Chapter 4

Transcendence results

We recall some basic definitions.

We call $\alpha \in \mathbb{C}$ *transcendental* if it is not algebraic, i.e., if it is not a zero of a non-zero polynomial from $\mathbb{Q}[X]$.

We call numbers $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ *algebraically independent* if there is no non-zero polynomial $P \in \overline{\mathbb{Q}}[X_1, \ldots, X_n]$ such that $P(\alpha_1, \ldots, \alpha_n) = 0$.

A single number $\alpha \in \mathbb{C}$ is algebraically independent if and only if it is transcendental. Indeed, if α is algebraic then there is a non-zero $P \in \mathbb{Q}[X]$ such that $P(\alpha) = 0$. Hence α is certainly not algebraically independent. Conversely, if α is not algebraically independent then there is a non-zero $P \in \overline{\mathbb{Q}}[X]$ such that $P(\alpha) = 0$. But this implies that α is algebraic.

It can be shown that $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ are algebraically independent if there is no non-zero $P \in \mathbb{Q}[X_1, \ldots, X_n]$ (so with coefficients in $\mathbb{Q}$ instead of $\overline{\mathbb{Q}}$) such that $P(\alpha_1, \ldots, \alpha_n) = 0$.

Given a subset S of $\mathbb{C}$, we define the *transcendence degree* of S, notation $\operatorname{trdeg} S$, to be the maximal number t such that S contains t algebraically independent elements.

4.1 The transcendence of e.

We define as usual $e^z = \sum_{n=0}^{\infty} z^n/n!$ for $z \in \mathbb{C}$. Further, $\overline{\mathbb{Q}} = \{\alpha \in \mathbb{C} : \alpha \text{ algebraic over } \mathbb{Q}\}$.

Theorem 4.1 (Hermite, 1873)**.** *e is transcendental.*

We assume that e is algebraic. This means that there are $q_0, q_1, \dots, q_n \in \mathbb{Z}$ with

$$q_0 + q_1 e + \cdots + q_n e^n = 0, \quad q_0 \neq 0. \tag{4.1}$$

Under this hypothesis, we construct $M \in \mathbb{Z}$ with $M \neq 0$ and $|M| < 1$ and obtain a contradiction. We need some auxiliary results. Of course we have to use certain properties of e. We use that $(e^z)' = e^z$.

Let $f \in \mathbb{C}[X]$ be a polynomial. For $z \in \mathbb{C}$ we define

$$F(z) := \int_0^z e^{z-u} f(u) du. \tag{4.2}$$

Here the integration is over the line segment from 0 to z. We may parametrize this line segment by $u = tz$, $0 \leqslant t \leqslant 1$. Thus,

$$F(z) = \int_0^1 e^{z(1-t)} f(zt) z dt.$$

Lemma 4.2. *Suppose f has degree m. Then*

$$F(z) = e^z \left(\sum_{j=0}^{m} f^{(j)}(0) \right) - \sum_{j=0}^{m} f^{(j)}(z).$$

Proof. Repeated integration by parts. □

Corollary 4.3. *Let f be as in Lemma 4.2. Then*

$$q_0 F(0) + \cdots + q_n F(n) = -\sum_{a=0}^{n} \sum_{j=0}^{m} q_a f^{(j)}(a).$$

Proof. Clear. □

Our aim is to show that for a suitable choice of f, the quantity $M := q_0F(0) + \cdots + q_nF(n)$ is a non-zero integer with $|M| < 1$. Note that Corollary 4.3 gives an identity with an analytic expression on the left-hand side, and an algebraic expression on the right-hand side. We prove that M is a non-zero integer by analyzing the right-hand side, and $|M| < 1$ by analyzing the left-hand side. For the latter, we need the following simple estimate.

Lemma 4.4. *Let $f \in \mathbb{C}[X]$ be any polynomial and let F be given by* (4.2). *Then for $z \in \mathbb{C}$ we have*

$$|F(z)| \leqslant |z| \cdot e^{|z|} \cdot \sup_{u\in\mathbb{C},\, |u|\leqslant|z|} |f(u)|.$$

Proof. We have

$$\begin{aligned} |F(z)| &\leqslant \int_0^1 |e^{z(1-t)}f(zt)z|dt \leqslant \int_0^1 e^{|z|}|z| \cdot |f(zt)|dt \\ &\leqslant |z| \cdot e^{|z|} \cdot \sup_{u\in\mathbb{C},\, |u|\leqslant|z|} |f(u)|. \end{aligned}$$

□

Let p be a prime number, which is chosen later to be sufficiently large to make all estimates work. We take

$$f(X) := \tfrac{1}{(p-1)!} \cdot X^{p-1}\left\{(X-1)(X-2)\cdots(X-n)\right\}^p. \tag{4.3}$$

In this case,

$$M = \sum_{a=0}^{n} q_aF(a) = -\sum_{a=0}^{n}\sum_{j=0}^{np+p-1} q_af^{(j)}(a). \tag{4.4}$$

Lemma 4.5. *We have*

$$f^{(p-1)}(0) = \{(-1)^n n!\}^p; \tag{4.5}$$

$$f^{(j)}(a) = 0 \quad \textit{for } a = 0,\dots,n,\ j = 0,\dots,p-1,\ (a,j) \neq (0,p-1); \tag{4.6}$$

$$f^{(j)}(a) \equiv 0 \pmod p \quad \textit{for } a = 0,\dots,n,\ j \geqslant p. \tag{4.7}$$

Proof. In general, if g is a polynomial of the shape $(X-a)^rh$ with $a \in \mathbb{C}$, $h \in \mathbb{C}[X]$, then $g^{(m)}(a) = 0$ for $m = 0,\dots,r-1$ and $g^{(r)}(a) = r!h(a)$. This implies (4.5), (4.6).

To prove (4.7), observe that for any $g = c_r X^r + \cdots + c_0 \in \mathbb{C}[X]$ and all $j \geqslant 0$ we have

$$\tag{4.8} \frac{1}{j!} g^{(j)} = c_r \tbinom{r}{j} X^{r-j} + c_{r-1} \tbinom{r-1}{j} X^{r-j-1} + \cdots + c_j.$$

In particular, since $(p-1)! f \in \mathbb{Z}[X]$ and the binomial coefficients are integers, we have for $j \geqslant p$, $a = 0, \ldots, n$ that $(p-1)! f^{(j)}/j! \in \mathbb{Z}[X]$, and so $f^{(j)}(a)/p \in \mathbb{Z}$. This implies at once (4.7). □

Lemma 4.6. *Assume that $p > |q_0 n|$. Then M is a non-zero integer.*

Proof. From (4.5) it follows that the term $q_0 f^{(p-1)}(0)$ is an integer not divisible by p, while all other terms $q_a f^{(j)}(a)$ in the right-hand side of (4.4) are integers that are either 0 or divisible by p. Hence M is an integer not divisible by p. □

Lemma 4.7. *For p sufficiently large, we have $|M| < 1$.*

Proof. By Lemma 4.4, we have for $a = 0, \ldots, n$,

$$|F(a)| \leqslant a \cdot e^{|a|} \cdot \sup_{|u| \leqslant a} |f(u)|.$$

For $a, b = 0, \ldots, n$, and $u \in \mathbb{C}$ with $|u| \leqslant a$ we have $|u - b| \leqslant |u| + |b| \leqslant 2n$. Hence

$$\sup_{|u| \leqslant a} |f(u)| \leqslant \frac{(2n)^{np+p-1}}{(p-1)!} \leqslant \frac{c^p}{(p-1)!},$$

say, where c is a constant independent of p, a, b. This implies

$$|M| \leqslant \sum_{a=0}^{n} |q_a F(a)| \leqslant \left(\sum_{a=0}^{n} q_a \cdot a \cdot e^a \right) \frac{c^p}{(p-1)!}.$$

For p sufficiently large this is < 1, since for any $c > 1$, $\frac{c^p}{(p-1)!} \to 0$ as $p \to \infty$. □

Summarizing, our assumption that e is algebraic implies that there is a quantity M, which is by Lemma 4.6 an non-zero integer, and by Lemma 4.7, of absolute value < 1. Since this is absurd, e must be transcendental.

