(2)$ being irrational is very similar to his proof for $\zeta(3)$. As in § 3.2, we will frequently make use of integrals over some multiple of the unit line $[0,1]$ in which our integral may be improper. Therefore, in order to make our arguments sound, we may instead consider the truncated interval $[0, 1-\varepsilon]$ and then let $\varepsilon \to 0^+$ to recover our original integral.

We start by considering

$$\iint_{[0,1]^2} \frac{x^k y^\ell}{1-xy}\, dxdy. \tag{4.6}$$

Then notice that if we assume $k > \ell$, it is a direct consequence of Theorem (3.13) and Equation (3.20)) that

$$\iint_{[0,1]^2} \frac{x^k y^\ell}{1-xy}\, dxdy = \frac{1}{k-1}\sum_{t=1}^{k-1}\frac{1}{t+1}.$$

Moreover, if we assume $1 \leq \ell < k \leq n$, then $k-\ell$ and $t+1$ must divide $\operatorname{lcm}(1,\ldots,n)$ for all $t \in \{\ell, \ell+1, \ldots, k-1\}$. Therefore we see that Equation (4.6)) has denominators dividing $\operatorname{lcm}(1,\ldots,n)^2$. We then have the following claim.

Theorem 4.7

The equality

$$\iint_{[0,1]^2} \frac{(xy)^k}{1-xy}\,dxdy = \zeta(2) - \sum_{t=1}^{k} \frac{1}{t^2}$$

holds.

Proof To see this, we take $k = \ell$ and $w = 0$ in Equation (3.20)). Then

$$\iint_{[0,1]^2} \frac{(xy)^k}{1-xy}\,dxdy = \sum_{m=0}^{\infty} \frac{1}{(k+m+1)^2}.$$

Upon making the substitution $t \mapsto k+m+1$ in our summation, we have the desired result. ■

Now that we have related our integral to the constant we're interested in, like in the case of $\zeta(3)$ we will tie in the Legendre polynomials (in this case however, we will only consider one Legendre polynomial).

Theorem 4.8

We have

$$\iint_{[0,1]^2} \frac{(1-y)^n P_n(x)}{1-xy}\,dxdy = \frac{\tilde{A}_n + \tilde{B}_n \zeta(2)}{\operatorname{lcm}(1,\ldots,n)^2}, \tag{4.7}$$

where $(\tilde{A}_n)$ and $(\tilde{B}_n)$ are sequences of integers.

Proof First recall the Binomial Theorem (3.12) and Equation (3.23)), which states that we may write

$$(1-y)^n P_n(x) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} y^k \sum_{\ell=0}^{n} (-1)^\ell \binom{n}{\ell} \frac{(n+\ell)!}{n!\ell!} x^\ell.$$

Therefore, we set

$$a_k = (-1)^k \binom{n}{k}, \qquad a_\ell = (-1)^\ell \binom{n}{\ell} \frac{(n+\ell)!}{n!\ell!},$$

such that $a_k \in \mathbb{Z}$ from Lemma (3.6) and $a_\ell \in \mathbb{Z}$ from the proof of Theorem (3.16). Therefore, we may write

$$\iint_{[0,1]^2} \frac{(1-y)^n P_n(x)}{1-xy}\,dxdy = \sum_{k=0}^{n} \sum_{\ell=0}^{n} a_k a_\ell \iint_{[0,1]^2} \frac{x^\ell y^k}{1-xy}\,dxdy.$$

We previously saw that when $k > \ell$, Equation (4.6)) has denominators dividing $\operatorname{lcm}(1,\ldots,n)^2$, and when $k = \ell$, results in a value involving $\zeta(2)$. Thus, since $a_k, a_\ell \in \mathbb{Z}$ and subsequently don't introduce any new denominators, we are done. ■

Now to simplify our integral representation in Equation (4.7)), we will apply a similar trick to what we did in Theorem (3.17).

Theorem 4.9

If we integrate n-times with respect to x, then

$$\iint_{[0,1]^2} \frac{(1-y)^n P_n(x)}{1-xy}\,dxdy = (-1)^n \iint_{[0,1]^2} \frac{(xy)^n (1-x)^n (1-y)^n}{(1-xy)^{n+1}}\,dxdy.$$

Proof Let $g(x) = x^n(1-x)^n$ and $f(x) = (1-y)^n/(1-xy)$. We therefore want to consider the integral

$$\frac{1}{n!}\int_0^1 g^{(n)}(x)f(x)\,dx.$$

We know from the proof of Theorem (3.17) that after n partial integrations with respect to x, this becomes

$$\frac{1}{n!}\int_0^1 g^{(n)}(x)f(x)\,dx = \frac{1}{n!}\left(\sum_{k=0}^{n-1}(-1)^k g^{(n-k-1)}(x)f^{(k-1)}(x)\Big|_0^1 + (-1)^n\int_0^1 g(x)f^{(n)}(x)\,dx\right), \tag{4.8}$$

where the entire summation vanishes since any derivative of g of order less than n vanishes at both zero and one. We then examine the derivative of f. We have

$$f'(x) = \frac{(1-y)^n y}{(1-xy)^2},$$

and that

$$f''(x) = \frac{2(1-y)^n y^2}{(1-xy)^3},$$

which leads us to guess

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

which can be shown via induction. Upon plugging $f^{(n)}$ into Equation (4.8)), we have the desired result. ■

We then of course have

$$\iint_{[0,1]^2}\frac{(1-y)^n P_n(x)}{1-xy}\,dxdy = \frac{\tilde{A}_n + \tilde{B}_n\zeta(2)}{\operatorname{lcm}(1,\ldots,n)^2}$$

as a result of Theorem (4.8) and the previous one. Now, in order to derive a contradiction at some point, we'd like to have some sort of numerical bound on our integrand.

Theorem 4.10

For x, y as defined throughout this analysis,

$$\frac{xy(1-x)(1-y)}{1-xy} \leq \frac{5\sqrt{5}-11}{2}.$$

Proof If we let $f(x,y) = xy(1-x)(1-y)/(1-xy)$, then one may quickly recognize $f(x,y) = f(y,x)$. We then would like to maximize this function. To do so, we consider the equations

$$\frac{\partial f}{\partial x} = 0, \qquad \frac{\partial f}{\partial y} = 0. \tag{4.9}$$

However, the x and y that satisfy the conditions $f(x,y) = f(y,x)$ and those in Equation (4.9)) are precisely those such that $x = y$. Thus we may simplify our analysis to

$$f(t) := f(t,t) = \frac{(1-t)^2t^2}{1-t^2}.$$

We then have

$$f'(t) = -\frac{2t(t^2+t-1)}{(1+t)^2}.$$

When we solve $f'(t) = 0$ for t, we find solutions $t_0 = 0$, $t_{1,2} = (-1 \pm \sqrt{5})/2$. We can then safely ignore t_0, since $f(0) = 0$ and so we check if the $t_{1,2}$ are maxima with $f(t_{1,2}) < 0$. Doing so will reveal that both t_1 and t_2 are maxima, but the greater of the two is $t_1 = (-1+\sqrt{5})/2$ for which $f(t_1) = (5\sqrt{5}-11)/2$, establishing the claimed inequality. ∎

Now, all that is left is for us to derive some sort of contradiction with all the pieces we've obtained.

Theorem 4.11

The number $\zeta(2)$ is irrational.

Proof From our bound in Theorem (4.10) and the fact that Equation (4.7)) is nonzero, we have

$$0 < |\tilde{A}_n + \tilde{B}_n\zeta(2)| \le \operatorname{lcm}(1,\dots,n)^2 \left(\frac{5\sqrt{5}-11}{2}\right)^n \iint_{[0,1]^2} \frac{dxdy}{1-xy}.$$

Then notice that if we take $k = \ell = 0$ in Equation (4.6)) then we have

$$0 < |\tilde{A}_n + \tilde{B}_n\zeta(2)| \le \operatorname{lcm}(1,\dots,n)^2 \zeta(2) \left(\frac{5\sqrt{5}-11}{2}\right)^n.$$

Then recall by Theorem (3.19) that there exists some $M(\varepsilon)$ such that whenever $n \ge M(\varepsilon)$, $\operatorname{lcm}(1,\dots,n)^2 < e^{n(2.01)} < 8^n$. We further have that

$$\frac{5\sqrt{5}-11}{2} < \frac{1}{10},$$

and $\zeta(2) < 2$. Therefore it follows that

$$0 < |\tilde{A}_n + \tilde{B}_n\zeta(2)| < 2\left(\frac{4}{5}\right)^n.$$

Now, for the sake of contradiction, let us suppose that $\zeta(2)$ is rational such that $\zeta(2) = a/b$ where $a \in \mathbb{Z}$ and $b \in \mathbb{Z}_{>0}$. Then we have that

$$\frac{1}{b} \le |b\tilde{A}_n + a\tilde{B}_n| < 2\left(\frac{4}{5}\right)^n.$$

Taking the limit as $n \to \infty$ implies that the integer $|b\tilde{A}_n + a\tilde{B}_n|$ is simultaneously bounded above by zero and strictly positive. ∎

5 A Way To Measure Irrationality

At this point, we have seen a few results on irrationality, but we would like to be able to say more than "the number x is irrational." Perhaps we could ask ourselves "*how* irrational is x?" To answer this question, we will consider what is known as an irrationality measure, or exponent, more specifically.

Definition 5.1: Irrationality Exponent(1)

The irrationality exponent of α, denoted by $\mu(\alpha)$ is the **largest** number μ for which there are infinitely many rationals p/q which satisfy

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

It is beneficial to also consider an alternative definition.

Definition 5.2: Irrationality Exponent(2)

The irrationality exponent of α, denoted by $\mu(\alpha)$ is the **smallest** number μ such that for every $\varepsilon > 0$, there are only finitely many rationals p/q which satisfy

$$\left|\alpha - \frac{p}{q}\right| < q^{-(\mu+\varepsilon)}.$$

These two definitions regard the same idea, but have their own benefits. They give us two different approaches to getting bounds on the irrationality exponent. If we are able to establish the inequality in Definition (5.1) with $\mu = t$, then we have the lower bound $\mu(a) \geq t$. If we instead are able to establish the inequality of Definition (5.2) with $\mu = t$, then we instead have the upper bound $\mu(\alpha) \leq t$.

The definition of an irrationality measure and irrationality exponent can differ from author to author, and, at times, both definitions are used interchangeably. However, the underlying idea is always the same: This allows us to measure how *far* α is from being rational. In particular, the irrationality exponent $\mu(\alpha)$ will give us information on how fast the error shrinks when approximating α by rationals with larger denominators. Therefore, the smaller $\mu(\alpha)$ is, the harder it is to approximate it with rationals and vice versa. Lastly, it is rather difficult to pin down the precise irrationality measure of a given number α, so we compromise by instead bounding it. We showcase the irrationality exponent with some examples we have already secretly seen and some new ones. First, consider the following theorem.

Theorem 5.1

Let $0 < \rho < \delta$. If there are infinitely many p_n/q_n for which

$$\left|\alpha - \frac{p_n}{q_n}\right| < q_n^{-(1+\delta)},$$

and the q_n are monotonically increasing where $q_n < q_{n-1}^{1+\rho}$, then for all $\varepsilon > 0$, the inequality

$$\left|\alpha - \frac{p}{q}\right| < q^{-\left(\frac{\delta+1}{\delta-\rho}+\varepsilon\right)}$$

has only finitely many solutions in integers p, q.

In order to make use of our theorem, notice that the requirement $q_n < q_{n-1}^{1+\rho}$ is the same as $\log q_n < (1+\rho)\log q_{n-1}$ or equivalently

$$\frac{\log q_n}{\log q_{n-1}} - 1 < \rho.$$

Therefore, if we have $\lim \frac{\log q_n}{\log q_{n-1}} = 1$, then for all $\varepsilon > 0$ there exists some N such that whenever $n \geq N$,

$$\frac{\log q_n}{\log q_{n-1}} - 1 < \varepsilon.$$

Thus if (q_n) is monotonically increasing and we set $\rho = \varepsilon$, then the hypothesis of Theorem (5.1) will be satisfied. This further implies the conclusion of Theorem (5.1) must also hold for all such $\rho > 0$ and so we would have the irrationality measure upper bound

$$\mu(\alpha) \leq \inf_{\rho \in (0,\delta)} \left(\frac{\delta + 1}{\delta - \rho} \right) = 1 + \frac{1}{\delta}.$$

We will now apply these ideas to the constants we dedicated the last two sections to. We first, however, consider a lemma we will need later.

Lemma 5.1

For all $\varepsilon > 0$, there exists some $K(\varepsilon)$ such that if $n \geq K(\varepsilon)$, then $e^{n(1-\varepsilon)} < \operatorname{lcm}(1, \ldots, n)$.

Theorem 5.2

The irrationality measure of $\zeta(3)$ satisfies

$$\mu(\zeta(3)) \leq 13.41 \ldots .$$

Proof Recall we have $q_n = 2b_n \operatorname{lcm}(1, \ldots, n)^3$ where (b_n) is defined as in Equation (3.5). As per our above discussion, we would like to show $\lim \frac{\log q_{n+1}}{\log q_n} = 1$ and (q_n) is monotonically increasing. For the latter, we know (b_n) is strictly increasing and nonnegative from Lemma (3.3). Furthermore, it is clear that $\operatorname{lcm}(1, \ldots, n)$ is nonnegative and strictly increasing. Thus (q_n) is monotonically increasing. We therefore need to show

$$\lim \frac{\log q_{n+1}}{\log q_n} = \lim \frac{\log 2 + \log b_{n+1} + 3 \log \operatorname{lcm}(1, \ldots, n+1)}{\log 2 + \log b_n + 3 \log \operatorname{lcm}(1, \ldots, n)} = 1.$$

Then recall from Theorem (3.7), $r_1^n \ll b_n \ll r_1^n$ which implies the existence of constants C_1 and C_2 such that

$$C_1 r_1^n \leq b_n \leq C_2 r_1^n$$

for sufficiently large n. We thus have $\log C_1 + n \log r_1 \leq \log b_n \leq \log C_2 + n \log r_1$ and so

$$\frac{\log C_1}{n \log r_1} + 1 \leq \frac{\log b_n}{n \log r_1} \leq \frac{\log C_2}{n \log r_1} + 1.$$

Therefore, we have $\lim \frac{\log b_n}{n \log r_1} = 1$ by the squeeze theorem. Moreover, by Theorem (3.9) and Lemma (5.1), for all $\varepsilon > 0$ there exist constants $K_1(\varepsilon)$ and $K_2(\varepsilon)$ such that whenever $n \geq \sup\{K_1(\varepsilon), K_2(\varepsilon)\}$ we have $e^{3n(1-\varepsilon)} < \operatorname{lcm}(1, \ldots, n)^3 < e^{3n(1+\varepsilon)}$. Thus, we have that $3n(1-\varepsilon) < 3 \log \operatorname{lcm}(1, \ldots, n) < 3n(1+\varepsilon)$ which implies that

$$\left| \frac{\log \operatorname{lcm}(1, \ldots, n)}{n} - 1 \right| < \varepsilon.$$

This is precisely the definition of the limit and therefore $\lim \frac{\log \operatorname{lcm}(1,\ldots,n)}{n} = 1$. All of this then allows us to write

$$\begin{aligned}\lim \frac{\log q_{n+1}}{\log q_n} &= \lim \frac{\log 2 + (n+1)\log r_1 + 3(n+1)}{\log 2 + n\log r_1 + 3n} \\ &= \lim \frac{\frac{\log 2}{n} + \left(1+\frac{1}{n}\right)\log r_1 + 3\left(1+\frac{1}{n}\right)}{\frac{\log 2}{n} + \log r_1 + 3} \\ &= 1\end{aligned}$$

Thus we are now justified in using Theorem (5.1) for which $\mu(\zeta(3)) \leq 1 + \frac{1}{\delta} = 13.41\ldots$, where δ comes from Equation (3.17). ■

We can similarly determine the exponent of $\zeta(2)$ and the story is essentially the same.

