ACTA ARITHMETICA
XIII (1968)

# A method in diophantine approximation (II)*

by

CHARLES F. OSGOOD (Washington, D. C.)

This paper applies the general method of approach developed in [2] to obtain further results about the diophantine approximation of values of certain functions. We shall repeat the theorem demonstrated in [2], calling it Theorem I here; state and prove Theorem II, which is a generalization of Theorem I; and obtain as a special case of Theorem II the principal result of this paper, Theorem III. Since all three theorems are quite abstractly stated we begin with a proposition which is a special case of Theorem III, followed by two detailed examples.

**Section I.** Let $D$ denote differentiation with respect to the complex variable $z$; let $l$ be an integer greater than one; and let each $g_j(z)$ $(1 \leqslant j \leqslant l)$ be a polynomial of degree less than $j$ with coefficients in the Gaussian field. Suppose that we are in a simply connected region $X$ where $a(z)$ is analytic and that $y_1, \ldots, y_n$ are $n \geqslant 1$ solutions of

$$(1)\qquad y = \sum_{j=1}^{l} g_j(z) D^j y + a(z)$$

which are analytic in some open disk $N \subseteq X$ with center $z_0$. Suppose $z_1$ belongs to $N$, $z_1$ is a Gaussian rational, and $0 \notin g_l(N)$. Let $\mathscr{C}$ be a differentiable path in $X$ with endpoints at $z_0$ which does not pass through any of the zeros of $g_l(z)$. Suppose that $\tilde{y}_1 \neq y_1, \ldots, \tilde{y}_n \neq y_n$ are the function elements analytic on $N$ obtained by extending $y_1, \ldots, y_n$ respectively about $\mathscr{C}$ and that the $y_j - \tilde{y}_j$ are linearly independent. Let

$$d = \max_i \frac{\deg g_i(z)}{i - \deg g_i(z)}.$$

Let $\|a\|$, for $a$ any complex number, denote the distance from $a$ to the nearest Gaussian integer. Let $(A_1, \ldots, A_n)$ denote any nonzero element of the cartesian product of the Gaussian integers with themselves $n$ times.

* This paper was supported in part by the National Science Foundation contract number GP 54-78 and in part by the National Bureau of Standards where the author is presently a postdoctoral fellow.

Proposition. *For each $\varepsilon > 0$ there exists a $c(\varepsilon) > 0$ such that*

$$\max_{0\leqslant i<l}\Big\{\Big\|\sum_{j=1}^{n} A_j D^i(y_j-\tilde{y}_j)(z_1)\Big\|\Big\} \geqslant c(\varepsilon)\min_j\{|A_j|^{-(d+\varepsilon)}\}$$

*for all $(A_1, \ldots, A_n)$.*

The proof of the Proposition is located after the proof of Theorem III.

Example I. Suppose that (1) is of the form

$$y = D\varphi(zD)y, \tag{2}$$

where $zD\varphi(zD)$ is a polynomial in $zD$ with coefficients in the Gaussian integers of degree $l > 1$ no two of whose roots differ by a rational integer. In this case the $l-1$ linearly independent non-entire solutions of (2) are

$$y_j = \sum_{n=0}^{\infty} \frac{z^{n+r_j}}{\prod_{k=1}^{n}\{\varphi(k+r_j)(k+r_j)\}} \qquad (1 \leqslant j \leqslant l-1),$$

where each $r_j$ is a zero of $\varphi(m)$. (See [3] for results about the entire solution of (2).) Now let $\mathscr{C}$ be a circular path winding once in the positive direction about the origin. Each $y_j - \tilde{y}_j = (1-e^{2\pi i r_j})y_j$; thus, the $y_j-\tilde{y}_j$ are a linearly independent set of functions over $C$. It would appear that we should apply the Proposition to the $y_j-\tilde{y}_j$. However, it is more interesting to apply it to the functions $Y_j - \tilde{Y}_j$, defined below.

