Transcendental Numbers (Draft)

Vincent Lin

Last Updated: December 5, 2023

Contents

Preface v

1 Preliminaries 1

1.1 Algebraic and Transcendental Numbers 1

1.2 Algebra Preliminaries 2

1.3 Analysis Preliminaries 7

1.4 Liouville Numbers 8

1.5 Transcendence of e and π 9

2 Lindemann-Weierstrass Theorem 11

3 Gelfond-Schneider Theorem 13

3.1 Background 13

3.2 Main Theorem 19

3.3 Consequences 19

4 Baker's Theorem 21

5 Schanuel's Conjecture 23

Preface

This uses the background information from 21441 and algebra and reviews it in order to discuss transcendental numbers. Then, it will introduce methods of explicitly constructing, finding, and detecting transcendental numbers. Several classes of transcendental numbers will be named and the methods of finding them will be discussed, including the Lindemann-Weierstrass theorem, Gelfond-Schneider theorem, and Baker's theorem. It concludes with a discussion of Schanuel's conjecture. The main sequence of theorems is primarily related to determining the transcendence of complex numbers using auxiliary functions, but there are many ways of finding transcendental numbers.

Chapter 1

Preliminaries

1.1 Algebraic and Transcendental Numbers

Def An **algebraic number** is $\alpha \in \mathbb{C}$ which is the root of a nonzero polynomial in $\mathbb{Q}[x]$. A **transcendental number** is a complex number that is not algebraic.

Note We will use $\mathbb{A}$ to denote the set of algebraic numbers and we will use $\mathbb{C} \setminus \mathbb{A}$ to denote the set transcendental numbers. We will use $\mathbb{R} \setminus \mathbb{A}$ to denote the set of real transcendental numbers.

To find examples of algebraic numbers, we can take any nonzero polynomial in $\mathbb{Q}[x]$ find its roots. By definition, these roots are algebraic numbers. For example, $\sqrt{2}$ is algebraic because it is a root of $x^2 - 2$. Also, i is algebraic because it is the root of $x^2 + 1$. All rational numbers are algebraic as well. Let $\frac{p}{q} \in \mathbb{Q}$ be rational, where $p, q \in \mathbb{Z}$ and q is nonzero. Then, it is the root of $x - \frac{p}{q}$. Therefore, $\mathbb{Q} \subseteq \mathbb{A}$.

What about transcendental numbers? Do they exist?

Theorem 1.1.1 *Yes, transcendental numbers exist.*

Proof Consider the set of algebraic numbers, which we will denote by $\mathbb{A}$. This set is countable. We will show this by forming a surjection

$$\phi : \left(\mathbb{N} \times \bigcup_{n \in \mathbb{N}} \mathbb{Q}^n\right) \to \mathbb{A}\,.$$

Note that $\mathbb{N} \times \bigcup\limits_{n \in \mathbb{N}} \mathbb{Q}^n$ is countable because it is the Cartesian product of a countable set with a countable union of countable sets. By the fundamental theorem of algebra, we know that a polynomial of degree k has k (not necessarily distinct) roots. Therefore, we can number the k roots from 1 to k for each polynomial. Thus, if we specify the nonzero rational coefficients $(a_0, \dots, a_k)$ and an index i for $i \in [k]$ to be the i-th root of the polynomial $a_k x^k + \cdots + a_0$, we get an algebraic number. We define the map $\phi(i, (a_0, \dots, a_k))$ to be the i-th root of $a_k x^k + \cdots + a_0$ if it exists. Otherwise, we map the input to zero. From the definition, we know that every algebraic number can be encoded this way since it is one of the roots of a nonzero rational polynomial. Thus, for all $a \in \mathbb{A}$, there must be an input x such that $\phi(x) = a$, making this a surjective map from a countable set to $\mathbb{A}$. This makes $\mathbb{A}$ countable. However, since $\mathbb{C}$ is uncountable, it cannot be the case that all complex numbers are algebraic. Thus, if a complex number is not algebraic, it must be transcendental. Furthermore, $\mathbb{R}$ is uncountable so there must be real transcendental numbers as well.

♦

If transcendental numbers exist, then can we find an example? To show that a number is algebraic, we just have to find a nonzero rational polynomial that it is a root of, evaluate that polynomial at that number, and verify that we get zero. However, checking if a number is transcendental is a very hard problem. The rest of this paper will discuss the transcendence of a large class of numbers and methods for determining transcendence.

1.2 Algebra Preliminaries

Before that, we will generalize the concept of transcendence to other fields. In order to do this, we will list several definitions. Let K, L be fields and let L/K be a field extension. We will introduce a few definitions for recall.

Recall

- $a \in L$ is **algebraic** over K if there is a nonzero polynomial $f \in K[x]$ such that $f(a) = 0$. Otherwise, a is **transcendental** over K. In other words, an algebraic number is a complex number that is algebraic over $\mathbb{Q}$ and a transcendental number is a complex number that is not algebraic over $\mathbb{Q}$.

