TRANSCENDENTAL NUMBER THEORY

HARIDAS CHOWDHURY

Abstract. This paper will start of with showing that transcendental numbers exist through Cantor's theorem as well as the origins of proving a number's transcendence with Lindemann's Theorem. Then, it will cover the transcendence of some well-known constants as well as show a theorem that will generalize them, specifically Lindemann-Weierstrass Theorem. It will also cover some of the implications of the transcendence of π in the solution to the problem of squaring the circle. It will also discuss Gelfond-Schneider's Theorem (which solved Hilbert's seventh problem) as well as a proof of Baker's Theorem, which generalizes this. This paper will then end off with a brief discussion of some of the more recent results in the area as well as some prominent problems that remain open.

1. Introduction

The search for numbers has been a quest in mathematics that dates back to ancient civilizations. We will go over a brief history here (see Cox [Cox04] for a more detailed history). The most basic level of numbers were natural numbers. However, with the introduction of addition and subtraction, it became necessary to expand this notion to integers. However, with multiplication and division, the notion of numbers was once again extended to include rational numbers. Multiplication and division also created the necessity for radicals, such as $\sqrt{2}$, which is a solution to $x^2 = 2$. This is an example of an irrational number (although we will not prove this here), which made it necessary to expand the set to the set of all real numbers.

However, there will still some equations which did not have any numeric solutions under the set of real numbers, even with this expanded notion, such as $x^2 + 1 = 0$, which prompted the discovery of imaginary numbers, thus expanding our set of numbers to now include all complex numbers. These types of numbers were found by finding solutions to these types of polynomial equations, and such numbers are called algebraic numbers.

Definition 1.1. An *algebraic number* is a number that is a root of some polynomial with integer coefficients.

However, mathematicians then began to wonder about the numbers that were not algebraic. In fact, it was not even known that they exist for a very long time. However, through a series of theorems and corollaries we will prove why such non-algebraic numbers, also known as transcendental numbers, exist.

Definition 1.2. A *transcendental number* is a number that is not a root of any polynomial with integer coefficients.

Theorem 1.3 (Cantor)**.** *The set of real numbers is uncountable.*

Proof. For the sake of contradiction, let's assume that the set of real numbers is countable. Then, we know there is some sequence of s_i such that

$$s_1 = 11100\ldots$$
$$s_2 = 11011\ldots$$
$$\ldots$$

where the subscript of the s_i are the natural numbers corresponding to some real number, which are expressed in binary. However, we can construct some number s whose nth digit is different than the nth digit of s_n, which means s is not part of the sequence s_i. However, we know that s is still a real number, but is not part of the constructed one-to-one correspondence, which is a contradiction to our assumption that the set of real numbers is countable.

Thus, the set of real numbers must be uncountable. ■

From this theorem, we conclude the following relatively direct corollary.

Corollary 1.4 (Cantor)**.** *The set of all complex numbers is uncountable.*

Proof. We know the set of real numbers is a subset of the set of complex numbers. By Theorem 1.1, we know that this subset is uncountable, which means that the entire set of complex numbers must also be uncountable. ■

Theorem 1.5 (Cantor)**.** *The set of algebraic numbers is countable.*

Proof. Assume $h = n + |a_1| + |a_2| + |a_3| + \cdots + |a_{n+1}|$ for some integer n and some sequence of integers a_i, where $a_1 \neq 0$. Given some arbitrary value of h, we can choose a finite number of ordered partitions of positive integers. Then, we assign the first number of the partition to be n, and for the mth number in the partition, we assign it to be $|a_{m-1}|$. We can then construct a polynomial with degree n and coefficients in the a_1. An example of one such partition would be if we took $h = 5 = 3 + 1 + 0 + 0 + 1$, which would correspond to the 8 polynomials $\pm 3x^3 \pm x^2 \pm 1$.

We know that the first number in the partition limits the number of terms, and that that there are only $h + 1$ possibilities for the value of n for any given h. Additionally, for any of these partitions, there are a most of 2^{n+1} possible corresponding polynomials, with each polynomial having at most n roots, all of which are algebraic by definition. Thus, for every value of h, there are a finite number of corresponding algebraic numbers. Note that since we are considering all possible polynomials with integer coefficients, we are also considering all possible algebraic numbers.

Let's define b_h as the number of distinct algebraic numbers we can get from some given h using the process above. Then, we can make each algebraic number that we get from $h = 1$ to correspond to a distinct number from 1 to b_1 and in general, assign each algebraic number from each height h to a distinct number from $1 + b_1 + b_2 + \cdots b_{h-1}$ to $b_1 + b_2 + \cdots b_h$. Thus, we have constructed a one-to-one correspondence between the algebraic numbers and the natural numbers, meaning the set of all algebraic numbers is countable. ■

Corollary 1.6 (Cantor)**.** *The set of transcendental numbers is uncountable.*

Proof. From 1.4 and 1.5, we know that the set of algebraic numbers is countable, which means it is a very small subset of the set of all complex numbers, which is uncountable. So, the remainder of the set of complex numbers must then be uncountable. However, since every non-algebraic real number is defined as transcendental, this means that the set of all transcendental numbers is uncountable. ■

So, now we know that not only do transcendental numbers exist, but they in fact make up the vast majority of all numbers, which further highlights their importance and makes them more intriguing. While we now know that transcendental numbers exist, showing that any given number is transcendental or algebraic is more difficult, and many techniques of doing so will be explored throughout this paper. However, one of the earliest methods of doing so was developed by Liouville from the definitions of algebraic and transcendental numbers. However, before we prove his theorem, let us first introduce the notion of a degree of an algebraic number.

Definition 1.7. The *degree of* α is the degree of the lowest degree polynomial with integer coefficients that has α as a root.

Now that we are equipped with this, let us take a look at Liouville's theorem.

Theorem 1.8 (Liouville)**.** *For every algebraic number α with degree $n > 1$, there exists some c such that*

$$\left|\alpha - \tfrac{p}{q}\right| > \tfrac{c}{q^n}$$

for all rationals $\frac{p}{q}$.

Proof. Since α is algebraic, we know it must be the root of some polynomial with degree n and integer coefficients $p(x)$. Then, by the mean value theorem, we have

$$p(\alpha) - p\left(\tfrac{p}{q}\right) = \left(\alpha - \tfrac{p}{q}\right) p'(\xi)$$

for some ξ between α and $\frac{p}{q}$. However, we know that $p(\alpha) = 0$. Substituting this value and taking the abolsute value of both sides gives us

$$\left|p\left(\tfrac{p}{q}\right)\right| = \left|\alpha - \tfrac{p}{q}\right| |p'(\xi)|$$

Note that if $\left|\alpha - \frac{p}{q}\right| > 1$, showing the result becomes trivial, as any $c < q^n$ would be sufficient. Thus, we can now consider the case where $\left|\alpha - \frac{p}{q}\right| \leq 1$, which would mean that $|\xi| < 1 + |\alpha|$. Since this limits the range of ξ, we know that there exists some constant c such that $|p'(\xi)| < \frac{1}{d}$ for some positive d. Thus, using our previous equation, we know that

$$\left|p\left(\tfrac{p}{q}\right)\right| < \tfrac{\left|\alpha - \frac{p}{q}\right|}{d}.$$

Multiplying both sides by d yields

$$d \cdot \left|p\left(\tfrac{p}{q}\right)\right| < \left|\alpha - \tfrac{p}{q}\right|.$$

Since polynomial $p(x)$ has integer coefficients, we know that $p\left(\frac{a}{q}\right)$ is some rational number with denominator q^n, which means

$$d \cdot \left|\tfrac{m}{q^n}\right| < \left|\alpha - \tfrac{p}{q}\right|.$$

Substituting in $c = d|m|$ for some constant c gives

$$\tfrac{c}{|q^n|} < \left|\alpha - \tfrac{p}{q}\right|,$$

which implies

$$\tfrac{c}{q^n} < \left|\alpha - \tfrac{p}{q}\right|,$$

allowing us to obtain our desired result. ■

Using this theorem, Liouville was able to construct one of the first known transcendental numbers, known as Liouville's constant. This number is

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

Intuitively, we see that Liouville's Theorem essentially states that any algebraic number can not be estimated too well by rational numbers. So, if the estimate with a rational number is too close, the number is transcendental. We see that with Liouville's constant, the power of 10 that we add becomes increasingly smaller, and it can not be approximated too well by the rationals. The proof of this constant's transcendence will not be concretely proven here, but it is directly a result of Liouville's theorem. However, we will now look at how the field of transcendental number theory applies to more well known constants and other fields of mathematics. Additionally, we will also look at a proof of Baker's Theorem and also discuss some of the more recent research done in the field as well as some unsolved problems in it.

