

# TRANSCENDENTAL NUMBERS

JEREMY BOOHER

## 1. INTRODUCTION

The Greeks tried unsuccessfully to square the circle with a compass and straightedge. In the 19th century, Lindemann showed that this is impossible by demonstrating that  $\pi$  is not a root of any polynomial with integral coefficients. Such numbers are said to be transcendental. Although the distinction between algebraic and transcendental numbers seems similar to the distinction between rational and irrational numbers, transcendental numbers are considerably more mysterious.

For example, it is a standard fact that the algebraic numbers are countable, while the transcendentals are uncountable. Although there are more transcendental numbers, it is surprisingly difficult to exhibit any transcendental number, let alone show a naturally occurring number like  $e$  or  $\pi$  is transcendental.<sup>1</sup>

We will first discuss Liouville's theorem, which shows that certain specially constructed numbers are transcendental. We then follow Hermite in proving that  $e$  is transcendental. The idea is to approximate the function  $e^z$  by rational functions, known as Padé approximants, and use these good approximations to show  $e$  cannot be algebraic. The actual proof is short and self-contained, but cryptic in isolation. We then illustrate how Padé approximants give a natural derivation of the continued fraction expansion of  $e$ , following Cohn [3]. Finally, we prove the Lindemann-Weierstrass theorem and as a consequence shows that  $\pi$  is transcendental, following the standard method as presented in [4] or [2].

## 2. APPROXIMATION AND TRANSCENDENTAL NUMBERS

We first prove a classical approximation theorem for algebraic numbers which will let us show that certain explicit, specially constructed numbers are transcendental.

**Theorem 1** (Liouville). *Let  $\beta$  be a solution to a polynomial  $f(x) = \sum_i a_i x^i$  of degree  $d$  with integral coefficients. Then there exists a constant  $\delta$ , depending on  $\beta$ , such that for any rational number  $\frac{p}{q} \neq \beta$  with  $q > 0$ , we have*

$$\left| \frac{p}{q} - \beta \right| > \frac{\delta}{q^d}.$$

*Proof.* We factor  $f(x) = a_d(x - \beta)^m(x - \alpha_1)^{m_1} \dots (x - \alpha_n)^{m_n}$  over the complex numbers. Let  $\gamma$  be the minimum distance between  $\beta$  and any of the  $\alpha_i$ . For any  $\delta < \gamma$ , no root  $\alpha_i$  of  $f(x)$  satisfies

$$|\alpha_i - \beta| < \frac{\delta}{q^d} < \gamma.$$

Thus we may assume that  $\frac{p}{q}$  is not a root, so  $q^d f(\frac{p}{q})$  is a non-zero integer. Thus

$$|q^d a_0 + q^{d-1} p a_1 + \dots + a_d p^d| \geq 1$$

---

*Date:* 2012, updated April 17, 2016.

<sup>1</sup>In contrast, it is elementary to find families of irrational numbers, like  $\sqrt{n}$  for  $n$  square-free.

and hence using the factorization that

$$(1) \quad \left| \frac{p}{q} - \beta \right|^m \geq \frac{1}{a_d q^d \left( \frac{p}{q} - \alpha_1 \right)^{m_1} \dots \left( \frac{p}{q} - \alpha_n \right)^{m_n}}.$$

If  $\left| \frac{p}{q} - \beta \right| > \gamma$  and  $\delta < \gamma$ , the proof is finished. So suppose that that  $\left| \frac{p}{q} - \beta \right| \leq \gamma$ . Under this assumption, letting  $\rho$  be the maximum of  $|\alpha_i - \beta|$  we see that

$$\left| \alpha_i - \frac{p}{q} \right| \leq |\alpha_i - \beta| + \left| \beta - \frac{p}{q} \right| \leq \rho + \gamma$$

This gives the estimate

$$\left| \frac{p}{q} - \alpha_1 \right|^{m_1} \dots \left| \frac{p}{q} - \alpha_n \right|^{m_n} \leq (\rho + \gamma)^{d-m}.$$

Taking  $\delta$  to be positive but less than both  $\gamma$  and  $(a_d(\rho + \gamma)^{d-m})^{-\frac{1}{m}}$  and using (1) gives

$$\left| \frac{p}{q} - \beta \right|^m \geq \frac{\delta^m}{q^d}$$

and hence after taking  $m$ th roots that

$$\left| \frac{p}{q} - \beta \right| \geq \frac{\delta}{q^{d/m}} \geq \frac{\delta}{q^d}.$$

Thus no rational number satisfies

$$\left| \frac{p}{q} - \beta \right| < \frac{\delta}{q^d}. \quad \square$$

This theorem can be used to exhibit an explicit transcendental number. Consider the number  $\beta = \sum_{k=1}^{\infty} 2^{-k!}$ . If it were algebraic, then it would be the root of a degree  $d$  polynomial. By Liouville's theorem, there would be a  $\delta > 0$  such that no rational number  $\frac{p}{q}$  satisfies

$$\left| \frac{p}{q} - \beta \right| < \frac{\delta}{q^d}.$$

This is not possible as the non-zero binary digits of  $\beta$  are far apart, so  $\sum_{k=1}^n 2^{-k!}$  will be too good an

approximation. More precisely, let  $q = 2^{K!}$  and  $p = \sum_{k=1}^K 2^{K!-k!}$ . Then we have

$$\left| \frac{p}{q} - \beta \right| = \sum_{k=K+1}^{\infty} 2^{-k!} \leq \sum_{k=(K+1)!}^{\infty} 2^{-k} = \frac{2^{-(K+1)!}}{1 - 2^{-1}} = 2q^{-(K+1)}.$$

If we pick  $K$  so  $K > d$  and  $2 \cdot 2^{-K!} < \delta$ , we get

$$\left| \frac{p}{q} - \beta \right| \leq \frac{2q^{-1}}{q^K} \leq \frac{\delta}{q^d}.$$

This contradicts Liouville's theorem, and hence  $\beta$  is transcendental.

