Journal of Approximation Theory **108**, 53–96 (2001)
doi:10.1006/jath.2000.3515, available online at http://www.idealibrary.com on IDEAL®

Asymptotic Error Estimates for L^2 Best Rational Approximants to Markov Functions

L. Baratchart

INRIA, 2004, route des Lucioles, B.P. 93, 06902 Sophia Antipolis Cedex, France
E-mail: baratcha@sophia.inria.fr

H. Stahl[1]

TFH-Berlin/FB II, Luxemburger Strasse 10, 13353 Berlin, Germany
E-mail: stahl@p-soft.de

and

F. Wielonsky

INRIA, 2004, route des Lucioles, B.P. 93, 06902 Sophia Antipolis Cedex, France
E-mail: wielonsky@sophia.inria.fr

Communicated by Guillermo López Lagomasino

Received November 12, 1999; accepted in revised form July 17, 2000; published online November 28, 2000

Let $f(z) = \int (t-z)^{-1}\, d\mu(t)$ be a Markov function, where μ is a positive measure with compact support in $\mathbb{R}$. We assume that $\text{supp}(\mu) \subset (-1, 1)$, and investigate the best rational approximants to f in the Hardy space $H_2^0(V)$, where $V := \{z \in \overline{\mathbb{C}} \mid |z| > 1\}$ and $H_2^0(V)$ is the subset of functions $f \in H_2(V)$ with $f(\infty) = 0$. The central topic of the paper is to obtain asymptotic error estimates for these approximants. The results are presented in three groups. In the first one no specific assumptions are made with respect to the defining measure μ of the function f. In the second group it is assumed that the measure μ is not too thin anywhere on its support so that the polynomials p_n, orthonormal with respect to the measure μ, have a regular nth root asymptotic behavior. In the third group the defining measure μ is assumed to belong to the Szegő class. For each of the three groups, asymptotic error estimates are proved in the L^2-norm on the unit circle and in a pointwise fashion. Also the asymptotic distribution of poles, zeros, and interpolation points of the best L^2 approximants are studied. © 2001 Academic Press

Key Words: best rational approximation in the L^2-norm on the unit circle; asymptotic error estimates; Markov's theorem.

[1] Research supported by the Deutche Forschungsgemeinschaft (AZ: Sta 299/8-2).

0021-9045/01 $35.00

1. INTRODUCTION

We consider best rational approximants to functions of the real Hardy space $H_2^0(V) := \{ f \in H_2(V) \mid f(\infty) = 0 \}$ with $V := \overline{\mathbb{C}} \backslash \overline{\mathbb{D}} = \{ z \in \overline{\mathbb{C}} \mid |z| > 1 \}$. This type of approximants are of particular interest in control theory, stochastic modeling and signal processing [3, 17, 23]. In the present paper, we are specifically concerned with the approximation of Markov functions, i.e., functions f of type

$$f(z) = f(\mu; z) := \int \frac{d\mu(t)}{t - z} \tag{1.1}$$

with *defining measure* μ that is positive and has compact support in $\mathbb{R}$. In order that $f \in H_2^0(V)$, we assume throughout that

$$\operatorname{supp}(\mu) \subset (-1, 1). \tag{1.2}$$

These approximants were studied in [6–9]. Besides the fact that they lend themselves to computation due to deep connections with the classical theory of orthogonal polynomials, it should be noticed that they occur in system theory as transfer functions of relaxation systems [10, 28, 40].

Let $\mathcal{P}_n$ denote the set of all complex polynomials of degree at most n, $\mathcal{R}_{m,n}$ the set of all rational functions of numerator and denominator degree at most m and n, respectively, and $\mathcal{R}^1_{m,n} \subseteq \mathcal{R}_{m,n}$ the subset of rational functions with all their poles in $\mathbb{D}$. Hence, $\mathcal{R}^1_{m,n} = \mathcal{R}_{m,n} \cap H_2^0(V)$, if $m < n$. By $R_n^* = R_n^*(f; \cdot)$ we denote a best rational approximant to f in $H_2^0(V)$ of order at most n, i.e., a rational function in $\mathcal{R}^1_{n-1,n}$ having minimal error with respect to the norm $\|\cdot\|$ defined by

$$\|g\| := \lim_{r \to 1+} \left[\frac{1}{2\pi} \int_0^{2\pi} |g(re^{it})|^2 \, dt \right]^{1/2} \quad \text{for} \quad g \in H^2(V). \tag{1.3}$$

Hence $R_n^* = R_n^*(f; \cdot) \in \mathcal{R}^1_{n-1,n}$ satisfies

$$\|f - R_n^*\| = \inf_{r \in \mathcal{R}^1_{n-1,n}} \|f - r\|. \tag{1.4}$$

A standard argument shows that for each pair of degrees $(n-1, n)$ there exists $R_n^* = R_n^*(f, \cdot) \in \mathcal{R}^1_{n-1,n}$ such that (1.4) holds. In general R_n^* is not unique (cf. [9]). In the remainder of the paper R_n^* will denote any best approximant of degree n. We study the behavior of the error function $f - R_n^*$ as $n \to \infty$. The results are presented in three groups: In the first one no special assumptions are made with respect to the defining measure μ beyond (1.2).

More specific asymptotic results are obtained in the second group, where it is assumed that the defining measure μ is not too thin anywhere on its support. It turns out that a good condition is to demand that the defining measure μ is such that the polynomials p_n orthonormal with respect to μ have a regular nth root asymptotic behavior (cf. [35, Chapter 3]). A sufficient condition for this property is, for instance, that $\mathrm{supp}(\mu)$ consists of a finite number of intervals and that the Radon–Nikodym derivative with respect to Lebesgue measure μ' of μ is positive almost everywhere on these intervals (the Erdős–Turán condition, cf. [35, Chapter 4.1]). In this second group we can prove asymptotic error estimates that are precise in an nth root sense.

In the third group it is assumed that the defining measure μ satisfies the Szegő condition, i.e., $\mathrm{supp}(\mu) = [a, b]$ and

$$\int_a^b \frac{\log \mu'(t)}{\sqrt{(t-a)(b-t)}}\, dt > -\infty, \tag{1.5}$$

where again μ' is the Radon–Nikodym derivative of μ. The Szegő condition is much stronger than the nth root regular behavior requires, and consequently, considerably stronger asymptotic error estimates can be proved in this last group.

Asymptotic estimates are presented in a norm and in a pointwise fashion. In the second group besides the L^2-norm also the L^∞-norm is considered. It turns out that under appropriate assumptions, best approximants in both norms have identical asymptotic error estimates in an nth root sense. Also, in the second group the asymptotic distribution of poles, zeros, and interpolation points of the best approximants R_n^* are studied. Using a slightly different approach that relies more on duality and the connection to n-widths, but again uses orthogonal polynomials with varying weights, more general results are proposed in [5] if μ satisfies the Szegő's condition. These results concern best meromorphic approximation with a prescribed number of poles in $L^p(\mathbb{T})$, $1 \leqslant p \leqslant \infty$, where $\mathbb{T}$ denotes the unit circle. They contain actually Theorem 8 of the present paper, since meromorphic approximation is just rational approximation when $p=2$, because of the orthogonality of $H_2^0(V)$ and the complex-conjugate space $\overline{H_2(V)}$ of $H_2(V)$ in $L^2(\mathbb{T})$.

The outline of the paper is as follows: All the results are stated and discussed in Section 2. In Section 3 preparatory results about best rational approximants in $H_2^0(V)$, rational interpolants to Markov functions, and potential theory are assembled. The proofs of the results stated in Section 2 are finally given in Sections 4 through 6.

2. MAIN RESULTS

In the first group of results no assumptions, except (1.2), are made with respect to the defining measure μ in (1.1). The concepts of the condenser capacity and condenser potentials are fundamental for the description of the asymptotic error estimates.

DEFINITION 2.1. The *logarithmic capacity* is denoted by $\operatorname{cap}(\cdot)$ (see [31, 32]). Let $K_1, K_2 \subseteq \overline{\mathbb{C}}$ be two disjoint, compact sets of positive capacity, and let the two sets be separated by the chain of smooth curves C, oriented so that K_1 is interior to C. There exists a function $p_{K_1, K_2}: \overline{\mathbb{C}} \backslash (\operatorname{Int}(K_1) \cup \operatorname{Int}(K_2)) \to \overline{\mathbb{R}}$ that is assumed to be subharmonic in a neighborhood of ∂K_1, superharmonic in a neighborhood of ∂K_2, harmonic in $\overline{\mathbb{C}} \backslash (K_1 \cup K_2)$, and there exists a constant $c > 0$ such that

$$p_{K_1, K_2}(z) = \begin{cases} 0 & \text{for quasi every} \quad z \in \partial K_1 \\ c & \text{for quasi every} \quad z \in \partial K_2, \end{cases} \tag{2.2}$$

and

$$\frac{1}{2\pi} \oint_C \frac{\partial p_{K_1, K_2}(\zeta)}{\partial n_\zeta} \, ds_\zeta = 1, \tag{2.3}$$

where $\partial/\partial n_\zeta$ denotes the outer normal derivative on C. The function p_{K_1, K_2} exists uniquely, and it is called the *condenser potential* of the condenser (K_1, K_2). The *condenser capacity* is defined as

$$\operatorname{cap}(K_1, K_2) = \frac{1}{c} \tag{2.4}$$

(cf. [1] or [32, Chapter VIII, Theorem 2.6]). Note that, by uniqueness, we have

$$p_{K_1, K_2}(z) + p_{K_2, K_1}(z) = \frac{1}{\operatorname{cap}(K_1, K_2)}, \qquad z \in \overline{\mathbb{C}} \backslash (K_1 \cup K_2).$$

A property is said to hold *quasi everywhere* on a set $S \subseteq \overline{\mathbb{C}}$ if it holds on S with possible exceptions on a subset of outer capacity zero (cf. [20, Chapter II, No. 6], or [35, Appendix I]). If $\operatorname{cap}(K_1) = 0$ or $\operatorname{cap}(K_2) = 0$, then by definition $p_{K_1, K_2}(z) := 0$ for all $z \in \overline{\mathbb{C}} \backslash (\operatorname{Int}(K_1) \cup \operatorname{Int}(K_2))$ and $\operatorname{cap}(K_1, K_2) := 0$.

Remarks. (1) In the present paper we consider only the two special condensers $(\operatorname{supp}(\mu), \mathbb{T})$, i.e., $K_1 = \operatorname{supp}(\mu)$ and $K_2 = \mathbb{T} = \partial \mathbb{D}$, and $(\operatorname{supp}(\mu), \operatorname{supp}(\mu)^{-1})$ with $\operatorname{supp}(\mu)^{-1}$ denoting the reflection of the set $\operatorname{supp}(\mu)$ across the unit circle $\mathbb{T}$.

(2) It follows from assumption (1.2) that $\operatorname{cap}(\operatorname{supp}(\mu), \mathbb{T}) < \infty$, and we have $\operatorname{cap}(\operatorname{supp}(\mu), \mathbb{T}) = 0$ if and only if $\operatorname{cap}(\operatorname{supp}(\mu)) = 0$.

(3) If $\operatorname{supp}(\mu)$ is an interval, i.e.,

$$\operatorname{supp}(\mu) = [a, b], \tag{2.5}$$

then rather explicit expressions can be given for the condenser potential $p_{[a,b],\mathbb{T}}$ and the condenser capacity $\operatorname{cap}([a, b], \mathbb{T})$ using elliptic integrals. Since we shall need these expressions at later places, we elaborate on case (2.5) in more detail in the next example.

