Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Advances in Mathematics 204 (2006) 448–480

ADVANCES IN Mathematics

www.elsevier.com/locate/aim

On rational approximation of algebraic functions

Julius Borcea*, Rikard Bøgvad, Boris Shapiro

Department of Mathematics, Stockholm University, SE-106 91 Stockholm, Sweden

Received 21 September 2004; accepted 9 June 2005

Communicated by Michael Hopkins
Available online 8 August 2005

Abstract

We construct a new scheme of approximation of any multivalued algebraic function $f(z)$ by a sequence $\{r_n(z)\}_{n\in\mathbb{N}}$ of rational functions. The latter sequence is generated by a recurrence relation which is completely determined by the algebraic equation satisfied by $f(z)$. Compared to the usual Padé approximation our scheme has a number of advantages, such as simple computational procedures that allow us to prove natural analogs of the Padé Conjecture and Nuttall's Conjecture for the sequence $\{r_n(z)\}_{n\in\mathbb{N}}$ in the complement $\mathbb{CP}^1\setminus\mathcal{D}_f$, where $\mathcal{D}_f$ is the union of a finite number of segments of real algebraic curves and finitely many isolated points. In particular, our construction makes it possible to control the behavior of spurious poles and to describe the asymptotic ratio distribution of the family $\{r_n(z)\}_{n\in\mathbb{N}}$. As an application we settle the so-called 3-conjecture of Egecioglu et al. dealing with a 4-term recursion related to a polynomial Riemann Hypothesis.

MSC: Primary 30E10; secondary 41A20; 42C05; 82B05

Keywords: Finite recursions; Asymptotic ratio distribution; Padé approximation

1. Introduction and main results

Rational approximants of analytic functions and the asymptotic distribution of their zeros and poles are of central interest in many areas of mathematics and physics. For

* Corresponding author.
E-mail addresses: julius@math.su.se (J. Borcea), rikard@math.su.se (R. Bøgvad), shapiro@math.su.se (B. Shapiro).

doi:10.1016/j.aim.2005.06.002

the class of algebraic functions these questions have attracted special attention due to their important applications ranging from the convergence theory of Padé approximants [30–35] and the theory of general orthogonal polynomials [16,26,36,38] to statistical mechanics [28,29], complex Sturm–Liouville problems [6], inverse scattering theory and quantum field theory [2].

The main purpose of this paper is to give a simple and direct construction of families of rational functions converging to a certain branch of an arbitrary (multivalued) algebraic function $f(z)$. While the usual Padé approximation requires the knowledge of the Taylor expansion of $f(z)$ at ∞, our scheme is based only on the algebraic equation satisfied by $f(z)$ and has therefore an essentially different range of applications. Indeed, let

$$P(y,z)=\sum_{i=0}^{k} P_{k-i}(z)y^{k-i} \tag{1.1}$$

denote the irreducible polynomial in (y,z) defining $f(z)$, that is, $P(f(z),z)=0$. Note that $P(y,z)$ is uniquely defined up to a scalar factor. Let us rewrite (1.1) as

$$-y^k=\sum_{i=1}^{k}\frac{P_{k-i}(z)}{P_k(z)}\,y^{k-i} \tag{1.2}$$

and consider the associated recursion of length $k+1$ with rational coefficients

$$-q_n(z)=\sum_{i=1}^{k}\frac{P_{k-i}(z)}{P_k(z)}\,q_{n-i}(z). \tag{1.3}$$

Choosing any initial k-tuple of rational functions $IN=\{q_0(z),\dots,q_{k-1}(z)\}$ one can generate a family $\{q_n(z)\}_{n\in\mathbb{N}}$ of rational functions satisfying (1.3) for all $n\geqslant k$ and coinciding with the entries of IN for $0\leqslant n\leqslant k-1$. The main object of study of this paper is the family $\{r_n(z)\}_{n\in\mathbb{N}}$, where $r_n(z)=\frac{q_n(z)}{q_{n-1}(z)}$.

In order to formulate our results we need several additional definitions. Consider first a usual recurrence relation of length $k+1$ with constant coefficients

$$-u_n=\alpha_1u_{n-1}+\alpha_2u_{n-2}+\dots+\alpha_ku_{n-k}, \tag{1.4}$$

where $\alpha_k\neq 0$.

Definition 1. The *asymptotic symbol equation* of recurrence (1.4) is given by

$$t^k+\alpha_1t^{k-1}+\alpha_2t^{k-2}+\dots+\alpha_k=0. \tag{1.5}$$

The left-hand side of the above equation is called the *characteristic polynomial* of recurrence (1.4). Denote the roots of (1.5) by $\tau_1, \ldots, \tau_k$ and call them the *spectral numbers* of the recurrence.

Definition 2. Recursion (1.4) and its characteristic polynomial are said to be of *dominant type* or *dominant* for short if there exists a unique (simple) spectral number $\tau_{\max}$ of this recurrence relation satisfying $|\tau_{\max}| = \max_{1 \leqslant i \leqslant k} |\tau_i|$. Otherwise (1.4) and (1.5) are said to be of *nondominant type* or just *nondominant*. The number $\tau_{\max}$ will be referred to as the *dominant spectral number* of (1.4) or the *dominant root* of (1.5).

The following theorem may be found in [37, Chapter 4].

Theorem 1. *Let $k \in \mathbb{N}$ and consider a k-tuple $(\alpha_1, \ldots, \alpha_k)$ of complex numbers with $\alpha_k \neq 0$. For any function $u : \mathbb{Z}_{\geqslant 0} \to \mathbb{C}$ the following conditions are equivalent*:

(i) $\sum_{n \geqslant 0} u_n t^n = \frac{Q_1(t)}{Q_2(t)}$, *where $Q_2(t) = 1 + \alpha_1 t + \alpha_2 t^2 + \cdots + \alpha_k t^k$ and $Q_1(t)$ is a polynomial in t whose degree is smaller than k.*
(ii) *For all $n \geqslant k$ the numbers u_n satisfy the $(k+1)$-term recurrence relation given by* (1.4).
(iii) *For all $n \geqslant 0$ one has*

$$u_n = \sum_{i=1}^{r} p_i(n) \tau_i^n, \tag{1.6}$$

where $\tau_1, \ldots, \tau_r$ are the distinct spectral numbers of (1.4) *with multiplicities $m_1, \ldots, m_r$, respectively, and $p_i(n)$ is a polynomial in the variable n of degree at most $m_i - 1$ for $1 \leqslant i \leqslant r$.*

Note that by Definition 2 the dominant spectral number $\tau_{\max}$ of any dominant recurrence relation has multiplicity one.

Definition 3. An initial k-tuple of complex numbers $\{u_0, u_1, \ldots, u_{k-1}\}$ is called *fast growing* with respect to a given dominant recurrence of form (1.4) if the coefficient $\kappa_{\max}$ of $\tau_{\max}^n$ in (1.6) is nonvanishing, that is, $u_n = \kappa_{\max} \tau_{\max}^n + \cdots$ with $\kappa_{\max} \neq 0$. Otherwise the k-tuple $\{u_0, u_1, \ldots, u_{k-1}\}$ is said to be *slow growing*.

Let $\mathcal{P}_k = \left\{a_k y^k + a_{k-1} y^{k-1} + \cdots + a_0 \mid a_i \in \mathbb{C}, 0 \leqslant i \leqslant k\right\}$ denote the linear space of all polynomials of degree at most k with complex coefficients.

Definition 4. The real hypersurface $\Xi_k \subset \mathcal{P}_k$ obtained as the closure of the set of all nondominant polynomials is called the *standard equimodular discriminant*. For any family

$$\Gamma(y, z_1, \ldots, z_q) = \left\{a_k(z_1, \ldots, z_q) y^k + a_{k-1}(z_1, \ldots, z_q) y^{k-1} + \cdots + a_0(z_1, \ldots, z_q)\right\}$$

of irreducible polynomials of degree at most k in the variable y we define the *induced equimodular discriminant* Ξ_Γ to be the set of all parameter values $(z_1, \ldots, z_q) \in \mathbb{C}^q$ for which the corresponding polynomial in y is nondominant. Given an algebraic function $f(z)$ defined by (1.1) we denote by Ξ_f the induced equimodular discriminant of (1.2) considered as a family of polynomials in the variable y.

Example 1. For $k = 2$ the equimodular discriminant $\Xi_2 \subset \mathcal{P}_2$ is the real hypersurface consisting of all solutions to $\varepsilon a_1^2 - 4a_0a_2 = 0$, where ε is a real parameter with values in $[1, \infty)$. More information on Ξ_k can be found in [10,12].

Definition 5. Given an algebraic function $f(z)$ defined by (1.1) and an initial k-tuple of rational functions $IN = \{q_0(z), \ldots, q_{k-1}(z)\}$ we define the *pole locus* $\Upsilon_{f,IN}$ associated with the data (f, IN) to be the union between the zero set of the polynomial $P_k(z)$ and the sets of all poles of $q_i(z)$ for $0 \leqslant i \leqslant k-1$.

As we explain in Section 3, one of the fascinating features of Padé approximants is the complexity of their convergence theory. The main challenges of this theory are the Padé (Baker–Gammel–Wills) Conjecture and Nuttall's Conjecture (cf. [2,35]). The general version of the former was recently disproved by Lubinsky [24]. Subsequently, Buslaev [15] constructed counterexamples to this conjecture for some special algebraic (hyperelliptic) functions. Nevertheless, a question of central interest in many applications is whether the Padé Conjecture could hold for certain classes of algebraic functions. The first main result of this paper shows that a natural analog of the Padé Conjecture is always true for our approximation scheme:

Theorem 2. *Let $f(z)$ be an algebraic function defined by* (1.1). *For any initial k-tuple of rational functions $IN = \{q_0(z), \ldots, q_{k-1}(z)\}$ there exists a finite set $\Sigma_{f,IN} \subset \mathbb{CP}^1 \setminus (\Xi_f \cup \Upsilon_{f,IN})$ such that*

$$r_n(z) = \frac{q_n(z)}{q_{n-1}(z)} \rightrightarrows y_{\mathrm{dom}}(z) \quad \text{in } \mathbb{CP}^1 \setminus \mathcal{D}_f \text{ as } n \to \infty,$$

where $y_{\mathrm{dom}}(z)$ is the dominant root of Eq. (1.2), *$\mathcal{D}_f = \Xi_f \cup \Upsilon_{f,IN} \cup \Sigma_{f,IN}$, and $\rightrightarrows$ stands for uniform convergence on compact subsets of $\mathbb{CP}^1 \setminus \mathcal{D}_f$.*

The set $\Sigma_{f,IN}$ consists precisely of those points $z \in \Omega$ such that the initial k-tuple $IN = \{q_0(z), \ldots, q_{k-1}(z)\}$ is slowly growing with respect to recurrence (1.3) evaluated at z (cf. Definition 3). This motivates the following definition.

Definition 6. The set $\Sigma_{f,IN}$ is called the *set of slow growth* associated with the data (f, IN).

Although $\Sigma_{f,IN}$ is the set of solutions of an implicit equation (cf. (2.13)) it turns out that the set of slow growth is actually a subset of a certain discriminantal set $\mathcal{S}$ which is given explicitly in terms of the data (f, IN) (see the proof of Lemma 6).

Remark 1. In the final stage of preparation of this paper the authors noticed that some weaker results similar to Theorem 2 appeared in [7,8]. The latter deal with pointwise limits of certain families of recursively defined polynomials and are often used in modern statistical physics, see, e.g., [28] and references therein.

The next result describes the rate of convergence of the sequence $\{r_n(z)\}_{n\in\mathbb{N}}$ to $y_{\mathrm{dom}}(z)$, which proves in particular the uniform convergence stated in Theorem 2. Given $0 < \varepsilon \ll 1$ set $\Theta_\varepsilon = \mathbb{CP}^1 \setminus \mathcal{O}_\varepsilon$, where $\mathcal{O}_\varepsilon$ is the ε-neighborhood of $\mathcal{D}_f = \Xi_f \cup \Upsilon_{f,IN} \cup \Sigma_{f,IN}$ in the spherical metric on $\mathbb{CP}^1$.

Theorem 3. *For any sufficiently small* $\varepsilon > 0$ *the rate of convergence of* $r_n(z) \rightrightarrows y_{\mathrm{dom}}(z)$ *in* Θ_ε *is exponential, that is, there exist constants* $\mathfrak{M} > 0$ *and* $q \in (0, 1)$ *such that* $|r_n(z) - y_{\mathrm{dom}}(z)| \leqslant \mathfrak{M}q^n$ *for all* $z \in \Theta_\varepsilon$.

