TRANSCENDENCE AND IRRATIONALITY PROOFS

LEE A. BUTLER

Abstract. This essay covers the history and proof of two famous mathematical results, one on the transcendence of a large class of numbers, and one on the irrationality of a single number.

1. Introduction

The study of equations forms an uninterrupted backbone of mathematical research over the last three thousand years. As Baker points out, “a history of the subject amounts, more or less, to a history of mathematics itself.”[3] To try to investigate the entirety of this subject would be folly, it is just too vast. Instead we will look at the solutions of certain equations, or rather the numbers which *cannot* be solutions of these equations.

We shall limit ourselves to polynomial equations in the ring $\mathbb{Z}[x]$ in this essay, for we will be studying the field $\overline{\mathbb{Q}}$ of algebraic numbers. In this essay we will look at two kinds of number: irrational numbers, which are those complex numbers that are not a root of any polynomial of degree 1 in $\mathbb{Z}[x]$, and the transcendental numbers, which are not the root of *any* polynomial in $\mathbb{Z}[x]$.

The discovery of irrational numbers far predates that of transcendentals. The first proof of the irrationality of a number is usually credited to Hippasus, a Pythagorean, who proved that $\sqrt{2}$ is irrational using a geometric approach. Due to the relative simplicity of the result there now exist a vast plethora of proofs.

When set theory was developed in the nineteenth century it was uncovered that “almost all” real numbers are irrational, that is, if we pick a real number at random the probability that it will be irrational is 1. This is actually quite intuitive if you consider the difference between rational and irrational numbers. We can define a real number x in the interval $(0, 1)$ by a sequence of digits $d_1, d_2, d_3, \dots$ which will be the decimal expansion of x, that is $x = 0.d_1d_2d_3\dots$. The condition that x is rational is equivalent to the condition that eventually this sequence enters some repeating pattern. But if we use some massive supercomputer to randomly generate the digits d_n then the thought of this computer after some finite time beginning to churn out the same series of numbers over and over from now until infinity is clearly absurd. So barring some miracle the random number will be irrational.

The first proof that displayed the transcendence of a number had to wait until 1844 when Joseph Liouville proved that the so called “Liouville numbers” were not algebraic[8]. More familiar numbers such as e and π had been suspected of being transcendental for some time, in 1755 Euler wrote the following about π:

> “It appears to be fairly certain that the periphery of a circle constitutes such a peculiar kind of transcendental quantity that it can in no way be compared with other quantities, either roots or other transcendentals.”

And indeed, thirty years after Lioville's result Hermite showed that e was transcendental, and in 1882 Lindemann backed up Euler's suspicions by proving the transcendence of π. Further proofs were given over the coming decades, but there are still many unknowns about transcendental numbers, and proving anything about them remains a great challenge.

2. A Transcendence Proof

2.1. e and π: The Lindemann-Weierstrass Theorem.

The irrationality of e was established by Euler in 1744 and that of π was proven by Johann Heinrich Lambert in 1761. Their transcendence was proved about a century later by Hermite and Lindemann respectively. Their work established that e^{α} is transcendental for any nonzero algebraic number α. The transcendence of e and π follow immediately from this. For e we can take $\alpha = 1$ while for π we proceed by contradiction. If π is algebraic then so is πi, but then $e^{\pi i} = -1$ should be transcendental, which clearly it is not.

A generalisation of the above result was given by Weierstrass in 1885, and is as follows.

Theorem 1 (Lindemann-Weierstrass theorem)**.** For any $N > 0$, given $N+1$ distinct algebraic numbers $\alpha_0, \alpha_1, \ldots, \alpha_N$, the numbers $e^{\alpha_0}, e^{\alpha_1}, \ldots, e^{\alpha_N}$ are linearly independent over $\overline{\mathbb{Q}}$.[1]

This result is very powerful indeed in establishing the transcendence of many numbers, since so many familiar objects in maths can be related to the exponential function somehow. The following are just a handful of results stemming straight from the above theorem.

Corollary 2.1. Given $N+1$ distinct nonzero algebraic numbers $\alpha_0, \alpha_1, \ldots, \alpha_N$ and any nonzero algebraic numbers $\beta_0, \beta_1, \ldots, \beta_N$, the number

$$\sum_{n=0}^{N} \beta_n e^{\alpha_n}$$

is transcendental.

Proof. Suppose the statement is false and $\sum_{n=0}^{N} \beta_n e^{\alpha_n}$ is in fact algebraic, say $\sum_{n=0}^{N} \beta_n e^{\alpha_n} = \gamma = \gamma e^0$. Now set $\alpha_{N+1} = 0$ and $\beta_{N+1} = -\gamma$, and we thus have $\sum_{n=0}^{N+1} \beta_n e^{\alpha_n} = 0$ for distinct algebraic α_i. But this contradicts the linear independence of the e^{α_i}, and so we must have that $\sum_{n=0}^{N} \beta_n e^{\alpha_n}$ is transcendental. □

Corollary 2.2. $\log(\alpha)$ is transcendental for any real algebraic number $\alpha \neq 0, 1$.

Proof. Suppose $\log(\alpha)$ is algebraic, then by the special case of Theorem 1 with $N = 1, \alpha_0 = \alpha, \alpha_1 = 0$, we have that $e^{\log(\alpha)}$ is transcendental. But $e^{\log(\alpha)} = \alpha \in \overline{\mathbb{Q}}$, thus we have a contradiction. □

[1] Some texts prefer to state the theorem in terms of α_i which are linearly independent over $\mathbb{Q}$ and then e^{α_i} being linearly independent over $\mathbb{Q}$. The two formulations are equivalent, so here we proceed with the statement as preferred in [3], [10], and [5].

Corollary 2.3. For any real nonzero algebraic number α, $\cos(\alpha)$ is transcendental.

Proof. Suppose $\cos(\alpha)$ is algebraic for some real nonzero algebraic number α, say $\cos(\alpha) = \beta$. Then we would have:

$$\cos(\alpha) = \frac{e^{i\alpha} + e^{-i\alpha}}{2i} = \frac{e^{i\alpha}}{2i} + \frac{e^{-i\alpha}}{2i} = \beta.$$

We can rewrite this as

$$\left(\frac{-i}{2}\right) e^{i\alpha} + \left(\frac{-i}{2}\right) e^{-i\alpha} + (-\beta)e^{0} = 0.$$

But $-i\alpha, 0, i\alpha$ are distinct algebraic numbers so $e^{-i\alpha}, e^{0}, e^{i\alpha}$ should be linearly independent over $\overline{\mathbb{Q}}$, hence we have a contradiction.

Alternatively this result follows immediately from Corollary 2.1 applied to the sum $\cos(\alpha) = \left(\frac{-i}{2}\right) e^{i\alpha} + \left(\frac{-i}{2}\right) e^{-i\alpha}$. □

Virtually identical proofs to the last one show that sin, sinh, and cosh applied to a real nonzero algebraic number gives a transcendental result.

I will strive here to give an elementary proof of Theorem 1 in the sense that I will use only results from Algebraic Number Theory rather than analytic results. First and foremost the following useful result is needed.

Lemma 2.1. Let $h(z) \in \mathbb{Z}[z]$ be given by $h(z) = \sum_{n=0}^{N} c_n z^n$ and let $\alpha_1, \alpha_2, \ldots, \alpha_N$ be the zeroes of $h(z)$. Now let $\mathcal{P}(z)$ be any polynomial with integer coefficients. Then the quantity

$$\mathcal{P}(\alpha_1) + \mathcal{P}(\alpha_2) + \ldots + \mathcal{P}(\alpha_N)$$

is a rational number whose denominator can be written as $c_N^{\deg(\mathcal{P})}$.

Proof. First we note that the polynomial $\mathfrak{P}\left(\alpha_1, \ldots, \alpha_N\right) = \mathcal{P}(\alpha_1) + \ldots + \mathcal{P}(\alpha_N)$ is symmetric, since permuting the α_i's only changes the order of addition on the right hand side, hence we can write it as a polynomial in the elementary symmetric polynomials[2] $\sigma_1, \sigma_2, \ldots, \sigma_N$ of $\alpha_1, \alpha_2, \ldots, \alpha_N$. Specifically we can write

$$\mathfrak{P}\left(\alpha_1, \ldots, \alpha_N\right) = \mathfrak{F}\left(\sigma_1, \ldots, \sigma_N\right) \in \mathbb{Z}[\sigma_1, \ldots, \sigma_N]$$

where $\deg\left(\mathfrak{F}\right) \leq \deg\left(\mathfrak{P}\right)$. We also know that

$$\begin{aligned} h(z) &= c_N z^N + c_{N-1} z^{N-1} + \ldots + c_1 z + c_0 \\ &= c_N \prod_{i=1}^{N} (z - \alpha_i) \\ &= c_N \left(z^N - \sigma_1 z^{N-1} + \ldots + (-1)^{N-1}\sigma_{N-1} z + (-1)^N \sigma_N\right). \end{aligned}$$

[2] Here we will suppress the arguments of the σ_i's, letting σ_i denote the more ungainly $\sigma_i(\alpha_1, \ldots, \alpha_N)$.

Comparing coefficients gives the well known equations relating the roots of a polynomial to its coefficients:

$$\sigma_k = (-1)^k \frac{c_{N-k}}{c_N}.$$

From this it follows that each of the elementary symmetric functions in the α_k's is a rational number with denominator c_N. Thus when we plug these rational numbers into $\mathfrak{F}(\sigma_1, \ldots, \sigma_N)$ we will get out a rational number (recall that $\mathfrak{F}$ is a polynomial with integer coefficients) with denominator $c_N^{\deg(\mathfrak{F})}$. But $\deg(\mathfrak{F}) \leq \deg(\mathfrak{P})$ so we can multiply the numerator and denominator of this rational number by the integer $c_N^{\deg(\mathfrak{P})-\deg(\mathfrak{F})}$ for the result of the Lemma.

□

Using this Lemma we can prove a weakened form of Theorem 1. That may not sound like a particularly good strategy; why prove a weak form of the Theorem when we could just aim straight for a proof of the full result? Well the weakened form actually leads directly to a proof of the Lindemann-Weierstrass theorem, and possesses a lot of useful symmetry which makes a proof far simpler. This weakened version is as follows[3].

Theorem 2. Let $A = \{\alpha_1, \alpha_2, \ldots, \alpha_N\}$ be a set of distinct nonzero algebraic numbers with the property that if $\alpha_n \in A$ then each conjugate of α_n is also in A. Suppose that $\beta_0, \beta_1, \ldots, \beta_N$ are nonzero integers satisfying the condition that if $\alpha_i, \alpha_j \in A$ are conjugate then $\beta_i = \beta_j$. Then:

$$\beta_0 + \sum_{n=1}^{N} \beta_n e^{\alpha_n} \neq 0.$$

Proof. Since proving that something is nonzero can be a bit tricky we will proceed by contradiction and assume that there exist nonzero integers $\beta_0, \beta_1, \ldots, \beta_N$ that satisfy all the hypotheses of the theorem, but which also satisfy

$$(\spadesuit) \qquad \beta_0 + \sum_{n=1}^{N} \beta_n e^{\alpha_n} = 0.$$

Next we simplify things by grouping the α_i's together into complete sets of conjugates. We can do this since the set A only contains such sets, so we may relabel them as

$$\alpha_{11}, \alpha_{12}, \ldots, \alpha_{1N_1}, \quad \alpha_{21}, \alpha_{22}, \ldots, \alpha_{2N_2}, \quad \ldots, \quad \alpha_{M1}, \alpha_{M2}, \ldots, \alpha_{MN_M}.$$

Here $\{\alpha_{m1}, \alpha_{m2}, \ldots, \alpha_{mN_m}\}$ is a complete set of conjugates, and our original set of algebraic numbers contains M such sets. If we relabel the β_n's in sympathy with the α_n's, so that the coefficient of $e^{\alpha_{mi}}$ in ($\spadesuit$) is β_{mi} then by our hypothesis we know that $\beta_{m1} = \beta_{m2} = \ldots = \beta_{mN_m}$, so we will label all these β_{mi} as simply β_m. As a final piece of notation in this proof we will let $\hat{\alpha}$ denote the element of

[3] The proof of this weakened version and, later, the full Lindemann-Weierstrass theorem follow a series of exercises in [5].

A with maximal absolute value. With all this notation we can rewrite our initial assumption (♠) as

$$\beta_0 + \sum_{m=1}^{M} \beta_m \left(\sum_{n=1}^{N_m} e^{\alpha_{mn}} \right) = 0.$$

Now let $f_m(z) = (z-\alpha_{m1})(z-\alpha_{m2})\cdots(z-\alpha_{mN_m})$. That is, $f_m(z)$ is the minimum polynomial, up to multiplication by an integer, for all the algebraic numbers in the set $\{\alpha_{m1}, \alpha_{m2}, \ldots, \alpha_{mN_m}\}$, and hence is in $\mathbb{Q}[z]$. Let $d_m > 0$ be the lowest common multiple of the denominators of all the coefficients of f_m, so $d_m f_m \in \mathbb{Z}[z]$.

We now define the auxiliary polynomial which will eventually lead to our contradiction. This polynomial is defined as

$$f(z) = (d_1 d_2 \cdots d_M)^p \, z^{p-1} f_1(z)^p f_2(z)^p \cdots f_M(z)^p$$

for an as yet undetermined prime number p. We can rewrite this as

$$f(z) = \sum_{j=p-1}^{(N+1)p-1} c_j z^j.$$

Since $f(z) = (d_1 f_1(z))^p \, (d_2 f_2(z))^p \cdots (d_M f_M(z))^p \, z^{p-1}$ and each bracketed polynomial is in $\mathbb{Z}[z]$ we have that $f \in \mathbb{Z}[z]$, so $c_j \in \mathbb{Z}$ for $p-1 \leq j \leq (N+1)p-1$.