EXAMPLE 2.6. Let $[a, b]^{-1}$ be the reflection of $[a, b]$ across the unit circle $\mathbb{T}$, and let φ be the conformal map of the ring domain $\overline{\mathbb{C}} \setminus ([a, b] \cup [a, b]^{-1})$ onto the annulus $\{r < |z| < 1/r\}$, $0 < r < 1$, with $\varphi(b) = r$. The restriction of φ to $\mathbb{D} \setminus [a, b]$ maps $\mathbb{D} \setminus [a, b]$ onto $\{r < |z| < 1\}$. Note also that $|\varphi|$ extends continuously to $[a, b]$ and $[a, b]^{-1}$, although the function φ itself has conjugate determination from above and below the cuts. It follows rather immediately from (2.2), (2.3), and (2.4) that

$$\operatorname{cap}([a, b], \mathbb{T}) = \frac{1}{\log 1/r}, \qquad \operatorname{cap}([a, b], [a, b]^{-1}) = \frac{1}{2 \log 1/r},$$

$$p_{[a,b],\mathbb{T}}(z) = \begin{cases} \log(|\varphi(z)|/r) & \text{for} \quad z \in \overline{\mathbb{D}} \\ \dfrac{1}{\log 1/r} & \text{for} \quad z \in \overline{\mathbb{C}} \setminus \overline{\mathbb{D}} \end{cases} \tag{2.6}$$

$$p_{[a,b],[a,b]^{-1}}(z) = \log(|\varphi(z)|/r) \qquad \text{for} \quad z \in \overline{\mathbb{C}}$$

(cf. [14, Chapter V, Sect. 1]). The number $1/r$ is also known as the modulus of the ring domain $\mathbb{D} \setminus [a, b]$. The mapping function φ can be expressed by the elliptic integral

$$\varphi(z) = \exp\left[\pi \frac{1 - ab}{2K} \int_1^z \frac{dt}{\sqrt{(t-a)(t-b)(1-ab)(1-bt)}}\right] \tag{2.7}$$

with integration along any path in $\overline{\mathbb{C}} \setminus ([a, b] \cup [a, b]^{-1})$, and K denotes the complete elliptic integral of the first kind

$$\begin{aligned} K := F\left(\frac{\pi}{2}, k\right) &= \int_0^1 \frac{dt}{\sqrt{(1-t^2)(1-k^2 t^2)}} \\ &= \frac{1 - ab}{2} \int_a^b \frac{dt}{\sqrt{(t-a)(b-t)(1-at)(1-bt)}} \end{aligned} \tag{2.8}$$

with modulus $k := (b-a)/(1-ab)$ (cf. [27, Chapter VI], or [18, Chapter 5]). The definite integral from b to 1 in (2.7) can be evaluated as

$$\int_b^1 \frac{dt}{\sqrt{(t-a)(t-b)(1-at)(1-bt)}} = \frac{1}{1-ab} \int_0^1 \frac{dt}{\sqrt{(1-t^2)(1-k'^2t^2)}}$$
$$= \frac{F\left(\frac{\pi}{2}, k'\right)}{1-ab} = \frac{K'}{1-ab} \tag{2.9}$$

with k' the conjugate modulus $k' = \sqrt{1-k^2} = \sqrt{(1-a^2)(1-b^2)}/(1-ab)$. From (2.6), (2.7), and (2.9) it follows that

$$\operatorname{cap}([a, b], \mathbb{T}) = \frac{2}{\pi} \frac{K}{K'}, \qquad \operatorname{cap}([a, b], [a, b]^{-1}) = \frac{1}{\pi} \frac{K}{K'}. \tag{2.10}$$

We note that if $\operatorname{supp}(\mu) \subseteq [a, b]$, then $\operatorname{cap}(\operatorname{supp}(\mu), \mathbb{T}) \leqslant \operatorname{cap}([a, b], \mathbb{T})$. Thus, (2.10) gives an explicit upper estimate for $\operatorname{cap}(\operatorname{supp}(\mu), \mathbb{T})$ and in the general case.

In the first group of results we have two theorems, the first one with an error estimate in norm and the second one with a pointwise version.

THEOREM 1. *If f is a Markov function* (1.1) *with defining measure μ satisfying* (1.2), *then for any sequence of best approximants $R_n^* = R_n^*(f; \cdot)$, $n = 1, 2, ...,$ we have*

$$\limsup_{n\to\infty} \|f - R_n^*\|^{1/2n} \leqslant \exp\left[\frac{-1}{\operatorname{cap}(\operatorname{supp}(\mu), \mathbb{T})}\right]. \tag{2.11}$$

Remarks. (1) Since the right hand side of (2.11) is smaller than 1, it follows from (2.10) that we always have geometric convergence in the L^2-norm (1.3). If $\operatorname{cap}(\operatorname{supp}(\mu)) = 0$, then the convergence even is faster than geometric. Theorem 3, below, will show that (2.11) is sharp.

(2) In the present paper only Markov functions f are approximated. For $\mathscr{K}$, a compact subset of $\mathbb{C}$ disjoint from $\mathbb{T}$, and f, a general analytic function in $\overline{\mathbb{C}} \backslash \mathscr{K}$, remarkable results of Parfenov and Prokhorov [29, 30] (formerly Gonchar's conjecture) assert that Theorem 1 remains true if $\operatorname{supp}(\mu)$ is replaced in (2.11) by $\mathscr{K}$ and lim sup replaced by lim inf. Results of the type given in Theorem 1 and also in the Theorems 3 and 4 have been proved in [15], but there the best rational approximants are defined on a real interval E disjoint from $\operatorname{supp}(\mu)$.

Back to the case of Markov functions: since R_n^* has all its poles in $\mathbb{D}$, we observe that (2.11) and the Cauchy formula imply that the convergence $R_n^* \to f$ holds throughout $\overline{\mathbb{C}} \backslash \overline{\mathbb{D}}$. But more can be proved.

THEOREM 2. *Let f be a Markov function* (1.1) *with defining measure μ satisfying* (1.2), *and let $I(\mu) = [a, b]$ denote the smallest interval containing* $\operatorname{supp}(\mu)$. *Then we have*

$$\limsup_{n \to \infty} |f(z) - R_n^*(z)|^{1/2n}$$

$$\leqslant \begin{cases} \exp[-p_{\operatorname{supp}(\mu),\, \mathbb{T}}(z)] \\ \qquad \textit{for} \quad z \in \overline{\mathbb{D}} \backslash I(\mu) \\ \max\left(\left|\dfrac{1-az}{a-z}\right|, \left|\dfrac{1-bz}{b-z}\right|\right) \exp\left[-\dfrac{1}{\operatorname{cap}(\operatorname{supp}(\mu), \mathbb{T})}\right] \\ \qquad \textit{for} \quad z \in \overline{\mathbb{C}} \backslash \overline{\mathbb{D}} \end{cases} \tag{2.12}$$

locally uniformly for $z \in \overline{\mathbb{C}} \backslash I(\mu)$, with $p_{\operatorname{supp}(\mu),\, \mathbb{T}}$ the condenser potential defined by (2.2) *and* (2.3).

Remark. Since the right-hand side of (2.12) is smaller than 1 for all $z \in \overline{\mathbb{C}} \backslash I(\mu)$, we have geometric convergence throughout $\overline{\mathbb{C}} \backslash I(\mu)$, and since $\max(|\frac{1-az}{a-z}|, |\frac{1-bz}{b-z}|) < 1$ for all $z \in \overline{\mathbb{C}} \backslash \overline{\mathbb{D}}$, the right-hand side of (2.12) is smaller in $\overline{\mathbb{C}} \backslash \overline{\mathbb{D}}$ than on $\mathbb{T}$. However, the estimate (2.12) may no be sharp on $\overline{\mathbb{C}} \backslash \overline{\mathbb{D}}$.

Next, we come to the results in the second group. The main difference with the first group is that the limes superior in (2.11) and (2.12) is replaced by proper limits and that the inequalities are replaced by equalities. Of course, this is only possible if the class of functions (1.1) is restricted by additional assumptions. These assumptions are concerned with the defining measure μ. It turns out that a natural condition is related with the asymptotic behavior of polynomials that are orthonormal with respect to μ.

DEFINITION 2.13. Let μ be a positive measure with compact support in $\mathbb{R}$. When $\operatorname{cap}(\operatorname{supp}(\mu)) > 0$, we say that the orthonormal polynomials with respect to μ have *regular nth root asymptotic behavior*, if the asymptotic relation

$$\lim_{n \to \infty} \frac{1}{n} \log |p_n(z)| = g_{\overline{\mathbb{C}} \backslash \operatorname{supp}(\mu)}(z, \infty) \tag{2.14}$$

holds locally uniformly for $z \in \mathbb{C} \backslash I(\mu)$. Here, by $g_D(z, x)$ we denote the Green function of the domain $D \subseteq \overline{\mathbb{C}}$. In case (2.14) is satisfied, we write $\mu \in \mathbf{Reg}$ (cf. [35, Chapter 3]). We also may say that μ is *regular* with respect to orthonormal polynomials.

There exist a number of criteria that ensure that $\mu \in \mathbf{Reg}$ (cf. [35, Chapter 4]). Most often used is the Erdős–Turán condition, which says that if $\operatorname{supp}(\mu)$ is the union of finitely many closed intervals and if the Radon–Nikodym derivative μ' of μ with respect to Lebesgue measure is positive almost everywhere on these intervals, then $\mu \in \mathbf{Reg}$. There exist weaker conditions, which however involve more complicated constructions.

The first two theorems in the present group are counterparts to Theorems 1 and 2. They are complemented by a theorem that estimates the L^∞-error and by a theorem about the asymptotic distribution of poles, zeros, and interpolation points of the H^2 approximants. The results are very analogous to those obtained in [15] for best rational approximants on an interval $[\alpha, \beta] \in \overline{\mathbb{R}} \backslash [-1, 1]$ if $[\alpha, \beta]$ is the reflection of $\operatorname{supp}(\mu)$ across $\mathbb{T}$.

THEOREM 3. *If f is a Markov function* (1.1) *with defining measure μ satisfying* (1.2) *and $\mu \in \mathbf{Reg}$, then we have*

$$\lim_{n \to \infty} \| f - R_n^* \|^{1/2n} = \exp\left[\frac{-1}{\operatorname{cap}(\operatorname{supp}(\mu), \mathbb{T})}\right]. \tag{2.15}$$

In the next two theorems we need the condenser potential $p_{\operatorname{supp}(\mu), \operatorname{supp}(\mu)^{-1}}$ of the condenser $(\operatorname{supp}(\mu), \operatorname{supp}(\mu)^{-1})$. Because of the symmetry of the condenser $(\operatorname{supp}(\mu), \operatorname{supp}(\mu)^{-1})$ with respect to $\mathbb{T}$, we have

$$\operatorname{cap}(\operatorname{supp}(\mu), \mathbb{T}) = 2 \operatorname{cap}(\operatorname{supp}(\mu), \operatorname{supp}(\mu)^{-1}) \tag{2.16}$$

and

$$p_{\operatorname{supp}(\mu), \operatorname{supp}(\mu)^{-1}}(z) = p_{\operatorname{supp}(\mu), \mathbb{T}}(z) \qquad \text{for} \quad z \in \overline{\mathbb{D}}. \tag{2.17}$$

THEOREM 4. *If f is a Markov function* (1.1) *with defining measure μ satisfying* (1.2) *and $\mu \in \mathbf{Reg}$, then we have*

$$\lim_{n \to \infty} |f(z) - R_n^*(z)|^{1/2n} = \exp[-p_{\operatorname{supp}(\mu), \operatorname{supp}(\mu)^{-1}}(z)], \tag{2.18}$$

locally uniformly for $z \in \overline{\mathbb{C}} \backslash (I(\mu) \cup I(\mu)^{-1})$, and, for any sequence $z_n \to z_0 \in I(\mu)^{-1}$, $z_n \in V = \overline{\mathbb{C}} \backslash \overline{\mathbb{D}}$ we have

$$\limsup_{n \to \infty} |f(z_n) - R_n^*(z_n)|^{1/2n} \leqslant \exp[-p_{\operatorname{supp}(\mu), \operatorname{supp}(\mu)^{-1}}(z_0)]. \tag{2.19}$$

Remark. A glance at (2.18) and (2.12) shows that, on $\overline{\mathbb{C}}\backslash\overline{\mathbb{D}}$, the right-hand side of (2.18) is necessarily not greater than that of (2.12). This will be established in the course of the proof. Note, already, that the right-hand sides are identical on $\overline{\mathbb{D}}\backslash I(\mu)$, by (2.17).

A restriction of (2.18) to the unit circle $\mathbb{T}$ yields with (2.16), (2.17), and (2.4) that

$$\lim_{n\to\infty} |f(z)-R_n^*(z)|^{1/2n} = \exp\left[\frac{-1}{\operatorname{cap}(\operatorname{supp}(\mu),\mathbb{T})}\right] \tag{2.20}$$

uniformly for $z\in\mathbb{T}$. Hence, the nth root of the approximation error is asymptotically circular on $\mathbb{T}$. As a consequence, we can deduce that the L^2- and L^∞-best approximants have identical asymptotic error estimates in a nth root sense. Let $R_{n,\infty}^*$ denote a best rational approximant with respect to the L^∞-norm, i.e., let $\|\cdot\|_{\mathbb{T}}$ denote the sup-norm on $\mathbb{T}$, and let $R_{n,\infty}^*\in\mathscr{R}_{n,n}^1$ satisfy (1.4) with the L^2-norm replaced by the norm $\|\cdot\|_{\mathbb{T}}$.

THEOREM 5. *Under the assumptions of Theorems* 3 *and* 4, *we have*

$$\lim_{n\to\infty} \|f-R_{n,\infty}^*\|_{\mathbb{T}}^{1/2n} = \exp\left[-\frac{1}{\operatorname{cap}(\operatorname{supp}(\mu),\mathbb{T})}\right]. \tag{2.21}$$

As an appendix to the results presented so far in the second group, we state and discuss a theorem about the asymptotic distribution of the poles and zeros of L^2-best approximants R_n^* and also of the zeros of the error functions $f-R_n^*$ in $\overline{\mathbb{C}}\backslash I(\mu)$. The zeros of the error function are interpolation points for the approximants R_n^*. The statement of the theorem demands some preparations.