Theorem 5.3

The irrationality exponent of $\zeta(2)$ satisfies

$$\mu(\zeta(2)) \leq 11.85\ldots.$$

Proof Recall we have $\tilde{q}_n = 2\tilde{b}_n \operatorname{lcm}(1,\ldots,n)^2$ where $(\tilde{b}_n)$ is defined as in Equation (4.1). As in the last Theorem, we would like to show $\lim \frac{\log \tilde{q}_{n+1}}{\log \tilde{q}_n} = 1$ and $(\tilde{q}_n)$ is monotonically increasing. Due to Lemma (4.2), we know $(\tilde{b}_n)$ is strictly increasing and nonnegative. Furthermore, as we saw before, $\operatorname{lcm}(1,\ldots,n)$ is nonnegative and strictly increasing. Thus $(\tilde{q}_n)$ is monotonically increasing. We therefore need to show

$$\lim \frac{\log \tilde{q}_{n+1}}{\log \tilde{q}_n} = \lim \frac{\log 2 + \log \tilde{b}_{n+1} + 2\log \operatorname{lcm}(1,\ldots,n+1)}{\log 2 + \log \tilde{b}_n + 2\log \operatorname{lcm}(1,\ldots,n)} = 1.$$

Then recall from Theorem (4.4), $\tilde{r}_1^n \ll \tilde{b}_n \ll \tilde{r}_1^n$ which implies the existence of constants C_1 and C_2 such that

$$C_1\tilde{r}_1^n \leq \tilde{b}_n \leq C_2\tilde{r}_1^n$$

for sufficiently large n. In a similar manner to that of Theorem (5.2) have $\lim \frac{\log \tilde{b}_n}{n\log \tilde{r}_1} = 1$. Thus it follows that

$$\begin{aligned}\lim \frac{\log \tilde{q}_{n+1}}{\log \tilde{q}_n} &= \lim \frac{\log 2 + (n+1)\log \tilde{r}_1 + 2(n+1)}{\log 2 + n\log \tilde{r}_1 + 2n} \\ &= \lim \frac{\frac{\log 2}{n} + \left(1+\frac{1}{n}\right)\log \tilde{r}_1 + 2\left(1+\frac{1}{n}\right)}{\frac{\log 2}{n} + \log \tilde{r}_1 + 2} \\ &= 1.\end{aligned}$$

Then Theorem (5.1) gives $\mu(\zeta(2)) \leq 1 + \frac{1}{\delta} = 11.85\ldots$, where δ comes from Equation (4.5). ■

As was alluded to earlier, finding the precise irrationality exponent of α is rather challenging so we typically compromise by simply finding an upper bound of $\mu(\alpha)$. However, one may be wondering "what about a lower bound," and "which numbers do we know the precise irrationality exponent of, if such a number exists." To address the first, German mathematician Johann Lejeune Dirichlet (1805 - 1859) has a rather pleasing answer, and, as a result of Dirichlet's response, the precise irrationality exponent can be written down for some numbers.

Definition 5.3: Fractional Part Of A Number

We shall make the definition $\operatorname{frac}(x) := x - \lfloor x \rfloor$.

Note that this function acts as one would think it does: it provides the fractional part of a number or the part of a number beyond the decimal point. For example, $\operatorname{frac}(\pi) = \pi - 3 \cdots = 0.14 \ldots$.

Principle 5.1: Pigeonhole Principle

Given N boxes and $N + 1$ items to put them in, at least one box must contain two items.

This is likely best seen with socks. Suppose you have white and black socks in a drawer. Then we have two different colors of sock (boxes) and so if we pull out three socks, we are guaranteed to have a matching pair.

Theorem 5.4: Dirichlet's Approximation Theorem

Let α be irrational. Then there are infinitely many rationals p/q such that

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

with $\gcd(p, q) = 1$.

Proof For any positive integer N, let us divide the interval $[0, 1)$ into sub-intervals of length $1/N$. We shall then let $n \in \{0, 1, \ldots, N\}$ and consider numbers of the form $\operatorname{frac}(\alpha n)$. There are $N + 1$ such numbers and there are N sub-intervals of the form $1/N$. Therefore, by The Pigeonhole Principle (5.1), it must be the case that at least one of these sub-intervals will have at least two n's for which $\operatorname{frac}(\alpha n)$ lies in. That is, there must exist two integers, s and t such that

$$|\operatorname{frac}(\alpha s) - \operatorname{frac}(\alpha t)| < \frac{1}{N}.$$

If we then expand the definition of $\operatorname{frac}(x)$ and set $q := |s - t|$, then there must exist integers p and q for which

$$|\operatorname{frac}(\alpha s) - \operatorname{frac}(\alpha t)| = |\alpha q - p| < \frac{1}{N}.$$

Then recall that $s, t \in \{0, 1 \ldots, N\}$ and so $q \leq N$. As a result we have

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

It then follows that

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

Lastly, since we demonstrated this process for some arbitrary integer N, it follows that there are infinitely many such approximations, completing the proof. ■

Corollary 5.1

If α is an irrational number, then $\mu(\alpha) \geq 2$.

This corollary may give an idea of how we could approach finding the exact irrationality measure of a number. Due to the ordering properties of the real line, if β is simultaneously greater than and less than or equal to b, then it must be the case that $\beta = b$ and so if we show the upper bound $\mu(\alpha) \leq 2$, then by Corollary (5.1), it must be the case that $\mu(\alpha) = 2$. We do just that.

Lemma 5.2

The number $\sqrt{2}$ fails to be rational.

Proof For the sake of contradiction, suppose that $\sqrt{2}$ is rational and can be written as $\frac{p}{q}$, where $p \in \mathbb{Z}$, $q \in \mathbb{Z}_{>0}$, and $\gcd(p, q) = 1$. If $\sqrt{2} = \frac{p}{q}$, then squaring both sides gives

$$2 = \left(\frac{p}{q}\right)^2.$$

Multiplying both sides by q^2 gives $2q^2 = p^2$. From this equation, we see that p^2 is even and thus p must also be even. Therefore, we can write $p = 2k$ for some integer k. Substituting $p = 2k$ into the equation $2q^2 = p^2$ gives $2q^2 = (2k)^2$. Lastly, if we divide both sides by 2, then $q^2 = 2k^2$. This shows that q^2 is even, which means that q is also even. Therefore, both p and q are even. However, this is a contradiction, because if both p and q are even, then they have a common factor of 2. This contradicts our assumption that $\gcd(p, q) = 1$. ■

Corollary 5.2

There does not exist a rational number p/q which is a root of the polynomial $p(x) = x^2 - 2$.

Now, by Lemma (5.2) and Corollary (5.1) we know $\mu(\sqrt{2}) \geq 2$.

Theorem 5.5

The irrationality exponent of $\sqrt{2}$ is two.

Proof Consider $f(x) = x^2 - 2$. We then have that

$$|f(p/q)| = \left|\frac{p^2}{q^2} - 2\right| = \frac{|p^2 - 2q^2|}{q^2}.$$

Then, by Corollary (5.2), $f(p/q) \neq 0$ for all rational p/q. Therefore we have

$$|f(p/q)| \geq \frac{1}{q^2}.$$

Now, by the mean value theorem, there must exist some ξ such that

$$\left|\frac{f(p/q) - f(\sqrt{2})}{\frac{p}{q} - \sqrt{2}}\right| = |f'(\xi)|,$$

which is $|f(p/q)| = |f'(\xi)| \cdot |p/q - \sqrt{2}|$. We then have that $f'(x) = 2x$, and since p/q approximates $\sqrt{2}$, we will assume their distance is no greater than one. Therefore, $\xi \in [\sqrt{2} - 1, \sqrt{2} + 1]$ and so $|f'(\xi)| \leq 2(\sqrt{2} + 1) =: \beta$. We therefore have

$$|p/q - \sqrt{2}| = \left|\frac{f(p/q)}{f'(\xi)}\right| \geq \frac{1}{\beta q^2}.$$

Now, for the sake of contradiction, suppose for all $\varepsilon > 0$, there exists infinitely many p/q for which

$$|\sqrt{2} - p/q| < q^{-(2+\varepsilon)}.$$

Then it follows that $(\beta q^2)^{-1} < q^{-(2+\varepsilon)}$ and so $q^\varepsilon < \beta$ for all $\varepsilon > 0$. However, with β fixed, as q increases, this inequality will become impossible at some point. Therefore we have $\mu(\sqrt{2}) \leq 2$. ■

It is easy to see that a similar line of arguments can be applied to any $\sqrt{p}$ for p prime. However, perhaps a possibly less clear fact is that the irrationality exponent of any algebraic irrational number also happens to be two[1].

5.1 The Arithmetic of The Natural Logarithm

We will now discuss our first generalization of the Riemann zeta function. This generalization is likely the most natural to the zeta function, in that it is less of a generalization, and more of an alternative way to express the zeta function. In fact, we have already secretly seen this function in § 4.1.

Definition 5.4: Dirichlet Eta Function

We define

$$\eta(s) := \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^s} = 1 - \frac{1}{2^s} + \frac{1}{3^s} - \cdots$$

for $\mathrm{Re}(s) > 0$.

In the proof of Theorem (4.1), we had actually proven that

$$\eta(2) = \frac{1}{2}\zeta(2).$$

More generally, we can relate the Riemann zeta function and the Dirichlet eta function with the following.

Theorem 5.6

For $\mathrm{Re}(s) > 1$,

$$\eta(s) = (1 - 2^{1-s})\zeta(s)$$

Proof By definition, we have

$$\eta(s) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^s} = 1 - \frac{1}{2^s} + \frac{1}{3^s} - \cdots$$

and

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots .$$

Therefore, we clearly have that

$$\eta(s) = \sum_{n \in \mathbb{O}_{>0}} \frac{1}{n^s} - \sum_{n \in \mathbb{E}_{>0}} \frac{1}{n^s}. \tag{5.1}$$

1. This follows by a theorem of German mathematician Klaus Roth (1925 – 2015).

Then notice that we may express the zeta function as

$$\zeta(s) = \sum_{n\in\mathbb{O}_{>0}} \frac{1}{n^s} + \sum_{n\in\mathbb{E}_{>0}} \frac{1}{n^s} = \sum_{n\in\mathbb{O}_{>0}} \frac{1}{n^s} + 2^{-s}\zeta(s).$$

If we then move the appropriate terms around in this equation, one finds

$$\sum_{n\in\mathbb{O}_{>0}} \frac{1}{n^s} = (1-2^{-s})\zeta(s) \tag{5.2}$$

and

$$\sum_{n\in\mathbb{E}_{>0}} \frac{1}{n^s} = 2^{-s}\zeta(s).$$

Lastly, plugging these into Equation (5.1) will yield the desired result. ■

If we then consider the Taylor series of the natural logarithm centered about $x = 0$, we have

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n.$$

Upon setting $x = 1$,

$$\log(2) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \eta(1).$$

Therefore we have that $\log(2) = \eta(1)$. We now spend the remainder of this subsection discussing the irrationality of $\log(2)$ and its subsequent irrationality measure. The reader is encouraged to fill in the details.

We start by defining the sequences

$$\breve{a}_n := \sum_{k=0}^{n}\sum_{m=1}^{k} \frac{(-1)^{m+1}}{m}\binom{n}{k}\binom{n+k}{k}, \qquad \breve{b}_n := \sum_{k=0}^{n}\binom{n}{k}\binom{n+k}{k}.$$

We then follow the methods seen in §§ 3.1, 4.1. The story goes like this: We first want to ensure $(\breve{a}_n)$ and $(\breve{b}_n)$ provide accurate rational approximations by showing they converge to $\log(2)$. We then show $(\breve{a}_n)$ and $(\breve{b}_n)$ satisfy the same linear recurrence relation with polynomial coefficients. This recurrence relation is then used to improve our first estimate to one dependent on $(\breve{b}_n)$. This estimate can further be improved by analyzing the growth of $(\breve{b}_n)$, and lastly, we define new integer sequences $(\breve{q}_n)$ and $(\breve{p}_n)$ derived from $(\breve{a}_n)$ and $(\breve{b}_n)$.

Theorem 5.7

The number which satisfies the equation $e^x = 2$ is irrational.

We shall also omit such a proof, however, the reader should verify that when going through the steps, one finds the "delta" which may be used in the Irrationality Criterion (3.1) to be $\delta = 0.276\ldots$. We will now use these results to discuss an upper bound for the irrationality exponent of $\log(2)$. This result is typically attributed to Indian-American mathematician Krishnaswani Alladi (1955–) and mathematician Michael Robinson [AR79].

Theorem 5.8

The irrationality exponent of $\log(2)$ *satisfies*

$$\mu(\log(2)) \leq 4.62\ldots.$$

Proof Let us define $\check{q}_n := \check{b}_n \operatorname{lcm}(1,\ldots,n)$. Then we would like to show $\lim \frac{\log \check{q}_{n+1}}{\log \check{q}_n} = 1$ and that $(\check{q}_n)$ is monotonically increasing. The latter should be clear from our previous work. Now,

$$\lim \frac{\log \check{q}_{n+1}}{\log \check{q}_n} = \lim \frac{\log \check{b}_{n+1} + \log(\operatorname{lcm}(1,\ldots,n+1))}{\log \check{b}_n + \log(\operatorname{lcm}(1,\ldots,n))}.$$

By analyzing the asymptotic growth of $(\check{b}_n)$, it should be easy to show that if we set $\check{r}_1 := (1+\sqrt{2})^2$, then $\check{r}_1^n \ll \check{b}_n \ll \check{r}_1^n$. We then have that $\lim \frac{\log \check{b}_n}{n \log \check{r}_1} = 1$. Therefore, we have

$$\begin{aligned}\lim \frac{\log \check{q}_{n+1}}{\log \check{q}_n} &= \lim \frac{(n+1)\log \check{r}_1 + (n+1)}{n \log \check{r}_1 + n} \\ &= \lim \frac{\left(1+\frac{1}{n}\right)\log \check{r}_1 + \left(1+\frac{1}{n}\right)}{\log \check{r}_1 + 1} \\ &= 1.\end{aligned}$$

Therefore, we may use Theorem (5.1) to find that $\mu(\log(2)) \leq 1 + \frac{1}{\delta} = 4.62\ldots$, where δ came from our previous discussion. ■

6 Irrationality Results After Zeta-3

We have now seen the irrationality of some zeta values. So, now what? An unfortunate shortcoming of Apéry's method is that attempts to generalize it to other odd values have been unsuccessful. However, there are typically two methods utilized to further progress our knowledge on the irrationality of positive odd zeta values, though they concern a linear combination of zeta values rather than singling one out as Beukers and Apéry did. Chronologically, the first successful crack at irrationality results beyond $\zeta(3)$ was that of British mathematician Keith Ball (1960–) and French mathematician Tanguy Rivoal (1972–) [BR01] in 2000 in which they proved that infinitely many odd zeta values are irrational. One year later Russian mathematician Wadim Zudilin (1970–) [Zud01] followed up by showing that at least one of $\zeta(5), \zeta(7), \zeta(9), \zeta(11)$ is irrational. A defining quality of these newer proofs is also that they use much higher grade machinery than that of Beukers or Apéry. From 2001 to now, a great deal of effort has been put into improving these results by using different sequences or more elementary methods, but a proof which singles out a particular odd zeta value has remained elusive since 1979. For the remainder of this section we will consider a proof of Zudilin [Zud18] which shows that at least one of $\zeta(5), \zeta(7), \ldots, \zeta(25)$ is irrational by *elementary means*, meaning it makes no use of heavier artillery than the prime number theorem and an asymptotic relationship we consider now.

Definition 6.1: Asymptotic

We say f is asymptotic to g as $x \to \infty$ if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$$

and we write $f(x) \sim g(x)$.

Theorem 6.1: Stirling's Formula

We have the asymptotic relationship

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

as $n \longrightarrow \infty$.

Then recall the *General Leibniz Rule* which tells us how to take the n-th derivative of a product of j functions.

Theorem 6.2: General Leibniz Rule

Given j *differentiable functions* $f_1, f_2, \ldots, f_j$,

$$(f_1 f_2 \ldots f_j)^{(n)} = \sum_{\ell_1+\ell_2+\cdots+\ell_j=n} \prod_{t=1}^{j} \frac{n!}{\ell_t!} f_t^{(\ell_t)},$$

where the sum is taken over all j*-tuples such that* $\ell_j \geq 0$ *and* $\ell_1 + \ell_2 + \cdots + \ell_j = n$.

