The purpose of these notes is to prove Lindemann's Theorem. The proof is adapted from Jacobson's book *Algebra I*, but I simplified some of the assumptions in order to make the proof easier. Also, I improved the proof somewhat.

1 The Main Result

Here is Lindemann's Theorem.

Theorem 1.1 *Let $u \neq 0$ be an algebraic number. Then e^u is transcendental.*

Theorem 1.1, applied to $u = 1$, immediately proves that e is transcendental. Here is another application.

Theorem 1.2 *π is transcendental.*

Proof: Suppose π is algebraic. Since $2i$ is also algebraic, $2\pi i$ is algebraic. But $e^{2\pi i} = 1$ and 1 is not transcendental. This contradicts Theorem 1.1. ♠

Rather than prove Theorem 1.1 directly, we'll prove a related result.

Theorem 1.3 *Let $\boldsymbol{F}$ denote the field of algebraic numbers. Suppose that $u_1, ..., u_k$ are distinct algebraic integers. Then the numbers $e^{u_1}, ..., e^{u_k}$ are linearly independent over $\boldsymbol{F}$.*

Let's first see how Theorem 1.3 implies Theorem 1.1. Suppose that u is some algebraic number and e^u is algebraic. Then $e^{ku} = (e^u)^k$ is algebraic for every integer k. We can choose k so that ku is an algebraic integer. Hence, without loss of generality, we can assume that u is an algebraic integer and $e^u = v$ is an algebraic number. But then we set $u_1 = 0$ and $u_2 = u$ and $v_1 = -v$ and $v_2 = 1$. We have

$$v_1 e^{u_1} + v_2 e^{u_2} = -v + e^u = 0.$$

This contradicts the fact that e^{u_1} and e^{u_2} are linearly independent over $\boldsymbol{F}$.

Remark: Jacobson proves Theorem 1.3 under the weaker assumption that $u_1, ..., u_k$ are just algebraic numbers and not necessarily algebraic integers. The stronger result in Jacobson is equivalent to the Lindemann-Weierstrass Theorem, a generalization of Lindemann's Theorem.

2 Outline of the Proof

Say that a *bad sum* is a nontrivial sum of the form

$$v_1 e^{u_1} + ... + v_n e^{u_n} = 0, \tag{1}$$

where $v_1, ..., v_n$ are algebraic numbers and $u_1, ..., u_n$ are algebraic integers. the content of Theorem 1.3 is that there are no bad sums. We will assume that there is a bad sum and derive a contradiction. Here is our first main result.

Lemma 2.1 (Step 1) *Suppose that there exists a bad sum. Then there exists a bad sum where* $v_1, ..., v_n \in \boldsymbol{Z}$.

Note that the n in Step 1 might be different from the n in Equation 1. The same thing is true for the remaining steps. We are just using n to denote a finite sum.

Suppose then that we have a bad sum in which all the v's are integers. We can find a normal extension K of $\boldsymbol{Q}$ such that $u_1, ..., u_n \in K$. Let $G = G(K, \boldsymbol{Q})$ denote the Galois group of K over $\boldsymbol{Q}$.

Lemma 2.2 (Step 2) *Suppose that there exists a bad sum as in Step 1. Then there exists a bad sum of the form* $v_1 T_1 + ... + v_n T_n$, *where*

$$T_k = \sum_{\phi \in G} e^{\phi(u_k)}, \tag{2}$$

and $v_1, ..., v_n \in \boldsymbol{Z}$.

Finally, here is the last of the algebraic steps.

Lemma 2.3 (Step 3) *Suppose that there exists a bad sum as in Step 2. Then we have a bad sum of the form*

$$v_0 + v_1 T_1 + ... + v_n T_n, \tag{3}$$

where $v_0 \in \boldsymbol{Z} - \{0\}$ *and the remaining terms are as in Step 2.*

We will work with the sum in Equation 3.

