MATH 195, SPRING 2015

TRANSCENDENTAL NUMBER THEORY

LECTURE NOTES

LENNY FUKSHANSKY

Contents

1. Notation and sets

We will start with some algebraic notation and definitions of all the necessary standard number sets. Our objective is to develop natural motivation for the introduction of different number sets. We leave many technical details to the exercises.

Definition 1.1. Let S be a set and let $* : S \times S \to S$ be a binary operation on S. It is implicit in this notation that S is closed under $*$, i.e. for every $a, b \in S$, $a * b \in S$. This operation is called **associative** if for any $a, b, c \in S$,

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

The operation $*$ is called commutative if $a * b = b * a$ for all $a, b \in S$. The set S with an associative binary operation $*$ on it is called a **semigroup**, we will denote the pair by $(S, *)$. A semigroup $(S, *)$ is called **abelian** if the operation $*$ is commutative on S. A semigroup $(S, *)$ is called a **monoid** if there exists an element $e \in S$, called identity, such that $e * a = a * e = a$ for every $a \in S$. A monoid $(S, *)$ is called a **group** if for any $a \in S$ there exists $b \in S$, called inverse of a and denoted a^{-1}, such that $a * b = b * a = e$.

Exercise 1.1. *Let $(S, *)$ be a group.*

(1) *Prove that identity e in S is unique.*
(2) *Prove that for every $a \in S$, inverse $a^{-1} \in S$ is unique.*
(3) *Prove the cancellation laws:*

$$\text{if } a * b = a * c \text{ then } b = c,$$

$$\text{if } b * a = c * a \text{ then } b = c.$$

A semigroup $(S, *)$ is said to be **generated** by a subset $C \subseteq S$ if every $a \in S$ can be described as a product (with respect to $*$) of some finite collection of elements in C, in any order and possibly with repetition. A **power** of an element $a \in S$ is a product of a with itself (with respect to $*$) some finite number of times. For example, we can define the set of **natural numbers**, which we denote by

$$\mathbb{N} := \{1, 2, 3, \dots\}$$

as the set of all powers of the element 1 under the operation $+$, which we call addition.

Exercise 1.2. *Prove that $(\mathbb{N}, +)$ is an abelian semigroup, but not a monoid.*

We introduce the additional element 0, defined by the property

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

for every $a \in \mathbb{N}$ and write

$$\mathbb{N}_0 = \mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \dots\}.$$

Exercise 1.3. *Prove that $(\mathbb{N}_0, +)$ is a monoid, but not a group.*

For each element $a \in \mathbb{N}$ we define its inverse under $+$, denoted by $-a$. Now the set

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$$

is called **integers**.

Exercise 1.4. *Prove that $(\mathbb{Z}, +)$ is an abelian group.*

Definition 1.2. Let R be a set with two associative binary operations $+, \cdot : R \times R \to R$ on it: we refer to $+$ as addition and to $\cdot$ as multiplication. $(R, +, \cdot)$ is called a **ring** if:

(1) $(R, +)$ is an abelian group.
(2) $(R, \cdot)$ is a semigroup.
(3) Multiplication distributes over addition, i.e.

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

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

The additive identity in R is denoted by 0. A ring $(R, +, \cdot)$ is said to have (multiplicative) **identity** if there exists an element $1 \in R$ such that $1 \cdot a = a \cdot 1 = a$ for all $a \in R$. A ring R is called **commutative** if its multiplication is commutative. For each element a in a ring R, we write $-a$ for its inverse under $+$. If R has identity 1, we write and a^{-1} for the inverse of $a \in R$ under $\cdot$, if it exists. Elements in R that have multiplicative inverses are called **units**, and the set of units is denoted by $R^\times$, i.e.

$$R^\times := \{a \in R : \exists\ b \in R \text{ such that } a \cdot b = 1\}.$$

Remark 1.1. For the operations of addition $+$ and multiplication $\cdot$, we will often write $a - b$ to denote $a + (-b)$ and a/b or $\frac{a}{b}$ to denote $a \cdot b^{-1}$, whenever b^{-1} exists.

Exercise 1.5. *Let $(R, +, \cdot)$ be a ring.*

(1) *Prove that $0 \cdot a = a \cdot a = 0$ for every $a \in R$.*
(2) *If R has identity, prove that $(R^\times, \cdot)$ is a group.*
(3) *Prove that $0 \in R$ is not a unit.*
(4) *Prove that $-a = (-1) \cdot a$ for every $a \in R$.*

Definition 1.3. A commutative ring $(R, +, \cdot)$ with identity is called a **field** if every $0 \neq a \in R$ has a multiplicative inverse, i.e. $R^\times = R \setminus \{0\}$. In other words, $(R, +)$ and $(R \setminus \{0\}, \cdot)$ are both abelian groups.

Definition 1.4. A subset T of a semigroup (respectively, monoid, group, ring, field) S is called a **sub-semigroup** (respectively, **sub-monoid**, **subgroup**, **subring**, **subfield**) if it is itself a semigroup (respectively, monoid, group, ring, field) under the same operation(s) as S.

Exercise 1.6. *Prove that $(\mathbb{Z}, +, \cdot)$ is a commutative ring with identity, but not a field. Furthermore, $\mathbb{Z}^\times = \{1, -1\}$.*

To construct an example of a field, we want to "invert" nonzero integers, i.e. to introduce the operation of division. This can be done by constructing the set of rational numbers. For this, we first need the notion of an equivalence relation.

Definition 1.5. A relation $\sim$ on a set S is called an **equivalence relation** if it is:

(1) *Reflexive:* $a \sim a$ for every $a \in S$,
(2) *Symmetric:* if $a \sim b$ then $b \sim a$ for all $a, b \in S$,
(3) *Transitive:* if $a \sim b$ and $b \sim c$ then $a \sim c$ for all $a, b, c \in S$.

For each $a \in S$ the set

$$S(a) := \{b \in S : a \sim b\}$$

is called the **equivalence class** of a. From the definition it is clear that any two equivalence classes in S with respect to $\sim$ are either equal or have empty intersection. This property allows to represent S as a disjoint union of equivalence classes under $\sim$.

Exercise 1.7. *Let us write*

$$\mathbb{Z}^2 := \{(a, b) : a, b \in \mathbb{Z}\},$$

and let

$$\mathbb{Z}^2_* := \left\{(a, b) \in \mathbb{Z}^2 : b \neq 0\right\}.$$

Define a relation $\sim$ on $\mathbb{Z}^2_*$ as follows:

$$(a, b) \sim (c, d) \text{ if } a \cdot d = b \cdot c.$$

Prove that this is an equivalence relation on $\mathbb{Z}^2_*$. An equivalence class in $\mathbb{Z}^2_*$ under this equivalence relation is called a **rational number**, *and the set of all such equivalence classes is denoted by $\mathbb{Q}$. Notice that $\mathbb{Z} \subset \mathbb{Q}$, since each $a \in \mathbb{Z}$ corresponds to the equivalence class $(a, 1)$.*

We usually denote an element of $\mathbb{Q}$ as $\frac{a}{b}$, where $a, b \in \mathbb{Z}$ with $b \neq 0$, which is precisely the equivalence class of (a, b) in $\mathbb{Z}^2_*$: indeed, $\frac{a}{b} = \frac{c}{d}$ if and only if $a \cdot d = b \cdot c$. By a certain abuse of notation, we also write $\frac{a}{b}$ for any choice of a representative of its equivalence class.

Exercise 1.8. *Define addition and multiplication on $\mathbb{Q}$ as follows:*

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd},$$
$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd},$$

where in the expression like $ad + bc$ we are using addition and multiplication operations on $\mathbb{Z}$.

(1) *Prove that these operations are well defined. In other words, if $\frac{a}{b} = \frac{a'}{b'}$ and $\frac{c}{d} = \frac{c'}{d'}$, then*

$$\frac{a}{b} + \frac{c}{d} = \frac{a'}{b'} + \frac{c'}{d'} \text{ and } \frac{a}{b} \cdot \frac{c}{d} = \frac{a'}{b'} \cdot \frac{c'}{d'}.$$

(2) *Prove that $(\mathbb{Q}, +, \cdot)$ is a field.*

Our next goal is to construct the set of real numbers $\mathbb{R}$. First we introduce a metric on $\mathbb{Q}$.

Definition 1.6. Each element of $\mathbb{Z}$ can be expressed as

$$1 + \cdots + 1$$

n times or

$$(-1) + \cdots + (-1)$$

also n times, i.e., it is either n-th power of 1 or of -1. In either case, we define its **absolute value** to be n. We denote absolute value of $a \in \mathbb{Z}$ by $|a|$. Notice that for each $a \in \mathbb{Z}$, $|a| \in \mathbb{N}_0$.

Definition 1.7. A **partial order** on an abelian group $(G, +)$ is a relation $\leq$, which is

(1) *Reflexive:* $a \leq a$ for every $a \in G$,
(2) *Antisymmetric:* if $a \leq b$ and $b \leq a$, then $a = b$,
(3) *Transitive:* if $a \leq b$ and $b \leq c$, then $a \leq c$.

This order is called **translation invariant** if $a \leq b$ implies $a + c \leq b + c$ for every $a, b, c \in G$. Write 0 for the identity in $(G, +)$. An element $0 \neq a \in G$ is called **positive** if $0 \leq a$ and **negative** otherwise.

Let us introduce a relation $<$ on $\mathbb{Z}$ by defining

$$\cdots < -3 < -2 < -1 < 0 < 1 < 2 < 3 < \ldots$$

We write $a \leq b$ if either $a < b$ or $a = b$ for $a, b \in \mathbb{Z}$. We write $a > b$ or $a \geq b$ if $b < a$ or $b \leq a$, respectively.

Exercise 1.9. *Prove that this relation is a translation invariant partial order on* $\mathbb{Z}$.

We can now extend this partial order to $\mathbb{Q}$. Let $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}$. We say that $\frac{a}{b} \leq \frac{c}{d}$ if and only if $a \cdot d \leq b \cdot c$ for any choice of representatives of the corresponding equivalence classes. Since $0 \in \mathbb{Z} \subset \mathbb{Q}$, we can talk about positive and negative rational numbers, same as we do about integers. We also extend the absolute value $|\ |$ to rationals by defining

$$\left|\frac{a}{b}\right| := |a| \cdot |b|^{-1} \in \mathbb{Q}.$$

Exercise 1.10. *Prove that the following is an equivalent definition of absolute value on* $\mathbb{Q}$*: for each* $r \in \mathbb{Q}$, $|r| = r$ *if* r *is positive and* $|r| = -r$ *if* r *is negative.*

Definition 1.8. Let $\{x_i\}_{i=1}^{\infty}$ be a sequence of rational numbers, we will denote it by (x_i) for brevity of notation. It is called a **Cauchy sequence** if for every positive rational number ϵ there exists a positive integer N such that for all integers $m, n > N$,

$$|x_m - x_n| < \epsilon.$$

Define addition and multiplication of Cauchy sequences componentwise:

$$(x_i) + (y_i) = (x_i + y_i),\ (x_i) \cdot (y_i) = (x_i \cdot y_i).$$

It is clear that these operations are commutative.

Exercise 1.11. *Prove that the sum and product of Cauchy sequences are again Cauchy sequences.*

Each rational number r can be identified with a constant Cauchy sequence $(r) := \{r, r, r, \dots\}$, and the zero Cauchy sequence is $(0) := \{0, 0, 0, \dots\}$. It is easy to see that it is the additive identity among Cauchy sequences, i.e. $(x_i) + (0) = (x_i)$. We can define the additive inverse of a Cauchy sequence (x_i) to be $-(x_i) := (-x_i)$: indeed, $(x_i) + (-(x_i)) = (x_i - x_i) = (0)$.

We now introduce an equivalence relation on Cauchy sequences: two sequences (x_i) and (y_i) are said to be equivalent if for every rational $\epsilon > 0$ there exists some positive integer N such that for all $i \geq N$

$$|x_i - y_i| < \epsilon.$$

Exercise 1.12. *Prove that this is indeed an equivalence relation.*

The set of all equivalence classes of Cauchy sequences of rational numbers is called the set of **real numbers**, denoted $\mathbb{R}$. From our above discussion, it is evident that $\mathbb{Q} \subset \mathbb{R}$.

Exercise 1.13. *Let us write $\langle (x_i) \rangle$ for the equivalence class of the Cauchy sequence (x_i). Prove that the operations of addition and multiplication on the set of Cauchy sequences induce analogous operations on the equivalence classes of Cauchy sequences, i.e.*

$$\langle (x_i) \rangle + \langle (y_i) \rangle = \langle (x_i + y_i) \rangle \,, \ \langle (x_i) \rangle \cdot \langle (y_i) \rangle = \langle (x_i \cdot y_i) \rangle \,.$$

By our previous arguments, we know that these operations are commutative and that $(\mathbb{R}, +)$ is a group. Prove that $(\mathbb{R} \setminus \{\langle (0) \rangle\}, \cdot)$ is a group, this way establishing that $(\mathbb{R}, +, \cdot)$ is a field.

We can define absolute value on $\mathbb{R}$ as follows: given $\langle (x_i) \rangle \in \mathbb{R}$, let

$$|\langle (x_i) \rangle| := \langle (|x_i|) \rangle \in \mathbb{R}.$$

It is also possible to introduce ordering on $\mathbb{R}$: we say that $\langle (x_i) \rangle \leq \langle (y_i) \rangle$ (respectively, $<$, $\geq$, or $>$) if for any two representatives of these equivalence classes, there exists a positive integer N such that for all $i \geq N$, $|x_i| \leq |y_i|$ (respectively, $<$, $\geq$, or $>$).

From here on, we will assume a less formal and more common notation of rational and real numbers, as we are used to from school, while always keeping in mind the actual meaning of these objects. For instance, the notion of a *convergent sequence* of rational numbers from calculus is simply a Cauchy sequence in our language, and the *limit* of a convergent sequence is simply the real number which is this Cauchy sequence. With this in mind, we will use the standard calculus notation.

Finally, we construct complex numbers. Define

$$\mathbb{C} := \{(a, b) : a, b, \in \mathbb{R}\} \,.$$

An element (a, b) of $\mathbb{C}$ will be called a **complex number** with its first coordinate a referred to as its **real part** and its second coordinate b as its **imaginary part**. We define addition $+$ and multiplication $\cdot$ on $\mathbb{C}$ as follows:

$$(a, b) + (c, d) = (a + c, b + d), \ (a, b) \cdot (c, d) = (ac - bd, bc + ad),$$

where in the expression like $ac - bd$ we are using addition and multiplication operations on $\mathbb{R}$.

Exercise 1.14. *Prove that $(\mathbb{C}, +, \cdot)$ is a field.*

We define an important functions on $\mathbb{C}$, called **conjugation**:

$$\_ : \mathbb{C} \to \mathbb{C}$$

by $\overline{(a,b)} := (a,-b)$. For each $(a,b) \in \mathbb{C}$, $\overline{(a,b)}$ is called the **conjugate** of (a,b). We can identify $\mathbb{R}$ with the subset of all complex numbers with imaginary part equal to 0, hence $\mathbb{R} \subset \mathbb{C}$. In fact,

$$\mathbb{R} = \left\{(a,b) \in \mathbb{C} : \overline{(a,b)} = (a,b)\right\}.$$

Geometrically speaking, $\mathbb{C}$ can be thought of the real plane $\mathbb{R}^2$ and $\mathbb{R}$ is identified with the horizontal axis in it. This definition, however, is somewhat formal and does not reflect the motivation for the construction of $\mathbb{C}$.

Indeed, notice that all of our other number sets thus far have been naturally motivated: natural numbers are introduced for counting purposes; integers are defined to give a full group (and ring) structure to natural numbers, in other words to define the subtraction operation; rational numbers allow for division, hence embedding integers into a field; real numbers are defined to include limits of rational Cauchy sequences. What is the purpose of complex numbers? Notice that there are polynomial equations with real coefficients that do not have real roots, like x^2+1. Introduction of the imaginary unit $i = \sqrt{-1}$ remedies this situation. The main feature of i is that it is linearly independent with 1 over reals, i.e. it cannot be expressed as any real multiple of the number 1. Geometrically this can be interpreted by identifying 1 with the unit vector $(1,0)$ in the real plane and i with the unit vector $(0,1)$. Complex numbers are then defined as $\text{span}_{\mathbb{R}}\{1,i\}$, which coincides with our definition above. Due to this interpretation, we can also use the notation $a+bi$ for a complex number previously denoted by (a,b).

In fact, one can view introduction of each next number set in terms of solving polynomial equations. For instance, an equation of the form $x+m=0$ with $m>0$ has coefficients in natural numbers, but no solution in natural numbers: it is however solvable over $\mathbb{Z}$. Similarly, an integral equation $mx+n=0$ does not necessarily have integer solutions, but has solutions over $\mathbb{Q}$. Then $x^2-2=0$ is only solvable over $\mathbb{R}$, but not over $\mathbb{Q}$, and finally $x^2+1=0$ only has complex roots. The idea of differentiating numbers in terms of their properties as solutions or non-solutions to polynomial equations with integer coefficients will be central to these notes.

Given a commutative ring R with identity, we will write $R[x]$ for the set of all polynomials in the variable x with coefficients in R. In other words, $R[x]$ is the set of all finite formal sums

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n,$$

where $n \in \mathbb{N}_0$, $a_0, \ldots, a_n \in R$, x can take any values in R, and x^k stands for the k-th power of x under multiplication on R. We can define addition and multiplication of polynomials as follows:

$$\sum_{i=0}^{n} a_i x^i + \sum_{j=0}^{m} b_j x^j := \sum_{k=0}^{\max\{n,m\}} (a_k + b_k) x^k,$$

where a_k or b_k for $k > n$ or $k > m$, respectively, is taken to be 0, and

$$\left(\sum_{i=0}^{n} a_i x^i\right) \cdot \left(\sum_{j=0}^{m} b_j x^j\right) := \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j x^{i+j}.$$

Exercise 1.15. *Prove that $R[x]$ under the above operations is a commutative ring with identity.*

An element $\alpha \in R$ is called a **root** of a polynomial $p(x) \in R[x]$ if $p(\alpha) = 0$. A field K is called **algebraically closed** if every polynomial in $K[x]$ has a root in K. Here is an important fact, that we give here without proof (we give a proof in Appendix B).

Theorem 1.1 (Fundamental Theorem of Algebra)**.** *The field of complex numbers $\mathbb{C}$ is algebraically closed. As a consequence, any polynomial $p(x) \in \mathbb{C}[x]$ of degree n has precisely n roots in $\mathbb{C}$, counted with multiplicity.*

2. Brief remarks on exponential and logarithmic functions

Before we continue, it will also be useful to define the notion of exponential and logarithmic functions. We give only an abbreviated and restrictive definition here; for a detailed treatment of this important topic, the reader may want to consult a good book on real and complex analysis, such as [5].

We start with some more algebraic notation.

Definition 2.1. A function $f : G \to H$ between two groups $(G, *)$ and $(H, \times)$ is called a **group homomorphism** if

$$f(a * b) = f(a) \times f(b)$$

for all $a, b \in G$. A function $f : R \to S$ between two rings $(R, +, \cdot)$ and $(S, *, \times)$ is called a **ring homomorphism** if

$$f(a + b) = f(a) * f(b), \;\; f(a \cdot b) = f(a) \times f(b)$$

for all $a, b \in R$. If R and S are both fields, we may refer to f as a **field homomorphism**. Notice, however, that it is possible for a ring

homomorphism to exist between two rings of which one is a field and one is not.

Exercise 2.1. *Let $f : G \to H$ be a homomorphism between groups $(G, *)$ and $(H, \times)$. Prove that f carries identity to identity, inverses to inverses, and powers to powers. In other words,*

$$f(e_G) = e_H, \ f(a^{-1}) = f(a)^{-1}, \ f(a^n) = f(a)^n$$

for all $a \in G$, $n \in \mathbb{N}$, where e_G, e_H are identity elements in G and H, respectively.

Exercise 2.2. *Let $f : R \to S$ be a homomorphism between rings $(R, +, \cdot)$ and $(S, *, \times)$. By the above exercise, it is clear that*

$$f(0_R) = 0_S, \ f(-a) = -f(a), \ f(na) = nf(a)$$

for all $a \in R$, $n \in \mathbb{N}$, where 0_R, 0_S are additive identity elements in R and S, respectively. Prove that in addition

$$f(1_R) = 1_S, \ f(a^{-1}) = f(a)^{-1}, \ f(a^n) = f(a)^n$$

for all $a \in R$ (provided a^{-1} exists in R), $n \in \mathbb{N}$, where 1_R, 1_S are multiplicative identity elements in R and S, respectively.

Exercise 2.3. *Use the properties of ring homomorphisms you just established to prove that if $f : \mathbb{Q} \to \mathbb{Z}$ is a ring homomorphism, then $f(a) = \pm 1$ for all $0 \neq a \in \mathbb{Q}$.*

Definition 2.2. A group or ring homomorphism $f : R \to S$ is called **injective** or **one-to-one** if

$$f(a) \neq f(b) \ \forall \ a \neq b \in R.$$

The homomorphism f is called **surjective** or **onto** if for every $b \in S$ there exists $a \in R$ such that $f(a) = b$. A homomorphism that is injective and surjective is called an **isomorphism**. An isomorphism always has an **inverse**, i.e. there exists an isomorphism $f^{-1} : S \to R$ such that $f^{-1}(f(a)) = a$ and $f(f^{-1}(b)) = b$ for all $a \in R$, $b \in S$, and this inverse is unique. If there exists an isomorphism between two groups or rings R and S, we say that they are **isomorphic**, denoted by writing $R \cong S$.