Moreover, we'll make use of the following function, which as we will see in Remark (6.1), generalizes the zeta function.

Definition 6.2: Polylogarithm Function

We define

$$\mathrm{Li}_s(z) := \sum_{k=1}^{\infty} \frac{z^k}{k^s}$$

when $|z| < 1$ for all s.

Remark 6.1: Special Values Of Li

We have that $\mathrm{Li}_1(z) = -\log(1-z)$, which comes from the Taylor expansion of the natural logarithm. If we instead had $z = 1$, then for $\mathrm{Re}(s) > 1$, we recover the typical Riemann zeta function. That is, for example, $\mathrm{Li}_2(1) = \zeta(2) = \pi^2/6$.

The general idea behind this proof is to construct two sequences r_n and $\hat{r}_n$ to be defined later, such that they result in a linear combination of $1, \zeta(3), \zeta(5), \ldots, \zeta(s)$ with rational coefficients for odd $s \geq 7$. However, proving the irrationality of at least one number from a set of numbers which contains a known irrational number is rather meaningless since we already knew at least one of said numbers was irrational. Therefore, not only do we want a sequence which results in a linear combination of the odd zeta values, but we also want the rational coefficient of $\zeta(3)$ in our sequences to be proportional. In the sequences we present, it turns out that the coefficient of $\zeta(3)$ in the sequence $\hat{r}_n$ is proportional to r_n with a factor of 7. That is to say, if we consider $7r_n - \hat{r}_n$, this will instead result in a sequence of a linear combination of $1, \zeta(5), \ldots, \zeta(s)$ with rational coefficients. If we then consider the common denominator of these numbers, say λ_n, it follows that $\lambda_n(7r_n - \hat{r}_n)$ is a linear combination of the odd zeta values other than 3 with integer coefficients.

Now comes the crux of the proof. We set $s = 25$ (for reasons which will become apparent later on) which will reveal that for sufficiently large n, $7r_n - \hat{r}_n > 0$. If we then suppose for the sake of contradiction that all odd zeta values (other than 3) in this range are rational and then show that $\lambda_n \operatorname{lcm}(1, \ldots, n)^{25}(7r_n - \hat{r}_n) \longrightarrow 0$ as $n \longrightarrow \infty$, then this, however, contradicts the fact that the sequence $\lambda_n(7r_n - \hat{r}_n)$ is a positive sequence of integers.

Consider the sequences

$$r_n = \sum_{t=1}^{\infty} R_n(t), \qquad \hat{r}_n = \sum_{t=1}^{\infty} R_n\left(t - \frac{1}{2}\right) \tag{6.1}$$

such that

$$R_n(t) = \frac{2^{6n} n!^{2-5} \prod_{j=0}^{6n} (t - n + \frac{j}{2})}{\prod_{j=0}^{n} (t + j)^{s+1}} \tag{6.2}$$

where we fix $s \geq 7$ and odd. Then we have the following lemma.

Lemma 6.1

A rational function of the form

$$S(t) = \frac{P(t)}{Q(t)} = \frac{P(t)}{(t - t_1)^{s_1} \ldots (t - t_q)^{s_q}} \tag{6.3}$$

where $\deg P(t) < \deg Q(t)$ *has the unique partial fraction decomposition*

$$S(t) = \sum_{j=1}^{q} \sum_{i=1}^{s_j} \frac{b_{ij}}{(t - t_j)^i}.$$

Proof Let $S(t) = \frac{P(t)}{Q(t)} = \frac{P(t)}{(t-t_1)^{s_1} \ldots (t-t_q)^{s_q}}$ such that $\deg P(t) < \deg Q(t)$. Then if we apply partial fraction decomposition to $S(t)$, we have

$$S(t) = \frac{A_1}{(t - t_1)^{s_1}} + \frac{A_2}{(t - t_2)^{s_2}} + \cdots + \frac{A_q}{(t - t_q)^{s_q}} = \sum_{j=1}^{q} \frac{A_j}{(t - t_j)^{s_j}},$$

for some constants A_j. We can then apply partial fraction decomposition again to the term $(t - t_j)^{-s_j}$ to get

$$\frac{1}{(t - t_j)^{s_j}} = \frac{B_1}{t - t_j} + \frac{B_2}{(t - t_j)^2} + \cdots + \frac{B_{s_j}}{(t - t_j)^{s_j}}.$$

Putting this into our previous formula will give

$$S(t) = \sum_{j=1}^{q} A_j \left(\frac{B_1}{t - t_j} + \frac{B_2}{(t - t_j)^2} + \cdots + \frac{B_{s_j}}{(t - t_j)^{s_j}} \right) = \sum_{j=1}^{q} A_j \sum_{i=1}^{s_j} \frac{B_i}{(t - t_j)^i} = \sum_{j=1}^{q} \sum_{i=1}^{s_j} \frac{L_{ij}}{(t - t_j)^i},$$

where $L_{ij} = B_i A_j$. We will omit such a verification of uniqueness. ■

Thus, we can express $R_n(t)$ in a similar manner.

Theorem 6.3

The rational function $R_n(t)$ *may be expressed as*

$$\sum_{i=1}^{s} \sum_{k=0}^{n} \frac{a_{ik}}{(t + k)^i}. \tag{6.4}$$

We may also explicitly determine the coefficients of the partial fraction decomposition found in Lemma (6.1).

Lemma 6.2

The coefficients found in Equation (6.3) are given by

$$b_{ij} = \frac{1}{(s_j - i)!} \lim_{t \to t_j} \frac{d^{s_j - i}}{dt^{s_j - i}} \left[(t - t_j^{s_j}) S(t) \right].$$

Proof Given

$$S(t) = \sum_{j=1}^{q} \sum_{i=1}^{s_j} \frac{b_{ij}}{(t - t_j)^i},$$

to determine the coefficients b_{ij}, let us fix a pole t_j of $S(t)$, which will be of order s_j. Then set $H(t) := (t - t_j)^{s_j} S(t)$, where H is analytic at t_j and thus admits a Taylor expansion about t_j. Expanding H in a Taylor series centered about t_j yields

$$H(t) = \sum_{m=0}^{\infty} \frac{H^{(m)}(t_j)}{m!} (t - t_j)^m. \tag{6.5}$$

Furthermore, we may write

$$H(t) = \sum_{i=1}^{s_j} b_{ij} (t - t_j)^{s_j - i} + G(t), \tag{6.6}$$

where

$$G(t) = \sum_{\substack{k=1 \\ k \neq j}}^{q} \sum_{i=1}^{s_k} \frac{b_{ik} (t - t_j)^{s_j}}{(t - t_k)^i}.$$

Then, since $t_k \neq t_j$, G is analytic at t_j and contains a factor of $(t - t_j)^{s_j}$. Since G contains a factor of $(t - t_j)^{s_j}$ we have $G^{(\tau)}(t_j) = 0$ for all $\tau = 0, 1, \ldots, s_j - 1$, and since $s_j - i \leq s_j - 1$ it follows that $G^{(s_j - i)}(t_j) = 0$. Now comparing the representations in Equations (6.5), (6.6) for H and using the fact that $G^{(s_j - i)}(t_j) = 0$, we see the coefficients are given by

$$b_{ij} = \frac{H^{(s_j - i)}(t_j)}{(s_j - i)!} = \frac{1}{(s_j - i)!} \frac{d^{s_j - i}}{dt^{s_j - i}} \left[(t - t_j)^{s_j} S(t) \right] \Bigg|_{t = t_j},$$

as desired. ■

We will then use this to show when we may be sure about clearing the denominators of the coefficients found in Equation (6.4) and subsequently Equation (6.3).

Lemma 6.3

Let $k_1, \ldots, k_q$ *be distinct numbers from the set* $\{0, 1, \ldots, n\}$, $s_1, \ldots, s_q$ *be positive integers, and* $s = s_1 + s_2 + \cdots + s_q$. *Then the coefficients in the expansion*

$$S(t) = \prod_{j=1}^{q} \frac{1}{(t + k_j)^{s_j}} = \sum_{j=1}^{q} \sum_{i=1}^{s_j} \frac{b_{ij}}{(t - k_j)^i}$$

satisfy $\operatorname{lcm}(1 \ldots, n)^{s-i} b_{ij} \in \mathbb{Z}$.

Proof Notice the symmetry of $S(t)$ with respect to j. Since the ordering of j does not matter, without loss of generality, we will show the claim holds for $\operatorname{lcm}(1,\dots,n)^{s-i}b_{i1}$ which shall imply the result. Thus we fix a pole of order s_1 at $t=-k_1$ in which we'll write $H(t) := (t+k_1)^{s_1}S(t)$, where $H(t)$ is analytic and nonzero at $t=-k_1$. As a result, we may also write

$$H(t) = \prod_{j=2}^{q}(t+k_j)^{-s_j}. \tag{6.7}$$

Now by Lemma (6.2),

$$b_{i1} = \frac{1}{(s_1-i)!}H^{(s_1-i)}(-k_1) = \frac{H^{(m)}(-k_1)}{m!}$$

where $m = s_1 - i$. Then, by the General Leibniz Rule (6.2),

$$\frac{H^{(m)}(t)}{m!} = \sum_{\ell_2+\ell_3+\cdots+\ell_q=m}\prod_{j=2}^{q}\frac{1}{\ell_j!}\frac{d^{\ell_j}}{dt^{\ell_j}}\frac{1}{(t+k_j)^{s_j}}.$$

Notice that

$$\frac{d}{dt}\frac{1}{(t+k_j)^{s_j}} = -\frac{s_j}{(t+k_j)^{s_j+1}}$$

and that

$$\frac{d^2}{dt^2}\frac{1}{(t+k_j)^{s_j}} = \frac{s_j(s_j+1)}{(t+k_j)^{s_j+2}},$$

which leads us to predict

$$\frac{d^{\ell_j}}{dt^{\ell_j}}\frac{1}{(t+k_j)^{s_j}} = \frac{(-1)^{\ell_j}s_j(s_j+1)\cdot\ldots\cdot(s_j+\ell_j-1)}{(t+k_j)^{s_j+\ell_j}}$$

which can be shown by induction. We therefore have

$$\frac{H^{(m)}(t)}{m!} = \sum_{\ell_2+\ell_3+\cdots+\ell_q=m}\prod_{j=2}^{q}\frac{(-1)^{\ell_j}s_j(s_j+1)\cdot\ldots\cdot(s_j+\ell_j-1)}{\ell_j!(t+k_j)^{s_j+\ell_j}}$$

If we then multiply up and down by $(s_j-1)!$, it then follows that

$$\frac{H^{(m)}(t)}{m!} = \sum_{\ell_2+\ell_3+\cdots+\ell_q=m}\prod_{j=2}^{q}\frac{(-1)^{\ell_j}}{(t+k_j)^{s_j+\ell_j}}\binom{s_j+\ell_j-1}{\ell_j}.$$

Now notice if $q=1$, then from Equation (6.7) we see that $H(t)=1$ in which $H^{(m)}(t)\in\{0,1\}$ and so

$$b_{i1} = \frac{H^{(m)}(-k_1)}{m!} \in \{0,1\} \subset \mathbb{Z}$$

which implies $\operatorname{lcm}(1,\dots,n)^{s-i}b_{i1}\in\mathbb{Z}$. Thus when $q=1$, the result is trivial and so we assume $q\geq 2$. We now have

$$b_{i1} = \frac{H^{(m)}(-k_1)}{m!} = \sum_{\ell_2+\ell_3+\cdots+\ell_q=m}\prod_{j=2}^{q}\frac{(-1)^{\ell_j}}{(k_j-k_1)^{s_j+\ell_j}}\binom{s_j+\ell_j-1}{\ell_j}. \tag{6.8}$$

Now, recall we wish to show that $\operatorname{lcm}(1\dots,n)^{s-i}b_{i1}\in\mathbb{Z}$. Since the integers are closed under products and sums this informs us that if the summand in Equation (6.8) is an integer after proper rescaling then we are done. With the binomial coefficients being integers, we only concern ourselves with the term $(k_j - k_1)^{s_j+\ell_j}$. More specifically, what remains to show is that

$$\operatorname{lcm}(1, \dots, n)^{s-i} \prod_{j=2}^{q} \frac{1}{(k_j - k_1)^{s_j+\ell_j}} \in \mathbb{Z}.$$

First, since $k_j - k_1$ is an integer with magnitude less than or equal to n, we have $k_j - k_1 \mid \operatorname{lcm}(1, \dots, n)$. As a result of this, multiplying the product of $(k_j - k_1)^{-s_j-\ell_j}$ by the factor

$$\operatorname{lcm}(1, \dots, n)^{\sum_{j=2}^{q}(s_j+\ell_j)}$$

will clear the denominators. By assumption and the definition of m,

$$\sum_{j=2}^{q}(s_j + \ell_j) = \sum_{j=2}^{q} s_j + \sum_{j=2}^{q} \ell_j = s - i,$$

from which the claim follows. ∎

Theorem 6.4

The coefficients a_{ik} satisfy $\operatorname{lcm}(1, \dots, n)^{s-i} a_{ik} \in \mathbb{Z}$.

It then follows that the function $R_n(t)$ satisfies a nice property which we will in turn use to prove a property about our coefficients a_{ik}.

Lemma 6.4

We have $R_n(-t-n) = -R_n(t)$.

Proof We start by expanding the definition of

$$R_n(-t-n) = \frac{2^{6n} n!^{2-5} \prod_{j=0}^{6n}(-t-n-n+\frac{j}{2})}{\prod_{j=0}^{n}(-t-n+j)^{s+1}} = \frac{(-1)^{6n+1} 2^{6n} n!^{2-5} \prod_{j=0}^{6n}(t+2n-\frac{j}{2})}{(-1)^{(s+1)(n+1)} \prod_{j=0}^{n}(t+n-j)^{s+1}}.$$

Then since $6n+1$ is always odd, $(-1)^{6n+1} = -1$, and since we fix s to be odd, $(-1)^{(s+1)(n+1)} = 1$. Thus

$$R_n(-t-n) = \frac{-2^{6n} n!^{2-5} \prod_{j=0}^{6n}(t+2n-\frac{j}{2})}{\prod_{j=0}^{n}(t+n-j)^{s+1}}.$$

We will then consider the numerator and denominator separately. In the numerator, consider the change of variables $j \mapsto 6n - k$ in which k ranges from $6n$ to 0. However, since the order of multiplication does not matter, we have

$$\prod_{j=0}^{6n}\left(t + 2n - \frac{j}{2}\right) = \prod_{k=0}^{6n}\left(t + -n + \frac{k}{2}\right).$$

Then for the denominator, consider $j \mapsto n - k$ in which k ranges from n to 0. We therefore have

$$\prod_{j=0}^{n}(t+n-j)^{s+1} = \prod_{k=0}^{n}(t+k)^{s+1}.$$

The claim follows directly after plugging these into the numerator and denominator respectively. ∎

Theorem 6.5

The coefficients a_{ik} satisfy $a_{ik} = (-1)^{i+1}a_{i(n-k)}$ and

$$\sum_{k=0}^{n} a_{ik} = 0$$

whenever i is even.

Proof Notice

$$R_n(-t-n) = \sum_{i=1}^{s}\sum_{k=0}^{n} \frac{a_{ik}}{(-t-n+k)^i} = \sum_{i=1}^{s}\sum_{k=0}^{n} \frac{(-1)^i a_{ik}}{(t+n-k)^i}.$$

If we then make the change of variables $n-k \mapsto p$, we have that p ranges from n to 0. Thus, upon switching the order of summation, we have

$$R_n(-t-n) = \sum_{i=1}^{s}\sum_{p=0}^{n} \frac{(-1)^i a_{i(n-p)}}{(t+p)^i}.$$

Therefore, by Lemma (6.4), it follows that

$$R_n(-t-n) = \sum_{i=1}^{s}\sum_{p=0}^{n} \frac{(-1)^i a_{i(n-p)}}{(t+p)^i} = -R_n(t) = \sum_{i=1}^{s}\sum_{p=0}^{n} \frac{-a_{ip}}{(t+p)^i}.$$

Then, by the uniqueness of partial fraction decompositions, $a_{ip} = (-1)^{i+1}a_{i(n-p)}$.

