La revue est consacrée à toutes les branches de l'*Arithmétique* et de la *Théorie des Nombres*, ainsi qu'aux *fonctions ayant de l'importance dans ces domaines*.

Prière d'adresser les textes dactylographiés à l'un des redacteurs de la revue ou bien à la Rédaction de

ACTA ARITHMETICA

Warszawa 1 (Pologne) ul. Śniadeckich 8.

La même adresse est valable pour toute correspondance concernant l'échange de Acta Arithmetica.

Les volumes IV et suivants de ACTA ARITHMETICA sont à obtenir chez

Ars Polona, Warszawa 5 (Pologne), Krakowskie Przedmieście 7.

Prix de ce fascicule 3.00 $

Les volumes I-III (reédits) sont à obtenir chez

Johnson Reprint Corp., 111 Fifth Ave., New York, N. Y.

PRINTED IN POLAND

WROCŁAWSKA DRUKARNIA NAUKOWA

ACTA ARITHMETICA
XII (1966)

A method in diophantine approximation*

by

CHARLES F. OSGOOD (Urbana, Ill.)

The object of this paper is to present a theorem which gives a method for obtaining results about the diophantine approximation of some values of certain functions into R^m where $m \geqslant 1$. The method is then applied in two corollaries and a number of examples to functions satisfying linear differential equations, both scalar and vector with not necessarily (real) analytic coefficients. Two examples deal with applications to non-linear functional equations.

Section I. Suppose that:

(I) y is a function from a set S to R^m (the m by 1 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) Φ is a function from M to M;

(VII) if f belongs to U_1 and g belongs to M, then gf belongs to U_1 and $Tgf = gTf+\Phi(g)f$;

(VIII) we have

$$y = \sum_{i=1}^{l} g_i T^i y \tag{1}$$

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 $n \geqslant 0$;

* Supported in part by the National Science Foundation.

(XI) there exists a $\delta > 0$ such that

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

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 $\Phi^j(g_i)(x_1)$ belongs to Q^{m^2} (Q denotes the rationals) for each $1 \leqslant i \leqslant l$, $0 \leqslant j < i$; 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

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

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

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

(a) $$\text{each } T^i y(x_1) = 0 \qquad (0 \leqslant i \leqslant l-1)$$

or

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

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

for all m by 1 matrices of integers P_i and positive integers q.

We defer the proof until later; instead, we now illustrate the theorem with two corollaries and several examples.

Definition. Given $\varepsilon_1 > 0$ and f a real valued function, by *f belongs to C^{l-1} ($l \geqslant 2$) on $[0, \varepsilon_1]$, $(0, \varepsilon_1]$ or $[0, \varepsilon_1)$* we shall mean that f belongs to C^{l-1} on $(0, \varepsilon_1)$ and that each of the indicated derivatives can be continuously extended to the endpoint(s).

Given ψ_1 and ψ_2 such that $1/\psi_1$ and ψ_2 belong to C^{l-1} on $[0, \varepsilon_1]$, we set

$$T = \psi_1 \frac{d}{dx} + \psi_2$$

and consider the functional equation

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

where here each g_i belongs to $C^i[0, \varepsilon_1)$ and each $\left(\psi_1 \frac{d}{dx}\right)^i g_i \equiv 0$. Pick x_1 belonging to $(0, \varepsilon_1)$ such that each $\left(\psi_1 \frac{d}{dx}\right)^j g_i(x_1)$ belongs to Q $(0 \leqslant j)$ and $g_l(x_1)$ is nonzero. Let $\deg g_i$ and d be as defined before.

Corollary I. *Under the above conditions if y satisfies the functional equation* (1) *on* $(0, \varepsilon_1)$, *y belongs to C^l on* $(0, \varepsilon_1)$, *and*

$$\lim_{x\to 0} \left[\left(\psi_1 \frac{d}{dx}\right)^j g_i\right] \left[\left(\psi_1 \frac{d}{dx} + \psi_2\right)^{i-j-1-k} y\right] = 0$$

for each $0 \leqslant j < i$, $k \geqslant 0$, and $i-j-1-k \geqslant 0$, then either

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

or

(b) *for every $\varepsilon > 0$ there exists $C(\varepsilon) > 0$ such that*

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

for all integers p_i and all positive integers q.

Proof. We apply the theorem:

(I) $S = (0, \varepsilon_1)$ and $y\colon (0, \varepsilon_1) \to R$;

(II) U is the space of all continuous functions from $(0, \varepsilon_1)$ to R;

(III) $T = \psi_1 \frac{d}{dx} + \psi_2$ and $U_i = C^i$ on $(0, \varepsilon_1)$;

(IV) y belongs to U_l;

(V) M is the vector space (over R) spanned by the $\left(\psi_1 \frac{d}{dx}\right)^j g_i$, $(1 \leqslant i \leqslant l, j \geqslant 0)$;

(VI) $$\Phi = \psi_1 \frac{d}{dx};$$

(VII) $$Tgf = gTf + \left(\psi_1 \frac{dg}{dx}\right) f;$$

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

(IX) $$T^{-1} h = \exp\left[-\int_0^x \frac{\psi_2}{\psi_1} ds\right] \int_0^x \exp\left[\int_0^t \frac{\psi_2}{\psi_1} ds\right] \frac{h(t)\,dt}{\psi_1},$$

for all h in TW, where W is the space of C^1 functions f on $(0, \varepsilon_1)$ satisfying $\lim_{x\to 0} f(x) = 0$.

If f belongs to W, then $T^{-1}Tf = f$. Hence $T^{-1}T/W = I/W$.