We list in the next lemma several assertions about the location of the zeros and poles of R_n^* and the zeros of $f-R_n^*$ that have been proved in [9, Proposition 5]. Actually, [9] only treats the case of rational functions with real coefficients, but it has been shown in [4] that best approximants to Markov functions have real coefficients if they are chosen among all rational functions of degree at most n with complex coefficients even.

LEMMA 2.22. *Let f be a Markov function* (1.1) *with defining measure μ satisfying* (1.2), *and assume that* $\operatorname{supp}(\mu)$ *contains infinitely many points. Then*

(i) *The numerator and denominator degrees of R_n^* are exactly $n-1$ and n, respectively*,

(ii) *All poles and zeros of R_n^* are simple and contained in $I(\mu)$* (*the smallest interval containing* $\operatorname{supp}(\mu)$),

(iii) *The* n *poles and* $n-1$ *zeros of* R_n^* *interlace,*

(iv) *The error function* $f-R_n^*$ *has a double zero at the reflection (across* $\mathbb{T}$*) of each pole of* R_n^**,*

(v) *In addition to the* $2n$ *zeros mentioned in* (iv), *the error function* $f-R_n^*$ *has a further zero at infinity,*

(vi) *Besides of the* $2n+1$ *zeros mentioned in* (iv) *and* (v), $f-R_n^*$ *has no other zeros in* $\overline{\mathbb{C}}\backslash I(\mu)$.

Remark. Since $f-R_n^*$ has a total number of $2n+1$ zeros in $\overline{\mathbb{C}}\backslash I(\mu)$, the rational function $R_n^* \in \mathscr{R}_{n-1,n}^1$ is uniquely determined by interpolation (cf. [35, Chapter 6.1]). (However, this does not mean that R_n^* is uniquely determined by its L^2-minimality.) The most important part of Lemma 2.22 is perhaps assertion (iv). It has first been proved for L^2-best approximants with fixed denominators by Walsh [39], for approximants with free poles it has been proved in [12] under the additional assumption that all poles of R_n^* are simple. The final form for R_n^* has been proved in [21], and its extension to stationary points of (1.4) in [2].

Let $P_n = P(R_n^*)$ and $Z_n = Z(R_n^*)$ denote the set of all poles and all zeros, respectively, of R_n^*, and $I_n = Z(f-R_n^*)\,|_{\overline{\mathbb{C}}\backslash I(\mu)}$ the set of all interpolation points of R_n^* outside of $I(\mu)$, taking account of multiplicities by repetitions. For the description of the asymptotic distributions we need the equilibrium distributions of condensers.

DEFINITION 2.23. Let $K_1, K_2 \subseteq \overline{\mathbb{C}}$ be two disjoint, compact sets of positive capacity. The condenser potential p_{K_2,K_1} can be represented as

$$p_{K_2,K_1}(z) = \int g_{\overline{\mathbb{C}}\backslash K_2}(z,x)\,d\omega_{K_1,K_2}(x), \tag{2.24}$$

with ω_{K_1,K_2} a probability measure on K_1 (cf. [20, Theorem 1.22′], or [32, Chapter VIII, Theorem 2.6]). The measure ω_{K_1,K_2} is the *condenser equilibrium distribution* on K_1 of the condenser (K_1, K_2). By interchanging the role of the two sets K_1 and K_2, a probability measure ω_{K_2,K_1} on K_2 is defined, which is the *condenser equilibrium distribution* on K_2.

Remark. The measure ω_{K_1,K_2} (and in an analogous way the measure ω_{K_2,K_1}) can also be defined by the minimal energy property

$$I(\omega_{K_1,K_2}) = \inf_{\omega} I(\omega), \qquad I(\omega) := \iint g_{\overline{\mathbb{C}}\backslash K_2}(x,y)\,d\omega(x)\,d\omega(y), \tag{2.25}$$

where the infimum extends over all probability measures ω with $\operatorname{supp}(\omega) \subseteq K_1$.

In our applications we are concerned only with the two condensers $(K_1, K_2) = (\mathrm{supp}(\mu), \mathbb{T})$ and $(K_1, K_2) = (\mathrm{supp}(\mu), \mathrm{supp}(\mu)^{-1})$. Because of the symmetry of the condenser $(\mathrm{supp}(\mu), \mathrm{supp}(\mu)^{-1})$ with respect to $\mathbb{T}$, $\omega_{\mathrm{supp}(\mu),\, \mathbb{T}} = \omega_{\mathrm{supp}(\mu),\, \mathrm{supp}(\mu)^{-1}}$, and $\omega_{\mathrm{supp}(\mu)^{-1},\, \mathrm{supp}(\mu)}$ is the image measure of $\omega_{\mathrm{supp}(\mu),\, \mathbb{T}}$ under the map $t \mapsto 1/t$.

EXAMPLE 2.26. In case $\mathrm{supp}(\mu)$ is an interval $[a, b] \subseteq (-1, 1)$, it is possible to give explicit representations for both measures $\omega_{[a,b],\, \mathbb{T}}$ and $\omega_{\mathrm{supp}(\mu)^{-1},\, \mathrm{supp}(\mu)}$. It follows from (2.6) and (2.7) (cf. [27, Chapter VI]) that

$$
\begin{aligned}
d\omega_{[a,b],\, \mathbb{T}}(x) &= \frac{(1-ab)\, dx}{2K\sqrt{(x-a)(b-x)(1-ax)(1-bx)}} \\
&\qquad \text{for} \quad x \in [a, b], \\
d\omega_{\mathbb{T},\, [a,b]}(\zeta) &= \frac{(1-ab)\, dt}{4K\, |\zeta - a|\, |\zeta - b|} \qquad \text{for} \quad \zeta = e^{it} \quad \zeta \in \mathbb{T}, \\
d\omega_{[a,b]^{-1},\, [a,b]}(x) &= \frac{(1-ab)\, dx}{2K\sqrt{(x-a)(b-x)(1-ax)(1-bx)}} \\
&\qquad \text{for} \quad x \in [a, b]^{-1}.
\end{aligned}
\tag{2.27}
$$

The square root is assumed to be positive, and K denotes the complete elliptic integral of the first kind given in (2.8).

THEOREM 6. *Let f be a Markov function* (1.1) *with defining measure μ satisfying* (1.2) *and $\mu \in \mathbf{Reg}$, then for any subinterval $[\alpha, \beta] \subseteq I(\mu)$ we have*

$$
\begin{aligned}
\lim_{n \to \infty} \frac{1}{n} \mathrm{card}(P(R_n^*) \cap [\alpha, \beta]) &= \lim_{n \to \infty} \frac{1}{n} \mathrm{card}(Z(R_n^*) \cap [\alpha, \beta]) \\
&= \omega_{\mathrm{supp}(\mu),\, \mathbb{T}}([\alpha, \beta]),
\end{aligned}
\tag{2.28}
$$

where $P(R_n^*)$ and $Z(R_n^*)$ are the sets of poles and zeros of R_n^*, respectively. For the set of interpolation points $I_n = Z(f - R_n^*)\,|_{\overline{\mathbb{C}} \setminus I(\mu)}$ and for each subinterval $[\alpha, \beta] \subseteq I(\mu)^{-1}$ we have

$$
\lim_{n \to \infty} \frac{1}{2n} \mathrm{card}(I_n \cap [\alpha, \beta]) = \omega_{\mathrm{supp}(\mu)^{-1},\, \mathrm{supp}(\mu)}([\alpha, \beta]). \tag{2.29}
$$

The measures $\omega_{\mathrm{supp}(\mu),\, \mathbb{T}}$ and $\omega_{\mathrm{supp}(\mu)^{-1},\, \mathrm{supp}(\mu)}$ are condenser equilibrium distributions on $\mathrm{supp}(\mu)$ *and* $\mathrm{supp}(\mu)^{-1}$, *respectively.*

Remarks. (1) The condition $\mu \in \mathbf{Reg}$ implies already the $\operatorname{cap}(\operatorname{supp}(\mu)) > 0$. Hence, the condition for the existence of the condenser equilibrium distributions $\omega_{\operatorname{supp}(\mu),\, \mathbb{T}}$ and $\omega_{\operatorname{supp}(\mu)^{-1},\, \operatorname{supp}(\mu)}$ is automatically satisfied.

(2) If condition $\mu \in \mathbf{Reg}$ is not satisfied, then an asymptotic distribution in the sense of (2.28) and (2.29) may not hold true.

The asymptotic relations (2.28) and (2.29) can be written in a more elegant form by using counting measures and weak convergence. For a finite set $S \subseteq \overline{\mathbb{C}}$ the *counting measure* v_S is defined as $\sum_{z \in S} \delta_z$, and a sequences of measures $\{\mu_n\}$ is said to *converge weakly* to a limit measure μ_0, written as $\mu_n \xrightarrow{*} \mu_0$, if $\int f\, d\mu_n \to \int f\, d\mu_0$ for any function f continuous on $\overline{\mathbb{C}}$. With these notions the relations (2.22) and (2.23) can be rewritten as

$$\begin{aligned} &\frac{1}{n} v_{P(R_n^*)} \xrightarrow{*} \omega_{\operatorname{supp}(\mu),\, \mathbb{T}} = \omega_{\operatorname{supp}(\mu),\, \operatorname{supp}(\mu)^{-1}}, \\ &\frac{1}{n} v_{Z(R_n^*)} \xrightarrow{*} \omega_{\operatorname{supp}(\mu),\, \mathbb{T}} = \omega_{\operatorname{supp}(\mu),\, \operatorname{supp}(\mu)^{-1}}, \\ &\frac{1}{2n} v_{I_n} \xrightarrow{*} \omega_{\operatorname{supp}(\mu)^{-1},\, \operatorname{supp}(\mu)}, \qquad \text{as} \quad n \to \infty. \end{aligned} \tag{2.30}$$

We now come to the third and last group of results. Here, a rather strong assumption is made with respect to the defining measure μ in (1.1): it has now to belong to the Szegő class. As a consequence we can prove rather precise asymptotic relations, so-called *power asymptotics* or strong asymptotics.

We start with some preparations, which includes the definition of a Szegő function that is specially adapted to the approximation problem. Originally, the Szegő function and also the Szegő condition have been introduced for the study of the asymptotic behavior of orthonormal polynomials defined on the circle $\mathbb{T}$ or on a real interval (cf. [37, Chapter XII]). In the present paper the definition of the Szegő function has to reflect the symmetry of the approximation problem with respect to $\mathbb{T}$, which is caused by the correspondence between the poles and the interpolation points of the approximants R_n^*.

DEFINITION 2.31. A positive measure μ with compact support belongs to the *Szegő class* on $\mathbb{R}$ (or satisfies the *Szegő condition* on $\mathbb{R}$) if in the Lebesgue decomposition

$$d\mu(x) = \mu'(x)\, dx + d\mu_s(x), \qquad x \in [a, b] := \operatorname{supp}(\mu), \tag{2.32}$$

the Radon–Nikodym derivative μ' satisfies (1.5), i.e.,

$$\int_a^b \frac{\log \mu'(x)}{\sqrt{(x-a)(b-x)}}\, dx > -\infty. \tag{2.33}$$

While the Szegő condition defined in (2.33) is identical with the classical one, the definitions of the geometric mean of the measure μ and the new Szegő function associated with μ possess some special features, which will be discussed in some remarks below.

DEFINITION 2.34. Let μ be a positive measure with compact support $\operatorname{supp}(\mu) \subseteq (-1, 1)$, and let

$$d\mu(x) = \frac{\dot{\mu}(x)\, dx}{\pi \sqrt{(x-a)(b-x)}} + d\mu_s(x), \qquad x \in [a, b] := I(\mu), \tag{2.35}$$

be the Lebesgue decomposition of μ with respect to the equilibrium distribution $d\omega_{[a,b]} := dx/(\pi \sqrt{(x-a)(b-x)})$, $x \in [a, b]$, on $[a, b]$. The measure μ_s is assumed to be singular with respect to the equilibrium distribution $\omega_{[a,b]}$, which is equivalent to being singular with respect to the Lebesgue measure. The *geometric mean* of μ is defined as

$$D(\mu) := \exp\left[\int \log \dot{\mu}(x)\, d\omega_{[a,b],\mathbb{T}}(x)\right], \tag{2.36}$$

where $\omega_{[a,b],\mathbb{T}}$ is the condenser equilibrium distribution of the condenser $([a, b], \mathbb{T})$ introduced in (2.24), and an explicit representation of which is given in (2.27).

LEMMA 2.37. *A measure μ with* $\operatorname{supp}(\mu) \subseteq [a, b] \subseteq (-1, 1)$ *belongs to the Szegő class on $[a, b]$ if, and only if, $D(\mu) > 0$.*

Remarks. (1) The two different types of equilibrium distributions $\omega_{[a,b]}$ and $\omega_{[a,b],\mathbb{T}}$ should not be mixed up. The density function $\dot{\mu}$ is defined with respect to the equilibrium distribution $\omega_{[a,b]}$, i.e., $\dot{\mu} = d\mu/d\omega_{[a,b]}$, but in (2.36) the integration of $\dot{\mu}$ is done with respect to $\omega_{[a,b],\mathbb{T}}$. Strictly speaking, $D(\mu)$ is a weighted geometric mean of $\dot{\mu}$ with respect to the condenser equilibrium distribution $\omega_{[a,b],\mathbb{T}}$ as weight. The use of this specific weight is justified by (2.44) in Theorem 7 and (2.46) in the corollary to Theorem 8, below.