2. Transcendence of e and π

While e was proven to be irrational by Euler, its transcendence remained a mystery until it was proven by Hermite.

Theorem 2.1 (Hermite)**.** *e is transcendental.*

Proof. To start, let's take $f(x)$ to be any polynomial with real coefficients and degree m. Then, we can define some function $I(t)$ as

$$I(t) = \int_0^t e^{t-u} f(u)du.$$

Using integration by parts, we get

$$\begin{aligned} I(t) &= \int_0^t e^{t-u} f(u)du \\ &= -e^{t-u} f(u)\Big|_0^t - \int_0^t -e^{t-u} f'(u)du \\ &= -e^0 f(t) + e^t f(0) + \int_0^t e^{t-u} f'(u)du \\ &= -f(t) + e^t f(0) + \int_0^t e^{t-u} f'(u)du. \end{aligned}$$

However, we see that the last term is much like our initial integral, but with the $f(u)$ replaced by $f'(u)$, which allows us to make a similar computation:

$$\begin{aligned} I(t) &= -f(t) + e^t f(0) + \int_0^t e^{t-u} f'(u)du \\ &= -f(t) + e^t f(0) - f'(t) + e^t f'(0) + \int_0^t e^{t-u} f''(u)du \\ &= -(f(t) + f'(t)) + e^t(f(0) + f'(0)) + \int_0^t e^{t-u} f''(u)du. \end{aligned}$$

Continuing this process indefinitely, we get

$$I(t) = -(f(t) + f'(t) + f''(t) + \cdots) + e^t(f(0) + f'(0) + f''(0) + \cdots).$$

However, since $f(x)$ has degree m, we know that every derivative after the mth derivative is equal to 0. Thus, we can further simplify our expression of $I(t)$ as

$$\begin{aligned} I(t) &= -(f(t) + f'(t) + f''(t) + \cdots + f^{(m)}(t)) + e^t(f(0) + f'(0) + f''(0) + \cdots + f^{(m)}(0)) \\ &= -\sum_{i=0}^{m} f^{(i)}(t) + e^t \sum_{k=0}^{m} f^{(k)}(0), \end{aligned}$$

where $f^{(j)}(x)$ denotes the jth derivative of $f(x)$. Additionally, if $\bar{f}(x)$ denotes the polynomial obtained by replacing all of the coefficients of f with their absolute values, then we have

$$|I(t)| \leq \left| \int_0^t e^{t-u} f(u)du \right| \leq |t|e^{|t|}\bar{f}(|t|),$$

which can be derived by observing the integral geometrically.

Now, for the sake of contradiction, let's assume that e is algebraic, in which case

$$q_0 + q_1 e + \cdots q_n e^n = 0$$

for some integers $n > 0, q_0 \neq 0, q_1, q_2, ..., q_n$. Now, let's define J as

$$J = q_0 I(0) + q_1 I(1) + \cdots q_n I(n),$$

where $I(t)$ is defined as it previously was with

$$f(x) = x^{p-1}(x-1)^p(x-2)^p \cdots (x-n)^p$$

for some large prime p. Then, we have

$$\begin{aligned}
J &= \sum_{j=0}^{n} q_j I(j) \\
&= \sum_{j=0}^{n} q_j \left(-\sum_{i=0}^{m} f^{(i)}(j) + e^j \sum_{k=0}^{m} f^{(k)}(0) \right) \\
&= -\left(\sum_{j=0}^{n} q_j \right) \left(\sum_{i=0}^{m} f^{(i)}(j) \right) + \left(\sum_{j=0}^{n} q_j e^j \right) \left(\sum_{k=0}^{m} f^{(n)}(0) \right) \\
&= -\sum_{j=0}^{n} \sum_{i=0}^{m} q_j f^{(i)}(j) + 0 \cdot \sum_{k=0}^{m} f^{(n)}(0) \\
&= -\sum_{j=0}^{n} \sum_{i=0}^{m} q_j f^{(i)}(j),
\end{aligned}$$

where $m = (n+1)p - 1$. Now, let's observe $f^{(i)}(j)$ more carefully. If we defined $g(x) = (x-1)(x-2)\cdots(x-n)$ and $h(x) = x^{p-1}$ we can rewrite $f(x) = h(x)g(x)^p$. If we take the derivative of this using the product rule, we have $f'(x) = h'(x) \cdot g(x)^p + h(x) \cdot (pg(x)^{p-1} \cdot g'(x))$. If we continue this process, we see that all of the terms in $f^{(i)}(j)$ will always have $g(j)^{p-i}$ as a factor. Then, for any $i < p$, all of the integers 1 through n will be a root of $f^{(i)}(j)$, which means that for any $j > 0$ in this case, $f^{(i)}(j) = 0$, since j can only go up to n in our summation above.

We also observe that since the lowest degree term in $f(x)$ has degree $p-1$ and every time we take the derivative, the exponent on this term decreases by 1, which means that for every $i < p-1$, $f^{(i)}(j)$ will have a factor of x, meaning for every $i < p-1$ and $j = 0$, $f^{(i)}(j) = 0$.

So, we know that $f^{(i)}(j) = 0$ if $i < p$ and $j > 0$ as well as if $i < p-1$ and $j = 0$. Then, all remaining terms in the expansion of J all have $i \geq p$, except for when $i = p-1$ and $j = 0$. In order to address the former case, let's once again consider what happens when we repeatedly take the derivative of $f(x)$. We notice that every nonzero term has $p!$ as a factor. However, for the remaining term, we notice that

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

Since $p!$ is a multiple of $(p-1)!$, J must be an integer multiple of $(p-1)!$ but not of $p!$, which means that we have $|J| \geq (p-1)!$. However, using our previous bound for $|I(t)|$ and our definition of J, we get that

$$|J| \leq \sum_{k=0}^{n} q_k |k| e^{|k|} \bar{f}(|k|).$$

Now, looking at $f(x)$, we see that the absolute value of each term in the expansion is at most n^m, and there are 2^m terms in the expansion, meaning $\bar{f}(|k|) \leq (2n)^m$. Now, let's take C to be a constant independent of p such that C is the maximum value of $q_k|k|e^{|k|}$ when k goes from 0 to n. Using this as well as the fact that $m = (n+1)p - 1$,

$$|J| \leq \sum_{k=0}^{n} q_k |k| e^{|k|} \bar{f}(|k|) \leq (n+1) \cdot C \cdot (2n)^m.$$

This is clearly an exponential function, which means we can crudely approximate it as

$$|J| \leq c^p$$

for some constant c independent of p. However, we also previously found the bound $|J| \geq (p-1)!$, which means that we must have $c^p \leq (p-1)!$. However, we can choose a sufficiently large p so that this is not true. This contradiction allows us to complete our proof that e is transcendental. ■

Definition 2.2. A *minimal polynomial* of an algebraic number α is the irreducible polynomial relatively prime integer coefficients, with the leading coefficient being positive, with lowest degree that has α as a root. Its roots are called its conjugates of α.

Now, another natural constant to wonder about is π, which is what Lindemann decided to do, as shown below.

Theorem 2.3 (Lindemann)**.** *π is transcendental.*

Proof. For the sake of contradiction, let's assume that π is algebraic. Then, this must mean that $\theta = \pi i$ is also algebraic. Let θ have degree d, let the conjugates be θ_1 (which we can define as θ), $\theta_2, \theta_3, \ldots, \theta_d$, and let l be the leading coefficient of the minimal polynomial. Then, since $1 + e^{\theta_1} = 1 + e^{\pi i} = 0$, we have that

$$(1 + e^{\theta_1})(1 + e^{\theta_2}) \cdots (1 + e^{\theta_d}) = 0$$

If we expand the product on the left, we get 2^d terms of the form where e^Θ where

$$\Theta = \epsilon_1\theta_1 + \epsilon_2\theta_2 + \cdots + \epsilon_d\theta_d$$

where all of the ϵ_j are either 0 or 1. Let's suppose that exactly n of the Θ are non-zero, which we can denote as $\alpha_1, \alpha_2, \ldots, \alpha_n$. Then, since there are 2^d total terms, this means that $2^d - n$ of them are 0. So, when we expand our previous product, we have

