THE APPROXIMATION TO ALGEBRAIC NUMBERS BY RATIONALS.

BY

F. J. DYSON

TRINITY COLLEGE, CAMBRIDGE.

I. The object of this paper is the proof of the following

Theorem. *If θ is any algebraic number of degree $n \geq 2$, and if there are an infinite number of rational fractions p/q such that*

$$|\theta - (p/q)| < q^{-\mu}, \tag{1}$$

then

$$\mu \leq \sqrt{(2n)}. \tag{2}$$

There is a famous theorem due to Siegel (**1**) which states that under the same hypotheses μ satisfies the weaker inequality

$$\mu \leq \min_{1 \leq s \leq n-1} \left(s + \frac{n}{s+1}\right) < 2\sqrt{n}. \tag{3}$$

It is easy to see that (2) is stronger than (3) for all $n > 2$, although the improvement is not great for small values of n.

The history of previous attempts to obtain a stronger result than Siegel's is a curious chapter of accidents, and will be briefly summarised here before proceeding to the proof of the theorem stated above. In the first place it is probable, and was conjectured by Siegel, that the correct conclusion to be drawn from the hypotheses of the theorem is

$$\mu \leq 2, \tag{4}$$

irrespective of the degree n of θ. In this direction, Siegel (**2**) proved that if (1) is satisfied for an infinite sequence of fractions p_i/q_i with

(5) $$\overline{\lim_{i \to \infty}} \left(\frac{\log q_{i+1}}{\log q_i} \right) < \infty,$$

then

(6) $$\mu \leq e(\log n + (\log n)^{-1}).$$

A proof of the full result (4) without the additional hypothesis (5) was attempted by Kuroda (**3**), but the argument was vitiated by a false premise which could not be dispensed with. Next, a valid proof of (4) was given by Schneider (**4**), but only with the hypothesis (5). Finally, Koksma (**5**) stated that he had been informed privately by Mahler that Schneider had proved (4) without the additional hypothesis. There is however no mention of this in the published work of Schneider, and Professor Mordell has recently ascertained from Mahler that Schneider's proof never appeared in print and was probably incomplete. The present paper provides what is, as far as is known to the author, the first improvement on Siegel's result (3) which covers the general case when (5) is not assumed to hold.

The proof of the theorem of this paper is divided into two parts, the second of which runs closely parallel to the proof of Siegel's theorem given in Landau (**7**). The first part includes the new idea by which (3) becomes improved to (2), and the conclusion of this part can conveniently be stated as a self-contained lemma which is of some interest in itself. In the statement of this lemma, and throughout the paper, $[x]$ will denote the greatest integer not exceeding x.

Lemma. *Let $R(x, y)$ be any polynomial with real or complex coefficients and of degree not exceeding u in x and s in y. Let $(n + 1)$ distinct real or complex numbers $x_0, x_1, \ldots, x_n$ and $(n + 1)$ distinct real or complex numbers $y_0, y_1, \ldots, y_n$ be given (with no restriction upon whether any of the x_i are equal or unequal to any of the y_i). Let $\delta, \lambda, t_0, t_1, \ldots, t_n$ be real numbers such that*

(7) $$\begin{cases} 0 < \delta < 1, & \lambda \geq 2\delta^{-1}, \quad s \leq \frac{1}{2} n \delta (u + 1); \\ 0 \leq t_i \leq s, & \lambda [t_i + 1] \leq u + 1, \quad i = 0, 1, \ldots, n. \end{cases}$$

Suppose that

(8) $$\left(\frac{\partial}{\partial x} \right)^{\mu} \left(\frac{\partial}{\partial y} \right)^{\nu} R(x, y) \bigg|_{x = x_i,\ y = y_i} = 0$$

for every three integers i, ν and μ such that

(9) $$0 \leq i \leq n, \quad 0 \leq \nu \leq t_i, \quad 0 \leq \mu \leq \lambda (t_i - \nu).$$

Then

$$(10) \qquad \lambda \sum_{i=0}^{n} (1 + [t_i])(t_i - \tfrac{1}{2}[t_i]) \leq (1 + \tfrac{1}{2} n(n+1)\delta)(s+1)(u+1).$$

The meaning of this lemma is as follows. The left side of (10) is approximately equal to the total number of conditions (8) satisfied by $R(x, y)$ at the points (x_i, y_i). The number of coefficients of $R(x, y)$ is $(u + 1)(s + 1)$, and this is the smallest number of conditions (8) which can possibly be inconsistent for a non-zero polynomial $R(x, y)$ with undetermined coefficients. Thus the lemma states that the maximum number of conditions of this type which can be satisfied by one polynomial cannot exceed the number which must be satisfied by some polynomial by a factor greater than $(1 + \frac{1}{2} n(n+1)\delta)$; and this factor can be made indefinitely near to unity by making the ratio u/s sufficiently large. The next section will be occupied with the proof of the lemma.