By inspection we see that

$$\begin{aligned} c_{p-1} &= \pm \left(d_1 d_2 \cdots d_M \alpha_{11} \alpha_{12} \cdots \alpha_{MN_M} \right)^p \\ &= \pm \left(d_1 d_2 \cdots d_M \alpha_1 \alpha_2 \cdots \alpha_N \right)^p \end{aligned}$$

We know that $d_i > 0$ for each i, and by hypothesis the α_j's are nonzero, so we can note for future reference that $c_{p-1} \neq 0$.

Now we want to prove that

$$\sum_{j=1}^{p-1} f^{(j)}(z) = \sum_{J=p-1}^{(N+1)p-1} \left(J! c_J \sum_{j=J-p+1}^{J-1} \frac{z^j}{j!} \right).$$

To verify this we note it's a simple matter to check that

$$f^{(j)}(z) = \sum_{J=p-1}^{(N+1)p-1} \frac{J!}{(J-j)!} c_J z^{J-j} \qquad \text{for } 1 \leq j \leq p-1.$$

Hence,

$$\begin{aligned} \sum_{j=1}^{p-1} f^{(j)}(z) &= \sum_{j=1}^{p-1} \left(\sum_{J=p-1}^{(N+1)p-1} \frac{J!}{(J-j)!} c_J z^{J-j} \right) \\ &= \sum_{J=p-1}^{(N+1)p-1} \left(J! c_J \sum_{j=1}^{p-1} \frac{z^{J-j}}{(J-j)!} \right) \\ &= \sum_{J=p-1}^{(N+1)p-1} \left(J! c_J \sum_{j=J-(p-1)}^{J-1} \frac{z^j}{j!} \right), \end{aligned}$$

which was what we wanted. We can also see that since $f(z) = g(z)(z-\alpha)^p$ for each $\alpha \in A$, where $g(z)$ is some polynomial dependent on α, then $f^{(j)}(\alpha) = 0$ for $1 \leq j \leq p-1$. So

$$\sum_{j=1}^{p-1} f^{(j)}(\alpha) = 0$$

for all $\alpha \in A$.

The fact that a sum involving $\dfrac{z^j}{j!}$ appears above suggests that we might now want to to look at the following product.

$$e^{\alpha} \sum_{J=p-1}^{(N+1)p-1} J! c_J = \sum_{J=p-1}^{(N+1)p-1} \left(J! c_J \sum_{j=0}^{\infty} \frac{\alpha^j}{j!} \right)$$

$$= \sum_{J=p-1}^{(N+1)p-1} \left(J! c_J \sum_{j=0}^{J-p} \frac{\alpha^j}{j!} \right) + \sum_{J=p-1}^{(N+1)p-1} \left(J! c_J \sum_{j=J-p+1}^{J-1} \frac{\alpha^j}{j!} \right) + \sum_{J=p-1}^{(N+1)p-1} \left(J! c_J \sum_{j=J}^{\infty} \frac{\alpha^j}{j!} \right).$$

The two minor results above tell us that the middle of these three sums is in fact $\sum_{j=1}^{p-1} f^{(j)}(\alpha) = 0$. So we have:

$$\begin{aligned} e^{\alpha} \sum_{J=p-1}^{(N+1)p-1} J! c_J &= \sum_{J=p-1}^{(N+1)p-1} \left(J! c_J \sum_{j=0}^{J-p} \frac{\alpha^j}{j!} \right) + \sum_{J=p-1}^{(N+1)p-1} \left(J! c_J \sum_{j=J}^{\infty} \frac{\alpha^j}{j!} \right) \\ &= \mathcal{P}_p(\alpha) + \mathcal{T}_p(\alpha) \end{aligned}$$

where $\mathcal{P}_p$ and $\mathcal{T}_p$ are the first and second sums respectively. (A polynomial and a tail.)

If you recall, the assumption contrary to the theorem that we were hoping to build a contradiction out of, (♠), was that

$$\beta_0 + \sum_{m=1}^{M} \beta_m \left(\sum_{n=1}^{N_m} e^{\alpha_{mn}} \right) = 0.$$

Let us multiply both sides of this identity by $\dfrac{J! c_J}{(p-1)!}$ and sum from $J = p-1$ to $(N+1)p-1$, so we get

$$\sum_{J=p-1}^{(N+1)p-1} \beta_0 \frac{J! c_J}{(p-1)!} + \sum_{J=p-1}^{(N+1)p-1} \frac{J! c_J}{(p-1)!} \left(\sum_{m=1}^{M} \beta_m \left(\sum_{n=1}^{N_m} e^{\alpha_{mn}} \right) \right) = 0.$$

Changing the order of summation gives

$$\beta_0 \sum_{J=p-1}^{(N+1)p-1} \frac{J!c_J}{(p-1)!} + \sum_{m=1}^{M} \beta_m \left(\frac{1}{(p-1)!} \sum_{J=p-1}^{(N+1)p-1} J!c_J \sum_{n=1}^{N_m} e^{\alpha_{mn}} \right) = 0.$$

But the double sum in the brackets is just the sum of $\mathcal{P}_p(\alpha_{mn})$ and $\mathcal{T}_p(\alpha_{mn})$:

$$\beta_0 \sum_{J=p-1}^{(N+1)p-1} \frac{J!c_J}{(p-1)!} + \sum_{m=1}^{M} \beta_m \left(\frac{1}{(p-1)!} \sum_{n=1}^{N_m} (\mathcal{P}_p(\alpha_{mn}) + \mathcal{T}_p(\alpha_{mn})) \right) = 0.$$

And taking the $\mathcal{T}_p(\alpha_{mn})$ to the other side leads us to

$$(\clubsuit)\ \beta_0 \sum_{J=p-1}^{(N+1)p-1} \frac{J!c_J}{(p-1)!} + \sum_{m=1}^{M} \beta_m \left(\sum_{n=1}^{N_m} \frac{\mathcal{P}_p(\alpha_{mn})}{(p-1)!} \right) = -\sum_{m=1}^{M} \beta_m \left(\sum_{n=1}^{N_m} \frac{\mathcal{T}_p(\alpha_{mn})}{(p-1)!} \right).$$

Clearly $\beta_0 \sum_{J=p-1}^{(N+1)p-1} \frac{J!c_J}{(p-1)!} \in \mathbb{Z}$ (recall that $\beta_0 \in \mathbb{Z}$). Now recall that

$$\mathcal{P}_p(z) = \sum_{J=p-1}^{(N+1)p-1} J!c_J \sum_{j=0}^{J-p} \frac{z^j}{j!}.$$

But the first term in this sum, when $J = p-1$, is an empty sum and thus equal to zero:

$$(p-1)!c_{p-1} \sum_{j=0}^{-1} \frac{z^j}{j!} = 0$$

So in fact

$$\mathcal{P}_p(z) = \sum_{J=p}^{(N+1)p-1} J!c_J \sum_{j=0}^{J-p} \frac{z^j}{j!} \in \mathbb{Z}[z].$$

We now want to prove that $p!$ divides every coefficient of $\mathcal{P}_p(z)$ so we can apply Lemma 2.1 to $\frac{\mathcal{P}_p}{p!}$ and f_m. Obviously $p! \mid J!$ for each $J \geq p$, but the problem is that the $\frac{1}{j!}$ terms in the inner sum may cancel out this divisibility property. The worst case is when j is maximal. If in this case $\frac{J!}{j!}$ is divisible by $p!$ then we will be okay for all the smaller values of j. If $J = p + r$ then this maximal value of j is $(p+r) - p = r$. Then we have

$$\frac{J!}{j!} = \frac{(p+r)!}{r!} = p! \binom{p+r}{r}$$

which is clearly divisible by $p!$, so we have verified the above claim and shown that $\frac{1}{p!}\mathcal{P}_p(z) \in \mathbb{Z}[z]$.

We now apply Lemma 2.1. In our case $h(z)$ is $d_m f_m(z)$ which has zeroes $\alpha_{m1}, \ldots, \alpha_{mN_m}$, and $\mathcal{P}(z)$ is $\frac{\mathcal{P}_p(z)}{p!}$ which is an integer polynomial by the above reasoning. We know that $\deg \mathcal{P}_p \leq Np - 1$ so by Lemma 2.1 for each m there is an integer $\tilde{a}_m$ such that

$$\sum_{n=1}^{N_m} \frac{\mathcal{P}_p(\alpha_{mn})}{p!} = \frac{\tilde{a}_m}{d_m^{Np-1}}.$$

Multiplying through by p and setting $a_m = p\tilde{a}_m$ we can see that for each m there is an integer a_m such that $p \mid a_m$ and

$$\sum_{n=1}^{N_m} \frac{\mathcal{P}_p(\alpha_{mn})}{(p-1)!} = \frac{a_m}{d_m^{Np-1}}.$$

With this result we can now see that (♣) becomes

$$\beta_0 \sum_{J=p-1}^{(N+1)p-1} \frac{J!c_J}{(p-1)!} + \sum_{m=1}^{M} \frac{a_m\beta_m}{d_m^{Np-1}} = -\sum_{m=1}^{M} \beta_m \left(\sum_{n=1}^{N_m} \frac{\mathcal{T}_p(\alpha_{mn})}{(p-1)!} \right).$$

We now define the integer $D = d_1 d_2 \cdots d_M$, and multiply both sides above by D^{Np} to get

$$\beta_0 D^{Np} \sum_{J=p-1}^{(N+1)p-1} \frac{J!c_J}{(p-1)!} + \sum_{m=1}^{M} a_m\beta_m d_m (D/d_m)^{Np} = -\sum_{m=1}^{M} \beta_m D^{Np} \left(\sum_{n=1}^{N_m} \frac{\mathcal{T}_p(\alpha_{mn})}{(p-1)!} \right).$$

Clearly D/d_m is an integer for each m so the left hand side of this identity is an integer, which we will denote $\mathcal{N}$. This integer $\mathcal{N}$ will eventually provide our contradiction, for we will show that for the right choice of p, $\mathcal{N}$ must be an integer satisfying $0 < |\mathcal{N}| < 1$, which is clearly impossible. First we will show it is nonzero.

Let us rewrite $\mathcal{N}$ as

$$\mathcal{N} = \beta_0 D^{Np} c_{p-1} + \left(\beta_0 D^{Np} \sum_{J=p}^{(N+1)p-1} \frac{J!c_J}{(p-1)!} + \sum_{m=1}^{M} a_m\beta_m d_m (D/d_m)^{Np} \right).$$

We found some time ago that $c_{p-1} \neq 0$. We also know that $p \mid a_m$ for each m, and clearly $p \mid J!$ for $J \geq p$. So p divides the integer in the brackets, while β_0, D, and c_{p-1} are all fixed nonzero integers. Thus if we choose $p > \max\{|\beta_0|, D, |c_{p-1}|\}$ then $p \nmid \beta_0 D^{Np} c_{p-1}$.

Putting this together we see that $\mathcal{N} \not\equiv 0 \pmod{p}$, thus we definitely have that $\mathcal{N} \neq 0$. If we now want to find a value of p such that $|\mathcal{N}| < 1$ then we need to look at the alternative way of defining $\mathcal{N}$, that is

$$\mathcal{N} = -\sum_{m=1}^{M} \beta_m D^{Np} \left(\sum_{n=1}^{N_m} \frac{\mathcal{T}_p(\alpha_{mn})}{(p-1)!} \right).$$

Using the triangle inequality gives

$$0 < |\mathcal{N}| \leq \sum_{m=1}^{M} \left(\sum_{n=1}^{N_m} \frac{\left|\beta_m D^{Np} \mathcal{T}_p(\alpha_{mn})\right|}{(p-1)!} \right).$$

If we can show that the sum on the right is strictly less than 1 then we will be done. So let us try to estimate the size of $\mathcal{T}_p(\alpha_{mn})$.

$$\begin{aligned}\mathcal{T}_p(z) &= \sum_{J=p-1}^{(N+1)p-1}\left(J!c_J\sum_{j=J}^{\infty}\frac{z^j}{j!}\right)\\ &= \sum_{J=p-1}^{(N+1)p-1}\left(c_J\sum_{j=0}^{\infty}\frac{J!}{(j+J)!}z^{j+J}\right).\end{aligned}$$

But

$$\frac{(j+J)!}{j!J!} = \binom{j+J}{j} \geq 1$$

so that

$$\frac{J!}{(j+J)!} \leq \frac{1}{j!}.$$

So

$$\begin{aligned}|\mathcal{T}_p(\alpha)| &= \left|\sum_{J=p-1}^{(N+1)p-1}\left(c_J\sum_{j=0}^{\infty}\frac{J!}{(j+J)!}\alpha^{j+J}\right)\right|\\ &\leq \sum_{J=p-1}^{(N+1)p-1}\left|c_J\sum_{j=0}^{\infty}\frac{J!}{(j+J)!}\alpha^{j+J}\right|\\ &\leq \sum_{J=p-1}^{(N+1)p-1}|c_J|\sum_{j=0}^{\infty}\frac{J!}{(j+J)!}\,|\alpha|^{j+J}\\ &\leq \sum_{J=p-1}^{(N+1)p-1}|c_J||\alpha|^J\sum_{j=0}^{\infty}\frac{|\alpha|^j}{j!}\\ &\leq |\hat{\alpha}|^{(N+1)p-1}\,e^{|\alpha|}\sum_{J=p-1}^{(N+1)p-1}|c_J|.\end{aligned}$$

