Applications of Padé approximations to diophantine inequalities in values of G-functions

D.V. Chudnovsky,*) G.V. Chudnovsky*)

Department of Mathematics
Columbia University
New York, N. Y. 10027

Introduction. In this lecture we study diophantine approximations to numbers represented as values of Siegel's G-functions [1]. The G-functions $f(x)$ are defined as solutions of linear differential equations over $\bar{\mathbb{Q}}(x)$ having an expansion at zero $f(x) = \Sigma_{n=0}^{\infty} a_n x^n$ with $a_n \in \bar{\mathbb{Q}}$ and $\overline{|a_n|} \leq c_f^n$, $\mathrm{denom}\{a_0,\ldots,a_n\} \leq c_f^n$ for a constant $c_f \geq 1$. These functions are important in the description of geometric objects (see §1 and [13]), and their values represent many classical constants (i.e. periods of algebraic varieties). Siegel [1] introduced G-functions and sketched a program of study of the arithmetic properties of values of G-functions at rational (algebraic) points near the origin. Some results along these lines were proved in [6]-[10], but under strong (G,C)-assumptions on linear differential equations satisfied by G-functions (the global nilpotence property,... etc. see §1). In this paper we prove the G-function results that Siegel sought, without any additional assumptions.

Our main results are collected in §1. § 1 also contains a discussion of G-functions, the (G,C)-property and its geometric and p-adic sense. Two of our key results are Theorems I and II on the absence of linear and algebraic relations between values of G-functions. The basis of all of our proofs is the method of Padé approximation of the second kind presented in §§2-4. The proof of Theorem I is presented in §5 and the proof of Theorem II is given in §§6-7. Another important result is Theorem III of §1, proved in §8, that any G-function is a (G,C)-function. As a consequence of this result and [11], any G-function

*) This work was supported in part by the U.S. Air Force under the Grant AFOSR-81-0190.

is a solution of a Fuchsian linear differential equation with rational exponents at regular singularities and the global nilpotence properties. For a discussion of the global nilpotence properties and the Grothendieck conjecture, see [4].

Our results on the diophantine inequalities for values of G-functions are particularly important for algebraic functions, when they imply effective bounds on solutions of diophantine equations. The relevant results for G-functions are formulated in §1 as Theorems V.

The method of their proof is the technique of graded Padé approximations [3], [19]. These theorems imply an effective version of a particular case of Schmidt's theorem [16], when algebraic numbers are values of algebraic G-functions near the origin [24]. Our results have the form of effective upper bounds on integral solutions of Norm-form equations, particularly Thue equations [19], depending on an integral parameter. The uniform bounds for Thue equations are established for the first time. We present a typical result of this form.

Theorem A: Let $n \geq 3$ and $F(x,y) \in \mathbb{Z}[x,y]$ be a polynomial of degree n in y, irreducible over $\mathbb{Q}[x,y]$. Let all real branches $y = y(x)$ of $F(x,y) \equiv 0$ have power series expansions in x^{-1} at $x = \infty$ with integral exponents (bounded from below) and with rational coefficients. Then the Thue equation depending on an integer parameter N:

$$f(X,Y;N) \overset{\text{def}}{=} Y^n \cdot F(N,\tfrac{X}{Y}) = A \tag{0.1}$$

has at most finitely many rationally parametrized solutions. These parametrized solutions have the form: $Y/X = P(N)/Q(N)$, A-fixed, for $P(x)$, $Q(x) \in \mathbb{Q}[x]$. Parametrized solutions can be determined as exceptionally good rational approximations $P(x)/Q(x)$ of a real branch $y = y(x)$ of $F(x,y) \equiv 0$ at $x = \infty$--and there are only finitely many such exceptionally good approximations [18]. With the exception of parametrized solutions, the Thue equation (0.1) has only finitely many integral solutions $(X,Y;N)$ for a fixed A. For any $\epsilon > 0$, and $N \geq N_1(\epsilon)$, the non-parametrized solutions X, Y of (0.1) are bounded from above as follows

$$\max(|X|,|Y|) \leq \gamma_0(\epsilon)|A|^{1/(n-2)-\epsilon} .$$

Here $\gamma_0(\epsilon) > 0$ is an effective constant depending on $\epsilon > 0$ and polynomially on the height of the polynomial $F(x,y)$.

§1. Siegel's G-functions.

Siegel had initiated in 1929 in [1] a program of study of arithmetic properties of values of analytic functions given as solutions of linear differential equations with additional arithmetic conditions on coefficients of their Taylor expansions. Two particular classes of functions were singled out in [1]. The first class of E-functions consists of entire functions $f(x) = \Sigma_{n=0}^{\infty} a_n x^n/n!$ such that $a_n \in \bar{\mathbb{Q}}$, $\overline{|a_n|} \leq n^{\epsilon n}$ and $\mathrm{denom}\{a_0,\ldots,a_n\} \leq n^{\epsilon n}$ for $n \geq n_0(\epsilon)$ and such that $f(x)$ satisfies a linear differential equation over $\bar{\mathbb{Q}}(x)$. Typical E-functions are the exponential function e^x and Bessel functions $J_\nu(x)$ with $\nu \in \mathbb{Q}$. For E-functions Siegel had proved very strong transcendence and algebraic independence results using the method of approximating forms [1], [2]. This method is essentially Padé approximation technique to solutions of linear differential equations [3]. The second class of functions, considered by Siegel in [1], called G-functions, consists of analytic functions $f(x) = \Sigma_{n=0}^{\infty} a_n x^n$ with $a_n \in \bar{\mathbb{Q}}$, $|a_n| \leq C^n$ and $\mathrm{denom}\{a_0,\ldots,a_n\} \leq C^n$ for some $C \geq 1$, such that $f(x)$ satisfies a linear differential equation over $\bar{\mathbb{Q}}(x)$. These functions are much more important for applications in diophantine geometry. This chapter is devoted to the discussion of geometric obstructions to Siegel's program to G-functions (expressed by p-curvature operators). We also present our new G-function theorems that overcome these obstructions and realize a large part of Siegel's program.

We use the standard notations of the algebraic number theory. For an algebraic number α and a complete set $\{\alpha_1 = \alpha,\ldots,\alpha_d\}$ of numbers algebraically conjugate to α, we denote by $\overline{|\alpha|} = \max\{|\alpha_1|,\ldots,|\alpha_d|\}$ the size of α. Also $\mathrm{den}(\alpha)$ denotes such a rational integer that $\mathrm{den}(\alpha)\cdot\alpha$ is an algebraic integer.

The product formula implies the following Liouville inequality: $|\mathrm{den}(\alpha)^d\cdot\alpha\cdot\overline{|\alpha|}^{d-1}| \geq 1$ where $\alpha \neq 0$ is an algebraic number of degree $\leq d$. Also we denote by $\mathrm{den}\{\alpha_0,\ldots,\alpha_n\}$ the common denominator of $\alpha_0,\ldots,\alpha_n$.

<u>Definition 1.1</u> (Siegel): Let $f(x) = \Sigma_{n=0}^{\infty} a_n x^n$ be a solution of a linear differential equation over $\bar{\mathbb{Q}}(x)$. $f(x)$ is called a G-function of $a_n \in \bar{\mathbb{Q}}$ and there exists a constant $C > 0$ such that $\overline{|a_n|} \leq C^n$ and the common denominator of $a_0,\ldots,a_n$ is at most C^n.

Remark 1.2: In fact, all coefficients a_n of the expansion of $f(x)$ belong to a fixed algebraic number field K. The field K is generated by the coefficients of a linear differential equation from $\bar{\mathbb{Q}}(x)[\frac{d}{dx}]$, satisfied by $f(x)$ and by first few a_n. For the purposes of this paper we can and will assume that $K = \mathbb{Q}$. Such a reduction to $K = \mathbb{Q}$ case can be achieved by considering simultaneously with $f(x) = \Sigma_{n=0}^{\infty} a_n x^n \in K[[x]]$ all functions $f^{(\sigma)}(x) = \Sigma_{n=0}^{\infty} a_n^{(\sigma)} x^n$ with $\sigma(a_n) = a_n^{(\sigma)}$ for isomorphic imbeddings $\sigma: K \to \mathbb{C}$. Then the functions $\mathrm{Sym}(f^{(\sigma)}(x): \sigma: K \to \mathbb{C})$, for all symmetric combinations Sym of $f^{(\sigma)}(x)$, are already G-functions with $K = \mathbb{Q}$ (i.e. one can replace $\bar{\mathbb{Q}}$ by $\mathbb{Q}$ in the Definition 1.1 above).

Obviously, algebraic functions are G-functions, because, by Eisenstein's theorem, for an algebraic function $f(x)$ with the expansion $f(x) = \Sigma_{n=0}^{\infty} a_n x^n$ with $a_n \in \bar{\mathbb{Q}}$, the common denominator of $a_0, \ldots, a_n$ divides $A \cdot B^n$ for appropriate integers A, B. Also, the class of G-functions is closed under integration, addition, multiplication and differentiation. In particular, solutions of Picard-Fuchs equations, including hypergeometric ${}_{m+1}F_m\left(\begin{matrix} a_1, \ldots, a_{m+1} \\ b_1, \ldots, b_m \end{matrix} \middle| x\right)$-functions with rational a_i, b_j also belong to the class of G-functions.

In [1], Siegel, while solving the problem of diophantine approximations to values of E-functions at algebraic points, made an indication that something similar could be done for values of G-functions. First of all, because G-functions have a finite radius of convergence and, obviously, their values at rational points are not necessarily irrational, there are natural restrictions on values of G-functions under consideration. Siegel proposed such conditions on (rational) points x, close to the origin for $x = \frac{p}{q}$, $|x| < |q|^{-\epsilon}$ or even $|x| < \exp(-\log|q|^{1/2+\epsilon})$ for $\epsilon > 0$ and large q, $|q| \geq q_0(\epsilon)$.

However, Siegel did not formulate any theorem on irrationality, measure of irrationality (or non-algebraicity of a bounded degree) for values of G-functions. Instead, he remarked that such theorems could be obtained, and gave a few examples: one concerning values of Abelian integrals and another dealing with values of particular hypergeometric functions. In fact, there are serious obstacles to any immediate attempts to extend Siegel's theory from E-functions to G-functions. The reason for this lies in the necessity fo bound the denominators of the coefficients of the expansion of a G-function $f(x)$ at points $x = t$, distinct from zero. To see the reason for this "global" condition on $f(x)$, we review briefly Siegel's method.

Siegel's method is based on construction of a system of approximating forms, or Padé approximations. These forms are constructed in the following way. Let $f_1(x) = f(x),\ldots,f_n(x)$ be a system of G-functions satisfying a system of the first order linear differential equations over $\bar{\mathbb{Q}}(x)$:

$$\frac{d}{dx}\vec{f}^{\,t} = A\cdot\vec{f}^{\,t}, \tag{1.1}$$

$\vec{f} = (f_1(x),\ldots,f_n(x))$ and $A = A(x) \in M(n,\bar{\mathbb{Q}}(x))$. Typically, $f_i(x) = (\frac{d}{dx})^{i-1}f(x)$, where $f(x)$ satisfies a scalar linear differential equation over $\bar{\mathbb{Q}}(x)$ of order n.

An approximating form for $f_1(x),\ldots,f_n(x)$ (or a remainder function in Padé-type approximation problem for $f_1(x),\ldots,f_n(x)$) has the form:

$$L(x) = P_1(x)f_1(x) +\ldots+ P_n(x)f_n(x).$$

Here $P_i(x)$ are polynomials $P_i(x) \in \bar{\mathbb{Q}}[x]$ of degrees at most D: $i = 1,\ldots,n$; where $L(x)$ has a zero at $x = 0$ of order at least $nD - [\epsilon D]$ for some $\epsilon > 0$.

Since $f_i(x)$: $i = 1,\ldots,n$ are G-functions, one can always find $P_i(x)$ with integral coefficients of sizes at most $C_1^{nD/\epsilon}$: $i = 1,\ldots,n$ for a constant $C_1 > 0$ (depending only on $f_1,\ldots,f_n$). This is achieved using the Thue-Siegel lemma [1], [5] being a version of Dirichlet's box principle.

Siegel's theory of approximating forms (developed by him for E-functions) predicts the existence of n linearly independent forms in $f_1(x),\ldots,f_n(x)$:

$$L_i(x) = P_{1,i}(x)f_1(x) +\ldots+ P_{n,i}(x)f_i(x), \tag{1.2}$$

where

$$L_i(x) = \frac{1}{k_i!}\cdot D(x)^{k_i}\cdot(\frac{d}{dx})^{k_i}L(x), \quad i = 1,\ldots,n$$

and $0 = k_0 <\ldots< k_n$ and $k_n \le \epsilon D + C_2$. Here $D(x)$ is a (polynomial) denominator of the elements of the matrix $A = A(x)$. In (1.2) all $P_{j,i}(x)$ $(i,j = 1,\ldots,n)$ are polynomials (from $\bar{\mathbb{Q}}[x]$). Let now the common denominators of the coefficients of the polynomials $P_{j,i}(x)$ $(i,j=1,\ldots,n)$ are not too big (like c_2^D for $c_2 > 0$). Then from the system (1.2) of n linearly independent forms one can immediately get (see the discussion in [3]) a nontrivial lower bound for a linear form in numbers $f_1(x_0),\ldots,f_n(x_0)$ with arbitrary integral coefficients

$$|q_1 f_1(x_0) + \ldots + q_n f_n(x_0)| > |q|^{-\lambda},$$

$|q| = \max(|q_1|,\ldots,|q_n|) \geq q_0$ and $\lambda > 0$ for a rational x_0 very close to the origin.

However, the sizes of the denominators of coefficients of $P_{j,i}(x)$ should grow as $\sim k_i!$ for large k_i. Indeed, to differentiate $L(x)$ k times we need to iterate the equation (1.1) k times. We get

$$(1.3) \qquad \left(\frac{d}{dx}\right)^k - A_k(x) \equiv 0 \ (\mathrm{mod}\ "\frac{d}{dx} - A(x)").$$

Here, in general, $k!$ does not divide the coefficients of polynomial entries of $D(x)^k \cdot A_k(x)$.

However, one sees that Siegel's method can be applied to G-functions $f_1(x),\ldots,f_n(x)$ satisfying equations (1.1), if the following additional (G,C)-assumption is imposed:

(G,C)-Assumption: We call functions $f_1(x),\ldots,f_n(x)$ (G,C)-functions, if they are G-functions, and, for a differential equation (1.1), that they satisfy, the common denominator of all coefficients of polynomial entries of matrices $\frac{1}{k!} D(x)^k \cdot A_k(x)$: $k = 0,1,\ldots,N$ grows not faster than c_2^N for a constant c_2.

This kind of assumption was first proposed by Galočkin [6]. (G,C)-assumption is correct for algebraic functions and solutions of Picard-Fuchs equations (because of their p-adic behavior, see later). This (G,C)-assumption is explicit in all nontrivial results on G-functions obtained since 1929 (Galočkin [6], Väänänen [7], Flicker [8], Bombieri [9], authors [10]).

It turns out, however, that the general G-function theory, as Siegel hoped for, and which is similar to the E-function theory can be constructed without any use of additional (G,C)-function assumptions. To see why this is so remarkable, one should realize that (G,C)-function assumption is an important p-adic condition and, unlike

G-function assumption has a direct geometric sense after reduction (mod p). One can visualize this by introducing p-curvature operators (Cartier, Grothendieck, Deligne, Katz, Dwork, see [11], [12], [13]):

$$\psi_p = (\frac{d}{dx}\cdot I - A)^p \pmod p.$$

In fact, (mod p), ψ_p is a linear operator and, in the notations of (1.3), [11]:

$$\psi_p = -A_p \pmod p.$$

Then the (G,C)-function assumption implies, in particular, that the operators ψ_p are nilpotent for almost all prime p. Moreover, the (G,C)-assumption can be reformulated in terms of the relation between the p-adic radius of convergence of solutions of (1.1) at a "generic point t" for almost all p [9]. According to Bombieri [9], equation (1.1) is of "arithmetic type" if $\Sigma_p\ \epsilon_p \log p < \infty$, where all solutions of (1.1), expanded at a "generic" point $x = t$, converge for $\mathrm{ord}_p(x-t) > \epsilon_p$. Then any equation of "arithmetic type" satisfies (G,C)-assumptions [9].