Now, if i is even, then $a_{ik} + a_{i(n-k)} = 0$. We then have

$$\sum_{k=0}^{n} a_{ik} = a_{i0} + a_{i1} + \cdots + a_{i(n-1)} + a_{in}. \tag{6.9}$$

If n were odd, then we could write Equation (6.9) as

$$\sum_{k=0}^{n} a_{ik} = (a_{i0} + a_{in}) + (a_{i(n-1)} + a_{in}) + \cdots + (a_{i(\frac{n+1}{2})} + a_{i(\frac{n-1}{2})}) = 0.$$

If n were even, then we could do the same as above, but instead end up with the leftover term $a_{i(\frac{n}{2})}$. However, when n is even, $\frac{n}{2} = n - \frac{n}{2}$, in which $2a_{i(\frac{n}{2})} = 0$ and therefore $a_{i(\frac{n}{2})} = 0$. ■

We will now shift our attention to the sequences previously introduced in Equation (6.1) by first considering a lemma.

Lemma 6.5

When $i \geq 2$,

$$(2^i - 1)\zeta(i) = \sum_{\ell=1}^{\infty} \frac{1}{(\ell - \frac{1}{2})^i}.$$

Proof We may write

$$\sum_{\ell=1}^{\infty} \frac{1}{(\ell-\frac{1}{2})^i} = \sum_{\ell=1}^{\infty} \frac{1}{\left(\frac{2\ell-1}{2}\right)^i} = 2^i \sum_{\ell=1}^{\infty} \frac{1}{(2\ell-1)^i}.$$

Then recall that

$$\sum_{\ell=1}^{\infty} \frac{1}{(2\ell-1)^i} = \sum_{\ell \in \mathbb{O}_{>0}} \frac{1}{\ell^i} = \zeta(i)(1-2^{-i})$$

by Equation (5.2). Finally, multiplying the last equation by 2^i will give the desired result. ∎

The next theorem is precisely what will justify our previous claim that $7r_n - \hat{r}_n$ will cause $\zeta(3)$ to vanish since $2^3 - 1 = 7$.

Theorem 6.6

For all $n \in \mathbb{Z}_{>0}$,

$$r_n = \sum_{i=2}^{s} \zeta(i) P_i + P_0, \qquad \hat{r}_n = \sum_{i=2}^{s} (2^i - 1)\zeta(i) P_i + \hat{P}_0,$$

where

$$P_i = \sum_{k=0}^{n} a_{ik}, \quad P_0 = -\sum_{i=1}^{s}\sum_{k=0}^{n}\sum_{\alpha=1}^{k} \frac{a_{ik}}{\alpha^i}, \quad \hat{P}_0 = \sum_{i=1}^{s}\sum_{k=0}^{m}\sum_{\ell=0}^{m-k} \frac{(-1)^i a_{ik}}{(\ell+\frac{1}{2})^i} - \sum_{i=1}^{s}\sum_{k=m+1}^{n}\sum_{\ell=1}^{k-m-1} \frac{a_{ik}}{(\ell-\frac{1}{2})^i}.$$

It then follows that $\operatorname{lcm}(1,\ldots,n)^{s-i} P_i,\ \operatorname{lcm}(1,\ldots,n)^s P_0,\ \operatorname{lcm}(1,\ldots,n)^s \hat{P}_0 \in \mathbb{Z}$.

Proof We will start with the sequence (r_n) and consider an auxiliary parameter z which will later be set to one. We then have

$$\begin{aligned} r_n(z) \coloneqq \sum_{t=1}^{\infty} R_n(t) z^t &= \sum_{t=1}^{\infty}\sum_{i=1}^{s}\sum_{k=0}^{n} \frac{a_{ik} z^t}{(t+k)^i} \\ &= \sum_{i=1}^{s}\sum_{k=0}^{n} a_{ik} z^{-k} \sum_{t=1}^{\infty} \frac{z^{t+k}}{(t+k)^i}. \end{aligned}$$

Then, in our last sum, making the substitution $t + k \mapsto \alpha$, α will range from $k+1$ to ∞, and so

$$\sum_{t=1}^{\infty} \frac{z^{t+k}}{(t+k)^i} = \sum_{\alpha=k+1}^{\infty} \frac{z^\alpha}{\alpha^i} = \operatorname{Li}_i(z) - \sum_{\alpha=1}^{k} \frac{z^\alpha}{\alpha^i},$$

We therefore have

$$\begin{aligned} r_n(z) &= \sum_{i=1}^{s}\sum_{k=0}^{n} a_{ik} z^{-k} \left(\operatorname{Li}_i(z) - \sum_{\alpha=1}^{k} \frac{z^\alpha}{\alpha^i} \right) \\ &= \sum_{i=1}^{s} \operatorname{Li}_i(z) \sum_{k=0}^{n} a_{ik} z^{-k} - \sum_{i=1}^{s}\sum_{k=0}^{n}\sum_{\alpha=1}^{k} \frac{a_{ik} z^{\alpha-k}}{\alpha^i}. \end{aligned}$$

Now, from Remark (6.1), at $z = 1$ and $i \geq 2$, the polylogarithm function is well defined with $\operatorname{Li}_i(1) = \zeta(i)$. However, $\operatorname{Li}_1(z) = -\log(1-z)$ is divergent as $z \to 1^-$. Therefore, as $z \to 1^-$, we may possibly have negative behavior in our first sum, so we will bring our attention there. Write

$$\lim_{z \to 1^-} \sum_{i=1}^{s} \operatorname{Li}_i(z) \sum_{k=0}^{n} a_{ik} z^{-k} = \lim_{z \to 1^-} \operatorname{Li}_1(z) \sum_{k=0}^{n} a_{1k} z^{-k} + \lim_{z \to 1^-} \sum_{i=2}^{s} \operatorname{Li}_i(z) \sum_{k=0}^{n} a_{ik} z^{-k}. \tag{6.10}$$

As per our discussion above, the second limit in Equation (6.10) is perfectly defined, and so we have

$$\lim_{z\to 1^-}\sum_{i=2}^{s}\mathrm{Li}_i(z)\sum_{k=0}^{n}a_{ik}z^{-k}=\sum_{i=2}^{s}\zeta(i)\sum_{k=0}^{n}a_{ik}.$$

We then analyze the first limit. Recall the numerator of $R_n(t)$ was of degree $6n+1$ and the denominator had degree at least $8n+8$. Therefore, at its lowest, when $n=1$, $R_n(t)=O(t^{-9})$ as $t\to\infty$. This implies there exists some constant C such that $\sum_t |R_n(t)|\le C\sum_t t^{-9}$, which is to say $r_n(z)$ converges to a finite value when $z=1$. As a result, it must be the case that

$$\sum_{k=0}^{n}a_{1k}=0$$

and so

$$\lim_{z\to 1^-}\sum_{k=0}^{n}a_{1k}z^{-k}=0,$$

because if not, the logarithmically divergent term would prevent $r_n(z)$ from having a finite limit as $z\to 1^-$. Now, if we wrote

$$\sum_{k=0}^{n}a_{1k}z^{-k}=\sum_{k=0}^{n}a_{1k}(z^{-k}-1),$$

since $z^{-k}-1=O(1-z)$ for z sufficiently close to 1, it follows that

$$\sum_{k=0}^{n}a_{1k}z^{-k}=O(1-z).$$

Therefore we have that there exists some constant C such that for z near 1,

$$\left|-\log(1-z)\sum_{k=0}^{n}a_{1k}z^{-k}\right|\le C(1-z)|\log(1-z)|.$$

It is then a standard result from Calculus that

$$\lim_{z\to 1^-}(1-z)|\log(1-z)|=0.$$

Therefore, it must be the case that

$$-\lim_{z\to 1^-}\log(1-z)\sum_{k=0}^{n}a_{1k}z^{-k}=0.$$

Now, for the second sum, this is well defined when $z=1$ and so

$$\lim_{z\to 1^-}\sum_{i=1}^{s}\sum_{k=0}^{n}\sum_{\alpha=1}^{k}\frac{a_{ik}z^{\alpha-k}}{\alpha^i}=\sum_{i=1}^{s}\sum_{k=0}^{n}\sum_{\alpha=1}^{k}\frac{a_{ik}}{\alpha^i}.$$

Putting all of this together then gives us

$$r_n=\lim_{z\to 1^-}r_n(z)=\sum_{i=2}^{s}\zeta(i)\sum_{k=0}^{n}a_{ik}-\sum_{i=1}^{s}\sum_{k=0}^{n}\sum_{\alpha=1}^{k}\frac{a_{ik}}{\alpha^i}.$$

Now, for $(\hat{r}_n)$, we will follow a similar procedure of introducing an auxiliary parameter z which will later be set to one. First, however, notice that in the definition of $R_n(t)$, if $t = -\frac{1}{2}$, then $R_n(t)$ vanishes (specifically when $j = 2n+1$). A similar argument holds for all $t = -\frac{1}{2}, -\frac{3}{2}, \dots, -n+\frac{1}{2}$. As a result, if we shift the lower index of summation down, but not past $-n+\frac{1}{2}$, in the definition of $(\hat{r}_n)$, it will remain unchanged. Furthermore, we specify the lower bound of this shift to be $-m$ where $m = \lfloor \frac{n-1}{2} \rfloor$. This is, in fact, not the furthest down we could shift our indices, however, this specific choice has some rather beneficial properties when it comes to our integrality conditions. However, this idea is best seen with a quick example.

Suppose $n = 5$. Then as noted above, we are certain $R_5(t)$ will vanish at $t = -\frac{1}{2}, -\frac{3}{2}, -\frac{5}{2}, -\frac{7}{2}, -\frac{9}{2}$. Moreover, by definition, $m = 2$. Therefore, our re-indexed summation will run from $-2, -1, 0, \dots$ and in particular, the inputs for $R_5(t-\frac{1}{2})$ will run from $-\frac{5}{2}, -\frac{3}{2}, -\frac{1}{2}, \dots$, which is to say we could have lowered our summation index more, but we also risk losing our claimed properties of m.

Now back to analyzing $\hat{r}_n$: We have that

$$\hat{r}_n(z) \coloneqq \sum_{t=-m}^{\infty} R_n\left(t - \frac{1}{2}\right) z^t = \sum_{t=-m}^{\infty} \sum_{i=1}^{s} \sum_{k=0}^{n} \frac{a_{ik} z^t}{(t+k-\frac{1}{2})^i}$$

Then if we take $t + k \mapsto \ell$, it follows that ℓ ranges from $k - m$ to ∞. Thus

$$\hat{r}_n(z) = \sum_{\ell=k-m}^{\infty} \sum_{i=1}^{s} \sum_{k=0}^{n} \frac{a_{ik} z^{\ell-k}}{(\ell-\frac{1}{2})^i} = \sum_{i=1}^{s} \sum_{k=0}^{n} a_{ik} z^{-k} \sum_{\ell=k-m}^{\infty} \frac{z^\ell}{(\ell-\frac{1}{2})^i}.$$

Now when $k \le m$, split $\sum_{\ell=k-m}^{\infty} = \sum_{\ell=k-m}^{0} + \sum_{\ell=1}^{\infty}$ since we of course have $k - m \le 0$. Then when $k > m$, we will split $\sum_{\ell=k-m}^{\infty} = \sum_{\ell=1}^{\infty} - \sum_{\ell=1}^{k-m-1}$. Therefore, we write

$$\hat{r}_n(z) = \sum_{i=1}^{s} \left[\sum_{k=0}^{m} \frac{a_{ik}}{z^k} \left(\sum_{\ell=k-m}^{0} \frac{z^\ell}{(\ell-\frac{1}{2})^i} + \sum_{\ell=1}^{\infty} \frac{z^\ell}{(\ell-\frac{1}{2})^i} \right) + \sum_{k=m+1}^{n} \frac{a_{ik}}{z^k} \left(\sum_{\ell=1}^{\infty} \frac{z^\ell}{(\ell-\frac{1}{2})^i} - \sum_{\ell=1}^{k-m-1} \frac{z^\ell}{(\ell-\frac{1}{2})^i} \right) \right].$$

Then concerning the sum over ℓ from $k - m$ to 0, consider $\ell \mapsto -w$ where w ranges from $m - k$ to 0 so that

$$\sum_{\ell=k-m}^{0} \frac{z^\ell}{(\ell-\frac{1}{2})^i} = \sum_{w=m-k}^{0} \frac{z^{-w}}{(-w-\frac{1}{2})^i}.$$

Of course, summing from $m - k$ to 0 is no different than from 0 to $m - k$ and w was just some letter, so

$$\hat{r}_n(z) = \sum_{i=1}^{s} \sum_{k=0}^{n} a_{ik} z^{-k} \sum_{\ell=1}^{\infty} \frac{z^\ell}{(\ell-\frac{1}{2})^i} + \sum_{i=1}^{s} \sum_{k=0}^{m} a_{ik} z^{-k} \sum_{\ell=0}^{m-k} \frac{(-1)^i z^{-\ell}}{(\ell+\frac{1}{2})^i} - \sum_{i=1}^{s} \sum_{k=m+1}^{n} a_{ik} z^{-k} \sum_{\ell=1}^{k-m-1} \frac{z^\ell}{(\ell-\frac{1}{2})^i}.$$

Notice that the last two terms are finite and well-defined when $z = 1$, however, the first term is not necessarily. In particular, when $i = 1$, the first term may be ill-behaved, so we write

$$\sum_{i=1}^{s} \sum_{k=0}^{n} a_{ik} z^{-k} \sum_{\ell=1}^{\infty} \frac{z^\ell}{(\ell-\frac{1}{2})^i} = \sum_{k=0}^{n} a_{1k} z^{-k} \sum_{\ell=1}^{\infty} \frac{z^\ell}{\ell-\frac{1}{2}} + \sum_{i=2}^{s} \sum_{k=0}^{n} a_{ik} z^{-k} \sum_{\ell=1}^{\infty} \frac{z^\ell}{(\ell-\frac{1}{2})^i}$$

and analyze

$$\lim_{z\to 1} \sum_{k=0}^{n} a_{1k} z^{-k} \sum_{\ell=1}^{\infty} \frac{z^\ell}{\ell-\frac{1}{2}}. \tag{6.11}$$

However, just like in the case of r_n, since $\sum |R_n|$ converges for $z = 1$, $\hat{r}_n(z)$ must converge to a finite value as $z \to 1$, which implies Equation (6.11) evaluates to zero. Therefore, we have that

$$\hat{r}_n = \lim_{z\to 1} \hat{r}_n(z) = \sum_{i=2}^{s} \sum_{k=0}^{n} a_{ik} (2^i - 1)\zeta(i) + \sum_{i=1}^{s} \sum_{k=0}^{m} a_{ik} \sum_{\ell=0}^{m-k} \frac{(-1)^i}{(\ell+\frac{1}{2})^i} - \sum_{i=1}^{s} \sum_{k=m+1}^{n} a_{ik} \sum_{\ell=1}^{k-m-1} \frac{1}{(\ell-\frac{1}{2})^i}.$$

Now, to see the integrality of our coefficients, $\mathrm{lcm}(1,\ldots,n)^{s-i}P_i \in \mathbb{Z}$ follows directly from Theorem (6.4). Then, for $\alpha \in \{1,2,\ldots,k\}$, we have $\alpha \mid \mathrm{lcm}(1,\ldots,n)$, which implies $\mathrm{lcm}(1,\ldots,n)^s \frac{1}{\alpha^i} \in \mathbb{Z}$ for all $1 \le i \le s$ and therefore $\mathrm{lcm}(1,\ldots,n)^s P_0 \in \mathbb{Z}$. We have multiple requirements for $\mathrm{lcm}(1,\ldots,n)^s \hat{P}_0 \in \mathbb{Z}$. We may express

$$\frac{1}{(\ell \pm \frac{1}{2})^i} = \frac{2^i}{(2\ell \pm 1)^i}$$

and therefore require $2\ell \pm 1 \le n$ so that $2\ell \pm 1 \mid \mathrm{lcm}(1,\ldots,n)$. In particular, we require $2\ell + 1 \le n$ for $0 \le k \le m$ and $2\ell - 1 \le n$ for $m+1 \le k \le n$, considering these cases separately. When $0 \le k \le m$, ℓ is largest for $\ell = m-k$, and when $m+1 \le k \le n$, ℓ is largest at $\ell = k-m-1$. Now our choice of m finally comes into play. Recall $m = \lfloor \frac{n-1}{2} \rfloor$ if and only if $m \le \frac{n-1}{2} < m+1$, or $2m \le n-1$. Therefore, when $0 \le k \le m$, $2\ell + 1 \le 2m - 2k + 1 \le n - 1 + 1 = n$. Similarly, when $m+1 \le k \le n$, we find that $2\ell - 1 \le 2k - 2m - 3 \le 2n - n - 2 < n$. Thus for any ℓ within our range, $\mathrm{lcm}(1,\ldots,n)^s(\ell \pm \frac{1}{2})^{-i} \in \mathbb{Z}$ and therefore $\mathrm{lcm}(1,\ldots,n)^s \hat{P}_0 \in \mathbb{Z}$. ■

Theorem 6.7

When $s \ge 7$ and odd, we have $\lim r_n^{1/n} = g(x_0) = \lim \hat{r}_n^{1/n}$ where

$$g(x) = \frac{64(x+3)^6(x+1)^{s+1}}{(x+2)^{2(s+1)}}$$

and x_0 is the unique positive zero of the polynomial $x(x+2)^{(s+1)/2} - (x-3)(x+1)^{(s+1)/2}$. We also have that $\lim \frac{r_n}{\hat{r}_n} = 1$.