Here we have assumed that $|\hat{\alpha}| \geq 1$ when pulling the α's out of the sum. If this is not the case then we could replace $|\hat{\alpha}|$ in all that follows by 1. Since that would simplify the rest of the proof slightly we will assume here that in fact $|\hat{\alpha}| \geq 1$. Hence to bound $\mathcal{T}_p(\alpha)$ we must find an upper bound for $\sum_{J=p-1}^{(N+1)p-1}|c_J|$. Recall that

$$f(z) = \sum_{j=p-1}^{(N+1)p-1} c_j z^j = D^p z^{p-1}(z-\alpha_1)^p(z-\alpha_2)^p\cdots(z-\alpha_N)^p.$$

Each bracketed factor could be written as

$$(z-\alpha_n)^p = \sum_{l=0}^{p} \binom{p}{l} (-\alpha_n)^{p-l} z^l.$$

This suggests that to find an upper bound for the sum of the $|c_J|$'s we should look at

$$\begin{aligned}\max_{l=0,1,\ldots,p} \left\{ \left| \binom{p}{l} (-\alpha)^{p-l} \right| \right\} &\leq \sum_{l=0}^{p} \binom{p}{l} |\hat{\alpha}|^{p-l} \\ &\leq |\hat{\alpha}|^p \sum_{l=0}^{p} \binom{p}{l} \\ &= (2\,|\hat{\alpha}|)^p .\end{aligned}$$

So each coefficient in these factors is at most $(2\,|\hat{\alpha}|)^p$ and there are N factors giving

$$|c_j| \leq D^p \prod_{n=1}^{N} (2\,|\hat{\alpha}|)^p = D^p\,(2\,|\hat{\alpha}|)^{Np} .$$

This bound may seem excessively poor, but the key point is that it is of the form K^p for some constant K. Plugging this bound into the sum we were considering gives

$$\sum_{J=p-1}^{(N+1)p-1} |c_J| \leq Np \left(2D^{1/N}\,|\hat{\alpha}|\right)^{Np} .$$

And so we can finally place a bound on $\mathcal{T}_p(\alpha)$ as follows.

$$\begin{aligned}|\mathcal{T}_p(\alpha)| &\leq |\hat{\alpha}|^{(N+1)p-1} e^{|\alpha|} Np \left(2D^{1/N}\,|\hat{\alpha}|\right)^{Np} \\ &\leq e^{|\alpha|} N^{p^2}\,|\hat{\alpha}|^{(N+1)p-1} \left(2D^{1/N}\,|\hat{\alpha}|\right)^{Np} \\ &= \frac{e^{|\alpha|}}{|\hat{\alpha}|} N^p D^p \left((2\,|\hat{\alpha}|)^N\right)^p \left(|\hat{\alpha}|^{N+1}\right)^p \\ &= K_1\,(K_2)^p\end{aligned}$$

Where

$$K_1 = \frac{e^{|\alpha|}}{|\hat{\alpha}|} \qquad\qquad K_2 = ND\left(2\,|\hat{\alpha}|^{2N+1}\right).$$

Again, the important thing is not the quality of these bounds, but the fact that they are of the form $K_1\,(K_2)^p$ and no worse.

We are now ready to show that our integer $\mathcal{N}$ is less than 1. Recall we had that

$$0 < |\mathcal{N}| \leq \sum_{m=1}^{M} \left(\sum_{n=1}^{N_m} \frac{\left|\beta_m D^{Np} \mathcal{T}_p(\alpha_{mn})\right|}{(p-1)!} \right).$$

Now let $B = \max\{|\beta_1|, |\beta_2|, \ldots, |\beta_M|\}$. Then in light of the bounds we've just found we can say that

$$0 < |\mathcal{N}| \leq BNK_1 \frac{\left(D^N K_2\right)^p}{(p-1)!}.$$

Each of the numbers B, N, K_1, D, and K_2 is constant and independent of p. And since $\frac{x^p}{p!} \to 0$ as $p \to \infty$ we can see that for sufficiently large primes the right hand side will be less than 1, thus we will have

$$0 < |\mathcal{N}| < 1$$

which is impossible. Hence the assumption that

$$\beta_0 + \sum_{n=1}^{N} \beta_n e^{\alpha_n} = 0$$

must have been false, and the theorem is proved to be true. □

At first glance, Theorem 2 may seem like such a diluted version of the real Lindemann-Weierstrass theorem that it will be of little use. After all, the Lindemann-Weierstrass theorem deals with an arbitrary collection of algebraic numbers and assures their linear independence over $\overline{\mathbb{Q}}$, while we have taken a far from arbitrary collection of algebraic numbers and shown that they are linearly independent over $\mathbb{Z}$. But the general case can in fact be poked and prodded until it is of the form used in our special case, and all this can be done without losing the information necessary to draw conclusions about the general case. All this will be done presently, but first we need to get a few small lemmas out of the way that will let us do the aforementioned prodding.

Lemma 2.2. Let $\rho_1, \rho_2, \ldots, \rho_L, \tau_1, \tau_2, \ldots, \tau_M \in \mathbb{C}$ be distinct and let $r_1, r_2, \ldots, r_L, t_1, t_2, \ldots, t_M \in \mathbb{C}$ be nonzero. If

$$\left(\sum_{\ell=1}^{L} r_\ell e^{\rho_\ell}\right)\left(\sum_{m=1}^{M} t_m e^{\tau_m}\right)$$

is expanded and like exponential terms are combined then the result is of the form

$$\sum_{n=1}^{N} s_n e^{\lambda_n}$$

for some $N \in \mathbb{N}, s_n \in \mathbb{C}$, and distinct $\lambda_n \in \mathbb{C}$ with $\lambda_n = \rho_\ell + \tau_m$. Moreover, at least one of the s_n is nonzero.

Proof. Clearly

$$\left(\sum_{\ell=1}^{L} r_\ell e^{\rho_\ell}\right)\left(\sum_{m=1}^{M} t_m e^{\tau_m}\right) = \sum_{\substack{1\leq \ell \leq L \\ 1 \leq m \leq M}} r_\ell t_m e^{\rho_\ell + \tau_m}.$$

Let $\mathcal{S} = \{\rho_\ell + \tau_m \mid 1 \leq \ell \leq L, 1 \leq m \leq M\}$, then $\#\mathcal{S} \geq 1$ and so we can write $\mathcal{S} = \{\lambda_1, \ldots, \lambda_N\}$ for distinct $\lambda_n \in \mathbb{C}$ and some $N \in \mathbb{N}$. Let

$$s_n = \sum_{\substack{\ell,m \text{ such that} \\ \rho_\ell + \tau_m = \lambda_n}} r_\ell t_m,$$

then

$$\left(\sum_{\ell=1}^{L} r_\ell e^{\rho_\ell} \right) \left(\sum_{m=1}^{M} t_m e^{\tau_m} \right) = \sum_{n=1}^{N} s_n e^{\lambda_n}.$$

Now look at $\{\Re(\rho_1), \ldots, \Re(\rho_L)\}$. This is a finite set of real numbers, and thus has a least element (or elements). If we say that $\{\rho_{\ell_1}, \ldots, \rho_{\ell_K}\}$ is the set of such elements with minimal real part then we can trivially note that all these elements have the same real part. By the same reasoning as above there is an element in this subset with minimal imaginary part, but this time the element must be unique. For suppose there were two elements in this subset with the same minimal imaginary part, then these two elements would be equal. But $\rho_1, \rho_2, \ldots, \rho_L$ are distinct so this could not happen.

Similarly we can find a minimal element out of $\tau_1, \tau_2, \ldots, \tau_M$. Let ρ_α and τ_α be these minimal elements, then there exists an element $\lambda_\alpha \in \mathcal{S}$ such that $\lambda_\alpha = \rho_\alpha + \tau_\alpha$ and no other two elements sum to give this element of $\mathcal{S}$. For suppose there were other elements ρ_β and τ_β such that $\lambda_\alpha = \rho_\beta + \tau_\beta$. Then $\Re(\rho_\alpha + \tau_\alpha) = \Re(\rho_\beta + \tau_\beta)$ and $\Im(\rho_\alpha + \tau_\alpha) = \Im(\rho_\beta + \tau_\beta)$, but this contradicts the minimality condition on ρ_α and τ_α, so there are no other such elements.

Since we have some λ_α uniquely expressible as $\rho_\alpha + \tau_\alpha$ we know that $s_\alpha = r_\alpha t_\alpha$, and since the r_ℓ and t_m are all nonzero we have that $s_\alpha \neq 0$, which is all we had left to prove. □

The above lemma tells us that we can multiply one exponential sum by another without everything going to zero if we have the appropriate exponents and coefficients. Hopefully that will be enough to ensure that when prodding our general case into something akin to the hypotheses in Theorem 2 we don't accidentally end up proving little more than $0 = 0$.

Next we want to see how to turn a set of arbitrary algebraic numbers into a set of algebraic numbers like those in our weak form of the Lindemann-Weierstrass theorem. Recall that in that version of the theorem our set of algebraic numbers had the useful property that if a number α was in the set then so were all conjugates of α. Rather than keep using this clunky description we will call a set of algebraic numbers with the property that if α appears n times then each conjugate of α appears n times a "conjugate complete" set of algebraic numbers.

We want to turn an arbitrary set of algebraic numbers into a conjugate complete one, and to that end the following lemma gives a useful criterion for conjugate completeness.

Lemma 2.3. The set $\{\alpha_1, \alpha_2, \dots, \alpha_L\} \subseteq \overline{\mathbb{Q}}$ is conjugate complete if and only if the polynomial $(z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_L)$ has rational coefficients.

Proof. First suppose $\{\alpha_1, \alpha_2, \dots, \alpha_L\}$ is conjugate complete, so we can split it into M conjugate complete sets in which any particular element appears only once. That is:

$$\{\alpha_1, \dots, \alpha_L\} = \{\alpha_{1_1}, \dots, \alpha_{1_{L_1}}\} \cup \dots \cup \{\alpha_{M_1}, \dots, \alpha_{M_{L_M}}\},$$

where each set $\{\alpha_{m_1}, \dots, \alpha_{m_{L_m}}\}$ is conjugate complete and no two elements are the same. So for each set like this the polynomial

$$(z - \alpha_{m_1})(z - \alpha_{m_2}) \cdots (z - \alpha_{m_{L_m}})$$

is the minimum polynomial (up to multiplication by an integer) of $\alpha_{m_1}, \dots, \alpha_{m_{L_m}}$, and thus is in $\mathbb{Q}[z]$. And the polynomial

$$(z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_L)$$

is just the product of these minimal polynomials, so is also in $\mathbb{Q}[z]$, as required.

Now suppose $f(z) = (z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_L) \in \mathbb{Q}[z]$. If any of the α_ℓ's are rational then we can remove them and what's left will still have rational coefficients. Moreover if α_ℓ is rational then $\{\alpha_\ell\}$ is conjugate complete so we can ignore all the rational α_ℓ's for now.

Having done so, relabel what remains and let $g(z) = (z - \alpha_1) \cdots (z - \alpha_K) \in \mathbb{Q}[z]$, where $\alpha_k \notin \mathbb{Q}$.

If g is irreducible then it must be the minimum polynomial of $\alpha_1, \dots, \alpha_K$, and thus $\{\alpha_1, \dots, \alpha_K\}$ would be a conjugate complete set and we would be done. If g is not irreducible then we can write

$$g = h_1^{e_1} h_2^{e_2} \cdots h_J^{e_J}$$

where each $h_j \in \mathbb{Q}[z]$ is irreducible and $e_j \geq 1$. Now we can relabel the α_i's *again* so that

$$h_j(z) = (z - \alpha_{j_1})(z - \alpha_{j_2}) \cdots (z - \alpha_{j_{L_j}}).$$

Now we note that $h_j \in \mathbb{Q}[z]$ is the minimum polynomial for $\alpha_{j_1}, \alpha_{j_2}, \dots, \alpha_{j_{L_j}}$, and so $\{\alpha_{j_1}, \alpha_{j_2}, \dots, \alpha_{j_{L_j}}\}$ is a conjugate complete set. Moreover,

$$\{\alpha_1, \alpha_2, \dots, \alpha_L\} = \bigcup_{j=1}^{J} \bigcup_{i=1}^{e_j} \{\alpha_{j_1}, \alpha_{j_2}, \dots, \alpha_{j_{L_j}}\},$$

that is our set of algebraic numbers is a union of conjugate complete sets, and thus is conjugate complete itself.

□

Now that we have a handy criterion for when a set is conjugate complete we can use this to manipulate our arbitrary set of algebraic numbers in the Lindemann-Weierstrass theorem into a conjugate complete set. The next lemma shows how.

Lemma 2.4. Let $\{\gamma_1, \ldots, \gamma_J\}$ be a set of algebraic numbers and $f(x_1, \ldots, x_J) = a_1x_1 + \ldots + a_Jx_J$ with $a_j \in \mathbb{Z}$ for $j = 1, \ldots, J$. Then the set

$$\{f(\tau_1, \ldots, \tau_J) \mid \tau_j \text{ is a conjugate of } \gamma_j \text{ for all } j\}$$

is conjugate complete.

Proof. By Lemma 2.3 this claim is equivalent to showing that the polynomial

$$p(z) = \prod_{\substack{\tau_j \text{ is a conjugate} \\ \text{of } \gamma_j \text{ for all } j}} (z - f(\tau_1, \ldots, \tau_J))$$

is in $\mathbb{Q}[z]$.