Obviously, the global nilpotence of (1.1) (i.e. the condition that ψ_p is nilpotent for almost all p) is a very restrictive arithmetic condition. It is widely suspected that all globally nilpotent equation in Dwork's phrase, "come from Geometry" [12]. The most known among these equations is the class of Picard-Fuchs equations for periods of algebraic varieties.

The G-function condition is, on the contrary, local and requires only power series expansion of <u>one</u> solution (<u>not n</u>) at <u>one</u> (<u>and not generic</u>) point. That is why the fulfillment of Siegel's program for G-functions is so important.

Our proof of G-function theorems relies on Padé approximation theory, but this time we are using Padé approximants of the second kind or Germanic polynomials (in the sense of Mahler [14]).

We present one of our main results on G-functions, that follow Siegel's program. The proof of this Theorem I is given in §5.

<u>Theorem I</u>: Let $f_1(x),\ldots,f_n(x)$ be G-functions with rational coefficients in their Taylor expansions, satisfying a first order linear

differential system (1.1) over $\mathbb{Q}(x)$, and such that functions 1, $f_1(x),\dots,f_n(x)$ are linearly independent over $\mathbb{Q}(x)$. Then for any $\epsilon > 0$ and arbitrary rational $r = a/b$ with (rational) integers a and b such that $|b|^{\epsilon} \geq c_3|a|^{(n+1)(n+\epsilon)}$, $r \neq 0$, the numbers 1, $f_1(r),\dots,f_n(r)$ are linearly independent over $\mathbb{Q}$; and for arbitrary rational integers $H_0,H_1,\dots,H_n$ we have

$$|H_0 + H_1f_1(\tfrac{a}{b}) + \dots + H_nf_n(\tfrac{a}{b})| > H^{-n-\epsilon},$$

with $H = \max(|H_0|,\dots,|H_n|)$, when $H \geq h_0$. Here $c_3 = c_3(f_1,\dots,f_n,\epsilon) > 0$, $h_0 = h_0(f_1,\dots,f_n,\epsilon,r) > 0$ are effective constants.

In general, we have

$$|H_0 + H_1f_1(\tfrac{a}{b}) + \dots + H_nf_n(\tfrac{a}{b})| > H^{\lambda-\epsilon}$$

with $\lambda = -n \log|b|/\log|b/a^{n+1}|$, whenever $|b| \geq c_4|a|^{n+1}$ and $H \geq h_1$ in the notations above, with effective constants $c_4 = c_4(f_1,\dots,f_n,n) > 0$ and $h_1 = h_1(f_1,\dots,f_n,n,r) > 0$.

The method of proof of Theorem I, and other similar results, is not based, like in Siegel's method for E-functions, on direct construction of approximating forms to $1,f_1(x),\dots,f_n(x)$. Rather we constuct Padé-type approximations of the second kind to $f_1(x),\dots,f_n(x)$. This system of Padé-type approximations takes the following form:

$$R_i(x) \stackrel{\text{def}}{=} Q(x)\cdot f_i(x) - P_i(x):\ i = 1,\dots,n,$$

where $P_1(x),\dots,P_n(x)$, $Q(x)$ are polynomials in x of degrees at most D, and such that

$$\operatorname{ord}_{x=0}R_i(x) \geq D + \frac{D}{n} - \epsilon D:$$

$i = 1,\dots,n$.

It is much easier to control denominators of derivatives of remainder functions $R_i(x)$ for Padé-type approximations of the second kind:

$$R_{i,k} = \frac{1}{k!}D(x)^k\cdot(\frac{d}{dx})^kR_i(x),$$

than for Padé-type approximants of the first kind. On the other hand there is a well-known duality principle that expresses Padé(-type) approximants of the first kind in terms of contiguous Padé (-type) approximants of the second kind, and vice versa [14]. This duality principle was axiomatized by Mahler [15] in his studies of integral points and successive minima of convex bodies in archimedian and non-archimedian metrics. (It is, in fact, one of the versions of Khintchine's transference principle, that corresponds to reciprocal parallelepipeds, see [16].) The duality principle allows us to pass from the remainder functions $R_{i,k}$ to a system of approximating forms with controllable denominators of coefficients of polynomial approximants. This is the key to proof of Theorem I and other similar results.

If one looks only on linear independence (irrationality) statements for values of G-functions, then restrictive conditions of Theorem I on $r = a/b$ can be considerably relaxed. We give an example of only one such result, where the (G,C)-assumption is not used.

Theorem II: Let $f_1(x),\dots,f_n(x)$ be G-functions satisfying matrix first order linear differential equations (1.1) over $\bar{\mathbb{Q}}(x)$, and such that functions $1, f_1(x),\dots,f_n(x)$ are algebraically independent over $\bar{\mathbb{Q}}(x)$. Then for any $t \geq 1$ there exists an effective constant $c_5 = c_5(f_1,\dots,f_n,t) > 0$ such that for any algebraic number $\xi \neq 0$ of degree $\leq t$, it follows from

$$|\xi| < \exp(-c_5\{\log\overline{|\xi|}\}^{\frac{4n}{4n+1}}), \tag{1.4}$$

that numbers

$$1, f_1(\xi),\dots,f_n(\xi)$$

are not related by an algebraic relation of degree $\leq t$ over $\mathbb{Q}(\xi)$.

The duality between Padé approximants of the first and the second kind enables us to settle a longstanding problem [10] on the relationship between G-functions and (G,C)-functions. It turns out that every G-function is, in fact, a (G,C)-function! Such a result fulfills another part in Siegel's program:

Theorem III: Let $f_1(x),\dots,f_n(x)$ be a system of G-functions satisfying a system of first order linear differential equations (1.1) over $\bar{\mathbb{Q}}(x)$. If $f_1(x),\dots,f_n(x)$ are linearly independent over $\bar{\mathbb{Q}}(x)$, then the functions $f_1(x),\dots,f_n(x)$ are (G,C)-functions.

Proof of Theorem III is presented below in §8. We remark that the condition of linear independence of $f_1(x),\dots,f_n(x)$ is clearly necessary, even in the case of scalar linear differential equations, when $f_i(x) = f^{(i-1)}(x)$: $i = 1,\dots,n$. E.g. one can consider $n = 2$ and the linear differential equation $f'' - f' = 0$ (two solutions: $f = 1$ and $f = e^x$). The function $f \equiv 1$ is a G-function, but the (G,C)-assumption is clearly false for the equation $f'' - f' = 0$.

Theorem III also implies that every equation of order n having a G-function solution (that does not satisfy an equation of order $n-1$ over $\bar{\mathbb{Q}}(x)$) is of "arithmetic type".

Our novel approach to G-functions, used in proofs of Theorems I-III, opens an opportunity to apply Padé approximations to the study of globally nilpotent linear differential equations. This new method is the basis of our results on the Grothendieck conjecture. These results are presented in [4].

It is known that bounds on linear forms in values of G-functions can be obtained in non-archimedian, as well as in achimedian metric, provided the (G,C)-assumption is met, see [7], [9]. Thus, we can use Theorem III to obtain results on simultaneous approximations in several metrics. However, it is easier to use directly the proofs of Theorems I and II. To formulate the kind of results we obtain, let us denote for an arbitrary G-function $f(x) = \Sigma_{n=0}^{\infty} a_n x^n$ ($a_n \in K$: $n = 0,1,\dots$) and a place v of K, by $f^{(v)}(x)$ the function defined on the completion of K corresponding to v. E.g. for the one (i-th) archimedian place v_i corresponding to the imbedding $\alpha \mapsto \alpha^{(i)}$ of $K \hookrightarrow \mathbb{C}$, $f^{(v_i)}(x) = \Sigma_{n=0}^{\infty} a_n^{(i)} x^n$. For a non-archimedian place v, $f^{(v)}(x)$ is defined on the completion K_v of K. E.g. the value of $f^{(v)}(x)$ at $x \in K_v$ is a v-adic number from K_v and it can be different from the value of $f(x)$, even when $x \in K \subset K_v$. Because $f(x)$ is a G-function, every function $f^{(v)}(x)$ has a nonzero radius of convergence in K_v.

In these notations results of Theorems I and II hold for any function $f^{(v)}(x)$. Namely, we have

Remark: If in Theorem II we consider K containing $\mathbb{Q}(\xi)$ and the field of coefficients of expansions of $f_i(x)$: $i = 1,\dots,n$, then the results of Theorem II holds for functions $f_1^{(v)}(x),\dots,f_n^{(v)}(x)$ instead of $f_1(x),\dots,f_n(x)$. One has only to replace $|\xi|$ in (1.4) by $|\xi|_v$ and (complex) numbers $f_1(\xi),\ldots,f_n(\xi)$ by v-adic ones: $f_1^{(v)}(\xi),\ldots,f_n^{(v)}(\xi)$ from K_v. The constant c_3 then depends on v as well. Similarly, in Theorem I, under the assumptions $|a/b|_v \le c_3 \max(|a|,|b|)^{\varepsilon/(n+1)(n+\varepsilon)-1}$, the v-adic numbers $1,f_1^{(v)}(r),\ldots,f_n^{(v)}(r)$ are linearly independent over $\mathbb{Q}$ and $|H_0 + H_1f_1(r) +\ldots+ H_nf_n(r)|_v > H^{-n-\varepsilon}$ with $H = \max(|H_0|,\ldots,|H_n|) \ge h_0$ and c_3, h_0 depending on v.

The results of Theorem II form is the kind of result Siegel had expected. One hopes, however, to go beyond the fulfillment of Siegel's program and to attempt to remove the condition (1.4) or, at least, to weaken it considerably. Though we cannot yet report an ultimate progress here, some progress has been achieved. We report one such result.

<u>Theorem IV</u>: Let $f_1(x),\ldots,f_n(x)$ be G-functions satisfying matrix first order linear differential equations (1.1) over $\bar{\mathbb{Q}}(x)$, and such that $f_1(x),\ldots,f_n(x)$ are algebraically independent over $\bar{\mathbb{Q}}(x)$. Then for any $d \ge 1$ and $\varepsilon > 0$ there exists $c_6 = c_6(f_1,\ldots,f_n,d,\varepsilon) > 0$ such that for any algebraic number $\xi \ne 0$ of degree $\le d$, from

$$|\xi| < \exp(-c_6\{\log\log H(\xi)\}^{1+\varepsilon}),$$

it follows that

$$f_1(\xi),\ldots,f_n(\xi)$$

are not related by an algebraic relation of degree $\le d$ over $\mathbb{Q}(\xi)$.

For applications of diophantine inequalities to diophantine equations one needs a version of Theorem I, when $f_1(x),\ldots,f_n(x)$ do not satisfy any more a first order system of linear differential equations (1.1), but instead are solutions of linear differential equations of an arbitrary order over $\bar{\mathbb{Q}}(x)$. Such results for functional diophantine approximations correspond to Kolchin's type problems [17] and to analogs of Schmidt's theorem [16] for solutions of linear differential equations proved in [18], [19]. Similar results for E-functions, particularly for sums of exponential functions were proved by authors using methods of graded Padé approximations [3]. We use methods of graded Padé approximations to prove results close to the best possible for values of G-functions satisfying linear differential equations of an arbitrary order. One of our results is the following.

Theorem V: Let $f_1(x),\dots,f_n(x)$ be G-functions with rational number coefficients of Taylor expansions, satisfying linear differential equations over $\mathbb{Q}(x)$. Then for any $\epsilon > 0$ and a rational number $r = a/b$, with integers a and b such that $|b| \geq c_7|a|^{n(n-1+\epsilon)}$, $c_7 = c_7(f_1,\dots,f_n,\epsilon) > 0$ we have the following lower bound for linear forms in $f_1(r),\dots,f_n(r)$. For arbitrary non-zero rational integers $H_1,\dots,H_n$ and $H = \max(|H_1|,\dots,|H_n|)$, if $H_1f_1(r) +\dots+ H_nf_n(r) \neq 0$, then

$$|H_1f_1(r) +\dots+ H_nf_n(r)| > |H_1 \dots H_n|^{-1}\cdot H^{1-\epsilon},$$

provided that $H \geq c_8$ with $c_8 = c_8(f_1,\dots,f_n,r,\epsilon) > 0$, and effective $c_7 > 0$, $c_8 > 0$.

Under the same assumptions on r, for linearly independent over $\mathbb{Q}(x)$ functions $1, f_1(x),\dots,f_n(x)$ and arbitrary rational integers $q, q_1,\dots,q_n$ we have:

$$|q_1\dots q_n|^{1+\epsilon}\cdot\|q_1\cdot f_1(r) +\dots+ q_n\cdot f_n(r)\| > 1$$

and

$$|q|^{1+\epsilon}\cdot\|qf_1(r)\| \cdots \|qf_n(r)\| > 1,$$

provided that $|q_1\dots q_n| > c_9$ and $|q| > c_9$. Here $\|\cdot\|$ is the distance to the nearest integer, and $c_9 = c_9(f_1,\dots,f_n,r,\epsilon) > 0$ is an effective constant.

In all results above we can also explicitly exhibit the dependence of constants c_8, and c_9 on r, namely on $|b|$. This is of particular importance in our applications to algebraic functions, where $r = 1/b$ with varying b. For example, under the assumptions of Theorem IV we have for rational integers $H_1,\dots,H_n$:

$$|H_1f_1(r) +\dots+ H_nf_n(r)| > |b|^{-n}\cdot H^{\lambda-\epsilon},$$

with $\lambda = -(n-1)\log|b|/\log|b/a^n|$, $H = \max(|H_0|,\dots,|H_n|)$ provided that $H \geq c_{10}(f_1,\dots,f_n,\epsilon)$.

Proof of Theorems V is based on graded Padé approximation methods developed by authors [19]-[20]. The essence of these methods consists in simultaneous approximations of all elements of graded submodules in Picard-Vessiot extensions of $\mathbb{C}(x)$ generated by linear differential equations satisfied by $f_1(x),\ldots,f_n(x)$. Namely, let under the assumptions of Theorem IV, functions $f_i(x)$ satisfy scalar linear differential equations over $\mathbb{Q}(x)$ of orders k_i: $i = 1,\ldots,n$. We introduce auxiliary variables $c_{i,j}$ $(j = 1,\ldots,k_i;\ i = 1,\ldots,n)$ and $\bar{c}_i = (c_{i,1},\ldots,c_{i,n})$: $i = 1,\ldots,n$.

<u>Definition 1.2</u> <u>(Of graded Padé approximations)</u>: Let $P_i(x|\bar{c})$: $i = 1,\ldots,n$ be polynomials in x of degrees at most D and in $\bar{c} = (\bar{c}_1,\ldots,\bar{c}_n)$, homogeneous in each group of variables $\bar{c}_j = (c_{j,1},\ldots,c_{j,k_j})$: $j \neq i$ of degree N, and in variables $\bar{c}_i = (c_{i,1},\ldots,c_{i,k_i})$ of degree $N-1$: $i = 1,\ldots,n$. Let the remainder function

$$R(x|\bar{c}) = \Sigma_{i=1}^{n} P_i(x|\bar{c})\cdot\{\Sigma_{j=1}^{k_i} c_{i,j}\cdot f_i^{(j-1)}(x)\}$$

has a zero at $x = 0$ of order at least t, for any choice of $\bar{c} = (\bar{c}_1,\ldots,\bar{c}_n)$. If $t \geq nD - \epsilon D$ and

$$t \geq \Sigma_{i=1}^{n} D\cdot\frac{\binom{N+k_i-2}{k_i-1}}{\binom{N+k_i-1}{k_i-1}} - \epsilon\frac{D}{\prod_{i=1}^{n}\binom{N+k_i-1}{k_i-1}},$$

then $P_i(x|\bar{c})$ are called Padé approximants and $R(x|\bar{c})$ is called a remainder function in the ϵ-graded Padé approximation problem with weight D of level N.

