ANNALES DE LA FACULTÉ DES SCIENCES DE TOULOUSE

HERBERT STAHL

Diagonal Padé approximants to hyperelliptic functions

Annales de la faculté des sciences de Toulouse 6e série, tome S5 (1996), p. 121-193

<http://www.numdam.org/item?id=AFST_1996_6_S5__121_0>

Annales de la Faculté des Sciences de Toulouse n° spécial Stieltjes, 1996

Diagonal Padé approximants to hyperelliptic functions

HERBERT STAHL[1]

RÉSUMÉ. — Nous étudions le problème de convergence des approximants de Padé diagonaux. Comme point de départ nous rappelons certains résultats généraux de convergence en capacité. Ensuite, une étude détaillée de la convergence de la suite des approximants de Padé diagonaux est proposée pour la classe des fonctions hyperelliptiques. L'étude de la localisation des pôles des approximants est d'une importance décisive. Après avoir éliminé certains pôles dits "*spurious*", nous montrons la convergence localement uniforme pour la suite des approximants ainsi modifiés. Ces résultats nous permettent de conclure que, sous certaines restrictions, la conjecture de Baker–Gammel–Wills est valable pour la classe des fonctions hyperelliptiques.

ABSTRACT. — The convergence of diagonal Padé approximants is investigated. Starting with a review of general results that are connected with convergence in capacity the investigation then concentrates on a detailed study of diagonal Padé approximants to hyperelliptic functions. Special emphasis is given to the study of spurious poles. Locally uniform convergence is proved after clearing diagonal Padé approximants from spurious poles, and it is shown that under certain conditions the Baker–Gammel–Wills conjecture holds true for hyperelliptic functions.

MOTS-CLÉS : Approximation rationnelle, approximation de Padé, *spurious* (faux) pôles, fonctions hyperelliptiques, la conjecture de Baker–Gammel–Wills.

KEY-WORDS : Rational approximation, Padé approximation, spurious poles, hyperelliptic functions, Baker–Gammel–Wills conjecture.

AMS Classification : 41A21, 41A25, 30E10.

---

(1) TFH-Berlin/FB 2; Luxemburgerstr. 10; 13353 Berlin; Germany.
Research supported by the Deutsche Forschungsgemeinschaft (AZ: Sta 299/8-1).
The paper was written while the author was visiting the Steklov Institute, Moskow.

1. Introduction

The central topic of the paper is the convergence of diagonal Padé approximants.

DEFINITION 1.1.— *Let the function f be analytic at infinity. The Padé approximant $[m/n]$ of degree $m, n \in \mathbb{N}$ to the function f developed at infinity is the rational function*

$$[m/n](z) := \frac{p_{mn}(1/z)}{q_{mn}(1/z)}, \tag{1.1}$$

where (p_{mn}, q_{mn}) is a pair of Padé polynomials $p_{mn} \in \mathcal{P}_m$, $q_{mn} \in \mathcal{P}_n$, $q_{mn} \not\equiv 0$, satisfying the relation

$$q_{mn}\left(\frac{1}{z}\right) f(z) - p_{mn}\left(\frac{1}{z}\right) = O(z^{-m-n-1}) \quad \textit{as } z \to \infty. \tag{1.2}$$

By $O(\cdot)$ we denote Landau's big "oh", and by $\mathcal{P}_n$ the set of all complex polynomials of degree at most n. It is not difficult to verify that the Padé approximant $[m/n]$ exists and is uniquely determined by (1.2). Uniqueness, however, is not guaranteed for the Padé polynomials p_{mn} and q_{mn}; they can always be multiplied by a non-zero constant, but there may also exist more essential non-uniqueness (cf. [Pe, chap. V], or [BaGM, chap. I]). In what is called the *normal case*, the Padé approximants $[m/n]$ have a contact with the function f at infinity of exact order $m+n+1$. In general this contact may be larger and, what is perhaps more surprising, also smaller than $m+n+1$. These apparent irregularities are only some of the difficulties that have to be dealt with in the convergence theory of Padé approximants. The possibility of so-called *spurious poles* is another more serious difficulty (cf. Definition 2.1, below). A closer study shows that the two phenomena are not independent. Spurious poles will play a major role in the present investigations.

Padé approximants are rational analogues of Taylor polynomials developed at infinity. Infinity as the point of development has been chosen in Definition 1.1 since it leads to simple notation, and Padé approximants are directly connected with (algebraic) continued fractions in this case (cf. [Pe, chap. V]). (In several respects the investigations in the paper can be considered as a continuation of earlier studies about the convergence of continued fractions.)

The Padé approximants $[m/n]$ are called *diagonal* if $m = n$. The diagonal sequence $\{[n/n]\}$ together with the shifted diagonal $\{[n-1/n]\}$ of the *Padé table* $\{[m/n]\}_{m,n\in\mathbb{N}}$ contain the convergents of the corresponding continued fraction of a function f (cf. [Pe, chap. V]). Corresponding continued fractions are the main object of investigation in T. J. Stieltjes' famous paper [Sti]; however, the emphasis in the present paper is somewhat different. More details will follow below.

The main difficulties for a convergence theory for Padé approximants are caused by the possibility of *spurious poles* of the Padé approximants. These are poles that cluster in the domain of analyticity of the function f, or they cluster at a pole of f with a total multiplicity that exceeds the order of the pole of the function f that has to be approximated. Thus, the presence of spurious poles makes locally uniform convergence impossible. It seems that the name "spurious", because of their unwanted nature, was coined by George Baker in the 60's. A more formal definition is given in the next section. Typically, these poles are paired with nearby zeros of the approximant, and they "cancel out" asymptotically as $n \to \infty$. Even in the case of a function f as simple as the square root of a polynomial of fourth order, the sequence of diagonal Padé approximants $[n/n]$, $n \in \mathbb{N}$, can have spurious poles clustering everywhere in $\mathbb{C}$ (cf. Theorem 6.6).

One way to circumvent the difficulties caused by spurious poles is to use a notion of convergence that allows for a certain number or a certain density of such poles. In the present paper this will be *convergence in capacity*. In the next section we assemble general results in this direction, which have been established in recent years. Although the notion of convergence in capacity is not strong enough for many purposes in analysis or in numerical applications, the results, nevertheless, give a good overview of the global convergence behavior of the approximants. They give us information about the shape of the domain in which diagonal Padé approximants typically converge. In contrast to discs, which are typical convergence domains of Taylor series, a geometric description of the convergence domain is more complicated, but nevertheless possible (cf. Theorems 2.3-2.5, below).

It seems that the mere knowledge of analyticity of the function f is practically never sufficient to ensure locally uniform convergence of diagonal or close-to-diagonal Padé approximants. Additional knowledge about the structure of the function always seems to be necessary. Thus, for instance in his fundamental paper [Sti], which is now seen as the starting point of the analytic theory of continued fractions, T. J. Stieltjes considered functions of the form

$$f(z) = \int \frac{\mathrm{d}\mu(x)}{x - z}, \tag{1.3}$$

where μ is a positive measure supported on the negative half-axis $\mathbb{R}_-$, and proved that the corresponding continued fraction converges locally uniformly in $\mathbb{C} \setminus \mathbb{R}_-$ if a certain condition is satisfied, which is equivalent to the moment problem of μ being determinate. Complementary results were proved by A. A. Markov and Hamburger (cf. [Kr] and [Ham]). In the Markov case the measure μ in (1.3) has compact support, while in the Hamburger case, it can have unbounded support in $\mathbb{R}$. In all three cases it is essential that the following three assumptions hold:

(i) f is of the form (1.3),

(ii) μ is a positive measure, and

(iii) $\operatorname{supp}(\mu) \subseteq \mathbb{R}$. In many methodological respects these results belong more to real than to complex analysis.

A rather different approach to studying and understanding the convergence of continued fractions was started by C. G. J. Jacobi [Ja] in 1830, continued by C. W. Borchardt [Bo] and G. H. Halphen [Ha] and brought to a certain conclusion (as far as the convergence aspect is concerned) by S. Dumas in his thesis [Du], supervised by A. Hurwitz. In the investigations, elliptical, hyperelliptical and algebraic functions play a basic role. Dumas investigated the development in continued fractions of the square root $\sqrt{(z-a_1)\cdots(z-a_4)}$ with $a_1, \ldots, a_4 \in \mathbb{C}$, $a_i \neq a_j$ if $i \neq j$. A central place in our investigation is taken by the development at infinity, *i.e.*, by

$$f(z) = \sqrt{(z-a_1)\cdots(z-a_4)} = z^n + b_1 z + b_0 + \cfrac{\alpha_0^2}{z - \beta_1 + \cfrac{\alpha_1^2}{z - \beta_2 + \cdots}}. \tag{1.4}$$

This development was also the object of investigations in [Ja] and was then generalized in [Bo] and [Ha]. In (1.4) the zeros $a_1, \ldots, a_4$ may be complex, and the complex nature of the function f already underlines the difference between the Stieltjes, Markov, and Hamburger theory. Dumas showed among other things that, with respect to convergence, there are three cases to be distinguished. The distinction depends on the arithmetic character of the integral

$$\int_C \frac{d\zeta}{\sqrt{(\zeta - a_1)\cdots(\zeta - a_4)}}, \tag{1.5}$$

where C is an integration path connecting infinity on the first sheet with infinity on the second sheet of the Riemann surface $\mathcal{R}$ defined by $y^2 = (z - a_1)\cdots(z - a_4)$. In the first two cases the continued fraction in (1.4) converges to f locally uniformly outside of finitely many arcs and a finite set of points. A totally different convergence behavior appears in the third case: there exist two infinite sets Σ_1, Σ_2, both dense in $\mathbb{C}$, and the continued fraction in (1.4) converges pointwise to f on Σ_1, but diverges at each point of Σ_2. Hence, the convergence behavior is rather irregular in the third case, and is not locally uniform anywhere in $\mathbb{C}$. More details about Dumas' results will be given in Section 6. For a more complete historic survey of the development of the theory from the time of Jacobi up to the 20's in the present century, we recommend Chapter 5 of [Br].

In a broad sense the material of Section 3 can be seen as a continuation and an extension of Dumas' investigations. Instead of the continued fractions (1.4) we now study diagonal Padé approximants, which are the same in substance, and instead of the square root of a fourth order polynomial, we study now the approximation of hyperelliptic functions (for a definition see (3.1) at the beginning of Section 3). A central place is taken by the investigation of spurious poles. Their number and distribution is studied, and it is shown that in case of a hyperelliptic function f, diagonal Padé approximants can have only a finite number of them. An upper bound is given for this number. In connection with the study of the distribution of spurious poles the Jacobi inversion problem on a compact Riemann surface plays a central role. As an application of this investigation it is possible to prove the Baker–Gammel-Wills conjecture with some restrictions: the function has to satisfy an additional condition, and it may be necessary to vary the point of development slightly away from infinity. It is probable that these restrictions are not really necessary and it may be possible to eliminate them. Baker–Gammel–Wills have conjectured in [BGW] that if a function f is meromorphic in the unit disc, then at least an infinite subsequence of diagonal Padé approximants (developed at the origin) will converge to f locally uniformly in the unit disc minus the poles of f. In general the conjecture may well be false. It turns out that in the case of hyperelliptic functions the conjecture does not only hold in discs, but also in the larger convergence domains that will be introduced in the next section, and which are typical for Padé approximants.

The outline of the paper is as follows: in the next section we survey general convergence results associated with convergence in capacity. In Section 3 we formulate and discuss new results about the convergence of diagonal Padé approximants to hyperelliptic functions. The results of Section 3 will be proved in Sections 4 and 5. Proofs that are connected with the Jacobi inversion problem, *i.e.*, the proofs of Theorem 3.3 and 3.8, are given in Section 5. In Section 6 a principal result from Dumas' dissertation is reproduced.

2. General Convergence Results

Convergence results are discussed that hold true for a large class of functions. These functions are characterized only by analyticity properties, however a price has to be paid for this generality: the form of convergence that can be proved is rather weak, in our case it is convergence in capacity. This concept of convergence allows spurious poles to cluster inside of the domain of convergence. Special emphasis will be given to functions with branch points. The section is closed by a more formal definition of spurious poles.

Results in the present section are connected with (logarithmic) capacity in several respects. The capacity is denoted by $\operatorname{cap}(\cdot)$ (for its definition see [StTo, Appendix I], or any other book on potential theory). One of the reasons why capacity plays such an omnipresent role in the subject stems from the fact that capacity is extremely well suited for measuring a filled lemniscate, *i.e.*, the set on which the modulus of a given monic polynomial is small. We start the review of general results by the

POMMERENKE–NUTTAL THEOREM ([Nu1], [Po2]).— *Let the function f be analytic (and single-valued) in the domain $\overline{\mathbb{C}} \setminus E$, where $E \subseteq \mathbb{C}$ is a compact set with* $\operatorname{cap}(E) = 0$. *Then for any compact set $V \subseteq \mathbb{C}$ and $\varepsilon > 0$ we have*

$$\lim_{n\to\infty} \operatorname{cap}\left\{z \in V \mid \left|(f - [n/n])(z)\right| > \varepsilon^n\right\} = 0. \tag{2.1}$$

Motivated by this Theorem *convergence in capacity* is defined as an analogue to convergence in measure.

DEFINITION 2.1.— *A sequence of functions* f_n, $n = 0, 1, 2, \ldots$, *is said to converge in capacity to* f *in a domain* $D \subseteq \overline{\mathbb{C}}$ *if for every* $\varepsilon > 0$ *and every compact set* $V \subseteq D \cap \mathbb{C}$ *we have* $\operatorname{cap}\{z \in V \mid |(f - f_n)(z)| > \varepsilon\} \to 0$ *as* $n \to \infty$.

From (2.1) it follows that the diagonal Padé approximants $[n/n]$ converge in capacity to f in $\mathbb{C}$ if the assumptions of the Nuttall–Pommerenke Theorem are satisfied. But even more, (2.1) shows that we have a convergence speed faster than geometric outside of some exceptional sets that become small in capacity as $n \to \infty$. It is not necessary to exclude the singularities of the function f from the convergence domain since $\operatorname{cap}(E) = 0$.

In [Po2] the Nuttall–Pommerenke Theorem has been proved not only for diagonal sequences of Padé approximants, but also for arbitrary sectorial sequences $\{[m/n]\}$, *i.e.*, for sequences with $\lambda > 0$ such that

$$\lambda n \leq m \leq \frac{n}{\lambda} \quad \text{as} \quad m, n \to \infty. \tag{2.2}$$

The theorem has been extended to fast approximable functions f in [Go]. These are functions that can be rationally approximated faster than geometrically on a set of positive capacity. In [Go], the connection between fast approximable functions and the property of single-valuedness has also been investigated.

The assumption $\operatorname{cap}(E) = 0$ is essential in the Nuttall–Pommerenke Theorem since in [Lu] and [Ra] it has been shown by counterexamples that if the function f has a set of singularities $\tilde{E} \subseteq \mathbb{C}$ of positive capacity, then it is no longer necessarily true that the diagonal sequence $\{[n/n]\}$ converges in capacity to f.

The situation is somewhat different if the function f has branch points, for instance if f is an algebraic function. In this case the set of singularities E may be of capacity zero, but the function f is not single-valued in $\overline{\mathbb{C}} \setminus E$. Since rational functions are single-valued, the function f cannot be approximated by $[n/n]$, even in capacity, throughout $\overline{\mathbb{C}} \setminus E$. Approximation is possible only in subdomains $D \subseteq \overline{\mathbb{C}}$ in which the function f is single-valued. Hence, in the case of functions with branch points, some cuts in $\overline{\mathbb{C}}$ are necessary in order to make f single-valued. It will turn out that Padé approximants determine such cuts in a very interesting way.

Let $D \subseteq \overline{\mathbb{C}}$ be a domain with $\infty \in D$ and $\operatorname{cap}(\partial D) > 0$. Then the Green function $g_D(z, w)$ exists in D (for a definition see [StTo, Appendix V]), and we define

$$G(z) = G_D(z) := \exp\left[-g_D(z, \infty)\right]. \tag{2.3}$$

It follows that $0 \leq G(z) < 1$ for all $z \in D$, $G(z) > 0$ for $z \in D \setminus \{\infty\}$, and we have $G(z) = 1$ for quasi every $z \in \partial D$ (cf. [StTo, Appendix V]). A property is said to hold *quasi everywhere* on a set $S \subseteq \overline{\mathbb{C}}$ if it holds for all $z \in S$ except on a subset of outer capacity zero. Capacity is usually defined only for bounded sets in $\mathbb{C}$. However, the definition of capacity zero can be extended to $\overline{\mathbb{C}}$ in an obvious way by a Moebius transform.

THEOREM 2.2 ([St3, Theorem 1.1 and 1.2]).— *Let $E \subseteq \overline{\mathbb{C}}$ be a compact set with $\operatorname{cap}(E) = 0$, let the function f be locally analytic in $\overline{\mathbb{C}} \setminus E$, assume that f has branch points, and let the branch, which is still denoted by f, be analytic at infinity. Then there exists a domain $D = D_f \subseteq \overline{\mathbb{C}}$, called the convergence domain, with $\infty \in D$, $\operatorname{cap}(\partial D) > 0$, and for all compact sets $V \subseteq D$ and $\varepsilon > 0$ we have*

$$\lim_{n\to\infty} \operatorname{cap}\left\{z \in V \mid \left|(f - [n/n])(z)\right| > \left(G(z) + \varepsilon\right)^{2n}\right\} = 0, \tag{2.4}$$

and if $0 < \varepsilon \leq G(z)$ for all $z \in V$, then also

$$\lim_{n\to\infty} \operatorname{cap}\left\{z \in V \mid \left|(f - [n/n])(z)\right| < \left(G(z) - \varepsilon\right)^{2n}\right\} = 0. \tag{2.5}$$

The limit (2.4) implies that the sequence $\{[n/n]\}$ converges in capacity to f in the domain D. In contrast to the Nuttall–Pommerenke Theorem the convergence is now only geometric (up to exceptions on the subsets that are allowed under convergence in capacity). The limit (2.5) shows that the value $G(z)$ is the exact convergence factor at the point $z \in D$, but again this holds only up to exceptional sets.

In [St3] Theorem 2.2 has been proved not only for diagonal sequences, but also for close-to-diagonal sequences $\{[n/n]\}$, *i.e.*, for sequences with numerator and denominator degrees satisfying

$$\frac{m}{n} \longrightarrow 1 \quad \text{as} \quad m, n \longrightarrow \infty. \tag{2.6}$$

A typical application of Theorem 2.2 is the approximation of algebraic functions f. In this case the set E of singularities is finite, and the diagonal

Padé approximants $[n/n]$ approximate one branch of f in a domain D that differs from $\overline{\mathbb{C}}$ by several cuts that connect branch points of f (cf. Theorem 2.4, below). For the special case of a hyperelliptic function f the convergence behavior of diagonal Padé approximants $[n/n]$ will be studied in more detail in the next section. For the results proved there Theorem 2.2 provides only a general frame.

In the next two theorems, among other things, a geometric description of the convergence domain D_f is given, which is independent of the approximation problem.

THEOREM 2.3

(a) If the function f is analytic at infinity, then there exists a domain $D \subseteq \overline{\mathbb{C}}$, which is uniquely determined by the following three conditions:

(i) $\infty \in D$ and f has a single-valued meromorphic continuation in D.

(ii) $\operatorname{cap}(\partial D) = \inf_{\tilde{D}} \operatorname{cap}(\partial \tilde{D})$, where the infimum extends over all domains $\tilde{D}$ that satisfy condition (i).

(iii) $D = \bigcup \tilde{D}$, where the union extends over all domains $\tilde{D}$ that satisfy the two conditions (i) and (ii).

(b) If the function f satisfies the assumptions of Theorem 2.2, then the domain determined by the conditions (i)–(iii) is identical with the convergence domain D_f of Theorem 2.2 up to a set of capacity zero.

Part (a) of Theorem 2.3 has been proved in [St1, Theorem 1 and 2], and part (b) has been proved in [St3, Theorem 1.4]. In [St1] condition (i) refers to analytic and not meromorphic continuation of f throughout D. However the difference is irrelevant since poles are isolated and denumerable, and therefore form a set of capacity zero. It may be interesting to note that the convergence domain is always dense in $\overline{\mathbb{C}}$, but more specific information is given in the next theorem, which contains a description of the topological structure of the complement F of the convergence domain D.

THEOREM 2.4

(a) Assume that the function f satisfies the assumptions of Theorem 2.2. Then the complement F of the convergence domain D has the structure

$$F := \overline{\mathbb{C}} \setminus D = F_0 \cup \bigcup_{j \in I} J_j, \tag{2.7}$$

where $F_0 \subseteq \mathbb{C}$ *is a compact set with* $\operatorname{cap}(F_0) = 0$, $F_0 \setminus E$ *consists of points that are isolated in* $\overline{\mathbb{C}} \setminus (E \cap F_0)$, *the* J_j, $j \in I$, *are open analytic arcs, and* $I \neq \emptyset$.

(b) *The Green function* $g_D(z, \infty)$ *has identical normal derivatives on both sides of the arcs* J_j, *i.e.,*

$$\frac{\partial}{\partial n_+} g_D(z, \infty) = \frac{\partial}{\partial n_-} g_D(z, \infty) \quad \textit{for all} \;\; z \in J_j, \; j \in I, \tag{2.8}$$

where $\partial/\partial n_+$ *and* $\partial/\partial n_-$ *denote the normal derivatives on both sides of the arcs* J_j.

Relation (2.8) is called the *symmetry property* of the Green function $g_D(z, \infty)$, and like-wise it is said that the domain D possesses the symmetry property if $\infty \in D$ and both properties (2.7) and (2.8) hold true. The symmetry property (2.8) implies a characterization of the analytic arcs J_j, $j \in I$, by quadratic differentials.

COROLLARY 2.5.— *Let* $h_D^*(z, \infty)$ *be the conjugate harmonic function to* $g_D(z, \infty)$ *(which is not single-valued) and define*

$$Q(z) := \left[\left(g_D(z, \infty) + i h_D^*(z, \infty) \right)' \right]^2 \quad \textit{for } z \in \mathbb{C} \setminus F_0 \,, \tag{2.9}$$

so that Q *is analytic in* $\overline{\mathbb{C}} \setminus (E \cap F_0)$ *and has a zero of order 2 at infinity. Let* $h_D^*(z, \infty)$ *be normalized in such a way that* $Q(z) z^2|_{z=\infty} > 0$. *Then the arcs* J_j, $j \in I$, *are trajectories of the quadratic differential* $Q(z)\,\mathrm{d}z^2$, *i.e.,*

$$Q(\alpha_j(t))\alpha_j'(t)^2 < 0 \quad \textit{for } t \in (0,1)\,, \; j \in I, \tag{2.10}$$

where $\alpha_j : [0,1] \to \mathbb{C}$ *is a smooth representation of the arc* J_j, $j \in I$.

Theorem 2.4 and Corollary 2.5 follow from [St2], Theorem 1 and the associated corollary. In case of an algebraic function f, the function Q in (2.9) is rational and all its poles lie at branch points of the function f, but not all branch points have to be poles of Q. The fact that the arcs J_j, $j \in I$, are trajectories of the quadratic differential $Q(z)\,\mathrm{d}z^2$ can be used for calculating the arcs J_j numerically.

Before we continue with the presentation of new results we will discuss a concrete example. The given function possesses two symmetries, which facilitates the determination of the convergence domain D. In general, the determination of D can cause serious problems and satisfactory methods for doing so apparently still do not exist.

Example. — We consider the function

$$f(z) := \sqrt{1 - \frac{2}{z^2} + \frac{9}{z^4}}. \tag{2.11}$$

The function has four branch points at $z_{1,\ldots,4} = \pm\exp(\pm i\pi/6)$ and a double pole at the origin. Thus, $E = \{z_1, \ldots, z_4, 0\}$. Let $D = D_f$ denote the convergence domain in accordance with Theorem 2.2. We have $D = \overline{\mathbb{C}} \setminus (\overline{J_1} \cup \overline{J_2})$, where J_1 and J_2 are two arcs that connect the points z_2 with z_3 and z_1 with z_4, respectively, (see figure 2.1). These arcs are trajectories of a quadratic differential, with

$$\frac{z^2}{z^4 - 2z^2 + 9}\, dz^2 \le 0 \quad \text{for } z \in J_j. \tag{2.12}$$

The function (2.3) can be defined in this case by

$$G_D(z) = \exp\left[-\mathrm{Re} \int_{z_1}^{z} \frac{\zeta\, d\zeta}{\sqrt{9 - 2\zeta^2 + \zeta^4}}\right]. \tag{2.13}$$

In figure 2.1 the poles of the Padé approximant [40/40] are shown. Close to the origin lie two simple poles that approximate the double pole of f

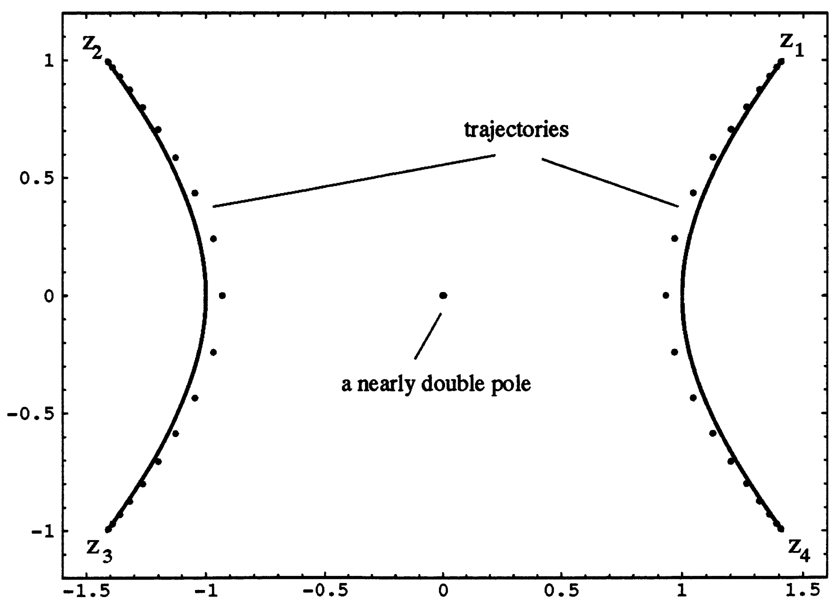

Fig. 2.1 The poles of the Padé approximant [40/40] together with the two trajectories J_1 and J_2 of the quadratic differential (2.12) that connect the pairs of points (z_1, z_4) and (z_2, z_3).

at the origin. The remaining 38 poles lie close to the arcs J_1 and J_2. It follows from the results in Section 6 that in this special case the diagonal Padé approximants $[n/n]$, $n \in \mathbb{N}$, converge locally uniformly to f in $\overline{\mathbb{C}}\setminus(\overline{J_1}\cup\overline{J_2}\cup\{0\})$, and locally uniformly in the spherical metric in $\overline{\mathbb{C}}\setminus(\overline{J_1}\cup\overline{J_2})$.