We can rewrite $p(z)$ as

$$\begin{aligned} p(z) &= \prod_{\substack{\tau_j \text{ is a conjugate} \\ \text{of } \gamma_j \text{ for} \\ 2\leq j\leq J}} \left(\prod_{\substack{\tau_1 \text{ is a conjugate} \\ \text{of } \gamma_1 \text{ for all } j}} (z - f(\tau_1, \ldots, \tau_J)) \right) \\ &= \prod_{\substack{\tau_j \text{ is a conjugate} \\ \text{of } \gamma_j \text{ for} \\ 2\leq j\leq J}} q(z). \end{aligned}$$

Now we see that if the conjugates of γ_1 are $\gamma_1 = \gamma_{1_1}, \gamma_{1_2}, \ldots, \gamma_{1_n}$ then

$$\begin{aligned} q(z) &= (z - f(\gamma_{1_1}, \tau_2, \ldots, \tau_J))(z - f(\gamma_{1_2}, \tau_2, \ldots, \tau_J)) \cdots (z - f(\gamma_{1_n}, \tau_2, \ldots, \tau_J)) \\ &= \prod_{i=1}^{n} \left(\underbrace{(z - a_2\tau_2 - \ldots - a_J\tau_J)}_{=:y} - a_1\gamma_{1_i} \right) \\ &= y^n - \sigma_1 y^{n-1} + \sigma_2 y^{n-2} - \ldots + (-1)^n \sigma_n. \end{aligned}$$

where here the σ_i's are the elementary symmetric polynomials of the conjugates of γ_1, and hence are all rational. So

$$y^n - \sigma_1 y^{n-1} + \sigma_2 y^{n-2} - \ldots + (-1)^n \sigma_n \in \mathbb{Q}[y]$$

and so $q(z) \in \mathbb{Q}[z, \tau_1, \ldots, \tau_J]$.

But the exact same argument works to eliminate any of the τ_j's, and so we must have that $p(z) \in \mathbb{Q}[z]$, and thus by Lemma 2.3, $\{f(\tau_1, \ldots, \tau_J)\}$ is conjugate complete. □

This Lemma tells us how to get a conjugate complete set out of an arbitrary one, but what we are really dealing with is the sum of e raised to such algebraic numbers, so if $\{\alpha_1, \ldots, \alpha_M\} \subset \overline{\mathbb{Q}}$ is a conjugate complete set then we call the sum

$e^{\alpha_1} + \ldots + e^{\alpha_M}$ a conjugate complete exponential sum. A nice property of these sums is as follows.

Lemma 2.5. The product of two conjugate complete exponential sums is a conjugate complete exponential sum.

Proof. Let $e^{\alpha_1} + e^{\alpha_2} + \ldots + e^{\alpha_L}$ and $e^{\beta_1} + e^{\beta_2} + \ldots + e^{\beta_M}$ be two conjugate complete exponential sums. Now,

$$\left(e^{\alpha_1} + e^{\alpha_2} + \ldots + e^{\alpha_L}\right)\left(e^{\beta_1} + e^{\beta_2} + \ldots + e^{\beta_M}\right) = \sum_{\substack{1\leq \ell \leq L \\ 1\leq m\leq M}} e^{\alpha_\ell + \beta_m},$$

so we need to show that $\{\alpha_\ell + \beta_m \mid 1 \leq \ell \leq L, 1 \leq m \leq M\}$ is a conjugate complete set of algebraic numbers.

But if we apply Lemma 2.4 to the polynomial $f(x_1, x_2) = x_1 + x_2$ then we see that $\{\alpha_\ell + \beta_m \mid 1 \leq \ell \leq L, 1 \leq m \leq M\}$ is indeed a conjugate complete set, and so

$$\sum_{\substack{1\leq \ell \leq L \\ 1\leq m\leq M}} e^{\alpha_\ell + \beta_m}$$

is a conjugate complete exponential sum. □

We now have all the ingredients needed to complete the proof of the Lindemann-Weierstrass theorem, starting with the general version stated in the hypotheses and coaxing it into the particular form we assumed in the weak version we proved earlier. So let us do that now.

Theorem 1 (Lindemann-Weierstrass theorem)**.** Given $M + 1$ distinct algebraic numbers $\alpha_0, \alpha_1, \ldots \alpha_M$, the numbers $e^{\alpha_0}, e^{\alpha_1}, \ldots e^{\alpha_M}$ are linearly independent over $\overline{\mathbb{Q}}$. That is, for any $\beta_0, \ldots, \beta_M \in \overline{\mathbb{Q}}$ not all zero,

$$\sum_{m=0}^{M} \beta_m e^{\alpha_m} \neq 0.$$

Proof. Rather than attempt to prove that for any collection of $M + 1$ algebraic numbers $\beta_0, \ldots, \beta_M$ not all zero we have $\sum_{m=0}^{M} \beta_m e^{\alpha_m} \neq 0$ we will proceed by contradiction and suppose that there do exist β_i for which

$$\sum_{m=0}^{M} \beta_m e^{\alpha_m} = 0 \tag{†}$$

and from this assumption we shall strive to create a contradiction, specifically a contradiction to Theorem 2.

The first step to this end is to multiply expression (†) by analogous expressions with each α_m replaced by each of its conjugates in all possible combinations:

$$\prod_{\substack{\rho_m \text{ a conjugate} \\ \text{of } \alpha_m \text{ for} \\ 0 \leq m \leq M}} (\beta_0 e^{\rho_0} + \beta_1 e^{\rho_1} + \ldots + \beta_M e^{\rho_M}) = 0.$$

This is a truly large product, if each α_i has degree n_i then the above product is composed of $\prod_{i=1}^{M} n_i$ terms, each of them being a sum of M terms. Nevertheless, once we multiply it out we will get something of the form

$$(\ddagger) \qquad b_0 E_0 + b_1 E_1 + \ldots + b_L E_L = 0$$

where the b_ℓ are integer polynomials in the β_m and the E_ℓ are conjugate complete exponential sums. This may seem implausible, but looking at the product should reveal it to be true. Each term in the sum we get when we multiply out the product will be of the form

$$g(\beta_1, \beta_2, \ldots, \beta_M) e^{r_1 + r_2 + \ldots + r_M}$$

where $g \in \mathbb{Z}[t_1, \ldots, t_M]$ is some polynomial and the r_i's are algebraic numbers from the set of all the conjugates of $\alpha_1, \alpha_2, \ldots, \alpha_M$.

If we apply Lemma 2.4 to the polynomial $x_1 + x_2 + \ldots + x_M$ and the set $\{\alpha_1, \alpha_2, \ldots, \alpha_M\}$ then we see that the terms in the above expression with the same coefficients will form conjugate complete exponential sums. And so the product multiplies out to $b_0 E_0 + b_1 E_1 + \ldots + b_L E_L$ as we stated. We can also apply Lemma 2.2 to this product to note that not all the b_ℓ's can be zero, so we have not simply given a long proof of the identity $0 = 0$. Without losing any generality we may as well assume that all the b_ℓ's are nonzero, since if any are zero we can remove them and relabel what's left.

In Theorem 2 the coefficients we were dealing with were integers, whereas at the moment our b_ℓ's are polynomials in algebraic numbers over the integers. But with another enormous product we can rectify this situation. In an analogous move to our last one we now multiply (‡) by all similar expressions where each b_ℓ is replaced by its conjugates to form the product

$$\prod_{\substack{\sigma_\ell \text{ a conjugate} \\ \text{of } b_\ell \text{ for} \\ 0 \leq \ell \leq L}} (\sigma_0 E_0 + \sigma_1 E_1 + \ldots + \sigma_L E_L) = 0.$$

Once again this product, once multiplied out, will simplify greatly to something looking like

$$\zeta_0 \mathcal{E}_0 + \zeta_1 \mathcal{E}_1 + \ldots + \zeta_K \mathcal{E}_K = 0,$$

where each ζ_k is a symmetric polynomial in a conjugate complete set of the b_ℓ's, and hence is rational, and each $\mathcal{E}_k$ is a product of E_ℓ's and so by Lemma 2.5 is a conjugate complete exponential sum. Moreover, by Lemma 2.2, not all the ζ_k can be zero, so we are still not down to the trivial identity $0 = 0$. Again we can remove any of the terms in the sum which are zero to assume that $\zeta_k \neq 0$ for all the k's. And since each ζ_k is rational we can multiply the last identity by the lowest common multiple of all their denominators and so assume without loss of generality that each ζ_k is in fact an integer.

As a final piece of housekeeping we may assume that all the exponents in any of the $\mathcal{E}_k$ are conjugate to one another, since any conjugate complete exponential sum can be simply split into a sum of conjugate complete exponential sums with this property. So our identity

$$(\bigstar) \qquad \zeta_0\mathcal{E}_0 + \zeta_1\mathcal{E}_1 + \ldots + \zeta_K\mathcal{E}_K = 0$$

now has $\zeta_k \in \mathbb{Z} \setminus \{0\}$ for all k and each $\mathcal{E}_k$ is an exponential sum with each exponent being conjugate to all others in that sum. The hypotheses of our weakened Lindemann-Weierstrass theorem are now almost entirely met, except that in that version we required one of the terms in the sum was just an integer, so in our case we require $\mathcal{E}_k = 1$ for some k.

If $\mathcal{E}_k = 1$ for any of the k then we are done, so let us assume that $\mathcal{E}_k \neq 1$ for $k = 0, 1, \ldots, K$. In particular we have

$$\mathcal{E}_0 = e^{\eta_1} + e^{\eta_2} + \ldots + e^{\eta_J}$$

where $\{\eta_1, \eta_2, \ldots, \eta_J\}$ is a conjugate complete set of nonzero algebraic numbers. Now let

$$\mathcal{E}_0' = e^{-\eta_1} + e^{-\eta_2} + \ldots + e^{-\eta_J},$$

and we may note that $\mathcal{E}_0'$ is also a conjugate complete exponential sum. Thus, by Lemma 2.5, $\mathcal{E}_0\mathcal{E}_0'$ is also a conjugate complete exponential sum, and moreover this product contains at least J copies of $e^0 = 1$ from the products $e^{\eta_j}e^{-\eta_j}$. So we may write

$$\mathcal{E}_0\mathcal{E}_0' = J + \mathcal{E}_0'',$$

where $\mathcal{E}_0''$ is a conjugate complete exponential sum. Now recall the final bit of housekeeping we did on the sum $\zeta_0\mathcal{E}_0 + \ldots + \zeta_K\mathcal{E}_K = 0$ to ensure that within distinct $\mathcal{E}_k$'s there were distinct exponents. In particular if $k \neq 0$ then none of the exponents in $\mathcal{E}_k$ will come from the set $\{\eta_1, \eta_2, \ldots, \eta_J\}$, and so $\mathcal{E}_k\mathcal{E}_0' = \mathcal{E}_k''$ will contain no e^0 terms, but will be a conjugate complete exponential sum by Lemma 2.5. So, multiplying the identity ($\bigstar$) through by $\mathcal{E}_0'$ gives us

$$J\zeta_0 + \zeta_0\mathcal{E}_0'' + \zeta_1\mathcal{E}_1'' + \ldots + \zeta_K\mathcal{E}_K'' = 0$$

with the ζ_k being nonzero integers and each $\mathcal{E}_k''$ a conjugate complete exponential sum with no terms of the form e^0. So $J\zeta_0$ is the only free integer, and it is nonzero. Thus this sum meets the hypotheses of Theorem 2 but contradicts that theorem's conclusion. So we are forced to conclude that our assumption (†) was false, and thence the Lindemann-Weierstrass theorem is true. □

2.2. **Later Results.**

The Lindemann-Weierstrass theorem was the pinnacle of the study of transcendence in the nineteenth century, but it was not the end of the story. Fifteen years after Weierstrass' proof David Hilbert delivered perhaps the most famous lecture in mathematical history at the International Congress of Mathematicians in Paris. In the talk he presented ten open problems in mathematics that he considered of the utmost importance to the development of the subject in the twentieth century.

The list, which would eventually be comprised of twenty three problems, is perhaps most famous now because of the appearance of Problem Number 8, better known as the Riemann hypothesis. This problem and Problem 12 (on extending Kronecker's theorem on abelian extensions of the rational numbers to any base number field) are the only two problems that have remained unsolved to this day[4]. At the time, though, Hilbert was hopeful that a solution to the Riemann hypothesis would be found soon, he even discussed follow-up problems in his lecture. A much harder problem, in his opinion, was Problem Number 7, which he described as follows.

> "Hermite's arithmetical theorems on the exponential function and their extension by Lindemann are certain of the admiration of all generations of mathematicians. Thus the task at once presents itself to penetrate further along the path here entered... I consider the proof of the following theorem very difficult:
>
> *The expression α^β, for an algebraic base α and an irrational algebraic exponent β, e.g., the number $2^{\sqrt{2}}$ or $e^\pi = i^{-2i}$, always represents a transcendental or at least an irrational number.*"[5]

Despite Hilbert's misgivings about the problem it was fully solved within his lifetime. In the 1930s Aleksandr Gelfond and Theodor Schneider independently proved that α^β is always transcendental for algebraic $\alpha(\neq 0, 1)$ and algebraic irrational β. The two numbers highlighted by Hilbert, e^π and $2^{\sqrt{2}}$, became known as Gelfond's constant and the Gelfond-Schneider constant respectively.

An extension to the theorem proposed by Gelfond was proved by Alan Baker in the 1960s, who also proved the definitive result in this area and earned a Fields Medal for his work. He proved that for non-zero algebraic numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$

[4] Several of the twenty three problems - and, indeed, the omitted twenty fourth problem - have been deemed too vague to admit a solution one way or the other.