4.2 The Lindemann-Weierstrass theorem

Lindemann proved in 1882 that e^α is transcendental for algebraic α, and Weierstrass proved in 1885 that if $\alpha_1, \dots, \alpha_n$ are algebraic numbers that are linearly independent over $\mathbb{Q}$, then $e^{\alpha_1}, \dots, e^{\alpha_n}$ are algebraically independent over $\mathbb{Q}$. The following result, due to A. Baker, is in fact equivalent to the Lindemann-Weierstrass Theorem.

Theorem 4.8. *Let $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in \overline{\mathbb{Q}}$. Suppose that $\alpha_1, \dots, \alpha_n$ are pairwise distinct, and that $\beta_1, \dots, \beta_n \neq 0$. Then*

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

Before proving this theorem, we state some corollaries.

Corollary 4.9. *(i) Let $\alpha \in \overline{\mathbb{Q}}$ be non-zero. Then e^α is transcendental.*
(ii) π is transcendental.

Proof. (i) Suppose that $e^\alpha =: \beta$ is algebraic. Then it follows that $1 \cdot e^\alpha - \beta \cdot e^0 = 0$, contradicting Theorem 4.8.
(ii) Suppose that π is algebraic. Then πi is algebraic. But $e^{\pi i} = -1$ is not transcendental, contradicting (i). □

Corollary 4.10. *(i) Let $\alpha \in \overline{\mathbb{Q}}$ and $\alpha \neq 0$. Then $\sin \alpha$, $\cos \alpha$, and $\tan \alpha$ are transcendental.*
(ii) Let $\alpha \in \overline{\mathbb{Q}}$ and $\alpha \neq 0, 1$. Then $\log \alpha$ is transcendental (for any choice of $\log \alpha$, i.e., any solution z of $e^z = \alpha$).

Exercise 4.1. *Prove Corollary 4.10.*

Corollary 4.11. *Let $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}$. Then the following two assertions are equivalent:*
(i) $\alpha_1, \dots, \alpha_n$ are linearly independent over $\mathbb{Q}$;
(ii) $e^{\alpha_1}, \dots, e^{\alpha_n}$ are algebraically independent.

Exercise 4.2. *Prove Corollary 4.11.*

We start with some preliminary comments on the proof of Theorem 4.8.

Our proof of the transcendence of e was by contradiction: we assumed that $q_0 + q_1 e + \cdots + q_n e^n = 0$ for certain rational integers $q_0, \dots, q_n$, and constructed from this a non-zero integer M with $|M| < 1$. To prove the Lindemann-Weierstrass Theorem, we may again proceed by contradiction and assume that $\beta_0 e^{\alpha_0} + \cdots + \beta_n e^{\alpha_n} = 0$. By following the transcendence proof of e, but replacing $0, 1, , \ldots, n$ by $\alpha_1, \ldots, \alpha_n$ and $q_0, \ldots, q_n$ by $\beta_1, , \ldots, \beta_n$, we obtain a non-zero algebraic integer M, not necessarily in $\mathbb{Q}$, such that $|M| < 1$. This is, however, not a contradiction. For instance, $\frac{1}{2}(1-\sqrt{5})$ is an algebraic integer of absolute value < 1.

Instead, we use the following observation. Let α be in the ring of integers of an algebraic number field K. Recall that the characteristic polynomial of α is $\chi_{\alpha,K}(X) = \prod_\sigma (X - \sigma(\alpha))$ where the product is over all embeddings $\sigma : K \hookrightarrow \mathbb{C}$ of K. This polynomial has its coefficients in $\mathbb{Z}$. So in particular the norm $N_{K/\mathbb{Q}}(\alpha) = \prod_\sigma \sigma(\alpha)$ is in $\mathbb{Z}$. If $\alpha \neq 0$, we have $N_{K/\mathbb{Q}}(\alpha) \neq 0$, hence $|N_{K/\mathbb{Q}}(\alpha)| \geqslant 1$.

So from the assumption that Theorem 4.8 is false we have to construct somehow a non-zero algebraic integer whose norm has absolute value < 1. If we just follow the transcendence proof of e without any modifications, we only get a non-zero algebraic integer M with $|M| < 1$, but we can not show that its norm has absolute value < 1. A priori, for some embedding σ, the quantity $|\sigma(M)|$ may be very large, and then $|N_{K/\mathbb{Q}}(M)|$ may be $\geqslant 1$.

To circumvent this, we deduce from the expression $\sum_{i=1}^n \beta_i e^{\alpha_i}$ a new expression $\sum_{i=1}^t \delta_i e^{\gamma_i}$, where the γ_i, δ_i satisfy certain symmetry conditions. These symmetry conditions allow to construct, under the hypothesis $\sum_{i=1}^t \delta_i e^{\gamma_i} = 0$, a non-zero algebraic integer having a norm with absolute value < 1. Thus, we obtain a weaker version of the Lindemann-Weierstrass Theorem, which asserts that under the said symmetry conditions, $\sum_{i=1}^t \delta_i e^{\gamma_i} \neq 0$. But as will be seen, this weaker version implies the general Lindemann-Weierstrass Theorem.

Theorem 4.12 ("Weak Lindemann-Weierstrass Theorem"). *Let $L \subset \mathbb{C}$ be a normal algebraic number field. Let $\gamma_1, \ldots, \gamma_t$, $\delta_1, \ldots, \delta_t \in L$ such that*

$$\gamma_1, \ldots, \gamma_t \text{ are distinct}, \quad \delta_1 \cdots \delta_t \neq 0,$$

and suppose moreover, that each $\tau \in \mathrm{Gal}(L/\mathbb{Q})$ permutes the pairs $(\gamma_1, \delta_1), \ldots, (\gamma_t, \delta_t)$. Then

$$\delta_1 e^{\gamma_1} + \cdots + \delta_t e^{\gamma_t} \neq 0.$$

We say that τ permutes the pairs $(\gamma_1, \delta_1), \ldots, (\gamma_t, \delta_t)$ if
$(\tau(\gamma_1), \tau(\delta_1)), \ldots, (\tau(\gamma_t), \tau(\delta_t))$ is a permutation of $(\gamma_1, \delta_1), \ldots, (\gamma_t, \delta_t)$.

We first prove the implication Theorem 4.12$\Longrightarrow$Theorem 4.8. After that, we prove Theorem 4.12.

Theorem 4.12 $\Longrightarrow$ Theorem 4.8. Assume that Theorem 4.8 is false. This means that there are $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in \overline{\mathbb{Q}}$ such that, $\alpha_1, \ldots, \alpha_n$ are distinct, $\beta_1, \ldots, \beta_n$ are non-zero, and

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

We derive from this a contradiction to Theorem 4.12.

Let L be the number field generated by $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$ and their conjugates. Then L is a normal number field. Let

$$\mathrm{Gal}(L/\mathbb{Q}) = \{\tau_1, \ldots, \tau_d\}.$$

Recall that if $\gamma \in L$, then the set $\{\tau_1(\gamma), \ldots, \tau_d(\gamma)\}$ contains all conjugates of γ. Clearly,

$$\prod_{i=1}^{d} \left(\tau_i(\beta_1) e^{\tau_i(\alpha_1)} + \cdots + \tau_i(\beta_n) e^{\tau_i(\alpha_n)}\right) = 0.$$

By expanding the product, we get

$$\sum_{i_1=0}^{n} \cdots \sum_{i_d=1}^{n} \tau_1(\beta_{i_1}) \cdots \tau_d(\beta_{i_d}) \cdot e^{\tau_1(\alpha_{i_1}) + \cdots + \tau_d(\alpha_{i_d})} = 0. \tag{4.9}$$

Each $\tau \in \mathrm{Gal}(L/\mathbb{Q})$ permutes the pairs $\big(\tau_1(\alpha_{i_1}) + \cdots + \tau_d(\alpha_{i_d}),\ \tau_1(\beta_{i_1}) \cdots \tau_d(\beta_{i_d})\big)$, since $\tau\tau_1, \ldots, \tau\tau_d$ is a permutation of $\tau_1, \ldots, \tau_d$.