(X) By assumption, $\lim_{x\to 0} \Phi^j(g_i)\, T^{i-j-1-k}y = 0$ for each $1 \leqslant i \leqslant l$, $0 \leqslant j < i$, $k \geqslant 0$, and $i-j-1-k \geqslant 0$. It follows that under these conditions each $\Phi^j(g_i)T^{i-j-1-k}y$ belongs to W. If $i-j-1-k < 0$, we see that the limit is, again, zero since $\Phi^j(g_i)$ belongs to $C^1[0, \varepsilon_1)$ and $\lim_{x\to 0} T^{-n}y = 0$ for each $n \geqslant 1$. Each $T^{-n}y$ $(n \geqslant 1)$ belongs to $C^1[0, \varepsilon_1)$; hence $\Phi^j(g_i)T^{i-j-1-k}y$ belongs to W for $1 \leqslant i \leqslant l$, $0 \leqslant j \leqslant 1$, and $k \geqslant 0$, as does each $T^{-n}y$ for $n \geqslant 0$.

$$\text{(XI)} \qquad |T^{-n}y(x_0)| \leqslant \left|\exp\left[-\int_0^{x_0} \frac{\psi_2}{\psi_1}\, ds\right]\right| \left(\max_{0\leqslant x\leqslant x_0} \left|\frac{1}{\psi_1}\right|\right)^n$$

$$\left(\max_{0\leqslant x\leqslant x_0} |y|\right) \left|\int_0^{x_0} \frac{(x_0-t)^{n-1}}{(n-1)!} \exp\left[\int_0^t \frac{\psi_2}{\psi_1}\, ds\right] dt\right| \leqslant \left(\frac{C(x_0)}{n}\right)^n.$$

By hypothesis both (XII) and (XIII) hold. The corollary then follows immediately from the theorem.

Example I. We wish to show that the hypotheses of Corollary I are satisfied under many circumstances. Let $\int_0^x \frac{ds}{\psi_1} = t$. Then

$$T = \psi_1 \frac{d}{dx} + \psi_2 = \frac{d}{dt} + \psi_2(t)$$

where we now regard ψ_2, y, and each $\Phi^j(g_i)$ as functions of t. (Since ψ_1 did not vanish on $[0, \varepsilon_1)$, we see that this change of variables is an l times differentiable homeomorphism.) We see that

$$\exp\left(\int_0^t \psi_2\, ds\right) Ty = \frac{d}{dt} \exp\left(\int_0^t \psi_2\, ds\right) y.$$

The identity $\Phi^i(g_i) \equiv 0$ says that $d^i g_i/dt^i \equiv 0$, hence each g_i is a polynomial in t of degree less than i. We obtain from (1) under these circumstances

$$\text{(2)} \qquad \exp\left(\int_0^t \psi_2\, ds\right) y = \sum_{i=1}^{l} g_i(t) \frac{d^i}{dt^i} \exp\left(\int_0^t \psi_2\, ds\right) y.$$

Suppose that zero is a regular singular point of the above analytic linear differential equation in

$$Y = \exp\left(\int_0^t \psi_2\, ds\right) y.$$

Suppose further that r of the roots of the indicial equation associated with the differential equation (2) in Y have real parts greater than $l-1$. By the construction of solutions about a regular singular point (see [2]) when the coefficient functions are polynomials there exists an r dimensional solution space over R of functions Y which satisfy (2), belong to C^∞ on $\left(0, \int_0^{\varepsilon_1} \frac{ds}{\psi_1}\right)$ in t, and satisfy

$$\lim_{t\to 0} \frac{d^n Y}{dt^n} = 0 \qquad (0 \leqslant n \leqslant l-1).$$

It follows that

$$0 = \lim_{t\to 0} \exp\left(\int_0^t \psi_2\, ds\right) T^n y = \lim_{t\to 0} T^n y = \lim_{x\to 0} T^n y.$$

Thus

$$\lim_{x\to 0} \Phi^j(g_i)\, T^{i-j-1-k} y = 0$$

for $1 \leqslant i \leqslant l$, $j < i$, $k \geqslant 0$, and $i-j-1-k \geqslant 0$. If we can find an x_1 such that each $\Phi^j g_i(x_1)$ belongs to Q and $g_l(x_1)$ does not vanish, we may apply the Corollary.

Example II. We treat in detail a specific equation, consider

$$y = \frac{d}{dx}\left(x \frac{d}{dx} - 4\right)^2 y = 9\frac{dy}{dx} - 5x \frac{d^2y}{dx^2} + x^2 \frac{d^3y}{dx^3}.$$

Here $T = d/dx$, $g_1 = 9$, $g_2 = -5x$, and $g_3 = x^2$. Two solutions of this equation which are linearly independent are $y_1 = x^4 + \ldots$, and $y_2 = (\ln x)y_1 + y_3$ where $y_3 = x^4 + \ldots$. Set $y = C_1 y_1 + C_2 y_2$ where C_1 and C_2 are arbitrary constants. Here

$$\lim_{x\to 0} \Phi^j(g_i)\, T^{i-j-1-k} y = 0, \qquad 1 \leqslant i \leqslant 3,\ j < i,\ k \geqslant 0,\ i-j-1-k \geqslant 0.$$

If $x_1 \geqslant 0$ is rational, then for every $\varepsilon > 0$ there exists $C(\varepsilon) > 0$ with

$$\max_{i=0,1,2} \left|\frac{d^i y}{dx^i} - \frac{p_i}{q}\right| > C(\varepsilon)\, q^{-(3+\varepsilon)},$$

by Corollary I. At any rational point x_1 larger than zero we have

$$\begin{vmatrix} y_1(x_1) & y_2(x_1) \\ y_1'(x_1) & y_2'(x_1) \end{vmatrix} \neq 0,$$

since otherwise there would exist nonzero constants C_1 and C_2 with $y = y' = 0$ and $y'' = 1$. Thus we may define y by

$$\begin{vmatrix} y_1(x_1) & y_2(x_1) \\ y_1'(x_1) & y_2'(x_1) \end{vmatrix} y = \begin{vmatrix} 1 & y_2(x_1) \\ 0 & y_2'(x_1) \end{vmatrix} y_1 + \begin{vmatrix} y_1(x_1) & 1 \\ y_1'(x_1) & 0 \end{vmatrix} y_2.$$

