Transcendental Numbers

by

Joseph Lipman

QUEEN'S PAPERS IN PURE AND
APPLIED MATHEMATICS – NO. 7

QUEEN'S UNIVERSITY, KINGSTON, ONTARIO
1966

PREFACE

These notes are intended to serve as an introduction to the subject of transcendental numbers. They should be accessible to the advanced undergraduate who knows what is meant by an algebraic number field K of finite degree over the rational field $\mathbb{Q}$, and by the norm from K to $\mathbb{Q}$ of an algebraic integer in K. The maximum principle for entire analytic functions of a complex variable is used once, in §7. The proof of one lemma (§11) is based on complex integration; however it is indicated there that the use of integration can be avoided. Power series with complex coefficients appear several times. Thus, while a first course in complex analysis is a desirable prerequisite, it is not absolutely essential.

The author gratefully acknowledges the support provided by NSF contract GP-4249 at Purdue University.

CONTENTS

Errata

p. 1: In equation (1) and throughout the proof of Theorem 1 read ''n'' for ''r''.

p. 10: (second line from bottom): $+ \frac{1}{2!} \beta^2 H d^2 \gamma^{d-2} + \ldots$

p. 11: for ''$\left(\frac{c_\alpha}{2\gamma}\right) d$'' read ''$\left(\frac{c_\alpha}{2\gamma}\right)^d$''

p. 12 (Theorem 4): for ''$\overline{|\alpha|}$'' read ''$\underline{|\alpha|}$''

(line 7): Read ''$\overline{|\beta|}$ and $H(\beta)$''

p. 23: (Theorem 7): $f_\nu^{(k)}(z_\lambda) = \left.\frac{d^k f_\nu(z)}{dz^k}\right|_{z=z_\lambda}$

p. 26: (Lemma 1b): $N > sM$

p. 31: (line 1, line 3): for ''$F^{(g)}$'' read ''$F^{(q)}$''

p. 37 (line 3): $g^{(k)} = \frac{d^k g}{dz^k} =$

p. 38 (line 11): for all i, $P_i \ll B(1 + X_1 + X_2 + \ldots + X_n)^{h+1}$

p. 40 (Theorem 9), and thereafter when called for by the context:

for ''Q'' read ''$\mathbb{Q}$'' (which denotes the field of rational numbers)

p. 52: $f = \frac{1}{1 - (\frac{g-1}{g})^{1/m+1}}$

p. 55 (line 1): for ''$|\bar{\alpha}|$'' read ''$\overline{|\alpha|}$''

(second last line): for ''prove'' read ''pave''

p. 61 (formula 10)
p. 62 (line 3) } for ''log zc'' read ''log 2c''

p. 65 (second last line): $\max_{\mu,q} |b_{\mu q}|$

p. 69 (third last line): for ''p_i'' read ''p_1''

p. 76 (line 5): for ''$\sum_{h=1}^{n}$'' read ''$= \sum_{h=1}^{n}$''

p. 77 (line 4): $\eta = |\frac{w}{2}|$

(line 5): for ''c_z'' read ''c_2''

Spelling

p. vi
p. 14 } Gelfond
p 44, 47 independence
p. 45
p. 47 interesting

p. 9 algebraic

p. 50 commas in formula

p. 53 for "(log" read "log"

p. 73 $Q_5(\ )$

INTRODUCTION

The subject of transcendental numbers was launched in 1844 with Liouville's discovery of the theorem which is now known as the Liouville approximation theorem. Liouville's theorem enabled him to give the first proof that there exist transcendental numbers, although it should be pointed out, Legendre had suspected half a century earlier that this might be the case.

An entirely different approach to the subject was taken by George Cantor, who published in 1874 his spectacular (at that time) result on the countability of the set of algebraic numbers. Along with his discovery that the complex numbers are uncountable, this showed that almost every number is transcendental. Cantor's methods were entirely non constructive, being based on his theory of countable and transfinite numbers, and there was some doubt among his contemporaries as to their validity. Indeed, the ridicule of Kronecker and his followers had its effect on Cantor's health. Nowadays, of course, Cantor's theorem is universally accepted.

It is one thing to construct transcendental numbers; it is another to investigate the transcendence of a specific number like e or π or Euler's constant.

In the former situation, the mathematician has considerable freedom; in the latter, as E.T. Bell puts it, it is the mathematician, and not the suspect, who takes orders. Thus it was that Hermite's proof for the transcendence of e (1873) and Lindemann's proof for the transcendence of π (1882) were considered to be among the greatest achievements of nineteenth century mathematics. Even Kronecker indicated, as best he could, his grudging admiration. He is reported to have asked Lindemann "of what use is your beautiful proof, since irrational numbers do not exist?"

In the ensuing years, great interest was attached to rendering the proofs for the transcendence of e and π as elementary (i.e. as free from the methods of analysis) as possible. Dozens of proofs were published, each a little more dependent on the particular properties of the exponential function than the one before, and therefore a little less capable of generalization. Over a long period, few new results were obtained.

New impetus was given to the subject by A. Gelfand and C.L. Siegel around 1930. Their ideas led to the independent solutions by Gelfand and Th. Schneider (1934) of Hilbert's seventh problem, now known as the Hilbert-Gelfand-Schneider theorem: if α and β are algebraic numbers, $\alpha \neq 0,1$, β irrational, then α^{β} is transcendental.*

---

* The question of the nature of logarithms of rational numbers with respect to a rational base was raised by Euler as early as 1748 [Introductio in Analysin Infinitorum]. Euler even states that such logarithms are either rational or transcendental, but the meaning of the word "transcendental", as used in those times, seems unclear.

Among the important results of the last thirty years have been the classification of transcendental numbers by Mahler (1932) and later Koksma (1939); Schneider's theorems on elliptic functions and Abelian integrals (1941) and the Thue-Siegel-Roth theorem on the approximation of algebraic numbers by rationals (1955).

At present, there remain many unsolved problems. For instance, it is not known whether any of the following numbers are transcendental (or even irrational): Euler's constant, $\zeta(2n+1)$, $\Gamma(x)$ for algebraic $x \neq 0, \pm 1, \pm 2, \ldots$. Since e and π are transcendental, at least one of the numbers $e\pi$, $e+\pi$, is transcendental; but neither has ever been proved irrational. Nor is it known whether or not e and π are algebraically dependent.

These are not the most important unsolved problems cf. [Schneider, 1], but they are interesting and easy to state. The answer to any of these problems seems remote. But there is always hope: Hilbert predicted that the transcendence of $2^{\sqrt{2}}$ would be settled only long after the Riemann hypothesis, and he was proven wrong.

As yet, there seems to be no aspect to the study of transcendental numbers which could be described as a general theory. However, there are some methods which are quite powerful in their own way, and we have tried in these notes to describe three of them.

Part I, entitled Approximation Methods, deals with generalizations and improvements of Liouville's theorem, with applications to the construction of transcendental numbers.

Part II, entitled Schneider's Theorem, is devoted to the proof of a theorem discovered by Schneider in 1949 [Schneider, 2] from which many results are obtained which previously required separate treatment; the transcendence of e and π, the Hilbert-Gelfand-Schneider theorem, and Schneider's results on elliptic functions, to name three. We do not give the theorem in its fullest generality (an account in English is given in [LeVeque, 2]) but rather in the simpler version which appears in [Schneider, 1] and whose proof illustrates all the principles of the proof of the general theorem.

Part III, entitled Algebraic Independence covers the theorem of Lindemann (1882) on the algebraic independence of $e^{\omega_1}, e^{\omega_2}, \ldots, e^{\omega_m}$ when $\omega_1, \omega_2, \ldots, \omega_m$ are linearly independent algebraic numbers. This theorem (from which, a propos, Lindemann obtained the transcendence of π as an immediate corollary) was remarkable in that it was for many years an isolated example of a theorem about algebraic independence. Not until Siegel's investigations of Bessel functions (1929) did there appear a method of any generality for proving algebraic independence. Siegel's method has proved fruitful, and is still more or less the only tool available in the study of algebraic independence. Unfortunately, its domain of application is rather limited.

Using Siegel's method as a starting point, A.B. Sidlovskii proved a theorem (1954) which, in some cases, reduces the question of algebraic independence of transcendental numbers over the field of rationals to the much simpler question of algebraic independence of transcendental functions over the field of rational functions. In §12.2, we state without proof the basic theorem, as well as a generalization of Lindemann's theorem which is an easy consequence.

The proof (following Mahler) of Lindemann's theorem which we give in Part III is by no means the simplest known. However, it does illustrate some of the principles of Siegel's method, and it has the added advantage of giving a transcendence measure for e^{ω} when ω is algebraic.

I. APPROXIMATION METHODS

§1. The Theorem of Liouville.

1.1 Theorem 1 (Liouville) Let α be an algebraic number of degree $n \geq 2$. There is a positive number c, depending only on α, such that the inequality

$$\left|\alpha - \frac{p}{q}\right| > \frac{c}{|q|^r} \qquad (1)$$

holds for all pairs of rational integers p, q $(q \neq 0)$.

Proof. Fix an irreducible polynomial $f(X)$ of degree n, with rational integer coefficients, such that $f(\alpha) = 0$. The derivative $f'(\alpha)$ does not vanish; therefore the definition of the derivative as a limit shows that for some $\delta > 0$

$$\left| \frac{f(p/q) - f(\alpha)}{(p/q) - \alpha} \right| < 2\ |f'(\alpha)|$$

whenever $|(p/q) - \alpha| \leq \delta$. Now $f(p/q) = b/q^r$ with $b \neq 0$ a rational integer, and so

$$|f(p/q) - f(\alpha)| = |f(p/q)| \geq \frac{1}{|q|^r}$$

If $|(p/q) - \alpha| \leq \delta$, we have therefore

$$\frac{1}{|q|^r} < 2|f'(\alpha)| \; |(p/q) - \alpha|$$

while if $|(p/q) - \alpha| > \delta$ then certainly $|(p/q) - \alpha| > \frac{\delta}{|q|^r}$
In either case (1) holds with

$$c = \min\left(\frac{1}{2|f'(\alpha)|}, \delta\right) \qquad \text{q.e.d.}$$

We shall give another proof in §2.

1.2 Liouville's theorem enables us to construct transcendental numbers. We define a <u>Liouville number</u> to be a non-rational number α such that for no pair $c > 0$, $n \geq 2$ is it true that the inequality (1) of theorem 1 holds for all rational p/q. It is evident that a Liouville number is real. Liouville's theorem shows that any Liouville number is transcendental.

We shall have examples of Liouville numbers in a moment, but first we derive an alternative definition. If ξ is a Liouville number, then for any positive (rational) integer N, there are actually infinitely many rationals p/q such that $0 < |\xi-(p/q)| < 1/|q|^N$; otherwise, if $c > 0$ were sufficiently small, we would have $0 < |\xi-(p/q)| < c/|q|^N$ for all p/q. Suppose, conversely, that ξ is a number with the property:

(P) For any positive integer N, there is a rational number p/q with $q \geq 2$ such that

$$0 < |\xi-(p/q)| < 1/q^N$$

Given $c > 0$, $n \geq 2$, choose m so that $(1/2^m) < c$, and set $N = n+m$; then for suitable p/q

$$|\xi-(p/q)| < \frac{1}{q^N} \leq \frac{1}{2^m} \cdot \frac{1}{q^n} < \frac{c}{q^n}$$

Moreover, ξ is not rational; for if ξ were rational, say $\xi = r/s$ $(s > 0)$, and if $0 < |\xi-(p/q)|$, then choosing N so that $2^{N-1} > s$, we would get

$$0 < \left| \frac{r}{s} - \frac{p}{q} \right| = \left| \frac{qr-ps}{sq} \right| \geq \frac{1}{sq} > \frac{1}{2^{N-1}q} \geq \frac{1}{q^N}$$

and so ξ could not enjoy the property (P) . We see then that (P) is a characterization of Liouville numbers.

An example of a Liouville number is

$$\zeta = \sum_{k=0}^{\infty} \frac{1}{10^{k!}}$$

To see this, set

$$\sum_{k=0}^{N} \frac{1}{10^{k!}} = \frac{p_N}{10^{N!}}$$

Then

$$0 < \left|\zeta - \frac{p_N}{10^{N!}}\right| = \left|\sum_{k=N+1}^{\infty} \frac{1}{10^{k!}}\right| < \frac{1}{10^{(N+1)!}}\left|\sum_{k=0}^{\infty} \frac{1}{10^{k}}\right| < \frac{1}{(10^{N!})^N}$$

Thus ζ satisfies (P) .

A similar argument shows that if $a \geq 2$ is a positive (rational) integer, $b_0 < b_1 < b_2 < \ldots$ is a sequence of positive integers with $\overline{\lim\limits_{k\to\infty}} (b_{k+1}/b_k) = \infty$, and $\{e_k\}_{k=0,1,2\ldots}$ is a <u>bounded</u> sequence of positive integers, then the numbers

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{a^{b_k}} \quad \text{and} \quad \sum_{k=0}^{\infty} \frac{e_k}{a^{b_k}}$$

are Liouville numbers.

1.3 Not all real transcendental numbers are Liouville numbers; in fact it can be shown that the Liouville numbers form a set of Lebesgue measure zero. The number π is transcendental (cf. §9) but not a Liouville number: Mahler has shown [Mahler, 1] that $|\pi-(p/q)| > q^{-42}$ when $q \geq 2$. We shall see in Ch. III that $e = 2.71828\ldots$ is a non-Liouville transcendental number. Suffice it for now to mention the following interesting result, also due to Mahler, concerning numbers whose decimal expansion can be described by an explicit formula:

Let $f(X)$ be a non-constant polynomial with rational coefficients such that $f(k)$ is a positive integer whenever k is. Then the number represented by the infinite decimal

$$0.f(1)f(2)f(3)\ldots$$

(formed by writing the numerals for $f(1), f(2), \ldots$ in succession after the decimal point) is transcendental, but is not a Liouville number. [Mahler, 2].

