UNIVERSITAT DE BARCELONA
Facultat de Matemàtiques i Informàtica

**Master in Advanced Mathematics**

**End of Master Thesis**

# Auxiliary Polynomials for Transcendence Results

**Author: Eduard Valcarce Dalmau**

**Director:** Dr. Martín Sombra
**Realized at:** Departament de Matemàtiques i Informàtica

**Barcelona,** September 2, 2024

## Abstract

The main goal of this work is to prove several transcendence results using auxiliary functions, and in doing so showcase their effectiveness in various contexts. The main theorems covered will be Hermite-Lindemann, Gelfond-Schneider, Schneider-Lang, and Baker's theorem. We will employ two different proof strategies with auxiliary polynomials: two similar ones for Hermite-Lindemann and Schneider-Lang, and a noticeably different one for Baker's theorem. Gelfond-Schneider will come as a corollary to Schneider-Lang.

We will ease into these theorems however, by first delving into the preliminary results and background knowledge requiered to understand their proofs. This includes but is not limited to derivations over number fields, valuation theory and height functions, and complex analysis. Furthermore, we will take a detour into ellipitic functions after proving the Schneider-Lang theorem due to independent interest, and to present a few applications of the Schneider-Lang theorem, as it is the most general one we will present.



2020 Mathematics Subject Classification. 11J86, 11J89, 11J81, 14H52.

## Acknowledgements

First of all, I would like to thank my tutor Dr. Martín Sombra for helping me narrow down the subject of this project, and supplying me with many books afterwards. He supported me in pursuing the endeavor of writing this project which had interested me for years.

I would also like to thank my parents for supporting me and being there when I needed them or simply wanted to have a chat. Just as important was my cat Luke, who always makes me smile and forces me to take a break every now and then.

Lastly, I want to thank my friends from university, the ones that I have kept in contact with from the double degree, as well as new ones from the master's degree. Particularly, I would like to thank my friends Àlex, Gerard, Oriol and Martín as they are the ones I bothered the most.
Thank you all.

## Agraïments

Primer de tot, voldria agrair al meu tutor, el Dr. Martín Sombra per ajudar-me a concretar el tema del TFM i deixar-me llibres per llegir sobre el tema. Gràcies a això he après molt sobre aquest tema que em portava anys intrigant.

També m'agradaria agrair als meus pares per donar-me suport i per estar allà tot i que nomes fos per xerrar. Igual d'important és el meu gat Luke, que sempre em posa un somriure a la cara i em fa prendre descansos ocasionalment.

Finalment, m'agradaría agrair als meus amics de la universitat, tant els de la carrera com els nous del màster. En particular m'agradaria agrair als meus amics Àlex, Gerard, Oriol i Martín.
Gràcies a tots.

# Contents

# Introduction

The possibility of certain classes of numbers not being algebraic dates back to 1748, when Euler *asserted* that $\log_a(b)$ was not algebraic for $a, b \in \mathbb{Q}$, with $b \neq a^c$ for any $c \in \mathbb{Q}$. Despite this, it would take almost a century until Liouville observed that algebraic numbers were "badly approximated" by rational numbers; this insight paved the way for a sufficient condition for a number to be transcendental, which was followed by the first concrete example of a transcendental number,

$$\sum_{n=1}^{\infty} \frac{1}{10^{n!}}.$$

These developments sparked some interest in the field, so it was not long before $e$ was proven to be transcendental by Hermite in 1873, notably by using a polynomial to mimic the function $e^x$. Shortly thereafter Lindemann expanded upon Hermite's idea to prove that for any non-zero algebraic number $\alpha$, $e^\alpha$ is transcendental, which proved the transcendence of $\pi$ due to Euler's formula. Despite these advances, the theory had not yet reached the more modern auxiliary polynomial techniques, and still relied on concrete polynomials. It was not until later that other renowned mathematicians such as Pólya began laying the foundation for the main proof technique we will present in this thesis, which likely inspired Gelfond in 1934 to prove the more famous Gelfond-Schneider theorem.

I well remember learning about transcendental numbers from a now 11 year old video and being fascinated with the concept, but it was not until meeting them again in the context of irrationality measures and Roth's theorem that I became truly interested in them. It was fun to recontextualize well known functions such as $\sin(x)$ as transcendental functions, and read about the many results within transcendental number theory, and their implied history from the many generalizations of previous results. In spite of all this, it was vexing to not know any of their proofs, and this latter part is what this thesis aims to address.

We will begin with some much needed preliminary results, and then give at least once transcendence result in each chapter. All of the main results will use proofs which rely in some way in auxiliary polynomials, and in the case of Hermite-Lindemann for example, we will purposefully not give the original proof, to instead give one that generalizes more easily and allows for a more seamless transition onto the Gelfond-Schneider theorem.

Chapter 1 will be all about laying the foundation for our proof of Hermite-Lindemann in chapter 2 with necessary results and definitions. We fittingly begin with the ever important Siegel's lemma, and move on to field extentions over the rationals to define the norm and trace of an algebraic element. A convenient way to bound algebraic elements will be necessary, so we also introduce the 'house' of an algebraic integer, which we use to close the chapter just as we started it, with Siegel's lemma, but this time for algebraic integers.

Chapter 2 is entirely about the proof of the Hermite-Lindemann theorem, which features three main lemmas that all work towards the goal of proving a non-zero polynomial has zeroes of infinite order. These three lemmas are the backbone of the main proof technique presented in this thesis, which we will employ again in chapter 3 for Gelfond-Schneider. We have also decided to include in this chapter a result that relates an upper

bound of a function $f$ with that of its derivatives over a disc of radius $R$. This is not so apparent now, but later in chapter 4 an important notion is that of the bounded growth order, which intuitively implies that part of the reason these functions are forced to take on transcendental values is the lack of algebraic values in the image of the function.

Chapter 3 is similar to chapters 1 and 2 together, as it begins with more groundwork, mainly the building up towards the Height of any algebraic number, as the house only works for algebraic integers which is not enough anymore. To reach heights, we begin by introducing valuations over field extensions $K/\mathbb{Q}$ and many results that arise from them, and then prove some properties of heights. Following a brief discussion of derivations over number fields, we move on to the proof of the Schneider-Lang theorem, and even stronger theorem than Gelfond-Schneider, which we use to prove the latter.

The Schneider-Lang theorem is in fact general enough that we then apply it to the Weirstrass $\wp$ function to obtain transcendence results, but before that we spend most of chapter 4 studying elliptic curves in general, and then specifically the properties of $\wp$ in order to apply the theorem.

Finally, chapter 5 is entirely devoted to the proof of Baker's theorem, which is a generalization of the previously seen Gelfond-Schneider theorem, and importantly also falls outside the scope of the Schneider-Lang theorem, which forces us to treat it separately. We take advantage of this by showing a different type of proof using auxiliary polynomials. Chapters 2 and 3 used a very similar method, but this proof relies on proving certain bounds for the value of the auxiliary polynomial and its derivatives at 0, and then showing there must be at least one incorrect bound within those.

# 1 Moving away from the rationals

Despite the title of this chapter, we feel obliged to begin with this classic lemma - the foundation of many results in diophantine approximation and transcendental number theory - which allows us in many instances to prove the existence of a polynomial with desirable properties. Before that, however, we must define the absolute value for vectors and matrices.

**Definition 1.1.** Let $A = (a_{ij})$ be a matrix with coefficients in $\mathbb{R}$, and let $z = (z_1, ..., z_n)$ be a vector with coefficients in $\mathbb{R}$, we define both of their absolute values to be, respectively

$$|A| = \max_{i,j}\{|a_{ij}|\} \qquad |z| = \max_{i}\{|z_i|\}.$$

**Lemma 1.2** (Siegel's Lemma)**.** Given an $M \times N$ matrix $A = (a_{ij})$ with coefficients in $\mathbb{Z}$, not all 0, and assuming $N > M$, then there exists a vector $z \in \mathbb{Z}^N$ satisfying

$$Az = 0, \qquad z \neq 0, \qquad |z| \leq (N|A|)^{\frac{M}{N-M}}$$

**Proof.** Let $Z = (N|A|)^{\frac{M}{N-M}}$, and $z = (z_1, ..., z_N) \in \mathbb{Z}^N$ be any vector such that $0 \leq z_i \leq Z$ $\forall i \in \{1, ..., N\}$. Given $|a_{jk}| \leq |A|, 0 \leq z_i \leq Z$, and the matrix is $M \times N$, $Az$ takes at most $(N|A|Z + 1)^M$ distinct values. This is because the number of values each entry in the vector $Az$ can take is bounded by $|Az_1 - Az_2| + 1 \leq |A|N|z_1 - z_2| + 1 \leq N|A|Z + 1$, and $Az \in \mathbb{Z}^M$. Now, due to the choice of $Z$ we have $(N|A|Z + 1)^M \leq (N|A|)^M(Z + 1)^M = Z^{N-M}(Z+1)^M < (Z+1)^N$, meaning the cardinality of the set of distinct possible vectors $z$ is strictly bigger than the cardinality of the set of distinct values $Az$ can take. Hence, there exist $z^{(1)} \neq z^{(2)}$ such that $Az^{(1)} = Az^{(2)}$, and $z := z^{(1)} - z^{(2)}$ satisfies the conditions of the lemma.

□

When proving the irrationality and even transcendence of numbers, a frequently exploited fact is that every non-zero integer has absolute value bigger or equal to 1. This is no longer the case when dealing with algebraic integers, as for instance $N = (\sqrt{5} - 2)^n \in \mathbb{Z}[\sqrt{5}]$ clearly doesn't satisfy $|N| \geq 1$ for all $n \geq 1$. Hence, it is in our best interest to develop a similar function for these algebraic integers that carries over nice properties that resemble how the absolute value works over the integers. To this end, we introduce the norm and trace, which will be used to define this measure, and later on to obtain a more general version of Siegel's lemma for algebraic integers.

**Definition 1.3.** Let $L/K$ be a finite dimensional field extension of degree $n$. The trace and norm of an element $x \in L$ are defined to be the trace and determinant, respectively, of the endomorphism of the $K$-vector space $L$

$$T_x : L \to L, \qquad T_x(\alpha) = x\alpha$$
$$Tr_{L/K}(x) = \mathrm{Tr}(T_x) \qquad N_{L/K}(x) = \det(T_x).$$

Note that this definition of $N_{L/K}(x)$ implies $N_{L/K} : L \longrightarrow K$, as $L$ is a $K$-vector space, hence all the coefficients in the matrix representation of $T_x$ are in $K$, and therefore $\det(T_x) \in K$. We want to find a formula for $Tr_{L/K}(x)$ which we will use later when dealing

with the discriminant of a number field, which we will define later. We begin by looking at the characteristic polynomial of $T_x$

$$f_x(t) = \det(tI_d - T_x) = t^n - a_1 t^{n-1} + \cdots + (-1)^n a_n \in K[t]$$

And notice that

$$a_1 = Tr_{L/K}(x) \qquad\qquad a_n = N_{L/K}(x)$$

From this observation, we can derive the following proposition.

**Proposition 1.4.** If $L/K$ is a separable, finite dimensional extension and $\sigma : L \to \overline{K}$ varies over the different $K$-embeddings of $L$ into an algebraic closure $\overline{K}$ of $K$, then we have, for a given $x \in L$

i) $f_x$ (t) $= \prod_\sigma (t - \sigma(x))$

ii) $Tr_{L/K}(x) = \sum_\sigma \sigma(x)$

iii) $N_{L/K}(x) = \prod_\sigma \sigma(x)$

**Proof.**
Given that $L/K$ is a separable, finite field extension we want to prove that

$$f_x(t) = (p_x(t))^d, \qquad \text{with} \qquad d = [L : K(x)] = [L : K]/[K(x) : K]$$

where $p_x(t)$ is the minimal polynomial of $x$ over $K$, which we write as

$$p_x(t) = t^m + c_1 t^{m-1} + \cdots + c_m, \qquad \text{with} \qquad m = [K(x) : K].$$

Consequently $\{1, x, ..., x^{m-1}\}$ is a basis of $K(x)/K$, and if we call the basis of $L/K(x)$ $\{\alpha_1, ..., \alpha_d\}$, then it is a classic result that the basis of $L/K$ is

$$\alpha_1, x\alpha_1, ..., x^{m-1}\alpha_1; ...; \alpha_d, x\alpha_d, ..., x^{m-1}\alpha_d,$$

and in this basis we have

$$T_x \begin{pmatrix} \alpha_1 \\ x\alpha_1 \\ \vdots \\ x^{m-1}\alpha_d \end{pmatrix} = \begin{pmatrix} x\alpha_1 \\ x^2\alpha_1 \\ \vdots \\ (-c_1 x^{m-1} - \cdots - c_m)\alpha_d \end{pmatrix}.$$

It is easy to see then that the matrix of $T_x(y) = xy$ with respect to this basis is only made of the same block repeated d times all throughout the diagonal. This block is

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ -c_m & -c_{m-1} & -c_{m-2} & \cdots & -c_1 \end{pmatrix},$$

and one can easily verify its characteristic polynomial is $p_x(t)$, as it should be, so the characteristic polynomial of the big matrix is $f_x(t) = (p_x(t))^d$ as expected.

Now, $L/K$ is a separable finite field extension, so $L/K(x)$ is also a separable finite field extension, and by the primitive element theorem we can write $L = K(x, y)$ for some $y \in L$. From this, we know that all the $K$-embeddings of $L$ are uniquely determined by the images of $x$ and $y$. From the degrees of each field extension, we can then state that the equivalence relation

$$\tau \sim \sigma \iff \tau(x) = \sigma(x)$$

partitions $\mathrm{Hom}_K(L, \overline{K})$ into $m = [K(x) : K]$ equivalence classes of $d = [L : K(x)]$ elements each. If we call $\sigma_1, ..., \sigma_m$ the representatives of each equivalence class, we see that

$$p_x(t) = \prod_{i=0}^{m} (t - \sigma_i(x))$$

which implies

$$f_x(t) = (p_x(t))^d = \prod_{i=0}^{m} (t - \sigma_i(x))^d = \prod_{i=0}^{m} \prod_{\sigma \sim \sigma_i} (t - \sigma(x)) = \prod_{\sigma \in \mathrm{Hom}_K(L, \overline{K})} (t - \sigma(x)).$$

This proves i), ii), and iii) due to Vietà's equations, which we alluded to before the beginning of the proposition.

□

We now define the function we alluded to before, called 'house'.

**Definition 1.5.** Let $K/\mathbb{Q}$ be a finite dimensional field extension, for every $x \in K$ we define its size, or house as

$$||x|| = \max_{\sigma} |\sigma(x)|$$

where the maximum is taken over all the embeddings $\sigma : K \longrightarrow \mathbb{C}$.

With this, we can now give a result that gives us a lower bound for the house, and the absolute value of an algebraic integer.

**Proposition 1.6.** Let $K/\mathbb{Q}$ be a finite dimensional field extension with $[K : \mathbb{Q}] = D$, and let $\omega \neq 0$ be an element in $\mathbb{Z}_K$, which is the notation we will use for the algebraic integers contained in the field $K$. Then

$$||\omega|| \geq 1 \qquad \text{and} \qquad |\omega| \geq ||\omega||^{-(D-1)}.$$

And in fact, fixed $D \geq 2$, the exponent $D - 1$ in the inequality is best possible

**Proof.**

Recall from definition 1.3, that $N_{L/K} : L \longrightarrow K$, and looking at the field extension $K/\mathbb{Q}$, we have $N_{K/\mathbb{Q}}(\omega) \in \mathbb{Q}$. On the other hand, we know from proposition 1.4 that the Norm is equivalent to the product of all the roots of the minimal polynomial of $\omega$, all of which are algebraic integers because they are roots of the same polynomial, so the norm is the product of algebraic integers, which is an algebraic integer that we know belongs to $\mathbb{Q}$, therefore $N_{K/\mathbb{Q}}(\omega) \in \mathbb{Z}$, and since $\omega \neq 0$, $|N_{K/\mathbb{Q}}(\omega)| \geq 1$. Now, using the same formula for the norm again we have

$$1 \leq |N_{K/\mathbb{Q}}(\omega)| = \Big| \prod_{\sigma} \sigma(\omega) \Big| = \Big| \omega \cdot \prod_{\sigma \neq \mathrm{Id}} \sigma(\omega) \Big| \leq |\omega| \cdot ||\omega||^{D-1}. \tag{1.1}$$

From the definition of $||\cdot||$, and $[K:\mathbb{Q}] = D$, meaning there's $D$ embeddings. Now, from this inequality it is then easy to see that since $||\omega|| < 1 \implies |\omega| < 1$, but we have $1 \leq |\omega| \cdot ||\omega||^{D-1}$, we must have $||\omega|| \geq 1$, and rearranging, $|\omega| \geq ||\omega||^{-(D-1)}$, too. We now see that the exponent $D-1$ is best possible.
Fix $D \geq 2$ and consider the polynomial $P(X) = X^D - tX + 2$, where $t$ is any even integer greater than 4. Now, 2 divides all but the leading coefficient of $P(X)$, and clearly 4 does not divide 2, so by Eisenstein's criterion $P(X)$ is irreducible, furthermore, it has a root $\omega = \omega_t$ satisfying $0 < \omega < 4/t$, as $P(0) = 2$ and $P(4/t) = (4/t)^D - 2 < 0$. We will consider this root $\omega$ for the rest of the proof as it will be necessary, and we will now show that $||\omega||^{D-1} \leq 2t$, by considering $y \in \overline{\mathbb{Q}}$ such that $|y^{D-1}| > 2t$, and proving that this restriction on $y$ implies it cannot be a root of the polynomial $P$:

$$|P(y) - 2| = |y| \cdot |(y^{D-1} - t)| > (2t)^{1/(D-1)}(|y^{D-1}| - t) > t \cdot (2t)^{1/(D-1)} > 4.$$

Clearly then $|P(y)| > 2$, and therefore $y$ can never be a root of $P$, so $||\omega||^{D-1} \leq 2t$. We can now use the bounds on the house and absolute value of $\omega$ together to show that $D-1$ is the best possible exponent. We will assume this isn't the case, hence there is some $\varepsilon > 0$ such that for every non-zero alebraic integer $\omega$, equation (1.1) looks like this instead:

$$1 \leq |\omega| \cdot ||\omega||^{D-1-\varepsilon}$$

However, from our bounds for $|\omega|$ and $||\omega||$ we have

$$1 \leq |\omega| \cdot ||\omega||^{D-1-\varepsilon} \leq (4/t) \cdot (2t)^{(D-1-\varepsilon)/(D-1)} \leq 8t^{-\varepsilon/(D-1)}.$$

Recalling that $t$ is any even integer, and can thus be as large as we desire, we reach a contradiction, as we have found an algebraic integer $\omega$ such that $|\omega| \cdot ||\omega||^{D-1-\varepsilon} < 1$ instead.

□

Now, choosing to remain in the realm of field extensions, we have the following definition. The concept of the discriminant will come up later when we extend Siegel's lemma to the algebraic integers.

**Definition 1.7.** Let $L/K$ be a finite dimensional field extension of degree $n$, and let $\beta_1, ..., \beta_n$ be elements of $L$. Their discriminant is defined to be

$$D(\beta_1, ..., \beta_n) = \det(Tr_{L/K}(\beta_i \beta_j))$$

From this definition, and using the previous lemma we now obtain a different formula for this discriminant.

**Proposition 1.8.** Let $K$ be a field of characteristic 0, and $L/K$ a finite dimensional field extension of degree $n$, further, let $\sigma_1, ..., \sigma_n$ be the $n$ distinct $K-$embeddings from $L$ onto $\overline{K}$. Then, for every basis $\{\beta_1, ..., \beta_n\}$ of $L$ as a $K-$vector space we have

$$D(\beta_1, ..., \beta_n) = \det(\sigma_i(\beta_j))^2 \neq 0$$

**Proof.**

We begin by proving the first equality:

$$\begin{aligned} D(\beta_1, ..., \beta_n) &= \det(Tr_{L/K}(\beta_i\beta_j)) \\ &= \det\left(\sum_{k=1}^{n} \sigma_k(\beta_i\beta_j)\right) \quad \text{by proposition 1.4} \\ &= \det\left(\sum_{k=1}^{n} \sigma_k(\beta_i)\sigma_k(\beta_j)\right) \\ &= \det(\sigma_k(\beta_i)) \cdot \det(\sigma_k(\beta_j)) \quad \text{(Transposing any one of the matrices)} \\ &= \det(\sigma_k(\beta_i))^2. \end{aligned}$$

And we will now show that this determinant can never be equal to 0. Since $L/K$ is a separable field extension, by the primitive element theorem, we can write $L = K(\alpha)$ for some $\alpha \in L$, that is, $L$ has a power basis, and in fact any basis $\{\beta_1, ..., \beta_n\}$ of $L$ as a $K-$vector space will be a linear combination of this power basis, i.e

$$M_{\mathbb{Q}} \begin{pmatrix} 1 \\ \alpha \\ \vdots \\ \alpha^{n-1} \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$$

for some $n \times n$ matrix $M_{\mathbb{Q}}$ with all entries in $\mathbb{Q}$. Because of this we have, transposing the matrices:

$$\det \begin{pmatrix} \sigma_1(\beta_1) & \cdots & \sigma_n(\beta_1) \\ \sigma_1(\beta_2) & \cdots & \sigma_n(\beta_2) \\ \vdots & \ddots & \vdots \\ \sigma_1(\beta_n) & \cdots & \sigma_n(\beta_n) \end{pmatrix} = \det(M_{\mathbb{Q}}) \cdot \det \begin{pmatrix} 1 & \cdots & 1 \\ \sigma_1(\alpha) & \cdots & \sigma_n(\alpha) \\ \vdots & \ddots & \vdots \\ \sigma_1(\alpha)^{n-1} & \cdots & \sigma_n(\alpha)^{n-1} \end{pmatrix}.$$

Where the rightmost determinant is the determinant of a transposed Vandermonde matrix, in this case equal to

$$\prod_{1 \le i < j \le n} (\sigma_j(\alpha) - \sigma_i(\alpha))$$

Since $M_{\mathbb{Q}}$ is essentially the change of variable matrix between two bases, its determinant is non-zero, and since $L/K$ is a separable field extension, the above product is non-zero too, so

$$\det(\sigma_k(\beta_i))^2 = \left(\det(M_{\mathbb{Q}}) \prod_{1 \le i < j \le n} (\sigma_j(\alpha) - \sigma_i(\alpha))\right)^2 \neq 0$$

as desired.

□

**Lemma 1.9** (Siegel's lemma for algebraic integers)**.** Let $K/\mathbb{Q}$ be a finite dimensional field extension with $[K : \mathbb{Q}] = D$, and consider its ring of algebraic integers, $\mathbb{Z}_K$, with basis elements $\omega_1, ..., \omega_D$, so we can write $\mathbb{Z}_K = \mathbb{Z}\omega_1 + \cdots + \mathbb{Z}\omega_D$. For positive integers $M$, $N$ with $DM < N$ and real $W \geq 1$, let $w_{mn}$ $(m = 1, \ldots, M; n = 1, \ldots, N)$ be elements of $\mathbb{Z}_K$ with $\|w_{mn}\| \leq W$. Then there are rational integers $x_1, \ldots, x_N$, not all zero, with

$$\sum_{n=1}^{N} w_{mn}x_n = 0 \qquad (m = 1, \ldots, M)$$

and

$$|x_n| \leq (D!\|\omega_1\| \cdots \|\omega_D\| NW)^{\frac{DM}{N-DM}} \qquad (n = 1, \ldots, N).$$

**Proof.**
Every $w \in \mathbb{Z}_K$ can be written as $w = \sum_{i=1}^{D} u_i w_i$ for integers $u_i$, so in particular we will do this for every $\omega_{mn}$ from the hypothesis, until we mention going back to our original $M$ equations from the hypothesis. We will drop the subindexes $mn$ for ease. Now, applying all of the $D$ embeddings, yields the following system of equations

$$\begin{pmatrix} \sigma_1(w_1) & \cdots & \sigma_1(w_D) \\ \sigma_2(w_1) & \cdots & \sigma_2(w_D) \\ \vdots & \ddots & \vdots \\ \sigma_D(w_1) & \cdots & \sigma_D(w_D) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_D \end{pmatrix} = \begin{pmatrix} \sigma_1(w) \\ \sigma_2(w) \\ \vdots \\ \sigma_D(w) \end{pmatrix}.$$

Now, the basis of $\mathbb{Z}_K$ is also a basis of $K$ as a $\mathbb{Q}-$vector space, so by proposition 1.8 , the above matrix is non-singular and we can obtain a bound for the $|u_i|$ by solving the above equation using Cramer's rule. This gives, for every $j \in \{1, ..., D\}$:

$$|u_j| \leq |u_j| \cdot ||w_j|| \leq D! \cdot ||w_1|| \cdots ||w_D|| \cdot ||w|| \leq D! \cdot ||w_1|| \cdots ||w_D|| \cdot W$$

By bounding the sum of the $D!$ products. Now, we recall our original system of $M$ equations, and since $w_{mn} \in \mathbb{Z}_K$, separating the $w_{mn}$ into each component ends up giving us $DM$ equations, with the components being the bounded $u_i$, so applying Siegel's lemma here finally gives us

$$|x_n| \leq (D!\|\omega_1\| \cdots \|\omega_D\| NW)^{\frac{DM}{N-DM}} \qquad (n = 1, \ldots, N).$$

As desired.

□

## 2 The Hermite-Lindemann theorem

We are finally set to prove this theorem. We will first prove that $z, e^z$ are algebraically independent, and then show there exists a polynomial such that $P(z, e^z)$ is identically 0, that is, $z, e^z$ are algebraically dependent, which is a contradiction that arises from assuming $e^\alpha = \beta$ is algebraic. To show this polynomial exists, and is identically zero, we must bound its coefficients well enough, so that under suitable assumptions they can be forced to be 0. Doing this, we will inductively show that the polynomial must be identically 0.
For the rest of this chapter we will let $\alpha$ be a non-zero algebraic number, and assume $\beta = e^\alpha$ is algebraic to eventually reach a contradiction. Furthermore, we will be considering the number field $K = \mathbb{Q}(\alpha, \beta)$, which by our previous hypothesis is a finite field extension over $\mathbb{Q}$. We will also write $D = [K : \mathbb{Q}]$.