Definition 7. Given a meromorphic function g in some open set $\Omega \subseteq \mathbb{C}$ we construct its (complex-valued) *residue distribution* v_g as follows. Let $\{z_m \mid m \in \mathbb{N}\}$ be the (finite or infinite) set of all the poles of g in Ω. Assume that the Laurent expansion of g at z_m has the form $g(z) = \sum_{-\infty<l\leqslant l_m} \frac{A_{m,l}}{(z-z_m)^l}$. Then the distribution v_g is given by

$$v_g = \sum_{m\geqslant 1} \left(\sum_{1\leqslant l\leqslant l_m} \frac{(-1)^{l-1}}{(l-1)!} A_{m,l} \frac{\partial^{l-1}}{\partial z^{l-1}} \delta_{z_m} \right), \tag{1.7}$$

where δ_{z_m} is the Dirac mass at z_m and the sum in the right-hand side of (1.7) is meaningful as a distribution in Ω since it is locally finite in Ω.

Remark 2. The distribution v_g is a complex-valued *measure* if and only if g has all simple poles, see [9, p. 250]. If the latter holds, then in the notation of Definition 7 the value of this complex measure at z_m equals $A_{m,1}$, i.e., the residue of g at z_m.

Definition 8. Given an integrable complex-valued distribution ρ in $\mathbb{C}$ we define its Cauchy transform $\mathcal{C}_\rho(z)$ as

$$\mathcal{C}_\rho(z) = \int_{\mathbb{C}} \frac{d\rho(\xi)}{z-\xi}.$$

It is not difficult to see that the distribution ρ itself may be restored from its Cauchy transform $\mathcal{C}_\rho(z)$ by

$$\rho = \frac{1}{\pi} \frac{\partial \mathcal{C}_\rho}{\partial \bar{z}}, \tag{1.8}$$

where $\mathcal{C}_\rho(z)$ and its $\frac{\partial}{\partial \bar{z}}$-derivative are understood as distributions. As is well-known, any meromorphic function g defined in the whole complex plane $\mathbb{C}$ is the Cauchy transform of its residue distribution ν_g if the condition $\int_{\mathbb{C}} d\nu_g(\xi) < \infty$ holds, see, e.g., [9, p. 261].

Definition 9. Given a family $\{\phi_n(z)\}_{n\in\mathbb{N}}$ of smooth functions defined in some open set $\Omega \subseteq \mathbb{C}$ one calls the limit $\Phi(z) = \lim_{n\to\infty} \frac{\phi_{n+1}(z)}{\phi_n(z)}$ the *asymptotic ratio* of the family, provided that this limit exists in some open subset of Ω. If $\{\phi_n(z)\}_{n\in\mathbb{N}}$ consists of analytic functions and ν_n denotes the residue distribution of the meromorphic function $\frac{\phi_{n+1}(z)}{\phi_n(z)}$ in Ω, then the limit $\nu = \lim_{n\to\infty} \nu_n$ (if it exists in the sense of weak convergence) is called the *asymptotic ratio distribution* of the family.

Our next main result describes the support and the density of the asymptotic ratio distribution associated with sequences of rational functions constructed by using our approximation scheme.

Theorem 4. *Let $f(z)$ be an algebraic function defined by* (1.1) *and fix a generic initial k-tuple of rational functions $IN = \{q_0(z), \dots, q_{k-1}(z)\}$. If ν_n denotes the residue distribution of $r_n(z)$ and $\nu = \lim_{n\to\infty} \nu_n$ is the asymptotic ratio distribution of the family $\{q_n(z)\}_{n\in\mathbb{N}}$ then the following holds*:

(i) *The support of ν does not depend on the set of slow growth $\Sigma_{f,IN}$. More precisely, the asymptotic ratio distribution ν vanishes in a sufficiently small neighborhood of each point in $\Sigma_{f,IN}$.*
(ii) *Suppose that there exists a nonisolated point $z_0 \in \Xi_f$ such that Eq.* (1.2) *considered at z_0 has the property that among its roots with maximal absolute value there are at least two with the same maximal multiplicity* (*compare with Definition* 10). *If the sequence $\{r_n(z_0)\}_{n\in\mathbb{N}}$ diverges then $\operatorname{supp} \nu = \Xi_f$.*
(iii) *One has*

$$\nu = \frac{1}{\pi}\frac{\partial y_{\text{dom}}}{\partial \bar{z}} \iff y_{\text{dom}}(z) = \int_{\mathbb{C}} \frac{d\nu(\xi)}{z-\xi}.$$

As a consequence of Theorems 2–4 we obtain the following asymptotic classification of the set of poles of the rational approximants $\{r_n(z)\}_{n\in\mathbb{N}}$ (see Fig. 1):

Corollary 1. *For any algebraic function $f(z)$ and any initial k-tuple IN of rational functions the set of all poles of the family $\{r_n(z)\}_{n\in\mathbb{N}}$ splits asymptotically into the following three types*:

(i) *The fixed part consisting of a subset of $\Upsilon_{f,IN} \setminus (\Xi_f \cup \Sigma_{f,IN})$.*
(ii) *The regular part tending asymptotically to the finite union of curves Ξ_f, that is, the induced equimodular discriminant of* (1.2) (*cf. Theorem* 2).
(iii) *The spurious part tending to the* (*finite*) *set of slow growth $\Sigma_{f,IN}$.*

The usage of the name “spurious” in Corollary 1(iii) follows the conventional way of describing poles of rational approximants to f that do not correspond to analytic

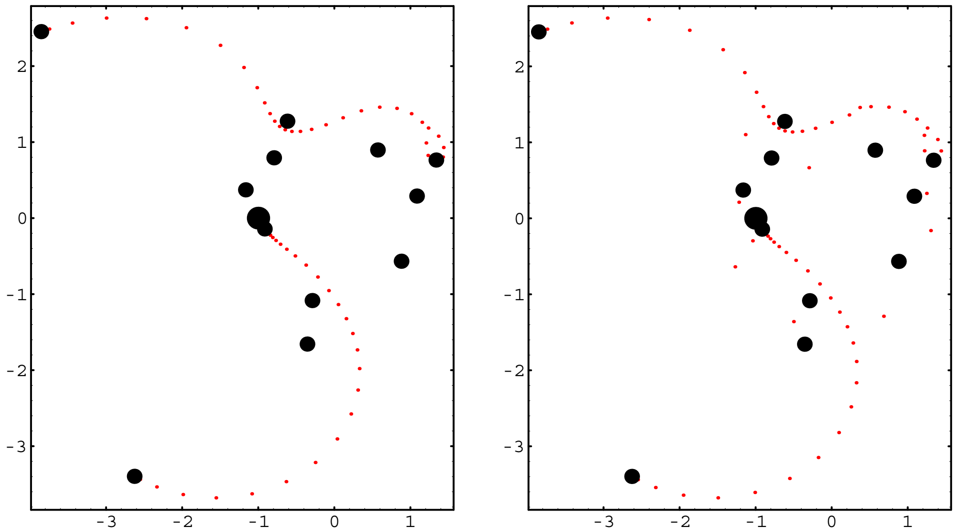

Fig. 1. Poles of $r_{31}(z)$ approximating the branch with maximal absolute value of the algebraic function $f(z)$ with defining equation $(z+1)y^3 = (z^2+1)y^2 + (z-5I)y + (z^3-1-I)$ for two different choices of initial triples.

properties of the function f. This terminology has been widely adopted in Padé approximation, see, e.g., [35]. The rigorous definition of spurious poles is rather technical, which is why we postpone it until Section 3.

Explanations to Fig. 1. The two pictures show the poles of $r_{31}(z)$ for the initial triples $p_{-2}(z) = p_{-1}(z) = 0$, $p_0(z) = 1$ and $p_{-2}(z) = z^5 + Iz^2 - 5$, $p_{-1}(z) = z^3 - z + I$, $p_0(z) = 1$, respectively (the second initial triple was picked randomly). The large fat point in these pictures is the unique pole $z = -1$ of f which in this case coincides with $\Upsilon_{f,IN}$ since the chosen initial triples are polynomial. The smaller fat points on both pictures are the branching points of $f(z)$. One can show that the curve segments belonging to Ξ_f can only end at these branching points, see [10,12]. Note that the poles form the same pattern on both pictures. This pattern is close to the curve segments constituting Ξ_f and contains an additional nine isolated points in the right picture. These nine isolated points form the set of slow growth in this case; one can also check numerically that the latter set is indeed a subset of the finite discriminantal set S described in the proof of Lemma 6.

The final result of the paper settles a natural analog of Nuttall's Conjecture (see Section 3 and [35]) for our approximation scheme:

Theorem 5. *For any algebraic function $f(z)$ and any initial k-tuple IN of rational functions there exists a finite upper bound for the total number of spurious poles of the approximants $\{r_n(z)\}_{n\in\mathbb{N}}$ associated with the data (f, IN).*

As it was pointed out in [35], there is strong evidence that there always exists some infinite subsequence of the sequence of diagonal Padé approximants to a given analytic function for which there are no spurious poles in a certain convergence domain. It is interesting to note that for the standard choice of initial k-tuple $IN = \{0, 0, \ldots, 0, 1\}$ in our approximation scheme none of the rational functions in the resulting sequence of approximants $\{r_n(z)\}_{n\in\mathbb{N}}$ has spurious poles, see Corollary 2.

Let us conclude this introduction with a few remarks on the approximation scheme proposed above and some related topics. In the present setting the pointwise convergence of the sequence of rational approximants $\{r_n(z)\}_{n\in\mathbb{N}}$ follows almost immediately from Theorem 1 (cf. Lemma 8). This is simply because our main recursion formula (1.3) may actually be viewed as a recurrence relation with constant coefficients depending algebraically on a complex variable z. However, in many important cases one has to invoke quite different arguments in order to derive similar (pointwise) convergence results. For instance, the major challenge in the theory of finite difference equations is related to recurrence relations with varying coefficients. In this case the fundamental theorems of Poincaré and Perron (see, e.g., [19, pp. 287–298]) serve as prototypes for many new developments. Some interesting generalizations of the Poincaré and Van Vleck theorems were recently obtained in [13,14]. Typical applications of this sort of results pertain to the ratio asymptotics of various kinds of orthogonal polynomials and may be found in the well-known paper [27]. We refer to [3,4] for more recent results in this direction. In particular, in loc. cit. the authors study the asymptotics of the zeros of a family of polynomials satisfying a usual 3-term recurrence relation with varying and stabilizing complex coefficients and prove that these zeros concentrate on an interval in the complex plane. We continue this line of research in a forthcoming publication [12], where we study the asymptotics of general families of functions satisfying functional recurrence relations with varying and stabilizing coefficients and obtain appropriate extensions of Poincaré's theorem.

This paper is organized as follows. In Section 2 we give all the proofs, Section 3 contains a short description of Padé approximants and their spurious poles while Section 4 is devoted to proving the 3-conjecture of Egecioglu, Redmond and Ryavec. Finally, in Section 5 we compare our approximation scheme with Padé approximation and discuss a number of related topics and open problems.

2. Proofs and further consequences

In order to prove Theorems 2 and 3 we need several preliminary results. Let us first fix the following notation:

Notation 1. Let Ω be a domain in $\mathbb{C}$ and $T : \Omega \ni z \mapsto T(z) = (t_{ij}(z)) \in M_k(\mathbb{C})$ be a $k \times k$ matrix-valued map. Denote by $\chi_T : \Omega \ni z \mapsto \chi_T(\zeta, z) \in \mathbb{C}[\zeta]$ the characteristic polynomial map associated with T, that is, for each $z \in \Omega$ the characteristic polynomial of $T(z)$ is given by the polynomial $\chi_T(\zeta, z)$ in the variable ζ with coefficients which are complex-valued functions of z. Let further $\Delta_T \subset \Omega$ be the discriminant surface associated with T, i.e., the set of all $z \in \Omega$ for which the (usual) discriminant of the polynomial $\chi_T(\zeta, z)$ vanishes. Finally, let Ξ_T denote the equimodular discriminant associated with T, that is, the induced equimodular discriminant of the family of polynomials $\{\chi_T(\zeta, z) \mid z \in \Omega\}$ (cf. Definition 4).

The following lemma is a consequence of well-known results on the diagonalization of analytic matrices which are based essentially on the implicit function theorem (see [20, p. 106]).

Lemma 1. *If the map T is analytic in Ω then the eigenvalues $\lambda_i(z)$, $1 \leqslant i \leqslant k$, of $T(z)$ are analytic functions in $\Omega \backslash \Delta_T$. Furthermore, there exists an analytic map $A : \Omega \backslash \Delta_T \ni z \mapsto A(z) \in GL_k(\mathbb{C})$ such that*

$$T(z) = A(z)\,\mathrm{diag}(\lambda_1(z), \ldots, \lambda_k(z))A(z)^{-1}, \quad z \in \Omega \setminus \Delta_T.$$

The map $A(z)$ is uniquely determined in $\Omega \setminus \Delta_T$ up to a nonvanishing complex-valued analytic function $\gamma(z)$. Moreover, its column vectors $A(z)^{(1)}, \ldots, A(z)^{(k)}$ are eigenvectors of $T(z)$ with eigenvalues $\lambda_1(z), \ldots, \lambda_k(z)$, respectively, while the row vectors of its inverse $A(z)^{-1}_{(1)}, \ldots, A(z)^{-1}_{(k)}$ are eigenvectors of the transpose $T(z)^t$ with eigenvalues $\lambda_1(z), \ldots, \lambda_k(z)$, respectively.