The exponents $\tau_1(\alpha_{i_1}) + \cdots + \tau_d(\alpha_{i_d})$ need not be distinct. We group together the terms with equal exponents. Let $\gamma_1, \ldots, \gamma_s$ be the distinct values among $\tau_1(\alpha_{i_1}) + \cdots + \tau_d(\alpha_{i_d})$ $(0 \leqslant i_1, \ldots, i_n \leqslant d)$, and for $k = 1, \ldots, s$, denote by J_k the set of tuples $(i_1, \ldots, i_d)$ such that

$$\tau_1(\alpha_{i_1}) + \cdots + \tau_d(\alpha_{i_d}) = \gamma_k.$$

Then (4.9) becomes

$$\sum_{k=1}^{s} \delta_k e^{\gamma_k} = 0, \quad \text{where } \delta_k = \sum_{(i_1, \ldots, i_k) \in J_k} \tau_1(\beta_{i_1}) \cdots \tau_d(\beta_{i_d}). \tag{4.10}$$

Notice that each $\tau \in \mathrm{Gal}(L/\mathbb{Q})$ permutes the pairs $(\gamma_1, \delta_1), \ldots, (\gamma_s, \delta_s)$. A priori, all coefficients δ_k might be 0. However, we show that there is a tuple $(i_1, \ldots, i_k)$ such that $\tau_1(\alpha_{i_1}) + \cdots + \tau_d(\alpha_{i_d})$ is different from all the other exponents. Thus, for some k, the set J_k has cardinality 1, and $\delta_k \neq 0$.

Define a total ordering on $\mathbb{C}$ by setting $\theta < \zeta$ if $\operatorname{Re}\theta < \operatorname{Re}\zeta$ or if $\operatorname{Re}\theta = \operatorname{Re}\zeta$ and $\operatorname{Im}\theta < \operatorname{Im}\zeta$. This ordering has the property that if θ_i, ζ_i are complex numbers with $\theta_i < \zeta_i$ for $i = 1, \dots, r$, then $\sum_{j=1}^r \theta_j < \sum_{j=1}^r \zeta_j$.

Since $\alpha_1, \dots, \alpha_d$ were assumed to be distinct, for each $\tau \in \operatorname{Gal}(L/\mathbb{Q})$, the numbers $\tau(\alpha_1), \dots, \tau(\alpha_d)$ are distinct. Hence for each $k \in \{1, \dots, d\}$, there is an index i_k such that $\tau_k(\alpha_{i_k}) > \tau_k(\alpha_j)$ for $j \neq i_k$. This implies

$$\tau_1(\alpha_{i_1}) + \cdots + \tau_d(\alpha_{i_d}) > \tau_1(\alpha_{j_1}) + \cdots + \tau_d(\alpha_{j_d})$$

for all tuples $(j_1, \dots, j_d) \neq (i_1, \dots, i_d)$, and so $\tau_1(\alpha_{i_1}) + \cdots + \tau_d(\alpha_{i_d})$ is distinct from the other exponents.

Assume without loss of generality that $\delta_1, \dots, \delta_t$ are the non-zero numbers among $\delta_1, \dots, \delta_s$. Then (4.10) becomes

$$\sum_{k=1}^{t} \delta_k e^{\gamma_k} = 0. \tag{4.11}$$

By construction, the numbers $\gamma_1, \dots, \gamma_t$ are distinct algebraic numbers. Further, $\delta_1, \dots, \delta_t$ are non-zero. As observed before, each $\tau \in \operatorname{Gal}(L/\mathbb{Q})$ permutes the pairs $(\gamma_1, \delta_1), \dots, (\gamma_s, \delta_s)$ from (4.10). But then τ permutes also the pairs with $\delta_k \neq 0$, i.e., τ permutes $(\gamma_1, \delta_1), \dots, (\gamma_t, \delta_t)$. Now Theorem 4.12 implies that (4.11) is false.

Thus, our assumption that Theorem 4.8 is false leads to a contradiction. □

Proof of Theorem 4.12. We follow the transcendence proof of e, with the necessary modifications. Before proceeding, we observe that there is no loss of generality to assume that $\delta_1, \dots, \delta_t$ are algebraic integers. Indeed, there is a positive $m \in \mathbb{Z}$ such that $m\delta_1, \dots, m\delta_t$ are algebraic integers (e.g, we may take for m the product of the denominators of $\delta_1, \dots, \delta_t$), and clearly, the conditions and conclusion of Theorem 4.12 are unaffected if we replace δ_i by $m\delta_i$ for $i = 1, \dots, t$.

Let $\gamma_1, \dots, \gamma_t$ be distinct algebraic numbers and $\delta_1, \dots, \delta_t$ non-zero algebraic integers from the normal number field L, such that each $\tau \in \operatorname{Gal}(L/\mathbb{Q})$ permutes the pairs $(\gamma_1, \delta_1), \dots, (\gamma_t, \delta_t)$. Assume that

$$\delta_1 e^{\gamma_1} + \cdots + \delta_t e^{\gamma_t} = 0. \tag{4.12}$$

Let again p be a prime number. Further, let l be a positive rational integer such that $l\gamma_1, \dots, l\gamma_t$ are all algebraic integers (e.g., the product of the denominators of $\gamma_1, \dots, \gamma_t$). For $k = 1, \dots, t$, define

$$f_k(X) := \tfrac{1}{(p-1)!} \cdot l^{tp}(X-\gamma_k)^{p-1}\prod_{j=1}^{t}(X-\gamma_j)^p,$$

$$F_k(z) := \int_0^z e^{z-u} f_k(u)du,$$

$$M_k := \delta_1 F_k(\gamma_1) + \cdots + \delta_t F_k(\gamma_t).$$

Then the following can be shown:

1) For each $\tau \in \mathrm{Gal}(L/\mathbb{Q})$ we have $\tau(M_1) \in \{M_1, \dots, M_t\}$;

2) for sufficiently large p, M_1 is a non-zero algebraic integer;

3) $|M_k| < 1$ for $k = 1, \dots, t$ and sufficiently large p.

From 1) and 3) it follows that $|N_{L/\mathbb{Q}}(M_1)| < 1$. But this contradicts 2).

Lemma 4.13. *(i) We have*

$$M_k = -\sum_{j=1}^{t}\sum_{m=0}^{tp-1}\delta_j f_k^{(m)}(\gamma_j) \quad \textit{for } k = 1, \dots, t.$$

(ii) For each $\tau \in \mathrm{Gal}(L/\mathbb{Q})$ we have $\tau(M_1) \in \{M_1, \dots, M_t\}$.

Proof. (i) This follows at once from Lemma 4.2 and our assumption $\sum_{j=1}^t \delta_j e^{\gamma_j} = 0$.

(ii) Let $\tau \in \mathrm{Gal}(L/\mathbb{Q})$. Then there is a permutation τ^* of $1, \dots, t$ such that

$$(\tau(\gamma_k), \tau(\delta_k)) = (\gamma_{\tau^*(k)}, \delta_{\tau^*(k)}) \ \text{ for } k = 1, \dots, t.$$

By applying τ to the coefficients of f_1, we obtain

$$l^{tp}(X-\tau(\gamma_1))^{p-1}\prod_{j=1}^{t}(X-\tau(\gamma_j))^p = l^{tp}(X-\gamma_{\tau^*(1)})\prod_{j=1}^{t}(X-\gamma_{\tau^*(j)})^p = f_{\tau^*(1)}.$$

Hence

$$\begin{aligned}\tau(M_1) &= -\frac{1}{(p-1)!}\sum_{j=1}^{t}\sum_{m=0}^{tp-1}\tau(\delta_j)f_{\tau^*(1)}^{(m)}(\tau(\gamma_j))\\ &= -\frac{1}{(p-1)!}\sum_{j=1}^{t}\sum_{m=0}^{tp-1}\delta_{\tau^*(j)}f_{\tau^*(1)}^{(m)}(\gamma_{\tau^*(j)}) = M_{\tau^*(1)}.\end{aligned}$$

□

Given two algebraic numbers α, β and a positive integer $m \in \mathbb{Z}$, we write $\alpha \equiv \beta \pmod{m}$ if $(\alpha - \beta)/m$ is an algebraic integer.