Now we have that $y(x_1) = 1$ and $y'(x_1) = 0$. Hence given $\varepsilon > 0$ there exists $C(\varepsilon) > 0$ with

$$|y''(x_1) - p/q| \geqslant C(\varepsilon) q^{-(3+\varepsilon)},$$

for all integers p and positive integers q, where

$$y''(x_1) = \frac{\begin{vmatrix} y_1''(x_1) & y_1'(x_1) \\ y_2''(x_1) & y_2'(x_1) \end{vmatrix}}{\begin{vmatrix} y_1(x_1) & y_1'(x_1) \\ y_2(x_1) & y_2'(x_1) \end{vmatrix}}.$$

Example III. Consider the nonlinear differential equation $y = x\left(\frac{d}{dx}+y\right)^2 y$. We shall show that there exists a solution

$$y_1 = \sum_{n=1}^{\infty} a_n x^n$$

which converges in an open disk of radius $\frac{1}{2}$ about the origin. The recurrence formula for the coefficients is gotten by simplifying to $y = x(y'' + 3yy' + y^3)$ and substituting $y = \sum_{n=0}^{\infty} a_n x^n$ for y. We obtain

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} n(n-1) a_n x^{n-1} + 3 \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} n a_n a_m x^{n+m} + \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\sum_{q=0}^{\infty} a_n a_m a_q x^{n+m+q+1}.$$

We must have that

$$a_n = (n+1) n a_{n+1} + 3 \sum_{m=0}^{n} m a_m a_{n-m} + \sum_{m+q\leqslant n-1} a_m a_q a_{n-m-q-1}.$$

Letting $n = 0$, we obtain $a_0 = 0$. Now set $a_1 = 1$. If $n > 0$ we define a_{n+1} by

$$a_{n+1} = \frac{1}{(n+1)(n)}\left[a_n - 3\sum_{m=0}^{n} m a_m a_{n-m} - \sum_{m+q\leqslant n-1} a_m a_q a_{n-m-q-1}\right].$$

Now $|a_0| = 0 < 2^0$ and $|a_1| = 1 < 2^1$. We shall show by induction that $|a_n| \leqslant 2^n$. If $n > 0$, then applying the induction hypothesis we obtain

$$\begin{aligned} |a_{n+1}| &\leqslant \frac{2^n}{(n+1)(n)}\left[1 + 3\sum_{m=0}^{n} m + \frac{1}{2}\sum_{m+q\leqslant n-1}\right] \\ &\leqslant \frac{2^n}{(n+1)(n)}\left[1 + \frac{3}{2}n^2 + \frac{3}{2}n + \frac{1}{2}\sum_{q=0}^{n-1}(n-1-q)\right] \\ &= \frac{2^n}{(n+1)(n)}\left[1 + \frac{3}{2}n^2 + \frac{3}{2}n + \frac{1}{2}(n)(n-1)\right] \leqslant 2^{n+1}. \end{aligned}$$

Now apply Corollary I to the linear equation $y = xT^2 y$, where $T = d/dx + y_1$ and $y = y_1$. We verify that y_1 belongs to C^∞ on $[0, \frac{1}{2}]$ and that $\Phi^2(x) \equiv d^2x/dx^2 \equiv 0$. Since y_1 vanishes at $x = 0$, we see that

$$\lim_{x\to 0} \Phi^j(x) T^{1-j-k} y_1 = 0$$

for $j < 2$, $k \geqslant 0$, and $1-j-k \geqslant 0$. For each rational number in $(0, \frac{1}{2})$ then it follows that one of the two numbers y_1 and $y_1' + y_1^2$ is irrational.

Example IV. Consider the equation

$$y = \int_0^x \frac{dt}{1+y}\left((1+y)\frac{d}{dx}\right)^2 y. \tag{3}$$

If there exists a solution y_1 of (3) which belongs to C^2 on $[0, \varepsilon_1)$ for some $\varepsilon_1 > 0$, satisfies $\lim_{x\to 0} y_1 = 0$, and does not equal -1 on $[0, \varepsilon_1)$, then we may apply Corollary I to the linear differential equation

$$y = \left(\int_0^x \frac{dt}{1+y_1}\right) T^2 y,$$

where $T = (1+y_1)d/dx$ and $y = y_1$. It will follow that one of

$$\int_0^x \frac{dt}{1+y_1},\quad y_1, \text{ and } (1+y_1)y_1'$$

is irrational at each point of $(0, \varepsilon_1)$.

We now construct a function y_1 which satisfies the above requirements. Consider the function

$$f(s) = \sum_{n=1}^{\infty} \frac{s^n}{n!(n-1)!}.$$

Suppose that we can find a solution y_1 of the equation

$$f\left(\int_0^x \frac{dt}{1+y}\right) = y$$

which belongs to C^2 on $[0, \varepsilon_1)$ and does not equal -1 on $[0, \varepsilon_1)$. Note that

$$f(s) = s\frac{d^2}{dt^2}f(s),$$

hence

$$f\left(\int_0^x \frac{dt}{1+y_1}\right) = \left(\int_0^x \frac{dt}{1+y_1}\right)\left[(1+y_1)\frac{d}{dx}\right]^2 f\left(\int_0^x \frac{dt}{1+y_1}\right)$$

$$y_1 = \left(\int_0^x \frac{dt}{1+y_1}\right)\left[(1+y_1)\frac{d}{dx}\right]^2 y_1.$$

We know that the non-linear differential equation $1/Y'-1 = f(Y)$ ha. an analytic solution Y_1 in a neighborhood of the origin with $Y_1(0) = 0$ (See [1]). Set $y_1 = 1/Y_1'-1 = f(Y_1)$. Since $Y_1(0) = f(0) = 0$, we have $y_1(0) = 0$. There exists $\varepsilon_1 > 0$ such that y_1 belongs to C^∞ on $[0, \varepsilon_1)$ and $y_1 \neq -1$ on $[0, \varepsilon_1)$. Finally, since $y_1 = 1/Y_1'-1$ and $y_1(0) = 0$, we have

$$y_1 = \int_0^x \frac{dt}{1+y_1}.$$

Therefore

$$y_1 = \frac{1}{Y_1'} - 1 = f(Y) = f\left(\int_0^x \frac{dt}{1+y_1}\right),$$

so y_1 satisfies (3) on $(0, \varepsilon)$. This completes example IV.