Let

$$y = \sum_{j=1}^{l-1} C_j(y_j-\tilde{y}_j)$$

be such that $D^i y(z_1) = B_i$ $(0 \leqslant i \leqslant l-2)$ where each $B_i$ is a Gaussian integer. It must be possible to find such a $y$, since otherwise the Wronskian of the $l-1$ functions $y_j-\tilde{y}_j$ would vanish at $z_1$ and then we could construct a linear combination (over $C$) of the $y_j-\tilde{y}_j$ which has a zero of order $l-1$ at $z_1$ and whose $(l-1)$-st derivative at $z_1$ is a Gaussian integer. This would violate the Proposition for $A_1 = 1$ and $A_2 = A_3 = \ldots = A_n = 0$. Therefore the $C_j$ exist and by Cramer's rule

$$C_j = \Delta_j(\Delta)^{-1},$$

where $\Delta$ is the Wronskian of the $y_j-\tilde{y}_j$ at $z_1$ and $\Delta_j$ differs from $\Delta$ in that the matrix of which it is the determinant has

$$\begin{pmatrix} B_0 \\ \cdot \\ \cdot \\ \cdot \\ B_{l-2} \end{pmatrix}$$



in the $j$th column. Then $y = \sum_j \Delta_j\cdot(\Delta)^{-1}\cdot(y_j-\tilde{y}_j)$. Expanding out each $\Delta_j$ and writing $y$ as a linear combination of the $B_j$ we have

$$y = \sum_{j=0}^{l-2} B_j(Y_j - \tilde{Y}_j),$$

for some set of functions $Y_j$ each of which satisfies (2). Since $(B_0, \ldots, B_{l-2})$ is an arbitrary element of the cartesian product of the Gaussian integers with themselves $l-1$ times we must have that

$$D^i(Y_j-\tilde{Y}_j)(z_1) = \delta_j^i.$$

Therefore the $Y_j - \tilde{Y}_j$ are linearly independent functions over $C$. We now apply the Proposition to the functions $Y_j$ and replace each $B_i$ by $A_{i+1}$ where $(A_1, \ldots, A_{l-1}) \neq (0, \ldots, 0)$. Our conclusion is that

$$\Big\|\sum_{j=0}^{l-2} B_j D^{l-1}(Y_j-\tilde{Y}_j)(z_1)\Big\| \geqslant c(\varepsilon)\min_j\{|A_j|^{-(l-1+\varepsilon)}\};$$

whence, for any Gaussian integer $A_l$

$$\Big|\sum_j \Delta_j(1-e^{2\pi i r_j})y_j(z_1) + A_l\Delta\Big| \geqslant c(\varepsilon)|\Delta|\min_j\{|A_j|^{-(l-1+\varepsilon)}\}.$$

Dividing out by a nonzero factor, we obtain the result that for an appropriate $c_1(\varepsilon) > 0$

$$\Big|\sum_j \Delta_j' D^{l-1}y_j(z_1) + A_l\Delta'\Big| \geqslant c_1(\varepsilon)\min_j\{|A_j|^{-(l-1+\varepsilon)}\},$$

where $\Delta_j'$ and $\Delta'$ are determinants which differ from $\Delta_j$ and $\Delta$, respectively, only in that $y_j(z_1)$ has been substituted for

$$y_j(z_1) - \tilde{y}_j(z_1) = (1-e^{2\pi i r_j})y_j(z_1)$$

throughout.

Thus

$$\left|\begin{vmatrix} A_1 & y_1(z_1) & \ldots & y_{l-1}(z_1) \\ \cdot & \cdot & \cdot & \cdot \\ A_l & D^{l-1}y_1(z_1) & \ldots & D^{l-1}y_{l-1}(z_1) \end{vmatrix}\right| \geqslant c_1(\varepsilon)\min_j\{|A_j|^{-(l-1+\varepsilon)}\},$$

as may be seen by expanding the above determinant by minors along the bottom row. Note that our condition that $(A_1, \ldots, A_{l-1}) \neq (0, \ldots, 0)$ may be replaced by the condition that $(A_1, \ldots, A_l) \neq (0, \ldots, 0)$, if we allow $c_1(\varepsilon)$ to be possibly smaller than above. If the $D^i y_j(z_1)$ are all real then for rational integers $A_i$ this is a best possible statement.

EXAMPLE II. Consider a function of the form

$$y = z^{-a}\exp\Big(\sum_{i=1}^{l-1} r_i z^{-i}\Big) \tag{3}$$

where $a$ and the $r_i$ are each rational, $a > l+1$, $a$ is not an integer, and $r_{l-1} \neq 0$.