For example, when $f(X) = X$ we get the number 0.1 2 3 4 5 6 7 8 9 10 11 12...; when $f(X) = \dfrac{X^2 + X}{2}$

we get 0.1 3 6 10 15 21 ...

§2. A Generalization of Liouville's Theorem

2.1 Liouville's theorem can be thought of as describing the size of $|P(\alpha)|$ for algebraic α, where P is the polynomial $qX-p$ with rational integral coefficients p,q. One may then ask about $|P(\alpha)|$ for arbitrary polynomials P with rational integer coefficients, say $P = a_oX^d + a_1X^{d-1} + \ldots + a_d$.

For such a P, we define the <u>height</u> $H = H(P)$ by

$$H = \max_j |a_j| \qquad (j = 0,1,\ldots,d)$$

An estimate for $|P(\alpha)|$ is given by the following theorem (of which theorem 1 is a corollary).

<u>Theorem 2</u>. Let α be an algebraic number, of degree n, and let $P = a_oX^d + a_1X^{d-1} + \ldots + a_d$ be a polynomial with rational integer coefficients, $P(\alpha) \neq 0$. Let $H(P)$

be the height of P . Then there is a positive number c_α , depending only on α , such that

$$|P(\alpha)| > \frac{(c_\alpha)^d}{H(P)^{n-1}} \qquad (2)$$

<u>Proof</u>. Let the conjugates of α be $\alpha_1(=\alpha), \alpha_2, \alpha_3, \ldots, \alpha_n$; and let $r = \max_i |\alpha_i|$. Clearly

$$\left|\frac{P(\alpha_i)}{H}\right| \leq 1+r+r^2+\ldots+r^d \leq (r+1)^d$$

Let b be a positive rational integer such that $b\alpha$ is an algebraic integer (such b always exist). Then $b^d P(\alpha_i)$ is a non-zero algebraic integer for $1 \leq i \leq n$, so that $N = b^d P(\alpha_1).b^d P(\alpha_2). \ldots .b^d P(\alpha_n)$ is a non-zero rational integer, and

$$H^{n-1}|P(\alpha)| = \left|\frac{N}{b^{nd}}\left(\frac{H}{P(\alpha_2)}\right)\left(\frac{H}{P(\alpha_3)}\right)\ldots\left(\frac{H}{P(\alpha_r)}\right)\right| \geq \frac{1}{b^{nd}} \cdot \frac{1}{(r+1)^{d(n-1)}}$$

Thus (2) holds with $c_\alpha = \frac{1}{2b^n} \cdot \frac{1}{(r+1)^{n-1}}$ q.e.d.

2.2 In the special case $d = 1$, i.e. when P has the form $qX-p$, and $H = \max(|p|,|q|)$, theorem 2 states

$$|q\alpha-p| > \frac{c_\alpha}{H^{n-1}} \tag{3}$$

We may deduce Liouville's theorem (theorem 1) as follows:

If $H = |q|$, then (3) gives $|\alpha-(p/q)| > c_\alpha/|q|^n$.

If $|\alpha-(p/q)| > 1$, then $|\alpha-(p/q)| > 1/|q|^n$.

If $H = |p|$ and also $|\alpha-(p/q)| \leq 1$, then $|p/q|^{n-1} \leq (1+|\alpha|)^{n-1} = \gamma$ (say), and combining this with (3), we get

$$\left|\alpha - \frac{p}{q}\right| > \frac{c_\alpha}{|q|\ |p|^{n-1}} \geq \frac{c_\alpha}{\gamma|q|^n}$$

In any case, theorem 1 holds with $c = \min(1, c_\alpha/\gamma)$.

§3. Another Generalization

3.1 There is a close connection between theorems concerning $|P(\alpha)|$ (cf. theorem 2) and theorems concerning the approximation of α by other algebraic numbers. Let ξ be an algebraic number of degree d with minimum polynomial

$P = a_o X^d + a_1 X^{d-1} + \ldots + a_d$, the a_i being chosen so as to be relatively prime rational integers (i.e. P is irreducible in $\mathbf{Z}[X]$). The height of ξ , $H(\xi)$ is then defined to be the height of P (cf. §2) . $H(\xi)$ is a positive rational integer.

Theorem 3. For any algebraaic number α of degree n , there is a positive number $\bar{c}_\alpha$, depending only on α , such that whenever $\xi \neq \alpha$ is an algebraic number of degree d and height $H(\xi)$, then

$$|\alpha - \xi| > \frac{(\bar{c}_\alpha)^d}{H(\xi)^n} \tag{4}$$

Proof. Since α has only finitely many conjugates, we may assume that ξ is not a conjugate of α . Then $P(\alpha) \neq 0$, where P is the minimum polynomial of ξ , with coefficients a_i chosen as above so that $H(\xi) = H(P) = H$ (say). Theorem 2 gives an inequality

$$|P(\alpha)| > \frac{(c_\alpha)^d}{H^{n-1}} \tag{5}$$

We may also assume that $\beta = |\alpha - \xi| < 1/d$, since otherwise (4) holds with $\bar{c}_\alpha = 1/2$. On comparing (5) with a suitable upper bound for $|P(\alpha)|$, involving β , we will get (4). To find such an upper bound, we use the Taylor expansion.

$$P(\alpha) = P(\xi) + (\alpha-\xi)P'(\xi) + \frac{1}{2!}(\alpha-\xi)^2P''(\xi)+\ldots+\frac{1}{d!}(\alpha-\xi)^dP^{(d)}(\xi)$$

Since, for $0 \leq j \leq d$,

$$P^{(j)}(\xi) = d(d-1)\ldots(d-j+1)a_o\xi^{d-j}+(d-1)(d-2)\ldots(d-j)a_1\xi^{d-j-1} + \ldots+ j!a_j$$

we have

$$\left|\frac{P^{(j)}(\xi)}{Hd^j}\right| \leq |\xi|^{d-j} + |\xi|^{d-j-1} +\ldots+ 1 \leq \gamma^{d-j}$$

where $\gamma = 2 + |\alpha| \geq 1 + (1/d) + |\alpha| > 1+|\xi|$ (since $|\alpha-\xi| < 1/d$). Also, $P(\xi) = 0$. Hence

$$|P(\alpha)| \leq \beta Hd\,\gamma^{d-1} + \frac{1}{2!}\beta^2Hd^2\,{}^{d-2} +\ldots+ \frac{1}{d!}\beta^dHd^d$$

$$\leq H\gamma^d\left(\frac{\beta d}{\gamma} + \left(\frac{\beta d}{\gamma}\right)^2 + \left(\frac{\beta d}{\gamma}\right)^3 +\ldots+ \left(\frac{\beta d}{\gamma}\right)^d\right)$$

$$\leq H\gamma^d . \frac{\beta d}{\gamma} (1 + \frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2^d})$$

$$\leq H\gamma^d . \frac{2\beta d}{\gamma} \tag{6}$$

where the second-last inequality holds because $(\beta d/\gamma) < 1/2$ (recall that $\beta d < 1$ and $\gamma > 2$). Combining (5) and (6), we get

$$|\alpha - \xi| = \beta > \frac{\gamma}{2d\gamma^d H} . \frac{(c_\alpha)^d}{H^{n-1}} > (\frac{c_\alpha}{2\gamma})^d . \frac{1}{H^n}$$

q.e.d.

§4. Construction of Transcendental Numbers

4.1 With the aid of theorem 3 it is now possible to find many new transcendental numbers by generalizing the construction of Liouville numbers given in §1.2. If α is an algebraic number whose conjugates are $\alpha_1 (=\alpha), \alpha_2, \ldots, \alpha_d$ then we set

$$\lceil \alpha \rceil = \max_i |\alpha_i| \ , \quad \lfloor \alpha \rfloor = \min_i |\alpha_i| \quad (i = 1, 2, \ldots, d)$$

<u>Theorem 4</u>. Let α be an algebraic number with $\overline{|\alpha|} > 1$. The power series

$$F(z) = \sum_{\nu=0}^{\infty} \frac{z^{\nu}}{\alpha^{\nu!}}$$

represents an entire function which assumes a transcendental value at every algebraic number $x \neq 0$.

The proof makes use of a relation between the numbers $\overline{|\beta|}$ and $H(\beta)$ associated with an algebraic number β . A rational integer $q > 0$ is called a <u>denominator</u> for β if $q\beta$ is an algebraic <u>integer</u>.

<u>Lemma</u>. Let β be an algebraic number of degree d and height $H(\beta)$, and let q be a denominator for β . Then $H(\beta) \leq 2^d q^d \overline{|\beta|}^d$.

<u>Proof</u>. Let $\beta_1 (=\beta), \beta_2, \ldots, \beta_d$ be the conjugates of β . Then $q\beta_j$ is an algebraic integer for each j , and the polynomial

$$(qX-q\beta_1)(qX-q\beta_2) \ldots (qX-q\beta_d) = \sum_{\nu=0}^{d} b_{\nu} X^{d-\nu} \qquad (7)$$

has rational integer coefficients; it follows immediately that $H(\beta) \leq \max_{\nu} |b_\nu|$. Multiplying out the left hand side of (7), we see that for all ν ,

$$|b_\nu| \leq q^{d-\nu}\binom{d}{\nu}\overline{|q\beta|}^{\nu} \leq 2^d q^d \overline{|\beta|}^{d} \qquad \text{q.e.d.}$$

We turn now to the <u>proof of theorem 4</u>. $F(z)$ is obviously entire. Let $x \neq 0$ be an algebraic number, and set

$$F_m(x) = \sum_{\nu=0}^{m} \frac{x^\nu}{\alpha^{\nu!}}$$

For all sufficiently large ν ,

$$\left| \frac{x^{\nu+1}}{\alpha^{(\nu+1)!}} \right| < \frac{1}{2} \left| \frac{x^\nu}{\alpha^{\nu!}} \right|$$

and it follows that

$$0 < |F(x) - F_m(x)| < 2 \left| \frac{x^{m+1}}{\alpha^{(m+1)!}} \right| \qquad (8)$$

for sufficiently large m .

$F_m(x)$ is algebraic of degree $\leq d = [\mathbb{Q}(x,\alpha) : \mathbb{Q}]$.

If also $F(x)$ is algebraic, then we shall see that theorem 3 gives a lower bound for $|F(x) - F_m(x)|$ which is incompatible with the inequality (8). The final conclusion must be that $F(x)$ is not algebraic.

Suppose then that $F(x)$ <u>is</u> algebraic, of degree n. If p is a common denominator for x and $(1/\alpha)$, then $p^m \cdot p^{m!}$ is a denominator for $F_m(x)$; moreover,

$$\overline{|F_m(x)|} < \sum_{\nu=0}^{\infty} \frac{\overline{|x|}^{\nu}}{\underline{|a|}\,\nu!} = \gamma \quad \text{(say)}$$

and so the lemma gives

$$H_m = H(F_m(x)) < 2^d (p^m \cdot p^{m!})^d \gamma^d \le \gamma_1 p_1^{m!}$$

where $\gamma_1 = 2^d \gamma^d$, $p_1 = p^{2d}$. Let m be so large that $F(x) \neq F_m(x)$ (cf. (8)). Theorem 3, with $\alpha = F(x)$, $\xi = F_m(x)$, $\bar{c}_\alpha = \bar{c}_{F(x)} = c$, shows now that

$$|F(x) - F_m(x)| > \frac{c^d}{(H_m)^n} > \frac{c^d}{(\gamma_1 p_1^{m!})^n} = \frac{\gamma_2}{p_2^{m!}} \qquad (9)$$

where $\gamma_2 = c^d/\gamma_1^n > 0$, and $p_2 = p_1^n$.

Comparison of (9) with (8) leads to the inequality

$$\frac{1}{(|\alpha|^{m+1})^{m!}} > \frac{\gamma_2}{2|x|^{m+1}} \cdot \frac{1}{p_2^{m!}} > (\gamma_3)^{m!}$$

for a suitable positive γ_3 not depending on m . Since $|\alpha| > 1$, such an inequality cannot hold for large m .

q.e.d.

4.2 A more general type of entire function taking transcendental values at algebraic $x \neq 0$ is described in theorem 5 below. The proof, being quite similar to that of theorem 4, will be left as an exercise (except for the following remark:

if $\sum_{\nu=0}^{\infty} \alpha_\nu x^\nu$ is an entire function, then

$\lim_{\nu\to\infty} |\alpha_\nu|^{1/\nu} = 0$, i.e. $\lim_{\nu\to\infty} \frac{\log|\alpha_\nu|}{\nu} = -\infty$) .

<u>Theorem 5</u>. Let K be an algebraic number field, of finite degree over the rationals, and let

$$G(z) = \sum_{\nu=0}^{\infty} \alpha_\nu z^\nu$$

be a power series with non-vanishing coefficients $\alpha_\nu \in K$ such that

1) $\lim\limits_{\nu\to\infty} \dfrac{|\alpha_{\nu+1}|}{|\alpha_\nu|} = 0$

2) There is a sequence $b_0, b_1, b_2, \ldots$ of positive rational integers such that for each ν the numbers $\alpha_0 b_\nu, \alpha_1 b_\nu, \ldots, \alpha_\nu b_\nu$ are algebraic integers (i.e. b_ν is a common denominator for $\alpha_0, \alpha_1, \ldots, \alpha_\nu$) and such that

$$\overline{\lim_{\nu\to\infty}} \frac{-\log|\alpha_{\nu+1}|}{\log b_\nu} = \infty$$

Then $G(z)$ represents an entire function, and $G(x)$ is transcendental for all algebraic $x \neq 0$ such that $\sum_{\nu=0}^{\infty} |\alpha_\nu x^\nu|$ converges (in particular for all algebraic $x = 0$ such that $(1/|x|) > \overline{\lim\limits_{\nu\to\infty}} |\alpha_\nu|^{1/\nu}$).

§5. The Thue-Siegel-Roth Theorem

5.1 The construction of transcendental numbers has depended so far on two things:

a) An approximation theorem setting restrictions on the approximability of one algebraic number by another.

b) The construction of a sequence of algebraic numbers which converges so rapidly to its limit that the restrictions imposed by a) are violated, unless the limit is transcendental.

