TRANSCENDENTAL NUMBERS AND A THEOREM OF

A. BAKER

Transcendental Numbers and a theorem of A. Baker

Cameron L. Stewart

Abstract

In this report I intend to survey the theory of transcendental numbers and investigate its most important recent result. That result is A. Baker's theorem, "Linear Forms in the Logarithms of Algebraic Numbers" first published in 1966.

I have divided my thesis into three chapters. In chapter I I outline the history of transcendental numbers and prove some of the more important results in the field. In chapter II I prove A. Baker's theorem. Finally in chapter III I discuss some of the applications and consequences of A. Baker's theorem.

Mathematics Master of Science

July 27,1972

Nombres transcendants et un théorème de A. Baker

Cameron L. Stewart

Résumé

Dans cet ouvrage je me propose de donner une vue d'ensemble des nombres transcendants et du résultat le plus récent qui s'y rattache. Ce résultat, dû à A. Baker parut pour la première fois en 1966 sous le titre "Linear Forms in the Logarithms of Algebraic Numbers".

Ma thèse comprend trois chapitres. Le premier chapitre esquisse l'histoire des nombres transcendants et prouve quelques résultats de la plus haute importance. Dans le chapitre deux je prouve le théorème de Baker. Enfin dans le chapitre trois je discute quelques applications et conséquences du théorème de Baker.

Mathematiques Maitrîse ès sciences

27 Juillet 1972

TRANSCENDENTAL NUMBERS AND A THEOREM OF A. BAKER

CAMERON L. STEWART

To my family

and to L., K., K., L., Poon, the polar bear, the giraffe, the thing and even the gorilla.

August 2, 1972

This thesis is presented in partial fulfillment of the requirements for the degree of Master of Science at McGill University, Montreal, Canada.

I would like to thank my advisor Professor Ian Connell for his help and support. I would also like to express my appreciation for the opportunities McGill University has afforded me.

I would also like to give my special thanks to my typist Miss Lois Campbell without whose patient help this thesis could not have been completed on time.

Table of Contents

Table of Contents

Notation

Z - the integers

Z^+ - the positive integers

$\mathbb{C}$ - the complex numbers

Q - the rational numbers

$||a||$ - the maximum of the absolute value of all the conjugates of a.

$|a|$ - the absolute value of a

$N(a)$ - the norm of a

$H(a)$ - the height of a, the maximum of the absolute values of the relatively prime integer coefficients in a's minimal defining polynomial.

$[a]$ - denotes the greatest integer less than or equal to a.

$g(x) = O(f(x))$ - $|f(x)| < Mg(x)$ for a constant $M > 0$

$f^{(n)}(z)$ - the n^{th}-derivative of $f(z)$

$\binom{n}{m}$ - $\dfrac{n!}{(n-m)!m!}$

CHAPTER I

A SURVEY OF THE THEORY OF TRANSCENDENTAL NUMBERS

1.1 Transcendental Numbers

In this chapter I will attempt to outline the main developments in the field of transcendental numbers. I will prove in detail both the Gelfond-Schneider theorem and Schneider's theorem. I hope to bring the reader to a position of appreciating the significance of a recent theorem of A. Baker [1] in the theory of transcendental numbers.

While A. Baker's result is applicable to a wide range of problems it is of particular importance to the theory of transcendental numbers for it gives the natural generalization of the Gelfond-Schneider theorem. The Gelfond-Schneider theorem states that if a is algebraic and not 0 or 1, and if b is algebraically irrational then a^b is transcendental. As a consequence of Baker's theorem we have the following: If $a_1,\ldots,a_n$ denote algebraic numbers that are neither 0 or 1 and if $b_1,\ldots,b_n$ denote algebraic numbers with $1, b_1,\ldots,b_n$ linearly independent over the rationals then $a_1^{b_1}\ldots a_n^{b_n}$ is transcendental.

Before commencing it might be in order to remind the reader that an algebraic number is a number that satisfies a polynomial equation $f(x) = a_m x^m + \ldots + a_1 x + a_o = 0$ where the a_i are rational integers and $a_m \neq 0$. A complex number that is not algebraic is transcendental. We see immediately that the algebraic numbers are countable.

1.2 Liouville's theorem and its improvements.

The first result in the theory is due to Liouville. He determined an estimate for how well a real algebraic number can be approximated by a rational number.

Liouville's Theorem. If r is a real algebraic number with minimum polynomial of degree $n>1$ over Q, then there exists an $m > 0$ depending only on r such that for all rational numbers $\frac{p}{q}$, $q > 0$ we have that $\left| r - \frac{p}{q} \right| > \frac{m}{q^n}$.

The theorem is very easily proved using only the mean value theorem. (Hardy and Wright, The Theory of Numbers, P 161).[2] This theorem now gives us a criterion for determining certain transcendental numbers. A Liouville number is a number $r \notin Q$ such that for all n there exists a rational number P_n/q_n such that $| r - P_n/q_n | < (\frac{1}{q_n})^n$. A Liouville number is transcendental and examples of such numbers are $\sum_{k=0}^{\infty} \frac{1}{10^{k!}}$, $\sum_{k=0}^{\infty} \frac{(-1)^k}{2^{k!}}$

Liouville published two papers in 1844 and 1851 [3] on the subject and thus launched the study of transcendental numbers.

It should be observed that almost all transcendental numbers are not Liouville numbers. Both π and e are not Liouville numbers with π satisfying $| \pi - P/q| > q^{-42}$ for $q > 2$. [4]

A simple generalization of Liouville's theorem appears in Joseph Lipman's book,"Transcendental Numbers". [5] . It is theorem I.

<u>Theorem I</u>. For any algebraic number r of degree n, there is a positive number $C(r)$ depending on r such that whenever $s \neq r$ is an algebraic number of degree d and height H then

$$|r - s| > \frac{C(r)^d}{H^n}$$

From the aforementioned theorem Lipman is able to deduce theorem II.

<u>Theorem II</u>. Let r be an algebraic number with $|r| > 1$. Then $F(z) = \sum_{k=0}^{\infty} \frac{z^k}{r^{k!}}$ is an entire function which assumes a transcendental value at every non-zero algebraic z .

The preceding nicely suggests the intimate relationship between entire functions and the study of transcendental numbers.

An improvement on Liouville's theorem for the determination of transcendental numbers is the following.

<u>Thue-Siegel-Roth theorem</u>

If r is an algebraic number and $\varepsilon > 0$ then the inequality $| r - \frac{p}{q} | > \frac{1}{q^{2+\varepsilon}}$ is true for all but a finite number of $\frac{p}{q}$.

This theorem allows one to prove that certain numbers defined by a sequence of rational approximations less rapid than that of Liouville numbers are transcendental.

(e.g. $\sum_{k=0}^{\infty} \frac{1}{10^{3^k}}$)

1.3 The theorems of Hermite and Lindemann

Twenty-two years after Liouville's paper was published the next advance in the subject was made. In 1873 Hermite published a proof that e was transcendental. [6] Hermite's proof involved the use of continued fractions. His approach stimulated Lindemann who modified Hermite's techniques and was able to establish Hermite's result as an immediate corollary of his famous theorem. Hermite's proof deserves the recognition it receives because it was the first proof of the transcendence of a naturally occuring number of analysis. Hermite's proof was altered and simplified by later mathematicians like Hilbert and Hurwitz. Elementary proofs can be found in Hardy and Wright "An Introduction to the Theory of Numbers", pp 172-173 [2] and Ivan Niven's "Irrational Numbers", pp 25-26 [7] .

After Hermite's breakthrough there was a 9-year pause before Lindemann gave a proof, in 1882, of the transcendence of π [8] . While Lindemann's objective was a proof of the transcendence of π he proved the following:

Lindemann's theorem B

If $a_1,\ldots,a_n$ are distinct algebraic numbers then $e^{a_1},\ldots,e^{a_n}$ are linearly independent over the rationals.

If π is algebraic then πi is algebraic. πi and $2\pi i$ are distinct yet $e^{\pi i}=-1$ and $e^{2\pi i}=1$ are not linearly independent over the rationals. We conclude that

π is transcendental.

Lindemann's original statement of proof may be reformulated.

Lindemann's Theorem A.

If $a_1,\ldots,a_n$ are algebraic numbers linearly independent over the rationals then $e^{a_1},\ldots,e^{a_n}$ are algebraically independent over the rationals.

We see that the two formulations are equivalent. Assume B, then if $a_1,\ldots,a_n$ are linearly independent over Q and $\sum_r c_r e^{a_1 r_1} e^{a_2 r_2} \ldots e^{a_n r_n} = 0$ where $r = (r_1,\ldots,r_n)$, $c_r \in Q$ and the sum is taken over only a finite number of r then $a_1 r_1 + \ldots + a_n r_n$ are distinct algebraic numbers for different n-tuples by linear independence and thus

$$\sum_r c_r e^{a_1 r_1 + \cdots + a_n r_n} = 0 .$$

This contradicts B.

Now assume A then let $a_1,\ldots,a_n$ distinct algebraic numbers, be written as linear combinations of a set of algebraic numbers $b_1,\ldots,b_t$ which are linearly independent over the rationals, i.e. $a_i = \sum_{k=1}^{t} d_k b_k$

Now $\sum_{r=1}^{n} c_r e^{a_r} = 0$ or $\sum_{r=1}^{n} c_r \prod_{k=1}^{t} e^{d_k b_k} = 0$

which contradicts the algebraic independence of $e^{b_1},\ldots,e^{b_t}$.

A generalized proof of Lindemann's may be found in

Ivan Niven's book "Irrational Numbers" [7] pp.117-130, and an elementary proof of the transcendence of π may be found in Hardy and Wright, "Introduction to the Theory of Numbers" pp.173-176 [2]. Niven proves: Given distinct algebraic numbers $a_1,\ldots,a_m$ the values $e^{a_1}, e^{a_2},\ldots,e^{a_m}$ are linearly independent over the field of algebraic numbers.

Some immediate consequences of the generalized Lindemann theorem are the transcendence of e^α, $\sin\alpha$, $\cos\alpha$, $\tan\alpha$, as well as the hyperbolic functions like sinh with argument α where α is a non-zero algebraic number. In addition, $\log\alpha$, $\arcsin\alpha$ and generally the inverse functions of those above are transcendental when α is algebraic and $\alpha \neq 0, 1$. These conclusions are all possible because of the above functions connection with the function e^z. e.g. Let α be algebraic, $\alpha \neq 0$. We then have $\cosh\alpha = \frac{1}{2}e^\alpha + \frac{1}{2}e^{-\alpha} = a$. If a is algebraic then $\frac{1}{2}e^\alpha + \frac{1}{2}e^{-\alpha} - ae^0 = 0$ contradicting the generalized Lindemann theorem. Therefore a is transcendental.

1.4 Siegel's E-functions

In the years following Lindemann's paper transcendence results were confined to simplifications of the proofs of Lindemann and Hermite. The results depended upon the specific behaviour of the function e^z and could not be generalized to a broader class of functions. Siegel in 1929 did generalize the theory by looking at functions that had an addition theorem

$f(x+y) = f(x) \cdot f(y)$ and satisfied a differential equation similar to $f^{(1)}(x) = f(x)$. He then considered functions satisfying only one of the above conditions. Siegel did considerable work on functions of the latter type, for example the E-functions. Definition $f(z)$ is an E-function if

$$f(z) = \sum_{n=0}^{\infty} c_n \frac{z^n}{n!} \quad \text{with}$$

1) All coefficients c_n belonging to the same algebraic field of finite degree over the rational number field.

2) If ε is a positive number then $||c_n|| = 0\ (n^{n\varepsilon})$ as $n \to \infty$.

3) There exists a sequence $q_o, q_1 \ldots$ of positive rational integers such that $q_n c_k$ is integral for $k = 0,1,\ldots n$ and $n = 0, 1, \ldots$ and that $q_n = 0\ (n^{n\varepsilon})$.

We have that the derivative of an E function is an E-function, that the E-functions form a ring and that e^z, any polynomial and the Bessel function $J_o(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n!)^2} \left(\frac{z}{2}\right)^{2n}$ are E-functions. Siegel proved that $J_o(z)$ and $J_o^{(1)}(z)$ are transcendental and algebraically independent if z is non-zero and algebraic.

1.5 Hilbert's seventh problem

In 1900, 29 years before Siegel began his study of E-functions, Hilbert drew up a list of 23 problems for which there appeared no suitable approach to a solution. Problem seven was to prove that if a and b are non-zero algebraic numbers with $a \neq 1$ then $\frac{\log b}{\log a}$ is either rational or transcendental. This is equivalent to the Gelfond-Schneider theorem stated earlier. No progress was made on the problem until 1929, when a partial solution was given by A.O. Gelfond. Gelfond showed that if a and b are algebraic, $a \neq 0, 1$, with b an imaginary quadratic irrationality then a^b is transcendental. In 1930 Kuzmin extended the proof to the case where b was a real quadratic irrationality. Thus the transcendence of $e^{\pi} = (-1)^{-i}$ and $2^{\sqrt{2}}$ were established by 1930. Finally in 1934 A. Gelfond, [10] and in 1935, Th. Schneider, [11] published independent complete proofs of Hilbert's seventh problem. I will prove the theorem by two methods, one due to Gelfond [12] and the other due to Siegel [13] .

1.6 Siegel's arithmetical lemmas

Before commencing the proofs I will prove two arithmetical lemmas due to Siegel. The first lemma will be used in proving Baker's theorem.

LEMMA 1. If n, m are integers such that $n > m > 0$ and $\sum_{j=1}^{n} a_{ij} x_j = 0 \quad i = 1, \ldots, m$ is a system of linear equations with integer coefficients a_{ij} such that $|a_{ij}| \leq A$, then there exists a non-trivial integral solution $x_1, \ldots x_n$ such that $|x_j| \leq (nA)^{m/(n-m)}$

We will consider the negative and positive a_{ij}'s separately to obtain the required estimate for B.

<u>Pf</u>:

We consider $x = (x_1,\dots,x_n) \neq (0,\dots,0)$ with $0 \leq x_j \leq B$.

Let D_i = sum of the negative a_{ij}'s and E_i = the sum of the positive a_{ij}'s.

Then $-D_i B \leq \sum_j a_{ij} x_j \leq E_i B$ so that $y_i = \sum_{j=1}^{n} a_{ij} x_j$ can assume only $(D_i + E_i) B + 1 \leq nAB + 1$ values and therefore $y = (y_1,\dots,y_m)$ can assume only $(nAB+1)^m$ values. There are $(B + 1)^n - 1$ different arguments $x = (x_1,\dots,x_n)$ that are admissible. Therefore if $(B + 1)^n - 1 \geq (nAB + 1)^m$ we will have either two distinct arguments x' and x'' with the same value under $M(a_{ij})$ or a non-trivial solution. The first case also gives rise to a solution, let $x = x' - x''$ and by linearity we have a non-trivial solution such that $|x_j| \leq B$.

To complete the lemma we check that $B = (nA)^{m/(n-m)}$ satisfies $(B + 1)^{n/m} > nAB+1$ and this is immediate.

<u>LEMMA 2</u>. Let K be a finite algebraic extension of Q of degree h.

If n, m are integers such that $n > m > 0$ and $\sum_{j=1}^{n} a_{ij} x_j = 0 \quad i = i,\dots,m$ is a system of linear equations with coefficients a_{ij} in I_k, the integral closure of Z in K, such that $\| a_{ij} \| < A$, then there exists a non-trivial solution $x = (x_1,\ldots,x_n)$, $x_i \in I_k$, for the system of linear equations such that $\| x_i \| < C\ (C\ nA)^{m/(n-m)}$

where C is a constant depending only on K.

<u>Pf</u>: Let $w_1,\ldots,w_h$ be an integral basis for K. Thus given

$\sum_{j=1}^{n} a_{ij}\, x_j = 0 \quad i = 1,\ldots m$ we may write $x_j = \sum_{k=1}^{h} b_{kj}\, w_k$

where the b_{kj} are rational integers. Therefore we have

$$0 = \sum_{j=1}^{n} \sum_{k=1}^{h} a_{ij}\, b_{kj}\, w_k \quad i = 1,\ldots,m$$

We may now write $a_{ij}\, w_k = \sum_{r=1}^{h} g_{ijkr}\, w_r$ where the g_{ijkr} are rational integers. We are thus trying to solve

$0 = \sum_{j=1}^{n} \sum_{k=1}^{h} \sum_{r=1}^{h} b_{kj}\, g_{ijkr}\, w_r \quad i=1,\ldots,m.$ But as the w_r are linearly independent this reducesto $0 = \sum_{j=1}^{n} \sum_{k=1}^{h} b_{kj}\, g_{ijkr}$

$i = 1,\ldots,m \quad r = 1,\ldots,h$ and by lemma 1 this has a non-trivial solution for $|b_{kj}| \leq (hnB)^{\frac{m}{n-m}}$ where $B = \max_{i,j,k,r} |g_{ijkr}|$.