A similar result to that of Theorem 2.2, but in a rather different setting, has been proved in [GoRa, sect. 3] (see also [St3, Theorem 1.7]). There it has been shown that if the domain $D \subseteq \overline{\mathbb{C}}$ possesses the symmetry property, *i.e.*, if the complement of D is of the form (2.7) and the Green function $g_D(z, \infty)$ satisfies the symmetry condition (2.8), and if further the function f is single-valued and analytic in D and has sufficiently nice boundary values on $\partial D \setminus F_0$ (for a precise statement of the last condition see [St3, Theorem 1.7]), then the conclusions (2.4) and (2.5) of Theorem 2.2 hold true. These results show that much less analyticity of the function f is needed than demanded in Theorem 2.2. The results in [St3] further show that the symmetry property (2.8) is typical for convergence domains of diagonal Padé approximants. Since in the present paper our main interest is the approximation of hyperelliptic functions, the additional analyticity assumed in Theorem 2.2 always holds. This argument, of course, also holds in the case of algebraic functions.

In all results discussed so far only convergence in capacity has been proved and discussed. Under assumptions as general as those in Theorem 2.2 locally uniform convergence is in general not true. The reason for this is that poles of the Padé approximants $[n/n]$ may cluster inside the convergence domain D. It has already been mentioned earlier that such poles are called spurious because of their unwanted nature. Their investigation will be one of the central topics in the next section. A formal definition of the spuriousness of poles has to be based on their asymptotic behavior, since only if n tends to infinity does the character of the poles of the approximants $[n/n]$ becomes fully clear.

We prepare the definition of spurious poles by discusing an interesting difference between rational and polynomial approximation that may shed light on the underlying problem. Let us consider the power series of a function f analytic at the origin. All zeros of the partial sums of this power series have to leave the disk of convergence as n tends to infinity, except for those zeros that approximate zeros of the function f. This behavior is a rather immediate consequence of the argument principle or Rouché's Theorem. In case of rational approximation such a conclusion can, however, not be drawn. Now, neither all zeros nor all poles of the approximants have to leave the convergence domain (for convergence in capacity). The argument principle or Rouché's Theorem only guarantees that besides those zeros and poles that approximate corresponding zeros and poles of the function f, the other poles and zeros have to appear in pairs so that they are neutral with respect to the argument principle, but they need not leave the convergence domain. Such pairs of poles and zeros really exist and can cluster inside the convergence domain. In a neighborhood of such clusterpoints uniform convergence is impossible. It is clear that in polynomial approximation such poles cannot exist.

DEFINITION 2.2.— *Let $[n/n]$ be a diagonal Padé approximant to the function f and let N be an infinite subsequence of $\mathbb{N}$. Spurious poles are defined in two circumstances.*

(i) For each $n \in N$ let the approximant $[n/n]$ have a pole at $z_n \in \overline{\mathbb{C}}$ and $z_n \longrightarrow z_0$ as $n \longrightarrow \infty$, $n \in N$. If the function f to be approximated is analytic in a neighborhood of z_0, and if the approximants $[n/n]$ converge in capacity to f in a neighborhood of z_0, then the poles of the $[n/n]$ at the points z_n, $n \in N$, are called spurious. If $z_0 = \infty$, then the convergence $z_n \longrightarrow z_0$ has to be understood in the spherical metric (for a definition see (3.14), below).

(ii) Let the function f have a pole of order k_0 at $z_0 \in \overline{\mathbb{C}}$ and let the total order of poles of the approximants $[n/n]$ near z_0 be $k_1 = k_{1n} > k_0$ for each $n \in N$, i.e., $[n/n]$ has poles at points z_{nj}, $j = 1, \ldots, m_n$, with total order k_{1n} and $z_{nj_n} \longrightarrow z_0$ as $n \longrightarrow \infty$, $n \in N$, for any selection of $j_n \in \{1, \ldots, m_n\}$. If the approximants $[n/n]$ converge in capacity to f in a neighborhood of z_0, then poles of order $k_{1n} - k_0$ of the approximant $[n/n]$, $n \in N$, are considered as being spurious.

It has already been mentioned that spurious poles of the approximants $[n/n]$ are paired with zeros of the same approximant, and asymptotically the distance between zeros and poles becomes arbitrarily small (for a more detailed description see Theorem 3.6, below).

In Theorem 1.8 of [St3] it has been shown that almost all poles of the Padé approximants $[n/n]$ converge to the boundary ∂D of the convergence domain. This implies that almost all poles are non-spurious.

THEOREM 2.6.— *Let the function f satisfy the assumptions of Theorem 2.2 and let k_n be the total number of spurious poles of the approximant $[n/n]$. Then*

$$\lim_{n\to\infty} \frac{k_n}{n} = 0. \tag{2.14}$$

The limit (2.14) gives a rather rough estimate of the possible number of spurious poles. It has been conjectured by J. Nuttall in [Nu2] among many other results that in case of an algebraic function f the number of spurious poles is bounded. In the next section such an upper bound is proved for hyperelliptic functions.

3. Approximation of Hyperelliptic Functions

The convergence of diagonal Padé approximants to a hyperelliptic function f is investigated. Central topics are the behavior of the approximants near poles of the function f, the number and the distribution of spurious poles of the approximants, the convergence after pole elimination, a proof of the Baker–Gammel–Wills conjecture under an additional condition, and the necessity to move the point of development away from infinity in some cases. All results stated and discussed in the present section will be proved in the Sections 4 and 5.

Throughout this section the function f is assumed to be hyperelliptic, analytic at infinity, and to have $2m$ branch points ($m \geq 1$) at $a_1, \ldots, a_{2m} \in \mathbb{C}$ with $a_i \neq a_j$ for $i \neq j$. Such a function can be represented by

$$f(z) = r_1(z) + r_2(z)y \quad \text{with} \quad y := \sqrt{(z-a_1)\cdots(z-a_{2m})} \tag{3.1}$$

and r_1 and r_2 rational functions. The sign of the square root should be chosen in such a way that $y = z^m + \ldots$ near infinity. Let $\mathcal{R}$ denote the Riemann surface defined by the equation $y^2 = (z-a_1)\cdots(z-a_{2m})$, and denote by $\pi : \mathcal{R} \longrightarrow \overline{\mathbb{C}}$ the canonical projection. The Riemann surface $\mathcal{R}$ is compact, has two sheets, and genus $g = m-1$. The function f can be lifted to $\mathcal{R}$, where it is a meromorphic function. In the present section we use the same symbol f for the function defined on $\mathbb{C}$ as well as its lifting to $\mathcal{R}$.

A hyperelliptic function f satisfies the assumptions of Theorem 2.2. Hence, there exists a convergence domain $D = D_f \subseteq \overline{\mathbb{C}}$, in which the sequence $\{[n/n]\}_{n\in\mathbb{N}}$ of diagonal Padé approximants converges to f in capacity. It is not difficult to verify that the complementary set $F = \overline{\mathbb{C}} \setminus D$ consists of finitely many analytic, closed arcs J_j, $j \in I$, (cf. Theorem 2.4, above). These arcs connect subsets of the set $\{a_1, \ldots, a_{2m}\}$ of branch points; each subset contains an even number of points. In the most typical situations there exist exactly m arcs J_j, $j = 1, \ldots, m$, each one connecting a pair of branch points.

Since the function f is single-valued in the convergence domain D (cf. Theorem 2.3, above), the lifting $\pi^{-1}(D)$ of D onto the Riemann surface $\mathcal{R}$ is an open set consisting of two domains B_1 and B_2. The complement $\mathcal{R}\setminus(B_1 \cup B_2)$ consists of a chain of closed curves separating B_1 and B_2, and it is denoted by Γ. Its projection $\pi(\Gamma) = F$ consists of the arcs J_j, $j \in I$. Thus, we have

$$\mathcal{R} = B_1 \cup \Gamma \cup B_2, \quad \pi(B_j) = D, \ j = 1, 2, \quad \pi(\Gamma) = F = \bigcup_{j\in I} J_j. \tag{3.2}$$

In the present section the function f is almost always considered as a function defined on $D \subseteq \overline{\mathbb{C}}$, however in some situations it is necessary to consider $f$as a function on $\mathcal{R}$. If this is the case, then it is assumed that the function f has identical values in D and B_1.

The first theorem in the present section is concerned with those poles of the Padé approximants $[n/n]$ that do not cluster on $F = \partial D$. Most interesting is the upper bound (3.3) for the number of spurious poles.

THEOREM 3.1.— *Let the function f be hyperelliptic and analytic at infinity.*

(i) Let f have a pole of order k at $z_0 \in D$. Then for each $n \in \mathbb{N}$ sufficiently large the Padé approximant $[n/n]$ has one or several poles near z_0 with a total order of at least k, and all these poles converge to z_0 as $n \longrightarrow \infty$. If the total order $k_1 = k_{1n}$ of these poles is larger than k, then $k_{1n} - k$ of the poles of the approximant $[n/n]$ near z_0 are spurious, the approximant $[n/n]$ has $k_{1n} - k$ zeros near z_0, and these zeros converge also to z_0 as $n \longrightarrow \infty$.

(ii) Let n_f^- denote the number of poles of the function f (lifted to $\mathcal{R}$) on Γ, and let $n_{r_2}^+$ and $n_{r_2}^-$ denote the number of zeros and poles of the

rational function r_2 *in the representation (3.1) on* F, *taking care of multiplicities. Then the Padé approximant* $[n/n]$, $n \in \mathbb{N}$, *has at most*

$$m - 1 + n_f^- + n_{r_2}^+ - n_{r_2}^- \tag{3.3}$$

spurious poles.

Remarks

(1) If the Padé approximant $[n/n]$ has poles of total order k_{1n} near z_0, and if k_{1n} is larger than the order k of the pole of f at z_0, then it already follows from Definition 2.1, (ii), that the surplus order $k_{1n} - k$ of poles has to be considered as spurious. This phenomenon is reflected by the fact that the approximant $[n/n]$ has one or several zeros near z_0 of the same total order $k_{1n} - k$.

(2) The poles of the approximant $[n/n]$ considered in (i) and (ii) of Theorem 3.1 are the only ones that do not cluster on $F = \partial D$. From (3.3) it follows that at most $k_0 + m - 1 + n_f^- + n_{r_2}^+ - n_{r_2}^-$ poles of $[n/n]$ can cluster outside of F, where k_0 denotes the total order of all poles of the function f in D.

(3) Of course, the upper bound (3.3) is much stronger than the estimate given in Theorem 2.6 under less restrictive assumptions. However, it seems that the upper bound (3.3) is in general not sharp. In order to have a sharp result we consider a more restricted situation in the next corollary. The corollary follows immediately from Theorem 3.1.

COROLLARY 3.2.— *Let the function* f *be hyperelliptic and analytic at infinity. Let the rational function* r_2 *be defined by representation (3.1), and assume that* r_2 *has no zeros or poles on* $F = \partial D$ *and that the meromorphic continuation of* f *has no poles on* F. *Then the Padé approximant* $[n/n]$, $n \in \mathbb{N}$, *has at most* $m - 1$ *spurious poles.*

Remark.— Note that $g = m - 1$ is the genus of the Riemann surface $\mathcal{R}$ associated with the function f. The assumptions of Corollary 3.2 imply $n_f^- = n_{r_2}^+ = n_{r_2}^- = 0$, and therefore the number (3.3) is equal to $m-1$ under these assumptions.

In principle we consider in the present paper only Padé approximants developed at infinity, but in the next theorem, where we show that the upper bound (3.3) is sharp under the assumptions of Corollary 3.2, it will be necessary also to consider Padé approximants developed at a point $\zeta_0 \in \mathbb{C}$ near infinity.

DEFINITION 3.1.— *For $\zeta_0 \in \mathbb{C}$ the Moebius transform*

$$\psi(z) := \frac{z}{1 - \zeta_0^{-1} z} \tag{3.4}$$

maps ζ_0 to ∞, and therefore the function $\tilde{f} := f \circ \psi^{-1}$ is analytic at infinity if f is analytic at ζ_0. Let $\widetilde{[n/n]}$, $n \in \mathbb{N}$, be the diagonal Padé approximant to $\tilde{f}$ developed at infinity, i.e., $\widetilde{[n/n]}$ is defined by (1.1) and (1.2). Then the rational function

$$[n/n](z) := \widetilde{[n/n]}(\psi(z)) = \frac{p_{nn}(z)}{q_{nn}(z)} \tag{3.5}$$

is the Padé approximant $[n/n]$ to f developed at ζ_0. From (1.1) and (1.2) it follows, after some calculations, that

$$(q_{nn} f - p_{nn})(z) = O\left((z - \zeta_0)^{2n+1}\right) \quad \text{as } z \to \zeta_0 \tag{3.6}$$

with p_{nn} and q_{nn} polynomials of degree not greater than n.

THEOREM 3.3.— *Let the function f be hyperelliptic and analytic at infinity, and let the assumptions of Corollary 3.2 be satisfied. For almost all $\zeta_0 \in \overline{\mathbb{C}}$ in a neighborhood of infinity as point of development the following holds true: for any selection of $m-1$ points $\pi_1, \ldots, \pi_{m-1} \in D$ there exists an infinite subsequence $N \subseteq \mathbb{N}$ such that the Padé approximants $[n/n]$, $n \in N$, developed at ζ_0 have $m-1$ spurious poles $\pi_{1n}, \ldots, \pi_{m-1,n} \in \mathbb{C}$ and*

$$\pi_{jn} \to \pi_j \quad \text{as } n \to \infty,\ n \in N \ \text{ for } j = 1, \ldots, m-1. \tag{3.7}$$

If some $\pi_j = \infty$, then the convergence (3.7) has to be understood in the spherical metric.

Remarks

(1) Theorem 3.3 shows that the upper bound (3.3) for the number of spurious poles is sharp under the assumptions of Corollary 3.2. For the proof of Theorem 3.3 the assumptions of Corollary 3.2 are essential. At the end of Section 5 the question of what may be true without these assumptions will be discussed.

(2) Since the points $\pi_1, \ldots, \pi_{m-1}$ are arbitrary, Theorem 3.3 shows that spurious poles of the sequence $\{[n/n]\}_{n\in\mathbb{N}}$ cluster everywhere in D, and

consequently the Padé approximants $[n/n]$, $n \in \mathbb{N}$, cannot converge locally uniformly to f on any subdomain of D. (From Theorem 2.2 we know that the approximants converge in capacity everywhere in D).

(3) At the end of Section 5 we will discuss which weaknesses of the method of the proof of Theorem 3.3 make it necessary to exclude certain points as points of development.

From Theorem 3.3 we know that under the assumptions of Corollary 3.2 the upper bound $m-1$ in (3.3) is sharp, however this bound is not sharp in general (cf. the discussion at the end of Section 5). At present, it is not fully clear what a sharp upper bound might be. That the zeros of r_2 on F have to be considered in (3.3) in some way is shown by

LEMMA 3.4.— *Let $0 < \alpha_1 < \alpha_2 < \pi$ be chosen in such a way that the numbers α_1, α_2, π are linearly independent over the rational numbers $\mathbb{Q}$, and choose $c_0, c_1, c_2 \in \mathbb{R}$ such that*

$$f(z) := (z - \cos\alpha_1)(z - \cos\alpha_2)\sqrt{z^2 - 1} - (z^3 + c_2 z^2 + c_1 z^1 + c_0) \quad (3.8)$$

is analytic at infinity. Then there exists an infinite subsequence $N \subseteq \mathbb{N}$ such that the Padé approximant $[n/n]$, $n \in N$, has two spurious poles. Further, for any $\zeta \in \mathbb{C} \setminus [-1,1]$ there exists an infinite subsequence $N \subseteq \mathbb{N}$ such that at least one spurious pole of $[n/n]$, $n \in N$, converges to ζ as $n \to \infty$, $n \in N$.

Remark.— It follows from the Theorems 2.2 and 2.3 that the Padé approximants $[n/n]$, $n \in N$, of the function f defined in (3.8) has $\overline{\mathbb{C}} \setminus [-1,1]$ as its convergence domain D. The Riemann surface $\mathcal{R}$ associated with the function f is of genus $1 - 1 = g = 0$. The two functions f and $r_2 := (\cdot - \cos\alpha_1)(\cdot - \cos\alpha_2)$ have no poles on Γ and $F = [-1,1]$, respectively, but r_2 has two zeros on $F = [-1,1]$, *i.e.*, $n_f^- = 0$, $n_{r_2}^- = 0$, and $n_{r_2}^+ = 2$. Thus, the number in (3.3) is equal to 2, and Lemma 3.4 shows that for the function (3.8) the upper bound (3.3) is sharp.

The function G_D in (2.3) has been defined with the help of the Green function $g_D(z, \infty)$ of the convergence domain D. Since in case of a hyperelliptic function f the complement $F = \overline{\mathbb{C}} \setminus D$ consists of finitely many arcs, it is a regular set with respect to the Dirichlet problem in D, and we have $g_D(z, \infty) = 0$ for all $z \in F = \partial D$. This implies that

$$\begin{aligned} &G_D(z) = 1 \quad \text{for all } z \in F\,, \quad G_D(\infty) = 0\,, \\ &\qquad\qquad \text{and } 0 < G_D(z) < 1 \quad \text{for } z \in D \setminus \{\infty\}. \end{aligned} \quad (3.9)$$

In the next theorem an estimate is proved for the speed with which poles of the approximants $[n/n]$ are attracted by poles of f in D as $n \to \infty$. It turns out that this speed is slower the nearer the poles of f lie to the boundary ∂D.

THEOREM 3.5.— *Let the function f be hyperelliptic and analytic at infinity, let $\pi_1, \ldots, \pi_k \in D$ be the poles of f in the convergence domain D, taking account of the order of the poles by repetition of points. Let further $\pi_{1n}, \ldots, \pi_{k_n n} \in \overline{\mathbb{C}}$ be poles of the approximant $[n/n]$ that converge to the poles of the function f in D in accordance with Theorem 3.1 (i). Then a selection of k poles π_{jn}, $j = 1, \ldots, k$, and their pairing with the poles $\pi_1, \ldots, \pi_k$ can be done in such a way that*

$$|\pi_{jn} - \pi_j| = O\left(\left(G_D(\pi_j) + \varepsilon\right)^{2n/\mathrm{st}(\,\mathrm{ord}(\pi_j) + k_{0j}\mathrm{st})}\right) \quad as\ n \to \infty\,,\ j = 1, \ldots, k, \tag{3.10}$$

where $\mathrm{ord}(\pi_j)$ *denotes the order of the pole of f at π_j, $\varepsilon > 0$ is arbitrary, and k_{0j} is the maximal number of spurious poles clustering at π_j. (From Theorem 3.1 (ii), we know that k_{0j} is bounded by the number (3.3)).*

Remark.— An appropriate selection of the poles $\pi_{1n}, \ldots, \pi_{kn}$ is necessary since near a pole π_j of the function f there may be a certain number of spurious poles of $[n/n]$ that may converge more slowly or even not at all to π_j.

With the same methods used in the proof of Theorem 3.5, it is possible to prove that the distance between spurious poles and the corresponding zeros of the approximant $[n/n]$ tends to zero as $n - \infty$ with a speed that also depends on the location of the limit point of the spurious poles in D.

THEOREM 3.6.— *Let the function f be hyperelliptic and analytic at infinity, and assume that there exists an infinite subsequence $N \subseteq \mathbb{N}$ such that the Padé approximant $[n/n]$, $n \in N$, has k spurious poles at $\pi_{1n}, \ldots, \pi_{kn} \in \mathbb{C}$ with repetion taking care of orders of poles, and assume that $\pi_{jn} - \pi_j \in D$ as $n - \infty$, $n \in N$, for $j = 1, \ldots, k$. Then to each spurious pole π_{jn} corresponds a zero ζ_{jn} of the approximant $[n/n]$, $j = 1, \ldots, k$, $n \in N$, such that*

$$|\pi_{jn} - \zeta_{jn}| = O\left(\left(G_D(\pi_j) + \varepsilon\right)^{2n/\mathrm{ord}_1(\pi_j)}\right) \quad as\ n \to \infty\,,\ n \in N\,,\ j = 1, \ldots, k, \tag{3.11}$$

where $\varepsilon > 0$ *is arbitrary and* $\mathrm{ord}_1(\pi_j)$ *denotes the number of poles of* $[n/n]$, $n \in N$, *spurious or non-spurious, converging to* π_j. *If* $\pi_j = \infty$, *then in (3.11) the Euclidean distance has to be replaced by the spherical metric.*

Next, we address the convergence problem. From Theorem 2.2 we know that the Padé approximants $[n/n]$, $n \in \mathbb{N}$, converge in capacity to f in the domain $D \subseteq \overline{\mathbb{C}}$. There are, in principle, two ways for obtaining locally uniform convergence in subdomains of D: either one removes spurious poles from the approximants $[n/n]$, $n \in \mathbb{N}$, by one or the other method, or one selects an infinite subsequence from $\{[n/n]\}_{n\in\mathbb{N}}$ that has no spurious pole on the subdomain of D, on which one wants to have locally uniform convergence. Of course, with the second strategy, the non-trivial question of whether or not such a subsequence exists, immediate arises. The assertion that such subsequences exist is known as the Baker–Gammel–Wills conjecture (more about this in Theorem 3.8).

For pole-removal we consider two possible procedures. The first one generates locally uniform convergence only in subdomains of D, where the function f is analytic. The elimination of spurious poles of $[n/n]$ near a pole of f is more complicated, since there, the two types of poles, spurious poles and poles that approximate a pole of f, are intermingled.

DEFINITIONS 3.2 (pole-clearing)

(i) *Let* r *be a rational function,* $D_0 \subseteq \overline{\mathbb{C}}$ *a subdomain of* $\overline{\mathbb{C}}$ *with a smooth boundary* ∂D_0 *such that* r *is analytic on* ∂D_0, $\infty \notin \partial D_0$, *and* $\partial D_0 \neq \emptyset$. *Then*

$$\hat{r}(z) := \frac{1}{2\pi i}\oint_{\partial D_0} \frac{r(\zeta)\,\mathrm{d}\zeta}{\zeta - z}, \quad z \in D_0, \tag{3.12}$$

is again a rational function. Its degree is not larger than that of r. *In (3.12) the integration path* ∂D_0 *has to be positively oriented if* $\infty \notin D_0$ *and negatively otherwise. The rational function* $\hat{r}$ *has no poles on* D_0, *and it is called pole-cleared on* D_0.

(ii) *Let* r *be a rational function, and let* $(\pi_j, \zeta_j) \in \overline{\mathbb{C}}^2$, $j = 1, \ldots, k$, *be* k *pairs of poles and zeros of* r. *Then*

$$\tilde{r}(z) := r(z)\prod_{j=1}^{k} \frac{z - \pi_j}{z - \zeta_j} \tag{3.13}$$

is called pole-cleared by factoring out pairs of poles and zeros.

Remarks

(1) Definition (3.12) is equivalent to the following procedure: represent r by partial fractions and drop all fractions with poles in D_0.

(2) For approximation purposes pole elimination by factoring out pairs of poles and zeros is helpful only if the distance between poles and zeros in each pair tends to zero with $n \to \infty$. Because of Theorem 3.6 this holds true in the case of spurious poles of Padé approximants $[n/n]$, $n \in \mathbb{N}$.

Let ϱ denote the spherical metric defined as

$$\varrho(z,w) := \begin{cases} \dfrac{2|z-w|}{\sqrt{(1+|z|^2)(1+|w|^2)}} & \text{if } z, w \neq \infty \\ \dfrac{2}{\sqrt{1+|z|^2}} & \text{if } w = \infty. \end{cases} \tag{3.14}$$

THEOREM 3.7.— *Let the function f be hyperelliptic and analytic at infinity.*

(i) Let D_0 be a subdomain of the convergence domain D with $\overline{D_0} \subseteq D$, assume that the function f is analytic on $\overline{D_0}$, and let the rational functions $\widehat{[n/n]}$, $n \in \mathbb{N}$, result from pole-clearing the Padé approximants $[n/n]$ on D_0. Then the sequence of pole-cleared approximants $\widehat{[n/n]}$ converge to f locally uniformly in D_0 as $n \to \infty$.

(ii) Let D_1 be an arbitrary subdomain of the convergence domain D with $\overline{D_1} \subseteq D$. Let $\pi_{1n}, \ldots, \pi_{k_n n}$ be the spurious poles of the Padé approximant $[n/n]$, $n \in \mathbb{N}$, on $\overline{D_1}$. In accordance with Theorem 3.1 (i) and Theorem 3.6 there exist corresponding zeros $\zeta_{1n}, \ldots, \zeta_{k_n n} \in \overline{\mathbb{C}}$ of $[n/n]$. Let $\widehat{[n/n]}$, $n \in \mathbb{N}$, be the rational functions resulting from factoring out the k_n pairs (π_{jn}, ζ_{jn}) of poles and zeros from the Padé approximants $[n/n]$ as done in (3.13). Then the sequence of approximants $\widehat{[n/n]}$ converges to f in the domain D_1 locally uniformly in the spherical metric as $n \to \infty$.