Now

$$\frac{y'}{y} = -az^{-1} + \sum_{i=1}^{l-2}(-i)r_i z^{-(i+1)} - (l-1)r_{l-1}z^{-l} \tag{4}$$

or

$$z^l y' = \Big(-az^{l-1} + \sum_{i=1}^{l-2}(-i)r_i z^{l-1-i} - (l-1)r_{l-1}\Big)y. \tag{5}$$

It is easily shown that the integrals

$$\frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} z^j y(z)e^{sz}dz = \frac{d^j}{ds^j}\Big(\frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} y(z)e^{sz}dz\Big)$$

exist for all real $s$ if $j = 0, 1, \ldots,$ or $l-1$, as does

$$\frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} z^l y'(z)e^{sz}dz = \frac{d^l}{ds^l}\Big(\frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} y'(s)e^{sz}dz\Big)$$

$$= \frac{d^l}{ds^l}\Big(-s\int_{1-i\infty}^{1+i\infty} y(z)e^{sz}dz\Big).$$

Set

$$Y(s) = \int_{1-i\infty}^{1+i\infty} y(z)e^{sz}dz,$$

which is the inverse Laplace transform of $y(z)$. Then

$$D^l(-s)Y(s) = \Big(-aD^{l-1} + \sum_{i=1}^{l-1}(-i)r_i D^{l-1-i} - (l-1)r_{l-1}\Big)Y(s). \tag{6}$$

We may put (6) in the same form as (1), then we shall apply the Proposition to a collection of functions consisting of the one function $KY(s)$, where $K$ is some appropriate nonzero number. The curve $\mathscr{C}$ will be a circular path once about the origin in the positive direction. First, however, we shall calculate $Y(s)$ explicitly and show that it may be defined as a multiplevalued function in the plane with $Y(s) - \tilde{Y}(s) \neq 0$.



Clearly

$$y(s) = \prod_i \big(z^{-a/(l-1)}\exp(r_i z^i)\big).$$

Note that each factor on the right side above has an inverse Laplace transform. An inverse Laplace transform takes an ordinary product of functions into the convolution product (denoted by $*$) of their inverse transform. Thus the first step is to calculate the transform of each factor, i.e.

$$z^{-a/(l-1)}\exp(r_i z^{-i}) = \sum_{n=0}^{\infty}\frac{r_i^n z^{-n-a/(l-1)}}{n!}.$$

We may transform the above series term by term. The only fact then about inverse Laplace transforms which we need to know is that $z^{-a}$ $(a > 0)$ goes into $s^{a-1}/\Gamma(a)$. Thus

$$\sum_{n=0}^{\infty}\frac{r_i^n s^{n+a/(l-1)-1}}{n!\,\Gamma\big(n+a/(l-1)\big)}$$

is the inverse Laplace transform of $z^{-a/(l-1)}\exp(r_i z^{-i})$. Using the identity

$$\frac{s^{a-1}}{\Gamma(a)} * \frac{s^{\beta-1}}{\Gamma(\beta)} = \int_0^s \frac{(s-t)^{a-1}}{\Gamma(a)}\cdot\frac{t^{\beta-1}}{\Gamma(\beta)}dt = \frac{s^{a+\beta-1}}{\Gamma(a+\beta)}$$

we obtain

$$Y(s) = s^{a-1}\sum_{n_1,\ldots,n_{l-1}=0}^{\infty}\frac{r_1^{n_1}\ldots r_{l-1}^{n_{l-1}} s^{\sum_{i=1}^{l-1} in_i}}{n_1!\ldots n_{l-1}!\,\Gamma\big(\sum in_i + a\big)}.$$

What we have shown then implies that

$$\Gamma(a)Y(s) = s^{a-1}\sum_{n_1,\ldots,n_{l-1}=0}^{\infty}\frac{r_1^{n_1}\ldots r_{l-1}^{n_{l-1}} s^{\sum_{i=1}^{l-1} in_i}}{n_1!\ldots n_{l-1}!\,\Gamma\big(\sum in_i + a\big)\big(\Gamma(a)\big)^{-1}}$$

is a solution of (6) if $a > l$. By analytic continuation with respect to $a$, if $a$ is not zero or a negative integer, $\Gamma(a)Y(s)$ is always a solution of (6). We therefore drop the restriction that $a > l+1$ and assume merely that $a$ and the $r_i$ are rational numbers, $a$ is not an integer, and $r_{l-1} \neq 0$. Set $s_1 = r_l \neq 0$. Then one of the numbers

$$D^p\Gamma(a)\big(Y(r_l) - \tilde{Y}(r_l)\big), \quad 0 \leqslant p \leqslant l-1,$$

is nonzero.