- $X \subseteq L$ is **algebraically independent** over K if for all $a_1, \dots, a_t$ distinct and all $f \in K[x_1, \dots, x_t]$, $f(a_1, \dots, a_t) = 0$ implies $f = 0$.

- If $a \in L$ is algebraic over K, the **minimal polynomial** of a over K, denoted m_a^K or m_a when the underlying field is implied, is the unique monic irreducible polynomial which generates the kernel of the evaluation map $\phi(f) = f(a)$ for $f \in K[x]$. The **degree** of $\alpha \in K$ is the degree of the minimial polynomial m_a and the **conjugates** of a are the roots of m_a.

- L/K is an **algebraic extension** if every element in L is algebraic over K.

- L is **algebraically closed** if every nonconstant single variable polynomial in L has a root in L.

- An **algebraic closure** of K is an algebraic extension of K that is algebraically closed.

Now we will consider some important facts about number fields:

Recall

- An **algebraic number field** or **number field** is an algebraic extension of $\mathbb{Q}$.
- An **algebraic integer** is $\alpha \in \mathbb{C}$ such that it is the root of a monic polynomial in $\mathbb{Z}[x]$.
- If K is a number field, $\mathcal{O}_K$ is the **ring of integers** of K and is the ring of all algebraic integers in K. It is a free $\mathbb{Z}$-module of rank $[K : \mathbb{Q}]$. If $\alpha \in K$, there is a positive $c \in \mathbb{Z}$ such that $c\alpha \in \mathcal{O}_K$.
- Let K be a number field. An **integral basis** is a $\mathbb{Q}$-basis for K that is also a $\mathbb{Z}$-basis for $\mathcal{O}_K$.

Lemma 1.2.1 *Let K be a number field and let $\alpha \in K$. The degree of α over $\mathbb{Q}$ divides the degree of K.*

Proof By the tower law, $[K : \mathbb{Q}(\alpha)][\mathbb{Q}(\alpha) : \mathbb{Q}] = [K : \mathbb{Q}]$. Thus, $[\mathbb{Q}(\alpha) : \mathbb{Q}] \mid [K : \mathbb{Q}]$.

♦

Note Let K be a number field of degree n over $\mathbb{Q}$ and let $\alpha \in K$. Then, let $\sigma_1, \dots, \sigma_n$ be the $\mathbb{Q}$-fixing embeddings $K \hookrightarrow \mathbb{C}$. We write $\|\alpha\|$ to denote $\max\limits_{i \in \{1,\dots,n\}} |\sigma_i(\alpha)|$.

Lemma 1.2.2 *Let $\alpha, \beta \in K$, a number field. Then, $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$ and $\|\alpha\beta\| \leq \|\alpha\|\|\beta\|$.*

Proof We first show this for addition.

$$\begin{aligned}
\|\alpha + \beta\| &= \max_{i \in \{1,\dots,n\}} |\sigma_i(\alpha + \beta)| \\
&= \max_{i \in \{1,\dots,n\}} |\sigma_i(\alpha) + \sigma_i(\beta)| \\
&\leq \max_{i \in \{1,\dots,n\}} |\sigma_i(\alpha)| + |\sigma_i(\beta)| \\
&\leq \max_{i \in \{1,\dots,n\}} |\sigma_i(\alpha)| + \max_{i \in \{1,\dots,n\}} |\sigma_i(\beta)| \\
&= \|\alpha\| + \|\beta\|
\end{aligned}$$

The proof for multiplication is similar by replacing $+$ with $\cdot$ instead.

♦

Lemma 1.2.3 *If K is a number field, it has an integral basis.*

Proof We know that $\mathcal{O}_K$ is a free $\mathbb{Z}$-module of rank $n := [K : \mathbb{Q}]$ so let $\{\beta_1, \ldots, \beta_n\}$ be a basis for $\mathcal{O}_K$ as a $\mathbb{Z}$-module. To show that this is a basis for K, we will show that this set $\mathbb{Q}$-spans K and is $\mathbb{Q}$-linearly independent.

First, we will show it is spanning. Let $\alpha \in K$. Then, find positive $c \in \mathbb{Z}$ such that $c\alpha \in \mathcal{O}_K$. Now, for $\lambda_1, \ldots, \lambda_n \in \mathbb{Z}$, we can write

$$c\alpha = \lambda_1\beta_1 + \cdots + \lambda_n\beta_n.$$

Therefore, since $c > 0$, and is an integer, we can divide by c to get

$$\alpha = \frac{\lambda_1}{c}\beta_1 + \cdots + \frac{\lambda_n}{c}\beta_n.$$

Since $\frac{\lambda_i}{c} \in \mathbb{Q}$, we know that the basis $\mathbb{Q}$-spans K. Now, we will show it is $\mathbb{Q}$-linearly independent. Suppose for contradiction

$$0 = \lambda_1\beta_1 + \cdots + \lambda_n\beta_n,$$

where $\lambda_i \in \mathbb{Q}$ but there is a nonzero λ_i. Then, we can multiply both sides by the least common multiple of the demoninators of the λ_i to get that the λ_i's are not $\mathbb{Z}$-linearly independent, which is a contradiction.