Liouville's theorem is a simple way to exhibit transcendental numbers. It is more work to show natural constants such as  $e$  and  $\pi$  are transcendental.

### 3. THE TRANSCENDENCE OF $e$

This section will establish that  $e$  is a transcendental number. We first present the elementary argument that  $e$  is irrational, and then look at the theory of Padé approximants to prove that  $e^r$  is irrational for  $r \in \mathbb{Q} \setminus \{0\}$ . We then simplify the proof and use its essential features to prove that  $e$  is transcendental.

**3.1. The irrationality of  $e$ .** It is very easy to show that  $e$  is irrational using its definition as

$$e := \sum_{n=1}^{\infty} \frac{1}{n!}.$$

Suppose  $e = \frac{a}{b}$  with  $a$  and  $b$  relatively prime integers. Then for  $n \geq b$ ,

$$N := n! \left( e - \sum_{k=0}^n \frac{1}{k!} \right)$$

is an integer because  $n!e$  and  $\frac{n!}{k!}$  are integers. However, estimating  $N$  shows that

$$0 < N = \frac{1}{n+1} + \frac{1}{(n+1)(n+2)} + \dots < \frac{1}{n+1} + \frac{1}{(n+1)^2} + \dots = \frac{1}{n}$$

which is impossible. Thus  $e$  is irrational.

**3.2. Interlude.** The above argument demonstrates the general method of proving that numbers are transcendental: first use analytic properties to approximating the number, and then suppose the number is algebraic and use the assumption to show the error is an integer between 0 and 1. Sometime it is more convenient to use the assumption to show the estimate is a non-zero multiple of a large integer but show the error cannot be that big.

The analytic content used to prove that  $e$  is irrational was the power series expansion of  $e^z$  around  $z = 0$ . Unfortunately, this does not generalize to showing  $e$  is transcendental, or even that related numbers like  $e^5$  are irrational. The approximation of  $e^z$  by a truncated power series is too weak to readily give proofs of these facts. Power series determine holomorphic functions which are the basic building blocks of the world of complex analysis. Meromorphic functions are the next simplest object, so we can hope to better approximate  $e^z$  using rational functions, just as ratios of integers provide better approximations to real numbers than integers do.

**3.3. Approximations by Rational Functions.** Hermite was able to use the idea of approximating  $e^z$  by rational functions to show that  $e$  is transcendental. The theory of approximating general power series was first studied systematically by Hermite's student Padé near the end of the nineteenth century, so these rational functions are called Padé approximants despite the special cases we care about being used earlier.

**Definition 2.** For an analytic function  $f(z)$  and pair  $(n, m)$  of non-negative integers, a Padé approximant of order  $(n, m)$  at 0 is a rational function  $R(z) = \frac{P(z)}{Q(z)}$  with  $\deg(P) \leq n$  and  $\deg(Q) \leq m$  such when expanded as power series around 0 we have

$$f(z) - R(z) = O(z^{n+m+1}).$$

Padé approximants are analogous to rational approximations of irrational numbers, so we hope for similarities with the theory of approximating real numbers by rationals. From the theory of continued fractions, we know:

**Fact 3.** *Let  $\alpha$  be an irrational number. For any  $n$ , there exists a rational number  $\frac{p}{q}$  with  $p$  and  $q$  relatively prime integers and  $q > n$  such that*

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

The quantity  $|p - q\alpha|$  is an estimate in the error of the approximation relative to the size of the denominator, and goes to zero as the size of the denominator becomes large. For a Padé approximant  $\frac{P(z)}{Q(z)}$  of  $f(z)$ , the analogous term is  $P(z) - Q(z)f(z)$  and being small means vanishing to a large order at  $z = 0$ .

We are mainly interested in Padé approximants for  $e^z$ , so instead of dealing with the general theory we will show that they are unique and then construct Padé approximants for  $e^z$  directly.

**Proposition 4.** *The Padé approximant of order  $(n, m)$  is unique up to scaling  $P$  and  $Q$ .*

*Proof.* Suppose  $\frac{P_1(z)}{Q_1(z)}$  and  $\frac{P_2(z)}{Q_2(z)}$  are both Padé approximants to the same function, both of order  $(n, m)$ . Then when expanding them as power series around 0, we have

$$\frac{P_1(z)}{Q_1(z)} - \frac{P_2(z)}{Q_2(z)} = O(z^{m+n+1})$$

Combining the fractions, we see that since  $Q_1(z)$  and  $Q_2(z)$  are nonzero at  $z = 0$ , the order of vanishing of  $P_1(z)Q_2(z) - P_2(z)Q_1(z)$  must be at least  $m + n + 1$ . But this is a polynomial of degree at most  $m + n$ , so it vanishes identically. Thus  $P_1 = cP_2$  and  $Q_1 = cQ_2$ .  $\square$

Since there are  $m + n + 2$  coefficients to pick and  $\frac{P(x)}{Q(x)}$  depends only on the polynomials up to scalars, we expect to be able to force  $m + n + 1$  coefficients to agree. So general existence theorems for Padé approximants are not surprising. However, we simply need to understand the approximants for  $e^z$ . One simple way to do this is to reformulate definition of Padé approximants in terms of complex analysis. Clearing the denominator in the definition, we can equally well require that the polynomials  $P(z)$  and  $Q(z)$  satisfy

$$P(z) - e^z Q(z) = O(z^{m+n+1})$$

as power series around 0, or equivalently that

$$\frac{P(z) - e^z Q(z)}{z^{m+n+1}}$$

is holomorphic at  $z = 0$ .