Lemma 2.4 (Step 4) *For any sufficiently large prime p, there is an integer $N \in \boldsymbol{Z} - p\boldsymbol{Z}$ and polynomial $F(x) \in \boldsymbol{Z}[x]$ such that*

$$|Ne^{\phi(u_i)} - F(\phi(u_i))| < 1/p,$$

for all u_i and all $\phi \in G$. Also, the coefficients of F are all divisible by p.

Now let's put the steps together. We pick some large prime p and multiply Equation 3 by N:

$$X = v_0 N + v_1 N T_1 + ... + v_n N T_n = 0. \tag{4}$$

Consider the related sum

$$Y = v_0 N + v_1 \sum_{\phi \in G} F(\phi(u_1)) + ... + v_n \sum_{\phi \in G} F(\phi(u_n)). \tag{5}$$

From Step 4, we have

$$|v_k N T_k - v_k \sum_{\phi \in G} F(\phi(u_k))| < \frac{M}{p}; \qquad M = \max(|v_1|, ..., |v_n|). \tag{6}$$

Subtracting X from Y term by term and using Equation 6, we get

$$|Y| = |Y - X| < \frac{nM}{p} < 1. \tag{7}$$

The last inequality holds when we pick p large enough. But each term

$$\sum_{\phi \in G} \frac{F(\phi(u_k))}{p} \tag{8}$$

is an algebraic integer that is fixed by all $\phi \in G$. Hence, this sum lies in $\boldsymbol{Q}$. The only algebraic integers in $\boldsymbol{Q}$ are ordinary integers. Hence the sum in Equation 8 is an integer! Therefore, all the summands of Y, after the first one, lie in $p\boldsymbol{Z}$. But the first summand of Y lies in $\boldsymbol{Z} - p\boldsymbol{Z}$ provided we take p large enough. Hence $Y \in \boldsymbol{Z} - p\boldsymbol{Z}$. In particular $|Y| \geq 1$. For p sufficiently large, Equation 7 says that $|Y| < 1$. This is a contradiction. Hence there are no bad sums.

This completes the proof, modulo the four steps above. Now we prove the four steps.

3 A Certain Ring

Let K be a finite normal extension of $\boldsymbol{Q}$. Let O_K be the ring of algebraic integers in K. We define a ring R, as follows. An element of R is a map $f : O_K \to K$ which is nonzero only at finitely many values. Given two elements $f_1, f_2 \in R$, we define $g = f_1 + f_2$ by the rule $g(a) = f_1(a) + f_2(a)$. Again, g is only nonzero at finitely many values, so $g \in R$. This makes R into an abelian group. We define $h = fg$ by the rule that

$$h(a) = \sum_{s+t=a} f(s)g(t). \tag{9}$$

Again h only takes on finitely many nonzero values. It is an easy but tedious exercise to check that these operations make R into a ring. For instance, the multiplication rule is associative and $(fg)h$ and $f(gh)$ both map a to

$$\sum_{r+s+t=a} f(r)g(s)h(t).$$

Here is a less obvious property.

Lemma 3.1 *R is an integral domain.*

Proof: This works for roughly the same reason that polynomial rings over fields are integral domains: The highest degree terms multiply together to get a result that isn't cancelled by anything else. We don't have the notion of degree here, but we can do something similar. We define an ordering on $\boldsymbol{C}$, as follows: $x_1 + iy_1 > x_2 + iy_2$ if and only if one of two things holds.

- $x_1 > x_2$.
- $x_1 = x_2$ and $y_1 > y_2$.

Our ordering has the following property: If $z_1 > z_1'$ and $z_2 > z_2'$ then $z_1 + z_2 > z_1' + z_2'$. Given nonzero $f, g \in R$, there are largest elements $s, t \in K$ such that $f(s) \neq 0$ and $g(t) \neq 0$. But then $fg(s+t) = f(s)g(t) \neq 0$. The point is that all other sums in Equation 9 are less than $s + t$ in the order. ♠

There is a map $\Psi : R \to \boldsymbol{C}$ given by

$$\Psi(f) = \sum_{a \in K} f(a)e^a. \tag{10}$$

This is a finite sum, so $\Psi(f)$ is a well-defined number.

