

Undergraduate Research Opportunity
Programme in Science

Transcendental Numbers

The study of π

Lau Wee Lip Jonathan ¹

Special Programme in Science

Dr. Lim Chong Hai

Department of Mathematics

National University of Singapore

Academic Year (2000/2001)

Semester 2

Singapore, June 21, 2001

¹Matric No. : 992932R02

Acknowledgements

First of all, I would like to thank my supervisor Dr. Lim Chong Hai for teaching and guiding me throughout my entire project. He has taken great efforts to ensure that the subject material is made coherent to me, and has spent lots of time correcting and suggesting improvements to my draft. Without his help, this report would not have been completed. I am particularly grateful to him for teaching me the topics on Rings and Fields, a subject essential to the understanding of transcendental numbers.

Secondly, I would like to thank my father, Lau Cher Chye. Without his patient probing into the initial draft, the report would be extremely incoherent.

Lastly, I would like to thank all my course-mates who helped me in one way or another. In particular, I would like to thank Pang Chin How Jeffrey for guiding me through some difficult concepts in field theory and polynomial rings.

Lau Wee Lip Jonathan

June 21, 2001

Introduction

The purpose of this project is to study two classes that all complex numbers are classified into, algebraic numbers and transcendental numbers. A number α is said to be *algebraic* if it is a root of the polynomial

$$f(x) = a_n x^n + \cdots + a_1 x + a_0, \quad f(x) \not\equiv 0$$

with rational coefficients. To be more precise, a complex number α is said to be *algebraic of degree n* if it is the root of a polynomial of degree n over \mathbb{Q} . A complex number α (which maybe a real number) that does not satisfy the above prerequisite is said to be *transcendental*. In other words, a transcendental number is not a root of any polynomial with algebraic coefficients which is not identically zero. One particular point to note is that the class of transcendental real numbers are irrational. This is because given any $\alpha \in \mathbb{Q}$, where $\alpha = p/q$, $(p, q) \in \mathbb{Z} \times \mathbb{N}$, it satisfies the polynomial

$$f(x) = x - \frac{p}{q}, \quad f(x) \in \mathbb{Q}[x].$$

This gives us the first important criterion to determine if a number is algebraic or transcendental. However, this does not mean that all irrational numbers are transcendental. This is a sufficient but not necessary condition. For example, the number $\sqrt{2}$ is an algebraic number of degree 2, satisfying the function, $f(x) = x^2 - 2$. The main crux of this report is studying methods of approximation of real numbers by real numbers finally leading to the approximation of algebraic numbers by algebraic numbers. This gives us the criterion in which a number τ must satisfy before deeming it as algebraic or transcendental. In fact Liouville (1809 - 1882) gave a criterion that any algebraic number of degree n must satisfy. It can be shown that the criterion limits the extent to which a real algebraic number of degree n can be approximated by rational numbers. The project goes on further to prove the the transcendence of π . We shall employ the method attributed to Hermite which was further developed by Lindemann in 1882 to prove the transcendence of π . In the process, Lindemann also provided basis for the proof of the impossibility of squaring the circle. This will also be touch upon in the project.

Approximation of algebraic numbers

We approach the problem by approximation methods to identify whether a number is rational or irrational. We define a real number α to have rational approximations of order $\varphi(q)$ if there exist a constant $c > 0$ depending only on α and the function $\varphi(q)$ such that the inequality

$$0 < \left| \alpha - \frac{p}{q} \right| < c\varphi(q)$$

has infinitely many solutions $(p, q) \in \mathbb{Z} \times \mathbb{N}$. We can show that all irrational numbers have rational approximations of order $1/q^2$. We introduce the concept of *irreducible polynomials* and use it to provide the first rigorous proof, by Liouville in 1844, for the existence of transcendental numbers. He proved: *If α is a root of an irreducible polynomial with rational coefficients of degree n , $n \geq 1$, then there exist a constant $c(\alpha) > 0$ such that the following inequality holds for any rational integers p and q , with $q > 0$ and $p/q \neq \alpha$:*

$$\left| \alpha - \frac{p}{q} \right| > \frac{c(\alpha)}{q^n}.$$

This criterion is of such nature that we can easily construct real numbers that violates it for every $n > 1$. Any such number will have to be transcendental. It also tells us that an algebraic number cannot be approximated “too closely” by rational numbers. By approximating algebraic numbers using algebraic numbers, we can obtain a stronger form of Liouville Theorem, and do multiple approximations for algebraic numbers.

Transcendence of π and Applications

The theorem developed by Liouville is not sufficient to prove the transcendence of π . In order to do so, we need to first prove the irrationality of π and use contradiction to show that π is transcendental. We make use of *Euler's Identity*:

$$e^{i\pi} + 1 = 0$$

to prove this case. This result will be applied to prove the impossibility of *squaring the circle*, or *quadrature of the circle* as it is sometimes called. This is one of the three classical

problems in Greek mathematics which were extremely influential in the development of geometry, the other 2 being doubling the cube and trisecting an angle. To approach the problem, we introduce the concept of *constructible numbers* and prove that they form a subfield in \mathbb{R} . By showing that all constructibles are algebraic numbers, we prove the result.

Contents

Acknowledgements	A
Introduction	I
1 Approximation of real and algebraic numbers	1
1. Approximation of real numbers by algebraic numbers	1
2. Simultaneous Approximation	8
3. Approximation of algebraic numbers by rational numbers	17
4. Approximation of algebraic numbers by algebraic numbers	23
2 Transcendence of π and applications	43
1. Transcendence of π	43
2. Squaring the circle	53
Bibliography	64

Chapter 1

Approximation of real and algebraic numbers

1. Approximation of real numbers by algebraic numbers

The following notations will be used in the report: \mathbb{N} is the set of natural numbers, \mathbb{Z} is the ring of rational integers, \mathbb{Z}^+ is the set of nonnegative rational integers, \mathbb{Q} is the field of rational numbers, \mathbb{R} the field of real numbers, and \mathbb{C} the field of complex numbers, and \mathbb{R}^m denotes the m -dimensional real Euclidean space.

Let $\alpha \in \mathbb{R}$, $p \in \mathbb{Z}$, $q \in \mathbb{N}$ and let $\varphi(q)$ be a function which is positive for all $q \in \mathbb{N}$. We consider the following definition:

Definition 1. *A real number α has rational approximation p/q of order $\varphi(q)$ if there exists a constant $c > 0$ depending only on α and the function $\varphi(q)$ such that the inequality*

$$0 < \left| \alpha - \frac{p}{q} \right| < c\varphi(q)$$

has infinitely many solutions $(p, q) \in \mathbb{Z} \times \mathbb{N}$.

A common choice for $\varphi(q)$ is the power function

$$\varphi(q) = \frac{1}{q^v}, \quad v > 0.$$

In this case, our interest lies in finding the different possible values for c and v . We will show for the special case of $v = 1$ and use *Dirichlet's Theorem* to prove for the case $v \geq 2$.

Suppose that α is a rational number: $\alpha = a/b$, $(a,b) \in \mathbb{Z} \times \mathbb{N}$, $\gcd(a,b) = 1$. From *Density Theorem*¹ we deduced that for any $q \in \mathbb{N}$ there exist $p \in \mathbb{Z}$ such that:

$$\frac{p}{q} < \frac{a}{b} < \frac{p+1}{q}.$$

Then we obtain the following relation:

$$0 < \left| \alpha - \frac{p}{q} \right| \leq \frac{1}{q}. \quad (1.1)$$

By allowing q to take on different values, we find that there exist an infinite set of fractions p/q , $p/q \neq \alpha$ that satisfies (1.1).

However for any fraction p/q , $p/q \neq a/b$, we get

$$\left| \alpha - \frac{p}{q} \right| = \left| \frac{a}{b} - \frac{p}{q} \right| = \frac{|aq - bp|}{bq} \geq \frac{1}{bq}. \quad (1.2)$$

The above equation is obvious since $a/b \neq p/q$, where $(a,b), (p,q) \in \mathbb{Z} \times \mathbb{N}$, therefore $|aq - bp| \geq 1$.

From (1.2), it can be seen that for any constant c with $0 < c < 1/b$ the inequality

$$0 < \left| \alpha - \frac{p}{q} \right| < \frac{c}{q}$$

has no solution p/q . This will lead us to the following theorem.

Theorem 1. *If $\varphi(q) > 0$ for all $q \in \mathbb{N}$, the inequality*

$$0 < \left| \alpha - \frac{p}{q} \right| < c\varphi(q)$$

has infinitely many solutions $(p,q) \in \mathbb{Z} \times \mathbb{N}$, and

$$\lim_{q \rightarrow +\infty} q\varphi(q) = 0$$

then α is an irrational number.

¹Bartle and Sherbert : Introduction to Real Analysis 3rd Edition Pg 42

Proof. Suppose α is a rational number. Then there exist $(a,b) \in \mathbb{Z} \times \mathbb{N}$ such that $\alpha = a/b$. Given that,

$$0 < \left| \alpha - \frac{p}{q} \right| < c\varphi(q)$$

for infinitely many solutions $(p,q) \in \mathbb{Z} \times \mathbb{N}$, and

$$\left| \alpha - \frac{p}{q} \right| = \left| \frac{aq - bp}{bq} \right| \geq \frac{1}{bq}.$$

We obtain the following result

$$c\varphi(q) > \frac{1}{qb} \implies q\varphi(q) > \frac{1}{cb}$$

for all $q \in \mathbb{N}$. This tells us that $q\varphi(q)$ is a divergent function or it does not converge to 0. However recall that:

$$\lim_{q \rightarrow +\infty} q\varphi(q) = 0.$$

This creates a contradiction in our argument. Hence we are forced to conclude that α must be irrational. □

One thing to note is that the order of approximation by rational numbers is not unique, i.e. for different real numbers the order of approximation is different. The next theorem aims to show that for any irrational number, it has rational approximation p/q of order $1/q^2$. To show this we rely on *Dirichlet's pigeon-hole principle*.

Definition 2 (Dirichlet's Pigeon-Hole Principle). *If n objects are placed in m boxes (or pigeon-holes) and if $n > m$, then some box will contain at least two objects.*

Mathematically, the principle asserts that if a set with n elements is the union of m of its subsets and if $n > m$, then some subset has more than one element.

Dirichlet's Theorem. *If $\alpha \in \mathbb{R}$, $t \in \mathbb{R}$, and $t \geq 1$, then there exist $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ such that*

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{qt}, \quad 0 < q \leq t. \tag{1.3}$$

Proof. Let $a \in \mathbb{R}$. Let $\{a\}$ and $[a]$ denote the fractional part and integer part of a respectively i.e. $a = [a] + \{a\}$. We define $T := [t] + 1$ and consider the $T + 1$ numbers:

$$0 = \{\alpha \cdot 0\}, \{\alpha \cdot 1\}, \dots, \{\alpha \cdot (T - 1)\}, 1. \tag{1.4}$$

where

$$\{\alpha x\} = \alpha x - [\alpha x], \quad 0 \leq \{\alpha x\} < 1, \quad (x = 0, 1, \dots, T-1).$$

Note : All the numbers in (1.4) are contained in the interval $0 \leq y \leq 1$. We divide this interval into T equal parts of length $1/T$:

$$\frac{k}{T} \leq y < \frac{k+1}{T}, \quad (k = 0, 1, \dots, T-2), \quad \frac{T-1}{T} \leq y \leq 1. \quad (1.5)$$

Each of the points in (1.4) must lie in *one of the intervals* in (1.5). We note that the number of points are *greater than* the number of intervals. Hence by *Dirichlet's Pigeon Hole Principle*, we come to the conclusion that one or more intervals in (1.5) must contain 2 or more points in (1.4). We then consider such an interval. Two possible scenarios will occur:

1. The interval is not the one on the extreme right. Suppose it contains the points $\{\alpha x_1\}$ and $\{\alpha x_2\}$ where $x_2 > x_1$. Then we have

$$|\{\alpha x_2\} - \{\alpha x_1\}| = |\alpha(x_2 - x_1) - ([\alpha x_2] - [\alpha x_1])| < \frac{1}{T} < \frac{1}{t}.$$

We define $q = x_2 - x_1$ and $p = [\alpha x_2] - [\alpha x_1]$. From the definition of T we see that $0 < q \leq T - 1 \leq t$. We then obtain the following result:

$$\begin{aligned} |\alpha q - p| &< \frac{1}{t}, \quad 0 < q \leq t. \\ \therefore \left| \alpha - \frac{p}{q} \right| &< \frac{1}{qt}. \end{aligned} \quad (1.6)$$

Further manipulation gives us the result in (1.3).

2. The interval is on the extreme right of (1.5). Suppose it contains the point $\{\alpha x_1\}$ as well as 1, where $x_1 \neq 0$. Then

$$|\{\alpha x_1\} - 1| = |\alpha x_1 - ([\alpha x_1] + 1)| \leq \frac{1}{T} < \frac{1}{t}.$$

As in the previous case we let $q = x_1$ and $p = [\alpha x_1] + 1$. We then get $0 < q \leq T - 1 \leq t$, and the equation (1.6) from which (1.3) will follow. \square

We shall consider 2 possible cases arising from (1.3):

Suppose $\alpha \in \mathbb{Q}$, $\alpha = a/b$, $(a, b) \in \mathbb{Z} \times \mathbb{N}$ and $\gcd(a, b) = 1$. Consider $t \geq b$. Recall from equation (1.2)

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

so if $t \geq b$ this means

$$|\alpha - \frac{p}{q}| \geq \frac{1}{bq} \geq \frac{1}{tq}.$$

Hence, together with (1.6), for (1.3) to be true, the only possible solution is $\alpha = p/q$.

Suppose $t < b$, then $q \leq t < b$. This tells us that there is a bound on the value of q . Hence there can only be a finite set of solutions (p, q) for (1.3). Thus p/q can be used as an approximation for α .

We will now prove a corollary that arises from (1.3)

Corollary. Suppose α is irrational, then there are infinitely many solutions (p, q) satisfying $|\alpha q - p| > 0$.

Proof. Suppose that there are finitely many solutions i.e $S = \{(p_1, q_1), (p_2, q_2), \dots, (p_n, q_n)\}$ is the solution set for (1.3). Then for all solutions in the set S , there exist a $(p_k, q_k) \in S$, $1 \leq k \leq m$ such that

$$|\alpha - \frac{p_k}{q_k}| = \beta$$

where β is the smallest value obtain from the solution set S when substituted into equation (1.3). Hence we obtain the following:

$$|\alpha - \frac{p}{q}| \geq \beta > 0 \quad \forall (p, q) \in S.$$

From (1.6), as $t \rightarrow \infty$, $\frac{1}{t} \rightarrow 0$. Hence by squeeze theorem we find that $|\alpha q - p| \rightarrow 0$. This would imply that there exist a $(p, q) \in \mathbb{Z} \times \mathbb{N}$ such that $\alpha = p/q$. However α is irrational. Hence this provides a contradiction. \square

Theorem 2. For any irrational number $\alpha \in \mathbb{R}$, the inequality

$$|\alpha - \frac{p}{q}| < \frac{1}{q^2} \tag{1.7}$$

has infinitely many solutions $(p, q) \in \mathbb{Z} \times \mathbb{N}$, and hence the set of denominators of the solutions to (1.7) is unbounded.

Proof. We apply Dirichlet's theorem to prove this problem. From (1.3) we note that for any solution (p, q) we have

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

From the corollary above, we see that if α is irrational, we have infinite solutions for (1.3). Hence (1.7) has an infinite set of solutions $(p, q) \in \mathbb{Z} \times \mathbb{N}$. Thus from **Definition 1** we see that any irrational number has approximations of order $1/q^2$. \square

From (1.7) we get

$$|\alpha q^2 - pq| < 1 \implies |\alpha q^2 - pq| = |\theta_q|, \quad |\theta_q| < 1.$$

Thus,

$$\begin{aligned} \alpha q^2 - pq &= \theta_q \\ \alpha &= \frac{p}{q} + \frac{\theta_q}{q^2} \end{aligned} \tag{1.8}$$

where $(p, q) \in \mathbb{Z} \times \mathbb{N}$, $|\theta_q| < 1$. Hence for any irrational number α , we can represent it in the form of (1.8).

Example 1. Consider the irrational number

$$\alpha = \sum_{k=1}^{\infty} \frac{1}{a^{k!}}, \quad a \in \mathbb{N}, \quad a \geq 2. \tag{1.9}$$

We define

$$\frac{p_n}{q_n} = \sum_{k=1}^n \frac{1}{a^{k!}}, \quad \text{where } q_n = a^{n!}, \quad p_n \in \mathbb{N}, \quad n = 1, 2, \dots$$

Then

$$r_n = \alpha - \frac{p_n}{q_n} = \sum_{k=1}^{\infty} \frac{1}{a^{k!}} - \sum_{k=1}^n \frac{1}{a^{k!}} = \sum_{k=n+1}^{\infty} \frac{1}{a^{k!}} > 0 \quad n = 1, 2, \dots$$

Hence

$$\begin{aligned} r_n &= \frac{1}{a^{(n+1)!}} + \frac{1}{a^{(n+2)!}} + \dots \\ &= \frac{1}{a^{(n+1)!}} \left(1 + \frac{1}{a^{(n+2)! - (n+1)!}} + \dots \right). \end{aligned}$$

We note that

$$(n+k)! - (n+1)! = (n+1)! \{(n+2) \dots (n+k) - 1\} > k, \quad k \in \mathbb{N}, \quad k \neq 1.$$

Hence,

$$\begin{aligned}
r_n &< \frac{1}{a^{(n+1)!}} \left(1 + \frac{1}{a} + \frac{1}{a^2} + \cdots \right) \\
&< \frac{1}{1 - \frac{1}{a}} \frac{1}{a^{(n+1)!}} \\
&= \frac{a}{a-1} \frac{1}{(a^{n!})^{n+1}} \\
&= \frac{a}{a-1} \frac{1}{(q_n)^{n+1}} \\
&= \frac{a}{a-1} \frac{1}{a^{n!}} \frac{1}{(q_n)^n} \\
&= \frac{1}{a^{n!-1}(a-1)} \frac{1}{(q_n)^n} \\
&< \frac{1}{(q_n)^n} \quad a \in \mathbb{N}, \quad a \geq 2.
\end{aligned}$$

Thus

$$0 < \alpha - \frac{p_n}{q_n} < \frac{1}{(q_n)^n} \quad n = 1, 2, \dots$$

This gives us the conclusion that for any non-negative $\nu \in \mathbb{R}$ the inequality

$$0 < \left| \alpha - \frac{p}{q} \right| < \frac{1}{q^{[\nu]+1}} < \frac{1}{q^\nu} \quad (1.10)$$

has a solution $(p, q) \in \mathbb{Z} \times \mathbb{N}$ because $[\nu] + 1 \in \mathbb{N}$.