[5] English translation of the lecture is available at http://aleph0.clarku.edu/ djoyce/hilbert/problems.html

such that $\log \alpha_1, \log \alpha_2, \ldots, \log \alpha_n$ are linearly independent over the rational numbers, the numbers $1, \log \alpha_1, \log \alpha_2, \ldots, \log \alpha_n$ are linearly independent over the algebraic numbers. This has several immediate consequences just as the Lindemann-Weierstrass theorem did. Most notably, perhaps, is that given $2n + 1$ non-zero algebraic numbers $\alpha_1, \ldots \alpha_n, \beta_0, \ldots \beta_n$ the number $e^{\beta_0} \alpha_1^{\beta_1} \cdots \alpha_n^{\beta_n}$ is transcendental.

The Gelfond-Schneider and Lindemann-Weierstrass theorems are both corollaries of a much more general result, Schanuel's conjecture. This result concerns the degree of an extension field over the rational numbers. This conjecture is almost fifty years old now and it is generally accepted that with current techniques a proof is unattainable[6]. If a proof was found, though, it would not only turn the aforementioned theorems into mere corollaries it would also prove that π and e are algebraically independent over $\overline{\mathbb{Q}}$. One consequence of this is that $e + \pi$ would be transcendental, a result that has gone unproved for centuries.

3. An Irrationality Proof

3.1. $\zeta(3)$: Apéry's theorem.

It would be unfair to say that mathematical progress in the area of irrationality and transcendence proofs had ground to a halt in the latter half of the twentieth century. But results like the Gelfond-Schneider theorem and Roth's theorem had been the pinnacle of their respective lines of inquiry. Certainly no one had announced a proof that a particularly famous number was irrational for quite some time. So it came as something of a shock to the mathematical community in 1979 when French mathematician Roger Apéry announced a proof that $\zeta(3)$ was irrational.

Before then he had liked to joke that he was the worst mathematician in France, since no famous theorem bore his name. But this result assured that both the theorem and the number itself would bear the name Apéry.

Apéry's approach is more reminiscent of the inequalities that Liouville dealt with rather than the complicated ways of constructing an integer in the interval $(0, 1)$ that many newer proofs rely on. That is not to say Apéry's proof is not complicated, it is just perhaps slightly more direct. It hinges on the following useful result.

Lemma 3.1. If there are infinitely many coprime solutions p, q of

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

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

Proof. We will deal with the contrapositive, proving that for rational α there are finitely many coprime integers p and q satisfying the inequality. So suppose that $\alpha = \frac{a}{b}$ for coprime a, b with $b > 0$. Then for $\frac{p}{q} \neq \alpha$ that satisfy $(\diamondsuit)$ we can immediately get the following inequality.

$$\begin{aligned}\frac{c}{q^{1+\delta}} &> \left|\alpha - \frac{p}{q}\right| \\ &= \left|\frac{a}{b} - \frac{p}{q}\right| \\ &= \left|\frac{aq - bp}{bq}\right| \\ &\geq \frac{1}{bq}.\end{aligned}$$

From this it follows that $q < (cb)^{1/\delta}$, hence there are only finitely many possibilities for q. Of course, there could be infinitely many choices for p, so we must check that too. Suppose we have p and q which satisfy the inequality of our hypothesis. Now consider fractions of the form $\frac{p+t}{q}$ for some integer variable t. If this fraction satisfies the inequality of our hypothesis then we can use the triangle inequality to deduce that

$$\frac{|t|}{q} = \left|\frac{t}{q} + \frac{p}{q} - \frac{a}{b} - \frac{p}{q} + \frac{a}{b}\right| \leq \left|\frac{p+t}{q} - \frac{a}{b}\right| + \left|\frac{p}{q} - \frac{a}{b}\right| < \frac{2c}{q^{1+\delta}}.$$

Hence $|t| < 2cq^{-\delta} \leq 2c$, so at best t is bounded by $-2[c]+1 \leq t \leq 2[c]-1$. So for any of our finitely many values of q there are at most $4[c]-1$ values of p which will satisfy the hypothesis. Hence there are only finitely many coprime solutions p, q as required. □

With this irrationality criterion in mind we should be looking for infinite sequences of integers (p_n) and (q_n) such that $\frac{p_n}{q_n} \to \zeta(3)$ about as fast as $\frac{1}{q_n^{1+\delta}} \to 0$. Apéry outlined a way of doing just this in 1978[1] and not long afterwards Alfred van der Poorten filled in most of the gaps. The following series of lemmas follow the outline given in [14] and the entertaining outline provided by [11].

First we need to define three sequences which will form the backbone of all that follows. For integers $0 \leq k \leq n$, define

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

and

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

As well as these three sequences we will also be using the Landau-Vinogradov symbols, also known as big O notation. We will say $f(x) = O\left(g(x)\right)$ or equivalently $f(x) \ll g(x)$ for some $g(x) \geq 0$ if:

$$\limsup_{x\to\infty} \frac{|f(x)|}{g(x)} \text{ is bounded,}$$

or if there exists some constant $C \geq 0$ such that $|f(x)| \leq Cg(x)$ for sufficiently large x. In this language Lemma 3.1's criterion becomes

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

With this notation we are now ready for the first step in Apéry's proof.

Lemma 3.2.

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \zeta(3) = \sum_{n=1}^{\infty} \frac{1}{n^3}.$$

Proof.

$$
\begin{aligned}
a_n &= \sum_{k=0}^{n} c_{n,k} \binom{n}{k}^2 \binom{n+k}{k}^2 \\
&= \sum_{k=0}^{n} \left(\sum_{m=1}^{n} \frac{1}{m^3} + \sum_{m=1}^{k} \frac{(-1)^{m-1}}{2m^3 \binom{n}{m}\binom{n+m}{m}} \right) \binom{n}{k}^2 \binom{n+k}{k}^2 \\
&= \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 \sum_{m=1}^{n} \frac{1}{m^3} + \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 \sum_{m=1}^{k} \frac{(-1)^{m-1}}{2m^3 \binom{n}{m}\binom{n+m}{m}} \\
&= b_n \sum_{m=1}^{n} \frac{1}{m^3} + \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 \sum_{m=1}^{k} \frac{(-1)^{m-1}}{2m^3 \binom{n}{m}\binom{n+m}{m}}. \qquad (1)
\end{aligned}
$$

For $1 \leq m \leq n$ we have

$$\binom{n}{m}\binom{n+m}{m} \geq \binom{n}{1}\binom{n+1}{1} = n(n+1) \geq n^2.$$

So

$$\frac{1}{\binom{n}{m}\binom{n+m}{m}} \leq \frac{1}{n^2}.$$

Hence

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

We know that $\sum_{m=1}^{k} \frac{1}{m^3} \to \zeta(3) \approx 1.202$ as $k \to \infty$. So $\frac{1}{n^2} \sum_{m=1}^{k} \frac{1}{m^3} \leq 2\frac{1}{n^2}$, say. That is,

$$\frac{1}{n^2} \sum_{m=1}^{k} \frac{1}{m^3} \ll \frac{1}{n^2}.$$

Putting this together with (1) gives

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

If we now write

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

then, by the integral test,

$$\sum_{m=n+1}^{\infty} \frac{1}{m^3} \leq \int_{n}^{\infty} \frac{1}{u^3}\, du = \frac{-1}{2u^2}\Big|_{n}^{\infty} = \frac{1}{2n^2} \ll \frac{1}{n^2}.$$

Putting this into (2) gives us

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

Rearranging this slightly gives

$$\frac{a_n}{b_n} - \zeta(3) = O\left(\frac{1}{n^2}\right) + O\left(\frac{1}{n^2}\right),$$

thence

$$\frac{a_n}{b_n} - \zeta(3) \ll \frac{1}{n^2} \to 0.$$

And so $\dfrac{a_n}{b_n} \to \zeta(3)$. □

It might be tempting at this point to proclaim that we have found the necessary sequences to apply Lemma 3.1, but the convergence we have proved above is not nearly fast enough[6]. We have to do better than this to prove the irrationality of $\zeta(3)$, and to do so we will need the following lemma.

Lemma 3.3. Let $B(n) = 34n^3 + 51n^2 + 27n + 5$. Then (a_n) and (b_n) satisfy the recursion formula:

$$(n+1)^3 u_{n+1} - B(n) u_n + n^3 u_{n-1} = 0, \quad n \in \mathbb{N}.$$

Proof. The mysterious $B(n)$ is in fact the n^{th} partial quotient in the continued fraction representation of $\zeta(3)$. Cohen proved this in 1979, but for now we concentrate on proving this lemma, which requires us to define the following number for integers $0 \le k \le n$.

$$\begin{aligned} \lambda_{n,k} &= \binom{n}{k}^2 \binom{n+k}{k}^2 \\ &= \frac{n!^2}{k!^2 (n-k)!^2} \cdot \frac{(n+k)!^2}{k!^2 n!^2} \\ &= \frac{(n+k)!^2}{k!^4 (n-k)!^2}. \end{aligned}$$

[6] A somewhat more grave problem with using these sequences is brought up later for those who have not yet noticed it.

If this looks familiar then it's probably because this number appears in the definition of both a_n and b_n. We also need the following definition:

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

We now claim the following:

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

If we explicitly write out the left hand side and divide this through by $\lambda_{n,k}$, which is nonzero by definition, then we find that this claim is equivalent to the following one:

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

So let us calculate the three quotients of the λ's explicitly.

$$\begin{aligned} \frac{\lambda_{n,k-1}}{\lambda_{n,k}} &= \frac{(n+k-1)!^2}{(k-1)!^4(n-k+1)!^2} \Big/ \frac{(n+k)!^2}{k!^4(n-k)!^2} \\ &= \frac{k!^4(n-k)!^2(n+k-1)!^2}{(k-1)!^4(n+k)!^2(n-k+1)!^2} \\ &= \frac{k^4}{(n+k)^2(n-k+1)^2}, \end{aligned}$$

$$\begin{aligned} \frac{\lambda_{n+1,k}}{\lambda_{n,k}} &= \frac{(n+k+1)!^2}{(k)!^4(n-k+1)!^2} \Big/ \frac{(n+k)!^2}{k!^4(n-k)!^2} \\ &= \frac{k!^4(n-k)!^2(n+k+1)!^2}{(k)!^4(n-k+1)!^2(n+k1)!^2} \\ &= \frac{(n+k+1)^2}{(n-k+1)^2}, \end{aligned}$$

$$\begin{aligned} \frac{\lambda_{n-1,k}}{\lambda_{n,k}} &= \left(\frac{\lambda_{n,k}}{\lambda_{n-1,k}}\right)^{-1} \\ &= \frac{(n-k)^2}{(n+k)^2}. \end{aligned}$$

Substituting these into (3) shows that our claim is equivalent to the daunting identity

$$4(2n{+}1)(2k^2{+}k{-}(2n{+}1)^2){-}4(2n{+}1)\left(2(k-1)^2+k-1-(2n+1)^2\right)\frac{k^4}{(n+k)^2(n-k+1)^2}$$
$$= (n+1)^3\frac{(n+k+1)^2}{(n-k+1)^2} - B(n) + n^3\frac{(n-k)^2}{(n+k)^2},$$

which in turn is equivalent to

$$4(2n{+}1)(2k^2{+}k{-}(2n{+}1)^2)(n{+}k)^2(n{-}k{+}1)^2{-}4(2n{+}1)\left(2(k-1)^2+k-1-(2n+1)^2\right)k^4$$
$$= (n{+}1)^3(n{+}k{+}1)^2(n{+}k)^2 - B(n)(n{+}k)^2(n{-}k{+}1)^2 + n^3(n{-}k)^2(n{-}k{+}1)^2.$$

Multiplying out both sides of this expression, either by hand or with a computer package, and then comparing coefficients shows that this identity does indeed hold. Hence our claim was right and we have

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

Let us now put $k = 0, 1, \ldots, n+1$ into this identity and sum.

$$\sum_{k=0}^{n+1}(B_{n,k} - B_{n,k-1}) = \sum_{k=0}^{n+1}\left((n+1)^3\lambda_{n+1,k} - B(n)\lambda_{n,k} + n^3\lambda_{n-1,k}\right)$$

which means

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

And since $\binom{n}{k} = 0$ if $n < k$ or $k < 0$,

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

But $\sum_{k=0}^{n+1}\lambda_{n+1,k} = \sum_{k=0}^{n+1}\binom{n+1}{k}^2\binom{n+1+k}{k}^2 = b_{n+1}$, and similarly for the other two sums. Thus we have

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

which proves half of the lemma.

To prove that a_n also satisfies this recurrence we first recall that

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

Hence we can immediately see that

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

More troublesome is the following claim for $1 \leq k \leq n$.

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

We will prove this by induction on k. The case $k = 1$ is simple enough. The right hand side becomes:

$$\frac{-(n-2)!}{n^2(n+1)!} = \frac{-1}{n^2(n+1)n(n-1)} = \frac{1}{n^3(1-n^2)}.$$

The left hand side, meanwhile, is:

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

Which was what we wanted.

Now suppose the claim is true for some $k < n$. We then have the following.

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

$$= \frac{(-1)^k k!^2(n-k-1)!}{n^2(n+k)!} + \frac{(-1)^k(k+1)!^2(n-k-1)!}{2(k+1)^3(n+k+1)!}$$

$$- \frac{(-1)^k(k+1)!^2(n-k-2)!(n-1)!}{2(k+1)^3(n-1)!(n+k)!}$$

$$= \frac{(-1)^k k!^2(n-k-2)!}{(n+k)!}\left[\frac{n-k-1}{n^2} + \frac{n-k-1}{2(n+k+1)(k+1)} - \frac{1}{2(k+1)}\right]$$

$$= \frac{(-1)^k k!^2(n-k-2)!}{(n+k)!}\left[\frac{2(n-k-1)(n+k+1)(k+1) + (n-k-1)n^2 - n^2(n+k+1)}{2n^2(n+k+1)(k+1)}\right]$$

Multiplying out the numerator inside the brackets, and simplifying leads to

$$c_{n,k+1} - c_{n-1,k+1} = \frac{(-1)^k k!^2(n-k-2)!}{(n+k)!}\left[\frac{-2(k+1)^3}{2n^2(n+k+1)(k+1)}\right]$$

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

which is precisely our claim for $k+1$. Thus we have proved the claim.