Lemma 2. *Under the assumptions of Lemma* 1 *the following holds*:

(i) *For any $\mathbf{x} = (x_{k-1}, \ldots, x_0)^t \in \mathbb{C}^k$ there exist complex-valued functions $\alpha_{ij}(z, \mathbf{x})$, $1 \leqslant i, j \leqslant k$, which are analytic in $\Omega \setminus \Delta_T$ such that*

$$T(z)^n \mathbf{x} = \left(\sum_{j=1}^{k} \alpha_{1j}(z, \mathbf{x})\lambda_j(z)^n, \ldots, \sum_{j=1}^{k} \alpha_{kj}(z, \mathbf{x})\lambda_j(z)^n \right)^t$$

$$\text{for } z \in \Omega \setminus \Delta_T \text{ and } n \in \mathbb{N}. \tag{2.1}$$

(ii) *For $z \in \Omega \setminus \Xi_T$ let $\lambda(z) = \lambda_1(z)$ denote the dominant eigenvalue of $T(z)$. There exist $\mathbb{C}^k$-valued analytic functions $\mathbf{u}(z) = (u_1(z), \ldots, u_k(z))^t$ and $\mathbf{v}(z) = (v_1(z), \ldots, v_k(z))^t$ defined on $\Omega \setminus \Delta_T$ such that*

$$\mathbf{v}(z)^t \mathbf{u}(z) = 1 \quad \text{for } z \in \Omega \setminus \Delta_T$$

$$\frac{T(z)^n}{\lambda(z)^n} \rightrightarrows \mathbf{u}(z)\mathbf{v}(z)^t \quad \text{in } \Omega \setminus (\Delta_T \cup \Xi_T) \text{ as } n \to \infty. \tag{2.2}$$

Thus, $\left\{\frac{T(z)^n}{\lambda(z)^n}\right\}_{n\in\mathbb{N}}$ *is a sequence of analytic matrices which converges locally uniformly in* $\Omega\setminus(\Delta_T\cup\Xi_T)$ *to* $\mathbf{u}(z)\mathbf{v}(z)^t$ *in the sup-norm on* $M_k(\mathbb{C})$.

(iii) *In the notations of* (i) *and* (ii) *one has*

$$\alpha_{i1}(z,\mathbf{x}) = u_i(z)\sum_{j=1}^{k} v_j(z)x_{k-j}, \quad 1\leqslant i\leqslant k, \tag{2.3}$$

for any $\mathbf{x} = (x_{k-1},\dots,x_0)^t \in \mathbb{C}^k$.

Proof. Part (i) is a direct consequence of Lemma 1. To prove (ii) recall the notations of Lemma 1 and set $\mathbf{u}(z) = A(z)^{(1)}$ and $\mathbf{v}(z) = A(z)^{-1}_{(1)}$, so that $\mathbf{v}(z)^t\mathbf{u}(z) = 1$ for all $z\in\Omega\setminus\Delta_T$. From Lemma 1 one gets

$$\frac{T(z)^n}{\lambda(z)^n} = A(z)\,\mathrm{diag}\left(1, \left(\frac{\lambda_2(z)}{\lambda(z)}\right)^n, \dots, \left(\frac{\lambda_k(z)}{\lambda(z)}\right)^n\right)A(z)^{-1} \tag{2.4}$$

for all $z\in\Omega\setminus(\Delta_T\cup\Xi_T)$ and $n\in\mathbb{N}$, which immediately proves (2.2) since by assumption $|\lambda_j(z)| < |\lambda(z)|$ if $2\leqslant j\leqslant k$. Finally, (2.3) follows from (i) and (ii) by elementary computations. $\square$

Recall now the polynomials $P_i(z)$, $0\leqslant i\leqslant k$, from (1.1) and define the rational functions

$$R_i(z) = -\frac{P_i(z)}{P_k(z)}, \quad 0\leqslant i\leqslant k-1. \tag{2.5}$$

Choosing any initial k-tuple of rational functions $IN = \{q_0(z),\dots,q_{k-1}(z)\}$ one can rewrite recursion (1.3) as

$$q_n(z) = \sum_{i=1}^{k} R_{k-i}(z)q_{n-i}(z) \quad \text{for } n\geqslant k \tag{2.6}$$

or, equivalently,

$$\mathbf{q}_n(z) = T(z)^{n-k+1}\mathbf{q}_{k-1}(z) \quad \text{for } n\geqslant k,$$

where

$$\mathbf{q}_m(z) = (q_m(z), q_{m-1}(z), \ldots, q_{m-k+1}(z))^t \in \mathbb{C}^k \quad \text{for } m \geqslant k-1$$

$$T(z) = \begin{pmatrix} R_{k-1}(z) & R_{k-2}(z) & R_{k-3}(z) & \ldots & R_0(z) \\ 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}. \tag{2.7}$$

We now apply Lemmas 1 and 2 to this specific case. It is easy to see that the characteristic polynomial of $T(z)$ is given by

$$\chi_T(\zeta, z) = \zeta^k - \sum_{i=0}^{k-1} R_{k-i-1}(z)\zeta^{k-i-1}. \tag{2.8}$$

Moreover, in view of Theorem 2, Notation 1 and Lemma 2(ii) one has

$$y_{\text{dom}}(z) = \lambda(z) \quad \text{for } z \in \mathbb{C} \setminus (\Upsilon_{f,IN} \cup \Xi_f) \text{ where } \Xi_f = \Xi_T. \tag{2.9}$$

Remark 3. Identity (2.9) implies in particular that $y_{\text{dom}}(z)$ is a nonvanishing analytic function in $\mathbb{C} \setminus (\Upsilon_{f,IN} \cup \Xi_f)$.

Remark 4. The discriminant Δ_T associated with the map $T(z)$ given by (2.7) is a finite subset of $\mathbb{C}$ since all the entries of $T(z)$ are rational functions of z (cf. (2.5)).

It turns out that in our specific case the vector-valued functions $\mathbf{u}(z)$ and $\mathbf{v}(z)$ defined in Lemma 2(ii) can be computed explicitly.

Lemma 3. *For $z \in \mathbb{C} \setminus (\Upsilon_{f,IN} \cup \Xi_f)$ let $\mathbf{u}(z)$ and $\mathbf{v}(z)$ be eigenvectors with eigenvalue $\lambda(z)$ of $T(z)$ and $T(z)^t$, respectively. There exist complex-valued functions $\gamma_1(z), \gamma_2(z)$ which are analytic and nonvanishing in $\mathbb{C} \setminus (\Upsilon_{f,IN} \cup \Xi_f)$ such that*

$$\mathbf{u}(z) = \gamma_1(z)(u_1(z), \ldots, u_k(z))^t \quad \text{where } u_i(z) = \lambda(z)^{k-i}, \ \ 1 \leqslant i \leqslant k,$$

$$\mathbf{v}(z) = \gamma_2(z)(v_1(z), \ldots, v_k(z))^t \quad \text{where } v_i(z) = \sum_{j=i}^{k} R_{k-j}(z)\lambda(z)^{i-j}, \ \ 1 \leqslant i \leqslant k.$$

In particular, the maps $\mathbf{u}(z)$, $\mathbf{v}(z)$, and $\mathbf{u}(z)\mathbf{v}(z)^t$ are all analytic in $\mathbb{C} \setminus (\Upsilon_{f,IN} \cup \Xi_f)$. Moreover, the functions $\gamma_1(z), \gamma_2(z)$ may be chosen so that $\mathbf{v}(z)^t\mathbf{u}(z) = 1$ for all $z \in \mathbb{C} \setminus (\Upsilon_{f,IN} \cup \Xi_f)$.

Proof. Note that $\mathbf{u}(z)$ and $\mathbf{v}(z)$ are analytic $\mathbb{C}^k$-valued functions in $\mathbb{C} \setminus (\Upsilon_{f,IN} \cup \Xi_f)$ since $\lambda(z)$ is a simple eigenvalue of $T(z)$ (and $T(z)^t$) for all z in this set (cf. the proof of Lemma 1). The explicit forms of $\mathbf{u}(z)$ and $\mathbf{v}(z)$ are easily obtained by elementary manipulations with the matrix $T(z)$ defined in (2.7). Note also that since $\lambda(z)$ is a root of the characteristic polynomial $\chi_T(\zeta, z)$ of $T(z)$ one has

$$\begin{aligned}\mathbf{v}(z)^t\mathbf{u}(z) &= \gamma_1(z)\gamma_2(z)\sum_{m=0}^{k-1}(k-m)R_m(z)\lambda(z)^m \\ &= \gamma_1(z)\gamma_2(z)\lambda(z)\,\frac{\partial}{\partial\zeta}\,\chi_T(\zeta, z)\Big|_{\zeta=\lambda(z)}.\end{aligned}$$

Therefore, if $\mathbf{v}(z)^t\mathbf{u}(z) = 0$ then $\lambda(z)$ must be a multiple root of the characteristic polynomial $\chi_T(\zeta, z)$ of $T(z)$. However, this contradicts the assumption $z \in \Xi_f$. Thus $\mathbf{v}(z)^t\mathbf{u}(z) \neq 0$ whenever $z \in \Xi_f$, which proves the last statement of the lemma. □

Henceforth, we assume that the complex-valued functions $\gamma_1(z), \gamma_2(z)$ defined in Lemma 3 are chosen so that $\mathbf{v}(z)^t\mathbf{u}(z) = 1$ for all $z \in \mathbb{C} \setminus (\Upsilon_{f,IN} \cup \Xi_f)$.

Remark 5. By Remark 3 all coordinates of $\mathbf{u}(z)$ are nonvanishing (analytic) functions in $\mathbb{C} \setminus (\Upsilon_{f,IN} \cup \Xi_f)$.

Lemma 4. *With the above notations one has*

$$\frac{q_n(z)}{\lambda(z)^n} \rightrightarrows \gamma_1(z)\gamma_2(z)u_1(z)\sum_{i=1}^{k} v_i(z)q_{k-i}(z) \quad \textit{in } \mathbb{C} \setminus (\Upsilon_{f,IN} \cup \Xi_f) \tag{2.10}$$

as $n \to \infty$.

Proof. Note that in view of recursion (2.6) all the functions $\frac{q_n(z)}{\lambda(z)^n}$, $n \in \mathbb{N}$, are analytic in $\mathbb{C} \setminus (\Upsilon_{f,IN} \cup \Xi_f)$ while Lemma 3 implies that the same is true for the function in the right-hand side of (2.10). From Lemma 2(i)–(iii) it follows that

$$\frac{q_n(z)}{\lambda(z)^n} \rightrightarrows \gamma_1(z)\gamma_2(z)u_1(z)\sum_{i=1}^{k} v_i(z)q_{k-i}(z) \quad \text{in } \mathbb{C} \setminus (\Upsilon_{f,IN} \cup \Xi_f \cup \Delta_T) \tag{2.11}$$

as $n \to \infty$. Since Δ_T is a finite subset of $\mathbb{C}$ (cf. Remark 4), Cauchy's integral formula shows that the local uniform convergence stated in (2.11) must actually hold even in an ε-neighborhood of the set $\Delta_T \setminus (\Upsilon_{f,IN} \cup \Xi_f)$, which proves (2.10). □

To further simplify the notations set

$$g_n(z) = \frac{q_n(z)}{\lambda(z)^n}, \quad n \in \mathbb{N} \quad \text{and} \quad g(z) = \gamma_1(z)\gamma_2(z)u_1(z) \sum_{i=1}^{k} v_i(z) q_{k-i}(z). \tag{2.12}$$

We can complement Lemma 4 with the following result:

Lemma 5. *Given* $0 < \varepsilon \ll 1$ *let* $\mathcal{V}_\varepsilon = \mathbb{CP} \setminus V_\varepsilon$, *where* V_ε *is the* ε*-neighborhood of* $\Upsilon_{f,IN} \cup \Xi_f$ *in the spherical metric. The rate of convergence of* $g_n(z) \rightrightarrows g(z)$ *in* $\mathcal{V}_\varepsilon$ *is exponential, that is, there exist constants* $\mathfrak{C}_\varepsilon$ *and* $\rho_\varepsilon \in (0,1)$ *such that* $|g_n(z) - g(z)| \leqslant \mathfrak{C}_\varepsilon \rho_\varepsilon^n$ *for all* $z \in \mathcal{V}_\varepsilon$.

Proof. By Lemma 2(i) and (2.4) we get an exponential rate of convergence for $g_n(z) \rightrightarrows g(z)$ in $\mathbb{CP} \setminus U_\varepsilon$, where U_ε is the ε-neighborhood of $\Upsilon_{f,IN} \cup \Xi_f \cup \Delta_T$ in the spherical metric. Since $g_n(z)$, $n \in \mathbb{N}$, and $g(z)$ are all analytic functions in $\mathbb{C} \setminus (\Upsilon_{f,IN} \cup \Xi_f)$ and Δ_T is a finite subset of $\mathbb{C}$, we get the desired conclusion again by Cauchy's integral formula. $\square$

The next lemma gives a complete description of the set of slow growth $\Sigma_{f,IN}$ that appears in Theorem 2 (cf. Definitions 3 and 6).