(2) Because of (2.35) we have $\dot{\mu} \equiv 1$ if $\mu = \omega_{[a,b]}$. The density $\dot{\mu}$ represents the deviation of the absolute continuous part of μ from the

equilibrium distribution $\omega_{[a,b]}$, and $\mu = \omega_{[a,b]}$ can be considered as the prototypical example for a defining measure in (1.1).

(3) From (2.32) and (2.35) it is immediate that $\dot{\mu}(x) = \pi\sqrt{(x-a)(b-x)}\,\mu'(x)$ for $x \in [a,b]$. From (2.27) it therefore follows that (2.33) is satisfied if, and only if, $\int \log \dot{\mu}(x)\, d\omega_{[a,b],\mathbb{T}}(x) > -\infty$ (cf. the proof of Lemma 2.37, below). This last condition guarantees that the Szegő function exists, which will be introduced now.

Definition 2.38. Let μ belong to the Szegő class with $\operatorname{supp}(\mu) = [a,b] \subseteq (-1,1)$. Then the function

$$D(\mu;z) := \exp\left[\sqrt{(z-a)(z-b)(1-az)(1-bz)} \times \frac{1}{2\pi}\int_a^b \frac{\log(\dot{\mu}(x)/D(\mu))}{\sqrt{(x-a)(b-x)(1-ax)(1-bx)}}\frac{(1-2xz+x^2)\,dx}{(x-z)(1-xz)}\right], \tag{2.39}$$

$z \in \overline{\mathbb{C}} \setminus ([a,b] \cup [a,b]^{-1})$, is called the *Szegő function* with respect to the measure μ and the condenser $([a,b],[a,b]^{-1})$, where the first square root has to be taken negative at $z=1$ and the second one positive for $x \in (a,b)$.

Lemma 2.40. (i) *The function $D(\mu;\cdot)$ is analytic and different from zero in $\overline{\mathbb{C}} \setminus ([a,b] \cup [a,b]^{-1})$.*

(ii) *For the increment of $\arg D(\mu;\cdot)$ along $\mathbb{T}$ we have $\Delta_{t=0}^{2\pi} \arg D(\mu; e^{it}) = 0$.*

(iii) *The function $D(\mu;z)$ has non-tangential boundary values almost everywhere on $[a,b] \cup [a,b]^{-1}$ from both sides, and we have*

$$|D(\mu; x \pm i0)|^2 = \begin{cases} \dfrac{\dot{\mu}(x)}{D(\mu)} & \text{for almost every} \quad x \in [a,b], \\[2ex] \dfrac{D(\mu)}{\dot{\mu}(1/x)} & \text{for almost every} \quad x \in [a,b]^{-1}. \end{cases} \tag{2.41}$$

(iv) *We have*

$$|D(\mu;z)|^2 = 1 \qquad \text{for all} \quad z \in \mathbb{T}. \tag{2.42}$$

Remarks. (1) The assertions (i), (ii), and (iii) in Lemma 2.40 determine the Szegő function $D(\mu,\cdot)$ uniquely, and they could be taken as definition for this function. In some sense they are also more instructive than Definition 2.38 itself. Lemma 2.40 will be proved in Section 5.

(2) We will see in the proof how the specific normalization of the density function $\hat{\mu}$ in (2.41) by the geometric mean $D(\mu)$ is necessary to achieve assertion (ii).

(3) The assertions (iii) and (iv) are the most important properties of the Szegő function $D(\mu;\ \cdot\ )$.

It seems that different approximation and extremality problems demand different Szegő functions. In some sense these functions have to be tailored for the specific needs. In [22] the asymptotic behavior of minimal Blaschke products has been investigated and in this problem lead to a Szegő function that is very similar to the one used here. Actually, it is identical up to two different normalizations. A definition of a Szegő function on an annulus along with a proof of its existence can also be found in [33, Theorem 9].

We are now prepared to state the two theorems that are the main results in the third group.

THEOREM 7. *Let f be a Markov function* (1.1) *with defining measure μ belonging to the Szegő class on* $[a, b] = \operatorname{supp}(\mu) \subseteq (-1, 1)$. *Set*

$$\rho = \rho_{[a,b]} := \exp\left[\frac{-1}{\operatorname{cap}([a,b], \mathbb{T})}\right]. \tag{2.43}$$

Then we have

$$\lim_{n\to\infty} \|f - R_n^*\|\, \rho_{[a,b]}^{-2n} = \sqrt{\frac{8K}{\pi(1-ab)}}\, D(\mu), \tag{2.44}$$

where K is the complete elliptic integral of the first kind with modulus $k = (b-a)/(1-ab)$.

Remark. Note that the constant $\rho_{[a,b]}$ depends only on the geometry of the problem, i.e., on $\operatorname{supp}(\mu) = [a, b]$, while $D(\mu)$ depends on the measure μ itself. If one takes the $2n$th root on both sides of (2.44), it is immediate that (2.44) implies (2.15).

As in Theorems 2 and 4, here also in the third group there exist strong asymptotic error estimates that hold throughout $\overline{\mathbb{C}}\setminus[a, b]$.

THEOREM 8. *Let f be a Markov function* (1.1) *with defining measure μ belonging to the Szegő class on* $[a, b] = \operatorname{supp}(\mu) \subseteq (-1, 1)$, *and let φ denote* *the conformal map* $\varphi\colon \overline{\mathbb{C}}\backslash([a,b]\cup[a,b]^{-1})\to\{\rho<|z|<1/\rho\}$ *defined as in* (2.7) *with* $\varphi(b)=\rho=\rho_{[a,b]}$ *and* $\rho_{[a,b]}$ *defined by* (2.43). *Then we have*

$$f(z)-R_n^*(z)=(1+o(1))\,\rho_{[a,b]}^{2n}\,\varphi(z)^{-2n}\,D(\mu)\,D(\mu;z)^2\,\frac{-2}{\sqrt{(z-a)(z-b)}},\tag{2.45}$$

when $o(1)$ *denotes Landau's little "oh", which holds locally uniformly for* $z\in\overline{\mathbb{C}}\backslash([a,b]\cup[a,b]^{-1})$.

Remarks. (1) In (2.45) the constant $\rho_{[a,b]}$, the function φ, and the function $[(z-a)(z-b)]^{-1/2}$ depend only on the geometry of the problem, while $D(\mu)$ and the Szegő function $D(\mu;\cdot)$ depend on the measure μ. If one takes the $2n$th root on both sides of (2.45), one arrives at (2.18) in Theorem 4.

(2) We have $|\rho_{[a,b]}\varphi(z)|<1$ for all $z\in\overline{\mathbb{C}}\backslash([a,b]\cup[a,b]^{-1})$. Therefore, (2.45) implies geometric convergence throughout $\overline{\mathbb{C}}\backslash([a,b]\cup[a,b]^{-1})$. The upper asymptotic estimate can be extended to $[a,b]^{-1}$ by the maximum principle.

Since $|\varphi(z)|=1$ for $z\in\mathbb{T}$, we deduce from Theorem 8 the following

COROLLARY. *Under the assumptions of Theorem* 7 *we have*

$$\lim_{n\to\infty}|f(z)-R_n^*(z)|\,\rho_{[a,b]}^{-2n}=\frac{2D(\mu)}{\sqrt{|(z-a)(z-b)|}},\tag{2.46}$$

uniformly for all $z\in\mathbb{T}$.

A comparison of (2.46) with (2.20) sheds light on a major difference between nth root and power asymptotics. From the limit (2.46) we learn that the error function $f-R_n^*$ is not asymptotically circular on $\mathbb{T}$. Hence, the L^2-best approximants R_n^* and the L^∞-best approximants $R_{n,\infty}^*$ have a different asymptotic behavior in the sense of power asymptotics. This difference is smoothed away by taking the nth root of the error function, as Theorem 5 shows.

All results stated so far in the present section hold not only for best rational approximants R_n^*, but also for all rational functions $\hat{R}_n\in\mathcal{R}_{n-1,n}^1$ for which the norm of the error $\|f-\hat{R}_n\|$ has zero derivative respective to the coefficients of $\hat{R}_n$. These functions will be called *stationary*, and their definition will be outlined below.

We consider the functional

$$\psi(q):=\left\|f-\frac{L_q}{q}\right\|=\inf_p\left\|f-\frac{p}{q}\right\|,\tag{2.47}$$

where the infimum extends over all $p \in \mathcal{P}_{n-1}$, while $q \in \mathcal{P}_n$ is a given monic polynomial of exact degree n with $Z(q) \subseteq \mathbb{D}$. In the sequel the set of all such polynomials q will be denoted by $\mathcal{P}_n^1$. The elements of $\mathcal{P}_n^1$ are parameterized by the n coefficients q_j, $j=1, ..., n$, of the monic polynomial $q(z)=z^n+q_1 z^{n-1}+\cdots+q_n \in \mathcal{P}_n^1$.

DEFINITION 2.48. A rational function

$$\hat{R}_n = \frac{L_q}{q}, \qquad q(z) = z^n + q_1 z^{n-1} + \cdots + q_n \tag{2.49}$$

is called a stationary point of (2.47) or a *stationary approximant* (in L^2) if

$$\frac{\partial}{\partial q_j} \psi(q) = 0 \qquad \text{for} \quad j = 1, ..., n. \tag{2.50}$$

The concept of stationary approximants $\hat{R}_n \in \mathcal{R}_{n-1,n}^1$ in H^2-best approximation has been developed and studied in [2] and [9]. All local minima of (2.47) are stationary points, and therefore stationary approximants are candidates for best rational approximants R_n^*. Of course, as usual, the relation between local and global minima is an intriguing problem. Fortunately, in the case of Markov functions, this relation is rather well understood.

In [9, Proposition 5], it has been shown that stationary points can be characterized by a weighted orthogonality relation for the denominators of the approximants $\hat{R}_n$'s. In [9] and [8] it has been shown that under certain conditions on μ there exists only one single stationary point $\hat{R}_n$ for a given n, and this is then the uniquely existing best rational approximant R_n^*.

In general, the best rational approximants R_n^* to a Markov function may not be unique (cf. [7]). The next theorem, however, guarantees that not only the different best approximants R_n^*, but also all stationary points $\hat{R}_n$ have the same asymptotic error estimates.

THEOREM 9. *Everywhere in the Theorems* 1 *through* 8 *the best rational approximants* R_n^* *can be replaced by any stationary approximant* $\hat{R}_n$ *of degree* n, *and all conclusions hold true for this larger family of functions.*

Remark. In [8] it has been shown that if the defining measure μ in the Markov function (1.1) belongs to the Szegő class, then there exists n_0 such that for every $n \geqslant n_0$ there exists only one stationary approximant $\hat{R}_n = R_n^*$. Thus, Theorem 9 states nothing new in this situation.

3. AUXILIARY RESULTS

In the present section a number of results and definitions are assembled, which will be fundamental for proofs in the Sections 4 and 5.

Let $\tilde{p}$ denote the *reversed* polynomial

$$\tilde{p}(z) := z^n \bar{p}\left(\frac{1}{z}\right), \tag{3.1}$$

where $p \in \mathscr{P}_n$. The zeros of $\tilde{p}$ are the reflections of the zeros of p on the unit circle $\mathbb{T}$. Notice that the definition of $\tilde{p}$ depends on n; usually it is clear from the context which n is meant. The next lemma has been proved in [9, Proposition 5].

LEMMA 3.2. *If f is a Markov function* (1.1) *with defining measure μ satisfying* (1.2) *and assume that* $\operatorname{supp}(\mu)$ *is not finite, then the assertion of Lemma* 2.22 *holds for all stationary rational functions $\hat{R}_n$, and not only for best rational approximants R_n^*, as stated in Lemma* 2.22.

Remarks. (1) Let q_n be the denominator of $\hat{R}_n$. It follows from assertion (i) of Lemma 2.22 that $\deg(q_n) = n$, and from assertions (iv) and (v) of the same lemma that the function $\hat{R}_n$ interpolates f in the zeros of the polynomial $\widetilde{q_n}^2$ in Hermite's sense. If one takes also into account the situation at infinity, then the total number of interpolation points is $2n+1$. Notice that at infinity both functions, f and $\hat{R}_n$, have a systematic zero. The function $\hat{R}_n$ is determined uniquely already by interpolation in the $2n$ zeros of $\widetilde{q_n}^2$.

(2) From assertion (ii) of Lemma 2.22 we learn that all zeros of q_n are contained in $I(\mu) \subseteq (-1, 1)$. Consequently all interpolation points are contained in $I(\mu)^{-1} \cup \{\infty\}$ which is a subset of $\overline{\mathbb{R}} \backslash [-1, 1]$. If the denominator polynomial q_n is chosen to be monic, then the polynomial q_n has only real coefficients.