Exercise 2.4. *Prove that isomorphism is an equivalence relation on the set of all groups, on the set of all rings, and on the set of all fields.*

Exercise 2.5. *Let $f : R \to S$ be a group or ring homomorphism. Define the* **kernel** *of f to be*

$$\mathrm{Ker}(f) := \{a \in R : f(a) = e_S\},$$

where e_S stands for the additive identity in S in case of rings. Prove that $\mathrm{Ker}(f)$ *is a subgroup or subring of R, respectively, and that f is injective if and only if* $\mathrm{Ker}(f) = \{e_R\}$. *If R is a field, prove that* $\mathrm{Ker}(f) = \{0_R\}$ *or R.*

Our main goal in this section is to define the *exponential function* $f_a : \mathbb{C} \to \mathbb{C}$ given by $f_a(x) = a^x$ for each base $a \in \mathbb{C}$ and outline some of its basic properties. We will do this in multiples steps. First assume that $0 \neq a \in \mathbb{C}$ and $b \in \mathbb{N}$, then

$$a^b := a \cdot \cdots \cdot a \text{ taken } b \text{ times}, \ a^0 := 1, \ 0^b := 0, \ a^{-b} := (a^{-1})^b.$$

If $b = \frac{m}{n} \in \mathbb{Q}_{>0}$ with $\gcd(m, n) = 1$ (a fraction can always be reduced) and a is a positive real number, then a^b is defined as the unique positive real root of the polynomial

$$x^n - a^m \in \mathbb{R}[x].$$

Exercise 2.6. *Let $p(x) = x^n - a^m$ for $n, m \in \mathbb{N}$ with* $\gcd(m, n) = 1$ *and $a \in \mathbb{R}_{>0}$, as above. Prove that $p(x)$ has precisely one positive real root.*

Remark 2.1. One important instance of a fractional power of a real number is the extension of the **absolute value** function to the field of complex numbers. Given a complex number $a = a_1 + ia_2$, where $a_1, a_2 \in \mathbb{R}$, notice that

$$a\overline{a} = (a_1 + ia_2)(a_1 - ia_2) = a_1^2 + a_2^2 \in \mathbb{R}.$$

Then $|a|$ is defined as the unique positive real root of the equation $x^2 - a\overline{a} = 0$, which can be denoted by writing

$$|a| = \sqrt{a_1^2 + a_2^2}.$$

In case a is real, i.e. $a_2 = 0$, this definition coincides with our previous definition of absolute value on real numbers.

We also define $a^{-b} := (a^{-1})^b$, same as for integer exponents. Now, if $b \in \mathbb{R}$, then there exists a rational Cauchy sequence $\{c_n\}_{n=1}^{\infty}$ converging to b, and so we define

$$a^b := \lim_{n \to \infty} a^{c_n}. \tag{1}$$

Remark 2.2. Equation (1) above needs some clarification. Consider the sequence $(a_n) = \{a^{c_n}\}_{n=1}^{\infty}$. Each element a_n of this sequence is a real number, and it can be shown that this sequence is a *Cauchy sequence*

of real number, meaning that for every positive *real* number ϵ there exists a positive integer N such that for all integers $m, n > N$,

$$|a_m - a_n| < \epsilon.$$

A famous result in analysis (can be found in most standard analysis books) is that a Cauchy sequence of real numbers converges to a real number. In the language of our Section 1 above, this means that every Cauchy sequence of real numbers is equivalent to some Cauchy sequence of rational numbers. A field with this property is called **complete**, and $\mathbb{R}$ is the most common example of a complete field. Hence a^b in (1) is precisely the equivalence class of the Cauchy sequence $\{a^{c_n}\}_{n=1}^{\infty}$.

Exercise 2.7. *Let $a, b \in \mathbb{R}_{>0}$, $c, d \in \mathbb{R}$. Prove that*

$$a^{c+d} = a^c a^d, \ (ab)^c = a^c b^c, a^{cd} = (a^c)^d. \tag{2}$$

Conclude that for each $a \in \mathbb{R}_{>0}$, $a \neq 1$, the **exponential map** $x \mapsto a^x$ *is an injective group homomorphism from $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$ (viewed as additive groups) to $\mathbb{R}^+ = \mathbb{R}_{>0}$ (viewed as a multiplicative group).*

In fact, the exponential map is an isomorphism of abelian groups $(\mathbb{R}, +)$ and $(\mathbb{R}^+, \cdot)$ (we do not prove it here, but a proof can be found in many standard algebra and analysis books). The inverse of this isomorphism is called the **logarithmic function** with base a, denoted $\log_a x$.

Next, let us recall a definition of e. Let $a \in \mathbb{R}_{>0}$ and consider the exponential function with base a, $f_a(x) = a^x$ for $x \in \mathbb{R}$. Notice that the derivative of this function at x is

$$f_a'(x) = \lim_{h \to 0} \frac{a^{x+h} - a^x}{h} = a^x \lim_{h \to 0} \frac{a^h - 1}{h},$$

and so

$$f_a'(0) = \lim_{h \to 0} \frac{a^h - 1}{h}.$$

This limit depends only on a, and there exists a unique value of a for which this limit is equal to 1. This value is called e. Hence e is the unique value of the base a for which the graph of $f_a(x)$ has slope $= 1$ at $x = 0$, as well as $f_a(x) = f_a'(x)$ for all x. It is also possible to define e in terms of its well-known properties:

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n = 2.71828\ldots,$$

as well as

$$\int_1^e \frac{1}{x}\, dx = 1.$$

This number is denoted by e in honor of Leonard Euler, who was first to prove its irrationality in 1737, although the number itself was first introduced by Jacob Bernoulli in 1683.

Recall from calculus the following power series expansions:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!},\ \cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!},\ \sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}. \tag{3}$$

Exercise 2.8. *Prove that the power series in* (3) *above converge for all* $x \in \mathbb{C}$.

Notice that one can treat these convergent power series expansions as definitions of e^x, $\cos x$, and $\sin x$ for any complex number x. In case of e^x, we choose to derive a formula that is easier to use.

Exercise 2.9. *Use expansions* (3) *to prove Euler's formula, established by him in 1740:*

$$e^{ix} = \cos x + i \sin x \tag{4}$$

for all $x \in \mathbb{C}$. *Furthermore, using Euler's formula, prove that any complex number* $a + bi$ *can be written as*

$$a + bi = |a + bi| e^{i\theta} = \sqrt{a^2 + b^2}\ e^{i\theta}$$

for some $\theta \in \mathbb{R}$. *Here* $\sqrt{a^2 + b^2}$ *is called the* **modulus** *and* θ *the* **argument** *of* $a + bi$, *denoted* $\arg(a + bi)$. *It is not hard to notice that modulus and argument identify the complex number uniquely.*

Remark 2.3. Notice that the argument of a complex number is not uniquely defined: it is easy to see from Euler's formula that if θ is equal to $\arg(a)$ then so is $\theta + 2\pi n$ for any $n \in \mathbb{Z}$. This problem leads to the general logarithmic function not actually being a function in the usual meaning of this word, but a *multivalued function* instead. We avoid this complication by restricting the argument: from here on, we will assume that

$$-\pi \leq \arg(a) < \pi \ \forall\ a \in \mathbb{C}$$

whenever it matters. Placing this restriction is usually called selecting the **principal branch**.

Now let $a \in \mathbb{C}$ and $b \in \mathbb{R}$. We can define

$$a^b = \left(|a| e^{i \arg(a)}\right)^b := |a|^b e^{ib \arg(a)} = |a|^b \left(\cos(b \arg(a)) + i \sin(b \arg(a))\right),$$

by Euler's formula. In other words, a^b is the complex number with modulus $|a|^b$ and argument $b \arg(a)$. It is easy to notice that the properties (2) of Exercise 2.7 apply to this situation as well.

We can now define the exponential for any base and exponent. We want our general definition to be consistent with all previous cases, which in particular means that me must have

$$(ab)^c = a^c b^c,\ a^{b+c} = a^b a^c,\ a^{bc} = (a^b)^c$$

for all $a, b, c \in \mathbb{C}$. Let $a, b \in \mathbb{C}$. It will be convenient to write $a = |a|e^{i\arg(a)}$, $b = b_1 + b_2 i$. Then

$$a^b = a^{b_1} a^{ib_2}.$$

We already know what a^{b_1} is, so it only remains to define

$$a^{ib_2} = |a|^{ib_2} \left(e^{i\arg(a)}\right)^{ib_2} = |a|^{ib_2} e^{i^2 b_2 \arg(a)} = |a|^{ib_2} e^{-b_2 \arg(a)}.$$

Now, $|a| \in \mathbb{R}^\times$, and hence by our discussion above

$$|a| = e^{\ln|a|},$$

where $\ln := \log_e$. Then

$$|a|^{ib_2} = \left(e^{\ln|a|}\right)^{ib_2} = e^{ib_2 \ln|a|}.$$

Thus we have

(5)

$$a^b := |a|^{b_1} e^{-b_2\arg(a)} e^{i(b_1\arg(a) + b_2 \ln|a|))} = |a|^{b_1 - \frac{b_2\arg(a)}{\ln|a|}} e^{i(b_1\arg(a) + b_2\ln|a|))}$$

for any $a = |a|e^{i\arg(a)}$ and $b = b_1 + ib_2$ in $\mathbb{C}$.

Now for every $a \in \mathbb{C}$, we have the exponential function with base a, $f_a : \mathbb{C} \to \mathbb{C}$ given by $f_a(x) = a^x$. It is a homomorphism of groups $(\mathbb{C}, +)$ and $(\mathbb{C}^\times, \cdot)$, which is surjective whenever $a \neq 0, \pm 1$, however is not injective: its kernel is equal to $\{2n\pi i : n \in \mathbb{Z}\}$ as can be seen from Euler's formula. Restricting the argument of x to the interval $[-\pi, \pi)$, as discussed above, we can define the inverse of f_a, the **logarithmic function**, denoted by $\log_a$: since f_a is surjective, for each $y \in \mathbb{C}$ there exists the unique $x \in \mathbb{C}$ with $\arg(x) \in [-\pi, \pi)$ such that $a^x = y$; define $\log_a(y)$ to be this x.

Remark 2.4. Unfortunately, our restriction of argument causes the logarithmic function not to be continuous. This difficulty can be overcome by introduction of a *Riemann surface* for the logarithm function, which is usually done in complex analysis (see, for instance, [5]). We choose not to do this here, since it would take us in a rather different direction from where we plan to go.

Exercise 2.10. *Derive a power series expansion for the exponential function $f_a(x) = a^x$ with base $a \in \mathbb{C}$, $a \neq 0, \pm 1$, which converges for all $x \in \mathbb{C}$.*

3. Basic properties of algebraic and transcendental numbers

We will start out with basic definitions and properties of algebraic and transcendental numbers.

Definition 3.1. A complex number α is called **algebraic** if there exists a nonzero polynomial $p(x)$ with integer coefficients such that $p(\alpha) = 0$. If α is not algebraic, it is called **transcendental**.

Remark 3.1. More generally, we will say that α is **algebraic over** K for some field K if there exists a polynomial $p(x) \in K[x]$ such that $p(\alpha) = 0$.

In other words, transcendental numbers are complex numbers that do not satisfy any polynomial equation with integer coefficients. We will write $\mathbb{A}$ for the set of all algebraic numbers and $\mathbb{T} := \mathbb{C} \setminus \mathbb{A}$ for the set of all transcendental numbers.

Examples of algebraic numbers are easy to construct. In fact, it is easily seen that every rational number $\frac{m}{n}$ is algebraic: it is the root of polynomial $p(x) = nx - m$. More generally, any number of the form $\left(\frac{m}{n}\right)^{1/k}$, where m, n are integers, $n \neq 0$, and k a positive integer is also algebraic: it is a root of the polynomial $p(x) = nx^k - m$. Notice that this example includes such instances as $\sqrt{2}$, $i = \sqrt{-1}$, and many others. These examples and the ease with which they can be constructed may give an impression that most complex numbers are algebraic. In fact, this is not true. Our first goal is to make this idea rigorous.

First let us introduce some additional notation. Recall that we write $\mathbb{Z}[x]$ for the ring of all polynomials with integer coefficients. We think of constants as polynomials of degree 0, and hence $\mathbb{Z} \subset \mathbb{Z}[x]$.
The **degree** of an algebraic number α is defined as

$$\deg(\alpha) := \min\{\deg(f(x)) : f(x) \in \mathbb{Z}[x],\ f(\alpha) = 0\}.$$

Exercise 3.1. *Let $\alpha \in \mathbb{A}$ and let $f(x), g(x) \in \mathbb{Z}[x]$ be two polynomials of degree $\deg(\alpha)$ such that $f(\alpha) = g(\alpha) = 0$. Prove that $f(x) = cg(x)$ for some constant c.*

Let $d = \deg(\alpha)$ and let $f(x) = \sum_{m=0}^{d} a_m x^m \in \mathbb{Z}[x]$ be a polynomial of degree d such that $f(\alpha) = 0$, $\gcd(a_0, \ldots, a_d) = 1$, and $a_d > 0$. By the exercise above, this polynomial is unique for each $\alpha \in \mathbb{A}$: it is called the **minimal polynomial** of α, denoted by $m_\alpha(x)$. A polynomial $p(x) \in \mathbb{Z}[x]$ is called **irreducible** if whenever $p(x) = f(x)g(x)$ for some $f(x), g(x) \in \mathbb{Z}[x]$ then either $f(x)$ or $g(x)$ is equal to ± 1.

Exercise 3.2. *Prove that $m_\alpha(x)$ is irreducible for each $\alpha \in \mathbb{A}$. Furthermore, prove that if $p(x) \in \mathbb{Z}[x]$ is such that $p(\alpha) = 0$, then $m_\alpha(x) \mid p(x)$, i.e. there exists some $g(x) \in \mathbb{Z}[x]$ such that $p(x) = m_\alpha(x)g(x)$.*

Remark 3.2. Minimal polynomial of a complex number can be defined over any subfield of $\mathbb{C}$, not only over $\mathbb{Q}$ or $\mathbb{Z}$. Specifically, if $\alpha \in \mathbb{C}$ and K is a subfield of $\mathbb{C}$, we define the **minimal polynomial of α over K** to be a polynomial $p(x)$ with coefficients in K and leading coefficient 1 of minimal degree such that $p(\alpha) = 0$, if such a polynomial exists. It is then not hard to see (the proofs are left to the reader) that $p(x)$ is irreducible over K and is unique. Furthermore, $p(x)$ divides (over K) any other polynomial with coefficients in K which has α as its root, analogously to the properties over $\mathbb{Z}$ described in Exercise 3.2 above.

A polynomial $p(x) \in \mathbb{Z}[x]$ is called **separable** if all of its roots are distinct. We will use the following fact without proof.

Fact 3.1. *Irreducible polynomials in $\mathbb{Z}[x]$ are separable.*

Definition 3.2. A set S is called **countable** if there exists a bijective (i.e., one-to-one and onto) map $f : \mathbb{N} \to S$.

Lemma 3.2. *Let $S_1, S_2, \dots$ be a collection of finite sets. Then their union*

$$S = \bigcup_{n=1}^{\infty} S_n$$

is countable.

Proof. For each $n \geq 1$, let a_n be the cardinality of S_n, and write

$$S_n = \{x_{n1}, \dots, x_{na_n}\}.$$

Then we can write

$$S = \{x_{11}, \dots, x_{1a_1}, x_{21}, \dots, x_{2a_2}, \dots\}.$$

Let y_m be the m-th element of S with respect to the above ordering, i.e. $y_m = x_{nj}$ for some n and j such that

$$a_1 + \cdots + a_{n-1} + j = m.$$

Then define $f : \mathbb{N} \to S$ by $f(m) = y_m$. This map is clearly a bijection, and hence S is countable. □

Lemma 3.3. *The set $\mathbb{N} \times \mathbb{N}$ is countable.*

Proof. Notice that

$$\begin{aligned}\mathbb{N}\times\mathbb{N} &= \{(m,n) : m,n\in\mathbb{N}\}\\ &= \{(m,n) : m,n\in\mathbb{N},\ m\le n\}\cup\{(m,n) : m,n\in\mathbb{N},\ m>n\}\\ &= \left(\bigcup_{n=1}^{\infty}\{(m,n) : m\le n\}\right)\cup\left(\bigcup_{m=1}^{\infty}\{(m,n) : n<m\}\right),\end{aligned}$$

which is a (countable) union of finite sets, and hence it is countable by Lemma 3.2 above. □

Lemma 3.4. *A countable union of countable sets is countable.*

Proof. Let $S_1, S_2, \dots$ be countable sets, say

$$S_n = \{x_{n1}, x_{n2}, \dots\},$$

and let

$$S = \bigcup_{n=1}^{\infty} S_n.$$

Then notice that there is a bijection $f : \mathbb{N}\times\mathbb{N} \to S$, given by $f(n,m) = x_{nm}$. By Lemma 3.3, $\mathbb{N}\times\mathbb{N}$ is countable, i.e. there exists a bijection $g : \mathbb{N} \to \mathbb{N}\times\mathbb{N}$. Since a composition of two bijections $f\circ g : \mathbb{N} \to S$ is again a bijection, we conclude that S is countable. □

Exercise 3.3. *Prove that any infinite subset of a countable set is countable. Use this fact to conclude that a superset of an uncountable set is uncountable.*

Lemma 3.5. *Let $m \ge 1$. The set*

$$\mathbb{Z}^m := \{\boldsymbol{a} = (a_1, \dots, a_m) : a_1, \dots, a_m \in \mathbb{Z}\}$$

is countable.

Proof. We argue by induction on m. First suppose that $m = 1$, then the set

$$\mathbb{Z} = \mathbb{N}\cup -\mathbb{N}\cup\{0\},$$

where $-\mathbb{N} = \{-x : x\in\mathbb{N}\}$. This is a union of two countable sets and one finite set, hence it is countable. Now suppose that the statement of the lemma is true for $m = d-1$. We prove it for $m = d$. Notice that

$$\mathbb{Z}^d = \left(\bigcup_{a\in\mathbb{N}_0}\{(\boldsymbol{x},a) : \boldsymbol{x}\in\mathbb{Z}^{d-1}\}\right)\cup\left(\bigcup_{a\in\mathbb{N}}\{(\boldsymbol{x},-a) : \boldsymbol{x}\in\mathbb{Z}^{d-1}\}\right),$$

Each set like $\{(\boldsymbol{x},a) : \boldsymbol{x}\in\mathbb{Z}^{d-1}\}$ or $\{(\boldsymbol{x},-a) : \boldsymbol{x}\in\mathbb{Z}^{d-1}\}$ for $a\in\mathbb{N}$ is in bijective correspondence with $\mathbb{Z}^{d-1}$, and hence is countable by induction hypothesis. Therefore $\mathbb{Z}^d$ is a countable union of countable sets, and hence is countable by Lemma 3.4. □

Remark 3.3. One can use Lemma 3.5 along with Exercise 1.7 to deduce that $\mathbb{Q}$ is a countable set. Indeed, Exercise 1.7 constructs rational numbers as the set of equivalence classes of the subset $\mathbb{Z}^2_*$ of $\mathbb{Z}^2$ under the specified equivalence relation. Identifying these equivalence classes with some choice of their representatives, we can view $\mathbb{Q}$ as a subset of $\mathbb{Z}^2$. Lemma 3.5 implies that $\mathbb{Z}^2$ is countable, and then Exercise 3.3 guarantees that $\mathbb{Q}$ is countable.

We will now prove a much stronger fact, namely countability of the set of all algebraic numbers, from which countability of $\mathbb{Q}$ follows yet again by Exercise 3.3.

Theorem 3.6. *The set $\mathbb{A}$ of algebraic numbers is countable.*

Proof. Notice that each $\alpha \in \mathbb{A}$ is a root of some polynomial in $\mathbb{Z}[x]$. Furthermore, each polynomial $p(x) \in \mathbb{Z}[x]$ has finitely many roots. For each $p(x) \in \mathbb{Z}[x]$, let R_p be the set of all roots of $p(x)$. Then

$$\mathbb{A} = \bigcup_{p(x)\in\mathbb{Z}[x]} R_p.$$

This union is not disjoint, i.e. roots may be repeated. Hence, if we just think of this union as a list of elements with repetition, then $\mathbb{A}$ is formally a subset of $\bigcup_{p(x)\in\mathbb{Z}[x]} R_p$. Now notice that each polynomial

$$p(x) = \sum_{n=0}^{d} a_n x^n \in \mathbb{Z}[x]$$

can be identified with its vector of coefficients $(a_0, \dots, a_d) \in \mathbb{Z}^{d+1}$, where $d = \deg(p(x))$. This defines a bijection between $\mathbb{Z}[x]$ and the set $\bigcup_{d\in\mathbb{N}_0} \mathbb{Z}^{d+1}$, which is a countable union of countable sets, hence is countable. Therefore $\bigcup_{p(x)\in\mathbb{Z}[x]} R_p$ is a countable union of finite sets, hence is countable, and so its subset $\mathbb{A}$ is also countable. □

Remark 3.4. In fact, we could rephrase the proof Theorem 3.6 in terms of just irreducible polynomials. In other words, there is a bijection between $\mathbb{A}$ and the *disjoint* union of sets of roots of all irreducible polynomials in $\mathbb{Z}[x]$. Since the set of irreducible polynomials is an infinite subset of the countable set $\mathbb{Z}[x]$, it is itself countable, hence we are done.

In contrast, let us consider the set of all real numbers.

Exercise 3.4. *Let $a < b$ be real numbers and let $I = [a, b]$ be a closed interval. Prove that I contains infinitely many real numbers.*

Theorem 3.7. *The set $\mathbb{R}$ of all real numbers is uncountable.*

Proof. Assume that $\mathbb{R}$ is countable. Then there exists some bijection $f : \mathbb{N} \to \mathbb{R}$. Let us write $x_n := f(n)$ for each $n \in \mathbb{N}$, so the image of f is the sequence $(x_n)_{n\in\mathbb{N}}$ of *distinct* real numbers, which is supposed to be equal to all of $\mathbb{R}$. We will reach a contradiction by showing that every sequence $(x_n)_{n\in\mathbb{N}}$ of distinct real numbers misses at least one $x \in \mathbb{R}$.