$$(2^d - n)e^0 + e^{\alpha_1} + e^{\alpha_2} + \cdots e^{\alpha_n} = 0.$$

Using the fact that $e^0 = 1$ and substituting in $q = 2^d - n$, we get

$$q + e^{\alpha_1} + e^{\alpha_2} + \cdots e^{\alpha_n} = 0,$$

from which we have

$$e^{\alpha_1} + e^{\alpha_2} + \cdots e^{\alpha_n} = -q.,$$

Now, let's define

$$J = I(\alpha_1) + I(\alpha_2) + \cdots I(\alpha_n),$$

where $I(t)$ is defined as in Theorem 2.1 with

$$f(x) = l^{np}x^{p-1}(x-\alpha_1)^p(x-\alpha_2)^p \cdots (x-\alpha_n)^p,$$

where p is a large prime. Recall that our simplified version of $I(t)$ was

$$I(t) = -\sum_{i=0}^{m} f^{(i)}(t) + e^t \sum_{k=0}^{m} f^{(k)}(0).$$

So, we have that

$$\begin{aligned}
J &= \sum_{j=1}^{n} I(\alpha_j) \\
&= -\sum_{j=1}^{n}\sum_{i=0}^{m} f^{(i)}(\alpha_j) + \sum_{j=1}^{n}\left(e^{\alpha_j}\sum_{k=0}^{m} f^{(k)}(0)\right) \\
&= -\sum_{j=1}^{n}\sum_{i=0}^{m} f^{(i)}(\alpha_j) + \left(\sum_{k=0}^{m} f^{(k)}(0)\right)\cdot\left(\sum_{j=1}^{n} e^{\alpha_j}\right) \\
&= -\sum_{i=0}^{m}\sum_{j=1}^{n} f^{(i)}(\alpha_j) - q\sum_{k=0}^{m} f^{(k)}(0),
\end{aligned}$$

where $m = (n+1)p - 1$, since that's the degree of the polynomial.

Let's take a closer look at the term

$$\sum_{j=1}^{n} f^{(i)}(\alpha_j)$$

We know this sum is symmetric over the $l\alpha_j$ and has integer coefficients. Then, using the Fundamental Theorem on Symmetric Functions, we know that it can be expressed as a polynomial with integer coefficients in the elementary symmetric functions of the $l\alpha_j$. However, we know that any function symmetric over the $l\alpha_j$ is also symmetric over all of the θ_j, as the α_j include all possible permutations of the summations of θ_j. Then, we know that the sum above is symmetric over the θ_j as well and can thus be represented as a polynomial over the elementary symmetric functions of the θ_j. We know that the θ_j are roots of a polynomial with integer coefficients, by definition. So, by Vieta's forumulas, all of the elementary symmetric functions over the θ_j are integers. Thus, our summation above is an integer.

Now, going back to our expression for J, we can use similar logic as we did in our proof of Theorem 2.1 to determine that the first term is divisible by $p!$ and thus $(p-1)!$. We can also use our logic from the same proof to determine that $f^{(j)}(0)$ is divisible by $p!$, and thus $(p-1)!$ for all $j \neq p-1$. When $j = p-1$, we know that

$$f^{(p-1)}(0) = (p-1)!(-l)^{np}(\alpha_1\alpha_2\cdots\alpha_n),$$

which is clearly divisible by $(p-1)!$, just like all of the other terms. Thus, we know that J is an integer divisible by $(p-1)!$ and thus $|J| \geq (p-1)!$. Additionally, we can use similar reasoning as we did in our proof of Theorem 2.1 to show that $|J| \leq c^p$ for some c independent of p as well. This would imply that $(p-1)! \leq c^p$, but this statement is not true for sufficiently large p. This creates a contradiction and thus proves our proof that π is transcendental. ■

This proof actually solved the very old problem of squaring the circle, which questioned if it was possible to construct a square with the exact same area as a given circle. Since π is transcendental and such a construction would require the construction of π, this was shown to be impossible.

Later on, Lindemann generalized the proof of the transcendence of both of these numbers, and this was later rigorously proved by Weirstrass, as shown below.

Theorem 2.4 (Lindemann and Weierstrass). *For any distinct algebraic numbers* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *and any non-zero algebraic numbers* $\beta_1, \beta_2, \ldots, \beta_n$, *we have*

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

Proof. For the sake of contradiction, let's assume that this is not true and that

$$\beta_1 e^{\alpha_1} + \beta_2 e^{\alpha_2} + \cdots + \beta_n e^{\alpha_n} = 0$$

Now, let's take the conjugates of each β_j to be $\beta_{j,1}$(which we can set equal to β_j), $\beta_{j,2}, \ldots, \beta_{j,k_j}$, where k_j is the degree of the minimal polynomial of β_j. Then, we consider terms of the following form:

$$\beta_{1,j_1} e^{\alpha_1} + \beta_{2,j_2} e^{\alpha_2} + \cdots + \beta_{n,j_n} e^{\alpha_n}$$

where each $j_i \leq k_i$. Now, if we take the product of all possible permutations of this, we will have a expression that's symmetric over the β_{k,j_i} for each constant k from 1 to n. This means that we can express our expression as a polynomial of elementary symmetric functions over the β_{k,j_i} for each constant k. Using Vieta's formulas and the fact that they are constants, we know that the coefficients would be integers, and would be of the form

$$B_1 e^{a_1} + B_2 e^{a_2} + \cdots B_n e^{a_n}$$

for all B_i being integers. Notice that this is in the same form as the expression $\beta_1 e^{\alpha_1} + \beta_2 e^{\alpha_2} + \cdots + \beta_n e^{\alpha_n}$. In fact, we know the expression is equal to 0, as $\beta_1 e^{\alpha_1} + \beta_2 e^{\alpha_2} + \cdots + \beta_n e^{\alpha_n}$ is one of the factors. Thus, without loss of generality, we can assume that $\beta_1, \beta_2, \ldots \beta_n$ are integers. Note that one of these integers has to be non-zero by definition.

Now, we let l be some positive integer and we let

$$f_i(x) = l^{np} \left((x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n) \right)^p / (x - \alpha_i),$$

where p is some large prime. The function $f_l(x)$ is defined in such a way that it involves the algebraic integers corresponding to the numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$. It is crucial to note that the behavior of this function as p increases gives rise to significant properties of the coefficients in the original equation. We now examine how these terms behave as p increases.

We define the quantities J_i for each i as follows:

$$J_i = \beta_1 L_i(\alpha_1) + \beta_2 L_i(\alpha_2) + \cdots + \beta_n L_i(\alpha_n) \quad \text{for each } i \quad (1 \leq i \leq n),$$

where $L_i(t)$ is the function defined in the proof of Theorem 1.2, with $f = f_t$. These quantities J_i play a pivotal role in our analysis.

We now arrive at the contradiction. From the properties of the function $f_l(x)$ and the algebraic integers involved, we find that the quantities $J_1, J_2, \ldots, J_n$ are rational integers that satisfy the following relationship:

$$|J_1 \ldots J_n| \leq \sum_{k=1}^{n} |\alpha_k \beta_k| e^{|a_k|} \leq cp,$$

where c is a constant independent of p. The key point here is that the inequalities are inconsistent if p becomes large enough, as we can use similar techniques as we did above to show that $|J_1 \cdots J_n| > (p-1)!$ by showing its divisibility. This inconsistency leads to a contradiction, thereby proving that our original assumption was false. ∎

3. Baker's Theorem

We begin with a theorem that solved Hilbert's seventh problem:

Theorem 3.1 (Gelfond and Schneider). *a^b is transcendental for any algebraic number a that is not 0 or 1 and for any irrational algebraic number b.*

We will not prove this here (see Baker [Bak75] for a proof), but we will discuss a generalization of it, namely Baker's Theorem, which is stated below.

Theorem 3.2 (Baker). *Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be non-zero algebraic numbers such that their natural logarithms are linearly independent over rational numbers. Then, for any non-zero algebraic numbers $\beta_1, \beta_2, \ldots, \beta_n$,*

$$\beta_1 \log \alpha_1 + \beta_2 \log \alpha_2 + \cdots + \beta_n \log \alpha_n$$

is transcendental.

The rest of this section presents a detailed proof of **Baker's Theorem**. We begin by assuming, for contradiction, that the theorem is false. Under this assumption, there exist algebraic numbers $\beta_0, \beta_1, \ldots, \beta_n$, not all zero, such that:

$$\beta_0 + \beta_1 \log \alpha_1 + \ldots + \beta_n \log \alpha_n = 0,$$

where $\alpha_1, \ldots, \alpha_n$ are algebraic.

Since not all β_i vanish, we may assume $\beta_n \neq 0$. Dividing by β_n, we define $\beta_j' = -\beta_j/\beta_n$ for $j = 0, \ldots, n-1$, and rewrite the equation as:

$$\beta_0 + \beta_1 \log \alpha_1 + \ldots + \beta_{n-1} \log \alpha_{n-1} = \log \alpha_n,$$

which exponentiates to:

$$e^{\beta_0} \alpha_1^{\beta_1} \cdots \alpha_{n-1}^{\beta_{n-1}} = \alpha_n. \tag{1}$$

This identity will be the basis for the argument to follow.