For general analytic functions, rational interpolants may not exist for every constellations of interpolation points. The existence problem is usually overcome by considering multipoint Padé approximants instead of the proper interpolants. However, for Markov functions and conjugate symmetric interpolation points, there are no difficulties of this kind. The convergence theory of multipoint Padé approximants will be a major tool in the proofs of results of Section 2. Before we can state the main result in this direction we have to introduce some notations.

Let a triangular *scheme* of interpolation points

$$\mathscr{A}=(a_{ij})_{\substack{j=1,\ldots,i\\ i=1,2,\ldots}}=\begin{pmatrix} a_{11} & & \\ a_{21} & a_{22} & \\ \vdots & \vdots & \ddots \end{pmatrix}, \qquad a_{ij}\in\overline{\mathbb{C}}, \tag{3.3}$$

be given and define the polynomial w_n by

$$w_n(z):=\prod_{j=1}^{n}\left(1-\frac{z}{a_{nj}}\right), \qquad n\in\mathbb{N}. \tag{3.4}$$

The *support* of the scheme $\mathscr{A}$ is defined as $\mathrm{supp}(\mathscr{A}):=\mathrm{Closure}\{a_{nj},\ n=1,2,\ldots,\ j=1,\ldots,n\}$, the *essential support* as the sets of accumulation points of $\mathrm{supp}(\mathscr{A})$, and it is denoted by $\mathrm{essensupp}(\mathscr{A})$. In order that the polynomial w_n in (3.4) is properly defined, we assume that all points a_{nj} in the interpolation scheme $\mathscr{A}$ are different from zero, which can always be achieved by a shift of the whole interpolation problem.

DEFINITION 3.5. Let the function f be defined on $\mathrm{supp}(\mathscr{A})$. For any pair of numbers $m,\ n\in\mathbb{N}$ there exists a pair of polynomials $p_{mn}\in\mathscr{P}_m$, $q_{mn}\in\mathscr{P}_n$, $q_{mn}\not\equiv 0$ that satisfies

$$(q_{mn}f-p_{mn})(z)=\begin{cases} O(w_{m+n+1}(z)) & \text{as}\quad z\to Z(w_{m+n+1})\\ O(z^{\deg(w_{m+n+1})-m-n-1}) & \text{as}\quad z\to\infty. \end{cases} \tag{3.6}$$

The rational function

$$R_{mn}=R_{mn}(f,\mathscr{A};\cdot):=\frac{p_{mn}}{q_{mn}} \tag{3.7}$$

is the *multipoint Padé approximant* of degree m, n. It is uniquely determined by (3.6), but in general it does not interpolate f in each of the $m+n+1$ points of the $m+n+1$th row A_{m+n+1} in the scheme $\mathscr{A}$ because (3.6) could read $0=0$ in case some a_{m+n+1j} is a zero of q_{mn}. Thus, in this case, we do not get a proper rational interpolant. If however a rational interpolant exists, then it is identical with R_{mn}. In (3.6) by $Z(p)$ we have denoted the set of all zeros of the polynomial p taking account of multiplicities, and by $O(\cdot)$ the Landau symbol "big oh."

In case of Markov functions (1.1) the multipoint Padé approximants have especially nice properties. If the interpolation scheme $\mathscr{A}$ is such that

$$w_n(z)=\overline{w_n(\bar{z})} \qquad \text{for all}\quad n\in\mathbb{N} \tag{3.8}$$

and

$$\operatorname{supp}(\mathscr{A}) \cap I(\mu) = \phi, \tag{3.9}$$

then the multipoint Padé approximants $R_{n-1,n}(f, \mathscr{A};\ \cdot)$ to Markov function f are proper rational interpolants (cf. [35, Lemma 6.1.2]).

In our analysis the interpolation points will be the zeros of the polynomial $\tilde{q}^2$, where q is the denominator of the stationary rational function $\hat{R}_n$. From the remarks to Lemma 3.2 we then know that the conditions (3.8) and (3.9) are satisfied for this situation.

It is a classical result that the denominators of finite sections of continued fractions are determined by orthogonality. The same is true for the denominators of multipoint Padé approximants, but now the orthogonality relation depends on n. In [35, Lemma 6.1.2], the following result has been proved.

Lemma 3.10. *A polynomial $q_n \in \mathscr{P}_n$, $q_n \not\equiv 0$, is the denominator of $R_{n-1,n}(f, \mathscr{A}; \cdot)$ if, and only if,*

$$\int x^l q_n(x) \frac{d\mu(x)}{w_{2n}(x)} = 0 \qquad \textit{for} \quad l = 0, ..., n-1. \tag{3.11}$$

For the interpolation error $f - R_{n-1,n}$ we have the representation

$$(f - R_{n-1,n}(f, \mathscr{A};\ \cdot))(z) = \frac{w_{2n}(z)}{q_n(z)^2} \int \frac{q_n(x)^2}{w_{2n}(x)} \frac{d\mu(x)}{x-z}. \tag{3.12}$$

Remarks. (1) Orthogonality (3.11) has been derived in [9] for stationary approximants $\hat{R}_n$ independently from the theory of multipoint Padé approximants.

(2) If the assumptions (3.8) and (3.9) are satisfied, then the measures $w_{2n}^{-1}\, d\mu$ are real and have no sign changes on $\operatorname{supp}(\mu)$.

(3) We note that (3.12) implies that the interpolation error $f - R_{n-1,n}$ is fully determined if the two polynomials q_n and w_{2n} are known. Since w_{2n} is usually given by the interpolation problem, the critical question is the determination of the denominator q_n.

Polynomials q_n satisfying relation (3.11) are called orthogonal with varying weights, and there is a growing literature on this subject. A comprehensive survey is contained in [32].

Relation (3.11) does not imply a specific normalization of the polynomials q_n. The most often used types are monic and orthonormal polynomials. In the later case we have $q_n(z) = \gamma_n z^n + \cdots \in \mathscr{P}_n$ with the leading coefficient $\gamma_n > 0$ determined by

$$\int q_n(x)^2 \frac{d\mu(x)}{w_{2n}(x)} = 1. \tag{3.13}$$

For multipoint Padé approximants to Markov functions asymptotic error estimates have been proved in the weak and strong sense. The formulation of these results demands some definitions.

Definition 3.14. We say that the interpolation scheme $\mathscr{A}$ has a probability measure $\alpha = \alpha(\mathscr{A})$ as its *asymptotic distribution* if

$$\frac{1}{n} \sum_{j=1}^{n} \delta_{a_{nj}} \xrightarrow{*} \alpha \qquad \text{as} \quad n \to \infty, \tag{3.15}$$

where, as in (2.30), $\xrightarrow{*}$ denotes convergence in the weak topology of measures defined on $\overline{\mathbb{C}}$.

For any domain $D \subseteq \overline{\mathbb{C}}$ and any positive measure α on D the *Green potential* is defined as

$$g(D, \alpha; z) := \int g_D(z, x)\, d\alpha(x). \tag{3.16}$$

As in (2.14), $g_D(z, x)$ denotes the Green function of the domain D. If $\operatorname{cap}(\partial D) = 0$, then by definition $g_D(z, x) = \infty$ for all z, $x \in D$. The next proposition has been proved in [35, Theorem 6.1.6], but is already contained in the earlier papers [16] and [24].

Proposition 3.17. *Let f be a Markov function and $\mathscr{A}$ an interpolation scheme with asymptotic distribution α that satisfies* (3.8) *and* (3.9).

(i) *We have*

$$\limsup_{n \to \infty} |f(z) - R_{n-1,n}(f, \mathscr{A}; z)|^{1/2n} \leqslant \exp[-g(\overline{\mathbb{C}} \backslash \operatorname{supp}(\mu), \alpha; z)] \tag{3.18}$$

locally uniformly for $z \in \overline{\mathbb{C}} \backslash I(\mu)$.

(ii) *If the defining measure μ of the Markov function* (1.1) *is regular in the sense of Definition* 2.13, *then, in* (3.18), *equality holds and we have a proper limit instead of the limes superior.*

The statement of strong asymptotic error estimates involves yet another Szegő function which is different from the one introduced in Definition 2.38.

DEFINITION 3.19. Let the defining measure μ in (1.1) belong to the Szegő class on $[a, b]$ introduced in Definition 2.31. With the Radon–Nikodym derivative $\dot{\mu} = d\mu/d\omega$ as introduced in (2.35) the *Szegő function* $G(\mu; \cdot)$ is defined as

$$G(\mu; z) := \exp\left[\frac{1}{2\pi}\sqrt{(z-a)(z-b)}\int_a^b \frac{\log \dot{\mu}(x)}{\sqrt{(b-x)(x-a)}}\frac{dx}{z-x}\right], \quad z \in \overline{\mathbb{C}}\backslash[a, b]. \tag{3.20}$$

We have

$$G(\mu; \infty) > 0 \tag{3.21}$$

if, and only if, μ belongs to the Szegő class. The properties of $G(\mu; \cdot)$ stated in the next lemma are a rather immediate consequence of (3.20), and they can be taken as defining properties of the Szegő function (cf. [19, p. 62] for the case of a disk, or [34, Assertion 2.13]).

LEMMA 3.22. *Let the measure μ belong to the Szegő class on $[a, b]$. Then*

(i) *$G(\mu; \cdot)$ is an outer function in the Hardy space H^2 on $\overline{\mathbb{C}}\backslash[a, b]$, it is analytic and different from zero in $\overline{\mathbb{C}}\backslash[a, b]$.*

(ii) *We have*

$$\lim_{y \to 0+} |G(\mu; x \pm iy)|^2 = \dot{\mu}(x) \tag{3.23}$$

for almost every $x \in [a, b]$.

Remark. It is a consequence of assertion (i) in Lemma 3.22 that $G(\mu; \cdot)$ has non-tangential boundary values almost everywhere in $[a, b]$ (cf. [19]). Hence, the limit in (3.23) exists for almost every $x \in [a, b]$.

By ψ we denote the conformal map $\psi: \overline{\mathbb{C}}\backslash[a, b] \to \mathbb{D}$ with $\psi(\infty) = 0$ and $\psi'(\infty) > 0$, which is given by

$$\psi(z) = \frac{2}{b-a}\left[z - \frac{b+a}{2} - \sqrt{(z-a)(z-b)}\right]. \tag{3.24}$$

Strong asymptotic error estimates have been proved in [38] under the assumption that the defining measure μ of the Markov function (1.1) is absolutely continuous (with respect to the Lebesgue measure) and that the points a_{nj} of the interpolation scheme $\mathscr{A}$ stay away from $[a, b]$ by a positive distance. In [25] and [11] it has been proved under the assumption (3.26) below, together with some other admissibility conditions. In the form stated here the proposition has been proved in [34, Theorem 2].

PROPOSITION 3.25. *Let the defining measure μ in the Markov function f belongs to the Szegő class on $[a, b]$, assume that the interpolation points a_{nj}, $j=1, ..., n$, $n \in \mathbb{N}$, of the scheme $\mathscr{A}$ satisfies* (3.8), (3.9), *and assume further that*

$$\lim_{n \to \infty} \sum_{j=1}^{2n} (1 - |\psi(a_{2n,\, j})|) = \infty. \tag{3.26}$$

Then we have

$$f(z) - R_{n-1,\, n}(f, \mathscr{A}; z)$$
$$= (-2 + o(1)) \frac{G(\mu; z)^2}{\sqrt{(z-a)(z-b)}} \prod_{j=1}^{2n} \frac{\psi(z) - \psi(a_{2n,\, j})}{1 - \overline{\psi(a_{2n,\, j})}\, \psi(z)} \quad \text{as} \quad n \to \infty \tag{3.27}$$

with $o(1)$ denoting the Landau symbol "little oh", which holds in (3.27) *on any compact subset of $\overline{\mathbb{C}} \backslash [a, b]$. The sign of the square root in* (3.27) *is chosen such that $\sqrt{(z-a)(z-b)} = z + O(1)$ as $z \to \infty$.*

We note that Proposition 3.25 is the strong version of the analogue of Markov's classical theorem about the convergence of continued fractions to functions of type (1.1). The analogy means here that multipoint Padé approximants and not sections of continued fractions are considered. Proposition 3.25 will be one of the main tools in the proof of the Theorems 7 and 8. In addition we need a result, which is closely related to Proposition 3.25, and is stated in the next proposition. A proof can be found in [34, Proposition 2.21], or in [25, Theorem 9].

PROPOSITION 3.28. *Let $q_n \in \mathscr{P}_n$ be the orthogonal polynomial defined by* (3.11) *and* (3.13). *Under the assumptions of Proposition* 3.25 *we have*

$$\frac{q_n^2}{|w_{2n}|}\, d\mu \xrightarrow{*} d\omega \qquad \text{as} \quad n \to \infty, \tag{3.29}$$

where $\xrightarrow{*}$ denotes the weak convergence of measures in $\overline{\mathbb{C}}$ and ω the equilibrium distribution $d\omega(x) = dx/(\pi \sqrt{(b-x)(x-a)})$, $x \in [a, b]$.