Indeed, let $(x_n)_{n\in\mathbb{N}}$ be such a sequence. We define a nested family of intervals as follows. Let $a_1 = \min\{x_1, x_2\}$ and $b_1 = \max\{x_1, x_2\}$. Since the elements of our sequence are all distinct, $a_1 < b_1$, and hence $I_1 := [a_1, b_1]$ is an interval, not a singleton. If I_1 contains only finitely many x_n's, then pick some $x \in I_1$ which is not one of these numbers (by Exercise 3.4 above, such x must exist), and we are done. Then assume I_1 contains infinitely many x_n's. Let y and z be the first two such elements, with respect to index, in the interior of I_1 and let $a_2 = \min\{y, z\}$, $b_2 = \max\{y, z\}$ so $a_2 < b_2$ and $I_2 := [a_2, b_2]$ is again an interval with non-empty interior such that $I_2 \subsetneq I_1$. Continue in the same manner to obtain a nested sequence of intervals:

$$\cdots \subsetneq I_n \subsetneq I_{n-1} \subsetneq \cdots \subsetneq I_2 \subsetneq I_1,$$

where each $I_n = [a_n, b_n]$ with $a_n < b_n$. Then notice that

$$a_1 < a_2 < \cdots < a_{n-1} < a_n < \cdots < b_n < b_{n-1} < \cdots < b_2 < b_1.$$

Therefore $(a_n)_{n\in\mathbb{N}}$ (respectively, $(b_n)_{n\in\mathbb{N}}$) is a monotone increasing (respectively, decreasing) sequence, which is bounded from above (respectively, below). By the Monotone Convergence Theorem (recall from Calculus), these sequences have limits, let us write

$$A := \lim_{n\to\infty} a_n, \; B := \lim_{n\to\infty} b_n.$$

It is clear that $A \leq B$, so the closed interval $I = [A, B]$ is not empty. Let $h \in I$, then $h \neq a_n, b_n$ for any $n \in \mathbb{N}$. In fact, we will show that $h \neq x_n$ for any $n \in \mathbb{N}$.

Suppose that $h = x_k$ for some $k \in \mathbb{N}$, so there are finitely many points in the sequence $(x_n)_{n\in\mathbb{N}}$ before h occurs, and hence only finitely many a_n's preceding h. Let a_d be the last element in the sequence $(a_n)_{n\in\mathbb{N}}$ preceding h. Since h cannot be equal to a_d, $a_d < h$, i.e. h is in the interior of I_d. Since it is contained in the limiting interval I, it must be contained in $I_{d+1} = [a_{d+1}, b_{d+1}]$ by our construction of the intervals. But this means that $a_d < a_{d+1} < h$, which contradicts our choice of a_d.

This shows that h is not an element of the sequence $(x_n)_{n\in\mathbb{N}}$, and hence at least one real number is not in this sequence. This means that $\mathbb{R}$ cannot be countable. $\square$

Remark 3.5. The fact of uncountability of reals was first established by Georg Cantor in 1874. In fact, Cantor presented at least three different proofs of this fact, including his famous diagonal argument (1891). Our proof of Theorem 3.7 above follows Cantor's first argument (1874).

Since $\mathbb{R} \subset \mathbb{C}$, we conclude that $\mathbb{C}$ is also uncountable, by Exercise 3.3. Now recall that $\mathbb{C} = \mathbb{A} \cup \mathbb{T}$, and $\mathbb{A}$ is countable. This means that $\mathbb{T}$, the set of transcendental numbers, is uncountable. Loosely speaking this means, that most complex numbers are in fact transcendental. Ironically, while constructing algebraic numbers is quite straightforward, as seen above, it is not at all easy to construct a transcendental number. Indeed, suppose we take a complex number α. To prove that it is algebraic, we can find its minimal polynomial $m_\alpha(x) \in \mathbb{Z}[x]$. Although this may be somewhat laborious, there are standard techniques in algebraic number theory that allow for such a construction. On the other hand, to prove that α is transcendental we would need to establish that α is not a root of *any* polynomial in $\mathbb{Z}[x]$. This kind of fact clearly requires some sort of indirect argument, which is the reason why it took mathematicians until mid-19th century to construct the first transcendental number. This construction, by Joseph Liouville, used the recently developed tools in the area of Diophantine Approximation. It is our next goal to develop the necessary tools and to present Liouville's construction.

4. Introduction to Diophantine Approximation: Dirichlet, Liouville, Roth

In Section 1 we constructed real numbers from rationals. Then with Theorem 3.7 we established that the set of real numbers $\mathbb{R}$ is uncountable. An implication of Theorem 3.6 together with Exercise 3.3 is that the set of rational numbers $\mathbb{Q}$ is a countable subset of $\mathbb{R}$. In other words, in a certain sense rational numbers appear to be sparse among the reals. On the other hand, it is always possible to find a rational number as close as we want to a given real number.

Theorem 4.1. *The set of rational numbers $\mathbb{Q}$ is* **dense** *inside of the set of real number $\mathbb{R}$, i.e. if $x < y \in \mathbb{R}$, then there exists $z \in \mathbb{Q}$ such that*

$$x < z < y.$$

Proof. Since $y - x > 0$, there must exist $n \in \mathbb{Z}$ such that

$$n(y - x) = ny - nx > 1,$$

so $nx + 1 < ny$. Let $m \in \mathbb{Z}$ be such that

$$m \leq nx + 1 < m + 1,$$

then we have

$$nx < m \leq nx + 1 < ny,$$

and hence

$$x < \frac{m}{n} < y.$$

Let $z = \frac{m}{n} \in \mathbb{Q}$, and this finishes the proof. □

Theorem 4.1 implies that, given a real number, we can approximate it arbitrarily well by rational numbers. For many purposes we may want to control how "complicated" the rational numbers we use for such approximations are, i.e. we may want to bound the size of their denominators. This is the starting point of the theory of Diophantine Approximation. The first result in this direction dates back to Dirichlet, and is proved with the use of Dirichlet's box principle; in fact, this is most likely the theorem to which this principle owes its name. For the rest of this section we follow [7].

Theorem 4.2 (Dirichlet, (1842))**.** *Let $\alpha \in \mathbb{R}$, and let $Q \in \mathbb{Z}_{>0}$. There exist relatively prime integers p, q with $1 \leq q \leq Q$ such that*

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q(Q+1)}. \tag{6}$$

Moreover, if α is irrational, then there are infinitely many rational numbers $\frac{p}{q}$ such that

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^2}. \tag{7}$$

Proof. If α is a rational number with denominator $\leq Q$, there is nothing to prove. Hence we will assume that either α is irrational, or it is rational with denominator $> Q$. Notice that

$$[0,1) = \bigcup_{i=1}^{Q+1} \left[\frac{i-1}{Q+1}, \frac{i}{Q+1}\right).$$

Consider the numbers $\{l\alpha\}$, $1 \leq l \leq Q+1$, where $\{\ \}$ denotes the fractional part function, i.e. $\{x\} = x - [x]$. These numbers lie in the interval $[0,1)$ and are distinct. Indeed, suppose that $\{l\alpha\} = \{m\alpha\}$ for some $1 \leq l < m \leq Q+1$, then $m\alpha - l\alpha$ is an integer, say

$$m\alpha - l\alpha = \alpha(m-l) = k \in \mathbb{Z},$$

and so $\alpha = k/(m-l)$, where $m-l \leq (Q+1)-1 = Q$, which contradicts our assumption.

Case 1. Suppose that each subinterval $\left[\frac{i-1}{Q+1}, \frac{i}{Q+1}\right)$ contains one of the numbers $\{l\alpha\}$, $1 \leq l \leq Q+1$. In particular, subintervals $\left[0, \frac{1}{Q+1}\right)$ and $\left[\frac{Q}{Q+1}, 1\right)$ contain such points, so at least one of them must contain some $\{l\alpha\}$ with $1 \leq l \leq Q$. Therefore, either

$$|l\alpha - [l\alpha]| \leq \frac{1}{Q+1}, \tag{8}$$

or

$$|l\alpha - [l\alpha] - 1| \leq \frac{1}{Q+1}. \tag{9}$$

This means that there exists an integer $1 \leq l \leq Q$ and an integer m equal to either $[l\alpha]$ or $[l\alpha]-1$, depending on whether (8) or (9) holds, such that

$$|l\alpha - m| \leq \frac{1}{Q+1}.$$

Let $d = \gcd(l, m)$, and let $p = \frac{m}{d}$ and $q = \frac{l}{d}$, then

$$|qd\alpha - pd| \leq \frac{1}{Q+1},$$

meaning that

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{qd(Q+1)} \leq \frac{1}{q(Q+1)},$$

proving (6) in this case.

Case 2. Now assume that one of the subintervals $\left[\frac{i-1}{Q+1}, \frac{i}{Q+1}\right)$ for some $1 \leq i \leq Q+1$ does not contain any of the numbers $\{l\alpha\}$, $1 \leq l \leq Q+1$. Since there are $Q+1$ such numbers and $Q+1$ subintervals, one of the subintervals must contain two such numbers, say $\left[\frac{j-1}{Q+1}, \frac{j}{Q+1}\right)$ for some $1 \leq j \leq Q+1$ contains $\{l\alpha\}$ and $\{m\alpha\}$ for some $1 \leq l < m \leq Q+1$. Therefore

$$|(m\alpha - [m\alpha]) - (l\alpha - [l\alpha])| = |(m-l)\alpha - ([m\alpha] - [l\alpha])| \leq \frac{1}{Q+1}.$$

Once again, let $d = \gcd((m-l), ([m\alpha] - [l\alpha]))$, and let $p = \frac{[m\alpha]-[l\alpha]}{d}$ and $q = \frac{m-l}{d}$, and so in the same way as above we obtain (6).

Exercise 4.1. *Prove that if* $\alpha = \frac{a}{Q+1}$ *for some integer* a *with*

$$\gcd(a, Q+1) = 1,$$

then there is equality in (6).

We can now derive (7) from (6): since $q \leq Q$,

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q(Q+1)} < \frac{1}{q^2}. \tag{10}$$

Now suppose that there are only finitely many rationals that satisfy (7), call them

$$\frac{p_1}{q_1}, \ldots, \frac{p_k}{q_k}.$$

Let

$$\delta = \min_{1 \leq i \leq k} \left|\alpha - \frac{p_i}{q_i}\right|,$$

then $\delta > 0$, since α is irrational. Let $Q \in \mathbb{Z}_{>0}$ be such that

$$\frac{1}{Q} < \delta.$$

By (10), there must exist $\frac{p}{q}$ with $1 \leq q \leq Q$ such that

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q(Q+1)} < \delta,$$

hence $\frac{p}{q} \notin \left\{\frac{p_1}{q_1}, \ldots, \frac{p_k}{q_k}\right\}$, which is a contradiciton. Thus there must be infinitely many such rationals. □

Remark 4.1. Notice that the argument that derives (7) from (6) is very similar to Euclid's proof of the infinitude of primes.

Hurwitz (1891) improved Dirichlet's bound (7) slightly by showing that for any irrational $\alpha \in \mathbb{R}$ there exist infinitely many distinct rational numbers $\frac{p}{q}$ such that

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{\sqrt{5}\ q^2}. \tag{11}$$

We will now show that in a certain sense (11) is best possible.

Lemma 4.3. *Let* $\alpha \in \mathbb{R}$ *be a quadratic irrational satisfying* $f(\alpha) = 0$*, where*

$$f(x) = ax^2 + bx + c$$

with $a, b, c \in \mathbb{Z}$ and $a > 0$. Write $D = b^2 - 4ac$ for the discriminant of f. Then for any real number $A > \sqrt{D}$, there are only finitely many rationals $\frac{p}{q}$ such that

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{Aq^2}. \tag{12}$$

Proof. We know that α is one of the roots of $f(x)$, then let β be the other one, i.e.

$$f(x) = a(x-\alpha)(x-\beta) = ax^2 - a(\alpha+\beta)x + a\alpha\beta,$$

meaning that $b = a(\alpha+\beta)$ and $c = a\alpha\beta$. Therefore

$$D = b^2 - 4ac = a^2(\alpha-\beta)^2.$$

Now suppose that for some $\frac{p}{q} \in \mathbb{Q}$ (12) holds. Notice that since $f(x)$ is a quadratic polynomial with irrational roots, then

$$0 \neq \left|f\left(\frac{p}{q}\right)\right| = \frac{|ap^2 + bpq + cq^2|}{q^2} \geq \frac{1}{q^2},$$

since $0 \neq ap^2 + bpq + cq^2 \in \mathbb{Z}$, hence $|ap^2 + bpq + cq^2| \geq 1$. Therefore

$$\begin{aligned}
\frac{1}{q^2} &\leq \left|f\left(\frac{p}{q}\right)\right| = a\left|\alpha - \frac{p}{q}\right|\left|\beta - \frac{p}{q}\right| \\
&< \frac{a}{Aq^2}\left|\beta - \frac{p}{q}\right| = \frac{a}{Aq^2}\left|\left(\alpha - \frac{p}{q}\right) + (\beta - \alpha)\right| \\
&\leq \frac{a}{Aq^2}\left|\alpha - \frac{p}{q}\right| + \frac{a}{Aq^2}|\beta - \alpha| < \frac{a}{A^2q^4} + \frac{\sqrt{D}}{Aq^2},
\end{aligned}$$

and subtracting $\frac{\sqrt{D}}{Aq^2}$ from both sides of the above inequality implies

$$\frac{1}{q^2}\left(1 - \frac{\sqrt{D}}{A}\right) < \frac{a}{A^2q^4}.$$

The left hand side of this inequality is not 0 since $A > \sqrt{D}$, and hence

$$q^2 < \frac{a}{A(A-\sqrt{D})}.$$

This implies that there are only finitely many possibilities for the denominator q, but for each such q there can be only finitely many p so that (12) holds. This completes the proof. □

Remark 4.2. Let $\alpha = \frac{1+\sqrt{5}}{2}$, then the corresponding polynomial

$$f(x) = x^2 - x - 1,$$

and its discriminant is $D = 5$. By Lemma 4.3, if $A > \sqrt{5}$ then there are only finitely many $\frac{p}{q} \in \mathbb{Q}$ such that

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

which proves that Hurwitz's bound (11) is best possible.

More generally, for every quadratic irrational α there exists a constant $C(\alpha) > 0$ such that for any $\frac{p}{q} \in \mathbb{Q}$

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{C(\alpha)}{q^2}. \tag{13}$$

In other words, quadratic irrationals are *badly approximable*.

Definition 4.1. An irrational number α is called **badly approximable** if there exists a positive real constant $C(\alpha)$ such that (13) holds for any $\frac{p}{q} \in \mathbb{Q}$.

As can be expected after the above discussion, algebraic numbers although are not necessarilly badly approximable, are certainly "worth" approximable than transcendental. This principle was first observed by Liouville in 1844.

Theorem 4.4 (Liouville)**.** *Let $\alpha \in \mathbb{R}$ be an algebraic number of degree $d = \deg(f) \geq 2$, where $f(x) \in \mathbb{Z}[x]$ is the minimal polynomial of α over $\mathbb{Q}$. Then there exists a positive real constant $C(\alpha)$ such that for any $\frac{p}{q} \in \mathbb{Q}$*

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{C(\alpha)}{q^d}. \tag{14}$$

Proof. Let

$$f(x) = \sum_{i=0}^{d} a_i x^i \in \mathbb{Z}[x].$$

Then, since $d \geq 2$ means that α is irrational, for each $\frac{p}{q} \in \mathbb{Q}$ we have

$$0 \neq q^d f\left(\frac{p}{q}\right) = \sum_{i=0}^{d} a_i p^i q^{d-i} \in \mathbb{Z}.$$

We can assume of course that $\left|\alpha - \frac{p}{q}\right| \leq 1$. Then, since $f(\alpha) = 0$,

$$\begin{aligned} 1 &\leq q^d \left|f\left(\frac{p}{q}\right)\right| = q^d \left|f(\alpha) - f\left(\frac{p}{q}\right)\right| = q^d \left|\int_{p/q}^{\alpha} f'(u)\, du\right| \\ &\leq q^d \left|\alpha - \frac{p}{q}\right| \max\{f'(u) : |\alpha - u| \leq 1\}. \end{aligned}$$

Then pick $C(\alpha) = (\max\{f'(u) : |\alpha - u| \leq 1\})^{-1}$, and the theorem follows. □

Liouville used his theorem to construct the first known example of a transcendental number.

Corollary 4.5 (Liouville)**.** *The number*

$$\alpha = \sum_{n=1}^{\infty} \frac{1}{a^{n!}}$$

is transcendental for any integer $a \geq 2$.

Proof. Let $a > 1$. For every $k \in \mathbb{Z}_{>0}$, let

$$p_k = a^{k!} \sum_{n=1}^{k} \frac{1}{a^{n!}},\ q_k = a^{k!} \in \mathbb{Z}.$$

Then

$$\left|\alpha - \frac{p_k}{q_k}\right| = \sum_{n=k+1}^{\infty} \frac{1}{a^{n!}} = \frac{1}{a^{(k+1)!}} \sum_{n=k+1}^{\infty} \frac{a^{(k+1)!}}{a^{n!}} < \frac{1}{a^{(k+1)!}} \sum_{n=0}^{\infty} \frac{1}{a^n}.$$

Clearly $\sum_{n=0}^{\infty} \frac{1}{a^n}$ is a convergent series, so let

$$\mathcal{C} = \sum_{n=0}^{\infty} \frac{1}{a^n},$$

and then we have

$$\left|\alpha - \frac{p_k}{q_k}\right| < \frac{\mathcal{C}}{a^{(k+1)!}} = \frac{\mathcal{C}}{q_k^{(k+1)}} < \frac{\mathcal{C}}{q_k^k}. \tag{15}$$

Suppose that α is rational, say $\alpha = c/d$ for some $c, d, \in \mathbb{Z}$. Then (15) implies that

$$|cq_k - dp_k| < \frac{\mathcal{C}d}{q_k^{k-1}}$$

for infinitely many p_k/q_k as above. The expression $\frac{\mathcal{C}d}{q_k^{k-1}}$ is < 1 for all large enough q_k. On the other hand, $|cq_k - dp_k|$ is a nonnegative integer,

which can be 0 for at most one p_k/q_k; hence $|cq_k - dp_k| \geq 1$ for infinitely many p_k/q_k. This is a contradiction, and so α cannot be rational.

Now suppose that α is algebraic of degree d. Then, by Theorem 4.4, there exists a constant $C(\alpha)$ such that

$$\left|\alpha - \frac{p_k}{q_k}\right| \geq \frac{C(\alpha)}{q_k^d},$$

for every $k \in \mathbb{Z}_{>0}$. However, if we take k large enough so that

$$\frac{\mathcal{C}}{q_k^k} < \frac{C(\alpha)}{q_k^d},$$

then (15) implies a contradiction; more specifically, we just need to take k large enough so that

$$k!(k-d) > \frac{\ln \mathcal{C} - \ln C(\alpha)}{\ln a}.$$

This completes the proof. □

Remark 4.3. Numbers that can be proved to be transcendental using Liouville's theorem are called **Liouville numbers**; they form a rather small set. In particular, e and π (which are transcendental, as we prove later in these notes) are not Liouville numbers, and neither are most transcendental numbers.

Theorem 4.4 implies that if α is an algebraic number of degree $d \geq 2$ and $\mu > d$, then there are only finitely many $\frac{p}{q} \in \mathbb{Q}$ with $\gcd(p,q) = 1$ such that

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^\mu}. \tag{16}$$

Indeed, suppose there were infinitely many rational numbers for which (16) holds. Let $C(\alpha)$ be the constant guranteed by Theorem 4.4. Let Q be an integer so that $C(\alpha) > \frac{1}{Q^{\mu-d}}$. Clearly there can be only finitely many $\frac{p}{q}$ with $\gcd(p,q) = 1$ for which (16) holds with $q \leq Q$, hence there must be infinitely many such rationals with $q > Q$. Suppose $\frac{p}{q}$ is one of them, then

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^\mu} < \frac{1}{Q^{\mu-d}q^d} < \frac{C(\alpha)}{q^d},$$

which contradicts (14). This proves finiteness of the number of solutions for (16).

For an algebraic number α of degree $d \geq 2$, what is the smallest possible μ for which (16) will have only finitely many solutions? Combining the discussion above with Dirichlet's theorem (Theorem 4.2),

we see that

$$2 \leq \mu \leq d + \delta,$$

for any $\delta > 0$. In 1908 Thue proved that $\mu \leq \frac{d+2}{2} + \delta$; in 1921 Siegel proved that $\mu \leq 2\sqrt{d} + \delta$. Dyson (1947) and Gelfond (1952) proved that $\mu \leq \sqrt{2d} + \delta$. The major breakthrough came with the famous theorem of Roth (1955) [4], for which he received a Fields medal in 1958.

Theorem 4.6 (Roth)**.** *Let α be an algebraic number. For any $\delta > 0$, there are only finitely many rationals $\frac{p}{q}$ with $\gcd(p, q) = 1$ such that*

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

Remark 4.4. Dirichlet's theorem shows that Roth's theorem is best possible, i.e. the exponent on q in the upper bound cannot be improved. Notice also that in case α has degree 2, Lemma 4.3 gives a better result. An outline of the proof of Roth's theorem can be found in [7]; complete versions of the proof can be found in [6], [2], and [4].

In other words, Roth's theorem implies that if α is algebraic, then the number of sufficiently good rational approximations to α is finite, so perhaps one can actually count them, although we are not quite ready to do this. If α is real, but not necessarily algebraic, there may be infinitely many good rational approximations to α, however we will now show that there are only finitely many of them within a finite interval. To prove a result of this sort, we will first need a certain "gap principle".

Definition 4.2. A set $S \subseteq \mathbb{R}$ is called a C**-set** for a real number $C > 1$ if for any two numbers m, n in S, $m \leq Cn$ and $n \leq Cm$.

Notice for instance that a C-set consisting of integers must be finite, although unless we know at least one of its elements, we cannot say anything about its cardinality.

Definition 4.3. A set $S \subseteq \mathbb{R}$ is called a γ**-set** for a real number $\gamma > 1$ if whenever $m, n \in S$ and $m < n$, then $\gamma m \leq n$.

Notice that a γ-set can be infinite, but it has a gap principle: its elements cannot be too close together, i.e., there is always a gap between them. A set $S \subseteq \mathbb{Z}_{>0}$ that is both a C-set and a γ-set will be called a (C, γ)**-set**. Notice that a (C, γ)-set is always finite. It is possible to estimate the cardinality of a (C, γ)-set without knowing anything about its elements.