II. The polynomial $R(x, y)$ may be expressed in the form

$$(11) \qquad R(x, y) = \sum_{j=0}^{l-1} f_j(x)\, g_j(y),$$

where the f_j are polynomials in x of degree u at most, the g_j are polynomials in y of degree s at most, and

$$(12) \qquad l \leq s + 1.$$

For example, such an expression is possible with

$$g_j(y) = y^j.$$

Let then an expression of the form (11) for $R(x, y)$ be chosen with the smallest possible value for l. If the $f_j(x)$ are linearly dependent, then for some value of i

$$f_i(x) = \sum_{j \neq i} c_j f_j(x),$$

and so

$$R(x, y) = \sum_{j \neq i} f_j(x) \{g_j(y) + c_j g_i(y)\},$$

which is impossible because $R(x, y)$ cannot be split into a sum of less than l terms of the form (11). Hence the $f_j(x)$ must be linearly independent, and by the same reasoning the $g_j(y)$ must be linearly independent also.

Suppose that for some integer k the determinant

$$(13) \qquad M_k = \left\| f_j^{(i)}(x) \right\|_{j=0,1,\ldots,k}^{i=0,1,\ldots,k}, \qquad 0 \leq k \leq l-1,$$

is identically zero. Then there exist rational functions $e_j(x)$ not all zero such that

$$(14) \qquad \sum_{j=0}^{k} e_j(x) f_j^{(i)}(x) = 0, \qquad i = 0, 1, \ldots, k.$$

Differentiating (14) and using (14) to simplify the result,

$$(15) \qquad \sum_{j=0}^{k} e_j'(x) f_j^{(i)}(x) = 0, \qquad i = 0, 1, \ldots, k-1.$$

If

$$(16) \qquad e_j'(x) = c(x)\, e_j(x), \qquad j = 0, 1, \ldots, k,$$

for some rational function $c(x)$ independent of j, then

$$e_j(x) = A_j \exp\left(\int c(x)\, dx\right), \qquad j = 0, 1, \ldots, k,$$

where the A_j are pure numbers, and this is impossible by (14) since the $f_j(x)$ are linearly independent. Hence (16) is false for any $c(x)$, and so (14) and (15) give two independent linear relations between the

$$f_j^{(i)}(x), \quad i = 0, 1, \ldots, k-1; \quad j = 0, 1, \ldots, k.$$

Eliminating $f_k^{(i)}(x)$ between (14) and (15) gives one linear relation between the

$$f_j^{(i)}(x), \quad i = 0, 1, \ldots, k-1; \quad j = 0, 1, \ldots, k-1;$$

and therefore the determinant M_{k-1} must also vanish identically. It is thus proved that if M_{k-1} is not zero then M_k is not zero. But

$$M_0 = f_0(x)$$

is not zero. Therefore M_k does not vanish for any k, and in particular

$$(17) \qquad L(x) = M_{l-1} = \left\| f_j^{(i)}(x) \right\|_{j=0,1,\ldots,l-1}^{i=0,1,\ldots,l-1}$$

is not identically zero. By an exactly similar argument,

$$(18) \qquad N(y) = \left\| g_j^{(i)}(y) \right\|_{j=0,1,\ldots,l-1}^{i=0,1,\ldots,l-1}$$

is not identically zero.

Let a set of polynomials $h_{ij}(y)$ and a set of integers d_{ij} be defined by the following inductive process. For each value of j in the range $0 \leq j \leq l-1$, $h_{ij}(y)$ is one of those linear combinations of the $g_k(y)$ which vanish to the least order, greater than any d_{im} with $m < j$, at $y = y_i$; d_{ij} is the order to which $h_{ij}(y)$ vanishes at $y = y_i$. These definitions can be satisfied in a variety of ways, and it is not necessary to enquire whether the $h_{ij}(y)$ and d_{ij} are uniquely fixed. The $h_{ij}(y)$ for a given value of i are linearly independent since they all vanish to different orders at $y = y_i$, and they are all linear combinations of the $g_k(y)$. Therefore

$$h_{ij}(y) = \sum_{k=0}^{l-1} C_{ijk}\, g_k(y), \tag{19}$$

where the determinant

$$D_i = \left\| C_{ijk} \right\|_{k=0,1,\ldots,l-1}^{j=0,1,\ldots,l-1} \tag{20}$$

is a number different from zero.