Lemma 3.2 *Ψ is a ring homomorphism.*

Proof: It is pretty obvious that Ψ is a group homomorphism. We compute

$$\Psi(fg) = \sum_{a\in K}(fg)(a)e^a =$$
$$\sum_{a\in K}\sum_{s+t=a} f(s)g(t)e^{s+t} =$$
$$\sum_{s,t\in K} f(s)g(t)e^{s+t} =$$
$$\sum_{s,t\in K} f(s)g(t)e^{s}e^{t} =$$
$$\Big(\sum_{s\in K} f(s)e^s\Big)\Big(\sum_{t\in K} g(t)e^t\Big). =$$
$$\Psi(f)\Psi(g).$$

The main point here is that $e^{s+t} = e^s e^t$. ♠

There are two more pieces of structure. Let $G = G(K, \boldsymbol{Q})$ be the Galois group of K over $\boldsymbol{Q}$. For any $\phi \in G$, the composition $\phi \circ f$ is also an element of R. This map has the action $\phi \circ f(a) = \phi(f(a))$. Similarly, the composition $f \circ \phi$ is an element of R.

4 Step 1

Suppose that we have a bad sum, as in Equation 1. We take the field K to be some finite normal extension that contains $u_1, ..., u_n, v_1, ..., v_n$.

Let N be the kernel of Ψ. If our bad sum exists, then N is nontrivial. In fact, N consists exactly in those elements which Ψ maps to bad sums.

Our bad sum gives us a nontrivial element $f \in N$. Consider the product

$$g = \prod_{\phi\in G}(\phi \circ f) \in N. \tag{11}$$

Since R is an integral domain, g is a nontrivial element of R. By construction $\phi \circ g = g$ for all $\phi \in G$. This is to say that $g(a)$ is fixed by all elements of G. But then $g(a) \in \boldsymbol{Q}$ for all $a \in O_K$. By construction $\Psi(g)$ is a bad sum with rational coefficients. We multiply through by a large integer to make all the coefficients integers. This completes Step 1.

5 Step 2

We keep the same notation. Suppose that $f \in N$ is such that $\Psi(f)$ is a bad sum with integer coefficients. We consider the product

$$g = \prod_{\phi \in G} (f \circ \phi) \in N. \tag{12}$$

By construction, $g \circ \phi = g$ for all $\phi \in G$. The map g assigns the same values to both a and $\phi(a)$ for all $\phi \in G$. Hence, in the bad sum $\Psi(g)$, the coefficient of e^a and the coefficients of $e^{\phi(a)}$ are the same for each $a \in O_K K$ and $\phi \in G$. By construction, these coefficients are integers. Hence, $\Psi(g)$ has exactly the form mentioned in Step 2.

6 Step 3

Say that an element $g \in R$ is *symmetric* if $g \circ \phi = g$ for all $\phi \in G$ and also g is integer valued. We established Step 2 by showing that the kernel N, if nonempty, contains a symmetric element. To complete Step 3, we just have to adjust g so that $g(0) \neq 0$.

Lemma 6.1 *The product of two symmetric elements is symmetric.*

Proof: Suppose that f and g are symmetric. Then, setting $s' = \phi^{-1}(s)$ and $t' = \phi^{-1}(t)$, we have

$$fg \circ \phi(a) = \sum_{s+t=\phi(a)} f(s)g(t) = \sum_{s'+t'=a} f(s)g(t) = \sum_{s'+t'=a} f(s')g(t') = fg(a).$$

Hence $fg \circ \phi = fg$. ♠

Given a symmetric $g \in N$, we choose some algebraic integer $a \in K$ such that $g(a) \neq 0$. We define h to be the symmetric element such that $h(-b) = g(a)$ if and only if $b = \phi(a)$ for some $\phi \in G$, and otherwise $h(b) = 0$. Finally, we set $f = gh$. By construction $f \in N$ and f is symmetric. We compute

$$f(0) = \sum_{s+t=0} g(s)h(t) = Cg(a)^2 \neq 0. \tag{13}$$

The only contributions from this sum arise when s lies in the G-orbit of a. The constant C is the number of points in the G-orbit of a. We have $f(0) \neq 0$ and $f \in N$ and f is symmetric. Hence $\Psi(f)$ is the kind of bad sum advertised in Step 3.