Remark.— In both parts of Theorem 3.7 nothing has been said about the speed of the convergence. It can be shown that the speed is geometric, and that the convergence factor is given by

$$q_j := \max_{z \in \overline{D_j}} G_D(z) + \varepsilon, \quad j = 0, 1, \tag{3.15}$$

with $\varepsilon > 0$ arbitrary, the maximum is taken over $\overline{D_0}$ in part (i), and over $\overline{D_1}$ in part (ii), and the function G_D has been defined in (2.3).

The last topic in the present section is concerned with the Baker–Gammel–Wills conjecture, which deals with the existence of infinite subsequences of $\{[n/n]\}$ that converge locally uniformly. The original conjecture in [BGW] has been formulated for a function f meromorphic in the unit disc $\mathbf{D}$ and for diagonal Padé approximants developed at the origin. We give here a formulation that is adapted to Padé approximants $[n/n]$ developed at infinity.

BAKER–GAMMEL–WILLS CONJECTURE

If the function f is meromorphic outside the closed disc $\overline{\mathbf{D}(r,0)}$ with radius $r > 0$ and center at the origin, then there exists an infinite subsequence $N \subseteq \mathbb{N}$ such that

$$[n/n] \to f \quad \text{as} \quad n \to \infty\,,\ n \in N, \tag{3.16}$$

locally uniformly in $\overline{\mathbb{C}} \setminus \overline{\mathbf{D}(r,0)}$, omitting poles of f.

We show that the conjecture holds true for hyperelliptic functions if some additional conditions are satisfied. Of course, discs are not typical for the problem. Therefore we first prove a more general result and then deduce the conjecture as a corollary.

THEOREM 3.8.— *Let the function f be hyperelliptic and analytic at infinity, and let the assumptions of Corollary 3.2 be satisfied. For almost all $\zeta_0 \in \mathbb{C}$ in a neighborhood of infinity as point of development the sequence of Padé approximants $[n/n]$, $n \in \mathbb{N}$, has the following property: there exists an infinite subsequence $N \subseteq \mathbb{N}$ such that*

$$\lim_{n\to\infty, n\in N} \varrho\big(f(z), [n/n](z)\big)^{1/(2n)} = G_D(z) \tag{3.17}$$

holds locally uniformly for $z \in D$, where D is the convergence domain (from Theorem 2.2) and ϱ is the spherical metric as defined in (3.14).

Remarks

(1) Since $G_D(z) < 1$ for $z \in D$, it follows from (3.17) that the subsequence $\{[n/n]\}_{n\in N}$ of diagonal Padé approximants converges to f locally uniformly in $D \setminus \{$ poles of $f\}$. The convergence holds true in the whole domain D in the spherical metric, and G_D gives the exact convergence factor.

(2) From (3.17) it further follows that there exists an infinite subsequence $N \subseteq \mathbb{N}$ such that the diagonal Padé approximants $[n/n]$, $n \in N$, have no spurious poles.

It has been shown in [Pol] that the projection ψ along radii onto a closed disc does not increase the capacity of a compact set $K \subseteq \mathbb{C}$, *i.e.*, $\operatorname{cap}(K) \geq \operatorname{cap}(\psi(K))$. From Theorem 2.2 and 2.3 it therefore follows that if the branch points $a_1, \dots, a_{2m}$ of the hyperelliptical function f are contained in the disc $\overline{\mathbf{D}(r,0)}$, then the whole set $F = \overline{\mathbb{C}} \setminus D \subseteq \overline{\mathbf{D}(r,0)}$. Hence, from Theorem 3.8 we deduce

COROLLARY 3.9.— *If the function f is hyperelliptic and analytic at infinity, and if it satisfies the assumptions of Corollary 3.2, then for almost every $\zeta_0 \in \overline{\mathbb{C}}$ near infinity as a point of development, the Baker–Gammel–Wills conjecture holds true for the sequence of diagonal Padé approximants $[n/n]$, $n \in \mathbb{N}$.*

At the end of Section 5 it will be discussed, whether the conditions and restrictions in Theorem 3.8 and Corollary 3.9 are really necessary.

4. Proofs of Results from Section 3, Part I

In the present section all results of Section 3 are proved, except the two Theorems 3.3 and 3.8. We start by introducing some terminology, and then study limit relations between zeros of the Padé polynomials Q_n and P_n and zeros and poles of the remainder function R_n. All three objects, the polynomials Q_n, P_n, and the remainder R_n, will be introduced in (4.8), below. After some preparation, Theorem 3.1 is proved, followed by the Theorems 3.5, 3.6 and 3.7. The section is closed by the proof of Lemma 3.4.

As before we denote by $\mathcal{R}$ the Riemann surface defined by the square root

$$y = y(z) = \sqrt{(z-a_1)\cdots(z-a_{2m})}, \quad a_i \neq a_j \text{ if } i \neq j. \tag{4.1}$$

The canonical mapping is denoted by $\pi : \mathcal{R} \longrightarrow \overline{\mathbb{C}}$, and as in (3.2) the surface $\mathcal{R}$ is broken down into the three sets B_1, B_2, and Γ such that $\mathcal{R} = B_1 \cup \Gamma \cup B_2$. The two domains B_1 and B_2 lie over the convergence domain D and we have $\pi(\Gamma) = F$. In order to have a complete decomposition of $\mathcal{R}$ in two sheets, we assume that the chain of curves Γ is broken down into two sets Γ_1 and Γ_2 such that $\Gamma = \Gamma_1 \cup \Gamma_2$ and over each point of $F \setminus \{a_1, \ldots, a_{2m}\}$ lies exactly one point of each set Γ_1 and Γ_2. On this basis we can define the two branches $\pi_j^{-1} : \overline{\mathbb{C}} \to B_j \cup \Gamma_j$, $j = 1, 2$, of the inverse π^{-1} of the canonical projection π. As a general rule we denote by z a point on the Riemann surface $\mathcal{R}$, and for $z \in \mathcal{R} \setminus \{a_1, \ldots, a_{2m}\}$ we consider z also as a local coordinate endowed with the Euclidean or, if necessary, with the spherical metric (3.14). In many situations it is useful to write $z^{(j)}$ for z if $z \in B_j \cup \Gamma_j$, $j = 1, 2$. Thus, $\infty^{(1)}$ and $\infty^{(2)}$ denotes infinity on the first and second sheet, respectively. As a rule, points on $\overline{\mathbb{C}}$ will be denoted by ζ. Further, we introduce the covering transformation $\varphi : \mathcal{R} \to \mathcal{R}$ as the map that satisfies $\pi \circ \varphi = \pi$ and $\varphi \neq \mathrm{id}_{\mathcal{R}}$. Thus, we have $\varphi(B_1) = B_2$, $\varphi(B_2) = B_1$, $\varphi(\Gamma) = \Gamma$, and the branch points $a_1, \ldots, a_{2m}$ are fixed points of φ.

From definition (2.3) of the function G_D, the fact that $F = \partial G$ consists of an union of finitely many analytic arcs, and from the symmetry property (2.8) in Theorem 2.4, it follows that there exists a function Φ locally analytic on $\mathcal{R} \setminus \{\infty^{(2)}\}$ and satisfying

$$\left|\Phi(z)\right| = \begin{cases} G_D\left(\pi(z)\right) & \text{for } z \in \overline{B_1} \\ 1/G_D\left(\pi(z)\right) & \text{for } z \in \overline{B_2}. \end{cases} \tag{4.2}$$

Indeed, $\log|\Phi(z)| = -g_D\left(\pi(z), \infty\right)$ for $z \in \overline{B_1}$ and $= g_D\bigl(\pi(z), \infty\bigr)$ for $z \in \overline{B_2}$. The symmetry property (2.8) guarantees that $\log\bigl|\Phi(z)\bigr|$ is harmonic throughout $\mathcal{R} \setminus \{\infty^{(1)}, \infty^{(2)}\}$.

The existence of the function Φ can also be proved directly without reference to the results in Section 2 by tools from the theory of compact Riemann surfaces (cf. [Sp, chap. 10.1]). Contrary to $|\Phi|$, the function Φ itself is in general not single-valued; it has a simple zero at $\infty^{(1)}$, a simple pole at $\infty^{(2)}$, and is different from zero elsewhere. These properties determine Φ up to a constant factor.

From (4.2) and (2.3) it follows that

$$\left|\Phi(z)\right| \begin{cases} < 1 & \text{for } z \in B_1, \\ = 1 & \text{for } z \in \Gamma, \\ > 1 & \text{for } z \in B_2. \end{cases} \tag{4.3}$$

For any $\varepsilon > 0$ we define an open neighborhood U of Γ by

$$U = U_\varepsilon := \left\{ z \in \mathcal{R} \mid \frac{1}{1+\varepsilon} < \left|\Phi(z)\right| < 1 + \varepsilon \right\}. \tag{4.4}$$

Let C_1 and C_2 denote the two chains of curves $\partial U \cap B_j$, $j = 1, 2$, and define $\tilde{U} := \pi(U)$, $\tilde{C} := \pi(C_1) = \pi(C_2)$.

Throughout the next two sections f is a hyperelliptic function defined on the Riemann surface $\mathcal{R}$, and it is analytic at $\infty^{(1)}$. Representation (3.1) of the function f lifted to $\mathcal{R}$ can be written as

$$f(z) = r_1\big(\pi(z)\big) + r_2\big(\pi(z)\big)y(z), \tag{4.5}$$

where r_1 and r_2 are rational functions defined on $\overline{\mathbb{C}}$, and the square root (4.1) is considered to be defined on $\mathcal{R}$ with

$$y = y(z) = z^m + O(z^{m-1}) \quad \text{as } z \to \infty^{(1)}. \tag{4.6}$$

We have $y \circ \varphi = -y$.

The asymptotic behavior ($n \to \infty$) of zeros and poles of certain sequences of functions is of central interest. For the description of poles and zeros we use *multisets*, *i.e.*, sets in which the same element can be repeated several times. By $Z(g)$ and $P(g)$ we denote the multiset of zeros and poles, respectively, of a meromorphic function g in its natural domain of definition. Thus, for instance, $P(f)$ is the set of all poles of the function f on $\mathcal{R}$ taking account of multiplicities by repetition of elements. By $S|_{D_0}$ we denote the restriction of a multiset S to elements that are contained in D_0, and by $\#(S)$ we denote the number of elements in S. All other set-theoretic symbols are used in the usual way. For sequences of finite sets a notion of convergence is defined in the following way: we write $S_n \to S$ as $n \to \infty$, if $\#(S_n) = \#(S)$ for $n \geq n_0$ and there exist bijections $\psi_n : S \to S_n$ such that

$$\sum_{x \in S} \operatorname{dist}\big(x, \psi_n(x)\big) \to 0 \quad \text{as } n \to \infty. \tag{4.7}$$

The distance in (4.7) is either the Euclidean or the spherical distance on $\overline{\mathbb{C}}$ or on the Riemann surface $\mathcal{R}$.

The defining relation (1.2) of Padé approximants lifted to $\mathcal{R}$ with the right-hand side denoted by R_n and the whole relation multiplied by z^n yields

$$\begin{aligned} &R_n(z) := f(z)Q_n\big(\pi(z)\big) - P_n\big(\pi(z)\big) = O(z^{-n-1}) \quad \text{as } z \to \infty^{(1)}, \\ &P_n(\zeta) := \zeta^n p_{nn}\left(\frac{1}{\zeta}\right), \quad Q_n(\zeta) := \zeta^n q_{nn}\left(\frac{1}{\zeta}\right), \end{aligned} \tag{4.8}$$

where R_n and f are functions on the Riemann surface $\mathcal{R}$ and the polynomials Q_n and P_n are functions on $\overline{\mathbb{C}}$, R_n is called the *remainder function*, and P_n and Q_n are the*Padé polynomials* with $[n/n] = P_n/Q_n$. The remainder function R_n has a zero of order at least $n+1$ at $\infty^{(1)}$ and at a pole of order not larger than n plus the order of a possible pole of f at $\infty^{(2)}$. From (4.8) it follows that

$$P(R_n)|_{\mathcal{R}\setminus\{\infty^{(2)}\}} \subseteq P(f). \tag{4.9}$$

Because of the special role of the two points $\infty^{(1)}$ and $\infty^{(2)}$ modified zero and pole sets are introduced for R_n by

$$\begin{aligned} \hat{Z}(R_n) &:= Z(R_n)|_{\mathcal{R}\setminus\{\infty^{(1)},\infty^{(2)}\}} \cup Z(z^{n+1}R_n)|_{\{\infty^{(1)}\}} \cup Z(z^{-n}R_n)|_{\{\infty^{(2)}\}}, \\ \hat{P}(R_n) &:= P(R_n)|_{\mathcal{R}\setminus\{\infty^{(2)}\}} \cup P(z^{-n}R_n)|_{\{\infty^{(2)}\}}. \end{aligned} \tag{4.10}$$

Thus, for instance, $\hat{Z}(R_n)$ contains $\infty^{(1)}$ only if R_n has a zero of order larger than $n+1$ at $\infty^{(1)}$.

Since the function R_n is meromorphic on the compact Riemann surface $\mathcal{R}$, the number of its poles and zeros on $\mathcal{R}$ are identical, and from (4.9) and (4.10) it follows that

$$\#(\hat{P}(R_n)) \leq \#(P(f)) \quad \text{and} \quad \#(\hat{Z}(R_n)) \leq \#(P(f)) - 1. \tag{4.11}$$

Thus, we know that both sequences of multisets $\{\hat{P}(R_n)\}_{n\in\mathbb{N}}$ and $\{\hat{Z}(R_n)\}_{n\in\mathbb{N}}$ contain only a bounded number of elements. Consequently, any infinite subsequence $N \subseteq \mathbb{N}$ contains an infinite subsequence, which again is denoted by N, such that there exist two multisets Z_R and P_R with elements from $\mathcal{R}$ and

$$\hat{Z}(R_n) \to Z_R, \quad \hat{P}(R_n) \to P_R \quad \text{as } n \to \infty,\ n \in N. \tag{4.12}$$

As distance function in the convergence (4.12) we use the spherical metric lifted to $\mathcal{R}$. From (4.11) and the definitions in (4.10) it follows that

$$\#(P_R) \leq \#(P(f)), \quad \#(Z_R) \leq \#(P(f)) - 1 \quad \text{and} \quad \#(P_R) = \#(Z_R) + 1. \tag{4.13}$$

The defining relation (1.2) for Padé approximants can be multiplied by any non-zero constant, and therefore the same is true for the remainder function R_n defined in (4.8). An appropriate normalization of R_n is essential for the result in the next lemma. Let the fixed point $z_0 \in$

$\Gamma \setminus (Z_R \cup P_R \cup \bigcup_{n\in N}(\hat{Z}(R_n) \cup \hat{P}(R_n)))$ be chosen. For the purpose of normalization we assume that

$$\left|R_n(z_0)\right| = 1 \quad \text{for } n \in N, \tag{4.14}$$

and in order to also fix the argument of the values of R_n, we further assume that in the development

$$\left(\frac{R_n}{r_2}\right)(z) = Az^k + O(z^{k-1}) \quad \text{as } z \to \infty^{(2)}, \tag{4.15}$$

the leading coefficient satisfies $A > 0$. We note that if in relation (4.8), and simultaneously in relation (1.2) for a given n among all admissible polynomials $Q_n \not\equiv 0$, a polynomial of minimal degree is chosen, then the condition (4.14) together with the condition $A > 0$ in (4.15) determines R_n uniquely. In the next lemma, however, such a perfect normalization is not necessary. It is enough that condition (4.14) is satisfied.

LEMMA 4.1.— *We have*

$$\lim_{n\to\infty, n\in N} \left|R_n(z)\right|^{1/n} = \left|\Phi(z)\right| \tag{4.16}$$

locally uniformly for $z \in \mathcal{R} \setminus \left(Z_R \cup P_R \cup \left\{\infty^{(1)}, \infty^{(2)}\right\}\right)$.

Proof.— For $z_1, z_2 \in \mathcal{R} \setminus \{z_0\}$, $z_1 \neq z_2$, we define the function $h(z_0, z_1, z_2; \cdot)$ by the following properties: The function is harmonic in $\mathcal{R} \setminus \{z_1, z_2\}$, $h(z_0, z_1, z_2; z_0) = 0$, and it has logarithmic poles at z_1 and z_2 with residues -1 and 1, respectively. The unique existence of such a function follows from [Sp, chap. 10.1]. It is not difficult to prove that if z_2 is kept fixed and z_1 tends to z_2, then

$$\lim_{z_1 \to z_2} h(z_0, z_1, z_2; z) = 0 \tag{4.17}$$

locally uniformly for $z \in \mathcal{R} \setminus \{z_2\}$. Let the functions h_n be defined as $h_n(z) := \log|R_n(z)| - n \log|\Phi(z)|$, and let h be the function that is harmonic in $\mathcal{R} \setminus (Z_R \cup P_R \cup \{\infty^{(1)}\})$, $h(z_0) = 0$, and h has logarithmic poles with residue -1 at the elements of $Z_R \cup \{\infty^{(1)}\}$ and with residue 1 at the elements of P_R, taking account of multiplicities of elements by adding up the residues. The existence of the functions h follows from the existence of the function $h(z_0, z_1, z_2; \cdot)$.

The convergence (4.12) implies that there exist bijections $\psi_n : Z_R \to \hat{Z}(R_n)$ and $\varphi_n : P_R \to \hat{P}(R_n)$, $n \in N$, such that

$$(h_n - h)(z) = \sum_{z_1 \in Z_R} h\big(z_0, \psi_n(z_1), z_1; z\big) - \sum_{z_2 \in P_R} h\big(z_0, \varphi_n(z_2), z_2; z\big) \quad (4.18)$$

and from the convergence in (4.12) together with (4.17) it then follows that $h_n \to h$ as $n \to \infty$, $n \in N$, locally uniformly in $\mathcal{R} \setminus \left(Z_R \cup P_R \cup \{\infty^{(1)}\}\right)$. The last limit together with the definition of h_n implies that

$$\lim_{n \to \infty, n \in N} \left| \frac{R_n(z)}{\Phi(z)^n} \right| = \exp h(z) \quad (4.19)$$

locally uniformly for $z \in \mathcal{R} \setminus \left(Z_R \cup P_R \cup \{\infty^{(1)}\}\right)$. Assertion (4.16) follows from (4.19). □

We note that in (4.19) we have proved a stronger asymptotic relation than that in (4.16). In (4.19) we have proved what is called power-asymptotics, while in (4.16) only n-th root asymptotics are stated. However, (4.16) is sufficient for the investigations in the present paper.

Using representation (4.5) and the equation $y \circ \varphi = -y$, we can eliminate the polynomial P_n from (4.8), which yields the representation

$$Q_n(\zeta) = \frac{R_n\left(\pi_1^{-1}(\zeta)\right) - R_n\left(\pi_2^{-1}(\zeta)\right)}{2 r_2(\zeta)\, y\left(\pi_1^{-1}(\zeta)\right)}, \quad \zeta \in \mathbb{C} \setminus F, \quad (4.20)$$

for the denominator polynomial Q_n. The parameter $\varepsilon > 0$ in definition (4.4) of the neighborhood $U = U_\varepsilon$ of Γ can be chosen so small that we have $S|_{\overline{U}} = S|_\Gamma$ for the multiset $S := Z_R \cup P_R \cup Z(f) \cup P(f) \cup \pi^{-1}\left(Z(r_2) \cup P(r_2)\right)$. Such a choice of $\varepsilon > 0$ is always possible since the set S is finite. The sets $\tilde{U}$, $\tilde{C}$, and C_j, $j = 1, 2$, are defined as above in connection with (4.4). It follows from (4.4), the limit (4.16) and the inequalities in (4.3) that for $n \in N$ sufficiently large the function $R_n \circ \pi_1^{-1}$ is uniformly smaller than $R_n \circ \pi_2^{-1}$ on $\tilde{C}$. From (4.20) and Rouché's Theorem we therefore deduce that the number of zeros of Q_n in $\tilde{U}$ depends only on the growth of the function $\arg(R_n \circ \pi_2^{-1} / r_2 y \circ \pi_1^{-1})$ on $\tilde{C} = \partial \tilde{U}$. This yields the formula

$$\begin{aligned} \#(Z(Q_n)|_{\tilde{U}}) &= n + 1 - m + \#(\hat{Z}(R_n)|_{B_1 \cup \Gamma}) - \#(\hat{P}(R_n)|_{B_1 \cup \Gamma}) \\ &\quad - \#(Z(r_2)|_F) + \#(P(r_2)|_F). \end{aligned} \quad (4.21)$$

We have used definition (4.10) and have taken into account that $\arg(y(z))$ grows by $2\pi m$ on the chain of curves C_2.

Considering relation (4.8) on B_1 and inserting representation (4.20) for Q_n yields a representation for the numerator polynomial P_n. We have

$$\begin{aligned} P_n(\zeta) &= f\left(\pi_1^{-1}(\zeta)\right) Q_n(\zeta) + R_n\left(\pi_1^{-1}(\zeta)\right) \\ &= f\left(\pi_1^{-1}(\zeta)\right) \frac{R_n\left(\pi_1^{-1}(\zeta)\right) - R_n\left(\pi_2^{-1}(\zeta)\right)}{2r_2(\zeta)\, y\left(\pi_1^{-1}(\zeta)\right)} + R_n\left(\pi_1^{-1}(\zeta)\right). \end{aligned} \tag{4.22}$$

With the same arguments as used for the derivation of (4.21) it follows from (4.22) that

$$\#(Z(P_n)|_{\tilde{U}}) = \#(Z(Q_n)|_{\tilde{U}}) + \#(Z(f)|_{B_2 \cup \Gamma}) - \#(P(f)|_{B_2 \cup \Gamma}). \tag{4.23}$$

As in the case of the remainder functions R_n in (4.10), so also for the denominator and numerator polynomials Q_n and P_n we introduce modified definitions of zero sets by

$$\begin{aligned} \hat{Z}(Q_n) &:= Z(Q_n) \cup Z(z^{-n} Q_n)|_{\{\infty\}} \\ \hat{Z}(P_n) &:= Z(P_n) \cup Z(z^{-n} P_n)|_{\{\infty\}}. \end{aligned} \tag{4.24}$$

The definition takes account of the possibility that the degrees of the polynomials P_n or Q_n may be less than n. With (4.24) we always have $\#(\hat{Z}(Q_n)) = \#(\hat{Z}(P_n)) = n$ for all $n \in \mathbb{N}$.

From (4.21), (4.11), and (4.23) we learn that only a finite number of zeros of the polynomials in the sequences $\{Q_n\}_{n\in N}$ and $\{P_n\}_{n\in N}$ can cluster outside of $F = \partial D$. Hence, we can assume that N contains an infinite subsequence of N, which we continue to denote by N, such that there exist finite multisets Z_Q and Z_P of points from the convergence domain $D = \overline{\mathbb{C}} \backslash F$ with

$$\hat{Z}(Q_n)|_{\overline{\mathbb{C}}\backslash \tilde{U}} \to Z_Q \quad \text{and} \quad Z(P_n)|_{\overline{\mathbb{C}}\backslash \tilde{U}} \to Z_P \quad \text{as } n \to \infty,\ n \in N. \tag{4.25}$$

If $\varepsilon > 0$ in the definition of $U = U_\varepsilon$ has been chosen sufficiently small, then it follows from (4.21), (4.11), and (4.23) that the sets Z_Q and Z_P have no elements in $\tilde{U}$. The elements of Z_Q and Z_P are the only cluster points of zeros of the polynomials Q_n and P_n outside of F as $n \to \infty$, $n \in N$.

From the equations (4.21), (4.11), (4.23), and the limits in (4.25) we deduce

LEMMA 4.2.— *We have*

$$\begin{aligned}\#(Z_Q) &= m-1-\#(Z_R|_{B_1\cup\Gamma})+\#(P_R|_{B_1\cup\Gamma})+\\&\quad+\#(Z(r_2)|_F)-\#(P(r_2)|_F),\\ \#(Z_P) &= \#(Z_Q)+\#(Z(f)|_{B_1})-\#(P(f)|_{B_1}).\end{aligned} \tag{4.26}$$

In the second equality of (4.26) we have used the equation $\#(Z(f)) = \#(P(f))$ and the decomposition (3.2).

The representations (4.20) and (4.22) together with the limit (4.16) and the inequalities in (4.3) yield

LEMMA 4.3.— *We have*

$$\left.\begin{aligned}&\lim_{n\to\infty, n\in N}\left|Q_n(z)\right|^{1/n}\\&\lim_{n\to\infty, n\in N}\left|P_n(z)\right|^{1/n}\end{aligned}\right\} = \left|\Phi\left(\pi_2^{-1}(z)\right)\right| \tag{4.27}$$

locally uniformly for $z\in\mathbb{C}\setminus(F\cup Z_Q)$ in the first limit, and locally uniformly for $z\in\mathbb{C}\setminus(F\cup Z_P)$ in the second one.

The limits in (4.16) and (4.27) together with relation (4.8), the inequalities in (4.3), and Rouché's Theorem show that in the neighborhood of each pole of the function $f\circ\pi_1^{-1}$ in the convergence domain $D=\overline{\mathbb{C}}\setminus F$, the denominator polynomial Q_n has zeros of at least the same order as the pole of $f\circ\pi_1^{-1}$. If the order of the zeros of Q_n is higher, then the numerator polynomial P_n has also to have one or several zeros in the same neighborhood with a total order that is equal to the difference. Analogous results hold true in neighborhoods of zeros of the function $f\circ\pi_1^{-1}$ in D with the role of the polynomials Q_n and P_n interchanged. In the next lemma these connections are expressed in terms of the limit sets Z_Q and Z_P.