Throughout the discussion, constants $c, c_1, c_2, \ldots$ denote positive quantities depending only on the α_i, β_j, and the branches of the logarithms. They are independent of any variable parameters introduced later. We also fix a sufficiently large integer h, chosen to exceed such constants when needed.

We record the following elementary fact. If α is algebraic and satisfies

$$A_0 \alpha^d + A_1 \alpha^{d-1} + \ldots + A_d = 0,$$

with $A_i \in \mathbb{Z}$ and $|A_i| \leq A$, then for any non-negative integer j, we can write:

$$(A_0 \alpha)^j = A_0^{(j)} + A_1^{(j)} \alpha + \ldots + A_{d-1}^{(j)} \alpha^{d-1},$$

with $A_m^{(j)} \in \mathbb{Z}$ and $|A_m^{(j)}| \leq (2A)^j$. This follows by induction using recurrence:

$$A_m^{(j)} = A_0 A_m^{(j-1)} - A_d \cdot m \cdot A_d^{(j-1)}, \quad 0 \leq m < d,\ j \geq d,$$

initialized with $A^{(j-1)} = 0$ where necessary.

Let d be the maximum degree of the minimal polynomials of $\alpha_1, \ldots, \alpha_n, \beta_0, \ldots, \beta_n$, and let a_i, b_j be their leading coefficients. Then, for any integer j, we can write:

$$(\alpha_r \alpha_s)^j = \sum_{s=0}^{d-1} a_s^{(j)} \alpha_s^r, \qquad (b_r \beta_r)^j = \sum_{t=0}^{d-1} b_t^{(j)} \beta_t^r, \tag{2}$$

with $|a_s^{(j)}|, |b_t^{(j)}| \le c_1^j$ for some constant c_1.

To simplify multivariable differentiation, we use the notation:

$$f_{m_0,\ldots,m_{n-1}}(z_0,\ldots,z_{n-1}) = \left(\frac{\partial}{\partial z_0}\right)^{m_0}\left(\frac{\partial}{\partial z_1}\right)^{m_1}\cdots\left(\frac{\partial}{\partial z_{n-1}}\right)^{m_{n-1}} f(z_0,\ldots,z_{n-1}),$$

where f is a sufficiently differentiable function of n complex variables. This will streamline later expressions involving partial derivatives.

Lemma 3.3. *Let M and N be integers such that $N > M > 0$, and let*

$$u_{ij} \quad (1 \le i \le M,\ 1 \le j \le N)$$

be a collection of integers with absolute values bounded by a fixed positive integer $U > 1$. Then there exists a nonzero integer vector $(x_1, x_2, \ldots, x_N)$, not all zero, such that each x_j is an integer with

$$|x_j| \le (NU)^{M(N-M)} \quad \textit{for all } j,$$

and this vector satisfies the system of linear homogeneous equations

$$\sum_{j=1}^{N} u_{ij}x_j = 0 \quad \textit{for all } 1 \le i \le M. \tag{3}$$

Proof. The strategy is to show that by choosing bounded integer vectors $(x_1,\ldots,x_N)$, there are more such vectors than there are possible outcomes for the left-hand side of the system (3). Hence, by the pigeonhole principle, two distinct input vectors must yield the same output vector, and their difference gives a nontrivial solution to the system.

We begin by defining a bound for the x_j's. Let

$$B = \left\lfloor (NU)^{M(N-M)} \right\rfloor,$$

where $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x. We now consider all integer vectors $(x_1,\ldots,x_N)$ such that each x_j satisfies

$$0 \le x_j \le B \quad \text{for all } j = 1, 2, \ldots, N.$$

There are exactly $(B+1)^N$ such vectors, since each of the N coordinates independently takes one of $B+1$ values.

For each such vector, compute the left-hand side of each equation in (3), that is,

$$y_i = \sum_{j=1}^{N} u_{ij}x_j \quad \text{for each } i = 1,\ldots,M.$$

We now estimate the possible range of values each y_i can take.

Let us define V_i to be the sum of the negative parts of the coefficients u_{ij}, and W_i to be the sum of the positive parts. That is,

$$V_i = \sum_{\substack{1\le j\le N\\ u_{ij}<0}} |u_{ij}|, \quad W_i = \sum_{\substack{1\le j\le N\\ u_{ij}>0}} u_{ij}.$$

Since $|u_{ij}| \le U$ and there are N such terms per row, we have the bound:

$$V_i + W_i \le NU.$$

For each x_j in the range $0 \le x_j \le B$, the quantity $y_i = \sum_j u_{ij} x_j$ satisfies:

$$-V_i B \le y_i \le W_i B.$$

Therefore, for each i, the number of possible values of y_i is at most $V_i B + W_i B + 1 \le NUB + 1$. Since there are M such values $y_1, \dots, y_M$, the total number of distinct M-tuples $(y_1, \dots, y_M)$ that can be produced by these equations is at most:

$$(NUB + 1)^M.$$

However, recall that we started with $(B+1)^N$ possible choices for $(x_1, \dots, x_N)$. We now compare the number of input vectors to the number of output vectors. If

$$(B+1)^N > (NUB + 1)^M,$$

then by the pigeonhole principle, there must be at least two distinct vectors $\mathbf{x} \neq \mathbf{x}'$ such that they both produce the same vector $\mathbf{y} = (y_1, \dots, y_M)$ under the linear system (3). That is, the difference $\mathbf{x} - \mathbf{x}'$ lies in the kernel of the matrix (u_{ij}).

Now observe that:

$$(B+1)^{N-M} > (NU)^M \quad \Rightarrow \quad (B+1)^N > (NU(B+1))^M \ge (NUB+1)^M,$$

which shows the inequality above holds when $B = \lfloor (NU)^{M(N-M)} \rfloor$. Therefore, such a nonzero integer solution must exist.

Taking the difference between the two distinct vectors that produce the same left-hand side yields a nontrivial solution to the system (3). Moreover, since each of the original vectors had entries in the range $[0, B]$, their difference results in a vector with entries bounded in absolute value by $B \le (NU)^{M(N-M)}$.

This completes the proof. ∎

Lemma 3.4. *There exist integers $p(\lambda_0, \dots, \lambda_n)$, not all zero, with absolute values bounded above by e^{h^3}, such that the auxiliary function*

$$\Phi(z_0, \dots, z_{n-1}) = \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_0, \dots, \lambda_n) z_0^{\lambda_0} e^{\lambda_0 \alpha_0} \alpha_1^{\lambda_1} \cdots \alpha_{n-1}^{\lambda_{n-1}}$$

satisfies the vanishing condition

$$\Phi_{m_0, \dots, m_{n-1}, l} = 0 \tag{4}$$

for all integers l in the range $1 \le l \le h$, and for all combinations of non-negative integers $m_0, \dots, m_{n-1}$ such that $m_0 + \cdots + m_{n-1} \le h^2$. Here, we define $\gamma_r = \lambda_r + \lambda_n \beta_r$ for $1 \le r < n$, and we set $L = \lfloor h^2 - 1/(4dn) \rfloor$.

Proof. In order to verify this lemma, it is sufficient—using relation (1) as a guiding principle—to construct the coefficients $p(\lambda_0, \dots, \lambda_n)$ in such a way that the following identity is satisfied:

$$\sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_0, \dots, \lambda_n) \cdot q(\lambda_0, \lambda_n, l) \cdot \alpha_1^{\lambda_1} \cdots \alpha_n^{\lambda_n} \cdot l^{\lambda_n} \cdot \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}} = 0 \tag{5}$$

for the specified ranges of l and m_r. The function $q(\lambda_0, \lambda_n, z)$ appearing in this expression is defined as

$$q(\lambda_0, \lambda_n, z) = \sum_{\mu_0=0}^{m_0} \binom{m_0}{\mu_0} \lambda_0(\lambda_0 - 1) \cdots (\lambda_0 - \mu_0 + 1)(\lambda_n \beta_0)^{m_0 - \mu_0} z^{\lambda_0 - \mu_0}.$$

This function arises from differentiating monomials and accounts for contributions from derivatives with respect to z_0.

To facilitate bounding the size of the coefficients and to make the equations integer-valued, we multiply both sides of (5) by a suitable integer factor:

$$P' = (a_1 \cdots a_n)^L \cdot b_0^{m_0} \cdots b_{n-1}^{m_{n-1}},$$

which eliminates denominators resulting from powers of algebraic numbers when substitutions are made.