Using the specialization of the remainder functions $R(x|\bar{c})$ and their x-derivatives, we prove Theorems IV-V, applying a version of Siegel's theorem, similar to our E-function results in [20].

§2. Padé-type approximants of the second kind.

In this chapter we study Padé-type approximations of the second kind to a system of functions satisfying linear differential equations with rational function coefficients. We start our presentation with Mahler's [14] definition of Padé approximants of the first and the second kind to an arbitrary system $f_1(x),\dots,f_m(x)$ of functions given by formal power series expansion at $x = 0$.

Definition 2.1: For m functions $f_1(x),\dots,f_m(x)$ given by formal power series at $x = 0$ and m non-negative integers $n_1,\dots,n_m$, we consider m polynomials $A_1(x),\dots,A_m(x)$ of degrees of at most $n_1,\dots,n_m$, respectively, such that the function

$$R(x) = A_1(x)f_1(x) + \dots + A_m f_m(x)$$

has a zero at $x = 0$ of the order of at least

$$\Sigma_{i=1}^{m} (n_i + 1) - 1.$$

The polynomials $A_i(x)$ are called Padé approximants of the first kind and are denoted by $A_i(x|n_1,\dots,n_m)$: $i = 1,\dots,m$. The function $R(x)$ is called the remainder function and is denoted by $R(x|n_1,\dots,n_m)$.

Definition 2.2: Let $f_1(x),\dots,f_m(x)$ be a formal power series and $n_1,\dots,n_m$ be non-negative integers. We say that the system of polynomials $(\mathcal{O}\!\!\mathcal{L}_1(x),\dots,\mathcal{O}\!\!\mathcal{L}_m(x))$ is the system of Padé approximants of the second kind to the system of functions $f_1(x),\dots,f_m(x)$ with weights $n_1,\dots,n_m$, if the following conditions are satisfied:

i) the polynomials $\mathcal{O}\!\!\mathcal{L}_1(x),\dots,\mathcal{O}\!\!\mathcal{L}_m(x)$ are not all zero;

ii) $\deg \mathcal{O}\!\!\mathcal{L}_i(x) \le \Sigma_{j=1, j\neq i}^{m} n_j = \sigma - n_i$ for $\sigma = n_1 + \dots + n_m$, $i = 1,\dots,m$;

iii) the order of zero of the function $\mathcal{O}\!\!\mathcal{L}_k(x)f_\ell(x) - \mathcal{O}\!\!\mathcal{L}_\ell(x)f_k(x)$ at $x = 0$ is at least $\Sigma_{i=1}^{m} n_i + 1 = \sigma + 1$; $k,\ell = 1,\dots,m$.

We denote Padé approximants of the second kind with weights $n_1,\dots,n_m$ as follows: $\mathcal{O}\!\!\mathcal{L}_i(x) = \mathcal{O}\!\!\mathcal{L}_i(x|n_1,\dots,n_m)$, $i = 1,\dots,m$.

Mahler [14] was the first to establish important duality relations between Padé approximants of the first and the second kind. These relations are a part of a more general duality principle from geometry of numbers (transference principles), see [21], [15], [16]. Mahler's relations [14] between Padé approximants of the first and the second kind can be represented in the matrix form as follows. We denote

$$A(x|n_1,\dots,n_m) = (A_i(x|n_1+\delta_{j1},\dots,n_m+\delta_{jm}))_{i,j=1}^{m};$$

$$\mathcal{A}(x|n_1,\dots,n_m) = (\mathcal{A}_i(x|n_1-\delta_{j1},\dots,n_m-\delta_{jm}))_{i,j=1}^{m}.$$

Then for (normalization) constants $c_1,\dots,c_m$:

$$A(x|n_1,\dots,n_m)\cdot\mathcal{A}(x|n_1,\dots,n_m)^t = \begin{pmatrix} c_1x^\sigma & & 0 \\ & \ddots & \\ 0 & & c_mx^\sigma \end{pmatrix}$$

with $\sigma = n_1+\dots+n_m$.

We refer the reader to [22] for further study of the relationship of Padé approximants of the first and the second kind.

For arithmetic applications, Padé approximants are not always convenient to use because of difficulties connected with denominators of their coefficients. That is why it is useful to apply Padé-type approximations, that are very similar to Padé approximations, with the difference that the order of zero of the remainder function is not the maximal possible. The theory of such Padé-type approximations is sufficiently developed, cf. [23].

We start with the introduction of new useful notations.

<u>Definition 2.3</u>: Let $g_i(x)$, $i = 1,\dots,n$ be functions regular at $x = 0$ and let M,D and D_0 be nonnegative integers. Let $Q(x)$ be a non-zero polynomial of degree of at most D_0. Then for every $i = 1,\dots, n$ there exists a unique polynomial

$$[Q\cdot g_i]_D \overset{\text{def}}{=} P_i(x)$$

of degree of at most D, such that $\text{ord}_{x=0}(Q(x)g_i(x) - P_i(x)) \geq D+1$.

To describe explicitly the polynomial $[Q\cdot g_i]_D$ we need the following simple lemma:

Lemma 2.4: Let $g_i(x) = \Sigma_{m=0}^{\infty} b_{m,i}x^m$, $i = 1,\ldots,n$, and let $Q(x) = \Sigma_{m=0}^{D_0} q_m x^m$. Then for $i = 1,\ldots,n$, the polynomial $P_i(x) = [Qg_i]_D$ has the form $P_i(x) = \Sigma_{m=0}^{D} p_{m,i}x^m$, where

$$(2.1) \qquad p_{m,i} = \Sigma_{\substack{k+\ell=m,\\ k\le D_0}} q_k \cdot b_{\ell,i}, \qquad m = 0,\ldots,D.$$

In particular, if $Q(x)$ has algebraic integer coefficients and a height $\le H_Q$, and Δ_D is a common denominator of algebraic numbers $b_{m,i}$, $m = 0,\ldots,D$; $i = 1,\ldots,n$, then the common denominator of coefficients of $P_i(x)$ divides Δ_D, and the height of $\Delta_D P_i(x)$ is bounded by $(D+1)H_Q C_D$, where $\max\{\Delta_D \cdot b_{m,i}: m \le D,\ i = 1,\ldots,n\} \le C_D$.

Proof of Lemma 2.4: First of all we notice that for an arbitrary polynomial $P(x) = \Sigma_{m=0}^{D} p_m x^m$, we have the following expansion of $Q(x)\cdot g_i(x) - P(x)$: $\Sigma_{m=0}^{\infty} x^m\{\Sigma_{k+\ell=m,k\le D_0} q_k \cdot b_{\ell,i} - p_m\}$. This and Definition 2.3 imply the expression (2.1) for the coefficients of $P_i(x) = [Q\cdot g_i]_D$. The second part of Lemma 2.4 follows immediately from (2.1).

Definition 2.5: Let D, D_0 and M be non-negative integers. Let $Q(x)$ be a non-zero polynomial of degree at most D_0 and let $P_i(x) \overset{\text{def}}{=} [Q\cdot g_i]_D$ $(i = 1,\ldots,n)$ for n functions $g_1(x),\ldots,g_n(x)$ regular at $x = 0$. If we now have

$$\text{ord}_{x=0}(Q(x)\cdot g_i(x) - P_i(x)) \ge D + M + 1$$

for every $i = 1,\ldots,n$, then the system of polynomials $(Q(x);P_1(x),\ldots,P_n(x))$ is called a system of Padé-type approximants of the second kind with weights D_0 and D and order M of approximation.

According to Dirichlet's box principle, Padé-type approximations of the second kind with weights D_0 and D and order M of approximation exist whenever $D_0 \ge nM$. We say briefly that the system $(Q;P_1,\ldots,P_n)$ of Padé-type approximants (of the second kind) has parameters (D_0,D,M).

We study now the Padé approximations (of the second kind) to a system of functions satisfying linear differential equations. This system of functions is denoted by $f_1(x),\ldots,f_n(x)$. We assume that all functions $f_1(x),\ldots,f_n(x)$ are regular at $x = 0$ and satisfy the following system of first order matrix linear differential equations

(2.2) $$\frac{d}{dx} f_i(x) = \Sigma_{j=1}^{n} A_{i,j}(x) f_j(x),$$

$i = 1,\dots,n$. Here $A_{i,j}(x)$ are rational functions $(i,j = 1,\dots,n)$ and we denote by $D(x)$ the (polynomial) common denominator of rational functions $A_{i,j}(x)$ $(i,j = 1,\dots,n)$. Let $d = \max\{\deg(D)-1, \deg(DA_{i,j});\ i,j = 1,\dots,n\}$. In applications, functions $f_1(x),\dots,f_n(x)$ are G-functions and $A_{i,j}(x) \in \mathbb{Q}(x), i,j = 1,\dots,n$.

Theorem 2.6: Let $f_1(x),\dots,f_n(x)$ satisfy a system of equations (2.2). Let $(Q(x);P_1(x),\dots,P_n(x))$ be Padé approximants to a system of functions $f_1(x),\dots,f_n(x)$ with parameters (D_0,D,M). Let $k \geq 0$ and $M \geq k(d+1)$.

Let us define

(2.3) $$Q^{\langle k\rangle}(x) = \frac{1}{k!}\cdot D(x)^k \cdot \left(\frac{d}{dx}\right)^k Q(x),$$

and $P_i^{\langle k\rangle}(x) \overset{\text{def}}{=} [Q^{\langle k\rangle}(x)\cdot f_i(x)]_{(D+kd)}$: $i = 1,\dots,n$. In these notations, the polynomials $(Q^{\langle k\rangle}(x);P_1^{\langle k\rangle}(x),\dots,P_n^{\langle k\rangle}(x))$ are Padé-type approximants to a system of functions $f_1(x),\dots,f_n(x)$ with parameters $(D_0 + kd, D + kd, M - k(d+1))$.

Proof: For $k = 0$, see Definitions 2.3 and 2.5. Let us assume that for all $k < K$ and polynomials $Q^{\langle k\rangle}(x)$ and $P_i^{\langle k\rangle}(x) = [Q^{\langle k\rangle}(x)\cdot f_i(x)]_{(D+kd)}$, $i = 1,\dots,n$, the function

(2.4) $$R_i^{\langle k\rangle}(x) \overset{\text{def}}{=} Q^{\langle k\rangle}(x)\cdot f_i(x) - P_i^{\langle k\rangle}(x).$$

has a zero at $x = 0$ order of at least $D + M - k + 1$, $i = 1,\dots,n$. We differentiate the equations (2.4) and take their linear combinations with appropriate polynomial coefficients. We have $\left(\frac{d}{dx}\right)^k Q(x) = k!D(x)^{-k}Q^{\langle k\rangle}(x)$. Thus

(2.5) $$Q^{\langle K\rangle}(x) = -\frac{(K-1)}{K}\cdot D'(x)\cdot Q^{\langle K-1\rangle}(x) + \frac{1}{K}\cdot D(x)\cdot \frac{d}{dx}(Q^{\langle K-1\rangle}(x)).$$

also, for $i = 1,\dots,n$ we have:

(2.6) $$\frac{d}{dx}(R_i^{\langle k\rangle}(x)) = \frac{d}{dx}(Q^{\langle k\rangle}(x))\cdot f_i(x) + Q^{\langle k\rangle}(x)\cdot \Sigma_{j=1}^{n} A_{i,j}(x) f_j(x) - \frac{d}{dx}(P_i^{\langle k\rangle}(x)).$$

Thus, from (2.5) and (2.6) it follows that

$$(2.7)\quad \begin{aligned}&-\frac{(K-1)}{K}D'(x)\cdot R_i^{\langle K-1\rangle}(x) + \frac{1}{K}D(x)\cdot\{\frac{d}{dx}(R_i^{\langle K-1\rangle}(x)) - \\ &- \Sigma_{j=1}^n A_{i,j}R_j^{\langle K-1\rangle}(x)\} = Q^{\langle K\rangle}(x)\cdot f_i(x) - P_{i,K}(x).\end{aligned}$$

Here, in (2.7),

$$\begin{aligned}P_{i,K}(x) &= -\frac{(K-1)}{K}D'(x)\cdot P_i^{\langle K-1\rangle}(x) + \frac{1}{K}D(x)\cdot\{\frac{d}{dx}(P_i^{\langle K-1\rangle}(x)) \\ &- \Sigma_{j=1}^n A_{i,j}(x)P_j^{\langle K-1\rangle}(x)\},\end{aligned}$$

and $P_{i,K}(x)$ is a polynomial in x of degree at most $D + (K-1)d + d = D + Kd$; $i = 1,\dots,n$. Since, by the induction hypothesis, $\mathrm{ord}_{x=0}R_i^{\langle k\rangle}(x) \geq D + M - k + 1$ for $k = 0,\dots,K-1$, we get from (2.7):

$$(2.8)\qquad \mathrm{ord}_{x=0}(Q^{\langle K\rangle}(x)\cdot f_i(x) - P_{i,K}(x)) \geq D + M - K + 1$$

for $i = 1,\dots,n$.

If $M \geq K(d+1)$, then $\mathrm{ord}_{x=0}(Q^{\langle K\rangle}(x)f_i(x) - P_{i,K}(x)) \geq D + Kd + 1$; $i = 1,\dots,n$. Because the degree of $Q^{\langle k\rangle}(x)$ is always at most $D_0 + kd$, the degree $P_i^{\langle k\rangle}(x)$ is at most $D + kd$, and the degree of $P_{i,K}(x)$ is at most $D + Kd$, we see that $P_{i,K}(x)$ is the unique polynomial of the form $[Q^{\langle K\rangle}(x)\cdot f_i(x)]_{D+Kd}$, $i = 1,\dots,n$.

This identification and (2.8) prove Theorem 2.6 for $k = K$.

We can express the formula for iterations of Padé approximants from Theorem 2.6 in the matrix form as follows:

Corollary 2.7: Under the assumptions of Theorem 2.6 in the matrix notations, $A = (A_{i,j}(x))_{i,j=1}^n$ and $I = (\delta_{i,j})_{i,j=1}^n$ we have the following recurrences. For $M \geq k(d+1)$ the polynomials $Q^{\langle k\rangle}(x), P_i^{\langle k\rangle}(x)$ and functions $R_i^{\langle k\rangle}(x)$ $(i = 1,\dots,n)$ satisfy

$$(2.9)\quad \begin{aligned}&Q^{\langle k\rangle}(x) = D(x)^k\cdot\frac{1}{k!}\cdot(\frac{d}{dx})^k Q(x), R_i^{\langle k\rangle}(x) = Q^{\langle k\rangle}(x)\cdot f_i(x) - P_i^{\langle k\rangle}(x), \\ &(P_1^{\langle k\rangle}(x),\dots,P_n^{\langle k\rangle}(x))^t = D(x)^k\cdot\frac{1}{k!}\cdot(\frac{d}{dx}\cdot I - A)^k\cdot(P_1(x),\dots,P_n(x))^t, \\ &(R_1^{\langle k\rangle}(x),\dots,R_n^{\langle k\rangle}(x))^t = D(x)^k\cdot\frac{1}{k!}\cdot(\frac{d}{dx}\cdot I - A)^k\cdot(R_1(x),\dots,R_n(x))^t,\end{aligned}$$

and $P_i^{\langle k\rangle}(x) = [Q^{\langle k\rangle}(x)\cdot f_i(x)]_{D+kd}$, $i = 1,\dots,n$.