This implies that for any solution for a given value of ν in (1.10) it works for a smaller value of ν as well. Hence for any $\nu \in \mathbb{R}$, $\nu > 0$, (1.10) has infinitely many solutions. \square

Theorem 3. *Given any positive function $\varphi(q)$ on the natural numbers, there exists an irrational number $\alpha \in \mathbb{R}$ such that the inequality*

$$\left| \alpha - \frac{p}{q} \right| < \varphi(q)$$

has infinitely many solutions $(p, q) \in \mathbb{Z} \times \mathbb{N}$.

Proof. Let us define a sequence $\{\gamma_n\}$, $\gamma_n \in \mathbb{R}$, $\gamma_n \geq 0$ such that the sequence of natural numbers

$$m_1 = 1 \quad m_n = \left[\log_2 \left(\frac{1}{\varphi(2^{m_{n-1}})} + \gamma_{n-1} \right) \right] + 1, \quad n = 2, 3, \dots$$

satisfies the condition $m_{n+1} > 2m_n$. We define

$$\alpha = \sum_{k=1}^{\infty} \frac{(-1)^k}{2^{m_k}} \quad (1.11)$$

and

$$q_n = 2^{m_n}, \quad p_n = q_n \sum_{k=1}^n \frac{(-1)^k}{2^{m_k}}, \quad n = 1, 2, \dots$$

We note that (1.11) is an alternating series with decreasing terms. Hence we get the inequalities

$$\begin{aligned} 0 &< \frac{1}{2^{m_{n+1}}} - \frac{1}{2^{m_{n+2}}} < \left| \alpha - \frac{p_n}{q_n} \right| < \frac{1}{2^{m_{n+1}}} \\ &< \frac{1}{2^{\log_2(1/\varphi(2^{m_n}))}} = \varphi(2^{m_n}) = \varphi(q_n), \quad n = 1, 2, \dots \end{aligned} \quad (1.12)$$

Since $m_{n+1} > 2m_n$, it follows from (1.12) that

$$0 < \left| \alpha - \frac{p_n}{q_n} \right| < \frac{1}{2^{m_{n+1}}} < \frac{1}{2^{2m_n}} = \frac{1}{q_n^2}$$

From theorem 1, α is irrational and we get Theorem 3 from (1.12). \square

2. Simultaneous Approximation

In this section, we will develop techniques for doing simultaneous approximation, that is, given a set of real numbers, is there a technique to allow us to detect whether if there are any irrational numbers in the set. Using methods developed in section 1, we will attempt to solve this problem. This will allow us to solve problems with greater ease especially computation types.

Theorem 4 (Generalized Dirichlet). *If $\alpha_1, \dots, \alpha_m$ are real numbers, $m \geq 1$, and if $t \in \mathbb{R}$, then there exist $a_1, \dots, a_m, b \in \mathbb{Z}$ such that*

$$\left| a_1\alpha_1 + \dots + a_m\alpha_m - b \right| < \frac{1}{t^m}, \quad 0 < \max_{1 \leq k \leq m} |a_k| < t. \quad (1.13)$$

Proof. We define $T := [t] + 1$ and consider the $T^m + 1$ numbers : the T^m fractional parts

$$\gamma = \{a_1\alpha_1 + \dots + a_m\alpha_m\}, \quad 0 \leq \gamma < 1, \quad \text{and} \quad 1. \quad (1.14)$$

We note that each a_k , where $0 \leq k \leq m$, has a choice of T numbers, i.e $0, \dots, T-1$. Since there are m of such a_k , the total number of changes possible is $T \cdot T \dots T \cdot T = T^m$. Given that γ is a fraction and 1 is also included, therefore there are $T^m + 1$ numbers. We divide the interval $0 \leq y \leq 1$ into T^m parts of length $1/T^m$.

$$\frac{k}{T^m} \leq y < \frac{k+1}{T^m}, \quad (k = 0, 1, \dots, T^m - 2), \quad \frac{T^m - 1}{T^m} \leq y \leq 1. \quad (1.15)$$

We make the observation that all the points in (1.14) are included in the range in (1.15). Since the number of intervals is lesser than the number of points, we apply the Pigeon Hole Principle to arrive at the conclusion that there must exist an interval in which 2 or more points are contained in it. We consider the following two cases:

1. The interval containing two of such points are not on the extreme right. Suppose the interval contains the points

$$\gamma_1 = \{a'_1\alpha_1 + \dots + a'_m\alpha_m\} \quad \text{and} \quad \gamma_2 = \{a''_1\alpha_1 + \dots + a''_m\alpha_m\}$$

where $\gamma_2 > \gamma_1$. Then

$$\begin{aligned} |\{a''_1\alpha_1 + \dots + a''_m\alpha_m\} - \{a'_1\alpha_1 + \dots + a'_m\alpha_m\}| &= |a_1\alpha_1 + \dots + a_m\alpha_m - b| \\ &< \frac{1}{T^m} \\ &= \frac{1}{([t] + 1)^m} < \frac{1}{t^m} \end{aligned} \quad (1.16)$$

where

$$\begin{aligned} a_k &= a''_k - a'_k \quad k = 1, \dots, m; \\ b &= [a''_1\alpha_1 + \dots + a''_m\alpha_m] - [a'_1\alpha_1 + \dots + a'_m\alpha_m] \end{aligned} \quad (1.17)$$

in which $0 < \max_{1 \leq k \leq m} |a_k| \leq T - 1 = [t] \leq t$

2. The interval considered is on the extreme right:

Suppose the interval contains the points

$$\gamma = \{a_1\alpha_1 + \dots + a_m\alpha_m\}, \quad 0 < \max_{1 \leq k \leq m} |a_k| \leq t \quad \text{and} \quad 1. \quad (1.18)$$

Then

$$\begin{aligned} |\{a_1\alpha_1 + \dots + a_m\alpha_m\} - 1| &= |a_1\alpha_1 + \dots + a_m\alpha_m - b| \\ &\leq \frac{1}{T^m} \\ &< \frac{1}{t} \end{aligned} \quad (1.19)$$

where $b = [a_1\alpha_1 + \dots + a_m\alpha_m] + 1$ which gives the form of equation (1.13). It is interesting to note that Dirichlet's Formula in section 1 is in fact a special case of the generalized Dirichlet's formula, i.e. $m = 1$. \square

Kronecker's Theorem. *If $\alpha_1, \dots, \alpha_m$ are arbitrary real numbers, $m \geq 1$, and if $t \in \mathbb{N}$, then there exist $p_1, \dots, p_m \in \mathbb{Z}$ and $q \in \mathbb{N}$ such that*

$$|\alpha_k - \frac{p_k}{q}| < \frac{1}{qt}, \quad (k = 1, \dots, m), \quad 0 < q \leq t^m. \quad (1.20)$$

Proof. In \mathbb{R}^m we consider the points

$$M_x(\{\alpha_1 x\}, \dots, \{\alpha_m x\}), \quad (x = 0, 1, \dots, t^m) \quad (1.21)$$

$$\text{where } \{\alpha_k\} = \alpha_k x - [\alpha_k x], \quad 0 \leq \{\alpha_k x\} < 1, \quad (k = 1, \dots, m);$$

and we consider the unit hypercube; i.e. the set of points

$$M(y_1, y_2, \dots, y_m), \quad 0 \leq y_k < 1, \quad k = 1, \dots, m; \quad (1.22)$$

On each axis, we divide each $0 \leq y_k \leq 1$, $k = 0, 1, \dots, m$ segments into t parts of $1/t$ and we use hyperplanes to further subdivide it into t^m small cubes., each of which consists of all the points $\mathbb{Q}(z_1, \dots, z_m)$ whose coordinates satisfy the equation

$$\frac{k_1}{t} \leq z_1 \leq \frac{k_1 + 1}{t}, \dots, \frac{k_m}{t} \leq z_m \leq \frac{k_m + 1}{t} \quad (1.23)$$

where the $k_i, i = 1, \dots, m$, are an m -tuple of integers from $\{0, 1, \dots, t - 1\}$. Since all the points in (1.21) are contained in the hyper unit cube, hence *each point* must be contained in *one of the smaller hyper cubes*. Furthermore, we note that the number of points exceed the number of hyper cubes. Hence by Dirichlet's Pigeon Hole Principle, there exist a hyper cube such that it contains at least 2 of the points in (1.21). We consider 2 such points:

$$M_{x_1}(\{\alpha_1 x_1\}, \dots, \{\alpha_m x_1\}), \quad M_{x_2}(\{\alpha_1 x_2\}, \dots, \{\alpha_m x_2\}), \quad x_2 > x_1.$$

Since these two points lie in the same hyper cube, it follows that the absolute difference between *any two corresponding coordinates* must be lesser than $1/t$. For example consider two points contained in a unit square. Then the absolute difference of each coordinate must be lesser than 1.

From above, we obtain the following result

$$|\{\alpha_k x_2\} - \{\alpha_k x_1\}| < \frac{1}{t}, \quad (k = 1, \dots, m).$$

Defining $q := x_2 - x_1$ and $p_k = [\alpha_k x_2] - [\alpha_k x_1]$ for $k = 1, \dots, m$, we obtain

$$|\alpha_k q - p_k| < \frac{1}{t}, \quad (k = 1, \dots, m), \quad 0 < q \leq t^m \quad (1.24)$$

from which equation (1.20) follows. Note that the case of which $m = 1$ gives us the Dirichlet's Theorem. \square

Corollary. *Under the assumptions in Kronecker's Theorem, if at least one of the $\alpha_1, \dots, \alpha_m$ is irrational, then the inequalities*

$$|\alpha_k - \frac{p_k}{q}| < q^{-1-\frac{1}{m}}, \quad (k = 1, \dots, m)$$

have an infinite set of solutions $p_1, \dots, p_m \in \mathbb{Z}$, $q \in \mathbb{N}$.

Proof. Let $t \in \mathbb{N}$ and from Kronecker's Theorem, we know that there exist $p_1, \dots, p_m \in \mathbb{Z}$ and $q \in \mathbb{N}$ such that

$$|\alpha_k - \frac{p_k}{q}| < \frac{1}{qt}, \quad (k = 1, \dots, m), \quad 0 < q \leq t^m.$$

Since

$$\begin{aligned} q &\leq t^m, \\ \implies q^{\frac{1}{m}} &\leq t. \\ \implies \frac{1}{q^{1/m}} &\geq \frac{1}{t}. \end{aligned}$$

Hence from Kronecker's Theorem, we have

$$|\alpha_k - \frac{p_k}{q}| < \frac{1}{qt} < \frac{1}{q^{1+1/m}}.$$

Suppose one or more α_k is irrational, where $k = 1, \dots, m$. Let us assume that there exist finite set of solutions $p_1, \dots, p_m \in \mathbb{Z}$, $q \in \mathbb{N}$. Without loss of generality, let us denote α_i to be irrational, where $1 \leq i \leq m$. Then

$$|\alpha_i - \frac{p_i}{q}| < \frac{1}{q^{1+1/m}}$$

has finitely many solutions $(p_i, q) \in \mathbb{Z} \times \mathbb{N}$.

However, let us recall from Theorem 2 that if a number $\alpha \in \mathbb{R}$ is irrational, then the inequality

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

has infinitely many solutions $(p, q) \in \mathbb{Z} \times \mathbb{N}$. Let us further note that

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

Hence,

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{q^{1+1/m}}$$

must have infinitely many solutions. Thus our previous assumption is wrong and we deduced the condition stated by the corollary. \square

Theorem 5. *If $\alpha_1, \dots, \alpha_m$ are arbitrary complex numbers, $m \leq 2$, and if $H \in \mathbb{N}$, then there exist numbers*

$$a_k \in \mathbb{Z}, \quad |a_k| \leq H, \quad (k = 1, \dots, m), \quad \max_{1 \leq k \leq m} |a_k| > 0 \quad (1.25)$$

such that the linear form

$$L = a_1\alpha_1 + \dots + a_m\alpha_m \quad (1.26)$$

satisfies the inequality

$$|L| \leq cH^{1-\tau m}, \quad (1.27)$$

where $\tau = 1$ if all the α_k are real and $\tau = 1/2$ if one or more of them is complex, and

$$c = \begin{cases} \sum_{k=1}^m |a_k| & , \quad \text{if } \tau = 1 \\ \sqrt{2} \sum_{k=1}^m |a_k| & , \quad \text{if } \tau = \frac{1}{2} \end{cases}$$

Proof. We consider the 2 cases : $H = 1$ and $H \geq 2$.

Case (1) : $H = 1$

From (1.26)

$$L = a_1\alpha_1 + \dots + a_m\alpha_m.$$

Then

$$\begin{aligned} |L| &= |a_1\alpha_1 + \cdots + a_m\alpha_m| \\ &\leq |a_1\alpha_1| + |a_2\alpha_2| + \cdots + |a_m\alpha_m|. \quad (\Delta \text{ inequality}) \end{aligned}$$

Since $|a_k| \leq 1$, ($k = 1, \dots, m$), we have

$$|L| \leq |\alpha_1| + |\alpha_2| + \cdots + |\alpha_m|.$$

Suppose all the α_k are real, then we get the form

$$\begin{aligned} |L| &\leq \sum_{k=1}^m |\alpha_k| \\ &= \sum_{k=1}^m |\alpha_k| (1)^{1-\tau m} \\ &= c. \end{aligned}$$

Hence the first part is shown.

Suppose some of the α_k are complex. We let $\alpha_k = \text{Re}(\alpha_k) + i\text{Im}(\alpha_k)$. Since $\mathbb{R} \in \mathbb{C}$, thus for each α_k ,

$$|\alpha_k| \leq |\text{Re}(\alpha_k)| + |\text{Im}(\alpha_k)|.$$

Recall from complex analysis, we have the following inequality:

$$\sqrt{2}|\alpha_k| \geq |\text{Re}(\alpha_k)| + |\text{Im}(\alpha_k)|.$$

Hence from (1.26) we get

$$\begin{aligned} |L| &\leq \sum_{k=1}^m |\alpha_k| \\ &\leq \sum_{k=1}^m \{|\text{Re}(\alpha_k)| + |\text{Im}(\alpha_k)|\} \\ &\leq \sqrt{2} \sum_{k=1}^m |\alpha_k| \\ &= \sqrt{2} \sum_{k=1}^m |\alpha_k| (1)^{1-\tau m}, \\ &= c, \quad \text{if } T = \frac{1}{2}. \end{aligned}$$

Case (2) : $H \geq 2$

Let us consider all possible linear forms of (1.26) where a_k independently run through all the integers satisfying the inequality below:

$$|a_k| \leq \left\lfloor \frac{H}{2} \right\rfloor, \quad (k = 1, \dots, m).$$

We note that $a_k \in \mathbb{Z}$. Hence the total number of such forms L is equal to

$$\left(2 \left\lfloor \frac{H}{2} \right\rfloor + 1 \right)^m. \quad (1.28)$$

Since we know that $|a_k| \leq \left\lfloor \frac{H}{2} \right\rfloor$, we get

$$\begin{aligned} |L| &= |a_1\alpha_1 + \dots + a_m\alpha_m| \\ &\leq |a_1||\alpha_1| + \dots + |a_m||\alpha_m|, \quad (\Delta \text{ inequality}) \\ &\leq (|\alpha_1| + |\alpha_2| + \dots + |\alpha_m|) \left\lfloor \frac{H}{2} \right\rfloor \\ &= \left(\sum_{k=1}^m |\alpha_k| \right) \left\lfloor \frac{H}{2} \right\rfloor. \end{aligned}$$

Hence $|L| \leq \gamma \left\lfloor \frac{H}{2} \right\rfloor$, $\gamma = \sum_{k=1}^m |\alpha_k|$. (1.29)

Excluding the case $\gamma = 0$, we consider the following 2 cases.

Case (1) : Suppose all of the $\alpha_1, \dots, \alpha_m$ are real numbers. Then all the values of L will be contained in the interval with endpoints $\pm\gamma \left\lfloor \frac{H}{2} \right\rfloor$, which has length $2\gamma \left\lfloor \frac{H}{2} \right\rfloor$. Let us divide the interval into

$$\left(2 \left\lfloor \frac{H}{2} \right\rfloor + 1 \right)^m - 1 \quad (1.30)$$

subintervals. We see that the total number of forms of L is greater than the number of subintervals. Hence by Dirichlet's Pigeon Hole Theorem, there exist a subinterval containing the values of two different forms L . We define the two forms of L as:

$$L' = a'_1\alpha_1 + \dots + a'_m\alpha_m, \quad L'' = a''_1\alpha_1 + \dots + a''_m\alpha_m. \quad (1.31)$$

Then from Dirichlet's Pigeon Hole theorem, we obtain the following inequality:

$$|L' - L''| \leq \frac{2\gamma \left\lfloor \frac{H}{2} \right\rfloor}{\left(2 \left\lfloor \frac{H}{2} \right\rfloor + 1 \right)^m - 1}.$$

Let us consider 2 sub-cases. Suppose that H is an even number, then $2 \left\lfloor \frac{H}{2} \right\rfloor = H$, and

$$\begin{aligned}
|L' - L''| &\leq \frac{\gamma H}{(H+1)^m - 1} \\
&= \frac{\gamma H}{1 + H + H^2 + \dots + H^m - 1} \quad (\text{binomial expansion}) \\
&= \frac{\gamma H}{H(1 + H + \dots + H^{m-1})} \\
&= \frac{\gamma}{(H+1)^{m-1}} \\
&< \frac{\gamma}{H^{m-1}} = \gamma H^{1-m}.
\end{aligned}$$

Suppose H is an odd number, then $2 \left\lfloor \frac{H}{2} \right\rfloor = H - 1$, and

$$\begin{aligned}
|L' - L''| &\leq \frac{\gamma(H-1)}{H^m - 1} \\
&= \frac{\gamma(H-1)}{(H-1)(H+1)^{m-1}} \quad (\text{binomial expansion}) \\
&= \frac{\gamma}{(H+1)^{m-1}} \\
&< \frac{\gamma}{H^{m-1}} = \gamma H^{1-m}.
\end{aligned}$$

Thus,

$$|L' - L''| < \gamma H^{1-m}.$$

We define $a_k = a'_k - a''_k$, $k = 1, \dots, m$. Then the form

$$L = L' - L'' = a_1 \alpha_1 + \dots + a_m \alpha_m$$

satisfies the conditions

$$|a_k| \leq |a'_k| + |a''_k| \leq 2 \left\lfloor \frac{H}{2} \right\rfloor \leq H, \quad \sum_{k=1}^m |a_k| > 0$$

and

$$|L| < c H^{1-m}, \quad c = \gamma.$$

Case (2) : At least one if the numbers a_1, \dots, a_m is complex.