Next, we express each power $\gamma_r^{m_r}$, for $1 \leq r < n$, using the binomial expansion:

$$\gamma_r^{m_r} = \sum_{\mu_r=0}^{m_r} \binom{m_r}{\mu_r} \lambda_r^{m_r-\mu_r} (\lambda_n \beta_r)^{\mu_r}.$$

Substituting these into (5), and recalling that the powers of α_r and β_r can be rewritten using identity (2), we arrive at a new expression involving only bounded powers of these algebraic quantities:

$$\sum_{s_1=0}^{d-1} \cdots \sum_{s_n=0}^{d-1} \sum_{t_0=0}^{d-1} \cdots \sum_{t_{n-1}=0}^{d-1} A(s,t) \alpha_1^{s_1} \cdots \alpha_n^{s_n} \beta_0^{t_0} \cdots \beta_{n-1}^{t_{n-1}} = 0,$$

where each $A(s,t)$ is a linear form in the $p(\lambda_0, \ldots, \lambda_n)$.

The coefficients $A(s,t)$ are explicitly given by:

$$A(s,t) = \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} \sum_{\mu_0=0}^{m_0} \cdots \sum_{\mu_{n-1}=0}^{m_{n-1}} p(\lambda_0, \ldots, \lambda_n) \cdot q' \cdot q'' \cdot q''',$$

where we define:

- $q' = \prod_{r=1}^{n} \left\{ a_r^{L-\lambda_r} \alpha_r^{(s_r)} \right\}$, which is clearly bounded by c_2^{Lh},
- $q'' = \prod_{r=1}^{n-1} \binom{m_r}{\mu_r} (b_r \lambda_n)^{m_r-\mu_r} \lambda_r^{(b_r^{(t_r)})}$,
- $q''' = \binom{m_0}{\mu_0} \lambda_0(\lambda_0 - 1) \cdots (\lambda_0 - \mu_0 + 1) \lambda_n^{m_0-\mu_0} b_0^{t_0} h_0^{\lambda_0-\mu_0}$.

The crux of the argument is now to bound the absolute value of the coefficients in each $A(s,t)$. We observe:

$$|q'| \leq c_2^{Lh},$$

$$|q''| \leq \prod_{r=1}^{n-1} (c_3 L)^{m_r},$$

$$|q'''| \leq 2^{m_0} (\lambda_0 b_0)^{\mu_0} (c_1 \lambda_n)^{m_0-\mu_0} \leq (c_3 L)^{m_0} hL,$$

using the facts that $\binom{m_r}{\mu_r} \leq 2^{m_r}$ and $l \leq h$.

Moreover, the total number of combinations of the m_r is bounded. Since each $m_r \geq 0$ and the sum of the m_r is at most h^2, we have:

$$(m_0 + 1)(m_1 + 1) \cdots (m_{n-1} + 1) \leq 2^{m_0 + \cdots + m_{n-1}} \leq 2^{h^2}.$$

Therefore, the total contribution to each coefficient in the $A(s,t)$ expressions is bounded by:

$$U = (2c_3 L)^{h^2} c_4^h.$$

Next, we count the number of linear equations and unknowns:

- The number of distinct sets $(l, m_0, \ldots, m_{n-1})$ is at most $h(h^2 + 1)^n$.

- Each such tuple leads to d^{2n} equations indexed by (s,t), so in total:
$$M \leq d^{2n} h(h^2+1)^n.$$
- On the other hand, the number of unknown coefficients is:
$$N = (L+1)^{n+1} > h^{2n+1} > 2M.$$

Because $N > 2M$, Lemma 3.3 ensures that the corresponding system of homogeneous linear equations has a nontrivial integer solution. Furthermore, the bound on the size of each coefficient of the system, along with the dimension count, allows us to conclude that there exists such a non-zero solution with:
$$|p(\lambda_0, \ldots, \lambda_n)| \leq NU \leq h^{2n+2}(2c_3L)^h c_4^h.$$

Now observe that this upper bound is of the order e^{h^3} for sufficiently large h. Indeed, since $L = \lfloor h^2 - 1/(4dn) \rfloor$, we have $L = O(h^2)$, and thus
$$NU \leq h^{2n+2}(2c_3h^2)^h c_4^h \leq e^{h^3}$$
as required.

This completes the proof of the lemma. ∎

Lemma 3.5. *Let $m_0, m_1, \ldots, m_{n-1}$ be non-negative integers such that*
$$m_0 + m_1 + \cdots + m_{n-1} \leq h^2,$$
and define the function
$$f(z) = \Phi_{m_0,\ldots,m_{n-1}}(z, z, \ldots, z). \tag{7}$$
Then, for any complex number z, the absolute value of $f(z)$ is bounded above as follows:
$$|f(z)| \leq c_6^{h^3 + L\log L},$$
where c_6 is a positive constant, and $L = \lfloor h^2 - 1/(4dn) \rfloor$ as before.

Moreover, for any positive integer l, we have the dichotomy: either $f(l) = 0$, or
$$|f(l)| > c_6^{-h^3 - L\log L}.$$

Proof. To begin, recall that the function $f(z)$ arises from evaluating the multivariate auxiliary function Φ—as constructed in Lemma 3.4—at the point $(z, \ldots, z)$, and then differentiating it appropriately with respect to its variables to the order prescribed by $(m_0, \ldots, m_{n-1})$.

By the construction from Lemma 3.4, the function $f(z)$ is given explicitly by a finite sum of the form
$$f(z) = P \sum_{\lambda_0=0}^{L} \sum_{\lambda_n=0}^{L} p(\lambda_0, \ldots, \lambda_n)\, q(\lambda_0, \lambda_n, z)\, \alpha_1^{\lambda_1} z^{\lambda_0} \cdots \alpha_n^{\lambda_n} z^{\lambda_n} \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}},$$
where:

- $p(\lambda_0, \ldots, \lambda_n)$ are the integer coefficients constructed in Lemma 3.4,
- $q(\lambda_0, \lambda_n, z)$ is a certain polynomial expression defined in that as well, capturing the effect of differentiation,
- $\gamma_r = \lambda_r + \lambda_n \beta_r$ for $1 \leq r < n$,
- P is a normalizing factor arising from the differentiation process, given by
$$P = (\log \alpha_1)^{m_1} (\log \alpha_2)^{m_2} \cdots (\log \alpha_{n-1})^{m_{n-1}}.$$

We now estimate the size of each factor appearing in this expression.

First, consider the bound for the function $q(\lambda_0, \lambda_n, z)$. By definition,

$$q(\lambda_0, \lambda_n, z) = \sum_{\mu_0=0}^{m_0} \binom{m_0}{\mu_0} \lambda_0(\lambda_0 - 1) \cdots (\lambda_0 - \mu_0 + 1)(\lambda_n \beta_0)^{m_0 - \mu_0} z^{\lambda_0 - \mu_0}.$$

Each term in this sum is bounded in absolute value, and since the binomial coefficient satisfies $\binom{m_0}{\mu_0} \leq 2^{m_0}$, and $\lambda_0, \lambda_n \leq L$, we have:

$$|q(\lambda_0, \lambda_n, z)| \leq (c_7 L)^{m_0} |z| \sum_{\mu_0=0}^{m_0} \binom{m_0}{\mu_0} \leq (2c_7 L)^{m_0} |z|,$$

for some constant c_7 depending on β_0.

Next, consider the product involving the algebraic numbers $\alpha_1, \ldots, \alpha_n$:

$$|\alpha_1^{\lambda_1} \cdots \alpha_n^{\lambda_n} z^{\lambda_0 + \lambda_n}| \leq c_8^{L \log L},$$

for some constant c_8, because each exponent λ_r is at most L, and there are at most $n+1$ such terms. Since $L \leq h^2$, we use $L \log L = O(h^2 \log h)$.

Additionally, for the product $\gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}}$, note that each γ_r is a linear combination of λ_r and $\lambda_n \beta_r$, and both λ_r and λ_n are at most L. Therefore,

$$|\gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}}| \leq (c_9 L)^{m_1 + \cdots + m_{n-1}},$$

for some constant c_9.

The number of terms in the sum is determined by how many multi-indices $(\lambda_0, \ldots, \lambda_n)$ exist with each $\lambda_r \leq L$, so the total number of terms is at most $(L+1)^{n+1} = O(h^{2(n+1)})$, which is less than h^{2n+2} for large h.