Here the differential operator $\frac{d}{dx} I - A$ acts on n-tuples of functions and $\cdot^t$ is the transposition operator.

<u>Proof</u>: Let us put $Q^{(k)}(x) = (\frac{d}{dx})^k Q(x)$ and let us define

$$
\begin{aligned}
&(P_1^{(k)}(x),\dots,P_n^{(k)}(x))^t \overset{\text{def}}{=} (\tfrac{d}{dx}\cdot I - A)^k \cdot (P_1(x),\dots,P_n(x))^t,\\
(2.10)\quad &\\
&(R_1^{(k)}(x),\dots,R_n^{(k)}(x))^t \overset{\text{def}}{=} (\tfrac{d}{dx}\cdot I - A)^k \cdot (R_1(x),\dots,R_n(x))^t,
\end{aligned}
$$

where $R_i(x) = Q(x)f_i(x) - P_i(x)$: $i = 1,\dots,n$. This means that, inductively, $P_i^{(0)}(x) \equiv P_i(x)$, $R_i^{(0)}(x) \equiv R_i(x)$ $(i = 1,\dots,n)$ and

$$
\begin{aligned}
&R_i^{(k+1)}(x) = \frac{d}{dx}(R_i^{(k)}(x)) - \Sigma_{j=1}^n A_{i,j}\cdot R_j^{(k)}(x),\\
(2.11)\quad &\\
&P_i^{(k+1)}(x) = \frac{d}{dx}(P_i^{(k)}(x)) - \Sigma_{j=1}^n A_{i,j}\cdot P_j^{(k)}(x)
\end{aligned}
$$

for $i = 1,\dots,n$ and $k \geq 0$. Since $f_1(x),\dots,f_n(x)$ satisfy the system (2.2), it follows from (2.11) that

$$(2.12)\qquad R_i^{(k)}(x) \equiv Q^{(k)}(x)\cdot f_i(x) - P_i^{(k)}(x)$$

for $i = 1,\dots,n$ and any $k \geq 0$. It follows from (2.10) that $\frac{1}{k!}\cdot D(x)^k\cdot P_i^{(k)}(x)$ is a polynomial in x of degree at most $D + kd$. Also, according to (2.3), $Q^{\langle k\rangle}(x) \equiv \frac{1}{k!} D(x)^k\cdot Q^{(k)}(x)$. The order of zero of $\frac{1}{k!} D(x)^k\cdot R_i^{(k)}(x)$ $(i = 1,\dots,n)$ at $x = 0$ is at least $D + M + 1 - k$, as follows from (2.10). Since $P_i^{\langle k\rangle}(x)$ is the only polynomial of degree $\leq D + kd$ such that $Q^{\langle k\rangle}(x)f_i(x) - P_i^{\langle k\rangle}(x)$ has a zero at $x = 0$ of order at least $D + M + 1 - k$ for $k(d+1) \leq M$, we get the identification

$$P_i^{\langle k\rangle}(x) = \frac{1}{k!}\cdot D(x)^k\cdot P_i^{(k)}(x),$$

$i = 1,\dots,n$. Thus $R_i^{\langle k\rangle}(x) = \frac{1}{k!}\cdot D(x)^k\cdot R_i^{(k)}(x)$: $i = 1,\dots,n$. Corollary 2.7 is proved.

<u>§3. Linear independence of Padé approximations of the second kind</u>.

In this chapter we prove linear independence of a system of Padé approximations of the second kind, constructed in Corollary 2.7 using the iteration of linear differential equations (2.2). This will allow us to construct in §4 a system of $n+1$ independent simultaneous rational approximations to numbers $f_1(x_0),\dots,f_n(x_0)$, whenever $x_0 \neq 0$ and x_0 is not a singularity of the system (2.2), i.e. $D(x_0) \neq 0$. For this we need a statement of linear independence of a system of Padé approximants constructed in Theorem 2.6:

Theorem 3.1: Let $f_1(x),\dots,f_n(x)$ satisfy a system of equations (2.2) and let functions $1, f_1(x),\dots,f_n(x)$ be linearly independent over $\mathbb{C}(x)$. Let $(Q(x);P_1(x),\dots,P_n(x))$ be Padé approximants to a system of functions $f_1(x),\dots,f_n(x)$ with parameters (D,D,M). Let for $k \geq 0$, the polynomials $(Q^{\langle k\rangle}(x);P_1^{\langle k\rangle}(x),\dots,P_n^{\langle k\rangle}(x))$ be Padé approximants defined as in Theorem 2.6 (see formulas (2.9)). Let

$$\Delta(x) = \det((Q^{\langle k\rangle}(x);P_1^{\langle k\rangle}(x),\dots,P_n^{\langle k\rangle}(x)) : k = 0,1,\dots,n).$$

Then for a sufficiently large M, $M \geq c_{13}$ the determinant $\Delta(x)$ is not identically zero. Here c_{13} is a constant depending only on the system of linear differential equations (2.2) and on the orders of zeroes of $f_1(x),\dots,f_n(x)$ at $x = 0$.

Proof: This result is dual to the well known results for approximating forms (i.e. for Padé-type approximants of the first kind). In our proof we use the arguments from [5, Chapter 11, Lemma 2]. Assume that $\Delta(x) \equiv 0$. Let $\ell \leq n$ be the integer such that the first ℓ columns $s_k = (Q^{\langle k\rangle}(x);\ P_1^{\langle k\rangle}(x),\dots,P_n^{\langle k\rangle}(x))^t$ $(0 \leq k \leq n)$ are linearly independent over $\mathbb{C}(x)$, but the $(\ell+1)$-st column is linearly dependent over $\mathbb{C}(x)$ on them. We denote by F the matrix formed by the first ℓ columns s_k $(0 \leq k \leq \ell-1)$ and by R and S we denote the matrices formed by the first ℓ rows and $n+1-\ell$ last rows of F, respectively. We assume, without loss of generality, that R is non-singular. Then, as is proved in [5, Ch. 11, Lemma 2], the degrees of the numerators and denominators of the rational function elements of the matrix SR^{-1} are bounded by c_{14}, where c_{14} depends only on the system (2.2). Denote by G the $\ell \times (n+1)$ matrix with ℓ rows $(f_i(x),0,\dots,-\delta_{i+1,j},\dots,0)$: $i = 1,\dots,\ell$; and let G_0 and G_1 denote the matrices formed by the first ℓ columns and $n+1-\ell$ last columns of G, respectively. Then for the $\ell \times \ell$ matrix $T \stackrel{\text{def}}{=} G.F$ we have $T_{i,j} = Q^{\langle j-1\rangle}(x)f_i(x) - P_i^{\langle j-1\rangle}(x)$, $i,j = 1,\dots,\ell$ and $T = G_0R + G_1S$, so that $TR^{-1} = G_0 + G_1SR^{-1}$. In view of Definitions 2.3, 2.5 and Theorem 2.6, all elements of T have orders of zero at $x = 0$ at least $D + M + 1 - \ell$. At the same time all (polynomial) elements of R have degrees in x at most $D + (\ell-1)d$. Hence $\det(TR^{-1})$ is a function in x with an order of zero at $x = 0$ of at least $\ell\{D+M+1-\ell-D-(\ell-1)d\} = \ell[M-(\ell-1)(d+1)\}$. On the other hand, $\det(G_0 + G_1SR^{-1}) \not\equiv 0$, since $1, f_1(x), \ldots, f_n(x)$ are linearly independent over $\mathbb{C}(x)$. Also the degrees of all rational functions elements of SR^{-1} are bounded by c_{15}. This implies that the order of zero at $x = 0$ of $\det(G_0 + G_1SR^{-1})$ is bounded by c_{16}, where c_{16} depends only on c_{15} and $f_1(x), \ldots, f_n(x)$. Hence, for $\ell\{M-(\ell-1)(d+1)\} > c_{16}$, or for a sufficiently large M, $\Delta(x)$ is not identically zero.

§4. Simultaneous rational approximations to values of G-functions.

In this chapter we use Theorem 3.1 on linear independence of Padé type approximations constructed in Theorem 2.6, to exhibit linearly independent simultaneous rational approximations to $1, f_1(x_0), \ldots, f_n(x_0)$ for $x_0 \neq 0$ and x_0 distinct from the singularities of the system (2.2).

Theorem 4.1: Under assumptions of Theorem 3.1 let, additionally, $x_0 \neq 0$ and $D(x_0) \neq 0$ (i.e. x_0 is distinct from the singularities of the equations (2.2)). Then there exist integers $k_0, \ldots, k_n$ such that $0 \leq k_0 < \ldots < k_n \leq D - nM + \frac{n(n+1)}{2}(d-1)$ and such that the following n+1 forms in the variables $y_0, \ldots, y_n$:

$$Q^{\langle k_j\rangle}(x_0)\cdot y_0 + \Sigma_{i=1}^{n} P_i^{\langle k_i\rangle}(x_0)y_i:\ i = 0, \ldots, n$$

are linearly independent.

Proof: From the definition of $\Delta(x)$ and the upper bounds of Theorem 2.6 on the degrees of $Q^{\langle k\rangle}(x)$, $P_i^{\langle k\rangle}(x)$, $i = 1, \ldots, n$, it follows that $\Delta(x)$ is a polynomial in x of degree of at most $D + nD + \frac{n(n+1)}{2}d$. On the other hand, $\Delta(x) = \det((Q^{\langle k\rangle}(x); P_1^{\langle k\rangle}(x)-Q^{\langle k\rangle}(x)\cdot f_1(x), \ldots, P_n^{\langle k\rangle}(x) - Q^{\langle k\rangle}(x)\cdot f_n(x)): k = 0, 1, \ldots, n)$. Since $\mathrm{ord}_{x=0}(P_i^{\langle k\rangle}(x)-Q^{\langle k\rangle}(x)\cdot f_i(x)) \geq M + D + 1 - k$, $i = 1, \ldots, n$, we have $\mathrm{ord}_{x=0}\Delta(x) \geq n(M+D)-\frac{n(n-1)}{2}$. Hence, if $\Delta(x) \not\equiv 0$, we have $\Delta(x) = x^a\cdot\Delta_0(x)$, where $\Delta_0(x)$ is a polynomial of degree of at most $D - nM + \frac{n(n+1)}{2}(d-1)-n$. Thus for any $x_0 \neq 0$, there exists a t, $0 \le t \le D - nM + \frac{n(n+1)}{2}(d-1)-n$, such that $\Delta^{(t)}(x_0) \neq 0$ but $\Delta^{(s)}(x_0) = 0$: $0 \le s < t$.

Let us introduce the following linear forms in n + 1 variables $z_0,\dots,z_n$:

$$\ell^{\langle k\rangle}(x;\bar{z}) \overset{\text{def}}{=} Q^{\langle k\rangle}(x)\cdot z_0 + \Sigma_{i=1}^{n} P_i^{\langle k\rangle}(x)\cdot z_i. \tag{4.1}$$

From the definition of $\Delta(x)$ we obtain

$$\Delta(x) z_i = \Sigma_{j=0}^{n} \ell^{\langle j\rangle}(\bar{z})\cdot \Delta_{ij}(x), \quad i = 0,1,\dots,n, \tag{4.2}$$

where $\Delta_{ij}(x)$ is the algebraic complement in $\Delta(x)$. We consider the system of linear differential equations conjugate to (2.2):

$$\frac{dz_0}{dx} = 0,$$

(4.3)

$$\frac{dz_i}{dx} = -\Sigma_{j=1}^{n} A_{ji}(x) z_j, \quad i = 1,\dots,n$$

with initial conditions $z_i(x_0) = y_i$, $i = 0,1,\dots,n$. Since x_0 is not a singularity of (2.2), such a solution $z = z_i(x)$, $i = 0,\dots,n$ exists. We have $\ell^{\langle k\rangle}(x;\bar{z}) = \frac{D(x)^k}{k!}\cdot m^{\langle k\rangle}(m;\bar{z})$ and

$$m^{\langle k\rangle}(x;\bar{z}) \overset{\text{def}}{=} Q^{(k)}(x)\cdot z_0 + \Sigma_{i=1}^{n} P_{i,k}^{1}(x) z_i.$$

According to (2.9), $Q^{(k)}(x) = (\frac{d}{dx})^k Q(x), (P_{1,k}^1(x),\dots,P_{n,k}^1(x))^t = (\frac{d}{dx}\cdot I - A)^k \cdot (P_1(x),\dots,P_n(x))^t$. Substituting $z_i = z_i(x)$ from (4.3) we see after differentiation that $\frac{d}{dx} m^{\langle k\rangle}(x;\bar{z}(x)) = m^{\langle k+1\rangle}(x;\bar{z}(x))$.

Consequently, $(\frac{d}{dx})^{k'} m^{\langle k\rangle}(x;\bar{z}(x)) = m^{\langle k+k'\rangle}(x;\bar{z}(x))$. Differentiating (4.2) t times with $z_i = z_i(x)$, $i = 0,\dots,n$ and substituting $x = x_0$, we obtain from $D(x_0) \neq 0$,

$$y_i = \Sigma_{s=0}^{n+t} \ell^{\langle s\rangle}(x_0;\bar{y})\cdot \Delta_i^s, \quad i = 0,1,\dots,n$$

with $\Delta_i^s = \Delta_i^s(x_0)$--rational in x_0, for arbitrary variables $\bar{y} = (y_0,\dots,y_n)$. Hence, among the n + t forms $\ell^{\langle s\rangle}(x_0;\bar{y}) = Q^{\langle s\rangle}(x_0)y_0 + \Sigma_{i=1}^{n} P_i^{\langle s\rangle}(x_0)\cdot y_i$ in variables $y_0,y_1,\dots,y_n$: $s = 0,\dots,n+t; n+t \le D - nM + \frac{n(n+1)}{2}(d-1)$, there exist $n+1$ linearly independent forms. Theorem 4.1 is proved.

Theorem 4.1 is applied in §5 to the proof of one of our key results, Theorem I on the linear independence of values of G-function formulated in §1.

§5. Proof of Theorem I.

To prove Theorem I we construct the system $(Q(x);P_1(x),\dots,P_n(x))$ of Padé approximants of the second kind to functions $f_1(x),\dots,f_n(x)$ with parameters $(D,D,[(\frac{1}{n}-\delta)D])$ for δ, $1/n > \delta > 0$ and a sufficiently large D. This construction is achieved using Thue-Siegel lemma. Then we apply results of Theorem 2.6 and statement on linear independence from Theorem 4.1.

Following the formulation of Theorem I we consider G-functions satisfying differential equations (2.2) over $\mathbb{Q}(x)$, i.e. $A_{i,j}(x) \in \mathbb{Q}(x)$ and $D(x) \in \mathbb{Z}[x]$. Also all coefficients of Taylor expansions of G-functions $f_1(x),\dots,f_n(x)$, satisfying equations (2.2) at $x = 0$ are assumed to be rational integers.

<u>Lemma 5.1</u>: Let $f_1(x),\dots,f_n(x)$ be a system of G-functions such that $f_i(x) = \Sigma_{m=0}^{\infty} a_{m,i}x^m, a_{m,i} \in \mathbb{Q}$ and for the common denominator Δ_m of $\{a_{0,i},\dots,a_{m,i}: i = 1,\dots,n\}$ we have $|\Delta_m \cdot a_{k,i}| \le C_0^m (k = 0,1,\dots,m)$ for some $C_0 > 1$ and all $m = 0,1,\dots$ and $i = 1,\dots,n$. Then for δ, $\frac{1}{n} > \delta > 0$, and an arbitrary positive integer D, there exists a system $(Q(x);P_1(x),\dots,P_n(x))$ of Padé approximants of the second kind to $f_1(x),\dots,f_n(x)$ with parameters $(D,D,[(\frac{1}{n}-\delta)D])$ such that $Q(x) \in \mathbb{Z}[x]$ and the height of $Q(x)$ is at most $D^{(1-\delta n)/\delta n} \cdot C_0^{(n+1-\delta n)(1-\delta n)D/\delta n^2}$.