Consider now the determinant

$$\Delta_i(x, y_i) = \left\| \left(\frac{\partial}{\partial x}\right)^{\mu} \left(\frac{\partial}{\partial y_i}\right)^{d_{ij}} R(x, y_i) \right\|_{j=0,1,\ldots,l-1}^{\mu=0,1,\ldots,l-1}. \tag{21}$$

By (11), (17), (19), (20) and the rule of multiplication of determinants,

$$\Delta_i(x, y_i) = L(x)\, D_i^{-1} \left\| \left(\frac{\partial}{\partial y_i}\right)^{d_{ij}} h_{ik}(y_i) \right\|_{k=0,1,\ldots,l-1}^{j=0,1,\ldots,l-1}. \tag{22}$$

The last determinant on the right of (22) is independent of x. The elements of the leading diagonal ($j = k$) are all non-zero, and the elements above the leading diagonal ($k > j$) are all zero. Hence the value of the determinant is a number E_i different from zero, and

$$\Delta_i(x, y_i) = E_i\, D_i^{-1} L(x). \tag{23}$$

By (8) and (9), the last element, and a fortiori each earlier element also, in column j of the determinant (21), vanishes at $x = x_i$ to the order

$$[\lambda(t_i - d_{ij})] - l + 2$$

at least. Hence the whole determinant vanishes at $x = x_i$ to the order

$$\sum_{j=0}^{l-1} \operatorname{Max}\{0, [\lambda(t_i - d_{ij}) - l + 2]\} \tag{24}$$

at least. Thus by (23) the non-zero polynomial $L(x)$ vanishes at $x = x_i$ to the order (24) at least. But the degree of $L(x)$ is lu at most, and the x_i are all distinct. Therefore

$$\sum_{i=0}^{n} \sum_{j=0}^{l-1} \text{Max}\,\{0, [\lambda(t_i - d_{ij}) - l + 2]\} \leq lu. \tag{25}$$

The determinant $N(y)$ is by (18), (19), and (20) equal to

$$D_i^{-1} \left\| h_{ij}^{(k)}(y) \right\|_{j=0,1,\ldots,l-1}^{k=0,1,\ldots,l-1}.$$

The general term in the expansion of the last determinant is

$$\pm\, h_{i0}^{(k_0)}(y)\, h_{i1}^{(k_1)}(y) \ldots h_{i,l-1}^{(k_{l-1})}(y), \tag{26}$$

where $k_0, k_1, \ldots, k_{l-1}$ is some permutation of the integers $0, 1, \ldots, l-1$. Now

$$h_{ij}^{(k_j)}(y)$$

vanishes at $y = y_i$ to the order $(d_{ij} - k_j)$ at least. Hence the product (26) vanishes at $y = y_i$ to the order

$$\sum_{j=0}^{l-1} (d_{ij} - k_j) = \sum_{j=0}^{l-1} (d_{ij} - j) \tag{27}$$

at least. Since (27) is independent of the particular term (26) chosen, the polynomial $N(y)$ vanishes at $y = y_i$ to the order (27) at least. Now it is possible to choose a set of linearly independent combinations $b_0(y), b_1(y), \ldots, b_{l-1}(y)$ of the $g_k(y)$, such that $b_0(y)$ is of degree s at most, $b_1(y)$ of degree $(s-1)$ at most, and so on. The degree of $N(y)$ is equal to the degree of the determinant

$$\left\| b_j^{(i)}(y) \right\|_{j=0,1,\ldots,l-1}^{i=0,1,\ldots,l-1},$$

which does not exceed the degree of a typical product

$$b_0^{(i_0)}(y)\, b_1^{(i_1)}(y) \ldots b_{l-1}^{(i_{l-1})}(y).$$

The degree of such a product does not exceed

$$\sum_{j=0}^{l-1} (s - j - i_j) = \sum_{j=0}^{l-1} (s - 2j) = l(s - l + 1). \tag{28}$$

The degree of the polynomial $N(y)$ is at most (28), and $N(y)$ vanishes to the orders (27) at the $(n+1)$ different values y_i. Therefore

$$\sum_{i=0}^{n}\sum_{j=0}^{l-1}(d_{ij}-j) \leq l(s-l+1). \tag{29}$$

Since for each i the d_{ij} are strictly increasing with j, the average value of $(d_{ij}-j)$ for $0 \leq j \leq b$ is a non-decreasing function of b. Thus (29) gives, for $0 \leq b \leq l-1$,

$$\sum_{i=0}^{n}\sum_{j=0}^{b}(d_{ij}-j) \leq (b+1)(s-l+1). \tag{30}$$

Let T be the largest $[t_i]$, and let b be the lesser of T and $(l-1)$. Then by (7)

$$0 \leq b \leq l-1, \qquad \lambda(T+1) \leq u+1.$$

Adding (30) multiplied by λ to (25), and using (12),

$$\sum_{i=0}^{n}\sum_{j=0}^{b} \text{Max}\,\{0, \lambda(t_i-j)\} \leq lu + (n+1)(b+1)(l-1) + \lambda(b+1)(s-l+1) \leq$$
$$\leq (u+1)\left(l + \frac{b+1}{T+1}(s-l+1)\right) + (n+1)ls. \tag{31}$$

