Diophantine approximation of values of G-functions at rational points

Jiacheng Xia[a], Advisor: Stéphane Fischler[b], Co-advisor: Tanguy Rivoal[c]

[a] *AAG M2, Department of Mathematics, Université Paris-Sud*
[b] *AAG, Department of Mathematics, Université Paris-Sud*
[c] *Institut Fourier, Université Grenoble Alpes*

Abstract

The field of Diophantine approximation deals with the approximation of real numbers by rational (algebraic, in more general sense) numbers, and is closely related to metric number theory and transcendental number theory as well as solutions of diophantine equations. There are two important classes of analytic functions, E-functions and G-functions. In this master thesis, we consider the measure of irrationality of the values of G-functions at rational points. Under further assumptions than usual on the G-function, Zudilin has improved the range of rational points for which diophantine approximation of function values at those points is valid, following the previous ineffective approach of Chudnovsky. In this paper we point out that the case for considering two functions are a little bit more tricky than that of $m \geq 3$ functions, and the assumption there can be weaker. We also discuss the idea and merit of applying graded Padé approximants to G-functions, which is the core of Zudilin's paper.

1. Introduction

Siegel (1929) has first defined E- functions and G- functions as follows.

Definition 1.1. A G-function $f(z)$ is a power series $f(z) = \sum_{n=0}^{\infty} a_n z^n$ such that the coefficients $a_n \in \overline{\mathbb{Q}}$ and there exists constant $C > 0$ such that for any $n \geq 1$:

(i) the maximum of the moduli of the algebraic conjugates of a_n is no more than C^n;

(ii) there is a sequence $d_n \in \mathbb{Z}$ with $|d_n| < C^n$, such that $d_n a_m \in \overline{\mathbb{Z}}$ for all $1 \leq m \leq n$;

(iii) $f(z)$ satisfies a homogeneous linear differential equation with coefficients in $\overline{\mathbb{Q}}(z)$.

An E-function is a power series $f(z) = \sum_{n=0}^{\infty} \frac{a_n}{n!} z^n$ with algebraic coefficients and satisfying the same conditions (i)-(iii).

Remark 1.2. For our interest in this particular paper, it is sufficient to consider the "rational" version of the definition. That is to say, by replacing $\overline{\mathbb{Q}}$ with $\mathbb{Q}$ for all that appear in the definition, and correspondingly by replacing "algebraic" with "rational". Note that there is also a little technical difference between this standard version and other mathematician's works, so we will recall the difference when necessary. From the definition, it is clear that the elementary functions e^z, $sinz$, $cosz$ are all E-functions but not G-functions. On the other hand, for examples of G-functions which are not E-functions, we give $Li_m(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^m}$, $m \geq 1$ and $g_n(z) = \sqrt[n]{1 - z^n}$, $n \in \mathbb{Z}_{\geq 2}$. Note that $Li_m(z)$ are G-functions because

$$l.c.m(1, 2, \cdots, n) = \prod_{p^r \leq n} p^r \leq n^{\pi(n)} \leq n^{\frac{2n}{\log n}} = e^{2n}$$

Email addresses: `philimathmus@gmail.com` (Jiacheng Xia), `Stephane.Fischler@math.u-psud.fr` (Advisor: Stéphane Fischler), `Rivoal@univ-grenoble-alpes.fr` (Co-advisor: Tanguy Rivoal)

where we use an estimation given by prime number theorem. These two kinds of functions expose our interest to pursue the irrationality of G-functions at some rational points: it is conjectured that the Riemann zeta function takes irrational value at all odd numbers ≥ 5, ($\zeta(3) \notin \mathbb{Q}$ is first shown by Apéry.) and this could be expressed also as $Li_{2k+1}(1) \notin \mathbb{Q}$. Even the Fermat's Last Theorem could be expressed as $g_n(z) \notin \mathbb{Q}$ at all rational points $0 < z < 1$ and for all $n \geq 3$. Although these two famous problems could not be studied from the aspect of G-functions (we will mention one reason later), it is still clear that studying the irrationality of the values of G-functions at some rational points is very interesting and meaningful. Last but not least, under our definition of E-functions it follows that they must be entire functions, while G-functions need not be.

It is essential to consider an arbitrary linear system of differential equations, namely we generalize the third condition in the definition to m G-functions, $f_j(z)$ $(j = 1, 2, \cdots, m)$ satisfying a system of linear differential equations

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}z} f_l(z) = Q_{l0}(z) + \sum_{j=1}^{m} Q_{lj}(z) f_j(z), && \forall l, \\ &Q_{lj}(z) \in \mathbb{C}(z), && \forall l, j. \end{aligned} \tag{1.1}$$

Denote by $T = T(z) \in \mathbb{C}[z]$ a least common denominator of all the rational functions $Q_{lj}(z)$.

Above is a general introduction to the inhomogeneous differential system for G-functions. Historically the first important difficult theorem on the relation between functional independence and their values' independence is on inhomogeneous differential system for E-functions, obtained by Shidlovskii in 1955.

Theorem 1.3 (First Fundamental theorem)**.** *Suppose the E-functions $f_j(z)$ $j = 1, 2, \cdots, m$ where $m \geq 2$, form a solution of the inhomogeneous system 1.1. Let α be an algebraic number such that $\alpha T(\alpha) \neq 0$. Then the numbers*

$$f_1(\alpha), \cdots, f_m(\alpha)$$

are algebraically independent over $\mathbb{Q}$ if and only if the functions $f_j(z)$ are algebraically independent over $\mathbb{Q}(z)$.

Zudilin's paper [5] is based on the systematic method in the first fundamental theorem's proof, which we call the Siegel-Shidlovskii Method and which can lead to the discussion of values of more general G-functions at rational points, under some further assumptions. In this particular mémoire following Zudilin's paper, we need to study graded Padé approximants of a given set of G-functions satisfying some further assumptions, and further investigate their higher order derivative forms with the help of certain recurrence formulas. Note that in fact the case $m = 2$ is a little bit more difficult than $m \geq 3$, since the assumptions could be weaker and for that purpose we need a special trick for $m = 2$ below in this section. However the case $m = 2$ is also convenient for writing more explicitly, so we mainly focus on this case in this mémoire, except from some remarks mentioning that the same argument works for all $m \geq 3$ in the proofs.

Definition 1.4. We fix our main problem here of rational version and define all the relevant notations in it. Let $f_1(z)$ and $f_2(z)$ be G-funtions (simplified into $\mathbb{Q}$ in place of $\overline{\mathbb{Q}}$ and $\mathbb{Z}$ in place of $\overline{\mathbb{Z}}$ in definition 1.1) satisfying a system of linear differential equations:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}z} f_1(z) =& Q_{10}(z) + Q_{11}(z) f_1(z) + Q_{12}(z) f_2(z) \\ \frac{\mathrm{d}}{\mathrm{d}z} f_2(z) =& Q_{20}(z) + Q_{21}(z) f_1(z) + Q_{22}(z) f_2(z) \end{aligned} \tag{1.2}$$

where $Q_{lj}(z) \in \mathbb{C}(z)$ and $1, f_1, f_2$ are linearly independent over $\mathbb{C}(z)$. We also consider its conjugate homogeneous part:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}z}a_1(z) =& S_{11}(z)a_1(z) + S_{12}(z)a_2(z)\\ \frac{\mathrm{d}}{\mathrm{d}z}a_2(z) =& S_{21}(z)a_1(z) + S_{22}(z)a_2(z)\end{aligned} \tag{1.3}$$

where $S_{jl}(z) = -Q_{lj}(z) \in \mathbb{Q}(z)$.

Then our first work is to show the omitted lemma in Zudilin's paper [5]:

Lemma 1.5. *In Eq. (1.2), for all l, j we have*

$$Q_{lj}(z) \in \mathbb{Q}(z).$$

The polynomial T can be chosen in $\mathbb{Z}[z]$ so that

$$T(z)Q_{lj}(z) \in \mathbb{Z}[z].$$

Proof. First we show that in fact in general if 1, y_l where $l = 1, 2, \cdots, m$ are linearly independent over $\mathbb{C}(z)$, then the linear differential system is uniquely determined. Suppose they satisfy another distinct system, then we immediately get that 1, y_l are linearly dependent which is a contradiction. Next, if we multiply the original equation by $T(z)$ and view the new system

$$\begin{aligned}&T(z)\frac{\mathrm{d}}{\mathrm{d}z}y_l(z) = T(z)Q_{l0}(z) + \sum_{j=1}^{m} T(z)Q_{lj}(z)y_j(z), && \forall l\\ &T(z)Q_{lj}(z) \in \mathbb{C}[z], && \forall l, j.\end{aligned} \tag{1.4}$$

as a linear dependence of the $m+1$ functions $\frac{\mathrm{d}}{\mathrm{d}z}y_l$, y_j where $j = 1, 2, \cdots, m$. We know such linear dependence has coefficients in $\mathbb{C}[z]$ and now we want to show that there exists a linear dependence of these $m+1$ functions with coefficients in $\mathbb{Z}[z]$. We argue as follows: Viewing the coefficients of the wanted polynomials, exactly $T(z)$ and $T(z)Q_{lj}(z)$ up to their highest order terms, as indeterminates and solve the homogeneous system of countably many linear equations in rational coefficients, with only finitely many indeterminates. since the original dependence gives a nontrivial complex solution, it follows by choosing a maximal linearly independent subsystem of these countably many linear equations (still homogeneous), that there exists an integer solution of the system which is to say, we obtain $T(z) \in \mathbb{Z}[z]$ and $T(z)Q_{lj}(z) \in \mathbb{Z}[z]$ satisfying Eq. (1.4), which might be different from the original ones. Finally note that the system is uniquely determined, so these $Q_{lj}(z) \in \mathbb{Q}(z)$ are exactly the original ones. □

The above discussions are valid for general G-functions or even any analytic functions $f_j(z)$ which are together with 1 linearly independent over $\mathbb{C}(z)$. Zudilin posed more elaborate assumptions in his paper to get a stronger result, one of which is the so called "cancellation of factorials" condition for the differential equations 1.2 and its homogeneous conjugate equations 1.3. This was first axiomatized by Galochkin in his paper [2]

Next we introduce the further assumptions of Zudilin and the main result in the dimension two case. First we should define a special class in the whole class of G-functions, which is called the class $G(\mathbb{Q}, C, \Phi)$.

Definition 1.6. Fix constants $C, \Phi \geq 1$. We say a set of functions $f_j(z) = \sum_{n=0}^{\infty} f_{j,n}z^n$ where $j = 1, 2, \cdots, m$ belongs to the class $G(\mathbb{Q}, C, \Phi)$ if and only if

(i) For all $j = 1, 2, \cdots, m$ and $n \geq 1$, $f_{j,n} \in \mathbb{Q}$;

(ii) For all $j = 1, 2, \cdots, m$ and any $\epsilon > 0$, $f_{j,n} = O(C^{(1+\epsilon)n})$ as $n \to \infty$;

(iii) There are natural numbers d_n satisfying for all $j = 1, 2, \cdots, m$ and for any $\epsilon > 0$, $d_n = O(\Phi^{(1+\epsilon)n})$ as $n \to \infty$, such that $d_n f_{j,l} \in \mathbb{Z}$ for all $j = 1, 2, \cdots, m$ and for all the indices $l = 1, 2 \cdots, n$.

Next we may pose the so called "cancellation of factorials" condition for the differential system Eq. (1.2) and its conjugate homogeneous part Eq. (1.3). Since Eq. (1.3) is homogeneous, it follows that

$$\begin{aligned} \frac{\mathrm{d}^n}{\mathrm{d}z^n} a_1(z) =& S_{11}^{[n]}(z) a_1(z) + S_{12}^{[n]}(z) a_2(z) \\ \frac{\mathrm{d}^n}{\mathrm{d}z^n} a_2(z) =& S_{21}^{[n]}(z) a_1(z) + S_{22}^{[n]}(z) a_2(z) \end{aligned} \tag{1.5}$$

where $S_{jl}^{[n]}(z) \in \mathbb{Q}(z)$, for all $j, l = 1, 2$ and for any $n \in \mathbb{N}$. Let $T_*(z) \in \mathbb{Z}[z]$ be the least common denominator of $S_{jl}(z) = -Q_{lj}(z)$, then $T_*(z) S_{jl}(z) \in \mathbb{Q}[z]$ and it is not hard to show by induction that $T_*^n(z) S_{jl}^{[n]}(z) \in \mathbb{Q}[z]$ and it is thus meaningful to talk about whether there exists a certain sequence of natural numbers satisfying geometrical growing rate to be the least common multiple of all the rational coefficients of $T_*^i(z) S_{jl}^{[i]}(z) \in \mathbb{Q}[z]$, for $i = 1, 2, \cdots, n$. Such important property is called the "cancellation of factorials" condition and we make the definition clear by denoting a new class for the differential equations.

Definition 1.7. Fix a constant $\Psi \geq 1$. We say a system of differential equations 1.2 with its conjugate homogeneous part 1.3 belongs to the class $G(\mathbb{Q}, \Psi)$, or satisfies the "cancellation of factorials" condition with factor Ψ, if and only if there is a sequence of natural numbers ψ_n satisfying for any $\epsilon > 0$, $\psi_n = O(\Psi^{(1+\epsilon)n})$ as $n \to \infty$, such that

$$\frac{\psi_n}{n!} T_*^i(z) S_{jl}^{[i]}(z) \in \mathbb{Z}[z]$$

for all $j, l = 1, 2$ and $i = 1, 2 \cdots, n$.

Remark 1.8. In fact there is another condition proposed by Bombieri in his paper written for a book [3] which plays the same role as the "cancellation of factorials" condition here. It is defined on a number field K, which equals to $\mathbb{Q}$ in our paper. Let $|\cdot|_v$ for each place v be the absolute value on $\mathbb{Q}$ where $\mathbb{Q}_v$ is the completion of $\mathbb{Q}$ with respect to v (so $v = p$ for a prime p or $v = \infty$). Let Ω_v be the completion of $\overline{\mathbb{Q}_v}$. Now we fix a point $x \in \Omega_v$ and consider the solutions of the system of homogeneous differential equations as formal power series in $(z - x)$ whose coefficients are all in Ω_v. We know that these formal power series would converge in Ω_v if $|z - x|_v$ is small enough. Thus we can define $r_v(x)$ to be the greatest radius for convergence and the system has m such solutions that are linearly independent over Ω_v. It is stated and proved there that if x is chosen in a technical way, then $r_v(x)$ does not depend on the choice of x. We can denote it by r_v. Bombieri's condition can be stated as follows:

$$\sum_{v \in V_0, r_v < 1} \log\left(\frac{1}{r_v}\right) < +\infty.$$

Here V_0 is the set of non-archimedean places on K, and in particular, all the primes when $K = \mathbb{Q}$. André has shown in his monograph [6] that in fact the factorial canceling properties of Galochkin and Bombieri's condition are equivalent.

From now on, for convenience we replace $T_*(z)$ by $T(z)$, so $T^n(z) S_{jl}(z) \in \mathbb{Q}[z]$ and we have upper bounds for degree and height of this polynomial:

Next we denote by t the following formula of a given differential system Eq. (1.2),

$$t = \max\Big(\deg T - 1, \max_{l,j}\big(\deg(TQ_{lj})\big)\Big).$$

We denote H by the following formula where $H(p)$ means the height for a given polynomial p.

$$H = \max\Big(\mathrm{H}(T), \max_{l,j}\big(\mathrm{H}(TQ_l)\big)\Big)$$

Lemma 1.9. *There is a constant $C_1 = C_1(t)$, such that*

$$\deg\left(T^n(z)S_{jl}^{[n]}(z)\right) \le tn$$

and

$$\mathrm{H}\left(T^n(z)S_{jl}^{[n]}(z)\right) \le C_1\big(2(t+1)^2 H\big)^n n!$$

for all $j, l = 1, 2$ and $n \in \mathbb{N}$.

We skip the proof of it, which is a routine after writing down the recurrence formula for $T^n(z)S_{jl}^{[n]}(z)$ between $(n+1)$ and n.

Then we can state the main theorem in Zudilin's paper [5].

Theorem 1.10 (Main Theorem)**.** *Let two functions $f_1(z)$ and $f_2(z)$ belong to the class $G(\mathbb{Q}, C, \Phi)$ and satisfy a system of linear differential equations 1.2 with its homogeneous conjugate part 1.3. Suppose that this system belongs to the class $G(\mathbb{Q}, \Psi)$. Assume furthermore that $1, f_1(z), f_2(z)$ are linearly independent over $\mathbb{C}(z)$ and these two functions $f_1(z)$ and $f_2(z)$ can not be written in the forms $a_1(z)+a_2(z)\sqrt{b_1(z)+b_2(z)\sqrt{P(z)}}$ and $a_3(z)+a_4(z)\sqrt{P(z)}+a_5(z)\sqrt{b_1(z)+b_2(z)\sqrt{P(z)}}$ where $a_i(z) \in \mathbb{Q}(z)$, $b_j(z) \in \mathbb{Q}[z]$ and $P(z) \in \mathbb{Q}[z]$ has only simple roots. (For the case $m \ge 3$ we only need to pose a stronger condition that $f_j(z)$ are algebraically independent over $\mathbb{C}(z)$, without such exclusion.)*

Consider an arbitrary rational point $\alpha = \frac{a}{b}$ with $a \in \mathbb{Z}$ and $b \in \mathbb{N}$ such that $\alpha T(\alpha) \neq 0$ where $T \in \mathbb{Z}[z]$ is a polynomial such that $T(z)Q_{ij}(z) \in \mathbb{Z}[z]$ in 1.2.