<u>Proof of Lemma 5.1:</u> In the proof we use Dirichlet's box principle in the form of Thue-Siegel's lemma [5]:

<u>Lemma 5.2</u>: Let M, N be rational integers with $N > M > 0$ and let $u_{ij} (1 \le i \le M,\ 1 \le j \le N)$ denote rational integers with absolute values at most U (≥ 1). Then there exist rational integers $x_1,\dots,x_N$ not all 0, with absolute values at most $(NU)^{M/(N-M)}$, such that $\Sigma_{j=1}^{N} u_{ij}x_j = 0$ $(1 \le i \le M)$.

Following Definition 2.3, for a given D, and $D+1$ undetermined coefficients $q_m (0 \leq m \leq D)$ of $Q(x) = \Sigma_{m=0}^{D} q_m x^m$, we put $P_i(x) \overset{\text{def}}{=} [Q \cdot f_i]_D$, $i = 1,\ldots,n$. Then the condition that $(Q(x);P_1(x),\ldots,P_n(x))$ are Padé approximants of the second kind of $f_1(x),\ldots,f_n(x)$ with parameters $(D,D,[(1/n-\delta)D])$ is equivalent to the following: $\mathrm{ord}_{x=0}(Q(x)f_i(x) - P_i(x)) \geq D + [(1/n - \delta)D] + 1$ for all $i = 1,\ldots,n$. The last condition, in view of the definition of $P_i(x)$, is equivalent to the system of linear equations:

$$(5.1) \qquad \Sigma_{k=0}^{D} q_k \cdot a_{m-k,i} = 0: \; m = D + 1,\ldots,D + [(\tfrac{1}{n} - \delta)D],$$

$i = 1,\ldots,n$. This is a system of at most $n \cdot [(1/n-\delta)D]$ equations in $D + 1$ unknowns q_m $(0 \leq m \leq D)$. The common denominator of all coefficients of equations (5.1) divides Δ_{D+M}, $M = [(1/n - \delta)D]$. According to the assumptions of Lemma 5.1, $\Delta_{D+M} \cdot a_{m,i}$ are rational integers $(m \leq D + M)$ $i = 1,\ldots,n$ of absolute value at most C_0^{D+M}. Hence, according to Lemma 5.2, there exists a nonzero polynomial $Q(x) \in \mathbb{Z}[x]$ of height at most $D^{(1-\delta n)/\delta n} \cdot C_0^{(n+1-\delta n)(1-\delta n)D/\delta n^2}$, satisfying all the conditions of Lemma 5.1.

From the discussion in Lemma 2.4 and Theorem 2.6, we obtain

Lemma 5.3: Let functions $f_1(x),\ldots,f_n(x)$ satisfy all assumptions of Lemma 5.1 and let $f_i(x) = \Sigma_{m=0}^{\infty} a_{m,i} x^m$, $i = 1,\ldots,n$ be a solution of a system (2.2) of differential equations with $|a_{m,i}| \leq C_1^m$ $(m = 0,1,\ldots,;\ i = 1,\ldots,n)$ for some $C_1 > 1$. Let $(Q(x);P_1(x),\ldots,P_n(x))$ be the system of Padé approximants of the second kind with parameters $(D,D,[(1/n - \delta)D])$ constructed in Lemma 5.1. If for $k \geq 0$, $kd < [(1/n - \delta)D]$, then $Q^{\langle k \rangle}(x) \in \mathbb{Z}[x]$, and the common denominator $\mathfrak{D}_k$ of the coefficients of polynomials $P_i^{\langle k \rangle}(x) \in \mathbb{Q}[x]$ $(i = 1,\ldots,n)$ is bounded by C_0^{D+kd}. Also, for $i = 1,\ldots,n$ and $|x| < C_1^{-1}$ we have

$$|Q^{\langle k \rangle}(x) f_i(x) - P_i^{\langle k \rangle}(x)| < H(D(x))^k H(Q(x)) \cdot 2^{D+kd} \cdot \frac{C_1}{C_1 - 1}$$
$$\times\ |C_1 x|^{D+M+1-k} / (1 - |xC_1|),$$

where $M = [(1/n-\delta)D]$ and $H(D(x))$, $H(Q(x))$ are the heights of the polynomial $D(x)$, $Q(x)$, respectively.

Proof: According to Theorem 2.6 we have

$Q^{\langle k\rangle}(x) = D(x)^k \cdot \frac{1}{k!}\cdot(\frac{d}{dx})^k Q(x)$, $P_i^{\langle k\rangle}(x) = [Q^{\langle k\rangle}(x)\cdot f_i(x)]_{(D+kd)}$. Since $D(x) \in \mathbb{Z}[x]$, $Q^{\langle k\rangle}(x) \in \mathbb{Z}[x]$ and $\deg(Q^{\langle k\rangle}(x)) \le D + kd$, $H(Q^{\langle k\rangle}(x)) \le H(D(x))^k \cdot H(Q(x))\cdot 2^{D+dk}$. Thus, if $Q^{\langle k\rangle}(x) = \Sigma_{j=0}^{D+kd} q_j^{\langle k\rangle}\cdot x^j$, then $P_i^{\langle k\rangle}(x) = \Sigma_{j=0}^{D+kd} x^j \Sigma_{m=0}^{\min(D+kd,j)} q_m^{\langle k\rangle}\cdot a_{j-m,i}$. This establishes the bound for the denominators $\mathfrak{D}_k$. Similarly, according to Lemmas 2.4 and Theorem 2.6 $Q^{\langle k\rangle}(x) f_i(x) - P_i^{\langle k\rangle}(x) = \Sigma_{m=D+M+1-k}^{\infty} x^m \cdot \Sigma_{j=0}^{D+kd} q_j^{\langle k\rangle} a_{m-j,i}$. Upper bounds on $H(Q^{\langle k\rangle}(x))$ and on $a_{m,i}$ prove Lemma 5.3.

We use Lemmas 5.1 and 5.3 to prove Theorem I. Under the assumptions of Theorem I, put

$$\ell = H_0 + H_1 f_1(r) + \ldots + H_n f_n(r).$$

For a fixed δ, $1/n(n+1) > \delta > 0$, we consider the Padé approximants $(Q(x);P_1(x),\ldots,P_n(x))$ to $f_1(x),\ldots,f_n(x)$ with parameters $(D,D,[(1/n-\delta)D])$ constructed in Lemma 5.1. Applying Theorem 4.1 we find an integer k_j, $0 \le k_j \le D - nM + n(n+1)(d-1)/2$ and $M = [(1/n-\delta)D]$, such that $H_0\cdot Q^{\langle k_j\rangle}(r) + \Sigma_{i=1}^n H_i P_i^{\langle k_j\rangle}(r) \ne 0$. Since H_i are rational integers and $r = a/b$, we obtain a non-zero rational integer $I = \mathfrak{D}_{k_j}\cdot b^{D+k_j d}\{H_0\cdot Q^{\langle k_j\rangle}(r) + \Sigma_{i=1}^n H_i P_i^{\langle k_j\rangle}(r)\}$. Thus $|I| \ge 1$. On the other hand, $I = \mathfrak{D}_{k_j}\cdot b^{D+k_j d}\cdot \{\ell\cdot Q^{\langle k_j\rangle}(r) - \Sigma_{i=1}^n H_i(Q^{\langle k_j\rangle}(r) f_i(r) - P_i^{\langle k_j\rangle}(r))\}$.

Now let $\delta n^2 = \epsilon$ for $\epsilon < 1/(d+2)$. Then for any $k \le D - nM + \frac{n(n+1)}{2}(d-1)$ and $D \ge c_{17}(n,d,\delta)$, we can combine estimates of Lemma 5.3 with an upper bound of $H(Q(x))$ from Lemma 5.1 to obtain

$$\begin{aligned} &|Q^{\langle k\rangle}(r) f_i(r) - P_i^{\langle k\rangle}(r)| < H(D(x))^{\frac{D\epsilon}{n}}\cdot 2^{(1+\frac{d\epsilon}{n})D} \\ &\times c_0^{\frac{(n+1)D}{\epsilon}}\cdot c_1^{(1+\frac{1}{n}-\frac{2\epsilon}{n})D}\cdot |r|^{(1+\frac{1}{n}-\frac{2\epsilon}{n})D}, \end{aligned} \tag{5.2}$$

whenever $|r| \le c_{18}$, $c_{18} = c_{18}(n,d,\epsilon)$. Similarly, Lemma 5.3 implies under the same assumptions on k:

$$|Q^{\langle k\rangle}(r)| \le c_0^{\frac{(n+1)}{\epsilon}D}\cdot H(D(x))^{\frac{D\epsilon}{n}}\cdot 2^{(1+\frac{d\epsilon}{n})D}. \tag{5.3}$$

We choose now the weight D as the smallest integer

$D \geq c_{17}(n,d,\delta)$ such that

$$(5.4) \qquad H\cdot n\cdot \mathfrak{D}_{k_j}\cdot |b|^{D+k_j d}\cdot \max\{|Q^{\langle k_j\rangle}(r)f_i(r)-P_i^{\langle k_j\rangle}(r)|:$$
$$i = 1,\dots,n\} < \frac{1}{2}.$$

Then $|I| \geq 1$ implies $|\ell| \geq |b|^{-D-k_j d}\cdot \mathfrak{D}_{k_j}^{-1}\cdot |Q^{\langle k_j\rangle}(r)|^{-1}/2$. We can represent the bound (5.4) in a form

$$(5.5) \qquad H\cdot c_{19}^{D}|a|^{(1+\frac{1}{n}-2\epsilon_0)D}\cdot |b|^{(2\epsilon_0+d\epsilon_0-\frac{1}{n})D} < 1,$$

where $c_{19} = c_{19}(n,D(x),\epsilon)$ and $\epsilon_0 = \frac{\epsilon}{n} < \frac{1}{n(d+2)}$. Then for any H, an integer $D \geq c_{17}$ exists whenever $|a|^{(1+1/n-2\epsilon_0)}\cdot |b|^{(2\epsilon_0+d\epsilon_0-1/n)} < c_{19}^{-1}$. The lower bound

$$|\ell| \geq |b|^{-(1+\epsilon_0 d)D}\, c_{20}^{D}$$

for $c_{20} = c_{20}(n,D(x),\epsilon) > 0$, together with the definition of D in (5.4), implies the assertions of Theorem I.

§6. The duality between Padé-type approximants of the first and second kind.

In this chapter we extend the duality principle (see §2) to exhibit the relationship between Padé-type approximants of the second kind and Padé-type approximants of the first kind to the system of functions 1, $f_1(x),\dots,f_n(x)$. We make our exposition a local one, associated with a nonsingular point $x_0 \neq 0$, and we construct $n + 1$ linearly independent approximating forms to $1,f_1(x_0),\dots,f_n(x_0)$. For this we start with a system $(Q(x);P_1(x),\dots,P_n(x))$ of Padé approximants with parameters (D,D,M) and with $x_0 \neq 0$ satisfying all assumptions of Theorem 4.1. We let $P_0^{\langle k\rangle}(x_0) \overset{\mathrm{def}}{=} Q^{\langle k\rangle}(x_0)$, and define rational functions $L_{p,j}$ $(p,j=0,\dots,n)$ in x_0 as solutions of the system of linear equations

$$(6.1) \qquad \Sigma_{p=0}^{n} L_{p,j}\cdot P_p^{\langle k_i\rangle}(x_0) = \delta_{i,j},\ i,j = 0,\dots,n,$$

where $0 \leq k_0 < \dots < k_n \leq \mathcal{K} \leq D - nM + n(n+1)(d-1)/2$ as in Theorem 4.1.

If we put $\nabla(x_0) \overset{\text{def}}{=} \det(P_j^{\langle k_i\rangle}(x_0): i,j = 0,1,\ldots,n)$, then $\nabla(x_0) \neq 0$ by Theorem 4.1 and $L_{p,j} = R_{p,j}(x_0)/\nabla(x_0)$, where $R_{p,j}(x_0)$ is a determinant from $P_q^{\langle k_i\rangle}(x_0)$, $q \neq p$ constructed by Cramer's rule from equations (6.1). Let $\beta = \min\{\mathrm{ord}_{x_0=0}(P_0^{\langle k_i\rangle}(x_0)),\ i = 0,\ldots,n\}$. Since $\beta \leq D$, we deduce from Theorem 2.6 that, whenever $M \geq \mathcal{K}(d+1)$, we have $\mathrm{ord}_{x_0=0}(P_j^{\langle k_i\rangle}(x_0)) \geq \beta$ for all $j = 1,\ldots,n$ and $i = 0,1,\ldots,n$. Thus we obtain

$$\mathrm{ord}_{x_0=0}(\nabla(x_0)) \geq \beta + n(M+D+1) - \Sigma_{i=1}^n k_i;$$

(6.2)

$$\mathrm{ord}_{x_0=0}(R_{p,j}(x_0)) \geq \beta + (n-1)(M+D+1) - \Sigma_{i=2}^n k_i,$$

$(p,j = 0,\ldots,n)$.

Let us denote, for simplicity, $f_0(x_0) \equiv 1$. If we put $w_j = \Sigma_{p=0}^n L_{p,j}\cdot f_p(x_0)$: $j = 0,\ldots,n$, then we have an identity

$$P_0^{\langle k_i\rangle}(x_0)\cdot w_j = \delta_{i,j} + \Sigma_{p=1}^n L_{p,j}(P_0^{\langle k_i\rangle}(x_0)f_p(x_0)$$

(6.3)

$$- P_p^{\langle k_i\rangle}(x_0)): \quad i,j = 0,1,\ldots,n.$$

Let us define $\xi = (n-1)(M+D)-(n-1)\mathcal{K}$ and let us put $v_j \overset{\text{def}}{=} w_j\cdot\nabla(x_0)\cdot x_0^{-\xi}$ for $j = 0,1,\ldots,n$. According to the estimate of $\mathrm{ord}_{x_0=0}(R_{p,j}(x_0))$ $(p,j = 0,\ldots,n)$ in (6.1) we have $v_j = \Sigma_{p=0}^n M_{p,j}(x_0)\cdot f_p(x_0)$ $(j = 0,\ldots,n)$, where $M_{p,j}(x_0)$ $(p,j = 0,\ldots,n)$ are polynomials in x_0. We can use (6.3) to estimate from below the order of zero of v_j at $x_0 = 0$: $\mathrm{ord}_{x_0=0}(v_j) \geq \min\{\mathrm{ord}_{x_0=0}(\nabla(x_0)),\ \mathrm{ord}_{x_0=0}(R_{p,j}(x_0)\cdot(P_0^{\langle k_i\rangle}(x_0)f_p(x_0)-P_p^{\langle k_i\rangle}(x_0))): i = 0,\ldots,n;\ p = 1,\ldots,n\} - \beta - \xi$. Using the estimates above in (6.1), we obtain $\mathrm{ord}_{x_0=0}(v_j) \geq M + D - \mathcal{K}$, $j = 0,\ldots,n$. Thus we deduce the following

<u>Theorem 6.1</u>: Let $f_1(x),\ldots,f_n(x)$ be solutions of (2.2), regular at $x = 0$, linearly independent with 1 over $\mathbb{C}(x)$. Let $(Q(x);P_1(x),\ldots,P_n(x))$ be a system of Padé approximants of $f_1(x),\ldots,f_n(x)$ with parameters (D,D,M) and let $x_0 \neq 0$ and $D(x_0) \neq 0$. Then for a sufficiently large

M, $M \geq \mathcal{K}(d+1)$, there are n+1 linearly independent forms $v_j = \Sigma_{p=0}^{n} M_{p,j}(x_0) f_p(x_0)$: $j = 0,\ldots,n$ in $f_0(x_0) \equiv 1$, $f_1(x_0),\ldots,f_n(x_0)$, with polynomial coefficients $M_{p,j}(x_0)$ $(p,j = 0,\ldots,n)$ such that

$$\deg_{x_0}(M_{p,j}(x_0)) \leq D - (n-1)M + \mathcal{K}(nd+n-1),$$

$$\mathrm{ord}_{x_0=0}(v_j) \geq D + M - \mathcal{K},$$

$p,j = 0,\ldots,n$. Here $\mathcal{K} \leq D - nM + n(n+1)(d-1)/2$. Moreover, if $Q^{\langle k_i\rangle}(x) \in \mathbb{Z}[x]$, $P_j^{\langle k_i\rangle}(x) \in \mathbb{Q}[x]$ and the common denominator of coefficients of $P_j^{\langle k_i\rangle}(x)$ is $\mathfrak{D}$$(j = 1,\ldots,n;\ i = 0,\ldots,n)$, while the heights of polynomials $Q^{\langle k_i\rangle}(x)$, $P_j^{\langle k_i\rangle}(x)$ are bounded by $H(j=1,\ldots,n; i=0,\ldots,n)$, then we can assume that $M_{p,j}(x_0) \in \mathbb{Z}[x_0]$ and $H(M_{p,j}(x_0)) \leq \mathfrak{D}^n \cdot H^n(p,j=0,\ldots,n)$.