Putting

$$b_i = \text{Min}\,\{[t_i], l-1\},$$

this becomes

$$\sum_{i=0}^{n}(1+b_i)(\lambda t_i - \tfrac{1}{2}\lambda b_i) \leq (u+1)\left(l + \frac{b+1}{T+1}(s-l+1)\right) + (n+1)ls. \tag{32}$$

Now it is an elementary principle that if $\alpha \geq \beta > 0$ and $\gamma \geq 0$ then $(\beta/\alpha) \leq \leq ((\beta+\gamma)/(\alpha+\gamma))$. Applying this with $\gamma = s + [t_i] + 1 - 2t_i$, which is positive by (7),

$$(2t_i - [t_i])/(2t_i - b_i) \leq (s+1)/(s+1+[t_i]-b_i);$$

and applying it with $\gamma = T - [t_i]$,

$$(1+[t_i])/(s+1+[t_i]-b_i) \leq (1+T)/(1+s+T-b_i).$$

These two inequalities combine to give

$$\frac{(1+[t_i])(2t_i-[t_i])}{(1+b_i)(2t_i-b_i)} \leq \frac{(1+T)(1+s)}{(1+b_i)(1+s+T-b_i)}. \tag{33}$$

In the same way, if $T \geq l-1$ the principle with $\gamma = s+1-l$ gives

$$l/(T+1) \leq (1+s)/(2+s+T-l),$$

which may be written

$$(34)\qquad 1 \leq \frac{(1+T)(1+s)}{l(2+s+T-l)}.$$

Now either $b_i = l-1$ or $b_i = [t_i]$. In the first case the right sides of (33) and (34) are the same, and in the second case the left sides are the same. If $T < l-1$ the second case holds for every value of i. Hence (33) and (34) together give for $i = 0, 1, \ldots, n$,

$$(35)\qquad \frac{(1+[t_i])(t_i - \frac{1}{2}[t_i])}{(1+b_i)(t_i - \frac{1}{2}b_i)} \leq \begin{cases} \dfrac{(1+T)(1+s)}{l(2+s+T-l)} & \text{if } T \geq l-1, \\ 1 & \text{if } T < l-1. \end{cases}$$

If $T < l-1$, then $b = T$, and (32) and (35) give

$$(36)\qquad \lambda \sum_{i=0}^{n} (1+[t_i])(t_i - \tfrac{1}{2}[t_i]) \leq (u+1)(s+1)\left\{1 + \frac{(n+1)ls}{(u+1)(s+1)}\right\}.$$

If $T \geq l-1$, then $b = l-1$, and (32) and (35) give

$$(37)\qquad \lambda \sum_{i=0}^{n} (1+[t_i])(t_i - \tfrac{1}{2}[t_i]) \leq (u+1)(s+1)\left\{1 + \frac{(1+T)(n+1)s}{(u+1)(2+s+T-l)}\right\}.$$

Substituting for l from (12) and for s from (7), (36) and (37) imply the truth of (10), whether $T < l-1$ or not. This completes the proof of the lemma.

III. The second part of the proof of the theorem, as explained at the beginning of this paper, starts at this point. In the first place, the proof may be reduced to the case in which θ is an algebraic integer. For suppose that any algebraic number θ satisfies the conditions of the theorem. There exists a rational integer N such that $N\theta$ is an algebraic integer, also of degree n, and by (1)

$$(38)\qquad |N\theta - (Np/q)| < Nq^{-\mu}$$

for an infinity of p/q. If ε is any positive number, then (38) implies

$$(39)\qquad |N\theta - (Np/q)| < q^{-\mu+\varepsilon}$$

for an infinity of Np/q. If the theorem is true for algebraic integers, (39) implies

$$\mu \leq \sqrt{(2n)} + \varepsilon,$$

and since ε is arbitrarily small

$$\mu \leq \sqrt{(2n)}.$$

The theorem is therefore true for any algebraic number θ, if it is proved for θ an algebraic integer.