Much as we did earlier, let us now define the following number.

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

If we note the following two facts

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

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

then, with a bit of algebra, we can rewrite $C_{n,k}$ in the following form,

$$C_{n,k} = \left[(n+1)^3\lambda_{n+1,k} - B(n)\lambda_{n,k} + n^3\lambda_{n-1,k}\right]c_{n,k}$$
$$+ (-1)^k k!^2\left(\frac{(n+1)^3\lambda_{n+1,k}(n-k)!}{(n+1)^2(n+1+k)!} - \frac{n^3\lambda_{n-1,k}(n-k-1)!}{n^2(n+k)!}\right).$$

Looking back a few pages we note that the expression in the square brackets is in fact $B_{n,k} - B_{n,k-1}$. So the above expression becomes

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

We now define another number which, having used $B_{n,k}$ and $C_{n,k}$, we will call $A_{n,k}$. It is given by

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

This number may not seem immediately useful, but it is important because with a little bit more algebraic legwork we arrive at the following fact.

$$A_{n,k} - A_{n,k-1} = (B_{n,k} - B_{n,k-1})\, c_{n,k} + B_{n,k-1} \frac{(-1)^{k-1}}{2k^3 \binom{n}{k} \binom{n+k}{k}} + \frac{5(-1)^{k-1}(2n+1)}{n(n+1)} \left[k\binom{n}{k}\binom{n+k}{k} + (k-1)\binom{n}{k-1}\binom{n+k-1}{k-1} \right]$$

While this may not immediately seem much to get excited about, if we let $\mathcal{S}_{n,k} = C_{n,k} - (A_{n,k} - A_{n,k-1})$ then, after a great deal of algebraic manipulation, we will find that $\mathcal{S}_{n,k} \equiv 0$. Or,

$$C_{n,k} = A_{n,k} - A_{n,k-1}.$$

We are now in familiar territory, and can repeat the dramatic conclusion of the first half of this proof. If we sum the above identity over $k = 0, 1, \ldots, n+1$ then we get:

$$\sum_{k=0}^{n+1} C_{n,k} = A_{n,n+1} - A_{n,-1} = 0.$$

So, inserting the definition of $C_{n,k}$,

$$\sum_{k=0}^{n+1} \left((n+1)^3 \lambda_{n+1,k} c_{n+1,k} - B(n) \lambda_{n,k} c_{n,k} + n^3 \lambda_{n-1,k} c_{n-1,k} \right) = 0,$$

which, recalling that $\lambda_{n,k} = 0$ if $k > n$, gives

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

But $\sum_{k=0}^{n} \lambda_{n,k} c_{n,k} = a_n$, and so

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

as required. □

To recap, then, we have defined two sequences, a_n and b_n, which we proved had the property that $\lim_{n\to\infty} \frac{a_n}{b_n} = \zeta(3)$. Now we have shown that they satisfy the rather bizarre recurrence relation laid out in Lemma 3.3. Why is this useful? Well it allows us to improve our knowledge on how fast the sequence $\frac{a_n}{b_n}$ converges. Before, we had the rather meagre fact that $\frac{a_n}{b_n} - \zeta(3) \ll \frac{1}{n^2}$. The next lemma will much improve on this.

Lemma 3.4.
$$\left|\zeta(3) - \frac{a_n}{b_n}\right| \ll \frac{1}{b_n^2}.$$

Proof. Remember that the Landau-Vinogradov symbol here means that we want to show $\limsup b_n^2 \left|\zeta(3) - \frac{a_n}{b_n}\right|$ is bounded.

We have just shown that

$$\begin{cases} n^3 a_n - B(n-1)a_{n-1} + (n-1)^3 a_{n-2} = 0 \\ n^3 b_n - B(n-1)b_{n-1} + (n-1)^3 b_{n-2} = 0 \end{cases}$$

for all $n \geq 2$.

Multiplying the first of these by b_{n-1}, the second by a_{n-1}, and then subtracting the second from the first gives us that

$$n^3(a_n b_{n-1} - b_n a_{n-1}) + (n-1)^3(a_{n-2}b_{n-1} - b_{n-2}a_{n-1}) = 0$$

or that

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

Letting $n \longmapsto n-1$ gives

$$(n-1)^3(a_{n-1}b_{n-2} - b_{n-1}a_{n-2}) = (n-2)^3(a_{n-2}b_{n-3} - a_{n-3}b_{n-2}).$$

Note that the left hand side of this last line is equal to the right hand side of the penultimate line, so we can say

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

Proceeding inductively we find that in fact

$$n^3(a_n b_{n-1} - b_n a_{n-1}) = (n-k)^3(a_{n-k}b_{n-k-1} - a_{n-k-1}b_{n-k})$$

for all $k < n$. If we then take the case $k = n-1$ we get the following identity,

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

or

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

But it is easy to check that $a_0 = 0, a_1 = 6, b_0 = 1$, and $b_1 = 5$. So we have:

$$a_n b_{n-1} - b_n a_{n-1} = \frac{6}{n^3} \qquad \forall n \in \mathbb{N}.$$

Let us now define $s_n = \zeta(3) - a_n/b_n$. For any natural number n we note that

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

We know from Lemma 3.2 that $s_n \to 0$ so we may write

$$\begin{aligned}
\zeta(3) - \frac{a_n}{b_n} &= s_n \\
&= s_n - s_{n+1} + s_{n+1} \\
&= s_n - s_{n+1} + s_{n+1} - s_{n+2} + s_{n+2} - s_{n+3} + \dots \\
&= \sum_{m=n+1}^{\infty} (s_{m-1} - s_m) \\
&= 6 \sum_{m=n+1}^{\infty} \frac{1}{m^3 b_{m-1} b_m}.
\end{aligned}$$

Clearly from its definition, (b_n) is a strictly increasing sequence of positive integers, so we may now say

$$\begin{aligned}
\left| \zeta(3) - \frac{a_n}{b_n} \right| &= 6 \sum_{m=n+1}^{\infty} \frac{1}{m^3 b_{m-1} b_m} \\
&\leq 6 \sum_{m=n+1}^{\infty} \frac{1}{m^3 b_n^2}
\end{aligned}$$

which gives us

$$b_n^2 \left| \zeta(3) - \frac{a_n}{b_n} \right| \leq 6 \sum_{m=n+1}^{\infty} \frac{1}{m^3} \leq 6\zeta(3) < 12.$$

Which was what we wanted.

□

We are now entering the final stages of Apéry's proof, so let us recap what we have done so far and what still needs to be done.

We are trying to use Lemma 3.1 to prove that $\zeta(3)$ is irrational. Hence we need to find sequences of integers (p_n) and (q_n) such that for some fixed $\delta > 0$ we have

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

So far we have found two sequences (a_n) and (b_n) such that

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

So you may think we are done. Taking $\delta = 1$ surely allows us to apply the lemma and conclude that $\zeta(3) \notin \mathbb{Q}$. But no! Lemma 3.1 required (p_n) and (q_n) to be sequences of integers. And while (b_n) is a sequence of integers - a fact we've used in the proofs above - the sequence (a_n) is not so lucky. For while $a_0 = 0$ and $a_1 = 6$, we only have to get to $n = 2$ to find that $a_2 = \frac{351}{4}$. And things don't get any better from there.

All is not lost, though, for if we could multiply a_n and b_n by the right factor then we could turn a_n into an integer while preserving the ratio $\frac{a_n}{b_n}$. Of course, if we multiply the two sequences by some factor dependent on n, say $\nu(n)$, then we will need to satisfy the tighter relation:

$$\left|\zeta(3) - \frac{\nu(n)a_n}{\nu(n)b_n}\right| \ll \frac{1}{\left(\nu(n)b_n\right)^{1+\delta}}.$$

This could pose a problem, but not if we can show that neither b_n nor this factor $\nu(n)$ grows *too* fast. So first let us place some bound on the growth of b_n.

Lemma 3.5. Let $\alpha = (1+\sqrt{2})^4$, then $\alpha^n \ll b_n \ll \alpha^n$.

Proof. We know from Lemma 3.3 that

$$n^3 b_n - B(n-1)b_{n-1} + (n-1)^3 b_{n-2} = 0.$$

If we write $B(n-1)$ out explicitly and divide through by $n^3 b_{n-1}$ then, after a little simplifying, we arrive at

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

Suppose that $\dfrac{b_{n+1}}{b_n} \to X$, then for large n the above identity tell us that

$$X - 34 + \frac{1}{X} = 0.$$

Or, more helpfully, that $X^2 - 34X + 1 = 0$. We know (b_n) is increasing so we take the positive root here and arrive at $X = 17 + 12\sqrt{2} = (1+\sqrt{2})^4 = \alpha$. So we have the estimate $\alpha^n \ll b_n \ll \alpha^n$. □

This kind of growth is called quasi-geometric, since it is approximately geometric, at least in the long term. With the previous two lemmas we have the following corollary.

Corollary 3.1.

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

This suggests our final line of attack. If we can find a factor $\nu(n)$ such that $\nu(n)a_n$ is an integer and $\nu(n)b_n \ll \alpha^{2n/(1+\delta)}$ then we will have

$$\left|\zeta(3) - \frac{\nu(n)a_n}{\nu(n)b_n}\right| \ll \frac{1}{\alpha^{2n}} \ll \frac{1}{\left(\nu(n)b_n\right)^{1+\delta}}.$$

Looking back at the definitions of a_n and $c_{n,k}$ we note that the problem seems to be the presence of $2m^3$ terms in the denominator. So a good value for our $\nu(n)$ might be $2\,\mathrm{lcm}\,[1,2,\ldots,n]^3$. With that guess in mind we now take the following useful estimate.

Lemma 3.6. For any $\varepsilon > 0$ there exists a natural number N such that

$$\mathrm{lcm}\,[1,2,\ldots,n] \leq \exp\left((1+\varepsilon)n\right)$$

for all $n > N$.

Proof. First note that if $p \mid \mathrm{lcm}\,[1,2,\ldots,n]$ for some prime p then $p \leq n$, but if $p \leq n$ then $p \mid \mathrm{lcm}\,[1,2,\ldots,n]$, in other words $p \mid \mathrm{lcm}\,[1,2,\ldots,n]$ if and only if $p \leq n$. So we can say that

$$\mathrm{lcm}\,[1,2,\ldots,n] = \prod_{p\leq n} p^{e_p}$$

for suitable e_p.

By their definition, then, we have that $p^{e_p} \leq n < p^{e_p+1}$, which is the same as saying:

$$e_p \leq \frac{\log n}{\log p} < e_p + 1.$$

But $e_p \in \mathbb{Z}$, so this is equivalent to saying

$$e_p = \left[\frac{\log n}{\log p}\right].$$

This proves that $\mathrm{ord}_p\left(\mathrm{lcm}\,[1,2,\ldots,n]\right) = \left[\frac{\log n}{\log p}\right]$, a fact that will be used later. For now we note that we can place this value into the formula for the lcm that we found a few lines above to find:

$$\begin{aligned}\operatorname{lcm}[1,2,\ldots,n] &= \prod_{p\leq n} p^{\left[\frac{\log n}{\log p}\right]} \\ &\leq \prod_{p\leq n} p^{\frac{\log n}{\log p}} \\ &= \prod_{p\leq n} n \\ &= n^{\pi(n)},\end{aligned}$$

where $\pi(n)$ is the prime counting function.

Now, in what is perhaps an unexpected move when we set out to prove that $\zeta(3)$ was irrational, we appeal to the prime number theorem. It says that for any $\varepsilon > 0$ there is a natural number N such that for all $n > N$ we have

$$\pi(n) \leq \frac{n}{\log n} + \varepsilon.$$

So for all $n > N$ we have

$$\begin{aligned}\log\left(\operatorname{lcm}[1,2,\ldots,n]\right) &= \pi(n)\log n \\ &\leq \frac{n}{\log n}\log n + \varepsilon\log n \\ &\leq n + \varepsilon n \\ &= (1+\varepsilon)n.\end{aligned}$$

Or, simply,

$$\operatorname{lcm}[1,2,\ldots,n] \leq \exp\left((1+\varepsilon)n\right).$$

□

This bound will prove very useful when we have to show that $\nu(n)b_n \ll \alpha^{2n/(1+\delta)}$ for some δ, but the previous proof also holds a detail that will be pivotal in checking that if we take $\nu(n) = 2\operatorname{lcm}[1,2,\ldots,n]^3$ then $\nu(n)a_n$ is an integer as required. Checking that detail is the crux of the next lemma.

Lemma 3.7. For integers $0 \leq k \leq n$,

$$2\operatorname{lcm}[1,2,\ldots,n]^3\binom{n+k}{k}c_{n,k} \in \mathbb{Z}.$$

Proof. Recall that

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

So if we can show that both

$$2\,\mathrm{lcm}\,[1,2,\ldots,n]^3\binom{n+k}{k}\frac{1}{m^3}$$

and

$$\frac{2\,\mathrm{lcm}\,[1,2,\ldots,n]^3\binom{n+k}{k}}{2m^3\binom{n}{m}\binom{n+m}{m}}$$

are integers for $1 \leq m \leq n$ then their sum over this range will also be an integer, as required.