Lemma 4.14. *let $k \in \{1, \ldots, t\}$. Then*

$$f_1^{(p-1)}(\gamma_1) = l^{tp}\left\{\prod_{k=2}^{t}(\gamma_1 - \gamma_k)\right\}^p, \tag{4.13}$$

$$f_1^{(j)}(\gamma_m) = 0 \quad \textit{for } m = 1, \ldots, t,\ j = 0, \ldots, p-1,\ (m, j) \neq (1, p-1), \tag{4.14}$$

$$f_1^{(j)}(\gamma_m) \equiv 0 (\operatorname{mod} p) \textit{ for } m = 1, \ldots, t,\ j \geqslant p. \tag{4.15}$$

Proof. The proofs of (4.13) and (4.14) are completely analogous to those of (4.5) and (4.6) in Lemma 4.5. We prove only (4.15). Let $m \in \{1, \ldots, t\}$ and $j \geqslant p$. Define

$$g(X) := f_1(X/l) = \tfrac{1}{(p-1)!} \cdot l(X - l\gamma_1)^{p-1}\prod_{k=2}^{t}(X - l\gamma_k)^p.$$

Then $(p-1)!g$ has algebraically integral coefficients. Using (4.8), one easily shows that the coefficients of $(p-1)!g_1^{(j)}/j!$ are algebraic integers. Hence for $j \geqslant p$, $g^{(j)}(l\gamma_j)/p$ is an algebraic integer, and therefore,

$$\frac{f_1^{(j)}(\gamma_m)}{p} = \frac{l^j g^{(j)}(l\gamma_m)}{p}$$

is an algebraic integer. This implies at once (4.15). □

Lemma 4.15. *For p sufficiently large, M_1 is a non-zero algebraic integer.*

Proof. An application of Lemma 4.14 gives

$$M_1 \equiv -\delta_1 A^p \,(\operatorname{mod} p) \quad \text{with } A := l^t \prod_{k=2}^{t} (\gamma_1 - \gamma_k).$$

Both δ_1, A are algebraic integers, hence M_1 is an algebraic integer. We prove that for sufficiently large p, $\delta_1 A^p/p$ is not an algebraic integer. Then necessarily, $M_1 \neq 0$.

Assume that $\delta_1 A^p/p$ is an algebraic integer. Let $b = N_{L/\mathbb{Q}}(\delta_1)$, $B = N_{L/\mathbb{Q}}(A)$. Then $b, B \in \mathbb{Z}$, and the norm $N_{L/\mathbb{Q}}(\delta_1 A^p/p) = bB^p/p^d$ is in $\mathbb{Z}$, where $d = [L : \mathbb{Q}]$. But this is impossible if $p > |bB|$. □

Lemma 4.16. *For p sufficiently large we have $|M_k| < 1$ for $k = 1, \ldots, t$.*

Exercise 4.3. *Prove this lemma.*

Thus our assumption that Theorem 4.12 is false implies the Lemmas 4.13, 4.15, 4.16, and these together give a contradiction. □

4.3 Other transcendence results

We give an overview of some other transcendence results, without proof. As usual, we define $e^z := \sum_{n=0}^{\infty} z^n/n!$ for complex numbers z. Given $\alpha, \beta \in \mathbb{C}$ we define $\alpha^\beta := e^{\beta \log \alpha}$ where $\log \alpha$ is any solution of $e^z = \alpha$. Recall that the latter equation has infinitely many solutions; if l_0 is one solution then the others are given by $l_0 + 2k\pi i$ with $k \in \mathbb{Z}$. This gives infinitely many possibilities $e^{\beta(l_0 + 2k\pi i)}$ for α^β. We agree that e^z will always be the above defined power series.

Theorem 4.17 (Gel'fond, Schneider, 1934)**.** *Let $\alpha, \beta \in \overline{\mathbb{Q}}$ with $\alpha \neq 0, 1$, $\beta \notin \mathbb{Q}$. Let $\log \alpha$ be any solution of $e^z = \alpha$. Then $\alpha^\beta := e^{\beta \log \alpha}$ is transcendental.*

Corollary 4.18. *Let $\alpha \in \overline{\mathbb{Q}}$ with $\alpha \notin \mathbb{Q}i$. Then $e^{\pi\alpha}$ is transcendental.*

Proof. Choose $\log(-1) = \pi i$. Then $e^{\pi\alpha} = e^{-i\alpha \log(-1)} = (-1)^{-i\alpha}$. □

Corollary 4.19. *Let α_1, α_2 be non-zero algebraic numbers. Take any solutions $\log \alpha_1$, $\log \alpha_2$ of $e^z = \alpha_1$, $e^z = \alpha_2$, respectively, and assume that these are linearly independent over $\mathbb{Q}$. Then for any two non-zero algebraic numbers β_1, β_2 we have*

$$\beta_1 \log \alpha_1 + \beta_2 \log \alpha_2 \neq 0.$$

Proof. Suppose $\beta_1 \log \alpha_1 + \beta_2 \log \alpha_2 = 0$. Put $\gamma := -\beta_2/\beta_1$. Then by assumption, $\gamma \notin \mathbb{Q}$, and

$$\alpha_2 = e^{\log \alpha_2} = e^{\gamma \log \alpha_1} = \alpha_1^{\gamma},$$

contradicting Theorem 4.17. □

In 1966, A. Baker proved the following far-reaching generalization.

Theorem 4.20 (A. Baker, 1966)**.** *Let $\alpha_1, \ldots, \alpha_n$ be non-zero algebraic numbers. For $i = 1, \ldots, n$ let $\log \alpha_i$ be any solution of $e^z = \alpha_i$, and assume that*

$$\log \alpha_1, \ldots, \log \alpha_n \text{ are linearly independent over } \mathbb{Q}.$$

Then for any non-zero algebraic numbers $\beta_1, \ldots, \beta_n$,

$$\beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n \text{ is transcendental.}$$

Definition. We say that non-zero complex numbers $\alpha_1, \ldots, \alpha_n$ are *multiplicatively dependent* if there are $x_1, \ldots, x_n \in \mathbb{Z}$, not all 0, such that

$$\alpha_1^{x_1} \cdots \alpha_n^{x_n} = 1.$$

Otherwise, $\alpha_1, \ldots, \alpha_n$ are called multiplicatively independent.

Corollary 4.21. *Let $n \geqslant 1$. Let $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in \overline{\mathbb{Q}}$ be such that*

$$\begin{aligned} &\alpha_1, \ldots, \alpha_n \neq 0, \quad \alpha_1, \ldots, \alpha_n \text{ are multiplicatively independent,} \\ &(\beta_1, \ldots, \beta_n) \notin \mathbb{Q}^n. \end{aligned}$$

Then $\alpha_1^{\beta_1} \cdots \alpha_n^{\beta_n}$ is transcendental. Here $\alpha_i^{\beta_i} := e^{\beta_i \log \alpha_i}$ where $\log \alpha_i$ is any solution of $e^z = \alpha_i$, for $i = 1, \ldots, n$.

Proof. Suppose that $\alpha_{n+1} := \alpha_1^{\beta_1} \cdots \alpha_n^{\beta_n} = e^{\beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n}$ is algebraic. Then we may choose $\log \alpha_{n+1}$ such that

$$\log \alpha_{n+1} = \beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n. \tag{4.16}$$

By Theorem 4.20, $\log \alpha_1, \ldots, \log \alpha_n$ and $\log \alpha_{n+1}$ are linearly dependent over $\mathbb{Q}$, that is, there are $x_1, \ldots, x_n, x_{n+1} \in \mathbb{Z}$, not all 0, such that

$$x_1 \log \alpha_1 + \cdots + x_n \log \alpha_n + x_{n+1} \log \alpha_{n+1} = 0. \tag{4.17}$$

Eliminating $\log \alpha_{n+1}$ from (4.16) and (4.17), we get

$$(x_{n+1}\beta_1 + x_1)\log \alpha_1 + \cdots + (x_{n+1}\beta_n + x_n)\log \alpha_n = 0.$$

Since $(\beta_1, \dots, \beta_n) \notin \mathbb{Q}^n$ we have $x_{n+1}\beta_i + x_i \neq 0$ for at least one $i \in \{1, \dots, n\}$. Applying again Theorem 4.20, we infer that there are $y_1, \dots, y_n \in \mathbb{Z}$, not all 0, such that

$$y_1 \log \alpha_1 + \cdots + y_n \log \alpha_n = 0.$$

Now we get

$$\alpha_1^{y_1} \cdots \alpha_n^{y_n} = e^{y_1 \log \alpha_1 + \cdots + y_n \log \alpha_n} = 1,$$

contrary to our assumption. □

The next exercise gives another application of Theorem 4.20.