In what follows it will be assumed that θ is an algebraic integer of degree n and satisfies the conditions of the theorem; θ will be a root of an equation

$$a(\theta) \equiv \theta^n + a_1 \theta^{n-1} + \cdots + a_n = 0 \tag{40}$$

with integer coefficients a_i. This equation, being the equation of lowest degree satisfied by θ, has n distinct real or complex roots

$$\theta = \theta_1, \theta_2, \theta_3, \ldots, \theta_n. \tag{41}$$

Let

$$A = \mathrm{Max}\{1, |a_1|, |a_2|, \ldots, |a_n|\}. \tag{42}$$

Two arbitrary positive integers s and t and a real number δ are now chosen, subject only to the conditions

$$t + 1 \leq s + 1 \leq \tfrac{1}{2} n(t + 1), \tag{43}$$

$$0 < \delta < (2n)^{-3}. \tag{44}$$

A fraction p_1/q_1 can be found which satisfies (1) and also

$$\log q_1 > 6\delta^{-2} t \log 4A. \tag{45}$$

A fraction p_2/q_2 can be found which satisfies (1) and also

$$\log q_2 > 2\delta^{-1} \log q_1. \tag{46}$$

Let

$$\lambda = \log q_2 / \log q_1, \tag{47}$$

$$u = [\tfrac{1}{2} n t (t + 1)\lambda/(s + 1)], \tag{48}$$

$$B = [q_1^{\delta u}]. \tag{49}$$

By (48) and (43)

$$u + 1 \geq \lambda t \geq \tfrac{1}{2}\lambda(t + 1) \geq \lambda(s + 1)/n. \tag{50}$$

By (46), (47) and (44)

$$\delta\lambda > 2, \qquad \lambda > 16 n^3. \tag{51}$$

By (49) and (45)

$$\log(B + 1) \geq 6 u t \delta^{-1} \log 4A. \tag{52}$$

By (51) and (50)

$$s < \tfrac{1}{2} n \delta(u + 1). \tag{53}$$

Consider now the set P of all polynomials

$$V(x,y) = \sum_{i=0}^{u} \sum_{j=0}^{s} v_{ij} x^i y^j,$$

where the v_{ij} are integers between 0 and B inclusive. The number of polynomials in P is

$$N = (B + 1)^{(s+1)(u+1)}. \tag{54}$$

Given one such polynomial $V(x,y)$, there is a set $D(V)$ of derived polynomials

$$\frac{1}{\alpha!\,\beta!} \left(\frac{\partial}{\partial x}\right)^{\alpha} \left(\frac{\partial}{\partial y}\right)^{\beta} V(x,y)$$

where α and β vary over all pairs of integers satisfying the conditions

$$0 \leq \beta \leq t - \delta, \qquad 0 \leq \alpha \leq \lambda(t - \beta - \delta).$$

$D(V)$ contains

$$N_0 = \sum_{\beta=0}^{t-1} [1 + \lambda(t - \beta - \delta)] \leq t + \tfrac{1}{2}\lambda t(t+1) - \delta t \lambda \tag{55}$$

polynomials. Each polynomial in $D(V)$ has coefficients which are obtained by multiplying one of the v_{ij} by two binomial coefficients, so that the coefficients of polynomials in $D(V)$ are integers not exceeding

$$B\, 2^{u+s}.$$

If $S(x,y)$ is a polynomial in $D(V)$, let $T_S(x)$ be the remainder when $S(x,x)$ is divided by the polynomial $a(x)$ of (40) by the ordinary long-division process. $T_S(x)$ is a polynomial in x of degree not exceeding $(n-1)$ and with integer coefficients. If $U_r(x)$ is the remainder after r steps of the long-division process, the coefficients of $U_{r+1}(x)$ are of the form $(u_1 - a_i u_0)$ where u_1 and u_0 are coefficients of $U_r(x)$ and a_i is a coefficient of $a(x)$. Hence the magnitude of the largest coefficient of $U_{r+1}(x)$ does not exceed $(1 + A)$ times that of $U_r(x)$. But $U_0(x) = S(x,x)$ and $T_S(x) = U_r(x)$ for some $r \leq u + s$. Therefore the coefficients of $T_S(x)$ are majorised by

$$M = (1 + A)^{u+s}(1 + s)\, B\, 2^{u+s}. \tag{56}$$

The number of coefficients of all the polynomials $T_S(x)$ with $S(x,y)$ in $D(V)$ does not exceed

$$n N_0 \leq nt + \tfrac{1}{2} nt(t+1)\lambda - \delta n t \lambda.$$

Each of these coefficients is a positive or negative integer majorised by M, and so can take only $(2M+1)$ distinct values. Hence the total number of possible distinct sets of values for the coefficients of all the $T_S(x)$ with $S(x,y)$ in $D(V)$ does not exceed

$$N' = (2M+1)^{nt+\frac{1}{2}nt(t+1)\lambda-\delta nt\lambda}.$$

By (56) and trivial inequalities

$$\log N' < nt(\tfrac{1}{2}\lambda(t+1)+1-\delta\lambda)(\log 2+(u+s)\log(2A+2)+\log(s+1)+\log(B+1))$$
$$< nt(\tfrac{1}{2}\lambda(t+1)+1-\delta\lambda)\log(B+1)+nt(1+\tfrac{1}{2}\lambda(t+1))(u+2s+2)\log 4A.$$