Lemma 4.7. *Let $C > 1$ and $\gamma > 1$, and suppose that $S \subseteq \mathbb{R}_{>0}$ is a (C, γ)-set. Then*

$$|S| \leq 1 + \frac{\ln C}{\ln \gamma}. \tag{17}$$

Proof. Clearly S is a finite set, so assume

$$S = \{m_0 < m_1 < \cdots < m_k\},$$

i.e. $|S| = k + 1$. Then for each $0 \leq i \leq k$,

$$m_i \geq m_0 \gamma^i,$$

and

$$C m_0 \geq m_k \geq m_0 \gamma^k.$$

Hence

$$k \leq \frac{\ln C}{\ln \gamma},$$

and (17) follows. □

Definition 4.4. Given $C > 1$, a **window of exponential width** C is an interval of real numbers x of type

$$w \leq x < w^C,$$

for some $w > 1$.

We can now use Lemma 4.7 to prove a bound on the number of good rational approximations to a real number α in a window of exponential width C for any $C > 1$. We will say that a rational number $\frac{p}{q}$ is *reduced* if $\gcd(p, q) = 1$.

Lemma 4.8. *Let $\alpha \in \mathbb{R}$, $\delta > 0$, and $C > 1$. Let $N_C(\alpha)$ be the number of reduced rational numbers $\frac{p}{q}$ such that*

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{2q^{2+\delta}} \tag{18}$$

and q is in a window of exponential width C. Then

$$N_C(\alpha) \leq 1 + \frac{\ln C}{\ln(1 + \delta)}. \tag{19}$$

Proof. Notice that if x, y are in a window of exponential width C, then

$$w \leq x < w^C \leq x^C, \ w \leq y < w^C \leq y^C,$$

for some $w > 1$, hence $x \leq y^C$ and $y \leq x^C$. Now suppose that $\frac{p_1}{q_1} \neq \frac{p_2}{q_2}$ are reduced fractions that satisfy (18) with $1 \leq q_1 \leq q_2$ in a window of exponential width C. Then

$$\begin{aligned} \frac{1}{q_1 q_2} &\leq \left| \frac{p_1}{q_1} - \frac{p_2}{q_2} \right| = \left| \left(\frac{p_1}{q_1} - \alpha \right) + \left(\alpha - \frac{p_2}{q_2} \right) \right| \\ &\leq \left| \alpha - \frac{p_1}{q_1} \right| + \left| \alpha - \frac{p_2}{q_2} \right| < \frac{1}{2q_1^{2+\delta}} + \frac{1}{2q_2^{2+\delta}} \leq \frac{1}{q_1^{2+\delta}}, \end{aligned}$$

and so

$$q_2 > q_1^{1+\delta}.$$

In other words, if $q_1 \leq q_2$ are denominators of the rational approximations $\frac{p_1}{q_1}$, $\frac{p_2}{q_2}$ satisfying the hypotheses of the lemma, then

$$\gamma \ln q_1 < \ln q_2,$$

where $\gamma = 1 + \delta$, i.e. logarithms of these denominators form a γ-set. On the other hand, if q_1, q_2 are in a window of exponential width C, then

$$\ln q_1 \leq C \ln q_2, \ \ln q_2 \leq C \ln q_1,$$

that is these logarithms also form a C-set, hence they form a (C, γ)-set, and by Lemma 4.7 the cardinality of this set is

$$\leq 1 + \frac{\ln C}{\ln \gamma} = 1 + \frac{\ln C}{\ln(1+\delta)},$$

but this is precisely the number $N_C(\alpha)$. This completes the proof. □

Remark 4.5. Suppose that $1 < A < B$ are given, and suppose that we want to know the number of reduced rational approximations $\frac{p}{q}$ to the real number α with

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

and $A \leq q \leq B$. Notice that denominators q lie in a window of exponential width $C = \frac{\ln B}{\ln A}$, since

$$A = e^{\ln A} \leq q \leq B = \left(e^{\ln A} \right)^{\frac{\ln B}{\ln A}},$$

and so by Lemma 4.8, the number of such approximations is

$$\leq 1 + \frac{\ln \left(\frac{\ln B}{\ln A} \right)}{\ln(1+\delta)}.$$

Definition 4.5. Let $\alpha \in \mathbb{R}$ and let $\delta > 0$. We will call $\frac{p}{q} \in \mathbb{Q}$ a **δ-approximation** to α if $q > 0$, $\gcd(p, q) = 1$, and

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

A method similar to the proof of Lemma 4.8 yields the following result; a proof of this can be found on p. 59 of [7].

Lemma 4.9. *Let $\alpha \in \mathbb{R}$, $\delta > 0$. The number of δ-approximations $\frac{p}{q}$ to α in a window $w \leq q \leq w^C$, where $w \geq 4^{1/\delta}$ is*

$$\leq 1 + \frac{\ln 2C}{\ln(1+\delta)}.$$

5. Some field theory

Our next goal is to develop some further properties of algebraic and transcendental numbers. For this we need to introduce some elements of field theory.

Definition 5.1. Let K and L be fields with the same addition and multiplication operations such that $K \subseteq L$. Then L is called a **field extension** of K, and K is called a **subfield** of L.

Exercise 5.1. *Suppose that L is a field extension of K. Prove that L is K-vector space. Its dimension is called the* **degree** *of this field extension, denoted by $[L : K]$.*

A classical example of field extensions comes from extending a subfield of $\mathbb{C}$ (often $\mathbb{Q}$) by some collection of complex numbers. Let $K \subseteq \mathbb{C}$ be a subfield, $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$, and define $K(\alpha_1, \ldots, \alpha_n)$ to be the smallest subfield of $\mathbb{C}$ with respect to inclusion that contains K and $\alpha_1, \ldots, \alpha_n$.

Exercise 5.2. *Let K and L be subfields of a field M. Prove that their intersection $K \cap L$ is also a subfield of M. Use this fact to conclude uniqueness of $K(\alpha_1, \ldots, \alpha_n)$ as defined above.*

Exercise 5.3. *Let $K \subseteq \mathbb{C}$ be a subfield, $\alpha, \beta \in \mathbb{C}$, and let $K_1 = K(\alpha)$, $K_2 = K(\beta)$, $L = K(\alpha, \beta)$. Prove that $L = K_1(\beta) = K_2(\alpha)$. Conclude that*

$$[L : K] = [L : K_1][K_1 : K] = [L : K_2][K_2 : K].$$

Definition 5.2. Let $K \subseteq \mathbb{C}$ and $\alpha \in \mathbb{C}$. We define

$$\begin{aligned} K[\alpha] &:= \operatorname{span}_K \left\{1, \alpha, \alpha^2, \ldots\right\} \\ &= \left\{ \sum_{m=0}^{n} a_m \alpha^m : a_0, \ldots, a_n \in K,\ n \in \mathbb{N}_0 \right\}, \end{aligned}$$

i.e., the set of all finite linear combinations of powers of α with coefficients from K. Then $K[\alpha]$ is a vector space over K, whose dimension

$\dim_K K[\alpha]$ is equal to the number of powers of α which are linearly independent over K.

Exercise 5.4. *Prove that* $\dim_{\mathbb{Q}} \mathbb{Q}[\sqrt{2}] = 2$.

Exercise 5.5. *Prove that* $K[\alpha] \subseteq K(\alpha)$ *for any subfield* $K \subseteq \mathbb{C}$ *and* $\alpha \in \mathbb{C}$.

We now establish some important properties of algebraic numbers.

Theorem 5.1. *Let* $\alpha \in \mathbb{C}$.

(1) *If* α *is transcendental, then* $\dim_{\mathbb{Q}} \mathbb{Q}[\alpha] = \infty$.
(2) *If* α *is algebraic of degree* n*, then* $\mathbb{Q}[\alpha] = \mathrm{span}_{\mathbb{Q}}\{1, \alpha, \dots, \alpha^{n-1}\}$, *and* $1, \alpha, \dots, \alpha^{n-1}$ *are linearly independent over* $\mathbb{Q}$. *Hence*

$$\dim_{\mathbb{Q}} \mathbb{Q}[\alpha] = n.$$

(3) $\mathbb{Q}[\alpha]$ *is a field if and only if* α *is algebraic.*
(4) *If* α *is algebraic, then* $\mathbb{Q}(\alpha) = \mathbb{Q}[\alpha]$.

Proof. To establish part (1), assume that $\dim_{\mathbb{Q}} \mathbb{Q}[\alpha] = n < \infty$. Then the collection of $n+1$ elements $1, \alpha, \dots, \alpha^n$ must be linearly dependent, i.e. there exist $c_0, \dots, c_n \in \mathbb{Q}$ such that

$$c_0 + c_1\alpha + \cdots + c_n\alpha^n = 0.$$

Clearing the denominators, if necessary, we can assume that $c_0, \dots, c_n \in \mathbb{Z}$, and hence α is a root of $\sum_{m=0}^{n} c_m x^m \in \mathbb{Z}[x]$, which means that it is algebraic.

To prove part (2), assume that α is algebraic of degree n and let

$$m_\alpha(x) = \sum_{m=0}^{n} a_m x^m \in \mathbb{Z}[x],$$

where $a_n \neq 0$ and $a_0 \neq 0$, since $m_\alpha(x)$ is irreducible. Since $m_\alpha(\alpha) = 0$, we have

$$\alpha^n = \sum_{m=0}^{n-1} \left(-\frac{a_m}{a_n}\right) \alpha^m. \tag{20}$$

Therefore any $\mathbb{Q}$-linear combination of powers of α can be expressed as a $\mathbb{Q}$-linear combination of $1, \alpha, \dots, \alpha^{n-1}$. Now suppose $1, \alpha, \dots, \alpha^{n-1}$ are linearly dependent, then there exist $c_0, \dots, c_{n-1} \in \mathbb{Q}$ such that

$$c_0 + c_1\alpha + \dots c_{n-1}\alpha^{n-1} = 0.$$

In fact, clearing the denominators if necessary, we can assume that $c_0, \ldots, c_{n-1} \in \mathbb{Z}$. But this means that α is a root of the polynomial

$$p(x) = \sum_{m=0}^{n-1} c_m x^m \in \mathbb{Z}[x],$$

which has degree $n-1$. This contradicts the assumption that $\deg(\alpha) = n$, hence $1, \alpha, \ldots, \alpha^{n-1}$ must be linearly independent, and so they form a basis for $\mathbb{Q}[\alpha]$ over $\mathbb{Q}$.

For part (3), assume first that $\alpha \in \mathbb{C}$ is algebraic. It is clear that $\mathbb{Q}[\alpha]$ is closed under addition and multiplication. We only need to prove that for any $\beta \in \mathbb{Q}[\alpha] \setminus \{0\}$, there exists $\beta^{-1} \in \mathbb{Q}[\alpha]$. By part (2), there exist $b_0, \ldots, b_{n-1} \in \mathbb{Q}$ such that

$$\beta = \sum_{m=0}^{n-1} b_m \alpha^m.$$

We want to prove the existence of

$$\gamma = \sum_{m=0}^{n-1} c_m \alpha^m \in \mathbb{Q}[\alpha] \tag{21}$$

such that $\beta\gamma = 1$. Let γ be as in (21) with coefficients $c_0, \ldots, c_{n-1}$ to be specified, then:

$$\beta\gamma = \sum_{m=0}^{n-1} \sum_{k=0}^{n-1} b_m c_k \alpha^{m+k} = \sum_{l=0}^{2n-2} \left(\sum_{m+k=l} b_m c_k \right) \alpha^l.$$

For each $l \geq n$, we can substitute (20) for α^n, lowering the power. After a finite number of such substitutions, we will obtain an expression

$$\beta\gamma = \sum_{l=0}^{n-1} f_l(c_0, \ldots, c_{n-1}) \alpha^l,$$

where $f_l(c_0, \ldots, c_{n-1})$ is a homogeneous linear polynomial in the variables $c_0, \ldots, c_{n-1}$ with coefficients depending on b_i's and a_i's, for each $0 \leq l \leq n-1$. Since we want $\beta\gamma = 1$, we set

$$\begin{aligned} f_0(c_0, \ldots, c_{n-1}) &= 1 \\ f_1(c_0, \ldots, c_{n-1}) &= 0 \\ \vdots \ \ldots \ \vdots \ \ldots \ \vdots \ \ldots &\vdots \\ f_{n-1}(c_0, \ldots, c_{n-1}) &= 0. \end{aligned} \tag{22}$$

This is a linear system of n equations in n variables, which can be written as $F\boldsymbol{c} = \boldsymbol{e}_1$, where F is the $n \times n$ coefficient matrix of linear

polynomials $f_0, \dots, f_{n-1}$, $\boldsymbol{e}_1 = (1, 0, \dots, 0)^t \in \mathbb{R}^n$ is the first standard basis vector in $\mathbb{R}^n$, and $\boldsymbol{c} = (c_0, \dots, c_{n-1})^t$. One can work out the explicit form of (22) to prove that F is a non-singular matrix, and hence (22) has a unique solution $\boldsymbol{c}$. Let γ be as in (21) with this choice of $\boldsymbol{c}$, then $\gamma = \beta^{-1} \in \mathbb{Q}[\alpha]$, and so $\mathbb{Q}[\alpha]$ is a field.

Now suppose α is not algebraic, i.e. it is transcendental. We show that $\mathbb{Q}[\alpha]$ is not a field. Assume it is, then $\alpha^{-1} \in \mathbb{Q}[\alpha]$, which means that

$$\alpha^{-1} = \sum_{m=0}^{n} a_m \alpha^m$$

for some $n \in \mathbb{N}_0$ and $a_0, \dots, a_n \in \mathbb{Q}$. Hence

$$1 = \alpha\alpha^{-1} = \sum_{m=0}^{n} a_m \alpha^{m+1},$$

and so

$$\sum_{m=0}^{n} a_m \alpha^{m+1} - 1 = 0.$$

This is a polynomial equation over $\mathbb{Q}$ satisfied by α, and multiplying through by the product of denominators of its coefficients we can obtain a polynomial equation over $\mathbb{Z}$ satisfied by α. This contradicts the assumption that α is transcendental. Hence $\mathbb{Q}[\alpha]$ cannot be a field.

Finally we establish part (4) by proving that $\mathbb{Q}[\alpha] = \mathbb{Q}(\alpha)$. First notice that $\mathbb{Q}[\alpha] \subseteq \mathbb{Q}(\alpha)$, since every $\mathbb{Q}$-linear combination of powers of α must be contained in any field containing $\mathbb{Q}$ and α. To show containment the other way, notice that, by part (3), $\mathbb{Q}[\alpha]$ is a field containing $\mathbb{Q}$ and α, and so it must contain $\mathbb{Q}(\alpha)$. □

Example 5.2. *We give an example of finding the inverse of an element of $\mathbb{Q}[\alpha]$ when α is algebraic. Consider*

$$\beta = 3^{2/3} + 2 \times 3^{1/3} - 2 \in \mathbb{Q}[3^{1/3}].$$

We look for

$$\beta^{-1} = a3^{2/3} + b3^{1/3} - c \in \mathbb{Q}[3^{1/3}].$$

Then we need

$$\begin{aligned} 1 &= \beta\beta^{-1} = a3^{4/3} + b3^{3/3} - c3^{2/3} + 2a3^{3/3} + 2b3^{2/3} - 2c3^{1/3} \\ &\quad -2a3^{2/3} - 2b3^{1/3} + 2c \\ &= (2b - 2a - c)3^{2/3} + (3a - 2c - 2b)3^{1/3} + (2c + 6a + 3b), \end{aligned}$$

in other words we are looking for $a, b, c \in \mathbb{Q}$ such that

$$2b - 2a - c = 0,\ 3a - 2c - 2b = 0,\ 2c + 6a + 3b = 1.$$

This system has a unique solution:

$$a = \frac{6}{61},\ b = \frac{7}{61},\ c = \frac{2}{61},$$

hence

$$\beta^{-1} = \frac{6}{61}3^{2/3} + \frac{7}{61}3^{1/3} - \frac{2}{61}.$$

An immediate consequence of Theorem 5.1 is an algebraic criterion for transcendence.

Corollary 5.3. *Let $\alpha \in \mathbb{C}$. Then α is transcendental if and only if $[\mathbb{Q}(\alpha) : \mathbb{Q}] = \infty$.*

Proof. If α is transcendental, then $\dim_{\mathbb{Q}} \mathbb{Q}[\alpha] = \infty$ by part (1) of Theorem 5.1. On the other hand, $\mathbb{Q}[\alpha] \subseteq \mathbb{Q}(\alpha)$ by Exercise 5.5. Hence $\mathbb{Q}(\alpha)$ must be an infinite-dimensional $\mathbb{Q}$-vector space, hence $[\mathbb{Q}(\alpha) : \mathbb{Q}] = \infty$.

Conversely, suppose that $[\mathbb{Q}(\alpha) : \mathbb{Q}] = \infty$. Assume, towards a contradiction, that α is algebraic of degree n. By part (4) of Theorem 5.1, we have $\mathbb{Q}(\alpha) = \mathbb{Q}[\alpha]$, but by part (2) of Theorem 5.1

$$\infty > n = \dim_{\mathbb{Q}} \mathbb{Q}[\alpha] = [\mathbb{Q}(\alpha) : \mathbb{Q}].$$

This is a contradiction, and hence α must be transcendental. □

Another important consequence is the following.

Theorem 5.4. *The set $\mathbb{A}$ of algebraic numbers is a field under the usual addition and multiplication of complex numbers.*

Proof. By Theorem 3.6, we know that $\mathbb{A}$ is countable, and so we can write

$$\mathbb{A} = \{\alpha_1, \alpha_2, \alpha_3, \dots\},$$

choosing an ordering on $\mathbb{A}$. For each $n \in \mathbb{N}$, define

$$K_n := \mathbb{Q}(\alpha_1, \dots, \alpha_n).$$

Let $n \in \mathbb{N}$ and let $\beta \in K_n$, then $\mathbb{Q}(\beta) \subseteq K_n$, which means that

$$[\mathbb{Q}(\beta) : \mathbb{Q}] \leq [K_n : \mathbb{Q}] < \infty,$$

and so β is algebraic, by Theorem 5.1. Therefore any element of any field K_n is in $\mathbb{A}$, and hence we have

$$\mathbb{Q} \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq \mathbb{A}.$$

Now let $0 \neq \beta, \gamma \in \mathbb{A}$, then there exist some integers $1 \leq k \leq n$ such that $\beta = \alpha_k$, $\gamma = \alpha_n$, and so $\beta, \gamma \in K_n$. Since K_n is a field, we have

$$\beta^{-1}, \gamma^{-1}, \beta \pm \gamma, \beta\gamma \in K_n \subseteq \mathbb{A}.$$

Therefore $\mathbb{A}$ is a field. □

An immediate implication of Theorem 5.4 is that a sum, a difference, a product, or a quotient of two algebraic numbers is again an algebraic number. This not always true for transcendental numbers, which is what we show next.

Lemma 5.5. *A sum or product of an algebraic number and a transcendental number is transcendental.*

Proof. Let $\alpha \in \mathbb{C}$ be algebraic and $\beta \in \mathbb{C}$ transcendental. Then $-\alpha$ and α^{-1} are algebraic. Suppose that $\alpha+\beta$ and $\alpha\beta$ are algebraic. Since sum and product of algebraic numbers are algebraic, we must have

$$\beta = (\alpha+\beta)+(-\alpha) = (\alpha\beta)\alpha^{-1} \in \mathbb{A},$$

which is a contradiction. Hence $\alpha+\beta$ and $\alpha\beta$ must be transcendental. □

Remark 5.1. A consequence of Lemma 5.5 is that, given one transcendental number β, we can produce infinitely many (but countably many) transcendental numbers:

$$\alpha \pm \beta,\ \alpha\beta,\ \alpha^{-1}\beta \ \forall\ 0 \neq \alpha \in \mathbb{A}.$$

Take, for instance, β to be a Liouville number.

Example 5.6. *Let $\alpha \in \mathbb{C}$ be algebraic and $\beta \in \mathbb{C}$ transcendental. Then $\alpha\beta, \alpha+\beta$ are transcendental by Lemma 5.5. On the other hand,*

$$\alpha = (\alpha+\beta)-\beta = \frac{\alpha\beta}{\beta}$$

is algebraic. Hence $\mathbb{T}$ is not a field.

To conclude this section, we introduce the notion of *algebraic independence*, which will be important later in the notes.

Definition 5.3. Let $\alpha, \beta \in \mathbb{C}$ be transcendental numbers. Then, as we know from Corollary 5.3,

$$[\mathbb{Q}(\alpha):\mathbb{Q}] = [\mathbb{Q}(\beta):\mathbb{Q}] = \infty.$$

These numbers are called **algebraically independent** if

$$[\mathbb{Q}(\alpha,\beta):\mathbb{Q}(\alpha)] = [\mathbb{Q}(\alpha,\beta):\mathbb{Q}(\beta)] = \infty.$$

More generally, a collection of transcendental numbers $\alpha_1, \dots, \alpha_n$ is algebraically independent if the degree of $\mathbb{Q}(\alpha_1, \dots, \alpha_n)$ over $\mathbb{Q}(S)$, where S is any proper subcollection of $\alpha_1, \dots, \alpha_n$, is equal to infinity. If K is a subfield of $\mathbb{C}$, then its **transcendence degree**, denoted $\operatorname{trdeg} K$, is the cardinality of a maximal (with respect to size) collection of algebraically independent elements in K.

Notice that no subcollection of each infinite collection of transcendental numbers mentioned in Remark 5.1 is algebraically independent. In other words, while we can construct infinitely many transcendental numbers given one, it is not so easy to construct algebraically independent transcendental numbers.

6. Number fields

We now need to introduce some further language of algebraic number theory.

Definition 6.1. Let $\alpha \in \mathbb{A}$, then the **algebraic conjugates** of α (also often called just conjugates) are all the roots of its minimal polynomial $m_\alpha(x)$. Since $m_\alpha(x)$ is irreducible, it must be separable by Fact 3.1, and hence α and its conjugates are all distinct.

Definition 6.2. A field extension K of $\mathbb{Q}$ is called **algebraic** if every element $\alpha \in K$ is algebraic. Further, $K/\mathbb{Q}$ is called a **finite** extension if $[K : \mathbb{Q}] < \infty$. A finite algebraic extension of $\mathbb{Q}$ is called a **number field**.