Let $\varepsilon < \frac{1}{t+3}$ be an arbitrary positive constant and $N = [\frac{1}{\varepsilon} - 1]$; Set the constant

$$C_0 = e^{\varepsilon(1-\log\varepsilon)}\big(2(t+1)^2 H\Psi\big)^{\varepsilon(1+\log N)}\Phi^{1+t\,\varepsilon}(C\Phi)^{\frac{N}{\varepsilon}(2-\frac{1}{N+1})}$$

and the constant

$$\eta_0 = \frac{(1+t\,\varepsilon)\log b + \log C_0}{\big(1-(t+3)\,\varepsilon\big)\log b - \log C_0 - (2-3\,\varepsilon)\log(C|a|)}$$

If α is such that $\eta_0 > 0$, that is,

$$b^{1-(t+3)\,\varepsilon} > C_0(C|a|)^{2-3\,\varepsilon},$$

then $f_j(\alpha)$ $(j = 1, 2)$ are irrational. Furthermore, for any $\eta > \eta_0$ and arbitrary $(p, q) \in \mathbb{Z} \times \mathbb{N}$, there is a constant $q_* = q_*(f_1, f_2; \alpha, \varepsilon, \eta)$, such that for any $q > q_*(f_1, f_2; \alpha, \varepsilon, \eta)$, the following inequality holds:

$$\left| f_j(\alpha) - \frac{p}{q} \right| > q^{-1-\eta}. \tag{1.6}$$

That is to say, in the language of Diophantine Approximation, the irrationality measures of $f_j(\alpha)$ $(j = 1, 2)$ are bounded above by $1 + \eta_0$.

Remark 1.11. Note that if $\eta_0 > 0$, then automatically we have $\eta_0 > 1$ from the formula of η_0, and this is consistent with the fact that every irrational number has irrationality measure at least 2, as the one dimensional case of our Corollary 2.2.

Also note that in the conclusion of this main theorem, it is not required that (p, q) should be coprime. Therefore, the conclusion part of "furthermore" implies that $f_j(\alpha)$ $(j = 1, 2)$ are irrational. So it suffices to prove the inequality 1.6.

Now we introduce the ideas to prove this main theorem. First we recall Siegel's lemma and some basic facts about Diophantine Approximants, and prove the existence of graded Padé approximants, which is the main tool of this paper, and which is a little bit different from the Padé approximation of first kind. After proving a lemma on the estimation of numerical forms, we obtain the central result on this side. On the other hand, we study the traditional Siegel-Shidlovskii's method, based on discussion of the Padé approximants of the first kind. After proving a series of linear algebraic results including the "Shidlovskii's lemma", we end by the proof of the non-vanishing property of functional determinant. Note that in this process we can not avoid an ineffective constant. Finally we apply the Siegel-Shidlovskii's method to show the full rank property of numerical matrix consisting of the coefficients appeared in the graded Padé approximants, in a similar way. Combining these two sides, we are given the conclusion of the inequality 1.6.

In fact, under the assumption of Theorem 1.10 we can find some special tricks for $m = 2$ case which is a little bit different from the $m \geq 3$ case. In our mémoire, all the proofs for the $m = 2$ case are just the same as that for $m \geq 3$ case, and only at somewhere shall we mention that the particular method works only for $m = 2$ case. Next we shall prove the first special trick for $m = 2$ when we compute functional determinant, which is neglected in Zudilin's paper [5].

Lemma 1.12. *Under the assumption of Theorem 1.10, there exists $l \in \{1, 2\}$ such that 1, $f_1(z)$, $f_2(z)$, and $f_l^2(z)$ are linearly independent over $\mathbb{C}(z)$. Without loss of certainty we can fix here $l = 2$.*

Proof. Suppose the contrary, we can write

$$f_j^2(z) = P_{j0}(z) + P_{j1}(z)f_1(z) + P_{j2}(z)f_2(z)$$

for $j = 1, 2$ where $P_{jk} \in \mathbb{C}(z)$. Now if $P_{12}(z) = 0$, then $f_1(z)$ is in some quadratic extension of $\mathbb{C}(z)$ as a meromorphic function. And if $P_{12}(z) \neq 0$, then $f_2(z)$ can be written as a quadratic polynomial of $f_1(z)$ with coefficients in $\mathbb{C}(z)$, and hence $f_1(z)$ must be in a quadratic extension of some quadratic extension of $\mathbb{C}(z)$ as a meromorphic function. However $f_1(z)$ and $f_2(z)$ satisfy Eq. (1.2). If the first case happens, that is to say $f_1(z) \in \mathbb{C}(z)(\sqrt{P(z)})$ for some $P(z) \in \mathbb{C}[z]$ with only simple roots, then we know that $\frac{\mathrm{d}}{\mathrm{d}z}f_1(z) \in \mathbb{C}(z)(\sqrt{P(z)})$ and hence from Eq. (1.2) that $f_2(z) \in \mathbb{C}(z)(\sqrt{P(z)})$. This is a contradiction with the assumption that 1, $f_1(z)$, and $f_2(z)$ are linearly independent over $\mathbb{C}(z)$. If the second case happens, This exactly gives the situation that we have excluded in Theorem 1.10, see also in section 7 about this exclusion. $\square$

2. Pigeonhole principle

The first result about upper bounds for Diophantine approximation is Dirichlet theorem (1842), in which pigeonhole principle plays an important role.

Theorem 2.1. *Consider the space $\mathbb{R}^d$ equipped with $|x| = max_{1 \leq i \leq d}|x_i|$. Given a point $x \in \mathbb{R}^d$ and $Q > 1$ an integer. Then there exists a pair $(p, q) \in \mathbb{Z}^d \times \mathbb{N}^*$ such that*

(i) $1 \leq q < Q^d$;

(ii) $|qx - p| \leq \frac{1}{Q}$.

Proof. Consider $Q^d + 1$ points in the cube $[0, 1]^d$: $\mathbf{0} = (0, 0, \cdots, 0)$, $\mathbf{1} = (1, 1, \cdots, 1)$,$\{kx\}$ where $k = 1, 2, \cdots, (Q^d - 1)$. Each point is contained in (at least) one of the cubes $\prod_{i=1}^d [\frac{k_i}{Q}, \frac{k_i+1}{Q}]$ where $k_i \in \{0, 1, \cdots, Q - 1\}$. By the Pigeonhole principle, there are at least two points in one cube. Note that they can not be $\mathbf{0}$ and $\mathbf{1}$. Discussion of three cases shows the existence of a pair (p, q) satisfying both (i) and (ii). $\square$

Corollary 2.2. *Given an arbitrary point $x \in \mathbb{R}^d \setminus \mathbb{Q}^d$, there are infinitely many irreducible pairs of (p, q) satisfying*

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^{1+\frac{1}{d}}}$$

Proof. For each $Q > 1$ we pick a pair (p, q) satisfying (i) and (ii) in the theorem. Write (r, s) for its irreducible form (so $s \leq q$), we have

$$\left|x - \frac{r}{s}\right| = \left|x - \frac{p}{q}\right| \leq \frac{1}{Qq} < \frac{1}{q^{1+\frac{1}{d}}} \leq \frac{1}{s^{1+\frac{1}{d}}}$$

Whence the irreducible pairs we construct satisfy the approximation upper bound. It suffices to show they are infinitely many. Now it concludes from the fact that the values $|sx - r| \leq |qx - p| \leq \frac{1}{Q}$ can be chosen arbitrarily small while they are all positive since $x \in \mathbb{R}^d \setminus \mathbb{Q}^d$. □

The fundamental theorem of the study in diophantine approximation (on linear forms) is due to Siegel, also with proof by the pigeonhole principle.

Theorem 2.3 (Siegel's lemma, 1929)**.** *Given a system of linear equations ($N > M$)*

$$\begin{aligned} &a_{11}X_1 + \cdots + a_{1N}X_N = 0 \\ &\cdots \\ &a_{M1}X_1 + \cdots + a_{MN}X_N = 0 \end{aligned}$$

where $a_{mn} \in \mathbb{Z}$ and not all zero, $|a_{mn}| \leq B$ for $1 \leq m \leq M$, $1 \leq n \leq N$. Then the system admits a nonzero integral solution $(X_s) \in \mathbb{Z}^N$ with each component absolute value bounded by $C = [(NB)^{\frac{M}{N-M}}]$.

Proof. Suppose the contrary. Consider the image of the nonzero integral points (X_s) under the linear mapping A defined by the coefficients of the system, where $0 \leq X_s \leq C$, then the number of possible values of such image is no more than $(NBC+1)^M - 1$ because zero is not in the image by the contrary supposition and the j-th component of the image is in the interval $[-N_jBC, (N - N_j)BC]$ where $0 \leq N_j \leq N$ is the number of negative terms of coefficients a_{jk}. Now that we are given $(C+1)^N - 1$ nonzero integral points (X_s), since from the definition of C we have $(C+1)^N > (NBC+1)^M$, by the pigeonhole principle at least two of them have the same image, whence their difference vector which is nonzero integral point with component absolute value bounded by C, satisfies the system, contradiction. □

Remark 2.4. Bombieri and Vaaler (1983) gave a sharper bound $C^* = (D^{-1}\sqrt{det(AA^T)})^{\frac{1}{N-M}}$ whose proof is through the method in geometry of numbers similar to the proof of the fundamental theorem of number theory: finite class number theorem and Dirichlet's unit theorem of algebraic number ring $\mathcal{O}_K$. In our topic diophantine approximation of values of G-function at rational points we only need to use the original Siegel's lemma.

Note that the results in this section rely wholly on pigeonhole principle, hence have nothing to do with explicit construction problems. For example, in Theorem 2.3, it is hard to construct explicitly such an integral solution even when $M = 1$: where we need to find a nonzero integral point with all coordinates either 0 or ± 1 in a hyperplane in $\mathbb{R}^N$ of dim $(N-1)$ and has a rational-coordinate vector as its normal vector, when N is large enough. This is nontrivial from direct constructional viewpoint. From now on M and N will be used in a totally different notation system which has nothing to do with that used in Siegel's lemma.

3. Graded Padé Approximants

Next we are going to define and prove the exitence of the graded Padé approximants of a set of given G-functions in a class $G(\mathbb{Q}, C, \Phi)$. The word "graded" comes from the graded ring $A = k[a_1, \cdots, a_n] = \oplus_{i=0}^{\infty} A_i$ where A_i consists of homogeneous polynomials of degree i. For convenience we denote by $\overline{\mathrm{k}} = (k_1, k_2, \cdots, k_m) \in \mathbb{Z}^m$ the power to the multi-indeterminate $\overline{\mathrm{a}} = (a_1, a_2, \cdots, a_m)$. By convention we define $\overline{\mathrm{a}}^{\overline{\mathrm{k}}} = \prod_{j=1}^{m} a_j^{k_j}$ and $|\,\overline{\mathrm{k}}\,| = \sum_{j=1}^{m} k_j$.

For an arbitrary natural number N, we define index set $\Omega = \Omega(m, N) = \{\overline{\mathrm{k}} : k_j \geq 0, \forall j, \quad |\,\overline{\mathrm{k}}\,| \in \{N-1, N\}\}$ and $\Theta = \Theta(m, N) = \{\overline{\mathrm{k}} : k_j \geq 0, \forall j, \quad |\,\overline{\mathrm{k}}\,| = N\}$. In this paper we focus on the special case where $m = 2$, so in particular we have $\omega = |\,\Omega\,| = 2N+1$ and $\theta = |\,\Theta\,| = N+1$. From now on we fix a set of G-functions $f_j(z)$ stated in Theorem 1.10, and we consider the graded form of degree N with functional coefficients and variables $\overline{\mathrm{a}}$:

$$R(z; \overline{\mathrm{a}}) = \sum_{\overline{\mathrm{s}} \in \Theta} \overline{\mathrm{a}}^{\overline{\mathrm{s}}}\, P_{\overline{\mathrm{s}}}(z) + \sum_{\overline{\mathrm{k}} \in \Omega \setminus \Theta} \overline{\mathrm{a}}^{\overline{\mathrm{k}}}\, P_{\overline{\mathrm{k}}}(z) \sum_{j=1}^{2} a_j f_j(z)$$

where $P_{\overline{\mathrm{s}}}, P_{\overline{\mathrm{k}}} \in \mathbb{C}(z)$. If we denote this homogeneous form by $\sum_{\overline{\mathrm{s}} \in \Theta} \overline{\mathrm{a}}^{\overline{\mathrm{s}}}\, R_{\overline{\mathrm{s}}}(z)$, comparing the coefficients of each $\overline{\mathrm{a}}^{\overline{\mathrm{s}}}$, we have

$$R_{\overline{\mathrm{s}}}(z) = P_{\overline{\mathrm{s}}} + \sum_{j=1}^{2} P_{\overline{\mathrm{s}} - \overline{\mathrm{e_j}}}(z) f_j(z).$$

where we adopt the convention of summation that if $\overline{\mathrm{s}} - \overline{\mathrm{e_j}} \notin \Omega \setminus \Theta$ in the index, then this particular summand is defined to be zero.

Definition 3.1 (Graded Padé approximants)**.** Let $f_1(z)$ and $f_2(z)$ be G-functions in the class $G(\mathbb{Q}, C, \Phi)$ defined in Definition 1.6. For a pair of natural numbers (N, M), we should denote the index set Ω and Θ discussed just above for N and $m = 2$. For an arbitrary positive constant $\varepsilon > 0$ we define the constant $C_2 = (C\,\Phi)^{\frac{\omega(\omega-\theta)}{\varepsilon\,\theta}} = (C\,\Phi)^{\frac{1}{\varepsilon} N (2 - \frac{1}{N+1})}$ which shall be used in the following lemma.

We say that a set of polynomials $P_{\overline{\mathrm{k}}}(z) \in \mathbb{Q}[z]$ with index $\overline{\mathrm{k}} \in \Omega$ is a set of graded Padé approximants of strength (N, M, ε) if and only it satisfies:

(i) $P_{\overline{\mathrm{k}}}(z) \in \mathbb{Q}[z]$ are not all identically zero function;

(ii) $\deg\big(P_{\overline{\mathrm{k}}}(z)\big) < M$ for all $\overline{\mathrm{k}} \in \Omega$;

(iii)
$$\mathrm{ord}_{z=0}\big(R_{\overline{\mathrm{s}}}(z)\big) \geq K = K(N, M, \varepsilon) = \left\lfloor \frac{(\omega - \varepsilon)M}{\theta} \right\rfloor = \left\lfloor \big(2 - \frac{1}{N+1} - \frac{\varepsilon}{\theta}\big) M \right\rfloor,$$

for all $\overline{\mathrm{s}} \in \Theta$, where

$$R_{\overline{\mathrm{s}}}(z) = P_{\overline{\mathrm{s}}}(z) + \sum_{j=1}^{2} P_{\overline{\mathrm{s}} - \overline{\mathrm{e_j}}}(z) f_j(z).$$

Remark 3.2. Note that this definition could be applied more generally. In fact, from the proof of next Lemma 3.3 we know clearly how to construct a set of graded Padé approximants of strength (N, M, ε) and it always exists for any arbitrary functions $f_1(z)$ and $f_2(z)$ with rational Taylor coefficients. But the point is that only graded Padé approximants are not interesting enough and will not help much, while the argument of proving the main theorem needs control of their heights, which is valid especially when we talk about G-functions in the class $G(\mathbb{Q}, C, \Phi)$, as stated in the next lemma. For this reason, we can think of this definition exclusively for them. On the other hand, graded Padé approximants can be applied to E-functions and other kinds of interesting functions when the upper bound of heights is modified.

Lemma 3.3. *Suppose $C\,\Phi > 1$. For any fixed pair (N, ε) in Definition 3.1 and another arbitrary $\varepsilon_1 > 0$, there is a constant $M_1 = M_1(N, \varepsilon_1)$, such that for any natural number $M > M_1$, there exists a set of graded Padé approximants $\{P_{\overline{\mathrm{k}}}(z)\}$, where $\overline{\mathrm{k}} \in \Omega$, of strength (N, M, ε) and with $P_{\overline{\mathrm{k}}}(z) \in \mathbb{Z}[z]$ for all $\overline{\mathrm{k}} \in \Omega \setminus \Theta$ satisfying the following height bounds:*

$$\mathrm{H}\big(P_{\overline{\mathrm{k}}}(z)\big) < C_2^{(1+\frac{\varepsilon_1}{2})M},$$

for all $\overline{\mathrm{k}} \in \Omega \setminus \Theta$, where C_2 is defined in Definition 3.1.