Now, using the bound from Lemma 3.4:

$$|p(\lambda_0, \ldots, \lambda_n)| \leq e^{h^3},$$

and putting all of the above estimates together, we conclude that each term in the sum is bounded in absolute value by

$$e^{h^3} \cdot (2c_7 L)^{m_0} |z| \cdot c_8^{L \log L} \cdot (c_9 L)^{m_1 + \cdots + m_{n-1}}.$$

Multiplying by the number of terms and the constant P, which only depends on the $\log \alpha_r$ and m_r, we get:

$$|f(z)| \leq h^{2n+2} \cdot e^{h^3} \cdot (c_{10} L)^{h^2} |z| \cdot c_8^{L \log L},$$

which is ultimately bounded by:

$$|f(z)| \leq c_6^{h^3 + L \log L},$$

for some sufficiently large constant c_6.

We now address the second claim. Suppose l is a positive integer and $f(l) \neq 0$. Define a new quantity

$$f' = \frac{P'}{P} f(l),$$

where P' is the normalizing factor introduced in the auxiliary function construction (see equation (6)). The quantity f' is then an algebraic number obtained by scaling $f(l)$ so that all coefficients become algebraic integers. In fact, f' is itself an algebraic integer, since both P' and P are constructed to clear denominators.

By construction, the degree of f' over $\mathbb{Q}$ is at most d^{2n}, since all the algebraic numbers involved (i.e., the α_r and β_r) lie in a number field of bounded degree.

Furthermore, each conjugate of f' (obtained by applying a Galois automorphism to all the α_r, β_r, and evaluating at the same integer l) has absolute value at most

$$|f'| \le c_{10}^{h^3+L\log L},$$

by the same bound as above applied to conjugates of $f(l)$.

Now, if $f' \neq 0$, then its algebraic norm (the product of all its Galois conjugates) is a non-zero integer and hence has absolute value at least 1. It follows that the geometric mean of the absolute values of its conjugates is also bounded below by 1, and hence the maximum must be at least

$$|f'| > c_{10}^{-(h^3+L\log L)d^{2n}}.$$

This lower bound transfers directly to $|f(l)|$, proving that

$$|f(l)| > c_6^{-h^3-L\log L},$$

for another constant c_6.

■

Lemma 3.6. *Let J be any integer such that $0 \le J \le (8n)^2$. Then the conclusion of equation (4) holds for all integers l satisfying $1 \le l \le h^{1+J/(8n)}$ and all non-negative integers $m_0, \dots, m_{n-1}$ such that*

$$m_0 + m_1 + \dots + m_{n-1} \le \frac{h^2}{2^J}.$$

Proof. We begin by establishing the base case for our induction. When $J = 0$, the conclusion follows directly from Lemma 3.4, which provides the vanishing of the auxiliary function under appropriate bounds.

Now suppose inductively that the lemma is valid for all integers up to some fixed K with $0 \le K < (8n)^2$. Our goal is to extend the result to $J = K + 1$, thereby completing the inductive step.

Let us introduce notation to clarify the quantities involved in our bounds. Define

$$R_J := \left\lfloor h^{1+\frac{J}{8n}} \right\rfloor \quad \text{and} \quad S_J := \left\lfloor \frac{h^2}{2^J} \right\rfloor \qquad \text{for } J = 0, 1, 2, \dots$$

as sequences of bounds on l and the total order of differentiation, respectively.

We now aim to show that if l is an integer such that

$$R_K < l \le R_{K+1},$$

and if $m_0, \dots, m_{n-1}$ are non-negative integers such that

$$m_0 + m_1 + \dots + m_{n-1} \le S_{K+1},$$

then the value of the auxiliary function $f(l)$ must be zero.

Recall from earlier that $f(z) = \Phi_{m_0,\dots,m_{n-1}}(z, \dots, z)$ is the specialization of the multivariate function Φ at the point $(z, \dots, z)$, and that this function is analytic. The function $f(z)$ can be thought of as a linear combination of exponential and algebraic terms, structured so that its derivatives vanish at many prescribed integer points.

To proceed, we analyze the derivatives of f at points $r < l$. Specifically, we consider

$$f_m^{(r)} := \left(\frac{d}{dz}\right)^m f(z)\bigg|_{z=r},$$

for all integers r such that $1 \le r \le R_K$, and for all integers m with $0 \le m \le S_{K+1}$. By the inductive hypothesis, each such $f_m^{(r)}$ vanishes.

Let us understand why this is the case. Since $f(z)$ is a specialization of a multivariate function $\Phi(z_0, \dots, z_{n-1})$, its m-th derivative at $z = r$ corresponds to a total m-th order mixed partial derivative of Φ at the point $(r, \dots, r)$. That is,

$$f_m^{(r)} = \left(\frac{\partial}{\partial z_0} + \cdots + \frac{\partial}{\partial z_{n-1}}\right)^m \Phi_{m_0,\dots,m_{n-1}}(z_0, \dots, z_{n-1})\bigg|_{z_0=\cdots=z_{n-1}=r}.$$

Expanding the differential operator via the multinomial theorem, this becomes

$$\sum_{j_0+\cdots+j_{n-1}=m} \frac{m!}{j_0! \cdots j_{n-1}!} \Phi_{m_0+j_0,\dots,m_{n-1}+j_{n-1}}(r, \dots, r).$$

Each term in the sum involves evaluation of Φ with total order

$$m_0 + j_0 + \cdots + m_{n-1} + j_{n-1} = m_0 + \cdots + m_{n-1} + m \le S_{K+1} + m \le 2S_{K+1} \le S_K,$$

so by the inductive assumption, every such term vanishes. Hence, we conclude that $f^{(m)}(r) = 0$ for all $r \le R_K$, and for all $m \le S_{K+1}$.

Now define the function

$$F(z) = (z-1)(z-2)\cdots(z-R_K),$$

and raise it to the S_{K+1}-th power to match the vanishing of derivatives up to order S_{K+1} at each $r \le R_K$. Then, $F(z)$ has a zero of multiplicity S_{K+1} at each such r, and hence the function

$$\frac{f(z)}{F(z)}$$

is analytic in a disk of radius slightly larger than R_K.

We now invoke the maximum modulus principle. Consider a closed disk $\mathcal{C}$ centered at the origin of radius

$$R := R_K + h^{1/(8n)}.$$

This disk contains all zeros of $F(z)$, and hence $f(z)/F(z)$ is analytic on and within $\mathcal{C}$. Then, for any point z on the boundary of $\mathcal{C}$, we have

$$|f(l)| \le \sup_{|z|=R} |f(z)| \cdot \frac{1}{|F(z)|}.$$

In other words, multiplying both sides by $|F(z)|$, we obtain the inequality

$$|\Theta f(l)| \le \Theta |f(z)|, \tag{8}$$

where θ and Θ denote the upper and lower bounds, respectively, of $|f(z)|$ and $|F(z)|$ on the circle $|z| = R$.

By Lemma 3.5, we know that the upper bound on $|f(z)|$ is at most

$$|f(z)| \le c_6^{h^3 + L\log R},$$

where $L \le h^2$ and $R \le 2h$, and thus $\log R = O(\log h)$.

Also, since each root of $F(z)$ lies at distance at least $h^{1/(8n)}$ from $l > R_K$, we estimate

$$|F(l)| \geq \left(\tfrac{1}{2}h^{1/(8n)}\right)^{R_K S_{K+1}}.$$

On the other hand, the lower bound on $|f(l)|$ provided by Lemma 3.5 (under the assumption that $f(l) \neq 0$) is:

$$|f(l)| > c_6^{-h^3 - L\log R}.$$

Putting these into inequality (8), we obtain the contradiction

$$c_6^{-h^3 - L\log R} \leq c_6^{h^3 + L\log R} \cdot \left(\tfrac{1}{2}h^{1/(8n)}\right)^{-R_K S_{K+1}}.$$

Taking logarithms and rearranging gives an inequality that becomes false for sufficiently large h, because the right-hand side grows faster than the left-hand side.

Therefore, the only consistent possibility is that $f(l) = 0$ for all l in the specified range. This completes the inductive step and hence the proof. ■

Lemma 3.7. *Define the function* $\phi(z) = \Phi(z, \ldots, z)$*, where* Φ *is as introduced previously. Then, for all integers* j *satisfying* $0 \leq j \leq h^{8n}$*, we have the estimate*

$$|\phi_j(0)| < \exp(-h^{8n}). \tag{9}$$

Proof. We begin by invoking the conclusion of Lemma 3.6, which asserts the vanishing of certain derivatives of the auxiliary function Φ when its arguments are evaluated at appropriate integer points. In particular, Lemma 3.6 ensures that for all integers l in a specified range, and all non-negative integers $m_0, \ldots, m_{n-1}$ with their total sum appropriately bounded, we have

$$\Phi_{m_0,\ldots,m_{n-1}}(l, \ldots, l) = 0.$$

Let us now define two auxiliary parameters that will help control the size of the derivatives and the region over which we estimate the function. Set:

$$X := h^{8n}, \qquad Y := \left\lfloor \frac{h^2}{2^{8n}} \right\rfloor.$$

These choices are made to ensure that X is large (exponentially in h) and Y is a corresponding bound on the order of differentiation.