It is clear from the above definition that every number field K is contained in $\mathbb{A}$. Furthermore, if $\alpha \in \mathbb{A}$, then $\mathbb{Q}(\alpha)$ is an algebraic extension of $\mathbb{Q}$, and $[\mathbb{Q}(\alpha) : \mathbb{Q}] = \deg(\alpha) < \infty$, hence $\mathbb{Q}(\alpha)$ is a number field. An element α in a number field K is called a **primitive element** if $K = \mathbb{Q}(\alpha)$, i.e. if α generates K over $\mathbb{Q}$. In fact, all number fields contain a primitive element.

Theorem 6.1 (Primitive Element Theorem)**.** *Let K be a number field. Then there exists $\alpha \in K$ such that $K = \mathbb{Q}(\alpha)$.*

Proof. Since K is a finite algebraic extension of $\mathbb{C}$, there must exist a finite collection of algebraic numbers $\alpha_1, \ldots, \alpha_n \in K$ such that $K = \mathbb{Q}(\alpha_1, \ldots, \alpha_n)$. Let $K_1 = \mathbb{Q}(\alpha_1)$, $K_2 = \mathbb{Q}(\alpha_1, \alpha_2) = K_1(\alpha_2)$, $\ldots$, $K = K_n = \mathbb{Q}(\alpha_1, \ldots, \alpha_{n-1}, \alpha_n) = K_{n-1}(\alpha_n)$. We can assume that no K_m equal to K_{m+1}, since otherwise we do not need α_{m+1} in the generating set. Hence we have

$$\mathbb{Q} \subsetneq K_1 \subsetneq K_2 \subsetneq \cdots \subsetneq K_{n-1} \subsetneq K_n = K.$$

Notice that it is sufficient for us to show that there exists $\beta_1 \in K$ such that $K_2 = \mathbb{Q}(\alpha_1, \alpha_2) = \mathbb{Q}(\beta_1)$: if this the case, then applying the same reasoning, we establish that

$$K_3 = K_2(\alpha_3) = \mathbb{Q}(\beta_1, \alpha_3) = \mathbb{Q}(\beta_2)$$

for some $\beta_2 \in K$, and continuing in the same manner confirm that $K = K_n = \mathbb{Q}(\beta_{n-1})$ for some $\beta_{n-1} \in K$.

Let $\deg(\alpha_1) = d$, $\deg(\alpha_2) = e$, and let

$$\alpha_1 = \alpha_{11}, \alpha_{12}, \dots, \alpha_{1d} \text{ and } \alpha_2 = \alpha_{21}, \alpha_{22}, \dots, \alpha_{2e}$$

be algebraic conjugates of α_1 and α_2, respectively. Since $m_{\alpha_1}(x)$ and $m_{\alpha_2}(x)$ in $\mathbb{Z}[x]$ are irreducible, they must be separable by Fact 3.1 above, and hence all α_{1n}'s and all α_{2m}'s are distinct. This means that for each $1 \leq n \leq d$, $1 < m \leq e$ the equation

$$\alpha_{1n} + t\alpha_{2m} = \alpha_{11} + t\alpha_{21} \tag{23}$$

has at most one solution t in $\mathbb{Q}$ (a solution t in $\mathbb{C}$ always exists, but it may not be in $\mathbb{Q}$). There are only finitely many equations (23), each having at most one solution, and hence we can choose $0 \neq c \in \mathbb{Q}$ which is not one of these solutions, then

$$\alpha_{1n} + c\alpha_{2m} \neq \alpha_{11} + c\alpha_{21}$$

for any $1 \leq n \leq d$, $1 < m \leq e$. Let

$$\beta_1 = \alpha_1 + c\alpha_2,$$

then $\beta_1 \neq \alpha_{1n} + c\alpha_{2m}$ for any $1 \leq n \leq d$, $1 < m \leq e$. We will now prove that $\mathbb{Q}(\beta_1) = \mathbb{Q}(\alpha_1, \alpha_2)$. It is clear that $\mathbb{Q}(\beta_1) \subseteq \mathbb{Q}(\alpha_1, \alpha_2)$, so we only need to show that $\mathbb{Q}(\alpha_1, \alpha_2) \subseteq \mathbb{Q}(\beta_1)$. For this, it is sufficient to prove that $\alpha_2 \in \mathbb{Q}(\beta_1)$, since then $\alpha_1 = \beta_1 - c\alpha_2 \in \mathbb{Q}(\beta_1)$, and hence $\mathbb{Q}(\alpha_1, \alpha_2) \subseteq \mathbb{Q}(\beta_1)$. Notice that

$$m_{\alpha_1}(\beta_1 - c\alpha_2) = m_{\alpha_1}(\alpha_1) = 0.$$

In other words, α_2 is a zero of the polynomial

$$p(x) := m_{\alpha_1}(\beta_1 - cx),$$

which has coefficients in $\mathbb{Q}(\beta_1)$. On the other hand, α_2 is also a root of its minimal polynomial $m_{\alpha_2}(x)$. The two polynomials $p(x)$ and $m_{\alpha_2}(x)$ have only one common root. Indeed, if $\xi \in \mathbb{C}$ is such that

$$p(\xi) = m_{\alpha_1}(\beta_1 - c\xi) = m_{\alpha_2}(\xi) = 0,$$

then ξ must be one of $\alpha_{21}, \dots, \alpha_{2e}$ and $\beta_1 - c\xi$ one of $\alpha_{11}, \dots, \alpha_{1d}$, i.e., for some $1 \leq n \leq d$, $1 \leq m \leq e$,

$$\xi = \alpha_{2m} \text{ and } \beta_1 - c\xi = \beta_1 - c\alpha_{2m} = \alpha_{1n},$$

which means that

$$\beta_1 = \alpha_{1n} + c\alpha_{2m} = \alpha_{11} + c\alpha_{21}.$$

This contradicts our choice of c unless $n = m = 1$.

Now let $h(x)$ be a minimal polynomial of α_2 over $\mathbb{Q}(\beta_1)$, as described in Remark 3.2. Since $p(x)$ and $m_{\alpha_2}(x)$ have coefficients in $\mathbb{Q}(\beta_1)$ and vanish at α_2, they must both be divisible by $h(x)$ over $\mathbb{Q}(\beta_1)$. This means that every root of $h(x)$ would be a common root of $p(x)$ and $m_{\alpha_2}(x)$, but we know that they have precisely one root in common. This means that $h(x)$ can have only one root, and hence is of degree 1. Thus

$$h(x) = x - \alpha_2,$$

which means that $\alpha_2 \in \mathbb{Q}(\beta_1)$. This completes the proof. □

An algebraic number α is called an **algebraic integer** if its minimal polynomial $m_\alpha(x) \in \mathbb{Z}[x]$ is *monic*, i.e. its leading coefficient is equal to 1. The set of all algebraic integers in a number field K is usually denoted by $\mathcal{O}_K$.

Exercise 6.1. *Prove that $\mathcal{O}_\mathbb{Q} = \mathbb{Z}$. Due to this property, elements of $\mathbb{Z}$ are often called* **rational integers**. *Prove also that $\mathbb{Z} \subseteq \mathcal{O}_K$ for any number field K.*

Let us define the set of all algebraic integers

$$\mathbb{I} = \{\alpha \in \mathbb{A} : m_\alpha(x) \text{ is monic}\}.$$

Let $\alpha \in \mathbb{I}$ and let $\deg(\alpha) = d$. Define

$$\mathbb{Z}[\alpha] := \left\{ \sum_{n=0}^{d-1} a_n \alpha^n : a_0, \ldots, a_{d-1} \in \mathbb{Z} \right\}.$$

Lemma 6.2. *Let $\alpha \in \mathbb{I}$ have degree d. Then $\mathbb{Z}[\alpha]$ is a commutative ring with identity under the usual addition and multiplication operations on complex numbers, which contains $\mathbb{Z}$. Rings like this are called* **ring extensions** *of $\mathbb{Z}$.*

Proof. The argument here bears some similarity with the proof of Theorem 5.1 above. It is clear that $\mathbb{Z} \subseteq \mathbb{Z}[\alpha]$, and hence $0, 1 \in \mathbb{Z}[\alpha]$. Also, if $\beta = \sum_{n=0}^{d-1} b_n \alpha^n \in \mathbb{Z}[\alpha]$, then $-\beta = \sum_{n=0}^{d-1} (-b_n)\alpha^n \in \mathbb{Z}[\alpha]$. Hence we only need to prove that for every $\beta, \gamma \in \mathbb{Z}[\alpha]$, $\beta + \gamma, \beta\gamma \in \mathbb{Z}[\alpha]$. Let

$$\beta = \sum_{n=0}^{d-1} b_n \alpha^n, \ \gamma = \sum_{n=0}^{d-1} c_n \alpha^n.$$

Then

$$\beta + \gamma = \sum_{n=0}^{d-1} (b_n + c_n)\alpha^n \in \mathbb{Z}[\alpha].$$

Since $\alpha \in \mathbb{I}$ of degree d, its minimal polynomial is monic of degree d, say

$$m_\alpha(x) = \alpha^d + \sum_{n=0}^{d-1} a_n x^n,$$

and $m_\alpha(\alpha) = 0$, meaning that

$$\alpha^d = -\sum_{n=0}^{d-1} a_n \alpha^n. \tag{24}$$

Now we have:

$$\beta\gamma = \sum_{n=0}^{d-1} \sum_{m=0}^{d-1} b_n c_m \alpha^{n+m},$$

and (24) can be used to express powers of α higher than $(d-1)$-st as linear combinations of lower powers of α with rational integer coefficients, hence ensuring that $\beta\gamma$ is a linear combination of the terms $1, \alpha, \dots, \alpha^{d-1}$ with coefficients in $\mathbb{Z}$. This means that $\beta\gamma \in \mathbb{Z}[\alpha]$ and completes the proof of the lemma. □

Exercise 6.2. *Let $\alpha \in \mathbb{I}$ have degree d and let $n \in \mathbb{N}_0$. Use* (24) *above to prove that the elements*

$$1, \alpha, \dots, \alpha^n$$

are linearly independent over $\mathbb{Z}$ if and only if $n < d$.

Exercise 6.2 guarantees that $1, \alpha, \dots, \alpha^{d-1}$ is a maximal linearly independent collection of powers of α over d, and we know that it spans $\mathbb{Z}[\alpha]$. Hence we call it a **basis** for $\mathbb{Z}[\alpha]$, which is an example of a **free $\mathbb{Z}$-module** or a **lattice**, that is an analogue of a vector space with coefficients in linear combinations of basis vectors coming from the ring $\mathbb{Z}$ instead of a field. The cardinality of such a basis, d in our case, is called the **rank** of the lattice $\mathbb{Z}[\alpha]$.

Lemma 6.3. *Let $\alpha \in \mathbb{C}$ be such that the additive abelian group generated by all powers of α is in fact finitely generated. Then $\alpha \in \mathbb{I}$.*

Proof. Let G be the additive abelian group generated by all powers of α, i.e.

$$G = \left\{ \sum_{n=0}^{k} a_n \alpha^n : k \in \mathbb{N}_0,\ a_0, \dots, a_k \in \mathbb{Z} \right\}.$$

Assume that G is finitely generated and let $v_1, \dots, v_m$ be a generating set for G. Since each v_n is a polynomial in α, there exists a positive integer ℓ which is the maximal power of α present in the representations

of $v_1, \ldots, v_m$. Then G is generated by $1, \alpha, \ldots, \alpha^\ell$. Since $\alpha^{\ell+1} \in G$, there must exist $a_0, \ldots, a_\ell \in \mathbb{Z}$ such that

$$\alpha^{\ell+1} = \sum_{n=0}^{\ell} a_n \alpha^n,$$

which means that α is a root of the polynomial

$$p(x) = x^{\ell+1} - \sum_{n=0}^{\ell} a_n x^n \in \mathbb{Z}[x].$$

By Exercise 3.2, we know that $m_\alpha(x) \mid p(x)$. Since $p(x)$ is a monic polynomial, it must be true that $m_\alpha(x)$ is also monic. Hence $\alpha \in \mathbb{I}$. $\square$

Theorem 6.4. *$\mathbb{I}$ is a commutative ring with identity under the usual addition and multiplication of complex numbers.*

Proof. We only need to prove that for any $\alpha, \beta \in \mathbb{I}$, $\alpha + \beta$ and $\alpha\beta$ are in $\mathbb{I}$. Notice that $\alpha+\beta$ and $\alpha\beta$ can be expressed as integral linear combinations of elements of the form $\alpha^m\beta^n$ for some nonnegative integers m, n, which means that

$$\alpha + \beta, \alpha\beta \in G := \mathbb{Z}[\alpha]\mathbb{Z}[\beta] \subset \mathbb{C}.$$

Exercise 6.3. *Prove that this G is a subgroup of $\mathbb{C}$ under the usual addition of complex numbers, and hence is an additive abelian group.*

Since α and β are algebraic integers, we know that $\mathbb{Z}[\alpha]$ and $\mathbb{Z}[\beta]$ are generated by only finitely many powers of α and β, respectively, say, it is $1, \alpha, \ldots, \alpha^k$ and $1, \beta, \ldots, \beta^\ell$. Then

$$G = \left\{ \left(\sum_{n=0}^{k} a_n \alpha^n \right) \left(\sum_{m=0}^{\ell} b_m \beta^m \right) : a_0, \ldots, a_k, b_0, \ldots, b_\ell \in \mathbb{Z} \right\},$$

and hence G is generated by all expressions of the form $\alpha^n\beta^m$ as an additive abelian group. Therefore G must also be finitely generated.

Fact 6.5. *Let G be a finitely generated additive abelian group, i.e., there exist $v_1, \ldots, v_k \in G$ such that for every $x \in G$,*

$$x = \sum_{n=1}^{k} a_n v_n$$

for some $a_1, \ldots, a_k \in \mathbb{Z}$. Let H be a subgroup of G. Then H is also finitely generated. This fact is established in Theorem A.4 in Appendix A below.

Since additive groups generated by all powers of $\alpha+\beta$ and $\alpha\beta$, respectively, are subgroups of G, the exercise above implies that they are also finitely generated. Now Lemma 6.3 guarantees that $\alpha+\beta$ and $\alpha\beta$ must be in $\mathbb{I}$. $\square$

Exercise 6.4. *Let A and B be subrings of the same ring R. Prove that $A \cap B$ is also a ring. Use this fact to prove that for any number field K, the set $\mathcal{O}_K$ of all algebraic integers in K is a commutative ring with identity.*

We now further study some properties of the ring of algebraic integers $\mathcal{O}_K$ of a number field K. First we observe that every element of K can be expressed as a fraction α/c, where α is an algebraic integer and c is a rational integer.

Lemma 6.6. *Let K be a number field and $\beta \in K$. Then there exists some $c \in \mathbb{N}$ such that $c\beta \in \mathcal{O}_K$. In fact, we can take c to be the leading coefficient of $m_\beta(x)$.*

Proof. Let $d = \deg(\beta)$ and let

$$m_\beta(x) = \sum_{n=0}^{d} a_n x^n \in \mathbb{Z}[x]$$

with $a_d > 0$. Notice that

$$p(x) := a_d^{d-1} m_\beta(x) = \sum_{n=0}^{d} a_n a_d^{d-1} x^n = \sum_{n=0}^{d} a_n a_d^{d-n-1} (a_d x)^n$$

has β as its root. Now

$$f(x) = \sum_{n=0}^{d} a_n a_d^{d-n-1} x^n = x^d + \sum_{n=0}^{d-1} a_n a_d^{d-n-1} x^n \in \mathbb{Z}[x]$$

is a monic polynomial, and $f(a_d\beta) = p(\beta) = 0$. This means that $a_d\beta \in \mathcal{O}_K$. Taking $c = a_d$ completes the proof of the lemma. $\square$

This lemma has some important corollaries.

Corollary 6.7. *A number field K can be described as*

$$K = \left\{ \frac{\alpha}{\beta} : \alpha, \beta \in \mathcal{O}_K, \beta \neq 0 \right\}.$$

Hence we can refer to K as the **field of fractions** *or* **quotient field** *of $\mathcal{O}_K$.*

Proof. Let

$$E := \left\{ \frac{\alpha}{\beta} : \alpha, \beta \in \mathcal{O}_K, \beta \neq 0 \right\}.$$

We need to prove that $E = K$. Lemma 6.6 implies that every $\beta \in K$ can be written as $\beta = \frac{\alpha}{c}$ for some $\alpha \in \mathcal{O}_K$ and $c \in \mathbb{Z}$. Since $\mathbb{Z} \subseteq \mathcal{O}_K$, we see that $\beta \in E$, hence $K \subseteq E$. Now suppose $\alpha/\beta = \alpha\beta^{-1} \in E$. Since $\alpha, \beta \in \mathcal{O}_K \subset K$, we must have $\beta^{-1} \in K$ and hence $\alpha\beta^{-1} \in K$, since K is a field. Therefore $E \subseteq K$, and thus $E = K$. □

Theorem 6.1 guarantees that a number field always has a primitive element. In fact, it always has a primitive element, which is an algebraic integer.

Corollary 6.8. *Let K be a number field. Then there exists $\alpha \in \mathcal{O}_K$ such that $K = \mathbb{Q}(\alpha)$.*

Proof. Let $\beta \in K$ be a primitive element. By Lemma 6.6, there exists an element $c \in \mathbb{Z}$ such that $\alpha := c\beta \in \mathcal{O}_K$. Since clearly $\mathbb{Q}(c\beta) = \mathbb{Q}(\beta)$, we are done. □

We can now define **embeddings** of a number field K into $\mathbb{C}$. Let $K = \mathbb{Q}(\alpha)$, then

$$d := \deg(\alpha) = [K : \mathbb{Q}].$$

Recall that

$$K = \mathbb{Q}(\alpha) = \mathbb{Q}[\alpha] = \mathrm{span}_{\mathbb{Q}}\{1, \alpha, \ldots, \alpha^{d-1}\},$$

and $1, \alpha, \ldots, \alpha^{d-1}$ are linearly independent over $\mathbb{Q}$. Let

$$\alpha = \alpha_1, \alpha_2, \ldots, \alpha_d$$

be the algebraic conjugates of α. For each $1 \leq n \leq d$, define a map $\sigma_n : K \to \mathbb{C}$, given by

$$\sigma_n \left(\sum_{m=0}^{d-1} a_m \alpha^m \right) = \sum_{m=0}^{d-1} a_m \alpha_n^m, \tag{25}$$

for each $\sum_{m=0}^{d-1} a_m \alpha^m \in K$.

Exercise 6.5. *Prove that each σ_n as defined above is an injective field homomorphism, and hence $K \cong \sigma_n(K)$ for each $1 \leq n \leq d$. Prove also that*

$$\mathbb{Q} = \{\beta \in K : \sigma_n(\beta) = \beta \ \forall \ 1 \leq n \leq d\}.$$

The embeddings $\sigma_1, \ldots, \sigma_d$ described above are, in fact, the *only* possible embeddings of K into $\mathbb{C}$.

Lemma 6.9. *Let $K = \mathbb{Q}(\alpha)$ be a number field of degree d over $\mathbb{Q}$. Let $\tau : K \to \mathbb{C}$ be an embedding, i.e. an injective field homomorphism. Then τ is one of the embeddings $\sigma_1, \dots, \sigma_d$ as defined in* (25).

Proof. First we will prove that $\tau(c) = c$ for each $c \in \mathbb{Q}$. Since τ is a field homomorphism, we must have $\tau(1) = 1$, and for each $a/b \in \mathbb{Q}$,

$$\tau(a/b) = \tau(a)\tau(b)^{-1} = a\tau(1)(b\tau(1))^{-1} = a/b.$$

Since $[K : \mathbb{Q}] = d$, we know that $\deg(\alpha) = d$, and so

$$K = \mathbb{Q}[\alpha] = \mathrm{span}_{\mathbb{Q}}\{1, \alpha, \dots, \alpha^{d-1}\}.$$

Let $\alpha = \alpha_1, \alpha_2, \dots, \alpha_d$ be the algebraic conjugates of α. Let $\beta = \sum_{n=0}^{d-1} c_n \alpha^n \in K$. Since τ is a field homomorphism,

$$\tau(\beta) = \sum_{n=0}^{d-1} \tau(c_n)\tau(\alpha)^n = \sum_{n=0}^{d-1} c_n \tau(\alpha)^n.$$

Hence we only need to show that $\tau(\alpha) = \alpha_n$ for some $1 \le n \le d$. Let

$$m_\alpha(x) = \sum_{m=0}^{d} b_m x^m \in \mathbb{Z}[x],$$

be the minimal polynomial of α. Then

$$m_\alpha(\alpha) = \sum_{m=0}^{d} b_m \alpha^m = 0,$$

and so

$$0 = \sum_{m=0}^{d} b_m \tau(\alpha)^m = m_\alpha(\tau(\alpha)).$$

Hence $\tau(\alpha)$ is a root of $m_\alpha(x)$, which means that $\tau(\alpha) = \alpha_n$ for some $1 \le n \le d$. Therefore $\tau = \sigma_n$ for some σ_n as in (25). This completes the proof. □

Exercise 6.6. *If $K = \sigma_n(K)$ for each $1 \le n \le d$, then the number field K is called* **Galois**. *In this case, prove that the set*

$$G := \{\sigma_1, \dots, \sigma_d\}$$

is a group under the operation of function composition. It is called the **Galois group** *of K over $\mathbb{Q}$, where $\mathbb{Q}$ is precisely the* **fixed field** *of G, as you just proved above. In this case, elements of G are called* **automorphisms** *of K over $\mathbb{Q}$.*

7. Function fields and transcendence

In this section, we briefly give another characterization of transcendental numbers. We start by defining polynomial rings in several variables. A monomial in the variables $x_1, \ldots, x_k$, $k \geq 1$, is an expression of the form

$$x_1^{m_1} x_2^{m_2} \cdots x_k^{m_k}, \tag{26}$$

where $m_1, \ldots, m_k \in \mathbb{N}_0$ with $m_1 + \cdots + m_k > 0$. Let R be a commutative ring with 1. Define $R[x_1, \ldots, x_k]$ to be the set of all finite linear combinations of 1 and all possible monomials as in (26) with coefficients from R.