To find  $p$  and  $q$ , the key is experience with evaluating integrals involving the exponential function multiplied by a polynomial. For example,

$$\int_0^1 x e^{z-xz} dx = \frac{-1}{z} x e^{z-xz} \Big|_0^1 + \int_0^1 \frac{e^{z-xz}}{z} dx = \frac{e^z - z - 1}{z^2}.$$

In general, if  $r(x)$  is a polynomial of degree  $n$ , then

$$\int_0^1 r(x) e^{z-xz} dx = \frac{P(z)e^z - Q(z)}{z^{n+1}}$$

where  $Q(z)$  and  $P(z)$  are polynomials obtained via repeated integration by parts. To find Padé approximants, we need the polynomials  $P(z)$  and  $Q(z)$  produced by integration by parts to have low degrees. This requires a special choice of  $r(x)$ .

**Proposition 5.** *For any polynomial  $r(x)$  of degree  $m + n$ ,*

$$\int_0^1 r(x) e^{z-xz} dx = \frac{F(z, 0)e^z - F(z, 1)}{z^{m+n+1}}$$

where  $F(z, x)$  is the sum

$$(2) \quad F(z, x) := r(x)z^{m+n} + r'(x)z^{m+n-1} + \dots + r^{(n+m)}(x).$$

In particular, for any  $m$  and  $n$ , there exist a polynomial  $r(x)$  of degree  $m + n$  such that

$$\int_0^1 r(x)e^{z-xz} dx = \frac{P(z)e^z - Q(z)}{z^{m+n+1}}$$

where  $\deg(P) \leq m$  and  $\deg(Q) \leq n$ .

*Proof.* The formula for  $F$  can be found by repeated integration by parts, but we can just verifying it directly:

$$\frac{d}{dx} (-F(z, x)e^{z-xz}) = \sum_{i=0}^{n+m} \left( r^{(i)}(x)z^{m+n+1-i}e^{z-xz} - r^{(i+1)}(x)z^{m+n-i}e^{z-xz} \right) = r(x)z^{m+n+1}e^{z-xz}.$$

Selecting  $r(x)$  so that  $r(0) = r'(0) = \dots = r^{(n)}(0) = 0$  and  $r(1) = r'(1) = \dots = r^{(m)}(1) = 0$  forces the degree of  $P(z) = F(z, 0)$  and  $Q(z) = F(z, 1)$  to be at most  $m$  and  $n$  respectively. To do this, take  $r(x) = x^n(1-x)^m$ .  $\square$

**Corollary 6.** For any pair  $(n, m)$  there exist a Padé approximant of order  $(n, m)$  for  $e^z$ .

**3.4. The Irrationality of Nonzero Rational Powers of  $e$ .** These basic ideas about Padé approximants provide a method (following [1]) to show powers of  $e$  are irrational.

**Proposition 7.** Let  $a$  be a non-zero rational number. Then  $e^a$  is irrational.

*Proof.* It suffices to prove this when the exponent is a positive integer, since powers of rational numbers are rational. Let  $e^p = \frac{a}{b}$  with  $a, b$  positive and relatively prime. By our analogy with approximation of rational numbers, we want to take  $z = p$  and use the fact that  $P(p) - Q(p)e^p$  is an estimate on the error. Combined with the hypothesis that  $e^p$  is rational, it will yield a contradiction. The error is closely related to the integrals of the previous section. If we set  $r(x) = x^n(1-x)^n$ , the relevant integral becomes

$$\int_0^1 r(x)e^{p-xp} dx = \frac{F(p, 0)e^p - F(p, 1)}{p^{2n+1}}.$$

Using the assumption that  $e^p$  is rational, a nice bound is

$$aF(p, 0) - bF(p, 1) = bp^{2n+1} \int_0^1 e^{p-px} x^n(1-x)^n dx \leq bp^{2n+1}e^p = ap^{2n+1}.$$

However, we know that  $r^{(k)}(0) = r^{(k)}(1) = 0$  for  $0 \leq k < n$ . Thus the only derivatives appearing in (??) are at least the  $n$ th derivative. Since  $r(x)$  has integral coefficients the  $n$ th and further derivatives at 0 and 1 must be multiples of  $n!$  by Taylor's formula. Thus  $aF(p, 1) - bF(p, 0)$  is a multiple of  $n!$ . It is not zero since the integral is clearly positive. If we chose  $n$  so that  $n! > ap^{2n+1}$  (possible since  $n! > e\left(\frac{n}{e}\right)^n$ ) we get a contradiction. Thus  $e^p$  is irrational.  $\square$

As we attempt to generalize this, there is one simplification to keep in mind. We argued that  $F(p, 1)$  and  $F(p, 0)$  are multiples of a large number ( $n!$  in this case). This happens because the first  $n$  derivatives of  $f(x)$  are zero at  $z = 0$  and  $z = 1$ . Since  $f$  has integral coefficients, Taylor's formula implies the higher derivatives are multiples of  $n!$ . It would be better to divide  $r(x)$  by  $n!$  so that  $F(p, 1)$  and  $F(p, 0)$  are integers and show the error is an integer between 0 and 1 than to try to keep track of divisibility. This will be implemented in the next section as we show  $e$  is transcendental.

