Today we will tackle one of the most elusive and challenging proofs to discover. It was conjectured that e was transcendental in 1768, but was not proven until Charles Hermite published his proof in 1873. Today we will explore that proof in depth. Although the proof may seem impossible to come up with, it is remarkably easy to follow for such a complex topic. Lets begin with a discussion of what is means for a number to be transcendental.

An *irrational number*, let's call it x, is a number which cannot be written in the form a/b for integers a and b. In other words, there are no integer solutions for a and b such that

$$a - bx = 0$$

We can view this linear equation as a polynomial of degree one with integer coefficients. If we consider some irrational, say $\sqrt{2}$, we cannot find integers a and b to satisfy these conditions. However, we can find that $\sqrt{2}$ is a solution to the second degree polynomial

$$x^2 - 2 = 0$$

We call any number that can be constructed as the solution to a polynomial (with integer coefficients) of any order *algebraic*. If a number cannot be expressed in this way, then it is *transcendental*, and in a sense "more irrational" than an algebraic number. Now we are ready to explore the proof. We will be performing a proof by contradiction.

Proof Assume that e is algebraic. This implied for some polynomial

$$\varphi(x) = c_0 + c_1 x + c_2 x^2 + ... + c_n x^n$$

for with c_0 to c_n all integers, we have $\varphi(e) = 0$. Note we will constrain our equation such that c_0 must be positive. If c_0 were 0 we could divide by a sufficiently large power of x, and if it is negative we multiply all coefficients by -1.

Lemma 1 *Suppose that e is a root of the polynomial* $\varphi(x) = c_0 + c_1 x + ... + c_n x^n$ *Let f be a polynomial and*

$$F(x) = \sum_{i=0}^{\infty} f^{(i)}(x),$$

then there exists real numbers $0 < \alpha_1, ..., \alpha_n < 1$ *such that*

$$c_0 F(0) + ... + c_n F(n) = c_1 \beta_1 + ... + c_n \beta_n$$

$$\beta_k = -k e^{k(1-\alpha_k)} f(k\alpha_k)$$

Where $f^{(i)}(x)$ denotes the i order derivative of f evaluated at x (similar to Taylor series form). This lemma seems to appear out of nowhere and make little sense, but if we prove it you will see how straightforward the idea is.

Since f is a finite polynomial, let us assume it is of degree r. This implies that

$$f^{(r+1)}(x) = f^{(r+2)}(x) = ... = 0 \tag{1}$$

So can rewrite our infinite sum F(x) as a finite sum and differentiate

$$\begin{aligned}F(x) &= f(x) + f^{(1)}(x) + f^{(2)}(x) + ... + f^{(r)}(x) \\ F'(x) &= f^{(1)}(x) + f^{(2)}(x) + ... + f^{(r)}(x) + f^{(r+1)}(x) \\ &= f^{(1)}(x) + f^{(2)}(x) + ... + f^{(r)}(x) \\ &= F(x) - f(x)\end{aligned} \tag{2}$$

Note that we were able to eliminate the higher order derivative terms using (1). A useful identity. Now we start to get a bit more random in our explanation. Lets define a new function

$$g(x) = e^{-x}F(x) \tag{3}$$

What happens if we differentiate? By product rule,

$$\begin{aligned}g'(x) &= -e^{-x}F(x) + e^{-x}F'(x) \\ &= e^{-x}(F'(x) - F(X)) \\ &= -e^{-x}f(x)\end{aligned} \tag{4}$$

Where we simplified the inside using (2). As we are still exploring, let's consider applying the Mean Value Theorem for derivatives to this function. We recall that

Theorem 1 *If $f : \mathbb{R} \to \mathbb{R}$, and f is differentiable, then on any interval (a, b), there exists a value c, with $a < c < b$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

We will remodel this theorem slightly to fit our needs. Note that since since $a < c < b$, we can express c as

$$c = a + \alpha(b - a)$$

For some value of $\alpha \in (0, 1)$. Therefore the new version of our MVT states

Theorem 2 *If $f : \mathbb{R} \to \mathbb{R}$, and f is differentiable, then on any interval (a, b), there exists a value α, with $0 < \alpha < 1$ such that*

$$f'(a + \alpha(b - a)) = \frac{f(b) - f(a)}{b - a}$$

We will now apply this modified MVT to our function f on the interval $[0, k]$. Therefore, we have a value $0 < \alpha < 1$ such that

$$\begin{aligned}g'(0 + \alpha(k - 0)) &= \frac{g(k) - g(0)}{k - 0} \\ g'(\alpha k) &= \frac{g(k) - g(0)}{k}\end{aligned}$$

We can now apply equation (3) to substitute for g and g'.

$$\begin{aligned}
-e^{-\alpha k}f(\alpha k) &= \frac{e^{-k}F(k) - e^{0}F(0)}{k} \\
e^{-k}F(k) - F(0) &= -ke^{-\alpha k}f(\alpha k) \\
F(k) - e^{k}F(0) &= -ke^{k(1-\alpha)}f(\alpha k) = \beta
\end{aligned}$$

This is reached by suitable algebraic simplification, as show. We define this value to be β, as the notation will serve us well later.