♦

Lemma 1.2.4 *For any nonzero polynomial $p(x) \in \mathbb{C}[x]$ with a root at $x = \alpha$ of multiplicity $m > 0$, $p'(x)$, has the root α with multiplicity $m - 1$. More generally, if R is an integral domain and $p \in R[x]$ has α with multiplicity m, p', the formal derivative, has root α with multiplicity $m - 1$ if* $\operatorname{char}(R)$ *is not a factor of m.*

Proof We can write $p(x)$ as $(x-\alpha)^m q(x)$ where $(x-\alpha) \nmid q(x)$. Now, taking the derivatives of both sides and using the product rule, we get

$$\begin{aligned} p(x) &= (x-\alpha)^m q(x) \\ p'(x) &= ((x-\alpha)^m)' q(x) + (x-\alpha)^m q'(x) \\ &= m(x-\alpha)^{m-1} q(x) + (x-\alpha)^m q'(x) \\ &= (x-\alpha)^{m-1} \big(mq(x) + (x-\alpha)q'(x)\big) \end{aligned}$$

Therefore, we get that $(x-\alpha)^{m-1} \mid p'(x)$ so $p'(x)$ has the root α of multiplicity at least $m-1$.

Now, we want to show that $p'(x)$ has root α with multiplicity at most $m-1$. We need to show that $x - \alpha$ does not divide $(mq(x) + (x-\alpha)q'(x))$. We can take this expression modulo $(x-\alpha)$ to get that it is equivalent to $mq(x)$. Note that $(x-\alpha) \nmid q(x)$, so we need $m > 0$. This is true when $\operatorname{char}(R)$ is not a factor of m (otherwise the entire expression is zero). Therefore, we have multiplicity at most $m-1$. Since $\operatorname{char}(\mathbb{C}) = 0$, we know this rule applies to complex polynomials.

♦

Def The square **Vandermonde matrix** of size n is the matrix $V = V(x_1, \ldots, x_n)$ is the $n \times n$ matrix

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix}$$

with nonzero complex elements x_i^j in the i-th row and $j+1$-th column column where $0 \le j \le n-1$ and $1 \le i \le n$.

Lemma 1.2.5 *Let $V = V(x_1, \dots, x_n)$ be the square Vandermonde matrix of size n. Then, the determinant of the square Vandermonde matrix is*

$$\prod_{1 \leq i < j \leq n} (x_j - x_i).$$

Therefore, the determinant vanishes if and only if $x_i = x_j$ for any $i \neq j$.

Proof For $1 \leq i \leq n$, let $E := \{e_i\}$ be the canonical basis for $\mathbb{C}^n$.

Let P_n be the $\mathbb{C}$-vector space of polynomials with degree less than n. Furthermore, for $1 \leq i \leq n$, let $A := \{a_i\}$ be the monomial basis for P_n where $a_i(x) = x^{i-1}$ and let $B := \{b_i\}$ be another monomial basis for P_n where $b_i(x) = \prod\limits_{j<i}(x - x_j)$. For example,

$$\begin{aligned} b_1 &= 1, \\ b_2 &= (x - x_1), \\ b_3 &= (x - x_1)(x - x_2), \\ &\vdots \\ b_n &= (x - x_1)(x - x_2) \dots (x - x_{n-1}). \end{aligned}$$

Consider the linear transformation $\psi : P_n \to \mathbb{C}^n$ via $\psi(p) = (p(x_1), \dots, p(x_n))$. We can think of this as a map from $p(x)$ to a column vector of length n.

$$p \mapsto \begin{pmatrix} p(x_1) \\ p(x_2) \\ \vdots \\ p(x_n) \end{pmatrix}$$

Let V be the matrix of ψ with respect to A and E. Now, let L be the matrix of ψ with respect to B and E. Finally, let U be the change of basis matrix from B to A. Then,

$$\begin{aligned} VU &= L \\ \det(VU) &= \det(L) \\ \det(V)\det(U) &= \det(L) \end{aligned}$$

Now, V, the matrix of ψ with respect to A and E is

$$\begin{pmatrix} | & | & \cdots & | \\ \psi(a_1) & \psi(a_2) & \cdots & \psi(a_n) \\ | & | & \cdots & | \end{pmatrix} = \begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix},$$

the square Vandermonde matrix of size n.

On the other hand, L, the matrix of ψ with respect to B and E is

$$\begin{pmatrix} | & | & \cdots & | \\ \psi(b_1) & \psi(b_2) & \cdots & \psi(b_n) \\ | & | & \cdots & | \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \cancelto{0}{(x_1 - x_1)} & \cdots & \cancelto{0}{(x_1 - x_1)}(x_1 - x_2)\dots(x_1 - x_{n-1}) \\ 1 & (x_2 - x_1) & \cdots & (x_2 - x_1)\cancelto{0}{(x_2 - x_2)}\dots(x_2 - x_{n-1}) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (x_n - x_1) & \cdots & (x_n - x_1)(x_n - x_2)\dots(x_n - x_{n-1}) \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & (x_2 - x_1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (x_n - x_1) & \cdots & (x_n - x_1)(x_n - x_2)\dots(x_n - x_{n-1}) \end{pmatrix}$$

Note that L is lower triangular so $\det(L)$ is the product of the entries along its diagonal, or $\prod\limits_{1 \le i < j \le n} (x_j - x_i)$.