Exercise 4.4. *Let $\alpha_1, \dots, \alpha_n,\ \beta_1, \dots, \beta_n \in \overline{\mathbb{Q}}$, and suppose that $\alpha_1, \dots, \alpha_n \neq 0$. For $i = 1, \dots, n$, let $\log \alpha_i$ be any solution of $e^z = \alpha_i$.*

(i) Assume that $\beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n \neq 0$. Prove that $\beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n$ is transcendental.

Hint. *Proceed by induction on n. In the induction step use Theorem 4.20.*

(ii) Let $\alpha_1, \dots, \alpha_n,\ \beta_1, \dots, \beta_n \in \overline{\mathbb{Q}}$ with $\alpha_1, \dots, \alpha_n \neq 0$ and let $\gamma \in \overline{\mathbb{Q}}$ with $\gamma \neq 0$. Put $\alpha_i^{\beta_i} := e^{\beta_i \log \alpha_i}$ for $i = 1, \dots, n$. Prove that $e^{\gamma}\alpha_1^{\beta_1} \cdots \alpha_n^{\beta_n}$ is transcendental.

There is a far reaching conjecture, due to Schanuel, which implies all results mentioned before and much more.

Schanuel's Conjecture. (1960's) *Let $x_1, \dots, x_n$ be any (not necessarily algebraic) complex numbers that are linearly independent over $\mathbb{Q}$. Then*

$$\mathrm{trdeg}(x_1, \dots, x_n, e^{x_1}, \dots, e^{x_n}) \geqslant n.$$

We give some examples of known cases.

Examples. 1. Let $x \in \mathbb{C}^*$. Then either x is transcendental, or x is algebraic and then by Lindemann's Theorem, e^x is transcendental. Hence $\mathrm{trdeg}(x, e^x) \geqslant 1$. Schanuel's Conjecture is still open for $n \geqslant 2$.

2. Let $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}$ and suppose they are linearly independent over $\mathbb{Q}$. By

Corollary 4.11 (a consequence of the Lindemann-Weierstrass Theorem), $e^{\alpha_1}, \dots, e^{\alpha_n}$ are algebraically independent. Hence $\mathrm{trdeg}(\alpha_1, \dots, \alpha_n, e^{\alpha_1}, \dots, e^{\alpha_n}) = n$.

We deduce some consequences of Schanuel's Conjecture which are still open.

Conjecture. *e and π are algebraically independent.*

Proof under the assumption of Schanuel's Conjecture. The transcendence degree of a set of complex numbers does not change if some algebraic numbers are added to or removed from it. Moreover, the transcendence degree of this set does not change if we multiply its elements with non-zero algebraic numbers. So by Schanuel's conjecture,

$$\mathrm{trdeg}(e, \pi) = \mathrm{trdeg}(e, \pi i) = \mathrm{trdeg}(1, \pi i, e, e^{\pi i}) = 2.$$

□

Conjecture. *Let $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}$ such that $\alpha_1, \dots, \alpha_n \neq 0$ and $\log \alpha_1, \dots, \log \alpha_n$ are linearly independent over $\mathbb{Q}$, where again $\log \alpha_i$ is any solution of $e^z = \alpha_i$ for $i = 1, \dots, n$. Then $\log \alpha_1, \dots, \log \alpha_n$ are algebraically independent.*

Proof under the assumption of Schanuel's Conjecture. We have

$$\mathrm{trdeg}(\log \alpha_1, \dots, \log \alpha_n) = \mathrm{trdeg}(\log \alpha_1, \dots, \log \alpha_n, \alpha_1, \dots, \alpha_n) = n.$$

□

The above conjecture implies that for every non-zero polynomial $P \in \overline{\mathbb{Q}}[X_1, \dots, X_n]$ we have $P(\log \alpha_1, \dots, \log \alpha_n) \neq 0$. Baker's Theorem 4.20 implies that this holds for linear polynomials $P \in \overline{\mathbb{Q}}[X_1, \dots, X_n]$, but even for quadratic polynomials P this is still open. For instance, the above conjecture implies that $\log 2 \cdot \log 3$ is transcendental, but as yet not even this very special case could be proved.

We mention some other consequences of Schanuel's conjecture, the deduction of which is left as an exercise.

Exercise 4.5. *Deduce the following from Schanuel's conjecture:*
(i) Let $\alpha \in \overline{\mathbb{Q}}$, $\alpha \notin i\mathbb{Q}$. Then π and $e^{\pi\alpha}$ are algebraically independent.
(ii) Let α, $\beta \in \overline{\mathbb{Q}}$ with $\alpha \notin \{0, 1\}$ and β of degree $d \geqslant 2$. Then $\alpha^{\beta}, \alpha^{\beta^2}, \dots, \alpha^{\beta^{d-1}}$ are algebraically independent. Here $\alpha^{\beta^j} = e^{\beta^j \log \alpha}$ with $\log \alpha$ any solution of $e^z = \alpha$.
(iii) Define the sequence $\{x_n\}_{n=1}^{\infty}$ by $x_1 = e$ and $x_n = e^{x_{n-1}}$ for $n \geqslant 2$, i.e., $x_2 = e^e$,

$x_3 = e^{e^e}$, etc. Then $x_1, \dots, x_N$ are algebraically independent for every $N \geqslant 1$.
(iv) Let $\alpha \in \overline{\mathbb{Q}} \setminus \{0,1\}$. Then $\log\alpha, \log\log\alpha$ are algebraically independent (for any solution $\log\alpha$ of $e^z = \alpha$ and any solution $\log\log\alpha$ of $e^z = \log\alpha$).
(v) Let p, q be two distinct prime numbers. Then for every irrational $x \in \mathbb{R}$, at least one of the numbers p^x, q^x is transcendental (here we just take the ordinary exponentiation of reals; for complex x this can also be deduced from Schanuel's conjecture but this is much harder).

Remark. The following has been proved.

In 1996, Nesterenko proved (among other things), that π, e^{π} and $\Gamma(\frac{1}{4})$ are algebraically independent. Recall that $\Gamma(x) = \int_0^{\infty} t^{x-1}e^{-t}dt$ for $x > 0$, that $\Gamma(n) = (n-1)!$ for every positive integer n, and that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

For α, β as in (ii), Diaz proved in 1989 that

$$\operatorname{trdeg}(\alpha^{\beta}, \alpha^{\beta^2}, \dots, \alpha^{\beta^{d-1}}) \geqslant [(d+1)/2]$$

where $[x]$ is the largest integer $\leqslant x$. This settles (ii) for $d = 3$.

In the 1960's, Lang and Ramachandra independently proved (among other things) that if p_1, p_2, p_3 are three distinct primes and x an irrational real, then at least one of the numbers p_1^x, p_2^x, p_3^x is transcendental.

4.4 Siegel's Lemma

In the next section we will prove a special case of the Gel'fond-Schneider Theorem, that is that if α, β are *real* algebraic numbers with $\alpha \neq 0, 1$, $\beta \notin \mathbb{Q}$, then $\alpha^{\beta} := e^{\beta \log\alpha}$ is transcendental, where $\log\alpha$ is just the usual real natural logarithm. In the present section we develop a tool which is very important in Diophantine approximation, the so-called *Siegel's Lemma,* which was formally stated for the first time by Siegel in 1929, but was known before. Essentially, it states that under certain hypotheses, a system of M homogeneous linear equations in N unknowns

$$(4.18) \qquad \left\{ \begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1N}x_N & = 0 \\ \vdots & & & & & \vdots \\ a_{M1}x_1 & + & \cdots & + & a_{MN}x_N & = 0 \end{array} \right.$$

has a non-trivial solution in integer coordinates, the absolute values of which are not too large.