We observe now that if $a = \sum_{i=1}^{h} g_i w_i$ then we have

$$(g_1,\ldots,g_h) \begin{pmatrix} w_1^{(1)} & \cdots & w_1^{(h)} \\ \vdots & & \vdots \\ w_h^{(1)} & \cdots & w_h^{(h)} \end{pmatrix} = (a,\ a^{(2)},\ldots,a^{(h)})$$

We assume $w_j^{(1)} = w_j$, $j = 1,\ldots,h$

where the superscript denotes a conjugate. The w_i form an integral basis. The determinant of $M(w_j^{(i)})$ is non-zero and so the g_i can be expressed as linear combinations of the $a^{(i)}$. Therefore $g_i \leq C_1 \; ||a||$ where C_1 depends upon K and the choice of basis. Therefore $B \leq C_1 A$ implies $|b_{kj}| \leq (C_o \; nA)^{\frac{m}{n-m}}$ and finally we may conclude that we have non-trivial solutions $X = (x_1, \ldots, x_n)$ such that

$||x_j|| \leq h.\max_k ||w_k||. \quad (C_o nA)^{\frac{m}{n-m}}$. This implies

$||x_j|| \leq C \; (CnA)^{\frac{m}{n-m}}$ where the constant C depends only on the field K and choice of basis.

1.7 The Gelfond-Schneider theorem

Theorem 1. If a is algebraic and $\neq$ 0 or 1, and b is algebraically irrational then a^b is transcendental.

Pf: We will assume that $c = a^b$ is algebraic and we will assume that a, b and c all lie in an algebraic number field K of degree h over Q. Let d be a rational integer such that da, db and dc are algebraic integers in K.

We now construct an entire function f(z) which by judicious use of Lemma 2 will have a large number of zeros and which will allow us to bring into play a, b and c. We revert to the exponential function with its useful addition and differential properties.

We define $f(z)$ as:

$$f(z) = \sum_{j=1}^{N} \sum_{i=1}^{N} r_{ij}\, e^{\log a\cdot(i+jb)z}$$

$$= \sum_{j=1}^{N} \sum_{i=1}^{N} r_{ij}\, a^{iz}\, a^{bjz}$$

We let log a be an arbitrary fixed determination of the logarithm, then $a^z = e^{z \log a}$. The r_{ij} are algebraic integers in K which will be defined in such a way that $f(z)$ has a large number of zeros.

(I) <u>Gelfond's proof</u>:

Gelfond uses lemma 2 to define r_{ij} such that $f^{(k)}(t) = 0\ ,\ 0 \leq k \leq \left[\frac{N^2}{\log N}\right]$ and $t = 0, 1, \ldots, [\frac{1}{2} \log N] = t_1$ with $||r_{ij}|| < d_o^{N^2}$ for some rational integer d_o, $0 \leq i,j \leq N$.

To see how this is done note that

$$f^{(k)}(t) = \sum_{j=1}^{N} \sum_{i=1}^{N} r_{ij}(\log a)^k (i+jb)^k .\, e^{\log a(i+jb)t}$$

Therefore

$$\alpha\ (k,t) = (\log a)^{-k}\ d^{3N^2}\ f^{(k)}(t)$$

$$= \sum_{j=1}^{N} \sum_{i=1}^{N} r_{ij}\ (i+jb)^k\ a^{it}\ c^{jt}\ .\ d^{3N^2}$$

$$= \sum_{j=1}^{N} \sum_{j=1}^{N} r_{ij}\ f_{ij}$$

where the f_{ij} are algebraic integers by definition of d.

By lemma 2 $||r_{ij}|| \leq U(UnA)^{m/n-m}$ where $m = \left[\frac{N^2}{\log N}\right] \cdot [\frac{1}{2}\log N]$,

$n = N^2$ and $A = \max\limits_{i,j} ||f_{ij}||$. But $m \leq \frac{1}{2} N^2$ and $n-m \geq \frac{1}{2} N^2$

therefore $||r_{ij}|| \leq u^2 N^2 A$. We now estimate A.

$$A \leq (d^3)^{N^2} \cdot (N+N||b||)^{\left[\frac{N^2}{\log N}\right]} \cdot ||a||^{N \cdot \frac{1}{2}[\log N]} \cdot ||c||^{[N + \frac{1}{2}[\log N]}$$

$\leq d_o^{N^2}$ for some rational integer d_o. We therefore have defined our entire function $f(z)$ up to our choice of N. The function has a large number of zeros and we will now show that it must have many more zeros at rational integer arguments. We note that the function is not identically zero however as that would imply that $f^{(k)}(z) = 0$ for $0 \leq k \leq N^2$.

Specifically $(\log a)^{-k} f^{(k)}(0) = \sum\limits_{j=1}^{N} \sum\limits_{i=1}^{N} r_{ij} (i+jb)^k = 0$

$0 \leq k \leq N^2$. This implies that matrix M

$$M = \begin{pmatrix} 1 & \cdots & 1 \\ 1+b & \cdots (i+jb)^k \cdots & \vdots \\ \vdots & \vdots & \vdots \\ (1+b)^{N^2-1} & \cdots & (n+nb)^{N^2-1} \end{pmatrix}$$ has determinant 0 .

This is so as the coefficients r_{ij} are not all 0 .

But M is a Vandermonde matrix and therefore

$i_1+j_1 b = i_2+j_2 b$ for $(i_1, j_1) \neq (i_2, j_2)$.

But b is assumed to be irrational so this is impossible. Therefore $f(z)$ is not identically zero.

To prove that a^b can not be algebraic we will show that our function $f(z)$ satisfies $f^{(k)}(z) = 0 , 0 \leq k \leq N^2$ and thus by the above argument we will arrive at a contradiction.

We now show that the number of integral zeros of $f(z)$ and their multiplicity may be increased by a consideration of the following integral form for $f^{(k)}(t)$.

Let C_1 be a circle containing t and C_2 be a circle containing C_1. Then

$$f^{(k)}(t) = \frac{k!}{2\pi i} \int_{C_1} \frac{f(z)}{(z-t)^{k+1}} \, dz$$

$$= \frac{k!}{(2\pi i)^2} \int_{C_1} \frac{dz}{(z-t)^{k+1}} \int_{C_2} \frac{f(y)}{y-z} \, dy$$

by repeated application of Cauchy's integral formula.

As $f(z)$ has zeros of multiplicity $\left[\frac{N^2}{\log N}\right] = u$ at $t = 0,\dots,t_1$ we have that

$$h(y) = \prod_{r=0}^{t_1} \left(\frac{z-r}{y-r}\right)^u f(y) \text{ is still an entire function}$$

of y with $h(z) = f(z)$. Thus we have that

$$\int_{C_2} \frac{h(y)}{y-z} \, dy = \int_{C_2} \frac{f(y)}{y-z} \, dy$$

and it is this fact that enables us to show that our original function must have zeros of multiplicity u for $t=0,\ldots,[\sqrt{N}]$.

Let C_1 be $|z| = N^{3/4}$ and let C_2 be $|y| = N$ then

$$f^{(k)}(t) = \frac{k!}{(2\pi i)^2} \int_{C_1} \frac{dz}{(z-t)^{k+1}} \int_{C_2} \frac{h(y)}{y-z}\, d(y)$$

for $t=0,\ldots,[\sqrt{N}]$ and $k=0,\ldots,u$.

We can now estimate the algebraic integer $\alpha(k,t) = d^{3N^2}(\log a)^{-k} f^{(k)}(t)$

$$\leq d^{3N^2}(\log a)^{-k} \cdot \frac{k!}{2\pi i^2} (N^{3/4} - \sqrt{N})^{-(k+1)} \cdot D \cdot N^{7/4}$$

where

$$D = \left(\frac{N^{3/4} + [\tfrac{1}{2}\log N]}{N - [\tfrac{1}{2}\log N]}\right)^u \cdot \left[\frac{\log N}{2}\right] \max_{|y|=N} |f(y)| \cdot \frac{1}{N - [\tfrac{1}{2}\log N]}$$

$$\leq d_1^{N^2} \cdot \left(\frac{N^{3/4} + [\tfrac{1}{2}\log N]}{N - [\tfrac{1}{2}\log N]}\right)^{\left[\frac{N^2}{\log N}\right][\frac{1}{2}\log N]}$$

$$\leq d_1^{N^2} \cdot N^{-\frac{1}{8}\cdot\frac{N^2}{2}}$$ for large enough N. But we know that the absolute value of the norm $\alpha(k,t) = d^{3N^2}(\log a)^{-k} f^{(k)}(t)$ must be 0 or greater than or equal to 1. We will now show that the norm must be zero.

$$\| d^{3N^2}(\log a)^{-k} f^{(k)}(t) \| \leq \sum_{i=1}^{N}\sum_{j=1}^{N} \| r_{ij}\| \; \|f_{ij}\|$$

$$\leq (N)^2 \cdot d_o^{N^2} \cdot d^{3N^2} (2N)^{\left[\frac{N^2}{\log N}\right]} \|b\|^{\left[\frac{N^2}{\log N}\right]} \cdot \|a\|^{N\sqrt{N}} \cdot \|c\|^{N\sqrt{N}}$$

$$\leq d_2^{N^2}$$ for some constant d_2.

Therefore $|N(\alpha(k,t))| \leq (h-1)\, d_2^{N^2} \cdot d_*^{N^2} \cdot N^{-\frac{1}{8}\frac{N^2}{2}}$

< 1 for N sufficiently large.

The basic ideas of the proof have now been revealed. We will now use the Cauchy integral form for $f^{(k)}(z)$ again, this time to reveal that $f(z)$ must have a zero of multiplicity N^2 at 0 .

$$f^{(k)}(0) = \frac{k!}{(2\pi i)^2} \int_{C_1} \frac{1}{z^{k+1}}\, dz \int_{C_2} \prod_{r=0}^{[\sqrt{N}]} \left(\frac{z-r}{\S-r}\right)^u \frac{f(\S)}{\S-z}\, d\S$$

where $C_1 = |z| = 1$ and $C_2 = |\S| = N$

Thus we have the algebraic integer

$$d^{3N^2}(\log a)^{-k} f^{(k)}(0) \leq d^{3N^2}(\log a)^{-k}\ k! \ .\ 1.\max_{|\S| = N} f(\S) \ .$$

$$.\ \frac{N}{N-1}\ .\ \left(\frac{1+\sqrt{N}}{N-\sqrt{N}}\right)^{\left[\frac{N^2}{\log N}\right]} .\ \sqrt{N}$$

$$\leq d_3^{N^2} \ .\ N^{2\cdot N^2} .\ N^{-\frac{1}{4} N^{\frac{9}{4}}}$$

But we may estimate $||\alpha(k,0)|| = ||\, d^{3N^2}(\log a)^{-k} f^k(0)\,||$ by considering the original definition of $f(z)$. We have

$$||\ \alpha(k,0)\,|| \leq (N+1)^2 \ .\ d^{3N^2}\ (d_4 N)^{N^2}$$

$$< d_5^{N^2} \ .\ N^{N^2}$$

So that

$$|\,N\ (\alpha(k,o))| < (h-1)\, d_5^{N^2}\ N^{N^2} \ .\ d_3^{N^2}\ N^{-\frac{1}{4}N^{\frac{9}{4}}}\ N^{2N^2}$$

and this approaches 0 as $N \to \infty$ so that for N sufficiently large we have that the norm of $\alpha(k,0)$ is less than 1 and hence $f^k(0) = 0$ for $0 \leq k \leq N^2$. But by the argument we have given earlier this is impossible. Therefore a^b is transcendental.

Gelfond gives an elementary proof of the transcendence of a^b for a, b real and algebraic, $a \neq 0, 1$ and b not rational in the book Elementary Methods in Analytic Number Theory. [14] The only analytic tool he brings into play is Rolle's theorem. Using Rolle's theorem we see that we would have the above theorem after the first application of Cauchy's integral form in the preceding proof. By Cauchy we concluded that $f(x)$ had $[\sqrt{N}] \cdot \left[\frac{N^2}{\log N}\right]$ zeros. With the above assumptions we may write $f(x) = \sum_{k=1}^{N^2} \beta_k e^{\alpha_k x}$ where the α_k are distinct and α_k, β_k, x are real for $k=1,\ldots,N^2$. Then $f(x)$ has at most N^2-1 zeros. Pf: By induction on k. True for $k = 1$. Assume true for $k < n$. Then if $\sum_{k=0}^{n} \beta_k e^{\alpha_k x}$ has $d > n-1$ zeroes then we consider $\frac{d}{dx} e^{-\alpha_o k} f(x) = \sum_{k=0}^{n-1} \gamma_k e^{(\alpha_k - \alpha_o)x}$.

By Rolle's this has $> n-2$ zeroes contradicting our induction hypothesis. Therefore when N is sufficiently large N^2-1 is less than $[\sqrt{N}] \cdot \left[\frac{N^2}{\log N}\right]$ and we have a contradiction. This proves the theorem for a, b real.

Siegel adapted Gelfond's method and was able to give a shorter proof. Siegel uses the same function as Gelfond but he needs to use Cauchy's integral form only once to obtain his solution. Siegel does not prove that $f(z)$ has a zero of high multiplicity at 0. He instead deduces a contradiction by considering the norm of a number obtained from the first non-zero valued derivative of $f(z)$ at an integer that is a zero of high multiplicity.

II Siegel's proof.

We will retain all the assumptions for the theorem made before the Roman numeral I indicating the start of what was specifically Gelfond's proof. We shall again use lemma 2 to determine r_{ij} such that $f^{(k)}(t) = 0$ this time for $k=0,\ldots,\frac{N^2}{4h+4}-1$ and $t=0,\ldots,2h+1$ We first define $m=2h+2$ and $n=\frac{N^2}{2m}$ where we assume that N^2 is an integral multiple of $2m$ and that $N > m$. Then we have

$$d^{n-1+2mN}(\log a)^{-k} f^{(k)}(t) = \sum_{j=1}^{N}\sum_{i=1}^{N} r_{ij}\,(i+jb)^k\, a^{(i+jb)t} \cdot d^{n-1+2mN}$$

$$= \sum_{j=1}^{N}\sum_{i=1}^{N} r_{ij}\, f_{ij} \quad \text{where the } f_{ij} \text{ are}$$

algebraic integers by definition of d.

By lemma 2 we can find r_{ij} such that $f_s^{(k)}(t) = 0$ $k=0,\ldots n-1$ and $t=0,\ldots m-1$ with $||r_{ij}|| \le V(VuA)^{\frac{s}{u-s}}$ where $u=N^2$, $s=mn$. Therefore $||r_{ij}|| \le V^2N^2A$ where $A = \max_{i,j} ||f_{ij}||$ as before and $A \leq N^2 d^{n-1+2mN} (N+N||b||)^n . ||a||^{mN} . ||c||^{mN}$

$$\leq d_1^{\,n} . \sqrt{n}^{\,n} = d_1^{\,n}\ n^{\frac{n}{2}}$$

Therefore $||r_{ij}|| \leq d_2^{\,n}\ n^{\frac{n}{2}}$ and $f(z)$ is not identically zero by the same argument we gave in the proof after Gelfond.

We now choose the integer p such that $f^{(p)}(t_o) \neq 0$ for some $t_o : 0 \leq t_o \leq m-1$ where $f^{(k)}(t) = 0$ for all $k<p$ and all $t : 0 \leq t \leq m-1$. We consider the number $\alpha = (\log a)^{-p} f^{(p)}(t_o)$. We know that $d_1^{\,p+2mN}\alpha$ is an algebraic integer and this gives us that $|N(\alpha)| > d_1^{\,-p}$. Also $||\alpha|| \leq N^2 d^{p-1+2mN} (N+N||b||)^p$ $||a||^{mN}\ ||c||^{mN} .\ d_2^{\,n}\ n^{n/2}$

$$\leq d_3^{\,p} (\sqrt{n})^p\ n^{\frac{n}{2}}$$

$$\leq d_3^{\,p}\ p^p$$

As in the previous proof we will make use of the fact that $|\alpha| \cdot ||\alpha||^{h-1} \geq |N(\alpha)|$.

This time we will show $|\alpha|$ is so small that a contradiction must arise. To estimate $|\alpha|$ we consider the entire function $h(z)$. $h(z) = \dfrac{p!\ f(z)}{(z-t_o)^p} \prod\limits_{\substack{r=0 \\ r \neq t_o}}^{m-1} \left(\dfrac{t_o - r}{z-r}\right)^p$ and $\alpha = (\log a)^{-p}\, h(t_o)$.

NB. $f(z) = \dfrac{(z-t_o)^p f^p(t_o)}{p!} + \{$ terms in $(z-t_o)^{p+i}$ $i > 0$ $\}$

by expanding f in a Taylor series.