Finally, consider the change of basis matrix U from basis B to basis A in P_n. The columns of U records the coefficients of the b_i polynomials after expanding. Note that since the b_i are monic, the diagonal of U consists of all 1s because the (i,i)-entry of U is the coefficient of x^{i-1} (the leading coefficient) in b_i. Therefore, U is of the form

$$\begin{pmatrix} 1 & \text{blah} & \text{blah} & \cdots & \text{blah} \\ 0 & 1 & \text{blah} & \cdots & \text{blah} \\ 0 & 0 & 1 & \cdots & \text{blah} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Thus, since U is upper triangular, $\det(U)$ is the product of the 1s along the diagonal so $\det(U) = 1$.

Plugging what we know about $\det(L)$ and $\det(U)$ into $\det(V)\det(U) = \det(L)$, we get that $\det(V) = \prod\limits_{1 \le i < j \le n} (x_j - x_i)$.

♦

1.3 Analysis Preliminaries

Now we will introduce some concepts from complex analysis. Consider the set of complex numbers $\mathbb{C}$ as a normed space equipped with the absolute value function $|\cdot| : \mathbb{C} \to \mathbb{R}$ via $|z| = \sqrt{z\bar{z}}$, where $\bar{z}$ is the complex conjugate of z. This norm induces a metric space, so we can define the distance between x and y as $d(x,y) := |x - y|$. Furthermore, this metric space induces a topological space so we can define open sets $U \subseteq \mathbb{C}$ such that all points

$z \in U$ are contained in a ball within U centered at z. In other words, for all $z \in U$, there is $\varepsilon > 0$ such that for all y where $d(z, y) < \varepsilon$, $y \in U$.

Def Let $f : \mathbb{C} \to \mathbb{C}$ be a function. The **limit** of f as $z \to z_0$ is L if and only if for all $\varepsilon > 0$, there is $\delta > 0$ such that $0 < |z - z_0| < \delta \implies |f(x) - L| < \varepsilon$. We write L as

$$\lim_{z \to z_0} f(z).$$

Def Let $U \subseteq \mathbb{C}$ be an open subset. Let $f : U \to \mathbb{C}$ be a function. The **derivative** of f at a point $z_0 \in \mathbb{C}$ is defined as

$$f'(z_0) := \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}.$$

If the derivative exists, the function is said to be **complex differentiable** or **differentiable** at z_0.

Def Let $f : \mathbb{C} \to \mathbb{C}$ be a function and let $U \subseteq \mathbb{C}$ be an open subset of $\mathbb{C}$. Then, f is **holomorphic** on U if it is complex differentiable on every point in U. Furthermore, if $U = \mathbb{C}$, then f is an **entire** function.

Theorem 1.3.1 *(Cauchy's Residue Theorem) Let $U \subseteq \mathbb{C}$ be a simply connected open set containing a finite number of points $\{a_1, \ldots, a_n\}$. Then, let U_0 be $U \setminus \{a_1, \ldots, a_n\}$. Let $f : U_0 \to \mathbb{C}$ be holomorphic on U_0. Furthermore, let γ be a closed rectifiable curve in U_0, $\mathrm{Res}(f, a_i)$ denote the residue of f at each a_k, and $I(\gamma, a_k)$ be the winding number of γ around a_k. Then,*

$$\oint_\gamma f(z)dz = 2\pi i \sum_{k=1}^{n} I(\gamma, a_k)\,\mathrm{Res}(f, a_k).$$

1.4 Liouville Numbers

Now we will start looking for our first transcendental number!

Def A **Liouville number** is a real number x such that for all $n \in \mathbb{N}$, there exists integers p, q with $q > 1$ such that

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

Essentially, Liouville numbers are real numbers that can be approximately really, really closely by a sequence of rationals of the form $\left\{\frac{p_i}{q_i}\right\}$, where $q_i > 1$, and the distance between x and $\frac{p_i}{q_i}$ is nonzero, but less than $\frac{1}{(q_i)^i}$.