Exercise 7.1. *Prove that $R[x_1, \ldots, x_k]$ is a commutative ring with identity under the standard operations of addition and multiplication on these multivariable polynomials.*

Let K be a field and $K[x_1, \ldots, x_k]$ be the polynomial ring in $k \geq 1$ variables with coefficients in K. Define

$$K(x_1, \ldots, x_k) = \left\{ \frac{p(x_1, \ldots, x_k)}{q(x_1, \ldots, x_k)} : p, q \in K[x_1, \ldots, x_k],\ q \neq 0 \right\},$$

where we say that

$$p(x_1, \ldots, x_k)/q(x_1, \ldots, x_k) = f(x_1, \ldots, x_k)/g(x_1, \ldots, x_k)$$

if and only if $p(x_1, \ldots, x_k)g(x_1, \ldots, x_k) = f(x_1, \ldots, x_k)q(x_1, \ldots, x_k)$; this can be viewed as an equivalence relation on the set of pairs of polynomials and then $K(x_1, \ldots, x_k)$ is the set of equivalence classes, analogously to construction of $\mathbb{Q}$ from $\mathbb{Z}$.

Exercise 7.2. *Write $\boldsymbol{x}$ for the variable vector $(x_1, \ldots, x_k)$, and prove that $K(\boldsymbol{x})$ is a field under the standard operations of addition and multiplication of rational functions:*

$$\frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} + \frac{f(\boldsymbol{x})}{g(\boldsymbol{x})} = \frac{p(\boldsymbol{x})g(\boldsymbol{x}) + f(\boldsymbol{x})q(\boldsymbol{x})}{q(\boldsymbol{x})g(\boldsymbol{x})}, \quad \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} \cdot \frac{f(\boldsymbol{x})}{g(\boldsymbol{x})} = \frac{p(\boldsymbol{x})f(\boldsymbol{x})}{q(\boldsymbol{x})g(\boldsymbol{x})}.$$

$K(x_1, \ldots, x_k)$ is called the **function field** *or* **field of rational functions** *in k variables over K, and is precisely the quotient field of the polynomial ring $K[x_1, \ldots, x_k]$.*

We can now give an alternative definition of algebraic independence.

Lemma 7.1. *A collection of numbers $\alpha_1, \ldots, \alpha_k \in \mathbb{C}$ is algebraically independent if and only if there does not exist any nonzero polynomial $p(x_1, \ldots, x_k) \in \mathbb{Q}[x_1, \ldots, x_k]$ such that*

$$p(\alpha_1, \ldots, \alpha_k) = 0. \tag{27}$$

Proof. Suppose that there exists some nonzero polynomial p satisfying (27). Define

$$f(x) = p(\alpha_1, \dots, \alpha_{k-1}, x).$$

Then $f(x) \in \mathbb{Q}(\alpha_1, \dots, \alpha_{k-1})[x]$ and $f(\alpha_k) = 0$. Let $d = \deg(f(x))$, then $1, \alpha_k, \dots, \alpha_k^d$ are linearly dependent over $\mathbb{Q}(\alpha_1, \dots, \alpha_{k-1})$. This means that

$$[\mathbb{Q}(\alpha_1, \dots, \alpha_k) : \mathbb{Q}(\alpha_1, \dots, \alpha_{k-1})] \leq d < \infty,$$

and hence $\alpha_1, \dots, \alpha_k$ are not algebraically independent. Thus, if the numbers $\alpha_1, \dots, \alpha_k$ are algebraically independent, then no nonzero polynomial p satisfying (27) can exist.

Conversely, suppose now that no nonzero polynomial p satisfying (27) exists. Suppose, towards a contradiction, that $\alpha_1, \dots, \alpha_k$ are algebraically dependent. Then, without loss of generality, we can assume that

$$[\mathbb{Q}(\alpha_1, \dots, \alpha_k) : \mathbb{Q}(\alpha_1, \dots, \alpha_{k-1})] < \infty.$$

Hence $1, \alpha_k, \dots, \alpha_k^d$ are linearly dependent over $\mathbb{Q}(\alpha_1, \dots, \alpha_{k-1})$ for some d. In other words, there exist $a_0, \dots, a_d \in \mathbb{Q}(\alpha_1, \dots, \alpha_{k-1})$ such that

$$\sum_{n=0}^{d} a_n \alpha_k^n = 0. \tag{28}$$

Notice that $a_0, \dots, a_d$ are rational functions in $\alpha_1, \dots, \alpha_{k-1}$, say

$$a_n = \frac{p_n(\alpha_1, \dots, \alpha_{k-1})}{q_n(\alpha_1, \dots, \alpha_{k-1})},$$

where $p_n(x_1, \dots, x_{k-1}), q_n(x_1, \dots, x_{k-1}) \in \mathbb{Q}[x_1, \dots, x_{k-1}]$. Write $\boldsymbol{x}$ for $(x_1, \dots, x_{k-1})$, $\boldsymbol{\alpha}$ for $(\alpha_1, \dots, \alpha_{k-1})$, and notice by (28) we have:

$$\sum_{n=0}^{d} p_n(\boldsymbol{\alpha}) \left(\prod_{m=0, m\neq n}^{d} q_m(\boldsymbol{\alpha}) \right) \alpha_k^n = 0.$$

Then define

$$p(x_1, \dots, x_k) = \sum_{n=0}^{d} p_n(\boldsymbol{x}) \left(\prod_{m=0, m\neq n}^{d} q_m(\boldsymbol{x}) \right) x_k^n \in \mathbb{Q}[x_1, \dots, x_k],$$

and notice that $p(\alpha_1, \dots, \alpha_k) = 0$. This contradicts our assumption, and hence $\alpha_1, \dots, \alpha_k$ must be algebraically independent. □

Let $\alpha_1, \dots, \alpha_k \in \mathbb{C}$, and consider a subfield $\mathbb{Q}(\alpha_1, \dots, \alpha_k) \subseteq \mathbb{C}$ generated by these elements. Let us write $\boldsymbol{\alpha}$ for the k-tuple $(\alpha_1, \dots, \alpha_k)$, and define the *evaluation map* $\varphi_{\boldsymbol{\alpha}} : \mathbb{Q}(x_1, \dots, x_k) \to \mathbb{Q}(\alpha_1, \dots, \alpha_k)$ given by sending $x_n \mapsto \alpha_n$ and extending to the rest of $\mathbb{Q}(x_1, \dots, x_k)$,

i.e., a rational function in $x_1, \ldots, x_k$ will map to its value at the point with $x_1 = \alpha_1, \ldots, x_k = \alpha_k$.

Theorem 7.2. *The following statements are equivalent:*

(1) *The map $\varphi_{\boldsymbol{\alpha}}$ is well-defined for all $f \in \mathbb{Q}(x_1, \ldots, x_k)$.*
(2) *The map $\varphi_{\boldsymbol{\alpha}}$ is an isomorphism of fields.*
(3) *The numbers $\alpha_1, \ldots, \alpha_k$ are algebraically independent.*

Proof. (1) $\Rightarrow$ (2): Let $f, g \in \mathbb{Q}(x_1, \ldots, x_k)$, then

$$\varphi_{\boldsymbol{\alpha}}(f+g) = (f+g)(\boldsymbol{\alpha}) = f(\boldsymbol{\alpha}) + g(\boldsymbol{\alpha}) = \varphi_{\boldsymbol{\alpha}}(f) + \varphi_{\boldsymbol{\alpha}}(g),$$

$$\varphi_{\boldsymbol{\alpha}}(fg) = (fg)(\boldsymbol{\alpha}) = f(\boldsymbol{\alpha})g(\boldsymbol{\alpha}) = \varphi_{\boldsymbol{\alpha}}(f)\varphi_{\boldsymbol{\alpha}}(g).$$

Hence $\varphi_{\boldsymbol{\alpha}}$ is a ring homomorphism. Suppose that $f \in \mathrm{Ker}(\varphi_{\boldsymbol{\alpha}})$, then $\varphi_{\boldsymbol{\alpha}}(f) = f(\boldsymbol{\alpha}) = 0$. We can write $f = g/h$, where $g, h \in \mathbb{Q}[x_1, \ldots, x_k]$ are polynomials in k variables with coefficients in $\mathbb{Q}$. Since $f(\boldsymbol{\alpha}) = 0$, we must have

$$g(\alpha_1, \ldots, \alpha_k) = 0.$$

Assume $g \neq 0$, then $1/g \in \mathbb{Q}(x_1, \ldots, x_k)$, however $\varphi_{\boldsymbol{\alpha}}$ is not defined at $1/g$. Hence we must have $g = 0$, meaning that $f = 0$. Therefore $\mathrm{Ker}(\varphi_{\boldsymbol{\alpha}}) = \{0\}$, and so $\varphi_{\boldsymbol{\alpha}}$ is injective. Finally, every element β of $\mathbb{Q}(\alpha_1, \ldots, \alpha_k)$ is a rational function in $\alpha_1, \ldots, \alpha_k$, which means that β is the value of some $f \in \mathbb{Q}(x_1, \ldots, x_k)$ at $\boldsymbol{\alpha}$. This proves surjectivity, and hence $\varphi_{\boldsymbol{\alpha}}$ is a field isomorphism.

(2) $\Rightarrow$ (3): If $\varphi_{\boldsymbol{\alpha}}$ is a field isomorphism, it must be well-defined as a function for each $f = g/h \in \mathbb{Q}(x_1, \ldots, x_k)$, where $g, h \in \mathbb{Q}[x_1, \ldots, x_k]$. This means that cannot exist a polynomial $p(x_1, \ldots, x_k) \in \mathbb{Q}[x_1, \ldots, x_k]$ such that $p(\boldsymbol{\alpha}) = 0$. Hence $\alpha_1, \ldots, \alpha_k$ are algebraically independent by Lemma 7.1.

(3) $\Rightarrow$ (1): Since $\alpha_1, \ldots, \alpha_k$ are algebraically independent, Lemma 7.1 implies that for any $0 \neq p \in \mathbb{Q}[x_1, \ldots, x_k]$, $p(\boldsymbol{\alpha}) \neq 0$. Then for any $f = g/h \in \mathbb{Q}(x_1, \ldots, x_k)$, where $g, h \in \mathbb{Q}[x_1, \ldots, x_k]$, $\varphi_{\boldsymbol{\alpha}}(f) = g(\boldsymbol{\alpha})/h(\boldsymbol{\alpha})$ is well-defined. □

Hence we have the following immediate characterization of transcendence and algebraic independence.

Corollary 7.3. *A collection of complex numbers $\alpha_1, \ldots, \alpha_k$ is algebraically independent if and only if $\mathbb{Q}(\alpha_1, \ldots, \alpha_k) \cong \mathbb{Q}(x_1, \ldots, x_k)$. In particular, $\alpha \in \mathbb{C}$ is transcendental if and only if $\mathbb{Q}(\alpha) \cong \mathbb{Q}(x)$.*

8. Hermite, Lindemann, Weierstrass

Arguably the two most famous transcendental numbers are e and π. Transcendence of e was originally established by Charles Hermite in 1873, and transcendence of π established in 1882 by Ferdinand von Lindemann by an extension of Hermite's technique. The much more general statement, from which these two results follow, was obtained by Karl Weierstrass in 1885. The most general form of the Hermite-Lindemann-Weierstrass Theorem is as follows.

Theorem 8.1. *Let $s \in \mathbb{N}$, $\alpha_1, \dots, \alpha_s$ be distinct algebraic numbers, and $d_1, \dots, d_s$ nonzero algebraic numbers. Then*

$$\sum_{k=1}^{s} d_k e^{\alpha_k} \neq 0.$$

In this section we will establish the famous results of Hermite, Lindemann, and Weierstrass. The general idea of the method used is similar in all three cases, however it will be easier to follow the development of this technique starting with transcendence of e, then π, and only then the general Theorem 8.1. Our exposition here closely follows [3]. We start with some preliminary observations.

Let $f(x)$ be a polynomial with complex coefficients, and let $F(x)$ be the polynomial obtained from $f(x)$ by replacing each coefficient of f with its absolute value. For a complex number t, define

$$I(t, f) := \int_0^t e^{t-u} f(u)\, du. \tag{29}$$

Then it is easy to see that

$$|I(t, f)| \leq |t| e^{|t|} F(|t|). \tag{30}$$

On the other hand, integrating by parts, we see that

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

If degree of $f(x)$ is equal to m, then iterating the above procedure m times, we obtain:

$$I(t, f) = e^t \sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m} f^{(j)}(t). \tag{31}$$

Theorem 8.2 (Hermite, 1873)**.** *The number e is transcendental.*

Proof. Working towards a contradiction, suppose that e is algebraic. Then there exist some integers $a_0, \ldots, a_n$, $n \geq 1$, such that

$$\sum_{k=0}^{n} a_k e^k = 0, \tag{32}$$

where $a_0, a_n \neq 0$. Let $p > |a_0|$ be a prime, and define a polynomial

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

Then degree of $f(x)$ is $m = (n+1)p - 1$ and each of the roots $x = 1, \ldots, n$ of $f(x)$ has multiplicity p and the root $x = 0$ has multiplicity $p-1$, which implies that

$$f^{(j)}(k) = 0 \ \forall\ 1 \leq k \leq n, 0 \leq j \leq p, \ f^{(j)}(0) = 0 \ \forall\ 0 \leq j \leq p-1. \tag{33}$$

With this notation, define

$$J := \sum_{k=0}^{n} a_k I(k, f),$$

where $I(k, f)$ is as in (29). Then, by (31) above, we have

$$\begin{aligned} J &= \sum_{k=0}^{n} \left(a_k e^k \sum_{j=0}^{m} f^{(j)}(0) - a_k \sum_{j=0}^{m} f^{(j)}(k) \right) \\ &= \sum_{j=0}^{m} \left(f^{(j)}(0) \sum_{k=0}^{n} a_k e^k \right) - \sum_{j=0}^{m} \sum_{k=0}^{n} a_k f^{(j)}(k) \\ &= -\sum_{j=0}^{m} \sum_{k=0}^{n} a_k f^{(j)}(k) = -\sum_{j=p-1}^{m} \sum_{k=0}^{n} a_k f^{(j)}(k), \end{aligned}$$

where the last line follows by (32) and (33). For $j = p-1$, the contribution from f is

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

hence, if $n < p$, then $f^{(p-1)}(0)$ is divisible by $(p-1)!$, but not by p. Now, for every $j \geq p$, then $f^{(j)}(0)$ and $f^{(j)}(k)$ for every $1 \leq k \leq n$ are divisible by $p!$. In other words, J is a nonzero integer divisible by $(p-1)!$, and so

$$|J| \geq (p-1)! \tag{34}$$

On the other hand, let $A = \max_{0 \leq k \leq n} |a_k|$, then (30) implies that

$$|J| \leq (n+1)A|I(k, f)| \leq n(n+1)Ae^n \max_{1 \leq k \leq n} F(k).$$

Notice that

$$\max_{1\leq k\leq n} F(k) = (2n)^{p-1}\,(2(n-1)!)^p = \frac{1}{n}((2n)!)^p,$$

and so

$$|J| \leq A(n+1)e^n((2n)!)^p. \tag{35}$$

Combining (34) and (35), we obtain:

$$(p-1)! \leq A(n+1)e^n((2n)!)^p,$$

which is certainly not true for sufficiently large p, and so we have a contradiction. □

To attempt the proof of transcendence of π, we need the notion of symmetric polynomials. Let $n \geq 1$ and define S_n to be the set of all permutations of the set of n elements $\{1, \ldots, n\}$.

Exercise 8.1. *Prove that S_n is a group under the operation of function composition.*

The group S_n is called the **symmetric group** on n letters. A polynomial $f(x_1, \ldots, x_n) \in \mathbb{Q}[x_1, \ldots, x_n]$ is called **symmetric** if for every $\tau \in S_n$,

$$f(x_1, \ldots, x_n) = f(x_{\tau(1)}, \ldots, x_{\tau(n)}).$$

Fact 8.3. *Let $\alpha \in \mathbb{A}$ be of degree n and let $\alpha = \alpha_1, \ldots, \alpha_n$ be algebraic conjugates of α. Let $f(x_1, \ldots, x_n) \in \mathbb{Q}[x_1, \ldots, x_n]$ be a symmetric polynomial. Then*

$$f(\alpha_1, \ldots, \alpha_n) \in \mathbb{Q}.$$

Moreover, if $\alpha \in \mathbb{I}$ and $f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$, then

$$f(\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}.$$

Let us also recall that π is half the circumference of a circle of radius 1, which is precisely the angle that the ray emanating from the origin through the point $(-1, 0)$ on the unit circle makes with the ray indicating the positive direction along the x-axis in the Cartesian plane. Hence

$$\cos\pi = -1, \ \sin\pi = 0.$$

Theorem 8.4 (Lindemann, 1882)**.** *The number π is transcendental.*

Proof. As in the proof of Theorem 8.2, suppose π is algebraic. Since we know that $i \in \mathbb{A}$ and $\mathbb{A}$ is a field, $\alpha = \pi i$ must also be algebraic. Let $d = \deg(\alpha)$ and let $\alpha = \alpha_1, \ldots, \alpha_d$ be conjugates of α. Let N be

the leading coefficient of $m_\alpha(x)$, then Lemma 6.6 implies that $N\alpha$ is an algebraic integer. By Euler's formula,

$$e^{\pi i} = -1,$$

and hence

$$(1 + e^{\alpha_1}) \cdots (1 + e^{\alpha_d}) = 0. \tag{36}$$

This product can be written as a sum of 2^d terms of the form e^θ, where

$$\theta = \varepsilon_1\alpha_1 + \cdots + \varepsilon_d\alpha_d,\ \varepsilon_k = 0, 1\ \forall\ 1 \le k \le d.$$

Suppose that exactly n of these numbers are nonzero, denote them $\beta_1, \ldots, \beta_n$. Let

$$h(x) = \prod_{\varepsilon_1=0}^{1} \cdots \prod_{\varepsilon_d=0}^{1} (x - (\varepsilon_1\alpha_1 + \cdots + \varepsilon_d\alpha_d))$$

and notice that $h(x)$ is symmetric in $\alpha_1, \ldots, \alpha_d$. Then Fact 8.3 implies that $h(x) \in \mathbb{Q}[x]$. Notice that the roots of $h(x)$ are $\beta_1, \ldots, \beta_d$ and 0, which has multiplicity $a = 2^d - n$. Clearing the denominators, this means that for some $C \in \mathbb{Z}$, $h(x) = Cx^a g(x)$, where $g(x) \in \mathbb{Z}[x]$ is the polynomial of degree n with roots $\beta_1, \ldots, \beta_n$. Now (36) implies that

$$(2^d - n)e^0 + e^{\beta_1} + \cdots + e^{\beta_n} = 0. \tag{37}$$

Let

$$f(x) = N^{np}x^{p-1}\prod_{k=1}^{n}(x - \beta_k)^p$$

for some large prime p, and let $I(t, f)$ for this choice of $f(x)$ be as in (29) above. Notice that degree of $f(x)$ is $m = (n + 1)p - 1$. Define

$$J := \sum_{k=1}^{n} I(\beta_k, f).$$

Then, by (31),

$$\begin{aligned} J &= \sum_{k=1}^{n}\left(e^{\beta_k}\sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m} f^{(j)}(\beta_k)\right) \\ &= \left(\sum_{k=1}^{n} e^{\beta_k}\right)\left(\sum_{j=0}^{m} f^{(j)}(0)\right) - \sum_{j=0}^{m}\sum_{k=1}^{n} f^{(j)}(\beta_k) \\ &= -(2^d - n)\left(\sum_{j=0}^{m} f^{(j)}(0)\right) - \sum_{j=0}^{m}\sum_{k=1}^{n} f^{(j)}(\beta_k), \end{aligned}$$

where the last equality follows by (37). Notice that $\sum_{k=1}^n f^{(j)}(\beta_k)$ is a symmetric polynomial in $N\beta_1, \dots, N\beta_n$ for each j. Furthermore, each $N\beta_k$ is a linear combination of algebraic integers $\alpha_1, \dots, \alpha_d$, and hence is an algebraic integer. Therefore, by Fact 8.3, for each $1 \leq j \leq m$, $\sum_{k=1}^n f^{(j)}(\beta_k) \in \mathbb{Z}$. Further, each β_k is a root of $f(x)$ of multiplicity p, which means that each derivative $f^{(j)}(\beta_k)$ vanishes for all $j < p$. For each $j \geq p$, $\sum_{k=1}^n f^{(j)}(\beta_k)$ is divisible by $p!$. Also,

$$f^{(p-1)}(0) = (p-1)!(-N)^{np}(\beta_1 \cdots \beta_n)^p,$$

which is not divisible by p provided that p is large (specifically, when $p > N\beta_1 \cdots \beta_n$). In addition, $f^{(j)}(0)$ is divisible by $p!$ for all $j \geq p$. Therefore K is divisible by $(p-1)!$, and hence

$$|J| \geq (p-1)!$$

On the other hand,

$$|J| \leq \sum_{k=1}^n |I(\beta_k, f)| \leq \sum_{k=1}^n |\beta_k| e^{|\beta_k|} F(|\beta_k|)$$

by (30) and $F(x)$ related to $f(x)$ is as above. Then we have

$$(p-1)! \leq |J| \leq AC^p$$

for some constants A and C. Taking p sufficiently large, we reach a contradiction. □

We are now ready to prove the Lindemann-Weierstrass Theorem.

Proof of Theorem 8.1. Towards a contradiction, suppose that there exist some algebraic numbers $d_1, \dots, d_s$, not all zero, such that

$$\sum_{k=1}^s d_k e^{\alpha_k} = 0. \tag{38}$$

Multiplying both sides by some N, by Lemma 6.6 we can assume that $d_1, \dots, d_s$ are algebraic integers. Let $K = \mathbb{Q}(d_1, \dots, d_s)$, $n = [K : \mathbb{Q}]$, and let $\sigma_k : K \to \mathbb{C}$ for $1 \leq k \leq n$ be embeddings of K. Notice that (38) implies that

$$\prod_{l=1}^n \left(\sum_{k=1}^s \sigma_l(d_k) e^{\alpha_k} \right) = 0. \tag{39}$$

The equation (39) can be written as

$$a_1 e^{\gamma_1} + \cdots + a_m e^{\gamma_m} = 0, \tag{40}$$

where each coefficient a_l is a sum of terms of the form $\sigma_m(d_k)$, which is invariant under each of the embeddings σ_m. Then Exercise 6.5 above

implies that $a_1, \ldots, a_m \in \mathbb{Q}$, and clearing the denominators, if necessary, we can assume that $a_1, \ldots, a_m \in \mathbb{Z}$. Further, we can assume that the set $\gamma_1, \ldots, \gamma_m$ contains all the conjugates of each of the γ_j's: if some of them are not there, they can always be included by choosing the corresponding a_l coefficient to be 0. Notice also that the exponents $\gamma_1, \ldots, \gamma_m$ are distinct algebraic numbers.