LEMMA 4.4.— *We have*

$$\begin{aligned}\pi(P(f)|_{B_1}) &= Z_Q\setminus Z_P,\\ \pi(Z(f)|_{B_1}) &= Z_P\setminus Z_Q.\end{aligned} \tag{4.28}$$

Lemma 4.4 shows that zeros of the denominator polynomials Q_n, $n\in N$, cluster at each pole of $f\circ\pi_1^{-1}$ in the convergence domain D with at least the same total order as that of $f \circ \pi_1^{-1}$. It follows from Definition 2.1, the relations (4.28), and the limits (4.25) that for the elements of the set $Z_Q \cap Z_P$ there exist three possibilities:

(i) they are cluster points of spurious poles of the approximants $[n/n]$, $n \in N$,

(ii) they are cluster points of common zeros of the Padé polynomials Q_n and P_n, $n \in N$, that cancel out in $[n/n]$, or

(iii) they are equal to $\infty^{(1)}$ and both Padé polynomials Q_n and P_n, $n \in N$, have at the same time degrees smaller than n.

From these considerations it follows that

$$\#\{\text{ spurious poles of } [n/n], n \in N\} \leq \#(Z_Q \cap Z_P). \tag{4.29}$$

We consider the partitioning

$$Z_Q = (Z_Q \setminus Z_P) \cup (Z_Q \cap Z_P), \tag{4.30}$$

and derive from the first equations in (4.26) and (4.28) that

$$\begin{aligned}\#(Z_Q \cap Z_P) &= \#(Z_Q) - \#(Z_Q \setminus Z_P)\\ &= m - 1 - \#\big(Z_R|_{B_1 \cup \Gamma}\big) + \#\big(P_R|_{B_1 \cup \Gamma}\big) + \#\big(Z(r_2)|_F\big)\\ &\quad - \#\big(P(r_2)|_F\big) - \#\big(P(f)|_{B_1}\big)\,.\end{aligned} \tag{4.31}$$

With (4.9), (4.10), and (4.12) it then further follows that

$$\begin{aligned}\#(Z_Q \cap Z_P) \leq m - 1 &- \#\big(Z_R|_{B_1 \cup \Gamma}\big) + \#\big(P(f)|_\Gamma\big)\\ &- \#\big(P(r_2)|_F\big) + \#\big(Z(r_2)|_F\big),\end{aligned} \tag{4.32}$$

which yields

LEMMA 4.5.— *We have*

$$\#(Z_Q \cap Z_P) \leq m - 1 + \#\big(P(f)|_\Gamma\big) - \#\big(P(r_2)|_F\big) + \#\big(Z(r_2)|_F\big) \tag{4.33}$$

Note that the number $\#\big(P(f)|_\Gamma\big) - \#\big(P(r_2)|_F\big)$ is always non-negative.

We are now prepared to start with the proofs of the Theorems 3.1, 3.5, 3.6 and 3.7.

Proof of Theorem 3.1

(i) First we consider only Padé approximants $[n/n]$ with $n \in N$, where $N \subseteq \mathbb{N}$ is an infinite subsequence as used in the Lemmas 4.1 through 4.5. Let $z_0 \in D$ be a pole of order k of the function $f \circ \pi_1^{-1}$. Since f is analytic at $\infty^{(1)}$ we have $z_0 \neq \infty$. From the first equation in (4.28) of Lemma 4.4 we know that near z_0 the denominator polynomials Q_n, $n \in N$, have exactly k zeros more than the numerator polynomials P_n, and all these zeros converge to z_0 as $n \to \infty$, $n \in N$. Consequently, the approximants $[n/n]$, $n \in N$, have poles of a total order at least k, which are converging to z_0 as $n \to \infty$, $n \in N$.

In the case that the total order of the poles of $[n/n]$ converging to z_0 is larger than k, it remains to show that $[n/n]$ possesses the zeros that are typically associated with spurious poles. If there are more than k poles (in the sense of total order), then the same holds true for zeros of the polynomials Q_n, $n \in N$. From (4.25) and the first equation in (4.28) of Lemma 4.4 it follows that the additional zeros of Q_n converge to points in $Z_Q \cap Z_P$. It then follows again from (4.25) that the numerator polynomials P_n, $n \in N$, have the same number of zeros converging to z_0. If some zeros of the polynomials P_n and Q_n are identical, then they cancel out in $[n/n]$, and the number of (spurious) poles and zeros of $[n/n]$ is reduced by the same number.

All conclusions are valid so far only for the subsequence $N \subseteq \mathbb{N}$ used in the Lemmas 4.1 through 4.5, but the construction of this subsequence guarantees that any potential, infinite, exceptional subsequence of $\mathbb{N}$ contains an infinite subsequence for which the conclusions hold true. Hence, all conclusions of part (i) are proved for the full sequence $\mathbb{N}$.

(ii) The part (ii) of the theorem follows directly from the inequalities (4.29) and (4.33). $\square$

Proof of Theorem 3.5 and 3.6 .— It turns out that it is best to prove both theorems together. Let $x \in D$ and let $f \circ \pi_1^{-1}$ have a pole of order $k_1 \geq 0$ at x. The case $k_1 = 0$ is not excluded, *i.e.*, $f \circ \pi_1^{-1}$ may be analytic at x. We shall show that the estimates (3.10) and (3.11) hold true for all poles of the approximants $[n/n]$, $n \in \mathbb{N}$, converging to x.

By $v_0 = v(f \circ \pi_1^{-1}, x)$ we denote the valuation of $f \circ \pi_1^{-1}$ at the point x, *i.e.*, v_0 is equal to the order of a zero of $f \circ \pi_1^{-1}$ at x or equal to the negative of the order of a pole of $f \circ \pi_1^{-1}$ at x. We have $v_0 \geq -k_1$. Set

$k_2 := v_0 + k_1$. Then we have $\min(k_0, k_1) = 0$. Let the infinite subsequence $N \subseteq \mathbb{N}$ be selected as in the Lemmas 4.1 through 4.5, and assume that also the convergence in (4.12) and (4.25) holds true. From Theorem 3.1 we know that for each $n \in N$ the denominator polynomial Q_n, $n \in N$, has zeros of total order $k_3 \geq k_1$ in a neighborhood of x, and the zeros converge to x as $n \to \infty$, $n \in N$. We can assume that the total order is constant for all $n \in N$ by choosing a subsequence. Set $k_4 := k_3 + v_0$ and denote the k_3 zeros of Q_n in a neighborhood of x by π_{jn}, $j = 1, \ldots, k_3$. According to Theorem 3.1 (i) the numerator polynomial P_n, $n \in N$, has $k_3 - k_1$ zeros in a neighborhood of x. If the function $f \circ \pi_1^{-1}$ has no pole, but a zero of order $k_2 = v_0$ at x, then P_n has in addition k_2 zeros in a neighborhood of x. The $k_3 - k_1 + k_2 = k_4$ zeros of P_n near x are denoted by ζ_{jn}, $j = 1, \ldots, k_4$. We define

$$q_n(\zeta) := (\zeta - x)^{k_2} \prod_{j=1}^{k_3} (\zeta - \pi_{jn}), \quad p_n(\zeta) := (\zeta - x)^{k_1} \prod_{j=1}^{k_4} (\zeta - \zeta_{jn}), \; n \in N, \tag{4.34}$$

and

$$\begin{aligned} \tilde{Q}_n &:= (\cdot - x)^{k_2} Q_n / q_n, \quad \tilde{P}_n := (\cdot - x)^{k_1} P_n / p_n, \\ \tilde{f} &:= (\cdot - x)^{k_1 - k_2} f \circ \pi_1^{-1}, \quad \tilde{R}_n := (\cdot - x)^{k_1} R_n \circ \pi_1^{-1}. \end{aligned} \tag{4.35}$$

In contrast with the functions f and R_n, the new functions $\tilde{f}$ and $\tilde{R}_n$ are defined on $\overline{\mathbb{C}}$. For $\varepsilon > 0$ sufficiently small the function $\tilde{f}$ is analytic and different from zero, the function $\tilde{R}_n$ analytic, and the polynomials $\tilde{P}_n$ and $\tilde{Q}_n$, $n \in N$, are different from zero on $\overline{\mathbf{D}(x, \varepsilon)}$ if N starts with n sufficiently large. Indeed, this is the case if the punctured disc $\overline{\mathbf{D}(x, \varepsilon)} \setminus \{x\}$ does not contain any element of Z_Q or Z_P and if the lifted set $\pi_1^{-1}\left(\overline{\mathbf{D}(x, \varepsilon)} \setminus \{x\}\right) \subseteq \mathcal{R}$ does not contain any element of $Z(f)$, $P(f)$, Z_R, or P_R.

By C we denote the circle $\partial \mathbf{D}(x, \varepsilon)$. Since all polynomials p_n and q_n, $n \in N$, are of identical degree $k_2 + k_3 = k_1 + k_4$, and since all zeros of q_n and p_n converge to x as $n \to \infty$, $n \in N$, we have

$$\lim_{n \to \infty, n \in N} |q_n(\zeta)|^{1/n} = 1, \qquad \lim_{n \to \infty, n \in N} |p_n(\zeta)|^{1/n} = 1 \tag{4.36}$$

uniformly for $\zeta \in C$. From (4.27) in Lemma 4.3 together with (4.35) and (4.36) we deduce that

$$\left. \begin{aligned} &\lim_{n \to \infty, n \in N} |\tilde{Q}_n(\zeta)|^{1/n} \\ &\lim_{n \to \infty, n \in N} |\tilde{P}_n(\zeta)|^{1/n} \end{aligned} \right\} = \left| \Phi(\pi_2^{-1}(\zeta)) \right| \tag{4.37}$$

uniformly for $\zeta \in C$. From (4.16) of Lemma 4.1 together with (4.35) we deduce that

$$\lim_{n\to\infty, n\in N} \left|\tilde{R}_n(\zeta)\right|^{1/n} = \left|\Phi\left(\pi_1^{-1}(\zeta)\right)\right| \tag{4.38}$$

uniformly for $\zeta \in C$. Inserting (4.35) into equation (4.8) yields

$$\frac{\tilde{Q}_n}{\tilde{P}_n}\tilde{f} - \frac{p_n}{q_n} = \frac{\tilde{R}_n}{q_n \tilde{P}_n}. \tag{4.39}$$

From the limits (4.36), (4.37), and (4.38) the asymptotic relation

$$\lim_{\substack{n\to\infty \\ n\in N}} \left|\frac{\tilde{R}_n}{q_n \tilde{P}_n}(\zeta)\right|^{1/n} = \frac{\left|\Phi\left(\pi_1^{-1}(\zeta)\right)\right|}{\left|\Phi\left(\pi_2^{-1}(\zeta)\right)\right|} = \left|\Phi\left(\pi_1^{-1}(\zeta)\right)\right|^2 \tag{4.40}$$

follows uniformly for $\zeta \in C$. In (4.40) equality (4.2) has been used. Since p_n and q_n are monic polynomials of identical degree, we have $(p_n/q_n)(\zeta) = 1 + O(\zeta^{-1})$ as $\zeta \to \infty$. Further, we know that all zeros of q_n lie inside of C and $\tilde{P}_n$ is different from zero on $\overline{\mathbf{D}(x,\varepsilon)}$ for $n \in N$ sufficiently large. Consequently we have

$$\frac{1}{2\pi i}\oint_C \left[\frac{\tilde{Q}_n}{\tilde{P}_n}\tilde{f} - \frac{p_n}{q_n}\right](t)\frac{dt}{t-\zeta} = 1 - \frac{p_n}{q_n}(\zeta) \quad \text{for } \zeta \in \overline{\mathbb{C}} \setminus \overline{\mathbf{D}(x,\varepsilon)}, \tag{4.41}$$

and with (4.39) and (4.40) we deduce that

$$\overline{\lim_{\substack{n\to\infty \\ n\in N}}} \left|1 - \frac{p_n}{q_n}(\zeta)\right|^{1/(2n)} \leq \sup_{\zeta\in C} \left|\Phi\left(\pi_1^{-1}(\zeta)\right)\right| \tag{4.42}$$

locally uniformly for $\zeta \in \overline{\mathbb{C}} \setminus \overline{\mathbf{D}(x,\varepsilon)}$.

Since the function $|\Phi|$ is continuous in $\mathcal{R} \setminus \{\infty^{(2)}\}$, for any $\delta > 0$ there exists $\varepsilon > 0$ such that the right-hand side of (4.42) is strictly smaller than $\left|\Phi\left(\pi_1^{-1}(x)\right)\right| + \delta$. With the first limit in (4.36) it therefore follows from (4.42) that

$$\begin{gathered} |q_n(\zeta) - p_n(\zeta)| \leq (c_0 + \delta)^{2n}, \quad c_0 := \left|\Phi\left(\pi_1^{-1}(x)\right)\right|, \\ \text{for } \zeta \in \overline{\mathbf{D}(x,\varepsilon)} \text{ and } n \in N. \end{gathered} \tag{4.43}$$

The sequence N has to start with n sufficiently large. The estimate (4.43) will be the major tool for finding an appropriate pairing of the $k_2 + k_3 = k_1 + k_4$ zeros of the polynomials p_n and q_n. Set

$$E_n := \left\{\zeta \in \mathbb{C} \mid |q_n(\zeta)| \le 2(c_0+\delta)^{2n}\right\}, \quad n \in N. \tag{4.44}$$

The lemmiscate E_n consists of the components $E_{1n}, \dots E_{l_n n}$. We show that in each component E_{ln} the two polynomials p_n and q_n have an identical number of zeros. Indeed, let us complete the two lists of zeros by setting $\zeta_{jn} := x$ for $j = k_3+1, \dots, k_3+k_2$ and the list of poles by setting $\pi_{jn} := x$ for $j = k_4+1, \dots, k_4+k_1$. All zeros ζ_{jn} and π_{jn} converge to x as $n \to \infty$, $n \in N$, $j = 1, \dots, k_2+k_3 = k_1+k_4$, and we have $E_n \subseteq \mathbf{D}(x, \varepsilon)$ for all $n \in N$ if N starts with n sufficiently large. Assume without loss of generality that $\pi_{1n}, \dots, \pi_{l_1 n n}$ are the only zeros of q_n in E_{1n}. From (4.43) and (4.44) we know that $|q_n(\zeta)| > |p_n(\zeta) - q_n(\zeta)|$ for all $\zeta \in \partial E_{1n}$. It therefore follows from Rouché's Theorem that p_n and q_n have the same number of zeros in E_{1n}. In the same way the conclusion follows for the other components of E_n.

Since q_n is a monic polynomial of degree $k_1 + k_4$, it follows from (4.44) that

$$\operatorname{cap}(E_{1n}) \le \operatorname{cap}(E_n) = 2^{1/(k_1+k_4)}(c_0+\delta)^{2n/(k_1+k_4)} \tag{4.45}$$

(cf. [La, chap. II, S4]). Further we know that for any continuum $K \subseteq \mathbb{C}$ we have $\operatorname{diam}(K) \le 4 \operatorname{cap}(K)$ (cf. [La, chap. II, S4]). Hence, for any pair of zeros π_{jn} and ζ_{jn}, $j = 1, \dots, k_1 + k_4$, in the same component E_{ln}, we have

$$|\pi_{jn} - \zeta_{jn}| \le (c_0+\delta)^{2n/(k_1+k_4)} \quad \text{for } j = 1, \dots, k_1+k_4\,,\ n \in N. \tag{4.46}$$

In the case $k_1 > 0$ let us consider the k_1 zeros $\zeta_{jn} = x$, $j = k_4+1, \dots, k_1+k_4$ of p_n. Using the definition of c_0 in (4.43) together with (4.2), it follows from (4.46) that there exist k_1 zeros of the denominator polynomial Q_n, which we number as π_{jn}, $j = 1, \dots, k_1$, and which are not at the same time zeros of the numerator polynomial P_n, and we have

$$|x - \pi_{jn}| \le (G_D(x)+\delta)^{2n/(k_1+k_4)} \quad \text{for } j = 1, \dots, k_1,\ n \in N. \tag{4.46}$$

For the remaining $k_4 = k_3 + v_0 = k_3 - k_1$ zeros π_{jn} of the polynomial Q_n near x there exist k_4 corresponding zeros of P_n in a neighborhood of x that can be numbered as ζ_{jn}, $j = k_1+1, \dots, k_1+k_4$, in such a way that

$$|\zeta_{jn} - \pi_{jn}| \le (G_D(x)+\delta)^{2n/(k_1+k_4)} \quad \text{for } j = k_1+1, \dots, k_4,\ n \in N. \tag{4.47}$$

From (4.46) the asymptotic estimate (3.10) in Theorem 3.5 follows for the pole of f at x and the subsequence N. From the construction of N in the preparation of the Lemmas 4.1 through 4.5 it is clear that any infinite sequence contains an infinite subsequence for which our conclusion holds true. Hence, exceptional subsequences are impossible and the conclusion is proved for the full sequence $\mathbb{N}$. Since x was an arbitrary pole of f in D, this concludes the proof of Theorem 3.5. Note that $\operatorname{ord}(\pi_j) + k_0 = k_1 + k_4 = k_2 + k_3$ if $x = \pi_j$.

From (4.47) we conclude in the same way that the asymptotic estimate (3.11) in Theorem 3.6 holds for all sequences of spurious poles that do not cluster at ∞. Note that $k_1 = 0$ has not been excluded, and this is the case if the function f is analytic at $x \in D \setminus \{\infty\}$. However, the case $x = \infty$ had been excluded in the analysis so far. This case can be treated in principle in the same way. One only has to use convergence in the spherical metric, and to do necessary adaptations in several formulas. □

Proof of Theorem 3.7.— From (1.1), (1.2), and (4.8) the formula

$$f \circ \pi_1^{-1} - [n/n] = \frac{R_n \circ \pi_1^{-1}}{Q_n} \tag{4.48}$$

for the approximation error follows. As already done in the proofs of the Theorems 3.1, 3.5 and 3.6, we first analyse the situation that the problem is restricted to an infinite subsequence $N \subseteq \mathbb{N}$, for which the Lemmas 4.1 through 4.5 hold true. Let V be an open set with $\overline{V} \subseteq D \setminus \{\infty\}$ and $\overline{V}$ does not contain any point of the sets Z_Q or $\pi\left(P(f)|_{B_1}\right)$. Since $\overline{V} \cap Z_Q = \emptyset$, the first limit in (4.27) of Lemma 4.3 holds uniformly on $\overline{V}$, and because of $\overline{V} \cap \pi(P_R|_{B_1}) = \emptyset$, it follows from the limit (4.16) in Lemma 4.1 together with (4.9) that

$$\varlimsup_{\substack{n\to\infty \\ n\in N}} \left| R_n(\pi_1^{-1}(\zeta)) \right|^{1/n} \leq \left| \Phi(\pi_1^{-1}(\zeta)) \right| \tag{4.49}$$

uniformly for $\zeta \in \overline{V}$. From (4.48) and the limits (4.27) and (4.49) it then follows that

$$\begin{aligned} \varlimsup_{\substack{n\to\infty \\ n\in N}} \left| (f \circ \pi_1^{-1} - [n/n])(\zeta) \right|^{1/(2n)} = \varlimsup_{\substack{n\to\infty \\ n\in N}} \left| \frac{R_n(\pi_1^{-1}(\zeta))}{Q_n(\zeta)} \right|^{1/(2n)} &\leq \\ \leq \left| \Phi(\pi_1^{-1}(\zeta)) \right| & \end{aligned} \tag{4.50}$$

uniformly for all $\zeta \in \overline{V}$.

(i) Let D_0 be a domain with the properties stated in part (i) of Theorem 3.7. We first assume that $\infty \notin \overline{D_0}$. Then there exists an open set V with $\overline{V} \subseteq D \setminus \{\infty\}$, $\partial D_0 \subseteq V$, and both sets $\overline{V \cup D_0}$ and $\overline{D_0}$ contain the same poles of the function $f \circ \pi_1^{-1}$. Let further V_n, $n \in N$, be open sets with $\overline{V_n} \subseteq V \cup D_0$, $\overline{D_0} \subseteq V_n$, let the approximant $[n/n]$ have the same set of poles on $\overline{V_n}$ and $\overline{D_0}$, let ∂V_n be smooth, and let there exist a constant c_1 such that length $(\partial V_n) \leq c_1$ for all $n \in N$. By $\widehat{[n/n]}$ we denote the approximant $[n/n]$ pole-cleared on V_n in accordance with Definition 3.2 (i). Because of our assumptions pole-clearing on V_n and on $\overline{D_0}$ has the same effect. From (3.13) it follows that

$$\left(f - \widehat{[n/n]}\right)(\zeta) = \frac{1}{2\pi i} \oint_{\partial V_n} (f - [n/n])(t) \frac{dt}{t - \zeta}, \quad \zeta \in \overline{D_0}. \tag{4.51}$$

Using definition (4.2) of the function Φ we deduce from (4.50) and (4.51) that

$$\overline{\lim_{\substack{n \to \infty \\ n \in N}}} \left| \left(f - \widehat{[n/n]}\right)(\zeta) \right|^{1/(2n)} \leq \sup_{\zeta \in \partial \overline{V}} G_D(\zeta) \tag{4.52}$$

uniformly for $\zeta \in \overline{D_0}$. Since the right-hand side of (4.52) is smaller than 1, this proves that $\widehat{[n/n]}$ converges to f uniformly on $\overline{D_0}$ as $n \to \infty$, $n \in N$. As the right-hand side of (4.52) does not depend on the selection of the subsequence N, the result is also proved for the full sequence $\mathbb{N}$. If $\infty \in \overline{D_0}$ formula (4.51) has to be changed in an obvious way.

(ii) Let now D_0 be a domain as assumed in part (ii) of Theorem 3.7. It is possible to choose two open sets V_1, V_2 with $\overline{V_j} \subseteq D$, $j = 1, 2$, $V_1 \cup V_2 \supseteq \overline{D_0}$, the two sets $\overline{V_1 \cup V_2}$ and $\overline{D_0}$ contain the same poles and zeros of the function $f \circ \pi_1^{-1}$, and both sets contain the same cluster points of zeros of $\{Q_n\}_{n \in N}$, *i.e.*, $Z_Q|_{\overline{V_1 \cup V_2}} = Z_Q|_{\overline{D_0}}$. Assume further that on $\overline{V_1}$ the function $f \circ \pi_1^{-1}$ has no poles, on $\overline{V_2}$ it has no zeros, and $\partial V_j \cap Z_Q = \emptyset$ for $j = 1, 2$. As in Theorem 3.7 let (π_{jn}, ζ_{jn}), $j = 1, \ldots, k_n$, $n \in N$, denote pairs of poles and zeros of $[n/n]$ that belong to spurious poles on $\overline{D_0}$, and define

$$g_n(\zeta) := \prod_{j=1}^{k_n} \frac{\zeta - \pi_{jn}}{\zeta - \zeta_{jn}}, \quad n \in N, \tag{4.53}$$

Since all poles and zeros considered in (4.53) cluster inside of V_1 or inside of V_2, and since their number is bounded, it follows from (3.11) in Theorem 3.6 that

$$g_n(\zeta) \to 1 \quad \text{as } n \to \infty,\ n \in N, \tag{4.54}$$

uniformly for $\zeta \in \partial V_1 \cup \partial V_2$. The asymptotic estimate proved in (4.50) holds uniformly on $\partial V_1 \cup \partial V_2$. With (4.54) and (3.14) in Definition 3.1 (ii), we therefore deduce that

$$\lim_{n\to\infty, n\in N} \widetilde{[n/n]} = f \circ \pi_1^{-1} \tag{4.55}$$

uniformly on ∂V_1, and since $f \neq 0$ on ∂V_2, we further deduce that

$$\lim_{n\to\infty, n\in N} \frac{1}{\widetilde{[n/n]}} = \frac{1}{f \circ \pi_1^{-1}} \tag{4.56}$$

uniformly on ∂V_2. The pole-cleared approximants $\widetilde{[n/n]}$ are analytic on $\overline{V_1}$. Hence, (4.55) hold uniformly on $\overline{V_1}$. Since $f \circ \pi_1^{-1}$ has no zeros on $\overline{V_2}$ the same is true for $\widetilde{[n/n]}$ if $n \in N$ is sufficiently large, and hence (4.56) holds uniformly on $\overline{V_2}$. Both limits (4.55) and (4.56) together prove uniform convergence on $\overline{D_0}$ in the spherical metric for the subsequence N. Convergence for the full sequence $\mathbb{N}$ follows in an obvious way as before. □

Proof of Lemma 3.4.— Set $r_2(\zeta) := (\zeta - \cos\alpha_1)(\zeta - \cos\alpha_2)$. It is well known (see for instance [StTo, Lemma 6.3.3]) that the defining condition (4.8) or (1.2) for the Padé polynomials P_n and Q_n implies that the denominator polynomial Q_n satisfies the orthogonality relation

$$\int_{-1}^{1} \zeta^l Q_n(\zeta) r_2(\zeta)\sqrt{1-\zeta^2}\,\mathrm{d}\zeta = 0 \quad \text{for } l = 0, \dots, n-1, \tag{4.57}$$

and each polynomial $Q_n \in \mathcal{P}_n$, $Q_n \not\equiv 0$, that satisfies (4.57), satisfied also (4.8) with an appropriately chosen polynomial $P_n \in \mathcal{P}_n$. Since the weight function $r_2(\zeta)\sqrt{1-\zeta^2}$ changes its sign on $[-1,1]$ only twice, the polynomial Q_n can have at most 2 zeros outside of $[-1,1]$. Therefore the Padé approximant $[n/n]$ can have at most 2 spurious poles. Note that $\overline{\mathbb{C}} \setminus [-1,1]$ is the convergence domain D from Theorem 2.2 for the sequence $\{[n/n]\}$.