Definition. If $g = (g_{jk})$ is a function into the m by m matrices over R, then by *g belongs to C^{l-1} $(l \geqslant 2)$ on $[0, \varepsilon_1]$, $(0, \varepsilon_1]$ or $[0, \varepsilon_1)$* we mean that each g_{jk} belongs to C^{l-1} on the appropriate set.

Now we show that Corollary I can be extended to the case of matrix coefficients. As before

$$T = \psi_1\frac{d}{dx} + \psi_2$$

where now ψ_1 and ψ_2 are matrix valued functions. We assume that ψ_1 takes on only nonsingular values and define ψ_3 by $\psi_3(x) = (\psi_1(x))^{-1}$. Assume that ψ_3 and ψ_2 belong to C^{l-1} on $[0, \varepsilon_1)$ and that each g_i $(0 < i \leqslant l)$ is a matrix valued function which belongs to C^i on $[0, \varepsilon_1)$. Assume that each g_i satisfies $(\psi_1 d/dx)^i g_i \equiv 0$ and that ψ_1 and ψ_2 both commute with each $(\psi_1 d/dx)^j g_i$. Define $\deg g_i$ and d as before. Choose x_1 belonging to $(0, \varepsilon_1)$ such that each $(\psi_1 d/dx)^j g_i(x_1)$ belongs to the m by m matrices over Q and $g_l(x_1)$ is nonsingular. With these assumptions we again consider the functional equation (1).

Corollary II. *Under the above conditions, if y is a function from $(0, \varepsilon_1)$ to R^m (the m by 1 matrices over R), y belongs to C^l on $(0, \varepsilon_1)$, y satisfies equation* (1) *on* $(0, \varepsilon_1)$, *and*

$$\lim_{x\to 0}\left[\left(\psi_1\frac{d}{dx}\right)^j(g_i)\right]\left[\left(\psi_1\frac{d}{dx}+\psi_2\right)^{i-j-1-k} y\right] = 0$$

(*for each* $0 \leqslant j < i$, $k \geqslant 0$, *and* $i-j-1-k \geqslant 0$), *then either*

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

or

(b) *for every $\varepsilon > 0$ there exists $C(\varepsilon) > 0$ such that*

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

for all m by 1 matrices P_i over the integers and positive integers q.

Proof. We apply the theorem.

(I) $S = (0, \varepsilon_1)$ and $y\colon (0, \varepsilon_1) \to R^m$;

(II) U is the space of all continuous functions from $(0, \varepsilon_1)$ to R^m;

(III) $T = \psi_1 d/dx + \psi_2$ and $U_i = C^i$ on $(0, \varepsilon_1)$;

(IV) y belongs to U_l;

(V) M is the vector space spanned by the $\Phi^j(g_i)(1 \leqslant i \leqslant l, j \geqslant 0)$;

(VI) $\Phi = \psi_1 d/dx$;

(VII) $Tgf = gTf + \varphi(g)f$ for all g in M and f in U_1.

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

(IX) $T^{-1}h = \alpha(x)\int_0^x (\psi_1(t)\alpha(t))^{-1}h(t)\,dt$ for all h belonging to TW,

where W is the subspace of U_1 consisting of all functions whose limit at $x=0$ is zero and $\alpha(x)$ is a certain m by m matrix valued function on $[0, \varepsilon_1)$. We choose $\alpha(x)$ to be any solution of

$$\left(\frac{d}{dx}+\psi_3\psi_2\right)\alpha(x) \equiv 0 \quad \text{on} \quad [0, \varepsilon_1)$$

which assumes nonsingular values on $[0, \varepsilon_1)$. (Such an $\alpha(x)$ exists since $\psi_3\psi_2$ is continuous on $[0, \varepsilon_1)$. See [1] where this result is proven.) If f belongs to W then $TT^{-1}f$ is defined and belongs to W. Note that $T^{-1}Tf-f$ belongs to W and also to the kernel of T. By the uniqueness of the solution of $Ty\equiv 0$ when $y(0)$ is known we see that $TT^{-1}f = f$ or $T^{-1}T/W = I/W$.

(X) The argument of Corollary I holds here.

(XI) $$|T^{-n}y(x_0)| \leqslant (m\max|\alpha(x)|)^n\left(m\max|(\alpha(x))^{-1}|\right)^n (m\max|\psi_3|)^n(m\max|y|)\left|\int_0^{x_0}\frac{(x-t)^{n-1}}{(n-1)!}dt\right| \leqslant \left(\frac{C(x_0)}{n}\right)^n,$$

where each maximum is taken on $[0, x_0]$.

Conditions (XII) and (XIII) hold by assumption. Thus Corollary II follows.