Now by (50) and (51)

$$u+2s+2 \leq u+2n(u+1)/\lambda < u(1+(4n/\lambda)) < 2u,$$
$$1-\delta\lambda < -\tfrac{1}{2}\delta\lambda, \qquad 1+\tfrac{1}{2}\lambda(t+1) \leq 1+\lambda t < \tfrac{3}{2}\lambda t.$$

Hence the inequality for $\log N'$ becomes

$$\log N' < nt(\tfrac{1}{2}\lambda(t+1)-\tfrac{1}{2}\delta\lambda)\log(B+1)+3nt^2\lambda u\log 4A.$$

Finally, by (52) and (48), this gives

$$\log N' < \tfrac{1}{2}nt\lambda(t+1)\log(B+1) < (s+1)(u+1)\log(B+1) = \log N,$$

and therefore

$$N' < N. \tag{57}$$

In view of (57) there must exist two distinct polynomials $V_1(x,y)$ and $V_2(x,y)$ in P such that the polynomials $T_S(x)$ derived from $V_1(x,y)$ as described above are all identically equal to the corresponding polynomials derived from $V_2(x,y)$. Let

$$R(x,y) = V_1(x,y) - V_2(x,y). \tag{58}$$

$R(x,y)$ is a polynomial of degrees not exceeding u in x and s in y; its coefficients are positive or negative integers of magnitude not exceeding B; it is not identically zero; and it has the property that the polynomials in z

$$\frac{1}{\alpha!\,\beta!}\left(\frac{\partial}{\partial x}\right)^{\alpha}\left(\frac{\partial}{\partial y}\right)^{\beta} R(x,y)\Big|_{x=y=z} \tag{59}$$

are exactly divisible by $a(z)$ for all pairs of integers α and β such that

$$0 \leq \beta \leq t-\delta, \qquad 0 \leq \alpha \leq \lambda(t-\beta-\delta). \tag{60}$$

The arbitrary polynomial $R(x, y)$ of the lemma will now be identified with the polynomial defined by (58). The numbers θ_i given by (41) are all distinct, and they are all irrational and therefore distinct from p_1/q_1 and p_2/q_2. The lemma can be applied with

(61) $$x_0 = p_1/q_1, \quad y_0 = p_2/q_2; \quad x_i = y_i = \theta_i, \quad i = 1, 2, \ldots, n.$$

By (59) and (60), the conditions (8) and (9) of the lemma are satisfied for $i = 1, 2, \ldots, n$ with

(62) $$t_i = t - \delta.$$

Suppose, if possible, that (8) and (9) are also satisfied for $i = 0$ with

(63) $$t_0 = (n + 1)(t + 1)\sqrt{(\tfrac{1}{2} n \delta)}.$$

By (44)

(64) $$t_0 \leq (n + 1)(t + 1)/4n \leq \tfrac{3}{8}(t + 1) \leq \tfrac{3}{4} t \leq t - \tfrac{1}{4} < t - \delta,$$

and by (50)

(65) $$\lambda[t_i + 1] \leq \lambda[t - \delta + 1] = \lambda t \leq u + 1, \quad i = 0, 1, \ldots, n.$$

The conditions of the lemma are then all satisfied, (7) being a consequence of (44), (51), (53), (43), (64) and (65). The lemma therefore leads to the result

$$\lambda(1 + [t_0])(t_0 - \tfrac{1}{2}[t_0]) + \tfrac{1}{2} n t \lambda (t - 2\delta + 1) \leq (1 + \tfrac{1}{2} n(n + 1)\delta)(s + 1)(u + 1).$$

By (48) this gives

(66) $$\lambda(1 + [t_0])(t_0 - \tfrac{1}{2}[t_0]) \leq \delta n t \lambda + \tfrac{1}{2} n(n + 1)\delta(s + 1)(u + 1) + s + 1.$$