Let us consider a square whose sides are of length $2\gamma \left\lfloor \frac{H}{2} \right\rfloor$, centered at the origin. We note that any form of L will be contained in the square. We divide the side of the square into M equal segments, with M satisfying the inequality:

$$\left(2 \left\lfloor \frac{H}{2} \right\rfloor + 1\right)^{m/2} - 1 \leq M < \left(2 \left\lfloor \frac{H}{2} \right\rfloor + 1\right)^{m/2}. \quad (1.32)$$

Using the segmented points, we draw lines to divide the square into M^2 parts. From (1.32) we arrive at the inequality

$$M^2 < \left(2 \left\lfloor \frac{H}{2} \right\rfloor + 1 \right)^m.$$

We see that the number of points is lesser than the number of squares, hence by the pigeon hole principle, we see that there must exist a square such that it contains two such forms of L . Note the distance between any two such forms is lesser than the length of the diagonal. Let us consider one such small square. The length of such a square is

$$\frac{2\gamma \left\lfloor \frac{H}{2} \right\rfloor}{M}.$$

Using Pythagorean Theorem, we know that the length of the diagonal must be

$$\frac{2\sqrt{2}\gamma \left\lfloor \frac{H}{2} \right\rfloor}{M}.$$

Hence for any $H \geq 2$ and $m \geq 2$, we obtain the following inequality:

$$|L' - L''| \leq \frac{2\sqrt{2}\gamma \left\lfloor \frac{H}{2} \right\rfloor}{M} \leq \frac{2\sqrt{2}\gamma \left\lfloor \frac{H}{2} \right\rfloor}{(2 \left\lfloor \frac{H}{2} \right\rfloor + 1)^{m/2} - 1}. \quad (1.33)$$

Suppose H is an even number, then $2\lfloor H/2 \rfloor = H$ and for $m \geq 2$, we applying binomial theorem as in the real case to obtain:

$$|L' - L''| \leq \frac{\sqrt{2}\gamma H}{(H + 1)^{m/2} - 1} \leq \sqrt{2}\gamma H^{1-\frac{m}{2}}.$$

Given that $m \geq 2$, $H \geq 1$,

$$\begin{aligned} H &\leq H^{m/2} & (\because \frac{m}{2} \geq 1) \\ \frac{1}{H^{m/2}} &\leq \frac{1}{H} \\ 1 - \frac{1}{H} &\leq 1 - \frac{1}{H^{m/2}}. \end{aligned}$$

Thus, we arrive at the inequality

$$\frac{1 - H^{-1}}{1 - H^{-m/2}} \leq 1.$$

Hence if H is odd, then $2\lfloor H/2 \rfloor = H - 1$ and we can arrive at the conclusion that

$$|L' - L''| \leq \frac{\sqrt{2}\gamma(H - 1)}{H^{m/2} - 1} = \sqrt{2}\gamma H^{1-\frac{m}{2}} \frac{1 - H^{-1}}{1 - H^{-m/2}} \leq \sqrt{2}\gamma H^{1-\frac{m}{2}}.$$

In both cases, we define $c = \sqrt{2}\gamma$ to arrive at the inequality (1.27). \square

3. Approximation of algebraic numbers by rational numbers

In this section, we will analyze the approximation of algebraic numbers by rational numbers. But first, we shall prove some lemmas and state some definitions.

Definition 3. *The polynomial $p(x) \in F[x]$ is irreducible if $p(x)$ is of positive degree and given any polynomial $f(x)$ in $F[x]$, then either $p(x) \mid f(x)$ or $p(x)$ is relatively prime to $f(x)$.*

In other words what the definition implies is that a polynomial of positive degree is said to be *irreducible* if it cannot be written as a product of two polynomials of positive degree.

Let us note that if $p(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$, $a_0 \neq 0$ is irreducible in $F[x]$, then $a_0^{-1}p(x)$ is also irreducible in $F[x]$. However, $a_0^{-1}p(x)$ has the advantage of being a monic polynomial. In other words, given any polynomial $p(x) \in F[x]$, we can trivially obtain a monic polynomial from $p(x)$.

Theorem 6. *Let $f(x) \in F[x]$ be of positive degree. Then either $f(x)$ is irreducible in $F[x]$ or $f(x)$ is the product of irreducible polynomials in $F[x]$. In fact, then,*

$$f(x) = ap_1(x)^{m_1}p_2(x)^{m_2} \cdots p_k(x)^{m_k}$$

where a is the leading coefficient of $f(x)$, $p_1(x), \dots, p_k(x)$ are monic and irreducible in $F[x]$, $m_1 > 0, \dots, m_k > 0$ and this factorization in this form is unique up to the order of $p_i(x)$.

Proof. We will prove it using induction. Suppose $\deg f(x) = 1$, then $f(x) = ax + b$ with $a \neq 0$. This is clearly irreducible on $F[x]$. Hence the case for $\deg f(x) = 1$ is true. Now let us suppose that the theorem holds true for all $g(x) \in F[x]$ such that $\deg g(x) < \deg f(x)$. Suppose $f(x)$ is irreducible, then there is nothing to prove. Suppose this is not so, then $f(x) = g(x)h(x)$ where $g(x), h(x) \in F[x]$ and $\deg g(x), \deg h(x) < \deg f(x)$. By induction since the theorem holds true for $g(x)$ and $h(x)$, hence $g(x), h(x)$ is irreducible or is the

product of irreducibles. This tells us that $f(x)$ is the product of irreducibles. Hence the first half of the theorem is proven.

As for the uniqueness, suppose the uniqueness for the product of polynomials holds for those with degree less than $\deg f(x)$. [Note that the uniqueness for degree 1 polynomials is clear]. Suppose

$$f(x) = ap_1(x)^{m_1}p_2(x)^{m_2} \cdots p_k(x)^{m_k} = aq_1(x)^{n_1}q_2(x)^{n_2} \cdots p_k(x)^{n_k}$$

where a is the leading coefficient of $f(x)$, where $p_i(x)$ and $q_j(x)$ are monic irreducibles and the m_i, n_j are all positive. Now, since $p_1(x)$ divides the right hand side, $p_1(x)$ divides some $q_j(x)$. Without loss of generality, we suppose $p_1(x) \mid q_1(x)$. Since $q_1(x)$ is irreducible, $p_1(x) = q_1(x)$. By induction, we have a unique factorization for $f(x)/p_1(x)$, and hence, our factorization for $f(x)$ is also unique. \square

Let $F = \mathbb{Q}$. We say that α is an algebraic number if α satisfies the equation $p(\alpha) = 0$ with $p(x) \in \mathbb{Q}[x]$. There exist an unique irreducible polynomial $f(x) \in \mathbb{Q}[x]$ with leading coefficient 1 which has α as a root. This polynomial is called the *minimal polynomial of α* , and its degree is called the *degree of α* , denoted $\deg \alpha$. If α is an algebraic number of degree n , then the roots $\alpha = \alpha_1, \alpha_2, \dots, \alpha_n$ of the minimal polynomial of α of the minimal polynomials of α are called the *conjugates* of α . Furthermore, a complex number is said to be an algebraic number if it is algebraic over \mathbb{Q} .

Example 2. The number $\sqrt{\sqrt{1 + \sqrt{7}}}$ is algebraic over \mathbb{Q} .

Proof.

Let $\alpha = \sqrt{\sqrt{1 + \sqrt{7}}}$. Then

$$\begin{aligned} \alpha^4 &= 1 + \sqrt{7}, \\ \alpha^4 - 1 &= \sqrt{7}, \\ (\alpha^4 - 1)^2 &= 7, \\ \alpha^8 - 2\alpha^4 - 6 &= 0. \end{aligned}$$

We define $f(x) = x^8 - 2x^4 - 6$. It is easy to verify that the other roots of $f(x)$ are:

$$-\sqrt{\sqrt{1 + \sqrt{7}}}, \quad \pm i\sqrt{\sqrt{1 + \sqrt{7}}}$$

$$\pm \frac{1}{\sqrt{2}}(1+i)\sqrt{\sqrt{\sqrt{7}-1}} \quad \text{and} \quad \pm \frac{1}{\sqrt{2}}(-1+i)\sqrt{\sqrt{\sqrt{7}-1}}.$$

Hence α is a root of $f(x) = x^8 - 2x^4 - 6$. □

Example 3. We can see that $\sqrt{2}$ is a root of $f(x) = x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$. Hence the conjugate of $\sqrt{2}$ are $\sqrt{2}$ and $-\sqrt{2}$. Furthermore, $\deg(\sqrt{2}) = \deg(x^2 - 2) = 2$. □

Lemma 0. Let $f(x) \in \mathbb{Q}[x]$, then

$$f(x) = \frac{u}{m}(a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0),$$

where u, m, a_n, \dots, a_0 are integers such that

$$\gcd(u, m) = 1 \quad \text{and} \quad \gcd(a_1, \dots, a_n) = 1.$$

Proof. Let

$$f(x) = \frac{A_n}{B_n} x^n + \cdots + \frac{A_0}{B_0}, \quad A_i, B_i \in \mathbb{Z}, \quad (i = 1, \dots, n).$$

Suppose

$$U = \gcd\left(A_1 \frac{B}{B_1}, A_2 \frac{B}{B_2}, \dots, A_i \frac{B}{B_i}, \dots, A_0 \frac{B}{B_0}\right)$$

where $B = B_0 \cdot B_1 \cdots B_n$, then

$$f(x) = \frac{U}{B}(a_n x^n + \cdots + a_0)$$

where $a_i \in \mathbb{Z}$. Clearing common factors in U and B yields the desired result. □

Definition 4. A commutative ring R is an integral domain if $a \cdot b = 0$ in R implies that $a = 0$ or $b = 0$.

A more detailed discussion of rings will be held in Section 4. For now we shall look at polynomial rings.

Gauss Lemma. Let $f(x) \in \mathbb{Z}[x]$ be a monic polynomial and $f(x) = a(x)b(x)$, where $a(x)$ and $b(x)$ are in $\mathbb{Q}[x]$. Then $f(x) = a_1(x)b_1(x)$, where $a_1(x), b_1(x)$ are monic polynomials in $\mathbb{Z}[x]$ and $\deg a_1(x) = \deg a(x)$ and $\deg b_1(x) = \deg b(x)$.

Proof. Suppose $f(x) = a_1(x)b_1(x)$ is in $\mathbb{Q}[x]$. Then from Lemma 0, we may write

$$a(x) = \frac{u_1}{v_1}a_1(x) \quad \text{and} \quad b(x) = \frac{u_2}{v_2}b_1(x)$$

with $a_1(x), b_1(x) \in \mathbb{Z}[x]$ and $u_1, u_2, v_1, v_2 \in \mathbb{Z}$. We write

$$f(x) = \frac{u}{v}a_1(x)b_1(x),$$

where u/v is the product of $(u_1u_2)/(v_1v_2)$ in its lowest term. Hence,

$$vf(x) = ua_1(x)b_1(x).$$

Suppose $v = 1$, then since $f(x)$ is monic, then $1 = ua'_nb'_m$, where a'_n and b'_m are the coefficients of $a_1(x)$ and $b_1(x)$ respectively. Therefore, $u = 1, a'_n = b'_m = 1$ (We may choose the coefficients to be positive integers). Therefore, $f(x)$ is a product of monic polynomials in $\mathbb{Z}[x]$.

Suppose $v \neq 1$. Then since v and u are relatively prime integers, there exist a prime p which divides v but not u . If

$$a_1(x) = a'_nx^n + a'_{n-1}x^{n-1} + \cdots + a'_0$$

then by our construction, we have $\gcd(a'_0, \dots, a'_n) = 1$ and hence there exist an i such that $p \nmid a'_i$. Similar there exist a j such that $p \nmid b'_j$, where

$$b_1(x) = b'_nx^n + b'_{n-1}x^{n-1} + \cdots + b'_0.$$

Now since $p \mid v$, $vf(x)$ is the zero polynomial in $(\mathbb{Z}/p\mathbb{Z})[x]$. However, since $p \nmid a'_i$ and $p \nmid b'_j$, the polynomial $a_1[x]$ and $b_1[x]$ are both nonzero polynomials in $(\mathbb{Z}/p\mathbb{Z})[x]$. Since $(\mathbb{Z}/p\mathbb{Z})[x]$ is an integral domain, this is impossible. Therefore $v = 1$. \square

This result tells us that *to determine if the monic polynomial $f(x) \in \mathbb{Z}[x]$ is irreducible in $\mathbb{Q}[x]$, it suffices to determine if $f(x)$ is irreducible in $\mathbb{Z}[x]$.*

Let us define

$$\overline{|\alpha|} = \max_{1 \leq k \leq n} |\alpha_k|$$

and call $\overline{|\alpha|}$ the *size* of the algebraic number α . We shall now prove Liouville's Theorem on the approximation of algebraic number by rational numbers.

Liouville's Theorem. *If α is a real algebraic number of degree n , $n \geq 1$, then there exist a constant $c = c(\alpha) > 0$ such that the following inequality holds for any $p \in \mathbb{Z}$ and $q \in \mathbb{N}$, $p/q \neq \alpha$*

$$\left| \alpha - \frac{p}{q} \right| > \frac{c}{q^n} \quad (1.34)$$

Proof. Suppose α is the root of the irreducible polynomial $f(x) = a_n x^n + \cdots + a_1 x + a_0$, $f(x) \in \mathbb{Z}[x]$, $a_n > 0$ and $\alpha = \alpha_1, \alpha_2, \dots, \alpha_n$ are conjugates of α . Then

$$|f(x)| = a_n |x - \alpha_1| |x - \alpha_2| \cdots |x - \alpha_n| = a_n |x - \alpha| \prod_{k=2}^n |x - \alpha_k|. \quad (1.35)$$

Let $p \in \mathbb{Z}$, $q \in \mathbb{N}$, $p/q \neq \alpha$. We consider two cases.

Case(1) : $\deg \alpha = n > 1$. Suppose we choose $(p, q) \in \mathbb{Z} \times \mathbb{N}$ such that $|\alpha - \frac{p}{q}| \geq 1$, then obviously,

$$\left| \alpha - \frac{p}{q} \right| \geq 1 \geq \frac{1}{q^n}. \quad (1.36)$$

Hence we just set $c = c(\alpha) = 1$ to obtain (1.34). Suppose we choose $(p, q) \in \mathbb{Z} \times \mathbb{N}$ such that $|\alpha - \frac{p}{q}| < 1$ then by Triangular Inequality,

$$\begin{aligned} \left| \left| \alpha \right| - \left| \frac{p}{q} \right| \right| &< \left| \alpha - \frac{p}{q} \right| < 1, \\ \implies -1 &< \left| \alpha \right| - \left| \frac{p}{q} \right| < 1, \\ \implies \left| \frac{p}{q} \right| &< \left| \alpha \right| + 1 \leq \overline{|\alpha|} + 1. \end{aligned} \quad (1.37)$$

We note that $f(\frac{p}{q}) \neq 0$ because $f(x)$ is irreducible on \mathbb{Q} . We substitute $x = p/q$ in (1.35) to obtain

$$\left| f\left(\frac{p}{q}\right) \right| = \frac{|a_n p^n + \cdots + a_1 p q^{n-1} + a_0 q^n|}{q^n} \geq \frac{1}{q^n}$$

Furthermore, substituting (1.37), we get

$$\prod_{k=2}^n \left| \frac{p}{q} - \alpha_k \right| \leq \prod_{k=2}^n \left(\left| \frac{p}{q} \right| + |\alpha_k| \right) < (2\overline{|\alpha|} + 1)^{n-1}.$$

This will allow us to obtain :

$$a_n \left| \frac{p}{q} - \alpha \right| (2\overline{|\alpha|} + 1)^{n-1} > \frac{1}{q^n}. \quad (1.38)$$

We define

$$c = c(\alpha) = \frac{1}{a_n (2\overline{|\alpha|} + 1)^{n-1}}.$$

to obtain the inequality in (1.34).

Case (2) : Suppose $\deg(\alpha) = n = 1$. Then there exist an $f(x) = a_0 + a_1x \in \mathbb{Z}[x]$ such that $f(\alpha) = a_0 + a_1\alpha = 0$. Hence $\alpha = -a_0/a_1$. Thus,

$$\left| \alpha - \frac{p}{q} \right| = \left| -\frac{a_0}{a_1} - \frac{p}{q} \right| = \left| \frac{a_0}{a_1} + \frac{p}{q} \right| \geq \frac{1}{a_1q} = \frac{c}{q}.$$

The assumption $\alpha = p/q$ holds for $n = 1$ only because for $n > 1$ $f(x)$ is irreducible in $\mathbb{Q}[x]$. Hence α must be irrational. \square

Corollary. *Under the conditions of Liouville's Theorem, there exists a constant $c = c(\alpha) > 0$ such that the inequality*

$$0 < \left| \alpha - \frac{p}{q} \right| < \frac{c}{q^n}$$

has no solutions $(p, q) \in \mathbb{Z} \times \mathbb{N}$.

This corollary is a direct consequence of Liouville's Theorem and it suggest the existence of transcendental numbers and the means of determining if a number in question is transcendental. This leads us to the following theorem.

Theorem 7. *Suppose that α is a real number such that for any $\nu \in \mathbb{R}, \nu > 0$, the inequality*

$$0 < \left| \alpha - \frac{p}{q} \right| < \frac{1}{q^\nu} \tag{1.39}$$

has an infinite set of solutions $(p, q) \in \mathbb{Z} \times \mathbb{N}$. Then α is transcendental.

Proof. We shall adopt proof by contradiction to show that α is transcendental. Suppose that α is an algebraic number that satisfies (1.39). Recall by Liouville's Theorem, there exist a constant $c(\alpha) > 0$ such that for any $(p, q) \in \mathbb{Z} \times \mathbb{N}, p/q \neq \alpha$, we have

$$\left| \alpha - \frac{p}{q} \right| > \frac{c(\alpha)}{q^n}. \tag{1.40}$$

From (1.39), we choose $\nu = n + 1$ and select a solution $(p, q) \in \mathbb{Z} \times \mathbb{N}$ such that $1/q < c(\alpha)$.