We apply the Proposition to the function

$$\frac{\Gamma(a)Y(s)}{\Gamma(a)D^p\big(Y(r_l) - \tilde{Y}(r_l)\big)} = \frac{Y(s)}{(1-e^{2\pi i a})D^pY(r_l)}.$$

As $d = (l-1)^{-1}$ we obtain

(7) $$\max_{j\neq p}\left\{\left\| A_1 \frac{D^j Y(r_l)}{D^p Y(r_l)}\right\|\right\} \geqslant c(\varepsilon)|A_1|^{-(1/(l-1)+\varepsilon)}.$$

Notice that the numbers being approximated are all real. Khintchine's transference principle (see [1]; p. 80) enables one to say (see [3]; p. 80) that

(8) $$\left\|\sum_{j\neq p} A_j \frac{D^j Y(r_l)}{D^p Y(r_l)}\right\| \geqslant c(\varepsilon)\min_{j\neq p}\{|A_j|^{-(l-1+\varepsilon)}\}$$

for all nonzero $(A_1, \ldots, A_{p-1}, A_{p+1}, \ldots, A_{l-1})$ belonging to the cartesian product of the integers with themselves $l-2$ times, i.e. $Z^{l-2}$. Multiplying by $r_l^{1-\alpha}\Gamma(\alpha)D^p y(r_l)$ we obtain

(9) $$\left|\sum_{j=1}^{l-1} A_j r_l^{1-\alpha} D^j \Gamma(\alpha) y(r_l)\right| \geqslant c_1(\varepsilon)\min_j\{|A_j|^{-(l-1+\varepsilon)}\}$$

for all nonzero $(A_1, \ldots, A_{l-1})$ belonging to $Z^{l-1}$. Line (9) is a best possible statement. Further, the numbers $r_l^{1-\alpha}D^j\Gamma(\alpha)y(r_l)$ may be replaced in (9) by the numbers $r_l^{1-\alpha+j}D^j\Gamma(\alpha)y(r_l)$, i.e.

$$\sum_{n_1,\ldots,n_l=0}^{\infty} \frac{r_1^{n_1}\ldots r^{n_{l-1}} r_l^{n_1+\ldots+(l-1)n_{l-1}}}{n_1!\ldots n_{l-1}!\,\Gamma(n_1+\ldots+(l-1)n_{l-1}+\alpha-j)\,(\Gamma(\alpha))^{-1}};$$

where $j = 0, 1, \ldots, l-1$ for a new constant $c(\varepsilon)$.

**Section II.** Now we present Theorem I, which was proven in [2].

Suppose that: (I) $y$ is a function from a set $S$ to $R^m$ (the $m$ by one matrices over $R$); (II) $U$ is a vector space of functions from $S$ to $R^m$ over the field $R$; (III) $T$ is a linear operator and $U_1 \supseteq U_2 \supseteq \ldots \supseteq U_l$ $(l \geqslant 2)$ are subspaces of $U$ such that $T^i$ is defined from $U_i$ to $U$ $(1 \leqslant i \leqslant l)$; (IV) $y$ belongs to $U_l$; (V) $M$ is a vector space over $R$ of functions from $S$ to the $m$ by $m$ matrices over $R$; (VI) $\Phi$ is a function from $M$ to $M$; (VII) if $f$ belongs to $U_1$ and $g$ belong to $M$, then $gf$ belongs to $U_1$ and $Tgf = gTf + \Phi(g)f$; (VIII) we have