3.5. **Transcendence of  $e$ .** Suppose  $e$  is algebraic, so there exists integers  $b_i \in \mathbb{Z}$  such that

$$b_0 + b_1 e + \dots + b_d e^d = 0.$$

We want to approximate  $e, e^2, \dots, e^d$  by rational functions, and then evaluate the error terms. One way to approximate the powers of  $e$  would simply be to use Padé approximants for each one. But they will be unrelated, and the argument becomes unnecessarily complicated.<sup>2</sup> The correct approach is a generalization of the problem of simultaneous approximation of irrationals.

**Fact 8** (Dirichlet). *Given  $\alpha_1, \dots, \alpha_r$  in  $\mathbb{R}$  and  $n > 0$ , there exist  $p_1, \dots, p_r, q \in \mathbb{Z}$  with  $0 < q \leq n$  such that*

$$\left| \frac{p_i}{q} - \alpha_i \right| < \frac{1}{qn^{\frac{1}{r}}}.$$

The key is to approximate the irrationals by fractions of the same denominator. As before, the quantity  $|p_i - q\alpha_i|$  is an estimate of the error relative the size of the denominator that becomes small as  $q$  becomes large.

By analogy, we need to approximate  $e, e^2, \dots, e^d$  by rational functions with the same denominator. We look at integrals of the form

$$\int_0^1 r(x) e^{z-xz} dx = \frac{F(z, 0)e^z - F(z, 1)}{z^{n+1}}$$

If we simply try plugging in different values of  $z$ , the denominator changes. To get the same denominator, we need to fix  $z$  and change the upper bound of integration. This will replace the  $e$  by  $e^m$  while keeping the  $F(z, 0)$  term. Therefore we are led to look at

$$\int_0^m r(x) e^{m-x} dx = F(1, 0)e^m - F(1, m)$$

as a measure of the error of approximating  $e^m$  by  $\frac{F(1, m)}{F(1, 0)}$ .

So to set up the proof, define

$$F(x) = r(x) + r'(x) + \dots + r^{\deg(r)}(x) \quad \text{and} \quad I(m) := \int_0^m r(x) e^{m-x} dx = F(0)e^m - F(m).$$

Since we want the degrees of  $F(j)$  to be small for  $j = 0, 1, 2, \dots, d$ , take

$$r(x) = \frac{1}{(n-1)!} x^{n-1} ((x-1)(x-2)\dots(x-d))^n.$$

This is very similar to the choice of  $x^n(1-x)^n$  from the irrationality argument, except we look at more points and we simplify by dividing by  $(n-1)!$ .

If we approximate  $e^m$  by  $\frac{F(m)}{F(0)}$ , the error is

$$0 + b_1 \left( e - \frac{F(1)}{F(0)} \right) + \dots + b_d \left( e^d - \frac{F(m)}{F(0)} \right).$$

Multiplying through by the common denominator  $F(0)$  gives

$$0 + b_1 I(1) + b_2 I(2) + \dots + b_d I(d).$$

as an estimate of the error. The term  $I(m) = F(m) - F(0)e^m$  corresponds to the error term  $b\alpha_i - a_i$  when we approximate the real number  $\alpha_i$  by  $\frac{a_i}{b}$ .

**Theorem 9.** *The number  $e$  is transcendental.*

<sup>2</sup>This approach is needed to prove the Lindemann-Weierstrass theorem in section 5.1. That proof uses the same ideas as in Hermite's proof but requires more work. The transcendence of  $e$  does follow immediately.

*Proof.* Suppose  $e$  is algebraic and satisfies

$$b_0 + b_1e + \dots + b_de^d = 0$$

where the  $b_i$  are integers. We have defined

$$\begin{aligned} r(x) &= \frac{1}{(n-1)!} x^{n-1} ((x-1)(x-2)\dots(x-d))^n \\ F(x) &= r(x) + r'(x) + \dots + r^{(nd-1)}(x) \\ I(m) &= \int_0^m e^{m-x} r(x) dx = F(0)e^m - F(m). \end{aligned}$$

We consider the error term

$$\begin{aligned} J &:= b_1I(1) + b_2I(2) + \dots + b_dI(d) \\ &= b_0(F(0) - F(0)) + b_1(F(0)e - F(1)) + b_2(F(0)e^2 - F(2)) + \dots + b_d(F(0)e^d - F(d)) \\ &= -(b_0F(0) + b_1F(1) + \dots + b_dF(d)). \end{aligned}$$

The first step is to show that  $J$  is a nonzero integer. We will do this when  $n$  is a prime larger than  $|b_0|$ . First, a lemma.

**Lemma 10.** *Let  $m = 0, 1, \dots, d$  and  $k$  be a positive integer. Then  $r^{(k)}(m)$  is an integer. Unless  $k = n - 1$  and  $m = 0$ ,  $r^{(k)}(m)$  is a multiple of  $n$ .*

*Proof.* The derivative  $r^{(k)}(m)$  is  $k!$  times the coefficient of  $(x - m)^k$  in the power series expansion of  $r(x)$  around  $m$ . By the definition of  $r(x)$ , it vanishes to order  $n - 1$  at 0 and to order  $n$  at  $1, 2, \dots, d$ . Therefore when  $r^{(k)}(m)$  is non-zero  $k \geq n - 1$  and hence the denominator is canceled out. Provided  $n$  is much larger than  $d$ ,  $r^{n-1}(0)$  is clearly the only derivative not a multiple of  $n$ .  $\square$

This immediately shows that  $F(m)$  is an integer, and hence  $J$  is an integer also. To show  $J$  is a nonzero integer, note that all of the derivatives appearing in the sums  $F(0), F(1) \dots F(d)$  are multiples of  $n$  except for  $r^{(n-1)}(0)$ . But since  $|b_0| < n$  and  $n$  is prime, the term  $b_0r^{(n-1)}(0)$  is not a multiple of  $n$ . Therefore  $J$  is a non-zero integer.