Sketch We start by noticing that when $t = 1$, in $R_n(t)$, we will have a zero in the numerator when $j = 2n-2$. Similarly when $t = 1/2$, $R_n(t)$ will vanish since we have a factor of zero when $j = 2n-1$. Similar analysis will reveal $R_n(t)$ vanishes at $t = \frac{1}{2}, \frac{2}{2}, \frac{3}{2}, \frac{4}{2}, \ldots, n - \frac{1}{2}, n$. Therefore we accordingly shift our indices of summation to

$$r_n = \sum_{t=n+1}^{\infty} R_n(t) = \sum_{k=0}^{\infty} c_k$$

and

$$\hat{r}_n = \sum_{t=n+1}^{\infty} R_n\left(t - \frac{1}{2}\right) = \sum_{k=0}^{\infty} \hat{c}_k$$

where $c_k = R_n(n+1+k)$ and $\hat{c}_k = R_n(n + \frac{1}{2} + k)$ which are strictly positive. We then have the asymptotic relationship

$$\frac{c_k}{\hat{c}_k} = \left(\prod_{j=0}^{6n} \frac{2k+2+j}{2k+1+j}\right)\left(\prod_{j=0}^{6n} \frac{n+k+\frac{1}{2}+j}{n+k+1+j}\right)^{s+1} \sim \frac{6n+2k+2}{2k+1}\left(\frac{n+k}{2n+k+1}\right)^{(s+1)/2}$$

as $n + k \to \infty$. Now, the idea is in order to determine the growth of r_n we will look to the sum $\sum c_k$. To find the largest term of r_n, we may examine the index k in which the behavior of c_k switches from increasing to decreasing, and the idea is that r_n's growth will be largely determined by the peak of $\sum c_k$. That is, we are looking for the index of k for which $\frac{c_k}{\hat{c}_k}$ is very close to one. One can show that as $n + k \to \infty$, we have the behavior

$$\frac{c_{k+1}}{c_k} = \frac{(k+3n+\frac{3}{2})(k+3n+2)}{(k+1)(k+\frac{3}{2})}\left(\frac{k+n+1}{k+2n+2}\right)^{s+1} \sim \left(\frac{k+3n}{k}\right)^2\left(\frac{k+n}{k+2n}\right)^{s+1}. \tag{6.12}$$

If we then set $f(x) := \frac{x+3}{x}\left(\frac{x+1}{x+2}\right)^{(s+1)/2}$, it follows that as $n+k\to\infty$, $\frac{c_{k+1}}{c_k}\sim f(\frac{k}{n})^2$. Therefore, the behavior of f can help us understand that of c_k. In its current state, trying to analyze the behavior of f would likely be a painful effort. However, with the simple definition $q := (s+1)/2 \geq 4$, the logarithmic derivative of f is

$$\begin{aligned}\frac{f'(x)}{f(x)} &= \frac{d}{dx}\left[\log(x+3)-\log(x)+q\left(\log(x+1)-\log(x+2)\right)\right]\\ &= \frac{1}{x+3}-\frac{1}{x}+q\left[\frac{1}{x+1}-\frac{1}{x+2}\right]\\ &= \frac{(q-3)(x^2+3)-6}{x(x+1)(x+2)(x+3)},\end{aligned}$$

which is much more pleasant to analyze! Now, if $f(x)^2>1$ then the terms c_k will be increasing and vice versa for $f(x)^2<1$. In particular, the largest term will occur where $f(x)$ is exactly 1. It is readily seen that $(q-3)(x^2+3)-6$ has a unique positive zero for each s which we will denote by x_1. One can then show that as x ranges from 0 to x_1, f will decrease monotonically from ∞ to $f(x_1)$. As x ranges from x_1 to ∞, f will increase monotonically from $f(x_1)$ to $\lim_{x\to\infty} f(x)$. Lastly, an important note is that $\lim_{x\to\infty} f(x)=1$, but this value is not attained for all sufficiently large x. Therefore, there is one and only one positive solution to $f(x)=1$ which we will denote by x_0. That is, x_0 is the solution to

$$\frac{x+3}{x}\left(\frac{x+1}{x+2}\right)^{(s+1)/2}=1$$

or the zero to $x(x+2)^{(s+1)/2}-(x+3)(x+1)^{(s+1)/2}$. Lastly, we already know $x_0>0$, but since $f(1)=4\left(\frac{2}{3}\right)^q<1$ for all q, it follows that $0<x_0<1$. We may then express the term c_k in terms of factorials to utilize Stirling's formula (6.1). We have

$$c_k = \frac{n!^{s-5}(6n+2k+2)!(n+k)!^{s+1}}{2(2k+1)!(2n+k+1)!^{s+1}},$$

so by Equation (6.12) and our above discussion, $\frac{c_{k+1}}{c_k}>1$ for indices $k<x_0n-\gamma\sqrt{n}$ and $\frac{c_{k+1}}{c_k}<1$ for indices $k>x_0n+\gamma\sqrt{n}$ where $\gamma>0$ is a constant coming from an application of Stirling's formula to c_k. Therefore the largest term c_k for a given n will occur near the index $k_0 := k_0(n)\sim x_0n$ where $|k-k_0|\leq\gamma\sqrt{n}$. This then suggests the asymptotic behavior of r_n shall be dictated by that of c_k for indices near k_0. A similar story holds for $\hat{r}_n$, which shall be dictated by that of $\hat{c}_k$ where k_0 is as above, $|k-k_0|\leq\hat{\gamma}\sqrt{n}$, and $\hat{\gamma}>0$. Then, applying Stirling's formula to c_k, we have that

$$\begin{aligned}\lim r_n^{1/n} &= \lim c_{k_0}^{1/n}\\ &= \lim \underbrace{\left(\frac{n}{e}\right)^{(s-5)}\left(\frac{6n+2k_0+2}{e}\right)^{\frac{6n+2k_0+2}{n}}\left(\frac{e}{2k_0+1}\right)^{\frac{2k_0+1}{n}}\left(\frac{n+k_0}{e}\right)^{\frac{(s+1)(n+k_0)}{n}}\left(\frac{e}{2n+k_0+1}\right)^{\frac{(s+1)(2n+k_0+1)}{n}}}_{T_n}\end{aligned}$$

Then to evaluate this, we first take the logarithm of T_n to get

$$\begin{aligned}\log T_n = \frac{1}{n}((s-5)n(\log n-1) &+ (6n+2k_0+2)(\log(6n+2k_0+2)-1)\\ &+ (2k_0+1)(1-\log(2k_0+1)) + (s+1)(n+k_0)(\log(n+k_0)-1)\\ &+ (s+1)(2n+k_0+1)(1-\log(2n+k_0+1)))\end{aligned} \tag{6.13}$$

We then look closer at the terms involving $\log n$ and n to ensure this limit does not diverge. Since $k_0 \sim x_0 n$ we can write $k_0 = x_0 n + o(n)$. Then we have

$$\begin{aligned}
&6n + 2k_0 + 2 = 6n + 2x_0 n + o(n) + 2 = 6n + 2x_0 n + o(n)\\
&2k_0 + 1 = 2x_0 n + o(n)\\
&n + k_0 = n + x_0 n + o(n)\\
&2n + k_0 + 1 = 2n + x_0 n + o(n).
\end{aligned}$$

Applying this to our logarithms will then give

$$\begin{aligned}
&\log(6n + 2k_0 + 2) = \log n + \log(6 + 2x_0 + o(1))\\
&\log(2k_0 + 1) = \log n + \log(2x_0 + o(1))\\
&\log(n + k_0) = \log n + \log(1 + x_0 + o(1))\\
&\log(2n + k_0 + 1) = \log n + \log(2 + x_0 + o(1)).
\end{aligned}$$

Then, plugging these into Equation (6.13) and combining like terms will show that the coefficient of $\log n$ and n turns out to be zero. Therefore, after factoring out an n from each term, it follows that

$$\begin{aligned}
\log T_n = {} & (s + 1)(2 + x_0 + o(1))(1 - \log n - \log(2 + x_0 + o(1)))\\
& + (s + 1)(1 + x_0 + o(1))(\log n + \log(1 + x_0 + o(1)) - 1)\\
& + (6 + 2x_0 + o(1))(\log n + \log(6 + 2x_0 + o(1) - 1)\\
& + (2x_0 + o(1))(1 - \log n - \log(2x_0 + o(1))) + (s - 5)(\log n - 1)
\end{aligned}$$

Thus, since $o(1) \to 0$ as $n \to \infty$, we have

$$\lim \log T_n = (6+2x_0)\log(6+2x_0) - (2x_0)\log(2x_0) + (s+1)(1+x_0)\log(1+x_0) - (s+1)(2+x_0)\log(2+x_0).$$

Finally, from here, we see that

$$\lim r_n^{1/n} \frac{(6 + 2x_0)^{6+2x_0}(1 + x_0)^{(s+1)(1+x_0)}}{(2x_0)^{2x_0}(2 + x_0)^{(s+1)(2+x_0)}} = \frac{2^6 2^{2x_0}(3 + x_0)^6(1 + x_0)^{s+1}}{2^{2x_0}(2 + x_0)^{2(s+1)}} f(x_0)^{2x_0} = g(x_0),$$

since $f(x_0) = 1$. It can then be shown that as $n+k \to \infty$, $\frac{\hat{c}_{k+1}}{\hat{c}_k} \sim \frac{c_{k+1}}{c_k}$ using what was found in Equation (6.12). As a result of this, it follows that $\lim c_{k_0}^{1/n} = \lim \hat{c}_{k_0}^{1/n}$ and subsequently $\lim r_n^{1/n} = g(x_0) = \lim \hat{r}_n^{1/n}$. What is left is to show that the limit of the ratio of r_n and $\hat{r}_n$ is one, which, in general, does not immediately follow from their limits agreeing. We have that

$$\begin{aligned}
\lim \frac{r_n}{\hat{r}_n} = \lim \frac{c_{k_0}}{\hat{c}_{k_0}} &= \lim \frac{6n + 2k_0 + 2}{2k_0 + 1}\left(\frac{n + k_0}{2n + k_0 + 1}\right)^{(s+1)/2}\\
&= \lim \frac{6n + 2x_0 n + o(n)}{2x_0 n + o(n)}\left(\frac{n + x_0 n + o(n)}{2n + x_0 n + o(n)}\right)^{(s+1)/2}\\
&= \frac{6 + 2x_0}{2x_0}\left(\frac{1 + x_0}{2 + x_0}\right)^{(s+1)/2}\\
&= 1,
\end{aligned}$$

completing the sketch. ⊠

Theorem 6.8

At least one of the numbers from the set $\{\zeta(5), \zeta(7), \zeta(9), \ldots, \zeta(23), \zeta(25)\}$ *is irrational.*

Proof Set $s = 25$. Then x_0 is the solution to $x(x+2)^{13} - (x+3)(x+1)^{13} = 0$ which is $x_0 = 0.000367\ldots$, and so

$$g(x_0) = \frac{64(x_0+3)^6(x_0+1)^{13}}{(x_0+2)^{52}}.$$

Then since $\lim \frac{r_n}{\hat{r}_n} = 1$, we have that $r_n = \hat{r}_n(1+o(1))$. If we then write $7r_n - \hat{r}_n = \hat{r}_n\left(7\frac{r_n}{\hat{r}_n} - 1\right)$, as $n \to \infty$, the term within the parentheses tends to $7 - 1 = 6$. Therefore, for sufficiently large n, $7r_n - \hat{r}_n = \hat{r}_n(6+o(1))$, and since $6 + o(1) > 1$ for sufficiently large n , we have that $7r_n - \hat{r}_n > 0$. Moreover, we may write

$$(7r_n - \hat{r}_n)^{1/n} = \hat{r}_n^{1/n}(6+o(1))^{1/n},$$

where $n \to \infty$ implies $(6+o(1))^{1/n} \to 1$ and $\hat{r}_n^{1/n} \to g(x_0)$ by Theorem (6.7). Therefore it follows that

$$\lim(7r_n - \hat{r}_n)^{1/n} = g(x_0). \tag{6.14}$$

Now, for the sake of contradiction, suppose every number from the set $\{\zeta(5), \zeta(7), \ldots, \zeta(25)\}$ is rational. Thus they must have some common divisor, which we will denote by λ_n, so that $\lambda_n \operatorname{lcm}(1,\ldots,n)^{25}(7r_n - \hat{r}_n)$ is a sequence of positive integers. Then by Equation (6.14), we have that $7r_n - \hat{r}_n \approx g(x_0)^n$ and can write

$$g(x_0)^n = e^{n\log g(x_0)} = e^{-\kappa n}$$

where $\kappa = -\log g(x_0) > 0$. Moreover, by Lemma (5.1) $\lim\left(\log\left(\operatorname{lcm}(1,\ldots,n)^{1/n}\right)\right) = 1$, and so $\operatorname{lcm}(1,\ldots,n) \approx e^n$. Thus the growth of $\lambda_n \operatorname{lcm}(1,\ldots,n)^{25}(7r_n - \hat{r}_n)$ as n gets large will roughly behave like $e^{25n}e^{-\kappa n} = e^{(25-\kappa)n}$. Therefore, as n gets large, if $\kappa > 25$, then our exponent will be negative and our terms will tend to zero. If instead $\kappa < 25$, then our exponent will be positive and this expression will blow up to infinity. Luckily, we already know κ, and it's promising. We have

$$\kappa = -\log g(x_0) = 25.2923\ldots.$$

Therefore, as discussed above, we have that $\lim \lambda_n \operatorname{lcm}(1,\ldots,n)^{25}(7r_n - \hat{r}_n) = 0$. This of course, however, is impossible, as a sequence of positive integers can never tend to zero as n gets large. Therefore, it must be the case that at least one of the numbers from the set $\{\zeta(5), \zeta(7), \ldots, \zeta(25)\}$ is irrational. ∎

6.1 An Informal Discussion

We briefly outline potential improvements to Zudilin's result, as a comprehensive study warrants a dedicated article. It became apparent to the author whilst writing § 7 that due to the relation between the Hurwitz[1] zeta function and the Riemann zeta function, that twists other than by half could bear fruit. We will consider a twist by a 3rd throughout this subsection. That is, we could consider

$$h_n := \sum_{t=1}^{\infty} R_n\left(t - \frac{1}{3}\right).$$

Upon setting $s = 7$, one finds

$$h_2 = C_1 + C_2\frac{\pi^3}{\sqrt{3}} + C_3\frac{\pi^5}{\sqrt{3}} + C_4\frac{\pi^7}{\sqrt{3}} + C_5\zeta(3) + C_6\zeta(5) + C_7\zeta(7)$$

1. See § 7.

where the $C_j \in \mathbb{Q}$. This is rather unpleasant though; we have polluted our linear combination with transcendental values. The true beauty however emerges once we consider the corresponding shift

$$\hat{h}_n := \sum_{t=1}^{\infty} R_n\left(t - \frac{2}{3}\right).$$

If we consider the same value of n and s, one finds

$$\hat{h}_2 = \hat{C}_1 - C_2\frac{\pi^3}{\sqrt{3}} - C_3\frac{\pi^5}{\sqrt{3}} - C_4\frac{\pi^7}{\sqrt{3}} + C_5\zeta(3) + C_6\zeta(5) + C_7\zeta(7)$$

where $C_1 \neq \hat{C}_1$. This is much more pleasant to discover as it directly implies $h_2 + \hat{h}_2 = C + 2C_5\zeta(3) + 2C_6\zeta(5) + 2C_7\zeta(7)$, where $C = C_1 + \hat{C}_1$.