(10) $$y = \sum_{i=1}^{l} g_i T^i y$$

where the $g_i$ belong to $M$ and each $\Phi^i(g_i) \equiv 0$; (IX) there exists a subspace $W$ of $U_1$ and a linear operator $T^{-1}$, defined from $TW$ to $U$ such that

$$T^{-1}T/W = I/W;$$



(X) $\Phi^j(g_i)T^{i-j-1-k}y$ is defined and belongs to $W$ for each $1 \leqslant i \leqslant l$, $j \geqslant 0$, and $k \geqslant 0$, as does each $T^{-n}y$ for $n \geqslant 0$; (XI) there exists a $\delta > 0$ such that

$$|T^{-n}y(x_0)| \leqslant \left(\frac{c(x_0)}{n}\right)^{n\delta} \qquad (n = 1, 2, \ldots),$$

where $c(x_0)$ is positive and independent of $n$, for each $x_0$ belonging to $S$; (XII) we are given $x_1$ belonging to $S$ such that each entry of each $\Phi^j g_i(x_1)$ belongs to $Q$ (the rationals) for $1 \leqslant i \leqslant l$, $0 \leqslant j \leqslant i$; and (XIII) $g_l(x_1)$ is nonsingular.

Now set

$$\deg g_i = \min_j\{j \mid \Phi^{j+1}(g_i) \equiv 0\}$$

and

$$d = \max_{l\geqslant i\geqslant 0}\left\{\frac{\deg g_i}{i - \deg g_i}\right\} \geqslant 0.$$

Definition. By the *absolute value of a matrix* we mean the maximum of the absolute values of its entries.

Theorem I. *Under conditions* (I)-(XIII) *either*

(a) *each* $T^i y(x_1) = 0$ $(0 \leqslant i \leqslant l-1)$

or

(b) *for each* $\varepsilon > 0$ *there exists* $c(\varepsilon) > 0$ *such that*

(11) $$\max_{0\leqslant i\leqslant l-1}\{|T^i y(x_1) - P_i/q|\} \geqslant c(\varepsilon)\,|q|^{-(1+d/\delta+\varepsilon)}$$

*for all $m$ by 1 matrices of integers $P_i$ and nonzero integers $q$.*

Lemma. *If we assume in Theorem I that $y$ takes values in the $m$ by 1 matrices over $\mathscr{C}$, that $U$ is a vector space over $\mathscr{C}$, that $M$ is a vector space of $m$ by 1 matrix valued functions over $\mathscr{C}$, that each $P_i$ is an $m$ by 1 matrix with Gaussian integral entries, that $q$ is a nonzero Gaussian integer, and that each $\Phi^j(g_i)(x_1)$ has Gaussian integral entries, then* (11) *still holds.*

Proof. The change to complex valued functions throughout the proof of Theorem I, along with the substitution of Gaussian integral $q$ for integral $q$ and Gaussian integral entries for integral entries in the $\Phi^j(g_j)(x_1)$ and the $P_i$, gives a proof of (11) under the conditions stated in the Lemma. (The only property of the integers which was used in the proof of Theorem I was that if $n$ is an integer and $n$ is not zero then $|n| \geqslant 1$.)

Definition. *By* $\|\alpha\|$, where $\alpha = (\alpha_{i,j})$ is a matrix, we mean $\max_{i,j}\{\|\alpha_{i,j}\|\}$.

Suppose that $y_1, \ldots, y_n$ are $n$ linearly independent $m$ by 1 matrix valued functions over the complex numbers which each satisfy the conditions of Theorem I, as modified by the Lemma, for the same set $S$;

the same operators $T, T^{-1}$, and $\Phi$; the same operator equation with the same coefficient functions $g_i(x)$; the same spaces $M$, $U$, $U_i$ $(1 \leqslant i \leqslant l)$ and $W$; and the same constants $c(x_0)$ and $\delta$.

Let $(A_1, \ldots, A_n)$ denote an arbitrary $n$-tuple of Gaussian integers.

THEOREM II. *For every* $\varepsilon > 0$

$$\max_{0 \leqslant i < l} \Big\{ \Big\| \sum_{j=1}^{n} A_j T^i y_j(x_1) \Big\| \Big\} \geqslant \min_j \{|A_j|^{-(d/\delta+\varepsilon)}\} \tag{12}$$

*except for a finite collection of* $(A_1, \ldots, A_n)$.