The equation

$$y = \sum_{i=1}^{l} g_i\left(\psi_1\frac{d}{dx}+\psi_2\right)^i y \tag{1}$$

may be transformed into the equation

$$y = \alpha(x)y_1 = \alpha(x)\sum_{i=1}^{l} g_i\left(\psi_1\frac{d}{dx}\right)^i y_1$$

where $\alpha(x)$ is as in the proof of Corollary II, *if $\alpha(x)$ can be chosen so as to commute with ψ_1 and each g_i*. In the general case there is no particular reason to believe that this is possible, or if possible, that it is easy to accomplish. It might be supposed that the g_i could be shown to be polynomials with matrix coefficients in $t = \int_0^x \psi_1^{-1}ds$. This does not appear to hold, however. We know that $(\psi_1 d/dx)^3 g_3 \equiv 0$, whence $(\psi_1 d/dx)^2 g_3 \equiv c_1$. Since ψ_1 commutes with each $\Phi^j(g_i)$, we have

$$\left(\psi_1\frac{d}{dx}\right)g_3 \equiv c_1t+c_2$$

and

$$\frac{d}{dx}g_3 \equiv c_1t\psi_1^{-1}+c_2\psi_1^{-1} \quad \text{so} \quad g_3 \equiv c_1\int_0^x t\psi_1^{-1}ds+c_2t+c_3.$$

Now $\int_0^x t\psi_1^{-1}ds = t^2/2$ if and only if $\psi_1^{-1}t = t\psi_1^{-1}$. The most that we know is that $\psi_1^{-1}(c_1t+c_2) = (c_1t+c_2)\psi_1^{-1}$ which is not good enough.

As we shall see we have reached the first place where the (possible) non-analyticity of our coefficients appears to represent a true generalization. For scalar equations of the type treated in Corollary I the change of dependent variable from y to Y given by

$$Y = \exp\left(\int_0^x \psi_2 ds\right)y$$

and the change of independent variable from x to t given by

$$t = \int_0^x \frac{ds}{\psi_1}$$

combine (see Example I of Corollary I) to yield an analytic differential equation in $Y(t)$ to which we may apply Corollary I. Calculating the d^iY/dt^i at

$$t_1 = \int_0^{x_1}\frac{ds}{\psi_1}$$

and the T^iy at x_1 $(0 \leqslant i \leqslant l-1)$, it becomes clear that each

$$\frac{d^iY}{dt^i}(t_1) = \exp\left(\int_0^{x_1}\psi_2 ds\right)T^iy(x_1).$$

Recalling that y is at most determined up to a multiplicative constant, it follows that we obtain the same number-theoretic information from the analytic differential equation for Y as from the original equation for y. Certainly, however, an analogous argument does not go through in the case of vector differential equations, for the reasons outlined above.

We may treat scalar analytic differential equations *in the plane* by writing them as vector differential equations. Given the equation

$$y = \sum_{i=1}^{l} g_i(z)\frac{d^iy}{dz^i},$$

we investigate the behavior of y at the point $\sigma+it\neq 0$ by substituting $\begin{pmatrix}\sigma & -t\\ t & \sigma\end{pmatrix}$ for z, $\begin{pmatrix}u(s)\\ v(s)\end{pmatrix}$ for $y=u+iv$, and $\begin{pmatrix}\sigma & -t\\ t & \sigma\end{pmatrix}^{-1}\dfrac{d}{ds}$ for $\dfrac{d}{dz}$ $(0\leqslant s<<\infty)$. This puts the equation in a form where it may be possible to apply Corollary II.

Several final remarks before proving the theorem. It seems very natural to generalize these results to functions of more than one variable and have T be, say, a partial differential operator. Suppose that

$$T=\frac{\partial}{\partial x_1}+\psi\frac{d}{\partial x_2}$$

where ψ is an m by m matrix valued function of x_1 and ψ commutes with each $T^j(g_i)$. If ψ, then $T^j(g_i)$, and the space U, are sufficiently "nice" one might hope to define T^{-1} by

$$T^{-1}=\left\{\sum_{n=0}^{\infty}\left(-E\psi\frac{\partial}{\partial x_2}\right)^n\right\}E$$

where

$$Eh=\int_0^{x_1}h(t,x_2)\,\partial t.$$

(Suppose that U consists of all functions f which are real analytic functions of x_1 and x_2 and which as analytic functions of z_1 and z_2 in $D\times C$, where D is an open disk about zero, satisfy

$$|f(z_1,z_2)|\leqslant\beta\exp(\alpha|z_2|),$$

for fixed $\alpha>0$ and $\beta>0$ depending on f. If we define $\|f\|=\inf\beta$, then U is a Banach space. Using the integral representation of $\partial f/\partial z_2$, we see that

$$\left|\frac{\partial f}{\partial z_2}\right|\leqslant\alpha\beta\exp(\alpha|z_2|),$$

hence $\partial/\partial z_2$ is a bounded operator. Suppose that $\psi=\psi(x_1)$ belongs to U. Then T^{-1} is a bounded operator from U to U. It is easy to show that $\|T^{-n}\|\leqslant B^n\|E^n\|$ for a constant $B>0$. W would consist of all f in U which vanish when $z_1=0$.

For the above choice of U and W one might also investigate the operator

$$T=\frac{\partial}{\partial z_1}+\psi\Delta_2$$

where ψ is as before and

$$\Delta_2 f(z_1,z_2)=f(z_1,z_2+1)-f(z_1,z_2).$$

If $g(z_1,z_2)$ belongs to M implies $g(z_1,z_2)=g(z_1)$ we have that

$$Tgf=gT(f)+\frac{\partial g}{\partial z_1}f$$

for all f in U. Here we may set

$$T^{-1}=\left\{\sum_{n=0}^{\infty}(-E\psi\Delta_2)^n\right\}E,$$

as Δ_2 is a bounded operator on U.)

Section II.