**Lemma 2.1.** The functions $z$ and $e^z$ are algebraically independent over $\mathbb{C}$.
**Proof.**
We want to show that there does not exist any non-zero polynomial $P(X, Y) \in \mathbb{C}[X, Y]$ such that $P(z, e^z)$ is equal to 0. We begin by assuming one such polynomial exists, and we will choose one such polynomial with minimal degree in $X$. We thus have

$$\sum_{i=0}^{n}\sum_{j=0}^{m} p_{ij}e^{iz}z^j = 0$$

which we can rewrite as

$$\sum_{i=0}^{n} e^{iz}\left(\sum_{j=0}^{m} p_{ij}z^j\right) = \sum_{i=0}^{n} e^{iz}Q_i(z) = Q_n(z)\sum_{i=0}^{n} e^{iz}R_i(z) = 0$$

for $Q_i \in \mathbb{C}[z]$ and $R_i = Q_i/Q_n \in \mathbb{C}(z)$, to conveniently obtain

$$e^{nz} + R_{n-1}(z)e^{(n-1)z} + ... + R_0(z) = 0. \tag{2.1}$$

Now, $e^{z+2\pi i} = e^z$ for all $z \in \mathbb{C}$, so replacing $z$ by $z + 2\pi i$ in the above equation yields another algebraic relationship:

$$e^{nz} + S_{n-1}(z)e^{(n-1)z} + ... + S_0(z) = 0 \tag{2.2}$$

with $S_i(z) = R_i(z + 2\pi i)$. Now subtracting (2.1) from (2.2) gives

$$(S_{n-1}(z) - R_{n-1}(z))e^{(n-1)z} + ... + (S_0(z) - R_0(z)) = 0. \tag{2.3}$$

If the leading coefficient is not equal to 0, we contradict the minimality of the polynomial's degree in $X$, so it must be 0, hence

$$Q_{n-1}(z + 2\pi i)/Q_n(z + 2\pi i) = Q_{n-1}(z)/Q_n(z). \tag{2.4}$$

Now, rewriting this as $F(z) = Q_{n-1}(z + 2\pi i)Q_n(z) = Q_{n-1}(z)Q_n(z + 2\pi i)$ and writing this polynomial as its product of linear factors, with $\alpha_i$ being the $r$ roots of $Q_{n-1}$, and $\beta_i$ being the $s$ roots of $Q_n$, we get

$$\begin{aligned} F(z) &= (z - (\alpha_1 - 2\pi i)) \cdots (z - (\alpha_r - 2\pi i))(z - (\beta_i)) \cdots (z - (\beta_s)) \\ &= (z - (\alpha_1)) \cdots (z - (\alpha_r))(z - (\beta_i - 2\pi i)) \cdots (z - (\beta_s - 2\pi i)). \end{aligned}$$

Since the coefficient of degree $deg(F(z)) - 1$ is fixed, this means that the number of $2\pi i$ terms in both representations is the same, that is, $2\pi i \cdot r = 2\pi i \cdot s$, therefore $r = s$ and $Q_n$ has the same degree as $Q_{n-1}$. This means that either $Q_n = Q_{n-1}$, or $Q_n$ does not divide $Q_{n-1}$. If $Q_n$ does not divide $Q_{n-1}$, because of (2.4), $R_{n-1} = Q_{n-1}/Q_n$ must have an infinite number of poles, as for any pole $\gamma$, $\gamma - 2\pi i$ would also be a pole. This cannot be the case for an element of $\mathbb{C}(z)$, therefore $Q_n = Q_{n-1}$.
We remain under the assumption that we are not contradicting the minimality of the original polynomial's degree in $X$, so we can apply the same reasoning as above to the coefficient of $e^{(n-2)z}$, and reach the same conclusions, until we reach the "constant" term. Since at this point $Q_n = Q_{n-1} = \cdots Q_1$, we have $S_0 = R_0$ too, and the same reasoning applies, so $Q_n = Q_{n-1} = \cdots Q_1 = Q_0$. This then implies that (2.1) becomes

$$e^{nz} + e^{(n-1)z} + ... + 1 = \frac{e^{(n+1)z} - 1}{e^z - 1} = 0$$

which is clearly false as $n \geq 1$, hence we have contradicted the minimality of the original polynomial's degree in $X$, and therefore $z$ and $e^z$ are algebraically independent.

□

For the following lemmas we will be considering the variables $L, S, T, T_1, S_1$. These will only be variables in the sense that they can vary to satisfy the various restrictions we will later impose upon them in order to prove the Hermite - Lindemann theorem, but otherwise in the context of the lemmas, they are all fixed quantities, even when we have to employ the lemmas repeatedly, we will consider the lemmas after fixing these values in order to get the desired results. On the other hand, the constants $c_i$ $i \in \mathbb{Z}$, are just constants that satisfy the inequalities, and are independent of the aforementioned "variables" $L, S, T, S_1, T_1$, though they depend on $\alpha$ and $\beta$. The same notation for the constants will be used for the rest of the thesis.

**Definition 2.2.** Given a function $\phi(z)$, analytic on a neighbourhood of $w \in \mathbb{C}$ and not identically 0 there, we will use the notation $\text{ord}_{z=w}\,\phi(z)$ for its (finite) order of vanishing at $w$. This is equivalent to saying that $\phi(z) = A_0(z-w)^g + A_1(z-w)^{g+1} + ...$ ($\phi$ is analytic on a neighbourhood of $w$) for $g = \text{ord}_{z=w}\,\phi(z)$ and $A_0 \neq 0$.

**Lemma 2.3.** For any $L \geq 2$, $S \geq 1$, $T \geq 1$ in $\mathbb{Z}$ with

$$(L+1)^2 \geq 2DST \tag{2.5}$$

there is $P$ in $\mathbb{Z}[X, Y]$, of degree at most $L$ in each variable, and with coefficients of absolute values at most $L^{3T}c_1^{LS}$ ($c_1 \in \mathbb{Z}$ for convenience later) such that

$$\phi(z) = P(z, e^z) \tag{2.6}$$

is not identically zero and

$$\text{ord}_{z=s\alpha}\,\phi(z) \geq T \quad (s = 1, \ldots, S). \tag{2.7}$$

**Proof.**
We again write $P(X, Y) = \sum_{i=0}^{L}\sum_{j=0}^{L} p_{ij}X^iY^j \in \mathbb{Z}[X, Y]$, and we have

$$\phi(z) = P(z, e^z) = \sum_{i=0}^{L}\sum_{i=0}^{L} p_{ij}z^ie^{jz}.$$

We then use Leibniz's rule on $\phi^{(t)}(s\alpha)$ to obtain the equations

$$\phi^{(t)}(s\alpha) = \sum_{i=0}^{L}\sum_{j=0}^{L} E_{ij}(s,t)p_{ij} = 0 \quad (s = 1,\dots,S;\ t = 0,1,\dots,T-1) \tag{2.8}$$

with

$$E_{ij}(s,t) = \sum_{u=0}^{t} \binom{t}{u} i(i-1)\cdots(i-u+1)(s\alpha)^{i-u} j^{t-u}\beta^{js}.$$

Now, to use the machinery we have presented thus far, we must make sure we are dealing with algebraic integers, so let $a, b \in \mathbb{N}$ be such that $a\alpha, b\beta \in \mathbb{Z}_K$, then $w_{ij} := a^L b^{LS} E_{ij}(s,t) \in \mathbb{Z}_K$ and bounding each term in $E_{ij}(s,t)$ gives

$$1 \leq ||w_{ij}|| \leq a^L b^{LS} \cdot 2^T L^T S^L \max\{1, ||\alpha||\}^L L^T \max\{1, ||\beta||\}^{LS} \leq c_2^{LS} L^{3T} = W,$$

where the max function is needed as $\alpha$ and $\beta$ may not be algebraic integers, and so their house may be smaller than 1. Further, $L \geq 2$ was used, as this bound (if it was not clear already) does not need to be too exact.
From the use of Leibnitz's rule we obtain $DST$ equations, with the known $w_{ij}$, and because of the double summation, we have $N = (L+1)^2$, using $N$ as we did in Siegel's lemma. Since $N \geq 2DST$, not only do we meet all the hypotheses to use Siegel's lemma for algebraic integers, we also know the exponent on the bound for the $|p_{ij}|$ will be $DST/((L+1)^2 - DST) \leq 1$, and in fact

$$|p_{ij}| \leq c_3(L+1)^2 W \leq c_1^{LS} L^{3T}.$$

Since $z$ and $e^z$ are algebraically independent, as shown in the previous lemma, $\phi(z)$ cannot be identically zero, which completes the proof.

□

We now use the foundation from the previous lemma to increase its scope into something we can actually use for our goals. The rough idea is that since $\phi^{(t)}(s\alpha)$ is small "up to" $t = T$ and $s = S$, we should be able to bound $\phi^{(t)}(s\alpha)$ a bit "further along". First, however, we prove a short lemma that we will use.
Before proceeding, we introduce the following notation: $|F|_R = \sup_{|z|=R} |F(z)|$ for $R \geq 0$.

**Lemma 2.4.** Let $\phi$ be an analytic function on an open set containing the disc $B(0,R) = \{z \in \mathbb{C} : |z| \leq R\}$, then, for $w \in \mathbb{C}$ such that $|w| < r < R$ we have

$$\frac{|\phi^{(m)}|_r}{m!} \leq \frac{|\phi|_R}{(R-r)^m}.$$

**Proof.**
First, $\phi(z)$ is by definition analytic on $B(0,R)$, so given $z \in B(0,R)$ and $w \in \mathbb{C}$ such that $|w| < r < R$, we can apply Cauchy's differentiation formula to obtain

$$\frac{\phi^{(m)}(w)}{m!} = \frac{1}{2\pi i}\int_{|z-w|=R-r} \frac{\phi(z)}{(z-w)^{m+1}} dz. \tag{2.9}$$

Now, from equation (2.9) we get

$$\begin{aligned}\left|\frac{\phi^{(m)}}{m!}\right|_r &\leq \frac{1}{2\pi}\left|\int_{|z-w|=R-r} \frac{\phi(z)}{(z-w)^{m+1}} dz\right|_r \\ &\leq \frac{1}{2\pi} \cdot \sup_{|w|=r} \int_{|z-w|=R-r} \left|\frac{\phi(z)}{(z-w)^{m+1}}\right| |dz|\end{aligned}$$

And by the maximum modulus principle, since
$\{z \in \mathbb{C} \mid \exists w \in \mathbb{C} \text{ with } |w| = r : |z - w| = R - r\} \subset B(0, R)$, we finally obtain

$$\frac{|\phi^{(m)}|_r}{m!} \leq \frac{1}{2\pi} \cdot \int_{|z-w|=R-r} \frac{|\phi|_R}{(R-r)^{m+1}} |dz| = \frac{|\phi|_R}{(R-r)^m}.$$

□

**Lemma 2.5.** Suppose that for some integers $S_1 \geq S$ and $T_1 \geq T$ we have

$$\text{ord}_{z=s\alpha}\, \phi(z) \geq T_1 \quad (s = 1, \ldots, S_1),$$

where $\phi(z)$ is the same as in lemma 2.3. Then the derivatives up to order $2T_1 - 1$ satisfy

$$|\phi^{(t)}(s\alpha)| \leq 2^{-S_1 T_1} c_7^{T_1} T_1^{2T_1} L^{3T_1} c_6^{LS_1} \quad (s = 1, \ldots, 2S_1;\ t = 0, 1, \ldots, 2T_1 - 1).$$

**Proof.**
We begin by dividing by the minimum multiplicity of each zero. Because of this, the function

$$\Phi(z) = \frac{\phi(z)}{(z-\alpha)^{T_1} \cdots (z - S_1\alpha)^{T_1}}$$

is entire, and using the maximum modulus principle again, we have

$$|\Phi|_{4S_1|\alpha|} \leq |\Phi|_{11S_1|\alpha|},$$

which in turn gives

$$(5S_1\alpha)^{-S_1T_1} |\phi|_{4S_1|\alpha|} \leq |\Phi|_{4S_1|\alpha|} \leq |\Phi|_{11S_1|\alpha|} \leq (10S_1\alpha)^{-S_1T_1} |\phi|_{11S_1|\alpha|}.$$

Hence

$$|\phi|_{4S_1|\alpha|} \leq 2^{-S_1T_1} |\phi|_{11S_1|\alpha|}.$$

Bounding now $|\phi|_{11S_1|\alpha|}$ we get

$$\begin{aligned} |\phi|_{11S_1|\alpha|} = \sup_{|z|=11S_1|\alpha|} \left| \sum_{i=0}^{L} \sum_{i=0}^{L} p_{ij} z^i e^{jz} \right| &\leq (L+1)^2 |p_{ij}| \max\{1, |11S_1\alpha|\}^L \max\{1, |e^{11S_1\alpha}|\}^L \\ &\leq (L+1)^2 (c_1^{LS_1} L^{3T_1}) c_4^L c_5^{LS_1} \leq c_6^{LS_1} L^{3T_1}. \end{aligned}$$

We will now use this to bound the derivatives, using lemma 2.4 with $m = t$, $r = 2S_1|\alpha|$, and $R = 2r$:

$$|\phi^{(t)}|_{2S_1|\alpha|} \leq t! \frac{|\phi|_{4S_1|\alpha|}}{(2S_1|\alpha|)^t} \leq (T_1!) c_7^{T_1} |\phi|_{4S_1|\alpha|} \leq T_1^{2T_1} c_7^{T_1} |\phi|_{4S_1|\alpha|} \leq 2^{-T_1S_1} T_1^{2T_1} c_7^{T_1} c_6^{LS_1} L^{3T_1}.$$

Obtaining the desired inequality.

□

This bound may look daunting at first, but because we can control $T_1$ and $S_1$, the term $2^{-T_1S_1}$ will in fact bring the bound close as close as necessary to 0.

**Lemma 2.6.** Given two positive integers $S_1 \geq S$, $T_1 \geq T$, and each of

$$\xi_{ts} = \phi^{(t)}(s\alpha) \quad (s = 1, \ldots, 2S_1;\ t = 0, 1, \ldots, 2T_1 - 1),$$

we have either $\xi_{ts} = 0$ or

$$|\xi_{ts}| \geq L^{-6DT_1} c_8^{-LS_1}.$$

**Proof.**
Writing out the expression for $\xi_{ts}$ using Leibnitz's rule as in lemma 2.3, we argue in the same way that $w_{ts} = a^L b^{LS_1} \xi_{ts} \in \mathbb{Z}_K$. We can therefore use the house:

$$\begin{aligned} ||w|| &\leq (L+1)^2 ||a^L b^{LS_1} E_{ij}(s,t)|| \cdot ||p_{ij}|| \leq a^L b^{LS_1} (L+1)^2 (c_2^{LS_1} L^{3T_1})(c_1^{LS_1} L^{3T_1}) \\ &\leq c_8^{LS_1} L^{6T_1} \end{aligned}$$

and using proposition 1.6 with this bound for the house, we obtain that for non-zero $w$

$$|w| \geq (c_8^{-DLS_1}) L^{-6DT_1}.$$

Hence, either $\xi_{ts} = 0$ or $|\xi_{ts}| \geq L^{-6DT_1} c_8^{-LS_1}$ as desired, where here $c_8$ is a different constant, but we are running out of single digit subindices.

□

With this, the preparations are done to prove the main theorem of this chapter. As mentioned earlier, we will use the bounds for the coefficients of the polynomial, which will be easier by parametrizing the various parameters in terms of an integer $n$. Recall that the parameters must satisfy

- $S_1 \geq S$
- $T_1 \geq T$
- $(L+1)^2 \geq 2DST$.

Choosing initially

$$L = \lfloor n^{3/4} \rfloor \qquad S = \lfloor n^{1/4} \rfloor \qquad T = n \qquad S_1 = S \qquad T_1 = T$$

we satisfy every condition for every $n$ big enough in terms of $\alpha$ and $\beta$, as recall $D$ depends on them.

**Theorem 2.7** (Hermite - Lindemann)**.** Let $\alpha \neq 0$ be algebraic over $\mathbb{Q}$. Then $e^\alpha$ is transcendental.

**Proof.**
We begin by using lemma 2.3 to obtain the existence of a polynomial $P(X, Y) \in \mathbb{Z}[X, Y]$ such that $P(z, e^z) = \phi(z)$ is not identically 0, and $\operatorname{ord}_{z=s\alpha} \phi(z) \geq T$ for $s = 1, \ldots, S$. At the same time, taking $S_1$ and $T_1$ as above, by lemma 2.5 we have

$$|\phi^{(t)}(s\alpha)| \leq 2^{-S_1 T_1} c_7^{T_1} T_1^{2T_1} L^{3T_1} c_6^{LS_1} \quad (s = 1, \ldots, 2S_1;\ t = 0, 1, \ldots, 2T_1 - 1).$$

Introducing our chosen values for the parameters, we get, for all the above values of $t$ and $s$, and for some constant $c_9$

$$|\phi^{(t)}(s\alpha)| \leq 2^{-n\lfloor n^{1/4} \rfloor} c_9^n n^{2n} = \left( \frac{c_9 n^2}{2^{\lfloor n^{1/4} \rfloor}} \right)^n.$$

Importantly, for a large enough $n$, we can make $c_9 n^2/2^{\lfloor n^{1/4} \rfloor}$ as small as possible, and in particular, smaller than $L^{-6D}c_6$ for $(s = 1, \ldots, 2S_1;\, t = 0, 1, \ldots, 2T_1 - 1)$ :

$$|\phi^{(t)}(s\alpha)| \geq L^{-6DT_1} c_8^{-LS_1} \geq (L^{-6D}c_8)^n > \left( \frac{c_9 n^2}{2^{\lfloor n^{1/4} \rfloor}} \right)^n \geq |\phi^{(t)}(s\alpha)|$$

so by lemma 2.6, $\xi_{ts} = 0$ for $(s = 1, \ldots, 2S_1;\, t = 0, 1, \ldots, 2T_1 - 1)$. This means that now our polynomial in $z, e^z$, $\phi(z)$ actually satisfies

$$\mathrm{ord}_{z=s\alpha}\, \phi(z) \geq T_1 \quad (s = 1, \ldots, S_1).$$

But this time for $T_1 = 2T$ and $S_1 = 2S$, and applying lemma 2.5 to $\phi(z)$ again with this newfound information, and then lemma 2.5, with the same considerations as before regarding the bounds, we obtain that in fact $\phi(z)$ actually satisfies

$$\mathrm{ord}_{z=s\alpha}\, \phi(z) \geq T_1 \quad (s = 1, \ldots, S_1).$$

But this time for $T_1 = 4T$ and $S_1 = 4S$. Repeating this same process inductively, doubling $S_1$ and $T_1$ each time, proves that $\phi(z)$ actually has zeroes of infinite order, and is in fact identically zero, contradicting lemma 2.3, where it was proven $\phi(z)$ is not identically zero. This contradiction proves that our initial assumption that $e^\alpha$ is algebraic was untrue, and this completes the proof of the Hermite - Lindemann theorem.

□

# 3 Gelfond - Schneider

## 3.1 Prelude to the Theorem: Heights and the Derivation

As mentioned, we are looking at this method of proof for these theorems, because of their simple generalization onto much stronger results. Before picking up where we left off in the previous chapter, we must look at the various absolute values over $\mathbb{Q}$, and from there build up some theory to meet the theoretical demands for the Gelfond - Schneider theorem. For the Hermite - Lindemann theorem it was enough to define the house, which only required the usual absolute value, but now we need deeper insights to define the heights, which work in even more general settings. Absolute values need no introduction, but we will have to start with all the possible absolute values over $\mathbb{Q}$ and build up from there.

**Definition 3.1.** Given an absolute value $|\cdot| : K \to \mathbb{R}$ defined over a field $K$, we say it is a non archimedean absolute value if it satisfies the ultrametric inequality, sometimes also called the strong triangle inequality:

$$|x+y| \leq \max\{|x|, |y|\} \qquad \forall x, y \in K$$

**Definition 3.2.** Given a prime $p$, and $x \in \mathbb{Q}^*$, let $\mathrm{ord}_p(x)$ be the unique integer such that $x$ can be written as

$$x = p^{\mathrm{ord}_p(x)} \cdot \frac{a}{b} \qquad\qquad a, b \in \mathbb{Z}, \;\; p \nmid ab$$

If $x = 0$, we set $\mathrm{ord}_p(x) = \infty$ by convention. With this map $\mathrm{ord}_p : \mathbb{Q} \to \mathbb{Z}$, we can define, for each prime $p$, the $p-$adic absolute value of $x \in \mathbb{Q}$ as

$$|x|_p = p^{-\mathrm{ord}_p(x)}.$$

This set of absolute values, along with the usual absolute value, $|\cdot|_\infty = |\cdot|$, and the trivial absolute value, is in a way the complete list of all the absolute values that can be defined over $\mathbb{Q}$, as we will see shortly. Before that we present a proposition that allows us to meaningfully distinguish between truly different absolute values.

**Proposition 3.3.** If two absolute values $|\cdot|_1$ and $|\cdot|_2$ define the same topology, they are called equivalent. Moreover, if they are defined over a field $K$, $|\cdot|_1$ and $|\cdot|_2$, are equivalent if and only if there exists a real number $s \in \mathbb{R}_{>0}$ such that, for all $x \in K$

$$|x|_1^s = |x|_2.$$

**Proof.**
If $|x|_1^s = |x|_2$ clearly the topologies they define are identical subsets of $K$, so all that's left is proving the other direction.
We begin the assault with the fact that for any absolute value $|\cdot|$ over $K$, and any $x \in K$, the inequality $|x| < 1$ is the same as the sequence $(x^n)_{n\in\mathbb{N}}$ converging to 0 in the topology defined by $|\cdot|$, so if the topologies are equivalent we have

$$|x|_1 < 1 \iff |x|_2 < 1.$$

Now fix $y \in K$ such that $|y|_1 > 1$, and let $x \in K^*$. Then $|x|_1 = |y|_1^\alpha$ for some $\alpha \in \mathbb{R}$. Now consider the sequence $\left(\frac{m_i}{n_i}\right)_i \subset \mathbb{Q}$, with $n_i > 0$, which converges to $\alpha$ from above, so we have

$$|x|_1 = |y|_1^\alpha < |y|_1^{m_i/n_i} \implies \left|\frac{x^{n_i}}{y^{m_i}}\right|_1 < 1 \implies \left|\frac{x^{n_i}}{y^{m_i}}\right|_2 < 1.$$

Now, working in the limit, this implies $|x|_2 \leq |y|_2^\alpha$, and using the same procedure, but with a sequence tending to $\alpha$ from below, we get $|x|_2 \geq |y|_2^\alpha$, and therefore $|x|_2 = |y|_2^\alpha$ too. This was independent of $x$, and therefore true for all $x \in K^*$, hence

$$\frac{\log|x|_1}{\log|y|_1} = \frac{\log|x|_2}{\log|y|_2} =: s,$$

and $|x|_1 = |x|_2^s$. Further $|y|_1 > 1 \iff |y|_2 > 1$, so $s > 0$.

□

After being able to meaningfully group potentially different absolute values into the same equivalence group according to the topology they induce, the following well known theorem tells us that we have indeed found all the absolute values over $\mathbb{Q}$.

**Theorem 3.4** (Ostrowski's First Theorem)**.** Every non trivial absolute value $\|\cdot\|$ on $\mathbb{Q}$ is equivalent to $|\cdot|_p$ for some prime $p$, or $p = \infty$.

**Proof.**

[5, Theorem 1]

Let $\mathbb{P}$ be the set of all primes, we now know every non trivial absolute value over $\mathbb{Q}$ is equivalent to some $|\cdot|_p$ for $p \in \mathbb{P} \cup \{\infty\}$. We will be using explicit notation for these absolute values from now on: given a number field $K$ with characteristic 0, we will write $M_K$ for the set of every non trivial absolute value on $K$, modulo equivalent absolute values. Over $\mathbb{Q}$, for example, we have a bijection between the two sets, $M_\mathbb{Q}$ and $\mathbb{P} \cup \{\infty\}$.

Knowing $M_\mathbb{Q}$ is not be enough however, as we need to extend this theory to more general number fields. As we will see shortly, doing this involves extending the absolute values over $\mathbb{Q}$, so our choice of representative for the absolute values over $\mathbb{Q}$ will be relevant henceforth. We will choose the representatives to be

$$|x|_p = p^{-\mathrm{ord}_p(x)} \quad \forall p \in \mathbb{P} \qquad\qquad |x|_\infty = \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases}.$$

With this, we can now begin expanding upon the foundation we have established thus far. A *place* or *valuation* $v$ is an equivalence class of non-trivial absolute values acting on a number field $K$, where two non-trivial absolute values belong to the same equivalence class if and only if they define the same topology over $K$. The absolute value in the equivalence class determined by the place $v$ is denoted by $|\cdot|_v$. Given the field extension $L/K$ and $v$ a place of $K$, then any place $w$ of $L$ such that the restriction of $|\cdot|_w$ over $K$ is a representative of $v$, is said to lie over $v$, or equivalently, we say that $w$ extends $v$. This is written as $w|v$, due to the fact that non-archimedean places in number fields correspond to prime ideals [9, (8.5) Proposition].

**Definition 3.5.** A valued field, $(K, |\cdot|_v)$, is a field with a valuation on it, where $K$ is a field, and $v$ a valuation.

**Definition 3.6.** Given a valued field $(K, |\cdot|_v)$, we call $(a_n)_{n\in\mathbb{N}} \subset K$ a Cauchy sequence if for every $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that

$$|a_n - a_m|_v < \varepsilon \quad \forall n, m \geq n_0.$$

**Definition 3.7.** A valued field $(K, |\cdot|_v)$ is called complete if every Cauchy sequence $(a_n)_{n\in\mathbb{N}} \subset K$ converges to an element $a \in K$, that is

$$\lim_n |a_n - a|_v = 0.$$

Given a number field with an absolute value $(K, |\cdot|_K)$, a completion of $K$, which we denote now by $(L, |\cdot|_L)$, is a number field with an absolute value, complete as a metric space and for which there exists an embedding $i : K \to L$ such that $i(K)$ is dense in $L$, with $|x|_K = |i(x)|_L \ \forall x \in K$. This completion is unique up to isomorphism, so in fact, with our notation, we say the *completion* of $K$ with respect to a place $v$ is the extension field $K_v$ with a place $w$ such that:

- $w|v$.
- The topology induced by $w$ on $K_v$ is complete.
- $K$ is a dense subset of $K_v$ in that same topology induced by $w$ (In this case the embedding would just be the identity).

The completion exists and is unique up to isometric isomorphisms [6, Chapter XII, §2, proposition 2.1]. Frequently, due to abuse of notation, the place $w$ is denoted $v$ as well, and often we will also simply write $|\cdot|$ when refering to $|\cdot|_\infty$. Completions can be quite mysterious, but for the infinity place they are well understood, as illustrated in the following theorem.

**Theorem 3.8** (Ostrowski's Second Theorem)**.** Let $(K, |\cdot|)$ be a complete valued field, with $|\cdot|$ an archimedean absolute value. Then there is an isomorphism $\varphi$ from $K$ onto $\mathbb{R}$ or $\mathbb{C}$ satisfying

$$|\varphi(a)| = |a|^s$$

for all $a \in K$, and some fixed $s \in (0, 1]$.

**Proof.**
[9, Chapter 2, §4, (4.2) Theorem]

In the following proposition we begin to build towards the degree formula, which is crucial when defining the height of an algebraic number $\alpha$, as it is used to prove that it is independent of the number field $K$ that contains $\alpha$.

**Proposition 3.9.** Let $K$ be a field with a fixed non trivial absolute value $|\cdot|_v$, and let $L/K$ be a finite degree field extension generated by a single element $\zeta$ with monic minimal polynomial $f(t)$ over $K$. Suppose that this polynomial can be decomposed into

$$f(t) = f_1^{n_1}(t) \cdots f_r^{n_r}(t)$$

where $f_i(t) \in K_v[t]$ are different irreducible monic factors, then for each $i \in \{1, 2, ..., r\}$ there exists an injective homomorphism of field extensions over $K$

$$\begin{aligned} \phi_i : L &\to K_i \cong K_v[t]/(f_i(t)) \subset \mathbb{C}_v \\ \zeta &\to t \end{aligned}$$

Now, for every $i$, there is a unique extension $|\cdot|_i$ of $|\cdot|_v$, and each one of these extensions is pairwise non-equivalent. Furthermore, $K_i$ is the completion of L with respect to $|\cdot|_i$ and the embedding $\phi_i$. Finally, for any absolute value $|\cdot|_w$ extending $|\cdot|_v$ to $L$, there is a unique $i$ such that $|\cdot|_i$ restricted to $L$ is equal to $|\cdot|_w$.

**Proof.**
[2, Proposition 1.3.1.]

□

**Corollary 3.10** (Degree formula)**.** If $L$ is a finite dimensional, separable field extension of $K$, then

$$\sum_{w|v} [L_w : K_v] = [L : K].$$

**Proof.**
By the primitive element theorem, we can write $L = K(\alpha)$ for some element $\alpha \in L$. Let $P(X)$ be the minimal polynomial of this same $\alpha$ over $K$. Then, given that $L/K$ is a separable field extension, we can factorize $P(X)$ over $K_v$ as

$$P(X) = P_1(X) \cdots P_n(X)$$

with $P(X)_1, ..., P(X)_n$ different irreducible polynomials. Now, the embeddings $\sigma : L \hookrightarrow \mathbb{C}_v$ correspond to the maps of $\alpha$ to the roots of $P(X)$, so by proposition 3.9 we have that for each $w \in M_L$ that has been extended from a given $v \in M_K$, $[L_w : K_v] = \text{degree}(P_j(X))$, where $j$ is the unique place such that $|\cdot|_w$ is equal to the restriction of $|\cdot|_j$ to $L$. Therefore

$$[L : K] = deg(P(X)) = \sum_{i=0}^{n} deg(P_i(X)) = \sum_{w|v} [L_w : K_v].$$

□

From this result, we can see why the terms $[L_w : K_v]$ are important for a field extension $L/K$, and absolute values $w \in M_L$, $v \in M_K$, $w|v$. This motivates the definition of the *local degree of $L/K$ in $w$*, and the *normalized absolute value associated to $v$*. They are, respectively

$$n_w = [L_w : K_v] \qquad\qquad \|x\|_w = |x|_w^{n_w}$$

There will not be confusion between the house $||x||$, and the normalized absolute value $\|x\|_w$, as can be seen right here. We also note that we may write $\|\cdot\|$ when referring to $\|\cdot\|_\infty$.