4. PROOF OF THE THEOREMS 1 THROUGH 6

There exists very close connection between the five Theorems 1–4 and 6, and it turns out to be best to prove all five theorems simultaneously. As stated in Theorem 9, all proofs are carried out for stationary approximants $\hat{R}_n$ instead of best rational approximants R_n^*.

In the present section the main tools in the proofs are of a potential-theoretic nature. Besides of the Green potentials $g(D, v; \cdot)$ defined in (3.16) and the condenser (equilibrium) potential p_{K_1, K_2} introduced in Definition 2.1, we also use logarithmic potentials, which are denoted by

$$p(v; z) := \int \log \frac{1}{|z-x|} \, dv(x), \tag{4.1}$$

where v is a measure on $\mathbb{C}$. The proofs of the theorems are prepared by three lemmas.

LEMMA 4.2. *For any probability measure v on $[a, b] \subseteq (-1, 1)$ we have*

$$g(\mathbb{D}, v; z) \geqslant \log \min \left(\left| \frac{1-az}{z-a} \right|, \left| \frac{1-bz}{z-b} \right| \right) \qquad \text{for} \quad z \in \mathbb{D}. \tag{4.3}$$

Proof. Using a Moebius transform $\psi: \mathbb{D} \to \mathbb{D}$ of the independent variable one can assume without loss of generality that $[a, b] = [-r, r]$, $0 < r < 1$. Then it is not difficult to verify that for all $c \in [-r, r]$ we have

$$\left| \frac{1+rz}{z+r} \right| \leqslant \left| \frac{1-cz}{z-c} \right| \qquad \text{for} \quad z \in \mathbb{D} \cap \{ w \mid Re(w) \geqslant 0 \}, \tag{4.4}$$

which implies that

$$\log \left| \frac{1+rz}{z+r} \right| \leqslant g(\mathbb{D}, v; z) \qquad \text{for} \quad z \in \mathbb{D} \cap \{ w \mid Re(w) \geqslant 0 \}. \tag{4.5}$$

In the same way we show that

$$\log \left| \frac{1-rz}{z-r} \right| \leqslant g(\mathbb{D}, v; z) \qquad \text{for} \quad z \in \mathbb{D} \cap \{ w \mid Re(w) \leqslant 0 \}. \tag{4.6}$$

The inequalities (4.5) and (4.6) imply (4.3). ∎

LEMMA 4.7. *Let $D \subseteq \overline{\mathbb{C}}$ be a domain such that $\partial D = K_1 \cup K_2$ with K_1 and K_2 two compact disjoint sets, and let further C be a chain of oriented smooth Jordan curves separating K_1 from K_2 in such a way that C has winding* *number* 1 *with respect to each point of* K_1 *and winding number* 0 *with respect to each point of* K_2. *If* u *is a harmonic function in* D *with*

$$\oint_C \frac{\partial}{\partial n} u(\zeta)\, ds_\zeta = 0, \tag{4.8}$$

where $\partial/\partial n$ *and* ds *are the normal derivative and the line element on* C, *respectively, then we have*

$$\sup_{z' \in K_1} \limsup_{z \to z',\, z \in D} u(z) \geqslant \inf_{z' \in K_2} \liminf_{z \to z',\, z \in D} u(z) \tag{4.9}$$

and the same relation holds with K_1 *and* K_2 *interchanged.*

Remarks. (1) Condition (4.8) means in potential theoretic terms that the total flux from K_1 and K_2 of the function u is zero.

(2) If equality holds in (4.9), then u is necessarily a constant. This stronger version of the lemma will not be needed and is also not proved here.

Proof. Assume that (4.9) is false. Then there exists $u_0 \in \mathbb{R}$ such that

$$\sup_{z' \in K_1} \limsup_{z \to z',\, z \in D} u(z) < u_0 < \inf_{z' \in K_2} \liminf_{z \to z',\, z \in D} u(z), \tag{4.10}$$

and we may assume that u_0 is a regular value of u, since critical values form a set of measure 0 (cf. [36, Theorem 3.1]). The set $C_0 = \{z \in D \mid u(z) = u_0\}$ then defines a closed 1-dimensional analytic manifold, and it has no boundary by (4.10) and the maximum principle. It can be oriented in such a way that C_0 has the properties of he chain C used in (4.8) and $\partial u/\partial n > 0$ everywhere on C_0. Because of this, we then have

$$\oint_{C_0} \frac{\partial u(\zeta)}{\partial n}\, ds_\zeta > 0. \tag{4.11}$$

Since the total flux from K_1 to K_2 is defined independently of the chain C or C_0 (cf. [32, Theorem II.1.1]), the inequality (4.11) contradicts assumption (4.8), which proves the lemma.

LEMMA 4.12. *Define* $S := \operatorname{supp}(\mu) \subseteq (-1, 1)$, $\Omega := \overline{\mathbb{C}} \backslash S$, *and assume that* $\operatorname{cap}(S) > 0$. *Then for every probability measure* ν *on* S^{-1} *we have*

$$\min_{\zeta \in \mathbb{T}} g(\Omega, \nu; \zeta) \leqslant \frac{1}{\operatorname{cap}(\mathbb{T}, S)} \leqslant \max_{\zeta \in \mathbb{T}} g(\Omega, \nu; \zeta). \tag{4.13}$$

In (4.13) *we have either proper inequalities or equalities simultaneously at both places. In the second case we have*

$$g(\Omega, \nu; \cdot) = p_{S, S^{-1}}, \tag{4.14}$$

and this case holds if and only if $\nu = \omega_{S^{-1}, S}$, *where* $\omega_{S^{-1}, S}$ *is the condenser equilibrium distribution on* S^{-1} *of the condenser* (S^{-1}, S) *introduced in Definition* 2.23.

Proof. From the symmetry of the domain $\overline{\mathbb{C}} \setminus (S \cap S^{-1})$ with respect to $\mathbb{T}$ we conclude that

$$p_{S, S^{-1}}(z) = \frac{1}{\operatorname{cap}(\mathbb{T}, S)} \qquad \text{for} \quad z \in \mathbb{T}. \tag{4.15}$$

We assume that

$$\frac{1}{\operatorname{cap}(\mathbb{T}, S)} \geqslant \max_{\zeta \in \mathbb{T}} g(\Omega, \nu; \zeta) \tag{4.16}$$

and show then that (4.14) holds true.

Because of (4.16) and (4.15) the difference

$$d(z) := p_{S, S^{-1}}(z) - g(\Omega, \nu; z) \tag{4.17}$$

satisfies $d(z) \geqslant 0$ for all $z \in \mathbb{T}$. The function d is bounded from above and from below in $\mathbb{D}$, and from the definitions of $p_{S, S^{-1}}$ and $g(\Omega, \nu, \cdot)$ we know that $d(z) = 0$ for quasi every $z \in S$. Because of the generalized minimum principle for harmonic functions (cf. [32, Theorem I.2.4]) we therefore can conclude that

$$d(z) \geqslant 0 \qquad \text{for all} \quad z \in \mathbb{D}. \tag{4.18}$$

Since both functions $p_{S, S^{-1}}$ and $g(\Omega, \nu; \cdot)$ are subharmonic in $\mathbb{D}$, it follows from the Riesz decomposition theorem (cf. [32, Theorem II.3.1]) that we can represent the function d as

$$d(z) = h(z) + g(\mathbb{D}, \sigma; z) \qquad \text{for} \quad z \in \overline{\mathbb{D}} \tag{4.19}$$

with h harmonic in $\mathbb{D}$ and σ a signed measure on S satisfying $\sigma(S) = 0$. Actually, we know that $\sigma = \hat{\nu} - \omega_{S, S^{-1}}$ with $\hat{\nu}$ the so-called balayage measure of ν out of Ω onto S.

If we now assume that $d \not\equiv 0$ on $\mathbb{T}$, then $h(z) > 0$ for all $z \in \mathbb{D}$, which implies the existence of $c \in \mathbb{R}$ such that

$$g(z) = g(\mathbb{D}, \sigma; z) \leqslant c < 0 \qquad \text{for} \quad z \in S. \tag{4.20}$$

Since $\sigma(S) = 0$, the total flux of g from $\mathbb{T}$ to S is zero, and hence Lemma 4.7 can be applied, which shows that (4.20) is impossible. Thus, we have shown that $d(z) = 0$ for all $z \in \mathbb{T}$. With the same argumentation as that leading to (4.18), we then conclude that $d(z) \leqslant 0$ for all $z \in \mathbb{D}$, which proves that

$$d(z) = 0 \qquad \text{for} \quad z \in \mathbb{D}. \tag{4.21}$$

Summarizing the chain of argumentation starting with assumption (4.16), we have shown that the second inequality in (4.13) is either a strict one or we have (4.21), and consequently also equality at both places in (4.13). We shall arrive at the same conclusions if we start the analysis by the first inequality in (4.13) instead of the second one.

It remains to show that (4.21) implies (4.14). Identity (4.14) is equivalent to $\nu = \omega_{S^{-1}, S}$. Since $\operatorname{supp}(\nu) \subseteq S^{-1}$, the Green potential $g(\Omega, \nu; \cdot)$ is harmonic in $\overline{\mathbb{C}} \backslash (S \cup S^{-1})$, and it follows therefore from reflection on $\mathbb{T}$ that (4.21) extends to $\overline{\mathbb{C}}$. Identity (4.14) then follows from (4.17). ▮

Simultaneous Proof of Theorems 1 *through* 4 *and* 6. Let S denote $\operatorname{supp}(\mu)$ and $\Omega := \overline{\mathbb{C}} \backslash S$. In the main stream of the proof we assume

$$\operatorname{cap}(S) > 0. \tag{4.22}$$

The case $\operatorname{cap}(S) = 0$ will then be discussed further below. Using a Moebius transform of the independent variable, if necessary, it can be assumed without los of generality that $0 \notin I(\mu) =: [a, b]$. Throughout the proof $\hat{R}_n = p_n / q_n$ denotes a stationary approximant of degree $n \in \mathbb{N}$ as introduced in Definition 2.48. From Lemma 2.22 in combination with Lemma 3.2 we know that the denominator q_n of $\hat{R}_n$ is of degree n and has only simple zeros, which are all contained in the interval $I(\mu)$. Note that because of $0 \notin I(\mu)$ we have always $\deg(\widetilde{q}_n) = n$. The error function $f - \hat{R}_n$ has the same zeros as $\widetilde{q}_n{}^2$ plus an additional zero at infinity. Thus, there are $2n$ interpolation points at finite distance, which are all contained in $I(\mu)^{-1}$, and, with the terminology of Definition 3.5 and Lemma 3.10, we have

$$w_{2n}(z) = \widetilde{q}_n(z)^2. \tag{4.23}$$

As a consequence we know from the error formula (3.12) in Lemma 3.10 that the interpolation error can be represented as

$$(f-\hat{R}_n)(z)=\frac{\tilde{q}_n(z)^2}{q_n(z)^2}\int\frac{q_n(t)^2}{\widetilde{q}_n(t)^2}\frac{d\mu(t)}{t-z}. \tag{4.24}$$

We introduce the constants

$$c_n:=\int\frac{q_n(t)^2}{\widetilde{q}_n(t)^2}\,d\mu(t), \qquad n\in\mathbb{N}, \tag{4.25}$$

and assume that the denominator polynomials q_n are monic, i.e., $q_n(z)=\prod_{j=1}^{n}(z-z_{jn})$.

By standard compactness arguments in $\mathbb{R}$ and in the sets of positive measures on $\mathbb{C}$ (Helly's selection Theorem), we know that any infinite subsequence $N_0\subseteq\mathbb{N}$ contains an infinite subsequence $N\subseteq N_0$ such that the two limits

$$\lim_{n\to\infty,\,n\in N}\frac{1}{2n}\log c_n:=c_0\in\mathbb{R}\cup\{-\infty\}, \tag{4.26}$$

$$\frac{1}{n}\sum_{j=1}^{n}\delta_{z_{jn}}=:\nu_n\xrightarrow{*}\nu \qquad \text{as} \quad n\to\infty, \quad n\in N, \tag{4.27}$$

hold true. From the orthogonality (3.11) in Lemma 3.10, which holds for q_n if w_{2n} is replaced by $\widetilde{q}_n{}^2$ in accordance with (4.23), it follows that q_n has in any interval of $I(\mu)\backslash\operatorname{supp}(\mu)$ at most one zero. As a consequence we see that ν is a probability measure with support in S. By $\tilde{\nu}$ we denote the image measure of ν under the mapping $t\mapsto 1/\bar{t}$. Thus, $\tilde{\nu}$ is a probability measure on S^{-1}. It follows from (3.1), (4.23), (4.27) and Definition 3.14 that $\tilde{\nu}$ is the asymptotic distribution of the interpolation points belonging to the sequence $\{\hat{R}_n\}_{n\in N}$.