Theorem 4.22 (Siegel's Lemma). *Assume that* $N > M > 0$*,* $A \geqslant 1$*, and*

$$a_{ij} \in \mathbb{Z}, \quad |a_{ij}| \leqslant A \quad \textit{for } i = 1, \dots, M,\ j = 1, \dots, N.$$

Then (4.18) *has a solution* $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{Z}^N \setminus \{\mathbf{0}\}$ *with*

$$\max_{1 \leqslant i \leqslant N} |x_i| \leqslant (NA)^{M/(N-M)}.$$

Proof. For $i = 1, \dots, M$, $\mathbf{x} \in \mathbb{Z}^N$, put $l_i(\mathbf{x}) := \sum_{j=1}^N a_{ij} x_j$ and let

$$-C_i := \sum_{j=1}^N \min(a_{ij}, 0), \quad D_i := \sum_{j=1}^N \max(a_{ij}, 0).$$

Notice that $C_i + D_i \leqslant NA$. Let B be a positive integer, and let $S_B := \{0, \dots, B\}^N$. For each $\mathbf{y} \in S_B$ we have

$$-C_i B \leqslant l_i(\mathbf{y}) \leqslant D_i B \quad \text{for } i = 1, \dots, M.$$

Notice that S_B has cardinality $(B+1)^N$. Further, if $\mathbf{y}$ runs through S_B, then $(l_1(\mathbf{y}), \dots, l_M(\mathbf{y}))$ runs through a set of cardinality at most

$$\prod_{i=1}^M (C_i B + D_i B + 1) \leqslant (NAB + 1)^M.$$

We choose B such that $(B+1)^N > (NAB+1)^M$. Then by the box principle, there are distinct $\mathbf{y}_1, \mathbf{y}_2 \in S_B$ with $l_i(\mathbf{y}_1) = l_i(\mathbf{y}_2)$ for $i = 1, \dots, M$. Take $\mathbf{x} = \mathbf{y}_1 - \mathbf{y}_2$. Then $\mathbf{x}$ satisfies (4.18) and $|x_i| \leqslant B$ for $i = 1, \dots, N$.

We finish our proof by showing that the choice $B = [(NA)^{M/(N-M)}]$ is valid. Indeed, with this choice of B we have $(B+1)^{N-M} > (NA)^M$, hence

$$(B+1)^N > (NA(B+1))^M > (NAB+1)^M.$$

□

We need a generalization where the coefficients a_{ij} are algebraic integers instead of just rational integers. To deduce this, we need some preparations.

Let K be an algebraic number field of degree d. Denote as usual by O_K its ring of integers. Assume K has r_1 real embeddings, and r_2 pairs of conjugate complex embeddings, so that $r_1 + 2r_2 = d$. We order the embeddings of K in $\mathbb{C}$ in such a way that $\sigma_1, \dots, \sigma_{r_1}$ are the real embeddings, and $\{\sigma_{r_1+1}, \sigma_{r_1+r_2+1} = \overline{\sigma_{r_1+1}}\}, \dots, \{\sigma_{r_1+r_2}, \sigma_{r_1+2r_2} = \overline{\sigma_{r_1+r_2}}\}$ are the pairs of conjugate complex embeddings. Define the map

$$\begin{aligned}&\varphi : K \to \mathbb{R}^d :\\ &\quad x \mapsto \big(\sigma_1(x), \dots, \sigma_{r_1}(x), \operatorname{Re}\sigma_{r_1+1}(x), \operatorname{Im}\sigma_{r_1+1}(x), \dots, \operatorname{Re}\sigma_{r_1+r_2}(x), \operatorname{Im}\sigma_{r_1+r_2}(x)\big).\end{aligned}$$

Further, we define the *house* of $x \in K$ by

$$\overline{|x|} := \max(|\sigma_1(x)|, \dots, |\sigma_d(x)|).$$

Lastly, we define $\|\mathbf{x}\|_\infty := \max_i |x_i|$ for $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$.

Lemma 4.23. *Let $\alpha \in O_K$ with $\|\varphi(\alpha)\|_\infty \leqslant \frac{2}{3}$. Then $\alpha = 0$.*

Proof. We prove that $|\sigma_i(\alpha)| < 1$ for $i = 1, \dots, n$. Then it follows that $|N_{K/\mathbb{Q}}(\alpha)| = \prod_{i=1}^d |\sigma_i(\alpha)| < 1$. But this is possible only if $\alpha = 0$, since otherwise, $N_{K/\mathbb{Q}}(\alpha)$ would be a non-zero integer, hence of absolute value at least 1.

Our assumption implies that $|\sigma_i(\alpha)| \leqslant \frac{2}{3} < 1$ for $i = 1, \dots, r_1$. Further, for $i = r_1 + 1, \dots, r_1 + r_2$ the real and imaginary parts of $\sigma_i(\alpha)$ have absolute values at most $\frac{2}{3}$. But this is then also true for their complex conjugates, which are $\sigma_i(\alpha)$ for $i = r_1 + r_2 + 1, \dots, r_1 + 2r_2 = d$. Hence $|\sigma_i(\alpha)| \leqslant \sqrt{(2/3)^2 + (2/3)^2} < 1$ for $i = r_1 + 1, \dots, d$. □

We consider again systems (4.18), but now the coefficients a_{ij} are from O_K.

Theorem 4.24. *Let $[K : \mathbb{Q}] = d$, let M, N be integers with $N > dM > 0$, let A be a real $\geqslant 1$, and suppose that*

$$a_{ij} \in O_K, \quad \overline{|a_{ij}|} \leqslant A \quad \text{for } i = 1, \dots, M, \; j = 1, \dots, N.$$

Then the system

$$(4.18) \qquad \left\{\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1N}x_N & = 0 \\ \vdots & & & & & \vdots \\ a_{M1}x_1 & + & \cdots & + & a_{MN}x_N & = 0 \end{array}\right.$$

has a solution $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{Z}^N \setminus \{\mathbf{0}\}$ *such that*

$$\max_{1 \leqslant i \leqslant N} |x_i| \leqslant (3NA)^{dM/(N-dM)}. \tag{4.19}$$

Proof. Write $l_i(\mathbf{x}) := a_{11}x_1 + \cdots + a_{1N}x_N$ for $i = 1, \dots, M$ and define the linear map $\psi : \mathbb{Z}^N \to \mathbb{R}^{Md}$ by

$$\psi(\mathbf{x}) := \big(\varphi(l_1(\mathbf{x})), \dots, \varphi(l_M(\mathbf{x}))\big).$$

This is well defined since $l_i(\mathbf{x}) \in O_K$ for $i = 1, \dots, M$. Note that for $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{Z}^N$, and for an embedding σ of K in $\mathbb{C}$ we have $\sigma(l_i(\mathbf{x})) = \sum_{j=1}^N \sigma(a_{ij})x_j$. Thus, we see that the components of $\psi(\mathbf{x})$ are linear expressions $b_1x_1 + \cdots + b_Nx_N$, where each coefficient b_j is either $\sigma_k(a_{ij})$ or the real or imaginary part of $\sigma_k(a_{ij})$, for some $k \in \{1, \dots, d\}$, $i \in \{1, \dots, M\}$. So clearly, all b_j have absolute value at most A.

Let B be a positive integer to be specified later and consider the vectors $\mathbf{y} \in S_B := \{0, \dots, B\}^N$. If $\mathbf{y} \in S$ then all coordinates of $\psi(\mathbf{y})$ have absolute value at most NAB, that is, $\psi(\mathbf{y})$ lies in the cube $[-NAB, NAB]^{Md}$.