Then

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{q^\nu} = \frac{1}{q^{n+1}} = \frac{1}{q} \left(\frac{1}{q^n} \right) < \frac{c(\alpha)}{q^n}.$$

This contradicts (1.40) and Liouville's Theorem. Hence we are forced to conclude that α is transcendental. \square

Theorem 8. *There exist transcendental numbers*

We can produce a real number τ such that it violates Liouville's Theorem. For example consider $\tau = 0.10100100000010\dots 010\dots$, where there are $m!$ 0's between the m^{th} 1 and the $(m+1)^{\text{th}}$ 1. Hence τ must be transcendental, since it violates Liouville's Theorem for all $n > 0$. Also from Example 1, we see that

$$\alpha = \sum_{n=1}^{\infty} \frac{1}{a^{n!}}, \quad a \in \mathbb{N}, \quad q \geq 2$$

has infinitely many solutions $(p, q) \in \mathbb{Z} \times \mathbb{N}$ for every non negative $\nu \in \mathbb{R}$. Hence α is transcendental. \square

4. Approximation of algebraic numbers by algebraic numbers

In this section, we shall introduce the concept of rings and fields to aid us in proving certain theorems and lemmas, We will start by defining some definitions.

Definition 5. *The height $H = H_{\Psi}$ of a polynomial $\Psi = \Psi(x) \in C[x]$ is the maximum modulus of the coefficients of the polynomial.*

For example let us consider the polynomial $\Psi(x) = x^2 - 10x + 1$. The *height* of the polynomial is 10.

A polynomial in $\mathbb{Z}[x]$ is called *primitive* if there is no integer greater than 1 which divides all of its coefficient. Suppose $f(x)$ is the minimal polynomial of an algebraic number α . By multiplying the least common denominator of its coefficient, we obtain a primitive polynomial $\varphi(x) \in \mathbb{Z}[x]$ which has α as a root. The height $H = H_{\alpha}$ of the algebraic number α is defined to be the height of the irreducible primitive polynomial $\varphi(x) \in \mathbb{Z}[x]$ which has α as a root. Lastly, an algebraic number α is said to be an *algebraic integer* if all of the coefficients of its minimal polynomial $f(x)$ are rational integers. We shall denote A to be the set of algebraic numbers.

Let us recall the definition of a ring. A non empty set R is said to be a *ring* if in R there are two operations $+$ and \cdot such that :

- (a) $a, b \in R$ implies that $a + b \in R$.
- (b) $a + b = b + a$ for $a, b \in R$.
- (c) $(a + b) + c = a + (b + c)$ for $a, b, c \in R$.
- (d) There exists an element $0 \in R$ such that $a + 0 = a$ for every $a \in R$.
- (e) Given $a \in R$, there exists a $b \in R$ such that $a + b = 0$.
- (f) $a, b \in R$ implies $a \cdot b \in R$.
- (g) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for $a, b, c \in R$.
- (h) $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(b + c) \cdot a = b \cdot a + c \cdot a$, for $a, b, c \in R$.

Suppose there exist an element $1 \in R$ such that $a \cdot 1 = 1 \cdot a = a$ for every $a \in R$. We then say that R is a *ring with unit*. Furthermore, a ring for which the commutative law of multiplication holds is known as a *commutative ring*. The set of algebraic integers form a ring which we shall denote \mathbb{Z}_A . We shall prove a lemma, which is of independent interest.

Lemma 1. *Suppose $\alpha \in A$, then there exist an $r \in \mathbb{N}$ such that $r\alpha \in \mathbb{Z}_A$. If α is a root of $\varphi(x) \in \mathbb{Z}[x]$, $\varphi(x) \not\equiv 0$, then we can choose r to be the modulus of the leading coefficient of $\varphi(x)$.*

Proof. Let us define

$$\varphi(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 \in \mathbb{Z}[x], \quad a_n \neq 0$$

such that $\varphi(\alpha) = 0$. Multiplying a_n^{n-1} to $\varphi(\alpha) = 0$, we get,

$$\begin{aligned} a_n^n \alpha^n + a_n^{n-1} a_{n-1} \alpha^{n-1} + a_n^{n-1} a_{n-2} \alpha^{n-2} + \cdots + a_n^{n-1} a_0 &= 0, \\ (a_n \alpha)^n + a_{n-1} (a_n \alpha)^{n-1} + a_{n-2} \alpha_n (a_n \alpha)^{n-2} + \cdots + a_0 a_n^{n-1} &= 0. \end{aligned}$$

We define $r = a_n$. Hence $r\alpha$ is a root of some $\Phi(x) \in \mathbb{Z}[x]$. □

Definition 6. *A commutative ring R with unit is called a field if every non-zero $a \in R$ has a multiplicative inverse, i.e there exist an element $a^{-1} \in R$ such that $a^{-1}a = 1$.*

A field is necessarily an integral domain for if $ab = 0$ and $a \neq 0$, then

$$b = 1 \cdot b = a^{-1}ab = a^{-1} \cdot 0 = 0.$$

We remark that a field may also be viewed as a commutative ring in which the non zero elements form a group under multiplication.

Let us now consider a relationship between two fields K and F , where $K \supset F$. We call K an *extension* or extension field of F , and we call F a *subfield* of K . The operations in F are those in K restricted to the elements of F . We say that K is a finite extension of F if, viewed as a vector space over F , $\dim_F(K)$ is finite. We shall write $\dim_F(K)$ as $[K : F]$ and call it the *degree of K over F* .

An algebraic number field is defined as the extension field of \mathbb{Q} obtained by adjoining an algebraic number θ , i.e, the set of numbers $\mathbb{Q}(\theta)$, where $\mathbb{Q}(x)$ runs through all the rational function in $\mathbb{Q}(x)$ whose denominator does not vanish at θ . We denote the algebraic number field as $\mathbb{Q}(\theta)$. We call θ the *generating element* or *primitive element* of the algebraic number field $\mathbb{Q}(\theta)$. The generating element θ is not unique, as different elements of the field can be choose to serve as a generating element. Each of such generating element have the same degree h . Since $K = \mathbb{Q}(\theta) \supset \mathbb{Q}$, we can view $K = \mathbb{Q}(\theta)$ as a vector space over \mathbb{Q} , i.e $\dim_{\mathbb{Q}}(K)$ is finite. We shall write $\dim_{\mathbb{Q}}(K)$ as $[K : \mathbb{Q}] = [\mathbb{Q}(\theta) : \mathbb{Q}]$ and denote it as the degree of an algebraic number field $\mathbb{Q}(\theta)$.

Let $p(x) \in \mathbb{Q}[x]$ be the minimum polynomial of θ . Given $f(x) \in \mathbb{Q}[x]$, then $f(x) = q(x)p(x) + r(x)$ where $q(x)$ and $r(x)$ are in $\mathbb{Q}[x]$ and $r(x) = 0$ or $\deg r(x) < \deg p(x)$ follows from the division algorithm. Hence , $f(\theta) = q(\theta)p(\theta) + r(\theta) = r(\theta)$, since $p(\theta) = 0$. Thus given any element $\alpha \in K$, we can express it as a polynomial expression in θ of degree at most $h - 1$ i.e

$$\alpha = r(\theta) = c_0 + c_1\theta + \cdots + c_{h-1}\theta^{h-1}, \quad c_i \in \mathbb{Q}, \quad (i = 0, 1, \dots, h - 1).$$

Suppose such an expression for α is not unique, then there exist $d_i \in \mathbb{Q}$, $i = 0, 1, \dots, h - 1$, $d_i \neq c_i$ such that $\alpha = d_0 + \cdots + d_{h-1}\theta^{h-1}$. We obtain the expression

$$(c_0 - d_0) + (c_1 - d_1)\theta + \cdots + (c_{h-1} - d_{h-1})\theta^{h-1} = 0.$$

But this contradicts the definition of the minimum polynomial. Hence all polynomial expressions of α must be unique.

We define the numbers

$$\alpha_i = r(\theta_i), \quad (i = 1, \dots, h) \quad (1.41)$$

as the conjugates of α in the field of K . They are the conjugates of α , but each conjugate is repeated $\frac{h}{\deg(\alpha)}$ times. Hence $\alpha_2, \dots, \alpha_n$ do not belong to K . For example, let us consider $\alpha = \sqrt[3]{2}$ whose minimum polynomial over \mathbb{Q} is $f(x) = x^3 - 2$. Then the roots of $f(x)$ is

$$\alpha, \omega\alpha, \omega^2\alpha, \quad \text{where } \omega = e^{\frac{2\pi i}{3}}.$$

Suppose $K = \mathbb{Q}(\sqrt[3]{2})$, then α is in K but $\omega\alpha, \omega^2\alpha \notin K$

If α is an element of K , we define its size $|\overline{\alpha}|$ by

$$|\overline{\alpha}| = \max_{1 \leq i \leq h} |r(\theta_i)|, \quad \alpha = r(\theta).$$

When considering an algebraic number field $K = \mathbb{Q}(\theta)$, we shall suppose that the numbering $\theta = \theta_1, \dots, \theta_h$ is fixed. Hence from (1.41) we also fix the numbering of the conjugates of any $\alpha \in K$. The set of algebraic integers contained in the field K form a ring which we shall denote as \mathbb{Z}_K . We denote the norm by $\mathcal{N}(\alpha)$:

$$\mathcal{N}(\alpha) = r(\theta_1) \cdots r(\theta_h), \quad \alpha = r(\theta).$$

We list some properties of the norm:

- (1) $\mathcal{N}(\alpha) \in \mathbb{Q}$ for all $\alpha \in K$;
- (2) $\mathcal{N}(\alpha) \in \mathbb{Z}$ for all $\alpha \in \mathbb{Z}_K$;
- (3) $\mathcal{N}(\alpha) = 0$ if and only if $\alpha = 0$;
- (4) $\mathcal{N}(\alpha) = \alpha^h$ if $\alpha \in \mathbb{Q}$;
- (5) $\mathcal{N}(\alpha\beta) = \mathcal{N}(\alpha)\mathcal{N}(\beta)$;
- (6) $\mathcal{N}(a\alpha) = a^h\mathcal{N}(\alpha)$ for all $a \in \mathbb{Q}$ and $\alpha \in K$.

For the rest of the discussion in this section, we shall use $\mathcal{N}(\alpha)$ to denote the norm of α in the algebraic field $\mathbb{Q}(\alpha)$ generated by α . If α is a root of a nonzero polynomial $f(x)$

over K , i.e., if $f(x) \in K[x]$, then $\alpha \in A$. The *degree of α over K* , denoted $\deg_K \alpha$, is the degree of the monic irreducible polynomial in $K[x]$ which has α as a root.

We shall now define the *symmetric polynomial* which will be used in the proof of a lemma.

Definition. Let V be a commutative ring with unit. A polynomial $F(\alpha_1, \dots, \alpha_n) \in V[\alpha_1, \dots, \alpha_n]$ is called a symmetric polynomial in $\alpha_1, \dots, \alpha_n$ if it does not change when $\alpha_1, \dots, \alpha_n$ are subjected to any permutation.

Let us further denote

$$\begin{aligned}\sigma_1 &= \alpha_1 + \cdots + \alpha_n, \\ \sigma_2 &= \alpha_1\alpha_2 + \alpha_1\alpha_3 + \cdots + \alpha_{n-1}\alpha_n, \\ &\vdots \\ \sigma_n &= \alpha_1 \cdots \alpha_n.\end{aligned}$$

as the *elementary symmetry polynomials* in $\alpha_1, \dots, \alpha_n$. Notice that they are up to sign equals to the coefficients of the polynomial $(x - \alpha_1) \cdots (x - \alpha_n)$. We shall now state the *symmetric polynomial theorem* which we will require in the proofing of *Lemma 2*.

Definition 7. The *symmetric polynomial theorem* states that any symmetric polynomial $F(\alpha_1, \dots, \alpha_n) \in V[\alpha_1, \dots, \alpha_n]$ can be expressed uniquely in the form $F(\alpha_1, \dots, \alpha_n) = H(\sigma_1, \dots, \sigma_n)$, where $H(\sigma_1, \dots, \sigma_n) \in V[\sigma_1, \dots, \sigma_n]$.

Lemma 2. Suppose that $\alpha \in A$, $\deg \alpha = n$, and $\alpha = \alpha_1, \dots, \alpha_n$ are the conjugates of α . Further suppose that

$$F = F(x_1, \dots, x_k; \alpha_1, \dots, \alpha_n) \in \mathbb{Q}[x_1, \dots, x_k; \alpha_1, \dots, \alpha_n], \quad k \geq 0,$$

and that, as a polynomial in $\alpha_1, \dots, \alpha_n$ with coefficients in $[x_1, \dots, x_k]$, F is a symmetric polynomial in $\alpha_1, \dots, \alpha_n$. Then

$$F = F(x_1, \dots, x_k) \in \mathbb{Q}[x_1, \dots, x_k];$$

and $F \in \mathbb{Q}$ in the case $k = 0$. In addition, if we also have $\alpha \in \mathbb{Z}_A$ and

$$F = F(x_1, \dots, x_k; \alpha_1, \dots, \alpha_n) \in \mathbb{Z}[x_1, \dots, x_k; \alpha_1, \dots, \alpha_n], \quad k \geq 0,$$

then it follows that

$$F = F(x_1, \dots, x_k) \in \mathbb{Z}[x_1, \dots, x_k];$$

and $F \in \mathbb{Z}$ in the case $k = 0$.

Proof. Let us consider F as a polynomial in $\alpha_1, \dots, \alpha_n$ with coefficients in $\mathbb{Q}[x_1, \dots, x_k]$. Since F is a symmetric polynomial in $\alpha_1, \dots, \alpha_n$, then by the symmetric polynomial theorem, we can write F as a polynomial H in the elementary symmetric polynomial $\sigma_1, \dots, \sigma_n$ with coefficients in $\mathbb{Q}[x_1, \dots, x_k]$.

Let us recall that the symmetric polynomial $\sigma_1, \dots, \sigma_n$ are equal (up to sign) to the coefficients of the minimum polynomial $f(x)$ of the algebraic number α and that $f(x) \in \mathbb{Q}(x)$. Hence each $\sigma_1, \dots, \sigma_n \in \mathbb{Q}$. This leads us to the fact that $H = H(x_1, \dots, x_k) \in \mathbb{Q}[x_1, \dots, x_k]$. Hence by the symmetric polynomial theorem, we come to the conclusion that

$$F = F(x_1, \dots, x_k) \in \mathbb{Q}[x_1, \dots, x_k].$$

In the case of $k = 0$, it just means that we substitute each x_1, \dots, x_k by a complex number and its conjugate and the sum of which will give us the result $F \in \mathbb{Q}$.

For the second part of the lemma, the proof is similar to that above except that we are considering algebraic integers which has its minimum polynomial $f(x)$ in $\mathbb{Z}[x]$. Hence the elementary symmetric polynomials $\sigma_1, \dots, \sigma_n$ are equal (up to sign) to the coefficients of the minimal polynomial $f(x)$ of the algebraic integer α . Thus they are elements of \mathbb{Z} . The rest of the proof follows from above. \square

We will now generalize the problem of approximating real numbers by rational numbers by studying the approximation of complex numbers by certain classes of algebraic numbers. That is, given a complex number α , we consider the modulus of the difference

$$|\alpha - \theta| \tag{1.42}$$

for different θ , where θ is an algebraic number. We are interested in the approximation of complex numbers by certain classes of algebraic numbers whose values are bounded by the following inequality

$$|\alpha - \theta| < \varphi(H, k),$$

where $H = H_\theta$ is the height of θ and $k = \deg \theta$. In particular, we want to investigate for which $\varphi(H, k)$ does the inequality has finite or infinite solutions in the algebraic number θ .

We shall consider approaching this problem by considering the behavior of $|P(\alpha)|$ for a complex number α and different polynomials $P(x) \in \mathbb{Z}[x]$ of degree k and height H . Suppose $P(\theta)$ is small, then given any α such that $\alpha \approx \theta$, $|P(\alpha)|$ will be arbitrary small. Conversely suppose if $|P(\alpha)|$ is small, then at least one of the differences $|\alpha - \theta_s|$, $1 \leq s \leq k$, is small where $\theta_1, \dots, \theta_k$ are the roots of $P(x)$. Hence the magnitude of $|P(\alpha)|$ characterizes the order of approximation of α by the algebraic number of height H and degree k .

For a fixed $\alpha \in \mathbb{C}$, we will like to find a lower bound

$$|P(\alpha)| > \varphi(H, k),$$

where $\varphi(H, k)$ is a function of the height H and the degree k of the polynomial $P(x)$, which holds for all polynomials $P(x) \in \mathbb{Z}[x]$, not necessarily irreducible, for which $P(\alpha) \neq 0$. We can generalize this problem to the case which involves several complex numbers $\alpha_1, \dots, \alpha_m$ and the polynomials $P(x_1, \dots, x_m) \in \mathbb{Z}[x_1, \dots, x_m]$.

We will now consider a generalization of Liouville's Theorem using the properties and criterion stated above. In particular we want to show that given any $\alpha \in A$, for all $P(x) \in \mathbb{Z}[x]$, $P(x)$ not necessary irreducible and $P(\alpha) \neq 0$ we have the following inequality,

$$|P(\alpha)| > \frac{c}{H^{n-1}}.$$

Let us consider for the case

$$P = P(x) = a_1x + a_0, \quad a_0 \in \mathbb{Z}, \quad a_1 \in \mathbb{N}, \quad H_P = H.$$

Let us recall from Liouville's Theorem that suppose $\alpha \in A$ then for any $(p, q) \in \mathbb{Z} \times \mathbb{N}$, there exist a $c = c(\alpha)$, $c(\alpha) > 0$ such that,

$$\left| \alpha - \frac{p}{q} \right| > \frac{c}{q^n}, \quad (p, q) \in \mathbb{Z} \times \mathbb{N}.$$

Given that,

$$P = P(x) = a_1x + a_0, \quad a_0 \in \mathbb{Z}, \quad a_1 \in \mathbb{N}, \quad H_P = H.$$

Then from Liouville's Theorem, we let $p = -a_0$, $q = a_1$. Then,

$$\begin{aligned} |\alpha - \frac{p}{q}| &> \frac{c}{q^n} \\ \implies |\alpha - \frac{-a_0}{a_1}| &> \frac{c}{a_1^n} \\ \implies |a_1\alpha + a_0| &> \frac{c}{a_1^{n-1}}. \end{aligned}$$

Therefore we get,

$$|P(\alpha)| > \frac{c}{a_1^{n-1}} \geq \frac{c}{H^{n-1}}.$$

Hence we obtain the desired inequality. \square

In general, this result is true for all polynomials $P(x) \in \mathbb{Z}[x]$ of any degree k . We shall show this result. But first we shall prove a lemma.