7 Step 4

We can find a polynomial $a(x) \in \boldsymbol{Z}[x]$ such that all the terms $\phi(u_j)$ are roots of a, and 0 is not a root of a.

Choose a prime p and consider the function

$$f(x) = \frac{1}{(p-1)!} x^{p-1} a(x)^p. \tag{14}$$

Next, we define

$$N = f^{(p-1)}(0) + f^{(p)}(0) + ...; \qquad F(x) = f^{(p)}(x) + f^{(p+1)}(x) + ... \tag{15}$$

These are finite sums because f is a polynomial.

Lemma 7.1 *If p is large enough, N is not divisible by p.*

Proof: We can write $f(x) = b_0 x^{p-1} + b_1 x^p + ...$, where $b_0 = a(0)^p/(p-1)!$. We have $f^{p-1}(0) = a(0)^p$ and all higher derivatives of f vanish at 0. If p is large then $a(0)^p$ is not divisible by p. ♠

Lemma 7.2 *$F(x) \in \boldsymbol{Z}[x]$ and all the coefficients are divisible by p.*

Proof: Since F is the sum of integer polynomials, $F(x) \in \boldsymbol{Z}[x]$. Note that $f^{(k)}(x)$ has all coefficients divisible by p as long as $k \geq p$. Hence the sum of these polynomials has all coefficients divisible by p. ♠

To finish our proof, it is convenient to introduce the function

$$G(x) = f(x) + f'(x) + f''(x)... \tag{16}$$

We have

$$G(0) = N; \qquad G(\phi(u_i)) = F(\phi(u_i)). \tag{17}$$

The reason this works is that the first $p-2$ derivatives of f vanish at 0 and the first $p-1$ derivatives of f vanish at each $\phi(u_i)$. So, to finish Step 4, we just have to prove that

$$|G(0)e^t - G(t)| < 1/p; \qquad \forall t = \phi(u_i). \tag{18}$$

The rest of the proof is devoted to the proof of Equation 18. Note that t might be a complex number here. On the first pass, you might want to just consider the case when t is always real. In this case, the derivatives we take are the ordinary derivatives. In the general case, the expression f' means df/dz, the complex derivative. The only difference between the general complex case and the real case is that you have to think a bit about why the starred inequality is true in the complex case.

Let

$$N = \max |\phi(u_i)| \tag{19}$$

where the max is taken over all possibilities. We have $|t| \leq N$.

Let $\psi(x) = e^{-x}G(x)$. We compute

$$\psi'(x) = -e^{-x}(G(x) - G'(x)) = -e^{-x}\Big(\sum_{i=0}^{\infty} f^{(i)}(x) - \sum_{i=1}^{\infty} f^{(i)}(x)\Big) = -e^{-x}f(x).$$

The sums are finite, because f is a polynomial. Our equation tells us that

$$|\psi'(x)| \leq e^N |f(x)|, \tag{20}$$

for all $x \in \boldsymbol{C}$ such that $|x| \leq N$. Letting B be the disk of radius N centered at the origin, we have

$$|G(t) - e^t G(0)| =$$

$$|e^t||\psi(t) - \psi(0)| \leq^*$$

$$te^t \max_B |\psi'| \leq$$

$$Ne^{2N} \max_B |f| \leq \frac{C^p}{(p-1)!},$$

where C is a constant that only depends on the original polynomial a and not on any properties of p. For p sufficiently large, this last bound is less than $1/p$. This finishes Step 4.