The first of these numbers is clearly an integer, since if $m \leq n$ then $m^3 \mid \mathrm{lcm}\,[1,2,\ldots,n]^3$. The second value is somewhat less trivially an integer. First we make the following claim:

Claim: $\mathrm{ord}_p\left(\binom{n}{m}\right) \leq \mathrm{ord}_p\left(\mathrm{lcm}\,[1,2,\ldots,n]\right) - \mathrm{ord}_p(m)$.

Proof of claim: De Polignac's formula tells us that

$$\mathrm{ord}_p\,(n!) = \sum_{k=1}^{\infty}\left[\frac{n}{p^k}\right].$$

So,

$$\begin{aligned}\mathrm{ord}_p\left(\binom{n}{m}\right) &= \mathrm{ord}_p\left(\frac{n!}{m!(n-m)!}\right)\\ &= \mathrm{ord}_p(n!) - \mathrm{ord}_p(m!) - \mathrm{ord}_p\left((n-m)!\right)\\ &= \sum_{k=1}^{\infty}\left(\left[\frac{n}{p^k}\right] - \left[\frac{m}{p^k}\right] - \left[\frac{n-m}{p^k}\right]\right)\\ &= \sum_{k=1}^{N}\left(\left[\frac{n}{p^k}\right] - \left[\frac{m}{p^k}\right] - \left[\frac{n-m}{p^k}\right]\right)\end{aligned}$$

where $N = \left[\frac{\log n}{\log p}\right] = \mathrm{ord}_p\left(\mathrm{lcm}\,[1,2,\ldots,n]\right)$, since if $k > \frac{\log n}{\log p}$ then $p^k > n$ and so $\left[\frac{n}{p^k}\right] = 0$. We now consider three cases.

Case 1 $p^k \mid m$. In this case we have $m = ap^k$ for some integer a. Suppose $n = qp^k + r$ with $0 \leq r < p^k$.

Then $\left[\frac{n}{p^k}\right] = q$, $\left[\frac{m}{p^k}\right] = a$ and $\left[\frac{n-m}{p^k}\right] = \left[\frac{(q-a)p^k + r}{p^k}\right] = q - a$.

So $\left[\frac{n}{p^k}\right] - \left[\frac{m}{p^k}\right] - \left[\frac{n-m}{p^k}\right] = q - a - q + a = 0.$

Case 2 $p^k \mid n-m$. So we have $n-m=bp^k$ for some integer b. Suppose again that $n=qp^k+r$ with $0 \leq r < p^k$.
Then

$$\left[\frac{n}{p^k}\right]=q, \quad \left[\frac{n-m}{p^k}\right]=b,$$

and

$$\left[\frac{m}{p^k}\right]=\left[\frac{n-(n-m)}{p^k}\right]=\left[\frac{(q-b)p^k+r}{p^k}\right]=q-b.$$

So $\left[\frac{n}{p^k}\right]-\left[\frac{m}{p^k}\right]-\left[\frac{n-m}{p^k}\right]=q-q+b-b=0.$

Case 3 $p^k \nmid m$, $p^k \nmid n-m$. Write $\begin{cases} m=\mu p^k+r_1, & 0<r_1<p^k \\ n-m=\nu p^k+r_2, & 0<r_2<p^k \end{cases}$, so

$$n=(n-m)+m=(\nu+\mu)p^k+r_1+r_2, \qquad 0<r_1+r_2<2p^k.$$

And $\left[\frac{m}{p^k}\right]=\mu, \quad \left[\frac{n-m}{p^k}\right]=\nu$, and

$$\left[\frac{n}{p^k}\right]=\begin{cases} \nu+\mu & \text{if } \ 0<r_1+r_2<p^k \\ \nu+\mu+1 & \text{if } \ p^k \leq r_1+r_2<2p^k \end{cases}.$$

So $\left[\frac{n}{p^k}\right]-\left[\frac{m}{p^k}\right]-\left[\frac{n-m}{p^k}\right]=\begin{cases} \nu+\mu-\mu-\nu=0 & \text{if } \ 0<r_1+r_2<p^k \\ \nu+\mu+1-\mu-\nu=1 & \text{if } \ p^k \leq r_1+r_2<2p^k \end{cases}$

From these three cases we see that $\operatorname{ord}_p\left(\binom{n}{m}\right)$ is a sum of $\operatorname{ord}_p(\operatorname{lcm}[1,\ldots,n])$ terms, each of them either 0 or 1. Moreover, at least $\operatorname{ord}_p(m)$ are zero corresponding to the $\operatorname{ord}_p(m)$ cases where $p^k \mid m$. Hence

$$\operatorname{ord}_p\left(\binom{n}{m}\right) \leq \operatorname{ord}_p(\operatorname{lcm}[1,2,\ldots,n])-\operatorname{ord}_p(m).$$

We can now prove that $\dfrac{\operatorname{lcm}[1,2,\ldots,n]^3\binom{n+k}{k}}{m^3\binom{n}{m}\binom{n+m}{m}}$ is an integer by showing that any given prime divides the numerator more times than it divides the denominator. If we write the denominator as

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

then we note that

$$\begin{aligned}\binom{n+m}{m}\binom{n+k}{k}^{-1} &= \frac{(n+m)!}{n!m!}\bigg/\frac{(n+k)!}{n!k!}\\ &= \frac{k!}{m!}\bigg/\frac{(n+k)!}{(n+m)!}\\ &= \frac{k!}{m!(k-m)!}\bigg/\frac{(n+k)!}{(n+m)!(k-m)!}\\ &= \binom{k}{m}\binom{n+k}{k-m}^{-1}.\end{aligned}$$

Now, looking at the denominator of our number and using the Claim from above as well as results from the previous lemma, we get

$$\begin{aligned}&\operatorname{ord}_p\left(m^3\binom{n}{m}\binom{n+m}{m}\binom{n+k}{k}^{-1}\right)\\ &= \operatorname{ord}_p\left(m^3\binom{n}{m}\binom{k}{m}\binom{n+k}{k-m}^{-1}\right)\\ &= 3\operatorname{ord}_p(m) + \operatorname{ord}_p\left(\binom{n}{m}\right) + \operatorname{ord}_p\left(\binom{k}{m}\right) - \operatorname{ord}_p\left(\binom{n+k}{k-m}\right)\\ &\le 3\operatorname{ord}_p(m) + \operatorname{ord}_p(\operatorname{lcm}[1,\dots,n]) - \operatorname{ord}_p(m) + \operatorname{ord}_p(\operatorname{lcm}[1,\dots,k]) - \operatorname{ord}_p(m)\\ &= \operatorname{ord}_p(m) + \left[\frac{\log n}{\log p}\right] + \left[\frac{\log k}{\log p}\right]\\ &\le 3\left[\frac{\log n}{\log p}\right]\end{aligned}$$

since $m \le k \le n$.

If we look at the numerator of our number then we simply get

$$\operatorname{ord}_p\left(\operatorname{lcm}[1,\dots,n]^3\right) = 3\left[\frac{\log n}{\log p}\right].$$

So any given prime divides the numerator of our number at least as often as it divides the denominator, or in other words our number is an integer. Thus

$$2\operatorname{lcm}[1,2,\dots,n]^3\binom{n+k}{k}c_{n,k}$$

is a sum of integers, hence is an integer itself.

□

We now have everything we need to prove the main result of this section.

Theorem 3 (Apéry's theorem). $\zeta(3)$ is irrational.

Proof. By Lemma 3.1 it suffices to show that there exist sequences of integers (p_n) and (q_n) such that, for some fixed $\delta > 0$, we have

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

for infinitely many n. So let us define

$$\begin{cases} p_n = 2\operatorname{lcm}[1,2,\ldots,n]^3\, a_n \\ q_n = 2\operatorname{lcm}[1,2,\ldots,n]^3\, b_n. \end{cases}$$

Since $b_n \in \mathbb{Z}$ for all n we know that $q_n \in \mathbb{Z}$. And by Lemma 3.7 we can also say that $p_n \in \mathbb{Z}$. Clearly $\frac{p_n}{q_n} = \frac{a_n}{b_n}$, so by Lemma 3.2 we know $\frac{p_n}{q_n} \to \zeta(3)$.

By Lemmas 3.5 and 3.6 we know that for any $\varepsilon > 0$ there exists a natural number N such that if $n > N$ then

$$\begin{aligned} q_n &= 2\operatorname{lcm}[1,2,\ldots,n]^3\, b_n \\ &\ll \alpha^n \exp\left((3+\varepsilon)n\right) \\ &= \alpha^n \alpha^{(3+\varepsilon)n/\log\alpha} \\ &= \alpha^{n\left(1+\frac{3+\varepsilon}{\log\alpha}\right)} \\ &= \alpha^{2n\left(\frac{\log\alpha+3+\varepsilon}{2\log\alpha}\right)}. \end{aligned}$$

Remember we wanted something of the form $q_n \ll \alpha^{2n/(1+\delta)}$, so we want

$$\frac{2\log\alpha}{\log\alpha + 3 + \varepsilon} = 1 + \delta$$

for some $\delta > 0$. Setting $\varepsilon = \frac{1}{3}$ shows then that $q_n \ll \alpha^{2n/(1+\delta)}$ where

$$\delta = \frac{3\log\alpha - 10}{3\log\alpha + 10} \approx 0.028 > 0$$

which will suffice. So we have that

$$\left|\zeta(3) - \frac{p_n}{q_n}\right| \ll \frac{1}{\alpha^{2n}} \ll \frac{1}{q_n^{1+\delta}}$$

for a positive constant δ, and so by Lemma 3.1, $\zeta(3)$ is irrational.

□

3.2. The source of Apéry's proof.

It was 2.00pm on a thursday afternoon in June 1978 when Roger Apéry gave a talk "Sur l'irrationalité de $\zeta(3)$." In this talk he claimed to have proofs that both $\zeta(2)$ and $\zeta(3)$ were irrational.[7]

The talk itself was exceedingly sketchy, Apéry made a series of increasingly outlandish claims that, if all true, would indeed confirm that $\zeta(3) \notin \mathbb{Q}$. However, Apéry's rather blasé approach to the proof and a general doubt about the result being proved meant that the talk did not spark huge interest at the time.

Scepticism was rife, but some mathematicians in the audience believed Apéry might have found a valid proof and set out to verify his claims. Three of these mathematicians, Henri Cohen, Hendrik Lestra, and Alfred van der Poorten spent an evening discussing the ideas surrounding Apéry's proof and checking some of his more outlandish numerical claims on their pocket calculators. They came away convinced that he was right, but they were unable to prove one crucial step, specifically Lemma 3.3 from above concerning the recursion relationship that the sequences (a_n) and (b_n) satisfy. At a conference in Helsinki in July 1978 Alfred van der Poorten pointed out to Apéry his groups' inability to prove this part, Apéry is reported to have considered this as more a compliment than a criticism.

After nearly two months of fruitless labour on the lemma, Cohen and van der Poorten showed the problem to German-American mathematician Don Bernhard Zagier, who quickly solved the corresponding recurrence quandary for the $\zeta(2)$ case. Using Zagier's idea, Cohen soon managed to prove the desired result for the sequences in the $\zeta(3)$ case. On August 18, just two months after Apéry's talk, Cohen delivered a lecture which finally explained how to complete the steps of Apéry's proof, and thus $\zeta(3)$ became Apéry's constant.

As van der Poorten himself put it, his and Cohen's work on the proof "constitutes a mystification rather than an explanation." However, following the talk Apéry himself took the podium to deliver a brief monologue on the state of the French language and then discuss the source of his ideas briefly. The most important identity Apéry used, hidden in the actual proof, is the following:

$$\zeta(3) = \frac{5}{2}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n^3\binom{2n}{n}}.$$

[7] Of course, it is well known that for any even integer n, $\zeta(n)$ is of the form $\frac{p}{q}\pi^n$, and thus is irrational. Apéry's proof that $\zeta(2) = \frac{\pi^2}{6}$ is irrational was interesting not only because it didn't use any properties of π, but also because it gave a better estimate on the irrationality measure[8] of $\zeta(2)$ and hence of π^2. In fact Apéry's proof improved the bound on the irrationality measure of π^2 to

$$\mu(\pi^2) \leq 11.85078\ldots.$$

[8] The irrationality measure of a number x, as the name suggests, is a measure of how irrational the number is, or rather of how well x can be approximated by rational numbers. It is also called the Liouville-Roth constant or irrationality exponent of x, denoted $\mu(x)$, and is the supremum of the real numbers μ such that

$$0 < \left|x - \frac{p}{q}\right| < \frac{1}{q^\mu}$$

is satisfied by infinitely many coprime pairs of integers p and q.