We can partition this cube into $(3NAB)^{Md}$ small cubes of side length $2/3$. Now suppose that the cardinality of S_B is larger than the number of small cubes, that is,

$$(B+1)^N > (3NAB)^{Md}. \tag{4.20}$$

Then there must be distinct $\mathbf{y}_1$, $\mathbf{y}_2 \in S_B$ such that $\psi(\mathbf{y}_1)$, $\psi(\mathbf{y}_2)$ lie in the same small cube. Let $\mathbf{x} := \mathbf{y}_1 - \mathbf{y}_2$. Then

$$\|\psi(\mathbf{x})\|_\infty = \|\psi(\mathbf{y}_1) - \psi_(\mathbf{y}_2)\|_\infty \leqslant \frac{2}{3}.$$

This implies $\|l_i(\mathbf{x})\|_\infty \leqslant \frac{2}{3}$ for $i = 1, \dots, M$. Since $l_i(\mathbf{x}) \in O_K$, by Lemma 4.23 we have $l_i(\mathbf{x}) = 0$ for $i = 1, \dots, M$. Further, $\|\mathbf{x}\|_\infty \leqslant B$.

It is easy to see that (4.20) is satisfied with $B = [(3NA)^{Md/(N-Md)}]$. Our theorem follows. □

4.5 The Gel'fond-Schneider Theorem

We prove the following theorem.

Theorem 4.25. *Let α, β be real algebraic numbers such that $\alpha > 0$, $\alpha \neq 1$ and $\beta \notin \mathbb{Q}$. Then α^β is transcendental.*

Here $\alpha^\beta = e^{\beta \log \alpha}$ with the usual natural logarithm for positive real numbers. The proof in the case that α, β are not both real or $\alpha < 0$ goes along the same lines, but with additional complications. Gel'fond and Schneider independently proved the above theorem, in the general case where α, β may be complex, with different proofs. We follow Schneider's proof.

We assume that $\gamma := \alpha^\beta$ is algebraic. Let $K = \mathbb{Q}(\alpha, \beta, \gamma)$, $d = [K : \mathbb{Q}]$. Recall that there are positive integers m_1, m_2, m_3 such that $m_1\alpha, m_2\beta, m_3\gamma$ are algebraic integers. Let $m := m_1 m_2 m_3$; then $m\alpha, m\beta, m\gamma$ are algebraic integers.

Let D_1, D_2, L be parameters with values taken from the positive integers,, which will be chosen optimally later. In what follows, $c_1, c_2, \ldots$ will be constants depending only on α, β, γ, and will be independent of D_1, D_2, L.

Lemma 4.26. *Assume that $D_1 D_2 \geqslant 2dL^2$. Then there are integers a_{ij} ($i = 0, \ldots, D_1 - 1, j = 0, \ldots, D_2 - 1$), not all zero, such that the function*

$$F(z) = F_{L,D_1,D_2}(z) = \sum_{i=0}^{D_1-1} \sum_{j=0}^{D_2-1} a_{ij} z^i \alpha^{jz} \tag{4.21}$$

has zeros $a + b\beta$ with $a, b = 1, \ldots, L$, and such that

$$|a_{ij}| \leqslant \exp\big(c_1(D_1 \log L + D_2 L)\big) \quad (i = 0, \ldots, D_1 - 1,\ j = 0, \ldots, D_2 - 1). \tag{4.22}$$

Proof. We have to find $a_{ij} \in \mathbb{Z}$, not all zero, such that $F(a + b\beta) = 0$ for $a, b = 1, \ldots, L$. Using $\alpha^{a+b\beta} = \alpha^a \gamma^b$, this translates into a system of L^2 linear equations in the $D_1 D_2$ unknowns a_{ij}:

$$\sum_{i=0}^{D_1-1} \sum_{j=0}^{D_2-1} a_{ij} (a + b\beta)^i \alpha^{aj} \gamma^{bj} = 0 \quad (a, b = 1, \ldots, L).$$

To apply Siegel's Lemma we want all coefficients of this system of equations to be algebraic integers. To this end, we multiply the equations with $m^{D_1 + 2LD_2}$ and obtain

$$\sum_{i=0}^{D_1-1} \sum_{j=0}^{D_2-1} a_{ij} m^{D_1+2LD_2} (a + b\beta)^i \alpha^{aj} \gamma^{bj} = 0 \quad (a, b = 1, \ldots, L). \tag{4.23}$$

Then indeed, the coefficients of system (4.23) are all algebraic integers. We estimate their houses. Put

$$H := 1 + \overline{|\alpha|} + \overline{|\beta|} + \overline{|\gamma|}.$$

Take a typical coefficient of (4.23), say $m^{D_1+2LD_2}(a+b\beta)^i\alpha^{aj}\gamma^{bj}$. Let $\sigma : K \hookrightarrow \mathbb{C}$ be an embedding of K. Then for the image under σ of this coefficient we have

$$\begin{aligned}
&|m^{D_1+2LD_2}(a+b\sigma(\beta))^i\sigma(\alpha)^{aj}\sigma(\gamma)^{bj}| \\
&\quad \leqslant m^{D_1+2LD_2}\big(L(1+|\sigma(\beta)|)\big)^{D_1}(1+|\sigma(\alpha)|)^{LD_2}(1+|\sigma(\gamma)|)^{LD_2} \\
&\quad \leqslant m^{D_1+2LD_2}L^{D_1}H^{D_1+2LD_2} \leqslant \exp\big(c_2(D_1\log L + D_2L)\big)
\end{aligned}$$

where the constant c_2 has been chosen large enough in terms of m, d and H. Now clearly the houses of the coefficients of system (4.23) are all bounded above by $\exp\big(c_2(D_1\log L + D_2L)\big)$. We are now in a position to apply Theorem 4.24, and conclude that system (4.23) has a solution in integers a_{ij}, not all zero, such that

$$|a_{ij}| \leqslant \big(3D_1D_2e^{c_2(D_1\log L + D_2L)}\big)^{dL^2/(D_1D_2-dL^2)} \leqslant \exp\big(c_1(D_1\log L + D_2L)\big),$$

choosing c_1 sufficiently large. Here we have used our assumption $D_1D_2 \geqslant 2dL^2$. □

We now choose the parameters D_1, D_2, L such that $D_1D_2 = 2dL^2$ and $D_1 = D_2L$ (to make $D_1\log L$ and D_2L about equal), i.e.

$$D_1 = \sqrt{2d}\cdot L^{3/2}, \quad D_2 = \sqrt{2d}\cdot L^{1/2} \tag{4.24}$$

(for instance, take $L = 2dM^2$, $D_1 = (2d)^2M^3$, $D_2 = 2dM$ for some positive integer M). Then the estimate in Lemma 4.26 becomes

$$|a_{ij}| \leqslant \exp\big(c_3L^{3/2}\log L\big). \tag{4.25}$$

We note that $F(z)$ is a so-called *exponential polynomial*, i.e., a function of the shape

$$E(z) = \sum_{k=1}^{r} p_k(z)e^{\gamma_k z},$$

where the $p_k(z)$ are non-zero polynomials, and the γ_k distinct numbers. We need a simple result on the number of zeros of such a function.

Lemma 4.27. *Assume that the γ_k and the coefficients of the p_k are all reals. Put $M := \sum_{k=1}^{r}(1+\deg p_k)$. Then $E(z)$ has at most $M-1$ zeros in $\mathbb{R}$.*

Exercise 4.6. *Prove this lemma.*
Hint. *Proceed by induction on M. Apply Rolle's Theorem, which asserts that if G is a differentiable real function and a, b are reals with $a < b$ and $G(a) = G(b) = 0$, then there is c with $a < c < b$ and $G'(c) = 0$.*

Notice that we can apply this lemma to our above function $F(z)$, thanks to our assumption that α, β are real and $\alpha > 0$. Thus, this lemma implies that $F(z)$ has at most $D_1 D_2 = 2dL^2$ zeros. We know already that $F(z)$ has the L^2 zeros $a + b\beta$ $(1 \leqslant a, b \leqslant L)$. These zeros are all different, since $\beta \notin \mathbb{Q}$.

We briefly sketch the idea how to derive a contradiction from this. Details are provided later. Here it is important that we have some freedom to choose the parameters D_1, D_2, L introduced above. Thus, we can choose L sufficiently large to make all estimates work.

Let $c := 1+[\sqrt{2d}]$. We show that for all sufficiently large L, we have $F(a+b\beta) = 0$ for all integers a, b with $1 \leqslant a, b \leqslant cL$. Thus, F has at least $c^2L^2 > 2dL^2 = D_1 D_2$ zeros, which contradicts Lemma 4.27.