Theorem 1.4.1 *(Liouville's Approximation Theorem) Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ be an algebraic number of degree n. Then, for any rational approximation $\frac{p}{q}$ to α, we have*

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{1}{q^n}$$

Proof Let f be the minimal polynomial of α over $\mathbb{Q}$. Note that $f'(\alpha) \neq 0$. This is because $\deg(f') < \deg(f)$ and f is the minimal polynomial. Furthermore, f is irreducible in $\mathbb{R}[x]$, so it cannot be the case that $(x - r) \mid f$ for a rational r. In other words, f has no rational roots. Now let $\frac{p}{q}$ be a rational number which we will use to approximate α and plug it into $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$.

$$\begin{aligned} f\left(\frac{p}{q}\right) &= a_n\left(\frac{p}{q}\right)^n + a_{n-1}\left(\frac{p}{q}\right)^{n-1} + \cdots + a_0 \\ &= \frac{C}{q^n} \end{aligned}$$

for some constant $C \neq 0$ as $\frac{p}{q}$ is not a root of f. Also, note that since we want a good approximation of α that is close to $\frac{1}{q^n}$ away, this rational must be at most distance 1 away from α.

♦

1.5 Transcendence of e and π

Lemma 1.5.1 *Let $f(x)$ be a real polynomial with degree m. Let*

$$I(t) = \int_0^t e^{t-x} f(x)dx.$$

Then,

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

where $f^{(j)}(t)$ is the j-th derivative of f with respect to x evaluated at t.

Proof Probably induction and integration by parts. I will fill this in later.

Chapter 2

Lindemann-Weierstrass Theorem

Theorem 2.0.1 *(Lindemann-Weierstrass Theorem) Suppose*

$$\alpha_1, \ldots, \alpha_n$$

are algebraic numbers that are linearly independent over $\mathbb{Q}$. Then,

$$e^{\alpha_1}, \ldots, e^{\alpha_n}$$

are linearly independent over the algebraic numbers. In other words, the extension field $\mathbb{Q}(e^{\alpha_1}, \ldots, e^{\alpha_n})$ has transcendence degree n over $\mathbb{Q}$.

Proof

Chapter 3

Gelfond-Schneider Theorem

3.1 Background

In this chapter, we will prove the Gelfond-Schneider Theorem. We will introduce two lemmas to help prove this.

Lemma 3.1.1 *(Siegel's Lemma) Consider the following system of m equations with n unknowns with $0 < m < n$:*

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

and $a_{i,j} \in \mathbb{Z}$. We write this compactly as $A\mathbf{x} = \mathbf{0}$. Let $a \in \mathbb{Z}$, $a > 0$, and $a \geq |a_{i,j}|$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$. Then, the system of equations has a nontrivial solution in $\mathbb{Z}^n$ for $\mathbf{x}$, such that for all x_j,

$$|x_j| < 1 + (na)^{\frac{m}{n-m}}.$$

Proof Let $\mathbf{x}$ be a column vector of height n and let $\mathbf{y}$ be a column vector of height m where $A\mathbf{x} = \mathbf{y}$. Let the entries of $\mathbf{y}$ be y_k where $1 \leq k \leq m$. A solution corresponds to finding an $\mathbf{x}$ such that $\mathbf{y}$ is the zero vector.

Def $\mathbf{x}$ is a **lattice point** if it is in $\mathbb{Z}^n$.

Note that if $\mathbf{x}$ is a lattice point, then since the entries of $A = (a_{i,j})$ are integers, $\mathbf{y}$ will also be a lattice point.

Let q be any positive integer. Let C be the n-dimensional hypercube centered at the origin of side length $2q$. Equivalently, this is all points with coordinates in $[-q, q]$ or points whose coordinate's absolute values are upper bounded by q. Now, let $\mathbf{x}$ range through all the $(2q+1)^n$ lattice points in C (each of the n coordinates is in $\{-q, -q+1, \ldots, 0, \ldots, q-1, q\}$, which is a set of $2q+1$ options). We can now upper bound each of the y_k coordinates in $\mathbf{y}$ via

$$\begin{aligned} |y_k| &= \left|\sum_{i=1}^{n} a_{k,j}x_j\right| \\ &\leq \sum_{i=1}^{n} |a_{k,j}x_j| \\ &= \sum_{i=1}^{n} |a_{k,j}||x_j| \\ &\leq \sum_{i=1}^{n} aq \\ &= naq \end{aligned}$$

Therefore, we know that the coordinates of $\mathbf{y}$ are in $\{-naq, -naq+1, \ldots, 0, \ldots, naq-1, naq+1\}$ and so there are $(2naq+1)^n$ possible lattice points $\mathbf{y}$ could be as $\mathbf{x}$ ranges through the lattice points of C.

As of now, we selected q to be any positive integer. However, to finish the proof of this lemma, we would like to use the pigeonhole principle to say that two of the inputs $\mathbf{x}', \mathbf{x}''$ in C correspond to outputs that coincide. That is to say, the $(2q+1)^n$ input points outnumber the $(2naq+1)^m$ possible output locations. In order to do this, we select q to be the unique integer such that:

$$(na)^{\frac{m}{n-m}} - 1 \leq 2q < (na)^{\frac{m}{n-m}} + 1.$$

This means $2q$ is the even integer in the interval of length 2. From the left inequality, we get $(na)^m \leq (2q+1)^{n-m}$. Now,

$$\begin{aligned} (2naq+1)^m &= \left(na \cdot \frac{2naq+1}{na}\right)^m \\ &= (na)^m \left(2q + \frac{1}{na}\right)^m \\ &< (na)^m (2q+1)^m \\ &\leq (2q+1)^{n-m}(2q+1)^m \\ &= (2q+1)^n. \end{aligned}$$