To prove this result we first note that for any sequence of integers $a_1, a_2, \ldots$ we have

$$\sum_{k=1}^{K} \frac{a_1 a_2 \cdots a_{k-1}}{(x+a_1)(x+a_2)\cdots(x+a_k)} = \frac{1}{x} - \frac{a_1 a_2 \cdots a_K}{x(x+a_1)\cdots(x+a_K)}.$$

If we let $x = n^2$, $a_k = -k^2$, and take $K = n-1$ we arrive at

$$\begin{aligned}\sum_{k=1}^{n-1} \frac{(-1)^{k-1}(k-1)!^2}{(n^2-1^2)\cdots(n^2-k^2)} &= \frac{1}{n^2} - \frac{(-1)^{n-1}(n-1)!^2}{n^2(n^2-1^2)\cdots(n^2-(n-1)^2)} \\ &= \frac{1}{n^2} - \frac{2(-1)^{n-1}}{n^2\binom{2n}{n}}.\end{aligned}$$

The last line follows from:

$$\begin{aligned}&\frac{(n-1)!^2}{(n^2-1^2)(n^2-2^2)\cdots(n^2-(n-1)^2)} \\ =&\frac{(n-1)!^2}{(n+1)(n-1)(n+2)(n-2)\cdots(n+n-1)(n-n+1)} \\ =&\frac{(n-1)!^2}{(2n-1)(2n-2)\cdots(n+1)(n-1)\cdots 2\cdot 1} \\ =&\frac{2n^2(n-1)!^2}{2n(2n-1)(2n-2)\cdots(n+1)n(n-1)\cdots 2\cdot 1} \\ =&2\frac{n!^2}{(2n)!} \\ =&2\binom{2n}{n}^{-1}.\end{aligned}$$

Now if we define

$$\epsilon_{n,k} = \frac{k!^2(n-k)!}{2k^3(n+k)!}$$

then we can note that

$$(-1)^k n\left(\epsilon_{n,k} - \epsilon_{n-1,k}\right) = \frac{(-1)^{k-1}(k-1)!^2}{(n^2-1^2)\cdots(n^2-k^2)}.$$

And so

$$\sum_{n=1}^{N}\sum_{k=1}^{n-1}(-1)^k\left(\epsilon_{n,k}-\epsilon_{n-1,k}\right)=\sum_{n=1}^{N}\frac{1}{n^3}-2\sum_{n=1}^{N}\frac{(-1)^{n-1}}{n^3\binom{2n}{n}} \tag{4}$$

$$=\sum_{k=1}^{N}(-1)^k\left(\epsilon_{N,k}-\epsilon_{k,k}\right)$$

$$=\sum_{k=1}^{N}\frac{(-1)^k}{2k^3\binom{N+k}{k}\binom{N}{k}}+\frac{1}{2}\sum_{k=1}^{N}\frac{(-1)^{k-1}}{k^3\binom{2k}{k}}. \tag{5}$$

If we now let $N \to \infty$ then we can see that the left hand term of (5) will tend to zero, and so putting the right hand side of (4) and (5) together gives

$$\zeta(3)=\frac{5}{2}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n^3\binom{2n}{n}}.$$

Looking at the above working we can now see that the sequence we used in the proof, $c_{n,k}$, differs from $\zeta(3)$ by the same amount as the above series differs. This is the starting point for Apéry's approach. The terms $c_{n,k}$ themselves don't converge on $\zeta(3)$ fast enough to assure its irrationality, so he "accelerated" the convergence with the sequences a_n and b_n which, roughly speaking, have the ratio $c_{n,k}$. And as we saw in this section, they do converge to $\zeta(3)$ quickly enough to prove that it is not a rational number.

The fact that Apéry's approach simplifies to work on $\zeta(2)$ is a byproduct of the pleasant identity

$$\zeta(2)=3\sum_{n=1}^{\infty}\frac{1}{n^2\binom{2n}{n}}.$$

However, it is now widely supposed that a similar proof will not work for $\zeta(5)$ nor any higher value of the zeta function evaluated at an odd integer. In the years since Apéry's proof much work has been done attempting to find a constant ξ_5 such that

$$\zeta(5)=\xi_5\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n^5\binom{2n}{n}}.$$

Alas, with extensive computer searching[2] it has been shown that if ξ_5 does exist and is algebraic with degree less than 25 then the coefficients of its minimum polynomial must have Euclidean norm exceeding 10^{383}, so proving the irrationality of $\zeta(5)$ and the higher zeta constants will almost certainly rely on different methods.

3.3. More proofs.

Alfred van der Poorten's paper on Apéry's proof was titled "A Proof that Euler Missed", and indeed the mathematics involved in the proof would certainly not have been beyond Euler[9], though the deeper reasoning which led up to the proof may have been.

While the proof may have eluded Euler, and indeed all other mathematicians until the 1970s, once Apéry had published his proof, others soon followed. Frits

[9] In fact it was Nick Katz who, after Cohen's talk in Helsinki, enthusiastically proclaimed "This is marvellous! It is something Euler could have done!"

Beukers was the first to find an alternative approach[4], with a very tidy proof involving integrals over the Legendre polynomials,

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

While it looks different to Apéry's proof, it is in fact much the same, simply finding integral expressions analogous to the sums in the original proof.

Much as Apéry's proof readily simplifies to prove the irrationality of $\zeta(2)$, so too does Beukers' proof. And the same problems that prevent one from extending Apéry's approach to cover $\zeta(5)$ or any higher zeta value also afflicts Beukers' method. In fact these hold true for all known proofs that $\zeta(3)$ is irrational.

The idea behind Beukers' proof is not to find an infinite sequence that satisfies the irrationality criterion of Lemma 3.1, but rather to use the assumption that $\zeta(3)$ is rational to construct a sequence that is bounded below by some strictly positive number, but which tends to zero, much as we did when proving the Lindemann-Weierstrass theorem. Beukers managed to prove the following identity:

$$\int_0^1\int_0^1 \frac{-\log(xy)}{1-xy}P_n(x)P_n(y)dxdy = \frac{A_n + B_n\zeta(3)}{\operatorname{lcm}[1,\ldots,n]^3}$$

for some integers A_n and B_n.[10]

Using some cunning partial integration, and assuming that $\zeta(3)$ is rational and equal to $\frac{a}{b}$, Beukers eventually arrived at the inequality

$$0 < \frac{1}{b} \leq |A_n + B_n\zeta(3)| \leq 4\left(\frac{4}{5}\right)^n,$$

which is clearly a contradiction since the right hand side tends to zero, so will eventually become less than $\frac{1}{b}$.

Beukers' proof is generally considered the neatest proof of Apéry's result, if not the most elementary. Certainly if one can manage the tricky integrations then the result follows far more quickly this way. A third proof of the result, due to Wadim Zudilin in 2002[17] is far more reminiscent of Apéry's approach, using many series and some creative telescoping of sums to finally construct an integer $\mathcal{I}$ in the range $0 < \mathcal{I} < 1$. His method is similar to an earlier proof of the result by Yuri Nesterenko[9], its starting point is the rational function

$$R_n(t) = \left(\frac{(t-1)(t-2)\cdots(t-n)}{t(t+1)(t+2)\cdots(t+n)}\right)^2.$$

From this function he formed a number F_n from the hypergeometric series given by

$$F_n = -\sum_{t=1}^{\infty} R_n'(t)$$

and showed that $F_n = u_n\zeta(3) - v_n$ where u_n and $\operatorname{lcm}[1,\ldots,n]^3\, v_n$ were integers.

[10] This is in fact closely related to Hadjicostas's formula

$$\int_0^1\int_0^1 \frac{1-x}{1-xy}\left[-\log(xy)\right]^s dxdy = \Gamma(s+2)\left[\zeta(s+2) - \frac{1}{s+1}\right]$$

though this formula was not conjectured nor proved until 2004[7].

Next he defined the rational function

$$\mathfrak{R}_n(t) = n!^2(2t+n)\frac{(t-1)\cdots(t-n)\cdot(t+n+1)\cdots(t+2n)}{(t(t+1)\cdots(t+n))^4}$$

and the corresponding hypergeometric series

$$\mathfrak{F}_n = \sum_{t=1}^{\infty}\mathfrak{R}_n(t).$$

With some clever use of recursion relationships Zudilin proved that $F_n = \mathfrak{F}_n$ for all n, and then used bounds found with $\mathfrak{F}_n$ to deduce that if $\zeta(3) = \frac{a}{b}$ then $b\,\mathrm{lcm}\,[1,\ldots,n]^3\,F_n$ is an integer, since it equals $\mathrm{lcm}\,[1,\ldots,n]^3\,u_n a - \mathrm{lcm}\,[1,\ldots,n]^3\,v_n b$, and it satisfies the inequality

$$0 < b\,\mathrm{lcm}\,[1,\ldots,n]^3\,F_n < 20b(n+1)^4 81^n(\sqrt{2}-1)^{4n}.$$

But $81(\sqrt{2}-1)^4 \approx 0.795 < 1$ so the right hand side tends to zero, contradicting the left hand side being an integer. And so $\zeta(3)$ must be irrational.

Most mathematicians are of the opinion that $\zeta(n)$ will be at least irrational, and most probably transcendental for all positive integers $n > 1$. It has been known for over a hundred years that $\zeta(2n)$ is transcendental, but it was only after Apéry's proof that any kind of progress was made on the arithmetic nature of the odd zeta constants.

Some authorities believe that a general result in this field is imminent, or at least a proof that some other zeta constant is irrational[11], and such optimism is not misplaced. In recent years using wholly different techniques to Apéry some new results have come to light.

In 2000 Tanguy Rivoal published a brief but extraordinary paper which proved that infinitely many of the numbers $\zeta(2n+1)$ are irrational[12]. The proof uses linear forms in values of the zeta function and estimates upon them to bound the dimension of a vector space spanned by values of the zeta function at odd integers. A direct consequence of this bound is the aforementioned result.

A year later, and with only a little more work, Rivoal[12] improved upon his previous bounds to show that in fact one of the nine numbers $\zeta(5), \zeta(7), \zeta(9), \ldots, \zeta(21)$ has to be irrational[13]. Using the same method Wadim Zudilin managed to remove $\zeta(21)$ from this list, and then reduce the statement even further to prove that one of the numbers $\zeta(5), \zeta(7), \zeta(9)$, or $\zeta(11)$ must be irrational[15][16].

Work on the problem did not stop there, of course, but hopes that Zudilin's approach would keep reducing the list until only one number remained did not come true, and five years on this is as far as anyone has got on proving the irrationality of the odd zeta values.

A seemingly related but frustratingly trickier problem than proving the irrationality of the odd zeta constants is doing the same for the Euler-Mascheroni

[11] [14] in particular is hopeful of such a result.

[12] And, independently, Zudilin.

constant γ, defined by

$$\gamma = \lim_{N\to\infty}\left(\sum_{n=1}^{N}\frac{1}{n} - \log N\right).$$

Euler first mentioned the number as being worthy of note in the 1730s. Two centuries later David Hilbert - who was not averse to setting challenging problems to his contemporaries - mentioned the irrationality of γ as an "unapproachable" problem, in the face of which mathematicians stood helpless. Even the eminent number theorist G. H. Hardy is said to have offered his Savilian Chair at Oxford to anyone who could prove the number was irrational. And more recently several mathematicians have been willing to make bets that the number is in fact transcendental, Conway and Guy among them.

Despite all this, very little progress has been made on proving the result, and those same mathematicians who would make bets on the number being irrational would also wager that a proof will not surface any time soon. Several technical criteria have been found that would imply the irrationality of γ, but perhaps the most solid result known so far on this subject is that if γ *is* rational then its denominator exceeds 10^{242080}.

The Euler-Mascheroni constant is not the only open problem in this area of mathematics. As was mentioned at the start of this essay, it is one of number theory's great paradoxes that while almost all real numbers are transcendental proving that any particularly number does not satisfy any polynomial equation is a most challenging problem. The Lindemann-Weierstrass and Gelfond-Schneider theorems may have dealt with large classes of numbers, but even they are powerless in the face of such deceptively simple numbers as $e\pi$ and $e+\pi$. Finding a way to deal with these, and similar, numbers almost certainly would, in the words of David Hilbert, "lead us to entirely new methods and to a new insight into the nature of special irrational and transcendental numbers."

References

[1] R. Apéry, *Irrationalité de $\zeta(2)$ et $\zeta(3)$*, Astérisque **61** (1979), pp. 11-13.

[2] D. H. Bailey, J. Borwein, N. Calkin, R. Girgensohn, R. Luke, and V. Moll, *Experimental Mathematics in Action*, 2007.

[3] A. Baker, *Transcendental Number Theory*, 1975.

[4] F. Beukers, *A note on the irrationality of $\zeta(2)$ and $\zeta(3)$*, Bull. London Math. Soc. **11** (1979), pp. 268-272.

[5] Edward B. Burger and Robert Tubbs, *Making Transcendence Transparent*, 2004.

[6] T. Y. Chow, *What is a Closed-Form Number*, Amer. Math. Monthly **106** (1999), pp. 440-448.

[7] P. Hadjicostas, *A Conjecture-Generalization of Sondow's Formula*, (2004), http://www.arxiv.org/abs/math.NT/0405423/

[8] J. Liouville, J. Math. Pures Appl. (1) **16** (1851), pp. 133-142.

[9] Y. V. Nesterenko, *A Few Remarks on $\zeta(3)$*, Mat. Zametki **59** (1996), pp. 865-880. English translation in Math. Notes **59** (1996), pp. 625-636.

[10] I. Niven, *Irrational Numbers*, 1956.

[11] A. van der Poorten, *A proof that Euler missed*, Math. Intelligencer **1** (1979), pp. 195-203.

[12] T. Rivoal, *La fonction zeta de Riemann prend une infinité de valuers irrationnelles aux entiers impairs*, Comptes Rendus Acad. Sci. Paris Sér. I Math. **331** (2000), pp. 267-270.

[13] T. Rivoal, *Irrationalité d'au moins un des neuf nombres $\zeta(5), \zeta(7), \ldots, \zeta(21)$*, (2001), http://arxiv.org/abs/math.NT/0104221/

[14] Jörn Steuding, *Diophantine Analysis*, 2005.

[15] W. Zudilin, *One of the eight numbers $\zeta(5), \zeta(7), \ldots, \zeta(17), \zeta(19)$ is irrational*, Mat. Zametki **70**:3 (2001), pp. 472-476.

[16] W. Zudilin, *One of the numbers $\zeta(5), \zeta(7), \zeta(9), \zeta(11)$ is irrational*, Uspekhi Mat. Nauk **56**:4 (2001), pp. 149-150.

[17] W. Zudilin, *An Elementary Proof of Apéry's Theorem*, (2002), http://arxiv.org/abs/math/0202159