Now we use Cauchy's integral formula to give us

$$h(t_o) = \frac{1}{2\pi i} \int_{|z| = \frac{p}{N}} \frac{h(z)}{z-t_o}\, dz \quad \text{where } |z| = \frac{p}{N} \text{ encloses } 0,\ldots,m-1.$$

Remembering that $|e^z| \le e^{|z|}$ we estimate $f(z)$ on $|z| = \frac{p}{N}$.

$$f(z) \le N^2 \cdot d_2^n \, n^{\frac{n}{2}} \cdot d_4{}^p \le p^{\frac{p}{2}} d_5{}^p$$

$$|z-r| \ge \frac{p}{2N} \Rightarrow |z-r|^{-p} \le \left(\frac{2N}{p}\right)^p$$

We therefore have that

$$\max_{|z| = \frac{p}{N}} |h(z)| \le p!\, p^{\frac{p}{2}} d_5{}^p \left(\frac{2N}{p}\right)^{p.m} \cdot (m-1)^{pm}$$

$$\le d_6^p \, p^{\frac{3}{2}p} \cdot \left(\frac{\sqrt{n}}{p}\right)^{mp}$$

$$\le d_6{}^p \, p^{\left(\frac{3-m}{2}\right)p}$$

and this implies that $|\alpha| \le d_7^p \, p^{\left(\frac{3-m}{2}\right)p}$

This is all we need. We now have $|N(\alpha)| \le p^{\left(\frac{3-m}{2}+h-1\right)p} \cdot \left(d_3{}^p p^p\right)^{h-1}$

$$|N(\alpha)| \le d_8{}^p \cdot p^{\left(\frac{3-m}{2}+h-1\right)p}$$

$$\le d_8{}^p \cdot p^{-\frac{1}{2}p}$$

and therefore as $|N(\alpha)| > d_1{}^{-p}$ we have that $d_9{}^p > p^{\frac{1}{2}p}$ and so $d_9 > p^{\frac{1}{2}}$. But $p > n$ and d_9 is independent of n and p. As we may choose n arbitrarily large we have a contradiction and the theorem is proved.

1.8 Schneider's theorem

We thus have seen two methods of proof for Hilbert's seventh problem. An english translation of Schneider's proof may be found in Siegel's monograph "Transcendental Numbers". Schneider's proof depends on the construction of an entire function out of the sum of products of polynomials and powers of a^x. This falls nicely into the scheme of Siegel's monograph.

The proofs of the Gelfond-Schneider theorem all follow a general pattern. By skillful use of Siegel's arithmetic lemma we can force a specially chosen entire function to have a large number of zeros. We can then choose a non-zero algebraic number defined from our entire function. We then use the Cauchy integral formula to determine the absolute value of our number. By consideration of the function we can determine the maximum of the absolute value of the conjugates of our number. Finally we use the fact that the norm of a non-zero algebraic number is greater than or equal to 1. By suitable comparison of these quantities we may establish our theorem.

Schneider was able to generalize this process and thus make the next large step forward in the field. He accomplished this with his paper in 1949. [15] In the years 1934-1949 Schneider had proved a number of interesting and important results on elliptic functions, periodic functions and elliptic integrals. His theorem of 1949 managed to consolidate these results, the Gelfond-Schneider theorem and the transcendence of e and π.

Schneider's proof may be found in English in the monograph by Lipman. [5] Before commencing the proof I will define the order of an entire and a meromorphic function.

<u>Definition</u> Let $f(z)$ be an entire function, then $f(z)$ is of order $\leq u$ if there exists a constant $c > 1$ such that for all large R

$$|f(z)| \leq c^{R^u} \quad \text{for} \quad |z| \leq R .$$

<u>Defintion</u> A meromorphic function is of order $\leq u$ if it is a quotient of entire functions of order $\leq u$.

<u>Schneider's theorem</u> 2 Given $f_1(z)$, $f_2(z)$ two meromorphic functions of order $\leq u$ and given distinct numbers $z_1,\ldots,z_m$ which are not poles of $f_1(z)$ or $f_2(z)$ and if the following conditions are satisfied:

(1) All $f_x^{(n)}(z_k)$ $(k=1,\ldots,m$, $x=1,2$ and $n= 0,1,\ldots)$ are algebraic and lie in a field K of degree s over the rationals.

There are natural numbers b, c and a constant $v > 0$ such that

(2) $b^{n+1} f_x^{(n)}(z_k)$ is an algebraic integer.

$(x=1,\ 2,\ \ k=1,\ldots,m\ ,\ n=0,\ 1,\ 2\ \ldots\ )$

(3) $||f_x^{(n)}(z_k)|| < c^{n+1} n^{vn}$

then if $m > 2u\ (\ (s-1)\ (2v+1) + v+\frac{3}{2}\ )$ the functions $f_1(z)$ and $f_2(z)$ are algebraically dependent.

Pf: We will assume the hypotheses of the theorem and that $f_1(z)$ and $f_2(z)$ are algebraically independent over Q.

We now define the meromorphic function $G(z)$.

$$G(z) = \sum_{i=o}^{r} \sum_{j=o}^{r} a_{ij} f_1^{i}(z) f_2^{j}(z) \qquad r = \sqrt{2mt}$$

$t>2m$ is picked so that $\sqrt{2mt}$ is an integer.

As before we choose the coefficients a_{ij} so that $G(z)$ has zeros of high multiplicity at $z_1, \ldots z_m$. But here our derivatives are more complicated than in the case of the derivatives of the function defined in the Gelfond-Schneider theorem.

$$G^{(n)}(z) = \sum_{i=0}^{r} \sum_{j=0}^{r} a_{ij} \left(f_1^i(z) f_2^j(z) \right)^{(n)}$$

$$\left(f_1^i(z) f_2^{j}(z)\right)^{(n)} = \sum_{e=0}^{n} \binom{n}{e} \left(f_1^i(z)\right)^{(e)} \left(f_2^j(z)\right)^{(n-e)}$$

and finally

$$\left(f_1^{i}(z)\right)^{(e)} = \Sigma f_1^{(e_1)}(z_1) f_1^{(e_2)}(z) \ldots f_1^{(e_i)}(z) \cdot c_e, c_e \text{ a constant,}$$

where $e_1 + e_2 + \ldots + e_i = e$ with $e_k \geq 0, k = 1, \ldots i$.

We may consider the conditions $G^{(n)}(z_k) = 0 \quad \substack{k=1,\ldots,m \\ n=0,\ldots,t-1}$ as a system of mt linear equations with algebraic coefficients in $(r+1)^2$ unknowns a_{ij}. To employ lemma 2 we must have our coefficients algebraic integers. By hypothesis $b^{n+1} f_x^{(n)}(z_k)$ is an algebraic integer ($x=1, 2, k=1, \ldots, m, n=0, 1, 2 \ldots$)

Thus $b^{n+2r} G^{(n)}(z_k) = 0 \quad \substack{k=1,\ldots,m \\ n=0,\ldots,t-1}$ is a system of mt linear equations with algebraic integer coefficients in $(r+1)^2$ unknowns a_{ij}.

By lemma 2 we may choose our unknowns a_{ij} such that

$$||a_{ij}|| \leq C\,(C\,(r+1)^2\,A)^{\frac{mt}{(r+1)^2 - mt}}$$

$$\leq C^2\,(r+1)^2\,A$$ where C is a constant

depending only on K.

$$A = \max_{n,i,j,k} ||\,(f_1^i\;f_2^j)^{(n)}\,(z_k)\,||\;\; b^{n+2r}$$

We now determine A.

$$A < b^{t+2r} \,.\, (2r)^t \,.\, c^{t+2r}.$$

$$.\; n_1^{vn_1} \,.\, n_2^{vn_2} \ldots n_k^{vn_k}$$

where $\sum_{i=1}^{k} n_i \leq t$

by the hypothesis that $||f_x^{(n)}(z_k)|| < c^{n+1}\, n^{vn}$

$x=1,2$, $k=1,\ldots,m$ $n=0, 1, 2,\ldots,$.

Therefore

$$A < b^{t+2r} \,.\, (2r)^t \,.\, c^{t+2r} \,.\, t^{vt}$$

$$< d_o^{\,t}\; t^{vt} \,.\, (2r)^t$$

$$< d_1^{\,t}\; t^{(\frac{1}{2}+v)t}$$

This finally shows us that we may choose our a_{ij} as algebraic integers such that $||a_{ij}|| \leq d_2^{\,t}\; t^{(\frac{1}{2}+v)t}$ and such that $G^{(n)}(z_k) \neq 0$ $k = 1,\ldots,m$, $n=0,\ldots,t-1$.

With this knowledge we may now precede as we did in the proof of the Gelfond-Schneider theorem .

We consider the first non-vanishing derivative say the p^{th}, of $G(z)$ for $z=z_k$ $k \in \{1,\ldots m\}$ such that $G^{(n)}(z_k) = 0$ for $0 \leq n < p$ and $k=1,\ldots,m$. We observe that $p \geq t$. We may assume without loss of generality that $G^{(p)}(z_1) \neq 0$

We may assume that p exists because we have assumed that $f_1(z)$ and $f_2(z)$ are algebraically independent and thus that $G(z)$ is not identically zero.

We will now estimate $||G^{(p)}(z_1)||$, $|N(G^{(p)}(z_1))|$ and $||G^{(p)}(z_1)||$ to establish a bound on m.

First $||G^{(p)}(z_1)|| \leq (r+1)^2 \cdot d_2{}^t \, t^{(\frac{1}{2}+v)t} \cdot$

$$\max_{i,j,k} ||(f_1{}^i f_2{}^j)^{(p)}(z_k)||$$

and $\max\limits_{i,j,k} ||(f_1{}^i f_2{}^j)^{(p)}(z_k)|| \leq c^{2r+p} \, p^{vp} \quad (2r)^p$.

$$\leq d_3^p \cdot p^{(\frac{1}{2}+v)p}$$

as $r < kt^{\frac{1}{2}} < k\, p^{\frac{1}{2}}$.

Thus

$$||G^{(p)}(z_1)|| \leq d_4{}^p \cdot t^{(\frac{1}{2}+v)t} \cdot p^{(\frac{1}{2}+v)t}$$

$$\leq d_4{}^p \cdot p^{(1+2v)p}$$

We easily get an estimate for $|N(G^{(p)}(z_1))|$ as we have that b^{2r+p}. $G^{(p)}(z_1)$ is an algebraic integer with norm $\neq 0$.

Thus we have $|N(G^{(p)}(z_1))| > \dfrac{1}{b^{(2r+p)s}}$

We now estimate $|G^{(p)}(z_1)|$. We wish to use Cauchy's integral form so that we must convert $G(z)$ to an entire function. As f_1 and f_2 are meromorphic of order $\leq u$ there is an entire function $h(z)$ of order $\leq u$ such that hf_1 and hf_2 are entire of order $\leq u$. We may assume that $h(z_1) \neq 0$. Thus $h^{2r}(z)\, G(z)$ is an entire function and we have that $H(z) = h^{2r}(z)\, G(z) \cdot \prod_{k=1}^{m} (z-z_k)^{-p}$ is an entire function also. By expanding $G(z)$ in a Taylor series in powers of $(z-z_1)$ we are able to conclude that

$$G^{(p)}(z_1) = p! \frac{H(z_1) \cdot \prod_{k=2}^{m} (z_1-z_k)^p}{h^{2r}(z_1)} .$$

and therefore $|G^{(p)}(z_1)| \leq |H(z_1)| \cdot p^p \cdot d_5^{p(m-1)} d_6^{2r}$

$\leq |H(z_1)|\, p^p \cdot d_7^p$ for some

constants d_5, d_6, d_7.

We now estimate $|H(z_1)|$ by means of the Cauchy integral formula.

$$H(z_1) = \frac{1}{2\pi i} \int_C \frac{H(z)}{z-z_1}\, dz$$

where C is the circle $|z| = R$, $R \geq 2 \max_k |z_k|$.

Then

$$|H(z_1)| \leq R \cdot \frac{1}{2R} \cdot \max_{|z|=R} \left| h^{2r}(z) . G(z) \cdot \prod_{k=1}^{m} (z-z_k)^{-p} \right|$$

$$\leq \tfrac{1}{2} \cdot d_8^{2rR^u} \cdot \max_{|z|=R} |G(z)| \cdot (2R)^{-pm}$$

and $\max_{|z|=R} | G(z) | < \max_{i,j} ||a_{ij}|| \cdot (r+1)^2 \cdot d_9^{rR^u} \cdot d_{10}^{rR^u}$

$$\leq d_2^{\,t} \cdot t^{(\frac{1}{2}+v)t} \cdot d_{11}^{rR^u}$$

We used in the above the fact that h^{2r}, $f_1^{\,r}$ and $f_2^{\,r}$ had order u. We can therefore conclude that

$$|G^{(p)}(z_1)| \leq d_8^{2rR^u} \cdot R^{-pm} \cdot d_2^{\,t}\, t^{(\frac{1}{2}+v)t} \cdot d_{11}^{rR^u} \cdot p^p \cdot d_7^{\,p}$$

$$\leq p^{(\frac{3}{2}+v)p} \cdot d_{12}^{\,p} \cdot (d_{13}^{\,r})^{R^u} \cdot R^{-pm}$$

We now let $R = r^{\frac{1}{u}}$. This is valid if we assume t is large enough

Then $|G^{(p)}(z_1)| \leq p^{(\frac{3}{2}+v)p} \cdot d_{12}^{\,p} \cdot d_{13}^{\,r^2} \cdot r^{-\frac{pm}{u}}$

$$\leq p^{(\frac{3}{2}+v)p} \cdot d_{14}^{\,p} \cdot p^{-\frac{mp}{2u}}$$

We now have an estimate for $|N(G^{(p)}(z_1))| \leq ||G^{(p)}(z_1)||^{s-1} \cdot |G^{(p)}(z_1)|$

$$\leq d_4^{\,p(s-1)}\, p^{(1+2v)\,p(s-1)} \cdot p^{(\frac{3}{2}+v)p} \cdot d_{14}^{\,p} \cdot p^{\frac{-mp}{2u}}$$

Comparing this to $|N(G^{(p)}(z_1))| \geq (b^{-1})^{(2r+p)s}$ and letting t, and thus p, approach ∞ we may conclude that

$$(1+2v)(s-1) + (\frac{3}{2}+v) - \frac{m}{2u} \geq 0$$

or in other words

$$m \leq 2u\,((1+2v)\cdot(s-1) + (\frac{3}{2}+v))$$

This concludes our theorem.

1.9 Corollaries of Schneider's theorem

We will now derive some of the results mentioned earlier as corollaries to this theorem.

Corollary 1.

If α is algebraic $\neq 0$, e^{α} is transcendental.

Pf: Given an algebraic number α assume that e^{α} is algebraic. This means that α and e^{α} lie in same field K of degree s over the rationals. We use the additive property of e^{z} and z and the differential property of e^{z} to state that $f_1^{(n)}(k\alpha)$, $f_2^{(n)}(k\alpha)$ belong to K for $k=1,\ldots$ and $n=0, 1,\ldots$ if we let $f_1(z)=z$, $f_2(z)=e^{z}$. f_1 and f_2 are meromorphic functions of order 0 and 1 respectively. z and e^{z} are algebraically independent. To see this we assume

$f(z)=\sum_{k=0}^{m} P_k(z)\, e^{kz}=0$ for some non-zero polynomials $P_k(z)$ with rational coefficients. Let the maximum degree of the $P_k(z)$ be s. We then consider $f^{(s)}(z)$. It has the form $f^{(s)}(z)=\sum_{k=1}^{m} P_k^{1}(z)\, e^{kz}=0$. Dividing by e^{z} we may write it in the form $\sum_{k=0}^{r-1} P_{k+1}^{1}(z)\, e^{kz}=0$. Where the $P_{k+1}^{1}(z)$ are non-zero polynomials with rational coefficients. We can continue this process until we have the absurd statement that a non-zero polynomial is zero.

We have that $b^{k} f_2^{(n)}(k\alpha)$ is an algebraic integer if be^{α} is an algebraic integer. b may be assumed rational. We also have that $||f_x^{(n)}(k\alpha)|| < (c^{k}) \cdot n^{0}$ for some rational integer c.

The $k\alpha$, $k = 1,2,\ldots$ are distinct as $\alpha\neq 0$.

By Theorem 2 we may only have the above conditions hold for $k \leq 2(s+1)$. However we can clearly let k be arbitrarily large. This contradicts Theorem 2 and thus e^{α} can not be algebraic for α algebraic. This allows us to conclude that both e and π are transcendental numbers.

Corollary 2. The Gelfond-Schneider Theorem.

If α is algebraic β algebraically irrational then α^{β} is transcendental. $(\alpha \neq 0,1)$

Pf: We consider the functions $f_1(z) = e^z$ and $f_2(z)=e^{\beta z}$.

These functions are algebraically independent as β is assumed algebraically irrational. To prove the algebraic independence of e^z and $e^{\beta z}$ we argue as we did in corollary 1. We show the algebraic independence of z and z^{β} and then specialize to e^z and $e^{\beta z}$. Both $f_1(z)$ and $f_2(z)$ are entire functions of order ≤ 1.

Now let α be an algebraic number, not 0 or 1, and consider $\log \alpha$ for some fixed determination of the logarithm ($\alpha^{\beta}=e^{\beta \log \alpha}$ for our determination of $\log \alpha$.) We assume $Q(\alpha,\beta,\alpha^{\beta}) = K$ is a field of finite degree s over the rationals.