Proof. Obviously the Lemma follows from Theorem II. On the other hand there does not seem to be an easy way of going from the Lemma, as stated, to Theorem II. However, we notice that in demonstrating Theorem I in [2] it was first shown that

$$\max_{0 \leqslant i \leqslant l-1} \{|T^i y(x_1) - P_i/q|\} \geqslant |q|^{-(1+d/\delta+\varepsilon)} \tag{13}$$

for all $q$ such that $|q| \geqslant q(\varepsilon)$. Examination of the steps involved reveals that $q(\varepsilon)$ is *uniform* for all $y$ satisfying a fixed set of hypotheses (I)-(XIII). (After deriving the above inequality $q(\varepsilon)$ was dropped in favor of a statement involving $c(\varepsilon)$, where $c(\varepsilon)$ depends upon how well the *particular numbers* $T^i y(x_1)$ can be approximated by fractions with denominators having absolute values less than $q(\varepsilon)$. Hence $c(\varepsilon)$ *depends upon* $y$.) Changing now to the hypotheses of the Lemma we again obtain (13) only with $y(x)$ an $m$ by 1 complex matrix valued function and each $P_i$ an $m$ by 1 matrix of Gaussian integers. Suppose that $y_1, \ldots, y_n$ satisfy the hypotheses of the Lemma. For each nonzero $(A_1, \ldots, A_n)$, then

$$y = \Big(\sum_{j=1}^{n} A_j y_j\Big)\big(n \max_j \{|A_j|\}\big)^{-1}$$

also satisfies the hypotheses of the Lemma and — hence — satisfies (13). We have for each Gaussian integer $A_0$

$$\max_i \Big| \sum_{j=1}^{n} A_j T^i y(x_1) \big(n \max_j \{|A_j|\}\big)^{-1} - A_0 \big(n \max_j \{|A_j|\}\big)^{-1} \Big| \geqslant n^{-(1+d/\delta+\varepsilon)} \min_j \{|A_j|^{-(1+d/\delta+\varepsilon)}\}$$

or

$$\max_i \Big\{ \Big\| \sum_{j=1}^{n} A_j T^i y(x_1) \Big\| \Big\} \geqslant n^{-(d/\delta+\varepsilon)} \min_j \{|A_j|^{-(d/\delta+\varepsilon)}\} \geqslant \min_j \{|A_j|^{-(d/\delta+\varepsilon/2)}\},$$

if $\max_j \{|A_j|\}$ is larger than some number depending on $\varepsilon$ but independent of $y_1, \ldots, y_n$. This proves Theorem II.



We now present the hypotheses of Theorem III. Suppose that (1)′ $y_1, \ldots, y_n$ are $n$ functions from a set $S$ to the $m$ by 1 matrices over $C$; (2)′ $U \supseteq U_1 \supseteq \ldots \supseteq U_l$ are vectors spaces of functions from $S$ to the $m$ by 1 matrices over $C$ with $y_1, \ldots, y_n$ belonging to $U_l$; (3)′ $T$ is a linear operator from $U_1$ to $U$ and $T^i$ is defined $(1 \leqslant i \leqslant l)$ from $U_i$ to $U$; (4)′ $M$ is a vector space over $C$ of functions from $S$ to the $m$ by $m$ matrices over $C$; (5)′ $\Phi$ is a function from $M$ to $M$; (6)′ if $f$ belongs to $U_1$ and $g$ to $M$ then $gf$ belongs to $U_1$ and $T(gf) = gTf + \Phi(g)f$; (7)′ $y_j = \sum_{i=1}^{l} g_i T^i y_j + a_j$ $(1 \leqslant j \leqslant n)$, for functions $a_j$ in $U_1$, where the $g_i$ belong to $M$ and each $\Phi^i(g_i) \equiv 0$; (8)′ there exists a linear operator $\sigma$ from $U$ to $U$ which is homogeneous with respect to elements of $M$ (i.e. $\sigma(gf) = g\sigma(f)$, if $g$ belongs to $M$) and which commutes with $T$; (9)′ there exists a linear operator $T^{-1}$: $U \to U_1$ such that $TT^{-1} = I$; (10)′ there exists a subspace $V$ of elements of $U_1$ which are left fixed pointwise by $\sigma$ such that $V \supset \operatorname{Ker} T$, each $a_j$ belongs to $V$, and $V$ is an invariant subspace under $T^{-1}$; (11)′ there exists a $\delta > 0$ such that for each $x_0$ belonging to $S$ there is a

$$c(x_0) > 0 \quad \text{with} \quad |T^{-n} y_j(x_0) - \sigma T^{-n} y_j(x_0)| \leqslant \left(\frac{c(x_0)}{n}\right)^{n\delta} \quad \text{for } n = 1, 2, \ldots;$$