Now by (51), (48) and (43),

$$\begin{aligned}
&\delta n t \lambda + \tfrac{1}{2} n(n + 1)\delta(s + 1)(u + 1) + s + 1 \\
&\quad \leq \delta n t \lambda + \tfrac{1}{2} n(n + 1)\delta u(s + 1) + \tfrac{1}{2} n(n + 1)\delta(s + 1) + \tfrac{1}{2}\delta\lambda(s + 1) \\
&\quad \leq \delta n t \lambda + \tfrac{1}{4} n^2(n + 1)\delta\lambda t(t + 1) + \tfrac{1}{2}\delta(\lambda + n(n + 1))\tfrac{1}{2} n(t + 1) \\
&\quad = \delta\lambda n\{\tfrac{1}{4} n(n + 1)t(t + 1) + t + \tfrac{1}{4}(t + 1) + n(n + 1)(t + 1)/4\lambda\} \\
&\quad < \delta\lambda n\{\tfrac{1}{4} n(n + 1)t(t + 1) + t + \tfrac{1}{2}(t + 1)\} \\
&\quad < \delta\lambda n\{\tfrac{1}{4} n(n + 1)t(t + 1) + \tfrac{3}{4}(t + 1)^2\} \leq \tfrac{1}{4}\delta\lambda n(t + 1)^2\{n^2 + n + 3\} \\
&\quad \leq \tfrac{1}{4}\delta\lambda n(t + 1)^2(n + 1)^2.
\end{aligned}$$

Therefore (66) implies

$$\tfrac{1}{2}\lambda t_0^2 < \lambda(1 + [t_0])(t_0 - \tfrac{1}{2}[t_0]) < \tfrac{1}{4}\delta\lambda n(t+1)^2(n+1)^2,$$

which contradicts (63). This means that the hypothesis that (8) and (9) are satisfied for $i = 0$ with t_0 given by (63) is untenable. That is to say, there exist two integers α and β such that

$$\left(\frac{\partial}{\partial x}\right)^{\alpha}\left(\frac{\partial}{\partial y}\right)^{\beta} R(x,y)\bigg|_{x=p_1/q_1,\ y=p_2/q_2} \neq 0, \tag{67}$$

while

$$\alpha + \lambda\beta \leq \lambda t_0. \tag{68}$$

Let one pair of integers α and β be chosen to satisfy (67) and (68), and let

$$R_0(x,y) = \frac{1}{\alpha!\,\beta!}\left(\frac{\partial}{\partial x}\right)^{\alpha}\left(\frac{\partial}{\partial y}\right)^{\beta} R(x,y). \tag{69}$$

By (59) and (60),

$$\left(\frac{\partial}{\partial x}\right)^{i}\left(\frac{\partial}{\partial y}\right)^{j} R_0(x,y)\bigg|_{x=y=\theta} = 0 \tag{70}$$

whenever

$$0 \leq j + \beta \leq t - \delta, \quad 0 \leq i + \alpha \leq \lambda(t - \beta - j - \delta).$$

By (68) this implies that (70) holds whenever

$$i + \lambda j \leq \lambda(t - \delta - t_0) = \lambda(t-\delta)(1-\varepsilon), \tag{71}$$

where ε is defined by the equation

$$\varepsilon(t-\delta) = t_0 = (n+1)(t+1)\sqrt{(\tfrac{1}{2}n\delta)}. \tag{72}$$

Now $R_0(x,y)$ has the Taylor expansion

$$R_0(x,y) = \sum_{i=0}^{u}\sum_{j=0}^{s} c_{ij}(x-\theta)^i(y-\theta)^j. \tag{73}$$

The general coefficient in the expansion is

$$c_{ij} = \frac{1}{i!\,j!}\left(\frac{\partial}{\partial x}\right)^{i}\left(\frac{\partial}{\partial y}\right)^{j} R_0(x,y)\bigg|_{x=y=\theta} \tag{74}$$

$$= \sum_{\varrho=0}^{u}\sum_{\sigma=0}^{s} a_{\varrho\sigma}\left(\frac{\varrho!}{i!\,\alpha!\,(\varrho-i-\alpha)!}\right)\left(\frac{\sigma!}{j!\,\beta!\,(\sigma-j-\beta)!}\right)\theta^{\varrho-i-\alpha+\sigma-j-\beta},$$

where the $a_{\varrho\sigma}$ are the coefficients of the polynomial $R(x, y)$. Remembering that p_1/q_1 and p_2/q_2 both satisfy (1), (73) and (47) together give

$$(75)\qquad \left| R_0\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right)\right| \leq \sum_{i=0}^{u}\sum_{j=0}^{s} |c_{ij}|\, q_1^{-\mu i} q_2^{-\mu j} = \sum_{i=0}^{u}\sum_{j=0}^{s} |c_{ij}|\, q_1^{-\mu(i+\lambda j)}.$$

But by (70), all the c_{ij} for which (71) is true are zero, and therefore (75) implies

$$(76)\qquad \left| R_0\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right)\right| \leq q_1^{-\mu\lambda(t-\delta)(1-\varepsilon)} \sum_{i=0}^{u}\sum_{j=0}^{s} |c_{ij}|.$$