Let us write $\gamma_j^{(l)}$ for the l-th conjugate of γ_j. Let t be a real variable and for each l define the conjugate function

$$A_l(t) := \sum_{k=1}^{m} a_k e^{\gamma_k^{(l)} t}.$$

We will use the fact that when the γ_k's are all distinct, the functions $A_l(t)$ are not identically zero. Define

$$B(t) = \prod_l A_l(t) = \sum_{k=1}^{M} b_k e^{\beta_k t},$$

where the product is over all the conjugate functions $A_l(t)$. Notice that $B(1) = 0$ by our original assumption. Since $a_1, \ldots, a_m \in \mathbb{Z}$, the coefficients $b_1, \ldots, b_M$ are also rational integers, not all equal to zero. Since $\beta_1, \ldots, \beta_M$ are algebraic numbers, let $N \in \mathbb{Z}$ be such that $N\beta_1, \ldots, N\beta_M$ are algebraic integers. For each $1 \leq r \leq M$, define a polynomial

$$f_r(x) = \frac{N^{Mp}}{x - \beta_r} \prod_{k=1}^{M} (x - \beta_k)^p,$$

where $p \in \mathbb{Z}$ is a prime. Let

$$f(x) = \sum_{r=1}^{M} f_r(x),$$

then coefficients of $f(x)$ are symmetric polynomials in the algebraic integers $N\beta_1, \ldots, N\beta_M$. On the other hand, this set of numbers contains all of their algebraic conjugates, since $\beta_1, \ldots, \beta_M$ were generated by $\gamma_1, \ldots, \gamma_m$, which included all the algebraic conjugates. Hence coefficients of $f(x)$ must be in $\mathbb{Z}$ by Fact 8.3.

Define

$$J_r := \sum_{k=1}^{M} b_k I(\beta_k, f_r)$$

for each $1 \leq r \leq M$ and let $J := J_1 \cdots J_M$. Let $m := \deg(f_r) = Mp-1$, and notice that by (31),

$$\begin{aligned} J_r &= \sum_{k=1}^{M} b_k \left(e^{\beta_k} \sum_{j=0}^{m} f_r^{(j)}(0) - \sum_{j=0}^{m} f_r^{(j)}(\beta_k) \right) \\ &= -\sum_{k=1}^{M} b_k \sum_{j=0}^{m} f_r^{(j)}(\beta_k), \end{aligned}$$

where the last equality follows from the assumption that $B(1) = 0$. Arguing analogously to our proofs of Theorems 8.2 and 8.4, we conclude that J is an algebraic integer which is fixed by all the embeddings of the number field $\mathbb{Q}(\beta_1, \ldots, \beta_M)$, hence it must be in $\mathbb{Z}$. Further, J is divisible by $(p-1)!$, but not by p for a sufficiently large p. In the opposite direction, each $|J_r|$ can be bounded by c_r^p for a suitable positive real c_r, and hence $|J|$ can be bounded by C^p for some constant C. Therefore,

$$(p-1)! \leq |J| \leq C^p,$$

which leads to a contradiction for a large enough p. This completes the proof. □

9. Corollaries of the Lindemann-Weierstrass Theorem and some further results and conjectures

In this section we discuss some consequences of Theorem 8.1. First notice that transcendence of e and π follow easily from the Lindemann-Weierstrass Theorem. Although we have already proved these facts separately, it is still worthwhile to see them derived as consequences of the Lindemann-Weierstrass Theorem. We present these derivations here.

Corollary 9.1. *e is transcendental.*

Proof. Suppose $e \in \mathbb{A}$. Then there exists some nonzero polynomial

$$p(x) = \sum_{k=0}^{n} a_k x^k \in \mathbb{Z}[x]$$

such that

$$p(e) = \sum_{k=0}^{n} a_k e^k = 0.$$

This, however, clearly contradicts Theorem 8.1. □

Corollary 9.2. *π is transcendental.*

Proof. Suppose π is algebraic. We know also that $i \in \mathbb{A}$, and hence $i\pi \in \mathbb{A}$ since $\mathbb{A}$ is a field. By Euler's formula,

$$e^{i\pi} = \cos\pi + i\sin\pi = -1,$$

and hence we have

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

which clearly contradicts Theorem 8.1. □

Theorem 8.1 has many other important consequences. Here are some of them.

Corollary 9.3. *Let $0 \neq \alpha \in \mathbb{A}$. Then the numbers e^{α}, $\ln\alpha$, $\sin\alpha$ and $\cos\alpha$ are transcendental.*

Proof. Suppose e^{α} is algebraic, say $\gamma = e^{\alpha} \in \mathbb{A}$. Then

$$e^{\alpha} - \gamma e^{0} = 0,$$

which contradicts Theorem 8.1. Hence e^{α} is transcendental. Now assume that $\ln\alpha$ is algebraic, then $e^{\ln\alpha} = \alpha$ would have to be transcendental, which is a contradiction. Furthermore, Euler's formula (4) implies that

$$\sin\alpha = \frac{1}{2i}\left(e^{i\alpha} - e^{-i\alpha}\right),\ \cos\alpha = \frac{1}{2}\left(e^{i\alpha} + e^{-i\alpha}\right),$$

and so

$$e^{0}\sin\alpha - \frac{1}{2i}e^{i\alpha} + \frac{1}{2i}e^{-i\alpha} = 0,$$

$$e^{0}\cos\alpha - \frac{1}{2}e^{i\alpha} - \frac{1}{2}e^{-i\alpha} = 0.$$

Now Theorem 8.1 implies that $\sin\alpha$, $\cos\alpha$ cannot be algebraic. □

Corollary 9.4. *Let $\alpha_1, \ldots, \alpha_n \in \mathbb{A}$ be linearly independent over $\mathbb{Q}$. Then the numbers*

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

are algebraically independent.

Proof. Suppose that $e^{\alpha_1}, \ldots, e^{\alpha_n}$ are algebraically dependent, then there exists some non-constant polynomial

$$p(x_1, \ldots, x_n) \in \mathbb{Q}[x_1, \ldots, x_n]$$

such that

$$p(e^{\alpha_1}, \ldots, e^{\alpha_n}) = \sum_{i_1,\ldots,i_n} a_{i_1,\ldots,i_n} e^{i_1\alpha_1 + \cdots + i_n\alpha_n} = 0,$$

where the coefficients $a_{i_1,\dots,i_n}$ are rational numbers, not all zero. Then Theorem 8.1 implies that the exponents

$$i_1\alpha_1 + \cdots + i_n\alpha_n$$

cannot be all distinct. Hence there exist some two distinct families of indices $i_1, \dots, i_n$ and $j_1, \dots, j_n$ such that

$$i_1\alpha_1 + \cdots + i_n\alpha_n = j_1\alpha_1 + \cdots + j_n\alpha_n,$$

in other words

$$\sum_{k=1}^{n} c_k\alpha_k = 0,$$

where not all of $c_k := i_k - j_k \in \mathbb{Z}$ are equal to zero. This contradicts the assumption that $\alpha_1, \dots, \alpha_n$ are linearly independent over $\mathbb{Q}$. □

In fact, it is easy to see that Corollary 9.4 is equivalent to the Lindemann-Weierstrass Theorem, i.e. it is a convenient reformulation of the famous result. A substantial strengthening of Corollary 9.4 is arguably the most important open problem in transcendental number theory.

Conjecture 9.1 (Schanuel's Conjecture)**.** *Let $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ be linearly independent over $\mathbb{Q}$. Then*

$$\operatorname{trdeg}(\mathbb{Q}(\alpha_1, \dots, \alpha_n, e^{\alpha_1}, \dots, e^{\alpha_n})) \geq n.$$

We now discuss some of the many remarkable implications of this conjecture. First we mention (a weak form of) the famous theorem of Alan Baker (1966) on linear independence of logarithms of algebraic numbers, for which he received a Fields Medal in 1970.

Theorem 9.5 (Baker's Theorem, 1966)**.** *Let*

$$\Lambda = \left\{ \ell \in \mathbb{C} : e^{\ell} \in \mathbb{A} \right\}.$$

If $\ell_1, \dots, \ell_n \in \Lambda$ are linearly independent over $\mathbb{Q}$, then they are algebraically independent (and hence linearly independent over $\mathbb{A}$).

Proof. Baker's theorem has been proved unconditionally, however the proof is quite complicated. Here we will only show how this result follows from Schanuel's Conjecture. Indeed, Schanuel's Conjecture implies that

$$\operatorname{trdeg}(\mathbb{Q}(\ell_1, \dots, \ell_n, e^{\ell_1}, \dots, e^{\ell_n})) \geq n.$$

Since $\ell_1, \dots, \ell_n \in \Lambda$, we know that $e^{\ell_1}, \dots, e^{\ell_n} \in \mathbb{A}$, which implies that

$$\operatorname{trdeg}(\mathbb{Q}(\ell_1, \dots, \ell_n)) = \operatorname{trdeg}(\mathbb{Q}(\ell_1, \dots, \ell_n, e^{\ell_1}, \dots, e^{\ell_n})) \geq n.$$

Hence $\ell_1, \dots, \ell_n$ are algebraically independent. □

In fact, a strong version of Baker's Theorem establishes transcendence of any nonzero linear combination of $\ell_1, \ldots, \ell_n$ with algebraic coefficients, which, in its turn, is a generalization and strengthening of the celebrated Gelfond-Schneider Theorem, established independently in 1934 by Alexander Gelfond and Theodor Schneider. Their theorem presented a solution to Hilbert's 7th Problem.

Theorem 9.6 (Gelfond-Schneider Theorem, 1934)**.** *Let $a, b \in \mathbb{A}$ be such that $a \neq 0, 1$ and $b \notin \mathbb{Q}$. Then $a^b \in \mathbb{T}$.*

Furthermore, Schanuel's Conjecture implies algebraic independence of e and π, which is currently an open problem, as well as a wide variety of other known results and open problems in transcendental number theory. We conclude with yet another famous open problem, which would follow from Schanuel's Conjecture.

Conjecture 9.2 (Schneider's Four Exponentials Conjecture)**.** *Let x_1, x_2 and y_1, y_2 be pairs of complex numbers linearly independent over $\mathbb{Q}$. Then at least one of the four numbers $e^{x_j y_k}$ where $1 \leq j, k \leq 2$ is transcendental.*

If the linearly independent pair y_1, y_2 in the conjecture above is replaced with the linearly independent triple y_1, y_2, y_3, then the conjecture becomes a theorem, known as the Six Exponentials Theorem. It is our next big goal to prove this result.

10. Siegel's Lemma

We now develop an important tool, which will be used to prove another celebrated transcendence result, the Six Exponentials Theorem. This tool is Siegel's Lemma, the simplest version of which was originally observed by Axel Thue in 1909 and then formally proved by Carl Ludwig Siegel in 1929. Our presentation here partially follows [7] and [3]. Let

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \ldots & a_{MN} \end{pmatrix}$$

be an $M \times N$ matrix with integer entries and rank equal to $M < N$. Define

$$\Lambda = \{\boldsymbol{x} \in \mathbb{Z}^N : A\boldsymbol{x} = \mathbf{0}\}.$$

Theorem 10.1 (Siegel's Lemma, version 1)**.** *With notation as above, there exists $\mathbf{0} \neq \boldsymbol{x} \in \Lambda$ with*

$$|\boldsymbol{x}| < 2 + (N|A|)^{\frac{M}{N-M}}, \tag{41}$$

where $|\boldsymbol{x}| = \max\{|x_n| : 1 \leq n \leq N\}$, $|A| = \max\{|a_{mn}| : 1 \leq m \leq M,\ 1 \leq n \leq N\}$.

Proof. Let $H \in \mathbb{Z}_{>0}$, and let

$$C_H^N = \{\boldsymbol{x} \in \mathbb{R}^N : |\boldsymbol{x}| \leq H\}$$

be the cube centered at the origin in $\mathbb{R}^N$ with sidelength $2H$. Then

$$|C_H^N \cap \mathbb{Z}^N| = (2H+1)^N.$$

Let $T_A : \mathbb{R}^N \to \mathbb{R}^M$ be a linear map, given by $T_A(\boldsymbol{x}) = A\boldsymbol{x}$ for each $\boldsymbol{x} \in \mathbb{R}^N$. Notice that for every $\boldsymbol{x} \in C_H^N$,

$$|T_A(\boldsymbol{x})| \leq N|A|H,$$

i.e. T_A maps C_H^N into $C_{N|A|H}^M \subseteq \mathbb{R}^M$, since $\mathrm{rk}(A) = M$. Now

$$|C_{N|A|H}^M \cap \mathbb{Z}^M| = (2N|A|H+1)^M.$$

Now let us choose H to be a positive integer satisfying

$$(N|A|)^{\frac{M}{N-M}} \leq 2H < (N|A|)^{\frac{M}{N-M}} + 2.$$

Then

$$\begin{aligned} |C_H^N \cap \mathbb{Z}^N| &= (2H+1)^N = (2H+1)^M(2H+1)^{N-M} \\ &\geq (2H+1)^M(N|A|)^M > (2N|A|H+1)^M \\ &= |C_{N|A|H}^M \cap \mathbb{Z}^M|. \end{aligned}$$

This means that T_A cannot be mapping $C_H^N \cap \mathbb{Z}^N$ into $C_{N|A|H}^M \cap \mathbb{Z}^M$ in a one-to-one manner. Hence, there must exist $\boldsymbol{x} \neq \boldsymbol{y} \in C_H^N \cap \mathbb{Z}^N$ such that $T_A(\boldsymbol{x}) = T_A(\boldsymbol{y})$, i.e.

$$T_A(\boldsymbol{x} - \boldsymbol{y}) = 0,$$

and so $\boldsymbol{x} - \boldsymbol{y} \in \Lambda$. On the other hand,

$$|\boldsymbol{x} - \boldsymbol{y}| \leq |\boldsymbol{x}| + |\boldsymbol{y}| \leq 2H < (N|A|)^{\frac{M}{N-M}} + 2,$$

and this finishes the proof. □

Notice that the main underlying idea in the proof of Siegel's Lemma was the pigeon hole principle. It is remarkable that the exponent $\frac{M}{N-M}$ in the upper bound of (41) cannot be improved. To see this, let for instance $M = N-1$ and for a positive integer R consider the $(N-1)\times N$ matrix

$$A = \begin{pmatrix} R & -1 & 0 & \dots & 0 & 0 \\ 0 & R & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & R & -1 \end{pmatrix}.$$

Then $|A| = R$, and every nonzero integer solution of the system of linear equations $A\boldsymbol{x} = \mathbf{0}$ must have $x_N = R^{N-1}x_1$. Therefore, if

$$\Lambda = \{\boldsymbol{x} \in \mathbb{Z}^N : A\boldsymbol{x} = \mathbf{0}\},$$

and $\mathbf{0} \neq \boldsymbol{x} \in \Lambda$, then

$$|\boldsymbol{x}| \geq R^{N-1} = |A|^{\frac{M}{N-M}}.$$

Siegel's Lemma-type results have been proved in a variety of considerably more general settings by a number of authors, employing quite sophisticated machinery from number theory and arithmetic geometry. However, the original motivation for Siegel's Lemma came from Diophantine approximation and transcendental number theory.

For our use, we will also need a basic version of Siegel's Lemma over number fields. We start with some additional algebraic notation. Let K be a number field of degree d with embeddings $\sigma_1, \ldots, \sigma_d$. For each $\alpha \in K$, define its **height**

$$H(\alpha) := \max\{|\sigma_k(\alpha)| : 1 \leq k \leq d\}.$$

An important fact about the ring of integers $\mathcal{O}_K$ is that it is a free $\mathbb{Z}$-module of rank d. In other words, $\mathcal{O}_K$ has a $\mathbb{Z}$-basis: there exists a linearly independent collection $\omega_1, \ldots, \omega_d \in \mathcal{O}_K$ such that

$$\mathcal{O}_K = \left\{ \sum_{k=1}^{d} a_k \omega_k : a_1, \ldots, a_d \in \mathbb{Z} \right\}.$$

Define the corresponding $d \times d$ basis matrix $W := (\sigma_\ell(\omega_k))_{1 \leq \ell, k \leq d}$. It is a standard fact in algebraic number theory that W is nonsingular. With this notation and information in mind, we can now prove our next result.

Theorem 10.2 (Siegel's Lemma, version 2)**.** *Let K be a number field of degree d, and let $A = (\alpha_{ij})$ be an $M \times N$ matrix of rank $M < N$ with entries $\alpha_{ij} \in \mathcal{O}_K$. Define*

$$H(A) := \max\{H(\alpha_{ij}) : 1 \leq i \leq M, 1 \leq j \leq N\}.$$

There exists a solution $\mathbf{0} \neq \boldsymbol{x} = (x_1, \ldots, x_N) \in \mathcal{O}_K^N$ to the homogeneous linear system $A\boldsymbol{x} = \mathbf{0}$ with

$$\max_{1 \leq j \leq N} H(x_j) < B_K(M, N) H(A)^{\frac{M}{N-M}}, \tag{42}$$

where $B_K(M, N)$ is some constant depending only on M, N and the number field K.

Proof. Let $\omega_1, \dots, \omega_d \in \mathcal{O}_K$ be a $\mathbb{Z}$-basis for $\mathcal{O}_K$, as described above, and let W be the corresponding basis matrix. Then for each entry α_{ij} of our matrix A, there exist $a_{ijk} \in \mathbb{Z}$, $1 \le k \le d$, such that

$$\alpha_{ij} = \sum_{k=1}^{d} a_{ijk}\omega_k.$$

Applying embeddings $\sigma_1, \dots, \sigma_d$ to the above equation, we obtain

$$\sigma_\ell(\alpha_{ij}) = \sum_{k=1}^{d} a_{ijk}\sigma_\ell(\omega_k)$$

for each $1 \le \ell \le d$, and hence

$$\boldsymbol{\alpha}_{ij} := (\sigma_1(\alpha_{ij}), \dots, \sigma_d(\alpha_{ij}))^t = W(a_{ij1}, \dots, a_{ijd})^t.$$

Since W is invertible, we have

$$\boldsymbol{a}_{ij} := (a_{ij1}, \dots, a_{ijd})^t = W^{-1}\boldsymbol{\alpha}_{ij}.$$

If we write $v_{k\ell}$ for the entries of W^{-1}, then

$$a_{ijk} = \sum_{\ell=1}^{d} v_{k\ell}\sigma_\ell(\alpha_{ij}),$$

and so

$$|a_{ijk}| \le d \max_{1\le\ell\le d} |v_{k\ell}\sigma_\ell(\alpha_{ij})| \le dC_K H(A), \tag{43}$$

where C_K is a constant depending only on the number field K such that $C_K \ge \max_{1\le k,\ell\le d} |v_{k\ell}|$.

Now suppose $\boldsymbol{x} \in \mathcal{O}_K^N$ is a nontrivial solution of the system $A\boldsymbol{x} = \boldsymbol{0}$, and write

$$\boldsymbol{x} = \left(\sum_{\ell=1}^{d} b_{1\ell}\omega_\ell, \dots, \sum_{\ell=1}^{d} b_{N\ell}\omega_\ell \right) \tag{44}$$

for some $b_{j\ell} \in \mathbb{Z}$ for $1 \le j \le N$, $1 \le \ell \le d$. Then i-th entry of the vector $A\boldsymbol{x}$ is

$$\sum_{j=1}^{N}\sum_{\ell=1}^{d}\sum_{k=1}^{d} a_{ijk}b_{j\ell}\omega_k\omega_\ell = 0.$$

Since $\omega_k\omega_\ell \in \mathcal{O}_K$, it can also be expressed as a linear combination of ω_m's with $\mathbb{Z}$-coefficients:

$$\omega_k\omega_\ell = \sum_{m=1}^{d} c_{k\ell m}\omega_m$$

for each $1 \leq k, \ell \leq d$, and hence we have

$$\sum_{m=1}^{d} \sum_{j=1}^{N} \sum_{\ell=1}^{d} \sum_{k=1}^{d} a_{ijk} b_{j\ell} c_{k\ell m} \omega_m = 0.$$

Since $\omega_1, \ldots, \omega_d$ are linearly independent over $\mathbb{Z}$, all the coefficients in the above equations must be zero, and hence we have a system of Md homogeneous linear equations with integer coefficients in the Nd variables $b_{j\ell}$:

$$\sum_{j=1}^{N} \sum_{\ell=1}^{d} \sum_{m=1}^{d} a_{ijk} b_{j\ell} c_{k\ell m} = 0,$$

for all $1 \leq i \leq M$, $1 \leq m \leq d$. Applying Theorem 10.1 along with (43), we see that there exists a solution with

$$\max_{j,\ell} |b_{j\ell}| \leq 2 + (Nd^2 C_K H(A))^{\frac{Md}{Nd-Md}},$$

and hence, by (44),

$$\max_{1 \leq j \leq N} H(x_j) \leq d \left(2 + (Nd^2 C_K H(A))^{\frac{M}{N-M}}\right) \max_{1 \leq \ell \leq d} H(\omega_\ell).$$

Since the choice of $\omega_1, \ldots, \omega_\ell$ depends only on K, the conclusion of the theorem follows. □

Recall that for any $\beta \in K$, there exists $c \in \mathbb{N}$ such that $c\beta \in \mathcal{O}_K$. In fact, for any collection $\beta_1, \ldots, \beta_n \in K$, let us define their **common denominator** to be

$$D(\beta_1, \ldots, \beta_n) = \min\{c \in \mathbb{N} : c\beta_k \in \mathcal{O}_K \ \forall \ 1 \leq k \leq n\}.$$

For an $M \times N$ matrix A with entries in K, we will write $D(A)$ for the common denominator of all of its entries, i.e.,

$$D(A) = D(\alpha_{ij} : 1 \leq i \leq M, 1 \leq j \leq N).$$

With this notation in mind, we have one more version of Siegel's lemma.