(12)′ the functions $y_j - \sigma y_j$ are linearly independent over $C$; (13)′ each $T^{l-1-k} y_j$ $(0 \leqslant k < \infty)$ belongs to $U_1$; each $\Phi^\gamma(g_i)(x_1)$ has Gaussian rational entries for $\gamma = 0, 1, \ldots$; and $g_l(x_1)$ is nonsingular.

Set

$$\deg g_i = \max_\gamma \{\gamma \,|\, \Phi^{\gamma+1}(g_i) \neq 0\}.$$

Set

$$d = \max_{l \geqslant i > 0} \left\{ \frac{\deg g_i}{i - \deg g_i} \right\}.$$

THEOREM III. *Under conditions* (1)′-(13)′, *for every* $\varepsilon > 0$

$$\max_{0 \leqslant i < l} \Big\| \sum_{j=1}^{n} A_j \big(T^i y(x_1) - T^i \sigma y(x_1)\big) \Big\| \geqslant \min_j \{|A_j|^{-(d/\delta+\varepsilon)}\}$$

*with the exception of at most a finite collection of* $(A_1, \ldots, A_n)$ *depending upon* $\varepsilon$.

Proof. We wish to satisfy (I)-(XIII) of Theorem II. Since the same letters are used in a number of cases in the statements of both theorems we shall place a bar over any symbol which is meant to refer to Theorem II. Thus (I) $\bar{S}$ is the set $S$ given in (1)′ above. (II) We set

$$\bar{U} = \frac{U}{V} \quad \text{and} \quad \bar{U}_i = \frac{U_i + V}{V}$$

under the definition

$$\{u\}(z) = u(z) - \sigma u(z)$$

where $\{u\}$ denotes the coset containing $u$. This is well defined as $\sigma$ leaves elements of $V$ fixed. (III) $\overline{T}\{u\} = \{Tu\}$. This is well defined since if $\{u_1\} = \{u_2\}$ then $u_1 = u_2 + v$, where $v$ belongs to $V$ and

$$Tu_1 - \sigma Tu_1 = Tu_2 - \sigma Tu_2 + (Tv - \sigma Tv).$$

The last term is zero as $\sigma$ leaves elements of $V$ fixed. Then $\overline{T}^i$ is defined from $\overline{U}_i$ to $\overline{U}$ for $1 \leqslant i \leqslant l$. (IV) The functions $\overline{y}_j = \{y_j\}$ $(1 \leqslant j \leqslant n)$ belong to $\overline{U}_l$. (V) Let $\overline{M} = M$. If $\overline{g} = g$ belongs to $\overline{M}$ and $\overline{f} = \{f\}$ belongs to $\overline{U}_1 = (U_1 + V)/V = U_1/V$ (as $V \subset U_1$), then $\overline{g}\overline{f} = \{gf\}$ which belongs to $U_1/V = \overline{U}_1$. (VI) $\overline{\Phi} = \Phi$. (VII) We note that $Tgf = \Phi(g)f + gTf$ implies that

$$Tgf - T\sigma(gf) = \Phi(g)(f - \sigma f) + g(Tf - T\sigma f) \quad \text{or} \quad \overline{T}\overline{g}\overline{f} = \overline{\Phi}(\overline{g})\overline{f} + \overline{g}\overline{T}\overline{f}.$$