Since

$$|a_{\varrho\sigma}| \leq B,$$

and the expressions

$$(\varrho!/i!\,\alpha!\,(\varrho - i - \alpha)!)$$

are terms in the multinomial expansion of

$$(1 + 1 + 1)^{\varrho},$$

the coefficients c_{ij} given by (74) are majorised by

$$B(u + 1)(s + 1)\, 3^{u+s} m^{u+s},$$

where $m = \text{Max}\ (1, |\theta|)$. Therefore (76) gives finally

$$(77)\qquad \left| R_0\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right)\right| \leq (u + 1)^2 (s + 1)^2 B(3m)^{u+s} q_1^{-\mu\lambda(t-\delta)(1-\varepsilon)}.$$

On the other hand, $R_0(x, y)$ is by (69) a polynomial of degrees not exceeding u and s in x and y respectively, and all its coefficients are integers. Therefore

$$R_0(p_1/q_1, p_2/q_2)$$

is a rational fraction with denominator not exceeding

$$q_1^u q_2^s.$$

This fraction is not zero, by (67). Therefore

$$(78)\qquad \left| R_0\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right)\right| \geq q_1^{-u} q_2^{-s} = q_1^{-u-\lambda s}.$$

Together, (77) and (78) give

$$\{u + \lambda s - \lambda\mu(t - \delta)(1 - \varepsilon)\}\log q_1 + \log B + (u + s)\log(3m) + 2\log((u + 1)(s + 1)) \geq 0.$$

Now

$$\log((u + 1)(s + 1)) \leq u + s.$$

Hence, substituting for $\log B$ from (49),

$$(79)\qquad \{u(1+\delta)+\lambda(s-\mu(t-\delta)(1-\varepsilon))\}\log q_1+(u+s)(2+\log(3m))\geq 0.$$

By (45), (44), (48) and (51) in turn,

$$n\delta^2\lambda(1+\delta s)\log q_1 > 6n\lambda t(1+\delta s)\log 4 > \tfrac{1}{2}n\lambda t+\lambda\delta s > u+s.$$

Therefore (79) implies

$$u(1+\delta)+\lambda(s-\mu(t-\delta)(1-\varepsilon))+n\delta^2\lambda(1+\delta s)(2+\log(3m)) > 0.$$

Hence, multiplying by $(s+1)/\lambda$ and using (48),

$$(80)\qquad \tfrac{1}{2}nt(t+1)(1+\delta)+(s+1)\{s-\mu(t-\delta)(1-\varepsilon)+n\delta^2(1+\delta s)(2+\log(3m))>0.$$

Since δ was chosen subject only to the condition (44), it is allowable to make δ tend to zero in (80). Of the quantities occurring in (80), all except ε are independent of δ, and ε tends to zero with δ according to (72). Thus in the limit (80) becomes

$$(81)\qquad \mu\leq\frac{s}{t}+\frac{1}{2}n\frac{t+1}{s+1}.$$

The inequality (81) is proved for any two positive integers s and t satisfying the conditions (43). In particular, the choice $t=1$ gives precisely Siegel's result (3). The best bound for μ deducible from (81) is however obtained by letting s and t tend together to infinity in such a way that the ratio s/t tends to $\sqrt{(\frac{1}{2}n)}$, which can always be done without violating the conditions (43). In the limit, as s and t tend to infinity in this way, (81) becomes (2), and so the theorem is proved.

IV. The fact that the bound (2) obtained for μ is proportional to the square-root of n is directly attributable to the two variables of the polynomial $R(x,y)$ of the lemma. If a corresponding lemma could be proved for a polynomial in three variables, a bound for μ would be obtained proportional to the cube-root of n. If a similar lemma could be proved for polynomials in an arbitrarily large number of variables, the full result (4) could be deduced. The present paper probably represents the limit of what can be done with two variables only. Further progress must wait upon a fundamental investigation of the properties of polynomials in a larger number of variables; such an investigation, with a proof of (4) as the final objective, would not be in any way a hopeless undertaking.

References.

(1) C. Siegel, *Math. Zeit.* 10, (1921), 173—213.
(2) ——, *Math. Ann.* 84, (1921), 80—99.
(3) S. Kuroda, *Proc. Imp. Acad. Japan,* 10, (1934), 619—622.
(4) T. Schneider, *Jour. Reine u. Ang. Math.* 175, (1936), 182—192.
(5) J. Koksma, *Diophantische Approximationen* (*Ergebnisse der Mathematik,* Bd. 4, Berlin 1937), 55.
(6) K. Mahler, *Proc. Akad. Wet. Amst.* 39, (1936), 633—640, 729—737.
(7) E. Landau, *Vorlesungen über Zahlentheorie,* (Leipzig 1927), Bd. 3, 41—55.