Corollary 10.3 (Siegel's Lemma, version 3)**.** *Let K be a number field of degree d, and let $A = (\alpha_{ij})$ be an $M \times N$ matrix of rank $M < N$ with entries $\alpha_{ij} \in K$. There exists a solution $\mathbf{0} \neq \boldsymbol{x} = (x_1, \ldots, x_N) \in \mathcal{O}_K^N$ to the homogeneous linear system $A\boldsymbol{x} = \mathbf{0}$ with*

$$\max_{1 \leq j \leq N} H(x_j) < B_K(M, N)(D(A)H(A))^{\frac{M}{N-M}}, \tag{45}$$

where $B_K(M, N)$ is the same constant as in Theorem 10.2 above.

Proof. Let $A' = D(A)A$, then A' is an $M \times N$ matrix with entries in $\mathcal{O}_K$, and $A\boldsymbol{x} = \mathbf{0}$ if and only $A'\boldsymbol{x} = \mathbf{0}$. Then apply Theorem 10.2 to the system $A'\boldsymbol{x} = \mathbf{0}$ while keeping in mind that $H(A') = D(A)H(A)$. □

11. The Six Exponentials Theorem

In this section we use Siegel's Lemma and Maximum Modulus Principle to prove the Six Exponentials Theorem. Our presentation here follows [3]. We start with another necessary piece of algebraic notation. Let K be a number field of degree d over $\mathbb{Q}$, and let $\sigma_1, \dots, \sigma_d$ be the embeddings of K into $\mathbb{C}$. For every $\alpha \in K$, define the **norm** of α over K to be

$$\mathrm{N}_K(\alpha) = \prod_{k=1}^{d} \sigma_k(\alpha),$$

and we write $\mathrm{N}(\alpha)$ for $\mathrm{N}_{\mathbb{Q}(\alpha)}(\alpha)$. It is not difficult to observe that

$$\mathrm{N}_K(\alpha) = \mathrm{N}(\alpha)^{[K:\mathbb{Q}(\alpha)]}.$$

Notice also that $\mathrm{N}(\alpha)$ is precisely the free coefficient of the minimal polynomial of α over $\mathbb{Q}$, and hence is a rational number. If $\alpha \in \mathcal{O}_K$, then the minimal polynomial of α over $\mathbb{Q}$ is equal to $m_\alpha(x)$, and hence $\mathrm{N}(\alpha) \in \mathbb{Z}$. This in particular implies that for every $\alpha \in \mathcal{O}_K$,

$$1 \leq |\mathrm{N}_K(\alpha)| = |\mathrm{N}(\alpha)|^{[K:\mathbb{Q}(\alpha)]} \leq |\mathrm{N}(\alpha)|^d \leq H(\alpha)^{d-1}|\alpha|, \tag{46}$$

since one of the embeddings $\sigma_1, \dots, \sigma_d$ is the identity map.

Theorem 11.1 (The Six Exponentials Theorem)**.** *Let $x_1, x_2 \in \mathbb{C}$ be linearly independent over $\mathbb{Q}$. Let $y_1, y_2, y_3 \in \mathbb{C}$ also be linearly independent over $\mathbb{Q}$. Then at least one of the six numbers $e^{x_i y_j}$ where $1 \leq i \leq 2$, $1 \leq j \leq 3$ is transcendental.*

Proof. Suppose that $e^{x_j y_k} \in \mathbb{A}$ for all $1 \leq j \leq 2$, $1 \leq k \leq 3$, and let K be a number field containing all of these numbers. Let $r \in \mathbb{N}$, $a_{ij} \in \mathcal{O}_K$ for all $1 \leq i, j \leq r$, and define

$$F(z) = \sum_{i=1}^{r} \sum_{j=1}^{r} a_{ij} e^{(ix_1 + jx_2)z} \tag{47}$$

for a variable $z \in \mathbb{C}$. Let $n \in \mathbb{N}$ and let $k_1, k_2, k_3 \in \mathbb{N}$ range between 1 and n. Then

$$F\left(\sum_{m=1}^{3} k_m y_m\right) = \sum_{i=1}^{r} \sum_{j=1}^{r} a_{ij} \exp\left((ix_1 + jx_2)(k_1 y_1 + k_2 y_2 + k_3 y_3)\right).$$

Since each $\exp\left((ix_1 + jx_2)(k_1 y_1 + k_2 y_2 + k_3 y_3)\right)$ is algebraic, setting each

$$F\left(\sum_{m=1}^{3} k_m y_m\right) = 0 \tag{48}$$

yields a system of n^3 equations with algebraic coefficients in the r^2 variables a_{ij}. We want to apply Siegel's Lemma to this system to obtain a small-height solution vector; for this we need $r^2 > n^3$. Let D be the common denominator of the six exponentials

$$\{e^{x_i y_j} : 1 \le i \le 2, 1 \le j \le 3\},$$

then the common denominator of the coefficients of the system (48) is bounded above by D^{6rn}, and heights of these coefficients are bounded above by $e^{c_0 rn}$ for some constant c_0. Now Theorem 10.3 guarantees that (48) has a solution vector with coordinates $a_{ij} \in \mathcal{O}_K$, not all zero, such that

$$\max_{i,j} H(a_{ij}) \le B_K(n^3, r^2) \left(D^{6rn} e^{c_0 rn}\right)^{\frac{n^3}{r^2 - n^3}}.$$

Then, choosing $r = 8n^{3/2}$, we ensure that

$$\max_{i,j} H(a_{ij}) \le B_K(n^3, 8n^{3/2}) \left(D^{48n^{5/2}} e^{8c_0 n^{5/2}}\right)^{\frac{1}{63}} \le e^{c_1 n^{5/2}} \tag{49}$$

for some appropriately chosen constant c_1. Then let $a_{ij} \in \mathcal{O}_K$ be a solution to (48) with $r = 8n^{3/2}$ satisfying (49), and let $F(z)$ be as in (47) for this choice of a_{ij}'s. Notice that $F(z)$ is not identically zero, since x_1, x_2 are linearly independent over $\mathbb{Q}$. Also notice that the set

$$S = \{k_1 y_1 + k_2 y_2 + k_3 y_3 : k_1, k_2, k_3 \in \mathbb{N}\}$$

is not discrete, since the numbers y_1, y_2, y_3 are linearly independent over $\mathbb{Q}$. Since $F(z)$ is not identically zero, it cannot vanish on a non-discrete set, and hence there must exist elements of S on which F is not zero. Let

$$s = \max\{t \in \mathbb{N} : F(k_1 y_1 + k_2 y_2 + k_3 y_3) = 0 \ \forall\ 1 \le k_i \le t\}.$$

Clearly, $s \ge n$. Define

$$w = k_1 y_1 + k_2 y_2 + k_3 y_3$$

with some $k_i = s + 1$ and all $1 \le k_i \le s + 1$ be such that $F(w) \ne 0$. Using (49), we can obtain an estimate on the height of $F(w)$:

$$H(F(w)) \le C_0^{n^{5/2} + (s+1)r} \le C_1^{s^{5/2}}$$

for some positive constants C_0, C_1. Observe also that $D^{6r(s+1)} F(w)$ is an algebraic integer. Then, by (46) we have:

$$1 \le \mathrm{N}_K(D^{6r(s+1)} F(w)) \le H(D^{6r(s+1)} F(w))^{[K:\mathbb{Q}]-1} |D^{6r(s+1)} F(w)|,$$

and so

$$|F(w)| \ge D^{-6r(s+1)[K:\mathbb{Q}]} H(F(w))^{-([K:\mathbb{Q}]-1)} \ge C_2^{-s^{5/2}}, \tag{50}$$

where C_2 is another constant independent of s.

Our next goal will be to arrive at a contradiction with (50) by obtaining an incompatible estimate for $|F(w)|$ from above. Notice that

$$F(w) = \lim_{z \to w} \left\{ F(z) \prod_{1 \leq k_1, k_2, k_3 \leq s} \left(\frac{w - (k_1 y_1 + k_2 y_2 + k_3 y_3)}{z - (k_1 y_1 + k_2 y_2 + k_3 y_3)} \right) \right\}. \tag{51}$$

The right hand side of the above identity is a holomorphic function that has s^3 factors in the product. Let R be a real number such that $|w| < R$ and

$$|z - (k_1 y_1 + k2 y_2 + k_3 y_3)| \geq R/2$$

for all z on the circle of radius R. Applying the Maximum Modulus Principle (specifically, Corollary B.2) to the right hand side of (51) on the disk of radius R, we conclude that it assumes its maximum value on the boundary, i.e. on the circle of radius R, and hence

$$|F(w)| \leq |F|_R (C_3 s/R)^{s^3},$$

for some constant C_3, where $|F|_R$, the maximum of $F(z)$ on the circle of radius R, can be estimated as follows:

$$|F|_R \leq C_4 e^{c_1 n^{5/2} + c_2 r R} r^2,$$

for some constants C_4, c_2, and with c_1 as above. Taking $R = s^{3/2}$, recalling that $r^2 = 64n^3$, and combining these inequalities yields:

$$|F(w)| \leq 64 n^3 C_4 e^{c_1 n^{5/2} + c_2 8 (ns)^{3/2}} \left(\frac{C_3}{\sqrt{s}} \right)^{s^3}.$$

Since $s \geq n$, taking n large will cause a contradiction with (50), hence completing the proof. □

Appendix A. Some properties of abelian groups

Here we briefly discuss some properties of abelian groups, in particular outlining a proof of the fact that any subgroup of a finitely generated abelian group is finitely generated. Throughout this section, we will mostly deal with a finitely generated abelian group G, written additively with $\mathbf{0}$ denoting the identity element and $n\boldsymbol{x}$, for $n \in \mathbb{Z}$ and $\boldsymbol{x} \in G$, denoting the n-th power of the element $\boldsymbol{x}$. A collection of elements $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ in an abelian group G is called **linearly independent** if whenever

$$n_1 \boldsymbol{x}_1 + \cdots + n_k \boldsymbol{x}_k = \mathbf{0}$$

for some $n_1, \ldots, n_k \in \mathbb{Z}$, then $n_1 = \cdots = n_k = 0$. A linearly independent generating set for an abelian group G is called a **basis**. An abelian group G is called **free** if it has a basis.

Exercise A.1. *Suppose that G is a free abelian group. Prove that the following property holds: whenever $n\boldsymbol{x} = \mathbf{0}$ for some $n \in \mathbb{Z}$ and $\boldsymbol{x} \in G$, then either $n = 0$ or $\boldsymbol{x} = \mathbf{0}$.*

The most common example of a finitely generated free abelian group is the group

$$\mathbb{Z}^k = \{\boldsymbol{x} = (x_1, \ldots, x_k) : x_1, \ldots, x_k \in \mathbb{Z}\}$$

under component-wise addition, where $k \in \mathbb{N}$. It is easy to notice that the set of vectors $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$, where $\boldsymbol{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{Z}^k$ with 1 in i-th position forms a basis for $\mathbb{Z}^k$: it is called the **standard basis** for $\mathbb{Z}^k$ and these vectors are called the **standard basis vectors**. In fact, it turns out that $\mathbb{Z}^k$ is the *only* example of a finitely generated free abelian group, up to isomorphism.

Lemma A.1. *Let G be a finitely generated free abelian group. Then $G \cong \mathbb{Z}^k$ for some $k \in \mathbb{N}$.*

Proof. Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ be a basis for G, then

$$G = \left\{ \sum_{i=1}^{k} n_i \boldsymbol{x}_i : n_1, \ldots, n_k \in \mathbb{Z} \right\}.$$

Define a map $\varphi : G \to \mathbb{Z}^k$, given by

$$\varphi\left(\sum_{i=1}^{k} n_i \boldsymbol{x}_i \right) = \sum_{i=1}^{k} n_i \boldsymbol{e}_i.$$

We leave it to the reader to check that this is a group isomorphism. □

Exercise A.2. *Suppose that $1 \leq k < m$. Prove that $\mathbb{Z}^k \ncong \mathbb{Z}^m$.*

Corollary A.2. *Let G be a finitely generated free abelian group. Then every basis in G has the same cardinality. This common cardinality is called the* **rank** *of G.*

Proof. Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ and $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ be two different bases for G. Then by the argument in the proof of Lemma A.1, $G \cong \mathbb{Z}^k$ and $G \cong \mathbb{Z}^m$. Now Exercise A.2 implies that $\mathbb{Z}^k \ncong \mathbb{Z}^m$ unless $k = m$. Recall from Exercise 2.4 that isomorphism is an equivalence relation on groups. Thus, since $G \cong \mathbb{Z}^k$ and $G \cong \mathbb{Z}^m$, we must have $\mathbb{Z}^k \cong \mathbb{Z}^m$. Hence $k = m$. □

Fact A.3. *Let H be a subgroup of a finitely generated free abelian group G of rank k. The H is also free abelian of rank $\leq k$.*

We do not present the proof of this fact here. A standard proof is along the lines of linear algebra, using Smith normal form for matrices, which constructs a basis for a subgroup starting with a basis for the group.

We now give some additional basic algebraic notation without proofs. We refer the reader to [1] for details. If G is an abelian group and H is a subgroup of G, then a **coset** of H in G is a set $\boldsymbol{x} + H$ where $\boldsymbol{x} \in G$. The group G can be represented as a disjoint union of all cosets of H in G. We write G/H for the set of such cosets, which is a group under the operation of addition of cosets:

$$(\boldsymbol{x} + H) + (\boldsymbol{y} + H) = (\boldsymbol{x} + \boldsymbol{y}) + H.$$

G/H is called the **quotient group** of G modulo H. The identity element in this group is the trivial coset $\mathbf{0} + H = H = \boldsymbol{x} + H$ for every $\boldsymbol{x} \in H$, and inverse of $\boldsymbol{y} + H$ is $-\boldsymbol{y} + H$ for every $\boldsymbol{y} \in G$. The **order** of G/H, i.e. its cardinality as a set (could be infinite) is called the **index** of H in G, and denoted by $|G : H|$. Suppose that G and E are two abelian groups and $\varphi : G \to E$ is a group homomorphism between them. Recall that $\mathrm{Ker}(\varphi)$ is a subgroup of G and $\varphi(G)$ is a subgroup of E. The First Isomorphism Theorem states that

$$G/\mathrm{Ker}(\varphi) \cong \varphi(G). \tag{52}$$

Finally, notice that a finitely generated group can only be isomorphic to another finitely generated group. We are now ready for the main result of this section.

Theorem A.4. *Let G be a finitely generated abelian group, and let H be a subgroup of G. Then H is finitely generated.*

Proof. Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ be a generating set for G, then every element $\boldsymbol{y} \in G$ is expressible as

$$\boldsymbol{y} = \sum_{i=1}^{k} n_i \boldsymbol{x}_i$$

for some $n_1, \ldots, n_k \in \mathbb{Z}$. Define a map $\varphi : \mathbb{Z}^k \to G$, given by

$$\varphi\left(\sum_{i=1}^{k} n_i \boldsymbol{e}_i\right) = \sum_{i=1}^{k} n_i \boldsymbol{x}_i.$$

We leave it to the reader to check that this is a group homomorphism. Let $K = \mathrm{Ker}(\varphi)$, then K is a subgroup of $\mathbb{Z}^k$, hence it is free abelian of rank $\ell \leq k$. Now H be a subgroup of G, then there exists a subgroup M of $\mathbb{Z}^k$ such that $\varphi(M) = H$; in other words, M is the pre-image

of H in $\mathbb{Z}^k$ under φ. Then M is also free abelian of rank $m \leq k$. Furthermore, M contains K: indeed, for every $\boldsymbol{x} \in K$, $\varphi(\boldsymbol{x}) = \mathbf{0} \in H$, hence $\boldsymbol{x} \in M$. Therefore $\ell \leq m$, and by (52),

$$H \cong M/K,$$

hence we only need to show that M/K is finitely generated.

By Lemma A.1 we know that $M \cong \mathbb{Z}^m$ and $K \cong \mathbb{Z}^\ell$. By viewing vectors in $\mathbb{Z}^\ell$ as m-tuples with last $m - \ell$ coordinates equal to 0, we can think of $\mathbb{Z}^\ell$ being contained in $\mathbb{Z}^m$. Hence we only need to show that $\mathbb{Z}^m/\mathbb{Z}^\ell$ is finitely generated. If $m = \ell$, then $\mathbb{Z}^m = \mathbb{Z}^\ell$ and so $\mathbb{Z}^m/\mathbb{Z}^\ell \cong \{\mathbf{0}\}$, the trivial group. Then assume that $m > \ell$. Considering the standard basis $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ for $\mathbb{Z}^m$, we can view $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_\ell$ as the standard basis for $\mathbb{Z}^\ell$ under its embedding into $\mathbb{Z}^m$. Then $\mathbb{Z}^m/\mathbb{Z}^\ell$ is isomorphic to $\mathbb{Z}^{m-\ell}$ via the map sending an element $\sum_{i=1}^m n_i \boldsymbol{e}_i + \mathbb{Z}^\ell$ in $\mathbb{Z}^m/\mathbb{Z}^\ell$ to $\sum_{i=m-\ell+1}^m n_i \boldsymbol{e}_i$ in $\mathbb{Z}^{m-\ell}$ (this is easily checked to be a group isomorphism). Now, $\mathbb{Z}^{m-\ell}$ is finitely generated, and hence we are done. □

Appendix B. Maximum Modulus Principle and Fundamental Theorem of Algebra

Our main goal here is to prove the Fundamental Theorem of Algebra. For this, we will use the Maximum Modulus Principle. We first need some basic notation from complex analysis. A **region** in $\mathbb{C}$ is a subset R of $\mathbb{C}$, which is open and connected. A function $f(z)$ on a region R is called analytic if for any $z_0 \in R$,

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n,$$

where $a_n \in \mathbb{C}$ for every $n \geq 0$ and the series is convergent to $f(z)$ in an an open neighborhood of z_0. It is a well-known fact that every holomorphic (i.e., complex-differentiable) function is analytic, and vice versa.

Theorem B.1 (Maximum Modulus Principle)**.** *Suppose $f(z)$ is a non-constant analytic function in a region R. Then the real-valued function $|f(z)|$ does not attain its maximum in R. In other words, if for some $z_0 \in R$, $|f(z)| \leq |f(z_0)|$ for all points $z \in R$, then $f(z)$ is constant on R.*

A proof of this theorem can be found in any book on complex analysis, for instance [5]. Here is an immediate consequence of Theorem B.1, which is very useful in applications.

Corollary B.2. *Let*

$$D_r = \{z \in \mathbb{C} : |z| \leq r\}$$

be the closed disk of radius r and let $f(z)$ be a continuous function on D_r, which is analytic on the open disk

$$D_r^o = \{z \in \mathbb{C} : |z| < r\}.$$

Then $f(z)$ assumes its maximum value on D_r on its boundary

$$\partial D_r = \{z \in \mathbb{C} : |z| = r\} = D_r \setminus D_r^o.$$

Proof. Since $f(z)$ is continuous and D_r is closed and bounded, $f(z)$ must have a maximum on D_r. On the other hand, since the open disk D_r^o is a region in $\mathbb{C}$, by Theorem B.1 $f(z)$ cannot have a maximum on D_r^o. Thus it must be assumed on the boundary. □

We will now derive an important consequence of this fundamental principle.

Theorem B.3 (Fundamental Theorem of Algebra, stated as Theorem 1.1 above)**.** *Any polynomial $p(x) \in \mathbb{C}[x]$ of degree n has precisely n roots in $\mathbb{C}$, counted with multiplicity. In other words, the field of complex numbers $\mathbb{C}$ is algebraically closed.*

Proof. Notice that it is sufficient to prove that any polynomial $p(x)$ of degree $n \geq 1$ has at least one root in $\mathbb{C}$. Suppose not, say $p(x) \in \mathbb{C}[x]$ of degree $n \geq 1$ has no complex roots. This means that $1/p(x)$ is an analytic (holomorphic) function. Notice that $1/p(x)$ tends to zero as $|x|$ tends to infinity. This means that for any $\alpha \in \mathbb{C}$ there exists an $r \in \mathbb{R}$ such that

$$1/|p(x)| < 1/|p(\alpha)|$$

for all $x \in \mathbb{C}$ with $|x| \geq r$. Now pick r large enough so that $|\alpha| < r$, and let D_r be the closed disk of radius r, as in Corollary B.2 above. Then $\alpha \in D_r$ and, since $1/|p(x)|$ is continuous, it assumes its maximum on D_r, specifically on its boundary, by Corollary B.2. Then there exists $\beta \in \partial D_r$ such that

$$1/|p(x)| \leq 1/|p(\beta)| \ \forall \ x \in D_r.$$

Now pick $t > r$ and D_t^o be the open disk of radius t. Then $D_r \subsetneq D_t^o$, and for all $x \in D_t^o \setminus D_r$,

$$1/|p(x)| < 1/|p(\alpha)| \leq 1/|p(\beta)|.$$

Hence $1/|p(x)|$ assumes its maximum on D_t^o at $x = \beta$. Since $1/p(x)$ is not a constant function (degree of $p(x)$ is > 0) and D_t^o is a region (it is open and connected), this violates the Maximum Modulus Principle. Hence $p(x)$ must have a zero in $\mathbb{C}$. □

References

[1] D. S. Dummit and R. M. Foote. *Abstract Algebra*. Wiley, 3rd edition, 2003.
[2] B. Edixhoven and J.-H. Evertse (Eds.). *Diophantine Approximation and Abelian Varieties*. Springer-Verlag, 1993.
[3] M. R. Murty and P. Rath. *Transcendental Numbers*. Springer, New York, 2014.
[4] K.F. Roth. Rational approximations to algebraic numbers. *Mathematika*, 2:1–20, 1955.
[5] W. Rudin. *Real and Complex Analysis*. McGraw-Hill Book Co., New York, 3rd edition, 1987.
[6] W. M. Schmidt. *Diophantine Approximation*. Springer-Verlag, 1980.
[7] W. M. Schmidt. *Diophantine Approximations and Diophantine Equations*. Springer-Verlag, 1991.

Department of Mathematical Sciences, Claremont McKenna College, 850 Columbia Avenue, Claremont, CA 91711
E-mail address: lenny@cmc.edu