If the approximation theorem can be improved, new transcendental numbers can be constructed by means of sequences which converge less rapidly than those previously required.

In 1955, K.F. Roth proved the following theorem, known for historical reasons as the Thue-Siegel-Roth theorem.

Theorem 6. If α is an algebraic number, and $\epsilon > 0$, then the inequality

$$|\alpha - \frac{p}{q}| > \frac{1}{q^{2+\epsilon}}$$

holds for all but a finite number of rational numbers p/q.

A more general statement, which improves theorem 3, has been given by LeVeque in [LeVeque, 2].*

Theorem 6'. Let α be an algebraic number. For any $\epsilon > 0$, the number of elements ξ lying in a fixed number field K (of finite degree), and satisfying

$$|\alpha - \xi| < \frac{1}{H(\xi)^{2+\epsilon}}$$

is at most finite.

We do not give the proof, which, though elementary, is long and intricate [see e.g. Leveque, 2]. However we can illustrate our opening remarks by pointing out that theorem 6 gives the transcendence of $\sum_{k=0}^{\infty} \frac{1}{10^{3^k}}$. The proof, left as an exercise, is almost the same as that given in §1.2 to show that $\sum_{k=0}^{\infty} \frac{1}{10^{k!}}$ is a Liouville number, with theorem 6 replacing theorem 1.

Similarly if $a \geq 2$ is a positive (rational) integer, $b_0 < b_1 < b_2 < b_3 < \ldots$ is a sequence of positive integers with $\overline{\lim_{k\to\infty}} \frac{b_{k+1}}{b_k} > 2$, and $\{e_k\}_{k=0,1,2\ldots}$ is a bounded

* A more exotic version of Roth's theorem can be found in S. Lang's book Diophantine Geometry (Interscience, 1962).

sequence of positive integers, then

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{a^{b_k}} \quad \text{and} \quad \sum_{k=0}^{\infty} \frac{e_k}{a^{b_k}}$$

are transcendental. This is clearly an improvement on the corresponding statement in §1.2.

These examples can be generalized as in §4. Using Theorem 6', the reader who has proved theorem 5 (§4) will have no difficulty in checking that for a given algebraic $x \neq 0$, theorem 5 remains true when the assumption $\overline{\lim_{\nu\to\infty}} \frac{-|\log|\alpha_{\nu+1}|}{\log b_\nu} = \infty$ is replaced by the weaker condition

$$\overline{\lim_{\nu\to\infty}} \frac{-\log|\alpha_{\nu+1}|}{\log|b_\nu|} > 2[K(x):\mathbb{Q}] .$$

5.2 Another process of convergence which should be mentioned is that of continued fractions. (See e.g. LeVeque [1] for notation and results). Let

$$\xi = a_0 + \frac{1}{a_1+} \frac{1}{a_2+} \cdots$$

be an <u>infinite</u> continued fraction, the a_k being positive rational integers. ξ is an irrational number. If p_k/q_k

is the k-th convergent to ξ, (i.e. $p_k/q_k = a_0 + \frac{1}{a_1 +} \frac{1}{a_2 +} + \ldots + \frac{1}{a_k}$) then

$$|\xi - \frac{p_k}{q_k}| < \frac{1}{a_{k+1}\, q_k^2} \tag{10}$$

Suppose that the a_k increase so rapidly with k that

$$\overline{\lim_{k\to\infty}} \frac{\log a_{k+1}}{\log q_k} = \infty$$

It follows immediately from (10) (and the fact that $q_k \geq 2$ for $k > 1$) that ξ is a Liouville number.

Here again Roth's theorem gives a better result: if

$$\overline{\lim_{k\to\infty}} \frac{\log a_{k+1}}{\log q_k} > 0 \tag{11}$$

then ξ is transcendental.

Example: If b is a positive integer, let $a_\nu = b^{2^\nu}$. Since $q_0 = 1$, $q_1 = a_1$, and $q_{k+1} = a_{k+1}q_k + q_{k-1}$ $(k > 0)$, condition (11) is easily verified; thus

$$b + \frac{1}{b^2 +} \frac{1}{b^4 +} \frac{1}{b^8 +} \ldots$$

is transcendental.

II. SCHNEIDER'S THEOREM

§6. Statement of the Theorem

6.1 We have already examined, in theorem 5, entire functions $f(z) = \Sigma \alpha_\nu z^\nu$ in which certain arithmetic restrictions were placed on the coefficients α_ν, i.e. on the derivatives of f at 0 (since $\alpha_\nu = (1/\nu!)f^{(\nu)}(0))$.

Schneider's theorem, in the form in which it will appear here, deals with two meromorphic functions f_1 and f_2, and a number of points $z_1, z_2, \ldots, z_m$. Arithmetic conditions are imposed on the derivatives of f_1 and f_2 at each of the points $z_1, z_2, \ldots, z_m$. These restrictions on f_1 and f_2 are, then, of a local nature. Global restrictions are placed on f_1 and f_2 in the form of a limitation on their order of growth (definition below). The theorem states that if all these restrictions are sufficiently numerous, or sufficiently stringent, then f_1 and f_2 are <u>algebraically dependent</u> over the rationals. Put another way: if f_1 and f_2 are algebraically <u>independent</u> over the rationals, then m is bounded by some finite constant (depending on the nature of the restrictions).

In typical applications (cf. §9) of Schneider's theorem a certain number ξ is assumed to be algebraic; this assumption imposes restrictions on two functions which are known not to be algebraically dependent, thereby violating Schneider's theorem. The contradiction establishes the transcendency of ξ.

6.2 Before stating the theorem, we recall some definitions. The notation "$f(z) = O(g|z|)$" describes the following situation: $g(x)$ is a <u>positive</u> real-valued function defined for all real $x \geq 0$; f is a function defined on some unbounded set of complex numbers; and for some constant C,

$$|f(z)| \leq Cg(|z|) \quad \text{for all } z \in S .$$

An entire function f is said to be of order $\leq \rho$ (where $\rho \geq 0$) if $f(z) = O(e^{|z|^\rho})$.* A meromorphic function

---

* Thus <u>order</u> measures the growth of $|f(z)|$ for $|z| \to \infty$ by comparing $\log(\max_{|z|=r}|f(z)|)$ (which is known to be a convex function of r) with r^ρ for various $\rho \geq 0$. It is known that if $f(z) = \sum_{\nu=0}^{\infty} a_\nu z^\nu$, then $\inf_{\mathrm{ord}(f)\leq\rho} \{\rho\} = \overline{\lim_{\nu\to\infty}} \frac{-\nu \log \nu}{\log |a_\nu|}$. We shall not use this fact.

f can always be written as a quotient $f = g/h$ of entire functions, h having zeros only at the poles of f. Under these circumstances, we say that $\mathrm{ord}(f) \leq \rho$ if $\mathrm{ord}(g) \leq \rho$ and $\mathrm{ord}(h) \leq \rho$. (The notation is self-explanatory.)

Recall that for an algebraic number α, $\overline{|\alpha|}$ is the greatest among the absolute values of the conjugates of α.

6.3 Theorem 7 (Schneider). Let $f_1(z)$, $f_2(z)$ be functions meromorphic in the whole plane, and having order $\leq \rho$. Let $z_1, z_2, \ldots, z_m$ be distinct points at which neither f_1 nor f_2 has a pole. For $\nu = 1,2$ let $f_\nu^{(0)}(z_\lambda) = f_\nu(z_\lambda)$ and

$$f_\nu^{(k)}(z_\lambda) = \left.\frac{\alpha^k f_\nu(z)}{dz^k}\right|_{z=z_\lambda} \quad (k = 1,2,\ldots) .$$

Assume that all the $f_\nu^{(k)}(z_\lambda)$ for $\nu = 1,2$; $\lambda = 1,2,\ldots,m$; $k = 0,1,2,\ldots$; lie in a field K of finite degree s over the rationals.

Assume further that there is a positive integer b such that

$$b^{k+1} f_\nu^{(k)}(z_\lambda) \text{ is an algebraic integer in } K \text{ for all } \nu,\lambda,k \qquad (1)$$

and an $\eta > 0$ such that

$$\overline{\left| f_\nu^{(k)}(z_\lambda) \right|} = 0(k^{\eta k}) \quad \text{for each} \quad \nu, \lambda \tag{2}$$

Then if f_1 and f_2 are algebraically independent over the rationals, we have

$$m \leq \rho \; (4s\eta - 2\eta + 2s + 1) \tag{3}$$

6.4 Before proceeding with the proof, we derive a preliminary arithmetic result concerning linear homogeneous equations. The reader who wishes to go on directly to the proof need only take note of the statement of lemma 1b below.

<u>Lemma 1a</u>. A homogeneous linear system

$$\sum_{j=1}^{N} b_{ij} X_j = 0 \qquad i = 1, 2, \ldots, M$$

with $N > M \geq 1$, and with rational integer coefficients, has a non-trivial solution in rational integers x_j such that, for all j,

$$|x_j| \leq 1 + (NB)^{M/(N-M)} \tag{a}$$

where $B = \max_{i,j} |b_{ij}|$.

Proof. We shall assume $B > 0$, since otherwise the lemma is trivial.

Let X be a positive integer. To each of the $(2X+1)^N$ integral N-tuples $(w_1, w_2, \ldots, w_N)$ with $|w_j| \leq X$ $(j = 1,2,\ldots,N)$, there is associated the integral M-tuple $(v_1, v_2, \ldots, v_m)$ defined by

$$v_i = \sum_{j=1}^{N} b_{ij} w_j \qquad (i = 1,2,\ldots,M)$$

Since $|v_i| \leq NBX$, at most $(2NBX + 1)^M$ of these M-tuples can be distinct. Thus if

$$(2NBX + 1)^M < (2X + 1)^N \qquad \text{(b)}$$

then two of the M-tuples must coincide, say those corresponding to the distinct N-tuples $(w'_1, w'_2, \ldots, w'_N)$, $(w''_1, w''_2, \ldots, w''_N)$, and so the system (a) has a non-trivial solution x_j with $|x_j| \leq 2X$, viz. $x_j = w'_j - w''_j$ $(j = 1,2,\ldots,N)$.

If we can find an X satisfying (b) and such that $2X \leq 1 + (NB)^{M/(N-M)}$ we are done. For $X = [\frac{1}{2} + \frac{1}{2}(NB)^{M/(N-M)}]$

("[ ]" denotes, as usual, "greatest integer") we have

$$2X > 2(\frac{1}{2} + \frac{1}{2} (NB)^{M/(N-M)}) - 2$$

whence $(2X + 1) > (NB)^{M/(N-M)}$ and

$$(2X + 1)^N > (NB)^M (2X + 1)^M > (2NBX + 1)^M$$

q.e.d.

<u>Lemma 1b</u>. A homogeneous linear system

$$\sum_{j=1}^{N} a_{ij} X_j = 0 \qquad i = 1,2,\ldots,M \qquad \text{(c)}$$

whose coefficients a_{ij} are algebraic integers, not all zero, in a field K of finite degree s over $\mathbb{Q}$, and for which $N > xM$, has a non-trivial solution in rational integers x_j such that

$$|x_j| \leq 2(\gamma N A)^{sM/(N-sM)} \qquad (j = 1,2,\ldots,N)$$

where $A = \max_{i,j} \lceil a_{ij} \rceil$, and γ is a constant depending only on K.

<u>Proof</u>. The algebraic integers a_{ij} may be represented in

the form

$$a_{ij} = \sum_{k=1}^{s} b_{ijk}u_k \qquad (d)$$

with rational integers b_{ijk} , $(u_1,u_2,\ldots,u_s)$ being a fixed integral basis for K . For this reason, the system (c) is equivalent to a system of sM equations

$$\sum_{j=1}^{N} b_{ijk} X_j = 0 \qquad \begin{cases} i = 1,2,\ldots,M \\ k = 1,2,\ldots,s \end{cases}$$

Thus, lemma 1b is a consequence of lemma 1a, provided we can show that

$$\gamma A \geq B = \max_{i,j,k} |b_{ijk}| (\geq 1)$$

for some constant γ .

To find γ , we recall that the bilinear form $(x,y) \longmapsto \text{Trace}(xy)$ from $K \times K$ into $\mathbb{Q}$ is non-degenerate, so that there are elements $v_1,v_2,\ldots,v_s$ of K satisfying Trace $(u_\sigma v_\tau) = \delta_{\sigma\tau}$ (Kronecker delta). It follows from (d) that

$$b_{ijk} = \text{Trace}\ (a_{ij}v_k) = \sum_{\theta} \theta(a_{ij}).\theta(v_k)$$

where θ runs through the automorphisms of $K/\mathbb{Q}$. Hence

$$B \leq \left(\sum_{k=1}^{s} \lceil v_k \rceil \right) A \qquad \text{q.e.d.}$$

§7. Proof of the theorem.

7.1 The strategy of the proof is as follows. Let $P = P(X,Y)$ be a polynomial in two variables, with undetermined rational integer coefficients, and of undetermined degree. We try to determine the coefficients in such a way that, while remaining within certain bounds, they give the function $F(z) = P(f_1(z), f_2(z))$ zeros of a high order, depending on the degree of P, at the points z_λ. This step turns out to be a straightforward application of lemma 1b (§6.4).

Once the existence of such coefficients is demonstrated, the local conditions (1) and (2) of §6.3 yield a lower bound for the first non-vanishing derivative of F at z_1. (Such a derivative exists since f_1 and f_2 are assumed to be algebraically independent). Then a global argument gives an upper bound involving m. Comparing these two bounds, and letting the degree of P tend to infinity, we obtain the desired bound on m .