With these definitions, Lemma 3.6 tells us that for all integers r in the interval $1 \leq r \leq X$, and for all integers m such that $0 \leq m \leq Y$, the derivatives

$$\phi_m(r) = \left(\frac{d^m}{dz^m}\phi(z)\right)\bigg|_{z=r}$$

vanish. This means that the function $\phi(z)$ has a zero of multiplicity at least Y at each point $z = r$ for $1 \leq r \leq X$.

To capture this structure algebraically, we define the function

$$E(z) := \prod_{r=1}^{X}(z - r)^Y = [(z-1)(z-2)\cdots(z-X)]^Y,$$

which is a polynomial of degree XY, vanishing to order Y at each of the integers $1, 2, \ldots, X$. Then the quotient $\phi(z)/E(z)$ is an entire function (i.e., analytic everywhere in the complex plane), because all zeros of $\phi(z)$ at $z = r$ are at least of multiplicity Y, and hence cancel exactly with those of $E(z)$.

Now, consider the circle Γ centered at the origin with radius

$$R := Xh^{1/(8n)}.$$

Since $R > X$, the function $\phi(z)/E(z)$ is analytic inside and on the boundary of this circle. Therefore, by the maximum modulus principle, we obtain a bound on $|\phi(w)|$ in terms of the values of $|E(w)|$ and the maximum and minimum of $|\phi(z)|$ and $|E(z)|$ on Γ.

Specifically, we let ξ and Ξ denote an upper bound for $|\phi(z)|$ and a lower bound for $|E(z)|$, respectively, for z on Γ. Then, for any point w inside the disk $|w| < X$, we have:

$$|\phi(w)| \leq \frac{\xi}{\Xi}|E(w)|.$$

We now estimate these quantities. First, on the circle of radius R, the modulus of $E(z)$ satisfies:

$$|E(z)| \leq (2X)^Y \quad \text{for } |z| \leq R,$$

since $|z - r| \leq |z| + |r| \leq 2X$ for each root $r = 1, \ldots, X$.

Next, for the lower bound, note that on Γ, the distance $|z-r| \geq \frac{1}{2}R$ for all $r \in \{1, \ldots, X\}$, since the circle has radius $R = Xh^{1/(8n)}$, which is sufficiently large compared to the location of the roots. Thus,

$$|E(z)| \geq \left(\tfrac{1}{2}R\right)^{XY} =: \Xi.$$

As for the upper bound ξ on $|\phi(z)|$, we appeal to Lemma 3.5, which gives

$$|\phi(z)| \leq c_6^{h^3+LR},$$

where L is a bound on the degree in the exponential components, and $R = Xh^{1/(8n)}$ as above.

Putting these together, we find:

$$|\phi(w)| \leq c_6^{h^3+LR} \cdot \left(\frac{1}{2}R\right)^{-XY}.$$

Now observe that $LR \leq h^{2n+2}$, and since

$$XY = h^{8n} \cdot \left\lfloor \frac{h^2}{2^{8n}} \right\rfloor \gg h^{2n+2},$$

we see that the exponent $h^3 + LR - XY \log R$ is negative and of large magnitude. Therefore, the right-hand side is smaller than $\exp(-XY)$, and hence

$$|\phi(w)| < e^{-XY}.$$

Finally, we use Cauchy's integral formula to estimate the coefficient $\phi_j(0)$. Since $\phi(z)$ is analytic in the disk of radius $R > 1$, we have

$$\phi_j(0) = \frac{j!}{2\pi i} \oint_\Lambda \frac{\phi(w)}{w^{j+1}}\, dw,$$

where Λ is the circle $|w| = 1$, traversed positively. On this circle, we have $|w| = 1$, and since $|\phi(w)| < e^{-XY}$, we deduce:

$$|\phi_j(0)| \leq j! \cdot e^{-XY}.$$

This final estimate implies that

$$|\phi_j(0)| < \exp(-h^{8n})$$

for all $j \leq h^{8n}$, completing the proof.

■

Lemma 3.8. *Consider any integers $t_1, t_2, \ldots, t_n$, not all simultaneously zero, with the property that their absolute values are bounded above by some positive integer T. Then the linear form in logarithms*

$$|t_1 \log \alpha_1 + t_2 \log \alpha_2 + \cdots + t_n \log \alpha_n|$$

is bounded below by an explicit positive quantity of the form

$$c_{11}^{-T},$$

where c_{11} is a positive constant depending only on the algebraic numbers involved.

Proof. To begin, for each j with $1 \leq j \leq n$, let us denote by a_j the leading coefficient of the minimal polynomial of α_j if $t_j > 0$, or the leading coefficient of the minimal polynomial of α_j^{-1} if $t_j < 0$. This choice ensures that the algebraic integers constructed will properly reflect the exponents t_j appearing in the linear form.

We now define the algebraic integer

$$\omega = a_1^{|t_1|} \cdots a_n^{|t_n|} \left(\alpha_1^{t_1} \alpha_2^{t_2} \cdots \alpha_n^{t_n} - 1 \right).$$

By construction, ω lies in the ring of algebraic integers, and its degree over $\mathbb{Q}$ is at most d^n, where d is the degree of the number field generated by the α_j.

Next, consider any conjugate of ω, which is obtained by applying an embedding of the number field into $\mathbb{C}$ that replaces each α_j by one of its conjugates. The absolute value of any such conjugate is bounded above by a constant of the form c_{12}^T, where c_{12} depends on the heights and absolute values of the conjugates of the α_j. This follows from the fact that raising α_j to the power t_j and multiplying by powers of the leading coefficients can only grow the magnitude exponentially in T.

There are two cases to consider:

- If $\omega = 0$, then by definition we have

  $$\alpha_1^{t_1} \alpha_2^{t_2} \cdots \alpha_n^{t_n} = 1,$$

  which implies that the linear form in logarithms

  $$\Omega = t_1 \log \alpha_1 + t_2 \log \alpha_2 + \cdots + t_n \log \alpha_n$$

  is an integral multiple of $2\pi i$. Since the logarithms $\log \alpha_j$ are assumed to be linearly independent over the rationals, Ω must actually be zero only if all t_j vanish, which contradicts the hypothesis that not all t_j are zero. Hence, in this scenario, the lemma holds trivially.
- Otherwise, if $\omega \neq 0$, then its norm (the product of all conjugates) is a nonzero integer and hence has absolute value at least 1. Consequently, the absolute value of ω itself must satisfy the inequality

  $$|\omega| \geq c_{12}^{-Tn},$$

  for some suitable constant c_{12}.

  On the other hand, using the well-known inequality for the exponential function,

  $$|e^z - 1| \leq |z| e^{|z|},$$

  valid for all complex z, we apply this to the expression inside ω:

  $$|\alpha_1^{t_1} \cdots \alpha_n^{t_n} - 1| = |e^{\Omega} - 1| \leq |\Omega| e^{|\Omega|}.$$

Assuming, without loss of generality, that $|\Omega| < 1$ (otherwise the lower bound is trivial), we then find a constant c_{13} such that

$$|\omega| \leq |\Omega| c_{13}^T.$$

Combining this upper bound with the lower bound on $|\omega|$, we deduce that

$$|\Omega| \geq c_{11}^{-T},$$

for some positive constant c_{11} depending only on the algebraic numbers and the degree d.

This completes the proof of the lemma, establishing a nontrivial explicit lower bound on the absolute value of any nonzero linear form in logarithms of algebraic numbers with bounded integer coefficients.

■

Lemma 3.9. *Consider positive integers R and S, and let $\sigma_0, \sigma_1, \ldots, \sigma_{R-1}$ be R distinct complex numbers. We define two important parameters:*

$$\sigma = \max\{1, |\sigma_0|, |\sigma_1|, \ldots, |\sigma_{R-1}|\},$$

which represents the maximum among 1 and the magnitudes of the σ_i, and

$$\rho = \min\{1, |\sigma_i - \sigma_j| : 0 \leq i < j < R\},$$

which is the minimum distance between any two distinct σ_i and σ_j, or 1, whichever is smaller. This minimum separation ρ measures how close together the points σ_i are in the complex plane, and ensures they are sufficiently apart for our construction.