We follow this with an important result on how an absolute value $v$ is extended from a complete field $K$ to a finite dimensional extension $K_v$, as we will need it to obtain a crucial relationship that we will need for the product formula.

**Theorem 3.11.** Let $(K, |\cdot|_v)$ be a complete valued field, with $|\cdot|_v$ a non trivial absolute value. Then $|\cdot|_v$ can be uniquely extended to an absolute value $|\cdot|_w$, of any given algebraic extension $L/K$, with $|\cdot|_w$ given by the formula

$$|x|_w = |N_{L/K}(x)|_v^{1/[L:K]}.$$

Moreover, if $L/K$ is a finite dimensional extension, then $(L, |\cdot|_w)$ is complete.

**Proof.**

We can reduce the general case to the finite dimensional case by taking the finite subextensions defined by the elements we are considering, and in fact every algebraic extension $L/K$ is the union of its finite subextensions.

We begin with the case wherein $|\cdot|_v$ is archimedean, and therefore by Ostrowski's second theorem, $K = \mathbb{R}$ or $\mathbb{C}$. $|N_{\mathbb{C}/\mathbb{R}}(z)| = |z\overline{z}| = |z|^2$, and $|N_{\mathbb{R}/\mathbb{R}}(z)| = |z|$, and we are essentially done with this case, so we will hereforward assume the valuation $|\cdot|_v$ to be non archimedean.

Let $\{u_1, ..., u_r\}$ be $K-$linearly independent elements of $L$, and let $\{a_n\}_n \subset L$ be a sequence of elements of the form $a_n = \sum_{i=1}^r a_{ni}u_i$, with the $a_{ni} \in K$. We will begin by showing that $\{a_n\}_n$ is a Cauchy sequence if and only if all the $\{a_{ni}\}_n ; i = 1, ..., r$ are Cauchy sequences. One direction of this is straightforward, if all the $\{a_{ni}\}_n$ are Cauchy sequences, then so is $\{a_n\}_n$. We will prove the opposite direction by induction on $r$, with the case $r = 1$ being obviously true. We will now look at the two possible cases for $\{a_{nr}\}_n$ in the $r^{\text{th}}$ inductive step. First assume $\{a_{nr}\}_n$ is Cauchy, then the sequence defined by $b_n = a_n - a_{nr}u_r = \sum_{i=1}^{r-1} u_i a_{ni}$ is Cauchy by induction hypothesis, and therefore $\{a_n\}_n$ is Cauchy too. Now we assume $\{a_{nr}\}_n$ is not Cauchy, by definition then there exists an $\varepsilon > 0$ such that for every $N \in \mathbb{Z}_{>0}$, there exist $p, q \in \mathbb{Z}$ such that $|a_{pr} - a_{qr}|_v > \varepsilon$. Since this is true for every $N$, we can build an arbitrarily large sequence $\{p_k, q_k\}_k$ of integers with $p_k < p_{k+1}$ and $q_k < q_{k+1}$ $\forall k$ such that $|a_{p_k r} - a_{q_k r}|_v > \varepsilon$ $\forall k$. From this bound we see that for every $k$, $(a_{p_k r} - a_{q_k r})^{-1}$ exists and belongs to $K$. We can therefore define the following sequence,

$$b_k = (a_{p_k r} - a_{q_k r})^{-1}(a_{p_k} - a_{q_k})$$

which is also Cauchy due to $\{a_n\}_n$ being Cauchy by hypothesis, and $(a_{p_k r} - a_{q_k r})^{-1} < \varepsilon^{-1}$. Furthermore, $\lim_n b_n = 0$ and

$$b_k = u_r + \sum_{i=1}^{r-1}(a_{p_k r} - a_{q_k r})^{-1}(a_{p_k i} - a_{q_k i})u_i = \sum_{i=1}^{r-1} b_{ki}u_i + u_r,$$

so $\left\{\sum_{i=1}^{r-1} b_{ki}u_i\right\}_n$ is a Cauchy sequence that tends to $-u_r$, and thus by induction hypothesis all the $\{b_{ni}\}_n$ are Cauchy sequences. Now, $K$ is complete with respect to $|\cdot|_v$ so $\lim_n b_{ni} = b_i \in K$, and we can take the limits of all the $b_{ni}$:

$$0 = \lim_n \sum_{i=1}^{r-1} b_{ki}u_i + u_r = \sum_{i=1}^{r-1} b_i u_i + u_r$$

which contradicts the $K-$linear independence of the $\{u_1, ..., u_r\}$, so $\{a_{nr}\}_n$ must be Cauchy. To finish this part of the proof, consider that by the $K-$linear independence of the $\{u_1, ..., u_r\}$, then $\lim_n a_{ni} = 0$ $\forall i$ $\iff$ $\lim_n a_n = 0$, so if we take $(u_1, ..., u_r)$ to be a $K-$basis of $L$, we see that $L$ is complete, as we can impose $r = [L : K]$. We keep

$r = [L : K]$, and now all that is left now is proving the existence and uniqueness of the extended absolute value.

We begin with the uniqueness, assuming the formula gives an absolute value. Let then $|x|_w = |N_{L/K}(x)|_v^{1/[L:K]}$. We will assume the equality does not hold, that is, there exists some $a = \sum_{i=1}^{r} a_i u_i$ such that $|a^r|_w \neq |N_{L/K}(a)|_v$, and we will further assume that $|a^r|_w < |N_{L/K}(a)|_v$ by substituting $a$ with $a^{-1}$ if necessary. Now let $b = a^r N_{L/K}(a)^{-1}$, so $|b|_w < 1$ and $N_{L/K}(b) = N_{L/K}(a)^r \cdot N_{L/K}(a)^{-r} = 1$. Since $|b|_w < 1$, the sequence $\{b^n\}_n$ tends to 0, but on the other hand if we write $b^n = \sum_{i=1}^{r} b_{ni} u_i$ , then $\lim_n b^n = 0$ implies, by the previous result, that $\lim_n b_{ni} = 0 \ \forall i$. Finally, the norm map $N_{L/K} : L \longrightarrow K$ is a polynomial map and therefore continuous, but $b_{ni} \to 0$ implies $1 = N(b^n) \to 0$, and this contradiction completes the proof.

We finish with the existence. For $|x|_w = |N_{L/K}(x)|_v^{1/r}$, it is clear that $|x|_w = 0 \iff x = 0$, and by the properties of the norm, $|x|_w |y|_w = |xy|_w \quad \forall x, y \in K$. Recall that the norm is the product over all the embeddings, which satisfy $\sigma(xy) = \sigma(x)\sigma(y)$ . For the ultrametric inequality, $|x+y|_w \leq \max\{|x|_w, |y|_w\}$, since we can assume at both terms are nonzero, we can divide by one of them, and all we have to prove is that

$$|x|_w \leq 1 \implies |x+1|_w \leq 1.$$

To prove this, consider the valuation ring of $K$, $\mathcal{V} = \{x \in K : |x|_v \leq 1\}$, and its integral closure in $L$, $\mathcal{O}$. From definition 1.3 and proposition 1.4, we can write

$$\mathcal{O} = \left\{x \in L : N_{L/K}(x) \in \mathcal{V}\right\},$$

so it suffices to prove that

$$|x|_w \leq 1 \implies |x+1|_w \leq 1,$$

which is true, hence $|x|_w = |N_{L/K}(x)|_v^{1/r}$ gives an absolute value over $L$.

□

Now, given $v \in M_{\mathbb{Q}}$ we have the following embeddings

$$\begin{array}{c} \mathbb{Q} \hookrightarrow \mathbb{Q}_v \hookrightarrow \overline{\mathbb{Q}_v} \hookrightarrow \mathbb{C}_v \\ |\cdot|_v \mapsto |\cdot|_{v_1} \mapsto |\cdot|_{v_2} \mapsto |\cdot|_{v_3} \end{array}$$

that extend $|\cdot|_v$ to a unique absolute value over $\mathbb{C}_v$, $|\cdot|_{v_3}$. This is because

i) The first embedding is a completion, therefore the absolute value over $\mathbb{Q}_v$ can be defined by $|x|_{v_1} = \lim_{n\to\infty} |x_n|_v$ for a succession $\{x_n\} \subset \mathbb{Q}$, $n \in \mathbb{N}$ with $x = \lim_{n\to\infty} x_n$.

ii) By Theorem 3.11, given $\mathbb{Q}_v$ is complete relative to $|\cdot|_{v_1}$.

iii) $\mathbb{C}_v$ is the completion of $\overline{\mathbb{Q}_v}$, so by the same logic as the first embedding, $|\cdot|_{v_3}$ is a unique extension of $|\cdot|_{v_2}$.

(For further detail one can check [9, Chapter 2; §3, §4, §8] and [5, Chapters 1 and 3]). Hence, $|\cdot|_{v_3}$ is a unique extension of $|\cdot|_v$. Furthermore, it is worth noting that we may assert more generally that an absolute value on a complete field admits a unique extension to its algebraic closure, since the latter is a union of finite dimensional extensions.

$\mathbb{C}_v$ is algebraically closed, therefore every finite field extension of $\mathbb{Q}$ can also be embedded into $\mathbb{C}_v$. Let $K$ be a finite field extension of $\mathbb{Q}$, and let the embeddings of $K$ into $\mathbb{C}_v$ be $\sigma_1, ..., \sigma_N$. Each such embedding can be used to define an absolute value on $K$, defined by

$$|x|_{\sigma_i} = |\sigma_i(x)|_v$$

which allows us to map the embeddings of $K$ into $\mathbb{C}_v$ to the absolute values:

$$\psi : \{\text{embeddings } K \hookrightarrow \mathbb{C}_v\} \longrightarrow \{w|v \ : w \in M_K, v \in M_{\mathbb{Q}}\}.$$

Now, in this thesis we will only study finite dimensional field extensions of characteristic 0, so by the primitive element theorem we can suppose $K/\mathbb{Q}$ is a finite dimensional field extension and $K$ is generated by a single element $\xi$. Therefore by proposition 3.9 the map $\psi$ is exhaustive and

$$\#\psi^{-1}(w) = n_w$$

as every absolute value $|\cdot|_w$ extending $|\cdot|_v$ is realized by exactly $n_w$ embeddings
This was the last result we needed in order to prove the product formula, which we will now prove after giving a definition.

**Proposition 3.12** (Product formula)**.** Let $K$ be a number field, and let $x \in K^*$. Then

$$\prod_{v \in M_K} \|x\|_v = 1$$

**Proof** Let $\mathbb{P}$ be the set of all primes, and we begin by showing that the proposition is true for $K = \mathbb{Q}$, as it will be needed afterwards. From the definition of the $p-$adic absolute value, it is easy to see that

$$\prod_{p \in \mathbb{P}} \|x\|_p = \frac{1}{|x|_\infty}$$

Thus,

$$\prod_{v \in M_{\mathbb{Q}}} \|x\|_v = \prod_{v \in M_{\mathbb{Q}}} |x|_v^1 = |x|_\infty \prod_{p \in \mathbb{P}} |x|_p = 1.$$

We will now prove for a number field $K$, let $x \in K^*$, therefore $N_{K/\mathbb{Q}}(x) \in \mathbb{Q}^*$ and we have

$$\begin{aligned}
1 &= \prod_{v \in M_{\mathbb{Q}}} |N_{K/\mathbb{Q}}(x)|_v \\
&= \prod_{v \in M_{\mathbb{Q}}} \prod_{\sigma \in \mathrm{Hom}(K, \mathbb{C}_v)} |\sigma(x)|_v \quad \text{by proposition 1.4} \\
&= \prod_{v \in M_{\mathbb{Q}}} \prod_{w|v,\, w \in M_K} |x|_w^{n_w} \quad \text{from the exhaustive definition for absolute values} \\
&= \prod_{v \in M_{\mathbb{Q}}} \prod_{w|v,\, w \in M_K} \|x\|_w = \prod_{v \in M_K} \|x\|_w
\end{aligned}$$

□

We will take this opportunity to define the following multiset (a set that allows repetition of elements), as we have begun taking products over the absolute values of a number field.

**Definition 3.13.** Let $K$ be a number field, and let every $v \in M_K$ have local degree $n_v$. We define the multiset $\overline{M_K}$ as the smallest multiset with the property that every $v \in M_K$ appears $n_v$ times in $\overline{M_K}$.

This will allow us to write

$$\prod_{v \in \overline{M_K}} |x|_v = \prod_{v \in M_K} |x|_v^{n_v},$$

which will allow us to talk about submultisets of all the places of a number field, as we may want to consider the full multiplicity of some places but not others. In particular, this happens in some uses of Liouville's inequality, which we introduce later. Having proved these properties, particularly the product formula, we now define the height of an algebraic number $\alpha$.

**Definition 3.14.** Given an algebraic number $\alpha \in K$, where $K/\mathbb{Q}$ is a finite dimensional field extension, we define its height, $H(\alpha)$, by

$$H(\alpha)^{[K:\mathbb{Q}]} = \prod_{v \in \overline{M_K}} \max\{1, |\alpha|_v\} = \prod_{v \in M_K} \max\{1, \|\alpha\|_v\}.$$

A key aspect of this definition is that it does not depend on the field $K$, that is, if we consider another field $K'$ such that $K \subset K'$, then

$$H(\alpha) = \prod_{v \in M_K} \max\{1, \|\alpha\|_v\}^{1/[K:\mathbb{Q}]} = \prod_{v \in M_{K'}} \max\{1, \|\alpha\|_v\}^{1/[K':\mathbb{Q}]}.$$

This is because of the degree formula from corollary 3.10. We now prove a couple of useful properties of this height.

**Proposition 3.15.** Let $\alpha \in K^*$ be an algebraic number, and $K/\mathbb{Q}$ a finite dimensional field extension. Then for any integer $n \in \mathbb{Z}$, we have

$$H(\alpha^n) = H(\alpha)^{|n|}.$$

In particular $H(\alpha) = H(\alpha^{-1})$.

**Proof**

Consider an integer $m \in \mathbb{Z}$. We will prove the statement in steps. To begin with, it is clear that for $m = 0$, $H(\alpha^0) = H(1) = 1 = H(\alpha)^0$, so consider now $m$ to be a positive integer. We have

$$H(\alpha^m)^{[K:\mathbb{Q}]} = \prod_{v \in \overline{M_K}} \max\{1, |\alpha^m|_v\} = \prod_{v \in \overline{M_K}} \max\{1, |\alpha|_v\}^m = H(\alpha)^{m[K:\mathbb{Q}]}.$$

Taking $[K : \mathbb{Q}]$th roots gives the result for positive integers. We will now use the product formula for this last step:

$$\begin{aligned} H(\alpha^{-1})^{[K:\mathbb{Q}]} &= \prod_{v \in \overline{M_K}} \max\{1, |\alpha^{-1}|_v\} = \prod_{v \in \overline{M_K}} \max\{1, |\alpha^{-1}|_v\} \prod_{v \in \overline{M_K}} |\alpha|_v = \prod_{v \in \overline{M_K}} \max\{1, |\alpha|_v\} \\ &= H(\alpha)^{[K:\mathbb{Q}]}. \end{aligned}$$

Taking again $[K : \mathbb{Q}]$th roots gives $H(\alpha^{-1}) = H(\alpha)$, which together with the result for positive integers, gives the general result.

□

The last of the properties we will show now goes goes by Liouville's inequality, and we will prove the more general version of it, even though its full scope will not be necessary.

**Proposition 3.16** (Liouville's inequality)**.** Let $K/\mathbb{Q}$ be a finitely dimensional field extension with $[K:\mathbb{Q}] = D$, $\alpha \in K^*$, and let $\mathcal{W} \subset \overline{M_K}$ be any submultiset of all the absolute values on $K$. Then

$$H(\alpha)^D \geq \prod_{v \in \mathcal{W}} \max\left\{\frac{1}{|\alpha|_v}, 1\right\}.$$

**Proof.**
By the product formula, we have

$$\prod_{v \in M_K} \|\alpha\|_v = 1,$$

From this, and the definition of the height of an element over $K$ we get

$$\begin{aligned} H(\alpha)^{[K:\mathbb{Q}]} &= \prod_{v \in M_K} \max\{\|\alpha\|_v, 1\} = \prod_{v \in M_K} \|\alpha\|_v \cdot \prod_{v \in M_K} \max\left\{\frac{1}{\|\alpha\|_v}, 1\right\} \\ &= \prod_{v \in \overline{M_K}} \max\left\{\frac{1}{|\alpha|_v}, 1\right\} \geq \prod_{v \in \mathcal{W}} \max\left\{\frac{1}{|\alpha|_v}, 1\right\}. \end{aligned}$$

□

Having defined the height of an algebraic number and given a few of its properties, we now go one step further and define the height of a linear form, as it will also be necessary moving forward. We will in fact be making immediate use of it, as it is used in the even more general version of Siegel's lemma we will present soon, which unlike lemma 1.9, allows linear forms with coefficients that are not algebraic integers. Let $\Lambda$ be a linear form with coefficients in the number field $K$, which we will write as

$$\Lambda = a_1 X_1 + \cdots a_N X_N \in K[X_1, ..., X_N].$$

We begin by defining

$$|\Lambda|_v = \max\big\{|a_1|_v, \cdots, |a_N|_v\big\},$$

and with this, we define its height to be

$$H(\Lambda)^{[K:\mathbb{Q}]} := \prod_{v \in M_K} \|\Lambda\|_v = \prod_{v \in \overline{M_K}} |\Lambda|_v.$$

It is worth mentioning that this definition can be proven to be independent of the field $K$ using the degree formula, from Corollary 3.10. We also remark that when we write the product over $v \in M_K$, we take the product over the places $v$ without taking into account possible repetitions, unlike in the rightmost product. We now show a good property of this height. Let $0 \neq \alpha \in K$, then by the product formula we get the following equality

$$H(\alpha\Lambda)^{[K:\mathbb{Q}]} = \prod_{v \in M_K} \|\alpha\Lambda\|_v = \left(\prod_{v \in M_K} \|\alpha\|_v\right) H(\Lambda)^{[K:\mathbb{Q}]} = H(\Lambda)^{[K:\mathbb{Q}]}.$$

This is intentional, the equations $\alpha\Lambda = 0$ and $\Lambda = 0$ are equivalent for a non-zero constant $\alpha \in K$, so we want them to have the same height. We also remark that because of the product formula, $H(\Lambda) \geq 1$, as we show here

$$H(\Lambda)^{[K:\mathbb{Q}]} = \prod_{v \in M_K} \|\Lambda\|_v \geq \prod_{v \in M_K} \|a_1\|_v = 1.$$

Where we choose a nonzero coefficient $a_i$, which we can assume is $a_1$ without loss of generality. For the sake of completeness, the absolute value of a polynomial $P(X) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$, is defined as

$$|P(X)|_v = \max\{|a_0|_v, |a_1|_v, ..., |a_n|_v\}.$$

As always, we will only write $|P(X)|$ to denote $|P(X)|_\infty$.

Now, our proof of a more general version of Siegel's lemma which we have been preparing throughout the chapter requires a way to count the elements of a field $K$, just as we did in the proof we give of Siegel's lemma, so we require one last preliminary result:

**Lemma 3.17.** Let $K/\mathbb{Q}$ be a finitely dimensional field extension, with $[K : \mathbb{Q}] = D$ and suppose that for each place $v$ we are given $R_v \in \mathbb{R}_{\geq 0}$, and $\mu_v \in K$, such that $R_\sigma = R_{\overline{\sigma}}$ and $\mu_\sigma = \mu_{\overline{\sigma}}$ for each archimedean $\sigma$, where $\overline{\sigma}$ denotes the complex conjugation of the embedding $\sigma$. Moreover, suppose $R_\mathfrak{p} \neq 1$ for at most finitely many non-archimedean $\mathfrak{p}$. Then there are at most

$$(2R^{1/[K:\mathbb{Q}]} + 1)^{[K:\mathbb{Q}]}$$

elements $\xi$ of $K$ with

$$|\xi - \mu_v|_v \leq R_v$$

for all $v$, where $R = \prod_{v \in \overline{M_K}} R_v < \infty$.

**Proof.**

We begin by defining the set $\Sigma$, the maximal set of embeddings associated to archimedean absolute values such that there are no two elements in $\Sigma$ that are complex conjugates of each other. Then for $\sigma \in \Sigma$, let $K_\sigma = \mathbb{R}$ if $\sigma(K) \subset \mathbb{R}$, and $K_\sigma = \mathbb{C}$ otherwise. With this, we define the cartesian product $\mathcal{F} = \prod_{\sigma \in \Sigma} K_\sigma (\cong \mathbb{R}^r \times \mathbb{C}^s$ for some $s, r \in \mathbb{Z}$, with $r + 2s = D$), and look at its elements $y \in \mathcal{F}$. As we have shown before, there is a correspondence between absolute values and embeddings, so we will write $y_\sigma$ for the corresponding real or complex variables, considered as coordinates of the $y \in \mathcal{F}$ we want to study. Moving on, we want to count the $\xi \in K$ such that $|\xi - \mu_v|_v \leq R_v$ for all $v$, so for those $\xi$ we define, accounting for all the embeddings, a $(r + s)-$dimensional ball, $B(\xi)$, defined coordinate wise as

$$B(\xi) = \left\{ y \in \mathcal{F} : |y_\sigma - \sigma(\xi)| < \frac{1}{2} R^{-1/D} R_\sigma \quad \forall \sigma \in \Sigma \right\}.$$

We will later count the number of these balls in order to count the number of $\xi$ that satisfy the hypotheses, so we must make sure they don't overlap, and that in fact $B(\xi) \cap K = \{\xi\}$. Consider therefore $\xi_1 \neq \xi_2$ both in $K$, and assume there exists some $t \in B(\xi_1) \cap B(\xi_2)$. From the definition of each ball and the triangular inequality, we get

$$|\xi_1 - \xi_2|_\sigma = |\sigma(\xi_1) - t_\sigma + t_\sigma - \sigma(\xi_2)| \leq |\sigma(\xi_1) - t_\sigma| + |\sigma(\xi_2) - t_\sigma| < R^{-1/D} R_\sigma$$

for all $\sigma \in \Sigma$, and since $R_\sigma = R_{\overline{\sigma}}$, $\mu_\sigma = \mu_{\overline{\sigma}}$, we also have

$$|\xi_1 - \xi_2|_{\overline{\sigma}} = |\overline{\sigma(\xi_1)} - \overline{\sigma(\xi_2)}| = |\sigma(\xi_1) - \sigma(\xi_2)| < R^{-1/D} R_{\overline{\sigma}}.$$

Hence, $|\xi_1 - \xi_2|_\sigma < R^{-1/D} R_\sigma$ for all archimedean absolute values.

Now, recall the additional condition on the $\xi \in K$ for which we defined the balls. Considering now non-archimedean absolute values, $\mathfrak{p}$, this time we have

$$|\xi_1 - \xi_2|_\mathfrak{p} = |\xi_1 - \mu_\mathfrak{p} + \mu_\mathfrak{p} - \xi_2|_\mathfrak{p} \leq \max\{|\xi_1 - \mu_\mathfrak{p}|_\mathfrak{p}, |\xi_2 - \mu_\mathfrak{p}|_\mathfrak{p}\} \leq R_\mathfrak{p}.$$

Taking the product over all the absolute values, bearing in mind the archimedean absolute value over $\mathbb{Q}$ is extended to $D$ absolute values on $K$ yields

$$\prod_{v\in\overline{M_K}} |\xi_1-\xi_2|_v < RR^{-1} = 1,$$

a contradiction due to the product formula, so $\xi_1 = \xi_2$, and therefore $B(\xi)\cap K = \{\xi\}$. Consider now $t \in B(\xi)$, then for this $t$ and archimedean $\sigma \in \Sigma$ we also have

$$|t_\sigma - \sigma(\mu_\sigma)| \le |t_\sigma - \sigma(\xi)| + |\sigma(\xi) - \sigma(\mu_\sigma)| < \Big(1+\frac{1}{2}R^{-1/D}\Big)R_\sigma,$$

that is, every $B(\xi)$ is contained in the ball defined by

$$\left\{t\in K : |t_\sigma - \sigma(\mu_\sigma)| \le \Big(1+\frac{1}{2}R^{-1/D}\Big)R_\sigma \quad \forall\sigma\in\Sigma\right\},$$

with $\mu_\sigma \in K$ also clearly satisfying $|\mu_\sigma - \mu_\sigma|_\sigma \le R_\sigma \quad \forall\sigma\in\Sigma$. This has given us the tools to find an upper bound for the $\xi$ with the conditions from the hypothesis. Consider the volume $V$ of the ball defined by $|y_\sigma| \le 1$, by the 2 dimensional recurrence relationship for the volume of an $n-$dimensional ball, the volume of each $B(\xi)$ is

$$V_\xi = V\prod_{\sigma\in\Sigma}\left(\frac{1}{2}R^{-1/D}R_\sigma\right)^{[K_\sigma:\mathbb{R}]} = V\left(\frac{1}{2}R^{-1/D}\right)^D\prod_{\sigma\in\Sigma}R_\sigma^{[K_\sigma:\mathbb{R}]},$$

while the volume of $B(\mu)$ is

$$V_\mu = V\left(1+\frac{1}{2}R^{-1/D}\right)^D\prod_{\sigma\in\Sigma}R_\sigma^{[K_\sigma:\mathbb{R}]}.$$

And finally, the maximum number of $\xi$ from the hypothesis is given by $B(\mu)/B(\xi) = (2R^{1/[K:\mathbb{Q}]}+1)^{[K:\mathbb{Q}]}$, as desired.

□

From now on, to differentiate between the two subsets of embeddings, one of which is in one-to-one correspondence with the archimedean places, the other with non archimedean places, we will do the following. On one hand, write $\mathfrak{p}$ for non-archimedean embeddings, choosing this letter due to the one-to-one correspondence between prime ideals and non-archimedean valuations, of which we have already chosen a representative. On the other hand, we will keep using $\sigma$ to refer to archimedean embeddings. We can now finally give an even more general version of Siegel's lemma, for linear forms with coefficients in a number field $K$.