Now, let us take a step back in order to understand why we are considering these new sequences in the first place. The main idea is that we want to improve our result on at least one of the numbers from $\{\zeta(5), \zeta(7), \ldots, \zeta(25)\}$ being irrational; that is, we would like to make this set smaller. Our previous sequences (r_n), $(\hat{r}_n)$ coupled with the identity from Lemma (6.5) allowed us to remove the $\zeta(3)$ term. Notice that we could have removed any other odd zeta value from our set. For example, if we instead considered $31r_n - \hat{r}_n$, then we would have effectively removed $\zeta(5)$ rather than $\zeta(3)$. However, it is important to note the consequence of this action: Our result would have lost all meaning[2], as we already know $\zeta(3)$ is irrational. The difficulty lies in the fact that since we are forced to remove the $\zeta(3)$ term, we also use up the only degree of freedom Lemma (6.5) afforded us. In the language of linear algebra, with only two sequences, we are asking for a solution to the system

$$\begin{aligned} A_1 + (2^3 - 1)A_2 &= 0 \\ A_1 + (2^s - 1)A_2 &= 0 \end{aligned}$$

where $s \in \{5, 7, \ldots, 25\}$. It is readily seen that this system offers only the trivial solution, which provides nothing of use to us. Recall that $A\vec{x} = \vec{0}$ will have a non-trivial solution provided that we have more variables than equations. Therefore, given two equations, we require an additional variable (sequence) to ensure a non-trivial solution. This is precisely why we considered (h_n) and $(\hat{h}_n)$.

We however, only showed that $h_2 + \hat{h}_2$ removed the transcendental factors, which does not imply this is the case for all n. For the remainder of this subsection, we will supply the ideas to remove an additional number from the set $\{\zeta(5), \zeta(7), \ldots, \zeta(25)\}$, as previously stated, a deep investigation warrants its own article. Our first necessary observation is the following.

Lemma 6.6

For all $\mathrm{Re}(s) > 1$,

$$\sum_{n=1}^{\infty} \frac{1}{(n - \frac{1}{3})^s} + \sum_{n=1}^{\infty} \frac{1}{(n - \frac{2}{3})^s} = (3^s - 1)\zeta(s).$$

Proof This can be seen by expanding the definitions

$$\sum_{n=1}^{\infty} \frac{1}{(n - \frac{1}{3})^s} = \sum_{n=1}^{\infty} \frac{3^s}{(3n - 1)^s} = 3^s\left(\frac{1}{2^s} + \frac{1}{5^s} + \frac{1}{8^s} + \cdots\right)$$

2. This would be analogous to rounding up 10 animals and one dog and then proclaiming at least one of them is a dog.

and

$$\sum_{n=1}^{\infty} \frac{1}{(n-\frac{2}{3})^s} = \sum_{n=1}^{\infty} \frac{3^s}{(3n-2)^s} = 3^s\left(1 + \frac{1}{4^s} + \frac{1}{7^s} + \cdots\right)$$

so that

$$\sum_{n=1}^{\infty} \frac{1}{(n-\frac{1}{3})^s} + \sum_{n=1}^{\infty} \frac{1}{(n-\frac{2}{3})^s} = 3^s\left(1 + \frac{1}{2^s} + \frac{1}{4^s} + \frac{1}{5^s} + \frac{1}{7^s} + \cdots\right). \tag{6.15}$$

It is easy to see that the terms within the parentheses of Equation (6.15) is the typical Riemann zeta function with the multiples of 3^{-s} removed. That is,

$$\begin{aligned} 3^s\zeta(s) &= 3^s\left(1 + \frac{1}{2^s} + \frac{1}{4^s} + \cdots\right) + 3^s \sum_{n=1}^{\infty} \frac{1}{(3n)^s} \\ &= 3^s\left(1 + \frac{1}{2^s} + \frac{1}{4^s} + \cdots\right) + \zeta(s), \end{aligned}$$

from which the claim follows after simplifications. ■

If we then take that the transcendental factors vanished when $n = 2$ for granted, this then gives us confirmation that the sequences (h_n) and $(\hat{h}_n)$ will allow us to remove an additional term. In particular, suppose we specifically wanted to remove $\zeta(5)$. Then we may set

$$\mathfrak{h}_n := h_n + \hat{h}_n = \sum_{t=1}^{\infty}\left(R_n\left(t - \frac{1}{3}\right) + R_n\left(t - \frac{2}{3}\right)\right),$$

and find the necessary coefficients A_1, A_2, A_3, for which

$$A_1 r_n + A_2 \hat{r}_n + A_3 \mathfrak{h}_n \in \mathbb{Q} + \mathbb{Q}\zeta(7) + \cdots + \mathbb{Q}\zeta(25).$$

To do so, we simply solve the system

$$\begin{aligned} A_1 + (2^3 - 1)A_2 + (3^3 - 1)A_3 &= 0 \\ A_1 + (2^5 - 1)A_2 + (3^5 - 1)A_3 &= 0 \end{aligned}$$

which has infinitely many solutions, perhaps the simplest of which is $A_1 = 37$, $A_2 = -9$, and $A_3 = 1$. Therefore,

$$37r_n - 9\hat{r}_n + \mathfrak{h}_n \in \mathbb{Q} + \mathbb{Q}\zeta(7) + \cdots + \mathbb{Q}\zeta(25), \tag{6.16}$$

and so we have removed an additional odd zeta value from our set! This result should build confidence in our method, however it is only half of the battle.

Although we were able to remove $\zeta(5)$ from our set, that also cost us the introduction of two new sequences; the subsequent at least one irrationality claim does not necessarily come for free. Therefore, we now discuss the extra necessary implications to justify that at least one of the numbers $\{\zeta(7), \zeta(9), \ldots, \zeta(25)\}$ is irrational.

Assuming we use the same method of proof as in Theorem(6.8), we will derive a contradiction by showing $37r_n - 9\hat{r}_n + \mathfrak{h}_n > 0$, assuming every number from $\{\zeta(7), \zeta(9), \ldots, , \zeta(25)\}$ is rational with common denominator $\tilde{\lambda}_n$, and then $\lim \tilde{\lambda}_n \operatorname{lcm}(1, \ldots, n)^{25}(37r_n - 9\hat{r_n} + \mathfrak{h}_n) = 0$. Now, in order to show $37r_n - 9\hat{r}_n + \mathfrak{h}_n > 0$, we can write $37r_n - 9\hat{r}_n + \mathfrak{h}_n = \hat{r}_n\left(37\frac{r_n}{\hat{r}_n} - 9 + \frac{\mathfrak{h}_n}{\hat{r}_n}\right)$. Then, we know from Theorem (6.7) that $\lim \frac{r_n}{\hat{r}_n} = 1$ and so if $\lim \frac{\mathfrak{h}_n}{\hat{r}_n} = L$, then the term in the parentheses will tend to $37 - 9 + L = 28 + L$ as

$n \to \infty$. Therefore, as long as $L \in \mathbb{R}$, for sufficiently large n we may write $37r_n - 9\hat{r}_n + \mathfrak{h}_n = \hat{r}_n(28 + L + o(1))$. In particular, with n sufficiently large, $37r_n - 9\hat{r}_n + \mathfrak{h}_n > 0$.

We of course would also like the control over our denominators to work out. If we write

$$(37r_n - 9\hat{r}_n + \mathfrak{h}_n)^{1/n} = \hat{r}_n^{1/n}(28 + L + o(1))^{1/n},$$

then since $\lim \hat{r}_n^{1/n} = g(x_0)$ and $\lim(28+L+o(1))^{1/n} = 1$, it then follows that $\lim(37r_n - 9\hat{r}_n + \mathfrak{h}_n)^{1/n} = g(x_0)$. From our previous work on the proof of Theorem (6.7), it therefore follows that $\lim \tilde{\lambda}_n \operatorname{lcm}(1,\dots,n)^{25}(7r_n - \hat{r}_n) = 0$, a contradiction. Therefore, all we would need to show is the limiting behavior of $\mathfrak{h}_n/\hat{r}_n$ works out the way we need it to and of course that the inclusions claimed in Equation (6.16) actually do hold.

7 Beyond The Riemann Zeta Function

As we have seen, there is quite a bit to say about the Riemann zeta function. As a matter of fact, we have really only just covered a small chunk of what one can say about this function. However, all this analysis on a rather simple-looking function may lead one to ask themselves about other functions of a similar flavor. This section will be dedicated to an attempt of answering the question "So what comes next?" We start out with a motivating example using some of the ideas we saw in § 1.3.

Lemma 7.1

Let $f(e^{i\theta})$ be piecewise continuous with an associated Fourier series $f(e^{i\theta}) \sim \sum c_n e^{in\theta}$. If $f(e^{i\theta})$ is differentiable at α, then at $\theta = \alpha$, the Fourier series of $f(e^{i\theta})$ will converge to $f(e^{i\theta})$.

Example 7.1: Fourier Series Redux

Show that

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = \frac{\pi}{4} \tag{7.1}$$

by examining the Fourier series of the function

$$f(e^{i\theta}) = \begin{cases} 1, & \pi < \theta < 2\pi \\ -1, & 0 < \theta < \pi \end{cases}.$$

Solution As we saw in § 1.3, the Fourier coefficients of this function are given by

$$c_n = \begin{cases} 0, & n \in \mathbb{E} \\ 2/\pi i n, & n \in \mathbb{O} \end{cases}.$$

Therefore, the associated Fourier series of f is

$$\sum_{n=-\infty}^{\infty} c_n e^{in\theta} = \sum_{n \in \mathbb{O}} \frac{2}{\pi i n} e^{in\theta}.$$

Then for $n \in \mathbb{O}$, notice $e^{in\theta} - e^{-in\theta} = 2i\sin(n\theta)$, so

$$\begin{aligned} \sum_{n=-\infty}^{\infty} c_n e^{in\theta} &= \frac{2}{\pi i}\left(\cdots - \frac{e^{-3i\theta}}{3} - e^{-i\theta} + e^{i\theta} + \frac{e^{3i\theta}}{3} + \dots\right) \\ &= \sum_{n \in \mathbb{O}_{>0}} \frac{4\sin(n\theta)}{\pi n}. \end{aligned}$$

Then notice at $\theta = \pi/2$, $f(e^{i\theta}) = 1$, in which $f'(e^{i\theta}) = 0$. Moreover, $f(e^{i\theta})$ is clearly piecewise continuous, and so by Lemma (7.1),

$$f(e^{i\theta}) = \sum_{k=0}^{\infty} \frac{4\sin((2k+1)\theta)}{\pi(2k+1)}$$

when $\theta = \pi/2$. That is

$$1 = \sum_{k=0}^{\infty} \frac{4(-1)^k}{\pi(2k+1)}$$

which verifies Equation (7.1). ♦

Now that we have seen Equation (7.1), perhaps a generalization to the Riemann zeta function we could consider is

$$\sum \frac{(-1)^m}{(2m+1)^s},$$

or perhaps the series of the form

$$\sum \frac{1}{(am+b)^s}.$$

We have two functions which are fundamentally related and perfectly match the kind of series we would like to concern ourselves with. We start with the latter which is named after German mathematician Adolf Hurwitz (1859 – 1919).

Definition 7.1: The Hurwitz Zeta Function

We define

$$\zeta(s,a) := \sum_{n=0}^{\infty} \frac{1}{(n+a)^s} = 1 + \frac{1}{n+a} + \frac{1}{(n+a)^2} + \frac{1}{(n+a)^3} + \cdots$$

where $a \in \mathbb{R}$ such that $0 < a \leq 1$ and $\mathrm{Re}(s) > 1$.

We then consider a class of functions which introduces a “twist” to the typical zeta function. To do so, we first consider a preliminary definition.

Definition 7.2: Dirichlet Character

Let $m \in \mathbb{Z}_{>0}$. A Dirichlet character modulo m is a function $\chi : \mathbb{Z} \to \mathbb{C}$ such that:

- $\chi(n+m) = \chi(n)$ for all integers n (so χ is periodic),
- $\chi(pq) = \chi(p)\chi(q)$ for all integers p, q (so χ is totally multiplicative),
- $\chi(n) = 0$ if $\gcd(n,m) \neq 1$, and if $\gcd(n,m) = 1$ then $\chi(n)$ is a complex number of modulus 1.

Definition 7.3: Principal Character

The principal Dirichlet character modulo m, denoted χ_1, is defined by

$$\chi_1(n) = \begin{cases} 1 & \text{if } \gcd(n,m) = 1, \\ 0 & \text{if } \gcd(n,m) \neq 1 \end{cases}.$$

A character which is not the principal character is called the **non-principal** character.

We use this definition to define the type of functions we saw in Example (7.1).

Definition 7.4: Dirichlet L-functions

We define

$$L(s,\chi) := \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \chi(1) + \frac{\chi(2)}{2^s} + \frac{\chi(3)}{3^s} + \cdots$$

where $\chi(n)$ is the Dirichlet character mod m and $\mathrm{Re}(s) > 1$.

As suggested earlier, the Hurwitz zeta function and Dirichlet L-function are fundamentally connected.

Lemma 7.2

The Dirichlet L-functions and Hurwitz zeta function are connected by

$$L(s,\chi) = \frac{1}{m^s} \sum_{r=1}^{m} \chi(r)\zeta\left(s, \frac{r}{m}\right). \tag{7.2}$$

Proof Let $\chi(n)$ be the Dirichlet character modulo m and thus consider the residue classes modulo m. We shall let $n = qm + r$ for $1 \le r \le m$ and $q \in \mathbb{Z}_{\ge 0}$. Thus we have

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \sum_{r=1}^{m} \sum_{q=0}^{\infty} \frac{\chi(qm+r)}{(qm+r)^s}.$$

The last equality holds because we are simply reordering our terms. That is, if $a_n = \chi(n)/n^s$, instead of adding $a_1 + a_2 + \ldots$, we are reordering and instead adding

$$(a_1 + a_{m+1} + a_{2m+1} + \ldots) + (a_2 + a_{m+2} + a_{2m+2} + \ldots) + \cdots + (a_n + a_{2n} + a_{3n} + \ldots).$$

Since we assume $\mathrm{Re}(s) > 1$, $L(s,\chi)$ is absolutely convergent, making such a reordering valid. We then have

$$\sum_{r=1}^{m} \sum_{q=0}^{\infty} \frac{\chi(qm+r)}{(qm+r)^s} = \frac{1}{m^s} \sum_{r=1}^{m} \sum_{q=0}^{\infty} \frac{\chi(r)}{\left(q + \frac{r}{m}\right)^s} = \frac{1}{m^s} \sum_{r=1}^{m} \chi(r)\zeta\left(s, \frac{r}{m}\right)$$

since $0 < \frac{r}{m} \le 1$. ∎

As one may guess due to the name, the Hurwitz zeta function is connected to the Riemann zeta function in that the Riemann zeta function is simply a special case of the Hurwitz zeta function.