We let $z_k=k \log \alpha, k=1,\ldots,m$ and note that the z_k are distinct.

(1) We have that $f_1^{(n)}(z)=f_1(z)=e^z$ and $f_2^{(n)}(z)=\beta^n e^{\beta z}$ and thus all $f_x^{(n)}(z_k)$ ($x=1, 2, k=1,\ldots,m$ and $n=0,1,2,\ldots$) are algebraic and lie in K.

(2) Let $b = \max (a_1, a_2, a_3)$ where a_1, a_2, a_3 are leading coefficients in the minimal defining polynomial of α, β and α^β respectively. We then have that b^m is a natural number such that $(b^m)^{n+1} \; f_x^{(n)}(z_k)$ is an algebraic integer

$(x=1, 2, k=1,\ldots,m, n=0, 1,2 \ldots)$

(3) Let $c = \max (|| \beta || , || \alpha ||^m, || \alpha^\beta ||^m)$ then

$$|| f_x^{(n)} (z_k) || < (c+1)^{n+1} . n^{0.n}$$

By Schneider's theorem we have that $m \leq 2s+1$. m was arbitrary, however, and we thus have a contradiction.

Therefore α^β must be transcendental.

I will now derive some results concerning the transcendence of values of the Weierstrass $\wp$ function from an alternate form of Schneider's theorem. Serge Lang proved the following theorem [16] in 1962.

<u>Theorem 2[1]</u>. Let K be a finite extension of the rational numbers. Let $f_1,\ldots f_N$ be meromorphic functions of order $\leq p$. Assume that the field $K (f_1,\ldots f_N)$ has transcendence degree ≥ 2 over K, and that the derivative $D = \frac{d}{dz}$ maps the ring $K [f_1,\ldots,f_N]$ into itself. Let $w_1,\ldots w_n$ be distinct complex numbers not lying among the poles of the f_i, such that $f_i(w_v) \varepsilon K$ for all $i=1,\ldots,N$ and $v=1,\ldots,m$. Then $m \leq 10p\,[K:Q]$.

Theorem 2[1] may be derived from Schneider's theorem quite easily. It is a somewhat more tractable form of the theorem.

The Weierstrass ρ function is defined as

$\rho(z) = \frac{1}{z^2} + \sum_{\omega\in\Omega-\{0\}} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}\right)$ where Ω is a discrete subgroup of the complex numbers with a base ω_1, ω_2 of vectors where $\mathrm{Im}\left(\frac{\omega_1}{\omega_2}\right) > 0$. The ρ-function is doubly periodic and meromorphic and in fact every Ω periodic function is of the form $F(\rho(z)) + G(\rho(z))\,\rho^{(1)}(z)$ where F and G are rational.

We have the following relation between $\rho^{(1)}(z)$ and $\rho(z)$:

$$(\rho^{(1)}(z))^2 = 4(\rho(z))^3 - 60a_2\,\rho(z) - 140a_4$$

where $a_2 = \sum_{\substack{\omega\in\Omega\\ \omega\neq 0}} \frac{1}{\omega^4}$ $\qquad a_4 = \sum_{\substack{\omega\in\Omega\\ \omega\neq 0}} \frac{1}{\omega^6}$

We also have $\rho^{(2)}(z) = 6(\rho(z))^2 - 30a_2$

and $\rho(z_1+z_2) = -\rho(z_1) - \rho(z_2) + \frac{1}{4}\left(\frac{\rho^{(1)}(z_1) - \rho^{(1)}(z_2)}{\rho(z_1) - \rho(z_2)}\right)^2$

therefore $\rho(2z) = -2\rho(z) + \frac{1}{4}\left(\frac{\rho^{(2)}(z)}{\rho^{(1)}(z)}\right)^2$

Corollary 3:

If b, a_2 and a_4 are algebraic and b is not a pole of $\rho(z)$ then $\rho(b)$ is transcendental.

Pf: We assume that b is not a pole of $\rho(z)$ and therefore also not a pole of $\rho^{(1)}(z) = \sum_{\omega \varepsilon \Omega} \frac{1}{(z-\omega)^3}$

We assume that $b, a_1, a_2, \rho(b)$ and $\rho^{(1)}(b)$ are algebraic. Therefore $K = Q\ (a_1, a_2, b, \rho(b), \rho^{(1)}(b))$ is a finite extension of Q. We have that $\rho(z)$ and z are algebraically independent over K and thus that $K\ (z, \rho(z), \rho^{(1)}(z)\ )$ has transcendence degree ≥ 2 over K.

As $\rho^{(2)}(z) = 6(\rho(z))^2 - 30a_2$ we have that the derivative D maps the ring $K[z, \rho(z), \rho^{(1)}(z)]$ into itself.

We may write $\rho(z) = -\left(\frac{\sigma^{(1)}(z)}{\sigma(z)}\right)^{(1)}$ where

$\sigma(z) = z \cdot \prod_{\omega \varepsilon \Omega - (0,0)} (1 - \frac{z}{w})\, e^{\frac{z}{w} + \frac{1}{2}(\frac{z}{w})^2}$ [17] from which we may conclude that $\rho(z)$ and $\rho^{(1)}(z)$ are meromorphic of order 2.

We now use the addition formula for $\rho(z)$ to reach a contradiction to theorem 2^1.

$$\#\quad \rho\ (2b) = -2\rho(b) + \tfrac{1}{4}\left(\frac{\rho^{(2)}(b)}{\rho^{(1)}(b)}\right)^2$$

$$= -2\rho(b) + \tfrac{1}{4}\left(\frac{6(\rho(b))^2 - 30a_2}{\rho^{(1)}(b)}\right)^2, \text{let } \psi = \tfrac{1}{2}(\rho(2b) + 2\rho(b))^{\frac{1}{2}}.$$

Therefore $\rho(2b)\ \varepsilon\ K$ and in fact $\rho(2^n b)\ \varepsilon\ K, n \varepsilon Z^+$.

We have an addition formula for $\rho^{(1)}(z)$ also, merely take derivative of $\#$.

$$2\rho^{(1)}(2z) = -2\rho^{(1)}(z) + \psi\left(\frac{\rho^{(1)}(z)^2 . 12\rho(z) - (6\rho(z)^2 - 30a_2)^2}{(\rho^{(1)}(z))^2}\right)$$

We thus have that $\rho^{(1)}(2^n b)\ \varepsilon\ K$ for $n\ \varepsilon\ Z^+$. If $2^n b$ is a pole for $n = n_o$ then for all $n \geq n_o, 2^n b$ is a pole. But then $2^n b+1$ is not a pole so by using the addition formula for $\rho(z_1+z_2)$ and $\rho^{(1)}(z_1+z_2)$ we find that $\rho^{(1)}(2^n b+1)$ and $\rho(2^n b+1)$ are elements of K for $n \in Z^+$.

From Theorem 2[1] we have $n \leq 10 \cdot 2\,[K : Q]$. Therefore either $\rho(b)$ or $\rho^{(1)}(b)$ is transcendental. But we have that $(\rho^{(1)}(z))^2 = 4(\rho(z))^3 - 60a_2\,\rho(z) - 140a_4$. As we assumed that a_2 and a_4 were algebraic we have that both $\rho(b)$ and $\rho^{(1)}(b)$ are transcendental.

I will state without proof 3 more corollaries of Schneider's theorem. [17]

<u>Corollary 4</u>. The 5 numbers $a_2, a_4, b, \rho(c), e^{cb}$ are not all algebraic.

<u>Corollary 5</u>. If τ is algebraic and not imaginary quadratic then $J(\tau)$ is transcendental. If τ is imaginary quadratic then $J(\tau)$ is algebraic.

J is the elliptic modular function.

<u>Corollary 6</u>. At least one of the seven numbers

$a_2, a_4, a_2^*, a_4^*, b, \rho(c), \rho^*(bc)$ is transcendental

if c and bc are not poles of ρ and ρ^* and if $\rho(z)$ and $\rho^*(bz)$ are algebraically independent.

1.10 Transcendence measures

Schneider's theorem was the last major result in the field prior to A. Baker's paper. Since 1950 the subject has been generalized and problems in the field have now been formulated using algebraic varieties and p-adic numbers. A. Brumer proved the p-adic generalization of Baker's theorem. [18] A great deal of work has been done on the measure of transcendence of a number by N. Feldman, A. Gelfond, K. Mahler, Morduhai-Boltovski, Siegel and others.

<u>Definition</u> A transcendence measure of a number a is a function $g(x,y)$ such that $g(H,d) \leq |P(a)|$ where P is a non-zero polynomial with integral coefficients of height $\leq H$ and degree $\leq d$.

<u>Definition</u> A measure of mutual transcendence of the numbers $a_1,\ldots,a_n$ is a function $g(x, y_1,\ldots,y_n)$ such that

$$g(H, d_1,\ldots,d_n) \leq | P(a_1,\ldots,a_n) | \text{ where}$$

$P(x_1,\ldots,x_n)$ is a non-zero polynomial in n indeterminates with integral coefficients of height $\leq H$ and degree $\leq d_i$ $i=1,\ldots,n$ in the x_i, $i=1,\ldots,n$ respectively. The measure of transcendence of a number gives us a quantitative criterion for the transcendence of that number. A. Baker's theorem gives us a modified mutual measure of transcendence for the numbers $\log a_1,\ldots,\log a_n$ where $a_1,\ldots a_n$ are algebraic numbers not equal to 0 or 1 and $\log a_1,\ldots,\log a_n$ are linearly independent over the rationals. It is "modified" as we only consider the polynomials $P(x_1,\ldots x_n) = b_1 x_1 + b_2 x_2 + \ldots + b_n x_n$ where the b_i, $i=1,..n$ are algebraic numbers, not all 0, with degrees at most d and height at most H. If A. Baker's theorem gave an unmodified transcendence measure we would have the algebraic independence of the logarithms of multiplicatively independent algebraic numbers.

The theory of transcendental numbers is in its infancy despite its chronological age of over 125 years. In reviewing those years I have discussed what I felt were the most important results. In the next chapter I intend to prove A. Baker's result. His theorem gives us the most recent important advance in the study of transcendental numbers. A. Baker's theorem, however, is applicable to problems outside the realm of transcendental numbers. I will discuss this in Chapter III.

CHAPTER II

A. BAKER'S THEOREM

"LINEAR FORMS IN THE LOGARITHMS OF ALGEBRAIC NUMBERS"

2.1 Baker's theorem

In this chapter I will prove A. Baker's theorem. I will follow his papers "Linear Forms in the Logarithms of Algebraic Numbers" (I) and (II). [19] I have mentioned Baker's theorem in the context of the study of transcendental numbers up to this point. It should be observed however, that his theorem on transcendental numbers is derived from the following theorem.

Theorem 3. If $a_1,\ldots,a_n$ are non-zero algebraic numbers such that $\log a_1,\ldots,\log a_n$ $(n \geq 2)$ are linearly independent over Q, and if $q > 2n+1$ and d is any positive integer, we have that there is an effectively computable number $C = C(n, a_1,\ldots a_n, q, d) > 0$ such that for all algebraic numbers $b_1,\ldots,b_n$ not all 0 with degrees at most d, we have $|b_1 \log a_1 + \ldots + b_n \log a_n| > Ce^{-(\log H)^q}$ where H denotes the maximum of the heights of $b_1,\ldots,b_n$.

Some of the consequences of this theorem will be discussed in Chapter III. In this chapter I am only interested in the extension of the Gelfond-Schneider theorem which may be obtained with Theorem 3. I will show this now.

2.2 The extension of the Gelfond-Schneider theorem

We have the following theorem as a weakened form of theorem 3.

Theorem 4. If $a_1,\ldots,a_n$ are non-zero algebraic numbers such that $\log a_1, \ldots, \log a_n$ are linearly independent over the rationals then they are linearly independent over the field of all algebraic numbers.

From theorem 4 one may now derive the following extension of the Gelfond-Schneider theorem.

Theorem 5. If $a_1, \ldots, a_n$ denote algebraic numbers other than 0 or 1 and if $b_1, \ldots, b_n$ denote algebraic numbers with $1, b_1, \ldots, b_n$ linearly independent over the rationals then $a_1^{b_1} \cdot a_2^{b_2} \cdot \cdot\cdot a_n^{b_n}$ is transcendental.

Pf: The proof will be done by induction. For $n = 1$ this is merely the Gelfond-Schneider theorem.
Assume the theorem is true for $n = k-1$
We will now assume it is not true for $n = k$ and reach a contradiction. Throughout the argument we will take suitable determinations of elements of the form a^b if z^b is a multivalued function (i.e. $a^b = e^{b \log a}$ for some determination of log a.)

Assume $a_1^{b_1} \ldots a_k^{b_k} = a_{k+1}$ where a_{k+1} is an algebraic number.

We then have $a_1^{b_1} \ldots a_k^{b_k} a_{k+1}^{b_{k+1}} = 1$ where $b_{k+1} = -1$.

Note that $b_1, \ldots, b_{k+1}$ are linearly independent over Q by assumption.

We now use ~~the converse of~~ theorem 4 to assert that $a_1^{c_1} \ldots a_k^{c_k} a_{k+1}^{c_{k+1}} = 1$ for rational numbers c_i, $i=1, \ldots, k+1$

with not all the $c_i = 0$. We may assume that $c_{k+1} \neq 0$ by permutation of the subscripts if necessary.

We then define $d_i = c_{k+1} b_i + b_{k+1} c_i$, $i=1,\dots,k$.

We have that the d_i are linearly independent over Q. This is because $b_1,\dots b_{k+1}$ are linearly independent over Q.

We also have

$$1 = (a^{b_1} \dots a_{k+1}^{b_{k+1}})^{c_{k+1}} \cdot (a_1^{c_1} \dots a_{k+1}^{c_{k+1}})^{-b_{k+1}} = a_1^{d_1} \dots a_k^{d_k}$$

Finally let $e_i = \dfrac{-d_i}{d_k}$, $i=1,\dots,k-1$

We have that $1, e_1,\dots e_{k-1}$ are linearly independent over Q and that

$$a_1^{e_1} \dots a_{k-1}^{e_{k-1}} = a_k$$

This contradicts the induction hypothesis and the proof is complete.

2.3 A simplified form of Baker's theorem

The rest of this chapter will be devoted to proving theorem 3. The theorem is quite complicated and for clarity it is split into 7 parts; a preliminary simplification, 5 lemmas, and the final proof. Rather than prove theorem 3 we will prove the following theorem.

Theorem 6. With the assumptions of theorem 3 we may conclude that there is a number $C_1 = C_1\ (n,a_1,\dots,a_n,q,d) \geq 1$ such that for all algebraic numbers $b_1,\dots,b_{n-1}$ with degrees at most d we have

$$| b_1 \log a_1 + \ldots + b_{n-1} \log a_{n-1} - \log a_n | \geq e^{-(\log H)^q}$$

where H is a number not less than C_1 and the heights of $b_1, \ldots, b_{n-1}$.

Theorem 6 is a simplified form of theorem 3. To show that Theorem 6 implies theorem 3 we need two minor lemmas.

Lemma III. If a is an algebraic number of degree d and height H then $| a | \leq dH$.

Pf: a has minimal polynomial $a_d x^d + \ldots + a_1 x + a_o = 0$ and either $| a | \leq 1$ in which case we are done or

$$| a | \leq |a_d \, a| = |a_{d-1} + a_{d-2} a^{-1} + \ldots + a_o a^{-d+1} |$$
$$\leq d \cdot H.$$

Lemma IV. If a and b are algebraic numbers with degrees at most d and heights at most H, then ab has degree at most d^2 and height H' where $\frac{\log H'}{\log H} \leq 4d^2(1 + \log d)$, if $H \geq 2$ and $\log H' \geq 2d^2 \log d$ if $H = 1$.

Pf: Let $a^{(i)}, b^{(j)}$ denote the conjugates of a, b and let c and e be the leading coefficients of their minimal polynomials. We then have that ab is a solution of

$$(ce)^{d^2} \cdot \prod_{i,j} (x - a^{(i)} b^{(j)}),$$ a polynomial with integer coefficients.

Thus ab has degree at most d^2.

We also have that ab has height $H' \leq (ce)^{d^2} ((dH)^2)^{d^2}$ by using lemma 3. Thus $H' \leq (ce)^{d^2} (d^2)^{d^2} \cdot (H^2)^{d^2}$

$$\log H' \leq d^2 (\log ce + 2 \log d + 2 \log H)$$

and for $H \geq 2$

$$\frac{\log H'}{\log H} \leq d^2 \left(2 + 2 \frac{\log d}{\log 2} + 2 \right)$$

$$\leq 4d^2 (1 + \log d)$$

while for $H = 1$, $\log H' \leq 2d^2 \log d$.

We now show that theorem 6 implies theorem 3.

Pf: Given $b_1, \ldots b_n$ algebraic numbers of degree at most d and height at most H with $b_n \neq 0$ we define $b_i' = \frac{-b_i}{b_n}$ for $i = 1, \ldots n-1$.