It is not difficult to verify that for any compact set $V\subseteq\Omega=\overline{\mathbb{C}}\backslash S$ we have

$$0<\max\left(\min_{z\in V,\,t\in S}\left|Re\left(\frac{1}{z-t}\right)\right|,\ \min_{z\in V,\,t\in S}\left|Im\left(\frac{1}{z-t}\right)\right|\right)$$

$$\leqslant\max_{z\in V,\,t\in S}\frac{1}{|z-t|}<\infty. \tag{4.28}$$

These estimates imply with (4.25) and $\{z_{1n},...,z_{nn}\}\subseteq I(\mu)$ that the limit

$$\lim_{n\to\infty,\,n\in N}\frac{1}{2n}\log\left|\frac{1}{c_n}\int\frac{q_n(t)^2}{\widetilde{q}_n(t)^2}\frac{d\mu(t)}{t-z}\right|=0 \tag{4.29}$$

holds locally uniformly for $z \in \Omega$ (cf. [35, Proof of Theorem 6.1.8]). From (4.24), (4.29), the limits (4.26) and (4.27), and the definition of $\widetilde{q}_n$ in (3.1) we conclude that the limit

$$\lim_{n \to \infty,\, n \in N} \frac{1}{2n} \log |(f - \hat{R}_n)(z)| = c_0 + \int \log \left| \frac{1 - tz}{z - t} \right| d\nu(t)$$

$$=: h(z) \tag{4.30}$$

holds locally uniformly for $z \in \overline{\mathbb{C}} \backslash (I(\mu) \cup I(\mu)^{-1})$. On $I(\mu)$ and $I(\mu)^{-1}$ limit (4.30) does in general not hold uniformly, however potential theory allows to give a meaning to the limit via the principle of descent, which will be done further below.

It is immediate that $|(1 - tz)/(z - t)| = 1$ for $z \in \mathbb{T}$ and $t \in S$. Thus, we deduce from (4.30) that

$$h(z) = c_0 \qquad \text{for} \quad z \in \mathbb{T}. \tag{4.31}$$

Since $g(\mathbb{D}, \nu; z) = \int \log(|1 - tz|/|z - t|)\, d\nu(t)$ it follows from (4.30) that

$$h(z) = \begin{cases} c_0 + g(\mathbb{D}, \nu; z) & \text{for} \quad z \in \overline{\mathbb{D}} \\ c_0 - g(\overline{\mathbb{C}} \backslash \overline{\mathbb{D}}, \tilde{\nu}; z) & \text{for} \quad z \in \overline{\mathbb{C}} \backslash \mathbb{D}. \end{cases} \tag{4.32}$$

The representation of h in the second line follows from the first line of (4.32) by reflection on $\mathbb{T}$.

In order to derive more properties of the function h we use Proposition 3.17. Since the $\hat{R}_n$, $n \in N$ are rational interpolants and $\tilde{\nu}$ is the asymptotic distribution of the interpolation points as $n \to \infty$, $n \in N$, we conclude from (3.18) in Proposition 3.17 that

$$h(z) \leqslant -g(\Omega, \tilde{\nu}; z) \qquad \text{for} \quad z \in \Omega. \tag{4.33}$$

In a first step we deduce from (4.31), (4.33),

$$c_0 \leqslant -g(\Omega, \tilde{\nu}; z), \qquad z \in \mathbb{T}.$$

From (4.13) in Lemma 4.12, we know that

$$-\max_{z \in \mathbb{T}} g(\Omega, \tilde{\nu}; z) \leqslant \frac{-1}{\operatorname{cap}(\mathbb{T}, S)};$$

hence

$$c_0 \leqslant \frac{-1}{\operatorname{cap}(\mathbb{T}, S)}. \tag{4.34}$$

Since the right-hand side is independent of the selection of the subsequences N_0 and N, this last inequality together with (4.30) and (4.31) implies that

$$\limsup_{n\to\infty} \frac{1}{2n} \log |(f-\hat{R}_n)(z)| \leqslant \frac{-1}{\operatorname{cap}(\mathbb{T}, S)} \tag{4.35}$$

holds uniformly for $z \in \mathbb{T}$, which proves Theorem 1 in case $\operatorname{cap}(S) > 0$, as it has been assumed in (4.22). If $\operatorname{cap}(S) = 0$, then $g_\Omega(z, x) = \infty$ for all $z, x \in \Omega$, and consequently $g(\Omega, \tilde{v}; z) = \infty$ for $z \in \Omega$. From (4.33) we then deduce that

$$h(z) = -\infty \qquad \text{for} \quad z \in \Omega, \tag{4.36}$$

which shows that (4.35) holds also in case of $\operatorname{cap}(S) = 0$.

Next we prove Theorem 2. We first show that

$$h(z) \leqslant -p_{S, S^{-1}}(z) \qquad \text{for} \quad z \in \mathbb{D}. \tag{4.37}$$

Indeed, if we consider the function

$$d(z) := -h(z) - p_{S, S^{-1}}(z), \tag{4.38}$$

then it follows from (4.15), (4.31) and (4.34) that $d(z) \geqslant 0$ for all $z \in \mathbb{T}$. From (4.33) and the properties of the Green potential $g(\Omega, \tilde{v}; \cdot)$ on S we deduce that

$$\limsup_{z\to z',\, z\notin S} h(z) \leqslant -\limsup_{z\to z',\, z\notin S} g(\Omega, \tilde{v}; z) = 0 \tag{4.39}$$

for quasi every $z' \in S$. Since $p_{S, S^{-1}}(z) = 0$ for quasi every $z \in S$, we see that also $d(z) \geqslant 0$ for quasi every $z \in S$. From (4.38) and (4.39) it is not difficult to deduce that $d(z)$ is bounded from below in $\mathbb{D} \setminus S$. Hence, by the generalized minimum principle of harmonic functions (cf. [32, Theorem I.2.4]) we conclude that $d(z) \geqslant 0$ for all $z \in \overline{\mathbb{D}}$, which proves (4.37).

Using representation (4.32) for h together with (4.3) in Lemma 4.2, we arrive at the following chain of identities and inequalities:

$$\begin{aligned} h(z) &= c_0 - g(\overline{\mathbb{C}} \setminus \overline{\mathbb{D}}, \tilde{v}; z) = c_0 - g(\mathbb{D}, v; 1/\bar{z}) \\ &\leqslant c_0 - \log \min\left(\left| \frac{1 - a/\bar{z}}{1/\bar{z} - a} \right|, \left| \frac{1 - b/\bar{z}}{1/\bar{z} - b} \right| \right) \\ &= c_0 + \log \max\left(\left| \frac{1 - az}{z - a} \right|, \left| \frac{1 - bz}{z - b} \right| \right) \end{aligned} \tag{4.40}$$

for $z \in \overline{\mathbb{C}} \setminus \mathbb{D}$.

The right-hand sides of (4.37) and (4.40) are independent of the selection of the subsequences N_0 or N, and therefore we conclude from (4.37), (4.40), (4.30), and (4.31) that

$$\limsup_{n\to\infty} \frac{1}{2n} \log |(f-\hat{R}_n)(z)| \leqslant \begin{cases} -p_{S,S^{-1}}(z) & \text{for} \quad z\in\overline{\mathbb{D}} \\ \dfrac{-1}{\operatorname{cap}(\mathbb{T},S)} + \log\max\left(\left|\dfrac{1-az}{z-a}\right|, \left|\dfrac{1-bz}{z-b}\right|\right) & \text{for} \quad z\in\overline{\mathbb{C}}\backslash\mathbb{D} \end{cases} \tag{4.41}$$

holds locally uniformly for $z\in\overline{\mathbb{C}}\backslash(S\cup S^{-1})$. Since the right-hand side of (4.41) is continuous in $\overline{\mathbb{C}}\backslash\overline{\mathbb{D}}$, it follows from the principle of descent in potential theory (cf. [32, Theorem I.6.8]) together with (4.24), (4.27), and (4.30) that the limit (4.41) holds also uniformly on S^{-1}. In this extended form (4.41) implies Theorem 2 in case $\operatorname{cap}(S)>0$, as was assumed in (4.22). If $\operatorname{cap}(S)=0$ it follows from (4.36) that also in this case Theorem 2 holds true.

In the remainder of the proof we assume that the measure $\mu\in\mathbf{Reg}$ in the sense of Definition 2.13. This implies that $\operatorname{cap}(S)>0$, and further it follows from part (ii) of Proposition 3.17 that instead of (4.33) we now have the stricter version

$$h(z) = -g(\Omega, \tilde{\nu}; z) \qquad \text{for} \quad z\in\Omega. \tag{4.42}$$

From (4.31) we then conclude that

$$c_0 = -g(\Omega, \tilde{\nu}; z) \qquad \text{for} \quad z\in\mathbb{T}. \tag{4.43}$$

Reflection on $\mathbb{T}$ further shows that

$$g(\Omega, \tilde{\nu}; z) = -2c_0 \qquad \text{for quasi every} \quad z\in S^{-1}. \tag{4.44}$$

A comparison of (4.44) with Definition 2.1 shows that

$$g(\Omega, \tilde{\nu}; \cdot) = p_{S,S^{-1}}, \tag{4.45a}$$

$$2c_0 = \frac{-1}{\operatorname{cap}(S^{-1},S)} = \frac{-2}{\operatorname{cap}(\mathbb{T},S)}, \tag{4.45b}$$

$$\tilde{\nu} = \omega_{S^{-1},S}, \qquad \nu = \omega_{S,S^{-1}} = \omega_{S,\mathbb{T}}. \tag{4.45c}$$

Since because of (4.45a) and (4.45b) the right-hand sides of (4.42) and (4.43) are independent of the subsequences N_0 and N, it follows from (4.30) and (4.31) that in (4.35) a proper limit and equality instead of an inequality holds true on $\mathbb{T}$. This proves Theorem 3.

In order to prove Theorem 4 we derive from (4.45a), (4.42), and (4.30) with the same arguments as that used in the proof of Theorem 3 that

$$\lim_{n\to\infty} \frac{1}{2n} \log |(f-\hat{R}_n)(z)| = -p_{S,\, S^{-1}}(z) \tag{4.46}$$

holds locally uniformly for $z \in \overline{\mathbb{C}} \setminus (S \cup S^{-1})$, which proves limit (2.18) in Theorem 4. Limit (2.19) then follows by the principle of descent in the same way as after (4.41) the asymptotic estimate (2.12) had been proved for $z \in S^{-1}$.

From (4.23) and the fact that the poles and zeros of $\hat{R}_n$ interlace, it follows from (4.27) and the observations after (4.27) that for the subsequence N and the same notation as used in (2.30) we have

$$\begin{aligned} &\frac{1}{n} \nu_{P(\hat{R}_n)} \xrightarrow{*} \nu, \\ &\frac{1}{n} \nu_{Z(\hat{R}_n)} \xrightarrow{*} \nu \qquad \text{as} \quad n \to \infty, \quad n \in N, \\ &\quad \frac{1}{2n} \nu_{I_n} \xrightarrow{*} \tilde{\nu}. \end{aligned} \tag{4.47}$$

From (4.45c) we know that the measures ν and $\tilde{\nu}$ do not depend on the selection of the subsequences N_0 and N. Therefore, (4.47) and (4.45c) imply the limits (2.30), which proves Theorem 6. ∎

Proof of Theorem 5. It is immediate that

$$\|f - R^*_{n,\infty}\|_{\mathbb{T}} \leqslant \max_{\zeta \in \mathbb{T}} |(f - R^*_n(\zeta))|. \tag{4.48}$$

On the other hand we also have

$$\|f - R^*_{n,\infty}\|_{\mathbb{T}} \geqslant \|f - R^*_{n,\infty}\| \geqslant \|f - R^*_n\| \geqslant \min_{\zeta \in \mathbb{T}} |(f - R^*_n(\zeta))|. \tag{4.49}$$

From (4.48), (4.49), (4.15) and limit (2.18) in Theorem 4 restricted to $\mathbb{T}$ the limit (2.21) of Theorem 5 follows. ∎

5. PROOFS OF THE THEOREMS 7 AND 8 AND RELATED RESULTS

We start the section with the proofs of the two Lemmas 2.37 and 2.40, which are related to the special Szegő function $D(\mu; \cdot)$ introduced in Definition 2.38.

Proof of Lemma 2.37. From (2.32) and (2.35) it follows that $\dot{\mu}(x) = \mu'(x)\,\pi\,\sqrt{(x-a)(b-x)}$, $x \in [a, b]$, and from the first line of (2.27) together with the definition of $\omega_{[a,b]}$ we know that $d\omega_{[a,b],\mathbb{T}}(x) = \pi(2K)^{-1}\,[(1-ax)(1-bx)]^{-1/2}\,d\omega_{[a,b]}(x)$, $x \in [a, b]$, where K denotes the complete elliptic integral of the first kind as given in (2.8). From both formulae we conclude that condition (2.33) holds true if, and only if,

$$\int \log \dot{\mu}(x)\, d\omega_{[a,b],\mathbb{T}}(x) > -\infty, \tag{5.1}$$

which proves Lemma 2.37. ∎

In the proof of Lemma 2.40, and also at later places, some properties of the solution of certain generalized Dirichlet problems are needed, and they are put togther in the next lemma. We note that in the lemma it is not assumed that the boundary function of the Dirichlet problem is continuous, which explains some of the precaucious formulations.