Proof of the theorem. From (X) we know that each $\Phi^j(g_i)T^{i-j-k-1}y$ belongs to the vector space W. Hence so does w where

$$w=\sum_{a=0}^{\deg g_i-j}(-1)^a\Phi^{j+a}(g_i)T^{i-j-k-a-1}y.$$

Since T is defined on $U_1\supset W$, we may apply T to w. Now each

$$\Phi^{j+a}(g_i)$$

belongs to M, each T^iy $(0\leqslant i<l-1)$ belongs to $U_{l-i}\subseteq U_1$, and each $T^{-n}y$ $(0\leqslant n<\infty)$ belongs to $W\subset U_1$. Thus we may apply the identity of (VII) to obtain

$$T\Phi^{j+a}(g_i)T^{i-j-k-a-1}y=\Phi^{j+a+1}(g_i)T^{i-j-a-1}y+\Phi^{j+a}(g_i)TT^{i-j-k-a-1}y.$$

If $i-j-k-a-1\geqslant 0$, then $TT^{i-j-k-a-1}y=T^{i-j-k-a}y$ by definition. If $i-j-k-a-1\leqslant -1$ then the function $T^{i-j-k-a}y$ belongs to W. It is easily verified that if $T/W\colon W\to TW$ and $T^{-1}\colon TW\to W$ satisfy $T^{-1}T/W=I/W$, then $T/WT^{-1}=I/TW$. Hence

$$TT^{i-j-k-a-1}y=TT^{-1}T^{i-j-k-a}y=T^{i-j-k-a}y$$

here also. It follows that we always have

$$T\Phi^{j+a}(g_i)T^{i-j-k-a-1}y=\Phi^{j+a+1}(g_i)T^{i-j-k-a-1}y+\Phi^{j+a}(g_i)T^{i-j-k-a}y.$$

Thus

$$Tw=-\sum_{a=1}^{\deg g_i-j}(-1)^a\Phi^{j+a}(g_i)T^{i-j-k-a}y+\\ +\sum_{a=0}^{\deg g_i-j}(-1)^a\Phi^{j+a}(g_i)T^{i-j-k-a}y=\Phi^j(g_i)T^{i-j-k}y.$$

Now using $T^{-1}Tw = w$, it follows that

$$(4)\qquad T^{-1}\Phi^j(g_i)T^{i-j-k}y = \sum_{\alpha=0}^{\deg g_i - j} (-1)^\alpha \Phi^{j+\alpha}(g_i)T^{i-j-k-\alpha-1}y.$$

We next show, by induction, that for each $n \geqslant 1$

$$(5)\qquad T^{-n}g_iT^iy = \sum_{\alpha=0}^{\deg g_i} (-1)^\alpha C_{i\alpha}^n \Phi^\alpha(g_i)T^{-n+(i-\alpha)}y$$

for nonnegative integers $C_{i\alpha}^n$ satisfying $C_{i\alpha}^n \leqslant n^\alpha$. If $n = 1$, then line (5) follows from (4) with $j = k = 0$. (Each $C_{i\alpha}^1 = 1 = 1^\alpha$.) We assume the induction statement for $n \geqslant 1$ and shall prove it for $n+1$. By (4) we see that

$$T^{-1}\Phi^\alpha(g_i)T^{i-\alpha-n}y = \sum_{\beta=0}^{\deg g_i-\alpha} (-1)^\beta \Phi^{\alpha+\beta}(g_i)T^{-(n+1)+(i-\alpha-\beta)}y.$$

Hence we obtain from (5), by applying T^{-1},

$$T^{-(n+1)}g_iT^iy = \sum_{\alpha=0}^{\deg g_i} (-1)^\alpha C_{i\alpha}^n \sum_{\beta=0}^{\deg g_i-\alpha} (-1)^\beta \Phi^{\alpha+\beta}(g_i)T^{-(n+1)+(i-\alpha-\beta)}y.$$

Setting $\gamma = \alpha+\beta$, we obtain

$$T^{-(n+1)}g_iT^iy = \sum_{\gamma=0}^{\deg g_i} (-1)^\gamma \Big(\sum_{\alpha=0}^{\gamma} C_{i\alpha}^n\Big)\Phi^\gamma(g_i)T^{-n+\gamma}y.$$

Set

$$(6)\qquad C_{i\gamma}^{n+1} = \sum_{\alpha=0}^{\gamma} C_{i\alpha}^n.$$

Then

$$C_{i\gamma}^{n+1} \leqslant \sum_{\alpha=0}^{\gamma} n^\alpha \leqslant \sum_{\alpha=0}^{\gamma} \binom{\gamma}{\alpha} n^\alpha = (n+1)^\gamma.$$

This proves (5). From above

$$C_{i\alpha}^n \leqslant n^\alpha = n^{\left(\frac{\alpha}{i-\alpha}\right)(i-\alpha)} \leqslant n^{\left(\frac{\deg g_i}{i-\deg g_i}\right)(i-\alpha)} \leqslant n^{d(i-\alpha)}.$$

We now define $C_{i\alpha}^n$ for $\deg g_i < \alpha \leqslant i-1$ to be zero. Hence our result (5) may be rewritten as

$$(7)\qquad T^{-n}g_iT^iy = \sum_{\alpha=0}^{i-1} (-1)^\alpha C_{i\alpha}^n \Phi^\alpha(g_i)T^{-n+(i-\alpha)}y$$

where each $C_{i\alpha}^n \leqslant n^{d(i-\alpha)}$.