7.2 To carry out the first step, let r be a rational integral multiple of $2sm$, and let

$$t = r^2/2sm \tag{4}$$

Let

$$F(z) = \sum_{i,j=1}^{r} c_{ij} f_1^i(z) f_2^j(z) \tag{5}$$

the c_{ij} being certain rational integers, as yet unknown. We shall require $F(z)$ to have a zero of order at least t at each z_λ , i.e.

$$F^{(k)}(z_\lambda) = 0 \qquad \begin{cases} k = 0,1,2,\ldots,t-1 \\ \lambda = 1,2,\ldots,m \end{cases} \tag{6}$$

and in fact we shall see that this requirement is fulfilled with certain integers c_{ij} , <u>not all zero</u>, satisfying

$$|c_{ij}| < \delta^t t^{(1/2+\eta)t} \tag{7}$$

where δ is a number independent of r and t .

The requirement (6), in view of (5), is

$$L_{k\lambda} = \sum_{i,j=0}^{r} c_{ij}(f_1^i f_2^j)^{(k)}(z_\lambda) = 0 \qquad \begin{cases} k = 0,1,2,\ldots,t-1 \\ \lambda = 1,2,\ldots,m \end{cases} \qquad (L)$$

(L) is a system of mt homogeneous linear equations in the r^2 unknowns c_{ij}. Before applying lemma 1b (§6.4) we must examine the coefficients of this system. By the product formula for differentiation, $(f_1^i f_2^j)^{(k)}(z_\lambda)$ may be expressed as a sum of $(i+j)^k$ terms, each of which has the form

$$f_1^{(k_1)} f_1^{(k_2)} \ldots f_1^{(k_i)} f_2^{(k_{i+1})} f_2^{(k_{i+2})} \ldots f_2^{(k_{i+j})}(z_\lambda)$$

with $k_1 + k_2 + \ldots + k_{i+j} = k$.

By condition (2) of theorem 7 (§6.3), and the definition (4) of t, it follows that, for some constants γ_1, γ_2,

$$\left|(f_1^i f_2^j)^{(k)}(z_\lambda)\right| < (i+j)^k \gamma_1^{i+j} k_1^{\eta k_1} k_2^{\eta k_2} \ldots k_{i+j}^{\eta k_{i+j}}$$

$$\leq (2r)^k \gamma_1^{2r} k^{\eta(k_1+k_2+\ldots+k_{i+j})}$$

$$< \gamma_2^t \, t^{t/2} \, t^{\eta t} \qquad (8)$$

Also, condition (1) of theorem 7 shows that

$$a_{ijk\lambda} = b^{i+j+k}(f_1^i f_2^j)^{(k)}(z_\lambda) \quad \text{is an integer of} \quad K \qquad (9)$$

so that the system (L) is equivalent to the system

$$b^{i+j+k} L_{k\lambda} = \sum_{i,j=0}^{r} a_{ijk\lambda} c_{ij} = 0 \quad \begin{cases} k=0,1,\ldots,t-1 \\ \lambda=1,2,\ldots,m \end{cases} \qquad (L')$$

in which the coefficients $a_{ijk\lambda}$ are algebraic <u>integers</u> satisfying

$$\overline{|a_{ijk\lambda}|} < b^{i+j+k} \gamma_2^t \, t^{(1/2+\eta)t} < \gamma_3^t \, t^{(1/2+\eta)t}$$

for a suitable constant γ_3 .

Now with $N = r^2$, $M = mt$ (so that $N > sM$ and $sM/(N-sM) = 1$) lemma 1b assures the existence of rational integers c_{ij} satisfying (L') , (hence satisfying (L)), and such that

$$|c_{ij}| < 2\gamma \, r^2 \, \gamma_3^t \, t^{(1/2+\eta)t} < \delta^t t^{(1/2+\eta)t}$$

for some constant δ , as desired.

7.3 Suppose $F^{(q)}(z_1)$ is a non-vanishing derivative of the function F at z_1 , so that $q \geq t$. We can estimate $|F^{(g)}(z_1)|$ from below in the following way. By (5), (7), and (8) of §7.2, we have

$$\overline{|F^{(g)}(z_1)|} < r^2 \delta^t t^{(1/2+\eta)t}(2r)^q \gamma_1^{2r} q^{\eta q}$$

Since $r = \sqrt{2sm}\ t^{1/2}$ and $t \leq q$, we get

$$\overline{|F^{(q)}(z_1)|} \leq \alpha_1 q^{(1+2\eta)q}$$

Here, and for the remainder of §7, $\alpha = \alpha(q)$ denotes a positive function of q which grows more slowly than any positive power of q^q; i.e. α is such that for any $\epsilon > 0$, $\lim_{q\to\infty} (\alpha(q)/q^{\epsilon q}) = 0$.

It follows that

$$|F^{(q)}(z_1)| \geq \frac{|\mathrm{Norm}_{K/\mathbb{Q}}(F^{(q)}(z_1))|}{\alpha_1^{s-1} q^{(1+2\eta)q(s-1)}}$$

By (5) and (9) of §7.2, we see that $b^{2r+q}F^{(q)}(z_1)$ is a non-zero algebraic integer. The norm of such an integer is ≥ 1 in absolute value; thus

$$|\mathrm{Norm}_{K/\mathbb{Q}} F^{(q)}(z_1)| \geq \frac{1}{b^{(2r+q)s}}$$

and so

$$|F^{(q)}(z_1)| > \frac{1}{\alpha_2 q^{(1+2\eta)(s-1)q}} \tag{10}$$

7.4 Now we seek an upper bound for $|F^{(q)}(z_1)|$ under the further condition that F has zeros of order $\geq q$ at each z_λ. F cannot vanish identically if f_1 and f_2 are algebraically independent, since at least one of the integers c_{ij} (cf. (5), §7.2) does not vanish. Hence there is an integer $q \geq t$ such that F has a zero of order $\geq q$ at each z_λ, while at some z_λ (which we may as well assume to be z_1) F has a zero of order <u>precisely</u> q. So the above condition may be assumed without loss of generality.

At z_1, $F(z)$ has the Taylor expansion

$$F(z) = \frac{F^{(q)}(z_1)}{q!}(z-z_1)^q + \ldots$$

so that

$$F^{(q)}(z_1) = \left.\frac{q!F(z)}{(z-z_1)^q}\right|_{z=z_1}$$

The hypotheses of theorem 7 include the existence of entire functions h_1, h_2, of order $\leq \rho$, with $h_1(z_1) \neq 0$, $h_2(z_1) \neq 0$, and such that $h_1 f_1$, $h_2 f_2$ are entire functions of order $\leq \rho$. If $H = h_1^r h_2^r$, then clearly HF is an entire function with zeros of order $\geq q$ at the z_λ. Thus, setting

$$G(z) = \frac{H(z)F(z)}{\prod_{\lambda=1}^{m} (z-z_\lambda)^q}$$

we see that G is an entire function, and that

$$|F^{(q)}(z_1)| = |G(z_1)| \left| \frac{q! \prod_{\lambda=2}^{m} (z_1-z_\lambda)^q}{H(z_1)} \right|$$

The second factor is of the type $\alpha_3 q^q$. It remains to find an upper bound for $|G(z_1)|$.

Either by Cauchy's integral formula, or by the maximum principle,

$$|G(z_1)| \leq \max_{|z|=R} |G(z)|$$

for any $R > |z_1|$. For all $R \geq 1 + \max_\lambda |z_\lambda|$, when $|z| = R$ the denominator $|\prod_{\lambda=1}^{m} (z-z_\lambda)^q|$ of $|G(z)|$ is $\geq (CR)^{mq}$, C being some positive number not depending on R.

As for the numerator $|H(z)\ F(z)|$, we see by (5) and (7) of §7.2, in conjunction with the fact that $h_1, h_2, h_1 f_1, h_2 f_2$ are all of order $\leq \rho$, that there is a constant C_1

such that for $|z| = R$,

$$|H(z)F(z)| < r^{2\delta t} t^{(1/2+\eta)t} (C_1 e^{R^\rho})^{2r}$$

$$= \alpha_4 q^{(1/2+\eta)q} e^{2rR^\rho}$$

Combining the preceding estimates, we find

$$|F^{(q)}(z_1)| < \frac{\alpha_5 q^{(3/2+\eta)q} e^{2rR^\rho}}{R^{mq}} \tag{11}$$

Take $R = q^{1/2\rho}$. If r is sufficiently large, then t is large, and since $q \geq t$, R will be large enough so that all the preceding estimates are valid. (11) becomes

$$|F^{(q)}(z_1)| < \alpha_6 \; q^{(3/2+\eta-(m/2\rho)q} \tag{12}$$

This is the estimate we were after.

7.5 Allowing r to tend to infinity, we see that the estimates (10) (§7.3) and (12) (§7.4) must hold simultaneously for arbitrarily large values of q . This is possible only if

$$3/2 + \eta - (m/2\rho) \geq -(1 + 2\eta)(s-1)$$

i.e.

$$m \leq \rho(4s\eta - 2\eta + 2s + 1) \qquad \text{q.e.d}$$

§8. A Useful Consequence

8.1 In applications of Schneider's theorem, the functions f_1 and f_2 are often such that the local arithmetic conditions (1) and (2) (§6.3) of that theorem can be deduced from analytic properties of f_1 and f_2, for example from the fact that they satisfy a certain kind of differential equation.

Suppose for instance that $f(z)$ is a meromorphic function which satisfies a differential relation of the form

$$\frac{d^n f}{dz^n} = f^{(n)} = P(f^{(n-1)}, f^{(n-2)}, \ldots, f', f)$$

where P is a polynomial in n variables with coefficients which are algebraic numbers. Let z_o be a point such that $f(z_o), f'(z_o), \ldots, f^{(n-1)}(z_o)$ are all algebraic numbers. Let K be the field generated over the rationals by the coefficients of P and by the numbers $f(z_o), f'(z_o), \ldots, f^{(n-1)}(z_o)$. Let U be an open neighbourhood of z_o on which f is defined and analytic. Finally, let $f_1, f_2, \ldots, f_n$ be the restrictions to U of $f, f', \ldots, f^{(n-1)}$ respectively. Then we have:

(i) K is an algebraic number field, of finite degree.

(ii) The $f_i(z)$ are functions defined and analytic on an open neighbourhood U of the point z_o.

(iii) $f_1(z_o), f_2(z_o), \ldots, f_n(z_o)$ all lie in K.

(iv) The derivation d/dz maps the ring $K[f_1, f_2, \ldots, f_n]$ into itself.

We shall see now that if K is a number field and $f_1(z), f_2(z), \ldots, f_n(z)$ are given functions such that (i),(ii), (iii),(iv) are satisfied, then <u>each</u> f_ν <u>satisfies all the (local) hypotheses of Schneider's theorem at</u> z_o.

(iv) says that there exist polynomials $P_1, P_2, \ldots, P_n \in K[X_1, X_2, \ldots, X_n]$ ($X_1, X_2, \ldots, X_n$ are indeterminates) such that

$$f'_i = \frac{df_i}{dz} = P_i(f_1, f_2, \ldots, f_n) .$$

Let D be the K-derivation of $K[X_1, X_2, \ldots, X_n]$ into itself such that $DX_i = P_i$. For any polynomial $Q \in K[X_1, X_2, \ldots, X_n]$ we have

$$DQ = \sum_{i=1}^{n} \frac{\partial Q}{\partial X_i} P_i \tag{13}$$

(This equation may be taken as the definition of D). If $g(z) = Q(f_1(z), f_2(z), \ldots, f_n(z))$, we see immediately that

$$g^{(k)} = \frac{d^k g}{dz} = D^k Q(f_1, f_2, \ldots, f_n) \qquad k = 0,1,2,\ldots$$

so that, by(iii), $g^{(k)}(z_o) \in K$ for all $k \geq 0$. For the special case $g = f_\nu$, this is one of the conditions of Schneider's theorem.

Now let

$$d = \max(\deg Q \ , \ \deg P_1, \ \deg P_2, \ldots, \deg P_n)$$

where "deg" stands for total degree. Let b be an integer such that, if α is any coefficient of Q or of one of the P_i , and if $a_1 + a_2 + \ldots + a_n \leq d$, then

$$b(\alpha f_1^{a_1} f_2^{a_2} \ldots f_n^{a_n} (z_o)) \text{ is an integer of } K .$$

An easy induction on k shows that $b^{k+1} D^k Q(f_1(z_o), \ldots, f_n(z_o)) = b^{k+1} g^{(k)}(z_o)$ is an integer of K $(k = 0,1,2,\ldots)$. For $g = f_\nu$, this is condition (1) of Schneider's theorem.

It remains to be shown that $\overline{|g^{(k)}(z_o)|} = O(k^{\eta k})$ for some $\eta > 0$. For this purpose, we recall the notion of <u>domination</u> of polynomials. Let $R = \sum_\sigma \alpha_\sigma M_\sigma$, $S = \sum_\sigma \beta_\sigma M_\sigma$ be two polynomials, with complex coefficients, in n variables $X_1, X_2, \ldots, X_n$, the M_σ being the various monomials in these variables. We say that S dominates R and write $R << S$ if, for all σ , β_σ is real and $|\alpha_\sigma| \leq \beta_\sigma$.

If $r = \deg R$, then $R << C(1 + X_1 + X_2 + \ldots + X_n)^r$ for some $C > 0$; in particular, if $P_1, P_2, \ldots, P_n$ are as above, there is a $B > 0$ and a positive integer h such that for all i $P_i << B(1 + X_1 + X_2 + \ldots + X_n)^{k+1}$. Note also that $\frac{\partial R}{\partial X_i} << rC(1 + X_1 + X_2 + \ldots + X_n)^{r-1}$. It is easily checked that the relation of domination is preserved under addition and multiplication. Hence, by (13),

$$DR << nrCB(1 + X_1 + X_2 + \ldots + X_n)^{r+h}$$

It follows, by induction on k , that

$$D^k R << Cn^k B^k (r)(r+h)\ldots(r+(k-1)h)(1+X_1+X_2+\ldots+X_n)^{r+kh}$$
$$<< C_R^k \, k!(1 + X_1 + X_2 + \ldots + X_n)^{r+kh}$$

for some $C_R > 0$.