Then by lemma IV the b_i', $i = 1, \ldots, n-1$ have degree at most d^2 and height at most H' where $\frac{\log H'}{\log H} \leq 4d^2(1+\log d)$ $= C_3(d)$ if $H \geq 2$ and $\log H' \leq 2d^2 \log d = C_3'(d) \leq C_3(d)$ if $H = 1$.

By theorem 6 we can state that

$$| b_1' \log a_1 + \cdots + b_{n-1}' \log a_{n-1} - \log a_n | \geq e^{-(\log H'')^{q^1}}$$

where $H'' = \max \{H', C\}$ and $2q^1 = q + 2n + 1$.

Thus $$\left| -\frac{1}{b_n} (b_1 \log a_1 + \cdots + b_{n-1} \log a_{n-1} + b_n \log a_n) \right|$$

$$\geq e^{-(\log H'')^{q^1}}$$

So $$| b_1 \log a_1 + \cdots + b_n \log a_n | \geq |b_n| e^{-(\log H'')^{q^1}}$$

$$\geq (dH)^{-1} e^{-(\log H'')^{q^1}}$$

As we have $e^{-(\log H'')^{q^1}} \geq e^{-(\log H')^{q^1}} \cdot e^{-(\log C_1)^{q^1}}$

as $C_1 \geq 1$ and $e^{-(\log H')^{q^1}} \geq e^{-C_3(d)^{q^1}} \cdot e^{-(C_3(d) \log H)^{q^1}}$

we can state that

$|b_1 \log a_1 + \ldots + b_n \log a_n| \geq C' (n, a_1, \ldots, a_n, q, d) . H^{-1} e^{-(C_3(d) \log H)^{q^1}}$

and if we let $C''(d,q)$ be the minimum of $\frac{1}{2} H^{-1} e^{-(C_3(d)\log H)^{q^1}}$.

$e^{(\log H)^q}$ considered as a function of H , $H \geq 1$, we have

$$| b_1 \log a_1 + \ldots + b_n \log a_n | > C' . C''(d,q) . e^{-(\log H)^q}$$

$$> C(n, a_1, \ldots, a_n, q, d) . e^{-(\log H)^q}$$

2.4 The proof of Baker's theorem

I will now restate Theorem 6 in full

Theorem 6. If $a_1, \ldots, a_n$ are non-zero algebraic numbers such that $\log a_1, \ldots \log a_n$ are linearly independent over the rationals ($n \geq 2$) and if $q > 2n+1$ and d is a positive integer we have that there is an effectively computable number $C_1 = C_1(n, a_1, \ldots, a_n, q, d) > 0$ such that for all algebraic numbers $b_1, \ldots, b_{n-1}$ with degrees at most d we have

$$| b_1 \log a_1 + \ldots + b_{n-1} \log a_{n-1} - \log a_n | \geq e^{-(\log H)^q}$$

where $H = \max_i \{ C_1, H(b_i) \}$.

Theorem 6 will be proved by assuming that there exists algebraic numbers $b_1, \ldots, b_{n-1}$ with degrees at most d such that (I) $| b_1 \log a_1 + \ldots + b_{n-1} \log a_{n-1} - \log a_n | < e^{-(\log H)^q}$ and deducing a contradiction. We will assume the hypotheses of theorem 6 as well as the above assumption in the proof of the 5 lemmas preceding the proof of the theorem. The lemmas represent certain stages of the proof. They are not independent.

Prior to lemma 3 I will state the following identities.

Identity 1. For $z \in ¢$ we have $|e^z - 1| =$

$$\leq | z | e^{|z|}$$

Let $z = b_1 \log a_1 + \ldots + b_{n-1} \log a_{n-1} - \log a_n$

Then by (I) we have

$$| a_1^{b_1} \ldots a_{n-1}^{b_{n-1}} - a_n | < | a_n | e^{-(\log H)^q + 1}$$

Identity 2. For any a that satisfies a polynomial equation $d_n x^n + \ldots + d_1 x + d_o = 0$ we may write

$$(d_n \ a)^j = g_o^{(j)} + g_1^{(j)} a + \ldots\ldots + g_{n-1}^{(j)} a^{n-1}$$

where $|g_i^{(j)}| \leq (2H(a))^j$ for all i

To see this we define the $g_i^{(j)}$ by the following recurrence relation.

$$g_i^{(j)} = d_n g_{i-1}^{(j-1)} - g_{n-1}^{(j-1)} d_i \quad (0 \leq i \leq n-1 , j \geq n)$$

with $g_{-1}^{(j-1)} = 0$

Note that

(i) $(d_n a)^j = (d_n a) (d_n a)^{j-1} = d_n a (g_o^{(j-1)} + g_1^{(j-1)} a + \ldots g_{n-2}^{(j-1)} a^{n-2})$
$+ d_n a^n g_{n-1}^{(j-1)}$

and (ii) $d_n a^n = -d_{n-1} a^{n-1} - d_{n-2} a^{n-2} \ldots - d_o$

Combining (i) and (ii) we verify the recurrence relation.

The recurrence relation gives our original inequality.

We are now ready to proceed with out theorem. As in Chapter I our approach will be the construction of a complex function, this time of several variables, which we will make small at integer values of the derivative of the function for partial derivatives of finite order, by judicious use of lemma 1. We do not make our main function and its derivatives zero for small integer values as in the Gelfond-Schneider theorem but instead make an auxiliary function that is very similar to our main function zero under the aforementioned conditions. Because of this similarity we can employ identity 1 to conclude our main function must be small. Our function and the conditions imposed upon it are more complicated than in the Gelfond-Schneider theorem as we are now working with a complex function of several variables. Lemma 3 merely shows that it is possible to define our function in the manner we described.

<u>Lemma 3</u>. Assume the hypotheses of theorem 6 and (I) then there are integers $p(\lambda_1,\ldots,\lambda_n)$, not all 0, with $|p(\lambda_1,\ldots\lambda_n)| \le e^{2hk}$ such that the function

$$f(z_1,\ldots,z_{n-1}) = \sum_{\lambda_1=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_1,\ldots,\lambda_n)\, a_1^{t_1 z_1} \cdots a_{n-1}^{t_{n-1} z_{n-1}}$$

where $t_i = \lambda_i + \lambda_n b_i$ and $L = [[h^{\delta}]^{1-\varepsilon}]$

where $h = [\log H]$, $\delta = \frac{1}{2}\left(1 + \frac{q}{2n+1}\right)$, $\varepsilon = \dfrac{1 - \frac{1}{\delta}}{2n}$

satisfies

(1) $\quad | f_{m_1,\ldots,m_{n-1}}(\ell, \ell, \ldots, \ell)| < e^{-\frac{1}{2}hq}$

for all integers $1 \le \ell \le h$ and all non-negative integers $m_1,\ldots,m_{n-1}$ with $m_1+\cdots+m_{n-1} \le [h^\delta]$

Note that $f_{m_1,\ldots m_{n-1}}(z_1,\ldots,z_{n-1}) = \dfrac{\partial^{m_1+\cdots+m_{n-1}}}{\partial z_1^{m_1}\cdots\partial z_{m_{n-1}}^{m_{n-1}}} f(z_1,\ldots z_{n-1})$

and that

$a_i^{t_i z} = e^{t_i z \log a_i}$ for some fixed determination of $\log a_i$.

Pf: For simplicity we denote $[h^\delta] = k$. We will apply lemma 1 to the function

$$g(\ell,m_1,\ldots m_{n-1}) = \sum_{\lambda_1=0}^{L}\cdots\sum_{\lambda_n=0}^{L} p(\lambda_1,\ldots,\lambda_n)\,(a_1^{\lambda_1}\cdots a_n^{\lambda_n})^{\ell}\cdot\prod_{i=1}^{n-1} t_i^{m_i}$$

to solve for $p(\lambda_1,\ldots\lambda_n)$ such that $g(\ell,m_1,\ldots,m_{n-1}) = 0$ for $1 \le \ell \le h$ and $m_1+\cdots+m_{n-1} \le k$ with $m_i \ge 0$.

After we have solved for the $p(\lambda_1,\ldots\lambda_n)$ we will show that this implies the lemma. We must therefore reduce the equations $g(\ell_1\, m_1,\ldots m_{n-1}) = 0$ to equations in integers only.

We let $c_1,\ldots,c_n$ and $d_1,\ldots,d_{n-1}$ denote the leading coefficients in the minimal polynomials of $a_1,\ldots,a_n$ and $b_1,\ldots,b_{n-1}$ respectively. We assume that the c_i and b_i are non-zero. We then have that

$$(c_i a_i)^j = \sum_{r=0}^{d-1} V_{ir}^{(j)}\, a_i^r \qquad (d_i b_i)^j = \sum_{u=0}^{d-1} W_{iu}^{(j)}\, b_i^u$$

where $V_{ir}^{(j)}$, $W_{iu}^{(j)} \varepsilon Z^+$ and $V_{ir}^{(j)} < C_5^j$, $W_{iu}^{(j)} < (2H)^j$ from identity (2).

Now consider $F = (c_1 c_2 \ldots c_n)^{L\ell} d_1^{m_1} \ldots d_{n-1}^{m_{n-1}} g(\ell, m_1, \ldots, m_{n-1})$

Writing $(d_p t_p)^{m_p} = \sum_{\mu_p=0}^{m_p} \binom{m_p}{\mu_p} (d_p \lambda_p)^{m_p-\mu_p} (\lambda_n d_p b_p)^{\mu_p}$

and substituting for $(c_i a_i)^j$ and $(d_i b_i)^j$ we have that

$$F = \sum_{\lambda_1=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_1, \ldots, \lambda_n) . \prod_{i=1}^{n} c_i^{(L-\lambda_i)\ell} . \sum_{r_i=0}^{d-1} V_{ir_i}^{(\lambda_i \ell)} a_i^{r_i}$$

$$. \left(\prod_{r=1}^{n-1} \sum_{\mu_r=0}^{m_r} \binom{m_r}{\mu_r} (d_r \lambda_r)^{m_r-\mu_r} . \lambda_n^{\mu r} . \sum_{g_r=0}^{d-1} W_{rg_r}^{(\mu_r)} b_r^{g_r} \right)$$

$$= \sum_{r_1=0}^{d-1} \cdots \sum_{r_n=0}^{d-1} \sum_{g_1=0}^{d-1} \cdots \sum_{g_{n-1}=0}^{d-1} G \; a_1^{r_1} \ldots a_n^{r_n} \; b_1^{g_1} \ldots b_{n-1}^{g_{n-1}}$$

where

$$G = \sum_{\lambda_1=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_1, \ldots \lambda_n) \sum_{\mu_1=0}^{m_1} \cdots \sum_{\mu_{n-1}=0}^{m_{n-1}} JK$$

$$J = \prod_{i=1}^{n} c_i^{(L-\lambda_i)\ell} . V_{ir_i}^{(\lambda_i \ell)}$$

and $$K = \prod_{r=1}^{n-1} \binom{m_r}{\mu_r} (d_r \lambda_r)^{m_r-\mu_r} \lambda_n^{\mu_r} W_{rg_r}^{(\mu_r)}$$

We therefore have that $F = 0$ and thus $g(\ell, m_1, \ldots, m_{n-1}) = 0$ if the d^{2n-1} equations $G=0$ hold for the stated range of the variables $\ell, m_1,\dots,m_{n-1}$. But these are linear equations in the $p(\lambda_1,\dots,\lambda_n)$ with integer coefficients. We prepare to use lemma 1 by estimating the absolute value of

$$\sum_{\mu_1=0}^{m_1}\dots\sum_{\mu_{n-1}=0}^{m_{n-1}} JK \text{ , the coefficients of } p(\lambda_1,\dots\lambda_n) \text{ .}$$

We have $|J| \leq C_6^{L\ell n} \leq C_6^{Lhn}$ as $\ell \leq h$.

and

$$|K| \leq \prod_{r=1}^{n-1} 2^{m_r} (HL)^{m_r-\mu_r} L^{\mu_r} (2H)^{\mu_r}$$

$$\leq \prod_{r=1}^{n-1} (4HL)^{m_r}$$

$$\leq (4HL)^k$$

and thus $\left| \sum_{\mu_1=0}^{m_1}\dots\sum_{\mu_{n-1}=0}^{m_{n-1}} JK \right|$

$$\leq (m_1+1)\dots(m_{n-1}+1) \cdot C_6^{Lhn} \cdot (4HL)^k$$

$$\leq (8HL)^k \cdot C_6^{Lhn}$$

We have at most $(k+1)^{n-1} h \cdot d^{2n-1}$ different equations $G=0$ corresponding to the $(k+1)^{n-1}.h$ distinct sets of integral values for $\ell, m_1,\dots,m_{n-1}$.

We also have $(L+1)^n$ unknowns $p(\lambda_1,\dots\lambda_n)$. Observe that $(L+1)^n > k^{n-n\varepsilon} > 2(k+1)^{n-1}. h.d^{2n-1}$ if H is assumed large enough. For this is true if $h^{1+\delta n\varepsilon}$ differs from h^δ by a power of h. In fact by our choice of δ and ε we have $1+\delta n\varepsilon = 1+\frac{\delta-1}{2} = \frac{\delta+1}{2} < \delta$ and we have our required inequality.

By lemma 1, therefore, with $N = (L+1)^n$, $M=(k+1)^{n-1}.h.d^{2n-1}$ and with $A = (8HL)^k.\ C_6^{Lhn}$ we have that there exists a non-trivial integral solution in the $p(\lambda_1,\ldots\lambda_n)$ to the equations $G = 0$ under consideration such that $|p(\lambda_1,\ldots,\lambda_n)| \leq N \cdot A$.

We have that $N = (L+1)^n \leq (k^{1-\varepsilon}+1)^n \leq k^n$, that $\log H < \frac{3}{2}h$ and that $nLh \leq nkh^{1-\delta\varepsilon}$

$$\leq nkh^{1-\frac{\delta-1}{2n}}$$ and therefore that

$$NA \leq k^n \ .\ (8ke^{\frac{3}{2}h})^k \ .\ C_6^{Lhn}$$

$$\leq e^{2hk}$$

for sufficiently large H.

We will now check that our function $g(\ell,m_1,\ldots,m_{n-1})$ vanishing for all non-negative integers ℓ such that $1 \leq \ell \leq h$ and $m_1,\ldots,m_{n-1}$ with $m_1+\ldots+m_{n-1} \leq k$ implies that $f(z_1,\ldots,z_{n-1})$ satisfies (1).

We note that $(\log a_1)^{m_1}\ldots(\log a_{n-1})^{m_{n-1}} g(\ell,m_1\ldots,m_{n-1})$ is the same as $f_{m_1,\ldots,m_{n-1}}(\ell,\ell,\ldots,\ell)$ if we substitute $a_1^{b_1}\ldots a_{n-1}^{b_{n-1}}$ for a_n in $g(\ell,m_1,\ldots,m_{n-1})$. We have that (3) $|a_1^{b_1}\ldots a_{n-1}^{b_{n-1}}| < |a_n| + 1$ from identity 1 and that

$$|E| = |(a_1^{b_1}\ldots a_{n-1}^{b_{n-1}})^{\lambda_n\ell} - a_n^{\lambda_n\ell}| \leq L\ .\ell.\ |a_n|e^{-(\log H)^{q+1}} \cdot (|a_n|+1)^{L\ell}$$

$$\leq C_6^{L\ell}\, e^{-h^q}$$

from identity 1, (3) and the inequality

$$| x^\lambda - y^\lambda | = |x - y| \cdot |x^{\lambda-1}y + \dots + xy^{\lambda-1}|$$

$$\leq \lambda |x - y| (|y| + 1)^\lambda \quad \text{for} \quad |x| < |y| + 1 .$$

As $|t_i| \leq 2dLH$ we have

$$|D| = |(\log a_1)^{m_1} \dots (\log a_{n-1})^{m_{n-1}} a_1^{\lambda_1 \ell} \dots a_{n-1}^{\lambda_{n-1}\ell} t_1^{m_1} \dots t_{n-1}^{m_{n-1}}|$$

$$\leq C_7^{k+L\ell} (2d\, LH)^k$$

$$< e^{2hk} \quad \text{if as before, we assume } H \text{ is}$$

sufficiently large.

Now $|f_{m_1,\dots m_{n-1}}(\ell,\dots,\ell)|$

$$= |f_{m_1,\dots,m_{n-1}}(\ell,\ell,\dots\ell) - g(\ell, m_1,\dots,m_{n-1}) \cdot \prod_{i=1}^{n-1} (\log a_i)^{m_i}|$$

$$\leq \sum_{\lambda_1=0}^{L} \dots \sum_{\lambda_n=0}^{L} |p(\lambda_1 \dots \lambda_n)| . |D| . |E|$$

$$\leq (L+1)^n \cdot e^{2hk} \cdot e^{2hk} \cdot C_6^{Lh} e^{-h^q}$$

$$\leq e^{-\frac{1}{2} h^q} \quad \text{for sufficiently large } H \text{ as}$$

$$Lh \leq h\,k \leq h^{1+\delta} < h^q . \qquad \text{Q.E.D.}$$

In lemma 3 we defined our function $f(z_1,\dots z_{n-1})$ and showed that its partial derivatives were subject to certain restrictions on their absolute value for integral values. In lemma 4 we determine a bound for the absolute value of $f(z_1,\dots z_{n-1})$ for certain of its partial derivatives while considering f as a function of one variable z .