Thus, we get that there are solutions $\mathbf{x}' \neq \mathbf{x}''$ such that $A\mathbf{x}' = A\mathbf{x}''$. Thus, $A(\mathbf{x}' - \mathbf{x}'') = \mathbf{0}$. Note that if we let x_j', x_j'' denote the j-th coordinate of $\mathbf{x}', \mathbf{x}''$ respectively,

$$|x_j' - x_j''| \leq |x_j'| + |x_j''| \leq q + q = 2q < (na)^{\frac{m}{m-n}} + 1,$$

so $\mathbf{x}' - \mathbf{x}''$ is a nontrivial solution to $A\mathbf{x} = 0$ we were looking for.

♦

Now, we will generalize the previous lemma.

Lemma 3.1.2 *Consider the following system of p equations with q unknowns with $0 < p < q$:*

$$\begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,q} \\ \vdots & \ddots & \vdots \\ \alpha_{p,1} & \cdots & \alpha_{p,q} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_q \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

and $\alpha_{i,j} \in K$, a number field. We write this compactly as $A\boldsymbol{\xi} = \mathbf{0}$. Let $a \in \mathbb{Z}$, $a > 0$, and $a \geq \|\alpha_{i,j}\|$ for all $1 \leq i \leq p$ and $1 \leq j \leq q$. Then, there is a constant $c > 0$ depending only on K and independent of p, q, and $\alpha_{i,j}$ such that the system of equations has a nontrivial solution in K^n for $\boldsymbol{\xi}$, such that for all ξ_j,

$$\|\xi_j\| < c + c(cqa)^{\frac{p}{q-p}}.$$

Proof Let $h := [K : \mathbb{Q}]$ and let $\{\beta_1, \ldots, \beta_h\}$ be an integral basis for K. For $\alpha \in \mathcal{O}_K$ we can write

$$\alpha = g_1\beta_1 + \cdots + g_h\beta_h$$

for $g_j \in \mathbb{Z}$. Let $\sigma_1, \ldots, \sigma_h$ be the h $\mathbb{Q}$-fixing embeddings from $K \hookrightarrow \mathbb{C}$ mapping elements in K to their conjugates. Write the image of $\sigma_i(\alpha)$ as $\alpha^{(i)}$ and similarly for the β's. Apply the embeddings to the previous equation, we get h new equations for each $i \in \{1, \ldots, h\}$ of the form

$$\alpha^{(i)} = g_1\beta_1^{(i)} + \cdots + g_h\beta_h^{(i)}$$

We can represent this system of equations using matrices:

$$\begin{pmatrix} \beta_1^{(1)} & \ldots & \beta_h^{(1)} \\ \vdots & \ddots & \vdots \\ \beta_1^{(h)} & \ldots & \beta_h^{(h)} \end{pmatrix} \begin{pmatrix} g_1 \\ \vdots \\ g_h \end{pmatrix} = \begin{pmatrix} \alpha^{(1)} \\ \vdots \\ \alpha^{(h)} \end{pmatrix}$$

Note that since the β's form a basis, $\Delta(\{\beta_1, \ldots, \beta_h\}) \neq 0$. Therefore, we can solve for the g_j's as a linear combination of the $\alpha^{(i)}$'s:

$$g_j = \lambda_1\alpha^{(1)} + \cdots + \lambda_h\alpha^{(h)}.$$

Let c_1 be a particular positive integer that will upper bound the first minors of the determininant $\left|\left(\beta_j^{(i)}\right)\right|$, which consequently upper bounds $|\lambda_1|+\cdots+|\lambda_h|$. This will be explained later. Now, we will take the absolute values of both sides to get:

$$\begin{aligned}
|g_j| &= \left|\lambda_1\alpha^{(1)} + \cdots + \lambda_h\alpha^{(h)}\right| \\
&\leq \left|\lambda_1\alpha^{(1)}\right| + \cdots + \left|\lambda_h\alpha^{(h)}\right| \\
&\leq |\lambda_1|\left|\alpha^{(1)}\right| + \cdots + |\lambda_h|\left|\alpha^{(h)}\right| \\
&\leq |\lambda_1|\|\alpha\| + \cdots + |\lambda_h|\|\alpha\| \\
&= \|\alpha\|\left(|\lambda_1| + \cdots + |\lambda_h|\right) \\
&\leq c_1\|\alpha\|
\end{aligned}$$

Note that since

$$\begin{pmatrix} g_1 \\ \vdots \\ g_h \end{pmatrix} = \begin{pmatrix} \beta_1^{(1)} & \dots & \beta_h^{(1)} \\ \vdots & \ddots & \vdots \\ \beta_1^{(h)} & \dots & \beta_h^{(h)} \end{pmatrix}^{-1} \begin{pmatrix} \alpha^{(1)} \\ \vdots \\ \alpha^{(h)} \end{pmatrix}$$

the c_1 bound on the λ coefficients is dependent only on the entries of $\left(\beta_j^{(i)}\right)^{-1}$ and thus c_1 only depends on the number field K; it is independent of α. From this, we have shown that for all $\alpha \in \mathcal{O}_K$, there is a constant c_1 dependent only on K and independent of α, such that if we express α as a linear combination of the integral basis, all coefficients have absolute value upper bounded by $c_1\|\alpha\|$.