The first two inequalities of Theorem 6.1 follow directly from the preceeding discussion. The linear independence of forms v_j: $j = 0,\ldots,n$ follows from $\nabla(x_0) \neq 0$ and equations (6.1) using Theorem 4.1, proved in §4. The last part of Theorem 6.1 is also obvious if one replaces $M_{p,j}(x_0)$ by $\mathfrak{D}^n \cdot M_{p,j}(x_0)$ in the definition of v_j above $(p,j = 0,\ldots,n)$.

§7. Proof of Theorem II.

Following Siegel's method of approximating forms [1], [2] we can apply Theorem 6.1 on the existence of Padé approximants of the first kind to the proof Theorem II from §1 on the absence of algebraic relations between values of G-functions at algebraic points. In fact, Theorem II is not the best result of this kind that we can prove. We present a complete proof of Theorem II to show how Padé approximants can give a proof of G-function theorem that Siegel [1] had envisioned.

Proof of Theorem II: In this theorem we use a separate numeration of constants, starting from c_{17} (do not confuse that with constants from the proof of Theorem I). Let N be a sufficiently large positive integer. We consider a new system of G-functions $\{f_1^{i_1}(x)\ldots f_n^{i_n}(x): 0 < i_1 +\ldots+ i_n \leq N\}$. Let us denote functions in this system by $F_1(x),\ldots,F_m(x)$, $m = \binom{N+n}{n} - 1$. Then $F_1(x),\ldots,F_m(x)$ are linearly independent over $\mathbb{Q}(x)$ and satisfy a system of linear differential equations of type (2.2), but with n replaced by m, and with the same common denominator $D(x)$ of its rational function coefficients. If C_0, C_1 denote constants introduced in Lemmas 5.1 and 5.3, respectively, for a system $f_1(x),\dots,f_n(x)$, then the corresponding constants for a new system $F_1(x),\dots,F_m(x)$ can be taken as C_0^N and C_1. This means that for $F_j(x) = \Sigma_{s=0}^{\infty} A_{s,j}x^s$ and the common denominator Δ_r of $A_{s,j}$: $s = 0,\dots,r$ and $j = 1,\dots,m$ we have: $|\Delta_r A_{s,j}| < C_0^{Nr}$ $(s=0,\dots,r)$ and we have $|A_{s,j}| < C_1^s$ for $s > s_0(N)$ and $j = 1,\dots,m$. Hence, for $1/m > \delta > 0$, we obtain a system $(Q(x);P_1(x),\dots,P_m(x))$ of Padé approximants of the second kind of functions $F_1(x),\dots,F_m(x)$ with parameters $(D,D,[(1/m-\delta)D]$, satisfying all the conditions of Lemmas 5.1 and 5.3 but with n replaced by m, C_0 replaced by C_0^N and f_i by F_j $(j = 1,\dots,m)$, with $D \geq D_2(F_1,\dots,F_m)$. We apply Theorem 6.1 to this system $(Q(x);\ P_1(x),\dots,P_m(x))$ of Padé approximants of $F_1(x),\dots,F_m(x)$, by choosing $x_0 = \xi$, for small $|\xi|$. This way we get a system of $m + 1$ linearly independent forms $u_j = \Sigma_{p=0}^{m} S_{p,j}(\xi)F_p(\xi)$: $j = 0,\dots,m$ in $F_0(\xi) = 1$, $F_1(\xi),\dots,F_m(\xi)$. Here $S_{p,j}(x) \in \mathbb{Z}[x]$, the degrees of polynomials $S_{p,j}(x)$ are bounded by $D\cdot\{1/m+\delta(m^2(d+1)-1)\} + c_{17}$, and heights by $\exp\{c_{18}DN/\delta\}$, when $\delta m^2(d+1) < 1/m$, $c_{17} = c_{17}(f_1,\dots,f_n,N,\delta) > 0$, $c_{18} = c_{18}(C_0,n) > 0$, $D \geq D_3\ (F_1,\dots,F_m)$ and $p,j = 0,1,\dots,m$. According to Theorem 6.1, the functions $\Sigma_{p=0}^{m} S_{p,j}(x)F_p(x)$ have zeroes at $x = 0$ of orders at least $D + (1/m-\delta)D - \delta mD - c_{19}$, $c_{19} = c_{19}(f_1,\dots,f_n,N,\delta) > 0$ for $j = 0,\dots,m$. Hence we obtain the following upper bounds on $|u_j|$:

$$(7.1)\qquad |u_j| < \exp\{c_{20}DN/\delta\}\cdot|\xi|^{D(1+1/m-\delta(m+1))-c_{19}},$$

$j = 0,\dots,m$ and $c_{20} = c_{20}(C_0,C_1,n) > 0$. Let us assume that there is a nontrivial algebraic relation $P \overset{\text{def}}{=} P(\xi,f_1(\xi),\dots,f_n(\xi)) = 0$ of degree $t' \leq t$ in $f_1(\xi),\dots,f_n(\xi)$ with coefficients from $\mathbb{Z}[\xi]$, not all zeroes, $P(x_0,x_1,\dots,x_n) \in \mathbb{Z}[x_0,\dots,x_n]$. Multiplying P by monomials $f_1^{i_1}(\xi)\ \dots\ f_n^{i_n}(\xi)$: $i_1 +\dots+ i_n \leq N-t'$, we obtain $m' \overset{\text{def}}{=} \binom{N-t'+n}{n}$ nontrivial linearly independent forms $L_\alpha(\xi)$ in $1,F_1(\xi),\dots,F_n(\xi)$ $(\alpha \leq m')$ with coefficients from $\mathbb{Z}[\xi]$. Hence, there are $m+1-m'$ linear forms u_{j_β}, among u_j, linearly independent from $L_\alpha(\xi)$ $(\beta = 1,\dots,m+1-m',\ \alpha = 1,\dots,m')$. For a determinant $R(\xi)$ formed from the coefficients of these $m + 1$

linearly independent forms in $1, F_1(\xi), \ldots, F_n(\xi)$, $R(\xi) \neq 0$ and $R(\xi)$ is a polynomial in ξ with rational integer coefficients of height at most $\exp\{c_{21}DN^n/\delta\}$, and of degree in ξ at most $c_{22}N^{n-1}D\{1/m + \delta(m^2(d+1)-1)\} + c_{23}$, where $c_{21} = c_{21}(C_0,C_1,n,t') > 0$, $c_{22} = c_{22}(n,t') > 0$, $c_{23} = c_{23}(f_1,\ldots,f_n,N,\xi) > 0$ and $N \geq N_0(n,t)$. Bounds (7.1) implies the following upper bound:

$$(7.2) \qquad |R(\xi)| < \exp\{c_{24}DN^n/\delta\}\cdot|\xi|^{D(1+1/m-\delta(m+2))},$$

where $c_{24} = c_{24}(C_0,C_1,n,t) > 0$, $N \geq N_1(n,t)$, $D \geq D_4(f_1,\ldots,f_n,N)$. On the other hand, we can apply the Liouville theorem [5] to bound from below $|R(\xi)|$ in terms of $H(\xi)$-the height of ξ. By choosing a sufficiently large N and putting $\delta = c_{25}N^{-3n}$, $c_{25} = c_{25}(C_0,C_1,n,d,t) > 0$, we obtain from Liouville's theorem: $\log|R(\xi)| > -c_{26}(N^{4n} + \log H(\xi)/N)$, for $c_{26} = c_{26}(C_0,C_1,n,d,t) > 0$ and $N \geq N_2(n,d,t)$. The contradiction between the two bounds proves Theorem II.

Conditions on $|\xi|$ from Theorem II can be considerably relaxed, without strengthening the assumptions of Theorem II. For example, using our results of §8, that G-functions satisfy (G,C)-property, we can replace the exponent $4n/(4n+1)$ in the bound on $|\xi|$ in Theorem II can be substituted by $n/(n+1)$. The case of algebraic relations between $1, f_1(x), \ldots, f_n(x)$ can be treated similarly to Theorem II. Also one gets lower bounds on polynomials in $f_1(\xi), \ldots, f_n(\xi)$ with coefficients from $\mathbb{Q}(\xi)$, similar to bounds of linear forms in Theorem I.

§8. The proof of (G,C)-property of an arbitrary G-function.

The global (G,C)-function assumption of §1 describes p-adic (for almost all p) properties of linear differential equations satisfied by G-functions. It is tempting to assume that local G-function condition for a single solution of a linear differential equation implies the (G,C)-property for the linear differential equation. Such a conjecture was suggested by the authors in [10]. It turns out that this local-global conjecture is true. Namely, if a linear differential equation of order n over $\bar{\mathbb{Q}}(x)$ has a G-function solution $f(x)$, which does not satisfy a linear differential equation of a smaller order, then the linear differential equation satisfy (G,C)-property. See Theorem

III of §1. (Note that the condition on $f(x)$ is essential, because one can always construct reducible linear differential equations of order n without the (G,C)-property and having a G-function solution $f(x)$, arising from a linear differential equation of order less than n.) The global result on G- and (G,C)-functions, that we prove below, is a counterpart of the Grothendieck conjecture that also relates local and global properties of linear differential equations, see [4]. Our methods of proof are based on Padé-type approximations of the second kind and on duality principles from §6.

We start with a more detailed formulation of (G,C)-property of §1. First of all, we have a system of first order matrix linear differential equations

$$L\bar{f}^t = 0 \tag{8.1}$$

for $L \overset{\text{def}}{=} \frac{d}{dx}\cdot I - A$ with $I = (\delta_{i,j})_{i,j=1}^n$, $A = (A_{i,j}(x))_{i,j=1}^n$ and $\bar{f} = (f_1(x),\ldots,f_n(x))$. In a coordinate form (8.1) is

$$\frac{d}{dx}\,\bar{f}^t = A\bar{f}^t \quad \text{or}$$

(8.2)

$$\frac{d}{dx}\,f_i(x) = \Sigma_{j=1}^n A_{i,j}(x) f_j(x): \quad i = 1,\ldots,n.$$

Here $A_{ij}(x)$ $(i,j = 1,\ldots,n)$ are rational functions, that we assume to belong to $\mathbb{Q}(x)$. We also denote by $D(x)$ the common denominator of all rational functions $A_{ij}(x)$ $(i,j= 1,\ldots,n)$, i.e. $D(x) \in \mathbb{Z}[x]$ and $D(x)\cdot A_{ij}(x)$ $(i,j = 1,\ldots,n)$ are polynomials with rational integer coefficients.

Differentiating linear differential equations (8.1) we obtain for an arbitrary $m \geq 0$ the following relation in the differential ring $Q(x)[\frac{d}{dx}]$:

$$\left(\frac{d}{dx}\right)^m\cdot I \equiv A_m (\mathrm{mod}\ \mathbb{Q}(x)[\tfrac{d}{dx}]\cdot L). \tag{8.3}$$

Here $A_m = (A_{i,j;m}(x))_{i,j=1}^n$ is an element of $M_n(\mathbb{Q}(x))$. Here $A_0 = I$ and $A_1 = A$. The recurrent formulas, connecting the matrices A_m $(m \geq 0)$ are the following:

$$A_{m+1} = A_m\cdot A + \frac{d}{dx}\,A_m. \tag{8.4}$$

It follows from (8.4) that

$$A_{ij,m}(x)\cdot D(x)^m \in \mathbb{Z}[x]:\ i,j = 1,\dots,m.$$

The relation (8.3) means that for an arbitrary solution $\bar{f} = (f_1(x),\dots,f_n(x))$ of (8.2) we have the following formula

$$(8.5)\qquad (\frac{d}{dx})^m f_i(x) = \Sigma_{j=1}^n A_{ij,m}(x)\cdot f_j(x),$$

$i = 1,\dots,n$ and $m \geq 0$.

The (G,C)-assumption of §1 or, equivalently, the (G,C)-property of a linear differential equation (8.1) means the following:

(G,C): There exists a constant $C \geq 1$ such that the common denominator D_N of the coefficients of the polynomials

$$\frac{1}{m!}\cdot D(x)^m\cdot A_{ij,m}(x):\quad i,j = 1,\dots,n$$

and $m = 1,\dots,N$ is bounded by C^N for any $N \geq 1$.

Let us start with a given solution $\bar{f}(x) = (f_1(x),\dots,f_n(x))$ of (8.1) such that $f_i(x)$ are G-functions and

$$(8.6)\qquad f_i(x) = \Sigma_{m=0}^\infty a_{m,i}x^m,\quad a_{m,i} \in \mathbb{Q}\quad (m = 0,1,\dots)$$

for $i = 1,\dots,n$. We denote by Δ_m the common denominator of $\{a_{0,i},\dots,a_{m,i}: i = 1,\dots,n\}$. According to the definition of G-functions, there exist two constants $C_0 > 1$ and $C_1 > 1$ such that

$$|\Delta_m| \leq C_0^m,\ |\Delta_m a_{k,i}| \leq C_0^m \qquad (k = 0,1,\dots,m)$$

(8.7) and

$$|a_{m,i}| \leq C_1^m$$

for any $m \geq 0$ and $i = 1,\dots,n$. (We note that $\Delta_m \in \mathbb{Z}$ and $\Delta_m\cdot a_{k,i} \in \mathbb{Z}$: $k = 0,\dots,m$ and $i = 1,\dots,n$.) We put, as above, $d = \max\{\deg(D(x)) - 1, \deg(D(x)A_{ij}(x)): i,j = 1,\dots,n\}$. We use results of §4 and Padé approximants of the second kind constructed there. Thus, let D be a sufficiently large integer and let $1/n > \delta > 0$. Summarizing Theorem 4.1 and 4.3 we obtain the following

<u>Corollary 8.1</u>: Under the assumptions above, there exists a system

$(Q(x);P_1(x),\dots,P_n(x))$ of Padé approximants of the second kind to $f_1(x),\dots,f_n(x)$ with parameters $(D,D,[(1/n-\delta)D])$ such that the following conditions are satisfied. Let $k \geq 0$ and $kd < [(1/n-\delta)D]$. Then $Q^{\langle k\rangle}(x) \in \mathbb{Z}[x]$ and the common denominator $\mathfrak{D}_K$ of the coefficients of polynomials $P_i^{\langle k\rangle}(x) \in \mathbb{Q}[x]$ $(i = 1,\dots,n)$; $k = 0,\dots,K$ is bounded by C_0^{D+kd}. The system $(Q^{\langle k\rangle}(x);P_1^{\langle k\rangle}(x),\dots,P_n^{\langle k\rangle}(x))$ is a system of Padé approximants to $f_1(x),\dots,f_n(x)$ with parameters $(D+kd,D+kd;[(1/n-\delta)D] - k(d+1))$. The heights of the polynomials $Q^{\langle k\rangle}(x)$; $P_1^{\langle k\rangle}(x),\dots,P_n^{\langle k\rangle}(x)$ are bounded as follows

$$H(Q^{\langle k\rangle}(x)) \leq H(D(x))^k \cdot H(Q(x)) \cdot 2^D (d+2)^k;$$

(8.8)

$$H(P_i^{\langle k\rangle}(x) \leq (D+1)C_1^D \cdot H(Q^{\langle k\rangle}(x)),$$

$i = 1,\dots,n$. Here $H(D(x))$ and $H(Q(x))$ are the heights of the polynomials of $D(x)$ and $Q(x)$, respectively, with

$$H(Q(x)) \leq D^{(1-\delta n)/\delta n} \cdot C_0^{(n+1-\delta n)(1-\delta n)D/(\delta n^2)}. \tag{8.9}$$

Proof: All statements of Corollary 8.1 are combinations of Theorems 4.1, 4.3 and Theorem 2.6. Also the bound for the height $H(Q(x))$ in (8.9) is contained in Theorem 4.1. We have to establish only the inequalities (8.8). The second inequality in (8.8) is a direct consequence of the representation (2.1) of the coefficients of the polynomial $P_i^{\langle k\rangle}(x) = [Q^{\langle k\rangle}(x) \cdot f_i(x)]_{(D+kd)}$ in terms of the coefficients of $Q^{\langle k\rangle}(x)$ and the expansion of $f_i(x)$ at $x = 0$: see the bounds (8.7). To prove the first inequality in (8.8) we need the definition (2.3): $Q^{\langle k\rangle}(x) = \frac{1}{k!} \cdot D(x)^k \cdot (\frac{d}{dx})^k Q(x)$. This expression implies: $H(Q^{\langle k\rangle}(x)) \leq H(D(x)^k) \cdot 2^{\deg(Q(x))} \cdot H(Q(x))$. This proves the inequalities (8.8) and Corollary 8.1.