Lemma 6. *Let* $IN = \{q_0(z), \ldots, q_{k-1}(z)\}$ *be a fixed initial k-tuple of rational functions as in* (2.6). *Then*

$$\Sigma_{f,IN} = \left\{ z \in \mathbb{C} \setminus (\Upsilon_{f,IN} \cup \Xi_f) \,\middle|\, \sum_{i=1}^{k} \sum_{j=i}^{k} \lambda(z)^{i-j} q_{k-i}(z) R_{k-j}(z) = 0 \right\}, \tag{2.13}$$

so that in particular $|\Sigma_{f,IN}| < \infty$.

Proof. The fact that $\Sigma_{f,IN}$ is given by (2.13) follows readily from Lemmas 2–4 and Remark 5. In order to show that $|\Sigma_{f,IN}| < \infty$ one can produce an explicit upper bound for $|\Sigma_{f,IN}|$ in the following way: set

$$\mu_{T,IN}(\zeta, z) = \sum_{i=1}^{k} \sum_{j=i}^{k} q_{k-i}(z) R_{k-j}(z) \zeta^{i-j}$$

and let $S(z)$ denote the resultant of the polynomials $\chi_T(\zeta, z)$ and $\mu_{T,IN}(\zeta, z)$ in the variable ζ. Let further $\mathcal{S} = \{z \in \mathbb{C} \mid S(z) = 0\}$. Since $S(z)$ is a rational function and $\Sigma_{f,IN} \subseteq \mathcal{S}$ it follows that $|\Sigma_{f,IN}| \leqslant |\mathcal{S}| < \infty$. $\square$

Both Theorems 2 and 3 are immediate consequences of Lemmas 4–6. The same ideas yield also a proof of Theorem 5. Actually, we can produce an asymptotically optimal bound for the number of spurious poles of the rational functions $r_n(z)$. To do this, let us introduce the following notation.

Notation 2. Given a sequence of rational functions $\{q_n(z)\}_{n\in\mathbb{N}}$ as in (2.6) denote by $Z_{sp}(q_n)$ the set of zeros of q_n which are contained in $\mathbb{C}\setminus(\Upsilon_{f,IN}\cup\Xi_f)$. The *total cardinality* $\|Z_{sp}(q_n)\|$ of $Z_{sp}(q_n)$ is the sum of the multiplicities of each of its elements as zeros of q_n. The *total cardinality* of $\Sigma_{f,IN}$ is $\|\Sigma_{f,IN}\| = \sum_{z\in\Sigma_{f,IN}} m(z)$, where $m(z)$ is the multiplicity of $z\in\Sigma_{f,IN}$ as zero of Eq. (2.13).

Lemma 7. *Let $IN=\{q_0(z),\dots,q_{k-1}(z)\}$ be a fixed initial k-tuple of rational functions. There exists $N\in\mathbb{N}$ such that $\|Z_{sp}(q_n)\|\leqslant\|\Sigma_{f,IN}\|$ for $n\geqslant N$. Thus,*

$$\|Z_{sp}(q_n)\|\leqslant \max(\|Z_{sp}(q_0)\|,\dots,\|Z_{sp}(q_{N-1})\|,\|\Sigma_{f,IN}\|)$$

for all $n\in\mathbb{N}$.

Proof. Recall (2.12) and note that since $|\Sigma_{f,IN}|<\infty$, it is enough to show that for any $z_0\in\Sigma_{f,IN}$ there exists $N_0\in\mathbb{N}$ and $\varepsilon_0>0$ such that $g_n(z)$ has (at most) $m(z_0)$ zeros in $\{z\in\mathbb{C}\mid |z-z_0|<\varepsilon_0\}$. But this is an immediate consequence of the uniform convergence established in Lemma 4 and Hurwitz's Theorem. □

An interesting consequence of the above results is that by using our scheme one can always find rational approximants $\{r_n(z)\}_{n\in\mathbb{N}}$ that have no spurious poles.

Corollary 2. *If f is an algebraic nonrational function and IN is chosen to be the standard initial k-tuple $\{0,0,\dots,1\}$ then $\Sigma_{f,IN}=\emptyset$. In particular, the rational approximants $\{r_n(z)\}_{n\in\mathbb{N}}$ corresponding to the standard initial k-tuple have no spurious poles.*

Proof. Indeed, if $\Sigma_{f,IN}\neq\emptyset$ and $z\in\Sigma_{f,IN}$ then by (2.8) and (2.13) the corresponding dominant root $y_{\mathrm{dom}}(z)=\lambda(z)$ has to satisfy

$$\lambda(z)^k-\sum_{i=0}^{k-1}R_{k-i-1}(z)\lambda(z)^{k-i-1}=0 \quad\text{and}\quad \sum_{j=1}^{k}R_{k-j}(z)\lambda(z)^{1-j}=0,$$

which gives $\lambda(z)=0$. However, the latter identity cannot be fulfilled by the dominant root of (1.3) if *f* is an algebraic nonrational function. □

Remark 6. An analog of Corollary 2 for 3-term recurrence relations satisfied by certain biorthogonal polynomials was proved in [12, Section 3].

Remark 7. Corollary 2 does not hold for arbitrarily chosen initial *k*-tuples consisting of constant functions, as one can easily see by considering, e.g., the initial *k*-tuple $IN=\{1,0,\dots,1\}$.

For the proof of Theorem 4 we need some additional notation and results.

Notation 3. Let $\mathfrak{Rec}_k$ be the k-dimensional complex linear space consisting of all $(k+1)$-term recurrence relations with constant coefficients of form (1.4). We denote by $\mathfrak{IN}_k$ the k-dimensional complex linear space of all initial k-tuples $(u_0, \dots, u_{k-1})$.

Recall the notion of recursion of nondominant type introduced in Definition 2.

Definition 10. A nondominant recurrence relation in $\mathfrak{Rec}_k$ with initial k-tuple $IN \in \mathfrak{IN}_k$ is said to be of *subdominant type* if the following conditions are satisfied. Let $\tau_1, \dots, \tau_s$, $s \leqslant k$, denote all distinct spectral numbers with maximal absolute value and assume that they have multiplicities $m_1, \dots, m_s$, respectively. Then there exists a unique index $i_0 \in \{1, \dots, s\}$ such that $m_i < m_{i_0}$ for $i \in \{1, \dots, s\} \setminus \{i_0\}$ and the initial k-tuple IN is *fast growing* in the sense that the degree of the polynomial p_{i_0} in (1.6) corresponding to τ_{i_0} is precisely $m_{i_0} - 1$. The number τ_{i_0} is called the *dominant spectral number* of this recurrence relation.

The following lemma is a simple consequence of Theorem 1.

Lemma 8. *In the above notation the following is true*:

(i) *The set of all slowly growing initial k-tuples with respect to a given dominant recurrence relation in* $\mathfrak{Rec}_k$ *is a complex hyperplane* $\mathcal{SG}_k$ *in* $\mathfrak{IN}_k$. *The set* $\mathcal{SG}_k$ *is called the* hyperplane of slow growth.

(ii) *For any dominant recurrence relation in* $\mathfrak{Rec}_k$ *and any fast growing initial k-tuple in* $\mathfrak{IN}_k$ *the limit* $\lim_{n\to\infty} \frac{u_{n+1}}{u_n}$ *exists and coincides with the dominant spectral number* $\tau_{\max}$, *that is, the* (*unique*) *root of the characteristic Eq.* (1.5) *with maximal absolute value.*

(iii) *Given a nondominant recurrence relation of subdominant type in* $\mathfrak{Rec}_k$ *and a fast growing initial k-tuple in* $\mathfrak{IN}_k$ *the limit* $\lim_{n\to\infty} \frac{u_{n+1}}{u_n}$ *exists and coincides with the dominant spectral number.*

(iv) *For any nondominant recurrence relation in* $\mathfrak{Rec}_k$ *which is not of subdominant type the set of initial k-tuples for which* $\lim_{n\to\infty} \frac{u_{n+1}}{u_n}$ *exists is a union of complex subspaces of* $\mathfrak{IN}_k$ *of positive codimensions. This union is called the* exceptional variety.

Proof. In order to prove (i) notice that the coefficient $\kappa_{\max}$ in Definition 3 is a nontrivial linear combination of the entries of the initial k-tuple $\{u_0, \dots, u_{k-1}\}$ with coefficients depending on $\alpha_1, \dots, \alpha_k$. Therefore, the condition $\kappa_{\max} = 0$ determines a complex hyperplane $\mathcal{SG}_k$ in IN_k. One can easily see that the hyperplane of slow growth is the direct sum of all Jordan blocks corresponding to the spectral numbers of a given recurrence (1.4) other than the leading one.

The assumptions of part (ii) together with (1.6) yield $u_n = \kappa_{\max}\tau_{\max}^n + \cdots$ for $n \in \mathbb{Z}_+$, where the dots stand for the remaining terms in (1.6) corresponding to the spectral numbers whose absolute values are strictly smaller than $|\tau_{\max}|$. Therefore, the quotient $\frac{u_{n+1}}{u_n}$ has a limit as $n \to \infty$ and this limit coincides with $\tau_{\max}$, as required. By definition $\tau_{\max}$ is a root of (1.5), which completes the proof of (ii). This last step can alternatively be carried out by dividing both sides of (1.4) by u_{n-k} and then letting $n \to \infty$. In view of Definition 10 the same arguments show that the assertion in (iii) is true as well.

For the proof of (iv) we proceed as follows. Take any nondominant recurrence relation of form (1.4) and let $\tau_1, \ldots, \tau_r$, $r \leqslant k$, be all its distinct spectral numbers with maximal absolute value. Thus, $|\tau_i| = |\tau_{\max}|$ if and only if $1 \leqslant i \leqslant r$. Choose an initial k-tuple $IT = \{u_0, \ldots, u_{k-1}\}$ and denote by $p_1, \ldots, p_r$ the polynomials in (1.6) corresponding to $\tau_1, \ldots, \tau_r$, respectively, for the sequence $\{u_n \mid n \in \mathbb{Z}_+\}$ constructed using the given recurrence with initial k-tuple IT as above. Assuming, as we may, that our recurrence relation is nontrivial we get from (1.6) that $\tau_i \neq 0$ if $1 \leqslant i \leqslant r$. We may further assume that the degrees $d_1, \ldots, d_r$ of the polynomials $p_1, \ldots, p_r$, respectively, satisfy $d_1 \geqslant \cdots \geqslant d_r$. A direct check analogous to the proof of part (ii) shows that if only the polynomial p_1 is nonvanishing then $\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = \tau_1$. If $r \geqslant 2$ and $s \in \{2, \ldots, r\}$ is such that $p_1, \ldots, p_s$ are all nonvanishing polynomials among $p_1, \ldots, p_r$ then using again (1.6) we get

$$\frac{u_{n+1}}{u_n} = \frac{p_1(n+1) + p_2(n+1)\left(\dfrac{\tau_2}{\tau_1}\right)^{n+1} + \cdots + p_s(n+1)\left(\dfrac{\tau_s}{\tau_1}\right)^{n+1} + o(1)}{p_1(n) + p_2(n)\left(\dfrac{\tau_2}{\tau_1}\right)^{n} + \cdots + p_s(n)\left(\dfrac{\tau_s}{\tau_1}\right)^{n} + o(1)}.$$

Since $\left|\dfrac{\tau_i}{\tau_1}\right| = 1$ and $\tau_i \neq \tau_1$, $2 \leqslant i \leqslant s$, it follows that if $d_1 = d_2$ then the expression in the right-hand side has no limit as $n \to \infty$. Therefore, if such a limit exists then $d_1 > d_2$, which gives us a complex subspace of IN_k of (positive) codimension equal to $d_1 - d_2$. Thus, the exceptional variety is a union of complex subspaces of IN_k of (in general) different codimensions. $\square$

Proof of Theorem 4. To prove the first statement notice that under the assumptions of the theorem the dominant root of the asymptotic symbol equation (1.2) is a well-defined analytic function in a sufficiently small neighborhood of $\Sigma_{f,IN}$. Therefore, the residue distribution (1.7) associated to this dominant root vanishes in a neighborhood of $\Sigma_{f,IN}$. Thus, the set $\Sigma_{f,IN}$ of slowly growing initial conditions can be deleted from the support of the asymptotic ratio distribution ν.

In order to show that under the nondegeneracy conditions stated in (ii) the support of ν must coincide with Ξ_f it is enough to prove that the sequence $\{r_n(z)\}_{n\in\mathbb{N}}$ diverges almost everywhere in Ξ_f. This is actually an immediate consequence of the fact that the nondegeneracy assumptions in (ii) imply that at almost every point $z \in \Xi_f$, recursion (1.3) is neither of dominant nor of subdominant type (cf. Definitions 2 and 10). Indeed, since the set $\Sigma_{f,IN}$ of slow growth is finite it follows from Lemma 8(iv) that the sequence $\{r_n(z)\}_{n\in\mathbb{N}}$ diverges for almost all $z \in \Xi_f$.