**Lemma 3.18.** Suppose $[K:\mathbb{Q}] = D$. For positive integers $M, N$ with $DM < N$ and real $H \ge 1$, let $\Lambda_m$ be linear forms with coefficients in $K$ and

$$H(\Lambda_m) \le \mathcal{H} \quad (m = 1,\dots,M).$$

Then there are rational integers $x_1,\dots,x_N$, not all zero, with

$$\Lambda_m(x_1,\dots,x_N) = 0 \quad (m = 1,\dots,M)$$

and

$$|x_n| \leq (N\mathcal{H})^{\frac{DM}{N-DM}} \quad (n = 1, \ldots, N).$$

**Proof.**
The method of proof resembles the one used for this same lemma over $\mathbb{Q}$, so we adequately reformulate the problem by considering the linear map $u : \mathbb{Q}^N \longrightarrow K^M$ defined by

$$u(\mathbf{x}) = (\Lambda_1(\mathbf{x}), ..., \Lambda_M(\mathbf{x}))$$

and solving for $u(\mathbf{x}) = 0$, with $\mathbf{x} = (x_1, ..., x_N)$.
We now fix $W \in \mathbb{R}_{>0}$, which we will soon choose adequately, and let $\mathcal{S}$ be the set of $\mathbf{x} \in \mathbb{Z}^N$ with $0 \leq x_1, ..., x_N \leq \lfloor W \rfloor$. Clearly $\#\mathcal{S} = (\lfloor W \rfloor + 1)^N$, and now we will use the previous lemma to bound the cardinality of the set $u(\mathcal{S})$. We fix $m$, and for every non archimedean embedding $\mathfrak{p}$, define $\mu_{\mathfrak{p}} = 0$, somewhat borrowing the notation from the previous lemma. For archimedean embeddings $\sigma$, we let $\mathbf{x_0} = (\lfloor W \rfloor/2, ..., \lfloor W \rfloor/2) \in \mathbb{Q}^N$, and define $\mu_\sigma = \Lambda_m(\mathbf{x_0})$ for every archimedean embedding of $K$. With this, and letting $\xi = \Lambda_m(\mathbf{x})$ we have, for non-archimedean absolute values first:

$$|\xi - \mu_{\mathfrak{p}}|_{\mathfrak{p}} = |\xi|_{\mathfrak{p}} \leq |\Lambda_m|_{\mathfrak{p}} =: R_{\mathfrak{p}}.$$

This inequality is due to the ultrametric inequality, and the fact that we are considering $\mathbf{x} \in \mathbb{Z}^N$, and the non-archimedean absolute value of an integer is bounded above by 1. Notice also that for only finitely many $\mathfrak{p}$, $R_{\mathfrak{p}} \neq 1$. For the archimedean absolute values we have

$$|\xi - \mu_\sigma|_\sigma = |\Lambda_m(\mathbf{x}) - \Lambda_m(\mathbf{x_0})|_\sigma = |\Lambda_m(\mathbf{x} - \mathbf{x_0})|_\sigma \leq N(\lfloor W \rfloor/2)|\Lambda_m|_\sigma =: R_\sigma.$$

We can now apply lemma 3.17 to count these $\xi$, with $R = \prod_{v \in \overline{M_K}} R_v = (N\lfloor W \rfloor/2)^D H(\Lambda)^D$ per the definition of $R$ from the lemma. By the hypothesis bound $H(\Lambda_m) \leq \mathcal{H}$, we have $R \leq (N\lfloor W \rfloor \mathcal{H}/2)^D$, hence the number of $\xi$ is bounded by

$$(2R^{1/D} + 1)^D = (N\mathcal{H}\lfloor W \rfloor + 1)^D \leq (N\mathcal{H}(\lfloor W \rfloor + 1))^D.$$

And since we obtained this considering only a fixed $m$, $\#u(\mathcal{S}) \leq (N\mathcal{H}(\lfloor W \rfloor + 1))^{DM}$. To ensure we lose injectivity, we want $\#u(\mathcal{S}) < \#S$, so we need

$$(N\mathcal{H}(\lfloor W \rfloor + 1))^{DM} < (\lfloor W \rfloor + 1)^N \iff \lfloor W \rfloor + 1 > (N\mathcal{H})^{\frac{DM}{N-DM}}.$$

Taking $W = (N\mathcal{H})^{\frac{DM}{N-DM}}$ suffices, and due to $u$ not being injective on $\mathcal{S}$, there exist at least two distinct $\mathbf{x}', \mathbf{x}'' \in \mathcal{S}$ such that $u(\mathbf{x}' = u(\mathbf{x}'')$, and therefore $u(\mathbf{x}) = 0$, with $\mathbf{x} = \mathbf{x}' - \mathbf{x}'' \neq 0$, and all components of $\mathbf{x}$ having absolute value smaller or equal to $W$.
□

We now define the concept of a derivation on a polynomial ring, and then look at a result we will make use of in the proof of the Gelfond - Schneider theorem. Derivations can be defined more in general over an algebra, but this suits out purposes perfectly well.

**Definition 3.19.** Given a field $\mathcal{K}$, a derivation on the polynomial ring $\mathcal{K}[X_1, ..., X_m]$ is a $\mathcal{K}-$linear map $\delta$ from $\mathcal{K}[X_1, ..., X_m]$ to itself that satisfies, for $P, Q \in \mathcal{K}[X_1, ..., X_m]$;

$$\delta(P + Q) = \delta(P) + \delta(Q) \qquad \delta(PQ) = P\delta(Q) + \delta(P)Q.$$

The rightmost condition is called Lebnitz's law, and furthermore, if $\mathcal{K}$ has a unit, 1, then since $\delta(1^2) = 2\delta(1)$, we have $\delta(1) = 0 \implies \delta(K) = 0$ for every $K \in \mathcal{K}$ by $\mathcal{K}-$linearity.

**Lemma 3.20.** Let $\delta$ be a derivation on $\mathcal{K}[X_1, \ldots, X_m]$. Then there are integers $a$, $b$ and polynomials $P_{l,k}$ in $\mathcal{K}[X_1, \ldots, X_m]$ with total degrees at most $a + bk$ such that for any $k \geq 0$, we have

$$\delta^k(X_l) = P_{l,k}(X_1, \ldots, X_m) \quad (l = 1, \ldots, m).$$

Further, if $\mathcal{K} = \mathbb{C}$ there is an integer $c$ such that the coefficients of $P_{l,k}$ have absolute values at most $k!c^k$.

**Proof.**

For simplicity, we will write $P_{l,0}(X_1, ..., X_m) = X_l$ and $P_{l,1} = P_l = \delta(X_l)$. We begin by noting the following recursive formula. We will write

$$P_{l,k}(X_1, ..., X_m) = P_{l,k} = \sum_{i_1=0}^{n_1} \cdots \sum_{i_m=0}^{n_m} p^{l,k}_{i_1 \cdots i_m} X^{i_1} \cdots X^{i_m}$$

And consider

$$\delta^k(X_l) = \delta(\delta^{k-1}(X_l)) = \delta(P_{l,k-1}) = \sum_{i_1=0}^{n_1} \cdots \sum_{i_m=0}^{n_m} p^{l,k-1}_{i_1 \cdots i_m} \delta(X_1^{i_1} \cdots X_m^{i_m}).$$

Now, using Leibnitz's law twice, and writing $X^{i_1 \cdots i_m} = X_1^{i_1} \cdots X_m^{i_m}$ we have

$$\delta(X_1^{i_1} \cdots X_m^{i_m}) = \sum_{r=1}^{m} \frac{X^{i_1 \cdots i_m} \delta(X_r^{i_r})}{X_r^{i_r}} = \sum_{r=1}^{m} \frac{X^{i_1 \cdots i_m} i_r X_r^{i_r - 1} \delta(X_r)}{X_r^{i_r}} = \sum_{r=1}^{m} \frac{\partial X^{i_1 \cdots i_m}}{\partial X_r} \delta(X_r),$$

which by $\mathcal{K}-$linearity implies

$$\delta^k(X_l) = \delta(P_{l,k-1}) = \sum_{r=1}^{m} \left[ \frac{\partial}{\partial X_r} \left( \sum_{i_1=0}^{n_1} \cdots \sum_{i_m=0}^{n_m} p^{l,k-1}_{i_1 \cdots i_m} X^{i_1, \ldots, i_m} \right) \delta(X_r) \right] = \sum_{r=1}^{m} \frac{\partial P_{l,k-1}}{\partial X_r} P_r,$$

that is,

$$\delta^k(X_l) = P_{l,k} = \sum_{r=1}^{m} \frac{\partial P_{l,k-1}}{\partial X_r} P_r. \tag{3.1}$$

We will begin by proving the bound on the degrees of the $P_{l,k}$. From equation (3.1), we get

$$\text{degree}(P_{l,k}) =: \mathcal{D}(P_{l,k}) = \max_r \left\{ \mathcal{D}\left( \frac{\partial P_{l,k-1}}{\partial X_r} P_r \right) \right\} = \max_r \left\{ \mathcal{D}\left( \frac{\partial P_{l,k-1}}{\partial X_r} \right) + \mathcal{D}(P_r) \right\}.$$

And since for polynomials in general we have that

$$\mathcal{D}\left( \frac{\partial P_{l,k-1}}{\partial X_r} \right) \leq \mathcal{D}(P_{l,k-1}),$$

then

$$\mathcal{D}(P_{l,k}) \leq \max_r \Big\{ \mathcal{D}(P_{l,k-1}) + \mathcal{D}(P_r) \Big\} = \mathcal{D}(P_{l,k-1}) + \max_r \Big\{ \mathcal{D}(P_r) \Big\}.$$

By induction then, we get, with $b := \max_r \big\{1, \mathcal{D}(P_r)\big\}$ :

$$\mathcal{D}(P_{l,k}) \leq 1 + bk.$$

Where the 1 comes from $\mathcal{D}(X_l)$. We shall hereafter write $d_k = \mathcal{D}(P_{l,k})$, and we now prove the bound on the coefficients of $P_{l,k}$, now with $\mathcal{K} = \mathbb{C}$.
Let $P, Q \in \mathbb{C}[X_1, ..., X_m]$, we define the length of $P$, $\mathcal{L}(P)$, to be the sum of the absolute values of all the coefficients of $P$. We then have in general

$$\mathcal{L}\Big(\frac{\partial P}{\partial X_i}\Big) \leq \mathcal{D}(P)\mathcal{L}(P) \quad \text{and} \quad \mathcal{L}(PQ) \leq \mathcal{L}(P)\mathcal{L}(Q).$$

So using equation (3.1) again we get

$$\mathcal{L}(P_{l,k}) \leq m \cdot \max_r \mathcal{L}\Big(\frac{\partial P_{l,k-1}}{\partial X_r}\Big) \max_r \mathcal{L}(P_r) \leq m\mathcal{D}(P_{l,k-1})\mathcal{L}(P_{l,k-1}) \max_r \mathcal{L}(P_r).$$

We define $l_1 := \max_r \mathcal{L}(P_r), d_k = \mathcal{D}(P_{l,k})$, and employ the above inequality inductively:

$$\mathcal{L}(P_{l,k}) \leq (ml_1)d_{k-1}(ml_1)d_{k-2}\mathcal{L}(P_{l,k-2}) \leq \cdots \leq (ml_1)^k \mathcal{L}(X_l) \prod_{r=0}^{k} d_r \leq (ml_1)^k \prod_{r=0}^{k} (2br)$$

with the rightmost term evaluating to $(2bml_1)^k k! = c^k k!$, and since $\mathcal{L}(P_{l,k})$ is an upper bound for the largest coefficient of $P_{l,k}$, we have our desired bound.

□

## 3.2 A familiar proof of Gelfond - Schneider

Now, with more advanced tools, we are ready to tackle this theorem employing a method very reminiscent of the one used to prove the Hermite - Lindemann theorem. We begin by giving the statement of the theorem despite not being able to give the proof yet, because we will be using its hypotheses for the following lemmas that are used in its proof.

**Definition 3.21.** Let $\rho > 0$, and consider only functions defined over $\mathbb{C}$. We say that an entire function, $f$, has strict growth order at most $\rho$ if there exist positive constants $c, C$ such that $|f|_R \leq cC^{R^\rho}$ for all $R \geq 0$. The growth order of $f$ is the infimum of all possible values of $\rho$. Furthermore, for meromorphic functions $F = g/h$, with $g, h$ entire functions, we will say that $F$ has growth order at most $\rho$ if both $g$ and $h$ have growth order at most $\rho$.

In the theorems of Hermite - Lindemann and Gelfond - Schneider, a $\rho = 1$ is sufficient, but for the rest of known cases we have $\rho = 2$. An example of this is the Weirstrass elliptic function in the following chapter. Now, the following theorem is noticeably more general than the Gelfond - Schneider theorem, and even has weaker hypotheses. It is a display of how powerful this proof strategy can be, and the diverse results it can unveil, but as we will show soon, it does imply Gelfond - Schneider.

**Theorem 3.22** (Schneider - Lang)**.** Let $f_1, \ldots, f_m$ be meromorphic functions of growth order at most $\rho > 0$, with at least two among them algebraically independent over $\mathbb{C}$. Suppose that for some number field $K$ with $[K : \mathbb{Q}] = D$ the derivatives $f_1', \ldots, f_m'$ lie in the ring $K[f_1, \ldots, f_m]$. Then there are at most $16\rho D$ complex numbers $w$ such that $f_1, \ldots, f_m$ are analytic at $w$ with values in $K$.

As mentioned, we will be following a similar proof strategy as in Hermite - Lindemann. We begin by considering a finite subset, $\mathcal{S}$, of the following set

$$\{w \in \mathbb{C} \,|\, f_1, ..., f_m \text{ are analytic at } w \text{ with values in } K\}.$$

We will write $S = \#\mathcal{S}$ for the cardinality of the subset, and we proceed by picking any $\mathbb{C}-$algebraically independent $f, g$ from the $f_1, ..., f_m$, and proving the existence of an auxiliary polynomial in $f, g$ that is not identically 0, and has zeroes in $\mathcal{S}$. We will then prove we can arbitrarily increase the multiplicities of the zeroes of the polynomial in $f, g$ provided $S$ is larger than a certain constant we will determine later. Due to this, assuming $S$ is larger than this constant will imply our auxiliary polynomial is identically 0, despite proving it is not identically 0, so $S$ must be smaller than this constant, which is in fact the result we want to prove.

Let us consider this subset $\mathcal{S}$ with cardinality $S$ for the following lemmas then. Again, as in the proof for the Hermite - Lindemann theorem, in the following lemmas we will be considering the variables $L, S, T, T_1, S_1$. These will only be variables in the sense that they can vary to satisfy the various restrictions we will later impose upon them in order to prove the Schneider - Lang theorem, or they may obey certain inequalities as is the case for $S$, but otherwise within the context of the lemmas, they are all fixed quantities. Even when we have to employ the lemmas repeatedly, we will consider the lemmas after fixing these values in order to get the desired results. On the other hand, the constants $C, C_i, c_i\ i \in \mathbb{Z}$, are just constants that satisfy the inequalities, and depend only on $f_1, ..., f_m$ and $\mathcal{S}$.

**Lemma 3.23.** Given the finite set $\mathcal{S} \subset \mathbb{C}$ we are considering, and two $\mathbb{C}-$algebraically independent meromorphic functions $f, g \in \{f_1, ..., f_m\}$ from theorem 3.22. For any $L \geq 2$, $T \geq 1$ in $\mathbb{Z}$ with

$$(L+1)^2 \geq 2DST \tag{3.2}$$

there is $P \in \mathbb{Z}[X, Y]$, of degree at most $L$ in each variable and with coefficients of absolute values at most $c_1^{L+T} L^T T^T$, such that

$$\phi(z) = P(f(z), g(z)) \tag{3.3}$$

is not identically zero and

$$\operatorname{ord}_{z=w} \phi(z) \geq T \quad \forall w \in \mathcal{S}. \tag{3.4}$$

**Proof.**
We consider again a polynomial

$$P(X, Y) = \sum_{i=0}^{L} \sum_{j=0}^{L} p_{ij} X^i Y^j$$

and we will show we can find $p_{ij} \in \mathbb{Z}$ that fit our criteria. To this end, we note that just as in (2.7) , equation (3.4) is also an equation of the type $\Lambda_{tw} = 0$, with coefficients $\alpha_{ijtw} = (d/dz)^t (f(z)^i g(z)^j)$ being evaluated at $z = w$. To be explicit we have

$$\Lambda_{tw} = \phi^{(t)}(w) = \sum_{i=0}^{L} \sum_{j=0}^{L} p_{ij} \alpha_{twij} = 0.$$

It is the heights of $\alpha_{twij}$ we will find a bound for in order to bound the heights of the $\Lambda_{tw}$ to use our lastest version of Siegel's lemma. To do this we will again use Lebniz's rule to differentiate, and we will write $f^{(n)} = \frac{d^n}{dz^n} f(z)$ for the $n-$th derivative of $f$, $\mathbf{t} = (t_1, ..., t_i)$, and $\mathbf{s} = (s_1, ..., s_j)$. We thus have

$$(d/dz)^t (f^i g^j) = \sum_{t_1 + \cdots + t_i + s_1 + \cdots + s_j = t} C(\mathbf{t}, \mathbf{s}) f^{(t_1)} \cdots f^{(t_i)} g^{(s_1)} \cdots g^{(s_j)}$$

Where each $t_\eta, s_\iota$ takes every integer value from 0 to $t$, and we write $f^i$ as the product of $f$ $i$ times for ease later on when substituting the derivations. Despite its appearance, this sum is the same as the usual derivative. Moreover, the constant is the multinomial coefficient

$$C(\mathbf{t},\mathbf{s}) = \frac{t!}{t_1!\cdots t_i! s_1!\cdots s_j!}.$$

Now, by the hypotheses of theorem 3.22, we have differential equations for all our functions

$$\frac{d}{dz}(f_l) = f_l' = P_l(f_1,...,f_m) \quad (l = 1,...,m),$$

which gives us a derivation on $K[f_1,...,f_m]$ given by $\frac{d}{dz}$, which by lemma 3.20 means all the $f^{(k)}, g^{(h)}$ are in $K[f_1,...,f_m]$, so the $\Lambda_{tw}$ do indeed have all coefficients in $K$, and we can find a solution to the system by the other hypothesis $(L+1)^2 \geq 2DST$, which is more than strictly necessary, but allows the bound for the $p_{ij}$ to have a Siegel exponent $\frac{DST}{(L+1)^2-DST} \leq 1$. Moving on, we write

$$f^{(k)} = F_k(f_1,...,f_m) \quad g^{(h)} = G_h(f_1,...,f_m),$$

both in $K[f_1,...,f_m]$ by lemma 3.20, and therefore the coefficients of the $\Lambda_{tw}$ are

$$\alpha_{twij} = H_{tij}(f_1(w),...,f_m(w))$$

for the polynomials

$$H_{tij}(f_1,...,f_m) = \sum_{t_1+\cdots+t_i+s_1+\cdots+s_j=t} C(\mathbf{t},\mathbf{s}) F_{t_1}\cdots F_{t_i} G_{s_1}\cdots G_{s_j}$$

as seen before. By lemma 3.20 again, we have a bound on their degrees, and therefore on the degree of the $H_{tij}$:

$$\mathcal{D}(H_{tij}) \leq (1+bt_1)+\cdots+(1+bt_i)+(1+bs_1)+\cdots+(1+bs_j) = (i+j)+bt \leq 2L+bT.$$

We will use this to find a bound for the $\alpha_{twij}$. We begin by considering the archimedian places:

$$|\alpha_{twij}|_\sigma = |\sigma(\alpha_{twij})|_\sigma \leq \mathcal{L}(\sigma(H_{tij}))c_2^{L+T}$$

Where the length arises naturally by considering the absolute values of all the terms that make up $\alpha_{twij}$, and $c_2^{L+T}$ comes from considering a bound for the largest term to the largest possible degree. Continuing, we have (omitting the indices of the sum for ease)

$$\mathcal{L}(\sigma(H_{tij})) \leq \sum C(\mathbf{t},\mathbf{s})\mathcal{L}(\sigma(F_{t_1}))\cdots\mathcal{L}(\sigma(F_{t_i}))\mathcal{L}(\sigma(G_{s_1}))\cdots\mathcal{L}(\sigma(G_{s_j})).$$

To bound this sum, recall that $\sum C(\mathbf{t},\mathbf{s}) = (i+j)^t \leq (2L)^T$. Moreover, consider the derivation obtained from $\sigma \circ (d/dz) = \sigma(d/dz)$, which is a derivation as it is simply applying $\sigma$ onto the resultant polynomial, so all the defining traits of a derivation are maintained. From lemma 3.20 again, from the bound on the degree and the absolute value of the coefficients we also get a bound on maximum number of terms the polynomials may have, thus

$$\mathcal{L}(\sigma(F_{t_i})) \leq (1+bt_1+1)^m t_1! c^{t_1} \leq t_1! c_3^{t_1}.$$

Joining these inequalities together yields

$$|\alpha_{twij}|_\sigma \leq (2L)^T \max_{t_1,...,t_i,s_1,...,s_j}\{t_1!,...,t_i!,s_1!,...,s_j!\} c_3^t c_2^{L+T} \leq (2L)^T t! c_4^{L+T} \leq c_5^{L+T} L^T T^T.$$

We now consider the non-archimedean places. In this case, even though we have developed the machinery of heights, it is useful to go back into the algebraic integers, as doing that will allow us to use the ultrametric inequality to its fullest potential. Consider a non-zero $\beta \in \mathbb{Z}$ such that $\beta f_l' = \beta P_l \in \mathbb{Z}_K[f_1, ..., f_m]$. From equation (3.1), it follows inductively that $\beta^k f_l^{(k)} = \beta^k P_{l,k} \in \mathbb{Z}_K[f_1, ..., f_m]$ too, hence $\beta^t \alpha_{twij} = \beta^t H_{tij}(f_1(w), ..., f_m(w)) \in \mathbb{Z}_K$ and by the ultrametric inequality, it suffices to bound the powers of the $f_l(w)$, if they're larger than 1, with their corresponding degrees. This yields

$$|\beta^t \alpha_{twij}|_{\mathfrak{p}} \leq \max\{1, |f_1(w)|_{\mathfrak{p}}\}^{2L+bT} \cdots \max\{1, |f_m(w)|_{\mathfrak{p}}\}^{2L+bT},$$

and since we are dealing with the non-archimedean absolute values, $|\beta|_{\mathfrak{p}} \leq 1$, hence

$$|\alpha_{twij}|_{\mathfrak{p}} \leq |\beta^{-T}|_{\mathfrak{p}} \max\{1, |f_1(w)|_{\mathfrak{p}}\}^{2L+bT} \cdots \max\{1, |f_m(w)|_{\mathfrak{p}}\}^{2L+bT} \leq |\beta^{-T}|_{\mathfrak{p}} M_{\mathfrak{p}}^{L+T}.$$

For a constant $M_{\mathfrak{p}}$ that depends on the place $\mathfrak{p}$, with $M_{\mathfrak{p}} \neq 1$ for finitely many $\mathfrak{p}$. Now, recall the formulas $H(\Lambda_{tw})^{[K:\mathbb{Q}]} = \prod_v |\Lambda_{tw}|_v$, and $|\Lambda_{tw}|_v = \max_{i,j} \{|\alpha_{twij}|_v\}$, the bounds for the $|\alpha_{twij}|_v$ give us a bound for $H(\Lambda_{tw})$:

$$H(\Lambda_{tw})^{[K:\mathbb{Q}]} \leq (c_5^{L+T} L^T T^T)^{[K:\mathbb{Q}]} |\beta|^T (\prod_{\mathfrak{p}} M_{\mathfrak{p}})^{L+T}.$$

Hence

$$H(\Lambda_{tw}) \leq (c_6^{L+T} L^T T^T).$$

We now apply Siegel's lemma 3.18, having $ST$ equations, and $(L+1)^2$ unknowns which implies the existence of the polynomial $P$ with the properties we want, and coefficients of absolute value at most

$$\left((L+1)^2 c_6^{L+T} L^T T^T\right)^{\frac{DST}{(L+1)^2 - DST}} \leq (L+1)^2 c_6^{L+T} L^T T^T \leq c_1^{L+T} L^T T^T.$$

□

The next lemma is intimately connected with the previous one, as they will both be part of the proof of the Schneider - Lang theorem, so all the notation carries forward.

**Lemma 3.24.** Suppose for some integer $T_1 \geq T$ that

$$\operatorname{ord}_{z=w} P(f(z), g(z)) = \operatorname{ord}_{z=w} \phi(z) \geq T_1 \quad \forall w \in \mathcal{S}.$$

Suppose further that for some $w_1 \in \mathcal{S}$ and some $t = 0, 1, \ldots, 2T_1 - 1$ we have

$$\operatorname{ord}_{z=w_1} \phi(z) \geq t.$$

Then for any $\mathcal{Y} \geq 1$ we have

$$|\phi^{(t)}(w_1)| \leq \mathcal{Y}^{-ST_1} c_6^{L+T_1} L^{T_1} T_1^{3T_1} C_1^{L\mathcal{Y}^\rho}.$$

**Proof.**
We will follow the same template as before, and divide $\phi$ by its zeroes, but this time also by its poles, to make the "new" $\phi$ entire. To be precise, there exists a non-zero complex function $h$ such that $h$, $hf$, and $hg$ are entire of growth order at most $\rho$. The existence of

this $h$ hinges upon a couple observations. First, the product of two entire functions $X, Y$ with growth orders at most $\rho_1$ and $\rho_2$ respectively is at most $\max\{\rho_1, \rho_2\}$, this is easily seen with the following inequalities:

$$|X(z)||Y(z)| \leq a_1 A_1^{R^{\rho_1}} a_2 A_2^{R^{\rho_1}} \leq b \max\{A_1, A_2\}^{R^{\rho_1}+R^{\rho_2}} \leq b \max\{A_1^2, A_2^2\}^{\max\{R^{\rho_1}, R^{\rho_2}\}}.$$

Secondly, since $f, g$ are meromorphic with growth order at most $\rho$, they are the quotient of two entire functions with growth order at most $\rho$, so if we write $f = r_1/s_1$ and $g = r_2/s_2$, we can take $h = s_1 s_2$, which will be entire of growth order at most $\rho$, and will also make $hf$ and $hg$ entire with growth order at most $\rho$. We note that this choice of $h$ also fulfills the property $h(w_1) \neq 0$, which will be necessary later. This is because by definition, $w_1 \in \mathcal{S}$, so $f(w_1), g(w_1) \in K$ and therefore $w_1$ cannot be a pole of $f$ or $g$, so it is not a zero of either $s_1$ or $s_2$. Now, with this $h$ we define $\psi(z) = h^{2L}(z)\phi(z)$ and

$$\Phi(z) = \frac{\psi(z)}{\prod_{w\in\mathcal{S}}(z-w)^{T_1}},$$

both entire. This then implies that for every positive real $R, R_1$, with $R \leq R_1$ we have

$$|\Phi|_R \leq |\Phi|_{R_1}.$$

Now, for our purposes, we will take $R = 1 + \sup_{w\in\mathcal{S}} |w|$ and $R_1 = (2\mathcal{Y}+1)R$, with $\mathcal{Y} \geq 1$ as in the hypothesis. We can further bound the previous inequality to obtain

$$\frac{|\psi(z)|_R}{(2R)^{ST_1}} \leq |\Phi|_R \leq |\Phi|_{R_1} \leq \frac{|\psi(z)|_{R_1}}{(2\mathcal{Y}R)^{ST_1}},$$

which implies

$$|\psi|_R \leq \mathcal{Y}^{-ST_1}|\psi|_{R_1}. \tag{3.5}$$

Which just as before will give us a way to bound the $|\psi|_R$ term, by cleverly choosing $\mathcal{Y}$. We now use lemma 2.4 again with $\phi = \psi$, $m = t$, and $r = R - 1 = \sup_{w\in\mathcal{S}} |w|$, to get the following inequality

$$|\psi^{(t)}(w_1)| \leq |\psi^{(t)}|_r \leq t!|\psi|_R$$

and then bound $|\psi|_R$ using lemma 3.23:

$$|\psi|_R \leq (L+1)^2 c_1^{L+T} L^T T^T \cdot \left(cC^{R_1^\rho}\right)^L$$

which due to $R$ being fixed yields

$$\begin{aligned}|\psi^{(t)}(w_1)| &\leq (2T_1)^{(2T_1)} \cdot \mathcal{Y}^{-ST_1} \cdot (L+1)^2 c_1^{L+T} L^T T^T \cdot cC^{((2\mathcal{Y}+1)R)^\rho L} \\ &\leq \mathcal{Y}^{-ST_1} c_7^{L+T_1} L^{T_1} T_1^{3T_1} C_1^{\mathcal{Y}^\rho L}.\end{aligned}$$

To finish the proof, recall that by hypothesis $\phi(w_1)^{(i)} = 0$ for all $i < t$, so $\psi^{(t)}(w_1) = \phi^{(t)}(w_1)h^{2L}(w_1)$, and since $h(w_1) \neq 0$,

$$|\phi^{(t)}(w_1)| \leq \mathcal{Y}^{-ST_1} c_6^{L+T_1} L^{T_1} T_1^{3T_1} C_1^{\mathcal{Y}^\rho L}.$$

□

The next and final step will be to find a lower bound for these same terms, that will force them to be 0 for our eventual choices of $T, S, L, \mathcal{Y}$. Again, the notation will be carried forward.