Remark 8.2: The denominator $\mathfrak{D}_K$ of the coefficients of polynomials $P_i^{\langle k\rangle}(x)$ $(i = 1,\dots,n)$; $k = 0,\dots,K$ divides the denominator Δ_{D+Kd} of $\{a_{m,i}: 0 \leq m \leq D + Kd,\ i = 1,\dots,n\}$.

We apply the Padé approximants of the second kind described in Corollary 8.1 to study the denominators of the coefficients of the polynomials $\frac{1}{m!} \cdot D(x)^m \cdot A_{ij,m}(x)$: $i,j = 1,\dots,n$ for $m \geq 0$.

According to Theorem 3.1, the determinant

$$\Delta(x) = \det((Q^{\langle k\rangle}(x);P_1^{\langle k\rangle}(x),\ldots,P_n^{\langle k\rangle}(x)): k = 0,1,\ldots,n)$$

is not identically zero, provided that D is a sufficiently large integer. For our purposes we need a slightly different determinant (a lower left $(n-1)\times(n-1)$ minor of $\Delta(x)$):

$$(8.10) \qquad \nabla(x) = \det(P_i^{\langle j-1\rangle}(x): i,j = 1,\ldots,n).$$

We have to prove that for sufficiently large D, $\nabla(x)$ is not identically zero. The proof of this statement is very similar to that of Theorem 3.1.

Lemma 8.3: Let $f_1(x),\ldots,f_n(x)$ satisfy a system of equations (8.2) and let functions $f_1(x),\ldots,f_n(x)$ be linearly independent over $\mathbb{C}(x)$. Let $(Q(x);P_1(x),\ldots,P_n(x))$ be Padé approximants to a system of functions $f_1(x),\ldots,f_n(x)$ with parameters (D,D,M). Let for $k \geq 0$, the polynomials $(Q^{\langle k\rangle}(x);P_1^{\langle k\rangle}(x),\ldots,P_n^{\langle k\rangle}(x))$ be Padé approximants defined as in Theorem 2.6 (see formulas (2.9)). Let

$$\nabla(x) = \det(P_i^{\langle j-1\rangle}(x): i,j = 1,\ldots,n).$$

Then for a sufficiently large M, $M \geq c_{21}$, the determinant $\nabla(x)$ is not identically zero. Here c_{21} is a constant depending only on the system of linear differential equations (8.2) and on the orders of zeroes of $f_1(x),\ldots,f_n(x)$ at $x = 0$.

Proof: Let us assume that $\nabla(x) \equiv 0$. Let $\ell \leq n-1$ be the integer such that the first ℓ columns $s_k = (P_1^{\langle k-1\rangle}(x),\ldots,P_n^{\langle k-1\rangle}(x))^t$ $(k=1,\ldots,n)$ are linearly independent over $\mathbb{C}(x)$, but the $\ell+1$ columns is linearly dependent on them over $\mathbb{C}(x)$. We denote by F the matrix formed by the first ℓ columns s_k $(k = 1,\ldots,\ell)$ and by R and S we denote the matrices formed by the first ℓ rows and $n - \ell$ last rows of F, respectively. We can assume without loss of generality that R is a non-singular matrix. Then, as it was proved in [5, Chapter 11, Lemma 2] the degrees of the numerators and denominators of the rational function elements of the matrix $S\cdot R^{-1}$ are bounded by c_{22}, where c_{22} depends only on the system (8.2). Let us denote by G the $\ell \times n$ matrix with ℓ rows $(f_i(x),0,\ldots,\underbrace{-f_1(x)}_{i},0,\ldots,0): i = 1,\ldots,\ell$; and let G_0 and G_1 denote the matrices formed by the first ℓ columns and $n-\ell$ last columns of G, respectively. Then for the $\ell \times \ell$ matrix $T \overset{\text{def}}{=} G\cdot F$ we have $T_{i,j} = P_1^{<j-1>}(x)\cdot f_i(x) - P_i^{<j-1>}\cdot f_1(x)$: $i,j = 1,\ldots,\ell$. Also $T = G_0R + G_1S$, so that $T\cdot R^{-1} = G_0 + G_1\cdot S\cdot R^{-1}$. According to Definitions 2.3 and 2.5 and Theorem 2.6, all elements of T have orders of zeroes at $x = 0$ at least $D + M + 1 - \ell$. The polynomial elements of R have degrees in x at most $D + (\ell-1)d$. Hence $\det(TR^{-1})$, as a function of x, has a zero at $x = 0$ of order at least $\ell\{D + M + 1 - \ell\} - \ell\{D + (\ell-1)d\} = \ell\{M - (\ell-1)(d+1)\}$. On the other hand, $\det(G_0+G_1SR^{-1}) \not\equiv 0$, because functions $f_1(x),\ldots,f_n(x)$ are linearly independent over $\mathbb{C}(x)$. On the other hand, the degrees of all rational function elements of $S\cdot R^{-1}$ are bounded by c_{22}. This implies that the order of zero at $x = 0$ of $\det(TR^{-1}) = \det(G_0+G_1SR^{-1})$ is bounded by c_{23}, where c_{23} depends only on c_{22} and $f_1(x),\ldots,f_n(x)$. Consequently, for $\ell\{M - (\ell-1)(d+1)\} > c_{23}$, or, equivalently, for a sufficiently large M, the determinant $\nabla(x)$ is not identically zero. Lemma 8.3 is proved.

In particular, the determinant $\nabla(x)$ in (8.10) is non-zero for Pade approximants of the second kind from Corollary 8.1, whenever D is sufficiently large.

For further exposition we need an auxiliary statement on the iterated action of the differential operator $L = \frac{d}{dx}I - A$. We consider here and everywhere below the action of L and its powers on $n \times n$ matrices ψ of functions. For convenience we denote

$$(8.11) \qquad \psi^{\{m\}} \overset{\text{def}}{=} L^m\cdot\psi\left(= \left(\frac{d}{dx}\, I - A\right)^m\psi\right).$$

We need the following identity:

Lemma 8.4: Let, as above, the matrix A_m be defined as in (8.3). Then for an arbitrary $n \times n$ matrix ψ and $m \geq 0$ we have

$$(8.12) \qquad \Sigma_{k=0}^{m}\binom{m}{k}\cdot(-1)^k\cdot\left(\frac{d}{dx}\right)^{m-k}\left(\psi^{\{k\}}\right) = A_m\cdot\psi.$$

Proof: The identity (8.12) is obviously correct for $m = 0$. Let us assume that it is true for a given $m \geq 0$ and let us prove it for $m + 1$. We have for an arbitrary matrix ψ,

$$\Sigma_{k=0}^{m}\binom{m}{k}(-1)^k\left(\frac{d}{dx}\right)^{m-k}\left(\psi^{\{k\}}\right) = A_m\psi.$$

Let us substitute here $\psi^{\{1\}}$ for ψ. It follows from (8.11) that $\psi^{\{m+1\}} = (\psi^{\{1\}})^{\{m\}}$. Thus we have

$$(8.13) \qquad \Sigma_{k=0}^{m}\binom{m}{k}(-1)^k\left(\frac{d}{dx}\right)^{m-k}(\psi^{\{k+1\}}) = A_m\psi^{\{1\}}.$$

We can also differentiate (8.12) with respect to x once and obtain:

$$(8.14) \qquad \Sigma_{k=0}^{m}\binom{m}{k}(-1)^k\left(\frac{d}{dx}\right)^{m+1-k}(\psi^{\{k\}}) = A_m'\psi + A_m\frac{d}{dx}\psi,$$

where $A_m' = \frac{d}{dx}(A_m)$. We note that $\psi^{\{1\}} = (\frac{d}{dx})\psi - A\psi$. Let us substract (8.13) from (8.14). We get

$$\Sigma_{k=0}^{m}\binom{m}{k}(-1)^k\left(\frac{d}{dx}\right)^{m+1-k}(\psi^{\{k\}})$$

$$- \Sigma_{k=1}^{m+1}\binom{m}{k-1}(-1)^{k+1}\left(\frac{d}{dx}\right)^{m+1-k}(\psi^{\{k\}})$$

$$= (A_m' + A_mA)\psi.$$

Since $\binom{m}{k} + \binom{m}{k-1} = \binom{m+1}{k}$, we get

$$\Sigma_{k=0}^{m+1}\binom{m+1}{k}(-1)^k\left(\frac{d}{dx}\right)^{m+1-k}(\psi^{\{k\}}) = (A_m' + A_mA)\psi.$$

From (8.4) it follows that $A_{m+1} = A_m' + A_mA$. This proves Lemma 8.4.

We need also an elementary algebraic number lemma:

Lemma 8.5: Let $D(x) \in \mathbb{Z}[x]$ and $P(x) \in \mathbb{Z}[x]$. Then for arbitrary non-negative integers s and ℓ the polynomial

$$\frac{1}{s!}\cdot D(x)^{s+\ell}\cdot\left(\frac{d}{dx}\right)^s\cdot(D(x)^{-\ell}\cdot P(x))$$

has(rational) integral coefficients.

Proof: We have

$$\frac{1}{s!}\cdot D(x)^{s+\ell}\cdot\left(\frac{d}{dx}\right)^s\cdot(D(x)^{-\ell}\cdot P(x))$$

$$= \Sigma_{i=0}^{s}\ D(x)^{s+\ell}\cdot\frac{1}{i!}\left(\frac{d}{dx}\right)^i(D(x)^{-\ell})\cdot\frac{1}{(s-i)!}\cdot\left(\frac{d}{dx}\right)^{s-i}(P(x)).$$

Since $P(x) \in \mathbb{Z}[x]$, for any $k \geq 0$ we have $\frac{1}{k!}\cdot(\frac{d}{dx})^k(P(x)) \in \mathbb{Z}[x]$. Thus, to prove Lemma 8.5, it is sufficient to prove that for any non-negative integers i and ℓ the polynomial

$$\frac{1}{i!}\cdot D(x)^{i+\ell}\cdot\left(\frac{d}{dx}\right)^i\left(D(x)^{-\ell}\right)$$

has integral coefficients. To show this, let us put $D(x) = c\cdot\prod_{j=1}^n(x-\alpha_j)$. Then

$$\frac{1}{i!}\cdot D(x)^{i+\ell}\cdot\left(\frac{d}{dx}\right)^i\left(c^{-\ell}\cdot\prod_{j=1}^r(x-\alpha_j)^{-\ell}\right)$$

$$= \frac{i}{i!}\cdot c^{i+\ell}\cdot\prod_{j=1}^r(x-\alpha_j)^{i+\ell}\cdot\Sigma_{k_1+\dots+k_r=i}\frac{i!}{k_1!\cdots k_r!}$$

$$\times\ c^{-\ell}\prod_{j=1}^r\left(\frac{d}{dx}\right)^{k_j}\left((x-\alpha_j)^{-\ell}\right)$$

$$= c^i\cdot\prod_{j=1}^r(x-\alpha_j)^{i+\ell}\cdot\Sigma_{k_1+\dots+k_r=i}\prod_{j=1}^r\frac{(-\ell)\dots(-\ell-k_j+1)}{k_j!}$$

$$\times\ (x-\alpha_j)^{-\ell-k_j}$$

$$= c^i\cdot\Sigma_{k_1+\dots+k_r=i}\prod_{j=1}^r\binom{\ell+k_j-1}{k_j}(-1)^{k_j}\cdot(x-\alpha_j)^{i-k_j}.$$

For any indices $j_1,\dots,j_t$: $1\le j_1<\dots<j_t\le r$, $c\cdot\alpha_{j_1}\cdots\alpha_{j_t}$ is an algebraic integer. Thus we conclude that $\frac{1}{i!}\cdot D(x)^{i+\ell}\left(\frac{d}{dx}\right)^i\left(D(x)^{-\ell}\right)$ is a polynomial with (rational) integer coefficients. Lemma 8.5 is proved.

We return to the system of Padé approximants $(Q^{\langle k\rangle}(x);P_1^{\langle k\rangle}(x),\dots,P_n^{\langle k\rangle}(x))$: $k\ge 0$, constructed in Corollary 8.1. We assume from now on that $f_1(x),\dots,f_n(x)$ are linearly independent over $\mathbb{C}(x)$. Let us use now the recurrence formula from Corollary 2.7:

$$(8.15)\qquad \begin{aligned}&(P_1^{\langle k\rangle}(x),\dots,P_n^{\langle k\rangle}(x))^t\\ &= \frac{1}{k!}\cdot D(x)^k\cdot\left(\frac{d}{dx}I - A\right)^k\cdot(P_1(x),\dots,P_n(x))^t.\end{aligned}$$

These formulas (8.15) are true whenever $kd < [(1/n-\delta)D]$. Following our agreement (8.11) on notations, we put

$$(8.16)\qquad \begin{aligned}&(P_1^{\{k\}}(x),\dots,P_n^{\{k\}}(x))^t\\ &\overset{\text{def}}{=} \left(\frac{d}{dx}I - A\right)^k\cdot(P_1(x),\dots,P_n(x))^t.\end{aligned}$$

Consequently for $kd < [(1/n-\delta)D]$,

(8.17) $$P_i^{\{k\}}(x) = k!\cdot D(x)^{-k}\cdot P_i^{\langle k\rangle}(x):\ i = 1,\ldots,n.$$

We note that $P_i^{\{k\}}(x)$ are rational functions, and not necessarily polynomials.