A connection between the polynomials Q_n and the Chebychev polynomials of the second kind

$$U_n(\zeta) = 2^{-n}\,\frac{\sin(n+1)\arccos\zeta}{\sin\arccos\zeta} = \zeta^n + \cdots \tag{4.58}$$

plays a basic role in the sequel. The polynomials U_n satisfy the orthogonality relation

$$\int_{-1}^{1} \zeta^l U_n(\zeta)\sqrt{1-\zeta^2}\,\mathrm{d}\zeta = 0 \quad \text{for } l = 0, \dots, n-1. \tag{4.59}$$

Let $(a_n, b_n) \in \mathbb{R}^2$ be the solution of the system

$$\begin{pmatrix} \sin(n+2)\alpha_1 & \sin(n+1)\alpha_1 \\ \sin(n+2)\alpha_2 & \sin(n+1)\alpha_2 \end{pmatrix} \begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} \sin(n+3)\alpha_1 \\ \sin(n+3)\alpha_2 \end{pmatrix}. \tag{4.60}$$

Since it has been assumed that the 3 numbers α_1, α_2, π are linearly independent over $\mathbb{Q}$, the values $\sin\alpha_1$ and $\sin\alpha_2$ are algebraically independent (cf. [Si, S12]), which implies that the determinant of the system (4.60) is different from zero, and hence the solution $(a_n, b_n) \in \mathbb{R}^2$ exists and is unique. From (4.60) together with (4.58) we deduce that the right-hand side of

$$(Q_n r_2)(\zeta) = (U_{n+2} + a_n U_{n+1} + b_n U_n)(\zeta) \tag{4.61}$$

has zeros at the two points $\cos\alpha_j$, $j = 1, 2$. The concrete form of the left-hand side of (4.61) then follows from a comparison of the two orthogonality relations (4.57) and (4.59).

In Lemma 4.6, below, it will be shown that the linear independence over $\mathbb{Q}$ of the three numbers α_1, α_2, π implies that the set

$$\{(a_n, b_n) \mid n = 1, 2, \ldots\} \tag{4.62}$$

in dense in $\mathbb{R}^2$. Let the conformal mapping $\psi : \overline{\mathbb{C}} \setminus [-1, 1] \to \overline{\mathbb{C}} \setminus \overline{\mathbf{D}(0, 1/2)}$ be defined by $\zeta \mapsto \psi(\zeta) := \left(\zeta + \sqrt{\zeta^2 - 1}\right)/2$. The polynomials U_n possess ratio asymptotics, *i.e.*, the limits

$$\lim_{n\to\infty} \frac{U_{n+1}(\zeta)}{U_n(\zeta)} = \psi(\zeta) \tag{4.63}$$

and

$$\lim_{n\to\infty} \frac{U_{n+2}(\zeta)}{U_n(\zeta)} = \psi(\zeta)^2 \tag{4.64}$$

hold true locally uniformly for $\zeta \in \overline{\mathbb{C}} \setminus [-1, 1]$. The identity (4.61) can be rewritten as

$$Q_n r_2 = U_n \left(\frac{U_{n+2}}{U_n} + a_n \frac{U_{n+1}}{U_n} + b_n \right). \tag{4.65}$$

Let $x \in \mathbb{C} \setminus [-1, 1]$ be given arbitrarily and define

$$a := -\psi(x) - \overline{\psi(x)}, \quad b := |\psi(x)|^2. \tag{4.66}$$

Because of the denseness of the set (4.62) in $\mathbb{R}^2$ there exists an infinite subsequence $N \subseteq \mathbb{N}$ such that

$$a_n \to a, \quad b_n \to b, \quad \text{as } n \to \infty,\ n \in N. \tag{4.67}$$

From the assertions (4.63) through (4.67) we then deduce that for each $n \in N$ the polynomial Q_n has a zero ζ_n such that

$$\zeta_n \to x \quad \text{as } n \to \infty,\ n \in N. \tag{4.68}$$

It remains to show that the zero ζ_n generates a spurious pole of the Padé approximant $[n/n]$. Indeed, if ζ_n is not a spurious pole, then ζ_n is also a zero of the numerator polynomial P_n in (4.8). It follows from (4.66) that with ζ_n also $\overline{\zeta_n}$ is a zero of Q_n and P_n, $n \in N$. We consider the polynomials

$$\begin{aligned} \tilde{Q}_n(\zeta) &:= Q_n(\zeta)/\left((\zeta - \zeta_n)(\zeta - \overline{\zeta_n})\right), \\ \tilde{P}_n(\zeta) &:= P_n/\left((\zeta - \zeta_n)(\zeta - \overline{\zeta_n})\right), \quad n \in N. \end{aligned} \tag{4.69}$$

Inserting them into (4.8) yield that

$$(\tilde{Q}_n f - \tilde{P}_n)(\zeta) = O(\zeta^{-n-1}) \quad \text{as } \zeta \to \infty. \tag{4.70}$$

From (4.70) it then follows, as at the beginning of the proof, that $\tilde{Q}_n$ satisfies the orthogonality relation (4.57). Since $\deg(\tilde{Q}_n) = n-2$ and $\tilde{Q}_n r_2$ satisfies relation (4.59), we have $\tilde{Q}_n r_2 = U_n$, which in turn implies with (4.58) that $\sin\left((n+1)\arccos(\cos\alpha_j)\right) = \sin(n+1)\alpha_j = 0$ for $j = 1, 2$ and $n \in N$. However, these last equations contradict the assumptions that the numbers α_1, α_2, and π are linearly independent over $\mathbb{Q}$. Hence, it is proved that ζ_n, $n \in N$, is a spurious pole of the approximant $[n/n]$. □

LEMMA 4.5.— *Let $(a_n, b_n) \in \mathbb{R}^2$, $n \in \mathbb{N}$, be the solution of the system (4.60), and assume that the numbers α_1, α_2, π are rationally independent. Then the set $\{(a_n, b_n) \mid n = 1, 2, \ldots\}$ is dense in $\mathbb{R}^2$.*

Proof.— After some trigonometric calculations it follows from (4.60) that

$$\begin{aligned} a_n &= 2\,\frac{\cos\alpha_1 \sin\alpha_2 \cot(n+2)\alpha_2 - \sin\alpha_1 \cos\alpha_2 \cot(n+2)\alpha_1}{\cos\alpha_2 - \cos\alpha_1 + \sin\alpha_1 \cot(n+2)\alpha_1 - \sin\alpha_2 \cot(n+2)\alpha_2} \\ b_n &= 1 + 2\,\frac{\cos\alpha_1 - \cos\alpha_2}{\cos\alpha_2 - \cos\alpha_1 + \sin\alpha_1 \cot(n+2)\alpha_1 - \sin\alpha_2 \cot(n+2)\alpha_2}. \end{aligned} \tag{4.71}$$

If in (4.71) the numbers a_n, b_n, $\cot(n+2)\alpha_1$, $\cot(n+2)\alpha_2$ are replaced by the variables a, b, x, y, respectively, then $(x,y) \mapsto (a,b)$ defines a mapping $\Psi : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \longrightarrow \overline{\mathbb{R}} \times \overline{\mathbb{R}}$, which is surjective and continuous if $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$ is endowed with the spherical metric. The surjectivity can be verified most easily by considering simultaneously the two expressions $a = a(x,y)$ and $(b-1)/a = \big(b(x,y)-1\big)/a(x,y)$.

Since the numbers $\alpha_1/2\pi$, $\alpha_2/2\pi, 1$ are linearly independent over $\mathbb{Q}$, it follows from Weyl's uniform distribution Theorem (cf. [Ch, chap. VIII]) that the set of simultaneous remainders

$$\Big\{(t_n, u_n) \in [0, 2\pi)^2 \mid (t_n, u_n) \equiv \big((n+2)\alpha_1, (n+2)\alpha_2\big) \\ \mathrm{mod}(2\pi, \ldots, 2\pi)\,, \quad n = 1, 2, \ldots\Big\} \tag{4.72}$$

is dense in $[0, 2\pi) \times [0, 2\pi)$. Hence, the set

$$\big\{\big(\cot(n+2)\alpha_1, \cot(n+2)\alpha_2\big) \mid n = 1, 2, \ldots\big\} \tag{4.73}$$

is dense in $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$, and the lemma follows from the properties of the mapping Ψ. □

5. Proofs of the Theorems 3.3 and 3.8

For the proofs of the Theorems 3.3 and 3.8 some elements of the theory of compact Riemann surfaces are essential. A fundamental role is played by the Jacobi inversion problem. With its help we can understand the asymptotic distribution of the zeros of the remainder function R_n introduced in (4.8). It turns out that in distinction to all poles and all other zeros of the remainder function R_n, the distribution of $m-1$ zeros of R_n is only very indirectly determined by properties of the function f. However, the concept of the Jacobi inversion problem allows us to understand the distribution of these last $m-1$ zeros. As a consequence we get information about the asymptotic distribution of spurious poles. The section is closed by a discussion of possible generalizations.

The terminology used in the theory of Riemann surfaces is based on the book [Sp] by G. Springer, where Chapter 10 is especially relevant. As before $\mathcal{R}$ denotes the concrete Riemann surface defined by the equation

$y^2 = (\zeta - a_1) \cdots (\zeta - a_{2m})$, and for all points on $\mathcal{R} \setminus \{a_1, \ldots, a_{2m}\}$ a local coordinate z is defined by the projection $\pi : \mathcal{R} \to \overline{\mathbb{C}}$. As before points on $\overline{\mathbb{C}}$ are denoted by ζ. As in (3.2) B_1 and B_2 denote the two domains in $\mathcal{R}$ lying over the convergence domain $D = \overline{\mathbb{C}} \setminus F$, we write $z^{(j)}$ if z lies in B_j, and π_1^{-1} and π_2^{-1} denote the two branches of the inverse mapping π^{-1} mapping D onto B_1 and B_2, respectively.

The Riemann surface $\mathcal{R}$ is of genus $g = m - 1$, and the g differentials

$$\omega_j := \frac{z^{j-1}\,\mathrm{d}z}{y(z)}, \quad j = 1, \ldots, g, \tag{5.1}$$

form a basis in the space of Abelian differentials of the first kind, where y is defined as in (4.1) (cf. [Sp, chap. 10.10]). In the previous section we have used multisets for the description of poles and zeros of meromorphic functions. In the Riemann surface theory it is usual to use *divisors* for this purpose: For a point $a \in \mathcal{R}$ by D_a we denote the elementary divisor, which is a mapping $D_a : \mathcal{R} \to \{0, 1\}$ defined by $D_a(a) = 1$ and $D_a(z) = 0$ for $z \in \mathcal{R} \setminus \{a\}$. Divisors form a multiplicative Abelian group generated by elementary divisors. Multiplication is defined by addition of its values. Thus, divisors assume their values in $\mathbb{Z}$, and they are different from zero only at finitely many points in $\mathcal{R}$. To each meromorphic function h on $\mathcal{R}$ a divisor (h) is associated by assigning a positive integer to each zero that is equal to its order and a negative integer to each pole of the function again equal to its order. Thus, if h has the zero set $Z(h) = \{z_1, \ldots, z_l\}$ and the pole set $P(h) = \{\pi_1, \ldots, \pi_l\}$, then the divisor (h) is given by

$$(h) := \frac{D_{z_1} \cdots D_{z_l}}{D_{\pi_1} \cdots D_{\pi_l}} = D_{z_1} \cdots D_{z_l} D_{\pi_1}^{-1} \cdots D_{\pi_l}^{-1}. \tag{5.2}$$

If γ is a simple arc on $\mathcal{R}$ issuing from $z_2 \in \mathcal{R}$ and leading to $z_1 \in \mathcal{R}$, then $\partial\gamma$ denotes the divisor

$$\partial\gamma := D_{z_1} D_{z_2}^{-1}, \tag{5.3}$$

and correspondingly a chain of arcs γ defines a divisor generated by more than two elementary divisors. In both cases, (h) and $\partial\gamma$, the sum of positive and negative values is equal, *i.e.*, the *degree* of these divisors $\deg(D) := \sum_{z \in \mathcal{R}} D(z)$ is equal to zero (cf. [Sp, chap. 10.4]).

Let the curves $\mathbf{a}_1, \ldots, \mathbf{a}_g$, $\mathbf{b}_1, \ldots, \mathbf{b}_g$ form a *homology basis* on $\mathcal{R}$, let $\varphi_1, \ldots, \varphi_g$ be the canonical basis of the Abelian differentials of the first kind associated with this homology basis. Such a canonical basis is defined by the following normalization of the period integrals:

$$\int_{\mathbf{a}_i} \varphi_j =: A_{ij} = \delta_{ij}, \quad \int_{\mathbf{b}_i} \varphi_j =: B_{ij}, \quad i, j = 1, \ldots, g, \tag{5.4}$$

and $(\mathrm{Im}(B_{ij}))_{i,j=1,\ldots,g}$ is a positive definite matrix. In many situations it is more comfortable to use the basis $\{\varphi_1, \ldots, \varphi_g\}$ instead of the differentials in (5.1). The two matrices (A_{ij}) and (B_{ij}) together form the *period matrix*. On $\mathbb{C}^g$ an equivalence relation

$$c \equiv c' \mod \mathrm{Per}(\varphi_1, \ldots, \varphi_g) \tag{5.5}$$

is introduced by defining (5.5) to hold true for $c = (c_1, \ldots, c_g)$, $c' = (c'_1, \ldots, c'_g) \in \mathbb{C}^g$ if and only if there exists $n_1, \ldots, n_{2g} \in \mathbb{Z}$, such that

$$c_i = c'_i + n_i + \sum_{j=1}^{g} n_{g+j} B_{ij} \quad \text{for } i = 1, \ldots, g. \tag{5.6}$$

As a consequence of the Riemann-Roch Theorem (cf. [Sp, chap. 10.5]) we have

PROPOSITION 5.1.— *Let l ($l \geq g = m-1$) points $\pi_1, \ldots, \pi_l$ and $l-g$ points $z_{g+1}, \ldots, z_l$ be given on $\mathcal{R}$.*

(a) It is always possible to choose g points $z_1, \ldots, z_g \in \mathcal{R}$ in such a way that a meromorphic function h on $\mathcal{R}$ exists with an associated divisor as defined in (5.2), i.e., $P(h) = \{\pi_1, \ldots, \pi_l\}$ and $Z(h) = \{z_1, \ldots, z_l\}$.

(b) The g points $z_1, \ldots, z_l \in \mathcal{R}$ of part (a) are uniquely determined by the given $2l-g$ points $\pi_1, \ldots, \pi_l$, $z_{g+1}, \ldots, z_l$ if and only if no non-trivial Abelian differential $\tilde{\varphi}$ of the first kind exists with zeros at the g points $z_1, \ldots, z_g$.

Remark.— The multiplicity of points in a divisor corresponds in the same way as the multiplicity of elements in a multiset with the order of zeros or poles of the associated function.

Proof

(a) Let D denote the divisor $D_{\pi_1} \cdots D_{\pi_l} D_{z_{g+1}}^{-1} \cdots D_{z_l}^{-1}$. The Riemann-Roch Theorem (cf. [Sp, Theorem 10.10]) says that

$$r\left(\frac{1}{D}\right) = \deg(D) + i(D) - g + 1, \tag{5.7}$$

where $\deg(D) = \sum_{z \in \mathcal{R}} D(z) = g$ is the degree of the divisor D, $i(D)$ the dimension of the space of all Abelian differentials on $\mathcal{R}$ with divisors that are multiples of D (for a detailed definition see [Sp, chap. 10.4]), and $r(1/D)$ is the dimension of the space $L(1/D)$ of all meromorphic functions h on $\mathcal{R}$ with a divisor (h) that is a multiple of $1/D$, *i.e.*, h can have poles only at $\pi_1, \ldots, \pi_l$ and has to have zeros at $z_{g+1}, \ldots z_l$. It follows immediately from (5.7) that $r(1/D) \geq 1$, and therefore there exists (if $l > g$) a non-constant meromorphic function h in $L(1/D)$. Let l_1 be the number of its zeros and poles. We have $l - g \leq l_1 \leq l$. Without loss of generality we assume h has poles at $\pi_1, \ldots, \pi_{l_1}$. Besides $z_{g+1}, \ldots, z_l$, there exist $l_1 - g$ zeros $z_1, \ldots z_{l_1 - g}$ of h. If we add to this list the points $\pi_{l_1+1}, \ldots, \pi_l$ as $z_{l_1-g+1}, \ldots, z_g$, then part (a) is proved. (If $l = g$, then the selection $z_j = \pi_j, j = 1, \ldots, g$, works).

(b) Assume that h and $\tilde{h}$ are two functions with the properties of the function h introduced in the proof of part (a). Let $z_1, \ldots, z_g$ and $\tilde{z}_1, \ldots, \tilde{z}_g$ be the points selected in connection with h and $\tilde{h}$, respectively. Set $h_1 := \tilde{h}/h$. Then we have

$$(h_1) = D_{\tilde{z}_1} \cdots D_{\tilde{z}_g} D_{z_1}^{-1} \cdots D_{z_2}^{-1} = \tilde{D}_0 / D_0, \tag{5.8}$$

where D_0 and $\tilde{D}_0$ are integral divisors of degree g. From the Riemann-Roch Theorem it follows that

$$r\left(\frac{1}{D_0}\right) = \deg(D_0) + i(D_0) - g + 1 = i(D_0) + 1. \tag{5.9}$$

Since D_0 is integral, $i(D_0)$ is the dimension of the space of all Abelian differentials φ of the first kind having zeros at the g points $z_1, \ldots, z_g$. If such a differential does not exist we have $i(D_0) = 0$, and $r(1/D_0) = 1$ in (5.9). As D_0 is an integral divisor, $L(1/D_0)$ contains the constant functions, and $r(1/D_0) = 1$ therefore implies that it only contains constant functions. Hence, $h_1 \equiv$ const., and consequently (5.8) implies that $z_j = \tilde{z}_j$ for $j = 1, \ldots, g$. It is clear that in case of $i(D_0) > 0$ the function h_1 is not necessarily constant, and therefore different sets of z_j's and $\tilde{z}_j$'s exist. □

From (5.1) we conclude that any Abelian differential ψ of the first kind can be represented as

$$\psi = \frac{P(z)\,\mathrm{d}z}{y} \tag{5.10}$$

with P a polynomial of degree at most $g-1$. Hence, the differential ψ can have zeros at more than $g-1$ points $z_1, \ldots, z_l \in \mathcal{R}$, $l \geq g$, if, and only if, at least two points lie over the same basis point, *i.e.*, there exists at least one point z_j with an image $\varphi(z_j)$ among the points $z_1, \ldots, z_l$, where φ is the covering transformation $\varphi : \mathcal{R} \to \mathcal{R}$, *i.e.*, $\pi = \pi \circ \varphi$ and $\varphi \neq \mathrm{id}_{\mathcal{R}}$. These observations are summarized in

LEMMA 5.2.— *For g points $z_1, \ldots, z_g \in \mathcal{R}$ there exists a non-trivial Abelian differential ψ of the first kind having zeros at $z_1, \ldots, z_g$ if and only if*

$$\{z_1, \ldots, z_g\} \cap \varphi(\{z_1, \ldots, z_g\}) \neq \emptyset. \tag{5.11}$$

The discrete set

$$\Gamma := \{c \in \mathbb{C}^g \mid c \equiv 0 \mod \mathrm{Per}(\varphi_1, \ldots, \varphi_g)\} \tag{5.12}$$

defines a lattice in $\mathbb{C}^g$ called the *period lattice* (cf. [Sp, chap. 10.8]), and

$$\mathrm{Jac}(\mathcal{R}) := \mathbb{C}^g/\Gamma \tag{5.13}$$

is called the *Jacobian variety* of the Riemann surface $\mathcal{R}$. Its definition depends on the choice of the basis $\varphi_1, \ldots, \varphi_g$, but any other choice leads to an isomorphic variety. Note that the use of the symbol Γ here has a different meaning from that in (3.2). There exist $2g$ vectors $\Gamma_1, \ldots, \Gamma_{2g} \in \mathbb{C}^g$ linearly independent over $\mathbb{R}$ and

$$\Gamma = \mathbb{Z}\Gamma_1 + \cdots + \mathbb{Z}\Gamma_{2g}. \tag{5.14}$$

For each $c \in \mathrm{Jac}(\mathcal{R})$ the *coordinates* $\gamma_i = \gamma_i(c) \in [0,1)$, $i = 1, \ldots, 2g$, are uniquely defined by

$$c \equiv \sum_{i=1}^{2g} \gamma_i \Gamma_i \mod \mathrm{Per}(\varphi_1, \ldots, \varphi_g). \tag{5.15}$$

In the *Jacobi inversion* problem properties of two rather similarly defined mappings $\tilde{J}$ and J are studied. The first mapping is defined as

$$\begin{aligned}&\tilde{J} : \mathcal{R}^g \longrightarrow \mathrm{Jac}(\mathcal{R}),\\ z = (z_1, \ldots, z_g) \mapsto &\left(\sum_{j=1}^{g} \int_{\infty^{(2)}}^{z_j} \varphi_j, \ldots, \sum_{j=1}^{g} \int_{\infty^{(2)}}^{z_j} \varphi_g \right) \operatorname{mod} \mathrm{Per}(\varphi_1, \ldots, \varphi_g).\end{aligned} \tag{5.16}$$

The definition depends on the reference point $\infty^{(2)} \in \mathcal{R}$. It is shown in [Sp, chap. 10.8], that $\tilde{J}$ is locally homeomorphic and that, what is more important, it is surjective.

One can consider $\tilde{J}$ as a mapping of divisors associated with chains γ of arcs that connect the g points $\infty^{(2)}, \ldots, \infty^{(2)}$ with g points $z_1, \ldots, z_g \in \mathcal{R}$, *i.e.*, divisors of the form

$$\partial\gamma = D_z := D_{z_1} \cdots D_{z_g} D_{\infty^{(2)}}^{-1} \cdots D_{\infty^{(2)}}^{-1}, \quad z = (z_1, \ldots, z_g). \tag{5.17}$$

The concept can be generalized by considering the set $\mathrm{Div}_0(\mathcal{R})$ of all divisors D with $\deg(D) = 0$, *i.e.*, divisors D that have the same number of positive and negative values. For each $D \in \mathrm{Div}_0(\mathcal{R})$ there exists a chain of arcs such that $\partial\gamma = D$, but, in contrast with (5.17), the number of points in the numerator and denominator is not necessarily equal to g. From Abel's Theorem (cf. [Sp, Theorem 10.7]) we know that for a given divisor $D \in \mathrm{Div}_0(\mathcal{R})$ there exists a meromorphic function h on $\mathcal{R}$ with $(h) = D$ if and only if for a chain γ of arcs with $\partial\gamma = D$ we have

$$\left(\int_\gamma \varphi_1, \ldots, \int_\gamma \varphi_g \right) \equiv 0 \quad \mathrm{mod}\ \mathrm{Per}(\varphi_1, \ldots, \varphi_g). \tag{5.18}$$

Divisors D that are defined by a meromorphic function h on $\mathcal{R}$, *i.e.*, $D = (h)$, are called *principal*, and the set of all principal divisors is denoted by $\mathrm{Div}_P(\mathcal{R})$. It is immediate that they form a subgroup of the Abelian group $\mathrm{Div}_0(\mathcal{R})$. The mapping

$$\begin{aligned}&J : \mathrm{Div}_0(\mathcal{R}) \longrightarrow \mathrm{Jac}(\mathcal{R})\\ &D = \partial\gamma \longmapsto \left(\int_\gamma \varphi_1, \ldots, \int_\gamma \varphi_g \right) \ \mathrm{mod}\ \mathrm{Per}(\varphi_1, \ldots, \varphi_g)\end{aligned} \tag{5.19}$$

can be seen as an extension of $\tilde{J}$ (see (5.16)). The mapping J is therefore also surjective. With the mapping J, Abel's Theorem can be formulated as

$\mathrm{Ker}(J) = \mathrm{Div}_P(\mathcal{R})$. Hence, passing to the factor group $\mathrm{Div}_0(\mathcal{R}) \setminus \mathrm{Div}_P(\mathcal{R})$ induces a bijective mapping $\hat{J} : \mathrm{Div}_0(\mathcal{R})/\,\mathrm{Div}_P(\mathcal{R}) \to \mathrm{Jac}(\mathcal{R})$, $[D] \mapsto J(D)$. In the sequel we shall use only the two maps $\tilde{J}$ and J.

We have seen that with the map J, Abel's Theorem can be rephrased as: Any $D \in \mathrm{Div}_0(\mathcal{R})$ belongs to $\mathrm{Div}_P(\mathcal{R})$ if and only if $J(D) = 0$. Note that by (5.13) addition and a topology is induced in $\mathrm{Jac}(\mathcal{R})$ from $\mathbb{C}^g$. The only conclusions we need from the Jacobi inversion problem are contained in the following

PROPOSITION 5.3.— *Let $G_1, \dots, G_g \subseteq \mathcal{R}$ be g open sets satisfying*

$$\overline{G_i} \cap \varphi(\overline{G_j}) = \emptyset \quad \text{for } i, j = 1, \dots, g,\ i \neq j. \tag{5.20}$$

Then there exists an open set $G_0 \subseteq \mathrm{Jac}(\mathcal{R})$ with the following property: If $2l - g$ points $\pi_1, \dots, \pi_l, z_{g+1}, \dots, z_l \in \mathcal{R}$, $l \geq g$, are given, the divisor D is defined as

$$D := D^g_{\infty^{(2)}} D_{z_{g+1}} \cdots D_{z_l} D^{-1}_{\pi_1} \cdots D^{-1}_{\pi_l} \in \mathrm{Div}_0(\mathcal{R}), \tag{5.21}$$

and if $z_1, \dots, z_g \in \mathcal{R}$ are g points selected in accordance with part (a) of Proposition 3.1, then

$$J(D) \in G_0, \tag{5.22}$$

implies that

$$z_j \in G_j \quad \text{for } j = 1, \dots, g \tag{5.23}$$

and the g points $z_1, \dots, z_g$ are uniquely determined on $\mathcal{R}$ by the condition of part (a) in Proposition 3.1.