**Lemma 3.25.** For any integer $T_1 \geq T$ and each of the

$$\xi_{tw} = \phi^{(t)}(w) \quad (w \in \mathcal{S};\, t = 0, 1, \ldots, 2T_1 - 1)$$

we have either $\xi_{tw} = 0$ or

$$|\xi_{tw}| \geq c_8^{-(L+T_1)} L^{-3DT_1} T_1^{-3DT_1}.$$

**Proof.**
$\xi_{tw}$ is again a linear form $\Lambda_{tw}$ which we must bound. Now, from lemma 3.23 we are familiar with $\xi_{tw}$, so let $\mathfrak{p}$ be a non-archimedian absolute value, we have:

$$|\xi_{tw}|_{\mathfrak{p}} \leq \max_{i,j} |p_{ij}|_{\mathfrak{p}} \max_{i,j} |\alpha_{twij}|_{\mathfrak{p}} \leq \max_{i,j} |\alpha_{twij}|_{\mathfrak{p}}.$$

For archimedean absolute values $\sigma$, on the other hand:

$$|\xi_{tw}|_{\sigma} \leq (L+1)^2 \max_{i,j} |p_{ij}|_{\sigma} \max_{i,j} |\alpha_{twij}|_{\sigma} \leq |C|_{\sigma} \max_{i,j} |\alpha_{twij}|_{\sigma}$$

Where $|C|_{\sigma} = |(L+1)^2 c_7^{L+2T_1} L^{2T_1} (2T_1)^{2T_1}|_{\sigma}$ is a constant larger than 1. We note that since we are now differentiating $2T_1$ times, we have adjusted the bounds accordingly for $|C|_{\sigma}$. Now, we can use the same bounds for $|\alpha_{twij}|_v$ that we used in lemma 3.23 to obtain bounds for $\max\{1, |\xi_{tw}|_v\}$, and taking the product over all the places, we thus obtain a bound for $H(\xi_{tw})$:

$$H(\xi_{tw}) \leq C c_1^{L+T} L^T T^T \leq c_9^{L+T_1} L^{2T_1} T_1^{2T_1} L^{T_1} T_1^{T_1} = c_9'^{L+T_1} L^{3T_1} T_1^{3T_1}.$$

Using now proposition 3.16, Liouville's inequality, with $\mathcal{W} = \{\infty\}$ we obtain either $\xi_{tw} = 0$ (we cannot use Liouville's inequality) or

$$\frac{1}{|\xi_{tw}|} \leq \max\left\{ \frac{1}{|\xi_{tw}|_{\infty}}, 1 \right\} \leq H(\xi_{tw})^D \leq c_8^{L+T_1} L^{3DT_1} T_1^{3DT_1},$$

that is $|\xi_{tw}| \geq c_8^{-(L+T_1)} L^{-3DT_1} T_1^{-3DT_1}$.

□

We now go back to the thorem, with the groundwork done to prove it.

**Proof of the Schneider - Lang theorem**
Recall the statement of the theorem, already given as theorem 3.22. We begin by considering the finite set $\mathcal{S}$ we described at the beginning of this section with cardinality $S$, and taking all the hypotheses from the Schneider - Lang theorem 3.22. As in Hermite - Lindemann, the term $\mathcal{Y}^{-ST_1}$ will have to beat every other term, with this time only the restrictions

- $T_1 \geq T$
- $(L+1)^2 \geq 2DST$.

With our setup established, we now assume $16\rho D < S < \infty$, and choose any $\mathbb{C}-$algebraically independent $f, g \in \{f_1, ..., f_m\}$. We then fix a sufficiently large $T$, and define $L := \lfloor\sqrt{2DST}\rfloor$, which satisfies $(L+1)^2 \geq 2DST$ and implies $L \leq c_{10}T \leq c_{10}T_1$. Employing now lemma 3.23 provides the existence of an auxiliary polynomial in $f, g$: $\phi(z) = P(f(z), g(z)) \in \mathbb{Z}[f, g]$ with the critical properties

- $\phi(z)$ is not identically 0
- $\operatorname{ord}_{z=w} \phi(z) \geq T \quad \forall w \in \mathcal{S}$.

We now fix $T_1 = T$, which clearly satisfies $T_1 \geq T$, and define $\mathcal{Y} := \left\lceil T_1^{1/2\rho} \right\rceil$. We now use lemma 3.24, with $w_1$ now being an element of $\mathcal{S}$ which we will call $w^*$, and $t = T$. This yields

$$|\phi^{(t)}(w^*)| \leq \mathcal{Y}^{-ST_1} c_6^{L+T} L^{T_1} T_1^{3T_1} C_1^{L\mathcal{Y}^\rho}.$$

On the other hand however, lemma 3.25 tells us that either $\phi^{(t)}(w^*) = 0$ or

$$|\phi^{(t)}(w^*)| \geq c_8^{-(L+T_1)} L^{-3DT_1} T_1^{-3DT_1}.$$

Our aim is to eventually force $\phi(z)$ to be a polynomial with zeroes of infinite order, so we want to arbitrarily increase the lower bound of $\operatorname{ord}_{z=w} \phi(z)$ for at least one $w \in \mathcal{S}$. This is equivalent to proving that for at least one $w \in \mathcal{S}$, $\phi^{(n)}(w) = 0$ for arbitrarily large $n \in \mathbb{N}$. Hence, we must make the two inequalities we have obtained for $\phi^{(t)}(w^*)$ contradict each other. Specifically, we want

$$\mathcal{Y}^{-ST_1} c_6^{L+T} L^{T_1} T_1^{3T_1} C_1^{L\mathcal{Y}^\rho} < c_8^{-(L+T_1)} L^{-3DT_1} T_1^{-3DT_1},$$

with the strict inequality as we must rule out an equality of all the terms. Now, from our definitions of $\mathcal{Y}$ and $L$, we have the following inequalities:

$$\mathcal{Y}^{-ST_1} c_7^{L+T_1} L^{T_1} T_1^{3T_1} C_1^{\mathcal{Y}^\rho L} \leq T_1^{-ST_1/2\rho} T_1^{T_1/2} T_1^{3T_1} C_2^{T_1} = \left(T_1^{-S/2\rho} T_1^{7/2} C_2\right)^{T_1},$$

$$c_8^{-(L+T_1)} L^{-3DT_1} T_1^{-3DT_1} \geq c_{11}^{T_1} T_1^{-3DT_1/2} T_1^{-3DT_1} = \left(T_1^{-9D/2} c_{11}\right)^{T_1}.$$

So it suffices to have

$$T_1^{-S/2\rho} T_1^{7/2} C_2 < T_1^{-9D/2} c_{11} \iff T_1^{-S/2\rho+7/2+9D/2} < \frac{c_{11}}{C_2} = C_3$$

Now, as we will see shortly, it is not enough that this is true for certain values of $T_1$, as we will want to increase the value of $T_1$ arbitrarily when we iteratively employ lemmas 3.24 and 3.25, therefore we must have

$$\frac{-S}{2\rho} + 7/2 + 9D/2 < 0,$$

which is equivalent to

$$S > (7 + 9D)\rho.$$

Since we have assumed $S > 16\rho D \geq (7 + 9D)\rho$, this condition is met. Furthermore, recall that when we began the proof we fixed a sufficiently large $T$. As we can see now, sufficiently large here means large enough such that $T^{-S/2\rho+7/2+9D/2} < C_3$, and therefore $T_1^{-S/2\rho+7/2+9D/2} < C_3$ for every value $T_1$ we choose, as they are all bounded below by $T$.

Now, because of our assumed bound on $S$, we have $|\phi^{(t)}(w^*)| \leq \mathcal{Y}^{-ST_1} c_6^{L+T} L^{T_1} T_1^{3T_1} C_1^{L\mathcal{Y}^\rho} < c_8^{-(L+T_1)} L^{-3DT_1} T_1^{-3DT_1} \leq |\phi^{(t)}(w^*)|$, a contradiction, so by lemma 3.25, this implies $\phi^{(t)}(w^*) = 0$, meaning we now have

$$\text{ord}_{z=w^*}\, \phi(z) \geq T + 1.$$

Notice that this procedure we just detailed for $w^*$ can be done for the rest of elements of $\mathcal{S}$, so in particular we repeat this exact procedure for the rest of the elements of $\mathcal{S}$, obtaining now

$$\text{ord}_{z=w}\, \phi(z) \geq T + 1 \quad \forall w \in \mathcal{S}.$$

We can keep repeating this process iteratively, as the hypotheses of lemmas 3.24 and 3.25 will always be met, and we made sure lemma 3.25 will always yield $\phi^{(t)}(w) = 0$. However, we can only choose $t$ to be as large as $2T_1 - 1$, so when we obtain

$$\text{ord}_{z=w}\, \phi(z) \geq 2T \quad \forall w \in \mathcal{S}$$

after the iterated uses of the lemmas, we must now define $T_1 = 2T$ before iteratively using both lemmas together. We keep doing this until we have to put $T_1 = 4T$, $T_1 = 8T$,..., making the orders of the zeroes of the auxiliary polynomial $\phi(z)$ arbitrarily large, so $\phi(z)$ must be identically 0, but in lemma 3.23 we proved $\phi(z)$ is not identically 0, a contradiction. This means our original assumption that $S > 16\rho D$ was wrong, and we therefore must have $S \leq 16\rho D$, which completes the proof of this theorem.

□

Now, before showcasing how this theorem implies Gelfond - Schneider, and even Hermite - Lindemann, we must prove the following result, which will give us a way to quickly prove the algebraic independence of $z$ and $e^z$, one of the hypotheses needed to use theorem 3.22 to prove the Gelfond - Schneider and Hermite - Lindemann theorems.

**Proposition 3.26.** Let $\lambda_1, ..., \lambda_m$ be different complex numbers, then $e^{\lambda_1 z}, ..., e^{\lambda_m z}$ are linearly independent over $\mathbb{C}(z)$.

**Proof.**

The proof will be by induction on $m$, with the case $m = 1$ being obviously true. Now assume we have rational functions $R_i(z) \in \mathbb{C}(z)$, non-zero by induction hypothesis, such that

$$R_1(z)e^{\lambda_1 z} + \cdots + R_m(z)e^{\lambda_m z} = 0.$$

Without loss of generality we can assume $R_m = 1$ and $\lambda_m = 0$, so this becomes $R_1(z)e^{\lambda_1 z} + \cdots + 1 = 0$, which we can differentiate to obtain yet another equation that is identically 0

$$(R_1' + R_1\lambda_1)e^{\lambda_1 z} + \cdots + (R_{m-1}' + R_{m-1}\lambda_{m-1})e^{\lambda_{m-1} z} = 0.$$

None of the coefficients of the $e^{\lambda_i z}$ are 0, as that would imply that $R_i'/R_i = -\lambda_i$ which for non-zero $\lambda_i$ implies $R_i$ is an exponential function, which is not in $\mathbb{C}(z)$, and if $\lambda_i = 0$, that would contradict the hypothesis that all the $\lambda_j$ are different, and we already have $\lambda_m = 0$. So all the coefficients are in $\mathbb{C}(z)$, but by induction hypothesis, $e^{\lambda_1 z}, ..., e^{\lambda_{m-1} z}$ are $\mathbb{C}(z)-$linearly independent, so the $R_i \in \mathbb{C}(z)$ cannot exist, which completes the proof.

□

We can now finally show how Gelfond - Schneider's theorem is a corollary of theorem 5.15, which highlights how general the proof method used for Hermite - Lindemann can become. It is part of the reason we are writing this thesis in the first place.

**Corollary 3.27** (Gelfond - Schneider)**.** Given an irrational algebraic number $\beta$, and an algebraic number $\alpha \neq 0, 1$. Let $\log(\alpha)$ be any non-zero determination of the logarithm, then $\alpha^\beta = e^{\beta \log(\alpha)}$ is a transcendental number.

**Proof.**

We begin by assuming that $\alpha^\beta = \gamma$ is algebraic, and therefore $K = \mathbb{Q}(\alpha, \beta, \gamma)$ is a finite dimensional number field over $\mathbb{Q}$. We now choose $f_1(z) = e^z$ and $f_2(z) = e^{\beta z}$, both have growth order at most $\rho = 1$, and they are algebraically independent over $\mathbb{C}$ by proposition 3.26. With $K$ as above, since $f_1'(z) = f_1(z)$ and $f_2'(z) = \beta f_2(z)$ both lie in the ring $K[f_1, f_2]$. Then, by theorem 3.22, there can only be at most $16[K : \mathbb{Q}]$ complex numbers such that $f_1$ and $f_2$ are analytic with values in $K$, in particular, there can only be finitely many such complex numbers. However, for every $n \in \mathbb{Z}_{>0} \subset \mathbb{C}$ and a non-zero determination of the logarithm where $\log(\alpha)$ is well defined, $z_n = n \log \alpha$ yields $f_1(z_n) = \alpha^n \in K$ and $f_2(z_n) = \gamma^n \in K$, and both $f_1, f_2$ are analytic at the $z_n$, which contradicts the theorem, hence we can't have $\gamma$ algebraic, which completes the proof of Gelfond - Schneider.

□

It is worth noting that using the same Theorem with the same argument works to prove Hermite - Lindemann too, this time choosing $f_1(z) = z$ and $f_2(z) = e^z$, whose algebraic independence over $\mathbb{C}$ we proved in lemma 2.1, and with $z_n = n\alpha$ for non-zero algebraic $\alpha$, $n \in \mathbb{Z}_{>0}$.

# 4 Elliptic functions

## 4.1 Introduction to elliptic functions

We fittingly begin the chapter by defining elliptic functions and presenting four results regarding the properties of general elliptic functions.

**Definition 4.1.** Let $f$ be a meromorphic function on $\mathbb{C}$, we say that $f$ is an *elliptic function* if there exist two $\mathbb{R}-$linearly independent complex numbers, $\omega_1, \omega_2 \in \mathbb{C}$, such that

$$f(z+\omega_1) = f(z) \quad \text{and} \quad f(z+\omega_2) = f(z) \quad \forall z \in \mathbb{C}.$$

They are *doubly-periodic* functions, and from now on we will assume that $\omega_1, \omega_2$ are the smallest periods of $f$, that is, for both $\omega_1$ and $\omega_2$, there does not exist any $\alpha \in (0,1)$ such that $f(z) = f(z+\alpha\omega_i) \quad \forall z \in \mathbb{C}$. Elliptic functions are doubly-periodic, and this property lends itself well to a tessellation of $\mathbb{C}$ with parallelograms, whose edges are a subgroup of the complex plane called a lattice $\Omega = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$. The parallelogram formed by joining in succession consecutive points of the set $0, \omega_1, \omega_2 + \omega_1, \omega_2, 0$ is called the fundamental parallelogram. To be precise, the fundamental parallelogram is the set

$$\big\{z \in \mathbb{C} \mid z = \alpha\omega_1 + \beta\omega_2 \ ; \ \alpha, \beta \in [0,1)\big\}.$$

The parallelograms that cover $\mathbb{C}$ are also called *meshes*, and clearly all the points in $z + \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$ occupy the same corresponding points in every mesh. These points are said to be *congruent* to each other, which is written as $z = z' \bmod \omega_1, \omega_2$, because $z' = z + \omega$, with $\omega \in \Omega$. From the definition of elliptic functions, it follows that an elliptic function attains the same value at on every set of congruent points, therefore its image over every mesh is the same. In the context of integraton, it is convenient to avoid integrating over a mesh when the elliptic function $f$ has a singularity in its boundary, so in that case, exploiting the periodic properties of $f$, the same exact integral can be evaluated over a translated mesh, which will not be another mesh (it would have the same issues we are trying to avoid). Such a parallelogram is called a cell. We follow this with the following definiton.

**Definition 4.2.** The set of poles, or zeroes, of an elliptic function $f$ in any given cell is called *irreducible*, and they will be referred to as the irreducible set of poles, or zeroes. Every other pole or zero of $f$ is congruent to one of the poles or zeroes of the irreducible set.

We will call any singularity that is not a pole, an *essential singularity*. We now present the four results.

**Proposition 4.3.** The number of zeroes and poles of an elliptic function $f \neq 0$ in any cell is finite.

**Proof.**

First, the number of poles is finite because otherwise the poles of $f$ would have a limit point, which implies that $f$ has an uncountable number of poles, contradicting the fact that $f$ is meromorphic. Now assume the number of zeroes of $f$ in any cell is not finite, then $1/f$ has infinitely many poles in the cell, and therefore the poles have a limit point, which must be an essential singularity. This essential singularity is also an essential singularity

of $f$, contradicting again that $f$ is meromorphic.

□

**Definition 4.4.** Given an elliptic function $f$, we define the *order* of $f$ as the number of poles of $f$ in a cell, counting multiplicity.

**Proposition 4.5.** The sum of the residues of an elliptic function $f$ at its poles in any cell is 0.

**Proof.**

Let $C$ be the cell we will consider, and $\partial C$ the contour of the cell, which has vertices at $t, t+\omega_1, t+\omega_1+\omega_2, t+\omega_2$, for $t \in \mathbb{C}$. The sum of the residues of $f$ at its poles in $C$ is given by

$$\frac{1}{2\pi i}\int_{\partial C} f(z)dz = \frac{1}{2\pi i}\left\{ \int_t^{t+\omega_1} + \int_{t+\omega_1}^{t+\omega_1+\omega_2} + \int_{t+\omega_1+\omega_2}^{t+\omega_2} + \int_{t+\omega_2}^{t} \right\} f(z)dz.$$

With the substitution $u+\omega_2 = z$ on the third integral, and $u+\omega_1 = z$ on the second integral, this is equivalent to

$$\frac{1}{2\pi i}\int_{\partial C} f(z)dz = \frac{1}{2\pi i}\int_t^{t+\omega_1} f(z)dz + \int_t^{t+\omega_2} f(u+\omega_1)du + \int_{t+\omega_1}^{t} f(u+\omega_2)du + \int_{t+\omega_2}^{t} f(z)dz.$$

Which is exactly 0, on account of the double periodicity of $f$.

□

**Theorem 4.6** (Liouville)**.** A holomorphic elliptic function $f$ is constant.

**Proof.**

Consider a cell $C$, since $f$ is holomorphic (no poles) and $C$ is contained in a compact subset of $\mathbb{C}$, $f$ is bounded on $C$. By the periodicity of $f$, it is then bounded on $\mathbb{C}$, and thus constant on $\mathbb{C}$.

□

**Corollary 4.7.** A non constant elliptic function $f$ has order at least 2.

**Proof.**

Assume $f$ has order at most 1, that is, it has at most a simple pole. Proposition 4.5 tells us that the residue at that pole is 0, so $f$ is holomorphic, but by Liouville's theorem, this implies $f$ is constant, a contradiction.

□

**Theorem 4.8.** Given an elliptic function $f$ and a constant $c \in \mathbb{C}$, the number of roots of the equation

$$f(z) = c$$

in any cell is constant, counting multiplicity. Moreover, it is independent of $c$.

**Proof.**

We begin by fixing a cell $C$, such that $f$ has no poles on its contour $\partial C$. From now on, when we refer to the number of zeroes or poles of $f$, we are taking their multiplicities into account too. Now, by Cauchy's argument principle, the difference between the number of zeroes and poles of $f(z)-c$ in the cell $C$ is given by

$$\frac{1}{2\pi i}\int_{\partial C} \frac{f'(z)}{f(z)-c}dz = \frac{1}{2\pi i}\int_{\partial C} g(z)dz,$$

with $g(z) = f'(z)/(f(z)-c)$. By the double periodicity of $f$ and our choice of $C$, we have $f'(z+\omega_1) = f'(z+\omega_2) = f'(z)$ for $z \in \partial C$, and since $f$ shares the exact same periodicity, $g(z+\omega_1) = g(z+\omega_2) = g(z)$ for $z \in \partial C$. Using the same argument as in proposition 6.4 when integrating over $\partial C$, we get

$$\frac{1}{2\pi i}\int_{\partial C} \frac{f'(z)}{f(z)-c} dz = 0.$$

So by Cauchy's argument principle, $f(z) - c$ has the same number of zeroes and poles. However, since $c$ is fixed, these are the same poles as $f(z)$, and the number of poles of $f(z)$ is fixed and independent of 0, which concludes the proof.

□

**Corollary 4.9.** Every elliptic function $f$ has the same number of zeroes and poles in a cell, counting multiplicity.

**Proof.**

Use the above theorem with $c = 0$.

□

We will now look at a particular class of elliptic functions.

## 4.2 The Weirstrass $\wp$ function

The Weirstrass elliptic function, denoted as $\wp(z)$, is a central object in the study of elliptic curves. As we know, the exponential is a periodic function that provides a group homomorphism $\mathbb{R}/\mathbb{Z} \to \mathbb{S}^1$. This notion can be extended in a way to $\mathbb{C}$ through the Weirstrass $\wp$ function, because $\wp$ is doubly-periodic and provides a way to define a group homomorphism from the complex torus to an elliptic curve, which we write as $\mathbb{T} \cong \mathbb{C}/\Omega \to E$. It is therefore not a big surprise that we have elliptic versions of the Hermite - Lindemann and Gelfond - Schneider theorems, but we must first lay the groundwork for that.

We will begin by looking at the Weirstrass form of a general elliptic curve

$$y^2 = 4x^3 - g_2 x - g_3. \tag{4.1}$$

We will always assume an elliptic curve in Weirstrass form when talking about an elliptic curve with invariants $g_2, g_3$. Now, we will assume three distinct roots for the cubic, which is equivalent to $g_2^3 - 27g_3^2 \neq 0$. This assumption allows us to compute two $\mathbb{R}-$linear independent complex numbers $\omega_1$ and $\omega_2$ called periods, which are completely characterized by $g_2$ and $g_3$ [12, Page 41]. The linear combination of these periods over $\mathbb{Z}$ forms a lattice, $\Omega = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$, which is consequently also characterized by $g_2, g_3$. On the other hand, the lattice also completely characterizes the values $g_2, g_3$ by means of the following formulae

$$g_2 = 60 \sum_{\omega \in \Omega \setminus \{0\}} \frac{1}{\omega^4} \quad \text{and} \quad g_3 = 140 \sum_{\omega \in \Omega \setminus \{0\}} \frac{1}{\omega^6} \tag{4.2}$$

([12, Page 41]). Since $g_2$ and $g_3$ completely determine the elliptic curve in Weirstrass form as seen by equation 6.1, we can equivalently determine the elliptic curve from its associated lattice ,$\Omega$, and the constants $g_2, g_3$.

For the rest of this chapter every time we consider a lattice it will be $\Omega = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$, with $\omega_1/\omega_2 \notin \mathbb{R}$, unless otherwise specified. Now, given a lattice $\Omega$, we can define the *Weierstrass Elliptic function* $\wp(z) = \wp_\Omega(z)$ as the following series

$$\wp(z) = \frac{1}{z^2} + \sum_{\omega \in \Omega \setminus \{0\}} \left( \frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right), \tag{4.3}$$

which is well defined and meromorphic, as we will now see. To do this, we will begin by showing that $\wp(z)$ is uniformly and absolutely convergent on every compact subset of $\mathbb{C}$ that does not contain any point of the lattice $\Omega$. Let $K$ be any such compact subset of $\mathbb{C}$, we will use Weirstrass' M-test, so we want to bound

$$\left| \frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right| = \left| \frac{z(2\omega - z)}{(z-\omega)^2 \omega^2} \right|.$$

Let $Z = \max_{z \in K} |z|$, we will first consider the sum on the subset of the lattice $W_1 = \{\omega \in \Omega : |\omega| \leq 2Z,\ \omega \neq 0\} \subset \Omega$. $W_1$ is compact (and finite) and $K \cap \Omega = \emptyset$, so the distance function $d(K, \cdot)$ attains a nonzero minimum over $W_1$, which we will call $A$. Because of this, we can bound

$$\left| \frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right| \leq \left| \frac{1}{(z-\omega)^2} \right| + \left| \frac{1}{\omega^2} \right| \leq \frac{1}{A^2} + \frac{1}{|\min\{\omega_1, \omega_2\}|^2} = \frac{1}{A^2} + \frac{1}{|\omega_{\min}|^2} =: C,$$

where $C$ is a constant. We now consider the rest of the sum on the subset of the lattice $W_2 = \{\omega \in \Omega : |\omega| > 2Z\} \subset \Omega$. Here we bound

$$\left| \frac{z(2\omega - z)}{(z-\omega)^2 \omega^2} \right| \leq \frac{|z|(|2\omega| + |z|)}{(|\omega| - |z|)^2 |\omega^2|} < \frac{|z|(5|\omega|/2)}{(|\omega|/2)^2 |\omega^2|} < \frac{(5|z|/2)}{(|\omega|/2)^2 |\omega|} = \frac{5|z|}{2|\omega|^3} \leq \frac{5Z}{|\omega|^3}.$$

We have therefore found a bound for each term in the sum defining $\wp(z)$ for all $z \in K$, so all that remains to use Weirstrass' M-test is showing that the sum of these bounds converges. $\#W_1$ is always bounded, so we only have to show that $\sum_{\omega \in W_2} \frac{1}{|\omega|^3}$ converges. It could be the case that $W_2 = \Omega \setminus \{0\}$, so instead all that remains is proving the convergence of

$$\sum_{\omega \in \Omega \setminus \{0\}} \frac{1}{|\omega|^3}.$$

To do this, consider the function $f : \mathbb{R}^2 \setminus \{(0,0)\} \longrightarrow \mathbb{R}$ defined by

$$f(n, m) = \frac{|n\omega_1 + m\omega_2|}{|n| + |m|}.$$

Notice that for every $t \in \mathbb{R}$, $f(tn, tm) = f(n, m)$, so we can restrict the domain of $f$ to the unit circle $\mathbb{S}^1$ without changing the image of $f$. $\mathbb{S}^1$ is a compact subset of $\mathbb{R}^2$, and since $f$ is continuous, it attains an infimum over the $\mathbb{S}^1$, which we will call $B$. This infimum is non-zero because $\omega_1$ and $\omega_2$ are $\mathbb{R}-$linearly independent, so going back to the integers now, we have a lower bound

$$|\omega| = |n\omega_1 + m\omega_2| \geq B(|n| + |m|),$$

which allows us to change the summation from $\omega \in W_2 \subset \Omega$ to adding over the possible values of $|n| + |m| = N$. There are $4N$ possible integer combinations $(n, m)$ that give $|n| + |m| = N$, hence

$$\sum_{\omega\in\Omega} \left|\frac{1}{\omega}\right|^3 \leq \frac{1}{B^3} \sum_{\substack{n,m\in\mathbb{Z} \\ (n,m)\neq(0,0)}} \frac{1}{(|n| + |m|)^3} \leq \frac{1}{B^3} \sum_{N=1}^{\infty} \frac{4N}{N^3} = \frac{4\pi^2}{6B^3} < \infty.$$

We can use the same argument for every compact subset of $\mathbb{C} \setminus \Omega$, so in particular, $\wp(z)$ is uniformly and absolutely convergent on every compact subset of $\mathbb{C} \setminus \Omega$. Now, for every finite subset $\Omega_0 \subset \Omega \setminus \{0\}$,

$$S_{\Omega_0}(z) = \sum_{\omega\in\Omega_0} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}\right)$$

is holomorphic over $\mathbb{C} \setminus \Omega$, so by the proven uniform and absolute convergence of $\wp(z)$, $\wp(z)$ is holomorphic too. This implies that given any open $U \subset \mathbb{C}$ such that $\overline{U} \cap \Omega = \emptyset$, $\wp'(z)$ exists on $U$ and is equal to

$$\wp'(z) = -2 \sum_{\omega\in\Omega} \frac{1}{(z-\omega)^3}. \tag{4.4}$$

On the other hand, it is clear by its series representation that $\wp(z)$ has poles of order 2 on $\Omega$, just as $S_{\Omega_0}(z)$ has poles on $\Omega_0$. From all this, we deduce that the Weirstrass $\wp$ function is meromorphic on $\mathbb{C}$.