Let us introduce an appropriate $n \times n$ matrix

(8.18) $$\psi \overset{\text{def}}{=} (P_i^{\{j-1\}}(x):\ i,j = 1,\ldots,n).$$

It is clear from the definition (8.16) that

$$(P_1^{\{k+\ell\}}(x),\ldots,P_n^{\{k+\ell\}}(x))^t$$

$$= (\frac{d}{dx} I-A)^k\cdot(P_1^{\{\ell\}}(x),\ldots,P_n^{\{\ell\}}(x))^t.$$

Thus, we get from (8.18),

(8.19) $$\psi^{\{m\}} = (P_i^{\{m+j-1\}}(x):\ i,j = 1,\ldots,n).$$

Let us use the identity from Lemma 8.4. We obtain

(8.20) $$\Sigma_{k=0}^m \binom{m}{k}\cdot(-1)^k\cdot(\frac{d}{dx})^{m-k}(\psi^{\{k\}}) = A_m\cdot\psi.$$

Let $(m+n-1)d < [(1/n-\delta)D]$. Then $P_i^{\{m+j-1\}}(x) = (m+j-1)!\cdot D(x)^{-m-j+1}\cdot P_i^{\langle m+j-1\rangle}(x)$ for $i,j = 1,\ldots,n$, according to (8.17). Let us put for any $k \geq 0$,

(8.21) $$P^{\{k\}} \overset{\text{def}}{=} \frac{1}{k!}\cdot D(x)^{k+n-1}\cdot\psi^{\{k\}}.$$

Thus, for $k \leq m$ we obtain:

(8.22)
$$P^{\{k\}} = (\frac{1}{k!}\cdot D(x)^{k+n-1}\cdot P_i^{\{k+j-1\}}(x):\ i,j = 1,\ldots,n)$$
$$= (\frac{(k+j-1)!}{k!}\cdot D(x)^{n-j}\cdot P_i^{\langle k+j-1\rangle}(x):\ i,j = 1,\ldots,n).$$

Thus we have, according to (8.20):

$$\Sigma_{k=0}^m \frac{m!}{(m-k)!k!}\cdot(-1)^k\cdot(\frac{d}{dx})^{m-k}(k!\cdot D(x)^{-k-n+1}\cdot P^{\{k\}}) = A_m\ \psi.$$

We can rewrite this identity as

$$\Sigma_{k=0}^{m}(-1)^{k}\cdot\frac{1}{(m-k)!}\cdot D(x)^{m+n-1}\cdot\left(\frac{d}{dx}\right)^{m-k}(D(x)^{-k-n+1}$$

(8.23)

$$\times\ p^{\{k\}}) = \frac{1}{m!}\cdot A_m\cdot D(x)^{m+n-1}\cdot\psi.$$

The identity (8.23) is the key element in the establishment of the bounds on the common denominator D_N of the coefficients of the polynomials $\frac{1}{m!}\cdot D(x)^m\cdot A_{ij,m}(x)$: $i,j = 1,\ldots,n$ and $m = 1,\ldots,N$.

According to the definition of matrix ψ in (8.18) and according to (8.17) we obtain

$$\det(D(x)^{n-1}\cdot\psi) = \{\Pi_{k=1}^{n-1}k!\cdot D(x)^k\}\cdot\det(P_i^{<j-1>}(x):\ i,j = 1,\ldots,n)$$

$$= \nabla(x)\cdot D(x)^{n(n-1)/2}\cdot\Pi_{k=1}^{n-1}k!\ .$$

Thus, according to Lemma 8.3, $\det(D(x)^{n-1}\cdot\psi)$ is non-zero for a sufficiently large D. We note now that in (8.23) in the left hand side we have a matrix with polynomial entries. Likewise, in the right hand side of (8.23), the matrices $A_m\cdot D(x)^m$ and $D(x)^{n-1}\cdot\psi$ have polynomial entries. If we denote by $Den(P;m)$ the common denominator of coefficients of all polynomial entries in the matrix $P^{\{k\}}$ for all $k = 0,\ldots,m$. It follows, from Lemma 8.5, that the left hand side of (8.23) is a matrix with entries that are polynomials with rational coefficients, whose common denominator divides $Den(P;m)$.

Let us invert now the matrix

(8.24) $$P \overset{def}{=} D(x)^{n-1}\cdot\psi.$$

According to (8.17) and (8.18), P is a matrix with polynomial entries. These polynomials have rational coefficients, whose common denominator, according to Corollary 8.1 and Remark 8.2, divides $\mathfrak{D}_{n-1}$. Here $\mathfrak{D}_{n-1}$ divides $\Delta_{D+(n-1)d}$ and $|\mathfrak{D}_{n-1}| \le |\Delta_{D+(n-1)d}| \le c_0^{D+(n-1)d}$, according to formula (8.7) and Remark 8.2. Thus $\mathfrak{D}_{n-1}\cdot P(= \mathfrak{D}_{n-1}\cdot D(x)^{n-1}\cdot\psi)$ has polynomial entries with rational integer coefficients, and

(8.25) $$\mathfrak{D}_{n-1}|\Delta_{D+(n-1)d}.$$

To invert the matrix P, we write $P = (P_{i,j}:\ i,j = 1,\ldots,n)$ and $P_{i,j} = (j-1)!D(x)^{n-j}\cdot P_i^{<j-1>}(x)$ $(i,j = 1,\ldots,n)$ according to (8.17). Let us denote by $M_{i,j}(x)$ the minor of $P_{i,j}$ in P and put

$N_{i,j}(x) = (-1)^{i+j} \cdot M_{i,j}(x)$ for $i,j = 1,\dots,n$. Then

$$P^{-1} = (N_{i,j}(x)/\det(P) \ : \ i,j = 1,\dots,n).$$

Thus the matrix

$$N \overset{\text{def}}{=} (N_{i,j}(x) : \ i,j = 1,\dots,n) \tag{8.26}$$

satisfy

$$P \cdot N = \det(P) \cdot I. \tag{8.27}$$

We see that the matrix N has as all its elements polynomials with rational coefficients, whose common denominator divides $\mathfrak{D}_{n-1}$. We can also estimate the sizes of polynomial entries of the matrix N using Corollary 8.1. According to (8.8) and (8.9), the sizes of polynomials $P_{i,j}$ are bounded as

$$H(P_{i,j}) \le e^{c_{24}(\delta) \cdot D} \ : \ i,j = 1,\dots,n \tag{8.28}$$

where $c_{24}(\delta)$ depends on n, δ, c_0, c_1 and $D(x)$ only. Consequently,

$$H(N_{i,j}(x)) \le e^{c_{25}(\delta) \cdot D} \ : \ i,j = 1,\dots,n, \tag{8.29}$$

for $c_{25}(\delta)$ depending on n, δ, C_0, C_1 and $D(x)$.

Finally, all elements of the matrix $\mathfrak{D}_{n-1}^{n} \cdot N$ are polynomials with integral coefficients. Similarly, according to (8.27), $\mathfrak{D}_{n-1}^{n} \cdot \det(P)$ is a polynomial with integral coefficients. Taking into account (8.25) the bounds (8.7) and (8.29), we obtain

$$\begin{aligned} &H(\mathfrak{D}_{n-1}^{n} \cdot \det(P)) \le e^{c_{26}(\delta) \cdot D}, \\ &H(\mathfrak{D}_{n-1}^{n} \cdot N_{i,j}(x)) \le e^{c_{27}(\delta) D} \ : \ i,j = 1,\dots,n. \end{aligned} \tag{8.30}$$

We multiply now both sides of identity (8.23) by $\mathfrak{D}_{n-1}^{n} \cdot \mathrm{Den}(P;m) \cdot N$. Here $\mathfrak{D}_{n-1}^{n}$, $\mathrm{Den}(P;m)$ are (rational) integers. We get:

$$\begin{aligned} &\Sigma_{k=0}^{m} (-1)^k \cdot \mathrm{Den}(P;m) \cdot \frac{1}{(m-k)!} \cdot D(x)^{m+n-1} \cdot \left(\frac{d}{dx}\right)^{m-k} \times (D(x)^{-k-n+1} \times P^{\{k\}}). \\ &\mathfrak{D}_{n-1}^{n} \cdot N = \frac{1}{m!} D(x)^m \cdot A_m \cdot \{\mathfrak{D}_{n-1}^{n} \cdot \det(P)\} \cdot \mathrm{Den}(P;m). \end{aligned} \tag{8.31}$$

According to the definition, $\mathrm{Den}(P;m)$ is the common denominator of coefficients of polynomial entries of $P^{\{k\}}$ for $k = 0,\dots,m$. Based on this we prove the following crucial

Proposition 8.6: The common denominator of coefficients of polynomial entries $\frac{1}{m!}\cdot D(x)^m \cdot A_{i,j,m}(x)$ of the matrix $\frac{1}{m!}\cdot D(x)^m\cdot A_m$ for $m = 0,\dots,[(1/n-\delta)D]/d - n$ (or $(m+n-1)d < [(1/n-\delta)D]$) is bounded by $e^{c_{28}(\delta)\cdot D}$. Here $c_{28}(\delta)$ is an effective constant that depends only on n, δ, C_0, C_1 and $D(x)$.

Proof: First of all we estimate $\mathrm{Den}(P;m)$. From the expression (8.22) and Corollary 8.1, Remark 2 we deduce that $\mathrm{Den}(P;m)$ divides $\Delta_{D+(m+n-1)d}$. Here

$$|\Delta_k| \le C_0^k$$

for any $k \ge 0$. If we denote $B \overset{\text{def}}{=} [(1/n-\delta)D]/d-n$, then $\mathrm{Den}(P;m)$ divides $\Delta_{D+(B+n-1)d}$ for any $m \le B$. Let us denote the polynomial

$$\mathfrak{D}_{n-1}^n\cdot \det(P)\cdot \Delta_{D+(B+n-1)d}$$

(cf. the right hand side of (8.31)), by $C(x)$ ($C(x)$ is independent of m). It follows from (8.30) that $C(x)$ is a polynomial with rational integer coefficients of sizes bounded by

$$H(C(x)) \le e^{c_{26}(\delta)\cdot D}\cdot C_0^{(1/n-\delta)D} \le e^{c_{29}(\delta)\cdot D}. \tag{8.32}$$

On the other hand, according to the definition of $\mathrm{Den}(P;m)$, according to Lemma 8.5 and because $\mathfrak{D}_{n-1}^n\cdot N$ is a matrix from $M_n(\mathbb{Z}[x])$, the matrix in the left hand side of

$$\begin{aligned} &\Sigma_{k=0}^m (-1)^k\cdot \Delta_{D+(B+n-1)d}\cdot \frac{1}{(m-k)!}\cdot D(x)^{m+n-1} \\ &\times \left(\frac{d}{dx}\right)^{m-k}\left(D(x)^{-k-n+1}\cdot P^{\{k\}}\right)\cdot \mathfrak{D}_{n-1}^n\cdot N \\ &= \frac{1}{m!}\cdot D(x)^m\cdot A_m(x)\cdot C(x), \end{aligned} \tag{8.33}$$

has polynomial entries with integral coefficients. Consequently, from (8.33) it follows that every polynomial entry $\frac{1}{m!}\cdot D(x)^m\cdot A_{i,j;m}(x)\cdot C(x)$ ($i,j = 1,\dots,n$) of $\frac{1}{m!}D(x)^m\cdot A_m(x)\cdot C(x)$ has rational integer coefficients.

We can use the Gauss lemma [21] in the following form:

Gauss Lemma 8.7: Let $f(x_1,\dots,x_m)$ be a polynomial with rational integer coefficients and $g(x_1,\dots,x_m)$ be a polynomial with rational coefficients. If $f\cdot g$ has integral coefficients, then the common denominator of coefficients of $g(x_1,\dots,x_m)$ divides the least common multiplier of coefficients of the polynomial $f(x_1,\dots,x_m)$.

From Lemma 8.7 it follows that the common denominator of coefficients of the polynomial $\frac{1}{m!}\cdot D(x)^m\cdot A_{i,j;m}(x)$: $i,j = 1,\dots,n$ and $m \le B$ divides the least common multiplier of the coefficients of $C(x)$. In particular, this common denominator is bounded by the height of the $C(x)$. From the bound on $H(C(x))$ in (8.32) we deduce Proposition 8.6.

If we now fix δ, $1/n > \delta > 0$, Proposition 8.6 implies the (G,C)-property as formulated above. One has to substitute $B = [(1/n-\delta)D]/d-n$ for N with fixed $\delta(1/n > \delta > 0)$. Then the constant C in the definition of (G,C)-property can be chosen as $e^{c_{28}(\delta)2d/(1/n-\delta)}$ for a sufficiently large N. This proves Theorem III from §1.

The crude estimates of Proposition 8.6 can be considerably improved to exhibit C in terms of C_0, n and d only.

References.

[1] C.L. Siegel, Über einige Anwendungen diophantischer Approximationen, Abh. Preuss. Akad. Wiss. Phys. Math. Kl. 1, 1929.

[2] C.L. Siegel, Transcendental Numbers, Princeton University Press, Princeton, 1949.

[3] G.V. Chudnovsky, Contributions to the Theory of Transcendental Numbers, Mathematical Surveys and Monographs, 19, American Mathematical Society, 1984, Chapter 5.

[4] D.V. Chudnovsky, G.V. Chudnovsky, Applications of Padé approximations to the Grothendieck conjecture on linear differential equations, see next paper, this volume.

[5] A. Baker, Transcendental Number Theory, Cambridge University Press, Cambridge, 1979.

[6] A.L. Galočhkin, Lower bounds of polynomials in the values of a certain class of analytic functions, Math. Sb. 95 (1974), 396-417.

[7] K. Väänänen, On linear forms of certain class of G-functions and p-adic G-functions, Acta Arith. 36 (1980), 273-295.

[8] Y.Z. Flicker, On p-adic G-functions, J. London Math. Soc. 15 (1977), 395-402.

[9] E. Bombieri, On G-functions, in Recent Progress in Analytic Number Theory (ed. by H. Halberstam and C. Hooly), Academic Press, N.Y., v.2, 1981, 1-67.

[10] G.V. Chudnovsky, Measures of irrationality, transcendence and algebraic independence.Recent progress, in Journees Arithmetiques 1980 (ed. by J.V. Armitage), Cambridge University Press, 1982, 11-82.

[11] N. Katz, Algebraic solutions of differential equations, Invent. Math., 18 (1972), 1-118.

[12] T. Honda, Algebraic differential equations, Symposia Mathematica v.24, Academic Press, N.Y., 1981, 169-204.

[13] B. Dwork, Arithmetic theory of differential equations, ibid., 225-243.

[14] K. Mahler, Perfect systems, Composito Math. 19 (1968), 95-166.

[15] K. Mahler, An analogue to Minkowski's geometry of numbers in a field of series, Ann. of Math. 42 (1941), 488-522.

[16] W.M. Schmidt, Diophantine Approximations, Lecture Notes Math. v. 785, Springer, 1980.

[17] E.R. Kolchin, Rational approximations to solutions of algebraic differential equations, Proc. Amer. Math. Soc. 10 (1959), 238-244.

[18] D.V. Chudnovsky, G.V. Chudnovsky, Rational approximations to solutions of linear differential equations, Proc. Nat'l. Acad. Sci. USA 80 (1983), 5158-5162.

[19] D.V. Chudnovsky, G.V. Chudnovsky, Padé approximations to solutions of linear differential equations and applications to diophantine analysis, Lecture Notes Math. v. 1052, Springer, 1984, 85-167.

[20] G.V. Chudnovsky, On some applications of diophantine approximations, Proc. Nat'l. Acad. Sci. USA 81 (1984), 1926-1930.

[21] J.W.S. Cassels, An Introduction to Diophantine Approximations, Cambridge University Press, Cambridge, 1957.

[22] J. Nuttal, Asymptotics of diagonal Hermite-Padé polynomials, J. Approximation Theory, 1985 (in press).

[23] C. Brezinski, Padé-type Approximations and General Orthogonal Polynomials, Birkhauser, Boston, 1980.

[24] G.V. Chudnovsky, The Thue-Siegel-Roth theorem for values of algebraic functions, Proc. Japan. Academy, Ser. A 59 (1983), 281-284.