Apply the operator T^{-n} now to the functional equation (1), term by term. We obtain

$$T^{-n}y = \sum_{i=1}^{l}\sum_{\alpha=0}^{i-1} (-1)^\alpha C_{i\alpha}^n \Phi^\alpha(g_i)T^{-n+(i-\alpha)}y,$$

which becomes, where $i-\alpha = \beta$,

$$T^{-n}y = \sum_{\beta=0}^{l}\Big(\sum_{\alpha=0}^{l-\beta} (-1)^\alpha C_{\alpha+\beta,\alpha}^n \Phi^\alpha(g_{\alpha+\beta})\Big)T^{-n+\beta}y.$$

Set

$$d_\beta^n(x_1) = \sum_{\alpha=0}^{l-\beta} (-1)^\alpha C_{\alpha+\beta,\alpha}^n \Phi^\alpha(g_{\alpha+\beta})(x_1).$$

Then

$$(8)\qquad T^{-n}y(x_1) = \sum_{\beta=1}^{l} d_\beta^n(x_1)T^{-n+\beta}y(x_1)$$

where

$$|d_\beta^n(x_1)| = \Big|\sum_{\alpha=0}^{l-\beta} (-1)^\alpha C_{\alpha+\beta,\alpha}^n \Phi^\alpha(g_{\alpha+\beta})(x_1)\Big| \leqslant \big(K_1(x_1)n\big)^{d\beta}$$

for some $K_1(x) \geqslant 1$. Now we may write for $n \geqslant 1$

$$\begin{bmatrix} T^{-n}y(x_1) \\ \vdots \\ T^{-n+(l-1)}y(x_1) \end{bmatrix} = \theta_n \begin{bmatrix} T^{-(n-1)}y(x_1) \\ \vdots \\ T^{-(n-1)+(l-1)}y(x_1) \end{bmatrix}$$

where θ_n is an l by l matrix with the identity (the entries are in Q^{m^2}, the m by m matrices over Q) in all of the $i+1$, i positions $(1 \leqslant i \leqslant l-1)$, a first row consisting of $\big(d_1^n(x_1), \ldots, d_l^n(x_1)\big)$, and zeros everywhere else. Thus where $\theta_n = (\theta_n^{ij})$, we have

$$(9)\qquad |\theta_n^{ij}| \leqslant \big(K_1(x_1)n^d\big)^{1+j-i}.$$

We now show that an inverse matrix exists for θ_n. This amounts to showing the existence of

$$\big(d_l^n(x_1)\big)^{-1}.$$

By definition,

$$d_l^n(x_1) = C_{l0}^n g_l(x_1).$$

Since by (XIII) $g_l(x_1)$ is nonsingular, $g_l(x) \not\equiv 0$ and we have $0 < \deg g_l(x)$. Therefore C_{l0}^n is defined inductively by line (6). Using (6) we have

$$C_{l0}^n = C_{l0}^{n-1} = \ldots = C_{l0}^0 = 1.$$

Then

$$d_l^n(x_1) = g_l(x_1)$$

which as noted above is nonsingular. So θ_n^{-1} exists.

Since each $\Phi^j g_i(x_1)$ is rational $(1 \leqslant i \leqslant l, j \geqslant 0)$, we may choose a positive integer $K_2(x_1)$ such that each $K_2(x_1) d_i^n(x_1)$ is an integer. Hence each entry in $K_2(x_1)\theta_n$ belongs to Z^{m^2}, the m by m matrices over the integers. Set

$$Y = \begin{bmatrix} y(x_1) \\ \vdots \\ T^{l-1} y(x_1) \end{bmatrix}, \quad T^{-n} Y = \begin{bmatrix} T^{-n} y(x_1) \\ \vdots \\ T^{-n+(l-1)} y(x_1) \end{bmatrix}, \quad \frac{P}{q} = q^{-1} \begin{bmatrix} P_0 \\ \vdots \\ P_{l-1} \end{bmatrix}$$

and

$$\Delta_n = \big(K_2(x_1)\big)^n \theta_n \theta_{n-1} \ldots \theta_1 .$$

Then we have

$$\big(K_2(x_1)\big)^n T^{-n} Y = \Delta_n Y$$

where Δ_n^{-1} exists and the entries of Δ_n belong to Z^{m^2}. We define the absolute value of a matrix with matrix entries to be the maximum of the absolute values of the entries. Now write

$$\big(K_2(x_1)\big)^n T^{-n} Y = \Delta_n \left(Y - \frac{P}{q}\right) + \Delta_n \frac{P}{q}$$

or

$$\Delta_n \left(Y - \frac{P}{q}\right) = -\left(\Delta_n \frac{P}{q} - \big(K_2(x_1)\big)^n T^{-n} Y\right),$$

which implies that

$$\left|\Delta_n \left(Y - \frac{P}{q}\right)\right| \geqslant \left|\Delta_n \frac{P}{q}\right| - \left|\big(K_2(x_1)\big)^n T^{-n} Y\right| .$$

Since Δ_n^{-1} exists, $\Delta_n P/q$ is the zero vector if and only if P is zero. We exclude this case. (There are two possibilities. If each $T^i y(x_1) = 0$, $0 \leqslant i \leqslant l-1$, we have nothing to prove. If some $T^i y(x_1)$ is nonzero, it suffices to prove the theorem for nonzero P.) Then $|\Delta_n P/q| \geqslant 1/q$ since each entry in Δ_n belongs to Z^{m^2}. Using also the bound for $|T^{-n} y(x_1)|$ in (XI), we see that there exists $K_3(x_1)$, a positive real number independent of n, such that

$$l|\Delta_n| \max_i \left| T^i y(x_1) - \frac{P_i}{q} \right| \geqslant \frac{1}{q} - \big(K_3(x_1)/n\big)^{\delta n}.$$

Choose n sufficiently large that $\big(K_3(x_1)/n\big)^{\delta n} \leqslant 1/2q$. It follows then that

$$\max_i \left| T^i y(x_1) - \frac{P_i}{q} \right| \geqslant (l2q|\Delta_n|)^{-1}. \tag{10}$$

To be definite in our choice of n we take n to be the first positive integer such that $\big(K_3(x_1)/n\big)^{\delta n} < 1/2q$. Since $K_3(x_1) \geqslant 1$, we must have $n > 1$ and therefore we may write

$$\big(K_3(x_1)/n\big)^{\delta n} < 1/2q \leqslant \big(K_3(x_1)/(n-1)\big)^{\delta(n-1)}. \tag{11}$$