Proof. Let's consider the problem as finding proper coefficients for the polynomials

$$P_{\overline{\mathrm{k}}}(z) = \sum_{n=0}^{M-1} P_{\overline{\mathrm{k}},n} z^n$$

for all $\overline{\mathrm{k}} \in \Omega$. First we will construct the rational coefficients $P_{\overline{\mathrm{k}},n}$ in such a way that $P_{\overline{\mathrm{k}}}(z)$ indeed form a set of Padé approximants of of strength (N, M, ε) of $f_1(z) = \sum_{n=0}^{\infty} f_{1,n} z^n$ and $f_2(z) = \sum_{n=0}^{\infty} f_{2,n} z^n$. Now we have that for all $\overline{\mathrm{s}} \in \Theta$,

$$\begin{aligned}
R_{\overline{\mathrm{s}}}(z) &= P_{\overline{\mathrm{s}}} + \sum_{j=1}^{2} P_{\overline{\mathrm{s}}-\overline{\mathrm{e_j}}}(z) f_j(z) \\
&= \sum_{l=0}^{M-1} P_{\overline{\mathrm{s}},l} z^l + \sum_{j=1}^{2} \sum_{l=0}^{\infty} \sum_{n=0}^{\min(l,M-1)} P_{\overline{\mathrm{s}}-\overline{\mathrm{e_j}},n} f_{j,l-n} z^l \\
&= \sum_{l=0}^{M-1} \Big(P_{\overline{\mathrm{s}},l} + \sum_{j=1}^{2} \sum_{n=0}^{l} P_{\overline{\mathrm{s}}-\overline{\mathrm{e_j}},n} f_{j,l-n} \Big) z^l + \sum_{l=M}^{K-1} \Big(\sum_{j=1}^{2} \sum_{n=0}^{M-1} P_{\overline{\mathrm{s}}-\overline{\mathrm{e_j}},n} f_{j,l-n} \Big) z^l \\
&\quad + \sum_{l=K}^{\infty} \Big(\sum_{j=1}^{2} \sum_{n=0}^{M-1} P_{\overline{\mathrm{s}}-\overline{\mathrm{e_j}},n} f_{j,l-n} \Big) z^l .
\end{aligned} \tag{3.1}$$

Note that in order to find $P_{\overline{\mathrm{k}},n} \in \mathbb{Q}$ such that $\operatorname{ord}_{z=0} R_{\overline{\mathrm{s}}}(z) \geq K$, it suffices to solve the equation for all $l = M, M+1, \cdots, K-1$:

$$\sum_{j=1}^{2} \sum_{n=0}^{M-1} P_{\overline{\mathrm{s}}-\overline{\mathrm{e_j}},n} f_{j,l-n} = 0 \tag{3.2}$$

and also set

$$P_{\overline{\mathrm{s}},l} = -\sum_{j=1}^{2} \sum_{n=0}^{l} P_{\overline{\mathrm{s}}-\overline{\mathrm{e_j}},n} f_{j,l-n} \tag{3.3}$$

if we get rational solutions $P_{\overline{\mathrm{s}}-\overline{\mathrm{e_j}},n}$ to Eq. (3.2), where $\overline{\mathrm{s}} - \overline{\mathrm{e_j}} \in \Omega \setminus \Theta$, then automatically we have $P_{\overline{\mathrm{s}},l} \in \mathbb{Q}$ for all $\overline{\mathrm{s}} \in \Theta$. However, Eq. (3.2) is a homogeneous linear equations with rational coefficients $f_{j,l-n}$ and unknowns $P_{\overline{\mathrm{s}}-\overline{\mathrm{e_j}},n}$. We can even expect to obtain integral solutions with well-controlled maximum value with the help of Siegel's lemma.

First recall the numbers d_n defined in Definition 1.6, here we just pick up d_K and multiply Eq. (3.2) by it, then we get a system of linear equations with integer coefficients

$$\sum_{j=1}^{2} \sum_{n=0}^{M-1} \Big(d_K f_{j,l-n} \Big) P_{\overline{\mathrm{s}}-\overline{\mathrm{e_j}},n} = 0 \tag{3.4}$$

for all $l = M, M+1, \cdots, K-1$ and for all $\overline{\mathrm{s}} \in \Theta$, so there are $\theta(K-M)$ equations altogether with $(\omega - \theta)M$ unknowns $P_{\overline{\mathrm{s}}-\overline{\mathrm{e_j}},n}$. Furthermore, the maximal absolute values of the coefficients of this system is bounded again, from Definition 1.6, by a function asymptotically in the class $O\Big((C\,\Phi)^{(1+\epsilon)K}\Big)$ as $K \to \infty$, where ϵ can be any positive real number. Let $\epsilon = \frac{\varepsilon_1}{4}$ where ε_1 is the given positive real number in the lemma, now we know all the coefficients are bounded by $\gamma(C\,\Phi)^{(1+\epsilon)K}$, where $\gamma = \gamma(\varepsilon_1)$ is a constant dependent only on ε_1.

To find $P_{\overline{\mathrm{k}}}(z) \in \mathbb{Z}[z]$ for all $\overline{\mathrm{k}} \in \Omega \setminus \Theta$ together with $P_{\overline{\mathrm{s}}}(z) \in \mathbb{Q}[z]$ for all $\overline{\mathrm{s}} \in \Theta$ already set in equality 3.3, satisfying the lemma, it suffices to find integral solutions to Eq. (3.4) with maximal absolute value $C_2^{(1+\frac{\varepsilon_1}{2})M}$. To conclude, we apply Siegel's lemma to Eq. (3.4) and obtain

$$\mathrm{H}\left(P_{\overline{\mathrm{k}}}(z)\right) \leq \Big((\omega - \theta)M\gamma(C\,\Phi)^{(1+\frac{\varepsilon_1}{4})K}\Big)^{\frac{\theta(K-M)}{(\omega-\theta)M-\theta(K-M)}}$$

To show the latter is less than $C_2^{(1+\frac{\varepsilon_1}{2})M}$ when M is large enough, it suffices to show

$$\Big((\omega-\theta)M\gamma\Big)^{\frac{\theta(K-M)}{(\omega-\theta)M-\theta(K-M)}} < (C\,\Phi)^{(1+\frac{\varepsilon_1}{2})\frac{\omega M}{\theta}\frac{\omega-\theta}{\varepsilon}-\frac{\theta(K-M)}{(\omega-\theta)M-\theta(K-M)}(1+\frac{\varepsilon_1}{4})K}$$

which is equivalent to

$$\gamma(\omega-\theta)M < (C\,\Phi)^{(1+\frac{\varepsilon_1}{2})\frac{\omega M}{\theta}\frac{\omega-\theta}{\varepsilon}\frac{(\omega-\theta)M-\theta(K-M)}{\theta(K-M)}-(1+\frac{\varepsilon_1}{4})K} \tag{3.5}$$

when M is large enough. Since

$$K-M=\left\lfloor\frac{(\omega-\varepsilon)M}{\theta}-M\right\rfloor=\left\lfloor\frac{(\omega-\theta-\varepsilon)M}{\theta}\right\rfloor,$$

we have $\theta(K-M)\le(\omega-\theta-\varepsilon)M$ which leads to the inequalities

$$\frac{(\omega-\theta)M-\theta(K-M)}{\theta(K-M)}\ge\frac{(\omega-\theta)M-(\omega-\theta-\varepsilon)M}{(\omega-\theta-\varepsilon)M}=\frac{\varepsilon}{\omega-\theta-\varepsilon}>\frac{\varepsilon}{\omega-\theta}.$$

Now we note that the exponential part appeared in the inequality 3.5 is larger than

$$(1+\frac{\varepsilon_1}{2})\frac{\omega M}{\theta}-(1+\frac{\varepsilon_1}{4})K,$$

and since $K=\left\lfloor\frac{(\omega-\varepsilon)M}{\theta}\right\rfloor\le\frac{(\omega-\varepsilon)M}{\theta}$, the exponent is finally larger than $\frac{M}{\theta}\Big(\frac{\varepsilon_1}{4}+\varepsilon(1+\frac{\varepsilon_1}{4})\Big)>\frac{M\,\varepsilon_1}{4\,\theta}$. To satisfy the inequality 3.5, it suffices to satisfy the following inequality

$$\gamma NM\le(C\,\Phi)^{\frac{M\,\varepsilon_1}{4\,\theta}}$$

which is obviously true for $M>M_1$ where $M_1=M_1(N,\varepsilon_1)$ is a constant. M_1 is not explicitly constructed, since γ is not explicit.

We are done of the proof of this lemma except the case $C=\Phi=1$, which should be excluded from our discussion since it will cause a contradiction: on the one hand the heights of these Padé approximants could only be zero under this assumption, on the other hand they can not all be zero polynomials.

□

Remark 3.4. From the proof of Lemma 3.3 we see that by modifying the proof and the lemma a little bit, we can also obtain nice control of the heights of Padé approximants of some class of functions $f_j(z)$, if these functions could have a good type of upper bound for the increase of the $d_K f_{j,l-n}$ in Eq. (3.4). For G-functions the type is exponential, and we expect this type to be greater than polynomial, and the type of final control we get of the heights is the same as that of $d_K f_{j,l-n}$ in Eq. (3.4). That is to say, for this purpose G-functions are of the very typical meaning when realizing an exponential type of control of the heights, which is crucial in the proof of the main theorem. On the other hand, we can not say that graded Padé approximants are totally designed for G-functions. In fact, in Zudilin's another paper [4] he proved a similar result on the measure of irrationality for the values of E-functions at rational points, and the technique used in this paper is also relied on a lemma similar to Lemma 3.3.

Next we can talk about differential operator acting on the graded form $R(z;\overline{\mathrm{a}})$. Suppose all the assumptions in the main theorem. In the following discussion, the reader can replace the field $\mathbb{C}(z)$ with the ring $\mathbb{C}[z]$ freely and the argument still holds, so we adopt the notion "$\mathbb{C}(z)$-module" here instead of "$\mathbb{C}(z)$- linear space". Denote by A the $\mathbb{C}(z)$-module generated by $\overline{\mathrm{a}}^{\overline{\mathrm{s}}}$ for all $\overline{\mathrm{s}}\in\Theta$ and $\overline{\mathrm{a}}^{\overline{\mathrm{k}}}\sum_{j=1}^{2}a_jf_j(z)$ for all $\overline{\mathrm{k}}\in\Omega\setminus\Theta$. So a particular graded form $R(z;\overline{\mathrm{a}})$ is nothing but an element in A. Note also that A is a free module and the generators form a basis, since 1, $f_1(z)$ and $f_2(z)$ are

$\mathbb{C}(z)$-linearly independent in the assumption of the main theorem. Now let us fix a system of linear differential equations in the main theorem, which has appeared as early as in Eq. (1.2), and construct a differential operator D associated to it, that is,

$$D = \frac{\partial}{\partial z} - \sum_{j=1}^{2} \Big(\sum_{l=1}^{2} Q_{lj}(z) a_l \Big) \frac{\partial}{\partial a_j}.$$

It is clear that $\mathbb{C}(z)$ itself is closed under the action of D, and D obeys Leibniz's rule, namely $D(u \cdot v) = Du \cdot v + u \cdot Dv$. With these two facts we claim that A is actually a $\mathbb{C}(z)[D]$-module. To check this conclusion, it suffices to check that the image of the generators $\bar{\mathrm{a}}^{\bar{\mathrm{s}}}$ and $\bar{\mathrm{a}}^{\bar{\mathrm{k}}} \sum_{j=1}^{2} a_j f_j(z)$ are still in A under the action of D. By the Leibniz's rule and Eq. (1.2), we're done of this routine and now it is natural to talk about $\mathcal{R}^{[n]}(z;\bar{\mathrm{a}}) = D^n R(z;\bar{\mathrm{a}})$ in the following way Eq. (3.6), where $\mathcal{P}^{[n]}_{\bar{\mathrm{s}}}(z)$ for all $\bar{\mathrm{s}} \in \Theta$ and $\mathcal{P}^{[n]}_{\bar{\mathrm{k}}}(z)$ for all $\bar{\mathrm{k}} \in \Omega \setminus \Theta$ are well defined as the $\mathbb{C}(z)$-coefficients of corresponding generators appeared the element $\mathcal{R}^{[n]}(z;\bar{\mathrm{a}})$ in A, since the generators of A is actually a basis.

$$\begin{aligned} \mathcal{R}^{[n]}(z;\bar{\mathrm{a}}) &= D^n R(z;\bar{\mathrm{a}}) \\ &= D^n \Big(\sum_{\bar{\mathrm{s}} \in \Theta} \bar{\mathrm{a}}^{\bar{\mathrm{s}}} P_{\bar{\mathrm{s}}}(z) + \sum_{\bar{\mathrm{k}} \in \Omega \setminus \Theta} \bar{\mathrm{a}}^{\bar{\mathrm{k}}} P_{\bar{\mathrm{k}}}(z) \sum_{j=1}^{2} a_j f_j(z) \Big) \\ &= \sum_{\bar{\mathrm{s}} \in \Theta} \bar{\mathrm{a}}^{\bar{\mathrm{s}}} \mathcal{P}^{[n]}_{\bar{\mathrm{s}}}(z) + \sum_{\bar{\mathrm{k}} \in \Omega \setminus \Theta} \bar{\mathrm{a}}^{\bar{\mathrm{k}}} \mathcal{P}^{[n]}_{\bar{\mathrm{k}}}(z) \sum_{j=1}^{2} a_j f_j(z) \\ &= \sum_{\bar{\mathrm{s}} \in \Theta} \bar{\mathrm{a}}^{\bar{\mathrm{s}}} \mathcal{R}^{[n]}_{\bar{\mathrm{s}}}(z), \end{aligned} \tag{3.6}$$

for all $n \in \mathbb{N}$. Then we have the recurrence relations directly from this definition, and we shall use these notations followed from the condition of main theorem in this whole paper.

Lemma 3.5.

$$\begin{aligned} \mathcal{P}^{[n+1]}_{\bar{\mathrm{k}}}(z) =& \frac{\mathrm{d}}{\mathrm{d}z} \mathcal{P}^{[n]}_{\bar{\mathrm{k}}}(z) - \sum_{l,j=1}^{2} (k_j - \delta_{lj} + 1) Q_{lj}(z) \mathcal{P}^{[n]}_{\bar{\mathrm{k}} - \overline{e_l} + \overline{\mathrm{e_j}}}(z) \\ & + (|\bar{\mathrm{k}}| - N + 1) \sum_{l=1}^{2} Q_{l0}(z) \mathcal{P}^{[n]}_{\bar{\mathrm{k}} - \overline{e_l}}(z), \\ & \forall \bar{\mathrm{k}} \in \Omega, \forall n \in \mathbb{N}; \\ \mathcal{R}^{[n+1]}_{\bar{\mathrm{s}}}(z) =& \frac{\mathrm{d}}{\mathrm{d}z} \mathcal{R}^{[n]}_{\bar{\mathrm{s}}}(z) - \sum_{l,j=1}^{2} (s_j - \delta_{lj} + 1) Q_{lj}(z) \mathcal{R}^{[n]}_{\bar{\mathrm{s}} - \overline{e_l} + \overline{\mathrm{e_j}}}(z), \\ & \forall \bar{\mathrm{s}} \in \Theta, \forall n \in \mathbb{N}. \end{aligned} \tag{3.7}$$

Proof. Immediate from the definition Eq. (3.6). □

Now we come to define the most important polynomials in this paper, let

$$P^{[n]}_{\bar{\mathrm{k}}}(z) = \frac{T^n(z)}{n!} \mathcal{P}^{[n]}_{\bar{\mathrm{k}}}(z)$$

for all $\bar{\mathrm{k}} \in \Omega$ and $n \in \mathbb{N}$, and let

$$R^{[n]}_{\bar{\mathrm{s}}}(z) = \frac{T^n(z)}{n!} \mathcal{R}^{[n]}_{\bar{\mathrm{s}}}(z)$$

for all $\bar{\mathrm{s}} \in \Theta$ and $n \in \mathbb{N}$.

Lemma 3.6. *$P_{\bar{k}}^{[n]}(z) \in \mathbb{Q}[z]$ and $\deg\left(P_{\bar{k}}^{[n]}(z)\right) \le M + tn - 1$; $\mathrm{ord}_{z=0}\left(R_{\bar{s}}^{[n]}(z)\right) \ge K - n$ where K is defined in Definition 3.1.*

Proof. By Eq. (3.7) we have a recurrence formula for $T^n(z)\mathcal{P}_{\bar{k}}^{[n]}(z)$:

$$\begin{aligned} T^{n+1}(z)\mathcal{P}_{\bar{k}}^{[n+1]}(z) =& T(z)\frac{\mathrm{d}}{\mathrm{d}z}\left(T^n(z)\mathcal{P}_{\bar{k}}^{[n]}(z)\right) - nT^{(1)}(z)T^n(z)\mathcal{P}_{\bar{k}}^{[n]}(z) \\ &- \sum_{l,j=1}^{2}(k_j - \delta_{lj} + 1)T(z)Q_{lj}(z)T^n(z)\mathcal{P}_{\bar{k}-\overline{e_l}+\overline{e_j}}^{[n]}(z) \\ &+ (|\bar{k}| - N + 1)\sum_{l=1}^{2} T(z)Q_{l0}(z)T^n(z)\mathcal{P}_{\bar{k}-\overline{e_l}}^{[n]}(z), \end{aligned} \tag{3.8}$$

for all $\bar{k} \in \Omega$ and $n \in \mathbb{N}$. Now from Definition 3.1 we know for all $\bar{k} \in \Omega$, $\mathcal{P}_{\bar{k}}^{[0]}(z) = P_{\bar{k}}^{[n]}(z) \in \mathbb{Q}[z]$ is of degree at most $M - 1$. We can easily show by induction on $n \in \mathbb{N}$ that $P_{\bar{k}}^{[n]}(z) = \frac{T^n(z)}{n!}\mathcal{P}_{\bar{k}}^{[n]}(z)$ are also in $\mathbb{Q}[z]$ and of degree at most $M + tn - 1$, since by each step the degree increases at most t. For the second half of the claim, we use Eq. (3.7) and $\mathcal{R}_{\bar{s}}^{[0]}(z) = R_{\bar{s}}(z)$ have order of zero at $z = 0$ at least K, for all $\bar{s} \in \Theta$, by induction we can easily show that $\mathrm{ord}_{z=0}\left(\mathcal{R}_{\bar{s}}^{[n]}(z)\right) \ge K - n$, whence $\mathrm{ord}_{z=0}\left(R_{\bar{s}}^{[n]}(z)\right) \ge K - n$. □

4. Siegel-Shidlovskii Method

In Shidlovskii's book [1] there is a linear algebra lemma which is crucial for the proof of lemma 2.1-2.4 in Zudilin's paper [5]. In this section we concentrate on the study of this lemma and some of its related results to prepare for the proof of Zudilin's lemmas in the next section. In this section we first recall the Padé approximants of the first kind, then study Siegel-Shidlovskii method on it, and apply this method to the graded Padé approximants discussed in section 3 to understand Zudilin's method in his paper and the merit of using graded Padé approximants over that of the first kind when we treat problems in Diophantine approximation, especially why is graded Padé approximants so efficient in the theory of G-functions.