Proof.— Set $G := G_1 \times \cdots \times G_g \subseteq \mathcal{R}^g$ and define $G_0 := -\tilde{J}(G) \subseteq \mathrm{Jac}(\mathcal{R})$. Since $\tilde{J}$ is locally homeomorphic, the set G_0 is open in $\mathrm{Jac}(\mathcal{R})$.

Assume that $J(D) \in G_0$ holds true with the divisor D defined in (5.21). From the definition of G_0 it follows that there exists $z = (z_1, \dots, z_g) \in G$ such that

$$\tilde{J}(z) = -J(D). \tag{5.24}$$

Comparing the definitions of $\tilde{J}$ and J in (5.16) and (5.19), respectively, shows that $\tilde{J}(z) = J(D_z)$ with the divisor D_z defined in (5.17). From (5.24) we deduce that

$$J(DD_z) = J(D) + J(D_z) = 0. \tag{5.25}$$

The definitions (5.17) and (5.21) together show that

$$DD_z = D_{z_1} \cdots D_{z_l} D_{\pi_1}^{-1} \cdots D_{z_l}^{-1} .$$

From Abel's Theorem and (5.25) we then know that there exists a meromorphic function h on $\mathcal{R}$ such that $(h) = DD_z$, which shows that the points $z_1, \ldots, z_g \in \mathcal{R}$ have the property required in part (a) of Proposition 5.1. Since we know that $z = (z_1, \ldots, z_g) \in G_1 \times \cdots \times G_g$, the assertions in (5.23) are satisfied, and it follows from condition (5.20), Lemma 5.2, and part (b) of Proposition 5.1 that the g points $z_1, \ldots, z_g$ are uniquely determined on $\mathcal{R}$, which completes the proof. □

Proposition 5.1 and 5.3 is used to analyse the remainder function R_n defined in (4.8). In this analysis the following three multisets play an important role:

$$\begin{aligned} PS_n &:= P(f)|_{\mathcal{R}\setminus\{\infty^{(2)}\}} \cup \pi_2^{-1}\big(P(z^{n+m} r_2)|_{\{\infty\}}\big) , \\ ZS_n &:= Z(R_n) \cup \big(P(f)|_{\mathcal{R}\setminus\{\infty^{(2)}\}}\big) \setminus P(R_n) \cup \pi_2^{-1}\big(Z(z^{-n} Q_n)|_{\{\infty\}}\big) , \\ CS_n &:= \Big(\varphi(P(f)|_{B_1}) \cup P(f)|_{B_2\setminus\{\infty^{(2)}\}}\Big) \setminus \pi_2^{-1}(P(r_2)) \\ &\qquad \cup \pi_2^{-1}(Z(r_2)|_{\mathbb{C}}) \cup \pi_1^{-1}(Z(z^{-n-1})) , \end{aligned} \tag{5.26}$$

where the functions f and r_2, the multisets $Z(\,\cdot\,)$ and $P(\,\cdot\,)$, the mappings φ and π_j^{-1}, $j = 1, 2$, are defined as in Section 4. The set PS_n contains all possible poles of R_n and possibly some additional points, the set ZS_n all zeros of R_n and possible some additional points, and the set CS_n contains the potential cluster points of the sets ZS_n. In the definitions in (5.26) special care is taken at the point $\infty^{(2)} \in \mathcal{R}$. Some basic facts are assembled in the next lemma.

LEMMA 5.4.— *We have*

$$P(R_n) \subseteq PS_n, \tag{5.27}$$

$$\pi_1^{-1}(Z(z^{-n-1})) \subseteq PS_n, \tag{5.28}$$

$$\pi_2^{-1}(Z(z^{-n-m+k_0})) \subseteq PS_n, \quad k_0 := \max(0, v(r_2, \infty)) , \tag{5.29}$$

$$\#(ZS_n) = \#(PS_n) = \#(P(R_n)) - \#(S_{0n}), \tag{5.30}$$

$$S_{0n} := \big(Z(f)|_{\mathcal{R}\setminus\{\infty^{(2)}\}}\big) \setminus P(R_n) \cup \pi_2^{-1}\big(Z(z^{-n} Q_n)|_{\{\infty\}}\big).$$

and if the assumptions of Corollary 3.2 are satisfied, then we further have

$$\#(ZS_n) - \#(CS_n) = m - 1 = g, \quad n \geq n_0, \tag{5.31}$$

$$\#(CS_n) = n + k_1, \quad k_1 := 1 + \#(P(f)|_{\mathcal{R}\setminus\{\infty^{(2)}\}}) - v(r_2, \infty). \tag{5.32}$$

Remark.— The constants k_0 and k_1 are independent of n. As before, the valuation of r_2 at ∞ is denoted by $v(r_2, \infty)$, *i.e.*, the order of a zero is counted positively and that of a pole negatively.

Proof.— Relation (4.8) and representation (4.5) imply that

$$R_n \circ \pi_2^{-1} = 2Q_n r_2 y \circ \pi_2^{-1} - R_n \circ \pi_1^{-1}. \tag{5.33}$$

Inclusion (5.27) is an immediate consequence of the definition of R_n in (4.8) and a consideration of (5.33) near $\infty^{(2)}$. Inclusion (5.28) follows from the definition of Padé polynomials in (1.1), (1.2), together with (4.8). Inclusion (5.29) is a consequence of the definition of PS_n in (5.26) and of equation (5.33) near $\infty^{(2)}$.

The first equality in (5.30) is a consequence of the fact that the number of poles and zeros of R_n on $\mathcal{R}$ is identical. In addition the sets PS_n and ZS_n have to be considered at $\infty^{(2)}$. The second equality then follows immediately.

If the assumptions of Corollary 3.2 are satisfied, then $P(f)|_{\mathcal{R}\setminus\{\infty^{(2)}\}} = P(f)|_{B_1} \cup P(f)|_{B_2\setminus\{\infty^{(2)}\}}$, $P(r_2) = P(r_2)|_D$, and $Z(r_2) = Z(r_2)|_D$ with $D \subseteq \overline{\mathbb{C}}$ the convergence domain. In order to prove (5.31) we first show that

$$\pi_2^{-1}(P(r_2)|_{\mathbb{C}}) \subseteq \varphi(P(f)|_{B_1}) \cup P(f)|_{B_2\setminus\{\infty^{(2)}\}}. \tag{5.34}$$

Indeed, (5.34) follows directly from representation (4.5) of f. Then it is best to use the first equality in (5.30) and consider the definitions of PS_n and CS_n on $\mathcal{R} \setminus \{\infty^{(2)}\}$ and at $\infty^{(2)}$ separately. Equation (5.32) is immediate. We note that we always have $k_1 \geq 0$. □

The three sets PS_n, ZS_n, and CS_n defined in (5.26) grow with n, but the dependence on n is not very complicated: except for variations of points in the set ZS_n the main change is the addition of the element $\infty^{(1)}$ or $\infty^{(2)}$

in each step $n \to n+1$. In order to have the interesting processes isolated in finite sets, we define

$$\begin{aligned} \widehat{PS} &:= PS_n \setminus \pi_2^{-1}(Z(z^{-n+k_0})), \\ \widehat{ZS}_n &:= Z_n \setminus \pi_1^{-1}(Z(z^{-n+k_0})), \quad n \geq k_0 \\ \widehat{CS} &:= CS_n \setminus \pi_1^{-1}(Z(z^{-n+k_0})), \end{aligned} \tag{5.35}$$

with k_0 introduced in (5.29). The first and the last set in (5.35) do not depend on n. If the assumptions of Corollary 3.2 are satisfied, then it follows (5.32), (5.30) and (5.31) in Lemma 5.4 that

$$\#(\widehat{PS}) = \#(\widehat{ZS}_n) = k_1 + k_0, \tag{5.36}$$

$$\#(\widehat{ZS}_n) = \#(\widehat{CS}) + g. \tag{5.37}$$

LEMMA 5.5.— *Let the assumptions of Corollary 3.2 be satisfied. For each $n \geq k_0$ it is possible to select g points $z_{1n}, \ldots, z_{gn} \in \mathcal{R}$ from the set $\widehat{ZS}_n$ in such a way that*

$$\widehat{ZS}_n \setminus \{z_{n1}, \ldots, z_{ng}\} \to \widehat{CS} \quad \textit{as } n \to \infty. \tag{5.38}$$

Remark.— The notion of convergence of multisets has been defined before (4.7).

Proof.— Before we start with the proof proper, we mention the special situation that zeros of Q_n lie exactly under poles of the function f. In this case poles of f are canceled out in (4.8), we then have a strict inclusion in (4.9), and

$$(P(f)|_{\mathcal{R}\setminus\{\infty^{(2)}\}}) \setminus P(R_n) \neq \emptyset. \tag{5.39}$$

This situation is excluded in the first step.

Equation (5.33) is basic for the proof of the lemma. Let $N \subseteq \mathbb{N}$ be an infinite subsequence such that the limits (4.12), (4.25), and also the limits (4.16) and (4.27) in Lemma 4.1 and 4.3 hold true. It follows from the limits (4.16) and (4.27) together with the inequalities in (4.3) that on the set $D \setminus (\{\infty\} \cup \pi(P_R))$ the last term in equation (5.33) is small in comparison with the first two terms. The set P_R has been defined in (4.12). Thus, we can apply Rouché's Theorem to (5.33) in the neighborhood of every point of $\widehat{CS}|_{B_2\setminus\{\infty^{(2)}\}}$.

From Lemma 4.5 we know that at each point of $\pi(P(f)|_{B_1})$ the denominator polynomial Q_n has a zero of at least the same order. Under the assumption that the set in (5.39) is empty it follows from (5.55) and Rouché's Theorem that in a neighborhood of a point $z \in \widehat{CS}|_{B_2}$ the set $\widehat{ZS}_n$ has at least the same number of elements as the multiplicity of z in $\widehat{CS}|_{B_2}$. That this also holds at $\infty^{(1)}$, follows from (5.26) and (4.8). From the assumption of Corollary 3.2 it follows that $Z(r_2)|_F = \emptyset$ and therefore $\widehat{CS}$ contains only points from $B_2 \cup \{\infty^{(1)}\}$, and consequently we have shown that near each point of $\widehat{CS}$ there lie points of $\widehat{ZS}_n$ with the necessary multiplicity. From (5.37) we know that $\widehat{ZS}_n$ contains exactly g points more than $\widehat{ZC}$. Hence, for each $n \in N$ it is possible to remove g points $z_{1n}, \ldots, z_{gn}$ from $\widehat{ZS}_n$, and for the remaining sets the convergence (5.38) holds true as $n \to \infty$, $n \in N$. Since $\widehat{CS}$ is independent of n, limit (5.38) hold for the full sequence $\mathbb{N}$.

If some zeros of Q_n cancel out poles of f, then an inspection of the consequences in equation (5.33) shows that in this situation the set $Z(R_n)$ may have less points than $\widehat{CS}$ near some points of $P(f)|_{B_2\setminus\{\infty^{(2)}\}}$. However, the missing points are contained in the set $P(f)_{B_2\setminus\{\infty^{(2)}\}} \subseteq \widehat{ZS}_n$. □

In the sequel it is assumed that the assumptions of Corollary 3.2 are always satisfied. From the definition of the two sets PS_n and ZS_n in (5.26), the two sets $\widehat{PS}$ and $\widehat{ZS}_n$ in (5.35), and from the description of the behavior of R_n at $\infty^{(2)}$ in (5.33) we deduce that the divisor of R_n is given by

$$\begin{aligned}(R_n) &= \prod_{z\in ZS_n} D_z \prod_{\pi\in PS_n} D_\pi^{-1} \\ &= \left(\prod_{z\in\widehat{ZS}_n} D_z \prod_{\pi\in\widehat{PS}} D_\pi^{-1} \right) \left(D_{\infty^{(1)}} D_{\infty^{(2)}}^{-1}\right)^{n-k_0} \\ &= \hat{D}_n \left(D_{\infty^{(1)}} D_{\infty^{(2)}}^{-1}\right)^{n-k_0}\end{aligned} \tag{5.40}$$

with the divisor $\hat{D}_n$ defined as

$$\hat{D}_n := \prod_{z\in\widehat{ZS}_n} D_z \prod_{\pi\in\widehat{PS}} D_\pi^{-1}. \tag{5.41}$$

It follows from (5.36) that $\hat{D}_n \in \mathrm{Div}_0(\mathcal{R})$. Set

$$\tilde{D}_n := D_{z_{1n}} \cdots D_{z_{gn}} D^{-g}_{\infty^{(2)}} \in \mathrm{Div}_0(\mathcal{R}), \tag{5.42}$$

where $z_{1n}, \ldots, z_{gn}$ are the points selected from $\widehat{ZS}_n$ in Lemma 5.5 for each $n \geq k_0$, and define further

$$\begin{aligned} D_n &:= (R_n)\tilde{D}_n^{-1} = D^g_{\infty^{(2)}} \left(\prod_{z \in \widehat{ZS}_n \setminus \{z_{1n}, \ldots, z_{gn}\}} D_z \prod_{\pi \in \widehat{PS}} D_\pi^{-1} \right) \\ &\quad \times \left(D_{\infty^{(1)}} D^{-1}_{\infty^{(2)}} \right)^{n-k_0}, \\ \hat{D} &:= D^g_{\infty^{(2)}} \prod_{z \in \widehat{CS}} D_z \prod_{\pi \in \widehat{PS}} D_\pi^{-1}. \end{aligned} \tag{5.43}$$

The second equality in the first line of (5.43) follows from (5.40). As a consequence of Lemma 5.5 we have

LEMMA 5.6.— *Let the assumptions of Corollary 3.2 be satisfied. Then we have*

$$\lim_{n \to \infty} J(\hat{D}_n \tilde{D}_n^{-1}) = J(\hat{D}), \tag{5.44}$$

$$\lim_{n \to \infty} \left(J(D_n) - J(D^n_{\infty^{(1)}} D^{-n}_{\infty^{(2)}}) \right) = J(\hat{D}) - J\left(D^{k_0}_{\infty^{(1)}} D^{-k_0}_{\infty^{(2)}} \right). \tag{5.45}$$

The convergence is understood in the topology on $\mathrm{Jac}(\mathcal{R})$ *induced from* $\mathbb{C}^g$.

Proof.— From (5.42) and (5.41) it follows that

$$\hat{D}_n \tilde{D}_n^{-1} = D^g_{\infty^{(2)}} \prod_{z \in \widehat{ZS}_n \setminus \{z_{1n}, \ldots, z_{ng}\}} D_z \prod_{\pi \in \widehat{PS}} D_\pi^{-1}. \tag{5.46}$$

The convergence (5.38) in Lemma 5.5 together with (4.7) and the definition of the mapping J in (5.19) gives (5.44). The limit (5.45) is a consequence of (5.44) together with the first line of (5.43). □

With Proposition 5.3 and the Lemmas 5.5 and 5.6 we have a rather good understanding about the asymptotic distribution of the zeros of the remainder function R_n as $n \to \infty$. With some simplifications it can be said that the distribution is understood if one understands the asymptotic distribution of the set

$$\left\{ J\left(D^n_{\infty^{(1)}} D^{-n}_{\infty^{(2)}}\right) \in \mathrm{Jac}(\mathcal{R}) \mid n = 1, 2, \ldots \right\} \tag{5.47}$$

in $\mathrm{Jac}(\mathcal{R})$.

With respect to the distribution of the set (5.47) in $\mathrm{Jac}(\mathcal{R})$ the analysis in the present paper is not optimal. This is also the reason why in the Theorems 3.3 and 3.8 it had to be permitted that the point of development for the Padé approximants $[n/n]$ is moved away from infinity in certain cases. For Theorem 3.8 this seems to be an unnecessary requirement. In the present paper the analysis of the distribution of the sequence of numbers in the set (5.47) is based solely on Weyl's Uniform Distribution Theorem, which is applicable only under certain assumptions.

LEMMA 5.7.— *Let* $\gamma_j = \gamma_j\big(J(D_{\infty^{(1)}} D^{-1}_{\infty^{(2)}})\big) \in [0,1)$, $j = 1,\ldots,2g$, *be the coordinates of the point* $J\big(D_{\infty^{(1)}} D^{-1}_{\infty^{(2)}}\big) \in \mathrm{Jac}(\mathcal{R})$ *as defined in (5.15), and assume that the elements of the set* $\{1, \gamma_1, \ldots, \gamma_{2g}\}$ *are linearly independent over* $\mathbb{Q}$. *Then the set (5.47) is dense in* $\mathrm{Jac}(\mathcal{R})$.

Proof.— The lemma is a direct consequence of Weyl's Uniform Distribution Theorem (cf. [Ch, chap. VIII]). Let us consider the linear mapping $\Psi : \mathrm{Jac}(\mathcal{R}) \longrightarrow \mathbb{R}^{2g}/\tilde{\Gamma}$ with $\tilde{\Gamma} := \mathbb{Z}e_1 + \ldots + \mathbb{Z}e_{2g}$, e_j the canonical unit vectors $e_j = (0,\ldots,1,\ldots,0)$, and the mapping Ψ defined by $\Psi(c) = \big(\gamma_1(c),\ldots,\gamma_{2n}(c)\big) \bmod (1,\ldots,1)$ for $c \in \mathrm{Jac}(\mathcal{R})$ and the $\gamma_j(c)$ the coordinates of c introduced before (5.15). Thus, we have $\Psi(J\big(D_{\infty^{(1)}} D^{-1}_{\infty^{(2)}}\big)) \equiv (\gamma_1,\ldots,\gamma_{2g}) \bmod (1,\ldots,1)$. Since it has been assumed that the numbers $1, \gamma_1, \ldots, \gamma_{2g}$ are rationally independent, we know from Weyl's uniform distribution Theorem that the set

$$\big\{n(\gamma_1,\ldots,\gamma_{2g}) \quad \mathrm{mod}\,(1,\ldots,1) \mid n = 1,2,\ldots\big\} \tag{5.48}$$

is uniformly distributed in $\mathbb{R}^{2g}/\tilde{\Gamma}$, and therefore also dense in $\mathbb{R}^{2g}/\tilde{\Gamma}$. The mapping Ψ is a continuous bijection. Hence, the denseness of the set (5.48) in $\mathbb{R}^{2g}/\tilde{\Gamma}$ implies the denseness of the set (5.47) in $\mathrm{Jac}(\mathcal{R})$. □

The assumption of Lemma 5.7 may not be satisfied for the Riemann surface $\mathcal{R}$ that is associated with a given hyperelliptic function f. However, with the help of the next lemma it is always possible to choose a new point of development ζ_0 near infinity for the Padé approximants $[n/n]$, $n \in \mathbb{N}$, such that the assumption of Lemma 5.7 is satisfied in the new situation.

LEMMA 5.8.— *For* $\zeta \in \overline{\mathbb{C}}$ *let* $z^{(j)} \in B_j \subseteq \mathcal{R}$, $j = 1, 2$, *denote the two points on* $\mathcal{R}$ *lying over* ζ, *i.e.,* $\pi(z^{(1)}) = \pi(z^{(2)}) = \zeta$, *and define the function* $c(\zeta)$ *with values in* $\mathrm{Jac}(\mathcal{R})$ *by*

$$c(\zeta) := \left(\int_{z^{(2)}}^{z^{(1)}} \varphi_1, \ldots, \int_{z^{(2)}}^{z^{(1)}} \varphi_g\right) \quad \mathrm{mod}\ \mathrm{Per}(\varphi_1,\ldots,\varphi_g). \tag{5.49}$$

Let $\gamma_j = \gamma_j(\zeta) := \gamma_j(c(\zeta))$, $j = 1, \ldots, 2g$, denote the coordinates of $c(\zeta)$ as defined before (5.15). Then for almost every $\zeta \in \overline{\mathbb{C}}$ in a neighborhood of ∞ (almost every with respect to planar Lebesgue measure) the elements of the set $\{1, \gamma_1(\zeta), \ldots, \gamma_2(\zeta)\}$ are linearly independent over $\mathbb{Q}$.

Proof.— Let $c = (c_1, \ldots, c_g) \in \mathbb{C}^g$. From (5.14) and (5.15) it follows that there exist $2g$ vectors $b_j = (b_{j1}, \ldots, b_{jg}) \in \mathbb{C}^g$, $j = 1, \ldots, 2g$, such that the coordinates $\gamma_j = \gamma_j(c) \in [0, 1)$ are given by

$$\gamma_j \equiv \operatorname{Re} \sum_{i=1}^{g} b_{ji} c_i \mod(1)\ j = 1, \ldots, 2g. \tag{5.50}$$

Hence, for the coordinates $\gamma_j(\zeta)$, $j = 1, \ldots, 2g$, of the function $c(\zeta)$ defined in (5.49) we have

$$\begin{aligned} &\gamma_j(\zeta) = \gamma_j\left(c(\zeta)\right) = \operatorname{Re} \sum_{i=1}^{g} b_{ji} \int_{z^{(2)}}^{z^{(1)}} \varphi_i = \operatorname{Re} \int_{z^{(2)}}^{z^{(1)}} \tilde{\varphi}_j \mod(1), \\ &\tilde{\varphi}_j := \sum_{i=1}^{g} b_{ji} \varphi_i. \end{aligned} \tag{5.51}$$

Since the vectors $\Gamma_1, \ldots, \Gamma_{2g}$ in (5.14) form a basis in $\mathbb{C}^g$ over $\mathbb{R}$, it follows from (5.50) that also the vectors $b_1, \ldots, b_{2g}$ form a basis in $\mathbb{C}^g$ over $\mathbb{R}$, and consequently none of the Abelian differentials $\tilde{\varphi}_1, \ldots, \tilde{\varphi}_{2g}$ and also no non-trivial linear combination of these differentials can vanish identically. Consequently, the same holds true for the coordinate functions $\gamma_1(\zeta), \ldots, \gamma_{2g}(\zeta)$, and these functions are harmonic in ζ near infinity. Let m denote the planar Lebesgue measure in the set $\Delta := \{\zeta \in \overline{\mathbb{C}} \mid |\zeta| \geq 1/\varepsilon\}$, $\varepsilon > 0$. We consider

$$\begin{aligned} G_1 &:= \{\zeta \in \Delta \mid \gamma_1(\zeta) \in \mathbb{Q}\} \\ &= \{\zeta \in \Delta \mid \lambda\{1, \gamma_1(\zeta)\} \text{ linearly dependent over } \mathbb{Q}\}, \end{aligned}$$

and generally

$$G_j := \{\zeta \in \Delta \mid \{1, \gamma_1(\zeta), \ldots, \gamma_j(\zeta)\} \text{ linearly dependent over } \mathbb{Q}\}, \quad j = 1, \ldots, g.$$

It is immediate that $G_j \subseteq G_{j+1}$ for $j = 1, \ldots, g-1$. The set $\{\zeta \in \Delta \mid \gamma_1(\zeta) = r\}$, $r \in \mathbb{Q}$, consists of level lines of a non-constant harmonic function. Hence, its Lebesgue measure is zero, and further $m(G_1) = 0$. All $\zeta \in G_{j+1} \setminus G_j$ lie on level lines defined by equations of the form

$$\gamma_{j+1}(\zeta) = (r_1\gamma_1 - \ldots - r_j\gamma_j)(\zeta), \quad r_1, \ldots r_j \in \mathbb{Q}, \tag{5.52}$$

and consequently also $m(G_{j+1} \setminus G_j) = 0$ for $j = 1, \ldots, g-1$. This proves that $m(G_g) = 0$. □

From the definition of the mapping J in (5.19) and from (5.49) we see that

$$J(D_{\infty^{(1)}} D^{-1}_{\infty^{(2)}}) = c(\infty). \tag{5.53}$$

If for a given hyperelliptic function f the associated Riemann surface $\mathcal{R}$ does not have the property that the elements of the set $\{1, \gamma_1, \ldots, \gamma_{2g}\} = \{1, \gamma_1(\infty), \ldots, \gamma_{2g}(\infty)\}$ are linearly independent over $\mathbb{Q}$, then Lemma 5.7 cannot be applied. However, we can use another point of development for the Padé approximants $[n/n]$. In Definition 3.1 diagonal Padé approximants $[n/n]$ developed at a point $\zeta_0 \neq \infty$ have been introduced by considering the function $\tilde{f} := f \circ \psi^{-1}$ with a Moebius transform

$$\psi(\zeta) := \frac{1}{1 - \zeta_0^{-1}\zeta} \tag{5.54}$$

mapping ζ_0 onto ∞. If $\widetilde{[n/n]}$ denotes the Padé approximant of $\tilde{f}$ developed at infinity, then from Definition 3.1, (3.5), we know that

$$[n/n] := \widetilde{[n/n]} \circ \psi \tag{5.55}$$

is the Padé approximant of f developed at ζ_0 (f is assumed to be analytic at ζ_0). Of course, $\tilde{f}$ is again an hyperelliptic function; it has branch points $\tilde{a}_j := \psi^{-1}(a_j)$, $j = 1, \ldots, 2m$. The Riemann surface $\tilde{\mathcal{R}}$ associated with $\tilde{f}$ is conformally equivalent to $\mathcal{R}$, and it is not difficult to see how all notions associated with $\mathcal{R}$ are caried over to those associated with $\tilde{\mathcal{R}}$, and vice versa. Thus, the convergence domain D of the sequence $\{[n/n]\}$ is equal to $\psi^{-1}(\tilde{D})$, where $\tilde{D}$ is the convergence domain of the sequence $\{\widetilde{[n/n]}\}$. Let $\tilde{\mathbf{a}}_1, \ldots, \tilde{\mathbf{a}}_g$, $\tilde{\mathbf{b}}_1, \ldots, \tilde{\mathbf{b}}_g$ be a homology basis on $\tilde{\mathcal{R}}$ and $\tilde{\varphi}_1, \ldots, \tilde{\varphi}_g$ the canonical basis of the Abelian differentials of the first kind on $\tilde{\mathcal{R}}$. It follows from the conformal equivalence of $\mathcal{R}$ and $\tilde{\mathcal{R}}$ that

$$\int_{z_0^{(2)}}^{z_0^{(1)}} \varphi_j = \int_{\tilde{\infty}^{(2)}}^{\tilde{\infty}^{(1)}} \tilde{\varphi}_j, \quad j = 1, \ldots, g, \tag{5.56}$$

Note that $\psi(\zeta_0) = \infty$, and $z_0^{(1)}$ and $z_0^{(2)}$ lie on $\mathcal{R}$ over the same basis point ζ_0.