Lemma 7.3: The Hurwitz Zeta Function Generalizes Riemann's Zeta Function

We have

$$\zeta(s,1) = \zeta(s).$$

Now, we previously claimed that these two functions were precisely the ones we are interested in, however, in their current state, they look nothing like the series requested. The next Lemma should assure the reader that these truly are the functions which satisfy our wishes.

Lemma 7.4

We may write down series of the form

$$\sum_{m=0}^{\infty} \frac{1}{(am+b)^s}, \qquad \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m+1)^s}$$

using the Hurwitz zeta and Dirichlet L-functions respectively.

Proof For the Hurwitz zeta function, first write

$$\frac{a^s}{(am+b)^s} = \frac{1}{\left(m+\frac{b}{a}\right)^s}.$$

Therefore, as long as $0 < b/a \leq 1$ or $b \leq a$, the sign of b is the same as that of a, and $a \neq 0$, we will have

$$\frac{1}{a^s}\zeta\left(s, \frac{b}{a}\right) = \sum_{m=0}^{\infty} \frac{1}{(am+b)^s}.$$

Now for the L-function, we must first consider the Dirichlet character. The factor of $(-1)^n$ implies we will have alternating behavior in our summand with even indices being positive and odd, negative. Moreover, the denominator $(2n+1)^s$ implies that we must skip all even denominators. Looking at the structure of $L(s,\chi)$, we are looking for a non-principal Dirichlet character which is zero for n even, and then alternates sign at odd indices. More specifically, in the definition of the L-function, when $n = 1$ we want to be positive, then we skip over $n = 2$, then when $n = 3$ we want to be negative, when $n = 4$ we should be zero, when $n = 5$ we want to be positive again, and so on. The Dirichlet character that has this precise behavior is the non-principal Dirichlet character modulo 4. We have that

$$\chi_4(n) = \begin{cases} 1, & n \equiv 1 \bmod 4 \\ 0, & n \in \mathbb{E} \\ -1, & n \equiv 3 \bmod 4 \end{cases}.$$

It then follows that

$$\begin{aligned} L(s,\chi_4) = \sum_{n=1}^{\infty} \frac{\chi_4(n)}{n^s} &= \chi_4(1) + \frac{\chi_4(2)}{2^s} + \frac{\chi_4(3)}{3^s} + \cdots \\ &= 1 + 0 - \frac{1}{3^s} + 0 + \frac{1}{4^s} + \cdots \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^s}, \end{aligned}$$

which is exactly what we wanted. ■

Since the Dirichlet L-function with character modulo 4 is of special interest to us, we will give it its own definition.

Definition 7.5: Dirichlet Beta Function

We set

$$L(s, \chi_4) \coloneqq \beta(s)$$

and call this the Dirichlet beta function.

Thus, as a result of Example (7.1) and Lemma (7.4), we have implicitly shown that

$$\beta(1) = \frac{\pi}{4}.$$

As in the Basel problem, this was a rather surprising result to see. Furthermore, simply due to the similarity of such a problem, perhaps we could expect analogous behavior of $\beta(s)$ to that of $\zeta(s)$. Thus, we shall spend the remainder of this section developing properties of the Dirichlet beta function.

7.1 Analytic Continuation

In this section, we would like to analytically continue the Dirichlet L-functions. We will, however, only concern ourselves with the Hurwitz zeta function since due to Lemma (7.2), if we are successful in analytically continuing the Hurwitz zeta function to the left of the line $\mathrm{Re}(s) = 1$, then this will subsequently imply the analytic continuation of the Dirichlet L-functions. To analytically continue the Hurwitz zeta function, we first consider the integral for $0 < a \leq 1$

$$\mathcal{I}(s, a) = \frac{1}{2\pi i} \int_{\mathcal{H}} \frac{(-z)^{s-1}}{e^{(a-1)z}(e^z - 1)}\, dz \tag{7.3}$$

over the same contour used in Figure (2). We then have the following theorem.

Theorem 7.1

For $\mathrm{Re}(s) > 1$,

$$\zeta(s, a) = -\Gamma(1 - s)\mathcal{I}(s, a). \tag{7.4}$$

We will not prove this theorem; however, its statement should be rather reasonable and is closely related to what we saw in Theorem (2.2). In fact, recall $\zeta(s, 1) = \zeta(s)$ and when we take $a = 1$ in Equation (7.3) we recover the contour integral we considered when analytically continuing the Riemann zeta function. Additionally, as we saw in § 2.1, $\Gamma(1 - s)$ is a meromorphic function of s, and $\mathcal{I}(s, a)$ is an entire function of s, so we will use Equation (7.4) to define $\zeta(s, a)$ beyond the line $\mathrm{Re}(s) = 1$, and then we will use Equation (7.2) to define $L(s, \chi)$ beyond $\mathrm{Re}(s) = 1$.

7.2 Functional Equation

We now direct our interest to deriving a functional equation for the Hurwitz zeta function, then we will relate this back to the Dirichlet beta function. We start by defining a new type of series.

Definition 7.6: Periodic Zeta function

We define

$$F(\alpha, s) \coloneqq \sum_{n=1}^{\infty} \frac{e^{2\pi i n\alpha}}{n^s} = e^{2\pi i\alpha} + \frac{e^{4\pi i\alpha}}{2^s} + \frac{e^{6\pi i\alpha}}{3^s} + \cdots$$

where $\alpha \in \mathbb{R}$ and $\mathrm{Re}(s) > 1$.

Theorem 7.2

The Hurwitz zeta function satisfies

$$\zeta(1-s,a) = \Gamma(s)(2\pi)^{-s}\left(e^{-\pi is/2}F(a,s) + e^{\pi is/2}F(-a,s)\right) \tag{7.5}$$

for $\mathrm{Re}(s) > 1$ *and* $0 < a \leq 1$.

Like the analytic continuation of the Hurwitz zeta function, we will not explicitly prove this theorem. However, notice the similarities between this claim and that of Theorem (2.2). Again taking $a = 1$ in Equation (7.5) will resolve to the case of Equation (2.4). Hence a similar argument (with minor modifications) to that found in the proof of Theorem (2.2) will justify the result.

Deriving the functional equation for the Dirichlet L-functions requires a little more work.

Definition 7.7: Gauss Sum

For a Dirichlet character mod m, set

$$G(n,\chi) := \sum_{r=1}^{m} \chi(r)e^{2\pi irn/m}.$$

Definition 7.8: Conjugate Character

Given a Dirichlet character, we denote its complex conjugate (to be understood in the typical sense) by $\overline{\chi}$.

The conjugate Dirichlet character satisfies a nice property. In the world of integers modulo m, $\gcd(n,m) = 1$ implies that n has some multiplicative inverse, say n^{-1}. Then,

$$\chi(n)\chi(n^{-1}) = \chi(nn^{-1}) = \chi(1) = 1.$$

To see why $\chi(1) = 1$, note that if χ were the principal character, then this would be immediate, so suppose otherwise. Notice that $\chi(1) = \chi(1 \cdot 1) = \chi(1)^2$ and so $\chi(1)^2 - \chi(1) = 0$. This can only have two possible solutions, either $\chi(1) = 0$ or $\chi(1) = 1$. If we have the second solution, we are done, so we'll show the first solution is impossible. If $\chi(1) = 0$, then this implies that $\gcd(1,m) \neq 1$. This itself is a contradiction and so $\chi(1) = 1$. Therefore,

$$\chi(n^{-1}) = \chi(n)^{-1}.$$

Now recall for a complex number z, we have $z\overline{z} = |z|^2$. Thus,

$$|\chi(n)|^2 = 1^2 = \chi(n)\overline{\chi}(n)$$

which implies $\chi(n)^{-1} = \overline{\chi}(n)$. We then have the following definition and lemma.

Definition 7.9: Primitive Character

We say that a Dirichlet character χ is primitive mod m if and only if for every divisor d of m where $0 < d < m$, there exists some $t \in \mathbb{Z}$ for which $t \equiv 1 \pmod{d}$, $\gcd(m,t) = 1$, and $\chi(t) \neq 1$.

Lemma 7.5

For a Dirichlet character mod m*, if* $\gcd(m,n) = 1$*, then* $G(n,\overline{\chi}) = \chi(n)G(1,\overline{\chi})$.

Proof Let χ be a Dirichlet character mod m and $\gcd(m,n) = 1$. We then have

$$G(n,\overline{\chi}) = \sum_{r=1}^{m} \overline{\chi}(r)e^{2\pi irn/m}.$$

Now, since $\gcd(n,m) = 1$, n is invertible mod m and consider $t \equiv rn \pmod{m}$. Since r ranges from 1 to m and multiplication by n simply permutes the residue classes, t will range from 1 to m as well. We also have that $r \equiv tn^{-1} \pmod{m}$ and so

$$\begin{aligned} G(n,\overline{\chi}) &= \sum_{t=1}^{m} \overline{\chi}(tn^{-1})e^{2\pi it/m} \\ &= \sum_{t=1}^{m} \overline{\chi}(t)\overline{\chi}(n^{-1})e^{2\pi it/m} \\ &= \chi(n) \sum_{t=1}^{m} \overline{\chi}(t)e^{2\pi it/m} \end{aligned}$$

where the last equations holds per our above discussion. ■

Remark 7.1

As it turns out, we may write $G(n,\overline{\chi}) = \chi(n)G(1,\overline{\chi})$ if and only if χ is a primitive character. However, we only need the converse for this section and omit such a proof.

Theorem 7.3

Suppose that χ *is a primitive character* mod m *and* $\mathrm{Re}(s) > 1$*. Then we have*

$$G(1,\overline{\chi})L(s,\chi) = \sum_{r=1}^{m} \overline{\chi}(r)F\left(\frac{r}{m}, s\right).$$

Proof On the right hand side, we have

$$\sum_{r=1}^{m} \overline{\chi}(r) \sum_{n=1}^{\infty} \frac{e^{2\pi irn/m}}{n^s} = \sum_{n=1}^{\infty}\sum_{r=1}^{m} \frac{\overline{\chi}(r)e^{2\pi irn/m}}{n^s} = \sum_{n=1}^{\infty} \frac{G(n,\overline{\chi})}{n^s}.$$

Then since χ is primitive, by Remark (7.1), $G(n,\overline{\chi}) = \chi(n)G(1,\overline{\chi})$ and therefore,

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

as desired. ■

From this, we may obtain the functional equation of the Dirichlet L-function.

Lemma 7.6

Given a Dirichlet character mod m*,*

$$\sum_{r=1}^{m} \chi(r) F\left(-\frac{r}{m}, s\right) = \chi(-1) \sum_{r=1}^{m} \chi(r) F\left(\frac{r}{m}, s\right).$$

Proof First notice that we may write $\chi(r) = \chi(-1)\chi(-r)$. Therefore we have

$$\sum_{r=1}^{m} \chi(r) F\left(-\frac{r}{m}, s\right) = \chi(-1) \sum_{r=1}^{m} \chi(-r) F\left(-\frac{r}{m}, s\right).$$

Now, recall that F is a periodic function with period one and χ is periodic with period m. That is to say,

$$\chi(-1) \sum_{r=1}^{m} \chi(-r) F\left(-\frac{r}{m}, s\right) = \chi(-1) \sum_{r=1}^{m} \chi(m-r) F\left(\frac{m-r}{m}, s\right).$$

Now consider the substitution $j \mapsto m - r$ which ranges from 0 to $m - 1$. We therefore have

$$\chi(-1) \sum_{r=1}^{m} \chi(m-r) F\left(\frac{m-r}{m}, s\right) = \chi(-1) \sum_{j=0}^{m-1} \chi(j) F\left(\frac{j}{m}, s\right).$$

Then simply applying the substitution $r \mapsto j$, we may shift our index of summation to run from 1 to m, and the claim immediately follows. ∎

Theorem 7.4

Let χ be any primitive Dirichlet character mod m*. Then*

$$L(1-s, \chi) = m^{s-1} \Gamma(s) (2\pi)^{-s} \left(e^{-i\pi s/2} + \chi(-1) e^{i\pi s/2}\right) G(1, \chi) L(s, \overline{\chi})$$

for all s.

Proof Starting with the left hand side,

$$\begin{aligned} L(1-s, \chi) &= m^{s-1} \sum_{r=1}^{m} \chi(r) \zeta\left(1-s, \frac{r}{m}\right) \\ &= m^{s-1} \sum_{r=1}^{m} \chi(r) \Gamma(s) (2\pi)^{-s} \left[e^{-i\pi s/2} F\left(\frac{r}{m}, s\right) + e^{i\pi s/2} F\left(-\frac{r}{m}, s\right)\right] \\ &= m^{s-1} \Gamma(s) (2\pi)^{-s} \sum_{r=1}^{m} \left[\chi(r) e^{-i\pi s/2} F\left(\frac{r}{m}, s\right) + \chi(r) e^{i\pi s/2} F\left(-\frac{r}{m}, s\right)\right]. \end{aligned}$$

If we then apply Lemma (7.6), we obtain

$$\sum_{r=1}^{m} \chi(r) F\left(-\frac{r}{m}, s\right) = \chi(-1) \sum_{r=1}^{m} \chi(r) F\left(\frac{r}{m}, s\right),$$

so it follows that

$$L(1-s,\chi) = m^{s-1}\Gamma(s)(2\pi)^{-s}\sum_{r=1}^{m}\left[\chi(r)F\left(\frac{r}{m},s\right)\left(e^{-i\pi s/2}+\chi(-1)e^{i\pi s/2}\right)\right].$$

The claim then follows by applying Theorem (7.3). ■

We then of course want a way to actually utilize this functional equation. In its current state, it does not seem to be all too useful because of the term $L(1-s,\chi)$. To amend this issue, like in the case of the Riemann zeta function, we first consider the change of variables $z \mapsto -z$ in the integral $\mathcal{I}(s,a)$. Doing so, we find that

$$\mathcal{I}(s,a) = -\int_{-\mathcal{H}}\frac{z^{s-1}e^{az}}{1-e^{z}}\,dz.$$

If we then set

$$\mathcal{J}(s,a) := \int_{-\mathcal{H}}\frac{z^{s-1}e^{az}}{1-e^{z}}\,dz,$$

then, Theorem (7.2) becomes

$$\zeta(s,a) = \Gamma(1-s)\mathcal{J}(s,a), \tag{7.6}$$

which leads us to the following.

Theorem 7.5

For all $n \in \mathbb{Z}_{>0}$

$$\zeta(1-n,a) = -\frac{B_n(a)}{n}.$$

Proof Setting $s = 1-n$ in Equation (7.6), $\zeta(1-n,a) = \Gamma(n)\mathcal{J}(1-n,a) = (n-1)!\mathcal{J}(1-n,a)$. Then we compute $\mathcal{J}$ with

$$\mathcal{J}(1-n,a) = \operatorname{Res}\left[\frac{z^{-n}e^{az}}{1-e^{z}},0\right] = -\operatorname{Res}\left[z^{-n-1}\frac{ze^{az}}{e^{z}-1},0\right] = -\operatorname{Res}\left[z^{-n-1}\sum_{m=0}^{\infty}\frac{B_m(a)}{m!}z^{m},0\right].$$

Then, since the residue at zero is precisely the coefficient of z^{-1}, we set $-n-1+m=-1$, which gives us $m=n$. Therefore, it follows that

$$\mathcal{J}(1-n,a) = -\frac{B_n(a)}{n!},$$

in which the claim follows. ■

To wrap up our analysis, we will relate all of this back to the Dirichlet beta function, in hopes of getting a meaningful formula in the way we did for the Riemann zeta function.