Lemma 3. *If*

$$f(x) = b_n x^n + \cdots + b_1 x + b_0, \quad f(x) \in \mathbb{C}[x], \quad b_n b_0 \neq 0, \quad H_{f(x)} = H$$

and $f(\alpha) = 0$, then the following inequality holds

$$\frac{|b_0|}{H + |b_0|} < |\alpha| < \frac{H + |b_n|}{|b_n|}, \quad |b_n \alpha| < 2H.$$

Proof. We shall consider 3 cases: $|\alpha| > 1$, $|\alpha| < 1$ and $|\alpha| = 1$. The case for $|\alpha| = 1$ is trivial. Let us note that for all b_i such that $i = 1, \dots, n$,

$$\begin{aligned} |b_i| &\leq H \\ \implies \frac{H}{|b_i|} &\geq 1 \\ \implies \frac{H}{|b_i|} + 1 &\geq 2 \\ \implies \frac{H + |b_i|}{|b_i|} &\geq 2 \quad \text{and} \quad \frac{|b_i|}{H + |b_i|} \leq \frac{1}{2}. \end{aligned}$$

Hence $|\alpha| = 1$ satisfies the inequality.

Suppose $|\alpha| < 1$. Then as in the case of $|\alpha| = 1$, the upper bound is easily proven as it is satisfied as shown above. We shall prove the lower bound. Given that α is a root of $f(x)$,

we get

$$\begin{aligned}
f(\alpha) &= b_n \alpha^n + \cdots + b_1 \alpha + b_0 = 0 \\
|b_0| &= |b_n \alpha^n + \cdots + b_1 \alpha| \\
&\leq H (|\alpha|^n + |\alpha|^{n-1} + \cdots + |\alpha|) \\
&= |\alpha| H (1 + \cdots + |\alpha|^{n-1}), \quad (\Delta \text{ inequality}) \\
&= H |\alpha| \frac{1 - |\alpha|^n}{1 - |\alpha|} < \frac{H |\alpha|}{1 - |\alpha|}.
\end{aligned}$$

Then we get,

$$\begin{aligned}
|b_0| - |b_0| |\alpha| &< H |\alpha| \\
|\alpha| (H + |b_0|) &> |b_0| \\
|\alpha| &> \frac{|b_0|}{H + |b_0|}.
\end{aligned}$$

Hence for $|\alpha| < 1$, it satisfies the inequality.

Let us consider the case for $|\alpha| > 1$. Using the fact that α is a root for $f(x)$, we get

$$\begin{aligned}
f(\alpha) = 0; \quad b_n \alpha^n + \cdots + b_1 \alpha + b_0 &= 0 \\
|b_n| |\alpha|^n &\leq H (|\alpha|^{n-1} + \cdots + 1) \\
&= H \frac{|\alpha|^n - 1}{|\alpha| - 1} < H \frac{|\alpha|^n}{|\alpha| - 1}.
\end{aligned}$$

Hence

$$\begin{aligned}
|b_n| &< \frac{H}{|\alpha| - 1} \\
|b_n| |\alpha| - |b_n| &< H \\
|\alpha| &< \frac{H + |b_n|}{|b_n|}.
\end{aligned}$$

The lower bound for the third case is quite trivial.

Let us recall that

$$|\alpha| > 1,$$

and that

$$|b_0| \leq H.$$

Hence we deduce the following:

$$\begin{aligned} \frac{|b_0|}{H + |b_0|} &< \frac{|b_0|}{|b_0| + |b_0|} \\ &= \frac{1}{2} \\ &< |\alpha|. \end{aligned}$$

Thus we get the lower bound for the third case.

Hence the first inequality is proven for all 3 cases.

The second inequality is quite obvious. From the upper bound we have,

$$\begin{aligned} |\alpha| &< \frac{H + |b_n|}{|b_n|} \\ |\alpha||b_n| &< H + |b_n| < H + H < 2H. \end{aligned}$$

Hence the second inequality is proven. \square

Corollary. *If α is a nonzero algebraic number of height h , then its size satisfies the bounds*

$$\frac{1}{2h} < \overline{|\alpha|} < h + 1 \leq 2h.$$

Proof. Let us recall the definition of an algebraic number. We say that α is an algebraic number if α satisfies the equation $p(\alpha) = 0$ with $p(x) \in \mathbb{Q}[x]$. Let $f(x)$ be a monic polynomial, $f(x) \in \mathbb{Q}[x]$. Then $b_n = 1$.

Hence $f(x) = x^n + b_{n-1}x^{n-1} + \dots + b_0$. Then from Lemma 3 we have

$$|\alpha| < \frac{h + |b_n|}{|b_n|} = \frac{h + 1}{1} = h + 1 < 2h.$$

Consider the polynomial $g(x) = \frac{1}{b_0}f(x)$. Then for all α_i , $i = 1, \dots, n$, such that $f(\alpha_i) = 0$ implies that $g(\alpha_i) = 0$ and from Lemma 3, we obtain,

$$\overline{|\alpha|} > \frac{|b'_0|}{h + |b'_0|} = \frac{1}{h + 1}, \quad \text{where } b'_0 = \frac{1}{b_0} \cdot b_0 = 1.$$

Recall that $h \geq 1$ from our definition of $f(x)$. Hence,

$$\overline{|\alpha|} > \frac{1}{h + 1} > \frac{1}{h + h} = \frac{1}{2h}.$$

Thus we obtain the inequality

$$\frac{1}{2h} < |\overline{\alpha}| < h + 1 \leq 2h.$$

□

We shall now state the theorem formally.

Theorem 9. *Suppose that $\alpha \in A$, $\deg \alpha \leq n$, $H_\alpha \leq h$ and $P = P(x) \in \mathbb{Z}[x]$, $\deg P \leq k$, $H_P \leq H$. Then either $P(\alpha) = 0$, or else the following inequality holds:*

$$|P(\alpha)| \geq \frac{c^k}{H^{n-1}}, \quad c = \frac{1}{3^{n-1}h^n}. \quad (1.43)$$

Proof. Without loss of generality, we may assume that $\deg \alpha = n$, $H_\alpha = h$, $\deg P = k$, $H_P = H$. Since $P(x) \in \mathbb{Z}[x]$, therefore $H \geq 1$. Given that $\alpha \in A$, this implies that there exist a $f(x) \in \mathbb{Z}[x]$ such that $f(\alpha) = 0$. The case for $P(\alpha) = 0$ is trivial. We shall suppose that $P(\alpha) \neq 0$. We define b_n as the leading coefficient in the primitive and irreducible polynomial $f(x) \in \mathbb{Z}[x]$ having α as a root. We shall consider possible 2 cases that arises.

Case (1) : $\deg \alpha = n = 1$

Suppose $\deg \alpha = n = 1$, then there exist $f(x) \in \mathbb{Z}[x]$, $f(x) = b_0 + b_1x$ such that

$$\begin{aligned} b_0 + b_1\alpha &= 0, \\ \alpha &= -\frac{b_0}{b_1}. \end{aligned}$$

Let us consider 2 possibilities. Suppose $P(x)$ is irreducible in $\mathbb{Z}[x]$. Then by Gauss Lemma $P(x)$ is irreducible in $\mathbb{Q}[x]$. Hence,

$$\begin{aligned} |P(\alpha)| &= \left| P\left(-\frac{b_0}{b_1}\right) \right| \\ &= \frac{|a_k b_0^k + \cdots + a_0 b_1^k|}{b_1^k} \\ &\geq \frac{1}{b_1^k} \\ &\geq \frac{1}{h^k} = \frac{c^k}{H^{1-1}}, \quad c = \frac{1}{3^{1-1}h^1}. \end{aligned}$$

Suppose $P(x)$ is reducible² in $\mathbb{Z}[x]$. Then from Theorem 6 in Section 3 we obtain

$$P(x) = P_{\sigma(1)}(x)P_{\sigma(2)}(x) \cdots P_{\sigma(j)}(x)$$

²This step is actually not necessary in the context of the proof.

where $P_{\sigma(i)}(x)$, $i = 1, \dots, j$ are irreducible in $\mathbb{Z}[x]$ and

$$\deg P_{\sigma(i)}(x) = m_i, \quad (i = 1, \dots, j); \quad \sum_{i=1}^j m_i = k.$$

Then applying Gauss Lemma to each irreducible polynomial, we obtain the same result as in the first case. Hence

$$|P(\alpha)| \geq \frac{1}{b_1^k} \geq \frac{1}{h^k}$$

which proves (1.43) for the case $n = 1$.

Case (2) $\deg \alpha > 1$.

Suppose $\deg \alpha = n > 1$. For each conjugate of α : $\alpha = \alpha_1, \alpha_2, \dots, \alpha_n$, let us define

$$\xi = b_n^k |P(\alpha)|.$$

From Lemma 1 we can deduce that $\xi \in \mathbb{Z}_A$. Hence,

$$\xi_i = b_n^k |P(\alpha_i)|, \quad \xi_i \in \mathbb{Z}_A \quad (i = 1, \dots, n). \quad (1.44)$$

Let $g(x)$ be the irreducible monic polynomial of ξ , $\deg g(x) = N$ and $g(x) \in \mathbb{Z}[x]$. Then,

$$\begin{aligned} g(x) &= x^N + C_{N-1}x^{N-1} + \dots + C_1x + C_0, \quad g(x) \in \mathbb{Z}[x], \quad C_0 \neq 0 \\ &= (x - \xi_1)(x - \xi_2) \dots (x - \xi_N), \end{aligned}$$

where $\xi = \xi_1, \xi_2, \dots, \xi_N$ are conjugates of ξ . Then

$$\begin{aligned} (-1)^N \xi_1 \cdot \xi_2 \dots \xi_N &= C_0 \\ \mathcal{N}(\xi) &= (-1)^N C_0 \\ |\mathcal{N}(\xi)| &= |C_0| \geq 1. \end{aligned} \quad (1.45)$$

Then from (1.44) and (1.45) we get,

$$|\mathcal{N}(\xi)| = b_n^k |P(\alpha)| \prod_{i=2}^n b_n^k |P(\alpha_i)| \geq 1. \quad (1.46)$$

Given any α_i , we have

$$\begin{aligned} |P(\alpha_i)| &\leq H(1 + |\alpha_i| + \dots + |\alpha_i|^k) \\ &\leq H(1 + |\alpha_i|)^k. \end{aligned} \quad (1.47)$$

Recall from Lemma 3

$$|\alpha_i| < \frac{h + |b_n|}{|b_n|}.$$

Hence

$$\begin{aligned} |P(\alpha_i)| &\leq H \left(\frac{h + 2|b_n|}{|b_n|} \right)^k < \frac{(3h)^k H}{|b_n|^k}, \\ |b_n|^k |P(\alpha_i)| &\leq (3h)^k H. \end{aligned} \tag{1.48}$$

From (1.46) we have

$$|P(\alpha)| \geq \frac{1}{b_n^k \prod_{i=2}^n |b_n|^k |P(\alpha_i)|}.$$

Applying (1.47) and (1.48) we get,

$$|P(\alpha)| \geq \frac{1}{b_n^k (3h)^{k(n-1)} H^{n-1}} \geq \frac{1}{(3^{n-1} h^n)^k H^{n-1}}.$$

Hence the case for $\deg \alpha = n > 1$ is proven. \square

Theorem 10. *If α is an algebraic number of degree n , $n \geq 1$, then there exists a constant $c = c(\alpha) > 0$ such that the following inequality holds for any algebraic number θ of degree k , $k \geq 1$, and height H for which $\theta \neq \alpha$.*

$$|\alpha - \theta| > \frac{c^k}{H^n}. \tag{1.49}$$

Proof. Suppose that θ is not equal to α and is a root of the irreducible polynomial

$$P(x) = a_k x^k + \cdots + a_1 x + a_0, \quad P(x) \in \mathbb{Z}[x], \quad a_k > 0.$$

Then three possible cases will arise:

Case (1) : $k = 1$

In this case we define

$$P(x) = a_1 x + a_0, \quad \text{where } P(\theta) = a_1 \theta + a_0 = 0, \quad a_1 > 0.$$

Then $\theta = -a_0/a_1$. From Liouville's Theorem we know that for all $(p, q) \in \mathbb{Z} \times \mathbb{N}$, there exist a $c = c(\alpha) > 0$ such that

$$\left| \alpha - \frac{p}{q} \right| > \frac{c(\alpha)}{q^n}, \quad \text{where } \deg \alpha = n.$$

Let us choose $p = -a_0$ and $q = a_1$. Then

$$|\alpha - \theta| = \left| \alpha - \frac{-a_0}{a_1} \right| > \frac{c(\alpha)}{a_1^n} > \frac{c(\alpha)}{H^n},$$

where $H = \max\{|a_1|, |a_0|\}$. We define

$$c = \sqrt[k]{c(\alpha)}.$$

Then we get the required inequality.

Case (2) $k > 1$ and $P(\alpha) = 0$:

Since $P(\alpha) = 0$, therefore $n = \deg \alpha = \deg \theta = k$, and θ is a conjugate of α . Let $\delta = \delta(\alpha)$ be any constant such that

$$0 < \delta < \min_{1 \leq i < j \leq n} |\alpha_i - \alpha_j|. \quad (1.50)$$

Then

$$\begin{aligned} |\alpha - \theta| > \delta &\geq \frac{(\sqrt[n]{\delta})^k}{H^n} \quad (\because k = n \text{ and } H \geq 1, H \in \mathbb{Z}) \\ &= \frac{c^k}{H^n}. \end{aligned} \quad (1.51)$$

Hence the inequality (1.51) shows that the theorem holds in the case $k > 1$ and $P(\alpha) = 0$.

Case 3 : $k > 1$ and $P(\alpha) \neq 0$

Let us consider 2 sub-cases. Suppose $|\alpha - \theta| \geq 1$, then

$$|\alpha - \theta| \geq 1 \geq \frac{1}{H^n} \quad (\because H \geq 1, H \in \mathbb{Z} \text{ and } c = 1). \quad (1.52)$$

Suppose that $|\alpha - \theta| < 1$, then by the triangular inequality,

$$\begin{aligned} ||\alpha| - |\theta|| &\leq |\alpha - \theta| < 1 \\ \implies |\theta| - |\alpha| &< 1 \\ \implies |\theta| &< |\alpha| + 1. \end{aligned} \quad (1.53)$$

Let us define

$$P_1(x) := \frac{P(x)}{x - \theta}.$$

Then given that $\alpha \neq \theta$ we can write

$$\begin{aligned} P_1(\alpha) &= \frac{P(\alpha)}{\alpha - \theta}. \\ \therefore |\alpha - \theta| &= \frac{|P(\alpha)|}{|P_1(\alpha)|}. \end{aligned} \tag{1.54}$$

Recall that:

$$P(x) = a_k x^k + \cdots + a_1 x + a_0, \quad P(x) \in \mathbb{Z}[x], \quad a_k > 0.$$

Hence

$$\begin{aligned} P_1(x) &= \frac{P(x)}{x - \theta} = \frac{P(x) - P(\theta)}{x - \theta} \quad \because P(\theta) = 0 \\ &= a_k \frac{x^k - \theta^k}{x - \theta} + a_{k-1} \frac{x^{k-1} - \theta^{k-1}}{x - \theta} + \cdots + a_1 \frac{x - \theta}{x - \theta}. \end{aligned}$$

Recall that:

$$x^n - \theta^n = (x - \theta)(x^{n-1} + \theta x^{n-2} + \cdots + \theta^{n-2} x + \theta^{n-1}).$$

Thus,

$$\begin{aligned} P_1(x) &= a_k(x^{k-1} + \theta x^{k-2} + \theta^2 x^{k-3} + \cdots + \theta^{k-2} + \theta^{k-1}) + \\ &\quad a_{k-1}(x^{k-2} + \theta x^{k-3} + \cdots + \theta^{k-2}) + \cdots + a_1 \\ &= \sum_{l=0}^{k-1} g_l(\theta) x^l, \end{aligned} \tag{1.55}$$

where

$$g_l(\theta) = a_{l+1} + a_{l+2}\theta + \cdots + a_k \theta^{k-l-1}, \quad (l = 0, 1, \dots, k-1). \tag{1.56}$$

Let us consider $g_l(\theta)$. From (1.53) and (1.56), we get

$$\begin{aligned} |g_l(\theta)| &< H(1 + |\theta| + \cdots + |\theta|^{k-l-1}) \\ &< H(1 + |\theta| + \cdots + |\theta|^k) \\ &\leq H(1 + |\theta|)^k \\ &< H(2 + |\alpha|)^k, \quad (l = 0, 1, \dots, k-1). \end{aligned} \tag{1.57}$$

Then from (1.55) and (1.57) we have

$$\begin{aligned} |P_1(\alpha)| &= |g_0(\theta)\alpha^0 + g_1(\theta)\alpha + \cdots + g_{k-1}(\theta)\alpha^{k-1}| \\ &\leq |g_0(\theta)\alpha^0| + \cdots + |g_{k-1}(\theta)\alpha^{k-1}| \end{aligned}$$

$$\begin{aligned}
&< H(2 + |\alpha|)^k(1 + \cdots + |\alpha|^{k-1}) \\
&< H(2 + |\alpha|)^k(1 + \cdots + |\alpha|^{k-1} + |\alpha|^k) \\
&< H(2 + |\alpha|)^k(1 + |\alpha|)^k.
\end{aligned}$$

From the Corollary to Lemma 3, we have

$$\frac{1}{2h} < \overline{|\alpha|} < h + 1 \leq 2h, \quad h \geq 1, \quad h \in \mathbb{Z},$$

where h is the height of α . Then,

$$|P_1(\alpha)| < H((h + 3)(h + 2))^k. \quad (1.58)$$

From Theorem 9, if $P(\alpha) \neq 0$, then

$$|P(\alpha)| \geq \frac{c^k}{H^{n-1}}, \quad c = \frac{1}{3^{n-1}h^n}.$$

Therefore from (1.54) we have

$$\begin{aligned}
|\alpha - \theta| &= \frac{|P(\alpha)|}{|P_1(\alpha)|} \\
&> \frac{1}{3^{n-1}h^n H^{n-1}} \cdot \frac{1}{H((2 + h)(h + 3))^k} \\
&= \frac{c^k}{H^n}, \quad \text{where} \quad c = \sqrt[k]{\frac{1}{3^{n-1}h^n((2 + h)(h + 3))^k}}, \quad c < 1. \quad (1.59)
\end{aligned}$$

Hence we obtain the desired inequality. \square

Theorem 11. *Suppose that α is a complex number such that for any $\nu \in \mathbb{R}$, $\nu > 0$ the inequality*

$$|\alpha - \theta| < \frac{1}{H_\theta^\nu}$$

has an infinite set of solutions in algebraic number θ of degree $\deg \theta \leq k$ and height H_θ . Then α is transcendental.