Identity (5.56) together with Lemma 3.8 implies that we always can assume that the assumptions of Lemma 5.7 hold true, if we are willing to move the point of development for the Padé approximants slightly away from infinity. Indeed, from Lemma 5.8 we know that if the assumptions of Lemma 5.7 are not satisfied, then we can choose a point ζ_0 near infinity as a new point of development for the Padé approximants $[n/n]$, $n \in \mathbb{N}$, such that the elements of the set $\{1, \gamma_1(\zeta_0), \dots, \gamma_{2g}(\zeta_0)\} = \{1, \tilde{\gamma}_1(\infty), \dots, \tilde{\gamma}_{2g}(\infty)\}$ are linearly independent over $\mathbb{Q}$. The analysis is then continued with the sequence $\{\widetilde{[n/n]}\}$ of Padé approximants to $\tilde{f}$, which are developed at infinity and for which the assumptions of Lemma 5.7 hold true.

So far we have only considered zeros (and poles) of the remainder function R_n. However, we are interested in spurious poles of the Padé approximants $[n/n]$, $n \in \mathbb{N}$. The next lemma provides the necessary connections.

LEMMA 5.9.— *Let g points $z_1, \dots, z_g \in \mathcal{R} \setminus \{\infty^{(2)}\}$ be given and assume that N is an infinite subsequence of $\mathbb{N}$ for which the Lemmas 5.1 through 5.5 hold true and assume that for each $n \in N$ g points $z_{1n}, \dots, z_{gn}$ have been selected from $\mathcal{R}$ in the same way as in Lemma 5.5 and that they satisfy*

$$z_{jn} \longrightarrow z_j \quad \text{as} \quad n \longrightarrow \infty, \ n \in N, \ j = 1, \dots, g, \tag{5.57}$$

where the convergence has to hold in the spherical metric if for some z_j we have $z_j = \infty^{(1)}$.

If for k $(0 \leq k \leq g)$ of the g points $z_1, \dots, z_g$ we have

$$z_1, \dots, z_k \in B_2 \quad \text{and} \quad z_j \in B_1 \setminus \{\varphi(z_1), \dots, \varphi(z_k)\}, \quad j = k+1, \dots, g, \tag{5.58}$$

then there exists an infinite subsequence of N, which we continue to denote by N, such that each Padé approximant $[n/n]$, $n \in N$, has excactly k spurious poles $\pi_{1n}, \dots, \pi_{kn} \in \mathbb{C}$ satisfying

$$\pi_{jn} \longrightarrow \pi(z_j) \quad \text{as } n \longrightarrow \infty, \ n \in N, \ j = 1, \dots, k. \tag{5.59}$$

Remark.— In (5.58) φ is the covering transformation on $\mathcal{R}$ and in (5.59) π denotes the canonical projection.

Proof

(a) We start with a preparatory consideration. Assume that $N \subseteq \mathbb{N}$ is an infinite subsequence such that for each $n \in N$ there exist k_2 spurious poles $\pi_{1n}, \ldots, \pi_{k_2 n} \in \mathbb{C}$ of the Padé approximant $[n/n]$ with

$$\pi_{jn} \longrightarrow \pi_1 \in D \quad \text{as } n \longrightarrow \infty,\ n \in N,\ j = 1, \ldots, k_2, \tag{5.60}$$

(D denotes the convergence domain). Let $k_3 \geq 0$ and $k_4 \geq 0$ be the order of the poles of the function f at the points $\pi_1^{-1}(\pi_1)$ and $\pi_2^{-1}(\pi_1)$, respectively, (the case $k_3 = 0$ or $k_4 = 0$ has not been excluded). From Theorem 3.1 (i) we know that the denominator polynomial Q_n, $n \in N$, has at least $k_2 + k_3$ zeros in a neighborhood of π_1 and these zeros converge to π_1 as $n \longrightarrow \infty$, $n \in N$. Starting as in the proof of Lemma 5.5 from equation (5.33), it is possible to show by Rouché's Theorem that the set ZS_n has k_2 elements more than the set CS_n in a neighborhood of $\pi_2^{-1}(\pi_1) \in B_2$ (both sets have been defined in (5.26)), and therefore these k_2 points have to be among those points of ZS_n that are selected in Lemma 5.5, for otherwise the convergence (5.38) would not be possible.

(b) We now start from the assumptions made in the lemma. We assume that k_5 limit points of the list $z_1, \ldots, z_g$ are identical with $z_1 \in B_2$. Without loss of generality let these be the points $z_1 = \ldots = z_{k_5} \in B_2$. Then (5.27) implies that

$$z_{jn} \longrightarrow z_1 \quad \text{as } n \longrightarrow \infty,\ n \in N \ \text{ for } \ j = 1, \ldots, k_5. \tag{5.61}$$

Let f have a pole of order $k_6 \geq 0$ at $\varphi(z_1) \in B_1$ ($k_6 = 0$ is not excluded). Then it follows from the definition of the sets ZS_n and CS_n in (5.26), the convergence (5.38) in Lemma 5.5, and an application of Rouché's Theorem to equation (5.33) as has been done in the proof of Lemma 5.5 that the denominator polynomial Q_n, $n \in N$, has $k_5 + k_6$ zeros in a neighborhood of $\pi(z_1)$ and these zeros converge to $\pi(z_1)$ as $n \longrightarrow \infty$, $n \in N$. From Theorem 3.1 (i) we know that k_6 zeros do not coincide with zeros of the numerator polynomial P_n, since the corresponding poles of $[n/n]$ converge to the pole of f at $\pi_1^{-1}(z_1)$. Let us denote the remaining k_5 zeros of Q_n near $\pi(z_1)$ by ζ_{jn}, $j = 1, \ldots, k_5$, $n \in N$. We have

$$\zeta_{jn} - \pi(z_1) \in D \quad \text{as } \ n \longrightarrow \infty,\ n \in N \ \text{ for } \ j = 1, \ldots, k_5. \tag{5.62}$$

In order to show that these zeros ζ_{jn} generate poles of $[n/n]$, which then are spurious, it is necessary to prove that these zeros are not zeros of the numerator polynomial P_n simultaneously.

Indeed, let us assume on the contrary that there exists an infinite subsequence $\tilde{N} \subseteq N$ such that for each $n \in \tilde{N}$ there exists at least one zero ζ_{jn}, $j \in \{1, \dots k_5\}$, which is simultaneously a zero of Q_n and P_n. It follows from (4.8) (and (1.2)) that in this case a linear factor can be factored out of P_n and Q_n without making (1.2) invalid. This implies that both points $\pi_1^{-1}(\zeta_{jn})$ and $\pi_2^{-1}(\zeta_{jn})$ are zeros of R_n and belong to the g points that have to be selected in Lemma 5.5. From the convergence (5.62) it then further follows that $z_1 \in B_2$ and $\varphi(z_1) \in B_1$ are limit points of $\pi_2^{-1}(\zeta_{jn})$ and $\pi_1^{-1}(\zeta_{jn})$ as $n \to \infty$, $n \in \tilde{N}$, and hence both points z_1 and $\varphi(z_1)$ belong to the list $z_1, \dots, z_g$. But this contradicts assumption (5.58). Hence, we know that all k_5 zeros ζ_{jn}, $j = 1, \dots, k_5$, of Q_n generate spurious poles of $[n/n]$ for n in a subsequence N that differs from the original one by only by finitely many elements.

The analysis so far has shown that there exist at least k spurious poles if k limit points out of the list $z_1, \dots, z_g$ lie in B_2. From part (a) we deduce that on the other hand there cannot exist more than k spurious poles. The limit (5.59) follows from (5.62) with $\zeta_{jn} = \pi_{jn}$ for $j = 1, \dots, k$ and $n \in N$. □

Proof of Theorem 3.3 .— We first assume that the $2g+1$ numbers $1, \gamma_1, \dots, \gamma_{2g}$ are linearly independent over $\mathbb{Q}$, where γ_j are the coordinates of the number $J\left(D_{\infty^{(1)}} D_{\infty^{(2)}}^{-1}\right) \in \mathrm{Jac}(\mathcal{R})$ introduced in Lemma 5.7. Further we know that the assumptions of Corollary 3.2 are satisfies, hence the lemmas 5.5, 5.6, and 5.9 can be used.

Define $z_j := \pi_2^{-1}(\pi_j) \in B_2$, $j = 1, \dots, g = m-1$, and let $G_1, \dots, G_g$ be open neighborhoods of these points which $\overline{G_j} \subseteq B_2$, $j = 1, \dots, g$. Since all $\overline{G_j}$ are contained in B_2, assumption (5.20) of Proposition 5.3 is satisfied, and therefore a non-empty open subset $G_0 \subseteq \mathrm{Jac}(\mathcal{R})$ exists with the properties stated in Proposition 5.3.

Because of the assumption of the rational independence of $1, \gamma_1, \dots, \gamma_{2g}$ we know from Lemma 5.7 that the set (5.47) is dense in $\mathrm{Jac}(\mathcal{R})$, and we can deduce from limit (5.45) in Lemma 5.6 that there exists an infinite subsequence $N \subseteq \mathbb{N}$ such that

$$J(D_n) \in G_0 \quad \text{for} \quad n \in N, \tag{5.63}$$

with divisors D_n defined as in (5.43).

Let $z_{1n}, \ldots, z_{gn} \in \mathcal{R}$, $n \in N$, be the g points selected in Lemma 5.5 such that convergence (5.38) holds true. Note that the divisor D_n is defined in (5.43) analogously to the divisor D in (5.21) of Proposition 5.3. The π_j's in D are now the elements of PS_n and the $z_{g+1}, \ldots, z_l$ the elements of $ZS_n \setminus \{z_{1n}, \ldots, z_{gn}\}$; (5.40) and (5.43) together then show the analogy between the definition of D_n and D. From Proposition 5.1 (a), we know that there exist g points $\tilde{z}_{1n}, \ldots, \tilde{z}_{gn} \in \mathcal{R}$ such that the two sets PS_n and $\{\tilde{z}_{1n}, \ldots, \tilde{z}_{gn}\} \cup \left(ZS_n \setminus \{z_{1n}, \ldots, z_{gn}\}\right)$ are pole- and zero-sets of a meromorphic function on $\mathcal{R}$. Because of (5.63) from Proposition 5.3 it follows that the selection of the g points $\tilde{z}_{1n}, \ldots, \tilde{z}_{gn}$ is unique. Since we know from (5.40) that the two sets PS_n and ZS_n are pole- and zero-sets (possibly plus some additional points that appear in PS_n and ZS_n simultaneously) of the meromorphic function R_n, as a first conclusion it follows that $z_{jn} = \tilde{z}_{jn}$, $j = 1, \ldots, g$, $n \in N$, is a possible choice in Proposition 5.1 (a). Because of the uniqueness it then follows as a second conclusion that this is the only choice possible. From (5.63) and (5.23) in Proposition 5.3 we then know that

$$z_{jn} = \tilde{z}_{jn} \in G_j \quad \text{for } j = 1, \ldots, g,\ n \in N. \tag{5.64}$$

If we repeat this consideration with shrinking neighborhoods G_j of the points z_j, then we can construct an infinite subsequence of $\mathbb{N}$, which we continue to denote by N, such that the selected sets $\{z_{1n}, \ldots, z_{gn}\}$, $n \in N$, in Lemma 5.5 satisfy

$$z_{jn} \longrightarrow z_j \quad \text{as } n \to \infty,\ n \in N \ \text{ for } j = 1, \ldots, g. \tag{5.65}$$

With Lemma 5.9 it then follows that each Padé approximant $[n/n]$, $n \in N$, has g spurious poles at $\pi_{jn} \in \mathbb{C}$, $j = 1, \ldots, g$, $n \in N$, and

$$\pi_{jn} \longrightarrow \pi(z_j) = \pi_j \quad \text{as } n \to \infty,\ n \in N \ \text{ for } j = 1, \ldots, g. \tag{5.66}$$

This proves the Theorem under the assumption that the $2g + 1$ numbers $1, \gamma_1, \ldots, \gamma_{2g}$ are linearly independent over $\mathbb{Q}$.

We know from Lemma 5.8 that if the linear independence over $\mathbb{Q}$ is not true, then we can choose almost any point ζ_0 in a neighborhood of infinity as a new point of development for the diagonal Padé approximants $[n/n]$ to f, and the diagonal Padé approximants $\widetilde{[n/n]}$ to the function $\tilde{f} := f \circ \psi^{-1}$ developed at infinity with ψ the Moebius transform (5.54) have the property that the points $1, \tilde{\gamma}_1, \ldots, \tilde{\gamma}_{2g}$ are rationally independent with $\tilde{\gamma}_j$ the coordinates $\tilde{\gamma}_j = \tilde{\gamma}_j(\infty) = \gamma_j(\zeta_0)$, $j = 1, \ldots, 2g$, that have been considered in the paragraph after (5.56). Hence, for the sequence $\{\widetilde{[n/n]}\}$ all the conclusions proved above hold true. In order to prove Theorem 3.3, one has to select a new point of development ζ_0 and to continue with the points $\tilde{\pi}_j := \psi(\pi_j)$, $j = 1, \ldots, g$, instead of the original points π_j. The results then proved for the Padé approximants $\widetilde{[n/n]}$ are pulled back to $[n/n]$ with the help of identity (5.55). □

Proof of Theorem 3.8 .— From a methodological point of view the proof is very similar to that of Theorem 3.3. We start again by assuming that the numbers $1, \gamma_1, \ldots, \gamma_{2g}$ are rationally independent and that the assumptions of Corollary 3.2 are satisfied. Next we choose g points $\zeta_1, \ldots, \zeta_g$ arbitrarily in the convergence domain $D \subseteq \overline{\mathbb{C}}$, but then we define $z_j := \pi_1^{-1}(\zeta_j)$, $j = 1, \ldots, g$, *i.e.*, the z_j are now contained in B_1. Let G_j be a neighborhood of z_j with $\overline{G_j} \subseteq B_1$, $j = 1, \ldots, g$. The important difference from the last proof is that the $\overline{G_j}$ and z_j are contained in B_1 and not in B_2. After this assumption the analysis is identical with that in the proof of Theorem 3.3, except that it now follows from Lemma 5.9 after (5.65) that each Padé approximant $[n/n]$, $n \in N$, has no spurious poles since all $z_1, \ldots, z_g \in B_1$. Since no spurious poles exist, a consideration of the limits in (5.66) is not necessary.

If the $2g + 1$ numbers $1, \gamma_1, \ldots, \gamma_{2g}$ are not rationally independent, then with the same arguments as applied in the proof of Theorem 3.3 it can be shown that the analysis can be repeated with a new point of development $\zeta_0 \in \mathbb{C}$.

We note that for the proof of Theorem 3.8 the limit (5.65) is not really necessary. It would be enough to know that the zeros $z_{1n}, \ldots, z_{gn}$ lie in B_1. This weaker condition would allow us to select a denser subsequence $N \subseteq \mathbb{N}$. However, in order to follow this path of proof the technical Lemma 5.9 would have to be formulated and proved in a different way. □

The section is closed by a discussion whether the assumptions made in Corollary 3.2 and also the special assumptions made in the Theorems 3.3 and 3.8 with respect to the point of development of the Padé approximants $[n/n]$ are really necessary. Actually, it seems that in Theorem 3.8 it should be possible to drop both assumptions completely. Further there is the question whether all or some results can be extended to a larger class of functions. A short discussion of these questions is organized in 4 subsections:

(i) There is some evidence that poles of the function f or poles and zeros of the rational function r_2 in the representation (4.5), which lie on $F = \overline{\mathbb{C}}/D$, play a different role if they lie on end points of the arcs J_j, $j \in I$, that form F. These endpoints usually are branch points of the function f. In the case of only two branch points the phenomenon can be illustrated by the behavior of the Jacobi polynomials $P_n^{(\alpha,\beta)}$ with parameters $\alpha = n_1 + 1/2$, $\beta = n_2 + 1/2$, n_1, $n_2 \in \mathbb{N}$. Padé approximants having these polynomials as denominators are free of spurious poles for all values $n_1, n_2 \in \mathbb{N}$.

(ii) If poles of the function f or zeros and poles of the rational function r_2 lie inside of the arcs J_j, $j \in I$, that form the set F, then these points should be considered as confluent pairs of branch points of f. Geometrically speaking, the Riemann surface $\mathcal{R}$ of the function f is considered as the limit situation of Riemann surfaces with a larger genus. Combining these considerations with that of subsection (i), it should be possible to find a sharp upper bound (3.3) in Theorem 3.1. If it is also possible to understand in this framework the rules that govern the distribution of spurious poles, then it should further be possible to prove the Baker–Gammel–Wills conjecture in the general form of Theorem 3.8 without the assumptions of Corollary 3.2.

(iii) The necessity to move to a point of development $\zeta_0 \neq \infty$ in Theorems 3.3 and 3.8 follows from the weakness of our analysis with respect to the distribution of the set (5.47) in the Jacobian variety $\mathrm{Jac}(\mathcal{R})$. However, it seems that this part of the analysis can be improved, and as a consequence Theorem 3.8 may hold true without excluding certain points of developments.

(iv) It is perhaps rather brave to conjecture that the results of Section 3 can be extended to algebraic functions f, *i.e.*, that a formula for a sharp bound (3.3) in Theorem 3.1 can be found and proved for all algebraic functions; further, that the rules which govern the distribution of spurious poles are understood; and, lastly, that the Baker–Gammel–Wills conjecture in the form of Theorem 3.8 holds true for all algebraic functions f. (The assumption that the function f should be analytic at infinity is of minor importance.) Of course so far it is not clear whether the results hold true in such a general form. We note that it has been conjectured by J. Nuttall that a finite bound (3.3) exists for algebraic functions f, and that in most situations this bound should be equal to genus of f.

6. A Result by S. Dumas Revisited

In his thesis [Du] Samuel Dumas investigated the development in continued fractions of the square root of a polynomial of fourth or third order with complex roots. He classified different situations, gave explicit formulae for the partial denominators of the continued fractions in terms of elliptic functions, and studied the convergence behavior of continued fractions. There, he cast light on a surprisingly strange behavior of the convergents of the continued fractions in certain situations. It is this last topic with which we are concerned in the present section. We consider Padé approximants instead of continued fractions and restrict ourself to developments at infinity. The analysis can be seen as an illustration of the more general results in Section 3 for the more explicit situation of a special elliptic function.

Let $a_1, \ldots, a_4 \in \mathbb{C}$ be 4 different points and define the function f by

$$f(z) := \sqrt{(z-a_1)\cdots(z-a_4)} - z^2 + z(a_1 + \cdots + a_4)/2. \tag{6.1}$$

The function is analytic at infinity. Let $[n/n]$, $n = 1, 2, \ldots$, be the diagonal Padé approximants to f developed at infinity. By $\mathcal{R}$ we denote the concrete Riemann surface defined by

$$y^2 = (z-a_1)\cdots(z-a_4) =: P(z). \tag{6.2}$$

As before $\pi : \mathcal{R} \longrightarrow \overline{\mathbb{C}}$ denotes the canonical projection. Let the two closed curves $\mathbf{a}$ and $\mathbf{b}$ form a homology basis on $\mathcal{R}$ and define

$$A := \oint_{\mathbf{a}} \frac{\mathrm{d}t}{\sqrt{P(t)}}, \quad \tau := \frac{1}{A}\oint_{\mathbf{b}} \frac{\mathrm{d}t}{\sqrt{P(t)}}. \tag{6.3}$$

We have $\tau \notin \mathbb{R}$, and it can be assumed that $\mathrm{Im}(\tau) > 0$. The set

$$\Gamma := \mathbb{Z} + \mathbb{Z}\tau \tag{6.4}$$

forms a lattice in $\mathbb{C}$. With some knowledge of elliptic functions, it is perhaps not surprising that the mapping

$$\begin{aligned} &u : \mathcal{R} \longrightarrow \overline{\mathbb{C}}/\Gamma \\ &z \longmapsto u(z) :\equiv \frac{1}{A}\int_{\infty^{(1)}}^{z} \frac{\mathrm{d}t}{\sqrt{P(t)}} \quad \mathrm{mod}\ (1, \tau) \end{aligned} \tag{6.5}$$

turns out to be a very helpful tool in the analysis. Because of (6.3) the right-hand side of (6.5) is well defined. The mapping u is bijective and the two surfaces $\mathcal{R}$ and $\overline{\mathbb{C}}/\Gamma$ are conformly equivalent. We have $u(\infty^{(1)}) = 0 \in \overline{\mathbb{C}}/\Gamma$. The image of $\infty^{(2)} \in \mathcal{R}$ under the mapping of u, *i.e.*,

$$v_\infty := u(\infty^{(2)}) \equiv \frac{1}{A}\int_{\infty^{(1)}}^{\infty^{(2)}} \frac{dt}{\sqrt{P(t)}} \mod (1,\tau) \tag{6.6}$$

plays an important role in the analysis. It turns out that its arithmetic character is decisive for the convergence behavior of the diagonal Padé approximants $[n/n]$, $n \in \mathbb{N}$.

As in (1.2) and (4.8) we have

$$R_n(z) = (fQ_n \circ \pi - P_n \circ \pi)(z) = \begin{cases} O(z^{-n-1}) & \text{as } z \to \infty^{(1)} \\ O(z^{n+2}) & \text{as } z \to \infty^{(2)}. \end{cases} \tag{6.7}$$

The functions R_n and f are analytic on $\mathcal{R} \setminus \{\infty^{(2)}\}$, and as before the Padé polynomials P_n, $Q_n \in \mathcal{P}_n$ are considered as functions defined on $\overline{\mathbb{C}}$. Since the function R_n has no poles outside of $\infty^{(2)}$, from (6.7) and the fact that a meromorphic function on a compact Riemann surface has an identical number of poles and zeros, it follows that the location of all poles and zeros of R_n is known except for one zero.

LEMMA 6.1.— *For each $n \in \mathbb{N}$ there exists a point $z_n \in \mathcal{R}$ such that one of the following three cases holds true:*

(i) $R_n(z_n) = 0$ *if* $z_n \in \mathcal{R} \setminus \{\infty^{(1)}, \infty^{(2)}\}$,

(ii) $R_n(z) = O(z^{-n-2})$ *as* $z \to \infty^{(1)}$ *if* $z_n = \infty^{(1)}$, *or*

(iii) $R_n(z) = O(z^{n+1})$ *as* $z \to \infty^{(2)}$ *if* $z_n = \infty^{(2)}$.

In any case $R_n(z) \neq 0$ *for* $z \in \mathcal{R} \setminus \{\infty^{(1)}, z_n\}$.

In case of an elliptic function Abel's Theorem tells us that the sum of the poles and the sum of the zeros in the v-plane are equal up to multiples of the periods 1 and τ. According to (6.7) the function $\tilde{R}_n(v) := R_n(u^{-1}(v))$ has $n+2$ poles at $v_\infty \in \overline{\mathbb{C}}/\Gamma$, $n+1$ zeros at $0 \in \overline{\mathbb{C}}/\Gamma$, and one zero at $u(z_n) \in \overline{\mathbb{C}}/\Gamma$. Thus, the next lemma follows directly from Abel's Theorem.