From Lemma 4.1 to Lemma 4.6 we adopt the notations in Shidlovskii's book [1] which is somehow classic but not related to the notations we used throughout the mémoire. They are original methods for proving Theorem 1.3, but our target is Theorem 1.10, so we need to modify it for our application in this section. From Lemma 4.8 to the end we come back to the notations we used for this mémoire.

Lemma 4.1 (Bounds for order of Padé approximants of the first kind)**.** *Given m analytic functions around 0, $f_j(z)$, $j = 1, 2, \cdots, m$, and let $n \in \mathbb{Z}_{\ge 1}$. Consider a nonzero linear form*

$$R(z) = \sum_{j=1}^{m} P_j f_j(z)$$

where $P_j(z) \in \mathbb{C}[z]$, $\deg P_j \le n$ with undetermined coefficients. Then there exists a constant $N = N(n, f_j)$ such that either $R(z) = 0$ or

$$\mathrm{ord}R(z) \le N$$

Furthermore, there is a nonzero linear form $R(z)$ satisfying $\deg P_j \le n$ and

$$\mathrm{ord}R(z) \ge m(n+1) - 1$$

Such linear form (not necessarily unique) is called Padé approximants of the first kind of the given analytic functions $f_j(z)$, $j = 1, 2, \cdots, m$.

Proof. The lower bound of the order of the nonzero linear form (when it is a Padé approximation) is obtained by comparing the number of indetermined coefficients and the order, in a system of homogeneous linear equations. The upper bound is easily obtained when P_j are all complex numbers by comparing the dimension of the linear space formed by such linear forms (which is m) and the fact that linear forms of distinct orders are linearly independent. To pass to the polynomial case, it suffices to apply the first step to $m(n+1)$ functions $f_{jk}(z) = z^k f_j(z)$ over $\mathbb{C}$ where $k = 0, 1, \cdots, n$, we are done. □

Remark 4.2. I believe that the upper bound is not efficient in view of the fact that there is no accurate description of the rank of the number matrix made by the coefficients vectors and their shift vectors of these given analytic functions around 0. For inhomogeneous linear forms, Zudilin's lemma 1.3 is on the construction of such a linear form with order as large as possible while the heights for the polynomials could be controlled, so we must use essentially Siegel's lemma.

Next we shall consider more generally linear forms with rational function coefficients. It is a convention that a rational function in $\mathbb{C}(z)$ is represented by coprime polynomials in the numerator and denominator, and we call the degree of it to be the sum of the degrees of the numerator and denominator, so we have the degree of the product/quotient of two rational functions does not exceed the sum of the degrees of these two functions.

Lemma 4.3 (Shidlovskii). *Let $f_1(z), \cdots, f_s(z)$ and $g_1(z), \cdots, g_m(z)$ be analytic functions around* 0 *and at least one of g_j is nonzero function. Then there exists a constant $N \in \mathbb{N}$ depending on these functions such that for all $\alpha_i, \beta_j \in \mathbb{C}$ such that $\sum_{j=1}^m \beta_j g_j(z)$ is a nonzero function, if*

$$h(z) = \frac{\sum_{i=1}^s \alpha_i f_i(z)}{\sum_{j=1}^m \beta_j g_j(z)} \in \mathbb{C}(z)$$

Then $\deg h \leq N$.

Proof. Suppose the $(s+m)$-tuple of functions $f_1(z), \cdots, f_s(z)$ and $g_1(z), \cdots, g_m(z)$ has $\mathbb{C}(z)$-rank r, and pick among them r linearly independent functions over $\mathbb{C}(z)$, say $h_k(z)$, $k = 1, 2, \cdots, r$. After rewriting the other functions as linear expansions of this basis, we simplify the expression into

$$h(z) = \frac{\sum_{k=1}^r A_k(z) h_k(z)}{\sum_{k=1}^r B_k(z) h_k(z)}$$

where $A_k(z)$, $B_k(z) \in \mathbb{C}(z)$ and the degree of them are bounded by a constant. Writing in $\mathbb{C}(z)$-linear combination of $h_k(z)$, we have the following equation

$$\sum_{k=1}^r (A_k(z) - h(z) B_k(z)) h_k(z) = 0$$

whence $A_k(z) = h(z) B_k(z)$ for all k. Moreover, since $B_k(z)$ are not all zero functions, we can pick one j such that $B_j(z)$ is nonzero function and hence $h(z) = \frac{A_j(z)}{B_j(z)}$ and the degree of $h(z)$ is bounded by a constant. □

Next we shall begin to consider linear forms in functions y_j which satisfy a certain system of linear ODEs. We first consider a homogeneous system,

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}z} y_l(z) = \sum_{j=1}^m Q_{lj}(z) y_j(z), \quad \forall l = 1, 2, \cdots, m, \\ &Q_{lj}(z) \in \mathbb{C}(z), \qquad\qquad\qquad\qquad \forall l, j. \end{aligned} \tag{4.1}$$

The polynomial $T(z)$ (least common denominator) has the same meaning as before. Define a differential operator

$$D = \frac{\partial}{\partial z} + \sum_{l=1}^{m} (\sum_{j=1}^{m} Q_{lj} y_j) \frac{\partial}{\partial y_l}$$

Then it is clear to see that the map of D followed by multiplication of T sends a linear form of y_j over $\mathbb{C}[z]$ to the same module. Let y_j $(j = 1, \cdots, m)$ be a solution of the system, and view the linear form

$$R(z) = \sum_{j=1}^{m} P_j y_j(z)$$

as a function of z, we then have $DR(z) = \frac{\mathrm{d}}{\mathrm{d}z} R(z)$ from the definition of D. Given an arbitrary linear form $R_1(z) = \sum_j P_{1j}(z) y_j(z)$ where $P_{1j}(z) \in \mathbb{C}[z]$, one can define recursively

$$R_k(z) = (T(z)D) R_{k-1}(z) = \sum_j P_{kj}(z) y_j(z) \tag{4.2}$$

(hence $P_{kj} \in \mathbb{C}[z]$). we have thus a recursive formula:

$$P_{kj} = T(\frac{d}{dz} P_{(k-1)j} + \sum_{l=1}^{m} P_{(k-1)l} Q_{lj})$$

Also we know if $y_j(z)$ is a solution of the ODEs then we can write $R_k(z) = T(z) R_{k-1}(z)$.

Lemma 4.4. *Let $R_1(z) = \sum_j P_{1j}(z) y_j(z)$ be an arbitrary linear form where $P_{1j} \in \mathbb{C}[z]$, $R_1(z), \cdots, R_l(z)$ defined as in 4.2. Suppose l is the maximal number such that $R_1(z), \cdots, R_l(z)$ are linearly independent over $\mathbb{C}[z]$, then the $\mathbb{C}[z]$-rank of $\{R_1(z), R_2(z), \cdots, \}$ is exactly l, as rank of $\mathbb{C}[z]$-module.*

Proof. From the assumption of the lemma, we can write $R_{l+1}(z)$ as a linear expansion of $R_1(z), R_2(z), \cdots, R_l(z)$. Apply $T(z)D$ to both sides $k-1$ times respectively, we can express $R_{l+k}(z)$ as a $\mathbb{C}[z]$-linear expansion of $R_1(z), \cdots, R_l(z)$. □

Then we pass to the discussion on determinant. Given $R_1(z)$, we can construct $R_k(z) = \sum_{j=1}^{m} P_{kj}(z) y_j(z)$ and

$$\Delta(z) = \det(P_{kj}(z))_{k=1,\cdots,m, j=1,\cdots,m.}$$

The determinant $\Delta(z)$ is useful for discussing the linear dependence of $R_1(z), \cdots, R_m(z)$. In fact, if there is a nonzero solution $(y_1(z), \cdots, y_m(z))$ of the differential system such that $R_1(z) = 0$, then we have $R_k(z) = 0$ for all k, since $\Delta y_j(z) = \sum_{k=1}^{m} \Delta_{kj}(z) R_k(z)$ where $\Delta_{kj}(z)$ is the cofactor of $P_{kj}(z)$ in the matrix, pick one nonzero $y_j(z)$ we have $\Delta(z) = 0$. By the definition of the determinant, we know that $R_1(z), \cdots, R_m(z)$ are linearly dependent over $\mathbb{C}(z)$. Conversely, we have the following more general lemma.

Lemma 4.5. *Let $R_1(z) = \sum_j P_{1j}(z) y_j(z)$ be an arbitrary linear form where $P_{1j}(z) \in \mathbb{C}[z]$, $R_1(z), \cdots, R_l(z)$ defined as in 4.2. Suppose $\{R_1(z), \cdots, R_m(z)\}$ has $\mathbb{C}(z)$-rank l, where $1 \leq l < m$, then there are at most and as many as $(m-l)$ $\mathbb{C}(z)$-linearly independent solutions of the differential system, say $y_{1s}(z), \cdots, y_{ms}(z)$ where $s = 1, \cdots, m-l$ such that $R_1(z) = 0$.*

Proof. From lemma 4.4 we know that $R_1(z), \cdots, R_l(z)$ are linearly independent while $R_1(z), \cdots, R_{l+1}(z)$ are linearly dependent, so there are $A_1(z) \cdots, A_{l+1}(z) \in \mathbb{C}[z]$ such that $A_{l+1}(z) \neq 0$ and

$$A_1(z) R_1(z) + \cdots + A_{l+1}(z) R_{l+l}(z) = 0$$

Denote by U the linear space over $\mathbb{C}$ formed by all the solutions $(y_1(z), \cdots, y_m(z))$ of the differential system. Denote by V the $\mathbb{C}$-linear space formed by all the solutions of the linear ODE:

$$A_1(z) y(z) + A_2(z) T(z) y^{(1)}(z) + A_3(z) T^2(z) y^{(2)}(z) + \cdots + A_{l+1}(z) T^l(z) y^{(l)}(z) = 0$$

Now we can define a linear map φ from U to V: given $(y_1(z)\cdots,y_m(z))\in U$, let the image of it under φ be just $R_1(y_1,\cdots,y_m)$. It is rountine to check this is a linear map and sends elements in U to V. Moreover, one can check that this map is surjective. Note that dim $U=m$ and dim $V=l$ by the theorem of linear ODE solutions, hence dim kerφ is exactly $m-l$. □

Lemma 4.6. *Given m analytic functions around 0, $\{f_j(z)\}$, where $j=1,\cdots,m$ satisfying the homogeneous differential system 4.1, and linearly independent over $\mathbb{C}(z)$. Set*

$$p=\min_{1\le j\le m}\mathrm{ord}_{z=0}\,f_j(z)$$

and

$$q=\max\Big(\deg\big(T(z)\big),\max_{l,j}\deg\big(T(z)Q_{lj}(z)\big)\Big).$$

Let $R_1=\sum_{j=1}^m P_{1j}y_j$ be an arbitrary nonzero linear form with $degP_{1j}\le n,\forall j$. Let l be the rank of $\{R_k(z)\}_{k=1,2\cdots}$ defined above, then for $y_j(z)=f_j(z)$,where $j=1,2,\cdots,m$, we have

$$\mathrm{ord}_{z=0}\,R_1(z)\le ln+C_0$$

where $C_0=\frac{qm(m-1)}{2}+m+N_1-1$ and N_1 is an ineffective constant on the functions f_j. Furthermore, if $l=m$, we have C_0 replaced by a constant $C_1=\frac{qm(m-1)}{2}+m+p-1$ which is effective.

Proof. We complete this proof in seven steps.

Step 1. (Interpret the rank l condition by a linear expansion of $R_k(z)$ in l terms for each k.)

Since $\{R_1(z),R_2(z)\cdots\}$ has rank $l\le m$, by lemma 4.4 we know the matrix $(P_{kj}(z))_{k=1,\cdots,l;j=1,\cdots,m}$ has full row rank, without loss of certainty we can assume that the first l columns are linearly independent, then we have the linear expansion

$$P_{kj}(z)=\sum_{i=1}^{l}P_{ki}(z)A_{ij}(z),$$

$$\forall k=1,\cdots,l;j=l+1,\cdots,m.$$

This gives us a linear expansion for R_k :

$$R_k(z)=\sum_{i=1}^{l}P_{ki}(z)u_i(z)$$

where

$$u_i(z)=y_i(z)+\sum_{j=l+1}^{m}A_{ij}(z)y_j(z)$$

When $l=m$, we just take $u_i(z)=y_i(z)$. It's important to note here $A_{ij}(z)\in\mathbb{C}(z)$ is dependent only on the choice of the linear form $R_1(z)$ when $l<m$.

Step 2. (Dependence of $A_{ij}(z)$ on a fundamental set of solutions of the differential system which contains a fundamental set of solutions of $R_1(z)=0$.)

Fix an arbitrary linear form $R_1(z)$. By lemma 4.5, one can pick $(m-l)$ linearly independent solutions of the differential system, $(y_{is}(z))_{i=1,\cdots,m;s=1,\cdots,m-l}$, so that $R_1(z)=0$ is satisfied. From the principal of expanding a linearly independent set to a basis, one can get a set of fundamental solutions of the differential system containing them, say $(y_{is}(z))_{i=1,\cdots,m;s=1,\cdots,m}$ which is a non-degenerate square matrix. Denote

$$u_{is}(z)=y_{is}(z)+\sum_{j=l+1}^{m}A_{ij}(z)y_{js}(z)$$

We have by the choice of (y_{is}) that

$$R_{ks}(z) = \sum_{i=1}^{l} P_{ki}(z)u_{is}(z) = 0, \quad \forall s \in \{1, \cdots, m-l\}, k \in \{1, \cdots, l\}. \tag{4.3}$$

Next we shall see the dependence of $(A_{ij}(z))$ on $(y_{is}(z))$. For any fixed $s \in \{1, \cdots, m-l\}$, Since this is a system of linear equations with l unknowns $u_{is}(z), i = 1, \cdots, l$, with the coefficients matrix $\big(P_{ki}(z)\big)$ being a non-degenerate matrix over $\mathbb{C}(z)$, it has only trivial solution,i.e.

$$y_{is}(z) + \sum_{j=l+1}^{m} A_{ij}(z)y_{js}(z) = 0, \quad \forall i, s. \tag{4.4}$$

Now we fix i and view it as an inhomogeneous system of $(m-l)$ linear equations with $(m-l)$ unknowns $A_{ij}(z), j = l+1, \cdots m$. If

$$\lambda(z) = \det\big(y_{js}(z)\big)_{j=l+1,\cdots,m;s=1,\cdots,m-l} \neq 0$$

Then the system has a unique solution (A_{ij}) and we can say that (A_{ij}) is dependent only on the choice of (y_{is}).

Step 3. (Proof of $\lambda(z) \neq 0$) If the determinant $\lambda(z) = 0$, that is,

$$\lambda(z) = \det\big(y_{js}(z)\big)_{j=l+1,\cdots,m;s=1,\cdots,m-l} = 0$$

then we will get the determinant of the whole matrix $(y_{is})_{i=1,\cdots,m;s=1,\cdots,m}$ be a zero function of z, which is contradiction of our choice of y_{is} as a fundamental set of solutions of the differential system. To see this, add each j-th row ($j = l+1, \cdots, m$) of the whole matrix multiplied by A_{ij} to the i-th row for $1 \leq i \leq l$, after completing this operation for each i, we will get a matrix with the same determinant but has a $l \times (m-l)$ bloc of zeroes in the upper left corner and a $(m-l) \times (m-l)$ square bloc with determinant 0 in the lower left part complementary to the upper left corner. Now we apply the general version of Laplace expansion theorem (see below) to compute the determinant of the whole matrix, it is equal to zero. We are done. For the convenience of readers, we add the general statement of Laplace expansion theorem in linear algebra, whose proof could be found in most ordinary textbooks.

Let $A = [a_ij]$ be an $n \times n$ matrix and S the set of k-element subsets of $\{1, 2, \cdots, n\}$, $H \in S$. Then the determinant of A can be expanded along the k rows identified by H as follows:

$$\det A = \sum_{L \in S} \epsilon^{H,L} a_{H,L} c_{H,L}$$

where $\epsilon^{H,L}$ is the sign of the permutation defined by H and L, which is equal to $(-1)^{\sum_{h\in H} h + \sum_{l \in L} l}$, $a_{H,L}$ is the determinant of the square submatrix of A obtained by deleting from A rows and columns with indices in H and L respectively, and $c_{H,L} = a_{H',L'}$ is the complement of $a_{H,L}$ with H' and L' being the complement of H and L respectively.