(i.e. $z_1 = z_2 = \ldots = z_{n-1} = z$). We also estimate $f(z_1,\ldots,z_{n-1})$ and certain of its partial derivatives for $z_1 = \ldots = z_{n-1} = \ell$ where ℓ is an integer bounded above by certain value. We did this in lemma 3 of course but in lemma 4 the range of ℓ is extended. Over this extended range of ℓ we prove that $|f_{m_1,\ldots,m_{n-1}}(\ell,\ell,\ldots\ell)|$ is either bounded above by a number or bounded below by a larger number. These two possibilities arise from the fact that $|N(\ )|$ of an algebraic number is either 0 or greater than or equal to 1. This technique of course was used repeatedly in Chapter I. The difference in this case is, as in lemma 3, that the algebraic number under consideration comes not from the function $f(z_1,\ldots z_{n-1})$ but from the auxiliary function $g(\ell, m_1,\ldots,m_{n-1})$. We obtain our result by employing identity 1. We will use the results of lemma 4 in estimating integrals that will arise in lemma 5.

<u>Lemma 4</u>. Assume the hypotheses of theorem 6 and (I) then for any non-negative integers $m_1,\ldots,m_{n-1}$ with $m_1+m_2+\ldots+m_{n-1} \leq k$ and any complex number z we have

(2) $|f_{m_1,\ldots,m_{n-1}}(z,\ldots,z)| \leq e^{4hk} C_9^{L|z|}$

and for any integer ℓ with $1 \leq \ell \leq h^{q-\delta+\frac{1}{2}\varepsilon\delta}$ either

(1) $|f_{m_1,\ldots,m_{n-1}}(\ell,\ldots,\ell)| < e^{-\frac{1}{2}h^q}$ or

(3) $|f_{m_1,\ldots,m_{n-1}}(\ell,\ldots,\ell)| > (e^{6hk} C_{10}^{L\ell})^{-d^{2n-1}}$

<u>Pf</u>: $f_{m_1,\ldots m_{n-1}}(z,\ldots,z) = (\log a_1)^{m_1}\cdots(\log a_{n-1})^{m_{n-1}} \sum_{\lambda_1=0}^{L}\cdots\sum_{\lambda_n=0}^{L}$

$$p\,(\lambda_1,\ldots,\lambda_n)\ .\prod_{i=1}^{n-1} a_i^{t_i z}\ .\prod_{i=1}^{n-1} t_i^{m_i}$$

From lemma 3 we recall that $|p\,(\lambda_1,\ldots\lambda_n)| < e^{2hk}$ and that $|t_i| < 2dLH$. We need only find an estimate for $\prod_{i=1}^{n-1} a_i^{t_i z}$. By our assumption (I) we have that

$$|z|\ \ |b_1 \log a_1+\ldots+b_{n-1} \log a_{n-1}| < |z|(\,|\log a_n|+e^{-(\log H)^q})$$

$$< \ |z|\ \ (|\log a_n| + 1\,)$$

and therefore that

$$|a_1^{b_1 z}\ a_2^{b_2 z}\ \ldots\ a_{n-1}^{b_{n-1} z}| < e^{\,|z|\ (|\log a_n| + 1)}\ .$$

We can now conclude that

$$|f_{m_1,\ldots,m_{n-1}}\ (z,z,\ldots z)| \ \le C_8^k\ (L+1)^n\ e^{2hk}.C_9^{L|z|}\ (2dLH)^k$$

$$\le\ e^{4hk}\ C_9^{L|z|}$$

We have thus proved our first statement. To prove our second statement we consider an algebraic integer defined by our function $g(\ell, m_1,\ldots,m_{n-1})$. We let $c_1,\ldots c_n$ and $d_1,\ldots d_{n-1}$ be defined as in lemma 3. We then have that $S = (c_1\ldots c_n)^{L\ell}\ .\ \prod_{i=1}^{n-1} d_i^{m_i}\ .\ g\ (\ \ell,\ m_1,\ldots m_{n-1})$ is an algebraic integer.

We note that S is always in the field $Q\ (a_1,\ldots,a_n,\ b_1,\ldots,b_{n-1})$ and thus we may conclude that the degree of S is less than or equal to d^{2n-1}. As any conjugate of S has absolute value at most

$$(L+1)^n e^{2hk} C_{10}^{L\ell} (2dLH)^{2k} \leq e^{6hk} C_{10}^{L\ell}$$

we may conclude that either $S = 0$ or

$$|S| \geq (e^{6hk} C_{10}^{L\ell})^{-d^{2n-1}+1}$$

This is just because if $S \neq 0$ then $|N(S)| \geq 1$.

We also have that

$$\left| \prod_{i=1}^{n-1} a_i^{t_i \ell} \prod_{i=1}^{n-1} t_i^{m_i} - \prod_{i=1}^{n} a_i^{\lambda_i \ell} \prod_{i=1}^{n-1} t_i^{m_i} \right|$$

$$\leq \left| \prod_{i=1}^{n-1} a_i^{\lambda_i \ell} \prod_{i=1}^{n-1} t_i^{m_i} \right| . \left| \prod_{i=1}^{n-1} a_i^{b_i \lambda_n \ell} - a_n^{\lambda_n \ell} \right|$$

and recalling the proof of lemma 2 and our estimate for E we conclude the above is

$$\leq C_{11}^{L\ell} (2dLH)^k C_{12}^{L\ell} e^{-h^q} \leq C_{13}^{L\ell} e^{2hk} e^{-h^q}$$

As $L \leq h^{\delta(1-\varepsilon)}$ and by assumption $\ell \leq h^{q+\delta\varepsilon\frac{1}{2}-\delta}$, we have $L\ell \leq h^{q-\frac{1}{2}\varepsilon\delta}$ and $hk \leq h^{1+\delta}$. We thus have that the above is $\leq e^{-\frac{3}{4}h^q}$ for sufficiently large h.

If we let $P = (c_1 \ldots c_n)^{L\ell} \prod_{i=1}^{n-1} d_i^{m_i}$ and

$$O = \prod_{i=1}^{n-1} (\log a_i)^{m_i}$$

then we have $|O^{-1} f_{m_1 \ldots m_{n-1}} (\ell, \ldots \ell) - P^{-1} S|$

$$\leq (L+1)^n e^{2hk} e^{-\frac{3}{4}h^q} \leq e^{-\frac{5}{8}h^q}$$

Now if $S = 0$ we have that

$$| f_{m_1 \cdots m_{n-1}}(\ell,..\ell) | \leq | O | e^{-\frac{5}{8} h^q} < e^{-\frac{1}{2}h^q}$$

and we have (1) the first of our two alternatives. If $S \neq 0$ then we have

$$| \frac{P}{O} f_{m_1,\ldots m_{n-1}}(\ell,\ldots,\ell) - S | \leq |P| e^{-\frac{5}{8}h^q}$$

$$|S| - |P| e^{-\frac{5}{8}h^q} \leq | \frac{P}{O} | | f_{m_1,\ldots,m_{n-1}}(\ell,\ldots,\ell) |$$

$$|\frac{O}{P}| \left(|S| - |P| e^{-\frac{5}{8} h^q} \right) \leq | f_{m_1,\ldots m_{n-1}}(\ell,\ldots,\ell) |$$

By observing that $|P| \leq C_{16}^{L\ell} \cdot (dH)^{+k}$

and $| \frac{O}{P} | > C_{16}^{-L\ell} (dH)^{-k} \cdot C_{17}^{-k}$

we may conclude that

$$| f_{m_1,\ldots,m_{n-1}}(\ell,\ldots,\ell) | \geq (e^{6hk} C_{10}^{L\ell})^{-d^{2n-1}} \quad \text{if } S \neq 0.$$

This then gives us (3) and the proof of the lemma is complete.

We now have the necessary estimates required for lemma 5. In lemma 5 we show that by increasing our restrictions on the partial derivatives of $f(z_1,\ldots,z_{n-1})$ we may relax our restrictions on the integer values $f(z,\ldots,z_{n-1})$ assumes such that with these restrictions f is small. By small I mean that f satisfies condition (1) of Lemma 3. We can, by this argument, then look at the function at integer values where we have no partial derivatives of the function to consider and thus show that our function is small for a large number of integer arguments, with the restriction that $z_1=z_2=\ldots=z_{n-1}$. We have ventured into the field of several complex variables to strengthen our estimate on f. Lemma 5 allows us to collapse to consideration of a function of a single complex variable again.

We prove lemma 5 by induction and Cauchy's residue theorem.

We observe that the hypotheses of lemma 4 are satisfied for our induction step and to complete our induction step we show that assumption of the second of the two alternatives of lemma 4 leads to a contradiction.

With lemma 5 we are able to conclude a modified version of Baker's theorem. In this version we must assume that $\log a_1,\ldots,\log a_n$ and $2\pi i$ are lineraly independent. We use an argument involving a Vandermonde determinant which is reminiscent of Gelfond's proof of the Gelfond-Schneider theorem. The inclusion of $2\pi i$ weakens the theorem and we can conclude only the following version of theorem 5.

<u>Theorem 5'</u> If $a_1,\ldots,a_n$ denote positive real algebraic numbers other than 0 or 1 and $b_1,\ldots,b_n$ denote real algebraic numbers with $1, b_1\ldots,b_n$ linearly independent over the rationals then $a_1^{b_1}\ldots a_n^{b_n}$ is transcendental.

The exclusion of $2\pi i$ from our hypotheses forces us to prove 2 more preliminary lemmas before we can finally prove our theorem. I will discuss this following the proof of lemma 5.

Lemma 5. Assume the hypotheses of theorem 6 and (I) then let J be any integer satisfying $0 \leq J < \tau$ where

$$\tau = 2\varepsilon^{-1} \left\{ \frac{(q-1)}{\delta} - 1 \right\} + 1$$

then condition (1) holds for all integers ℓ with $1 \leq \ell \leq hk^{\frac{1}{2}\varepsilon J}$ and each set of non-negative integers $m_1,\ldots,m_{n-1}$ with $m_1+\cdots+m_{n-1} \leq \frac{k}{2^J}$

Pf: We will use induction. From lemma 3 we have that for $J = 0$ our lemma is true. We now let N be an integer such that $0 \leq N < \tau-1$ and we assume that the lemma is true for $J = 0,\ldots,N$. We will now prove that the lemma is true for $J = N+1$

We let $S = \left[\frac{k}{2^{N+1}}\right]$, $R_N = \left[hk^{\frac{1}{2}\varepsilon N}\right]$ and $R_{N+1} = \left[hk^{\frac{1}{2}\varepsilon(N+1)}\right]$. We now need only prove that for integer ℓ with $R_N < \ell \leq R_{N+1}$ and any non-negative integers $m_1,\ldots,m_{n-1}$ with $m_1+\cdots+m_{n-1} \leq S$ we have condition (1) satisfied.

We also define

$$f(z) = f_{m_1,\ldots m_{n-1}}(z,\ldots z,\ldots,z)$$

and thus we have

$$f^{(m)}(z) = \sum_{\substack{j_1=0 \\ j_1+\dots+j_{n-1}=m}}^{m} \cdots \sum_{j_{n-1}=0}^{m} \frac{m!}{j_1!\dots j_{n-1}!} f_{m_1+j_1,\dots,m_{n-1}+j_{n-1}}(z,z,\dots,z)$$

We therefore conclude that $| f^{(m)}(\ell) | < n^m . e^{-\frac{1}{2}h^q}$ for $1 < \ell \leq R_N$ and $0 \leq m \leq S$. This is so as $m_1 + j_1 + \dots + m_{n-1} + j_{n-1} \leq 2\,S < \frac{k}{2^N}$ which means that $| f_{m_1+j_1\dots,m_{n-1}+j_{n-1}}(\ell,\dots,\ell) | < e^{-\frac{1}{2}h^q}$ by the induction hypothesis.

We now define $F(z) = ((z-1)\dots\dots(z-R_N))^{S+1}$ and we let C and C_r be circles in the complex plane described in the positive sense, such that C is defined by $|z| = R_{N+1} \log h$ and C_r is defined by $|z-r| = \frac{1}{2}$. By Cauchy's residue theorem we have

$$\frac{1}{2\pi i} \int_C \frac{f(z)}{(z-\ell)F(z)} dz = \frac{f(\ell)}{F(\ell)} + \frac{1}{2\pi i} \sum_{r=1}^{R_N} \int_{C_r} \frac{f(z)}{(z-\ell)\,F(z)} dz$$

and $\displaystyle\int_{C_r} \frac{f(z)}{(z-\ell)\,F(z)} dz = 2\pi i \frac{1}{S!} \frac{d^S}{dz^S} \frac{(z-r)^{S+1} f(z)}{(z-\ell)\,F(z)}$

evaluated at r.

$$= \frac{2\pi i}{S\,!} \sum_{i=0}^{S} \binom{S}{i} f^{(i)}(z) \left(\frac{(z-r)^{S+1}}{(z-\ell)\,F(z)} \right)^{(S-i)}$$

evaluated at r , by Liebnitz's rule for the derivative of a product.

$$= \sum_{i=0}^{S} \frac{f^{(i)}(z)}{i!} \int_{C_r} \frac{(z-r)^{i}}{(z-\ell)\,F(z)} dz.$$

We note that we are able to use Cauchy's theorem because the function

$$f(z) = (\log a_1)^{m_1} \cdots (\log a_{n-1})^{m_{n-1}} \sum_{\lambda_1=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_1,\dots,\lambda_n) \cdot \prod_{i=1}^{n-1} t_i^{m_i} \cdot$$

$$e^{z\cdot(t_1 \log a_1+\dots t_{n-1}\log a_{n-1})}$$

is entire for a fixed determination of $t_1 \log a_1+\dots+t_{n-1}\log a_{n-1}$.

We now derive our lemma by comparing estimates of both sides of the equation

$$(5) \quad \frac{1}{2\pi i}\int_C \frac{f(z)}{(z-\ell)F(z)}\,dz = \frac{f(\ell)}{F(\ell)} + \frac{1}{2\pi i}\sum_{r=1}^{R_N}\sum_{i=0}^{S} \frac{f^{(i)}(z)}{i!}\int_{C_r} \frac{(z-r)^i}{(z-\ell)}\,\frac{dz}{F(z)}$$

We have that as $\ell \le R_{N+1} \le hk^{\frac{1}{2}\varepsilon(N+1)}$

$$\le hk^{\frac{1}{2}\varepsilon\tau} \le h^{1+\frac{1}{2}\delta\varepsilon\tau} \le h^{\frac{1}{2}\varepsilon\delta+q-\delta}$$

either (1) $|f(\ell)| < e^{-\frac{1}{2}h^q}$ or (3) $|f(\ell)| > (e^{6hk}C_{10}^{L\ell})^{-d^{2n-1}}$

by lemma 4. We will use the above equation to show that the second possibility leads to contradiction. This will suffice to prove our lemma.

Assume (3) therefore, in what follows.

We then have that the absolute value of the double sum in (5) denote it by A satisfies.

$$|A| \le R_N \cdot (S+1) \cdot n^k\, e^{-\frac{1}{2}h^q} \cdot \max_{|z-r|=\frac{1}{2}} \left|\frac{(z-r)^i}{(z-\ell)\;F(z)}\right|$$

$$\le hk^{\frac{1}{2}\varepsilon\tau} \cdot (k+1) \cdot n^k \cdot e^{-\frac{1}{2}h^q} \cdot \quad 4^{S+1}$$

$$\le e^{-\frac{1}{4}h^q}$$

for sufficiently large h.

We now estimate $\frac{f(\ell)}{F(\ell)}$ the first term on the right side of (5). We have

$$|F(\ell)| \le \ell^{R_N(S+1)} \le R_{N+1}^{R_N(S+1)}$$

and recalling that $R_{N+1} \le h^{q-\delta+\frac{1}{2}\varepsilon\delta}$ we check that $R_N(S+1)$ is less than or equal to $h^{-\xi} \cdot h^q$ where $\xi > 0$ to conclude that $|F(\ell)| \le e^{\frac{1}{8}h^q}$ for h sufficiently large.

We have $R_N(S+1) \le hk^{\frac{1}{2}\varepsilon N+1}$

$$= h^{1+\delta\frac{1}{2}\varepsilon(\tau-1)+\delta} \cdot h^{\delta\frac{1}{2}\varepsilon(N-(\tau-1)}$$

$$= h^q \cdot h^{-\xi} \qquad \text{where } \xi > 0.$$

We thus conclude $|F(\ell)| \le e^{\frac{1}{8}h^q}$.

Now by (3)

$$|f(\ell)| > (e^{6hk}C_{10}^{L}R_{N+1})^{-d^{2n-1}}$$

$$> (e^{6hk}C_{10}h^{q-\frac{1}{2}\varepsilon\delta})^{-d^{2n-1}}$$

$$> 2e^{-\frac{1}{8}h^q}$$

and so $\left|\frac{f(\ell)}{F(\ell)}\right| \ge \frac{|f(\ell)|}{|F(\ell)|} \ge 2e^{-\frac{1}{4}h^q}$.