LEMMA 6.2.— *In* $\overline{\mathbb{C}}/\Gamma$ *we have*

$$u(z_n) = (n+2)v_\infty \quad \textit{for} \quad n \in \mathbb{N}. \tag{6.8}$$

In the analysis we need some tools from the Weierstrass approach to the theory of elliptic functions, particulary the Weierstrass functions $\wp$, ζ, and σ. As a general reference we use [HuCo] or [Hi], but any other book on elliptic functions can be used. As usual, $\wp$ denotes the Weierstrass $\wp$-function, which is a doubly-periodic function, defined in the v-plane $\mathbb{C}$ with periods $(1, \tau)$, and having a double pole at the origin and at all equivalent points. The function $\zeta(v)$ is meromorphic in the v-plane, and defined by $\zeta'(v) = -\wp(v)$ up to an additive constant, which can be determined by the functional equation $\zeta(-v) = -\zeta(v)$. The function ζ is not periodic, but it satisfies the functional equation $\zeta(v + n_1 + n_2\tau) = \zeta(v) + n_1\eta_1 + n_2\eta_2$ for all $v \in \mathbb{C}$ and $n_1, n_2 \in \mathbb{Z}$ with two constants $\eta_1, \eta_2 \in \mathbb{C}$ that satisfy the Legendre relation $\eta_1 + \eta_2\tau = 2\pi i$ (cf. [HuCo, II, 1, S6]). Finally, the σ-function is an entire function defined by

$$\left(\log \sigma(v)\right)' = \frac{\sigma'(v)}{\sigma(v)} = \zeta(v) \tag{6.9}$$

and satisfying the functional equation

$$\begin{aligned} &\sigma(v + n_1 + n_2\tau) = \\ &\quad = (-1)^{n_1+n_2+n_1n_2} \exp\left[(n_1\eta_1 + n_2\eta_2)\big(v + (n_1 + n_2\tau)/2\big)\right] \sigma(v) \end{aligned} \tag{6.10}$$

for $v \in \mathbb{C}$ and $n_1, n_2 \in \mathbb{Z}$ (cf. [HuCo, II, 1, S9]). The Weierstrass σ-function is very useful for representing doubly periodic meromorphic functions, *i.e.*, meromorphic functions defined on $\overline{\mathbb{C}}/\Gamma$. The function σ has a simple zero at the origin and at all equivalent points in the v-plane. In the next lemma we show that with the help of the σ-function it is possible to give a constructive representation of the function Φ defined in (4.2). We recall that the function Φ has played a dominant role in the proofs of most results in Section 3. If the function Φ is defined in a constructive way, then this also implies a constructive procedure for determining the convergence domain D and the convergence factor G_D (defined in (2.3)),

LEMMA 6.3

(i) There exists a constant $c \in \mathbb{C}$ such that the function

$$\tilde{G}(v) := e^{c(v - v_\infty/2)} \frac{\sigma(v)}{\sigma(v - v_\infty)}, \quad v \in \mathbb{C}, \tag{6.11}$$

satisfies the functional equation

$$\left|\tilde{G}(v + n_1 + n_2\tau)\right| = \left|\tilde{G}(v)\right| \quad \textit{for all} \;\; v \in \mathbb{C},\; n_1, n_2 \in \mathbb{Z}, \tag{6.12}$$

and there exists $c_0 \in \mathbb{C}$ *with* $|c_0| = 1$ *such that*

$$\Phi(z) = c_0 \tilde{G}(u(z)) \quad \textit{for} \quad z \in \mathcal{R}. \tag{6.13}$$

(ii) Let $B_1,\ B_2 \subseteq \mathcal{R}$ *be the two domains lying over the convergence domain* $D \subseteq \overline{\mathbb{C}}$. *Then we have*

$$\begin{aligned} B_1 &= \left\{ z \in \mathcal{R} \mid e^{\operatorname{Rest}(c(u(z)-v_\infty/2)\mathrm{st})} \left|\sigma(u(z))\right| < \left|\sigma(u(z) - v_\infty)\right| \right\} \\ B_2 &= \left\{ z \in \mathcal{R} \mid e^{\operatorname{Rest}(c(u(z)-v_\infty/2)\mathrm{st})} \left|\sigma(u(z))\right| > \left|\sigma(u(z) - v_\infty)\right| \right\}. \end{aligned} \tag{6.14}$$

Proof

(i) From (6.10) and (6.11) we deduce that

$$\tilde{G}(v + n_1 + n_2\tau) = \exp\left[c(n_1 + n_2\tau) + (n_1\eta_1 + n_2\eta_2)v_\infty\right] \tilde{G}(v). \tag{6.15}$$

Hence, (6.12) holds if $\operatorname{Re}(n_1(c + \eta_1 v_\infty) + n_2(c\tau + \eta_2 v_\infty)) = 0$, which is equivalent to the system of equation

$$\begin{pmatrix} 1 & 0 \\ \operatorname{Re}(\tau) & -\operatorname{Im}(\tau) \end{pmatrix} \begin{pmatrix} \operatorname{Re}(c) \\ \operatorname{Im}(c) \end{pmatrix} = -\begin{pmatrix} \operatorname{Re}(\eta_1 v_\infty) \\ \operatorname{Re}(\eta_2 v_\infty) \end{pmatrix} \tag{6.16}$$

for $\operatorname{Re}(c)$ and $\operatorname{Im}(c)$. Since $\tau \notin \mathbb{R}$, it follows that $c \in \mathbb{C}$ exists and is uniquely determined by (6.16). In order to prove (6.13), we have to verify that $\left|\tilde{G}(v)\ \tilde{G}(v + v_\infty)|_{v=0}\right| = 1$, since we know from (6.5) that $\left|u'(\infty^{(1)})\right| = \left|u'(\infty^{(2)})\right|$ and that the definition of Φ has been based on (4.2).

Indeed, from (6.9) and the antisymmetry $\zeta(-v) = -\zeta(v)$ it follows that

$$\log\left(\sigma(v_\infty)\right) - \log\left(\sigma(-v_\infty)\right) = \int_{-v_\infty}^{v_\infty} \zeta(v)\,\mathrm{d}v = 0. \tag{6.17}$$

Hence, from (6.11) we have

$$\left|\tilde{G}(v)\ \tilde{G}(v + v_\infty)|_{v=0}\right| = \left.\frac{\sigma(v)}{\sigma(v - v_\infty)} \frac{\sigma(v + v_\infty)}{\sigma(v)}\right|_{v=0} = \frac{\sigma(v_\infty)}{\sigma(-v_\infty)} = 1. \tag{6.18}$$

(ii) The identities (6.14) are an immediate consequence of (4.3) together with (6.11) and (6.13). □

The next lemma shows, how the special zero z_n of R_n is connected with spurious poles of the Padé approximants $[n/n]$.

LEMMA 6.4.— *Assume that $N \subseteq \mathbb{N}$ is an infinite subsequence such that*

$$z_n \longrightarrow z_0 \in \mathcal{R} \quad \textit{as} \;\; n \longrightarrow \infty,\; n \in N. \tag{6.19}$$

(i) If $z_0 \in B_0 \setminus \{\infty^{(2)}\} \subseteq \mathcal{R}$, then there exists an infinite subsequence of N, which we continue to denote by N, such that for each $n \in N$ the denominator-polynomial Q_n has a zero and the Padé approximant $[n/n]$ a spurious pole at $\pi_n \in \mathbb{C}$, and we have

$$\pi_n \longrightarrow \pi(z_o) \quad \textit{as} \;\; n \longrightarrow \infty,\; n \in N. \tag{6.20}$$

(ii) If $z_0 \in B_1 \subseteq \mathcal{R}$, then the Padé approximants $[n/n]$, $n \in N$, have no spurious poles.

Proof.— From limit (6.19) and Lemma 6.1 together with (6.7) we deduce that

$$\left|R_n(z)\right|^{1/n} \longrightarrow \left|\tilde{G}\big(u(z)\big)\right| \quad \text{as} \;\; n \longrightarrow \infty,\; n \in N, \tag{6.21}$$

locally uniformly for $z \in \mathcal{R} \setminus \{z_0, \infty^{(2)}\}$. Let $\varphi : \mathcal{R} \longrightarrow \mathcal{R}$ be the covering transformation on $\mathcal{R}$, and let $\pi_j^{-1} : D \longrightarrow B_j$, $j = 1, 2$, denote the two branches of the inverse π^{-1}. From (6.7) and (6.1) analogously to (5.33) we have

$$R_n\Big(\pi_2^{-1}(\zeta)\Big) + 2\sqrt{P(\zeta)}\; Q_n(\zeta) = R_n\Big(\pi_1^{-1}(\zeta)\Big)\,, \quad \zeta \in D. \tag{6.22}$$

Because of the convergence (6.21) and the inequalities in (6.14) it follows from Rouché's Theorem that Q_n has a zero in D approximately below the place where R_n has a zero in B_2.

(i) If $z_0 \in B_2 \setminus \{\infty^{(2)}\}$, then there exists a zero $\pi_n \in D$ of Q_n near z_n for $n \in N$ sufficiently large, and because of (6.19) $\pi_n \longrightarrow \pi(z_0) \in D$ as $n \longrightarrow \infty$, $n \in N$. If π_n were a zero of P_n at the same time, then it would follow from (6.7) that R_n had zeros at $\pi_n \in B_2$ and at $\varphi(\pi_n) \in B_1$. But this is impossible because of Lemma 6.1. Thus, $[n/n]$ has a pole at π_n, and the convergence (6.20) shows that this pole is spurious.

(ii) If $z_0 \in B_1$, then $z_n \in B_1$ for $n \in N$ sufficiently large. Hence, it follows from Lemma 6.1 that R_n has no zeros on B_2, and consequently $[n/n]$, $n \in N$, has no spurious pole. □

DEFINITION 6.1.— *The two numbers $r_1, r_2 \in [0,1)$ satisfying*

$$v_\infty \equiv r_1 + r_2\tau \mod (1,\tau) \tag{6.23}$$

are called the coordinates of the point $v_\infty \in \overline{\mathbb{C}}/\Gamma$. In the same way we define $(r_{1n}, r_{2n}) \in [0,1) \times [0,1)$ by

$$u(z_n) \equiv r_{1n} + r_{2n}\tau \mod (1,\tau) \tag{6.24}$$

as coordinates of $u(z_n)$, $n \in \mathbb{N}$.

LEMMA 6.5.— *Let $\{r\} := r - [r] \in [0,1)$ denote the integer-remainder of $r \in \mathbb{R}$. We have*

$$r_{jn} = \{(n+2)r_j\}, \quad j = 1,2,\ n = 1,2,\ldots \tag{6.25}$$

With respect to the distribution of the set

$$\{(r_{1n}, r_{2n}) \in [0,1) \times [0,1) \mid j = 1,2,\ldots\} \tag{6.26}$$

we can distinguish three cases:

(i) If $(r_1, r_2) \in \mathbb{Q} \times \mathbb{Q}$, then the set (6.26) is contained in a finite lattice in $[0,1) \times [0,1)$.

(ii) If $(r_1, r_2) \notin \mathbb{Q} \times \mathbb{Q}$, but the 3 numbers $1, r_1, r_2$ are linearly dependent over $\mathbb{Q}$. then the set (6.26) is contained in a finite number of parallel lines in $[0,1) \times [0,1)$.

(iii) If the 3 numbers $1, r_1, r_2$ are linearly independent over $\mathbb{Q}$, then the set (6.26) is dense in $[0,1) \times [0,1)$.

Proof.— Relation (6.25) is an immediate consequence of (6.24), (6.8) in Lemma 6.2, and (6.23). From Weyl's Uniform Distribution Theorem (cf. [Ch, chap. VIII]) it follows that if $\alpha \in (0,1)$ is irrational, then the set $A_1 := \{\{n\alpha\} \mid n = 1,2,\ldots\}$ is dense and uniformly distributed in $[0,1)$. If α is rational, then the set A_1 forms a finite, 1-dimensional lattice in $[0,1)$. Part (i) follows from the last assertion.

The assumption of part (ii) can be subdivided in three cases: (a) $r_1 \in \mathbb{Q}$, $r_2 \notin \mathbb{Q}$, (b) $r_1 \notin \mathbb{Q}$, $r_2 \in \mathbb{Q}$, (c) $r_1, r_2 \notin \mathbb{Q}$ and $r_1/r_2 \in \mathbb{Q}$. In the cases (a) and (b) the assertion of (ii) follows immediately from Weyl's Uniform Distribution Theorem. We consider case (c): there exists $\alpha \in \mathbb{Q}$ with $r_2 = \alpha r_1$, and hence

$$\begin{aligned} r_{2n} &\equiv (\alpha r_{1n} + \alpha\,[(n+2)r_1]) \mod (1) \\ &= \alpha r_{1n} + \zeta_n, \end{aligned} \tag{6.27}$$

where ζ_n can assume only finitely many rational values. The assertion of part (ii) in case (c) follows directly from (6.27).

The assertion of (iii) follows directly from Weyl's Uniform Distribution Theorem. □

Dumas proved the next theorem in the language of continued fractions, which especially in its third subsection shows how poor and disappointing the convergence behavior of the diagonal Padé approximants $[n/n]$, $n \in \mathbb{N}$ can be.

THEOREM 6.6 (Dumas). — *The set*

$$S := \pi\Big(\{\zeta \in \mathcal{R} |\ |\tilde{G}(u(z))| = 1\}\Big) \tag{6.28}$$

consists of two arcs γ_1 and γ_2 connecting two branch points each. With the same classification in subcases as used in Lemma 6.5 we have:

(i) If $(r_1, r_2) \in \mathbb{Q} \times \mathbb{Q}$, then there exist finitely many points $\zeta_1, \ldots, \zeta_{n_1} \in \overline{\mathbb{C}}$ such that

$$\lim_{n\to\infty} [n/n](\zeta) = f(\zeta) \tag{6.29}$$

locally uniformly for $\zeta \in \overline{\mathbb{C}} \setminus (S \cup \{\zeta_1, \ldots, \zeta_{n_1}\})$.

(ii) If $(r_1, r_2) \notin \mathbb{Q} \times \mathbb{Q}$ but the numbers $1, r_1, r_2$ are rationally dependent, then there exist finitely many arcs $\gamma_3, \ldots, \gamma_{n_2}$ in $\overline{\mathbb{C}}$ such that the limit (6.29) holds true locally uniformly for $\zeta \in \overline{\mathbb{C}} \setminus (S \cup \gamma_3 \cup \cdots \cup \gamma_{n_2})$.

(iii) If the three numbers $1, r_1, r_2$ are rationally independent, then there exist two infinite sets Σ_1, $\Sigma_2 \in \overline{\mathbb{C}}$, both dense in $\overline{\mathbb{C}}$, and the limit (6.29) holds point-wise for each $\zeta \in \Sigma_1$ and each $\zeta \in \Sigma_2$ is a cluster point of poles of $[n/n]$, $n \in \mathbb{N}$.

Proof.— From (6.14) and (6.11) we deduce that the set $\mathcal{R}\setminus(B_1 \cup B_2) = \left\{z \in \mathcal{R} \mid |\tilde{G}(u(z))| = 1\right\}$ consists of two closed arcs, the branch points $a_1, \ldots, a_4$ belong to this set, and $\pi(B_1) = \pi(B_2)$. The convergence domain for the diagonal Padé approximants $[n/n]$, $n \in \mathbb{N}$, is given by

$$D = \overline{\mathbb{C}} \setminus S. \tag{6.30}$$

Let $N \subseteq \mathbb{N}$ be an infinite subsequence such that the limit (6.19) holds true. From (6.14) and (6.11) we know that

$$\left|\tilde{G}\Big(u(\pi_1^{-1}(\zeta))\Big)\right| < \left|\tilde{G}\Big(u(\pi_2^{-1}(\zeta))\Big)\right| \quad \text{for } \zeta \in D. \tag{6.31}$$

Hence, (6.21) and (6.22) imply that

$$\lim_{n\to\infty} |Q_n(\zeta)|^{1/n} = \left|\tilde{G}\Big(u(\pi_2^{-1}(\zeta))\Big)\right| \tag{6.32}$$

locally uniformly for $\zeta \in D \setminus \{\pi(z_0)\}$ if the limit point z_0 in (6.19) lies in B_2 or locally uniformly for $\zeta \in D$ if $z_0 \notin B_2$. From (6.7) it follows that

$$(f - [n/n])(\zeta) = \frac{R_n(\pi_1^{-1}(\zeta))}{Q_n(\zeta)}. \tag{6.33}$$

From the proof of Lemma 6.3 we know that $\left|\tilde{G}(u(z))\tilde{G}(u(\varphi(z)))\right| = 1$ for all $z \in \mathcal{R}$ and φ denoting the covering transformation of $\mathcal{R}$, *i.e.*, $\pi \circ \varphi = \pi$ and $\varphi \neq \mathrm{id}_{\mathcal{R}}$. We deduce from this together with (6.33), (6.21), and (6.32) that

$$\begin{aligned}\lim_{\substack{n\to\infty\\ n\in N}} |(f - [n/n])(\zeta)|^{1/n} &= \frac{\left|\tilde{G}\Big(u(\pi_1^{-1}(\zeta))\Big)\right|}{\left|\tilde{G}\Big(u(\pi_2^{-1}(\zeta))\Big)\right|} \\ &= \left|\tilde{G}\Big(u(\pi_1^{-1}(\zeta))\Big)\right|^2 < 1\end{aligned} \tag{6.34}$$

locally uniformly for $\zeta \in D \setminus \{\pi(z_0)\}$ if $z_0 \in B_2$ and locally uniformly for $\zeta \in D$ else. We note that the right-hand side of (6.34) is independent of the selection of the subsequence N, but the limit point z_0 in (6.19) depends on N.

Define $\tilde{B}_j := u(B_j) \subseteq \overline{\mathbb{C}}/\Gamma$, $j = 1, 2$, and the maps $\tilde{\pi}_j : \tilde{B}_j \longrightarrow D$ by $t \longmapsto \pi(u^{-1}(t))$, which are conformal bijections between the $\tilde{B}_j$ and D,

$j = 1, 2$, respectively. Let $\tilde{A}$ denote the set (6.26) inbeded into $\overline{\mathbb{C}}/\Gamma$, and define

$$A := \tilde{\pi}_2(\tilde{A} \cap \tilde{B}_2). \tag{6.35}$$

(i) From Lemma 6.5 we know that if $(r_1, r_2) \in \mathbb{Q} \times \mathbb{Q}$, then the set $\tilde{A}$ is finite, and consequently also the set A, *i.e.*, $A = \{\zeta_1, \ldots, \zeta_{n_1}\}$. From Lemma 6.4 we then know that all spurious poles of the Padé approximants $[n/n]$, $n \in \mathbb{N}$, have to cluster at points of A. The conclusion of part (i) then follows from (6.34).

(ii) If the assumptions of part (ii) are satisfied, then we know from Lemma 6.5 that the set $\tilde{A}$ is contained in a finite number of parallel lines in $\overline{\mathbb{C}}/\Gamma$. Let $\tilde{\gamma}_3, \ldots, \tilde{\gamma}_{n_2}$ be the non-empty intersections of these lines with the domain $\tilde{B}_2$, and let the images under $\tilde{\pi}_2$ be denoted by $\gamma_2, \ldots, \gamma_{n_2}$. Then the conclusion in part (ii) follows in exactly the same way as in part (i).

(iii) If the numbers $1, \gamma_1, \gamma_2$ are rationally independent, then we know from Lemma 6.5 that the set $\tilde{A} \cap \tilde{B}_2$ is dense in $\tilde{B}_2$, and hence the set A is dense in $\overline{\mathbb{C}}$. From Lemma 6.4 it then follows that each $z \in A$ is a cluster point of spurious poles of the Padé approximants $[n/n]$ as $n \to \infty$. This proves the first conclusion in part (iii) with $\Sigma_1 := A$.

In order to prove the second conclusion in part (iii) we need some preparation. Let $K, K_1 \subseteq D \setminus \{\infty\}$ be two compact sets with $K \subseteq \mathrm{Int}(K_1)$. For $\varepsilon > 0$ sufficiently small we define

$$A_n := \left\{ \zeta \in K \;\middle|\; \left| \frac{R_n(\pi_1^{-1}(\zeta))}{Q_n(\zeta)} \right| \geq (1 - \varepsilon)^{2n} \right\}. \tag{6.36}$$

Let $N \subseteq \mathbb{N}$ be an infinite subsequence such that for $n \in N$ a zero π_n of Q_n exists with $\pi_n \in K_1$. For $n \in \mathbb{N} \setminus N$ sufficiently large we have $A_n = \emptyset$. Define $\tilde{Q}_n := Q_n/(\cdot - \pi_n)$ for $n \in N$. From the limit (6.34) and from (6.36) we deduce that for all $n \in N$ sufficiently large and all $\zeta \in A_n$ we have

$$|\zeta - \pi_n| \leq (1 - \varepsilon)^{-2n} \left| \frac{R_n(\pi_1^{-1}(\zeta))}{\tilde{Q}_n(\zeta)} \right| \leq c_0(K)^{2n} \tag{6.37}$$

with

$$c_0(K_1) := \sup_{\zeta \in K_1} \left[(1 - 2\varepsilon)^{-1} \left| \tilde{G}\left(u(\pi_1^{-1}(\zeta)) \right) \right| \right]. \tag{6.38}$$

Hence, we have

$$m(A_n) \leq \pi c_0(K)^{2n} \tag{6.39}$$

for $n \in \mathbb{N}$ sufficiently large, where m is the planar Lebesgue measure. Since $c_0(K) < 1$, for $\tilde{A}_n := \bigcup_{n' \geq n} A_{n'}$ we have

$$m(\tilde{A}) \leq \frac{\pi}{1 - c_0(K)^2} c_0(K)^{2n} \quad \text{for} \quad n \geq n_0. \tag{6.40}$$

From (6.36) and (6.33) it follows that for each $\zeta \in D \setminus \tilde{A}_n$ the limit (6.29) hold true point-wise. From (6.40) we then deduce that in D point-wise convergence hold true almost everywhere with respect to planar Lebesgue measure. This proves the second conclusion in part (iii). □

We remark that in part (iii) of Theorem 6.6 less has been stated than has been proved. The convergence almost everywhere, which has been proved, also follows from convergence in capacity, which has been proved in Theorem 2.2 (by much less elementary means).

It is not clear whether there exists an infinite subset $\Sigma_2 \subseteq \mathbb{C}$ dense in $\mathbb{C}$ such that the sequence $[n/n]$, $n \in \mathbb{N}$, diverges at every point $\zeta \in \Sigma_2$. By a more detailed investigation of part (i) in Theorem 6.6 than given here, it could be shown that pointwise convergence holds true in (6.29) for all $\zeta \in D = \mathbb{C} \setminus S$, however, in neighborhoods of the points $\zeta_1, \ldots, \zeta_{n_1}$ the convergence is not uniform.

From the Lemmas 6.2, 6.4, and 6.5 it can be deduced that in the case of the function (6.1) an infinite subsequence $N \subseteq \mathbb{N}$ always exists such that the diagonal Padé approximants $[n/n]$, $n \in \mathbb{N}$, converge locally uniformly to f in the domain $D = \mathbb{C} \setminus S$, *i.e.*, that for these special functions f the Baker–Gammel–Wills conjecture can be proved in the form of Theorem 6.8 without any restriction or additional assumption. Actually, it can further be shown that the subsequence N can have a denseness in $\mathbb{N}$ of $1/2$.

Acknowledgement

The ideas underlying the material in Section 3 are based on discussions the author has had with J. Nuttall.

References

[BaGM] BAKER (G. A.) JR. and GRAVES-MORRIS.— *Padé Approximants Parts I and Part II*, Encycl. Math. Vols **13** and **14**. Cambridge University Press, Cambridge (1981).

[BGW] BAKER (G. A.) JR., GAMMEL (J. L.) and WILLS (J. G.).— *An investigation of the applicability of the Padé approximant method*, J. Math. Anal. Appl. **2** (1961), pp. 405-418.

[Br] BREZINSKI (C.).— *History of Continued Fractions and Padé approximants*, Series in Comp. Math. **12**, Springer-Verlag, Heidelberg (1991).

[Bo] BORCHARDT (C. W.).— *Application des transcendantes abéliennes á la théorie des fractions continues*, J. Reine Angew. Math. **48** (1854), pp. 69-104.

[Ch] CHANDRASEHHARAN (K.).— *Analytic Number Theory*, Springer-Verlag, Berlin (1968).

[Du] DUMAS (S.).— *Sur le développement des fonctions elliptiques en fractions continues*, Thésis, Univ. Zürich, (1908).

[Go] GONCHAR (A. A.).— *A local condition for single-valuedness of analytic functions*, Mat. Sbornik, **89** (131) (1972); English transl. in Math. USSR Sb. **18** (1972), pp. 151-167.

[GoRa] GONCHAR (A. A.) and RAKHMANOV (E. A.).— *Equilibrium distributions and degree of rational approximation of analytic functions*, Mat. Sb. **134** (1987); English transl. in Math. USSR Sb. **62** (1989), pp. 305-48.

[Ham] HAMBURGER (H.).— *Über eine Erweiterung des Stieltjesschen Momentenproblems* I, II, III, Math. Ann. **81** (1920). pp. 235-319, Math. Ann. **82** (1921), pp. 120-164, 168-187.

[Ha] HALPHEN (G.-H.).— *Traité des fonctions elliptiques et de leurs applications*, Tome II. Gauthier-Villars, Paris (1888).

[Hi] HILLE (E.).— *Analytic Function Theory*. Ginns and Company, Boston (1962).

[HuCo] HURWITZ (A.) and COURANT (R.).— *Funktionentheorie*, Springer-Verlag. Berlin (1929).

[Ja] JACOBI (C.-G.-J.).— *Note sur une nouvelle application de l'analyse des functions elliptiques a l'algèbre*, J. Reine Angew. Math. **7** (1830), pp. 41-43.

[Kr] KREIN (M. G.).— *The ideas of P. L. Chebychev and A. A. Markov in the theory of limiting values of integrals and their further development*, Amer. Math. Soc. Transl. Series **12** (1959), pp. 1-121.

[La] LANDKOF (N. S.).— *Foundations of Modern Potential Theory*, Springer-Verlag Berlin (1972).

[Lu] LUBINSKY (D. S.).— *Divergence of complex rational approximants*. Pacific J. Math. **108** (1983), pp. 141-153.

[Nu1] NUTTALL (J.).— *The convergence of Padé approximants of meromorphic functions*, J. Math. Anal. Appl. **31** (1970), pp. 129-140.

[Nu2] NUTTALL (J.).— *Asymptotics of diagonal Hermite-Padé polynomials*. J. Approx. Theory **42** (1984), pp. 299-386.

[Pe] PERRON (O.). — *Die Lehre von den Kettenbrüchen*, Teubner, Stuttgard (1913).

[Po1] POMMERENKE (CH.). — *Ueber die Kapazitaet ebener Kontinuen*, Math. Ann. **139** (1959), pp. 64-75.

[Po2] POMMERENKE (CH.). — *Padé approximants and convergence in capacity*, J. Math. Anal. Appl. **41** (1973), pp. 775-780.

[Ra] RAKHMANOV (E. A.). — *On the convergence of Padé approximants in classes of holomorphic functions*, Math. USSR Sb. **40** (1980), pp. 149-155.

[Si] SIEGEL (C. L.). — *Transzendente Zahlen*, Bibliographisches Institut, Mannheim (1967).

[Sp] SPRINGER (G.). — *Introduction to Riemann Surfaces*, Chelsea publ. cop. New York (1981).

[St1] STAHL (H.). — *Extremal domains associated with an analytic function I, II*, Complex Variables Theory Appl. **4** (1985), pp. 311-324, 325-338.

[St2] STAHL (H.). — *The structure of extremal domains associated with an analytic function*, Complex Variables Theory Appl. **4** (1985), pp. 339-354.

[St3] STAHL (H.). — *The convergence of Padé approximants to functions with branch points*, Submitted to J. Approx. Theory.

[Sti] STIELTJES (T.-J.). — *Recherches sur les fractions continues*, Ann. Fac. Sci. Toulouse **8** (1894), pp. 1-122.

[StTo] STAHL (H.) and TOTIK (V.). — *General Orthogonal Polynomials*, Encycl. Math. Vol. **43**, Cambridge University Press, Cambridge, (1992).