Applying this result to the particular case $R = Q$, we see that

$$|\overline{g^{(k)}(z_o)}| = |\overline{D^k Q(f_1(z_o),\ldots,f_n(z_o))}|$$

$$\leq C_Q^k\, k!(1 + |\overline{f_1(z_o)}| + |\overline{f_2(z_o)}| + \ldots + |\overline{f_n(z_o)}|)^{r+kh}$$

Since $k! \leq k^k$, we conclude that

$$|\overline{g^{(k)}(z_o)}| = O(k^{\eta k}) \quad \text{for any} \quad \eta > 1 .$$

For $g = f_\nu$, this is condition (2) of §6.3.

Note, finally, that if condition (3) of §6.3 holds for all $\eta > 1$, then $m \leq \rho(6s-1)$.

8.2 In light of the foregoing discussion, Schneider's theorem implies the following [cf. Lang, 1]:

__Theorem 7'__. Let $f_1(z), f_2(z), \ldots, f_n(z)$ be functions meromorphic in the whole plane, and of order $\leq \rho$. Let K be an algebraic number field, of finite degree s, and suppose that the derivation d/dz maps the ring $K[f_1, f_2, \ldots, f_n]$ into itself.

Let $z_1, z_2, \ldots, z_m$ be distinct points, at which no f_ν has a pole, and assume that $f_\nu(z_\lambda) \in K$ $(\nu = 1,2,\ldots,n$; $\lambda = 1,2,\ldots,m)$.

If some two of $f_1, f_2, \ldots, f_n$ are algebraically independent over the rationals, then $m \leq \rho(6s-1)$.

8.3 All our applications of theorem 7' will be to sets of functions for which, in addition to the properties of theorem 7', a certain type of addition theorem holds. To cover all cases, we state:

<u>Theorem 8</u>. Let K be a number field of finite degree over Q. Let $f_1(z), f_2(z), \ldots, f_n(z)$ $(n \geq 2)$ be functions meromorphic in the whole plane and of order $\leq \rho < \infty$. Suppose that the derivation d/dz maps the ring $K[f_1, f_2, \ldots, f_n]$ into itself, and that some two of the f_i are algebraically independent over K. Assume further that there exist rational functions $R_i \in K(X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_n)$ $(i = 1,2,\ldots,n)$ such that, for all x,y, and $i = 1,2,\ldots,n$,

$$f_i(x+y) = R_i(f_1(x), f_2(x), \ldots, f_n(x), f_1(y), f_2(y), \ldots, f_n(y))$$

(we do not exclude the possibility that $f_i(x+y) = \infty$) .

Then, for any α distinct from zero and from the poles of the f_i , at least one of the numbers

$$f_1(\alpha), f_2(\alpha), \ldots, f_n(\alpha)$$

is transcendental over K .

[The meaning of the R_i may be made clearer by an example: take $K = Q$, $f_1(z) = z$, $f_2(z) = \cos(z)$, $f_3(z) = \sin(z)$, $R_1 = X_1+Y_1$, $R_2 = X_2Y_2-X_3Y_3$, $R_3 = X_2Y_3 + X_3Y_2$] .

Theorem 8 is an immediate consequence of theorem 7', for if all the numbers $f_i(\alpha)$ are algebraic, we may suppose without loss of generality that they lie in K , and then because of the R_i the same is true of all the numbers $f_i(\lambda\alpha)$ $(i = 1,2,\ldots,n;\ \lambda = 1,2,3,\ldots)$ those finitely many λ being omitted for which $\lambda\alpha$ is a pole of one of the f_i .

§9. Applications

9.1 Some classical results on transcendence are special cases of theorem 8. As a first example, take

$f_1(z) = z$, $f_2(z) = e^z$, $K = Q$. Here we may set $R_1 = X_1 + Y_1$, $R_2 = X_2 Y_2$. If the functions z , e^z were algebraically dependent, there would be an identity

$$z^n P_o(e^z) + z^{n-1} P_1(e^z) + \dots + P_n(e^z) = 0 \tag{14}$$

the P_i being polynomials with coefficients in K , $P_o \not\equiv 0$. Choosing a ξ such that $P_o(e^\xi) \neq 0$, and setting $z = \xi + 2\mu\pi i$ $(\mu = 0,1,2,\dots)$ in (14), we find that all the numbers $\xi + 2\mu\pi i$ are roots of the polynomial $z^n P_o(e^\xi) + z^{n-1} P_1(e^\xi) + \dots + P_n(e^\xi)$. This is absurd. Thus, the conditions of theorem 8 are satisfied and we have:

<u>For any</u> $\alpha \neq 0$, <u>at least one of</u> α, e^α <u>is transcendental</u>.

In particular (for $\alpha = 1$) e <u>is transcendental</u>. Since $\beta = e^{\log \beta}$, we have: <u>for algebraic</u> $\beta \neq 0$ <u>any non-zero value of</u> $\log \beta$ <u>is transcendental</u>. When $\beta = 1$, this gives: π <u>is transcendental</u>.

(For a generalization of these results (and of many of those in this section) to arbitrary "group varieties" cf. [Lang; 1].)

As a further consequence: if $\beta \neq 0$ is an algebraic number, and if $\varphi(z)$ is any function defined on a subset S of the complex plane, such that $\varphi(z)$ and $e^{\beta z}|S$ are algebraically dependent over the field of all algebraic numbers, then $\varphi(\alpha)$ is transcendental for any non-zero algebraic $\alpha \in S$.

For, if there is an identity

$$(e^{\beta z})^n P_o(\varphi(z)) + (e^{\beta z})^{n-1} P_1(\varphi(z)) + \ldots + P_n(\varphi(z)) = 0$$

for all $z \in S$, where the P's are <u>relatively prime</u> polynomials with algebraic coefficients, and if $\varphi(\alpha)$ is algebraic, then setting $z = \alpha$, we see that $e^{\beta\alpha}$ is an algebraic number (note that $P_i(\varphi(\alpha)) \neq 0$ for at least one i).

In particular: <u>for algebraic</u> $\alpha \neq 0$, cos (α) <u>and</u> sin (α) <u>are transcendental</u>. (Take $\beta = \sqrt{-1}$).

9.2 Let β be an irrational algebraic number, let $K = Q(\beta)$, $f_1(z) = e^{\beta z}$, $f_2(z) = e^z$, $R_1 = X_1Y_1$, $R_2 = X_2Y_2$. To apply theorem 8, we need only check that $e^{\beta z}$ and e^z are algebraically independent. Suppose not. Arguing as in §9.1, we find a number ξ such that some polynomial has as roots all the numbers $e^{\beta(\xi+2\mu\pi i)} = e^{\beta\xi}e^{2\mu\beta\pi i}$ ($\mu = 1,2,3,\ldots$) . These numbers are distinct because β is irrational, and we have an absurdity. Thus theorem 8 (with δ in place of α) gives:

If $\delta \neq 0$, and if β is an irrational algebraic number, then at least one of $e^{\beta\delta}$, e^{δ} is transcendental.

Setting $e^{\delta} = \alpha$, we have:

Theorem 9. (Hilbert-Gelfand-Schneider). If α, β are algebraic numbers, $\alpha \neq 0,1$; β irrational, then α^{β} is transcendental.

If $\alpha(\neq 0,1)$ and $\alpha^{\beta} = \gamma$ are both algebraic then $\beta = \log \gamma / \log \alpha$ must be either rational or transcendental. The theorem may therefore be stated as: the quotient of the logarithms of two algebraic numbers is either rational or transcendental.

9.3 The following examples involve the Jacobi elliptic functions sn(z), cn(z), dn(z) cf. [E.T. Copson: Theory of Functions of a Complex Variable, Oxford, 1935; or any book on elliptic functions.] These functions are meromorphic in the whole plane, and of order ≤ 3 . By standard formulae, for example

$$\frac{d}{dz}\,\mathrm{sn}(z) = \mathrm{cn}(z)\mathrm{dn}(z)$$

$$\mathrm{sn}(x+y) = \frac{\mathrm{sn}(x).\mathrm{cn}(y).\mathrm{dn}(y)+\mathrm{sn}(y).\mathrm{cn}(x).\mathrm{dn}(x)}{1 - k^2\mathrm{sn}^2(x)\mathrm{sn}^2(y)}$$

etc; k being a constant which we assume to be algebraic, the conditions of theorem 8 are fulfilled when $f_1(z) = z$ $f_2(z) = \mathrm{sn}(z)$, $f_3(z) = \mathrm{cn}(z)$, $f_4(z) = \mathrm{dn}(z)$, $K = Q(k)$. (The algebraic independent of $f_1(z)$, $f_2(z)$ results from the periodicity of $f_2(z)$, as in §9.1). Hence, <u>for any algebraic</u> α <u>distinct from</u> 0 <u>and from the poles of</u> $\mathrm{sn}(z)$, $\mathrm{sn}(\alpha)$ <u>is transcendental</u>. (Recall that $\mathrm{cn}(z) = (1-\mathrm{sn}^2(z))^{1/2}$, $\mathrm{dn}(z) = (1-k^2\mathrm{sn}^2(z))^{1/2}$).

For the elliptic integral of the first kind, we have

$$\int_\alpha^\beta (1-z^2)^{-1/2}(1-k^2z^2)^{-1/2}dz = v-u$$

where $\mathrm{sn}(v) = \beta$, $\mathrm{sn}(u) = \alpha$. If α and β are algebraic, the above addition formula for $\mathrm{sn}(x+y)$ shows that $\mathrm{sn}(v-u)$ is algebraic unless $v-u$ is a pole of $\mathrm{sn}(z)$. The above result becomes: <u>the value of an elliptic integral of the first kind, with algebraic</u> k , <u>evaluated between distinct algebraic limits</u> α, β , <u>is either transcendental or a pole of</u> $sn(z)$. In particular, <u>the periods of the integral (and hence also the poles of</u> $sn(z)$) <u>are transcendental</u>.

In a similar way, we may treat the elliptic integral of the second kind

$$\int_{\alpha}^{\beta}(1-z^2)^{-1/2}(1-k^2z^2)^{1/2}dz = E(v)-E(u)$$

where $sn(v) = \beta$, $sn(u) = \alpha$. The function E satisfies the relations

$$E(x) + E(y) - E(x+y) = k^2 sn(x)sn(y)sn(x+y)$$

and $d/dz(E(z)) = dn^2(z)$.

Theorem 8 applies with $f_1(z) = E(z)$, $f_2(z) = sn(z)$, $f_3(z) = cn(z)$, $f_4(z) = dn(z)$, $K = Q(k)$. The details are left to the reader. We find that if $sn(v)$, $sn(u)$ are algebraic, then so is $sn(v-u)$, whence $E(v-u)$ is transcendental, i.e. $E(v) - E(u) + k^2 sn(u)sn(v-u)sn(v)$ is transcendental. Hence: <u>The value of an elliptic integral of the second kind, with algebraic</u> k , <u>between distinct algebraic</u>

limits, is transcendental. One geometric consequence is that the circumference of an ellipse whose axes have algebraic length, is transcendental.

Several results concerning the Weierstrass elliptic functions are given in [Schneider; 1]. For example: if α is algebraic and not a pole of $\wp$ then $\wp(\alpha)$ is transcendental. In analogy with the Hilbert-Gelfand-Schneider theorem: if $\wp(z)$ and $\wp(\beta z)$ are algebraically independent, β being algebraic, then one of $\wp(\alpha)$, $\wp(\beta\alpha)$ is transcendental for any $\alpha \neq 0$. For the elliptic modular function J : if τ is an algebraic number which is not imaginary-quadratic, then $J(\tau)$ is transcendental.

($J(\tau)$ is known to be algebraic when τ is imaginary-quadratic).

These results depend on theorem 8 and the addition theorem for the $\wp$-function.

9.4 Schneider has applied his methods to functions of more than one variable, and this has led to results about Abelian integrals and functions. An interest example is

$$\underline{B}(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}dx = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

Schneider's result is: $\underline{B}(p,q)$ <u>is transcendental for all rational, non-integral</u> p,q. For $p = q = 1/2$, we have, once again, the transcendence of π.

III. ALGEBRAIC INDEPENDENCE

§10. Transcendence Properties of the Exponential Function.

10.1 In 1882, F. Lindemann proved his celebrated theorem on the algebraic independence of exponentials:

Theorem 10. Let $w_1, w_2, \ldots, w_m$ be algebraic numbers, linearly independent over the field of rationals Q. Then $e^{w_1}, e^{w_2}, \ldots, e^{w_m}$ are algebraically independent over Q.

For the case $m = 1$, the theorem states: e^{α} is transcendental for algebraic $\alpha \neq 0$. We have already proved this fact (§9), and pointed out that it implies the transcendence of π (take $\alpha = \sqrt{-1}\,\pi$).

Another form (in fact, the original form) of Lindemann's theorem is:

Let $w_1, w_2, \ldots, w_n$ be distinct algebraic numbers. Then $e^{w_1}, e^{w_2}, \ldots, e^{w_n}$ are linearly independent over the rationals.

To see this, let $\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_m$ be linearly independent algebraic numbers such that each w_i is a linear combination

$$w_i = \sum_{j=1}^{m} q_{ij}\bar{w}_j \qquad (i = 1,2,\ldots,n)$$

with rational integers q_{ij}. Choose an integer q such that $q + q_{ij} \geq 0$ for all i,j. Then any non-trivial <u>linear</u> relation $\sum_{i=1}^{n} a_i e^{w_i} = 0$ among the numbers $e^{w_i} = (e^{\bar{w}_1})^{q_{i1}}(e^{\bar{w}_2})^{q_{i2}} \ldots (e^{\bar{w}_m})^{q_{im}}$ $(i = 1,2,\ldots,n)$ would become, after multiplication by $(e^{\bar{w}_1})^q(e^{\bar{w}_2})^q \ldots (e^{\bar{w}_m})^q$, a non-trivial <u>polynomial</u> relation among the $e^{\bar{w}_j}$ $(j = 1,2,\ldots,m)$, contradicting Lindemann's theorem. Conversely, if $w_1,w_2,\ldots,w_m$ are <u>linearly independent</u> over Q, then any non-trivial polynomial relation

$$\sum_{(q)} a_{(q)}(e^{w_1})^{q_1}(e^{w_2})^{q_2},\ldots,(e^{w_m})^{q_m} = 0 \quad (q) = (q_1,q_2,\ldots,q_m)$$

would be a non-trivial linear relation among the numbers $e^{w_1q_1+w_2q_2+\ldots+w_mq_m}$, whose exponents $w_1q_1+w_2q_2+\ldots+w_mq_m$ are <u>distinct</u> for distinct m-tuples $(q_1,q_2,\ldots,q_m)$.