This shows us that the right side of (5) has absolute value at least $\frac{1}{2}\left|\frac{f(\ell)}{F(\ell)}\right|$.

We now will estimate the left hand side of (5). Let V be the upper bound of $|f(z)|$ and W be the lower bound of $|F(z)|$ on the circle C. Then as $2\ |z-\ell|$ is greater than the radius of C for z on C we have

$$\frac{2V}{W} > \frac{1}{2} \frac{|f(\ell)|}{|F(\ell)|}$$

or
$$4V\ |f(\ell)|^{-1} > W\ |F(\ell)|^{-1}$$

Now we have

$$W \geq (\tfrac{1}{2} R_{N+1} \log h)^{R_N(S+1)}$$

and therefore $\quad W\ |F(\ell)|^{-1} \geq (\tfrac{1}{2} \log h)^{R_N(S+1)}$

From the first assertion of lemma 4 we have

$$V \leq e^{4hk}\, C_9^{\,L\, R_{N+1} \log h} .$$

and therefore $\quad V\ |f(\ell)|^{-1} \leq \left(e^{6hk}\, C_{18}^{\,LR_{n+1} \log h}\right)^{d^{2n-1}+1}$

Combining the three preceding results and taking logarithms we find

$$\log 4 + (d^{2n-1} + 1)\ (6hk + LR_{N+1} \log h \log C_{18})$$

$$\geq R_N\ (S+1)\ (-\log 2 + \log\log h)$$

or for sufficiently large h either

(i) $\quad C_{19}\ L\ R_{N+1} \log h \geq R_N\ (S+1) \log\log h \quad$ or

(ii) $\quad C_{19}\ hk \geq R_N\ (S+1)\ \log\log h\ $ depending on whether $N > 0$ or $N = 0$ respectively.

But $\ hk \leq h^{1+\delta}\ $ and $\ LR_{N+1} \leq h^{\delta(1-\varepsilon)+1+\delta\frac{1}{2}\varepsilon\ (N+1)}$

$$\leq h^{1+\delta+\delta\frac{1}{2}\varepsilon(N-1)}$$

while $R_N(S+1) \geq \frac{1}{2} hk^{\frac{1}{2}\varepsilon N} . k . 2^{-N-1}$

$$\geq C_{20} h^{1+\delta+\delta\frac{1}{2}\varepsilon N}$$

substituting these quantities into either equations (i) or (ii) we reach a contradiction. We have thus proved lemma 5.

As was mentioned earlier, lemmas 3, 4 and 5 suffice to prove theorem 5' . We need, however, two additional lemmas to prove theorem 3. Lemma 6 contains the result that will be used to prove our theorem. Lemma 7 is a simple lemma that is not connected with the chain of lemmas that culminate in lemma 6. In lemma 7 we employ the linear independence of $\log a_1,\ldots,\log a_n$ and the fact that $|N(\ )|$ of a non-zero algebraic integer is ≥ 1 to derive an estimate we will use once in the final proof of theorem.

In lemma 6 we use lemma 5 and the techniques of lemma 5 to improve our estimate on $g(z) = f(z,\ldots,z)$ at certain integer values of z. We then use Cauchy's residue theorem to establish an estimate for the derivatives of $g(z)$ of high order at the point $z = 0$. The estimate will be sufficiently strong for us to establish a contradiction and thus our theorem. Lemma 6 is the fourth step in the improvement of our estimate on the behaviour of the function $f(z_1,\ldots,z_{n-1})$. It follows, as do the others, from the lemmas preceding it.

<u>Lemma 6</u>. With the hypotheses of theorem 6 and (I) we have $\log |g^{(j)}(0)| < - h^q/\log h$ for $1 \leq j \leq k^n$

Pf: We define $X = [\frac{1}{2} h^{q-\delta}]$ and $Y = [k/(8q \log h)]$ and we note that $[hk^{\frac{1}{2}\epsilon N}] \geq X$ and $[k/2^N] \geq Y$ where $\tau > N > [\tau]$ if τ is not an integer or $N = \tau - 1$ if τ is an integer.

We therefore have from lemma 5 that $|f_{m_1,\ldots,m_{n-1}}(\ell,\ldots,\ell)| < e^{-\frac{1}{2}h^q}$ for non-negative integers $m_1,\ldots,m_{n-1}$, ℓ such that $1 \leq \ell \leq x$ and $m_1+\ldots+m_{n-1} \leq Y$. This allows us to conclude that as

$$g^{(m)}(\ell) = \left(\frac{\partial}{\partial z_1} + \ldots + \frac{\partial}{\partial z_{n-1}}\right)^m f(z_1,\ldots,z_{n-1})$$

evaluated at ℓ. We have

$$|g^{(m)}(\ell)| < n^m e^{-\frac{1}{2}h^q} \leq n^k e^{-\frac{1}{2}h^q}$$

for integers m, ℓ such that $1 \leq \ell \leq X$ and $0 \leq m \leq Y$.

We proceed as in lemma 5. We consider circles in the complex plane, described in a positive sense, where C, C_o, and C_i, $i = 1,\ldots,X$ are $|z| = X \log h$, $|z| = \frac{1}{2}$, and $|z-i| = \frac{1}{2}$ respectively. We also define the function $E(z) = ((z-1)\ldots(z-X))^{Y+1}$.

Our purpose is to calculate an upper bound for $|g(y)|$ where $y \in C_o$.

As in lemma 5 we have from Cauchy's residue theorem

$$\frac{1}{2\pi i}\int_C \frac{g(z)\,dz}{(z-y)E(z)} = \frac{g(y)}{E(y)} + \frac{1}{2\pi i}\sum_{i=1}^{X}\sum_{m=0}^{Y}\frac{g_m(i)}{m!}\int_{C_i}\frac{(z-i)^m\,dz}{(z-y)E(z)}$$

and designating the double sum by B we have

$$|B| < X(Y+1)\, n^k e^{-\frac{1}{2}h^q} \cdot 4^{Y+2}$$
$$< (4n)^{k+2} \cdot h^q \cdot e^{-\frac{1}{2}h^q}$$
$$< e^{-\frac{1}{4}h^q}$$

We let S be the maximum of $|g(z)|$ on C and T be the minimum of $E(z)$ on C. We thus have that

$$|g(y)| \leq (2.S.T^{-1} + e^{-\frac{1}{4}h^q})\, |E(y)|$$

By lemma 4 condition (2) we have

$$|S| < e^{4hk}\, C_9^{LX\log h}$$

We also have that $|T| \geq (\frac{1}{2} X \log h)^{X(Y+1)}$ for h large enough and
$$|E(y)| \leq (x+1)^{X(Y+1)}$$
$$\leq (2X)^{(X(Y+1)}$$
$$\leq (2X)^{2XY} \quad \text{by inspection.}$$

Therefore we have

$$|g(y)| \leq 2\, e^{4hk} C_9^{LX \log h} (\tfrac{1}{4} \log h)^{-X(Y+1)}$$
$$+ e^{-\frac{1}{4}h^q} \cdot (2X)^{2XY}$$

and as $L\, X \leq h^{\delta(1-\varepsilon)} \cdot \frac{1}{2}\, h^{q-\delta} < h^{q-\delta\varepsilon}$

and $X(Y+1) \geq \frac{1}{4}h^{q-\delta} . h^{\delta} / (16q \log h) \geq h^q/(64q \log h)$

$$|g(y)| \leq (\log h)^{-\frac{3}{4} X(Y+1)} + e^{-\frac{1}{4}h^q} \cdot (2X)^{2XY}$$

$$2\, X\, Y \log 2X \leq h^{q-\delta} \cdot \frac{h^{\delta}}{8q \log h} \cdot (q-\delta) \log h$$
$$\leq h^q/8$$

So
$$|g(y)| \leq (\log h)^{-\frac{3}{4} X\,(Y+1)} + e^{-\frac{1}{8}h^q}$$
$$\leq (\log h)^{-\frac{1}{2}X(Y+1)}$$

We now will specialize our estimate to $g^{(j)}(0)$ by using Cauchy's residue theorem.

$$\frac{j!}{2\pi i} \int_{C_0} \frac{g(y)}{y^{j+1}} \, dy = g^{(j)}(0) \text{ and for } 0 \leq j \leq k^n$$

we have

$$|g^{(j)}(0)| < j!\, 4^j (\log h)^{-\frac{1}{2}X(Y+1)}$$
$$< k^{k^{n+1}} (\log h)^{-\frac{1}{2}X(Y+1)}$$

But

$$\tfrac{1}{2} X (Y+1) \log\log h \geq \frac{h^q}{128 q \log h} . \log\log h$$

$$\geq 2\delta h^{\delta(n+1)} \log h \geq 2k^{n+1} \log k$$

as $\delta (2n+1) < q$ by assumption .

Thus

$$|g^{(j)}(0)| < (\log h)^{-\frac{1}{4}X(Y+1)}$$

for $0 \leq j \leq k^n$ and the lemma is proved.

<u>Lemma 7</u>. Assume the hypotheses of theorem 6. If $t_1, \ldots t_n$ are non-zero integers with $\max_i |t_i| \leq T$, then

$$| t_1 \log a_1 + \ldots + t_n \log a_n | > C_{21}^{-T}$$

<u>Pf</u>: As in lemma 3 we let $c_1, \ldots, c_n$ represent the leading coefficients in the minimal defining polynomials of $a_1, \ldots, a_n$ respectively. We then consider $\alpha = c_1^{t_1} \ldots c_n^{t_n} (a_1^{t_1} \ldots a_n^{t_n} - 1)$, an algebraic integer of degree at most d^{2n-1}.

We have that the absolute value of any conjugate of α is less than C_{22}^T. Thus, if $\alpha \neq 0$, $|\alpha| > C_{22}^{-(d^{2n-1})T}$ by consideration of the norm of α . If $\alpha = 0$ we have that $|t_1 \log a_1 + \ldots + t_n \log a_n|$ is a non-zero multiple of $2\pi i$ from the linear independence over the rationals of $\log a_1, \ldots, \log a_n$. This case clearly yields our theorem.

As $|e^z-1| \leq 2|z|$ for $|z| \leq \frac{1}{2}$ we have

$$|a_1^{t_1}\ldots a_n^{t_n}-1| \leq 2 \mid t_1 \log a_1+\ldots +t_n\log a_n \mid$$

for $\mid t_1 \log a_1+\ldots+t_n \log a_n \mid \leq\frac{1}{2}$.

Therefore

$$C_{22}^{-(d-1^{2n-1})T} < \mid \alpha \mid \leq \mid c_1^{t_1}\ldots c_n^{t_n}|.|a_1^{t_1}\ldots a_n^{t_n}-1|$$

$$\leq 2C_{23}^{T} \mid t_1 \log a_1+\ldots+t_n \log a_n \mid$$

and thus the lemma is complete.

We are now ready to finish the proof of the theorem. For the final proof we shall need lemmas 6 and 7. We will show that the assumption (I) must be false. We will estimate a Vandermonde determinant whose elements are of the form $(t_1 \log a_1 +\ldots+t_n\log a_n)^j$ in two ways to arrive at our contradiction. We derive one estimate with the aid of lemma 7. Our second estimate comes from consideration of certain sums of the original elements of the Vandermonde determinant. These sums will have values close to $g^{(j)}(0)$. We can then employ our estimate for $g^{(j)}(0)$ and the property of determinants that a multiple of one row of a determinant added to another row leaves the determinant unchanged. We are thus able to procure our contradiction and the theorem will be proved.

<u>The proof of Theorem 6 and thus also of Theorem 3.</u>

<u>Pf</u>: We order the $p(\lambda_1,\ldots,\lambda_n)$ by associating them with $r = \lambda_n\lambda_{n-1}\ldots\lambda_1.00$ number expressed in base $L+1$. We let $p_r = p(\lambda_1,\ldots,\lambda_n)$, $\theta_r=\lambda_1\log a_1+\ldots+\lambda_n \log a_n$ and $\Theta_j = \sum_{r=0}^{R} p_r \theta_r^j \quad 0 \leq j \leq R$ where $R = (L + 1)^n-1$.

We consider

$$|g^{(j)}(0) - \theta_j| = |\sum_{\lambda_1=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_1,\ldots,\lambda_n)$$

$$((t_1 \log a_1 + \ldots + t_{n-1} \log a_{n-1})^j$$

$$- (\lambda_1 \log a_1 + \ldots + \lambda_n \log a_n)^j)|$$

Note that $(a+b)^j - (a+c)^j = (a+c+b-c)^j - (a+c)^j$

$$= \binom{j}{1}(b-c)\ (a+c)^{j-1} + \binom{j}{2}(b-c)^2\ (a+c)^{j-2} \ldots + (b-c)^j$$

and letting $a = \lambda_1 \log a_1 + \ldots + \lambda_{n-1} \log a_{n-1}$

$b = \lambda_n(b_1 \log a_1 + \ldots + b_{n-1} \log a_{n-1})$ and $c = \lambda_n \log a_n$

we conclude that

$$|(t_1 \log a_1 + \ldots + t_{n-1} \log a_{n-1})^j - (\lambda_1 \log a_1 + \ldots + \lambda_n \log a_n)^j|$$

$< (C_{25}L)^R \cdot e^{-h^q}$ for sufficiently large h by employing (I) and estimating θ_r by $|\theta_r| < C_{24}L$.

Thus we have

$$|g^{(j)}(0) - \theta_j| < (L+1)^n \cdot e^{2hk} \cdot (C_{25}L)^R \cdot e^{-h^q}$$

$$< (h^{\delta(1-\varepsilon)}+1)^n \cdot e^{2h^{1+\delta}} \cdot (C_{25} h^{\delta(1-\varepsilon)})^{h^{n\delta}} e^{-h^q}$$

$$\leq e^{-\frac{1}{2}h^q}$$

and so $|\theta_j| < e^{-\frac{1}{2}h^q} + |g^{(j)}(0)|$

which gives $\log |\theta_j| < -\frac{1}{2}h^q/\log h$

We now define Δ as the Vandermonde determinant of order $(L+1)^n$ with $\theta_r^{\,j}$ in the $(r+1)$ -th row and $(j+1)$ -th column. Thus $\Delta = \prod_{0 \leq r<s\leq R} (\theta_s - \theta_r)$.

From lemma 7 and the assumption that $\log a_1, \ldots, \log a_n$ are linearly independent over the rationals we have

$$|\Delta| \geq (C_{26}{}^{-L})^{(R+1)^2}$$

and thus

$$\log |\Delta| \geq -C_{27} L (L+1)^{2n}$$
$$\geq - C_{27} h^{\delta(2n+1)}$$

We now derive an upper bound for $\log |\Delta|$. We know that at least one of the p_r is non-zero. Assume $p_i \neq 0$ then

$$\Delta = p_i^{-1} \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & \theta_1 & \cdots & \theta_1^R \\ \vdots & \vdots & & \vdots \\ \theta_0 & \theta_1 & \cdots & \theta_R \\ \vdots & \vdots & & \vdots \\ 1 & \theta_R & \cdots & \theta_R^R \end{vmatrix} = p_i^{-1} \Delta^*$$

as the determinant of a matrix is unchanged if a multiple of one row is added to another row. The i^{th} row of Δ^* is obtained from the i^{th} row of Δ by multiplying first by p_i and then adding p_j times the j^{th} row to the i^{th} row for $j = 0, \ldots R$, $j \neq i$.

As $|\theta_r| < C_{28} k$ and $\log |\Theta_j| < -\frac{1}{2}h^q/\log h$ we have

$$\log |\Delta| \leq \log p_i^{-1} + \log ((L+1)^n !)$$
$$+ R \log (C_{28}k) - \tfrac{1}{2} h^q/\log h$$

and as $(L+1)^n ! \leq k^{nk^n} \leq h^{\delta n h^{\delta n}}$ and $q > \delta(2n+1)$ we have that

$$\log|\Delta| \leq -\tfrac{1}{4}h^q/\log h \quad \text{if } h \text{ is sufficiently large.}$$

But comparing this with our previous estimate for log $|\Delta|$ we see that we have

$$-C_{27}h^{\delta(2n+1)} \leq -\tfrac{1}{4}h^q/\log h$$

But we have that $q > \delta(2n+1)$ by assumption and thus we have a contradiction. The theorem is complete.

N.B. The constant C referred to in Theorem 3 is effectively computable. The calculation of the constant has been omitted from the proof. We have instead assumed our estimates hold for h sufficiently large. We can, however, determine those values explicitly.

CHAPTER III

CONSEQUENCES OF A. BAKER'S THEOREM

3.1 Applications of Baker's theorem

In this chapter I will briefly outline some of the improvements and consequences of A. Baker's theorem.