Finally, in order to prove (iii) notice that by Definitions 7–8 any rational function r is the Cauchy transform of its residue distribution v_r and that by Theorem 2 one has $r_n(z) = \frac{q_n(z)}{q_{n-1}(z)} \rightrightarrows y_{\text{dom}}(z)$ in $\mathbb{C} \setminus \mathcal{D}_f$. Therefore, if v_n denotes the residue distribution of r_n then the Cauchy transform $\mathcal{C}_v$ of $v := \lim_{n\to\infty} v_n$ equals y_{dom}. It then follows from the standard relation between a distribution and its Cauchy transform (see (1.8)) that

$$v = \frac{1}{\pi}\frac{\partial \mathcal{C}_v}{\partial \bar{z}} = \frac{1}{\pi}\frac{\partial y_{\text{dom}}}{\partial \bar{z}},$$

which completes the proof of the theorem. $\square$

Proof of Corollary 1. The assertion is a direct consequence of the previous results since Theorem 3 implies that for any sufficiently small $\varepsilon > 0$ the complement Θ_ε of the ε-neighborhood $\mathcal{O}_\varepsilon$ of $\mathcal{D}_f = \Xi_f \cup \Upsilon_{f,IN} \cup \Sigma_{f,IN}$ in the spherical metric contains no poles of $r_n(z)$ for sufficiently large n. Note, in particular, that by Theorem 4(iii) the regular part of the poles of the family $\{r_n(z)\}_{n\in\mathbb{N}}$ must tend to Ξ_f in such a way that $\lim v_n = v = \frac{1}{\pi}\frac{y_{\text{dom}}(z)}{\partial \bar{z}}$. $\square$

3. A brief excursion around Padé approximation

In order to make this paper self-contained and to be able to compare our approximation scheme with the classical Padé scheme we give in this section a brief account of Padé approximants. For a detailed overview of this subject we refer to, e.g., [2,35]. As in Section 1, we denote by $\mathcal{P}_k$ the linear space of all complex polynomials of degree at most k.

Definition 11. Let $f(z)$ be a function analytic at infinity with power series expansion $f(z) = \sum_{i=0}^{\infty} c_i z^{-i}$. For each pair $(m, n) \in \mathbb{N}^2$ there exist two polynomials $p_{m,n} \in \mathcal{P}_m$ and $q_{m,n} \in \mathcal{P}_n \setminus \{0\}$ such that

$$q_{m,n}\left(\frac{1}{z}\right) f(z) - p_{m,n}\left(\frac{1}{z}\right) = \mathcal{O}(z^{-m-n-1}) \tag{3.1}$$

as $z \to \infty$. The rational function

$$[m/n](z) := \frac{p_{m,n}(1/z)}{q_{m,n}(1/z)} \tag{3.2}$$

is called the (m, n)-Padé approximant to the function $f(z)$ (expanded at ∞).

The notation introduced in Definition 11 will be used throughout this section. Padé approximants can be seen as a generalization of Taylor polynomials to the field of rational functions. Note that in view of (3.1) these approximants are well-defined and exist uniquely for any function $f(z)$ that has a power series expansion at ∞. Padé approximants are also closely related to (generalized) orthogonal polynomials. As we shall now explain, this connection is especially simple in the case of *diagonal Padé approximants*, that is, approximants of the form $[n/n](z)$. To this end let us define the reverse denominator polynomial Q_n of $q_{n,n}$ by

$$Q_n(z) = z^n q_{n,n}\left(\frac{1}{z}\right).$$

The following results may be found in [35, Lemma 2.2].

Lemma 9. *A polynomial $q_{n,n} \in \mathcal{P}_n \setminus \{0\}$ is the denominator of the Padé approximant $[n/n]$ to f if and only if the reverse polynomial Q_n of $q_{n,n}$ satisfies the* (*generalized*) *orthogonality relation*

$$\int_C \zeta^i Q_n(\zeta) f(\zeta)\, d\zeta = 0, \quad i \in \{0, \ldots, n-1\}, \tag{3.3}$$

where C is a closed contour containing all the singularities of $f(z)$ (*so that $f(z)$ is analytic on C and in its exterior*).

Condition (3.3) is a (generalized) orthogonality relation for the polynomial family $\{Q_n(z)\}_{n\in\mathbb{N}}$ on the contour C with respect to the complex-valued measure $f(\zeta)\, d\zeta$. The fact that the quadratic form

$$(g, h)_f := \int_C g(\zeta)h(\zeta)f(\zeta)\, d\zeta$$

diagonalized by the family $\{Q_n(z)\}_{n\in\mathbb{N}}$ is in general neither Hermitian nor positive-definite accounts for the rather chaotic behavior of the zero loci of the polynomials $Q_n(z)$, $n \in \mathbb{N}$, see [30,31,35] and the results below.

Remark 8. Note that the generalized orthogonality relation (3.3) and Favard's theorem imply that if the polynomials in the family $\{Q_n(z)\}_{n\in\mathbb{N}}$ are assumed to be monic then they satisfy a three-term recurrence relation of the form

$$zQ_n(z) = Q_{n+1}(z) + \alpha_n Q_n(z) + \beta_n Q_{n-1}(z) \tag{3.4}$$

with complex-valued coefficients α_n and β_n, see [16]. Apparently, the chaotic behavior of the zero loci of the polynomials $Q_n(z)$, $n \in \mathbb{N}$, is inherited by the families of coefficients $\{\alpha_n\}_{n\in\mathbb{N}}$ and $\{\beta_n\}_{n\in\mathbb{N}}$, as these do not necessarily converge, see Section 5.2.3.

Probably the most important general theorem that applies to functions meromorphic in the complex plane is that of Nuttall–Pommerenke asserting that if f is meromorphic in $\mathbb{C}$ and analytic at ∞ then the sequence of diagonal Padé approximants $\{[n/n](z)\}_{n\in\mathbb{N}}$ converges to $f(z)$ in planar measure (cf. [2,34]). The convergence of the diagonal Padé approximants to $f(z)$ in the case when $f(z)$ is a (multivalued) algebraic function was studied in detail in [34]. As pointed out in *loc. cit.*, the following important result was proved already in [17].

Theorem 6. *Let $f(z)$ be a locally meromorphic function in some domain $\mathbb{CP}^1 \setminus E$, where E is a compact subset of $\mathbb{C}$ with* $\mathrm{cap}(E) = 0$. *There exists a domain $D_f \subseteq \mathbb{CP}^1$ such that the sequence of diagonal Padé approximants $\{[n/n](z)\}_{n\in\mathbb{N}}$ converges in capacity to $f(z)$ in D_f but it does not converge in capacity to $f(z)$ in any domain $\widetilde{D} \subseteq \mathbb{CP}^1$ with* $\mathrm{cap}(\widetilde{D} \setminus D) > 0$.

The domain D_f in Theorem 6 is uniquely determined up to a set of zero capacity and is called the *convergence domain* of $f(z)$. (Concerning the notion of (logarithmic) capacity and convergence in capacity one may consult, e.g., [36].) The following improvement of Theorem 6 was obtained in [35] by building on the results of [32].

Theorem 7. *Let $f(z)$ be a locally meromorphic function in some domain $\mathbb{CP}^1 \setminus E$, where E is a compact subset of $\mathbb{C}$ with* $\mathrm{cap}(E) = 0$. *The convergence domain D_f of $f(z)$ is uniquely determined (up to a set of capacity zero) by the following two conditions*:

(i) *D_f is a subdomain of $\mathbb{CP}^1$ containing ∞ such that $f(z)$ has a single-valued meromorphic continuation throughout D_f*;

(ii) $\mathrm{cap}([\mathbb{CP}^1 \setminus D_f]^{-1}) = \inf_{\widetilde{D}} \mathrm{cap}([\mathbb{CP}^1 \setminus \widetilde{D}]^{-1})$, *where the infimum ranges over all domains $\widetilde{D} \subseteq \mathbb{CP}^1$ satisfying* (i). (*Here for a given set $S \subset \mathbb{CP}^1$ we denote by S^{-1} the set of all $z \in \mathbb{CP}^1$ such that $z^{-1} \in S$*.)

In other words, Theorem 7 asserts that the convergence domain D_f of $f(z)$ is characterized by the property that its boundary ∂D_f has minimal capacity.

Definition 12. Given a function $f(z)$ analytic near ∞ we say that a compact set $K \subseteq \mathbb{CP}^1$ is *admissible* for $f(z)$ if $f(z)$ has a single-valued continuation throughout $\mathbb{CP}^1 \setminus K$.

Using the terminology introduced in Definition 12 one may rephrase Theorem 7 as follows: given a locally meromorphic function $f(z)$ in some domain $\mathbb{CP}^1 \setminus E$, where E is a compact subset of $\mathbb{C}$ with $\mathrm{cap}(E) = 0$, there exists an unique admissible compact set $K_0 \subseteq \mathbb{CP}^1$ such that

$$\mathrm{cap}(K_0) = \inf_{K} \mathrm{cap}(K),$$

where the infimum is taken over all admissible compact sets $K \subseteq \mathbb{CP}^1$.

Let $g_D(z, w)$ denote the Green's function of a domain $D \subset \mathbb{CP}^1$ with $\operatorname{cap}(\partial D) > 0$. The following result may be found in [32, Part III, Theorem 1 and Lemma 5].

Theorem 8. *In the above notation the minimal set K_0 is the union of a compact set E_0 with analytic Jordan arcs. The set E_0 has the property that the analytic continuation of $f(z)$ in $\mathbb{CP}^1 \setminus K_0$ has a singularity at every boundary point of E_0 while the analytic arcs in $K_0 \setminus E_0$ are trajectories of the quadratic differential $T_f(z)\, dz^2$ given by*

$$T_f(z) = \left\{ \frac{\partial}{\partial z} g_{\mathbb{CP}^1 \setminus K_0}(z, \infty) \right\}^2.$$

Moreover, if $f(z)$ is algebraic then $T_f(z)$ is a rational function.

Theorem 9. *Let $f(z)$ be an algebraic nonrational function which is analytic near ∞. If all singularities of $f(z)$ are contained in a set $W = \{w_1, \dots, w_n\} \subset \mathbb{C}$ then there exist a subset W' of W with $|W'| = n' \geqslant 2$ and a set $Z \subset \mathbb{C}$ with (not necessarily distinct) $n' - 2$* points such that

$$T_f(z) = \frac{\Pi_{z_j \in Z}(z - z_j)}{\Pi_{w_j \in W'}(z - w_j)} = c \left\{ \frac{\partial}{\partial z} g_{\mathbb{CP}^1 \setminus K_0}(z, \infty) \right\}^2.$$

Here, c is an appropriate complex constant and $z \in \mathbb{CP}^1 \setminus K_0$, where K_0 is the union of trajectories of the quadratic differential $T_f(z)\, dz^2$ and a set of isolated points contained in W.

Theorem 9 shows that most of the poles of the diagonal Padé approximants $[n/n](z)$ tend to K_0 when $n \to \infty$. However, not all of them do! Indeed, even in the case when $f(z)$ is algebraic one cannot hope for a better convergence type than convergence in capacity (which is only slightly stronger than convergence almost everywhere with respect to Lebesgue measure). As shown by the following example of Stahl, uniform convergence in $\mathbb{CP}^1 \setminus K_0$ fails in a rather dramatic way.

Example 2. Consider the function

$$f(z) = \frac{(z - \cos \pi\alpha_1)(z - \cos \pi\alpha_2)}{\sqrt{z^2 - 1}} - z + (\cos \pi\alpha_1 + \cos \pi\alpha_2),$$

where $1, \alpha_1, \alpha_2$ are rationally independent numbers. Then the reverse denominators $\{Q_n(z)\}_{n \in \mathbb{N}}$ satisfy the generalized orthogonality relation

$$\int_{-1}^{1} Q_n(x) Q_m(x) \frac{(x - \cos \pi\alpha_1)(x - \cos \pi\alpha_2)}{\pi\sqrt{1 - x^2}}\, dx = 0 \quad \text{for } m \neq n.$$

Even though each polynomial $Q_n(z)$ has at most two zeros outside $[-1, 1]$, the zeros of $Q_n(z)$ cluster everywhere in $\mathbb{C}$ as $n \to \infty$ (see [35, Example 2.4]).

A naive definition of *spurious poles* of Padé approximants would be that these are the poles which are not located near K_0. A more rigorous definition is as follows (cf. [35, Definition 4.1]).