Step 4. (Express explicitly A_{ij} in terms of y_{is} for $i = 1, \cdots, l; s = 1, \cdots, m-l$)

Now we shall express the dependence shown in step 2 and step 3 in an explicit formula. Consider the matrix $(y_{is})_{i=1,\cdots,l;s=1,\cdots,m-l}$ and replace the j-th row ($j = l+1, \cdots, m$)by the i-th row $i = 1, \cdots, l$, then we get a new matrix and denote the determinant of it by λ_{ij}. If we multiply the linear equations with unknowns A_{ij} and coefficients y_{is} by the adjoint matrix of (A_{ij}), we see that

$$A_{ij} = -\frac{\lambda_{ij}}{\lambda}$$

for any $i = 1, \cdots, l$ and $j = l+1, \cdots, m$. Till now our expressions for A_{ij} is still dependent on a particular choice of solutions for a fixed R_1. To show the whole lemma we need to construct a formula independent of the choice of R_1.

Step 5. (Formula of A_{ij} on a fixed fundamental set of solutions of the differential system)

Let (x_{ik}) $(i, k = 1, \cdots, m)$ be a fixed fundamental set of solutions to the differential system. Consider it as the whole matrix, then we have $m_1 = \binom{m}{m-l}$ minors of size $(m-l) \times (m-l)$ in the last $(m-l)$ rows. Denote them by $\varphi_1, \cdots, \varphi_{m_l}$ as functions of z. If we replace the j-th row $(j = l+1, \cdots, m)$ by the i-th row $(i = 1, \cdots, l)$ of the whole matrix, we will get another m_1 minors in the last $(m-l)$ rows, denote them by $\varphi_{ij1}, \cdots, \varphi_{ijm_1}$. Since both (x_{ik}) and (y_{is}) are fundamental solutions of the linear system, we can write

$$(y_{is}) = (x_{ik})(c_{ks})$$

as multiplication of matrix where $(c_{ks}) \in GL_m(\mathbb{C})$. Then we can compute λ and λ_{ij} by Binet-Cauchy formula respectively. Binet-Cauchy formula states that if A and B are $m \times n$ and $n \times m$ matrix respectively, then we have

$$\det(AB) = \sum_{S \in \binom{[n]}{m}} \det\left(A_{[m],S}\right) \det\left((B_{S,[m]}\right)$$

where $[n]$ refers to the set $\{1, 2, \cdots, n\}$, and with the convention that if $n < m$ then the sum is added over zero term, hence zero. Finally we get the formula for A_{ij}:

$$A_{ij}(z) = \frac{\sum_{k=1}^{m_1} \alpha_{ijk} \varphi_{ijk}(z)}{\sum_{k=1}^{m_1} \beta_k \varphi_k(z)}$$

where α_{ijk} and β_k are homogeneous forms of complex constant entries in the matrix (c_{ks}). Moreover $\sum_{k=1}^{m_1} \beta_k \varphi_k(z)$ is a nonzero function of z.

Step 6. (Upper bounds for the degree of A_{ij})

Since

$$A_{ij}(z) = \frac{\sum_{k=1}^{m_1} \alpha_{ijk} \varphi_{ijk}(z)}{\sum_{k=1}^{m_1} \beta_k \varphi_k(z)}$$

with nonzero denominator and $A_{ij}(z) \in \mathbb{C}(z)$, from lemma 4.3 we know that there is a constant N depending only on the analytic functions $\varphi_{ijk}(z)$ and $\varphi_k(z)$ such that $\deg A_{ij}(z) \leq N$. But these functions depend only on our fixed solutions $\big(x_{ik}(z)\big)$. When we take different $R_1(z)$, only the value of l and the corresponding minors would change, but there are only finitely many such minors of all kinds of sizes within the fixed matrix $\big(x_{ik}(z)\big)$, so eventually we have an upper bound of $\deg\big(A_{ij}(z)\big)$ only dependent on the differential system, which is determined also by the original analytic functions f_j where $j = 1, \cdots, m$. Hence there is a constant $N_0 = N_0(f_1(z), \cdots, f_m(z))$ such that $\deg\big(A_{ij}(z)\big) \leq N_0$. If we substitute f_j in place of y_j and write

$$u_i(z) = f_i(z) + \sum_{j=l+1}^{m} A_{ij}(z) f_j(z)$$

and then denote $T_i(z)$ as a least common denominator of $A_{ij}(z)$ for $j = l+1, \cdots, m$. (When $l = m$, set $T_i(z) = 1$). Denote by $H_i(z) = T_i(z) u_i(z)$ a linear form of $f_j(z)$ with coefficients in $\mathbb{C}[z]$. Clearly H_i is nonzero function due to the linearly independence condition of $f_j(z)$. Finally from the upper bound of $A_{ij}(z)$ we will get an upper bound for the order of the linear form $H_i(z)$. Since $\deg\big(T_i(z)\big)$, $\deg\big((T_i(z)A_{ij}(z)\big) \leq \sum_{j=l+1}^{m} \deg\big(A_{ij}(z)\big) \leq (m-l)N_0$, by Lemma 4.1, there exists a constant N_1, dependent only on the functions $f_j(z)$, such that $\operatorname{ord}\big(H_i(z)\big) \leq N_1, \forall i$. Also note that when $l = m$, we have $\min_i \operatorname{ord}_{z=0}(H_i(z)) = p$.

Step 7. (Conclusion of the proof by passing to the order of the determinant of (P_{ki}))

Now we consider $\Delta_0(z) = \det(P_{ki}(z))$ again. Let $\Delta_{ki}(z)$ denote the cofactor of $P_{ki}(z)$ in the determinant. Then we have immediately $\Delta_0(z)u_j(z) = \sum_{k=1}^{l} \Delta_{kj}(z)P_k(z)$ for all $j = 1, \cdots, l$. Multiplying both sides by $T_j(z)$ we get

$$\Delta_0(z)H_j(z) = T_j(z)\sum_{k=1}^{l} \Delta_{kj}(z)P_k(z)$$

On the other hand, from the recursive formula

$$P_{ki}(z) = T(z)\Big(\frac{\mathrm{d}}{\mathrm{d}z}P_{(k-1)i}(z) + \sum_{j=1}^{m} P_{(k-1)j}(z)Q_{ji}(z)\Big)$$

we get

$$\deg(P_{ki}(z)) \leq n + q(k-1)$$

whence

$$\mathrm{ord}(\Delta_0(z)H_j(z)) \leq ln + \sum_{k=1}^{l}(k-1) + N_1 = ln + \frac{ql(l-1)}{2} + N_1$$

Finally from $R_k(z) = T(z)\frac{\mathrm{d}}{\mathrm{d}z}R_{k-1}(z)$ we have

$$\mathrm{ord}(R_k(z)) \geq \mathrm{ord}(R_1(z)) - k + 1$$

Comparing with the relation

$$\Delta_0(z)H_j(z) = T_j(z)\sum_{k=1}^{l} \Delta_{kj}(z)P_k(z)$$

we conclude that

$$\mathrm{ord}(\Delta_0(z)H_j(z)) \geq \mathrm{ord}(R_1(z)) - l + 1$$

while

$$\mathrm{ord}(\Delta_0(z)H_j(z)) \leq ln + \frac{ql(l-1)}{2} + N_1$$

whence

$$\mathrm{ord}(R_1(z)) \leq ln + \frac{qm(m-1)}{2} + m - 1 + N_1$$

and when $l = m$ we have $N_1 = p$. Set $C_0 = \frac{qm(m-1)}{2} + m - 1 + N_1$ and $C_1 = \frac{qm(m-1)}{2} + m - 1 + p$, we are done with the proof of the whole lemma.

□

Remark 4.7. This particular technique is essential in proving the lemma 2.1-2.2 in Zudilin's paper [5] and even many important bounds and calculations among them are the same. Roughly speaking, the idea of this kind of proof belongs to "deg → ord"- type, where one needs to first get an upper bound for the degree of polynomial coefficients, then one would get lower and upper bound of the order of linear forms and determinants. In our paper, effective control of the lower bound of the order in the construction of graded Padé approximation in section 3 is the core of the applying graded Padé approximation to G-functions and is also the merit of this method.

Now we turn to treat similar problems for graded Padé approximants. All the notations are followed from section 3.

Theorem 4.8. *For a given natural number N, there exists a constant $M_3 = M_3(N)$, for all $M > M_3(N)$, we have*

$$\Delta(z) = \det\left(P_{\overline{\mathrm{k}}}^{[n]}(z)\right) \neq 0.$$

where $n = 0, 1, \cdots, \omega - 1$, and $\overline{\mathrm{k}} \in \Omega$.

The proof of this theorem is one of the major tasks in this section. We will see the trick applied for the case $m = 2$ which is however invalid for $m \geq 3$. We will also apply Siegel-Shidlovskii method in a systematic way to prove this theorem, along this way there are several details especially imitated from the proof of Lemma 4.6.

Proof. We shall take the following steps:

Suppose the determinant is zero function. That is to say the rank of the square matrix is smaller than ω.

Recall the ω functional forms

$$R^{[n]}(z; \overline{\mathrm{a}}) = \sum_{\overline{\mathrm{s}} \in \Theta} \overline{\mathrm{a}}^{\overline{\mathrm{s}}} P_{\overline{\mathrm{s}}}^{[n]}(z) + \sum_{\overline{\mathrm{k}} \in \Omega \setminus \Theta} \overline{\mathrm{a}}^{\overline{\mathrm{k}}} P_{\overline{\mathrm{k}}}^{[n]}(z) \sum_{j=1}^{2} a_j f_j(z)$$

for $n = 0, 1 \cdots, \omega - 1$. Recall also that the generators $\overline{\mathrm{a}}^{\overline{\mathrm{s}}}$ and $\overline{\mathrm{a}}^{\overline{\mathrm{k}}}\left(\sum_{j=1}^{2} a_j f_j(z)\right)$ are linearly independent over $\mathbb{C}(z)$. Therefore we can view $R^{[n]}(z; \overline{\mathrm{a}})$ as a functional linear form and translate our language into the $\mathbb{C}(z)$-rank of $\{R^{[0]}(z; \overline{\mathrm{a}}), R^{[1]}(z; \overline{\mathrm{a}}), \cdots, R^{[\omega-1]}(z; \overline{\mathrm{a}})\}$ is equal to $\tilde{\omega}$ where $1 \leq \tilde{\omega} < \omega$. Repeat and imitate the proof of Lemma 4.4, this implies the $\mathbb{C}(z)$-linear independence of $\{R^{[n]}(z; \overline{\mathrm{a}})\}$ for $n = 0, 1 \cdots \tilde{\omega} - 1$, and one can choose $\tilde{\omega}$ linearly independent columns in the matrix

$$\left(P_{\overline{\mathrm{k}}}^{[n]}(z)\right)$$

where $n = 0, 1 \cdots, \tilde{\omega} - 1$ and $\overline{\mathrm{k}} \in \Omega$. Denote by $\tilde{\Omega} \subset \Omega$ the index set of these columns. We can imitate step 1 in the proof of Lemma 4.6 and write down in ordered indices the linear dependence expression such that every element is written as a linear combination of those terms appeared only before it.

$$P_{\overline{\mathrm{k}}'}^{[n]}(z) = \sum_{\overline{\mathrm{k}} \in \tilde{\Omega}} P_{\overline{\mathrm{k}}}^{[n]}(z) D_{\overline{\mathrm{k}}, \overline{\mathrm{k}}'}(z) \tag{4.5}$$

for $n = 0, 1, \cdots, \tilde{\omega} - 1$ and $\overline{\mathrm{k}}' \in \Omega \setminus \tilde{\Omega}$ where $D_{\overline{\mathrm{k}}, \overline{\mathrm{k}}'}(z) \in \mathbb{C}(z)$ and $D_{\overline{\mathrm{k}}, \overline{\mathrm{k}}'}(z) = 0$ for all $\overline{\mathrm{k}} > \overline{\mathrm{k}}'$, $\overline{\mathrm{k}} \in \tilde{\Omega}$, $\overline{\mathrm{k}}' \in \Omega \setminus \tilde{\Omega}$. This is for the purpose of "triangular adjustment" for combinatorial technique on the matrix.

Then by our choice and Lemma 3.6 we have

$$\Delta = \Delta(z) = \det\left(P_{\overline{\mathrm{k}}}^{[n]}(z)\right)_{n=0,1,\cdots,\tilde{\omega}-1; \overline{\mathrm{k}} \in \tilde{\Omega}}$$

is a nonzero polynomial in $\mathbb{Q}[z]$ of degree

$$\deg \Delta < \sum_{n=0}^{\tilde{\omega}-1} (M + tn) = \tilde{\omega} M + \frac{\tilde{\omega}(\tilde{\omega} - 1)t}{2}. \tag{4.6}$$

Now, for any solution of Eq. (1.3), $a_j(z)$ where $j = 1, 2$, it is a routine (where we indeed need to define the conjugate system before for cancellation in the computation) to check that $\overline{\mathrm{a}}^{\overline{\mathrm{k}}}$ for all $\overline{\mathrm{k}} \in \Theta$ and $\overline{\mathrm{a}}^{\overline{\mathrm{k}}}\left(\sum_{j=1}^{2} a_j f_j(z)\right)$ for all $\overline{\mathrm{k}} \in \Omega \setminus \Theta$ form exactly a solution of the following system of homogeneous linear differential equations, denoted by $x_{\overline{\mathrm{k}}}$ where $\overline{\mathrm{k}} \in \Omega$, and this also partly explains the origin of our graded Padé approximants.

$$\frac{\mathrm{d}}{\mathrm{d}z} x_{\overline{\mathrm{k}}} = -\sum_{l,j=1}^{2} k_j Q_{lj}(z) x_{\overline{\mathrm{k}} - \overline{\mathrm{e_j}} + \overline{\mathrm{e_l}}} + (N - |\,\overline{\mathrm{k}}\,|) \sum_{l=1}^{2} Q_{l0}(z) x_{\overline{\mathrm{k}} + \overline{\mathrm{e_l}}} \tag{4.7}$$

for all $\overline{\mathrm{k}} \in \Omega$. Note that this is a system of order ω in the language of linear differential equations. Then we just repeat and imitate Step 2 to Step 4 of the proof of Lemma 4.6 and prove in the same way that

$$D_{\overline{\mathrm{k}},\overline{\mathrm{k}}'}(z) = -\frac{\lambda_{\overline{\mathrm{k}},\overline{\mathrm{k}}'}(z)}{\lambda(z)}$$

for $\overline{\mathrm{k}} \in \tilde{\Omega}$ and $m^{'} \in \Omega \setminus \tilde{\Omega}$ where $\lambda(z) \neq 0$, $\lambda(z)$ and $\lambda_{\overline{\mathrm{k}},\overline{\mathrm{k}}'}(z)$ are minors of size $\omega - \tilde{\omega}$ of some matrix formed by fundamental solutions of the system Eq. (4.7). Then we continue imitating Step 5 of the proof of Lemma 4.6 and get a formula of $D_{\overline{\mathrm{k}},\overline{\mathrm{k}}'}(z)$ expressed in fixed fundamental solutions of Eq. (4.7), and repreat the same treatment of Step 6 of the proof of Lemma 4.6 and use the famous Shidlovskii's lemma Lemma 4.3 to get the conclusion that there exists a constant C_7 depending only on our system Eq. (4.7) such that $\deg D_{\overline{\mathrm{k}},\overline{\mathrm{k}}'}(z) \leq C_7$, $\deg D(z) \leq C_7$ and $\deg \big(D(z) D_{\overline{\mathrm{k}},\overline{\mathrm{k}}'}(z)\big) \leq C_7$ where $D(z)$ is the least common denominator of $D_{\overline{\mathrm{k}},\overline{\mathrm{k}}'}(z)$. But this is eventually depending only on N and our original systemEq. (1.2), which is determined uniquely by $f_1(z)$ and $f_2(z)$, so we have $C_7 = C_7(f_1, f_2, N)$.