We will now show that $\wp(z)$ is an elliptic function. All that remains is to prove that it is doubly periodic with periods $\omega_1$ and $\omega_2$. We begin by noting that $-\Omega = \Omega$, and therefore

$$\wp(-z) = \frac{1}{(-z)^2} + \sum_{\omega\in\Omega\setminus\{0\}} \left(\frac{1}{(-z-\omega)^2} - \frac{1}{\omega^2}\right) = \frac{1}{z^2} + \sum_{\omega\in\Omega\setminus\{0\}} \left(\frac{1}{(z+\omega)^2} - \frac{1}{(-\omega)^2}\right) = \wp(z).$$

So $\wp(z)$ is even. Now, clearly by its definition above, $\wp'(z)$ is periodic with periods $\omega_1$ and $\omega_2$. Furthermore, $\wp'(z)$ is absolutely convergent for all $z \in \mathbb{C} \setminus \Omega$, as we showed $\sum_{\omega\in\Omega} |\omega|^{-3} < \infty$ ([13, Page 270] for further details). Because of this there must exist constants $a, b \in \mathbb{C}$ such that for all $z \in \mathbb{C}$,

$$\wp(z + \omega_1) = \wp(z) + a \quad \text{and} \quad \wp(z + \omega_2) = \wp(z) + b.$$

But we know $\wp(z)$ is even, so setting $z = -\omega_1/2$ in the first equation and $z = -\omega_2/2$ in the second equation, we get

$$\wp(\omega_1/2) = \wp(-\omega_1/2) + a \quad \text{and} \quad \wp(\omega_2/2) = \wp(-\omega_1/2) + b.$$

Hence, $a = b = 0$, and $\wp(z)$ is doubly periodic with periods $\omega_1$ and $\omega_2$. We will now see the relationship between $\wp(z)$ and elliptic curves.

Consider the function

$$f(z) = \wp(z) - \frac{1}{z^2} = \sum_{\omega \in \Omega \setminus \{0\}} \left( \frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right).$$

It is holomorphic on a neighbourhood of the origin, as it is $\wp(z)$ with it's pole at the origin removed, and therefore it is analytic in that same neighbourhood. Hence, by Taylor's theorem, we can compute the coefficients of its power series centred on $z = 0$. We note that we can differentiate under the summation sign because of the previously proven uniform and absolute convergence of the series. The $n-$th derivative of $f$ is thus

$$f^{(n)}(z) = (-1)^n (n+1)! \sum_{\omega \in \Omega \setminus \{0\}} \frac{1}{(z-\omega)^{n+2}},$$

which evaluated at 0 for $n = 2, 4$ yields

$$f^{(2)}(0) = 6 \sum_{\omega \in \Omega \setminus \{0\}} \frac{1}{\omega^4} = g_2/10, \quad f^{(4)}(0) = 120 \sum_{\omega \in \Omega \setminus \{0\}} \frac{1}{\omega^6} = 6g_3/7.$$

We are cheating a bit, having already given the expressions for $g_2$ and $g_3$, but this will serve as a way to verify these formulas. Moreover, by the symmetry of $\Omega$, $f^{(n)}(0) = 0$ for odd $n$. We can therefore write

$$\wp(z) = \frac{1}{z^2} + \frac{g_2 z^2}{20} + \frac{g_3 z^4}{28} + O(z^6),$$

and differentiating

$$\wp'(z) = \frac{-2}{z^3} + \frac{g_2 z}{10} + \frac{g_3 z^3}{7} + O(z^5).$$

We want to get rid of the terms with negative degree, so we square $\wp'(z)$ and cube $\wp(z)$ to obtain

$$\wp(z)^3 = \frac{1}{z^6} + \frac{3g_2}{20z^2} + \frac{3g_3}{28} + O(z^2) \quad , \quad \wp'(z)^2 = \frac{4}{z^6} - \frac{2g_2}{5z^2} - \frac{4g_3}{7} + O(z^2).$$

Then

$$\wp'(z)^2 - 4\wp(z)^3 = -\frac{g_2}{z^2} - g_3 + O(z^2),$$

and from our previously computed Laurent series for $\wp(z)$, we finally obtain

$$G(z) = \wp'(z)^2 - 4\wp(z)^3 + g_2\wp(z) + g_3 = O(z^2).$$

Now, the sum of meromorphic funtions is meromorphic, so $G(z)$ is meromorphic, and in particular, since all the added meromorphic functions have the same poles, the poles of $G(z)$ are exactly $\Omega$. Furthermore, $G(z)$ is an elliptic function, as it clearly satisfies $G(z + \omega_1) = G(z + \omega_2) = G(z) \ \forall z \in \mathbb{C}$. On the other hand, $G(z)$ is analytic by construction, and therefore holomorphic, but by Liouville's theorem 4.6, $G(z)$ must then be constant. Taking the limit as $z$ tends to 0 reveals this constant to be 0, so $G(z) = 0$. Therefore we have the following order 1 differential equation

$$\wp'(z)^2 = 4\wp(z)^3 - g_2\wp(z) - g_3. \tag{4.5}$$

This implies that for every $z \in \mathbb{C} \setminus \Omega$, the map $C(z) = (\wp(z), \wp'(z))$ gives a point on the cubic curve (4.1), which is how we define the previously alluded to map from $\mathbb{C} \setminus \Omega$ to

$E$. It is worth noting that the poles of $\wp$ and $\wp'$ are mapped onto the infinity point of the elliptic curve. Due to the periodicity of $\wp$, this map is not one to one, but given suitable conventions regarding the boundary of the fundamental parallelogram (or equivalently, any parallelogram), taking values of $z$ contained in the fundamental parallelogram does make $C(z)$ a one to one map. So the fundamental parallelogram is mapped one to one onto the complex points of the curve (4.1), furthermore, the map $z \mapsto C(z)$ satisfies the following property [12, Page 43]

$$C(z_1 + z_2) = C(z_1) + C(z_2)$$

[12, Page 43], which establishes a homomorphism from the additive group of complex numbers, $\mathbb{G}_a$, to the group of complex points of the cubic, $E(\mathbb{C})$, with its corresponding sum. This property also motivates the addition formula for $\wp$ given by

$$\wp(z_1 + z_2) = -\wp(z_1) - \wp(z_2) + \frac{1}{4}\left(\frac{\wp'(z_1) - \wp'(z_2)}{\wp(z_1) - \wp(z_2)}\right)^2 \tag{4.6}$$

[15, Page 463], where we are back to only using the usual sum over the complex numbers. We will also take the time to present some other identities that we will require later due to the hypothesis of the Schneider - Lang theorem that $f_i' \in K[f_1, ..., f_m]$. To begin with, note that we cannot use equation (4.6) in the case $z_1 = z_2$, so we present the doubling formula for the Weirstrass $\wp$ function

$$\wp(2z) = \frac{16\wp(z)^4 + 8g_2\wp(z)^2 + 32g_3\wp(z) + g_2^2}{16\wp'(z)^2} \tag{4.7}$$

[7, Page 245]. We cannot use this formula when $\wp'(z) = 0$, so we must know what the zeroes of $\wp'$ are. $\wp'$ is an elliptic function too, and because it clearly has poles of order 3 (see equation (4.4)), it must also have 3 zeroes in every mesh, counting multiplicity. Now, from its definition in equation (4.4), it is clear that $\wp'$ has the same periods as $\wp$. Furthermore

$$\wp'(-z) = \sum_{\omega\in\Omega} \frac{1}{(-z-\omega)^3} = -\sum_{\omega\in\Omega} \frac{1}{(z+\omega)^3} = -\wp'(z) \qquad (\text{as } -\Omega = \Omega),$$

so $\wp'$ is an odd function, and because of it's double periodicity too we have

$$-\wp'(\omega_1/2) = \wp'(-\omega_1/2) = \wp'(\omega_1 - \omega_1/2) = \wp'(\omega_1/2),$$

and analogously for $\omega_2$, hence $\wp'(z)$ has a zero on each of its half periods. There is another zero which we can find using the same method

$$-\wp'\left(\frac{\omega_1+\omega_2}{2}\right) = \wp'\left(-\frac{\omega_1+\omega_2}{2}\right) = \wp'\left(\omega_1+\omega_2-\frac{\omega_1+\omega_2}{2}\right) = \wp'\left(\frac{\omega_1+\omega_2}{2}\right),$$

so the zeroes of $\wp'$ are exactly the elements of $\frac{1}{2}\Omega \setminus \Omega$, and thus the doubling formula is valid for all $z \notin \left(\frac{1}{2}\Omega \setminus \Omega\right)$. We now go on a bit of a detour, because we must prove $\wp$ has a finite growth order $\rho$, again, due to the hypotheses of the Schneider - Lang theorem.

**Lemma 4.10.** The Weirstrass $\wp$ function defined over any lattice $\Omega$ has growth order at most $\rho = 2$.

**Proof.**

Just as in the proof for the absolute and uniform convergence of $\wp(z)$, we will begin by considering any compact subset of $\mathbb{C}$, which we will call $K$. Let $Z = \max_{z\in K} |z|$, $\omega_{min} = \min\{\omega_1, \omega_2\}$, and let $W_1 = \{\omega \in \Omega : |\omega| \leq 2Z\}$, $W_2 = \{\omega \in \Omega : |\omega| > 2Z\}$ be two disjoint subsets of $\Omega$. We begin by noting that since $\#W_1$ is finite, we can bound the finite part of the product like so

$$z \prod_{\omega\in W_1} \Big(1-\frac{z}{\omega}\Big) \exp\left(\frac{z}{\omega}+\frac{1}{2}\Big(\frac{z}{\omega}\Big)^2\right) \leq Z\left(1+\frac{Z}{|\omega_{min}|}\right)^{\#W_1} \exp\left(\frac{Z\#W_1}{\omega_{min}}+\frac{\#W_1}{2}\Big(\frac{Z}{\omega_{min}}\Big)^2\right).$$

Now, the order of the product of two functions is the maximum of the order of both functions, as we established in lemma 3.24, and as we will prove now, the order of a polynomial is 0. Consider, for $r \in \mathbb{R}_{>0}$, $n \in \mathbb{N}$

$$e^r = (e^{r/n})^n \geq (1 + r/n)^n \geq 1 + r^n/n^n.$$

With $r = n|z|^{1/n}$ this becomes $e^{n|z|^{1/n}} \geq 1 + |z|$, so for any degree 1 polynomial with complex coefficients we have

$$|az + b| \leq |a||z| + |b| \leq \max\{|a|, |b|\}(|z| + 1) \leq c_1 e^{n|z|^{1/n}}$$

with $c_1 = \max\{|a|, |b|\}$, so for $\rho = 1/n$, and $C = e^n$, $|az+b| \leq c_1 C^{|z|^\rho}$. Since we can make $n$ arbitrarily large, the growth order of any linear factor is 0, and therefore the order of any polynomial is 0. On the other hand there exist positive constants $c_2, c_3$ such that

$$\exp\left(\frac{Z\#W_1}{\omega_{min}} + \frac{\#W_1}{2}\Big(\frac{Z}{\omega_{min}}\Big)^2\right) \leq c_2 c_3^{Z^2}.$$

So this part of the product is bounded over $K$, and in particular has growth order 2. We now deal with the rest of the product. Let

$$\sigma_{W_2}(z) = \prod_{\omega\in W_2} \Big(1 - \frac{z}{\omega}\Big) \exp\left(\frac{z}{\omega} + \frac{1}{2}\Big(\frac{z}{\omega}\Big)^2\right).$$

We will show its absolute and uniform convergence using Weirstrass M-test again, this time applied to

$$\log \sigma_{W_2}(z) = \sum_{\omega\in W_2} \Big(\log(1 - z/\omega) + z/\omega + \frac{1}{2}(z/\omega)^2\Big).$$

We will begin by bounding every term of the sum. For any $\varepsilon \in (0, 1/2)$, we have

$$|\log(1 - z/\omega) + z/\omega + \frac{1}{2}(z/\omega)^2| = \left|\sum_{i=3}^{\infty} \frac{z^i}{i \cdot \omega^i}\right| \leq \frac{1}{3}\Big|\frac{z}{\omega}\Big|^{2+\varepsilon} \left|\sum_{i=3}^{\infty} \Big(\frac{z}{\omega}\Big)^{i-2-\varepsilon}\right|.$$

For $\omega \in W_2$ we have $\left|\frac{z}{\omega}\right| < \frac{1}{2}$, hence

$$\frac{1}{3}\Big|\frac{z}{\omega}\Big|^{2+\varepsilon} \left|\sum_{i=3}^{\infty} \Big(\frac{z}{\omega}\Big)^{i-2-\varepsilon}\right| \leq \frac{1}{3}\Big|\frac{z}{\omega}\Big|^{2+\varepsilon} \sum_{i=1}^{\infty} \Big(\frac{1}{2}\Big)^{i-\frac{1}{2}} = \frac{\sqrt{2}Z^{2+\varepsilon}}{3}\Big|\frac{1}{\omega}\Big|^{2+\varepsilon}.$$

Going back to our original sum, this yields

$$|\log \sigma_{W_2}(z)| \leq \frac{\sqrt{2}Z^{2+\varepsilon}}{3} \sum_{\omega \in W_2} \left|\frac{1}{\omega}\right|^{2+\varepsilon},$$

and the same argument we used before works here to prove the convergence of this sum, as $\sum_{N=1}^{\infty} \frac{4N}{N^{2+\varepsilon}}$ converges. Therefore we have absolute and uniform convergence by Weirstrass' M-test, and since we are taking the product of holomorphic functions for each $\omega \in \Omega$, then $\sigma(z)$ is holomorphic over $K$ too. This can be done for any $K \subset \mathbb{C}$, so in fact $\sigma(z)$ is entire. Additionally, notice that the previous inequality implies

$$\sigma_{W_2}(z) \leq c_4 c_5^{Z^{2+\varepsilon}},$$

for some positive constants $c_4, c_5$. Hence $\sigma_{W_2}(z)$ has growth order 2, and therefore $\sigma(z)$ has growth order 2 as well, due to being the product of two functions of growth order 2. Now that we have done all this work, we will connect this Weirstrass sigma function to the Weirstrass $\wp$ function. Their relationship is as follows:

$$\wp(z) = \frac{\sigma'(z)^2 - \sigma(z)\sigma''(z)}{\sigma(z)^2} \tag{4.8}$$

[7, Page 245]. Recall now lemma 2.4 that allows us to give the following bounds for any $R, r \in \mathbb{R}_{>0}$: $|\sigma'(z)|_r \leq |\sigma(z)|_{r+1}$, and $|\sigma''(z)|_R \leq 2|\sigma(z)|_{R+1}$. These two bounds imply that the growth order of $\sigma'(z)$ and $\sigma''(z)$ is also 2, hence the growth orders of the numerator and denominator of equation (4.8) are at most 2, and by our definition of growth order for meromorphic functions, the growth order of $\wp(z)$ is 2.

□

Now, with all the results we have proven thus far, we are almost ready to use the Schneider - Lang theorem to prove that given non-zero algebraic $\alpha \in \mathbb{C}$, $\wp(\alpha)$ is transcendental, the so called elliptic Hermite - Lindemann theorem. The last prerequisite before that is proving the algebraic independence of $z$ and $\wp(z)$, as clearly from what we have seen, $\wp(z)$ and $\wp'(z)$ for instance are not algebraically independent.

**Lemma 4.11.** The functions $z$ and $\wp(z)$ are algebraically independent over $\mathbb{C}$.

**Proof.**

$\wp(z)$ has two periods but only one will be necessary for the proof, as we will mimic the one for the algebraic independence of $z$ and $e^z$. Assume first that there's a polynomial with coefficients $c_{ij} \in \mathbb{C}$ and minimal degree on $\wp(z)$ such that

$$P(z, \wp(z)) = \sum_{i=0}^{L_1} \sum_{j=0}^{L_2} c_{ij} z^i \wp(z)^j = 0.$$

Since we have chosen a polynomial with minimal degree on $\wp(z)$ (assuming it exists), $P_{L_2} \neq 0$, and we can rewrite this as

$$0 = \sum_{j=0}^{L_2} \wp(z)^j P_j(z) = P_{L_2}(z) \sum_{j=0}^{L_2} \wp(z)^j R_j(z) \implies \sum_{j=0}^{L_2} \wp(z)^j R_j(z) = 0,$$

where $P_j(z) \in \mathbb{C}[z]$ and $R_j(z) = P_j(z)/P_{L_2}(z) \in \mathbb{C}(z)$. Now, exploiting the periodicity of $\wp(z)$, with periods assumed to be $\omega_1$ and $\omega_2$, we can obtain another equation

$$\sum_{j=0}^{L_2} \wp(z+\omega_1)^j R_j(z+\omega_1) = \sum_{j=0}^{L_2} \wp(z)^j R_j(z+\omega_1) = 0.$$

Subtracting these equations, we get

$$\sum_{j=0}^{L_2-1} \wp(z)^j (R_j(z) - R_j(z+\omega_1)) = 0 \tag{4.9}$$

Which would contradict our assumption that the previous polynomial was of minimal degree on $\wp(z)$ and such that $P(z, \wp(z)) = 0$, unless all the coefficients of 6.7 cancel. This would mean that we have $R_j(z+\omega_1) = R_j(z)$, and since $\omega_1 \neq 0$ this implies that the reational functions $R_j$ have infinitely many zeroes and poles, which cannot be the case, and so we reach a contradiction, so $z$ and $\wp(z)$ are algebraically independent.

□

We are now ready to prove the so called elliptic Hermite - Lindemann theorem, which implies a theorem proved by Siegel and later expanded upon by Schneider where it is proven that assuming the invariants $g_2$ and $g_3$ are algebraic implies that the periods of the elliptic function $\wp(z)$ are transcendental [14][Theorems 17,18].

**Theorem 4.12.** Assume the invariants $g_2$ and $g_3$ are algebraic, and that $\alpha \neq 0$ is also algebraic. Then $\alpha$ is not contained in the lattice $\Omega$ associated to the elliptic curve with invariants $g_2, g_3$. Furthermore, $\wp(\alpha)$ is transcendental.

**Proof.**

We begin by showing that $\alpha \notin \Omega$. We will assume $\alpha \in \Omega$, and consider the functions $f_1(z) = z$, $f_2(z) = 1/\wp(z)$ and $f_3(z) = f_2'(z)$, and the field $K = \mathbb{Q}(\alpha, g_2, g_3)$. Note that the following relationship holds

$$f_3'(z) = -2g_3 f_2(z)^3 - \frac{3}{2} g_2 f_2(z)^2 + 2 \tag{4.10}$$

[7, Page 246]. By lemma 4.11 we have the algebraic independence of $f_1$ and $f_2$ over $\mathbb{C}$, and by equation (4.10) it is clear that $f_i'(z) \in K[f_1, f_2, f_3]$. Furthermore, by combining lemmas 4.10 and 2.4, we see that the growth order of $f_1(z)$ is 0, and for $f_2(z)$ and $f_3(z)$ its 2. We can therefore employ the Schneider - Lang theorem. $\wp(z)$ is meromorphic with poles of order 2 on $\Omega$ due to its double periodicity, and $\wp'(z)$ is meromorphic with poles of order 3 on $\Omega$ due for the same reason, so $f_2(z) = 1/\wp(z)$ and $f_3(z) = -\wp'(z)/\wp(z)^2$ have zeroes of order 2 and 1, respectively, on $\Omega$, and more importantly, they are analytic on $\Omega$ too. Considering now $z_n = n\alpha \in \Omega$, with $n \in \mathbb{Z}_{>0}$, we have $f_1(z_n) = n\alpha \in K$, and $f_1$ clearly analytic at $z_n$, $f_2(z_n) = f_3(z_n) = 0 \in K$, and both analytic on $z_n$. However, we have infinite $z_n$ for which this is the case, contradicting the theorem, hence $\alpha \notin \Omega$. Importantly, we can also assume $q\alpha \in \Omega$, for any fixed $q \in \mathbb{Q}$, and letting $s \in \mathbb{Z}_{>0}$ be the denominator of $q$, we can consider this time $z_n = nsq\alpha \in \Omega$. With the same reasoning as before, we reach a contradiction with the Schneider - Lang theorem, and therefore even rational multiples of $\alpha$ cannot belong to $\Omega$.

We now assume $\wp(\alpha)$ is algebraic, consider the functions $f_1(z) = z, f_2(z) = \wp(z), f_3(z) = \wp'(z)$, and let $K = \mathbb{Q}(g_2, g_3, \alpha, \wp(\alpha))$. $f_1$ and $f_2$ are algebraically independent by 4.11 again, and by equation (4.5), and lemma 4.10, $f_1, f_2$ and $f_3$ have bounded growth orders. Moreover, $f_1', f_2', f_3' \in K[f_1, f_2, f_3]$ because we have

$$f_3'(z) = 6f_2(z)^2 - \frac{1}{2} g_2$$

[7, Page 245]. Before continuing, recall the doubling formula (equation (4.7)).

$$\wp(2z) = \frac{16\wp(z)^4 + 8g_2\wp(z)^2 + 32g_3\wp(z) + g_2^2}{16\left(4\wp(z)^3 - g_2\wp(z) - g_3\right)} = \frac{16\wp(z)^4 + 8g_2\wp(z)^2 + 32g_3\wp(z) + g_2^2}{16\wp'(z)^2}.$$

The resultant of the top and bottom polynomials in $\wp$ is $2^{28}(g_2^3 - 27g_3^2)^2$ which is not equal to 0, as all throughout the chapter we assume the cubic polynomial for the elliptic curve has 3 different roots. This implies that if the denominator vanishes at some $z_0 \in \mathbb{C}$, the numerator does not vanish. Differentiating the doubling formula then yields the following monstruous equation for $\wp'(2z)/\wp'(z)$, which we will now use.

$$\frac{\wp'(2z)}{\wp'(z)} = \frac{64\wp(z)^6 - 80g_2\wp(z)^4 - 320g_3\wp(z)^3 - 20g_2^2\wp(z)^2 - 16g_2g_3\wp(z) + g_3^2 - 32g_2^3}{32\wp'(z)^4} \tag{4.11}$$

[7, Page 245].

We are now ready to continue with the proof. Consider this time $z_n = 2^n\alpha$ for $n \in \mathbb{Z}_{>0}$, and if necessary, $z_n = 2^n\alpha$ for $n > n_0$. Clearly $f_1$ is analytic at every $z_n$, and $f_1(z_n) \in K$, futhermore, since we proved before that $\mathbb{Q}\alpha \cap \Omega = \emptyset$, $z_n/2^k$ is not a pole of the doubling formula for every $k \in \mathbb{Z}$, as recall that the zeroes of $\wp'(z)$ are $(\frac{1}{2}\Omega \setminus \Omega)$. Moreover, by repeated use of the doubling formula it is clear that $f_2(z_n) \in K$, and $f_2$ is analytic at $z_n$. For $f_3$ we have

$$f_3(z_n) = \wp'(2^n\alpha) = \wp'(\alpha)\prod_{r=0}^{n-1}\frac{\wp'(2^{r+1}\alpha)}{\wp'(2^r\alpha)},$$

so again using that $\mathbb{Q}\alpha \cap \Omega = \emptyset$ and the formula for $\wp'(2z)/\wp'(z)$ we see that $f_3(z_n) \in K$ and $f_3(z)$ is analytic on $z_n$. This all is true for every positive integer $n$, so we again contradict the Schneider - Lang theorem, and therefore $\wp(\alpha)$ is transcendental.

□

Moving onwards, we will also need the algebraic independence of $\wp(z)$ and $\wp(\beta z)$, but in this case we need a stronger hypothesis on $\beta$, because depending on the lattice $\Omega$, $\wp(z)$ and $\wp(\beta z)$ may in fact be algebraically dependent (linearly even) for non rational $\beta$. For example, if we fix $\Omega = \omega_0(\mathbb{Z} + i\mathbb{Z})$, for some $\omega_0 \in \mathbb{C}$, then

$$\wp(iz) = \frac{-1}{z^2} + \sum_{\omega\in\Omega\setminus\{0\}}\left[\frac{-1}{(z+i\omega)^2} - \frac{1}{\omega^2}\right] = \frac{-1}{z^2} - \sum_{\omega\in\Omega\setminus\{0\}}\left[\frac{1}{(z-(-i\omega))^2} - \frac{1}{(-i\omega)^2}\right] = -\wp(z).$$

As $-i\Omega = \Omega$. This example provides some insight into what type of restriction we should place on $\beta$.

**Lemma 4.13.** Consider the functions $\wp(z)$ and $\wp(\beta z)$, as well as their associated lattice $\Omega$, with basis elements $\omega_1, \omega_2$. If $\beta \notin \mathbb{Q}(\omega_1/\omega_2)$, then $\wp(z)$ and $\wp(\beta z)$ are algebraically independent over $\mathbb{C}$.

**Proof.**

Let $\mathcal{K}$ be the field of meromorphic functions, and assume there is some non-zero polynomial $P \in \mathbb{C}[X, Y]$ such that $P(\wp(z), \wp(\beta z)) = 0$. Now, since $\wp(z)$ is not constant, and we have in fact seen $\wp(z)$ is transcendental for algebraic $z$, the polynomial $Q(Y) = P(\wp(z), Y)$ is not identically zero. Then, for any $\omega \in \Omega$ and any $m \in \mathbb{Z}$,

$$0 = P(\wp(z + m\omega), \wp(\beta z + \beta\omega m)) = Q(\wp(\beta z + \beta\omega m)).$$

Since $Q$ is a polynomial, it has finitely many zeroes, and therefore there exist two different integers $m', m''$ such that $\wp(\beta z+\beta m'\omega) = \wp(\beta z+\beta m''\omega)$. $\wp(z)$ is not injective, but we can compare the poles of both functions, which implies $\beta z + \beta m'\omega - (\beta z + \beta m''\omega) = \omega' \in \Omega$. Taking $\omega = r\omega_1 + s\omega_2$ and $\omega' = a\omega_1 + b\omega_2$ this yields

$$\beta = \frac{a\omega_1 + b\omega_2}{(m' - m'')(r\omega_1 + s\omega_2)} = \frac{a\omega_1/\omega_2 + b}{(m' - m'')(r\omega_1/\omega_2 + s)} \in \mathbb{Q}(\omega_1/\omega_2),$$

a contradiction.

□

We are now ready for the elliptic analogue of Gelfond - Schneider.

**Theorem 4.14.** Assume that the invariants $g_2, g_3$ are algebraic, and suppose that the lattice $\Omega$ they correspond to has basis elements $\omega_1, \omega_2$. Further, assume that $\beta$ is algebraic over $\mathbb{Q}$, but does not belong to $\mathbb{Q}(\omega_1/\omega_2)$. Then for every complex $u \notin \Omega$ such that $\wp(u)$ is algebraic, $\beta u \notin \Omega$ and $\wp(\beta u)$ is transcendental.

**Proof.**

As before, we will begin by proving that $\beta u \notin \Omega$ by contradiction, so assume $\beta u \in \Omega$. We will split this into 2 cases for simplicity, instead of treading through iterated uses of the addition formula (4.6). We first assume that $u = \omega/m$ for some $\omega \notin m\Omega$ and $m \in \mathbb{Z}_{>0}$. We will use the functions $f_1(z) = 1/\wp(z), f_2(z) = 1/\wp(\beta z), f_3(z) = f_1'(z), f_4(z) = f_2'(z)$, with $f_1, f_2$ satisfying the hypothesis of algebraic independence by the previous lemma. All the growth orders are bounded by some $\rho$, and for $K = \mathbb{Q}(g_2, g_3, \beta)$; $f_1', f_2', f_3', f_4' \in K[f_1, f_2, f_3, f_4]$ by equation (4.10). This time we choose $z_n = nu$ for $n$ such that $m|n$, so $z_n \in \Omega$, and by the assumption that $\beta u \in \Omega$, we also have $\beta z_n = \frac{n}{m}\beta u \in \Omega$. Because of this, $f_1(z)$ and $f_2(z)$ have zeroes of order 2 on every $z_n$, and are therefore analytic there. Similarly,

$$f_3(z) = \frac{-\wp'(z)}{\wp(z)^2} \quad \text{and} \quad f_4(z) = \frac{-\beta\wp'(\beta z)}{\wp(\beta z)^2}$$

have zeroes of order 1 on every $z_n$, and are therefore analytic there, too. So all the functions are analytic at $z_n$ for every $n$ that is divisible by $m$, and

$$f_1(z_n) = f_2(z_n) = f_3(z_n) = f_4(z_n) = 0 \in K.$$

Clearly there is an unbounded number of $z_n$ that satisfy these conditions, contradicting the Schneider - Lang theorem, and therefore $\beta u \notin \Omega$ in this case.