The second step is to estimate the integral. If  $y$  is in the interval  $[0, d]$  then

$$|r(y)| \leq \frac{1}{(n-1)!} (2d)^{n-1} (2d)^n \dots (2d)^n = \frac{1}{(n-1)!} (2d)^{dn-1}.$$

As  $(n-1)!$  grows faster than  $c^n$ , this is less than any  $\epsilon > 0$  for sufficiently large  $n$ . But then

$$|I(m)| = \left| \int_0^m r(x) e^{m-x} dx \right| \leq m e^m \epsilon, \text{ so}$$

$$\begin{aligned} |J| &\leq |b_1I(1)| + |b_2I(2)| + \dots + |b_dI(d)| \\ &\leq \epsilon (b_1 \cdot e + b_2 \cdot 2e^2 + \dots + b_d \cdot de^d) \end{aligned}$$

which can be made less than 1 by appropriate choice of  $\epsilon$ . This contradicts the fact that  $J$  is a non-zero integer. Therefore  $e$  is transcendental.  $\square$

**Corollary 11.** *For any non-zero rational number  $a$ ,  $e^a$  is transcendental.*

*Proof.* Suppose  $\beta = e^a = e^{\frac{p}{q}}$  is algebraic where  $p$  and  $q$  are integers. We may assume both are positive. Then  $e$  satisfies the equation  $z^p - \beta^q$  which has algebraic coefficients. Since the algebraic numbers form an algebraically closed field,  $e$  is algebraic, a contradiction.  $\square$

**Remark 12.** The choice to use the different exponents  $n - 1$  and  $n$  in

$$r(x) = x^{n-1} ((x-1)(x-2)\dots(x-d))^n$$

simplifies Hermite's original by making it clear that  $J$  is a non-zero integer.

4. AN APPLICATION OF PADÉ APPROXIMANTS TO THE CONTINUED FRACTION EXPANSION OF  $e$ 

Continued fractions of rational numbers and of quadratic irrationals are finite and periodic respectively. There are also surprising patterns in the continued fraction expansions of transcendental numbers like  $e$ . Padé approximants provide a way to find the continued fraction expansion of  $e$ . (Another way is to follow Euler and look at a differential equation called the Riccati equation.)

A few minutes with a calculator will suggest that the continued fraction expansion of  $e$  is  $[2, 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, \dots]$ . (Sometimes this is re-written as  $[1, 0, 1, 1, 2, 1, 1, 4, 1, 1, \dots]$  to extend the pattern.) Using the standard notation for continued fractions, for  $i \geq 1$  we set  $a_{3i} = 2i$ ,  $a_{3i+2} = a_{3i+1} = 1$ . Letting  $[a_1, a_2, \dots, a_i] = \frac{p_i}{q_i}$ , in general we have

$$p_{i+2} = a_{i+2}p_{i+1} + p_i \quad \text{and} \quad q_{i+2} = a_{i+2}q_{i+1} + q_i.$$

In this case, we can recursively calculate

$$\begin{aligned} p_{3n} &= 2np_{3n-1} + p_{3n-2}, & q_{3n} &= 2nq_{3n-1} + q_{3n-2} \\ p_{3n+1} &= p_{3n} + p_{3n-1}, & q_{3n+1} &= q_{3n} + q_{3n-1} \\ p_{3n+2} &= p_{3n+1} + p_{3n}, & q_{3n+2} &= q_{3n+1} + q_{3n}. \end{aligned}$$

The question is whether  $\lim_{n \rightarrow \infty} [a_0, a_1, a_2, \dots, a_n] = \lim_{n \rightarrow \infty} \frac{p_n}{q_n}$  equals  $e$ . To check this, we can just check that

$$\lim_{n \rightarrow \infty} q_n e - p_n = 0.$$

Using the connection with Padé approximants, this should be given by an integral  $\int_0^1 r(x)e^x dx$ .

Again, we have a choice of what to use for  $r(x)$ . If we take  $r(x) = \frac{x^n(1-x)^n}{n!}$ , the integral evaluates to some of the convergents of  $e$ . Denote the integral by  $I(n)$ . For example,  $n = 1$  gives  $3 - e$ ,  $n = 2$  gives  $7e - 19$ , and  $I(3) = 193 - 71e$ . The second, fifth, and eighth convergents to  $e$  are  $3$ ,  $19/7$ , and  $193/71$ . Where are the other convergents coming from? Our choice of  $r(x)$  gives a good approximation because the degree of the numerator and denominator are balanced. We would still expect a good approximation if the degrees were  $n$  and  $n + 1$ . Define  $J(n)$  and  $K(n)$  to be

$$J(n) := \int_0^1 \frac{x^{n+1}(1-x)^n}{n!} e^x dx \quad \text{and} \quad K(n) := \int_0^1 \frac{x^n(1-x)^{n+1}}{n!} e^x dx.$$

**Lemma 13.** *The integrals are related to the convergents of  $e$  as follows:*

$$\begin{aligned} I(n) &= (-1)^n (q_{3n-1}e - p_{3n-1}) \\ J(n) &= (-1)^{n+1} (q_{3n}e - p_{3n}) \\ K(n) &= (-1)^n (q_{3n+1}e - p_{3n+1}). \end{aligned}$$

*Proof.* We will prove the results by showing that the integrals satisfy the same initial conditions and recurrence relations as the convergents. Now we have that

$$\begin{aligned} I(1) &= \int_0^1 x(1-x)e^x dx = 3 - e \\ J(1) &= \int_0^1 x^2(1-x)e^x dx = 3e - 8 \\ K(1) &= \int_0^1 x(1-x)^2 e^x dx = 11 - 4e. \end{aligned}$$