Next we shall start from definition of $R^{[n]}_{\overline{\mathrm{s}}}(z)$ and write down

$$R^{[n]}_{\overline{\mathrm{s}}}(z) = P^{[n]}_{\overline{\mathrm{s}}}(z) + \sum_{j=1}^{2} P^{[n]}_{\overline{\mathrm{s}} - \overline{\mathrm{e_j}}}(z) f_j(z) = \sum_{\overline{\mathrm{k}} \in \Omega} P^{[n]}_{\overline{\mathrm{k}}}(z) x_{\overline{\mathrm{k}},\overline{\mathrm{s}}}(f)$$

for all $\overline{\mathrm{s}} \in \Theta$ where the functions $x_{\overline{\mathrm{k}},\overline{\mathrm{s}}}(f)$ are defined to be $\delta_{\overline{\mathrm{k}},\overline{\mathrm{s}}}$ when $\overline{\mathrm{k}} \in \Theta$ and $\sum_{j=1}^{2} \delta_{\overline{\mathrm{k}} + \overline{\mathrm{e_j}},\overline{\mathrm{s}}} f_j(z)$ when $\overline{\mathrm{k}} \in \Omega \setminus \Theta$. Now we plug in Eq. (4.5) for all $\overline{\mathrm{k}} \in \Omega \setminus \tilde{\Omega}$ and get

$$R^{[n]}_{\overline{\mathrm{s}}}(z) = \sum_{\overline{\mathrm{k}} \in \tilde{\Omega}} P^{[n]}_{\overline{\mathrm{k}}}(z) \tilde{x}_{\overline{\mathrm{k}},\overline{\mathrm{s}}}(f)$$

for all $n = 0, 1, \cdots, \tilde{\omega} - 1$ and $\overline{\mathrm{s}} \in \Theta$ where

$$\tilde{x}_{\overline{\mathrm{k}},\overline{\mathrm{s}}}(f) = x_{\overline{\mathrm{k}},\overline{\mathrm{s}}}(f) + \sum_{\overline{\mathrm{k}} \in \Omega \setminus \tilde{\Omega}} D_{\overline{\mathrm{k}},\overline{\mathrm{k}}'} x_{\overline{\mathrm{k}}',\overline{\mathrm{s}}}(f)$$

Next we are going to state without detailed proof (for the case $m \geq 3$) that for any $\tilde{\Omega} \subset \Omega$ there exist subset $\Omega_1 \subset \tilde{\Omega}$ and $\Theta_1 \subset \Theta$ with $|\,\Omega_1\,| = |\,\Theta_1\,| = \tilde{\theta}$ such that

$$\det \Big(\tilde{x}_{\overline{\mathrm{k}}',\overline{\mathrm{s}}}(f)\Big)_{\overline{\mathrm{k}} \in \Omega_1; \overline{\mathrm{s}} \in \Theta_1} \neq 0$$

as a function of z and such that

$$\frac{\tilde{\omega}}{\tilde{\theta}} \leq \frac{\omega}{\theta} - \frac{1}{\theta(N + m - 1)} \tag{4.8}$$

The detailed proof of this fact can be found in Proposition 2.1 of the paper [4], and we need to mention the idea of this technical proof and modify it for our case $m = 2$. First of all, fix an arbitrary $\Omega_1 \subset \tilde{\Omega}$ and consider an injection $\sigma : \Omega_1 \longrightarrow \Theta$ with the image $\sigma(\Omega_1) = \Theta_1$, our goal is to modify it to satisfy the requirement. We first suppose each term appeared in the minor as a product, $x_{\overline{\mathrm{k}},\sigma(\overline{\mathrm{k}})}$, to be nonzero, that is to say from our definition for $x_{\overline{\mathrm{k}},\overline{\mathrm{s}}}(f)$, that for each $\overline{\mathrm{k}} \in \Omega_1$, we have $\sigma(\overline{\mathrm{k}}) = \overline{\mathrm{k}}$

if $\overline{\mathrm{k}} \in \Omega_1 \cap \Theta$, and there exists a $j = j(\overline{\mathrm{k}}, \sigma) \in \{1, 2\}$ such that $\sigma(\overline{\mathrm{k}}) = \overline{\mathrm{k}} + \overline{\mathrm{e_j}}$ if $\overline{\mathrm{k}} \in \Omega_1 \setminus \Theta$, and hence we have

$$\prod_{\overline{\mathrm{k}} \in \Omega} x_{\overline{\mathrm{k}}, \sigma \overline{\mathrm{k}}} = \prod_{\overline{\mathrm{k}} \in \Omega_1 \setminus \Theta} f_j(\overline{\mathrm{k}}, \sigma).$$

Imitate Step 5 in the proof of Lemma 4.6 and apply Binet-Cauchy formula again to the minor

$$\begin{aligned} &\det \left(\tilde{x}_{\overline{\mathrm{k}}', \overline{\mathrm{s}}}(f) \right)_{\overline{\mathrm{k}} \in \Omega_1; \overline{\mathrm{s}} \in \Theta_1} \\ &= \sum_{\Omega' \subset \Omega : |\Omega'| = |\Omega_1|} \det \left(\tilde{D}_{\overline{\mathrm{k}}, \overline{\mathrm{k}}'} \right)_{\overline{\mathrm{k}} \in \Omega_1; \overline{\mathrm{k}}' \in \Omega'} \det \left(\tilde{x}_{\overline{\mathrm{k}}', \overline{\mathrm{s}}} \right)_{\overline{\mathrm{k}}' \in \Omega'; \overline{\mathrm{s}} \in \Theta_1} \end{aligned} \tag{4.9}$$

where $\tilde{D}_{\overline{\mathrm{k}}, \overline{\mathrm{k}}'}$ is defined to be $\delta_{\overline{\mathrm{k}}, \overline{\mathrm{k}}'}$ if $\overline{\mathrm{k}}' \in \tilde{\Omega}$ and $D_{\overline{\mathrm{k}}, \overline{\mathrm{k}}'}$ if $\overline{\mathrm{k}}' \in \Omega \setminus \tilde{\Omega}$. Then one can show that the summand corresponding to $\Omega' = \Omega_1$ has actually the largest multidegree among all the nonzero summands and is nonzero, hence

$$\det \left(\tilde{x}_{\overline{\mathrm{k}}', \overline{\mathrm{s}}}(f) \right)_{\overline{\mathrm{k}} \in \Omega_1; \overline{\mathrm{s}} \in \Theta_1} \neq 0.$$

The details are based on combinatorial technique of lexicographic order which we have mentioned as "triangular adjustment" at the beginning of our proof here. Note that this argument essentially relies on the assumption that $f_j(z)$ are algebraically independent over $\mathbb{C}(z)$, (for all $m \geq 2$) together with the following fact (for the proof of the inequality 4.8): For any nonempty subset $\tilde{\Omega}$ of Ω, there is a subset $\Omega_1 \in \tilde{\Omega}$ and an injection σ from Ω_1 to Θ such that

(i) $\sigma(\overline{\mathrm{k}}) = \overline{\mathrm{k}}$ if $\overline{\mathrm{k}} \in \Omega_1 \cap \Theta$, and there exists a $j = j(\overline{\mathrm{k}}, \sigma) \in \{1, 2\}$ such that $\sigma(\overline{\mathrm{k}}) = \overline{\mathrm{k}} + \overline{\mathrm{e_j}}$ if $\overline{\mathrm{k}} \in \Omega_1 \setminus \Theta$;

(ii) If $\overline{\mathrm{k}} \in \Omega_1$ and $\overline{\mathrm{k}} + \overline{\mathrm{e_j}} < \sigma(\overline{\mathrm{k}})$, then we have $\overline{\mathrm{k}} + \overline{\mathrm{e_j}} \in \sigma(\Omega_1)$ and $\sigma^{-1}(\overline{\mathrm{k}} + \overline{\mathrm{e_j}}) > \overline{\mathrm{k}}$.

We skip the proof of this fact here and now we are going to investigate the trick for the case $m = 2$ to avoid the strong assumption of algebraic independence. We pick $\Omega_1 \in \tilde{\Omega}$ a satisfying this property and set $\Theta_1 = \sigma(\Omega_1)$. We will luckily find that $\det \left(\tilde{x}_{\overline{\mathrm{k}}', \overline{\mathrm{s}}}(f) \right)_{\overline{\mathrm{k}} \in \Omega_1; \overline{\mathrm{s}} \in \Theta_1}$ is in a product form, and up to a sign, equal to

$$\prod_{\overline{\mathrm{k}} \in \Omega_1 : \sigma(\overline{\mathrm{k}}) = \overline{\mathrm{k}} + \overline{e_2}} \tilde{x}_{\overline{\mathrm{k}}, \overline{\mathrm{k}} + \overline{e_2}} \cdot \prod_{\overline{\mathrm{k}} \in \Omega_1 : \overline{\mathrm{k}} - \overline{e_2} \in \Omega_1} \left(\tilde{x}_{\overline{\mathrm{k}} - \overline{e_2}, \overline{\mathrm{k}} - \overline{e_2} + \overline{e_1}} - \tilde{x}_{\overline{\mathrm{k}}, \overline{\mathrm{k}} - \overline{e_2} + \overline{e_1}} \tilde{x}_{\overline{\mathrm{k}} - \overline{e_2}, \overline{\mathrm{k}}} \right). \tag{4.10}$$

We can prove this by permuting the columns of the matrix $\left(\tilde{x}_{\overline{\mathrm{k}}', \overline{\mathrm{s}}}(f) \right)_{\overline{\mathrm{k}} \in \Omega_1; \overline{\mathrm{s}} \in \Theta_1}$ in such a way that the "diagonal" could be adjusted to consist of elements $\tilde{x}_{\overline{\mathrm{k}}, \sigma(\overline{\mathrm{k}})}$ for $\overline{\mathrm{k}} \in \Omega_1$ and note that if $\overline{\mathrm{k}} > \overline{\mathrm{s}}$ for some $\overline{\mathrm{k}} \in \tilde{\Omega}$ and $\overline{\mathrm{s}} \in \Theta$, then $\tilde{x}_{\overline{\mathrm{k}}, \overline{\mathrm{s}}} = 0$. Discuss in two situations $\overline{\mathrm{k}} \in \Omega \setminus \Theta$ and $\overline{\mathrm{k}} \in \Theta$, we find that the "diagonal" consists of either dimension one blocks or dimension two blocks which gives us the desired formula. Now we may apply Lemma 1.12 to conclude. From Eq. (4.10), $\det \left(\tilde{x}_{\overline{\mathrm{k}}', \overline{\mathrm{s}}}(f) \right)_{\overline{\mathrm{k}} \in \Omega_1; \overline{\mathrm{s}} \in \Theta_1}$ is a product with terms being either linear combination of $f_1(z)$, $f_2(z)$ and 1 or linear combination of $f_2^2(z)$,$f_1(z)$, $f_2(z)$ and 1, with coefficients in $\mathbb{C}(z)$, hence nonzero by Lemma 1.12.

Now we denote by $D(z)$ the least common denominator of the rational functions $D_{\overline{\mathrm{k}}, \overline{\mathrm{k}}'}(z)$ and by $\chi(z)$ the determinant

$$\det \left(D(z) \tilde{x}_{\overline{\mathrm{k}}, \overline{\mathrm{s}}}(f) | \delta_{\overline{\mathrm{k}}, \overline{\mathrm{s}}'} \right)_{\overline{\mathrm{k}} \in \tilde{\Omega}; \overline{\mathrm{s}} \in \Omega_1, \overline{\mathrm{s}}' \in \tilde{\Omega} \setminus \Omega_1}.$$

The last step is to deduce a contradiction from the aspect of order of zero of $\Delta(z)$ and $\chi(z)$. Recall that we have $\deg D_{\overline{\mathrm{k}},\overline{\mathrm{k}}'}(z) \leq C_7$, $\deg D(z) \leq C_7$ and $\deg\big(D(z)D_{\overline{\mathrm{k}},\overline{\mathrm{k}}'}(z)\big) \leq C_7$, and the functions $D(z)\tilde{x}_{\overline{\mathrm{k}},\overline{\mathrm{s}}}(f)$ are linear forms in 1, $f_1(z)$ and $f_2(z)$ with coefficients in $\mathbb{C}[z]$ of degree at most C_7. Therefore, we have $\deg(\chi(z)) \leq \tilde{\theta}C_7$ where by $\deg(\Delta(z))$ we mean the maximal degree of the coefficients in $\mathbb{C}[z]$ of the monomials in $f_j(z)$ appeared in the linear expansion of $\chi(z)$. And we also have $\chi(z) = D^{\tilde{\theta}}(z)\det\Big(\tilde{x}_{\overline{\mathrm{k}}',\overline{\mathrm{s}}}(f)\Big)_{\overline{\mathrm{k}}\in\Omega_1;\overline{\mathrm{s}}\in\Theta_1}$, hence $\chi(z)$ is a nonzero function. By Lemma 4.1, we have $\mathrm{ord}_{z=0}\big(\chi(z)\big) \leq C_8$ where $C_8 = C_8(f_1(z), f_2(z), \tilde{\theta}, \tilde{\theta}C_7) = C_8(f_1(z), f_2(z), N)$ is believed by me to be inefficient, see Remark 4.2. This C_8 is the only constant in this mémoire which is not easy to be very efficient.

Now we multiply the matrix $P^{[n]}_{\overline{\mathrm{k}}}(z)_{n=0,1,\cdots,\tilde{\omega}-1;\overline{\mathrm{k}}\in\tilde{\Omega}}$ on the right by the marix

$$\Big(D(z)\tilde{x}_{\overline{\mathrm{k}},\overline{\mathrm{s}}}(f)|\delta_{\overline{\mathrm{k}},\overline{\mathrm{s}}'}\Big)_{\overline{\mathrm{k}}\in\tilde{\Omega};\overline{\mathrm{s}}\in\Omega_1,\overline{\mathrm{s}}'\in\tilde{\Omega}\setminus\Omega_1},$$

and we will get the matrix

$$\Big(D(z)R^{[n]}_{\overline{\mathrm{s}}}(z)|P^{[n]}_{\overline{\mathrm{s}}'}(z)\Big)_{n=0,1,\cdots,\tilde{\omega}-1;\overline{\mathrm{s}}\in\Omega_1,\overline{\mathrm{s}}'\in\tilde{\Omega}\setminus\Omega_1}$$

whose determinant is $\Delta(z)\chi(z)$ and whose first $\tilde{\theta}$ columns consist of functions with order at $z=0$ at least $K-\tilde{\omega}+1$ by Lemma 3.6, so

$$\tilde{\theta}(K-\tilde{\omega}) < \mathrm{ord}_{z=0}\big(\Delta(z)\chi(z)\big) = \mathrm{ord}_{z=0}\big(\Delta(z)\big) + \mathrm{ord}_{z=0}\big(\chi(z)\big) \leq \mathrm{ord}_{z=0}\big(\Delta(z)\big) + C_8$$

According to Eq. (4.6) we have the following contradiction when $M > M_3$ for a constant $M_3 = M_3(N)$, which is inefficient due to the deficiency of C_8.

$$\begin{aligned} 0 &\leq \deg\big(\Delta(z)\big) - \mathrm{ord}_{z=0}\big(\Delta(z)\big) < \tilde{\omega}M + \frac{\tilde{\omega}(\tilde{\omega}-1)t}{2} - \tilde{\theta}K + \tilde{\theta}\tilde{\omega} + C_8 \\ &\leq \tilde{\omega}M - \frac{\tilde{\theta}}{\theta}(\omega-\varepsilon)M + C_9 \leq \frac{\tilde{\theta}(\omega-\varepsilon-\frac{\varepsilon}{N+1})M}{\theta} - \frac{\tilde{\omega}(\omega-\varepsilon)M}{\theta} + C_9 \\ &= -\frac{\tilde{\theta}\,\varepsilon\,M}{\theta(N+1)} + C_9 \leq -\frac{\varepsilon\,M}{\theta(N+1)} + C_9 \end{aligned} \tag{4.11}$$

where we have used Eq. (4.8) and hence there is a contradiction when $M > M_3$ for a constant M_3. We are done of proving this theorem. □

Remark 4.9. We hope to modify the proof of Theorem 4.8 to obtain a more efficient M_3, that is to say, we need to obtain a more efficient C_8. This is most hopeful when $m=2$ because there we have a product formula for the determinant. Even there I think the upper bound of order of linear form appeared in each factor of that product formula will essentially use Lemma 4.1 and hence still not so good. Therefore I believe that in order to do so, we need to add more requirement on the functions $f_1(z)$ and $f_2(z)$ in such a way that their Taylor coefficients should obey a very strict linear-algebraic property, but in general I just believe that it is hopeless to get a more efficient C_8 than what Zudilin has achieved.

Lemma 4.10. *For all $M > M_3$, there exists a positive constant $C_{10} = C_{10}(N)$, such that*

$$\Delta(z) = z^{\mathrm{ord}_{z=0}\Delta}\Delta_0(z),$$

where $\Delta_0(z)$ is a nonzero polynomial and $\deg\big(\Delta_0(z)\big) < \varepsilon\,M + C_{10}$.

Proof. We have

$$\deg\big(\Delta_0(z)\big) = \deg\big(\Delta(z)\big) - \mathrm{ord}_{z=0}\big(\Delta(z)\big) < \omega\,M + \frac{\omega(\omega-1)t}{2} - \theta(K-\omega) + C_8 \leq \varepsilon\,M + C_10$$

Note that since C_8 is not very efficient, neither is C_{10}.

□

5. Main results for numerical forms

The purpose of this section is to prove the following lemmas on numerical estimation and their corollaries. Comparing with functional argument, this section is relatively easy since the rational points we consider in this section satisfy the assumption of Theorem 1.10, in particular their numerators are much smaller than their denominators. These results are all stated for function values at a certain point α and thus will directly lead to the final proof of the main theorem. Let us adopt and fix all the notations and conditions in the main theorem. Recall the definition of $P_{\overline{\mathrm{k}}}^{[n]}(z)$ and $R_{\overline{\mathrm{s}}}^{[n]}(z)$ in section 3, for all $\overline{\mathrm{k}} \in \Omega$, $\overline{\mathrm{s}} \in \Theta$ and $n \in \mathbb{N}$. Fix $l \in \{1,2\}$ and $\overline{\mathrm{k}}^* = (N-1)\,\overline{\mathrm{e_l}}$, $\overline{\mathrm{s}}^* = N\,\overline{\mathrm{e_l}}$. Denote by w_n the least common denominator of all the rational coefficients of the polynomials $P_{\overline{\mathrm{k}}^*}^{[n]}(z) \in \mathbb{Q}[z]$ for all $n \in \mathbb{N}$.

Lemma 5.1. *Let $\varepsilon_1 > 0$ and M_1 be chosen the same as that in Lemma 3.3. For any $\varepsilon_1 > 0$, there exists a constant $M_2 = M_2(N, \varepsilon_1) \geq M_1$, such that for all $M > M_2$, we have the following estimation for all $n < \frac{K-M}{t+1} \in \mathbb{N}$ and all $\alpha \in \mathbb{R}$ such that $|\alpha| < \frac{1}{2C}$:*

$$\begin{aligned}
&w_n \leq C_3\,\Psi^{(1+\frac{\varepsilon_1}{2})n(1+logN)},\\
&\left|P_{\overline{\mathrm{k}}^*}^{[n]}(\alpha)\right| \leq C_4\binom{M}{n}(2(t+1)^2H)^{n(1+logN)}C_2^{(1+\varepsilon_1)M},\\
&\left|R_{\overline{\mathrm{s}}^*}^{[n]}(\alpha)\right| \leq C_5\binom{M}{n}(2(t+1)^2H)^{n(1+logN)}C_2^{(1+\varepsilon_1)M}(C|\alpha|)^{K-n},\\
&d_{M+tn}w_nP_{\overline{\mathrm{s}}^*}^{[n]}(z) \in \mathbb{Z}[z],
\end{aligned} \tag{5.1}$$

where $C_3 = C_3(N,\varepsilon_1) > 0$, $C_4 = C_4(N) > 0$ and $C_5 = C_5(N) > 0$ are constants.