This is where the unexpected magic will slowly start to appear. We will perform this MVT process for g on the interval $[0, k]$ for $k = 1, 2, ...n$ where we recall n to be the degree of polynomial φ, for which $\varphi(e) = 0$. Each time we perform the MVT we will get a unique value of α which we will call α_k, and a corresponding β_k. We will find for each k,

$$F(k) - e^{k}F(0) = -ke^{k(1-\alpha_k)}f(k\alpha_k) = \beta_k$$

We will multiply both sides by the corresponding coefficient of φ on the x^k term:c_k. Notice, this will give us

$$c_kF(k) - c_ke^{k}F(0) = c_k\beta_k \tag{5}$$

Finally, we can construct the final equation and prove the lemma. To do this, we simply add the result from (5) together for each value $k = 1, 2, ..., n$. We then get

$$c_1F(1) + c_2F(2) + ... + c_nF(n) - F(0)(c_1e + ... + c_ne^n) = c_1\beta_1 + ... + c_n\beta_n$$

The term in parentheses can be simplified greatly since

$$\begin{aligned}
\varphi(e) &= 0 \\
&\implies c_0 + c_1e + ... + c_ne^n = 0 \\
-c_0 &= c_1e + ... + c_ne^n
\end{aligned}$$

We can substitute this equation into the parentheses to show that

$$c_0F(0) + c_1F(1) + c_2F(2) + ... + c_nF(n) = c_1\beta_1 + ... + c_n\beta_n \tag{6}$$

Now that wasn't so hard, was it? This is the equation we will use to create out contradiction. With a particularly clever choice of polynomial, f, due to Hermite, we can eventually show that the RHS is a non-zero integer, while the right can be constrained to have a magnitude less than 1.

So now that the groundwork is laid out, we can see the genius of Hermite in his choice of our polynomial. Consider a prime number, p (why it must be prime will be apparent later), such that $p > n, c_0$. Since both $nandc_0$ are finite, the infinitude of the prime numbers guarantees that such a prime exists. Then we define,

$$f(x) = \frac{x^{p-1}(1-x)^p(2-x)^p...(n-x)^p}{(p-1)!} \tag{7}$$

Just as a reminder we will recall that we call the degree of this polynomial to be r.

Our task is now to begin the contradiction by proving the LHS of (6) is a non-zero integer. We start by proving it is an integer.

This can be shown by demonstrating that for every m such that $m = 1, 2, ..., n$ we will have $F(m)$ be divisible by p, implying that $C_m F(m)$ is an integer.

Lemma 2 *If g is a polynomial with integer coefficients and*

$$h(x) = \frac{g(x)}{(p-1)!},$$

then for $i \geq p$, $h^{(i)}(x)$ is a polynomial with integer coefficients, each of which is divisible by p.

To save space, I will not provide a full proof of this result, rather a firm intuition. Nevertheless, it is an easy result to grasp conceptually. Essentially, all terms of order $i - 1$ or lower will disappear, while all terms of order i or greater will only need a factor of $p!$ to guarantee all coefficients will be an integer divisible by p, which will come by repeated differentiation. g is of the form

$$c_0 + c_1 x + ... + c_i x^i + c_{i+1} x^{i+1} + ...c_n x^n$$

Differentiating i times,

$$c_i i! + c_{i+1} \frac{(i+1)!}{1!} x + c_{i+2} \frac{(i+2)!}{2!} x^2 + ... + c_n \frac{(n!)}{(n-i)!} x^{n-i}$$

Since $n > i$ and $i \geq p$, we know that $p < n$. We can see this implies that there is a factor of $p!$ on every term. When we divide by $(p-1)!$, we can rewrite $p! = p(p-1)!$ to cancel the denominator, making each term an integer divisible by p. That is my best intuitive explanation of the lemma, if you have any questions feel free to comment but understanding of this is not of any particular importance to the proof, only the application.

Let's move right along and apply this to our function. We note that, is $i \geq p$, for any integer m we will have $f^{(i)}(m)$ is also an integer divisible by p (since each term of the polynomial expansion is an integer multiple of p, and we can factor it out). So,

$$F(m) = f^{(1)}(m) + f^{(2)}(m) + ... + f^{(r)}(m)$$

We notice that for $m = 1, 2, ..., n$, m is a root of multiplicity p to our function f. This is where we will introduce the third and final lemma to this proof.

Lemma 3 *Suppose m is a root of multiplicity f with multiplicity p. Then we have*

$$g^{(k)}(m) = 0 \ \forall k = 0, 1, ..., p-1$$

Since this proof is already so long, this lemma's will be provided without proof as it is a very simple concept. Instead, I offer an example (Ah, yes. The most rigorous form of proof).

If $f(x) = x^3$, we notice 0 is a root of multiplicity 3. Furthermore, the first and second derivatives evaluated at 0 yield 0, while the third is equal to 6.