Now, we write every ξ_i in terms of the integral basis for $i = 1, \dots, q$:

$$\xi_i = x_{i,1}\beta_1 + \cdots + x_{i,h}\beta_h$$

Note that we now get the following equations (in matrix form):

$$\begin{pmatrix} x_{1,1} & \dots & x_{1,h} \\ \vdots & \ddots & \vdots \\ x_{q,1} & \dots & x_{q,h} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_h \end{pmatrix} = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_q \end{pmatrix}$$

Finding a bound on the ξ_i's is therefore related to finding a bound on the integers $x_{i,j}$. Consider the inital system of equations we were working with:

$$\begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,q} \\ \vdots & \ddots & \vdots \\ \alpha_{p,1} & \cdots & \alpha_{p,q} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_q \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

We will substitute the $\boldsymbol{\xi}$ vector with the previous matrix equation to get:

$$\begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,q} \\ \vdots & \ddots & \vdots \\ \alpha_{p,1} & \cdots & \alpha_{p,q} \end{pmatrix} \begin{pmatrix} x_{1,1} & \dots & x_{1,h} \\ \vdots & \ddots & \vdots \\ x_{q,1} & \dots & x_{q,h} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_h \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Written in equation form, we can think of this substitution as:

$$\sum_{i=1}^{q} \alpha_{k,i}\xi_i = \sum_{i=1}^{q}\sum_{j=1}^{h} \alpha_{k,i}\beta_j x_{i,j} = \mathbf{0}$$

Note that since $\alpha_{k,i}\beta_j$'s are in $\mathcal{O}_K$, we will also write them in terms of the integral basis and integer coefficients again. Specifically, we will write

$$\alpha_{k,i}\beta_j = \sum_{r=1}^{h} m_{k,i,j,r}\beta_r$$

for $k = 1, \ldots, p$, $i = 1, \ldots, q$, and $j = 1, \ldots, h$, where $m_{k,i,j,r} \in \mathbb{Z}$.

We therefore have converted our entire system of p equations to be written in terms of the integral basis and integer coefficients:

$$\sum_{i=1}^{q}\sum_{j=1}^{h}\sum_{r=1}^{h} m_{k,i,j,r} x_{i,j}\beta_r = \mathbf{0}$$

where $k = 1, \ldots, p$.

Note that since $\{\beta_1, \ldots, \beta_h\}$ forms an integral basis, it is $\mathbb{Q}$-linearly independent so if some linear combination of them sum to $\mathbf{0}$, it must be the case that the integer coefficients

$$\sum_{i=1}^{q}\sum_{j=1}^{h} m_{k,i,j,r} x_{i,j} = \mathbf{0}$$

for all $k = 1, \ldots, p$, $r = 1, \ldots, h$. The $x_{i,j}$'s are the unknown variables because they are the coefficients of the ξ_i's which were the original unknown variables, but the $m_{k,i,j,r}$'s come from the $\alpha_{k,i}$'s. We have therefore converted the problem from a system of equations in terms of algebraic integers to ph equations and qh unknowns ($x_{i,j}$'s) for integers. To finish this proof we will use Siegel's lemma so we need to find an upper bound on the absolute values of the coefficients $m_{k,i,j,r}$.

Recall For $\alpha \in \mathcal{O}_K$ expressed as a linear combination of the integral basis, the coefficients g_j have absolute value strictly upper bounded by $c_1\|\alpha\|$ for c_1 dependent only on K and not α.

If we use $\alpha_{k,i}\beta_j$ for α, the coefficients are exactly $m_{k,i,j,r}$ so we know that

$$\begin{aligned} |m_{k,i,j,r}| &< c_1\|\alpha_{k,i}\beta_j\| \\ &\leq c_1\|\alpha_{k,i}\|\|\beta_j\| \\ &\leq c_1 a\|\beta_j\| \end{aligned}$$

where a is the upper bound on the $\|\alpha_{k,i}\|$ in the statement of the lemma we are trying to prove.

Now, pick $c_2 \in \mathbb{R}$, such that

1. $\max_j c_1\|\beta_j\| \leq c_2 < 1 + \max_j c_1\|\beta_j\|$.
2. $c_2 a \in \mathbb{Z}$.