Furthermore, integration by parts gives that

$$I(n+1) = \int_0^1 \frac{x^{n+1}(1-x)^{n+1}}{(n+1)!} e^x dx = - \int_0^1 \frac{-x^{n+1}(1-x)^n + x^n(1-x)^{n+1}}{n!} e^x dx = J(n) - K(n).$$



In addition,

$$\begin{aligned}
 J(n) &= \int_0^1 \frac{x^{n+1}(1-x)^n}{n!} e^x dx \\
 &= - \int_0^1 \frac{(n+1)x^n(1-x)^n - nx^{n+1}(1-x)^{n-1}}{n!} e^x dx \\
 &= - \int_0^1 \frac{(2n+1)x^n(1-x)^n - nx^n(1-x)^{n-1}}{n!} e^x dx \\
 &= -(2n+1)I(n) + J(n-1) \\
 &= -2nI(n) + K(n-1)
 \end{aligned}$$

Finally, elementary algebra shows that

$$K(n) = \int_0^1 \frac{x^n(1-x)^{n+1}}{n!} e^x dx = \int_0^1 \frac{x^n(1-x)^n}{n!} e^x dx - \int_0^1 \frac{x^{n+1}(1-x)^n}{n!} e^x dx = I(n) - J(n).$$

In particular, these relations among the integrals are the same recurrence relations as the numerator and denominator of the convergents.  $\square$

It is now easy to derive the continued fraction expansion of  $e$ .

**Theorem 14.** *We have that  $e = [2, 1, 2, 1, 1, 4, 1, 1, \dots]$ .*

*Proof.* It suffices to show that  $\lim_{n \rightarrow \infty} p_n - q_n e = 0$ . But this expression is given by  $I(m)$ ,  $J(m)$ , or  $K(m)$  where  $m \approx n/3$ . But as  $n \rightarrow \infty$  the size of all three integrals goes to 0 since the interval is a constant length,  $x^j(1-x)^k e^x$  is bounded, and the denominator  $n!$  goes to infinity. Therefore the limit is 0, so  $[2, 1, 1, 4, 1, 1, 6, \dots]$  gives a continued fraction expansion of  $e$ .  $\square$

## 5. MORE GENERAL TRANSCENDENCE RESULTS

It is possible to prove that  $\pi$  is irrational and transcendental using arguments analogous to those for  $e$ . This is not surprising, as  $e^{\pi i} = -1$  connects the two. However, it is also a consequence of a much more general theorem which Lindemann proved for  $n = 1$  and Weierstrass proved in general.

**Theorem 15** (Lindemann-Weierstrass). *Let  $\beta_1, \dots, \beta_n$  be algebraic numbers not all zero, and  $\alpha_1, \dots, \alpha_n$  be distinct algebraic numbers. Then*

$$\beta_1 e^{\alpha_1} + \beta_2 e^{\alpha_2} + \dots + \beta_n e^{\alpha_n} \neq 0.$$

**Corollary 16.** *The number  $\pi$  is transcendental.*

*Proof.* Suppose  $\pi$  is algebraic. Then so is  $\pi i$ , and taking  $\alpha_1 = 0, \alpha_2 = \pi i$  and  $\beta_1 = \beta_2 = 1$  we have that

$$1 + e^{\pi i} \neq 0. \quad \square$$

**Corollary 17.** *If  $\alpha$  is a non-zero algebraic integer,  $e^\alpha$ ,  $\sin(\alpha)$ , and  $\cos(\alpha)$  are transcendental.*

*Proof.* Suppose  $e^\alpha$  satisfies a polynomial equation of degree  $n$ . Then letting  $\alpha_j = j\alpha$  for  $0 \leq j \leq n$  and  $\beta_j$  be the (algebraic) coefficient of  $x^j$ , the theorem produces a contradiction. For the other parts, write the trigonometric functions in terms of exponentials.  $\square$

The proof of Lindemann's theorem has two components. The first is an analytic argument similar to the proof that  $e$  is transcendental that establishes the theorem in the special case that the  $\beta_i$  are rational and that there are integers  $n_0 = 0 < n_1 < n_2 < \dots < n_r = n$  such that

$$(3) \quad \beta_{n_t+1} = \beta_{n_t+2} = \dots = \beta_{n_{t+1}} \text{ and } \alpha_{n_t+1}, \alpha_{n_t+2}, \dots, \alpha_{n_{t+1}} \text{ are all of the conjugates of } \alpha_{n_t+1}.$$

The second component is algebraic in nature, and uses ideas from the theory of symmetric polynomials to extend the first part to the general theorem.

**5.1. The Analytic Part.** With the hypothesis (3) on the  $\alpha_i$  and the  $\beta_i$  given above, fix a  $b$  for which the  $b\alpha_i$  is an algebraic integer for each  $i$ . We need to approximate  $e^{\alpha_i z}$  by a quotient of rational functions, but the argument is now more complicated because the  $\alpha_i$  are no longer just  $1, 2, \dots, d$ . Trying a similar approach, let

$$(4) \quad r_i(x) = b^{n(p)}(x - \alpha_1)^p \dots (x - \alpha_n)^p / ((p-1)!(x - \alpha_i))$$

where  $p$  is a large prime number. Define

$$I_i(\alpha) = \int_0^\alpha r_i(z) e^{\alpha-z} dz.$$

This is a generalization of the  $I(m)$ , and is well defined independent of the path from 0 to  $z$  because the integrand is holomorphic. Via integration by parts, we can evaluate this integral as

$$(5) \quad I_i(\alpha) = F_i(0)e^\alpha - F_i(\alpha) \quad \text{where} \quad F_i(z) = \sum_{k=0}^{np-1} r_i^{(k)}(z).$$