Proof. Suppose that α were an algebraic number of degree n . From Theorem 10, there would exist a constant $c = c(\alpha) > 0$ such that the following inequality holds for any algebraic number θ of degree k , $k \geq 1$, and height H for which $\theta \neq \alpha$.

$$|\alpha - \theta| > \frac{c(\alpha)^k}{H^n}. \quad (1.60)$$

Let us set $\nu = n + 1$ in (1.60) and choose an algebraic number θ such that $1/H\theta < c(\alpha)^k$.

Then,

$$\begin{aligned} 0 < |\alpha - \theta| &< \frac{1}{H^{n+1}\theta} \\ &= \frac{1}{H\theta} \cdot \left(\frac{1}{H^n}\right) \\ &< \frac{c(\alpha)^k}{H^n}. \end{aligned}$$

This contradicts Theorem 10 for algebraic numbers. Thus α must be transcendental. \square

Let us now consider a generalization of Theorem 9 to the case of a polynomial in several algebraic numbers. We make use of an algebraic number field $K = \mathbb{Q}(\theta)$ containing the algebraic numbers $\alpha_1, \dots, \alpha_m$.

Theorem 12. *Suppose that $\alpha_1, \dots, \alpha_m$ are algebraic numbers, and h is the degree of an algebraic number field $K = \mathbb{Q}(\theta)$ which contains these numbers. Then there exist a constant $c = c(\alpha_1, \dots, \alpha_m) > 0$ such that for any polynomial*

$$P = P(z_1, \dots, z_m) \in \mathbb{Z}[z_1, \dots, z_m], \quad \deg_{\mathbb{Z}} P \leq k, \quad H_P \leq H,$$

either $P(\alpha_1, \dots, \alpha_m) = 0$ or else

$$|P(\alpha_1, \dots, \alpha_m)| \geq \frac{c^k}{H^{h-1}}. \quad (1.61)$$

Without loss of generality, let us assume that $\deg_{\mathbb{Z}} P = k$, $H_P = H$ and $\deg_{\mathbb{Q}} \mathbb{Q}(\theta) = n$. Let $\alpha_i = \alpha_i^{(1)}, \dots, \alpha_i^{(h)}$ be the conjugates in the field K , $i = 1, \dots, m$, and let $r \in \mathbb{N}$ be chosen such that from Lemma 1, $r\alpha_i \in \mathbb{Z}_K$. Hence by definition and properties of \mathbb{Z}_K

$$\xi = r^k P(\alpha_1, \dots, \alpha_m) \in \mathbb{Z}_K. \quad (1.62)$$

Suppose $h = 1$. Then $\alpha_i = \alpha_i^{(1)}$, $i = 1, \dots, m$. From (1.62) we know that $\xi \in \mathbb{Z}_K$. Hence there must exist a minimum irreducible polynomial $f(x) \in \mathbb{Z}[x]$ such that $f(\xi) = 0$. Since $\deg \xi = 1$,

$$\therefore f(x) = x + c_0, \quad c_0 \in \mathbb{Z}, \quad \xi \neq 0.$$

Since $\xi \in \mathbb{Z}$,

$$\begin{aligned} \therefore \xi &= -c_0 \\ \implies |\xi| &\geq 1. \end{aligned}$$

Hence we get

$$\begin{aligned} |\xi| = r^k |P(\alpha_1, \dots, \alpha_m)| &\geq 1, \\ |P(\alpha_1, \dots, \alpha_m)| &\geq \frac{1}{r^k}. \end{aligned} \tag{1.63}$$

We define $c = \frac{1}{r}$. Hence

$$|P(\alpha_1, \dots, \alpha_m)| \geq \frac{1}{r^k} = \frac{c^k}{H^{(1-1)}}, \quad n = 1.$$

Therefore we get the form of (1.61) for $h = 1$.

Now let us suppose that $h > 1$. Let us denote

$$A_j = P(\alpha_1^{(j)}, \dots, \alpha_m^{(j)}), \quad (j = 1, \dots, h).$$

Let us recall that two algebraic numbers α and β are said to conjugates if and only if there exist an isomorphism

$$\sigma : \mathbb{Q}(\alpha) \longrightarrow \mathbb{Q}(\beta)$$

such that

$$\sigma(\alpha) = \beta.$$

Let $\alpha_i^{(q)}$ be conjugates of α_i , $i = 1, \dots, m$. There exist \mathbb{Q} -isomorphisms

$$\sigma_1 : \alpha_1^{(q)} \longrightarrow \alpha_1^{(l)}, \dots, \dots, \sigma_m : \alpha_m^{(q)} \longrightarrow \alpha_m^{(l)}$$

and σ_i fixes every $\alpha_j^{(l)}$ when $i \neq j$, i.e. $\sigma_i(\alpha_j^{(l)}) = \alpha_j^{(l)}$,

Let us denote $\sigma = \sigma_m \circ \sigma_{m-1} \circ \dots \circ \sigma_2 \circ \sigma_1$. Then

$$\begin{aligned} \sigma(A_q) &= \sigma_m \circ \sigma_{m-1} \circ \dots \circ \sigma_2 \circ \sigma_1(P(\alpha_1^{(q)}, \sigma_2^{(q)}, \dots, \sigma_m^{(q)})) \\ &= \sigma_m \circ \sigma_{m-1} \circ \dots \circ \sigma_3 \circ \sigma_2(P(\alpha_1^{(l)}, \sigma_2^{(q)}, \dots, \sigma_m^{(q)})) \\ &= \sigma_m \circ \sigma_{m-1} \circ \dots \circ \sigma_4 \circ \sigma_3(P(\alpha_1^{(l)}, \sigma_2^{(l)}, \dots, \sigma_m^{(q)})) \\ &\vdots \\ &= P(\alpha_1^{(l)}, \sigma_2^{(l)}, \dots, \sigma_m^{(l)}) \\ &= A_l. \end{aligned}$$

Hence the numbers A_1, \dots, A_n are conjugates in the field K .

We define

$$\begin{aligned}\xi_j &= r^k P(\alpha_1^{(j)}, \dots, \alpha_m^{(j)}), \quad (j = 1, \dots, h) \\ &= r^k A_j, \quad \xi_j \in \mathbb{Z}_k.\end{aligned}$$

Then from Lemma 1, we conclude that

$$\xi = \xi_1, \dots, \xi_n$$

are conjugates of ξ . Since $\xi_j \in \mathbb{Z}_K$ there exist a $f(x) \in \mathbb{Z}[x]$ such that

$$f(x) = x^h + d_{n-1}x^{n-1} + \dots + d_0 \quad \text{and} \quad f(\xi_j) = 0,$$

$$d_y \in \mathbb{Z}, \quad (y = 0, \dots, n-1); \quad d_0 \neq 0.$$

Let us recall that

$$\begin{aligned}\mathcal{N}(\xi) &= (-1)^n \xi_1 \xi_2 \dots \xi_h = (-1)^h d_0. \\ \therefore |\mathcal{N}(\xi)| &= |\mathcal{N}(r^k A_1)| = r^{kh} |A_1 \dots A_n| = |d_0| \geq 1.\end{aligned}$$

Therefore

$$|A_1| \geq \left(r^{kh} \prod_{j=2}^h |A_j| \right)^{-1}. \quad (1.64)$$

Let us note that

$$P(z_1, \dots, z_m) = \gamma_1 z_1^{k_{11}} \dots z_m^{k_{m1}} + \gamma_2 z_1^{k_{12}} \dots z_m^{k_{m2}} + \dots + \gamma_k z_1^{k_{1k}} \dots z_m^{k_{mk}} \quad (1.65)$$

$$= \sum_{0 \leq k_1 + \dots + k_m \leq k} c_{k_1, \dots, k_m} z_1^{k_1} \dots z_m^{k_m}, \quad c_{k_1, \dots, k_m} \in \mathbb{Z} \quad (1.66)$$

where $\gamma_\eta \in \mathbb{Z}$, $\eta_j = k_{i_j} + \dots + k_{m_j}$, $\forall j$, $(1 \leq j \leq k)$.

Then from (1.65) and the definition of A_j , we have

$$|A_j| = |\gamma_1| |\alpha_1^{(j)}|^{k_{11}} \dots |\alpha_m^{(j)}|^{k_{m1}} + \dots + |\gamma_k| |\alpha_1^{(j)}|^{k_{1k}} \dots |\alpha_m^{(j)}|^{k_{mk}}$$

where

$$0 \leq k_{1_w} + \dots + k_{m_w} \leq k, \quad (1 \leq w \leq k).$$

Then,

$$\begin{aligned}
|A_j| &\leq H(|\alpha_1|^{k_{1_1}} \cdots |\alpha_m|^{k_{m_1}} + \cdots + |\alpha_1|^{k_{1_k}} \cdots |\alpha_m|^{k_{m_k}}) \\
&\leq H \left(1 + \sum_{i=1}^m |\alpha_i| \right)^k \\
&= \rho_0^k H, \quad \rho_0 = 1 + \sum_{i=1}^m |\alpha_i|.
\end{aligned} \tag{1.67}$$

From (1.64) and (1.67) we obtain:

$$\begin{aligned}
|P(\alpha_1, \dots, \alpha_m)| &> \frac{1}{(r^h \rho_0^{h-1})^k H^{h-1}} \\
&= \frac{\rho^k}{H^{h-1}}, \quad \rho = \frac{1}{r^h \rho_0^{h-1}}.
\end{aligned} \tag{1.68}$$

Hence for the case of $h > 1$, we obtain the desired inequality. Using the Corollary to Lemma 3, we can further refine (1.68) to get:

$$|P(\alpha_1, \dots, \alpha_m)| > \frac{\rho_1^k}{H^{h-1}}, \quad \rho_1 = \frac{1}{r^h \left(1 + \sum_{i=1}^m h_i \right)}$$

where h_i is the height of the algebraic number α_i , $1 \leq i \leq m$. □

Chapter 2

Transcendence of π and applications

In this chapter, we shall investigate and prove the transcendence of π . It should be noted from the previous chapter that the class of real transcendental numbers are irrational. Hence, our strategy is to first prove the irrationality of π and using analytic methods and theorems and lemmas from Chapter 1 to show the transcendence of π . However it should be noted that proving the transcendence of π is sufficient to show that π is irrational. Lastly, we will apply the transcendence property of π to show the impossibility of squaring the circle.

1. Transcendence of π

Lemma 1. *For any fix c , $c \in \mathbb{R}$, we have $\lim_{n \rightarrow \infty} \frac{c^n}{n!} = 0$.*

Proof. The case for $c = 0$ is obvious. We need only to consider the case for which $c \in \mathbb{R}^+$. The proof for $c < 0$ is exactly the same for $c > 0$. Let us observe that

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

Let us fix $M > c$, $M \in \mathbb{Z}$. Then for any $n > M$,

$$a_n = \frac{c \cdot c \cdots c \cdot c}{n(n-1) \cdots (M+1)} a_M < \frac{c}{n} a_M.$$

Hence by squeeze theorem, $\lim_{n \rightarrow \infty} a_n = 0$. □

Theorem 13. π is irrational.

Proof. Suppose on the contrary that π is rational, i.e. there exist $a, b \in \mathbb{N}$ such that $\pi = a/b$.

Let us define a function $f(x)$ such that

$$f(x) = \frac{x^n(a - bx)^n}{n!}, \quad \pi = \frac{a}{b}. \quad (2.1)$$

Then $f(x)$ is a polynomial of degree $2n$ with rational coefficients and that given any x such that $0 < x < a/b$, $f(x) \neq 0$. Expanding the function we obtain

$$f(x) = \frac{a_0x^n + a_1x^{n+1} + \dots + a_nx^{2n}}{n!}$$

where

$$a_0 = a^n, a_1 = -na^{n-1}b, \dots, a_i = \frac{(-1)^i n!}{i!(n-i)!} a^{n-i} b^i, \dots, a_n = (-1)^n b^n$$

are integers.

Let us observe the function of $f(x)$.

$$\begin{aligned} f(x) &= \frac{b^n x^n \left(\frac{a}{b} - x\right)^n}{n!} \\ &= \frac{b^n x^n (\pi - x)^n}{n!} \\ &= f(\pi - x). \end{aligned}$$

Hence

$$f^{(i)}(x) = (-1)^i f^{(i)}(\pi - x), \quad \forall x. \quad (2.2)$$

Recall that $\deg f(x) = 2n$. This means that for all $m > 2n$, $f^{(m)}(x) = 0$. Furthermore, from the definition of $f(x)$, for all $i < n$, $f^{(i)}(0) = 0$. Hence we shall consider i for $n \leq i \leq 2n$. We observe that

$$f^{(i)}(0) = \frac{i! a_{i-n}}{n!}, \quad n \leq i \leq 2n. \quad (2.3)$$

Since $i > n$, this implies that $i!/n!$ is a positive integer and a_{i-n} is also an integer. Hence $f^{(i)}(0)$ is an integer for all non negative i . Since $f^{(i)}(x) = (-1)^i f^{(i)}(\pi - x)$, hence $f^{(i)}(\pi)$ must be an integer for all non negative i . Let us define

$$F(x) = f(x) - f^{(2)}(x) + \dots + (-1)^n f^{(2n)}(x).$$

Recall that since $f^{(m)}(x) = 0$ if $m > 2n$. we deduce

$$\begin{aligned}\frac{d^2F}{dx^2} &= F''(x) = f^{(2)}(x) - f^{(4)}(x) + \cdots + (-1)^{n-1}f^{(2n)}(x) \\ &= -F(x) + f(x).\end{aligned}\tag{2.4}$$

Hence

$$\begin{aligned}\frac{d}{dx}(F'(x) \sin x - F(x) \cos x) &= F''(x) \sin x + F'(x) \cos x \\ &\quad - F'(x) \cos x + F(x) \sin(x) \\ &= (F''(x) + F(x)) \sin x = f(x) \sin x.\end{aligned}\tag{2.5}$$

Thus

$$\begin{aligned}\int_0^\pi f(x) \sin x dx &= [F'(x) \sin x - F(x) \cos x]_0^\pi \\ &= F(\pi) + F(0).\end{aligned}$$

Since we have proven for all non-negative i , $f^{(i)}(0)$ and $f^{(i)}(\pi)$ are integers, hence

$$F(\pi) + F(0) \in \mathbb{Z}.$$

Let us consider $f(x) \sin x$. Note that for all $0 < x \leq \pi$, $0 < \sin x \leq 1$. Hence,

$$0 < f(x) \sin x \leq f(x) = \frac{x^n(a - bx)^n}{n!} \leq \frac{\pi^n a^n}{n!}.\tag{2.6}$$

Let us define $u = a\pi$. Recall the Taylor Series of e^u :

$$e^u = \sum_{n=0}^{\infty} \frac{u^n}{n!}.$$

Since the function $\sum_{n=0}^{\infty} \frac{u^n}{n!}$ converges, therefore from Lemma 1,

$$\lim_{n \rightarrow \infty} \frac{u^n}{n!} = 0.$$

Let us choose a n large enough such that

$$\frac{u^n}{n!} < \frac{1}{\pi}.$$

Then from (2.6) we get

$$\begin{aligned}0 &< \int_0^\pi f(x) \sin x dx \leq \int_0^\pi \frac{u^n}{n!} \\ \therefore 0 &< \int_0^\pi f(x) \sin x dx \leq \frac{\pi u^n}{n!} < 1.\end{aligned}$$

But this contradicts the result that

$$\int_0^\pi f(x) \sin x dx = F(\pi) + F(0) \in \mathbb{Z}.$$

Hence π must be irrational. □

Hermite's Identity. Let $f(x)$ be a degree ν polynomial with real coefficients and set

$$F(x) = f(x) + f'(x) + \cdots + f^{(\nu)}(x). \quad (2.7)$$

Then

$$e^x \int_0^x f(t)e^{-t} dt = F(0)e^x - F(x). \quad (2.8)$$

Proof. Let us consider

$$\int_0^x f(t)e^{-t} dt.$$

Integrating by parts, we get

$$\int_0^x f(t)e^{-t} dt = f(0) - f(x)e^{-x} + \int_0^x f'(t)e^{-t} dt. \quad (2.9)$$

Repeating this process $\nu + 1$ times, we get

$$\int_0^x f(t)e^{-t} dt = F(0) - F(x)e^{-x}.$$

Hence we get the equation in (2.8). □

Lemma 2. If $g(x) \in \mathbb{Z}[x]$, then for any $k \in \mathbb{N}$, all the coefficients of the k -th derivative $g^{(k)}(x)$ are divisible by $k!$.

Proof. Let us consider the polynomial x^s , $s > 0$. Suppose $k > s$, then the k -th derivative of x^s is 0. Suppose $1 \leq k \leq s$, then the k -th derivative of x^s is

$$k! \binom{s}{k} x^{s-k},$$

in which the binomial coefficient

$$\binom{s}{k}$$

is an integer. Hence the lemma is proven. □

Lemma 3 (Leibniz's Formula). *Let f and g be analytic in a domain D . Then the n th derivatives of the product fg is given by:*

$$(fg)^{(n)}(z) = \sum_{m=0}^n \binom{n}{m} f^{(n-m)}(z)g^{(m)}(z) \quad \forall z \in D,$$

where $f^{(0)} = f$, $g^{(0)} = g$ and

$$\binom{n}{m} = \frac{n!}{(n-m)!m!}$$

is the m th binomial coefficient.

Proof. We prove the above Lemma using induction.

Let f, g be analytic in a domain D .