Theorem 7.6

We have

$$\beta(s) = \frac{\pi^{s}2^{s-2}\left(\zeta\left(1-s,\frac{1}{4}\right)-\zeta\left(1-s,\frac{3}{4}\right)\right)}{\Gamma(s)\sin\left(\frac{\pi s}{2}\right)} \tag{7.7}$$

for all s. In particular, when $n \in \mathbb{Z}_{>0}$,

$$\beta(2n-1) = \frac{\chi_4(2n-1)\pi^{2n-1}2^{2n-3}\left(B_{2n-1}\left(\frac{3}{4}\right) - B_{2n-1}\left(\frac{1}{4}\right)\right)}{(2n-1)!} \tag{7.8}$$

Proof Recall that we defined the beta function to be the Dirichlet L-function with associated non-principal character modulo 4 which we will denote by χ_4. First note that this character has the property that $\chi_4 = \overline{\chi}_4$, and $\chi_4(-1) = -1$. Thus, by Theorem (7.4), we have

$$\beta(1-s) = 4^{s-1}\Gamma(s)(2\pi)^{-s}\left(e^{-i\pi s/2} - e^{i\pi s/2}\right)G(1,\chi_4)\beta(s).$$

However, we can make numerous simplifications. First,

$$G(1,\chi_4) = \sum_{r=1}^{4}\chi_4(r)e^{2\pi i r/4} = \chi_4(1)e^{\pi i/2} + \chi_4(3)e^{3\pi i/2} = 2i.$$

Furthermore, we can write

$$e^{-i\pi s/2} - e^{i\pi s/2} = -i\sin\left(\frac{\pi s}{2}\right) - i\sin\left(\frac{\pi s}{2}\right) = -2i\sin\left(\frac{\pi s}{2}\right).$$

So, our equation simplifies to

$$\begin{aligned}\beta(1-s) &= 4^{s-1}\Gamma(s)(2\pi)^{-s}\left(-2i\sin\left(\frac{\pi s}{2}\right)\right)(2i)\beta(s)\\ &= 4^{s}\Gamma(s)(2\pi)^{-s}\sin\left(\frac{\pi s}{2}\right)\beta(s).\end{aligned}$$

Therefore, if we move some things around, we find

$$\beta(s) = \frac{\beta(1-s)}{\Gamma(s)\sin\left(\frac{\pi s}{2}\right)}\left(\frac{\pi}{2}\right)^{s}. \tag{7.9}$$

If we would then like to further simplify this, Lemma (7.2) can be used to write

$$\begin{aligned}\beta(1-s) = L(1-s,\chi_4) &= 4^{s-1}\sum_{r=1}^{4}\chi_4(r)\zeta\left(1-s,\frac{r}{4}\right)\\ &= 4^{s-1}\left(\zeta\left(1-s,\frac{1}{4}\right) - \zeta\left(1-s,\frac{3}{4}\right)\right).\end{aligned}$$

Plugging this into Equation (7.9) will then produce (7.7), where the Hurwitz zeta terms may be computed using Theorems (7.5, 2.6).

Now notice that for $n \in \mathbb{Z}_{>0}$, we may write $\chi_4(2n-1) = \sin\left(\frac{\pi(2n-1)}{2}\right)$. Furthermore, if we express the gamma function as a factorial, then we have $\Gamma(2n-1) = (2n-2)!$. Our last observation is that by Theorem (7.5),

$$\zeta\left(-2n,\frac{1}{4}\right) - \zeta\left(-2n,\frac{3}{4}\right) = \frac{B_{2n-1}\left(\frac{3}{4}\right) - B_{2n-1}\left(\frac{1}{4}\right)}{2n-1}.$$

Putting all this together yields

$$\begin{aligned}\beta(2n-1) &= \frac{\chi_4(2n-1)\pi^{2n-1}2^{2n-3}\left(B_{2n-1}\left(\frac{3}{4}\right)-B_{2n-1}\left(\frac{1}{4}\right)\right)}{(2n-2)!\cdot(2n-1)} \\ &= \frac{\chi_4(2n-1)\pi^{2n-1}2^{2n-3}\left(B_{2n-1}\left(\frac{3}{4}\right)-B_{2n-1}\left(\frac{1}{4}\right)\right)}{(2n-1)!},\end{aligned}$$

as desired. ■

We can then finally use this result to confirm what we found in Example (7.1), just as we did for the Basel problem all the way in § 1.3.

Corollary 7.1

We have $\beta(1) = \frac{\pi}{4}$.

Proof Here we take $n = 1$ in Equation (7.8). As a result, we have

$$\beta(1) = \frac{\left(B_1\left(\frac{3}{4}\right)-B_1\left(\frac{1}{4}\right)\right)\pi}{2}.$$

Then we compute

$$B_1(x) = \sum_{m=0}^{\infty}\binom{1}{m}B_n x^{n-m} = \binom{1}{0}B_0x + \binom{1}{1}B_1 = x - \frac{1}{2},$$

by Theorem (2.6). Therefore, we have

$$\beta(1) = \frac{\left(\frac{3}{4}-\frac{1}{2}-\frac{1}{4}+\frac{1}{2}\right)\pi}{2} = \frac{\pi}{4},$$

confirming what we found in Example (7.1). ■

7.3 A Discussion on The Arithmetic Of the Dirichlet Beta Function

At this point, a natural follow up inquiry would be the arithmetic nature of the beta function for even integer input. Upon inspection of Equation (7.7), we quickly realize we are in trouble again. Our chances of using this specific equation to evaluate the beta function at say 2 is spoiled by that sine in the denominator. To keep a long story short, as it turns out, the behavior of the beta function is mirrored by that of what we saw in the zeta function, but for even integers! A rather tantalizing fact is that we know less of the arithmetic behavior for the beta function than the zeta function. Due to Apéry and Beukers' we are familiar with the irrationality of $\zeta(3)$ and $\zeta(2)$ (without π). However, like in the case of higher odd zeta values, as of the time of writing, any attempts to generalize these results to the beta function have been futile.

We are not completely empty-handed though. Due to Zudilin [Zud19], we can *at least* say something on the arithmetic nature of the beta function.

Theorem 7.7: Zudilin, 2019

At least one of the numbers from the set $\{\beta(2), \beta(4), \beta(6), \beta(8), \beta(10), \beta(12)\}$ is irrational.

The method of proof, however, relies on some rather intricate techniques and will not be presented in this paper. We will, however, mention that considering the results in § 6 suggests a fruitful method could be examining sequences similar to $\hat{r}_n$. In fact, if we specifically define an alternating version of $\hat{r}_n$ as

$$\tilde{r}_n := \sum_{t=1}^{\infty} (-1)^{t+1} R_n\left(t - \frac{1}{2}\right),$$

then one can see that

$$\tilde{r}_2 = 137536 - 27264\,\beta(2) - 71424\,\beta(4) - 41984\,\beta(6),$$

so it shouldn't be too difficult to convince oneself that sequences defined by similar alternating rational functions to that found in Equations (6.1), (6.2) should allow one to draw conclusions on the state of irrationality of at least one element from some set.

8 Closing Remarks and Further Directions

The arithmetic properties of the Riemann zeta function and its generalizations remain a key interest in number theory. While this article has focused on specific special values and foundational results, there are several natural avenues for further inquiries.

- **Irrationality of Dirichlet L-functions**
  Beyond the Riemann zeta function, the arithmetic of special values of Dirichlet L-functions is an active subject of research. A notable recent development is the work of Australian-American mathematician Frank Calegari (1975–), Bulgarian mathematician Vesselin Dimitrov (1986–), and Chinese mathematician Yunqing Tang (1989–) [CDT24], who demonstrated that the Dirichlet L-function with non-principal Dirichlet character modulo 3 is irrational. That is,

  $$L(2, \chi_3) = 1 - \frac{1}{2^2} + \frac{1}{4^2} - \frac{1}{5^2} + \cdots$$

  is irrational. Their approach suggests that the techniques of Apéry might be extended to L-functions of different characters, potentially yielding new irrationality results for specific $L(k, \chi)$ where $k \geq 2$.

- **Refining Beta Value Sets**
  As discussed in § 7.3, current methods allow us to conclude that at least one value within a set of six Dirichlet beta values is irrational. A primary goal for future work is the reduction of this set size (ideally towards the isolation of a single irrational value). Furthermore, there is a distinct lack of *elementary methods* in literature; establishing the irrationality of at least one value in a small set utilizing elementary means would be a significant contribution.

- **Improving Irrationality Exponents**
  The quantitative side of these results, such as bounding the irrationality exponent $\mu(\alpha)$, offers a separate set of challenges. While we have established the upper bound $\mu(\zeta(3)) \leq 13.41\ldots$ in § 5, this serves as a historical baseline rather than being the tightest upper bound. The current record, $\mu(\zeta(3)) \leq 5.51\ldots$, is currently held by French mathematician Georges Rhin and Italian mathematician Carlo Viola (1943–) [RV01]. Finding novel constructions of rational functions and sequences or optimizing the so-called saddle-point method remains a viable path for improving these exponents.

- **Arithmetic of p-adic Zeta and L-functions**
  Shifting from the Archimedean case to the p-adic setting alters our concept of irrationality and

re-frames our arithmetic questions. One could naturally extend the inquires of this paper to the p-adic Riemann zeta function and its p-adic generalizations. A good introduction to the theory of p-adic special functions is given in [JW25]. The structure of the p-adic Riemann zeta function and its generalization is much less explored making it a promising path. See, for example, [LSZ25], where Chinese mathematician Li Lai, German mathematician Johannes Sprang, and Zudilin proved that the 2-adic zeta function at 5 is irrational. We can also prove similar things to that seen in § 6. For example, Lai and Sprang [LS26] prove that for any given prime p, infinitely many p-adic zeta values $\zeta_p(s)$ are irrational. As a last remark, one could show that at least one from a small set is irrational, which is exactly what Lai did in [Lai25]. It does however seem that the literature is more thorough for the p-adic Riemann zeta function in comparison to its generalizations. Thus, a more in-depth review of these generalizations could be fruitful. In particular, a result similar to Theorem (7.7) for the p-adic beta function does not seem to appear.

- **Arithmetic Nature of Derivatives of Special Functions**
  While the bulk of the literature focuses on values of $\zeta(s)$ and $L(s, \chi)$, one could ask a similar set of questions about the arithmetic nature of their derivatives. For instance, the derivative of the Riemann zeta function at even integers, such as $\zeta'(2)$, is intrinsically related to other mathematical constants of interest to number theorists[1]. Whether these derivatives yield irrational or transcendental values remains an open question and presents numerous challenges. Future research could focus on constructing specialized *Padé approximations* or *modular forms* engineered to capture the structural behavior of $\zeta'(s)$ or $\beta'(s)$. In particular, we refer the reader to the work of Beukers [Beu87] for a deeper understanding of Padé approximations and modular forms.

- **Re-framing Irrationality via Differential Equations**
  As a final remark, the arithmetic results covered throughout this article can be re-contextualized through the lens of differential equations. The linear combinations and recurrence relations found in Apéry-like proofs are often deeply connected to the so-called *Picard-Fuchs equations*, certain families of algebraic varieties, or generally, to the theory of G-functions. By re-framing irrationality questions to specific differential systems, one can leverage powerful geometric machinery. Investigating the monodromy groups and differential Galois groups associated with these systems offers a more structured pathway to understanding *why* certain special values possess specific irrationality properties. For a study of the Picard-Fuchs equations, we refer the reader to a paper by Beukers [Beu83] which discusses this and the generating functions of the Apéry sequences we defined in previous sections. For the theory of G-functions, monodromy, and Galois groups, we refer to the work of French mathematician Yves André (1959–) [And89].

References

[AR79] K. Alladi and M. L. Robinson. "On certain irrational values of the logarithm". In: *Number Theory Carbondale 1979*. Ed. by Melvyn B. Nathanson. Vol. 751. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1979, pp. 1–9. ISBN: 978-3-540-34852-8. DOI: 10.1007/BFb0062698. URL: https://doi.org/10.1007/BFb0062698.

[And89] Yves André. *G-Functions and Geometry*. Vol. E13. Aspects of Mathematics. Braunschweig: Friedr. Vieweg & Sohn, 1989. ISBN: 978-3-528-06318-4.

[Apé79] Roger Apéry. "Irrationalité de $\zeta(2)$ et $\zeta(3)$". In: *Astérisque* 61 (1979). Luminy Conference on Arithmetic, pp. 11–13. ISSN: 0303-1179.

1. In particular, $\zeta'(2)$ is related to the infamous Euler-Mascheroni constant.

[Apo76] Tom M. Apostol. *Introduction to Analytic Number Theory*. Undergraduate Texts in Mathematics. New York, Heidelberg, Berlin: Springer-Verlag, 1976. ISBN: 0-387-90163-9.

[BR01] Keith Ball and Tanguy Rivoal. “Irrationalité d’une infinité de valeurs de la fonction zéta aux entiers impairs”. In: *Inventiones Mathematicae* 146.1 (2001), pp. 193–207. DOI: 10.1007/s002220100168.

[Beu79] Frits Beukers. “A note on the irrationality of $\zeta(2)$ and $\zeta(3)$”. In: *Bulletin of the London Mathematical Society* 11.3 (1979), pp. 268–272. ISSN: 0024-6093. DOI: 10.1112/blms/11.3.268. URL: https://doi.org/10.1112/blms/11.3.268.

[Beu83] Frits Beukers. “Differential equations associated with $\zeta(2)$ and $\zeta(3)$”. In: *Séminaire de Théorie des Nombres de Bordeaux* 12.4 (1983), pp. 1–11.

[Beu87] Frits Beukers. “Irrationality proofs using modular forms”. In: *Astérisque* 147 (1987), pp. 271–283.

[CDT24] Frank Calegari, Vesselin Dimitrov, and Yunqing Tang. *The linear independence of* 1, $\zeta(2)$*, and* $L(2, \chi_{-3})$. 2024. arXiv: 2408.15403 [math.NT]. URL: https://arxiv.org/abs/2408.15403.

[Gam01] Theodore Gamelin. *Complex Analysis*. Undergraduate Texts in Mathematics. Springer New York, 2001. ISBN: 9780387950938.

[JW25] Joaquín Rodrigues Jacinto and Chris Williams. “An introduction to p-adic L-functions”. In: *Essential Number Theory* 4 (2025), pp. 101–216. DOI: 10.2140/ent.2025.4.101.

[Lai25] Li Lai. “On the irrationality of certain 2-adic zeta values”. In: *International Journal of Number Theory* 21.1 (2025), pp. 207–235. DOI: 10.1142/S1793042125500113.

[LS26] Li Lai and Johannes Sprang. “Many p-adic odd zeta values are irrational”. In: *Michigan Mathematical Journal* (2026). Advance Publication, pp. 1–36. DOI: 10.1307/mmj/20236490.

[LSZ25] Li Lai, Johannes Sprang, and Wadim Zudilin. *A note on the irrationality of* $\zeta_2(5)$. 2025. arXiv: 2505.05005 [math.NT]. URL: https://arxiv.org/abs/2505.05005.

[Poo79] Alfred van der Poorten. “A proof that Euler missed... Apéry’s proof of the irrationality of $\zeta(3)$”. In: *The Mathematical Intelligencer* 1.4 (1979). An informal report, pp. 195–203. ISSN: 0343-6993. DOI: 10.1007/BF03028234. URL: https://doi.org/10.1007/BF03028234.

[RV01] Georges Rhin and Carlo Viola. “The group structure for $\zeta(3)$”. eng. In: *Acta Arithmetica* 97.3 (2001), pp. 269–293. DOI: 10.4064/aa97-3-4. URL: http://eudml.org/doc/278880.

[Ste05] Jürgen Steuding. *Diophantine Analysis*. Discrete Mathematics and Its Applications. Taylor & Francis, 2005. ISBN: 9781584884828. URL: https://books.google.com/books?id=bUD8phVTSj8C.

[Zud18] Wadim Zudilin. “One of the odd zeta values from $\zeta(5)$ to $\zeta(25)$ is irrational. By elementary means”. In: *SIGMA. Symmetry, Integrability and Geometry. Methods and Applications* 14 (2018), Paper No. 028, 8. ISSN: 1815-0659. DOI: 10.3842/SIGMA.2018.028. URL: https://doi.org/10.3842/SIGMA.2018.028.

[Zud19] Wadim Zudilin. “Arithmetic of Catalan’s constant and its relatives”. In: *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg* 89.1 (2019), pp. 45–53. ISSN: 0025-5858. DOI: 10.1007/s12188-019-00203-w. URL: https://doi.org/10.1007/s12188-019-00203-w.

[Zud01] Wadim V. Zudilin. “One of the numbers $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, $\zeta(11)$ is irrational”. In: *Uspekhi Matematicheskikh Nauk* 56.4(340) (2001), pp. 149–150. DOI: 10.4213/rm715.