To prove that $F(a+b\beta) = 0$ for all integers a, b with $1 \leqslant a, b \leqslant cL$, we proceed as follows. For such a, b, the number $A := m^{D_1+2D_2cL} F(a + b\beta)$ is an algebraic integer. Using an analytic argument, we show that $|A|$ is very small. Further, by a trivial estimate we show that if σ is an embedding of K different from the identity, then $|\sigma(A)|$ is not too large. It will follow that $|N_{K/\mathbb{Q}}(A)| = |\prod_\sigma \sigma(A)| < 1$, where the product is over all embeddings, the identity included. But then, $F(a+b\beta) = A = 0$, since the norm of a non-zero algebraic integer is a non-zero element of $\mathbb{Z}$.

We now work out the details. We need a few facts from complex analysis. Recall that an *entire function* is a function $f : \mathbb{C} \to \mathbb{C}$ that is everywhere analytic, i.e., $f'(z) = \lim_{w\to z} \frac{f(w)-f(z)}{w-z}$ exists for every $z \in \mathbb{C}$. The following two lemmas are standard, and their proofs can be found in any textbook on complex analysis.

Lemma 4.28. *Let f be an entire function and $a \in \mathbb{C}$ a zero of f. Then there is an entire function g such that $f(z) = g(z) \cdot (z - a)$ for $z \in \mathbb{C}$.*

Lemma 4.29 (Maximum Modulus Principle)**.** *Let f be an entire function. For $R > 0$, define*

$$|f|_R := \sup_{z\in\mathbb{C}, |z|=R} |f(z)|.$$

Then for every $z \in \mathbb{C}$ with $|z| \leqslant R$ we have $|f(z)| \leqslant |f|_R$, i.e., $|f(z)|$ attains its maximum on the disk $|z| \leqslant R$ on the boundary of that disk.

As a consequence of these two lemmas we obtain the following estimate, which implies that if an entire function has many zeros in a disk $|z| \leqslant R$, then it is everywhere small on that disk.

Lemma 4.30. *Let f be an entire function and $a_1, \dots, a_r$ distinct zeros of f. Let R, T be reals such that $|a_i| \leqslant R$ for $i = 1, \dots, r$ and $T \geqslant 3R$. Then*

$$|f(z)| \leqslant |f|_T \big(3R/T\big)^r \quad \textit{for all } z \in \mathbb{C} \textit{ with } |z| \leqslant R.$$

Proof. By Lemma 4.28, there is an entire function g such that

$$f(z) = g(z)(z - a_1) \cdots (z - a_r) \quad \text{for } z \in \mathbb{C}.$$

Let $z \in \mathbb{C}$ with $|z| \leqslant R$. On the one hand, by Lemma 4.29,

$$|f(z)| \leqslant |g(z)| \prod_{i=1}^{r} (|z| + |a_i|) \leqslant |g(z)|(2R)^r \leqslant |g|_T (2R)^r,$$

on the other hand, we have for $w \in \mathbb{C}$ with $|w| = T$,

$$|g(w)| = \frac{|f(w)|}{|w - a_1| \cdots |w - a_r|} \leqslant |f(w)| \cdot (3/2T)^r,$$

since $|w - a_i| \geqslant |w| - |a_i| \geqslant T - R \geqslant \frac{2}{3}T$. Hence $|g|_T \leqslant |f|_T (3/2T)^r$. Our lemma follows. □

Proof of Theorem 4.25. Let $c := 1 + [\sqrt{d}]$ and put

$$R := (1 + |\beta|)cL, \quad T := 3eR = 3e(1 + |\beta|)cL.$$

Choose integers a, b with $1 \leqslant a, b \leqslant cL$. Then $a + b\beta$ lies inside the disk $|z| \leqslant R$. We first estimate $|F(a + b\beta)|$ by applying Lemma 4.30. A simple application of the triangle inequality gives

$$|F|_T \leqslant \sum_{i=0}^{D_1 - 1} \sum_{j=0}^{D_2 - 1} |a_{ij}| T^i |(1 + |\alpha|)^{Tj} \leqslant D_1 D_2 \exp(c_3 L^{3/2} \log L) \cdot T^{D_1} (1 + |\alpha|)^{T D_2}.$$

Here we have used $|a_{ij}| \leqslant \exp(c_3 L^{3/2} \log L)$ for all i, j. Using our choices $D_1 = \sqrt{2d}L^{3/2}$, $D_2 = \sqrt{2d}L^{1/2}$, $T = 3e(1+|\beta|)cL$, we see that all terms have exponent of order at most $L^{3/2} \log L$. We thus obtain

$$|F|_T \leqslant \exp\big(c_4 L^{3/2} \log L\big).$$

Recall that by its very construction, F has the L^2 distinct zeros $u + v\beta$ with $u, v = 1, \ldots, L$ inside the disk $|z| \leqslant R$. So by Lemma 4.30, using that $3R/T = e^{-1}$,

$$|F(a+b\beta)| \leqslant |F|_T e^{-L^2} \leqslant \exp\big(c_4 L^{3/2} \log L - L^2\big).$$

We multiply $F(a+b\beta)$ with $m^{D_1+2D_2cL}$ to obtain an algebraic integer, i.e.,

$$A(a,b) = A_{D_1,D_2,L}(a,b) := m^{D_1+2D_2cL} \sum_{i=0}^{D_1-1} \sum_{j=0}^{D_2-1} a_{ij}(a+b\beta)^i(\alpha^a\gamma^b)^j.$$

Clearly, $m^{D_1+2D_2cL} \leqslant \exp\Big((\sqrt{2d} \cdot L^{3/2} + 2\sqrt{2d}L^{1/2} \cdot cL) \cdot \log m\Big) \leqslant \exp\big(c_5 L^{3/2}\big)$. Hence

$$|A(a,b)| \leqslant \exp\Big(c_6 L^{3/2} \log L - L^2\Big).$$

We estimate the absolute values of the other conjugates of $A(a,b)$. Let σ be an embedding of $K = \mathbb{Q}(\alpha, \beta, \gamma)$ in $\mathbb{C}$. Let $H := 1 + \overline{|\alpha|} + \overline{|\beta|} + \overline{|\gamma|}$. Then by the triangle inequality,

$$\begin{aligned} |\sigma(A(a,b))| &\leqslant m^{D_1+2D_2\cdot cL} \sum_{i=0}^{D_1-1} \sum_{j=0}^{D_2-1} |a_{ij}|(a+b|\sigma(\beta)|)^i(|\sigma(\alpha)|^a|\sigma(\gamma)|^b)^j \\ &\leqslant m^{D_1+2D_2\cdot cL} D_1 D_2 \cdot \exp\big(c_3 L^{3/2} \log L\big) \cdot (cL)^{D_1} H^{D_1+2D_2\cdot cL} \\ &\leqslant \exp\Big(c_7 L^{3/2} \log L\Big). \end{aligned}$$

So altogether we have an upper bound for $|A(a,b)|$ and upper bounds for $|\sigma(A(a,b))|$ for the other $d-1$ embeddings σ of K not equal to the identity. By taking their product we obtain

$$|N_{K/\mathbb{Q}}(A(a,b))| = |N_{K/\mathbb{Q}}(A_{D_1,D_2,L}(a,b)| \leqslant \exp\Big(c_6 L^{3/2} \log L + (d-1)c_7 L^{3/2} \log L - L^2\Big).$$

This estimate is valid for all positive integers L, D_1, D_2 with $D_1 = \sqrt{2d}L^{3/2}$, $D_2 = \sqrt{2d}L^{1/2}$ and all integers a, b with $1 \leqslant a, b \leqslant cL$. In the course of our argument, we did not impose any other restrictions on L, D_1, D_2. Now we choose L large enough, to make $L^2 > (c_6 + (d-1)c_7)L^{3/2} \log L$. Then $|N_{K/\mathbb{Q}}(A(a,b))| < 1$ for all $a, b = 1, \ldots, cL$. Since the norm of a non-zero algebraic integer is a non-zero rational integer, this must imply $A(a,b) = 0$, or equivalently, $F(a + b\beta) = 0$ for all $a, b = 1, \ldots, cL$. As explained above, this contradicts Lemma 4.27. Our proof of Theorem 4.25 is complete. □