Observe that $\log\big(K_3(x_1)/n\big)^{\delta n}$ is asymptotic to $\log\big(K_3(x_1)/(n-1)\big)^{\delta(n-1)}$. Thus given $\varepsilon_2 > 0$, there exists N such that if our $n \geqslant N > 1$ we have

$$\big(K_3(x_1)/n\big)^{\delta n} \geqslant \big(K_3(x_1)/(n-1)\big)^{\delta(n-1)(1+\varepsilon_2)}.$$

Therefore, if $n \geqslant N$

$$\big(K_3(x_1)/n\big)^{\delta n} \geqslant \big(K_3(x_1)/(n-1)\big)^{\delta(n-1)(1+\varepsilon_2)} \geqslant (2q)^{-(1+\varepsilon_2)}. \tag{12}$$

If $n \leqslant K_3(x_1)$ then line (11) cannot hold. Thus $n > K_3(x_1)$. For $n > K_3(x_1)$ the extreme left-hand side of (11) decreases in a strictly monotone manner to zero as $n \to \infty$. Therefore, if we restrict ourselves to values of q which satisfy $\big(K_3(x_1)/N\big)^{\delta N} > 1/2q$, we must have $n > N$ and we may use (12). We restrict ourselves to values of q in this range in what follows.

At this point an estimate of $|\Delta_n|$ in terms of n is needed. The desired inequality is

$$|\Delta_n| \leqslant \big(K_4(x_1) n^d\big)^n \tag{13}$$

for some positive number $K_4(x_1)$ independent of n. We shall demonstrate that the theorem follows from (13) and end the proof of the theorem by showing (13). We have that

$$|\Delta_n| \leqslant \big(K_4(x_1) K_3^d(x_1)\big)^n \big(n/K_3(x_1)\big)^{nd},$$

so there exists $K_5(x_1) > 0$, independent of n, such that

$$|\Delta_n| \leqslant K_5(x_1) \big(n/K_3(x_1)\big)^{n(d+\varepsilon_2)},$$

as may be seen by consideration of the orders of growth of the functions involved. Now apply (12) to obtain

$$|\Delta_n| \leqslant K_5(x_1)(2q)^{(d+\varepsilon_2)(1+\varepsilon_2)/\delta}$$

which we use in (10). Then

$$\max_i |T^i y(x_1) - P_i/q| \geqslant [K_5(x_1) l (2q)^{1+(d+\varepsilon_2)(1+\varepsilon_2)/\delta}]^{-1},$$

if q is sufficiently large. We see that there exists $K_6(x_1) > 0$ such that

$$\max_i |T^i y(x_1) - P_i/q| \geqslant K_6(x_1) q^{-1-(d+\varepsilon_2)(1+\varepsilon_2)/\delta}.$$

Choose ε_2 such that $(d+\varepsilon_2)(1+\varepsilon_2) = d+\delta\varepsilon$. Then

$$(15) \qquad \max_i |T^i y(x_1) - P_i/q| \geqslant K_6(x_1) q^{-(1+d/\delta+\varepsilon)}.$$

Corresponding to the finite number of q's for which (12) does not hold find the minimal value of

$$\max_i |T^i y(x_1) - P_i/q|$$

and call it $K_7(x_1)$. If $K_7(x_1)$ is not zero let $C(\varepsilon)$ be the smaller of $K_7(x_1)$ and $K_6(x_1)$. Then

$$(16) \qquad \max_i |T^i y(x_1) - P_i/q| \geqslant C(\varepsilon) q^{-(1+d/\delta+\varepsilon)}$$

for all P/q with $q > 0$. If $K_7(x_1)$ is zero, then $Y = P/q_1$ for some q_1 small enough that (12) and, hence, (15) does not apply. But then obviously

$$Y = \frac{mP}{mq_1}$$

for $m = 1, 2, \ldots$. At some point (15) applies, and we have a contradiction. Thus (16) holds, and except for (13) the theorem has been proven.

Now to establish (13). Recall that

$$\Delta_n = \big(K_2(x_1)\big)^n \theta_n \theta_{n-1} \ldots \theta_1.$$

Hence we would be through if we could show $|\theta_n \ldots \theta_1| \leqslant \big(K_8(x_1) n^d\big)^n$, for some positive constant $K_8(x_1)$ independent of n. We shall establish by induction a much stronger statement that the term in the ith row and jth column of $\theta_n \ldots \theta_1$ is less in absolute value than

$$K_9(x_1)\big(K_1(x_1) lmn^d\big)^{n+j-i}$$

for some positive $K_9(x_1)$ independent of n. This would imply that

$$|\theta_n \ldots \theta_1| \leqslant K_9(x_1)\big(K_1(x_1) lmn^d\big)^{n+l-1} \leqslant \big(K_8(x_1) n^d\big)^n,$$

for some positive $K_8(x_1)$ independent of n. We choose $K_9(x_1)$ large enough that our statement is true for $n = 1, 2, \ldots, l-1$. Assuming the induction hypothesis for $n \geqslant l-1$, we shall show it for $n+1$. Recall line (9) now that

$$|\theta_n^{ij}| \leqslant \big(K_1(x_1) n^d\big)^{1+j-i}$$

where $\theta_n = (\theta_n^{ij})$. Then the i, jth term of $\theta_{n+1} \ldots \theta_1$ has absolute value less than or equal to

$$m \sum_{k=1}^{l} \big(K_1(x_1)(n+1)^d\big)^{(1+k-i)} K_9(x_1)\big(K_1(x_1) lmn^d\big)^{n+j-k}$$
$$\leqslant K_9(x_1)\big(K_1(x_1) lmn(n+1)^d\big)^{n+1+j-i}.$$

This completes the proof (13) and of the theorem.

References

[1] Witold Hurewicz, *Lectures on ordinary differential equations*, Cambridge, Mass., 1958.

[2] E. G. Poole, *Introduction to the theory of linear differential equations*, New York 1960.

UNIVERSITY OF ILLINOIS

Reçu par la Redaction le 14. 1. 1966