Let P_n be the statement " $(fg)^{(n)}(z) = \sum_{m=0}^n \binom{n}{m} f^{(n-m)}(z)g^{(m)}(z)$, $\forall z \in D$ "

Consider P_1 . The case for $n = 1$ is just the product rule. Hence,

$$\begin{aligned} (fg)^{(1)}(z) &= fg^{(1)}(z) + f^{(1)}g(z) \\ &= \sum_{m=0}^1 \binom{1}{m} f^{(1-m)}(z)g^{(m)}(z) \end{aligned}$$

Thus we prove the case for $n = 1$. Suppose P_k is true for some $k \in \mathbb{Z}^+$. Then we get

$$(fg)^{(k)}(z) = \sum_{m=0}^k \binom{k}{m} f^{(k-m)}(z)g^{(m)}(z) \quad \forall z \in D,$$

We differentiate the equation with respect to z to get:

$$\frac{d}{dz} ((fg)^{(k)}(z)) = \frac{d}{dz} \left(\sum_{m=0}^k \binom{k}{m} f^{(k-m)}(z)g^{(m)}(z) \right).$$

Expanding the equation, we get

$$\begin{aligned} \frac{d}{dz} \left\{ f^{(k)}g + \binom{k}{1} f^{(k-1)}g^{(1)} + \binom{k}{2} f^{(k-2)}g^{(2)} + \dots + \binom{k}{k-1} f^{(1)}g^{(k-1)} + fg^{(k)} \right\} = \\ \left\{ f^{(k+1)}g + \binom{k}{0} f^{(k)}g^{(1)} \right. \\ \left. + \binom{k}{1} f^{(k)}g^{(1)} + \binom{k}{1} f^{(k-1)}g^{(2)} \right\} \end{aligned}$$

$$\begin{aligned}
& + \binom{k}{2} f^{(k-1)} g^{(2)} + \binom{k}{2} f^{(k-2)} g^{(3)} \\
& + \binom{k}{3} f^{(k-2)} g^{(3)} + \binom{k}{3} f^{(k-3)} g^{(4)} \\
& \quad \vdots \\
& + \binom{k}{k-1} f^{(2)} g^{(k-1)} + \binom{k}{k-1} f^{(1)} g^{(k)} \\
& + f^{(1)} g^{(k)} + f^{(0)} g^{(k+1)} \}.
\end{aligned}$$

But let us recall that:

$$\begin{aligned}
\binom{k}{m} + \binom{k}{m-1} &= \frac{k!}{(k-m)!m!} + \frac{k!}{(k-m+1)!m!} \\
&= \frac{k!}{(k-m)!m!} \left(1 + \frac{m}{k-m+1} \right) \\
&= \frac{k!}{(k-m)!m!} \left(\frac{k-m+1+m}{k-m+1} \right) \\
&= \frac{(k+1)!}{(k-m+1)!m!} \\
&= \frac{(k+1)!}{(k+1-m)!m!} \\
&= \binom{k+1}{m}.
\end{aligned}$$

Summing the terms, we get

$$\begin{aligned}
(fg)^{(k+1)}(z) &= f^{(k+1)}g + \binom{k+1}{1} f^{(k)} g^{(1)} + \dots + \binom{k+1}{k} f^{(1)} g^{(k)} + fg^{(k+1)} \\
&= \sum_{m=0}^k \binom{k}{m} f^{(k-m)}(z) g^{(m)}(z), \quad \forall z \in D.
\end{aligned}$$

Hence P_{k+1} is true.

P_k is true $\implies P_{k+1}$ is true.

$P_1 \implies P_2 \implies P_3 \dots$

Therefore using induction, we proved Lemma 3.

Theorem 14. π is transcendental (Lindemann).

Consider the following equation (*Euler's Identity*)

$$e^{\pi i} + 1 = 0. \tag{2.10}$$

Let us assume that π is an algebraic number. Then there must exist a $f(x) \in \mathbb{Q}[x]$ such that $f(\pi) = 0$. Since π is algebraic, therefore $\gamma = \pi i$ is also algebraic.

Let $\nu = \deg \gamma$ and let $\gamma = \gamma_1, \dots, \gamma_\nu$ be the conjugates of γ . Let us note that since

$$\begin{aligned} e^\gamma + 1 &= 0; \\ \implies (1 + e^{\gamma_1})(1 + e^{\gamma_2}) \cdots (1 + e^{\gamma_\nu}) &= 0; \\ \implies \prod_{i=1}^{\nu} (1 + e^{\gamma_i}) &= 0; \quad i = (0, 1, \dots, \nu). \end{aligned} \tag{2.11}$$

We shall now prove by induction that

$$\prod_{i=1}^{\nu} (1 + e^{\gamma_i}) = \sum_{\epsilon_1=0}^1 \cdots \sum_{\epsilon_\nu=0}^1 e^{\epsilon_1 \gamma_1 + \cdots + \epsilon_\nu \gamma_\nu}.$$

Let P_ν be the statement:

$$\prod_{i=1}^{\nu} (1 + e^{\gamma_i}) = \sum_{\epsilon_1=0}^1 \cdots \sum_{\epsilon_\nu=0}^1 e^{\epsilon_1 \gamma_1 + \cdots + \epsilon_\nu \gamma_\nu}; \quad \forall \nu \in \mathbb{Z}^+.$$

Note that P_1 is obviously true. We shall prove P_2 .

Consider P_2 where $\gamma = \gamma_1, \gamma_2$ are conjugates of γ . Then,

$$\begin{aligned} (1 + e^{\gamma_1})(1 + e^{\gamma_2}) &= 1 + e^{\gamma_1} + e^{\gamma_2} + e^{\gamma_1 + \gamma_2} \\ &= \sum_{\epsilon_1=0}^1 \sum_{\epsilon_2=0}^1 e^{\epsilon_1 \gamma_1 + \epsilon_2 \gamma_2}. \end{aligned}$$

Therefore P_2 is true.

Assume that P_k is true for some $k \in \mathbb{Z}^+$.

$$\therefore \prod_{i=1}^k (1 + e^{\gamma_i}) = \sum_{\epsilon_1=0}^1 \cdots \sum_{\epsilon_k=0}^1 e^{\epsilon_1 \gamma_1 + \cdots + \epsilon_k \gamma_k}$$

where $\gamma_1, \dots, \gamma_k$ are conjugates of γ .

Let γ_{k+1} be a conjugate of γ such that $\gamma_{k+1} \neq \gamma_i, i = 1, \dots, k$. Hence,

$$\begin{aligned} (e^{\gamma_{k+1}} + 1) \prod_{i=1}^k (1 + e^{\gamma_i}) &= 0; \\ (e^{\gamma_{k+1}} + 1) \sum_{\epsilon_1=0}^1 \cdots \sum_{\epsilon_k=0}^1 e^{\epsilon_1 \gamma_1 + \cdots + \epsilon_k \gamma_k} &= 0; \end{aligned}$$

$$\begin{aligned}
& \sum_{\epsilon_1=0}^1 \cdots \sum_{\epsilon_k=0}^1 e^{\epsilon_1 \gamma_1 + \cdots + \epsilon_k \gamma_k + \gamma_{k+1}} + \sum_{\epsilon_1=0}^1 \cdots \sum_{\epsilon_k=0}^1 e^{\epsilon_1 \gamma_1 + \cdots + \epsilon_k \gamma_k} = 0; \\
& \sum_{\epsilon_1=0}^1 \cdots \sum_{\epsilon_k=0}^1 e^{\epsilon_1 \gamma_1 + \cdots + \epsilon_k \gamma_k + 1 \cdot \gamma_{k+1}} + \sum_{\epsilon_1=0}^1 \cdots \sum_{\epsilon_k=0}^1 e^{\epsilon_1 \gamma_1 + \cdots + \epsilon_k \gamma_k + 0 \cdot \gamma_{k+1}} = 0; \\
& \sum_{\epsilon_1=0}^1 \cdots \sum_{\epsilon_k=0}^1 \sum_{\epsilon_{k+1}=0}^1 e^{\epsilon_1 \gamma_1 + \cdots + \epsilon_k \gamma_k + \epsilon_{k+1} \gamma_{k+1}} = 0.
\end{aligned}$$

Thus,

$$\prod_{i=1}^{k+1} (1 + e^{\gamma_i}) = \sum_{\epsilon_1=0}^1 \cdots \sum_{\epsilon_k=0}^1 \sum_{\epsilon_{k+1}=0}^1 e^{\epsilon_1 \gamma_1 + \cdots + \epsilon_k \gamma_k + \epsilon_{k+1} \gamma_{k+1}}.$$

Hence by induction, we get

$$\prod_{i=1}^{\nu} (1 + e^{\gamma_i}) = \sum_{\epsilon_1=0}^1 \cdots \sum_{\epsilon_{\nu}=0}^1 e^{\epsilon_1 \gamma_1 + \cdots + \epsilon_{\nu} \gamma_{\nu}}. \quad (2.12)$$

Let m denote the number of non-zero exponents in (2.12) and $a = 2^{\nu} - m$ be the number of exponents that are zero, $a \geq 1$. Let us define $\alpha_1, \dots, \alpha_m$ as the non-zero exponents. Hence from (2.12)

$$\begin{aligned}
a \cdot e^0 + e^{\alpha_1} + \cdots + e^{\alpha_m} &= 0, \\
a + e^{\alpha_1} + \cdots + e^{\alpha_m} &= 0.
\end{aligned} \quad (2.13)$$

Consider the polynomial

$$\varphi(x) = \prod_{\epsilon_1=0}^1 \cdots \prod_{\epsilon_{\nu}=0}^1 \left(x - (\epsilon_1 \gamma_1 + \cdots + \epsilon_{\nu} \gamma_{\nu}) \right).$$

We observe that $\varphi(x)$ is a symmetrical polynomial in $\gamma_1, \dots, \gamma_{\nu}$ with coefficients in $\mathbb{Z}[x]$. Hence by Lemma 2 of Chapter 1, $\varphi(x) \in \mathbb{Q}[x]$. From (2.13) the roots of the degree 2^{ν} polynomial $\varphi(x)$ are $\alpha_1, \dots, \alpha_m$ and 0 with multiplicity a . [This is because there are a possible ways of getting 0 exponents in (2.13)]

Hence the deg m polynomial $x^{-a} \varphi(x) \in \mathbb{Q}[x]$ has precisely the numbers $\alpha_1, \dots, \alpha_m$ as its roots. Let $r \in \mathbb{N}$ be the *least common denominator* of the coefficients of this polynomial.

We define

$$\psi(x) = \frac{r}{x^a} \varphi(x) = b_m x^m + \cdots + b_1 x + b_0, \quad \psi(x) \in \mathbb{Z}[x], \quad b_m > 0, \quad b_0 \neq 0$$

has precisely $\alpha_1, \dots, \alpha_m$ as its roots.

Let us recall Hermite's Identity:

$$\int_0^x f(t)e^{-t} dt = F(0) - F(x)e^{-x}$$

where $F(x) = f(x) + f^{(1)}(x) + \dots + f^{(\nu)}(x)$, $\deg(f(x)) = \nu$

Let $x = \alpha_1, \dots, \alpha_m$. Let us consider α_k , $1 \leq k \leq m$. Then by Hermite's Identity we get,

$$\begin{aligned} \int_0^{\alpha_k} f(t)e^{-t} dt &= F(0) - F(\alpha_k)e^{-\alpha_k}, \\ e^{\alpha_k} \int_0^{\alpha_k} f(t)e^{-t} dt &= e^{\alpha_k}F(0) - F(\alpha_k). \end{aligned}$$

Then summing up all the conjugates of α , we get

$$\sum_{k=1}^m e^{\alpha_k} \int_0^{\alpha_k} f(t)e^{-t} dt = F(0) \sum_{k=1}^m e^{\alpha_k} - \sum_{k=1}^m F(\alpha_k).$$

Recall from (2.13)

$$\sum_{k=1}^m e^{\alpha_k} = -a.$$

Therefore we get

$$-aF(0) - \sum_{k=1}^m F(\alpha_k) = \sum_{k=1}^m e^{\alpha_k} \int_0^{\alpha_k} f(t)e^{-t} dt.$$

Let us define

$$f(x) = \frac{1}{(n-1)!} b_m^{nm-1} x^{n-1} \psi^n(x) \tag{2.14}$$

$$= \frac{1}{(n-1)!} b_m^{(m+1)n-1} x^{n-1} (x - \alpha_1)^n \dots (x - \alpha_m)^n. \tag{2.15}$$

Let us note that

$$\begin{aligned} \deg f(x) &= n - 1 + nm \\ &= (m+1)n - 1 \geq n - 1. \end{aligned}$$

Hence from (2.14), we apply Leibniz Formula to get:

$$f^{(l)}(0) = 0, \quad (l = 0, 1, \dots, n-2), \quad f^{(n-1)}(0) = b_m^{m-1} b_0^n$$

and from Hermite's Identity and the definition of $f(x)$, we get

$$F(0) = \sum_{l=n-1}^{(m+1)n-1} f^{(l)}(0) = b_m^{mn-1} b_0^n + nA, \quad A \in \mathbb{Z} \quad (\because \varphi(x) \in \mathbb{Z}[x]). \quad (2.16)$$

Since α_k is a root of $f(x)$ of multiplicity n , therefore

$$f^{(l)}(\alpha_k) = 0, \quad (l = 0, 1, \dots, n-1), \quad (k = 1, \dots, m).$$

From Lemma 2, the l th derivative of $x^{n-1}\psi^n(x)$ has integer coefficients which are divisible by $n!$. Thus for $l > n$, the coefficients of $f^{(l)}(x)$ are integers divisible by nb_m^{mn-1} . Thus,

$$\begin{aligned} F(\alpha_k) &= \sum_{l=n}^{(m+1)n-1} f^{(l)}(\alpha_k) \\ &= nb_m^{mn-1} \Phi(\alpha_k), \quad (k = 1, \dots, m), \quad \Phi(x) \in \mathbb{Z}[x]. \end{aligned} \quad (2.17)$$

Let us consider $\beta_k = b_m \alpha_k$, $k = 1, \dots, m$. From Lemma 1 of Chapter 1, we can construct a polynomial of degree m in $\mathbb{Z}[x]$ such that β_k form the roots of the polynomial. Hence β_k is an algebraic integer. We can write

$$b_m^{mn-1} \Phi(\alpha_k) = H(\beta_k), \quad H(x) \in \mathbb{Z}[x].$$

We note that $\Phi(x)$ is a symmetrical polynomial. Hence by using Lemma 2 of Chapter 1, we conclude

$$\sum_{k=1}^m b_m^{mn-1} \Phi(\alpha_k) = \sum_{k=1}^m H(\beta_k) = B. \quad (2.18)$$

From (2.16), (2.17), (2.18) we get

$$aF(0) + \sum_{k=1}^m F(\alpha_k) = ab_0^n b_m^{mn-1} + n(aA + B) \in \mathbb{Z}. \quad (2.19)$$

Let us choose $n \in \mathbb{N}$ such that

$$\gcd(n, b_0 b_m) = 1, \quad n > a. \quad (2.20)$$

Since the right hand side of (2.19) is an integer and if it satisfies (2.20), then we arrive at the conclusion

$$\left| aF(0) + \sum_{k=1}^m F(\alpha_k) \right| \geq 1 \quad (2.21)$$

since $ab_0^n b_m^{mn-1}$ is not a factor of n .

Let us consider a circle $|x| < R$, $R > 0$, $x \in \mathbb{C}$. We choose the circle such such that all the points $\alpha_1, \dots, \alpha_m$ are contained within the circle. Let us denote

$$\max_{|x| \leq R} |b_m^m \psi(x)| = C$$

in which C is independent of n .

From Hermite's Identity, we obtain

$$\begin{aligned} \left| \sum_{k=1}^m e^{\alpha_k} \int_0^{\alpha_k} f(x) e^{-x} dx \right| &\leq \sum_{k=1}^m \left| \int_0^{\alpha_k} |f(x)| |e^{(\alpha_k - x)}| dx \right| \\ &\leq \frac{R^{n-1} e^R}{(n-1)!} C^n \sum_{k=1}^m \left| \int_0^{\alpha_k} dx \right| \\ &\leq m e^R \frac{(RC)^n}{(n-1)!}. \end{aligned} \tag{2.22}$$

Let us consider the following expression

$$\frac{(RC)^n}{(n-1)!}. \tag{2.23}$$

From Lemma 1, we know that

$$\lim_{n \rightarrow \infty} \frac{(RC)^n}{(n-1)!} = \lim_{n \rightarrow \infty} \frac{(RC)^n}{n!} = 0.$$

This implies that there exist an n_0 such that for a sufficiently large n such that $n \geq n_0$ and n satisfies (2.20), we get

$$\frac{(RC)^n}{(n-1)!} < \frac{1}{m e^R}. \tag{2.24}$$

Hence from (2.22) and (2.24)

$$\left| \sum_{k=1}^m e^{\alpha_k} \int_0^{\alpha_k} f(x) e^{-x} dx \right| < 1.$$

However, this is a contradiction to (2.21). Therefore we are forced to conclude that π is transcendental. □

2. Squaring the circle

In this section, we will consider the problem of *squaring the circle* or *quadrature of the circle* as written in some journals. We approach the problem by asking ourselves whether

it is possible, by using only compass and a straight edge, to obtain a square whose area is equal to that of a circle. We will use the result from the previous section that π is transcendental to show that it is impossible to square the circle.

Theorem 15. *Let $L \supset K \supset F$ be three fields such that both $[L : K]$ and $[K : F]$ are finite. Then L is a finite extension of F and $[L : F] = [L : K][K : F]$.*

Proof. We shall prove that L is a finite extension of F by exhibiting a finite basis of L over F . In doing so, we shall obtain the stronger result that $[L : F] = [L : K][K : F]$.