Proof. Before proving this metric conclusion for functional values, we shall first state a fact about height of product of polynomials. Suppose $p_i(z)$ where $i = 1, 2, \cdots, n$ are n polynomials in $\mathbb{R}[z]$. Let $p(z) = \prod_{i=1}^n p_i(z)$, then we have

$$\mathrm{H}\left(p(z)\right) \leq (\deg p(z) + 1)^n \prod_{i=1}^n \mathrm{H}\left(p_i(z)\right) \tag{5.2}$$

The proof of it is a routine and we skip it here. Now recall the definition of $P_{\overline{\mathrm{k}}}^{[n]}(z)$ for $\overline{\mathrm{k}} \in \Omega$ and $n \in \mathbb{N}$ and we have from Eq. (3.7) that

$$\sum_{\overline{\mathrm{k}}\in\Omega\setminus\Theta} \overline{\mathrm{a}}^{\overline{\mathrm{k}}}\,P_{\overline{\mathrm{k}}}^{[n]}(z) = \frac{T^n(z)}{n!}D^n\Big(\sum_{\overline{\mathrm{k}}\in\Omega\setminus\Theta} \overline{\mathrm{a}}^{\overline{\mathrm{k}}}\,P_{\overline{\mathrm{k}}}(z)\Big) \tag{5.3}$$

for $n \in \mathbb{N}$. Hence $P_{\overline{\mathrm{k}}^*}^{[n]}(z)$ is nothing but the coefficient of $\overline{\mathrm{a}}^{\overline{\mathrm{k}}^*}$ appeared in the functional form

$$\frac{T^n(z)}{n!}D^n\Big(\sum_{\overline{\mathrm{k}}\in\Omega\setminus\Theta} \overline{\mathrm{a}}^{\overline{\mathrm{k}}}\,P_{\overline{\mathrm{k}}}(z)\Big),$$

Therefore it suffices to compute the coefficient of $\overline{\mathrm{a}}^{\overline{\mathrm{k}}^*}$ appeared in the functional form

$$D^n\big(\,\overline{\mathrm{a}}^{\overline{\mathrm{k}}}\,P_{\overline{\mathrm{k}}}(z)\big).$$

Then we use an important property of graded Padé approximants, that if $a_1(z)$ and $a_2(z)$ form a solution to the system Eq. (1.3), then we have

$$D^n\big(\,\overline{\mathrm{a}}(z)^{\overline{\mathrm{k}}}P_{\overline{\mathrm{k}}}(z)\big) = \frac{\mathrm{d}^n}{\mathrm{d}z^n}\big(\,\overline{\mathrm{a}}(z)^{\overline{\mathrm{k}}}P_{\overline{\mathrm{k}}}(z)\big).$$

Therefore we can plug Eq. (1.5) in and compute this coefficient of $\overline{\mathrm{a}}^{\overline{\mathrm{k}}^*}$ in $D^n\big(\overline{\mathrm{a}}^{\overline{\mathrm{k}}}\,P_{\overline{\mathrm{k}}}(z)\big)$ for each fixed $\overline{\mathrm{k}} \in \Omega \setminus \Theta$, by the Leibniz rule for higher order derivatives of multi-factor product, and it is equal to

$$\sum_{n_1,\cdots,n_N \in \mathbb{N}; n_1+\cdots+n_N=n} \frac{n!}{n_1!\cdots n_N!} \prod_{i=1}^{k_1} S_{1l}^{[n_i]}(z) \prod_{i=k_1+1}^{N-1} S_{2l}^{[n_i]}(z) P_{\overline{\mathrm{k}}}^{(n_N)}(z). \tag{5.4}$$

So that we have the coefficient of $\overline{\mathrm{a}}^{\overline{\mathrm{k}}^*}$ in

$$\frac{T^n(z)}{n!} D^n\Big(\sum_{\overline{\mathrm{k}}\in\Omega\setminus\Theta} \overline{\mathrm{a}}^{\overline{\mathrm{k}}}\,P_{\overline{\mathrm{k}}}(z)\Big),$$

is equal to

$$P_{\overline{\mathrm{k}}^*}^{[n]}(z) = \sum_{\overline{\mathrm{k}}\in\Omega\setminus\Theta} \sum_{n_1,\cdots,n_N \in \mathbb{N}; n_1+\cdots+n_N=n} p(\overline{\mathrm{k}}; n_1, \cdots, n_N) \tag{5.5}$$

where

$$p(\overline{\mathrm{k}}; n_1, \cdots, n_N) = \prod_{i=1}^{k_1} \frac{T^{n_i}(z) S_{1l}^{[n_i]}(z)}{n_i!} \prod_{i=k_1+1}^{N-1} \frac{T^{n_i}(z) S_{2l}^{[n_i]}(z)}{n_i!} T^{n_N}(z) \frac{P_{\overline{\mathrm{k}}}^{(n_N)}(z)}{n_N!} \tag{5.6}$$

Within one summand of Eq. (5.5) we can assume without loss of certainty that $n_1 \geq n_2 \geq \cdots \geq n_N$, and so we have

$$n_i \leq \Big\lfloor \frac{n}{i} \Big\rfloor$$

for $i = 1, 2, \cdots, N$. Now apply this inequality and Lemma 1.9 together with Definition 1.7 and our core result Lemma 3.3, we can easily show that $\frac{P_{\overline{\mathrm{k}}}^{(n_N)}(z)}{n_N!} \in \mathbb{Z}[z]$ and $\prod_{j=i}^{N} \lfloor \frac{n}{i} \rfloor p(\overline{\mathrm{k}}; n_1, \cdots, n_N) \in \mathbb{Z}[z]$ when $n < \frac{M}{2}$ and $\overline{\mathrm{k}} \in \Omega \setminus \Theta$. And we have estimation by the fact at the beginning of this proof that

$$\begin{aligned} &\deg\big(p(\overline{\mathrm{k}}; n_1, \cdots, n_N)\big) \leq M + (2t+1)n - 1 \\ &\mathrm{H}\big(p(\overline{\mathrm{k}}; n_1, \cdots, n_N)\big) \leq \big(M + (2t+1)n\big)^{N+1} C_1^N \big(2(t+1)^2 H\big)^{n(1+\log N)} \binom{M}{n} C_2 (1 + \frac{\varepsilon_1}{2}) M. \end{aligned} \tag{5.7}$$

To estimate the upper bound for w_n, we can use Eq. (5.5) to get

$$w_n \leq \prod_{j=i}^{N} \Big\lfloor \frac{n}{i} \Big\rfloor p(\overline{\mathrm{k}}; n_1, \cdots, n_N),$$

and by the condition of "cancellation of factorials" Definition 1.7 we obtain

$$w_n \leq C_3\, \Psi^{(1+\frac{\varepsilon_1}{2})n(1+logN)}.$$

where C_3 depends only on N and ε_1, for $n < \frac{M}{2}$. From Eq. (5.5) and Eq. (5.7) we obtain

$$\mathrm{H}\big(P_{\overline{\mathrm{k}}^*}^{[n]}(z)\big) \leq C_6 (3M)^{2N} \big(2(t+1)^2 H\big)^{n(1+\log N)} \binom{M}{n} C_2 (1 + \frac{\varepsilon_1}{2}) M \tag{5.8}$$

where $C_6 = C_6(N)$, for all $n < \frac{M}{2}$. Now for α whose absolute value is less than 1, by Lemma 3.6 and by Eq. (5.8) we are done with

$$\Big| P_{\overline{\mathrm{k}}^*}^{[n]}(\alpha) \Big| \leq C_4 \binom{M}{n} (2(t+1)^2 H)^{n(1+logN)} C_2^{(1+\varepsilon_1)M}$$

where $C_4 = C_4(N)$, when $M > M_2{}'(N, \varepsilon_1)$. The last two assertions follow from taking the Taylor expansion of $f_l(z)$ and use the fact that

$$\deg\left(P_{\overline{\mathrm{s}}^*}^{[n]}(z)\right) \leq M + tn - 1 < K - n \leq \mathrm{ord}_{z=0}\, R_{\overline{\mathrm{s}}^*}^{[n]}(z)$$

when $n < \frac{K-M}{t+1}$ and also use the assumption $\alpha < \frac{1}{2C}$. The estimation is valid when $M > M_2^{''}(N, \varepsilon_1)$, so we need to take $M_2 = \max\left(M_2^{'}(N, \varepsilon_1), M_2^{''}(N, \varepsilon_1)\right)$. This final estimation uses convergence of power series when the ratio is less than 1, which is simple but very useful for metric number theory.

□

Remark 5.2. In the last overview section we will see that this lemma is the core of metric number theoretic aspect when applying graded Padé approximants to a certain class of G-functions, that is to say, the merit of graded Padé approximants is to obtain the estimation in this lemma, and all the other properties like nonzero functional determinant in Theorem 4.8 are not exclusive for graded Padé approximants. For example we have already seen in section 4 that Siegel- Shidolovskii method gives us a similar result for Padé approximants of the first kind. So we can say that what we have obtained in this lemma is one of the exclusive merits for applying graded Padé approximants to G-functions Eq. (5.1), and for other functions like E-functions we also have estimation for the upper bound of $\left|P_{\overline{\mathrm{k}}^*}^{[n]}(\alpha)\right|$ and $\left|R_{\overline{\mathrm{s}}^*}^{[n]}(\alpha)\right|$, but in different accuracy. One should also note that only the estimation of $\left|R_{\overline{\mathrm{s}}^*}^{[n]}(\alpha)\right|$ needs the sharper assumption $|\alpha| < \frac{1}{2C}$, while $\left|P_{\overline{\mathrm{k}}^*}^{[n]}(\alpha)\right|$ only needs $|\alpha| < 1$.

Next we turn to linear algebraic result for numerical forms, whose proof is developed from Siegel-Shidlovskii method studied systematically in section 4.

Lemma 5.3. *Let α be a rational number stated in Theorem 1.10 satisfying the same condition there. Let $\varepsilon < \frac{1}{t+3}$ be an arbitrary positive constant stated in Theorem 1.10. Let C_{10} and M_3 be the constant stated in Lemma 4.10. Then for all $M > M_3$, the $\omega + \lfloor \varepsilon M \rfloor + C_{10}$ by ω numerical matrix*

$$\left(P_{\overline{\mathrm{k}}}^{[n]}(\alpha)\right)$$

where $n = 0, 1, \cdots, \omega + \lfloor \varepsilon M \rfloor + C_{10}$ and $\overline{\mathrm{k}} \in \Omega$, has exactly rank ω, that is to say, of full rank.

Proof. We have $\alpha T(\alpha) \neq 0$. Suppose $\mathrm{ord}_{z=\alpha}\left(\Delta(z)\right) = r$, then by Lemma 4.10 we have $r \leq \deg\left(\Delta_0(z)\right) < \varepsilon M + C_{10}$, so it suffices to show the lemma for the matrix with indices $n = 0, 1, \cdots, \omega + r - 1$ and $\overline{\mathrm{k}} \in \Omega$. We are going to apply the same method above Lemma 4.5 to the $\mathbb{C}(z)$- linear forms $R^{[n]}(z; \overline{\mathrm{a}})$ with generators $\overline{\mathrm{a}}^{\overline{\mathrm{s}}}$ and $\overline{\mathrm{a}}^{\overline{\mathrm{k}}} \sum_{j=1}^{2} a_j f_j(z)$. Recall that A is the $\mathbb{C}(z)$-module consisting of these linear forms and A is of dimension ω. This is a classical linear-algebraic method useful when determinant and differential operator are involved. We can view $R^{[n]}(z; \overline{\mathrm{a}})$ playing the role of $R_k(z)$ in the discussion of Shidlovskii method. Denote these ω generators by $F_{\overline{\mathrm{k}}}$ where $\overline{\mathrm{k}} \in \Omega$, so we have

$$R^{[n]}(z; \overline{\mathrm{a}}) = \sum_{\overline{\mathrm{k}} \in \Omega} P_{\overline{\mathrm{k}}}^{[n]}(z) F_{\overline{\mathrm{k}}}$$

where we just view $F_{\overline{\mathrm{k}}}$ as generator or one element of A. Then we repeat what we have done just before Lemma 4.5 and obtain the following identity for functional determinant:

$$\Delta(z) F_{\overline{\mathrm{k}}} = \sum_{n=0}^{\omega-1} \Delta_{n,\overline{\mathrm{k}}}(z) R^{[n]}(z; \overline{\mathrm{a}}) \tag{5.9}$$

where $\Delta_{n,\overline{\mathrm{k}}}(z)$ is the cofactor of $P_{\overline{\mathrm{k}}}^{[n]}(z)$ in the functional determinant. Now we can apply the d-ifferential operator D to both sides of Eq. (5.9), and since A is closed under the action of D, we have

$$\Delta^{(1)}(z) F_{\overline{\mathrm{k}}} + \Delta(z) L_{1,\overline{\mathrm{k}},0} = \sum_{n=0}^{\omega} M_{1,\overline{\mathrm{k}},n}(z) R^{[n]}(z; \overline{\mathrm{a}})$$

where $L_{1,\overline{k},0} \in A$ and $M_{1,\overline{k},n}(z) \in \mathbb{C}(z)$. Repeat this process r times, we obtain

$$\Delta^{(r)}(z)F_{\overline{k}} + \sum_{d=0}^{r-1} \Delta^{(d)}(z)L_{r,\overline{k},d} = \sum_{n=0}^{\omega+r-1} M_{r,\overline{k},n}(z)R^{[n]}(z;\overline{a}) \tag{5.10}$$

where $L_{r,\overline{k},d} \in A$ and $M_{r,\overline{k},n}(z) \in \mathbb{C}(z)$. Since $\Delta(z)$ has zero at $z = \alpha$ of order r, we have $\Delta^{(d)}(\alpha) = 0$ for $d = 0, 1, \cdots, r-1$ and $\Delta^{(r)}(\alpha) \neq 0$. Plug these values in Eq. (5.10), we will get

$$\Delta^{(r)}(\alpha)F_{\overline{k}} = \sum_{n=0}^{\omega+r-1} M_{r,\overline{k},n}(\alpha)R^{[n]}(\alpha;\overline{a}) \tag{5.11}$$

Recall that $R^{[n]}(z;\overline{a})$ is actually an element in the module A with a basis $F_{\overline{k}}$, with coefficients $P^{[n]}_{\overline{k}}(z)$. Under this basis Eq. (5.11) can also be written in the matrix form:

$$\Big((\Delta^{(r)}(\alpha))^{-1}M_{r,\overline{k},n}(\alpha)\Big)\Big(P^{[n]}_{\overline{k}}(\alpha)\Big) = \mathrm{Id}_{\overline{k}\in\Omega;\overline{k}\in\Omega}$$

where $\overline{k} \in \Omega$ and $n = 0, 1, \cdots, \omega+r-1$. We can directly see the full rank property of the two matrices in the product, or we can apply Binet-Cauchy formula to conclude here, which gives what we want to prove. □

Corollary 5.4. *Let α, ε, C_{10}, M_3 be the same as in Lemma 5.3, and fix as in Lemma 5.1 $\overline{s}^* = N\,\overline{e_1}$ and $\overline{k}^* = (N-1)\,\overline{e_1}$. then for all $M > M_3$, the $\omega + \lfloor\varepsilon M\rfloor + C_{10}$ by 2 numerical matrix*

$$\begin{pmatrix} P^{[n]}_{\overline{s}^*}(\alpha) & P^{[n]}_{\overline{k}^*}(\alpha) \end{pmatrix}$$

where $n = 0, 1, \cdots, \omega + \lfloor\varepsilon M\rfloor + C_{10}$, has exactly rank 2, that is to say, of full rank.

Proof. Immediate from Lemma 5.3. □

Remark 5.5. The reason we choose two special indices $\overline{s}^* = N\,\overline{e_1}$ and $\overline{k}^* = (N-1)\,\overline{e_1}$ is very clear: they will give the simplest expansion formula.

Now we are prepared for the proof of the main theorem, where only Lemma 5.1 and Corollary 5.4 are needed.