We now assume again $\beta u \in \Omega$, but $u$ is not a fraction of a period. Recall that we have assumed $\wp(u)$ algebraic, hence $\wp'(u)^2 \in \mathbb{Q}(\wp(u))$, and therefore $\wp'(u)$ cannot be transcendental as there exists a polynomial $P(X^2)$ such that $P((\wp'(u))^2) = 0$. $\wp'(u)$ is therefore algebraic and we can consider $K = \mathbb{Q}(g_2, g_3, \beta, \wp(u), \wp'(u))$, plus the functions $f_1(z) = \wp(z)$, $f_2(z) = 1/\wp(\beta z)$, $f_3(z) = f_1'(z)$, $f_4(z) = f_2'(z)$. Their growth orders are all bounded by some $\rho$, $f_1$ and $f_2$ are algebraically independent by the previous lemma, and because of equation (4.10) and

$$\wp''(z) = 6\wp(z)^2 - \frac{1}{2}g_2 \tag{4.12}$$

[7, Page 245], $f_1', f_2', f_3', f_4' \in K[f_1, f_2, f_3, f_4]$. Finally, let $z_n = 2^n u$ for $n \in \mathbb{Z}_{>0}$, since $\beta u \in \Omega$ but $u \notin \Omega/m \ \forall m \in \mathbb{Z}$, ($\beta z_n \in \Omega$ but $z_n \notin \Omega$) $\forall n \in \mathbb{Z}_{>0}$. Recall $f_4(z) = \frac{\beta\wp'(\beta z)}{\wp(\beta z)^2}$,

so $f_2$ and $f_4$ are analytic on $z_n$ for every $n$, and in fact equal to $0 \in K$ there. On the other hand, to see that $f_1(z_n), f_3(z_n) \in K$ we will need the doubling formulas, but it is otherwise clear that they are both analytic on $z_n \notin \Omega$. $\wp(u)$ and $\wp'(u)$ are algebraic, so by the doubling formula (4.7), $\wp(2u) \in \mathbb{Q}(\wp(u))$. Repeating this process inductively yields $\wp(z_n) \in K$ for all $n \in \mathbb{Z}_{>0}$. For $f_3$ now, again $\wp'(u), \wp(u)$ are algebraic, so the ratio $\wp'(2u)/\wp'(u)$ from equation (4.11) is algebraic, and iteratively applying

$$f_3(z_n) = \wp'(2^n u) = \wp'(u) \prod_{r=0}^{n-1} \frac{\wp'(2^{r+1}u)}{\wp'(2^r u)}$$

for every power of 2 implies $\wp'(z_n) \in K$ for all $n \in \mathbb{Z}_{>0}$. Again, we have an unbounded number of $z_n$, but this contradicts the Schneider-Lang theorem, hence $\beta u \notin \Omega$.

The second part of the proof is almost identical, step by step. This time we begin by assuming that $\wp(\beta u)$ is algebraic. Notice that just as before, this implies $\wp'(\beta u) \in \mathbb{Q}(\wp(\beta u))$. *From here* we split the proof into two cases again. In the first case we assume $u = \omega/m$ for some $\omega \notin m\Omega$. We will consider $K = \mathbb{Q}(g_2, g_3, \beta, \wp(u), \wp(\beta u))$, and use $f_1(z) = 1/\wp(mz), f_2(z) = \wp(\beta z), f_3(z) = f_1'(z), f_4(z) = f_2'(z)$. $f_1$ and $f_2$ are algebraically independent because $\beta/m \notin \mathbb{Q}(\omega_1/\omega_2)$, and all four functions have a bounded growth order. Now, by differential equation (6.12) and

$$\left(\frac{1}{\wp(z)}\right)'' = -2g_3\frac{1}{\wp(z)^3} - \frac{3g_2}{2}\frac{1}{\wp(z)^2} + 2 \tag{4.13}$$

[7, Page 246], $f_1', f_2', f_3', f_4' \in K[f_1, f_2, f_3, f_4]$. We then consider $z_n = 2^n u$, so $mz_n \in \Omega$ and $\beta u \notin \Omega$, and therefore all four functions are analytic on $z_n$ for all $n \in \mathbb{Z}_{>0}$ with

$$f_1(z_n) = f_3(z_n) = 0 \quad f_2(z_n) = \wp(2^n \beta u) \in K \quad f_4(z_n) = \beta\wp'(2^n \beta u) \in K.$$

We know $f_2(z_n)$ and $f_4(z_n)$ belong to $K$ by using the doubling formulas (4.7) and (4.11) in the same way they have been used before. Now, this all contradicts the Schneider-Lang theorem that says there should only be finitely many such $z_n$, so we reach a contradiction in this case.

Finally, we now assume $u$ is not the fraction of a period. We will consider $K = \mathbb{Q}(g_2, g_3, \beta, \wp(u), \wp(\beta u))$, and use $f_1(z) = \wp(z), f_2(z) = \wp(\beta z), f_3(z) = f_1'(z), f_4(z) = f_2'(z)$. $f_1$ and $f_2$ are algebraically independent because $\beta \notin \mathbb{Q}(\omega_1/\omega_2)$, and all four functions have a bounded growth order. Again $f_1', f_2', f_3', f_4' \in K[f_1, f_2, f_3, f_4]$, and we consider $z_n = 2^n u$. Since $\beta z_n, z_n \notin \Omega \ \forall n \in \mathbb{Z}_{>0}$, $f_1, f_2, f_3, f_4$ are analytic on $z_n$, and since $\wp'(\beta u) \in \mathbb{Q}(\wp(\beta u))$ and $\wp'(u) \in \mathbb{Q}(\wp(u))$, by the doubling equations (4.7) and (4.11)

$$f_1(z_n), f_2(z_n), f_3(z_n), f_4(z_n) \in K \ \ \forall n \in \mathbb{Z}_{>0}.$$

As usual, this contradicts the Schneider-Lang theorem, because there is an unbounded number of these $z_n$'s, and since we have reached a contradiction in both cases, our original assumption that $\wp(\beta u)$ is algebraic must be wrong, hence, $\wp(\beta u)$ is transcendental.

□

We will now show a specific use of these previous two theorems, which consists on proving the transcendence of any $u \in \mathbb{C} \setminus \Omega$ such that $\wp(u)$ is algebraic. These numbers classically arise from definite integrals, which includes integrals that evaluate to interesting constants. This comes from the fact that $\wp(z)$ parametrizes the Weirstrass form of an elliptic curve by turning it into a differential equation, in a sense:

$$\Big(\frac{d\wp(u)}{du}\Big)^2 = 4\wp(u)^3 - g_2\wp(u) - g_3.$$

After substituting $x = \wp(u)$, and specifying the path plus the choice of square root, the differential equation yields

$$\int_{\wp(u^*)}^{\wp(u)} \frac{1}{\sqrt{4x^3 - g_2x - g_3}}dx = u - u^*.$$

To give an example of a concrete use of this, we can choose $\Omega = \omega_0(\mathbb{Z} + i\mathbb{Z})$ for simplicity, as pairing $i\omega$ with $\omega$ implies $g_3 = 0$, on account of $(i\omega)^6 = -\omega^6$. Notice that this implies $g_2 \neq 0$ because we are still assuming 3 distinct roots in the cubic equation, and in fact by fixing $\omega_0$ we can make $g_2$ take on any value. We will choose $g_2 = 4$, which allows us to factorize $4x^3 - 4x = 4x(x-1)(x+1)$. Now, recall that the zeroes of $\wp'(z)$ are precisely $\big(\frac{1}{2}\Omega \setminus \Omega\big)$, and for this particular elliptic curve we have

$$\wp'(z)^2 = 4\wp(z)(\wp(z) - 1)(\wp(z) + 1).$$

This implies the existence of two different $u, u^* \in \big(\frac{1}{2}\Omega \setminus \Omega\big)$ such that $\wp(u) = 1$ and $\wp(u^*) = 0$. Furthermore, by the double periodicity of $\wp$, $u, u^*$ are two different elements of $\big\{\frac{\omega_0}{2}, \frac{i\omega_0}{2}, \frac{\omega_0 + i\omega_0}{2}\big\}$, so $u - u^* \in \frac{1}{2}\Omega$, and there exist $n, m \in \big\{-\frac{1}{2}, 0, \frac{1}{2}\big\}$ such that

$$u - u^* = (n + mi)\omega_0.$$

This is crucial, because while we know $u$ and $u^*$ are transcendental because $\wp(u)$ and $\wp(u^*)$ are algebraic, we would not be able to say whether or not $u - u^*$ is transcendental (to this day we do not even know whether $e + \pi$ is transcendental), but since $u - u^*$ is the product of an algebraic number $n + im$ and $\omega_0$, we know $u - u^*$ is transcendental. Recall that we know $\omega_0$ is transcendental because by the elliptic Hermite-Lindemann, if $g_2$ and $g_3$ are algebraic, then for any algebraic $\alpha \neq 0$, we have $\alpha \notin \Omega$. All that remains is evaluating the integral, to know what value we have proven to be transcendental. In this case, with the substitution $t = x^2$ we get

$$\int_0^1 \frac{1}{\sqrt{4x^3 - 4x}}dx = \frac{-i}{4}\int_0^1 \frac{1}{t^{3/4}(1-t)^{1/2}} = \frac{-i}{4}\mathrm{B}\Big(\frac{1}{4}, \frac{1}{2}\Big),$$

where $\mathrm{B}(\cdot, \cdot)$ is the beta function. The beta function is real for these inputs, and therefore $n = 0$ and $m = \frac{1}{2}$ or $m = -\frac{1}{2}$. The sign is not relevant in this case however, and we could simply choose $\omega_0 = \frac{1}{2}\mathrm{B}\big(\frac{1}{4}, \frac{1}{2}\big)$, as the lattice $\Omega$ would remain unaffected. Regardless of these details though, we have obtained the transcendence of $\mathrm{B}\big(\frac{1}{4}, \frac{1}{2}\big)$. Exploring this result a bit more, it is well known that for the Beta formula we have

$$\mathrm{B}(z_1, z_2) = \frac{\Gamma(z_1)\Gamma(z_2)}{\Gamma(z_1 + z_2)},$$

where $\Gamma(z)$ is the gamma function. Hence

$$\frac{\Gamma(1/4)\Gamma(1/2)}{\Gamma(3/4)} = \frac{\Gamma(1/4)\sqrt{\pi}}{\Gamma(3/4)}$$

is transcendental too. We can now also use Euler's reflection formula for the gamma function, $\Gamma(1-z)\Gamma(z) = \frac{\pi}{\sin(\pi z)}$ which implies $\Gamma(3/4)\Gamma(1/4) = \pi\sqrt{2}$, so

$$\frac{\Gamma(1/4)^2}{\sqrt{2\pi}}$$

is transcendental, and since $\sqrt{2}$ is algebraic, we also obtain the transcendence of

$$\frac{\Gamma(1/4)^2}{\sqrt{\pi}}.$$

# 5 Baker's Theorem

## 5.1 Introduction and corollaries

We now return to the study of more elementary functions, in particular a type of function that falls outside the scope of the powerful Schneider - Lang theorem: the logarithm. This occurs because $\frac{d}{dx}\log(x) = \frac{1}{x}$, so we can never have a set of functions $f_i$ including the logarithm such that $f_i' \in K[f_i]$, one of the hypothesis in the Schneider-Land theorem. Nonetheless, the Gelfond - Schneider theorem can be applied to obtain the following corollary:

**Corollary 5.1.** Given non-zero algebraic numbers $\alpha_1, \alpha_2$, with $\alpha_2 \neq 1$, and $\frac{\log(\alpha_1)}{\log(\alpha_2)} \notin \mathbb{Q}$, then

$$\beta_1 \log(\alpha_1) + \beta_2 \log(\alpha_2) \neq 0$$

for all $\beta_1, \beta_2 \in \overline{\mathbb{Q}}$. That is, if $\log(\alpha_1)$ and $\log(\alpha_2)$ are $\mathbb{Q}-$linearly independent, then they are $\overline{\mathbb{Q}}-$linearly independent too.

**Proof.**
Let $r = \frac{\log(\alpha_1)}{\log(\alpha_2)} \notin \mathbb{Q}$, and assume that $r$ is algebraic. Then $\alpha_2^r = \alpha_1$, but by the Gelfond - Schneider theorem, this implies $\alpha_1$ is transcendental, a contradiction.

□

The Gelfond - Schneider theorem is only strong enough for the linear independence of two logarithms, however, this gives hope that a more general version of this result will be even stronger than the Gelfond - Schneider theorem. This more general result is Baker's theorem.

**Theorem 5.2** (Baker)**.** Given non-zero algebraic numbers $\alpha_1, ..., \alpha_n$ such that their logarithms $\log(\alpha_1), ..., \log(\alpha_n)$ are $\mathbb{Q}-$linearly independent, then $1, \log(\alpha_1), ..., \log(\alpha_n)$ are $\overline{\mathbb{Q}}-$linearly independent, too. Where $\overline{\mathbb{Q}}$ are the algebraic numbers.

Before presenting the proof of this theorem, we will state without proof two of its corollaries, to show some its consequences.

**Theorem 5.3.** For any non-zero algebraic numbers $\alpha_1, ..., \alpha_n$, not all 1, and any algebraic numbers $\beta_0, \beta_1, ..., \beta_n$, with $\beta_0 \neq 0$, we have

$$\beta_0 + \beta_1 \log(\alpha_1) + \cdots \beta_n \log(\alpha_n) \neq 0.$$

Hence, this value is always transcendental.

**Theorem 5.4.** For any non-zero algebraic numbers $\alpha_1, ..., \alpha_n, \beta_0, \beta_1, ..., \beta_n$,

$$e^{\beta_0} \alpha_1^{\beta_1} \cdots \alpha_n^{\beta_n}$$

is transcendental.

## 5.2 Necessary lemmas for Baker's theorem

As usual, we will need a few lemmas to prove the theorem. This section is dedicated to the proof of these lemmas, before moving on to the proof of the theorem. As we did for Schneider - Lang, we will also take the hypotheses of Baker's theorem into account for

the lemmas, as we will need them, and proving the theorem is the whole purpose of the lemmas in the first place. The method of proof strongly resembles the one used before for the Hermite - Lindemann and Schneider - Lang theorems, so we again start by assuming that Baker's theorem is false, i.e given the hypotheses of the theorem, there exist algebraic numbers $\beta_0, \beta_1, ..., \beta_n$, not all 0, such that

$$\beta_0 + \beta_1 \log(\alpha_1) + \cdots \beta_n \log(\alpha_n) = 0.$$

Note that we can assume all $\alpha_i \neq 1$ without loss of generality, as they do not have any relevance to the theorem otherwise. Our job henceforth is deriving a contradiction from this. To begin with, since not all $\beta_i$ are 0, we can assume without loss of generality $\beta_n = -1$, which implies

$$e^{\beta_0} \alpha_1^{\beta_1} \cdots \alpha_{n-1}^{\beta_{n-1}} = \alpha_n. \tag{5.1}$$

Again, in this chapter we will denote by $c, c_1, c_2, ...$ positive constants which depend only on the $\alpha_i, \beta_i$, and the determinations of the logarithms, and we will reserve the use of the letter $h$ for constants which will be chosen to be as large as needed.

**Lemma 5.5.** Given an algebraic number $\alpha$ of degree $d$ over $\mathbb{Q}$ with minimal polynomial

$$Q(X) = A_0 X^d + A_1 X^{d-1} + ... + A_{d-1} X + A_d \in \mathbb{Z}[X]$$

(Which is also the minimal polynomial of $\alpha$ over $\mathbb{Q}$ by Gauss' lemma.) Then for every $r \geq 0$, it is possible to write

$$(A_0 \alpha)^r = A_{d-1}^{(r)} \alpha^{d-1} + A_{d-2}^{(r)} \alpha^{d-2} + ... + A_1^{(r)} \alpha + A_0^{(r)},$$

where the terms $A_i^{(r)}$ are integers satisfying $|A_i^{(r)}| \leq (2|Q|)^r$ for all $i \in \{0, 1, ..., d-1\}$.

**Proof.**

The proof will be done by induction on $r$. Clearly the statement is true for $0 \leq r \leq d-1$ since we can take $A_i^{(r)} = A_0^r \delta_{ir}$. It is also true for $r = d$ since we can use the minimal polynomial, but it is not necessary for the proof. We assume now that the lemma is true for $r = l$, and will prove its thus true for $r = l + 1$. By hypothesis:

$$\begin{aligned} A_0^{l+1} \alpha^{l+1} &= A_0 \alpha \cdot \sum_{i=0}^{d-1} A_i^{(l)} \alpha^i = A_0 A_{d-1}^{(l)} \alpha^d + A_0 \alpha \sum_{i=0}^{d-2} A_i^{(l)} \alpha^i \\ &= A_{d-1}^{(l)} \sum_{i=0}^{d-1} -A_{d-i} \alpha^i + A_0 \sum_{i=0}^{d-2} A_i^{(l)} \alpha^{i+1} = A_{d-1}^{(l)} \sum_{i=0}^{d-1} -A_{d-i} \alpha^i + A_0 \sum_{i=1}^{d-1} A_{i-1}^{(l)} \alpha^i, \end{aligned}$$

where on the last step we have employed $Q(\alpha) = 0$. Merging the sums we now have:

$$A_0^{l+1} \alpha^{l+1} = \sum_{i=1}^{d-1} \left\{ (A_0 A_{i-1}^{(l)} - A_{d-i} A_{d-1}^{(l)}) \alpha^i \right\} - A_d A_{d-1}^{(l)},$$

meaning we can set, for all $i \in \{1, ..., d-1\}$

$$A_i^{(l+1)} = A_0 A_{i-1}^{(l)} - A_{d-i} A_{d-1}^{(l)} \qquad \text{and} \qquad A_0^{(l+1)} = -A_d A_{d-1}^{(l)}.$$

And now to finish up we are ready to come up with a bound for the coefficients in the case $r = l + 1$. By induction hypothesis we have

$$|A_i^{(l+1)}| \leq 2|Q| \cdot \max_i \{|A_i^{(l)}|\} \leq 2|Q| \cdot (2|Q|)^l \leq (2|Q|)^{l+1}.$$

□

Applying this lemma on the $\alpha_1, ..., \alpha_n, \beta_0, ..., \beta_{n-1}$, assuming $d$ is the maximum of their degrees, we can write

$$(a_i\alpha_i)^j = \sum_{s=0}^{d-1} a_{is}^{(j)}\alpha_i^s, \quad (b_i\beta_i)^j = \sum_{s=0}^{d-1} b_{is}^{(j)}\beta_i^s, \tag{5.2}$$

where the $a_i, b_i$ are the respective leading coefficients of the minimal polynomials of the $\alpha_i, \beta_i$. We will say all the integers $b_{is}^{(j)}, a_{is}^{(j)}$ are bounded by $c_1^j$. Furthermore, we will keep this notation for the rest of the proof, when we come back to use (5.2). Before continuing, for brevity, we will write

$$f_{m_0,...,m_{n-1}}(z_0, ..., z_{n-1}) = \frac{\partial^{m_0+\cdots+m_{n-1}}}{\partial^{m_0} z_0 \cdots \partial^{m_{n-1}} z_{n-1}} f(z_0, ..., z_{n-1}).$$

We can now proceed with the construction of the auxiliary function $\Phi$.

**Lemma 5.6.** There are integers $p(\lambda_0, ..., \lambda_n)$, not all 0, with absolute values at most $e^{h^3}$, such that the function

$$\Phi(z_0, ..., z_{n-1}) = \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_0, ..., \lambda_n) z_0^{\lambda_0} e^{\lambda_n\beta_0 z_0} \alpha_1^{\gamma_1 z_1} \cdots \alpha_{n-1}^{\gamma_{n-1} z_{n-1}},$$

where $\gamma_r = \lambda_r + \lambda_n\beta_r$ $(1 \le r < n)$ and $L = \lfloor h^{2-\frac{1}{4n}} \rfloor$, satisfies

$$\Phi_{m_0,...,m_{n-1}}(l, ..., l) = 0 \tag{5.3}$$

for all integers $1 \le l \le h$, and all non-negative integers $m_0, ..., m_{n-1}$ with $m_0 + \cdots + m_{n-1} \le h^2$. We remark now, that we will be using this same $\Phi(z, ..., z)$ in the following lemmas too, as it will be used in the proof of the theorem. Therefore, unless specified otherwise, $\Phi(z, ..., z)$ refers to this function we have just defined.

**Proof.**

We begin by applying Leibniz's rule. Note, however, that because of equation (5.1), it suffices to determine the $p(\lambda_0, ..., \lambda_n)$ such that

$$\sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_0, ..., \lambda_n) q(\lambda_0, \lambda_n, l) \alpha_1^{\lambda_1 l} \cdots \alpha_{n-1}^{\lambda_{n-1} l} \alpha_n^{\lambda_n l} \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}} = 0 \tag{5.4}$$

for all the values of $l, m_1, ..., m_{n-1}$ we are considering. Notice that we do not consider the $\log(\alpha_i)$ terms we obtain from differentiating as they are not zero and factor outside the sum. Moreover, the new factor in the sum is given in general by

$$q(\lambda_0, \lambda_n, z) = \sum_{\mu_0=0}^{m_0} \binom{m_0}{\mu_0} \frac{\lambda_0!}{(\lambda_0 - \mu_0)!} (\lambda_n\beta_0)^{m_0-\mu_0} z^{\lambda_0-\mu_0}. \tag{5.5}$$

We now proceed multiplying (5.4) by

$$P' = (a_1 \cdots a_n)^{Ll} b_0^{m_0} \cdots b_{n-1}^{m_{n-1}}. \tag{5.6}$$

Writing

$$\gamma_r^{m_r} = \sum_{\mu_r}^{m_r} \binom{m_r}{\mu_r} \lambda^{m_r-\mu_r} (\lambda_r\beta_r)^{\mu_r},$$

it becomes clear we multiplied by $P'$ in order to substitute (5.2) into (5.4) to obtain perhaps the worst set of equations I have ever seen in my life:

$$\sum_{s_1=0}^{d-1} \cdots \sum_{s_n=0}^{d-1} \sum_{t_0=0}^{d-1} \cdots \sum_{t_{n-1}}^{d-1} A(\mathbf{s},\mathbf{t}) \alpha_1^{s_1} \cdots \alpha_n^{s_n} \beta_0^{t_0} \cdots \beta_{n-1}^{t_{n-1}} = 0, \tag{5.7}$$

for all the values of $l, m_1, ..., m_{n-1}$ we are considering. While this does not look all that bad right now, we have

$$A(\mathbf{s},\mathbf{t}) = \sum_{\lambda_1=0}^{L} \cdots \sum_{\lambda_n=0}^{L} \sum_{\mu_0=0}^{m_0} \cdots \sum_{m_{n-1}}^{m_{n-1}} p(\lambda_0, ..., \lambda_n) q' q'' q''', \tag{5.8}$$

with $q', q'', q'''$ given by

$$q' = \prod_{r=1}^{n} a_r^{(L-\lambda_r)l} a_{rs_r}^{(\lambda_r l)},$$

$$q'' = \prod_{r=1}^{n-1} \binom{m_r}{\mu_r} (b_r \lambda_r)^{m_r-\mu_r} \lambda_n^{\mu_r} b_{rt_r}^{(\mu_r)},$$

$$q''' = \binom{m_0}{\mu_0} \frac{\lambda_0!}{(\lambda_0-\mu_0)!} \lambda_n^{m_0-\mu_0} b_n^{\mu_0} l^{\lambda_0-\mu_0} b_{0t_0}^{(m_0-\mu_0)}.$$

Way back in $A(\mathbf{s},\mathbf{t})$, we used $\mathbf{s} = (s_1, ..., s_n)$ and $\mathbf{t} = (t_0, ..., t_{n-1})$, and assumed $d$ is the maximum of the degrees of the $\alpha_i$ and $\beta_j$ over $\mathbb{Q}$.

Now, clearly if we have $A(\mathbf{s},\mathbf{t}) = 0$ for all the $d^{2n}$ possibilities of $(\mathbf{s},\mathbf{t})$, then all the equations in (5.7), and equivalently (5.4) are satisfied. The equations for $A(\mathbf{s},\mathbf{t}) = 0$ are linear equations with integer coefficients given by the sums of $q'q''q'''$, with the $p(\lambda_0, ..., \lambda_n)$ being the unknowns. All that is left is bounding the coefficients of these equations to use Siegel's lemma 1.2. Recall all the coefficients from (5.2) are bounded by $c_1$, plus by definition $l \leq h$, and $\binom{m_r}{\mu_r} \leq 2^{m_r}$, hence

$$|q'| \leq \prod_{r=1}^{n} a^{(L-\lambda_r)l} c_1^{\lambda_r l} \leq c_2^{Lh},$$

$$|q''| \leq \prod_{r=1}^{n-1} c_3^{m_r} L^{m_r-\mu_r} L^{\mu_r} \leq \prod_{r=1}^{n-1} (c_3 L)^{m_r},$$

$$|q'''| \leq 2^{m_0} \lambda_0^{\mu_0} \lambda_n^{m_0-\mu_0} b_n^{\mu_0} h^L c_1^{m_0-\mu_0} \leq (c_3 L)^{m_0} h^L.$$

Now, as mentioned before, the coefficients are given by

$$\sum_{\mu_0=0}^{m_0} \cdots \sum_{m_{n-1}}^{m_{n-1}} q'q''q'''.$$

The number of terms in this sum is $\prod_{r=0}^{n-1}(m_r+1) \leq \prod_{r=0}^{n-1} 2^{m_r} \leq 2^{h^2}$ by hypothesis. We can therefore bound the absolute value of the coefficients by

$$2^{h^2} (c_3 L)^{m_0+\cdots+m_{n-1}} c_2^{Lh} h^L \leq (2c_3 L)^{h^2} c_4^{Lh}.$$

We therefore have $(L+1)^{n+1}$ unknowns, and $d^{2n}$ equations $A(\mathbf{s},\mathbf{t})=0$ for every possible combination of $(l, m_0, ..., m_{n-1})$. $l \leq h$ and $m_r \in \{0, ..., h^2\}$, so we therefore have at the absolute most $d^{2n}h(h^2+1)^n$ equations. All that's left is confirming we can use Siegel's lemma 1.2 with $U=(2c_3L)^{h^2}c_4^{Lh}$, $M=d^{2n}h(h^2+1)^n$, $N=(L+1)^{n+1}$. Recall $L=\lfloor h^{2-\frac{1}{4n}} \rfloor$ by hypothesis, so for large enough $h$ we have

$$N > (h^{2-\frac{1}{4n}})^{n+1} \geq h^{2n+\frac{3}{2}} \geq 2d^{2n}h(h^2+1)^n \geq 2M.$$

Notice that with this we also get $\frac{M}{N-M} < 1$. In any case, with this, by Siegel's lemma 1.2 this system of equations can be solved non trivially with the integers $p(\lambda_0, ..., \lambda_n)$ having absolute value bounded by

$$NU \leq h^{2n+2}(2c_3L)^{h^2}c_4^{Lh} \leq e^{h^3},$$

for sufficiently large $h$, as we wanted to prove.

□

**Lemma 5.7.** Let $m_0, ..., m_{n-1}$ be any non-negative integers satisfying $m_0 + \cdots + m_{n-1} \leq h^2$, and let

$$f(z) = \Phi_{m_0,...,m_{n-1}}(z, ..., z).$$

Then, for any $z$ we have $|f(z)| \leq c_5^{h^3+L|z|}$, furthermore, for any positive integer $l$, either $f(l)=0$ or $|f(l)| > c_6^{h^3-Ll}$.