Suppose that $[L : K] = m$ and $[K : F] = n$ then L has a basis

$$v_1, v_2, \dots, v_m$$

over K , and K has a basis

$$w_1, w_2, \dots, w_n$$

over F . We shall prove that the mn elements $v_i w_j$ where $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$ constitute a basis of L over F . Let $a \in L$. Since v_1, v_2, \dots, v_m form a basis of L over K , we have

$$a = k_1 v_1 + \dots + k_m v_m, \quad k_1, k_2, \dots, k_m \in K. \quad (2.25)$$

Since w_1, w_2, \dots, w_n is a basis of K over F , we can express each k_i as

$$k_i = f_{i1} w_1 + f_{i2} w_2 + \dots + f_{in} w_n \quad (2.26)$$

where

$$f_{ij} \in F, \quad (i = 1, 2, \dots, m), \quad (j = 1, 2, \dots, n).$$

Substituting (2.25) in (2.26), we get,

$$\begin{aligned} a &= (f_{11} w_1 + f_{12} w_2 + \dots + f_{1n} w_n) v_1 \\ &\quad + \dots + (f_{m1} w_1 + f_{m2} w_2 + \dots + f_{mn} w_n) v_m. \end{aligned} \quad (2.27)$$

Rearranging this equation, we obtain,

$$a = f_{11} w_1 v_1 + f_{12} w_2 v_1 + \dots + f_{ij} w_j v_i + \dots + f_{mn} w_n v_m. \quad (2.28)$$

This tells us that L is a finite extension of F and hence $[L : F]$ is defined and finite and that $[L : F] \leq mn$. Suppose $a = 0$. Then for some $b_{ij} \in F$, we have

$$b_{11}w_1v_1 + b_{12}w_2v_1 + \cdots + b_{ij}w_jv_i + \cdots + b_{mn}w_nv_m = 0.$$

Let us denote

$$c_i = \sum_{j=1}^n b_{ij}w_j, \quad (i = 1, \dots, m). \quad (2.29)$$

Then we get

$$c_1v_1 + c_2v_2 + \cdots + c_mv_m = 0.$$

But since v_1, \dots, v_m are basis of F over K , therefore

$$c_1 = c_2 = \cdots = c_m = 0.$$

Hence

$$\sum_{j=1}^n b_{ij}w_j = 0, \quad (i = 1, 2, \dots, m).$$

But w_1, w_2, \dots, w_n are basis of K over L . Therefore

$$b_{ij} = 0, \quad (i = 1, \dots, m), \quad (j = 1, \dots, n). \quad (2.30)$$

Hence the set

$$S = \{v_iw_j \mid (i = 1, \dots, m), (j = 1, \dots, n)\}$$

is a linearly independent set and it forms a basis of F over L . It is easy to see that there are nm elements in S . Hence we conclude that,

$$[L : F] = nm = [L : K][K : F]. \quad (2.31)$$

This proves the theorem. □

Corollary. If $L \supset K \supset F$ are three fields such that $[L : F]$ is finite, then $[K : F]$ is finite and divides $[L : F]$.

Proof. Since $L \supset K$, K cannot have more linearly independent elements over F than L . From Linear Algebra, we know that $[L : F]$ is the size of the largest set of linearly independent elements in L over F , therefore we have $[K : F] \in [L : F]$. Since L is finite

dimensional over F and $K \supset F$, therefore L must be finite dimension over K . From theorem 15 we have

$$[L : F] = [L : K][K : F].$$

Hence $[K : F]$ divides $[L : F]$. □

We shall state the following theorem without proof. The proof is beyond the scope of this report.

Theorem 16. *Let $K \supset F$ and suppose that a in K is algebraic over F of degree n . Then $F(a)$, the field obtained by adjoining a to F , is a finite extension of F , and*

$$[F(a) : F] = n.$$

This theorem tells us that if $p(x)$ is the minimal polynomial of a , then $[F(a) : F] = \deg p(x) = n$.

We shall now consider the problem of *constructibility*. We define a straight-edge to be a straight line with no quantitative or metric properties attributed to it. Suppose we are given a line segment - to which we assign length 1 - then all other length that we get are obtained from this straight edge by employing straight-edge and a compass techniques. We shall now introduce the concept of a *constructible length*

Definition 8. *A nonnegative real number, b , is said to be a constructible length if, by a finite number of applications of the straight-edge and compass and the points of intersection obtained between the lines and circles constructed, we can construct a line segment of length b , starting from the line segment we assigned length 1.*

We claim the following properties:

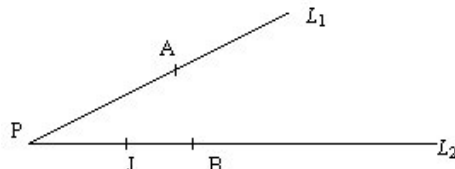
Property 1. If a and b are constructible lengths, then so is $a + b$.

Proof. This property can be easily seen from the application of the straight-edge and compass. Without loss of generality, let us assume that $a > b$. We define the straight-edge AB to be of length a and CD to be of length b . We start up drawing a perpendicular bisector to AB and label the corresponding straight edge as EF . From EF , we again

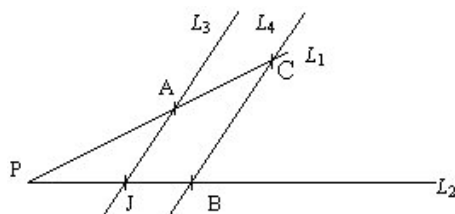
construct a perpendicular bisector, hence extending the straight edge AB . From B , we apply the compass to get the required length $a + b$. The case for $a - b$ is similar. Again assuming that $a > b$, we denote AB as the straight edge whose length is a . By applying the compass at B , we can get the length $a - b$. \square

Property 2. If a and b are constructible lengths, then so is $a \cdot b$.

Proof. The case in which $a = 0$ or $b = 0$ is quite trivial. Hence we may assume that $a \neq 0$ and $b \neq 0$. Consider the following diagram,



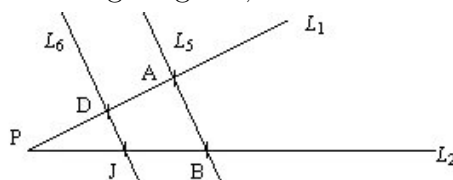
Let L_1 and L_2 be two distinct lines intersecting at P , such that PA has length a and PB has length b . We place J on L_2 such that PJ has length 1. We define L_3 to be the line passing through A and J and L_4 to be the line parallel to L_3 passing through B . Let C be the point of intersection between L_1 and L_4 . This will give us the following diagram,



All these constructions can be done using the straight edge and compass. Using the property of similar triangles, we deduce that the length of PC is $a \cdot b$. Hence $a \cdot b$ is constructible if a and b is constructible. \square

Property 3. If a and b are constructible and $b \neq 0$, then a/b is constructible.

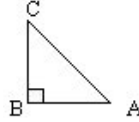
Proof. Let us consider the following diagram,



We denote P , A , B , L_1 , and L_2 as in Property 2. Let L_5 be the line passing through A and B and let L_6 be the line passing through J . Suppose D is the point of intersection between L_6 and L_1 , then by the property of similar triangles, we deduce that line of PD

is a/b . □

This result tells us that all nonnegative rational numbers are constructible lengths, since they are quotients of nonnegative integers, of which were deduced to be of constructible length from Property 3. We would be tempted to say that the class of irrational numbers are not constructible. This is definitely untrue as we can construct $\sqrt{2}$ by constructing a right-angle triangle using the compass and straight edge.



We define AB and BC to be of length 1 each and by Pythagorean Theorem, AC is of length $\sqrt{2}$. We can, in fact, extend this result to construct the following construction.

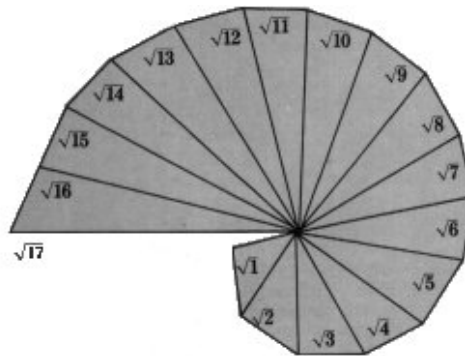


Figure 2.1: Greek Geometrical Construction.

We will like to consider the case for negative constructible lengths. We define the following to get around this problem

Definition 9. *The real number a is said to be a constructible number if $|a|$, the absolute value of a , is a constructible length.*

Theorem 17. *The constructible numbers form a subfield of the field of real numbers*

Proof. Let us denote Γ to be the set of constructible numbers. We define the length 1 to be the unit element in Γ . From Properties 1 to 3, we note that $1 \in \Gamma$.

1) From Property 1, we know that if a, b are constructible, then $a + b$ is constructible. Therefore $a, b \in \Gamma$ implies $a + b \in \Gamma$.

2) From Property 1, it is obvious that if a, b are constructible, then $a + b = b + a$.

3) Let a, b and $c \in \Gamma$. Then $d = a + b$ is constructible from Property 1. Hence $d + c$ is constructible and using field properties defined on \mathbb{R} , we get

$$d + c = (a + b) + c = a + b + c. \quad (2.32)$$

Similar, let us denote $e = b + c$. From Property 1, we know that e is constructible. Hence $a + e$ is also constructible and using field properties defined on \mathbb{R} , we get

$$a + e = a + (b + c) = a + b + c. \quad (2.33)$$

From (2.32) and (2.33), we can conclude that

$$(a + b) + c = a + (b + c).$$

4) We take 0 to be the zero element. Hence 0 is in Γ . Then from Property 1, we note that

$$a + 0 = a = 0 + a$$

for all constructible numbers a .

5) We consider using vectorial methods. Let us fix one direction as positive and the direction directly opposite as negative. From Property 1, we can find a b such that the magnitude of b is that of a but running in an opposite direction as oppose to a . Hence,

$$a + b = a + (-a) = 0.$$

6) Suppose $a, b \in \Gamma$, then from Property 2, $a \cdot b \in \Gamma$.

7) Let $a, b \in \Gamma$. Then by Property 2, $a \cdot b$ and $b \cdot a$ are in Γ . Recall that since \mathbb{R} is a field therefore given any $x, y \in \mathbb{R}$, $x \cdot y = y \cdot x$. Since $\Gamma \subset \mathbb{R}$, therefore we can conclude, using the field properties of \mathbb{R} , that $a \cdot b = b \cdot a$. Hence all elements in Γ obey the commutative law of multiplication.

8) Let a, b and $c \in \Gamma$. Then from Property 2, we can deduce that $a \cdot (b \cdot c)$ and $(a \cdot b) \cdot c$ are in Γ . Recall that $\Gamma \in \mathbb{R}$. From the field properties of \mathbb{R} we know that if x, y and $z \in \mathbb{Z}$, then $x \cdot (y \cdot z)$ and $(x \cdot y) \cdot z \in \mathbb{R}$ and $x \cdot (y \cdot z) = (x \cdot y) \cdot z$. Hence, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

9) From Property 1 and 2, we see that if a , b and c are constructible numbers, then $a \cdot b$, $a \cdot c$ and $a \cdot (b + c)$ must be constructible. Note given any x , y and $z \in \mathbb{R}$, we have

$$x \cdot (y + z) = x \cdot y + x \cdot z.$$

Since $\Gamma \in \mathbb{R}$, therefore given any a , b and $c \in \Gamma$, we have

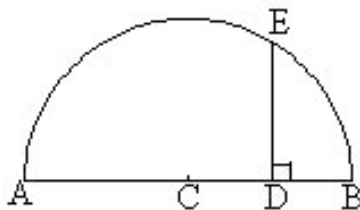
$$a \cdot (b + c) = a \cdot b + a \cdot c.$$

10) From Property 3, we can see that for all $a \in \Gamma$, there exist a a^{-1} such that $a \cdot a^{-1} = 1$.

Hence Γ is a subfield of \mathbb{R} . □

Theorem 18. *Suppose $a \geq 0$ is a constructible number, then \sqrt{a} is constructible.*

Proof. Let us consider the following diagram



It is a semicircle of radius $(a + 1)/2$. We define AD to be of length a and DB to be of length 1. From Property 1 and Property 3, the semicircle AB is constructible. Using geometry and Pythagorean Theorem, we deduce that DE is of length \sqrt{a} . Hence \sqrt{a} is constructible. □

We want to show that a constructible number must be an algebraic number. Let Γ be the field of constructible numbers, and let Γ_0 be a subfield of Γ . We define the *plane* of Γ_0 as the set of points (a, b) in the real Euclidean plane whose coordinates a and b are in Γ_0 . Let (a, b) and (c, d) be in the plane of Γ_0 . From coordinate geometry, the straight line joining (a, b) and (c, d) has the equation

$$\frac{y - b}{x - a} = \frac{b - d}{a - c}. \tag{2.34}$$

We can rearrange the equation to get the form

$$ux + vy + w = 0 \tag{2.35}$$

where u , v and w are in Γ_0 . We consider the following lemma.

Lemma 4. *Given two lines $l_1 : u_1x + v_1y + w_1 = 0$ and $l_2 : u_2 + v_2y + w_2 = 0$, where u_1, v_1, w_1 and u_2, v_2, w_2 are in Γ_0 , then either they are parallel or their point of intersection is a point in Γ_0 .*

Proof. We shall consider the case for which l_1 and l_2 intersect. The case in which l_1 and l_2 are parallel is quite obvious. Let l_1 and l_2 intersect at (a, b) . Since l_1 and l_2 are contained in the plane of Γ_0 , therefore for all points contained in l_1 and l_2 , they are contained in Γ_0 . Hence $(a, b) \in \Gamma_0$. and their point of intersection lie in the plane of extension of Γ_0 of degree 1. \square

Let us consider a circle whose radius r is in Γ_0 and whose (a, b) is in the plane of Γ_0 . From coordinate geometry, the equation of a circle is

$$(x - a)^2 + (y - b)^2 = r^2, \quad (2.36)$$

which when further expanded, gives rise to the form

$$x^2 + y^2 + dx + cy + f = 0, \quad (2.37)$$

where d, e, f are in Γ_0 . We assume that the circle intersects a line l_1 in the plane of Γ_0 , $l_1 : ux + vy + w = 0$. To obtain the points of intersection, we solve the simultaneous equation. Suppose $v \neq 0$, then

$$y = -\frac{ux + w}{v}. \quad (2.38)$$

Substituting (2.37) in (2.38), we get a quadratic equation i.e. $x^2 + s_1x + s_2$, involving the x - coordinate c of the intersection point. Recall from the quadratic formula,

$$c = \frac{-s_1 \pm \sqrt{s_1^2 - 4s_2}}{2}. \quad (2.39)$$

If the line l_1 and the circle intersect on the real plane, then $s = s_1^2 - 4s_2 \geq 0$. If $\Gamma_1 = \Gamma_0(\sqrt{s})$, then we see that c lies in Γ_1 . If $\sqrt{s} \in \Gamma_0$, then $\Gamma_1 = \Gamma_0$; otherwise $[\Gamma_1 : \Gamma_0] = 2$. Let d be the y -coordinate of intersection. Since $d = (-uc + w)/v$, therefore d must be in Γ_1 . Hence (c, d) lies in Γ_1 and $[\Gamma_1 : \Gamma_0] = 1$ or 2 . We can apply similar methods for the case of $v = 0$ and $u \neq 0$.

Finally, let us consider the case of the intersection between two circles, $x^2 + y^2 + dx + ey + f = 0$ and $x^2 + y^2 + gx + hy + f = 0$. By doing a simultaneous equation, we will obtain the

equation of line l_2 , where l_2 is the line passing through the points of intersection. We define l_2 as $l_2 : (d - g)x + (e - h)y + (f - k) = 0$. To find the points of intersection, we just need to do a simultaneous equation between l_2 and one of the circles. This will give us $[L_1 : L_0] = 1$ or 2 as before,

Let us now consider the construction of a . By properties of construction using the straight edge and compass, we begin all construction in the plane of \mathbb{Q} . From before, for any application of the straight edge and the compass, their intersections will give an extension of degree 1 or 2 in \mathbb{Q} . To get a , we need to move from the plane of \mathbb{Q} to L_1 . We get $[L_1 : \mathbb{Q}] = 1$ or 2 . If we get a from this construction we stop the process. If not we carry on to the plane in L_2 and $[L_2 : L_1] = 1$ or 2 . We stop if we get a and repeat the process if we don't. Suppose we obtain a at the L_n plane, then we get a finite sequence $\mathbb{Q} = L_0 \subset L_1 \subset \dots \subset L_m$.

From Theorem 15, $[L_n : \mathbb{Q}] = [L_n : L_{n-1}][L_{n-1} : L_{n-2}] \dots [L_1 : \mathbb{Q}]$. Since each of the $[L_i : L_{i-1}] = 1$ or 2 , $i = 1, \dots, n$, therefore the value of $[L_n : \mathbb{Q}]$ is that of power 2. Since $a \in L_n$, we want to show that $\mathbb{Q}(a)$ is a subfield of L_n . Let us recall the definition of $\mathbb{Q}(a)$. We define

$$\mathbb{Q}(a) = \left\{ \frac{f(a)}{g(a)} \mid f, g \in \mathbb{Q}[x], g(a) \neq 0 \right\}.$$

Since $f(a), g(a) \in L_n$, this implies that $f(a)/g(a) \in L_n$. We can easily show that $\mathbb{Q}(a)$ is a field. Since $f(a)/g(a) \in L_n$ for all $f, g \in \mathbb{Q}[x]$, therefore $\mathbb{Q}(a) \subset L_n$. From the Corollary to Theorem 15, we have $[\mathbb{Q}(a) : \mathbb{Q}] = 2^m$ for some nonnegative integer m . From Theorem 16, we conclude that the minimum polynomial for a over \mathbb{Q} must have degree a of power 2. This condition is *necessary* for constructibility. This leads us the following theorem:

Theorem 19. *In order that the real number a be constructible, it is necessary that $[\mathbb{Q}(a) : \mathbb{Q}]$ be a power of 2. Equivalently, the minimal polynomial of a over \mathbb{Q} have degree of power 2.*

In other words, if a is constructible, it must be algebraic.

Theorem 20. *It is impossible to square the circle.*

Let us construct a circle of radius 1. Suppose that it is possible to square the circle, then we can construct a square of side $\sqrt{\pi}$. This would imply that $\sqrt{\pi}$ is constructible.

From Property 3, this gives us the notion that π is constructible. But from Theorem 14, we know that π is transcendental. Hence from Theorem 19, it cannot be constructible. This provides a contradiction to our argument. Hence, we are forced to conclude that it is impossible to square the circle. \square

Bibliography

- [1] Andrei Borisovich SHIDLOVSKII, *Transcendental Numbers*, Walter de Gruyter, Berlin
- [2] Robert G. BARTLE & Donald R. SHERBERT, *Introduction to Real Analysis 3rd Edition*, John Wiley & Sons, Inc.
- [3] I. NIVEN, “A Simple Proof That π is Irrational”, Bulletin of American Mathematical Society, vol. 53 (1947), p.509
- [4] I.N. HERSTIEN, *Abstract Algebra 3rd Edition*, John Wiley & Sons, Inc.
- [5] David M BURTON, *Elementary Number Theory, 3rd Edition*, Wm. C.Brown Publishers
- [6] *Squaring a circle*, web-page, http://www.cut-the-knot.com/impossible/sq_circle.html
- [7] *Squaring the circle*, web-page, http://www-groups.dcs.st-andrews.ac.uk/~history/HistTopics/Squaring_the_circle.html
- [8] *A history of π* , web-page, http://www-groups.dcs.st-andrews.ac.uk/~history/HistTopics/Pi_through_the_ages.html
- [9] *A chronology of π* , web-page, http://www-groups.dcs.st-andrews.ac.uk/~history/HistTopics/Pi_chronology.html