Note that this construction is possible since $a \in \mathbb{Q}$ and $a \geq 1$. While this construction makes c_2 technically depend on a, it is dependent on it in a trivial way and the magnitude of c_2 is bounded. This is because the bound to use to use the previous lemma is an integer. We will later further upper bound this with another constant so that the dependence disappears completely. Note that with c_2, we know that $|m_{k,i,j,r}| < c_1 a\|\beta_j\| \leq c_2 a$. We can now apply Siegel's lemma with m, n, a in its statement replaced by $ph, qh, c_2 a$ respectively to get that there is a nontrivial solution $x_{i,j}$ such that

$$|x_{i,j}| < 1 + (qhc_2 a)^{\frac{ph}{qh-ph}} = 1 + (hc_2 qa)^{\frac{p}{q-p}}$$

Now that we have our bound on $|x_{i,j}|$, recall that

$$\xi_i = x_{i,1}\beta_1 + \cdots + x_{i,h}\beta_h$$

so we can say

$$\|\xi_i\| < h \cdot \max_j \|\beta_j\| \left(1 + (hc_2 qa)^{\frac{p}{q-p}}\right).$$

Select c which exceeds $h\|\beta\|$ and hc_2 (note this makes c dependent only on K) so that

$$\|\xi_i\| < c + (cqa)^{\frac{p}{q-p}}.$$

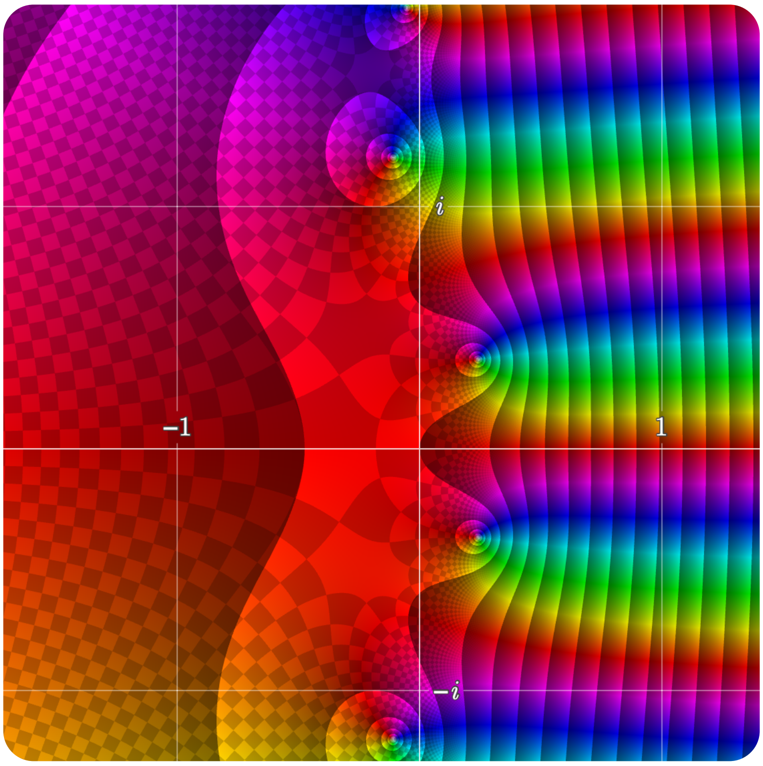
From the conclusion of Siegel's lemma, the system of equations has a nontrivial solution with this bound.

♦

3.2 Main Theorem

Theorem 3.2.1 *(Gelfond-Schneider Theorem) Let α and β be algebraic numbers such that $\alpha \notin \{0, 1\}$ and $\beta \in \mathbb{C} \setminus \mathbb{Q}$. Then, α^β is transcendental.*

Proof Assume for contradiction that as per the conditions of the theorem, $\alpha, \beta \in \mathbb{A}$ where $\alpha \neq 0, 1$ and $\beta \notin \mathbb{Q}$, but $\alpha^\beta \in \mathbb{A}$.



3.3 Consequences

Chapter 4

Baker's Theorem

Theorem 4.0.1 *(Baker's Theorem) Let* $\mathbb{L} = \left\{\lambda \in \mathbb{C} : e^{\lambda} \in \overline{\mathbb{Q}}\right\}$. *Then, if*

$$\lambda_1, \ldots, \lambda_n \in \mathbb{L}$$

are linearly independent over $\mathbb{Q}$*,*

Chapter 5

Schanuel's Conjecture

Conjecture 5.0.1 *(Schanuel's Conjecture) Suppose we have n complex numbers*

$$z_1, \ldots, z_n$$

that are linearly independent over $\mathbb{Q}$. Then, $\mathbb{Q}(z_1, \ldots, z_n, e^{z_1}, \ldots, e^{z_n})$ has transcendence degree at least n over $\mathbb{Q}$.

Index

algebraic, 2
algebraic closure, 2
algebraic extension, 2
algebraic integer, 3
algebraic number, 1
algebraic number field, 3
algebraically closed, 2
algebraically independent, 2

complex differentiable, 8
conjugates, 2

degree, 2
derivative, 8
differentiable, 8

entire, 8

holomorphic, 8

integral basis, 3

lattice point, 13
limit, 8
Liouville number, 8

minimal polynomial, 2

number field, 3

ring of integers, 3

transcendental, 2
transcendental number, 1

Vandermonde matrix, 5