**Lemma 18.** *Let  $m = 1, \dots, n$  and  $k$  be a positive integer. Then  $r_i^{(k)}(\alpha_m)$  is an algebraic integer. Unless  $k = p-1$  and  $m = i$ ,  $r_i^{(k)}(m)$  is  $p$  times an algebraic integer.*

This is essentially the same as Lemma 10. Finally define

$$J_i = \beta_1 I_i(\alpha_1) + \beta_2 I_i(\alpha_2) + \dots + \beta_n I_i(\alpha_n).$$

We will give bounds on  $|J_1 \dots J_n|$  using the hypothesis that

$$\beta_1 e^{\alpha_1} + \dots + \beta_n e^{\alpha_n} = 0.$$

Using (5) on  $J_i$ , the terms involving powers of  $e$  in  $J_i$  disappear and we have

$$J_i = -\beta_1 F_i(\alpha_1) - \beta_2 F_i(\alpha_2) - \dots - \beta_n F_i(\alpha_n).$$

By the lemma, all of the terms in the sums for  $F_i(\alpha_j)$  are  $p$  times an algebraic integer except for the term  $r_i^{(p-1)}(\alpha_i)$ . Now recognize that

$$r_i^{(p-1)}(\alpha_i) = b^{np} \prod_{\substack{1 \leq k \leq n \\ k \neq i}} (\alpha_i - \alpha_k)^p = (g'(\alpha_i))^p$$

where

$$g(x) = \prod_{k=1}^n (bx - b\alpha_k) \in \mathbb{Z}[X].$$

If we now look at the product  $J_1 J_2 \dots J_n$ , all but one term in the product will be  $p$  times an algebraic integer. This term is the product of the terms from each of the  $J_i$  that are not multiples of  $p$ , and so is

$$g'(\alpha_1)^p g'(\alpha_2)^p \dots g'(\alpha_n)^p.$$

It is a rational integer, since all of the  $g'(\alpha_i)$  are algebraic integers and it is invariant under the action of the Galois group since the  $\alpha_i$  include complete sets of conjugates. Since  $g'(\alpha_1)g'(\alpha_2) \dots g'(\alpha_n)$  is independent of  $p$ , for large  $p$  it cannot be a multiple of  $p$ . Thus  $J_1 J_2 \dots J_n$  is a non-zero algebraic integer.

We will next show that  $J_1 J_2 \dots J_n$  is rational by showing it is symmetric in  $\alpha_{n_t+1}, \dots, \alpha_{n_{t+1}}$  (recall this is the grouping of the  $\alpha_i$  into sets of conjugates). So let  $\sigma$  be any permutation of  $n_t + 1, \dots, n_{t+1}$  and extend it by the identity to be a permutation on  $1, 2, \dots, n$ . Now

$$J_1 J_2 \dots J_n = (-1)^n \prod_{i=1}^n (\beta_1 F_i(\alpha_1) + \beta_2 F_i(\alpha_2) + \dots + \beta_n F_i(\alpha_n)).$$

Unraveling the definition of  $F_i(\alpha_j)$ , we see that applying  $\sigma$  to it gives  $F_{\sigma(i)}(\alpha_{\sigma(j)})$ . But then applying  $\sigma$  to the product and using the fact that  $\beta_i \in \mathbb{Q}$ , we obtain

$$\prod_{i=1}^n (\beta_1 F_{\sigma(i)}(\alpha_{\sigma(1)}) + \dots + \beta_n F_{\sigma(i)}(\alpha_{\sigma(n)})) = \prod_{i=1}^n (\beta_{\sigma^{-1}(1)} F_i(\alpha_1) + \dots + \beta_{\sigma^{-1}(n)} F_i(\alpha_n))$$

by re-indexing. By hypothesis  $\beta_{n_t+1} = \beta_{n_t+2} = \dots = \beta_{n_{t+1}}$ , so the product is invariant. Therefore it is a non-zero rational integer.

However, estimating the integral as in the proof of the transcendence of  $e$  gives that

$$|J_1 \dots J_n| \leq \prod_{i=1}^n \sum_{k=1}^n |\beta_k| |\alpha_k| e^{|\alpha_k|} |r_i(\alpha_k)| < 1$$

for large enough  $p$  as  $(p-1)!$  grows faster than  $c^p$ . But a non-zero rational integer cannot have absolute value less than 1, so our hypothesis that

$$\beta_1 e^{\alpha_1} + \dots + \beta_n e^{\alpha_n} = 0$$

leads to a contradiction. □

**5.2. The Algebraic Part.** We now reduce from the general case of Lindemann's theorem to the case (3) when the  $\beta_i$  are rational and that there are integers  $n_0 = 0 < n_1 < n_2 < \dots < n_r = n$  such that

$$\beta_{n_t+1} = \beta_{n_t+2} = \dots = \beta_{n_{t+1}} \quad \text{and} \quad \alpha_{n_t+1}, \alpha_{n_t+2}, \dots, \alpha_{n_{t+1}}$$

are a complete set of conjugates.