We note in passing that the converse of Lindemann's theorem is trivial; for if $q_1w_1+q_2w_2+\ldots+q_mw_m = 0$, then

$$(e^{w_1})^{q_1}(e^{w_2})^{q_2} \dots (e^{w_m})^{q_m} - 1 = 0 .$$

10.2 We are going to prove somewhat more than Lindemann's theorem, augmenting its qualitative statement by a quantitative estimate - a transcendence measure - of "how algebraically independent" $e^{w_1}, e^{w_2}, \dots, e^{w_m}$ actually are.

Let $M = (M_1, M_2, \dots, M_m)$ be an m-tuple of positive rational integers. If $q = (q_1, q_2, \dots, q_m)$ is another m-tuple of integers, we write $0 \leq q \leq M$ to mean $0 \leq q_j \leq M_j$ for all $j = 1,2,\dots,m$. Let $P(X_1, X_2, \dots, X_m)$ be a non-zero polynomial with rational integer coefficients

$$P = \sum_{0 \leq q \leq M} b_q X^q \qquad (X^q = X_1^{q_1} X_2^{q_2} \dots X_m^{q_m})$$

We say then that P is of degree $\leq M$, and that $H = \max_q |b_q|$ is the <u>height</u> of P.

With this notation, the refined version of theorem 10 is:

<u>Theorem 11</u>. Let $w_1, w_2, \dots, w_m$ be algebraic numbers which are linearly independent over Q, and let $g = [Q(w_1, w_2, \dots, w_m):Q]$. There is a number T, depending

only on g and m, such that if $P(X_1,\ldots,X_m)$ is a polynomial of degree $\leq M$ and of height H then

$$|P(e^{w_1},e^{w_2},\ldots,e^{w_m})| \geq H^{-TM_1M_2\ldots M_m}$$

provided that H is sufficient large. More exactly, for $g > 1$ and

$$f = \frac{1}{1-(\frac{g-1}{1})^{1/m+1}}$$

the statement holds with $TM_1M_2\ldots M_n$ replaced by

$$f^{m+1}(\prod_{k=1}^{m}(M_k+\frac{1}{f}))-1$$

while the case $g = 1$ is described in theorem 12 below.

<u>Remark</u>: For some idea of the size of f when $g > 1$, we may use the binomial expansion of $(1-\frac{1}{g})^{1/m+1}$ to get the series

$$f = (m+1)g\{1+\frac{1}{g}(\frac{m}{2!(m+1)})+\frac{1}{g^2}(\frac{m(2m+1)}{3!(m+1)(m+1)})+\ldots\}^{-1}$$

from which it follows at once that

$$(m+1)g > f > (m+1)g(1-\frac{1}{g}) = (m+1)(g-1)$$

For the case $g = 1$ we will be even more precise.

Theorem 12. Let $w \neq 0$ be a rational number. Let $P(X)$ be a non-zero polynomial in one variable X, with rational integer coefficients; let n be the degree of P and let H be its height. There is a number c_w, depending only on w, and a number $H_o(n)$, depending on n, such that whenever $H > H_o(n)$ it holds that

$$|P(e^w)| > H^{-n-c_w n^3 \left(\frac{\log(n+1)}{\log\log H}\right)}$$

This result is not the best one known. Mahler has shown that the factor n^3 in the exponent can be reduced to n^2, cf. [Schneider, 1, p.88]. Theorem 12 has the advantage for us that it can be obtained with almost no effort from the proof of theorems 10 and 11.

In general, a function $\varphi(n,t)$ is called a transcendence measure for the transcendental number ξ if for each positive integer n there is a number c_n such that for any polynomial P of degree n and height H,

$$|P(\xi)| > c_n\, \varphi(n,H)$$

A theorem giving a measure of transcendence of a number

ξ is a refinement of the assertion that ξ is transcendental. Theorem 12 implies that when $w \neq 0$ is rational $t^{-n-\epsilon}$ is a transcendence measure for e^w if $\epsilon > 0$. (On the other hand, Popken has shown that $t^{-n+\epsilon}$ is not a transcendence measure [Math.Z., 1929]). Taking $n = 1$, we see that e^w is not a Liouville number (cf. §1). Similarly, theorem 11 shows that e^w is not a Liouville number for any algebraic w.

For a discussion of the important topic of transcendence measure and the related classifications of transcendental numbers due to Mahler and Koksma, the reader is referred to [Schneider, 1].

10.3 The proof of theorems 10-12 which we shall present depends on the existence, for each positive integer p, of "linearly independent" linear combinations of the form $r = \sum_{k=1}^{n} a_k e^{w_k}$ whose coefficients a_k behave "reasonably well" (see below) with respect to p, and whose absolute value $|r|$ becomes very small for large p.

For the rest of §10, "p" will denote a variable integer, and "c" will denote a positive number whish is independent of p.

Recall that for algebraic α , $|\overline{\alpha}|$ is defined to be the greatest among the absolute values of the conjugates of α . Our preceding informal remarks are made explicit in:

<u>Lemma 2</u>. Let $w_1, w_2, \ldots, w_n$ be distinct algebraic numbers, and let $K = Q(w_1, w_2, \ldots, w_n)$. For each $p = 1,2,3,\ldots$ there exist n linear combinations

$$r_h(p) = \sum_{k=1}^{n} a_{hk}(p) e^{w_k} \qquad (h = 1,2,\ldots,n) \qquad (1)$$

with coefficients $a_{hk}(p)$ in K , such that the $n \times n$ matrix (a_{hk}) is <u>non-singular</u>, and such that for all h,k

1) $c_1^p . p! a_{hk}$ is an integer of K

2) $\lceil a_{hk} \rceil < c_2^p$

3) $|r_h| < \frac{c_3}{p!^n}$

c_1, c_2, c_3 being positive integers which do not depend on p .

We postpone the proof until §11.

10.4 The next three sections (§§10.4,10.5,10.6) prove the way for the proof of theorems 10-12. Let $w_1, w_2, \ldots, w_n$

$(n \geq 2)$ be distinct algebraic numbers, let $t < n$ be a positive integer, and let

$$L_h = \sum_{k=1}^{n} b_{hk} e^{w_k} \qquad (h = 1,2,\ldots,t) \tag{2}$$

be t linear combinations of $e^{w_1},\ldots,e^{w_n}$, with rational integer coefficients b_{hk}. We assume that the linear forms $\sum_k b_{hk} Z_k$ are linearly independent, i.e. that the matrix (b_{hk}) has maximum possible rank t. (We may say then that the L_h are "linearly independent" as linear combinations). Our main task is to find a relation between the size of the b's and the size of the L's. This information will be obtained by estimating the size of a determinant Δ which we now define.

Again let $K = Q(w_1, w_2, \ldots, w_n)$. Since, for any p, the matrix (a_{hk}) of lemma 2 is non-singular, we may assume (after renumbering the rows of (a_{hk}), if necessary) that the determinant

$$\Delta = \begin{vmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{t1} & \cdots & b_{tn} \\ a_{t+1,1} & \cdots & a_{t+1,n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

does not vanish.

Δ is bihomogeneous of degree $n-t$ in the a's and of degree t in the b's , therefore by lemma 2, $(c_1^p.p!)^{n-t}\,\Delta$ is integral, and also, with $b = \max\limits_{h,k} |b_{hk}|$, $\overline{|\Delta|} \leq n!\; c_2^{p(n-t)} b^t$. Hence, if $g = [K:Q]$,

$$1 \leq |\mathrm{Norm}_{K/Q}\ ((c_1^p.p!)^{n-t}\Delta)|$$

$$= |(c_1^p.p!)^{(n-t)g}\ \mathrm{Norm}_{K/Q}(\Delta)|$$

$$\leq (c_1^p.p!)^{(n-t)g}(n!c_2^{p(n-t)}b^t)^{g-1}|\Delta|$$

In other words,

$$|\Delta| \geq \frac{1}{n!^{g-1}c_4^p\ p!^{(n-t)g}\ b^{t(g-1)}} \qquad (3)$$

with $c_4 = c_1^{(n-t)g} \ c_2^{(n-t)(g-1)}$.

10.5 Now let us find an upper bound for $|\Delta|$. Let $B_1, B_2, \ldots, B_t, A_{t+1}, \ldots, A_n$ be the cofactors of the first column of Δ . By (1) (§10.3) and (2) (§10.4)

$$\Delta \, e^{w_1} = \sum_{h=1}^{t} B_h L_h + \sum_{h=t+1}^{n} A_h r_h \tag{4}$$

As before, lemma 2 entails

$$|B_h| \leq (n-1)! \, c_2^{p(n-t)} b^{t-1}$$

$$|A_h| \leq (n-1)! \, c_2^{p(n-t-1)} b^{t}$$

and also

$$|r_h| \leq \frac{c_3}{p!^n}$$

So if $\lambda = \max_h |L_h|$, (4) gives

$$|\Delta| \, |e^{w_1}| \leq t(n-1)! c_2^{p(n-t)} b^{t-1} \lambda + (n-t)(n-1)! c_2^{p(n-t-1)} b^t (c_3/p!^n)$$

and therefore

$$|\Delta| \le n!(c_2^{n-t})^p b^t c_5(\frac{\lambda}{b} + \frac{1}{p!^n}) \tag{5}$$

where $c_5 = 1 + (c_3/|e^{w_1}|)$.

10.6 We repeat: L_h being as in (2) (§10.4), λ and b are defined by

$$\lambda = \max|L_h| \qquad\qquad b = \max_{h,k} |b_{hk}|$$

Also, $g = [Q(w_1,\ldots,w_n):Q]$.

Our two estimates (3) (§10.4) and (5) (§10.5) for $|\Delta|$ show that

$$\frac{1}{n!^{g-1} c_4^p p!^{(n-t)g} b^{t(g-1)}} \le n!(c_2^{n-t})^p \, b^t \, c_5(\frac{\lambda}{b} + \frac{1}{p!^n})$$

Setting $c = n!^g \, c_4 \, c_2^{n-t} \, c_5 = n!^g \, (c_1c_2)^{(n-t)g} \, c_5 \ge 1$, and $s = n-(n-t)g$, we deduce

$$\frac{\lambda}{b} > \frac{1}{p!^{(n-t)g}} (\frac{1}{c^p \, b^{tg}} - \frac{1}{p!^s}) \tag{6}$$

This is the basic inequality. For example, if $t = 1$ and if $w_i = i$ for $i = 1,2,\ldots,n$, so that $g = 1$, then taking p sufficiently large we see that $\lambda > 0$; <u>in other words we have the transcendence of</u> e . Essentially the same argument proves theorem 10; however some preparation is necessary because (6) says nothing at all unless $s > 0$, (i.e. $n-(n-t)g > 0$) , or, equivalently,

$$t > n\left(\frac{g-1}{g}\right) \tag{7}$$

At this point the reader may proceed directly to §10.7 where the proof of Lindemann's theorem is completed.

For theorems 11 and 12, however, we must extract from (6) more information about the relation between λ and b . We shall do so by making a good choice for the as yet unspecified integer p . The final result is:

<u>Lemma 3</u>. If the relation (6) holds for all integers $p \geq 1$, where n,t,g,b are positive integers such that $t < n$ and $s = n-(n-t)g > 0$, and λ and c are real numbers with $c \geq 1$; and if b is sufficiently large, say $\log b > 2^8 c^6 n^2$, then, with $\sigma = ntg/s$, we have

$$\lambda > b^{1-\sigma-\left(\frac{5\sigma \log 2c}{\log \log b}\right)} \tag{8}$$

This inequality underlies theorems 11 and 12. For example, we can deduce theorem 12 as follows:

Let $w \neq 0$ be a rational number, and let $w_i = (i-1)w$ $(i = 1,2,\ldots,n$, with $n \geq 2)$. Let $P(X) = b_1+b_2X+\ldots+b_nX^{n-1}$ $(b_n \neq 0)$ be a polynomial of degree $n-1$, with rational integer coefficients. Put $L_1 = P(e^w) = \sum_{k=1}^{n} b_k e^{w_k}$, and let $t = 1$, $g = 1$. Then we have the situation of §§10.4-10.6, with $b = \max_k |b_k|$ = height of P, and $\lambda = |L_1|$, and $\sigma = n$.

After proving lemma 2 in §11, we shall see that under the present circumstances, there is a number $\bar{c}_w$ such that

$$\log 2c \leq \bar{c}_w(n-1)^2 \log n \tag{9}$$

where c is as in (6). Theorem 12 follows directly from (9) and lemma 3.

We will prove lemma 3 by showing:

(<u>A</u>) For given $\epsilon > 0$, if b is sufficiently large, say

$$\log \log b \geq \left(3 + \frac{5\sigma}{2\epsilon}\right) \log zc + \log 2n \tag{10}$$

then there exists a positive integer p such that

(i) $c^p b^{tg} > \frac{1}{2}(p!)^s$

(ii) $(p!)^n \leq b^{\sigma+\epsilon}$

For ϵ take the number $5\sigma \log 2c/\log\log b$. If we assume that

$$\log\log b > 2(3\log 2c + \log 2n) = \log(2^8 c^6 n^2) \quad (11)$$

then we have

$$\epsilon \geq \frac{(5/2)\ \sigma \log 2c}{\log\log b - 3\log 2c - \log 2n}$$

so that (10) holds, and (A) asserts the existence of p satisfying (i) and (ii). For such a p, (6) gives

$$\frac{\lambda}{b} > \frac{1}{p!^n} \geq b^{-\sigma-\epsilon}$$

Thus (A) implies lemma 3.