Definition 13. Let f be a function satisfying the conditions of Definition 11. Let further $\mathcal{N} \subseteq \mathbb{N}$ be an infinite sequence and $\{[n/n]\}_{n\in\mathcal{N}}$ be the corresponding subsequence of diagonal Padé approximants to f. We define *spurious poles* in two different situations:

(i) Assume that for each $n \in \mathcal{N}$ the approximant $[n/n]$ has a pole at $z_n \in \mathbb{CP}^1$ such that $z_n \to z_0$ as $n \to \infty$, $n \in \mathcal{N}$. If f is analytic at z_0 and the approximants $[n/n]_{n\in\mathcal{N}}$ converge in capacity to f in some neighborhood of z_0 then the poles of the approximants $[n/n]$ at z_n, $n \in \mathcal{N}$, are called *spurious*. In case $z_0 = \infty$, the convergence $z_n \to z_0$ has to be understood in the spherical metric.

(ii) Let the function $f(z)$ have a pole of order k_0 at z_0 and assume that for each $n \in \mathcal{N}$ the total order of poles of the approximant $[n/n]$ near z_0 is $k_1 = k_{1,n} > k_0$. Assume further that $[n/n]$ has poles at $z_{n,j}$, $j \in \{1, \dots, m_n\}$, of total order $k_{1,n}$ and that for any selection of $j_n \in \{1, \dots, m_n\}$ one has $z_{n,j_n} \to z_0$ as $n \to \infty$, $n \in \mathcal{N}$. Then poles of order $k_1 - k_0$ out of all poles of the approximants $[n/n]$ near z_0, $n \in \mathcal{N}$, are called *spurious*.

With this definition we can finally formulate the Baker–Gammel–Wills (Padé) Conjecture and some related questions.

Padé (Baker–Gammel–Wills) Conjecture. *If the function f is meromorphic in the unit disk $\mathbb{D}$ and analytic at* 0, *and if $\{[n/n]\}_{n\in\mathbb{N}}$ denotes the sequence of Padé approximants to f expanded at the origin, then there exists an infinite subsequence $\mathcal{N} \subseteq \mathbb{N}$ such that $[n/n](z) \rightrightarrows f(z)$ in $\mathbb{D} \setminus \{$poles of $f\}$ as $n \to \infty$, $n \in \mathcal{N}$.*

The Padé Conjecture is widely regarded as the main challenge in the theory of Padé approximation. The importance of this conjecture is due to the homographic invariance property of diagonal Padé approximants and its many consequences, see [2]. A major breakthrough was recently made by Lubinsky [24] who showed that the general version of the Padé Conjecture stated above is false. Subsequently, Buslaev [15] constructed counterexamples to this conjecture for some special types of algebraic (hyperelliptic) functions. However, the important question whether the Padé Conjecture could hold for certain classes of algebraic functions is still open. A key step in this direction would be a deeper understanding of the asymptotic distribution of spurious poles of diagonal Padé approximants. To this end Nuttall proposed the following conjecture (cf. [35]).

Nuttall's Conjecture. *Let f be an algebraic function which is analytic at ∞. Then there exists an upper bound for the number of spurious poles* (*in the sense of total order*) *for all Padé approximants* $[n/n]$, $n \in \mathbb{N}$.

Note that Theorems 2 and 5 show that natural analogs of both the Padé Conjecture and Nuttall's Conjecture are always true for the rational approximants to algebraic functions constructed by using our approximation scheme.

4. Application: The 3-conjecture of Egecioglu, Redmond and Ryavec

The following interesting conjecture appeared in [18]—where it was nicknamed the 3-conjecture—in connection with counting problems for various combinatorial objects such as alternating sign matrices with vertical symmetries and n-tuples of nonintersecting lattice paths.

Conjecture 1. *All polynomials belonging to the sequence $\{p_n(z)\}_{n\in\mathbb{N}}$ defined by the 4-term recursion*

$$p_n(z) = zp_{n-1}(z) - Cp_{n-2}(z) - p_{n-3}(z),$$

where

$$p_{-2}(z) = p_{-1}(z) = 0, \quad p_0(z) = 1 \quad \text{and} \quad C \in \mathbb{R}, \tag{4.1}$$

have real zeros if and only if $C \geqslant 3$. Moreover, if $C > 3$ then the zeros of $p_{n+1}(z)$ and $p_n(z)$ are interlacing for all $n \in \mathbb{N}$.

Theorem 1 of [18] proves the sufficiency part as well as the last statement in the above conjecture. We shall now fully settle the 3-conjecture by showing that $C \geqslant 3$ is indeed a necessary condition for the validity of this conjecture.

Proposition 1. *If $C < 3$ and $\{p_n(z)\}_{n\in\mathbb{N}}$ is the sequence of polynomials defined by recursion* (4.1) *then there exists $N \in \mathbb{N}$ such that not all zeros of $p_N(z)$ are real.*

The proof of Proposition 1 is based on the ideas developed in the previous sections and goes as follows. We study the behavior of the branching points of the curve Γ_C given by solutions t of the asymptotic symbol equation

$$t^3 - zt^2 + Ct + 1 = 0 \tag{4.2}$$

considered as a branched covering of the z-plane (cf. Definition 1). More precisely, we show that for any $C < 3$ the curve Γ_C has two complex conjugate branching points and the induced equimodular discriminant Ξ_C of (4.2) is a real semialgebraic curve ending at these branching points (cf. [10,12]). Therefore, if $\varepsilon > 0$ there exists some sufficiently large $N \in \mathbb{N}$ such that for all $n \geqslant N$ the polynomial $p_n(z)$ has (complex conjugate) roots lying in the ε-neighborhoods of these branching points.

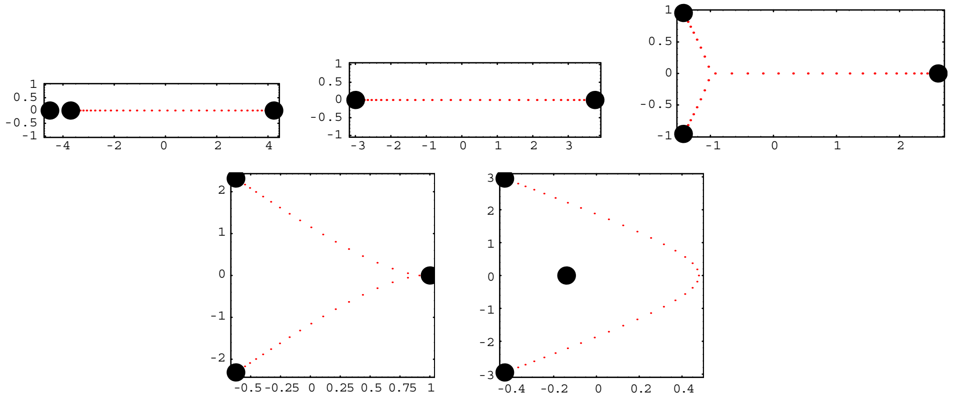

Fig. 2. Zeros of $p_{41}(z)$ satisfying (4.1) for $C = 4, 3, 1, -1, -2$.

Explanations to Fig. 2. The five pictures show the qualitative changes of the induced equimodular discriminant Ξ_C and the branching points of the curve Γ_C when the parameter C runs from ∞ to $-\infty$. The fat points are the branching points of Γ_C while the thin points illustrate the zeros of the polynomial $p_{41}(z)$. The first picture depicts the typical behavior for $C > 3$. In this case, the branching points of Γ_C are all real and the equimodular discriminant Ξ_C is the segment joining the two rightmost branching points. If $C = 3$, then the two leftmost branching points of Γ_C coalesce. For $-1 < C < 3$ there are one real and two complex conjugate branching points of Γ_C and the induced equimodular discriminant Ξ_C is the *Y*-shaped figure joining these three points. If $C = -1$, then the triple point occurring in the previous case coincides with the (unique) real branching point of Γ_C. Finally, if $C < -1$ then one has a stable configuration consisting of the equimodular discriminant Ξ_C joining the two complex conjugate branching points of Γ_C and the remaining (real) branching point of Γ_C lying outside Ξ_C.

Straightforward computations using Mathematica™ yield the following result:

Lemma 10. *The branching points of Γ_C satisfy the equation*

$$4z^3 + C^2z^2 - 18Cz - 27 - 4C^3 = 0. \tag{4.3}$$

The discriminant of the polynomial in z occurring in the left-hand side of (4.3) *is given by*

$$64(C-3)^3(C^2+3C+9)^3. \tag{4.4}$$

Lemma 11. *The behavior of the induced equimodular discriminant Ξ_C near a branching point of Γ_C changes if and only if the branching point becomes degenerate, i.e., the absolute value of the double root of the asymptotic symbol equation corresponding* *to the branching point coincides with the absolute value of the remaining root. This happens precisely for those values of C for which there exist z, τ, θ such that the relation*

$$t^3 - zt^2 + Ct + 1 = (t-\tau)^2(t - \tau e^{i\theta}) \tag{4.5}$$

is satisfied for all $t \in \mathbb{C}$. If C is real this occurs only for $C = 3$ and $C = -1$.

Proof. The first part of the lemma is obvious and can be found in, e.g., [10]. To prove the statement about the values of C consider the system for the coefficients of (4.5):

$$\begin{cases} z = \tau(2 + e^{i\theta}), \\ C = \tau^2(1 + 2e^{i\theta}), \\ 1 = -\tau^3 e^{i\theta}. \end{cases}$$

The last equation gives $\tau = -e^{i(-\theta+2k\pi)/3}$, where $k \in \{0, 1, 2\}$. Substituting this expression in the second equation one gets

$$C = e^{-2i(\theta+4k\pi)/3} + 2e^{i(\theta+4k\pi)/3}.$$

Elementary computations now show that C is real if and only if $\theta = (3m - 4k)\pi$ for some $m \in \mathbb{Z}$, which gives either $C = 3$ (in which case $z = -3$ and $\tau = -1$) or $C = -1$ (in which case $z = 1$ and $\tau = 1$). $\square$

Proof of Proposition 1. Let us first show that it suffices to check that the following three conditions hold for any value of $C \in (-\infty, 3)$:

(i) the curve Γ_C has two (distinct) complex conjugate branching points;
(ii) at each of these two complex branching points of Γ_C the absolute value of the double root of the corresponding asymptotic symbol equation (4.2) exceeds the absolute value of the remaining root;
(iii) at a generic point $z_0 \in \Xi_C$ the roots of the asymptotic symbol equation (4.2) with maximal absolute value have distinct arguments and the sequence $\left\{\frac{p_n(z_0)}{p_{n-1}(z_0)}\right\}_{n\in\mathbb{N}}$ diverges.

Indeed, if (i) and (ii) are true then the induced equimodular discriminant Ξ_C must necessarily pass through (actually, end at) the two complex conjugate branching points of Γ_C (cf. [10,12]). In particular, there exist some nonisolated point $\zeta \in \Xi_C$ and $\varepsilon > 0$ such that $\{z \in \mathbb{C} \mid |z - \zeta| < \varepsilon\} \subset \mathbb{C} \setminus \mathbb{R}$. Moreover, if (iii) holds then Theorem 4(ii) implies that $\Xi_C = \operatorname{supp} \nu$, where ν denotes the asymptotic ratio distribution of the family $\{p_n(z)\}_{n\in\mathbb{N}}$. Therefore, there exists some sufficiently large $N \in \mathbb{N}$ such that for all $n \geqslant N$ the polynomial $p_n(z)$ has at least one root in $\{z \in \mathbb{C} \mid |z - \zeta| < \varepsilon\} \subset \mathbb{C} \setminus \mathbb{R}$, as required.