(VIII) $\overline{y}_j = \sum_{i=1}^{l} \overline{g}_i \overline{T}^i \overline{y}_j$ from $(7)'$ by the same type of argument as in (VII) above. (IX) Set $\overline{W} = \overline{U}_1$. Suppose $\{u\}$ belongs to $\overline{U}_1$. We define $\overline{T}^{-1}$ by $\overline{T}^{-1}\{u\} = \{T^{-1}u\}$. This is well defined, as if $u_1 = u_2 + v$ then

$$T^{-1}u_1 = T^{-1}u_2 + T^{-1}v$$

where $T^{-1}v$ belongs to $V$. Since $TT^{-1} = I$, by $(9)'$, we have for $u$ in $U_1$ that $T^{-1}Tu = u + v_1$ where $v_1$ belongs to $\ker T \subseteq V$. Thus $\overline{T}^{-1}\overline{T} = \overline{I}$. (X) $\overline{U}_1 = \overline{W}$ is closed under multiplication from $\overline{M}$ and $\overline{T}^{(l-1-k)}\overline{y}_j$ belongs to $\overline{U}_1$ by $(13)'$ for each nonnegative integer $k$. (XI) By $(11)'$ we have

$$|\overline{T}^{-n}\overline{y}_j(x_0)| = |T^{-n}y_j(x_0) - \sigma T^{-n}y_j(x_0)| \leqslant \left(\frac{c(x_0)}{n}\right)^{n\delta}.$$

Now (XII) and (XIII) follow from $(13)'$.

By $(12)'$ the $y_j - \sigma y_j$ are linearly independent functions. Thus the $\overline{y}_j$ are linearly independent. Then by Theorem II we conclude that

$$\max_i \left\| \sum_{j=1}^{n} A_j\big(T^i y_j(x_1) - T^i \sigma y_j(x_1)\big) \right\| \geqslant \min\{|A_j|^{-(d/\delta+\varepsilon)}\}$$

with but a finite number of exceptions depending upon $\varepsilon$ but independent of $y_1, \ldots, y_n$. This proves Theorem III.

Proof of the Proposition. Let us apply Theorem III. Let $S$ be $N$. The space $U$ is the space of all functions which are analytic in $N$ and which can be extended analyticly around $\mathscr{C}$ any positive integral number of times to be defined again (possible differently) on $N$. Let $U_i = U$ $(1 \leqslant i \leqslant l)$. Let $T = d/dz$. Let $M$ be all polynomials in $z$ over $C$. Now $\Phi = d/dz$. Clearly $y_1, \ldots, y_n$ belong to $U_l$. Parts $(6)'$ and $(7)'$ hold trivially. If $f$ belongs to $U$ by $\sigma f$ we denote the analytic continuation of $f$ once around $\mathscr{C}$. Thus $(8)'$ holds. We set

$$T^{-n}f = \int_{z_0}^{z} \frac{(z - z_0 - t)^{n-1}}{(n-1)!} f(t)\,dt,$$

where the path is the ray from $z_0$ to $z$. By analytic continuation of the above expression we obtain

$$\sigma T^{-n}y(z) = \int_{z_0}^{z} \frac{(z - z_0 - t)^{n-1}}{(n-1)!} f(t)\,dt,$$

where the path is once around $\mathscr{C}$ followed by the ray from $z_0$ to $z$. Now $(11)'$ follows easily (using Sterling's approximation) with $\delta = 1$. For $(10)'$ we let the subspace $V$ consist of all functions analytic on $X$. Parts $(12)'$ and $(13)'$ were either explicitly assumed or are obvious. The conclusion obtained from Theorem III is the Proposition, so we are through.

We conclude with the remark that, in the case of vector differential equations where $T = \psi_1(t)\dfrac{d}{dt} + \psi_2(t)$ for possibly nonanalytic $m$ by $m$ matrix valued functions $\psi_1$ and $\psi_2$ (see [2]), it should be possible to apply Theorem III if $\psi_1$ and $\psi_2$ are periodic functions of period $\tau$. Then we would wish to set $\sigma y(t) = y(t+\tau)$. If we translate the procedure for complex scalar differential equations into 2 dimensional real vector notation with an appropriate parameter $t$ this is what our procedure would amount to anyway.

## References

[1] J. W. S. Cassels, *An Introduction to Diophantine Approximation*, Cambridge 1957.
[2] C. F. Osgood, *A method in diophantine approximation*, Acta Arith. 12 (1966), pp. 111-129.
[3] — *Some theorems on diophantine approximation*, Trans. Amer. Math. Soc. 123 (1) (1966), pp. 64-87.

UNIVERSITY OF ILLINOIS, URBANA ILLINOIS
NATIONAL BUREAU OF STANDARDS, WASHINGTON, D.C.

*Reçu par la Rédaction le 28. 1. 1967*