To prove (A), we begin with:

(B) Let α, β, γ be real numbers, $\alpha > 0$, $\beta > 1$, $\gamma > 1$. If β is sufficiently large, say $\log\log\beta \geq$ $\geq (\alpha+1)\log\gamma$, then for any positive integer p such that $\gamma^p \geq \beta$, we have $p! > \beta^\alpha$.

Proof: By Stirling's formula, $(p!)^{1/p} > \frac{1}{e}\cdot p$, and by assumption $\frac{1}{e}.p \geq \frac{1}{e}\frac{\log \beta}{\log \gamma}$ (log γ being positive); also, since $\alpha > 0$, we have $\gamma^{\alpha} \geq \beta^{\alpha/p}$. So it will be enough to show that

$$\frac{1}{e}\ \frac{\log \beta}{\log \gamma} \geq \gamma^{\alpha}$$

i.e. that

$$\log \log \beta - (1+\log \log \gamma) \geq \alpha \log \gamma$$

But $(1 + \log x) \leq x$ for all positive x, in particular for $x = \log \gamma$, and so it is sufficient that

$$\log \log \beta - \log \gamma \geq \alpha \log \gamma$$

q.e.d.

Using (B), we show

(C) Let $\delta > 0$, let $B > 1$ and let $c \geq 1$. If B is sufficiently large, say

$$\log \log (B^{\delta}) \geq (3 + \tfrac{1}{\delta})\log 2c \tag{12}$$

then there is a positive integer p such that

(iii) $B^{\delta} > (2c)^{p}$

(iv) $B^{1+2\delta} \geq p! > B^{1+\delta}$

and consequently

(v) $p! > B(2c)^p \geq 2Bc^p$

Proof: Let p be the largest positive integer such that $p! \leq B^{1+2\delta}$. By (B), with $\alpha = \frac{1+2\delta}{\delta}$, $\beta = B^\delta$, $\gamma = 2c$, it follows that $(2c)^p < B^\delta$ provided that

$$\log\log(B^\delta) \geq (\frac{1+2\delta}{\delta} + 1)\log 2c = (3 + \frac{1}{\delta})\log 2c .$$

Now suppose $p! \leq B^{1+\delta}$. Then

$$(p+1)! = (p+1)p! \leq (2c)^p B^{1+\delta} < B^\delta B^{1+\delta} = B^{1+2\delta}$$

contradicting the definition of p. q.e.d.

Conditions (ii) and (i) of A are given respectively by (iv) and (v) above, with

$$B = b^{tg/s} = b^{\sigma/n} \qquad \delta = \frac{\epsilon s}{2tgn} = \frac{\epsilon}{2\sigma}$$

For these values of B, δ, (12) becomes

$$\log\log b + \log(\epsilon/2n) \geq (3 + \frac{2\sigma}{\epsilon})\log 2c \tag{13}$$

Since $-\log \epsilon \leq \frac{1}{e}\cdot\frac{1}{\epsilon}$, (13) is implied by

$$\log\log b \geq (3 + \frac{2\sigma}{\epsilon})\log 2c + \log 2n + \frac{1}{e}\cdot\frac{1}{\epsilon} \tag{14}$$

and since σ is ≥ 2 (as is easily checked), so that

$\frac{1}{2}\sigma \log 2c > \frac{1}{e}$, (14) is implied by (10) of (A) . This completes the proof of lemma 3.

10.7 We shall now prove theorem 10. Let $w_1, w_2, \ldots, w_m$ be algebraic numbers, linearly independent over Q . Let $M = (M_1, M_2, \ldots, M_m)$ be an m-tuple of positive integers, and let $q = (q_1, q_2, \ldots, q_m)$ be a "variable" m-tuple of non-negative integers. Let $P = P(X_1, X_2, \ldots, X_m)$ be a non-zero polynomial

$$P = \sum_{0 \leq q \leq M} a_q X^q \qquad (X^q = X_1^{q_1} X_2^{q_2} \ldots X_m^{q_m})$$

with rational integer coefficients a_q . Our object is to show that $L = P(e^{w_1}, e^{w_2}, \ldots, e^{w_m}) \neq 0$.

Let $\nu = (\nu_1, \nu_2, \ldots, \nu_m)$ be an m-tuple of non-negative integers, and for $0 \leq \mu \leq \nu$ let

$$X^{\mu} P = \sum_{0 \leq q \leq M+\mu} b_{\mu q} X^q$$

The coefficients $b_{\mu q}$ are rational integers and $\max_{\mu, q} |b_{\mu q}| = \max_q |a_q| = H$ (the height of P). Setting $w_q = q_1 w_1 + q_2 w_2 + \ldots + q_m w_m$, we see that the substitution of e^{w_i} for X_i $(i = 1,2,\ldots,m)$ turns X^q into e^{w_q} ; hence

$$L_\mu = e^{w_\mu} L = \sum_{0 \leq q \leq M+\mu} b_{\mu q} e^{w_q}$$

If $\lambda = \max|L_\mu|$, then clearly $L = 0$ if and only if $\lambda = 0$.

Now the w_q are distinct for distinct values of q , since $w_1, w_2, \ldots, w_m$ are linearly independent. Moreover, the matrix $(b_{\mu q})$ has maximum possible rank, that is to say if $\sum_\mu c_\mu b_{\mu q} = 0$ for all q , then $c_\mu = 0$ for all μ (because if $Q = \Sigma\, c_\mu X^\mu$, then $QP = \sum_{\mu q} c_\mu b_{\mu q} X^q \equiv 0$, whence $Q \equiv 0$). In other words we are dealing with "linearly independent" linear combinations of the e^{w_q} , and we have exactly the sort of situation discussed in §§10.4-10.6. As soon as we satisfy the condition (7) (§10.6) - which concerns the number t of distinct L_μ as compared to the number n of distinct e^{w_q} - then, choosing a sufficiently large value for p in (6) (§10.6), we find that $\lambda > 0$. Hence $L \neq 0$, as required

Now $t = \prod_{k=1}^{m} (\nu_k + 1)$, while $n = \prod_{k=1}^{m} (M_k + \nu_k + 1)$. Set $g = [K:Q]$ where K is the field generated over Q by the various w_q ; clearly $K = Q(w_1, w_2, \ldots, w_m)$ so that g

depends only on $w_1, w_2, \ldots, w_m$. The condition (7) is that $t > n(g-1)/g$, i.e.

$$\prod_{k=1}^{m} (\nu_k+1) > \frac{g-1}{g} \prod_{k=1}^{m} (M_k + \nu_k + 1) \tag{15}$$

This inequality <u>will hold</u> for sufficiently large ν , for example if $1 > (\frac{g-1}{g})^{1/m} (\frac{M_k}{\nu_k+1} + 1)$ for all k . Thus (7) can be satisfied, and the proof is complete.

10.8 Theorem 11 results from a proper choice of the ν_k . We assume $g > 1$, the case $g = 1$ having been covered by theorem 12. Put $h = (g/g-1)^{1/m+1}$ and, for each k , let

$$\nu_k = \left[\frac{M_k}{h-1} \right]$$

where "[ ]" denotes the "greatest integer" function. Then (15) (§10.7) is satisfied, since $\nu_k+1 > M_k/(h-1)$, so that

$$\frac{n}{t} = \prod_{k=1}^{m} \frac{M_k+\nu_k+1}{\nu_{k+1}} = \prod_{k=1}^{m} (1 + \frac{M_k}{\nu_k+1}) < h^m < h^{m+1} = \frac{g}{g-1}$$

Previously (lemma 3, §10.6) we have defined σ by

$$\sigma = \frac{ntg}{n-(n-t)g} = \frac{n}{1-\frac{n}{t}\cdot\frac{g-1}{g}}$$

Now, with our choice of ν_k,

$$1-\frac{n}{t}\cdot\frac{g-1}{g} > 1 - h^m \circ h^{-(m+1)} = 1 - \frac{1}{h} = f^{-1}$$

where $f = (1-h^{-1})^{-1}$ is as in theorem 11. Also

$$n = \prod_{k=1}^{m}(M_k+\nu_k+1) \leq \prod_{k=1}^{m}(M_k+\frac{M_k}{h-1}+1) = \prod_{k=1}^{m}(\frac{h}{h-1}M_k+1) = \prod_{k=1}^{m}(fM_k+1)$$

Thus

$$\sigma < f \prod_{k=1}^{m}(fM_k+1) = f^{m+1}\prod_{k=1}^{m}(M_k + \frac{1}{f}) = \tau \quad (\text{say}) \qquad (16)$$

(We may remark here that our choice of ν_k was dictated by the desire to make the upper bound τ in (16) as small as possible.)

By the definition of λ and L (§10.7), there is a positive number α, depending only on ν and on $w_1, w_2, \ldots, w_m$, such that $|L| \geq \alpha\lambda$. Combining this remark with (8) of lemma 3 (§10.6), (where now $b = H$, the height of the polynomial P, cf. §10.7), we have

$$|L| > \alpha\, H^{1-\sigma-\frac{5\sigma\log 2c}{\log\log H}} \qquad (17)$$

for sufficiently large H. Let $\epsilon = \tau - \sigma > 0$. Choose

H so large that (17) holds and also so that

$$\frac{5\sigma \log 2c}{\log \log H} \leq \frac{\epsilon}{2} \quad \text{and} \quad \alpha H^{\epsilon/2} \geq 1$$

Then (17) gives

$$|L| > \alpha H^{1-\tau+\epsilon-(\epsilon/2)} = H^{1-\tau}(\alpha H^{\epsilon/2}) \geq H^{1-\tau}$$

q.e.d.

§11. Arithmetic Properties of Green's Functions

11.1 Our aim now is to prove lemma 2 (§10.3). The proof rests on the properties of the complex integral

$$R(x) = \frac{1}{2\pi i} \int_C \frac{e^{xz}}{Q(z)} \, dz$$

where $Q(z) = (z-w_1)^{p_i}(z-w_2)^{p_2} \ldots (z-w_n)^{p_n}$ for certain positive integers $p_1, p_2, \ldots, p_n$, and C is a circle containing $w_1, w_2, \ldots, w_n$ in its interior. We shall see that

$R(x) = \sum_{k=1}^{n} P_k(x)e^{w_k x}$ where the P_k are polynomials of degree $p_k - 1$, and that, with $v = \sum_{k=1}^{n} p_k$, $R(x)$ has a Taylor expansion beginning with $(1/(v-1)!)x^{v-1}$. The sought-after linear combinations r_h will be defined to be $R_h(1) = \sum_{k=1}^{n} P_{hk}(1)e^{w_k}$, the $R_h(x)$ being suitable integrals of the above type.

It is easily seen that $R(x)$ is the unique solution of the differential equation

$$Q(D)y = 0 \qquad (D = \frac{d}{dx})$$

with initial conditions $y(0) = y'(0) = \ldots = y^{(v-2)}(0) = 0$, $y^{(v-1)}(0) = 1$; in other words, $R(x)$ is the Green's function associated with the linear operator $Q(D)$. As such, $R(x)$ can be studied without the use of any integration whatsoever. The treatment based on complex integration, however, though not "elementary", is most economical.

11.2 By Cauchy's residue theorem, $R(x) = \sum_{k=1}^{n} \rho_k(x)$ where $\rho_k(x)$ is the residue of $e^{xz}/Q(z)$ at the point w_k. For $\rho_k(x)$ we have the two representations

$$\rho_k(x) = \frac{1}{2\pi i} \int_{C_k} \frac{e^{xz}}{Q(z)}\, dz = \frac{1}{(p_k-1)!} \left(\frac{(z-w_k)^{p_k} e^{xz}}{Q(z)} \right)^{(p_k-1)} \Bigg|_{z=w_k}$$

Here C_k is a circle about w_k such that w_j lies outside C_k whenever $j \neq k$; and for any $G(z)$, $G^{(p_k-1)}(z) = D^{p_k-1}G(z)$ where D is the operator d/dz .

An easy induction on s shows that for any non-negative integer s , and any meromorphic function $f(z)$, $D^s(e^{xz}f) = e^{xz}(D+x)^s f$. Using this fact to carry out the indicated differentiation in the second representation of $\rho_k(x)$, we find that $\rho_k(x) = e^{w_k x} P_k(x)$, where

$$P_k(x) = \frac{1}{(p_k-1)!} \sum_{t=0}^{p_k-1} \binom{p_k-1}{t} \left\{ D^t \left. \frac{(z-w_k)^{p_k}}{Q(z)} \right|_{z=w_k} \right\} x^{p_k-1-t}$$

We see from this that P_k is a polynomial of degree <u>exactly</u> p_k-1 , and that the coefficients of the powers of x in P_k are sums of terms of the form.

$$\frac{N}{(p_k-1)! \prod_{\substack{j=1 \\ j\neq k}}^{n} (w_k-w_j)^{p_j + \sigma_j}} \qquad \begin{cases} N \text{ a rational integer} \\ \sigma_j \geq 0 \\ \sigma_1+\sigma_2+\ldots+\sigma_n \leq p_k-1 \end{cases} \tag{18}$$

From (18) we draw two conclusions about $P_k(1)$. First of all, if d is a rational integer such that

$$\frac{d}{\prod_{\substack{j,k=1 \\ j\neq k}}^{n} (w_k-w_j)}$$

is an algebraic integer, and if $v = \sum_k p_k$, then

$$(p_k-1)!\; d^v P_k(1) \quad \text{is integral} \tag{19}$$

Secondly, if we regard $P_k(1)$ as a function of $w_1, w_2, \ldots, w_n$, we see that any conjugate $\theta P_k(1)$ of $P_k(1)$ (θ being an automorphism of $Q(w_1, w_2, \ldots, w_n)/Q$) is the <u>same</u> function of $\theta w_1, \theta w_2, \ldots, \theta w_n$. Hence, if we set $\bar{w}_k = \theta w_k$ and $\bar{Q}(z) = \prod_{k=1}^{n} (z-\bar{w}_k)$, then

$$\bar{\rho}_k(1) = e^{\bar{w}_k} . \theta P_k(1) = \frac{1}{2\pi i} \int_{\bar{C}_k} \frac{e^z}{\bar{Q}(z)} dz \tag{20}$$

where $\bar{C}_k$ is a circle about $\bar{w}_k$ with all $\bar{w}_j (j \neq k)$ in its exterior.