6. Conclusion of the proof of the Main Theorem

From the discussion in Remark 1.11, it suffices to prove the inequality 1.6 to conclude. Fix all the parameters in the assumption of Theorem 1.10, the proof has no difference written in the general case or the special case of $m = 2$. We will replace m with 2 to accord with the statement in our Theorem 1.10. For a given $\eta > \eta_0$ and $\varepsilon > 0$, we can choose a constant

$$C_{11} = b^{\varepsilon_1}(C|\alpha|)^{-\varepsilon\,\varepsilon_1}C_0^{1+\varepsilon_1}$$

only a little bit larger than C_0 by choosing a small constant $\varepsilon_1 = \varepsilon_1(\eta, \varepsilon, \alpha)$ such that

$$\begin{aligned} &\frac{(1+t\,\varepsilon)\log b + \log C_{11}}{\big(1-(t+3)\,\varepsilon\big)\log b - \log C_{11} - (2-3\,\varepsilon)\log(C|a|)} \\ <&\frac{\eta+\eta_0}{2} = \eta_1. \end{aligned} \tag{6.1}$$

Note that here ε_1 is efficient since all the related formula is given explicitly. Now we are going to pick a lower bound M_* such that for all $M > M_*$ and all natural numbers $n < \varepsilon M + C_{10} + \omega$, we will get a nice type of the upper bound of $d_{M+tn}w_n b^{M+tn}\left|P^{[n]}_{\overline{k}^*}(\alpha)\right|$ and $d_{M+tn}w_n b^{M+tn}\left|R^{[n]}_{\overline{s}^*}(\alpha)\right|$, and such that we could combine Corollary 5.4 and Lemma 5.1 properly. To do this, first we need $M_* \geq M_2 = M_2(N, \varepsilon_1)$ to apply Lemma 5.1 and $M_* \geq M_3$ to apply Corollary 5.4. Next we need a nice estimation of $\binom{M}{n}$ appeared in Lemma 5.1. Since

$$\binom{M}{n} = \frac{(M-n+1)(M-n+2)\cdots(M-1)M}{n!} < \frac{M^n}{(n/e)^n} = \left(\frac{eM}{n}\right)^n.$$

and the real function $f(x) = \left(\frac{eM}{x}\right)^x$ has derivative $\left(\frac{eM}{x}\right)^x\Big(\log\left(\frac{eM}{x}\right) - 1\Big)$, so $f(x)$ is increasing when $0 < \mathbf{x} < M$, and thus

$$\left(\frac{eM}{n}\right)^n \leq f(\varepsilon\, M + C_{10} + \omega) < \left(\frac{e}{\varepsilon}\right)^{\varepsilon\, M + C_{10} + \omega}$$

when $n < \varepsilon\, M + C_{10} + \omega$. On the other hand, If $M > \frac{(2+\varepsilon_1)(C_{10}+\omega)}{\varepsilon\,\varepsilon_1}$, then we have $\varepsilon\, M + C_{10} + \omega < (1 + \frac{\varepsilon_1}{2})(\varepsilon\, M + C_{10} + \omega) \leq (1+\varepsilon_1)\,\varepsilon\, M$ (by the way, to ensure that $\varepsilon\, M + C_{10} + \omega < M$, we need $\varepsilon_1 \leq \frac{1}{\varepsilon} - 1$ but this is trivial since at beginning we can choose ε_1 to be sufficiently small.) and so

$$\binom{M}{n} < e^{(1-\log\varepsilon)(1+\varepsilon_1)\,\varepsilon\, M}.$$

Moreover, since $n < \varepsilon\, M + C_{10} + \omega < (1+\varepsilon_1)\,\varepsilon\, M$ and $K = \left\lfloor \frac{(\omega-\varepsilon)M}{\theta} \right\rfloor > \frac{(\omega-\varepsilon)M}{\theta} - 1$, and $N = \left\lfloor \frac{1}{\varepsilon} - 1 \right\rfloor > \frac{1}{\varepsilon} - 2$, so we have

$$\begin{aligned} K - n &> \frac{(\omega-\varepsilon)M}{\theta} - 1 - (1+\varepsilon_1)\,\varepsilon\, M \\ &= \big(2 - \frac{1}{N+1} - \frac{\varepsilon}{\theta} - \varepsilon - \varepsilon\,\varepsilon_1\big)M - 1 \\ &> \big(2 - \frac{1}{\frac{1}{\varepsilon}-1} - \frac{\varepsilon}{\frac{1}{\varepsilon}-1} - \varepsilon - \varepsilon\,\varepsilon_1\big)M - 1 \\ &\geq (2 - 3\,\varepsilon - \varepsilon\,\varepsilon_1)M \end{aligned} \tag{6.2}$$

when $M \geq \frac{1}{(\frac{1}{2}+\varepsilon)\,\varepsilon}$, here for convenience we can use another constant M_4 to settle down all the "explicit tiny errors" in the computation so that when $M > M_4$ there is no trouble to explain these problems along the way, at this moment $M_4 = M_4(\varepsilon)$. Now we assume that $M^* \geq \frac{\log(C_3C_4C_5)}{\varepsilon_1 \log b}$, and modify the constant M_4 into $M_4(\varepsilon, N, t)$ so that for all $M > M_4$, $\varepsilon\, M + C_{10} + \omega \leq \frac{K-M}{t+1}$. For such M and $n < \varepsilon\, M + C_{10} + \omega$ we can apply Lemma 5.1 to conclude that

$$\begin{aligned} d_{M+tn}w_n b^{M+tn}\left|P^{[n]}_{\overline{k}^*}(\alpha)\right| &< \big(b^{1+t\,\varepsilon}(C|\alpha|)^{\varepsilon\,\varepsilon_1}C_{11}\big)^M < (b^{1+t\,\varepsilon}C_{11})^M, \\ d_{M+tn}w_n b^{M+tn}\left|R^{[n]}_{\overline{s}^*}(\alpha)\right| &< \big(b^{1+t\,\varepsilon}(C|\alpha|)^{\varepsilon\,\varepsilon_1}C_{11}\big)^M (C|\alpha|^{K-n}) \\ &< (b^{1+t\,\varepsilon}C_{11}C|\alpha|^{2-3\,\varepsilon})^M. \end{aligned} \tag{6.3}$$

Note that the inequalities 6.3 hold for an additional condition, that is to modify M_4 again into $M_4(\varepsilon, N, t, \Phi, \Psi, \varepsilon_1)$ (but we know ε_1 is already chosen at the beginning as $\varepsilon_1(\eta, \varepsilon, \alpha)$)which is still explicit but very complicated to write down, and 6.2 is relatively simple routine to check after this error-neglect for $M > M_4$. In the second inequality we actually used 6.2.

This is the moment for conclusion, we now fix the constant

$$M_* = M_*(N, \varepsilon_1, \alpha) = \max\{M_2, M_3, M_4, \frac{(2+\varepsilon_1)(C_{10}+\omega)}{\varepsilon\,\varepsilon_1}, \frac{\log(C_3C_4C_5)}{\varepsilon_1 \log b}\}.$$

and set the constant

$$q_* = \max\{\frac{1}{2}e^{\frac{(1+t\,\varepsilon)\log b+\log C_{11}}{\eta_1}M_*}, e^{\frac{(\eta_1+1)\log 2+(1+t\,\varepsilon)\log b+\log C_{11}}{\eta-\eta_1}}\}.$$

Now let

$$M = \left\lfloor \frac{\eta_1 \log(2q)}{(1+t\,\varepsilon)\log b + \log C_{11}} \right\rfloor + 1,$$

we can easily check that $M > M_*$ for all $q > q_*$, and

$$M((1+t\,\varepsilon)\log b + \log C_{11}) < \eta \log q - \log 2$$

for all $q > q_*$, that is to say, we obtain sufficiently large and also appropriate M by choosing sufficiently large q, and hence the efficiency of the main theorem is totally determined by the efficiency of M_*, which is an important observation when we analyze the whole proof. Finally we consider a numerical form $r = qf_l(\alpha) - p$ for all natural numbers $q > q_*$ and $p \in \mathbb{Z}$. Apply Corollary 5.4, since $M > M_*$, there exists a natural number $n < \varepsilon M + C_{10} + \omega$ such that

$$qP^{[n]}_{\bar{\mathrm{s}}^*}(\alpha) + pP^{[n]}_{\bar{\mathrm{k}}^*}(\alpha) \neq 0.$$

That is to say, in writing

$$P^{[n]}_{\bar{\mathrm{k}}^*}(\alpha)r = -\big(qP^{[n]}_{\bar{\mathrm{s}}^*}(\alpha) + P^{[n]}_{\bar{\mathrm{k}}^*}(\alpha)\big) + q\big(P^{[n]}_{\bar{\mathrm{s}}^*}(\alpha) + P^{[n]}_{\bar{\mathrm{k}}^*}(\alpha)f_l(\alpha)\big), \tag{6.4}$$

one can use the common argument in Diophantine approximation that a nonzero integer has absolute value larger or equal to 1. By Lemma 5.1, $d_{M+tn}w_nP^{[n]}_{\bar{\mathrm{s}}^*}(z) \in \mathbb{Z}[z]$. By Lemma 3.6, $\deg\big(d_{M+tn}w_nP^{[n]}_{\bar{\mathrm{s}}^*}(z)\big) \leq M + tn - 1$. Combining these two facts, We have

$$d_{M+tn}w_nb^{M+tn}P^{[n]}_{\bar{\mathrm{s}}^*}\big(\frac{a}{b}\big) \in \mathbb{Z}\,.$$

On the other hand, we also have $\deg\big(d_{M+tn}w_nP^{[n]}_{\bar{\mathrm{k}}^*}(z)\big) \leq M + tn - 1$ and w_n is defined to be the least common denominator of all the rational coefficients of the polynomials $P^{[n]}_{\bar{\mathrm{k}}^*}(z) \in \mathbb{Q}[z]$, so

$$d_{M+tn}w_nb^{M+tn}P^{[n]}_{\bar{\mathrm{k}}^*}\big(\frac{a}{b}\big) \in \mathbb{Z}\,.$$

After multiplying $d_{M+tn}w_nb^{M+tn}$ to both sides of Eq. (6.4), and taking the absolute value on both sides, with the recurrence relation $R^{[n]}_{\bar{\mathrm{s}}^*}(\alpha) = P^{[n]}_{\bar{\mathrm{s}}^*}(\alpha) + P^{[n]}_{\bar{\mathrm{k}}^*}(\alpha)f_l(\alpha)$, we have

$$d_{M+tn}w_nb^{M+tn}\left|P^{[n]}_{\bar{\mathrm{k}}^*}(\alpha)r\right| = \left|\Pi + qd_{M+tn}w_nb^{M+tn}R^{[n]}_{\bar{\mathrm{s}}^*}(\alpha)\right| \tag{6.5}$$

where $\Pi \in \mathbb{Z}\backslash\{0\}$. From the triangular inequality $|A+B| \geq |A| - |B|$, one has that the right hand side of Eq. (6.5) is at least

$$1 - qd_{M+tn}w_nb^{M+tn}R^{[n]}_{\bar{\mathrm{s}}^*}(\alpha).$$

which is furthermore greater than

$$1 - q\Big(b^{1+t\,\varepsilon-(2-3\,\varepsilon)}C_{11}C|a|^{2-3\,\varepsilon}\Big)^M$$

by the estimation 6.3. Finally we obtain this right hand side is larger than $\frac{1}{2}$ which is to say that

$$|r| > \frac{1}{2}\big(d_{M+tn}w_nb^{M+tn}\left|P^{[n]}_{\bar{\mathrm{k}}^*}(\alpha)\right|\big)^{-1},$$

again by 6.3 we have

$$\begin{aligned}|r| &> \frac{1}{2}(b^{1+t\,\varepsilon}C_{11})^{-M} = \frac{1}{2}e^{-M(\log b(1+t\,\varepsilon)+\log C_{11})}\\ &> \frac{1}{2}e^{-(\eta\log q-\log 2)} = q^{-\eta}\end{aligned} \tag{6.6}$$

so we are done with the wanted inequality Eq. (1.6).

7. Overview of the proof by applying the graded Padé approximants

In this section we will review the idea and crucial techniques in the proof of the main theorem, in particular the reason of applying graded Padé approximants to a set of G-functions in the class $G(\mathbb{Q}, C, \Phi)$.

First of all, we apply Siegel's lemma to construct a graded Padé approximants with the order o of zero at $z = 0$ of each linear form as large as possible while controlling the heights of these approximants as small as possible. This step is purely a procession of solving linear equations in rational coefficients, and how well we can do at this step is totally determined by two things: The first one is the type of functions $f_j(z)$, no matter they are E-functions or G-functions, we can apply graded Padé approximants to them, and even apply Padé approximants of first kind, or of second kind to them, in that it is just a procession of solving linear equations, but the difference lies in that the result we obtain is dependant on precise data, like C and Φ of the class $G(\mathbb{Q}, C, \Phi)$, and it varies from E-functions to G-functions. The second point is the number of linear forms. For the first kind approximants we only have one linear form as the "target", but for the graded approximants of homogeneous degree N we have $\theta = \binom{N+m-1}{m-1}$ (here when $m = 2$ it is $N+1$) different "targets" $R_{\overline{\mathrm{s}}}(z)$ to be simultaneously worked out, and N can be arbitrarily large.

Secondly, with a certain kind of Padé approximants, one can always play "differential determinant" on it. Since f_j satisfy a system of linear differential equations, and it is easy to define a differential operator from it such that the type of linear forms are closed under this operator, so that we can always define objects like $P^{[n]}_{\overline{\mathrm{k}}}(z) \in \mathbb{Q}[z]$ and $R^{[n]}_{\overline{\mathrm{s}}}(z)$ and work out linear algebraic results on it. Siegel-Shidlovskii method gives us the standard way to treat the determinant of a functional matrix formed by $P^{[n]}_{\overline{\mathrm{k}}}(z)$, as well as the systematic theory to prove the full rank property of a numerical matrix formed by $P^{[n]}_{\overline{\mathrm{k}}}(\alpha)$, for n up to a upper bound which is not so efficient for computation since it is relied on the linear algebraic property of the matrix formed by Taylor coefficients of these $f_j(z)$. All of these aspects could be treated in a united way, and is not exclusive for graded Padé approximants or the class $G(\mathbb{Q}, C, \Phi)$.

Thirdly, when we look up the details of section 6 it is not surprising to find that except from some "constant-chasing" technical parts or "supporting arguments", what is unique for graded Padé approximants applying to $G(\mathbb{Q}, C, \Phi)$ is exactly the type of bounds appeared in Lemma 5.1, where $C_2^{(1+\varepsilon_1)M}$ is the main term. This is not surprising since the main term is already given by Lemma 3.3, what's really surprising, or say what's really the creative idea of applying Padé approximants is, the balance between the bounds type there with one term $(C|\alpha|)^{K-n}$ and the upper bound of the degree of the polynomial $P^{[n]}_{\overline{\mathrm{s}}^*}(z)$, namely $M + tn - 1$. The "asymptotic ratio", namely some kind of limit of $\frac{K}{K-M}$, where by "some kind" we mean that M in fact represents the degree upper bound of $P^{[n]}_{\overline{\mathrm{k}}}(z)$, and luckily in our paper it equals 2. This exactly produces the highlight of Zudilin's paper: to improve the scope of rational points $\alpha = \frac{a}{b}$ which are valid for the main theorem. We know that Zudilin has improved the relative scope to the extent of "order-2", expressed in more details as $\eta_0 > 0$, where the classical result is only for "order-$(m+1)$".

Last but not least, the difficulty is how to improve the Padé approximants application to let the limit of $\frac{K}{M}$ when $M \to \infty$ as large as possible. Here Zudilin used graded Padé approximants and obtained 2, but is it possible for us to get $2 + \delta$ for some $\delta > 0$? This is not easy to approach, since here M actually represents asymptotic notation of $M + tn - 1$, the order upper bound for $P^{[n]}_{\overline{\mathrm{k}}}(z)$. This is the so called "balance" problem when we talk about how to obtain a large value for $\frac{K}{M}$. If n need to be as large as cM (where c is not an arbitrary small constant like ε), so that we can ensure the full-rank property of numerical matrix, then the ratio could be smaller even if we take larger K with respect to M. This is the sophisticated reason to choose graded Padé approximants for obtaining our main theorem.

Finally let me conclude a little bit for the excluded situation in our main theorem. In fact, most of our discussions made for $\mathbb{C}(z)$ also can be worked out for $\mathbb{Q}(z)$, and in particular, in Eq. (4.5) we have

$D_{\overline{\mathrm{k}},\overline{\mathrm{k}}'}(z) \in \mathbb{Q}(z)$, and therefore in Eq. (4.10) we only require 1, $f_1(z)$, $f_2(z)$ and $f_2^2(z)$ to be linearly independent over $\mathbb{Q}(z)$ to make our proof valid. Hence we can replace $\mathbb{C}(z)$ with $\mathbb{Q}(z)$ in Lemma 1.12 and in the proof of Lemma 1.12, thus making the excluded situation in Theorem 1.10 so special, as the values of such excluded functions at rational points are algebraic numbers or degree at most 4, in particular, by Roth's theorem we know the measure of irrationality of such values are exactly 2 if they are irrational, so even without such exclusion, we can say they almost satisfy the original conclusion in Zudilin's paper.

Acknowledgement

It is my great pleasure to thank Tanguy Rivoal for motivating me studying the Diophantine approximation to the values of G-functions at certain rational points, and giving me lots of suggestion on choosing the problems and main target, even on the details about writing style, as well as encouraging me a lot in more systematic study along the way. I also thank well Stéphane Fischler for assisting this project in many aspects.

I gratefully acknowledge financial support from the ALGANT Erasmus Mundus Program in Algebra, Geometry, and Number Theory, and from AAG of Université Paris-Sud.

References

[1] A.B.Shidlovskii. *Transcendental Numbers*. Walter de Gruyter, 1989.

[2] A.I.Galochkin. Estimates from below of polynomials in the values of analytic functions of a certain class. *Sbornik: Mathematics*, 95(3):396–417, 1974.

[3] E.Bombieri. *On G-functions, In: Recent Progress in Analytic Number Theory, vol 2*. Academic Press, 1981.

[4] V.V.Zudilin. On rational approximation of values of a certain class of entire functions. *Sbornik: Mathematics*, 186(4):555–590, 1995.

[5] V.V.Zudilin. On a measure of irrationality for values of g-functions. *Isvestiya: Mathematics*, 60 (1):91–118, 1996.

[6] Y.André. *G-functions and Geometry*. Vieweg+Teubner Verlag, 1989.