"Linear Forms in the Logarithms of Algebraic Numbers" has evolved in four stages. The third stage [20] gives us the inhomogeneous form of Theorem 3. It serves to settle further questions in the field of transcendental numbers. I will discuss this in more detail shortly. The fourth version [21] of the original theorem was intended to sharpen the estimates that were likely to arise in the solution, obtained with Theorem 3, of certain Diophantine equations. Its statement follows.

Theorem 7. If $a_1,\ldots,a_n$ $n \geq 2$ are non-zero algebraic numbers with heights and degrees less than H and d respectively and if rationals integer $b_1,\ldots,b_n$ exist, with absolute values at most T such that

$$0 < |b_1 \log a_1+\ldots+b_n \log a_n| < e^{-\delta T}$$

where $\log a_i$ denotes the principal value of the logarithm and $0 < \delta \leq 1$ then

$$T < (4^{n^2} \delta^{-1} d^{2n} \log H)^{(2n+1)^2}$$

To illustrate the efficacy of theorem 7 A. Baker in conjunction with H. Davenport used it to prove that the only solutions in positive integers x, y, z of the equations $3x^2-2=y^2$, $8x^2-7=z^2$ are given by $x=1$ and $x=11$. [22]

A. Baker has succeeded in employing theorem 3 or a variation of it in the determination of upper bounds for the integral solution of the following diophantine equations, $y^2 = ax^3 + bx^2 + cx + d$ [23], more specifically $y^2 = x^3 + k$ [24], more generally $y^m = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_o$ [25] and finally $f(x,y) = m$ [26] where f is an irreducible binary form of degree at least 3. The preceding diophantine equations are assumed to have integral coefficients. While Thue, Siegel or Roth had worked with the diophantine equations mentioned and had proved that the equations could have only a finite number of integral solutions their methods shared at least one property, non-effectiveness. The importance of the effectively computable constant C in theorem 3 is now brought to light. It is the effective nature of theorem 3 that allows Baker to exhibit explicit upper bounds for the integral solutions of the aforementioned equations. This in principle reduces the above problems to a finite amount of computation. The bounds obtained however, are generally astronomical and without refinement they leave computation that can't feasibly be done even with the aid of a computer. By following Baker's method, adapting it to the specific case under consideration and employing certain lemmas from the study of diophantine approximation or continued fractions the bounds can, however, be reduced to a more practical size.

This was done, for example, by W. J. Ellison, F. Ellison, P. Pesek, C. C. Stahl and D. S.Stall [27] in determining the integral solutions of the equation $y^2 = x^3-28$.

As a particular consequence of Baker's work on the diophantine equation $f(x,y) = m$, where f is an irreducible binary form with integer coefficients and degree at least 3, we have the first effective improvement on the accuracy with which a rational number can approximate a real algebraic number since Liouville's theorem of 1844.

Finally we have that theorem 3 can be used to determine an upper bound for $d \varepsilon Z^+$ where $Q(\sqrt{-d})$ is an imaginary quadratic field of class number 1 or with the added restriction $d \not\equiv 3 \pmod 8$ class number 2. [28]

This settles the conjecture of Gauss that $Q(\sqrt{-d})$ with $d = 1,2,3,7,11,19,43,67,163$ are the only imaginary quadratic fields of class number 1. H. Stark and K. Heegner also proved the above conjecture for class number 1. Their method involved the use of elliptic modular functions. Their proof has been adapted by Kenku and Weinberger to the case of class number 2. Baker uses Theorem 7 to show that if $d \not\equiv 3 \pmod 8$ then $d < 10^{500}$ for class number 2. The cases where $d < 10^{500}$ were tested by Ellison, Pesek, Stall and Lunnon [29] and the fields of class number 2 were then given by $d = 5, 6, 10, 13, 15, 22, 37, 58$.

3.2 An effective improvement of Liouville's theorem.

While Liouville's theorem has gone through many improvements culminating in the Thue-Siegel-Roth theorem these improvements are all non-effective. We invoke the following theorem of A. Baker, [26] which is a development of the theorem 3, to establish the first generally effective improvement of Liouville's theorem.

Theorem 8. If $f(x,y) = m$ is an irreducible binary form with integer coefficients and degree $n \geq 3$, $m \varepsilon Z^+$ then all integer solutions x,y of $f(x,y) = m$ satisfy $\max(|x|,|y|) < Ce^{(\log m)^q}$ where $q > n$ and C is an effectively computable number depending only on n, q and the coefficients of f.

This enables us to prove theorem 9.

Theorem 9. If α is an algebraic number with degree $n \geq 3$, then there exists a constant $C = C(\alpha,q) > 0$, where $q \geq n$, such that

$$|\alpha - \frac{P}{S}| > \frac{C\, e^{(\log S)^{1/q}}}{S^n}$$

where $P/S \varepsilon Q$ and $S > 0$.

Pf: We let $f(x)$ be the minimal polynomial of α. We then have that

$$|f(\tfrac{P}{S})| = |f(\tfrac{P}{S}) - f(\alpha)| = f^{(1)}(c)|\alpha-\tfrac{P}{S}| < C_1(\alpha)|\alpha-\tfrac{P}{S}|$$

where $C_1(\alpha)$ is effectively computable by the mean value theorem. We now note that $m = S^n |f(\frac{P}{S})|$ is an irreducible binary form of degree $n \geq 3$ with integer coefficients and thus by theorem 8 we have that

$$* \qquad \frac{m}{S^n} C_1^{-1}(\alpha) < |\alpha - \frac{P}{S}| \text{ where}$$

$$S < C_2(\alpha,q)\ e^{(\log m)^q}.$$

Thus

$$m > e^{(\log S - \log C(\alpha,q))^{\frac{1}{q}}} > C_3(\alpha,q)\ e^{(\log S)^{\frac{1}{q}}}$$

where $C_3(\alpha,q)$ is computable. Substituting for m in * we have our desired inequality.

Despite the very powerful result we used, our inequality has not been very significantly improved from Liouville's theorem. We would like an effective equivalent of the Thue-Siegel-Roth theorem.

3.3 The integral solutions of $y^2 = x^3 - 28$

We have, from the result of Baker mentioned earlier [24], that all integral solutions of $y^2 + k = x^3$ satisfy $\max\{|x|,|y|\} \leq \exp\{10^{10}|k|^{10^4}\}$. Thus with a finite amount of computation we could discover all the integral solutions of $y^2 = x^3-28$. We can simplify the computation by the following method [27] which I shall outline.

We reduce $y^2 + 28 = x^3$ to the following.

Thue equations: (i) $\pm 4 = 3X\,Y + XY^2 - Y^3$

(ii) $X^3 - 12XY^2 - 12Y^3 = \pm 1$

by considering $y^2+28=(y+2\sqrt{-7})(y-2\sqrt{-7})$ in the field $Q(\sqrt{-7})$. We note that a prime in $Q(\sqrt{-7})$ which divides $(y+2\sqrt{-7})$ and $(y-2\sqrt{-7})$ must be either $\sqrt{-7}$, $\frac{1+\sqrt{-7}}{2}$, $\frac{1-\sqrt{-7}}{2}$. This yields the above 2 equations. Case (i) gives the two solutions $x = 37$, $y = \pm 225$. We will use Baker's result [21] to deal with case (ii).

We let $f(X,Y) = X^3-12XY^2-12Y^3$ and we work in $Q(\theta)$ where $f(\theta,1) = 0$. An integral basis of $Q(\theta)$ is $\{1, \theta, \theta^2/2\}$ and a pair of fundamental units are $\eta_1 = -7-4\theta + 3\theta^2/2$ and $\eta_2 = 11 + \theta - \theta^2$. Then if we have $a, b \in Z$ such that $f(a,b) = \pm 1$ then $(a-\theta^{(1)}b)(a-\theta^{(2)}b)\cdot(a-\theta^{(3)}b)=\pm 1$, where the bracketed superscript indicates a conjugate. We thus have that $(a-\theta^{(1)}b)$ is a unit and so $=\pm\eta_1^{d_1}\eta_2^{d_2}$ for $d_1, d_2 \in Z$. It is this representation that is crucial for we then have

$$\log|a-\theta^{(i)}b| = d_1 \log|\eta_1^{(i)}| + d_2 \log|\eta_2^{(i)}|$$

for $1 \leq i \leq 3$. Letting $H = \max\{|d_1|, |d_2|\}$ we are able to deduce that

$$\left| d_1 \log\left|\frac{\eta_1^{(j)}}{\eta_1^{(k)}}\right| + d_2 \log\left|\frac{\eta_2^{(j)}}{\eta_2^{(k)}}\right| - \log|m| \right| < e^{-H\cdot(.404)}$$

where $m = \frac{\theta^{(\ell)}-\theta^{(j)}}{\theta^{(k)}-\theta^{(\ell)}}$ for $\{\ell,j,k\} = \{1,2,3\}$ and for $H > 20$.

To apply Baker's theorem, theorem 7, we calculate the heights and degrees of $\dfrac{\eta_1^{(j)}}{\eta_1^{(k)}}$, $\dfrac{\eta_2^{(j)}}{\eta_2^{(k)}}$ and m.

We have that the maximum height is 27236 and the maximum degree is 6. From theorem 7 we then conclude that $H < 10^{563}$. This is an immense improvement over our initial inequality $\max\{ |x|,|y| \} \leq \exp\{ 10^{10}|28|^{10^4}\}$ but it is still impractical. To reduce our upper bound for H we use the following lemma of Davenport. This result from diophantine approximation theory was also used in [22] . Accordingly I will quote it.

<u>Lemma 8</u>. If θ and β are given real numbers and $m,B > 6$ are given integers and if p,q are integers such that $1 \leq q \leq BM$, $|\theta q-p| \leq 2(BM)^{-1}$ then if $||q\beta|| \geq 3B^{-1}$ there is no solution of the inequality $|b_1\theta + b_2-\beta| \leq K^{-|b_1|}$ in integers b_1, b_2 with $\log(B^2M) / \log K \leq |b_1| \leq M$.

By letting $\theta = \log\left|\dfrac{\eta_1^{(j)}}{\eta_1^{(k)}}\right| \Big/ \log\left|\dfrac{\eta_2^{(j)}}{\eta_2^{(k)}}\right|$ and $\beta = \dfrac{\log|m|}{\log\left|\dfrac{\eta_2^{(j)}}{\eta_2^{(k)}}\right|}$

$K = e^{0.404}$, $M = 10^{563}$ and $B = 10^{33}$ and applying our lemma twice we find $|d_1| \leq 44$. This is now an easily handled upper bound and we can determine all integral solutions of $X^3-12XY^2-12Y^3 = \pm 1$. These in turn give the integral solutions $(x,y) = (4, \pm 6)$ and $(8, \pm 22)$. Combining these with

$(x,y) = (37,\pm 225)$ from case (i) we have all the integral solutions to $y^2 = x^3-28$.

3.4 The inhomogeneous form of theorem 3.

The final topic I will discuss in this report is the following generalization of Baker's theorem. [20]

Theorem 10. If $a_1,\ldots,a_n$ and $b_o, b_1,\ldots,b_n$ are non-zero algebraic numbers, then

$$| b_o + b_1 \log a_1 + \ldots + b_n \log a_n | > Ce^{-(\log H)^q}$$

where $q > n + 1$, d and H are respectively the maximum of the degrees and heights of $b_o, b_1,\ldots,b_n$ and $C = C(n, a_1,\ldots,a_n, q, d)$ is an effectively computable number.

N.B. The theorem holds for any determination of the logarithm of a_i, $i = 1,\ldots,n$ but the constant C depends on those determinations.

Baker's method of proof of theorem 10 follows quite closely his proof of theorem 3. It applies to the case $b_o=0$ if we strengthen out hypotheses slightly. We thus can state the following improvement of theorem 3.

Theorem 11. If $a_1,\ldots,a_n$ and $b_1,\ldots,b_n$ are non-zero algebraic numbers with either $\log a_1,\ldots,\log a_n$ or $b_1,\ldots,b_n$ linearly dependent over the rationals then

$$| b_1 \log a_1 + \ldots + b_n \log a_n| > Ce^{-(\log H)^q}$$

where $q > n$, d and H are respectively the maximum of the degrees and heights of $b_1, \ldots b_n$ and $C = C(n, a_1, \ldots a_n, q, d) > 0$ is an effectively computable number.

As with theorem 3, theorem 10 gives us two substantial results in the theory of transcendental numbers. We have that if $a_1, \ldots, a_n$ and $b_1, \ldots b_n$ are non-zero algebraic numbers then $b_1 \log a_1 + \ldots + b_n \log a_n$ is transcendental or zero. From this we may conclude that $\pi + \log a = i \log(-1) + \log a$ is transcendental if a is a non-zero algebraic number. We thus have $\int_0^1 \frac{dx}{1+x^3} = \frac{1}{3}(\log 2 + \frac{\pi}{\sqrt{3}})$ is a transcendental number and this settles a question raised by Siegel. [13]

We can also say that if $a_1, \ldots, a_n$, b_o, $b_1, \ldots, b_n$ are non-zero algebraic numbers then $e^{b_o} a_1^{b_1} \ldots a_n^{b_n}$ is transcendental. If it were algebraic we would have a contradiction to theorem 10.

Bibliography

[1] A. Baker, Linear Forms in the Logarithms of Algebraic Numbers, Mathematika 13 (1966) , pp.204-216.

[2] G. H. Hardy and E. M. Wright, The Theory of Numbers, Oxford, 3rd edition 1954.

[3] J. Liouville, Sur les classes très étèndues de quantités dont la valeur n'est ni algébrique ni même réductible à des irrationelles algébriques, Comptes Rendus Acad. Sci. Paris 18 (1844), 883-885, 910-111 ; Journal Math. Pures et Appl., 16 (1851), 133-142.

[4] K. Mahler, On the Approximation of π . Proc. Akad. Wetensch. Ser. A. 56, (1953), pp.30-42.

[5] J. Lipman, Transcendental Numbers, Queen's Papers in Pure and Applied Mathematics - No. 7, 1966.

[6] C. Hermite, Oeuvres Volume III, pp.150-181.

[7] I. Niven, Irrational Numbers, The Carus Mathematical Monographs, Number 11, 1956.

[8] F. Lindemann, Ueber die Zahl π, Math. Annalen, 20 (1882), pp.213-225.

[9] R. Kuzmin, On a New Class of Transcendental Numbers, Izvestiya Akad. Nauk SSSR Ser. matem., 7 (1930), pp.585-597.

[10] A. Gelfond, On Hilbert's Seventh Problem, Doklady Akad. Nauk SSSR, 2 (1934), pp.1-6.

[11] Th. Schneider, J. reigne angew. Math., 172 (1935) pp. 65-69.

[12] A. Gelfond, Transcendental and Algebraic Numbers, translated by Leo F. Boron, Dover, New York, (1960).

[13] C. L. Siegel, Transcendental Numbers, Annals of Math. Princeton (1949).

[14] A. O. Gelfond and Yu V. Linnik, Elementary Methods in Analytic Number Theory, Rand McNally and Company, (1965), pp.232-239.

[15] Th. Schneider, Ein Satz uber ganzwertige Funktionen als Prinzip fur Transzendenzbeweise. Math. Annalen 121, pp.131-140 (1949-50).

[16] Serge Lang, Transcendental points on group varieties, Topology, (1963).

[17] Th. Schneider, Einfuhrung in die Transzendenten Zahlen, Berlin, Springer-Verlag, (1957).

[18] A. Brumer, On the units of algebraic number fields, Mathematika 14, pp.121-124, (1967).

[19] A. Baker, Linear Forms in the Logarithms of Algebraic Numbers, II, Mathematika 14 (1967), pp.102-107.

[20] A. Baker, Linear Forms in the Logarithms of Algebraic Numbers, III, Mathematika 14 (1967) pp.220-228.

[21] A. Baker, Linear Forms in the Logarithms of Algebraic Numbers, IV, Mathematika 15 (1968), pp.204-216.

[22] A. Baker and H. Davenport, Quarterly J. Math. Oxford (2), 20 (1969), pp.129-137.

[23] A. Baker, The Diophantine Equation $y^2 = ax^3+bx^2+cx+d$ J. London Math. Soc. 43, pp.1-9.

[24] A. Baker, Contributions to the theory of Diophantine Equations, II The Diophantine Equation $y^2 = x^3+k$. Phil. Trans. Royal Soc. London A 263, pp.193-208.

[25] A. Baker, Bounds for the solutions of the hyperelliptic equation, Proc. Comb. Phil. Soc. (1969), 65, pp.439-444.

[26] A. Baker, Contributions to the theory of Diophantine Equations, I On the representation of integer by binary forms. Phil. Trans. Royal Soc. London A 263, pp.173-191.

[27] W. S. Ellison, F. Ellison, J. Pesek, C. E. Stahl and D. S. Stall, The Diophantine Equation $y^2+k=x^3$, J. Number theory 4, (1972) pp.107-117.

[28] A. Baker, A remark on the class number of quadratic fields, Bull. London Math. Soc. 1 (1969), pp.98-102.

[29] W. J. Ellison, J. Pesek, D. S. Stall and W. F. Lunnon, A postscript to a paper of A. Baker, Bull. London Math. Soc., 3 (1971), pp.75-78.