To verify conditions (i)–(iii) note first that if $C < 3$ then (4.4) implies that all three branching points of Γ_C are necessarily distinct. Since (4.3) is a polynomial equation with real coefficients depending continuously on C it is enough to check that (4.3) has (distinct) complex conjugate roots for an arbitrarily chosen value of C in the interval $(-\infty, 3)$. This is easily done numerically for, e.g., $C = 0$. Notice next that by using a similar continuity argument together with Lemma 11 it is enough to verify condition (ii) for $C = -1$ and for arbitrarily chosen values of C in the intervals $(-\infty, -1)$ and $(-1, 3)$, respectively. This is again an easy numerical test for, e.g., $C = -1$, $C = 0$, and $C = 2$. Finally, in order to check (iii) note that at any point $z_0 \in \Xi_C$ which is not a branching point of Γ_C the roots of the asymptotic symbol equation (4.2) with maximal absolute value have necessarily distinct arguments since these roots must be distinct. Fix any such z_0 and denote by $\tau_i(z_0)$, $1 \leqslant i \leqslant 3$, the (distinct) zeros of the corresponding asymptotic symbol equation (4.2), so that $\tau_i(z_0) \neq 0$ for $1 \leqslant i \leqslant 3$. By (4.1) and (1.6) there exist complex constants $K_i(z_0)$, $1 \leqslant i \leqslant 3$, such that

$$p_n(z_0) = \sum_{i=1}^{3} K_i(z_0)\tau_i(z_0)^n \quad \text{for } n \in \{-2, -1, 0\} \cup \mathbb{N}.$$

Using the initial values $p_{-2}(z) = p_{-1}(z) = 0$ and $p_0(z) = 1$ together with the above formula for $p_n(z_0)$, one can easily check that $K_i(z_0) \neq 0$ for $1 \leqslant i \leqslant 3$. Assume that $|\tau_1(z_0)| = |\tau_2(z_0)| \geqslant |\tau_3(z_0)|$. Then

$$\frac{p_n(z_0)}{p_{n-1}(z_0)} = \tau_1(z_0) \frac{K_1(z_0) + K_2(z_0)\left(\dfrac{\tau_2(z_0)}{\tau_1(z_0)}\right)^n + K_3(z_0)\left(\dfrac{\tau_3(z_0)}{\tau_1(z_0)}\right)^n}{K_1(z_0) + K_2(z_0)\left(\dfrac{\tau_2(z_0)}{\tau_1(z_0)}\right)^{n-1} + K_3(z_0)\left(\dfrac{\tau_3(z_0)}{\tau_1(z_0)}\right)^{n-1}}$$

and since $\left|\dfrac{\tau_2(z_0)}{\tau_1(z_0)}\right| = 1$ and $\tau_2(z_0) \neq \tau_1(z_0)$ it follows that the right-hand side of the above identity has no limit as $n \to \infty$, which proves (iii). $\square$

5. Comparison of approximation schemes and open problems

5.1. Approximation via the algebraic scheme versus Padé approximation

As we already pointed out in the introduction, the usual Padé approximation and the approximation scheme introduced in this paper have an essentially different range of applications. Nevertheless, when applied to an arbitrarily given algebraic function f whose defining equation is known, the scheme for approximating the branch with maximal modulus of f proposed above has a number of advantages compared to the usual Padé approximation scheme for *f*:

(1) While Padé approximation requires the knowledge of the Taylor expansion at ∞ of $f(z)$, our scheme uses only the defining algebraic equation for $f(z)$.
(2) The regular poles of Padé approximants concentrate on the union of certain trajectories of the quadratic differential described in Section 3. While it is in general rather difficult to understand the structure of these trajectories, the induced equimodular discriminant Ξ_f has a more transparent definition and is also easier to study. Interesting examples of equimodular discriminants relevant to the behavior of Potts models on sequences of finite graphs tending to certain infinite lattices have been studied in a number of papers by Sokal and his coauthors, see e.g., [28,29]. They also developed two approaches for finding Ξ_f explicitly. One is a more or less direct approach while the other is based on a resultant type equation which can also be found in [10].
(3) In general, Padé approximants have spurious poles with uncontrolled behavior which makes uniform convergence impossible. Our scheme has only a finite number of spurious poles tending to a quite understandable finite set $\Sigma_{f,IN}$. In addition to that, the latter scheme has also a well-controlled (exponential) rate of convergence.
(4) The denominators of the rational approximants $\{r_n(z)\}_{n\in\mathbb{N}}$ constructed in this paper satisfy a simple recursion with fixed rational coefficients. By contrast, the denominators of Padé approximants satisfy a 3-term recurrence relation with varying coefficients which are difficult to calculate and whose behavior is apparently rather chaotic (see Section 5.2.3).

5.2. Related topics and open problems

In this section, we discuss several questions pertaining to our approximation scheme that seem to have important applications not only to the convergence theory of Padé approximants [30–35] but also to complex Sturm–Liouville problems [6], modern statistical physics [28,29] and the theory of (general) orthogonal polynomials [16,26,36,38].

5.2.1.

Let $\{p_n(z)\}_{n\in\mathbb{N}}$ be a family of monic complex polynomials. In addition to the asymptotic ratio distribution ν introduced in Definition 9 one can associate to this family an asymptotic probability measure μ defined as follows.

Definition 14. To a degree d_n complex polynomial $p_n(z)$ we associate a finite probability measure μ_n called the *root counting measure* of $p_n(z)$ by placing the mass $\frac{1}{d_n}$ at any of its simple roots. If some root of $p_n(z)$ is multiple we place at this point a mass equal to the multiplicity of the root divided by d_n. The limit $\lim_{n\to\infty}\mu_n$ (if is exists in the sense of the weak convergence of measures) is called the *asymptotic root counting measure* of the family $\{p_n(z)\}_{n\in\mathbb{N}}$ and denoted by μ.

It is clear from Definitions 9 and 14 that for any given polynomial family $\{p_n(z)\}_{n\in\mathbb{N}}$ the supports of ν and μ coincide.

Problem 1. Find the relation between the asymptotic root counting measure μ and the asymptotic ratio distribution ν for general polynomial families. Is it true that ν depends only on μ, i.e., can two polynomial families with the same asymptotic root counting measure have different asymptotic ratio distributions?

A more concrete question about the asymptotic root counting measure μ may be formulated as follows. Take a sequence $\{(Q_{1,n}(z), \ldots, Q_{k,n}(z))\}_{n\in\mathbb{N}}$ of k-tuples of polynomials and consider the polynomial family $\{p_n(z)\}_{n\in\mathbb{N}}$ satisfying the recurrence relation

$$p_{n+1}(z) = \sum_{i=1}^{k} Q_{i,n}(z)p_{n-i}(z)$$

with initial k-tuple, e.g., $p_{-k+1}(z) = p_{-k+2}(z) = \cdots = p_{-1}(z) = 0,\ p_0(z) = 1$. We say that the polynomial family $\{p_n(z)\}_{n\in\mathbb{N}}$ belongs to the (*generalized*) *Nevai class* if the sequence $\{(Q_{1,n}(z), \ldots, Q_{k,n}(z))\}_{n\in\mathbb{N}}$ converges pointwise (and coefficientwise) to a fixed k-tuple of polynomials $\{(\widetilde{Q}_1(z), \ldots, \widetilde{Q}_k(z))\}$ when $n \to \infty$. Recall from [26] that the (ordinary) Nevai class of orthogonal polynomials consists of all polynomial families $\{p_n(z)\}_{n\in\mathbb{N}}$ satisfying a 3-term recursion of the form

$$zp_n(z) = a_{n+1}p_{n+1}(z) + b_n p_n(z) + a_n p_{n-1}(z),$$

where $a_n > 0$, $b_n \in \mathbb{R}$ for $n \in \mathbb{N}$ and $\lim_{n\to\infty} a_n = a \geqslant 0$, $\lim_{n\to\infty} b_n = b \in \mathbb{R}$. Theorem 5.2 of [26] claims that the asymptotic root counting measure exists for any orthogonal polynomial family in Nevai's class and its density is given by $\mu(x) = \frac{1}{\pi\sqrt{(\beta-x)(x-\alpha)}}$, where $\alpha = b - 2a,\ \beta = b + 2a$. A classical result of Szegö states that in this case the density of ν equals $\nu(x) = \frac{2\sqrt{(\beta-x)(x-\alpha)}}{\pi}$.

Problem 1′. Calculate the asymptotic root-counting measure μ (if it exists) for any polynomial family in the generalized Nevai class.

Problem 1′ is important and nontrivial already in the context of the present paper, namely, for recurrences with fixed polynomial or rational coefficients. Quite recently, we were informed by Sokal that the asymptotic root-counting measure μ is conjecturally given by the distributional derivative of $\log|y_{\mathrm{dom}}(z)|$.

5.2.2.

Let us consider a finite recurrence relation with fixed polynomial coefficients and arbitrarily prescribed initial values of the form

$$p_{n+1}(z) = \sum_{i=1}^{k} Q_i(z)p_{n-i}(z). \tag{5.1}$$

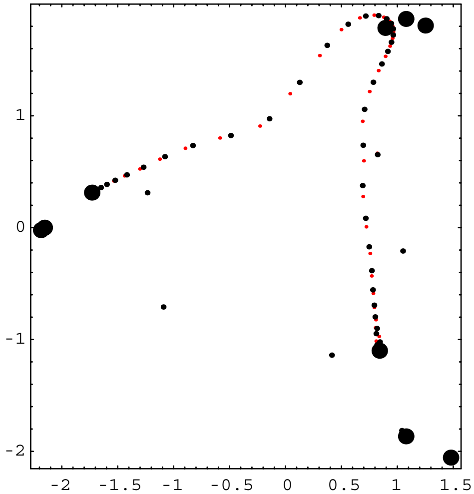

Fig. 3. Zeros of $p_{40}(z)$ and $p_{41}(z)$ satisfying the 4-term recurrence relation $p_{n+1}(z) = (z+1-I)p_n(z) + (z+1)(z-I)p_{n-1}(z) + (z^3+10)p_{n-2}(z)$ with $p_0(z) = z^6 - z^4 + I$, $p_1(z) = z - I + 2$ and $p_2(z) = (2+I)z^2 - 8$.

Denote by Ξ_Q the induced equimodular discriminant of the asymptotic symbol equation of (5.1) viewed as a family of polynomials in the variable z (cf. Definitions 1 and 4). The following fact was observed by the authors in extensive computer experiments and needs both a mathematical ground and a conceptual explanation.

Conjecture 2. *For any polynomial family* $\{p_n(z)\}_{n\in\mathbb{N}}$ *satisfying a finite recurrence relation of form* (5.1) *such that* $\deg p_n(z) = n$ *for* $n \in \mathbb{N}$ *the zeros of any two consecutive polynomials* $p_{n+1}(z)$ *and* $p_n(z)$ *interlace along the curve* Ξ_Q *for all sufficiently high degrees* n.

An example illustrating the interlacing phenomenon conjectured above is shown in Fig. 3. The large fat points in this picture are as usual the branching points. Normal size fat points are the zeros of $p_{41}(z)$ while thin points represent the zeros of $p_{40}(z)$. One notices a typical interlacing pattern between these two sets of zeros.

A simple case of this phenomenon was studied in [18] (see Conjecture 1). Similar interlacing properties for complex zeros along specified curves were noticed in e.g., [6]. Some caution is actually required when defining the interlacing property since the zeros of $p_n(z)$ do not lie exactly on Ξ_Q. In the case when Ξ_Q is a smooth curve we may proceed as follows. Identify some sufficiently small neighborhood $N(\Xi_Q) \subset \Omega$ of Ξ_Q with the normal bundle to Ξ_Q by equipping $N(\Xi_Q)$ with a projection onto Ξ_Q along the fibers which are small curvilinear segments transversal to Ξ_Q. Then one can say that two sets of points in $N(\Xi_Q)$ *interlace* if their projections on Ξ_Q interlace in the usual sense. If Ξ_Q has singularities one should first remove some sufficiently small neighborhoods of these singularities and proceed in the above way on the remaining part of Ξ_Q. The above conjecture then says that for any sufficiently small neighborhood $N(\Xi_Q)$ of Ξ_Q and any of its identifications with the normal bundle to Ξ_Q there exists n_0 such that the interlacing property for the roots of $p_n(z)$ and $p_{n+1}(z)$ holds for all $n \geqslant n_0$. It is worth mentioning that the restriction $\deg p_n(z) = n$ is apparently unnecessary.

5.2.3.

Consider the following analog of recursion (1.3) with variable coefficients

$$-q_n(z) = \sum_{i=1}^{k} \frac{P_{k-i,n}(z)}{P_{k,n}(z)}\, q_{n-i}(z),$$

where $P_{j,n}(z) \in \mathbb{C}[z]$, $0 \leqslant j \leqslant k-1$, $n \in \mathbb{N}$. Assume that the sequence of k-tuples $\{\big(P_{k,n}(z), P_{k-1,n}(z), \ldots, P_{0,n}(z)\big)\}_{n\in\mathbb{N}}$ converges (coefficientwise) uniformly on compact subsets of $\mathbb{C}$ to $\{\big(\widetilde{P}_k(z), \widetilde{P}_{k-1}(z), \ldots, \widetilde{P}_0(z)\big)\}$. As in Section 1 we set $r_n(z) = \frac{q_n(z)}{q_{n-1}(z)}$ for $n \in \mathbb{N}$ and we denote by $\widetilde{y}_{\mathrm{dom}}(z)$ the dominant root of the limit equation

$$\sum_{i=0}^{k} \widetilde{P}_{k-i}(z) y^{k-i} = 0.$$

Problem 2. Is it true that the sequence of rational functions $\{r_n(z)\}_{n\in\mathbb{N}}$ converges uniformly (in the complement of some set of zero Lebesgue measure) to $\widetilde{y}_{\mathrm{dom}}(z)$?

Several results in this direction were obtained in [12]. Note that an affirmative answer to Problem 2 would imply in particular that the coefficients of the 3-term recurrence relation (3.4) can have no limits.

5.2.4.

Let $f(z)$ be an algebraic function and $\{r_n(z)\}_{n\in\mathbb{N}}$ be a sequence of rational approximants to $f(z)$ given by Theorem 2. It would be interesting to know what is the order of contact (in the sense of (3.1)) of $r_n(z)$ and $f(z)$ at ∞. More generally, in which cases does the sequence $\{r_n(z)\}_{n\in\mathbb{N}}$ coincide (up to a finite number of terms) with the sequence $\{[n/n](z)\}_{n\in\mathbb{N}}$ of diagonal Padé approximants to $f(z)$?

5.2.5.