We can use this to simplfy the expression for $F(m)$, since this lemma implies $f^{(1)}(m), f^{(2)}(m), ..., f^{(p-1)}(m) = 0$. Therefore

$$F(m) = f^{(p)}(m) + ... + f^{(r)}(m)$$

And since each of these terms are integers divisible by p as we demonstrated using Lemma 2, $F(m)$ is an integer divisible by p. Now when we look at 6,

$$c_0F(0) + \underbrace{c_1F(1) + c_2F(2) + ... + c_nF(n)}_{=\text{An integer divisible by } p} = c_1\beta_1 + ... + c_n\beta_n \tag{8}$$

This accounts for everything except the $F(0)$ term. Now, if that term is an integer then the entire LHS is certainly an integer, but how do we show that it is non-zero (which will be of critical importance)? Well, if the $c_0F(0)$ term is not divisible by p, then the expression cannot be 0. Think about it this way, for the expression to be 0, then

$$c_0F(0) = -\left(c_1F(1) + c_2F(2) + ... + c_nF(n)\right)$$

However, this implies that the LHS and RHS are both divisible by p, which is an obvious contradiction to our assumption. Furthermore, since we chose p such that $p > c_0$, we know that c_0 cannot be divisible by p. So to prove that the LHS of 6 is a non-zero integer, we simply must show that $F(0)$ is an integer not divisible by p.

If $m = 0$, 0 is a root of multiplicity $p - 1$ in our polynomial f. By Lemma 3,

$$f^{(1)}(m) = f^{(2)}(m) = ... + f^{(p-2)}(m) = 0$$

So,

$$F(0) = f^{(p-1)}(m) + \underbrace{f^{(p)}(m) + ... + f^{(r)}(m)}_{\text{All integers divisible by p}}$$

What about the first term? Recall that

$$f(x) = \frac{x^{p-1}(1-x)^p(2-x)^p...(n-x)^p}{(p-1)!}$$

If we expand out,

$$f(x) = \frac{(n!)^px^{p-1}}{(p-1)!} + \frac{a_0x^p}{(p-1)!} + \frac{a_1x^{p+1}}{(p-1)!} + ...$$

Now, when we differentiate $p - 1$ times all of the higher order terms will still contain a power of x, so when we plug in $x = 0$ they will all disappear.

Therefore, $f^{(p-1)}(0)$ is just the derivative of the first term evaluated at 0. Then we have,

$$f^{(p-1)}(0) = (n!)^p$$

Now, since $n < p$, (as we have chose), that means that there is no factor of $n!$ that is equal to p. Furthermore, no product of the factors of n can equal p since we have chosen p to be prime (finally it all makes sense!). Thus, this term is an integer not divisible by p, meaning that the entire LHS of (6) is an integer not divisible by p, which we have already established implies it is also nonzero. So,

$$\underbrace{c_0F(0) + c_1F(1) + c_2F(2) + ... + c_nF(n)}_{\text{=A nonzero integer}} = c_1\beta_1 + ... + c_n\beta_n$$

Now, if we can prove

$$|c_1\beta_1 + ... + c_n\beta_n| < 1$$

then we have an obvious contradiction because the RHS is either 0 or not an integer, and the LHS is a non-zero integer. Lets do it

Recall

$$\begin{aligned}\beta_k &= -ke^{k(1-a_k)}f(ka_k)\\ &= -ke^{k(1-a_k)}\frac{(ka_k)^{p-1}(1-ka_k)^p(2-ka_k)^p...(n-ka_k)^p}{(p-1)!}\end{aligned}$$

We know that $0 < a_k < 1$ and $k \leq n$, so we can set up the useful inequalities that $e^{k(1-a_k)} < e^n$, $k(ka_k)^{p-1} < n^p$, and finally $(1-ka_k)^p(2-ka_k)^p...(n-ka_k)^p < (n!)^p$

Therefore we can bound the magnitude of β_k as a function of p such that

$$\begin{aligned}\lim_{p\to\infty}|\beta_k| &< \lim_{p\to\infty}\frac{e^n n^p(n!)^p}{(p-1)!}\\ &= e^n\lim_{p\to\infty}\frac{(n\cdot n!)^p}{(p-1)!}\\ &= 0\end{aligned}$$

Perfect! We are in the home stretch. Using the triangle inequality,

$$|c_1\beta_1 + ... + c_n\beta_n| < |c_1\beta_1| + |c_2\beta_2| + ... |c_n\beta_n|$$

Now if we choose a prime number large enough so that each $c_i\beta_i < 1/n$, then we will reach a satisfying conclusion? How can we do this? Well think of it as an epsilon delta proof, since the limit approaches 0 we can find a delta for any epsilon, and we simply want the value to be within $1/n$ of the limit, which is completely justifiable. Now, this will imply

$$\begin{aligned}|c_1\beta_1 + ... + c_n\beta_n| &< \underbrace{\frac{1}{n} + ... + \frac{1}{n}}_{\text{n times}}\\ &= 1\end{aligned}$$

Hence, a contradiction. e cannot be the solution to a polynomial with integral coefficients, so it must be transcendental.