**Proof.**

We have an explicit formula for $f(z)$:

$$f(z) = P\sum_{\lambda_0=0}^{L}\cdots\sum_{\lambda_n=0}^{L} p(\lambda_0, ..., \lambda_n)q(\lambda_0, \lambda_n, z)\alpha_1^{\lambda_1 z}\cdots\alpha_n^{\lambda_n z}\gamma_1^{m_1}\cdots\gamma_{n-1}^{m_{n-1}}, \tag{5.9}$$

with $q(\lambda_0, \lambda_n, z)$ defined as in lemma 5.6 and

$$P = (\log(\alpha_1))^{m_1}\cdots(\log(\alpha_{n-1}))^{m_{n-1}},$$

as alluded to in the previous lemma. Notice that we do not have a logarithm of $\alpha_n$, because of how $\Phi$ is defined. From equation (5.5) for $q(\lambda_0, \lambda_n, z)$ we have

$$\begin{aligned}|q(\lambda_0, \lambda_n, z)| &\leq \sum_{\mu_0=0}^{m_0}\binom{m_0}{\mu_0}\lambda_0^{\mu_0}\lambda_n^{m_0-\mu_0}\beta_0^{m_0-\mu_0}\max\{1,|z|\}^{\lambda_0}\\ &\leq (c_7L)^{m_0}\max\{1,|z|\}^{L}\sum_{\mu_0=0}^{m_0}\binom{m_0}{\mu_0} \leq (2c_7L)^{m_0}|z|^L,\end{aligned}$$

where we have been able to change $\max\{1,|z|\}^L$ by $|z|^L$ by using the fact that we will only ever take integer values of $z$, so the only possible problem is at $z=0$, but there both the bound and $q(\lambda_0, \lambda_n, 0)$ are equal to zero, so the bound remains valid. On the other hand, we also have

$$|\alpha_1^{\lambda_1 z}\cdots\alpha_n^{\lambda_n z}| \leq c_8^{L|z|}, \quad |P\gamma_1^{m_1}\cdots\gamma_{n-1}^{m_{n-1}}| \leq (c_9L)^{m_1+\cdots+m_{n-1}},$$

where from the definition of $\gamma_r$ it is clear it can be bounded by some constant multiplied by $L$. Now, the number of terms in the sum for $f(z)$ is at the absolute most $(L+1)^{n+1} = (\lfloor h^{2-1/4n} \rfloor + 1)^{n+1} \leq h^{2(n+1)}$ for large enough $h$, so we obtain

$$|f(z)| \leq h^{2h+2}(c_9L)^{m_1+\cdots+m_{n-1}}c_8^{L|z|}(2c_7L)^{m_0}|z|^L e^{h^3} \leq c_5^{h^3+L|z|}$$

as $m_0 + \cdots + m_{n-1} \leq h^2$. We now move on to the second result of this lemma. Recall $\gamma_r = \lambda_r + \lambda_n \beta_r$ and take $P'$ as defined in equation (5.6), then

$$f_l = \frac{P'}{P} f(l) = P' \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_0, ..., \lambda_n) q(\lambda_0, \lambda_n, l) \alpha_1^{\lambda_1 l} \cdots \alpha_n^{\lambda_n l} \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}}$$

is an algebraic integer of degree at most $d^{2n}$, as $P' \alpha_1^{\lambda_1 l} \cdots \alpha_n^{\lambda_n l} \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}}$ is the product of $2n-1$ algebraic integers, each of degree at most $d$. Moreover, with the same estimates we just did to bound $|f(z)|$, any conjugate of $f_l$ obtained from substituting any $\alpha_i, \beta_j$ by any of their conjugates also has absolute value bounded by $c_{10}^{h^3+Ll}$. By proposition 1.6 then, if $f_l \neq 0$, then $||f_l||^{-d^{2n}} \leq |f_l|$ and since we also have $||f_l|| \leq c_{10}^{h^3+Ll}$, we obtain the result

$$\left(c_{10}^{-(h^3+Ll)}\right)^{d^{2n}} = c_6^{-h^3-Ll} \leq |f_l|.$$

□

**Lemma 5.8.** Let $J$ be any integer such that $0 \leq J \leq (8n)^2$. Then (5.3) holds for all integers $l$ with $1 \leq l \leq \lfloor h^{1+J/8n} \rfloor$ and all non-negative integers $m_0, ..., m_{n-1}$ with $m_0 + \cdots + m_{n-1} \leq \lfloor h^2/2^J \rfloor$.

**Proof.**

We will prove this via strong induction. The result is true for $J = 0$ by lemma 5.6, so we now assume that given an integer $0 \leq K < (8n)^2$, the lemma is true for all $J \in \{0, ..., K\}$, and using this we will prove the lemma for $J = K+1$. Let $R_J = \lfloor h^{1+J/8n} \rfloor, S_J = \lfloor h^2/2^J \rfloor$, and notice that there is an overlap between the values of $l$ and $m_0, ..., m_{n-1}$ from an inductive step to the next, in particular, increasing the value of $J$ reduces the possible combinations of $m_0, ..., m_{n-1}$ such that $m_0 + \cdots + m_{n-1} \leq h^2/2^J$, but increases the possible values of $l$. It therefore suffices to prove that for any integer $l$ such that $R_K < l \leq R_{K+1}$, and any set of non-negative integers $m_0, ..., m_{n-1}$ with $m_0 + \cdots + m_{n-1} \leq S_{K+1}$, we have $f(l) = 0$, where $f(z)$ is as defined in (5.9). By induction hypothesis, $f(r) = 0$ for all integers $r, m$ with $1 \leq r \leq R_K$ and $0 \leq m_0 + \cdots + m_{n-1} = m \leq S_{K+1} < S_K$. Now, let $j_0, ..., j_{m-1}$ be non-negative integers such that $j_0 + \cdots + j_{n-1} = m$ too. We define $f_m(r)$ to be

$$\left( \frac{\partial^{j_0}}{\partial^{j_0} z_0} + \cdots + \frac{\partial^{j_{n-1}}}{\partial^{j_{n-1}} z_{n-1}} \right) \Phi_{m_0,...,m_{n-1}}(z_0, ..., z_{n-1})$$

evaluated at $z_0 = ... = z_{n-1} = r$, which is equal to

$$\sum_{j_0+\cdots+j_{n-1}=m} \frac{m!}{j_0! \cdots j_{n-1}!} \Phi_{m_0+j_0,...,m_{n-1}+j_{n-1}}(r, ..., r).$$

The derivatives here are all 0, because $j_0 + \cdots + j_{n-1} + m_0 + \cdots + m_{n-1} = 2m \leq 2S_{K+1} \leq S_K$, so $f_m(r) = 0$. Thus,

$$F(z) = [(z-1) \cdots (z-R_K)]^{S_{K+1}}$$

divides $f(z)$, so $f(z)/F(z)$ is a regular function on the closed disc $D$ centred on the origin and with radius $R = R_{K+1} h^{1/8n}$. Now, by the maximum modulus principle

$$|f(l)| \cdot \inf_{z \in \partial D} |F(z)| \leq |F(l)| \cdot |f(l)| \leq |F(l)| \cdot \sup_{z \in \partial D} |f(z)|. \tag{5.10}$$

Furthermore;

- $\inf_{z\in\partial D}|F(z)| \geq \prod_{i=1}^{R_K}(R-i) \geq \left(\frac{1}{2}R\right)^{R_K S_{K+1}}$,
- by lemma 5.7, $\sup_{z\in\partial D}|f(z)| \leq c_5^{h^3+LR}$,
- $|F(l)| \leq R_{K+1}^{R_K S_{K+1}}$,
- by lemma 5.7, either $f(l)=0$ or $|f(l)| > c_6^{-h^3-Ll}$.

Note that the bounds on the infimum of $|F(z)|$ do not contradict the bound on $|F(l)|$ because we do not have a "minimum modulus principle" for $F$ because $F(z)=0$ inside $D$. Now, assuming $f(l)\neq 0$ and joining these bounds with (5.10) yields

$$\left(\frac{1}{2}h^{1/8n}\right)^{R_K S_{K+1}} \leq (c_6c_5)^{h^3+LR}.$$

On the other hand

$$LR \leq h^{2-1/4n}h^{1+K/8n} < h^{3+K/8n} \leq 2^{K+2}S_{K+1}2R_K = 2^{K+3}S_{K+1}R_K,$$

which implies

$$\left(\frac{1}{2}h^{1/8n}\right)^{R_K S_{K+1}} \leq (c_6c_5)^{2^{K+4}S_{K+1}R_K} \implies h \leq (2c_6c_5)^{2^{(8n)^2+4}8n}.$$

Recall $n$ is fixed, as at the beginning of the proof of Baker's theorem we consider any number $n$ of algebraic numbers, but this choice at the start fixes $n$. $h$ is not fixed however, and can be as large as desired, which leads to a contradiction here, and therefore $f(l)=0$. The lemma follows by induction.

□

**Lemma 5.9.** Let $\phi(z)=\Phi(z,...,z)$, then

$$|\phi_j(z)| < e^{-h^{8n}} \quad (0\leq j\leq h^{8n}) \tag{5.11}$$

**Proof.**
This proof will be for the most part almost identical to the one in the previous lemma. By lemma 5.8 we know that (5.3) holds for all integers $l$ and non-negative integers $m_0,...,m_{n-1}$ such that $0\leq l\leq X$ and $m_0+\cdots+m_{n-1}\leq Y$, for $X=h^{8n}$ and $Y=\lfloor h^2/2^{(8n)^2}\rfloor$. Hence, just as in the previous lemma, we have $\phi_m(r)=0$ for integers $m,r$ such that $0\leq r\leq X$ and $0\leq m\leq Y$. Defining now

$$E(z)=[(z-1)\cdots(z-X)]^Y$$

it again follows that $E$ divides $\phi$ and $\phi(z)/E(z)$ is regular on the closed disc $\Gamma$, centred on the origin and with radius $R=Xh^{1/8n}$. By the maximum modulus principle, with a similar argument as in the previous lemma, we have, for every $\omega$ with $|\omega|<X$:

$$|\phi(\omega)| \leq \frac{\sup_{z\in\partial\Gamma}|\phi(z)|}{\inf_{z\in\partial\Gamma}|E(z)|}|E(\omega)|.$$

Moreover, $|E(\omega)|\leq(2X)^{XY}$, $\inf_{z\in\partial\Gamma}|E(z)| \geq \left(\frac{1}{2}R\right)^{XY}$ and $\sup_{z\in\partial\Gamma}|\phi(z)| \leq c_5^{h^3+LR}$. Thus

$$|\phi(\omega)| \leq c_5^{h^3+LR}\left(\frac{1}{4}h^{1/8n}\right)^{-XY},$$

and since

$$LR \leq h^{8n+2} \leq 2^{(8n)^2+2}XY,$$

then

$$|\phi(\omega)| \leq \left(\frac{1}{4}h^{1/8n}c_5^{-2^{(8n)^2+3}}\right)^{-XY}.$$

As usual, we can make $h$ as large as we want, so in particular we can make $h$ large enough so that

$$|\phi(\omega)| \leq e^{-XY}.$$

To finish, by Cauchy's formulae,

$$\phi_j(0) = \frac{j!}{2\pi i}\int_{|\omega|=1}\frac{\phi(\omega)}{\omega^{j+1}}d\omega,$$

which implies

$$|\phi_j(0)| \leq \frac{j^j}{2\pi}2\pi|\phi(\omega)| \leq j^j e^{-XY}.$$

Now, we thus have

$$|\phi_j(0)| \leq j^j e^{-X}e^{-(Y-1)X} \leq X^X e^{-(Y-1)X}e^{-X} = (Xe^{-(Y-1)})^X e^{-X}, \tag{5.12}$$

so if we verify $Xe^{-(Y-1)} < 1$ for large enough $h$ we are done. Indeed,

$$e^{8n\log(h)}e^{-(Y-1)} \leq e^{8n\log(h)-(h^2/2^{(8n)^2}-2)} < 1$$

for sufficiently large $h$ and thus

$$|\phi_j(0)| \leq e^{-X} = e^{-h^{8n}} \quad (0 \leq j \leq h^{8n}).$$

□

We will prove this following lemma as a bit of a short break, and because we will need it in the following lemma.

**Lemma 5.10.** For all $z \in \mathbb{C}$, we have

$$|e^z - 1| \leq e^{|z|} - 1 \leq |z|e^{|z|}.$$

**Proof.**
We have

$$|e^z - 1| = \left|\sum_{i=1}^{\infty}\frac{z^n}{i!}\right| \leq \sum_{i=1}^{\infty}\frac{|z|^n}{i!} = e^{|z|} - 1.$$

For the following part of the lemma we can simply consider $t \in \mathbb{R}_{\geq 0}$. We know that $e^{-t} \geq 1 - t$ for all $t \in \mathbb{R}$, and in particular for $t \in \mathbb{R}_{\geq 0}$. The second inequality

$$te^t \geq e^t - 1$$

follows from there, and all that's left is letting $t = |z|$.

**Lemma 5.11.** Given any integers $t_1, ..., t_n$, not all 0, and with absolute values at most $T$, we have, for the algebraic numbers $\alpha_1, ..., \alpha_n$

$$|t_1 \log(\alpha_1) + \cdots + t_n \log(\alpha_n)| > c_{11}^{-T}.$$

**Proof.**
Let $a_j \quad (1 \leq j \leq n)$ be the leading coefficient in the minimal polynomial of $\alpha_j$ or $\alpha_j^{-1}$ accordingly, depending on whether $t_j \geq 0$ or $t_j < 0$ respectively, so that each $a_j/\alpha_j$ is an algebraic integer. This implies that

$$\omega = a_1^{|t_1|} \cdots a_n^{|t_n|}(\alpha_1^{t_1} \cdots \alpha_n^{t_n} - 1)$$

is an algebraic integer with degree at most $d^n$ (the degree of each $\alpha_j$ is bounded by $d$). Now, any conjugate of $\omega$ obtained by substituting arbitrary conjugates of $\alpha_1, ..., \alpha_n$ has absolute value at most $c_{12}^T$. We now consider two cases. If $\omega = 0$, then $\alpha_1^{t_1} \cdots \alpha_n^{t_n} = 1$ and thus

$$\Omega = t_1 \log(\alpha_1) + \cdots + t_n \log(\alpha_n)$$

must be a non-zero multiple of $2\pi i$, as $\Omega \neq 0$ by the hypothesis of linear independence from Baker's theorem. In this case the lemma is trivially true as $|\Omega| \geq 2\pi$. On the other hand, if $\omega \neq 0$, by proposition 1.6, $|\omega| \geq ||\omega||^{-(d^n-1)} \geq c_{12}^{-Td^n}$. Using now the previous lemma, we have

$$a_1^{|t_1|} \cdots a_n^{|t_n|}|e^{\Omega} - 1| = |\omega| \leq a_1^{|t_1|} \cdots a_n^{|t_n|}|\Omega|e^{|\Omega|} = |\Omega|e^{|\Omega|}c_{13}^T.$$

Since we can further assume that $|\Omega| < 1$, as otherwise the lemma is trivially true, $e^{|\Omega|}$ is not arbitrarily large, and from the previous inequalities we get

$$c_{12}^{-Td^n} c_{13}^{-T} e^{-|\Omega|} \leq |\Omega|,$$

so $c_{11}^{-T} < |\omega|$ for some constant $c_{11}$, as desired.

□

**Lemma 5.12.** Let $R, S$ be positive integers strictly larger than 1, and let $\sigma_0, ..., \sigma_{R-1}$ be distinct complex numbers. We define $\sigma$ as the maximum of $1, |\sigma_0|, ..., |\sigma_{R-1}|$ and define $\rho$ as the minimum of 1 and the $|\sigma_i - \sigma_j|$ with $0 \leq i < j < R$. Then, for any integers $r, s$ with $0 \leq r < R$ and $0 \leq s < S$, there exist complex numbers $w_i$ $(0 \leq i < RS)$ with absolute values at most $(8\sigma/\rho)^{RS}$ such that the polynomial

$$W(z) = \sum_{j=0}^{RS-1} w_j z^j$$

satisfies $W_j(\sigma_i) = 0$ for all $i, j$ with $0 \leq i < R$, $0 \leq j < S$ except $i = r$ and $j = s$, where $W_s(\sigma_r) = 1$.
**Proof.**
The polynomial we are looking for is given by

$$W(z) = \frac{-1}{2\pi i s!} \int_{C_r} \frac{(\zeta - \sigma_r)^s U(z)}{(\zeta - z)U(\zeta)} d\zeta,$$

where

$$U(z) = [(z - \sigma_0) \cdots (z - \sigma_{R-1})]^S$$

and $C_r$ is a circle with centre $\sigma_r$ and a sufficiently small radius, in particular smaller than $\rho$ and $|z-\sigma_r|$ for $z \neq \sigma_r$ (notice $z$ is fixed when computing the integral). With our integrand and our definition of $C_r$ we can apply the Residue theorem, so we lay the groundwork:

$$\lim_{|\zeta|\to\infty} |\zeta| \left| \frac{(\zeta-\sigma_r)^s U(z)}{(\zeta-z)U(\zeta)} \right| = 0$$

so the residue at infinity is 0, and therefore since the sum of all the residues is 0, we have

$$\int_{C_r} \frac{(\zeta-\sigma_r)^s U(z)}{(\zeta-z)U(\zeta)} d\zeta = -\sum_{\substack{j=0 \\ j\neq r}}^{R-1} \int_{C_j} \frac{(\zeta-\sigma_r)^s U(z)}{(\zeta-z)U(\zeta)} d\zeta - 2\pi i \mathrm{Res}\left( \frac{(\zeta-\sigma_r)^s U(z)}{(\zeta-z)U(\zeta)}, z \right),$$

where $C_j$ is a circle about $\sigma_j$ with sufficiently small radius (again smaller than $\rho$ and $|z-\sigma_j|$ for $z \neq \sigma_j$, for each $j$). On the other hand

$$\mathrm{Res}\left( \frac{(\zeta-\sigma_r)^s U(z)}{(\zeta-z)U(\zeta)}, z \right) = \lim_{\zeta\to z} (\zeta-z) \frac{(\zeta-\sigma_r)^s U(z)}{(z-z)U(\zeta)} = (\zeta-\sigma_r)^s,$$

so using now the residue theorem yields

$$W(z) = \frac{(z-\sigma_r)^s}{s!} + \frac{U(z)}{2\pi i s!} \sum_{\substack{j=0 \\ j\neq r}}^{R-1} \int_{C_j} \frac{(\zeta-\sigma_r)^s}{(\zeta-z)U(\zeta)} d\zeta.$$

By definintion, $z \notin C_j$, so the sum over $j$ is a rational function of $z$, and regular at $z = \sigma_r$. Further, since $U(z)$ has a 0 of order $S$ at $z = \sigma_r$, it is clear that $W_j(\sigma_r) = 1$ if $j = s$ and 0 otherwise, including $W_j(\sigma_i) = 0$ for every other possible pair $i, j$.

We will now consider another equivalent formulation of $W(z)$ to prove the rest of its properties. By Cauchy's integral formulae we have

$$\frac{-1}{2\pi i s!} \int_{C_r} \frac{(\zeta-\sigma_r)^s U(z)}{(\zeta-z)U(\zeta)} d\zeta = \frac{-1}{2\pi i s!} \cdot \frac{2\pi i}{t!} \frac{d^t}{d\zeta^t} \left( \frac{(\zeta-\sigma_r)^S U(z)}{(\zeta-z)U(\zeta)} \right) \Bigg|_{\zeta=\sigma_r},$$

where $t = S - s - 1$. Note that

$$\frac{(\zeta-\sigma_r)^S}{(\zeta-z)U(\zeta)} = \frac{1}{(\zeta-z)\prod_{\substack{i=0 \\ i\neq r}} (\zeta-\sigma_i)^S},$$

so

$$W(z) = \frac{(-1)^{t-1}}{s!} U(z) \sum_{\substack{j_0+\cdots+j_{R-1}=t \\ j_i \geq 0 \ \forall i}} v(j_0, ..., j_{R-1}) (\sigma_r - z)^{-j_r-1},$$

where

$$v(j_0, ..., j_{R-1}) = \prod_{\substack{i=0 \\ i\neq r}}^{R-1} \binom{S+j_i-1}{j_i} (\sigma_r - \sigma_i)^{-S-j_i}.$$

Now, $t = S-s-1 \geq S-1 \geq j_r$, so $j_r+1 \in \{1, ..., S\}$, and therefore $W(z)$ is a polynomial, as $(\sigma_r - z)^{j_r+1}$ always divides $U(z)$. Moreover, $W(z)$ has degree at most $RS-1$. From $(\sigma_r - z)^{j_r+1} | U(z)$ we also notice that $W(z)$ has a zero of order $S$ at $z = \sigma_i$, $i \neq r$, so

$W_j(\sigma_i) = 0$ for all $j < S$. Now, for the absolute values of the coefficients of $W(z)$, notice first that by definition of $v(j_0, ..., j_{R-1})$ we have

$$|v(j_0, ..., j_{R-1})| \leq \prod_{\substack{i=0 \\ i \neq r}}^{R-1} 2^{S+j_i-1} \rho^{-S-j_i} = \frac{2^{(R-1)(S-1)}}{\rho^{S(R-1)}} \prod_{\substack{i=0 \\ i \neq r}}^{R-1} \frac{2^{j_i}}{\rho^{j_i}} \leq \frac{2^{(R-1)(S-1)}}{\rho^{S(R-1)}} \frac{2^{S-1}}{\rho^{S-1}} = \frac{2^{R(S-1)}}{\rho^{RS-1}}$$

so $|v(j_0, ..., j_{R-1})| \leq (2/\rho)^{RS}$. On the other hand, the coefficients of $U(z)(\sigma_r - z)^{-j_r-1}$ have absolute values at most $(\sigma + 1)^{RS}$, and the sum over the non-negative $j_0, ..., j_{R-1}$ terms such that their sum is equal to $t$ has at the absolute most $S^R$ terms as $j_i \in \{0, 1, ..., S-1\}$. Hence the coefficients of $W(z)$ have absolute values at most

$$S^R(\sigma + 1)^{RS}(2\rho)^{RS} \leq 2^{RS}(2\sigma)^{RS}(2\rho)^{RS} = (8\sigma/\rho)^{RS}.$$

□

### 5.3 Proof of Baker's theorem

Recall the statement of the theorem, already given as Theorem 5.2. To prove it, we will see that the inequalities (5.11) cannot all be true, and this contradiction will prove the theorem. To begin with, let $L = \lfloor h^{2-\frac{1}{4n}} \rfloor$ for some $h$ as in lemma 5.6, let also $S = L + 1, R = S^n$, and note that every integer $0 \leq i < RS$ has a unique base $S$ representation given by

$$i = \lambda_0 + \lambda_1 S + \cdots + \lambda_n S^n,$$

where the $\lambda_0, ..., \lambda_n$ are integers between 0 and $L$ inclusive. For every such $i$ we define

$$\nu_i = \lambda_0, \quad p_i = p(\lambda_0, ..., \lambda_n),$$

and we put

$$\psi_i = \lambda_1 \log(\alpha_1) + \cdots + \lambda_n \log(\alpha_n).$$

By lemma 5.6 and equation (5.1) then, we have

$$\phi(z) = \sum_{i=0}^{RS-1} p_i z^{\nu_i} e^{\psi_i z}. \tag{5.13}$$

Note that this is where the $\overline{\mathbb{Q}}-$linear independence of 1 with the logarithms of the algebraic numbers appears, because equation (5.1) comes from asuming there exist algebraic $\beta_i$ such that $\beta_0 + \beta_1 \log(\alpha_1) + \cdots + \beta_n \log(\alpha_n) = 0$. Continuing now, by lemma 5.11, any two $\psi_i$ which correspond to distinct sets $\lambda_1, ..., \lambda_n$ differ by at least $c_{11}^{-L}$, as $L$ is an upper bound of $|\lambda_1|, ..., |\lambda_n|$. Now, note that there are exactly $S^n = R$ distinct $\psi_i$ because of the abscence of $\lambda_0$; we will denote these $\psi_i$ in some order by $\sigma_0, ..., \sigma_{R-1}$. Let $\rho$ and $\sigma$ be defined as in lemma 5.12, then clearly $\sigma \leq c_{14}L$ and $\rho \geq c_{15}^{-L}$.

Let now $t \in \mathbb{Z}_{\geq 0}$ be any suffix such that $p_t \neq 0$, let $s = \nu_t$, let $r$ be the suffix for which $\psi_t = \sigma_r$, and let $W(z)$ be the polynomial given by lemma 5.12. By the properties of $W(z)$ proven in that very lemma, using $W(z)$ as a Kronecker delta of sorts we have

$$p_t = \sum_{i=0}^{RS-1} p_i W_{\nu_i}(\psi_i),$$

further, by Leibniz's theorem, we have

$$W_{\nu_i}(\psi_i) = \sum_{j=0}^{RS-1} j(j-1)\cdots(j-\nu_i+1)w_j\psi_i^{j-\nu_i} = \sum_{j=0}^{RS-1} w_j \left[\frac{d^j}{dz^j}(z^{\nu_i}e^{\psi_i z})\right]_{z=0},$$

which combined with equation (5.13) yields

$$p_t = \sum_{j=0}^{RS-1} w_j\phi_j(0).$$

Now, $RS \leq h^{(2-\frac{1}{4n})(n+1)} \leq h^{2n+2}$, and by lemma 5.9 now, this implies the inequalities in (5.11) must be true for $0 \leq j \leq RS$. Moreover, by lemma 5.12 we have

$$|w_j| \leq (8\sigma/\rho)^{RS} \leq (8c_1 4Lc_1^L 5)^{RS} \leq c_{16}^{h^{2n+2}}.$$

However, $p_t$ is a non-zero intger, so $|p_t| \geq 1$, but from lemma 5.9 and this bound above we have

$$1 \leq |p_t| \leq RSc_{16}^{h^{2n+4}}e^{-h^{8n}}$$

which implies

$$0 \leq \log RS + c_{17}h^{2n+4} - h^{8n}.$$

This is clearly not possibly for large enough $h$, so at least one of the previously proven inequalities cannot be true, and this contradiction completes the proof of the theorem.

□

# References


[1] Baker, A. *Transcendental number theory.* Cambridge university press, 1975.

[2] Bombieri, E., and Gubler, W. *Heights in Diophantine geometry.* No. 4. Cambridge university press, 2006.

[3] Hindry, M., and Silverman, J. H. *Diophantine geometry: an introduction*, vol. 201. Springer, 2013.

[4] Jacobson, N. *Basic Algebra II: Second Edition.* Dover Books on Mathematics. Dover Publications, 2012.

[5] Koblitz, N. *p-adic Numbers, p-adic Analysis, and Zeta-Functions*, vol. 58. Springer, 2012.

[6] Lang, S. *Algebra.* Graduate Texts in Mathematics. Springer, 2005.

[7] Masser, D. *Auxiliary polynomials in number theory*, vol. 207. Cambridge University Press, 2016.

[8] Milne, J. S. Algebraic number theory (v3.08), 2020.

[9] Neukirch, J. *Algebraic number theory*, vol. 322. Springer, 2013.

[10] Siegel, C. L. Über die perioden elliptischer funktionen. *Journal für die reine und angewandte Mathematik* (1932).

[11] Silverman, J. H. *The arithmetic of elliptic curves*, vol. 106. Springer, 2009.

[12] Silverman, J. H., and Tate, J. T. *Rational points on elliptic curves*, vol. 9. Springer, 1992.

[13] Stein, E. M., and Shakarchi, R. *Princeton lectures in analysis.* Princeton University Press Princeton, 2003.

[14] Waldschmidt, M. Elliptic functions and transcendence. surveys in number theory, 2008.

[15] Whittaker, E. T., and Watson, G. N. *A course of modern analysis: an introduction to the general theory of infinite processes and of analytic functions; with an account of the principal transcendental functions.* University press, 1920.