Under these conditions, for any pair of integers r and s satisfying $0 \leq r < R$ and $0 \leq s < S$, there exist complex coefficients w_j, for $j = 0, 1, \ldots, RS - 1$, with absolute values bounded above by $(8\rho/\sigma)^{RS}$, such that the polynomial

$$W(z) = \sum_{j=0}^{RS-1} w_j z^j$$

exhibits the following interpolation properties:

$$W^{(j)}(\sigma_i) = 0 \quad \text{for all } 0 \leq i < R,\ 0 \leq j < S \quad \text{except for } (i, j) = (r, s),$$

and at the exceptional point,

$$W^{(s)}(\sigma_r) = 1.$$

In other words, $W(z)$ is a polynomial of degree at most $RS - 1$ that vanishes with multiplicity S at each σ_i except at σ_r, where the s-th derivative of W equals 1, and all lower order derivatives vanish. This polynomial thus acts as a specialized interpolating polynomial focused on the s-th derivative at σ_r.

Proof. To explicitly construct such a polynomial, we use an integral representation based on Cauchy's integral formula for derivatives. Define

$$U(z) = (z - \sigma_0)(z - \sigma_1) \cdots (z - \sigma_{R-1})^S,$$

which is a polynomial vanishing at each σ_i with multiplicity S. This polynomial $U(z)$ encodes the zeros and their multiplicities, serving as a foundational building block for $W(z)$.

The polynomial $W(z)$ can then be written as

$$W(z) = \frac{(-1)^s}{s!}\frac{1}{2\pi i}\int_{C_r} \frac{(\zeta-\sigma_r)^s U(z)}{(\zeta-z)U(\zeta)}\,d\zeta,$$

where C_r is a positively oriented (counterclockwise) circle centered at σ_r with sufficiently small radius $\delta < \rho$ so that C_r encloses only σ_r and excludes the other σ_i. The choice of the contour ensures the integrand is analytic on and inside C_r except at $\zeta = z$ and $\zeta = \sigma_r$.

By Cauchy's differentiation formula, this integral representation guarantees that $W(z)$ is a polynomial of degree at most $RS-1$ with the desired vanishing conditions on derivatives. To verify the interpolation properties, we use the residue theorem and properties of $U(z)$.

Next, the proof leverages an alternative representation for $W(z)$, obtained by considering the behavior of the integral as $|\zeta| \to \infty$. Since $U(\zeta)$ grows like ζ^{RS} for large ζ, the integrand multiplied by $|\zeta|^\ell$ tends to zero for sufficiently large ℓ. Applying Cauchy's residue theorem at infinity, we obtain an expression for $W(z)$ as a sum of integrals over small circles C_j around the other points σ_j (for $j \neq r$):

$$W(z) = \frac{(z-\sigma_r)^s}{s!}\frac{1}{2\pi i}\sum_{\substack{j=0\\ j\neq r}}^{R-1}\int_{C_j} \frac{(\zeta-\sigma_r)^s}{\zeta - z}\frac{U(z)}{U(\zeta)}\,d\zeta.$$

This representation makes it clear that $W(z)$ is a rational function, regular at each σ_j for $j \neq r$, since $U(z)$ has zeros of order S at these points. Hence, $W^{(j)}(\sigma_i) = 0$ for all $i \neq r$ and $j < S$, establishing the required vanishing of derivatives at those points.

Furthermore, by examining the integrand near σ_r and using the integral representation, it follows that $W^{(s)}(\sigma_r) = 1$, and all lower order derivatives vanish there. This confirms the interpolation condition at the point σ_r.

Finally, the lemma asserts that the coefficients w_j of the polynomial $W(z)$ are uniformly bounded in magnitude by $(8\rho/\sigma)^{RS}$. This bound is established by analyzing the size of the terms in the integral and the sum over indices, noting the finite number of terms involved and the size constraints on σ_i and their separations. In particular, the number of terms in the sums is at most S^R, and the magnitudes of the rational function components are controlled by powers of σ and ρ. Combining these observations yields the stated bound on the coefficients.

Thus, the lemma constructs a carefully controlled polynomial $W(z)$ which serves as an interpolating polynomial for derivatives at prescribed points, with explicit bounds on its coefficients. This tool is fundamental in applications requiring precise polynomial approximations with derivative constraints at multiple points. ■

We now proceed to demonstrate that the inequalities labeled as (9) in Lemma 3.7 cannot all hold simultaneously. Establishing this impossibility will lead directly to the proof of Theorem 2.1 by contradiction.

To start, we introduce the notation $S = L+1$ and $R = S^n$. Any integer i in the range $0 \le i < RS$ can be uniquely expressed in the base S expansion form:

$$i = \lambda_0 + \lambda_1 S + \cdots + \lambda_n S^n,$$

where each λ_j is an integer satisfying $0 \le \lambda_j \le L$.

For each such integer i, we define the quantities

$$v_i = \lambda_0, \quad p_i = p(\lambda_0, \ldots, \lambda_n),$$

and also set

$$\psi_i = \lambda_1 \log \alpha_1 + \cdots + \lambda_n \log \alpha_n.$$

With these definitions, the function $\phi(z)$ can be written explicitly as

$$\phi(z) = \sum_{i=0}^{RS-1} p_i z^{v_i} e^{\psi_i z}. \tag{10}$$

Using Lemma 3.8, we know that any two distinct values of ψ_i, corresponding to different multi-indices $\lambda_1, \ldots, \lambda_n$, differ by at least a positive constant c_{11}^{-L}. There are exactly R such distinct values, which we denote by $\sigma_0, \sigma_1, \ldots, \sigma_{R-1}$.

If we let ρ and σ be as defined in Lemma 3.9, then it follows that

$$\sigma \leq c_{14}^L, \quad \rho \geq c_{15}^{-L}.$$

Next, choose any suffix t such that the coefficient $p_t \neq 0$. Define $s = v_t$, and let r be the index for which $\psi_t = \sigma_r$. Denote by $W(z)$ the polynomial constructed in Lemma 3.9. By the properties given in that lemma, the following equality holds:

$$p_t = \sum_{i=0}^{RS-1} p_i W_i(\psi_i).$$

Applying Leibniz's rule for differentiation, we have

$$W_r(\psi_i) = \sum_{j=0}^{RS-1} j(j-1)\cdots(j-v_i+1) w_j \psi_i^{j-v_i} = \sum_{j=0}^{RS-1} w_j \left[\frac{d^r}{dz^r} \left(z^i e^{\psi_i z} \right) \right]_{z=0},$$

and substituting this into equation (10), we get

$$p_t = \sum_{j=0}^{RS-1} w_j \phi_j(0).$$

Since $RS \leq h^{2n+2}$, Lemma 3.7 guarantees that inequality (9) is valid for all indices i with $0 \leq j \leq RS$. Moreover, by Lemma 3.9, the coefficients w_j satisfy

$$|w_j| \leq (8\sigma/\rho)^{RS} \leq (8c_{14}Lc_{15}^L)^{RS} \leq c_{16}^{h^{2n+4}}.$$

Because $|p_t| \geq 1$, it follows that

$$0 \leq \log RS + c_{17} h^{2n+4} - h^{8n}.$$

However, this inequality cannot hold if the parameter h is chosen sufficiently large, leading to a contradiction. Therefore, the assumption that all inequalities (9) hold is false, which completes the proof of the theorem.

4. Recent Research and Open Problems

While there are any newer results in this field, we will mention one of the most important, namely Nesterenko's.

Definition 4.1. A set of elements is *algebraically independent* over a field if there is no nontrivial polynomial relation among them with coefficients in that field.

Theorem 4.2 (Nesterenko)**.** *π and e^π are algebraically independent over the rationals.*

We will not go too in depth into this theorem, but this is here merely to point out that even some relatively simple-looking results are still being shown in recent times. In fact, some other open problems in the field include the algebraic independence of e and π as well as the transcendence of Euler's constant γ. Both of these, along with many other open problems, would greatly contribute to the field. To see more open problems, see Waldschmidt [Wal22].

5. Acknowledgments

The author would like to thank Simon Rubenstein-Salzedo for providing the environment to make write this paper possible as well as Benjamin Vakil for providing the resources and help for this paper as well as general support throughout the process

References

[Bak75] Alan Baker. *Transcendental Number Theory*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1975.

[Cox04] David A. Cox. *Making Transcendence Transparent: An Intuitive Approach to Classical Transcendental Number Theory*, volume 117 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2004.

[Wal22] Michel Waldschmidt. Transcendental numbers: Open problems. `https://webusers.imj-prg.fr/~michel.waldschmidt/articles/pdf/TNTOpenPbs.pdf`, 2022. Lecture slides, available online.