A natural problem arising in the present context is the study of the stratification of the space of algebraic functions according to the topological structure of the induced equimodular discriminant Ξ_f. More precisely, consider the space $\mathcal{A}_k$ of all algebraic equations of degree k. An equation $P(y,z) \in \mathcal{A}_k$ is called *branching nondegenerate* if all its branching points in the z-plane are simple. As noted in Section 4 the endpoints of the curves in Ξ_f can only occur at the branching points of $P(y,z)$ (cf. [10,12]). Thus, the topological structure of Ξ_f can change if and only if either (i) a pair of branching points collapses and $P(y,z)$ becomes branching degenerate or (ii) a 4-tuple of values of $f(z)$ have the same (maximal) absolute value. (Note that for a given z only the branching points for which $|y(z)|$ is maximal are relevant to Ξ_f.) What can be said about this stratification? In particular, how many top-dimensional strata are there and what are their (co)homology groups? A similar stratification related to certain quadratic differentials was studied in [5].

5.2.6.

Well-known results on diagonal Padé approximants at ∞ going back to Markov [25] claim that if the analytic function under consideration is the Cauchy transform of a positive measure supported on a real interval then the poles of all Padé approximants belong to this interval and thus both the Padé Conjecture and Nuttall's Conjecture are trivially valid for such functions (see also [15]). On the other hand, we saw in Section 3 that already algebraic functions which are obtained as the Cauchy transform of a real but not necessarily positive measure supported on an interval might violate Nuttall's Conjecture. It is therefore natural to investigate the Padé Conjecture and Nuttall's Conjecture for algebraic functions representable as Cauchy transforms of positive measures supported on compact subsets of $\mathbb{C}$. An intriguing class of such functions was recently discovered in [11]. Let $k \in \mathbb{N}$ and consider the algebraic curve Γ given by the equation

$$\sum_{i=0}^{k} Q_i(z)w^i = 0, \tag{5.2}$$

where $Q_i(z) = \sum_{j=0}^{i} a_{i,j} z^j$ with $\deg Q_i \leqslant i$ for $0 \leqslant i \leqslant k$. The curve Γ is called of *general type* if the following two nondegeneracy requirements are satisfied:

(i) $\deg Q_k(z) = k$;
(ii) all roots of the (characteristic) equation

$$a_{k,k} + a_{k-1,k-1}t + \cdots + a_{0,0}t^k = 0 \tag{5.3}$$

have pairwise distinct arguments (in particular, 0 is not a root of (5.3)).

In [11] we showed that all the branches of an algebraic curve Γ of general type vanish at ∞ and that each such branch is representable (up to a constant factor) as the Cauchy transform of a certain compactly supported positive measure. We believe that the answer to the following question is affirmative.

Problem 3. Is it true that all the branches of an arbitrary algebraic curve Γ of general type satisfy both the Padé Conjecture and Nuttall's Conjecture near ∞?

5.3. The geometry of the algebraic scheme and multidimensional continuous fractions

Let us finally discuss certain similarities between the scheme for approximating algebraic functions proposed in this paper and some recent developments in multidimensional continuous fractions, see [1,21–23]. Recall first that the conjugates of an algebraic number ξ are by definition all algebraic numbers other than ξ satisfying the irreducible algebraic equation with integer coefficients defining ξ. Note that Theorem 2 has an obvious analog for algebraic numbers.

Definition 15. An algebraic number ξ is called *dominant* if its absolute value is strictly larger than the absolute values of all its conjugates.

Clearly, any dominant algebraic number ξ is necessarily real. Assume additionally that the polynomial with integer coefficients $P_\xi(t) = \alpha_0 t^k + \alpha_1 t^{k-1} + \dots + \alpha_k$ defining ξ is monic, that is, $\alpha_0 = 1$. (This assumption is implicit in all the references quoted above where the defining polynomial is always the characteristic polynomial of a matrix with integer coefficients.) Consider the $k \times k$ matrix with integer coefficients given by

$$M_\xi = \begin{pmatrix} -\alpha_1 & -\alpha_2 & -\alpha_3 & \dots & -\alpha_k \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

The characteristic polynomial of M_ξ coincides with $P_\xi(t)$. Set $\mathbf{v}_0 = (1, 0, \dots, 0)^t$ and let $\mathbf{v}_n = \left(v_n^{(k)}, v_n^{(k-1)}, \dots, v_n^{(1)}\right)^t := M_\xi^n \mathbf{v}_0 \in \mathbb{Z}^k$ for $n \in \mathbb{N}$. Using the same methods as in Section 2 one can easily show that the sequence of vectors $\{\mathbf{v}_n\}_{n\in\mathbb{N}}$ thus defined enjoys the following property:

Lemma 12. *For any $i \in \{1, \dots, k-1\}$ one has* $\lim_{n\to\infty} \frac{v_n^{(i+1)}}{v_n^{(i)}} = \xi$.

Matrices with integer coefficients whose characteristic polynomials have only distinct positive zeros are called *hyperbolic*. Such matrices are the main ingredient in the vast program of studying multidimensional continuous fractions suggested by Arnold (following the ideas of F. Klein). More precisely, assuming that M_ξ is a hyperbolic matrix one gets a decomposition of $\mathbb{R}^k$ into 2^k orthants spanned by all possible choices of directions given by the k one-dimensional eigenspaces of M_ξ. The *sail* of such an orthant is the convex hull of the set of all integer points in $\mathbb{R}^k$ contained in the orthant. If all the eigenspaces of M_ξ are irrational then Dirichlet's theorem on unities implies that the subgroup $G_{M_\xi} \subset GL_k(\mathbb{Z})$ acting on $\mathbb{R}^k$ and preserving all the eigenspaces of M_ξ is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}^{k-1}$. The combinatorial properties of the sails defined by M_ξ and the action of the group G_{M_ξ} on these sails are the subject of interesting and delicate generalizations of Lagrange's theorem on the periodicity of continuous fractions for quadratic irrationalities. A natural problem that arises in this context is as follows.

Problem 4. Let ξ be a dominant algebraic number and assume that its associated matrix M_ξ is hyperbolic. Study the behavior of the sequence $\{\mathbf{v}_n\}_{n\in\mathbb{N}}$ with respect to the union of the sails defined by M_ξ.

Acknowledgements

Communication with Alan Sokal has been of great help in relating our results to those of [7,8]. The authors are obliged to the anonymous referees for their constructive comments and for pointing out several references.

References

[1] V. Arnold, Continuous Fractions, MCCME, 2001, 40pp. (in Russian).
[2] G. Baker, P. Graves-Morris, Padé Approximants, I and II, Encyclopedia of Mathematics and Its Applications, vols. 13 and 14, Addison-Wesley, Reading, MA, 1981.
[3] D. Barrios Rolanía, G. López Lagomasino, Ratio asymptotics for polynomials orthogonal on arcs of the unit circle, Constr. Approx. 15 (1999) 1–31.
[4] D. Barrios Rolanía, G. López Lagomasino, E. Torrano, Distribution of zeros and the asymptotics of polynomials that satisfy three-term recurrent relations with complex coefficients, Mat. Sb. 184 (1993) 63–92 (in Russian) (translation in Russian Acad. Sci. Sb. Math. 80 (1995) 309–333).
[5] Yu. Baryshnikov, On Stokes sets, in: New Developments in Singularity Theory, Cambridge, 2000, NATO Science Series II Mathematical and Physical Chemistry, vol. 21, Kluwer Academic Publishers, Dordrecht, 2001, pp. 65–86.
[6] C.M. Bender, S. Boettcher, V.M. Savage, Conjecture on the interlacing of zeros in complex Sturm-Liouville problems, J. Math. Phys. 41 (2000) 6381–6387.
[7] S. Beraha, J. Kahane, N.J. Weiss, Limits of zeroes of recursively defined polynomials, Proc. Natl. Acad. USA 72 (1975) 4209.
[8] S. Beraha, J. Kahane, N.J. Weiss, Limits of zeros of recursively defined families of polynomials, in: G.-C. Rota (Ed.), Studies in Foundations and Combinatorics, Advances in Mathematics Supplementary Studies, vol. 1, Academic Press, New York, 1978, pp. 213–232.
[9] C. Berenstein, R. Gay, Complex variables, An introduction, Graduate Texts in Mathematics, vol. 125, Springer, New York, 1991.
[10] N. Biggs, Equimodular curves, Discrete Math. 259 (2002) 37–57.
[11] J. Borcea, R. Bøgvad, B. Shapiro, Asymptotics of polynomial eigenfunctions for linear ordinary differential operators, I: homogenized spectral pencils, in preparation.
[12] J. Borcea, B. Shapiro, On the asymptotic ratio of recursively defined families of functions, in preparation.
[13] V.I. Buslaev, On Poincaré's theorem and its applications to problems of the convergence of continued fractions, Mat. Sb. 189 (1998) 13–28 (in Russian) (translation in Sb. Math. 189 (1998) 1749–1764).
[14] V.I. Buslaev, Remarks on Poincaré and Van Vleck theorems, Univ. of North Carolina Industr. Math. Inst. Res. Reports 99:14, 1999.
[15] V.I. Buslaev, On the Baker–Gammel–Wills conjecture in the theory of Padé approximants, Mat. Sb. 193 (2002) 25–38 (in Russian) (translation in Sb. Math. 193 (2002) 811–823).
[16] T.S. Chihara, An Introduction to Orthogonal Polynomials, Gordon & Breach, New York, 1978.

[17] S. Dumas, Sur le développement des fonctions elliptiques en fractions continues, Thèse, Faculté de Philosophie, Université de Zurich, 1908.

[18] Ö. Egecioglu, T. Redmond, C. Ryavec, From a polynomial Riemann hypothesis to alternating sign matrices, Electron. J. Combin. 8 (2001), Research Paper 36, 51pp. (electronic).

[19] A.O. Gel'fond, Differenzenrechnung, Hochschulbücher für Mathematik, vol. 41, VEB Deutscher Verlag der Wissenschaften, Berlin, 1958.

[20] T. Kato, Perturbation theory for linear operators, Die Grundlehren der mathematischen Wissenschaften, vol. 132, Springer, New York, 1966.

[21] E. Korkina, La periodicité des fractions continues multidimensionnelleś, C.R. Acad. Sci. Paris 319 (1994) 777–780.

[22] E. Korkina, The simplest 2-dimensional continued fraction, J. Math. Sci. 82 (1996) 3680–3685.

[23] G. Lachaud, Polyèdre d'Arnol'd et voile d'un cône simplicial: analogues du théorème de Lagrange, C.R. Acad. Sci. Paris 317 (1993) 711–716.

[24] D. Lubinsky, Rogers-Ramanujan and the Baker–Gammel–Wills (Padé) conjecture, Ann. of Math. (2) 157 (2003) 847–889.

[25] A.A. Markoff [Markov], Deux démonstrations de la convergence de certaines fractions continues, Acta Math. 19 (1895) 93–104.

[26] P. Nevai, Orthogonal Polynomials, 4th ed., Memoirs of American Mathematical Society, vol. 213, 1979.

[27] E.A. Rakhmanov, The asymptotic behavior of the ratio of orthogonal polynomials, Mat. Sb. 103(145) (1977) 237–252 (in Russian); E.A. Rakhmanov, Mat. Sb. 118(160) (1982) 104–117, 143 (in Russian).

[28] J. Salas, A. Sokal, Transfer matrices and partition-function zeros for antiferromagnetic Potts models I, J. Statist. Phys. 104 (2001) 609–699 http://arxiv.org/archive/cond-mat/0004330.

[29] A. Sokal, Chromatic roots are dense in the whole complex plane, Combin. Probab. Comput. 13 (2004) 221–261 http://arxiv.org/archive/cond-mat/0012369.

[30] H. Stahl, On the divergence of certain Padé approximants and the behavior of the associated orthogonal polynomials, in: Polynômes Orthogonaux et Applications, Lecture Notes in Mathematics, vol. 1171, Springer, Heidelberg, 1983, pp. 321–330.

[31] H. Stahl, Divergence of diagonal Padé approximants and the asymptotic behavior of orthogonal polynomials associated with nonpositive measures, Constr. Approx. 1 (1985) 249–270.

[32] H. Stahl, Extremal domains associated with an analytic function, I–III, Complex Var. Theory Appl. 4 (1985) 311–354.

[33] H. Stahl, Diagonal Padé approximants to hyperelliptic functions, Ann. Fac. Sci. Toulouse, no spécial Stieltjes, 1996, pp. 121–193.

[34] H. Stahl, The convergence of Padé approximants to functions with branch points, J. Approx. Theory 91 (1997) 139–204.

[35] H. Stahl, Spurious poles in Padé approximation, J. Comput. Appl. Math. 99 (1998) 511–527.

[36] H. Stahl, W. Totik, General orthogonal polynomials, Encyclopedia of Mathematics, vol. 43, Cambridge University Press, Cambridge, 1992.

[37] R. Stanley, Enumerative Combinatorics, vol. 1, Wadsworth Brooks/Cole Math. Ser., The Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, 1986.

[38] G. Szegö, Orthogonal Polynomials, 4th ed., Coll. Publ., vol. XXIII, American Mathematical Society, Providence, RI, 1975.