Let

$$\eta = \frac{1}{2} \min_{j \neq k, \theta} |\theta w_k - \theta w_j|$$

the minimum being taken over all pairs $j \neq k$ and <u>all</u> automorphisms θ . For $\bar{C}_k$ in (20) take the circle with center $\bar{w}_k$ and radius η . Then we get

$$|\theta P_k(1)| = \left| \frac{1}{2\pi i} \int_{\bar{C}_k} \frac{e^{z-\bar{w}_k}}{\bar{Q}(z)} dz \right| \leq \frac{1}{2\pi} \cdot \frac{e^{\eta}}{\eta^v} \cdot 2\pi\eta$$

Thus,

$$|P_k(1)| \leq \frac{e^{\eta}}{\eta^{v-1}} \qquad (k = 1,2,\ldots,n) \qquad (21)$$

11.3 As soon as we define r_h , (19) and (21) will give us conditions 1) and 2) of lemma 2. For condition 3), we must estimate the size of $R(x)$.

Once again, let $v = \sum_k p_k$. We have

$$\frac{1}{Q(z)} = \frac{1}{z^v \prod_{j=1}^{n} (1 - \frac{w_j}{z})^{p_j}}$$

$$= \sum_{s=0}^{\infty} \frac{Q_s(w_1,w_2,\ldots,w_n)}{z^{s+v}}$$

where $Q_s(X_1,X_2,\ldots,X_n)$ is a homogeneous polynomial, of degree s , with non-negative rational integer coefficients. Note that Q_o is the constant 1 . If $\mu = \max(|w_1|,|w_2|,\ldots,|w_n|)$, and $\epsilon > 0$, then the above series for $Q(z)$ converges uniformly for $|z| \geq \mu + \epsilon$.

Let C be the circle with center 0 and radius $\mu + \epsilon$. Because of the uniform convergence,

$$R(x) = \sum_{s=0}^{\infty} \frac{1}{2\pi i} \int_C \frac{e^{xz}}{z^{s+v}} dz \cdot Q_s(w_1, w_2, \ldots, w_n)$$

$$= \sum_{s=0}^{\infty} \frac{x^{s+v-1}}{(s+v-1)!} Q_s(w_1, w_2, \ldots, w_n)$$

$$= \frac{x^{v-1}}{(v-1)!} + \ldots \qquad (22)$$

Since Q_s has non-negative coefficients,

$$|Q_s(w_1, w_2, \ldots, w_n)| \leq Q_s(\mu, \mu, \ldots, \mu)$$

Hence

$$|R(x)| \leq \sum_{s=0}^{\infty} \frac{|X|^{s+v-1}}{(s+v-1)!} Q_s(\mu, \mu, \ldots, \mu)$$

$$= \frac{1}{2\pi i} \int_C \frac{e^{|X|z}}{(z-\mu)^v} dz$$

$$= \frac{|X|^{v-1}}{(v-1)!} e^{|X|\mu}$$

In particular

$$|R(1)| \leq \frac{e^{\mu}}{(v-1)!} \qquad (23)$$

11.4 Now let p be a positive integer, and let

$$p_{hk} = p + \delta_{hk} \text{ (Kronecker } \delta) \quad (1 \leq h \leq n, 1 \leq k \leq n)$$

Let $R_h(x)$ be the function $R(x)$ of §11.1 when $p_1 = p_{h1}$, $p_2 = p_{h2}, \ldots, p_n = p_{hn}$. Note that $v_h = \sum_{k=1}^{n} p_{hk} = np+1$. Also

$$R_h(x) = \sum_{k=1}^{n} P_{hk}(x) e^{w_k x} \tag{24}$$

the P_{hk} being certain polynomials. Let $a_{hk} = P_{hk}(1)$ and set

$$r_h = R_h(1) = \sum_{k=1}^{n} a_{hk} \, e^{w_k}$$

By (19), $p! d^{np+1} a_{hk}$ is integral, and (1) of lemma 2 follows directly.

Similarly (21) gives condition (2), and (23) gives condition (3) (since $(v-1)! = (np)! > p!^n$).

It remains to be seen that the matrix (a_{hk}) is non-singular. Consider the matrix $(P_{hk}(x))$. In the expansion of its determinant $D(x)$, each term is a polynomial of degree $< np$ <u>except for</u> the term $P_{11}(x)P_{22}(x)\ldots P_{nn}(x)$, whose degree is np . [Recall that $P_{hk}(x)$ is of degree <u>precisely</u> $(p_{hk}-1)$]. Thus $D(x)$ is a polynomial of degree $np(=v-1)$.

On the other hand, (24) implies that

$$D(x)\ e^{w_1 x} \quad \sum_{h=1}^{n} A_h(x) R_h(x)$$

where the $A_h(x)$ are the cofactors of the first column of $(P_{hk}(x))$. Since, by (22), $R_h(x) = \frac{x^{v-1}}{(v-1)!} + \ldots$, and $e^{w_1 x} = 1 + w_1 x + \ldots$, it follows that $D(x)$ <u>has no terms of degree</u> $< v-1$. This means that $D(x) = cx^{np}$, $c \neq 0$.

Thus $\det(a_{hk}) = \det(P_{hk}(1)) = D(1) = c \neq 0$, and (a_{hk}) is non-singular.

This completes the proof of lemma 2.

11.5 For theorem 12, we are interested in the situation in which $w \neq 0$ is a rational number, and $w_i = (i-1)w$ $(i = 1,2,\ldots,n$, with $n \geq 2)$. A close look at (18) shows that, if d is the numerator of w , and if $u = \max_{j,k} (p_j + p_k)$, then

$$(p_k-1)!(n-1)!^u \, d^{(n-1)u} P_k(1) \quad \text{is an integer .}$$

By our definition of the r_h (§11.4), it follows that for the number c_1 of lemma 2 we may use $((n-1)!d^{n-1})^3$.

Similarly, taking $\eta = |w/z|$ in (21), we may set $c_z = (d')^n$ for some positive integer d' depending only on w ; and in view of (23), we may put $c_3 = (d'')^n$ where d'' also depends only on w .

For the number c of inequality (6), (§10.6) (where now $t = g = 1$), we have then

$$\log 2c \leq \bar{c}_w (n-1) \log (n!) \leq \bar{c}_w (n-1)^2 \log n$$

for some number $\bar{c}_w$ depending only on w . As we have seen in §10.6, this implies theorem 12.

Remark: By exercising more care in the choice of c_1, c_2, c_3 , and by a slightly different treatment of inequality (6), one can replace the factor log (n+1) in theorem 12 by unity without changing the basic argument.

§12 Siegel's Method

12.1 In the introduction, we have referred to Siegel's work on the algebraic independence of transcendental numbers. Some of the basic ideas of Siegel's method were illustrated in our proof of Lindemann's theorem. We will describe briefly, without proof, the main features of the method. Full details are in [Siegel; 1].

The construction of the "approximation form" $R(x) = \sum_{k=1}^{n} P_k(x)e^{w_k x}$ given in §11 depended on the special properties of the exponential function. Siegel has given a construction which applies to a large class of functions - the so-called E functions - which includes the exponential functions. Siegel's construction is purely algebraic, being based on a lemma similar to lemma 1b, §6.

The function $f(x)$ is an E-function if

$$f(x) = \sum_{\nu=0}^{\infty} a_\nu \frac{x^\nu}{\nu!}$$

where the coefficients a_ν are such that

1) All the a_ν belong to a single algebraic number field of finite degree over Q.

2) $|\overline{a_\nu}| = O(\nu^{\epsilon\nu})$ for any $\epsilon > 0$, as $\nu \to \infty$.

3) There is a sequence of positive integers $\{b_\nu\}$ such that for all ν, the numbers $a_0 b_\nu, a_1 b_\nu, \ldots, a_\nu b_\nu$ are algebraic integers and $b_\nu = O(\nu^{\epsilon\nu})$ for any $\epsilon > 0$ as $\nu \to \infty$.

One verifies that E-functions are entire, that the E functions form a ring, and that the derivative of an E function is again an E-function.

The problem is the following: given m E-functions $E_1, E_2, \ldots, E_m$, and an algebraic α, investigate the transcendence and algebraic independence of the numbers $E_1(\alpha), E_2(\alpha), \ldots, E_m(\alpha)$.

Additional assumptions have to be made on the E_k, namely that they satisfy a system of homogeneous differential equations

$$\frac{dE_j}{dx} = \sum_{k=1}^{m} Q_{jk} E_k \qquad (j = 1,2,\ldots,m)$$

where the Q_{jk} are rational functions with algebraic coefficients; and also that the E_k satisfy a certain complicated independence condition called <u>normality</u>.

Under these assumptions, Siegel showed that $E_1(\alpha), E_2(\alpha), \ldots, E_m(\alpha)$ are indeed transcendental and algebraically independent, as long as α is neither 0, nor a pole of any $Q_{k\ell}$.

Siegel's proof makes use of "approximation forms" in a manner which, when suitably specialized, yields a proof of Lindemann's theorem very much like the one in §10. Other applications of Siegel's method to particular E-functions require a careful study of the algebraic and analytic properties of the functions under consideration, because of the need to establish "normality". Siegel has carried out such investigations in several instances.

The best known result of this work concerns the Bessel function

$$J_0(x) = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu}{(\nu!)^2} \left(\frac{x}{2}\right)^{2\nu}$$

which is easily seen to be an E-function. For any algebraic $x \neq 0$, Siegel showed that $J_0(x)$, and $J_0'(x)$, are transcendental and algebraically independent.

One very pretty consequency of further work on Bessel functions is: the continued fraction $a_1 + \frac{1}{a_2+} \frac{1}{a_3+} \cdots$ is transcendental whenever the integers $a_1, a_2, a_3, \ldots$ form an arithmetic progression. Thus $1 + \frac{1}{2+} \frac{1}{3+} \ldots$ is transcendental; as is $\frac{1}{2+} \frac{1}{6+} \frac{1}{10+} \ldots$. The latter continued fraction converges to $\frac{e-1}{e+1}$ (see e.g. [Hermite; 1]) , whence, again, e is transcendental.

12.2 We close with some recent developments.

In 1954, A.B. Sidlovskii, elaborating on Siegel's work, proved the following theorem: [see Math. Reviews 21, #1295].

Let $E_1, E_2, \ldots, E_m$ be E-functions satisfying

$$\frac{dE_j}{dx} = Q_{oj} + \sum_{k=1}^{m} Q_{jk} \cdot E_k \quad (j = 1, 2, \ldots, m)$$

where the Q's are rational functions, and let $\alpha \neq 0$ be an algebraic number distinct from the poles of these rational functions.

Then the numbers $E_1(\alpha), \ldots, E_m(\alpha)$, are algebraically independent over the rationals if and only if the functions $E_1, E_2, \ldots, E_m$ are algebraically independent over the field of rational functions.

From this fundamental theorem, Sidlovskii deduced a generalization of Lindemann's theorem: [Soviet Mathematics, v. 2, #3, (May, 1961) pp. 841-844].

Suppose the transcendental E-function $e(z)$ is a solution of the first order linear differential equation

$$P(z)\,y' + Q(z)y = R(z)$$

where $P, Q, R,$ are polynomials. Let $w_1, w_2, \ldots, w_m$ be algebraic numbers, linearly independent over Q, and distinct from the zeros of P. Then $e(w_1), e(w_2), \ldots, e(w_m)$ are algebraically independent over Q.

In another paper [Math. Reviews 26, #1285] V. A. Oleinikov applied Sidlovskii's theorem to the confluent hypergeometric function.

$$K_{\alpha,\beta}(z) = \sum_{\nu=0}^{\infty} \frac{\beta(\beta+1)\ldots(\beta+\nu-1)}{\alpha(\alpha+1)\ldots(\alpha+\nu-1)} \frac{z^\nu}{\nu!} \quad (\alpha, \beta \neq 0, -1, -2, \ldots)$$

Oleinikov's result is that if α, β are rational, $\beta \neq 1, 2, \ldots$, $\beta - \alpha \neq 0, 1, 2, \ldots$ then $K_{\alpha,\beta}(z)$ and $K'_{(\alpha,\beta)}(z)$ are algebraically independent for every non-zero algebraic value of z.

For further results, see Sidlovskii's paper in [Amer. Math. Soc. Translations (2) 27 (1963) pp. 191-230].

REFERENCES

Only a few references are given here. An extensive bibliography appears in [Schneider, 1].

HERMITE, Ch. [1]: Sur la fonction Exponentielle (1873) Oeuvres III, 150-181.

LANG, S. [1]: Transcendental Points on Group Varieties, Topology 1, (1962), 313-318.

LEVEQUE, W. [1]: Topics in Number Theory, Vol.1, Addison-Wesley (1956)

[2]: ----- Vol.2.

MAHLER, K. [1]: On the Approximation of π. Proc. Akad. Wetensch. Ser.A. 56, (1953), 30-42.

[2]: Uber die Dezimalbruchentwicklung gewisser Irrationalzahlen. Mathematika B, Zutphen, 6 (1937), 22-36.

SCHNEIDER, Th.[1]: Einführung in die Transzendenten Zahlen, Springer-Verlag (1957)

[2]: Ein Satz uber ganzwertige Funktionen als Prinzip fur Transzendenzbeweise, Math. Ann. 121 (1949) 131-140.

SIEGEL, C.L. [1]: Transcendental Numbers. Annals of Math. Studies, 16, Princeton, 1949.

****