To reduce to the case that the  $\beta_i$  are rational, we will inductively construct a sequence of non-zero polynomials  $F_i$  with rational coefficients in  $2n+1-i$  variables such that  $F_i(\beta_i, \dots, \beta_n, e^{\alpha_1}, \dots, e^{\alpha_n}) = \beta_1 e^{\alpha_1} + \dots + \beta_n e^{\alpha_n}$ . To construct the  $F_i$ , let  $F_1(x_1, \dots, x_n, y_1, \dots, y_n) = x_1 y_1 + \dots + x_n y_n$ . In general, let

$$F_i(x_i, \dots, x_n, y_1, \dots, y_n) := \prod_{\beta'} F_{i-1}(\beta', x_i, \dots, x_n, y_1, \dots, y_n)$$

where the product ranges over all  $\beta'$  that are conjugate to  $\beta_{i-1}$ . The definition is invariant under permuting the conjugates of  $\beta_{i-1}$ , so  $F_i$  has rational coefficients. By induction, if  $F_i$  has degree  $N_i$  the coefficient of  $y_1^{N_i}$  is always non-zero.

Next we expand out the polynomial  $F_{n+1}(e^{\alpha_1}, \dots, e^{\alpha_n})$  (of degree  $N$ ) in the form

$$\beta'_1 e^{\alpha'_1} + \dots + \beta'_m e^{\alpha'_m}$$

We wish to show that at least one of the coefficients  $\beta'_i$  is non-zero. Without loss of generality, we may assume that  $\alpha_1$  has the largest absolute value of the  $\alpha_i$ . The exponents  $\alpha'_i$  that arise are

those of the form  $\sum_{j=1}^n c_j \alpha_j$  where  $\sum_j c_j = N$ . The only way to get an exponent of  $N\alpha_1$  is to pick

$c_1 = N$ ,  $c_j = 0$  for  $j \neq 1$  since  $|\alpha_1|$  is the largest of all  $|\alpha_j|$ . Since the coefficient of  $y_1$  is non-zero, this implies that the term  $e^{N\alpha_1}$  has a non-zero coefficient. Thus we are in the case where the  $\beta'_i$  are rational, the  $\alpha'_n$  are algebraic, and not all of the  $\beta'_i$  are zero. This expression is zero if the original expression was, so it suffices to prove Lindemann's theorem when all of the  $\beta_i$  are rational. Multiplying through by the common denominator, we may also assume the  $\beta'_i$  are rational integers.

We now will further restrict to the case when all of the conjugates of  $\alpha_i$  appear with the same rational coefficient. Pick a polynomial with integral coefficients that has all of the  $\alpha_i$  as roots. Let  $\alpha_{n+1}, \dots, \alpha_N$  be the additional roots of this polynomial. Extend the definition of  $\beta_i$  to be zero for  $n < i \leq N$ , and consider the product

$$\prod_{\sigma \in S_N} (\beta_1 e^{\alpha_{\sigma(1)}} + \dots + \beta_N e^{\alpha_{\sigma(N)}}).$$

If

$$\beta_1 e^{\alpha_1} + \dots + \beta_n e^{\alpha_n} = 0$$

the entire product is zero. If we were to expand the product, it would again be of the form

$$\beta'_1 e^{\alpha'_1} + \dots + \beta'_m e^{\alpha'_m}.$$

Now if  $\beta e^\alpha$  arises as a term in this product,  $\alpha = c_1 \alpha_1 + \dots + c_N \alpha_N$  where  $c_1 + c_2 + \dots + c_N = N!$  and  $\beta$  is a product of rational integers. There will also be terms with exponent  $c_1 \alpha_{\sigma(1)} + \dots + c_N \alpha_{\sigma(N)}$  for any permutation. These terms will run through a complete set of conjugates for  $\alpha$  (along with other things) but have the same value of  $\beta$ . By grouping terms in this way, it is clear that the coefficient for each conjugate  $e^{\alpha'}$  will be the same. Furthermore, at least one will be non-zero: in each product

$$(\beta_1 e^{\alpha_{\sigma(1)}} + \dots + \beta_N e^{\alpha_{\sigma(N)}})$$

pick the non-zero term with the largest value of  $\alpha_i$ , where the ordering is done first by real part and then by imaginary part. This term in the expansion will be non-zero, and cannot be canceled by anything else because each of the exponents is the largest possible. This completes the reduction to the case where (3) holds which we already proved using the analytic argument.  $\square$

## 6. LATER RESULTS

The two results presented here were known in the 19th century. Over the course of the 20th century, the integrals and use auxiliary functions  $r(x)$  were generalized considerably yielding more powerful results. The results are proven using the same strategies, but more complicated functions. Details and proofs of the following can be found in Baker [2].

**Theorem 19** (Gelfond-Schneider). *Let  $\alpha$  be an algebraic number not equal to 0 or 1, and  $\beta$  an algebraic number that is not rational. Then  $\alpha^\beta$  is transcendental.*

**Theorem 20** (Baker). *If  $\beta_0, \beta_1, \dots, \beta_n$  and  $\alpha_1, \dots, \alpha_n$  are algebraic numbers, then*

$$\beta_0 + \beta_1 \log(\alpha_1) + \dots + \beta_n \log(\alpha_n)$$

*is either zero or transcendental.*

Baker's work also allows lower bounds to be constructed which suffice to solve the class number one problem.

Some other results and approaches are presented in Lang [5].

## REFERENCES

1. Martin Aigner and Günter M. Ziegler, *Proofs from The Book*, fifth ed., Springer-Verlag, Berlin, 2014, Including illustrations by Karl H. Hofmann. MR 3288091
2. Alan Baker, *Transcendental number theory*, second ed., Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1990. MR 1074572 (91f:11049)
3. Henry Cohn, *A short proof of the simple continued fraction expansion of  $e$* , Amer. Math. Monthly **113** (2006), no. 1, 57–62. MR 2202921 (2006j:11008)
4. Michael Filaseta, *Lecture notes for math 785*, Available at <http://www.math.sc.edu/~filaseta>.
5. Serge Lang, *Introduction to transcendental numbers*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1966. MR 0214547 (35 #5397)