LEMMA 5.2. *Let R be the ring domain $\mathbb{D}\backslash[a, b]$ and $\omega_{R,z}$ the harmonic measure representing $z \in R$ on ∂R. Let further $u \in L^1(\omega_{[a,b]})$ be a function on $[a, b]$. On ∂R a boundary function v is defined by $v(x \pm i0) := u(x)$ for $x \in [a, b]$ and $v(z) = 0$ for $z \in \mathbb{T}$.*

(i) *The function v is integrable with respect to every harmonic measure $\omega_{R,z}$, $z \in R$.*

(ii) *A generalized solution h of the Dirichlet problem on R with boundary function v exists and is unique if the representation.*

$$h(z) = \int v\, d\omega_{R,z}, \qquad z \in R, \tag{5.3}$$

is demanded to hold true. The function h is harmonic in $\mathbb{D}\backslash[a, b]$ and on $\mathbb{T}$, it has non-tangential boundary values from both sides almost everywhere on $[a, b]$, we have

$$\lim_{y \to 0+} h(x \pm iy) = u(x) \qquad \text{for almost every} \quad x \in [a, b], \tag{5.4}$$

and we have $h(z) = 0$ for all $z \in \mathbb{T}$.

(iii) *The total flux across $\mathbb{T}$ is given by*

$$\frac{1}{2\pi} \oint_{\mathbb{T}} \frac{\partial}{\partial n} h(\zeta)\, ds_\zeta = \operatorname{cap}([a, b], \mathbb{T}) \int u(x)\, d\omega_{[a,b],\mathbb{T}}(x) \tag{5.5}$$

with $\partial/\partial n$ denoting the normal derivative on $\mathbb{T}$, ds the line element, $\omega_{[a,b],\mathbb{T}}$ the equilibrium distribution on $[a,b]$ of the condenser $([a,b],\mathbb{T})$ as introduced in Definition 2.23, *and $\omega_{[a,b]}$ the (ordinary) equilibrium distribution on $[a,b]$.*

Remark. Since in (5.4) it only has been demanded that h and u agree on $[a,b]$ almost everywhere, the solution of the Dirichlet problem would not be unique if representation (5.3) were not demanded. For a definition of the harmonic measures $\omega_{R,z}$, $z\in R$, we refer to [32, Appendix B], or [31, Chapter 4, Section 3]. The measure $\omega_{R,z}$ is supported on ∂R and depends harmonically on z for fixed Borel set $B\subseteq\partial R$, i.e., $\omega_{R,z}(B)$ is an harmonic function of z.

Proof. (i) As reference to a proof of existence and properties of generalized solutions to a Dirichlet problem we use Appendix A of [32]. In step V of the proof of Theorem A.3.1 in [32] it is proved that for any pair of points z_1, $z_2\in R$ there exists a constant $c_{z_1z_2}<\infty$ such that $\omega_{R,z_1}\leqslant c_{z_1z_2}\omega_{R,z_2}$; this mutual absolute continuity of harmonic measures is also proved in [31, Corollary 4.3.5]. Let $\omega_{R,z}|_{[a,b]}$ denotes the restriction of $\omega_{R,z}$ to $[a,b]$. Since $\omega_{[a,b]}$ is equal to the harmonic measure $\omega_{\overline{\mathbb{C}}\setminus[a,b],\infty}$ of the domain $\overline{\mathbb{C}}\setminus[a,b]$ at infinity (see e.g. [31, Theorem 4.3.14]), it follows from the subordination principle for harmonic measure ([31, Theorem 4.3.8]) applied to the injection $R\to\overline{\mathbb{C}}\setminus[a,b]$ and from the mutual absolute continuity of the measures $\omega_{\overline{\mathbb{C}}\setminus[a,b],z}$ that $\omega_{R,z}|_{[a,b]}\leqslant c_z\omega_{[a,b]}$. From the last inequality and the assumption that $u\in L^1(\omega_{[a,b]})$ it follows that the boundary function v is integrable with respect to each $\omega_{R,z}$, $z\in R$.

(ii) It is proved in [32, Theorem A.3.1], that it follows from part (i) that the function h defined in (5.3) satisfies what is called the Perron–Wiener–Brelot solution of the generalized Dirichlet problem. The other assertions in Section (ii) follow from the definition of the Perron–Wiener–Brelot solution (cf. [32, Appendix 2]). A proof of (5.4) follows independently from Theorem I.5.3 in [13].

(iii) From Green's formula we know that for two functions h, g harmonic in R we have

$$\oint_{\partial R}\left(g\frac{\partial h}{\partial n}-h\frac{\partial g}{\partial n}\right)ds=0. \tag{5.6}$$

We choose $g=p_{[a,b],\mathbb{T}}$, where $p_{[a,b],\mathbb{T}}$ is the equilibrium potential introduced in Definition 2.1. We have $p_{[a,b],\mathbb{T}}(z)=0$ for $z\in[a,b]$, and

$$d\omega_{[a,b],\mathbb{T}}(x)=\frac{1}{\pi}\frac{\partial}{\partial n}p_{[a,b],\mathbb{T}}(x)\,dx, \qquad x\in[a,b]. \tag{5.7}$$

Indeed,

$$p_{[a,b],\mathbb{T}}(z) = \frac{1}{\operatorname{cap}([a,b],\mathbb{T})} - p_{\mathbb{T},[a,b]}(z),$$

and

$$\begin{aligned} p_{\mathbb{T},[a,b]}(z) &= \int \log\left|\frac{1-zt}{z-t}\right| d\omega_{[a,b],\mathbb{T}}(t) \\ &= \int \log|1-zt|\, d\omega_{[a,b],\mathbb{T}}(t) + \int \log\frac{1}{|z-t|}\, d\omega_{[a,b],\mathbb{T}}(t). \end{aligned}$$

The first term of the previous sum is harmonic across $[a,b]$ while the second equals the logarithmic potential of the measure $\omega_{[a,b],\mathbb{T}}$. Hence, equality (5.7) follows from [32, Chapter II, Theorem 1.5]. With (5.3) and (5.6) it then follows that

$$\begin{aligned} \frac{1}{\operatorname{cap}([a,b],\mathbb{T})} \oint_{\mathbb{T}} \frac{\partial}{\partial n} h(\zeta)\, ds_\zeta &= 2\int u(x) \frac{\partial}{\partial n} p_{[a,b],\mathbb{T}}(x)\, dx \\ &= 2\pi \int u(x)\, d\omega_{[a,b],\mathbb{T}}(x), \end{aligned} \tag{5.8}$$

which proves (5.5). ∎

An example of a generalized solution to the Dirichlet problem is the restriction to $\mathbb{D}\backslash[a,b]$ of $\log|G(\mu,z)|$, where G is the Szegő function given in Definition 3.19. Specifically, we have the following lemma.

Lemma 5.9. *Let again R be the ring domain $\mathbb{D}\backslash[a,b]$ and $\omega_{R,z}$ the harmonic measure representing $z \in R$ on ∂R. Define v_G to be* $\log|G|$ *on $\mathbb{T}$ and* $\log(\dot{\mu}(x))/2$ *for $x \in [a,b]$. Then*

$$\log|G(\mu,z)| = \int v_G\, d\omega_{R,z}, \qquad z \in R. \tag{5.10}$$

Proof. Assume first that $\log(\dot{\mu}(x))$ satisfies an Hölder condition on $[a,b]$. Then, from the Plemelj formulae applied to (3.20), we have that

$$\lim_{z\to x} \log|G(\mu,z)| = \tfrac{1}{2}\log\dot{\mu}(x), \qquad x \in (a,b).$$

Moreover, the integral in the right-hand side of (3.21) is of the order of $|z-a|^{-1/2}$ (resp. $|z-b|^{-1/2}$) as z tends to a (resp. to b), cf. [26]. Therefore, $\log|G|$ remains bounded as z approaches $[a,b]$, and since it is regular across $\mathbb{T}$, it is a bounded harmonic function that converges to v_G on $(a, b) \cup \mathbb{T}$, hence nearly everywhere on ∂R. Thus, it is indeed a solution to the generalized Dirichlet problem (cf. [32, Chapter 1, Lemma 2.6] or [31, Corollary 4.2.6]) so that (5.10) holds. Since Hölder continuous functions are dense in $L^1(\omega_{[a,b]})$, the general case follows now by density, recalling that $\omega_{R,z}$ is absolutely continuous with respect to $\omega_{[a,b]}$, for each $z \in R$. ∎

Proof of Lemma 2.40. Because of the possibility to adjust the problem via a Moebius transform $\psi: \overline{\mathbb{C}} \to \overline{\mathbb{C}}$, which leaves $\mathbb{T}$ invariant, we can assume without loss of generality that

$$0 < a, \tag{5.11}$$

which implies that $[a, b] \subseteq (0, 1)$ and therefore $[a, b]^{-1}$ is the interval $[1/b, 1/a] \subseteq \mathbb{R}_+$. Let u be the generalized solution of the Dirichlet problem in the ring domain $R := \overline{\mathbb{C}} \setminus ([a, b] \cup [a, b]^{-1})$ with boundary function

$$u(x) := \begin{cases} \frac{1}{2} \log(\dot{\mu}(x)/D(\mu)) & \text{for} \quad x \in [a, b] \\ -\frac{1}{2} \log(\dot{\mu}(1/x)/D(\mu)) & \text{for} \quad x \in [a, b]^{-1}. \end{cases} \tag{5.12}$$

The function $\dot{\mu}$ and the constant $D(\mu)$ have been introduced in Definition 2.34. Because of the Szegő condition (2.33) we know that $\log(\dot{\mu}(\cdot)/D(\mu)) \in L^1(\omega_{[a,b]})$, i.e., $D(\mu) > 0$, and the existence of the generalized solution u follows from Lemma 5.2 if one takes into consideration reflection on $\mathbb{T}$. We have $u(z) = u(\bar{z})$ and $u(1/z) = -u(z)$ for $z \in R$, which implies that

$$u(z) = 0 \qquad \text{for} \quad z \in \mathbb{T}. \tag{5.13}$$

From Lemma 5.2 we know that non-tangential boundary values exist almost everywhere on $\partial R = [a, b] \cup [a, b]^{-1}$, and we have

$$u(x \pm i0) = u(x) \qquad \text{for almost every} \quad x \in [a, b] \cup [a, b]^{-1}. \tag{5.14}$$

As a consequence of (5.5) in Lemma 5.2 and (2.36) in Definition 2.34 we deduce that

$$\begin{aligned} \frac{1}{2\pi} \oint_{\mathbb{T}} \frac{\partial}{\partial n} u(\zeta)\, dz_\zeta &= \operatorname{cap}([a, b], \mathbb{T}) \int_a^b u(x)\, d\omega_{[a,b],\mathbb{T}}(x) \\ &= \frac{1}{2} \operatorname{cap}([a, b], \mathbb{T}) \left[\int \log \dot{\mu}(x)\, d\omega_{[a,b],\mathbb{T}}(x) - \log D(\mu) \right] \\ &= 0. \end{aligned} \tag{5.15}$$

Because of this identity the conjugate functions v to u are single-valued in R, and consequently

$$g_1(z) := u(z) + iv(z), \qquad \text{for} \quad z \in R, \quad v(1) := 0, \tag{5.16}$$

is an analytic (single-valued) function in R. We define

$$g_2(z) := \frac{g_1(z)}{\sqrt{(z-a)(z-b)(1-az)(1-bz)}} \tag{5.17}$$

and assume that for $z = 1$ the square root in (5.17) is negative, which implies that for $z = x + i0$, $x \in [a, b]$, the square root is negative imaginary. By the superscripts $^+$ and $^-$ we mark the upper and lower boundary values on $[a, b] \cup [a, b]^{-1}$. It is immediate that on $[a, b] \cup [a, b]^{-1}$ we have $u^+ = u^-$ and $v^+ = -v^-$. For $x \in [a, b]$ we have $u^+(x) = -u^+(1/x)$ and $v^+(x) = -v^-(1/x) = v^+(1/x)$. From (5.16) and (5.17) we then deduce that

$$x^2 \operatorname{Re} g_2^+(x) = x^2 \operatorname{Re} g_2^-(x) = -\operatorname{Re} g_2^+(1/x) = -\operatorname{Re} g_2^-(1/x) \qquad \text{and}$$
$$x^2 \operatorname{Im} g_2^+(x) = -x^2 \operatorname{Im} g_2^-(x) = \operatorname{Im} g_2^+(1/x) = -\operatorname{Im} g_2^-(1/x) \tag{5.18}$$

for $x \in [a, b]$. Further, we have $g_2(\infty) = g_2'(\infty) = 0$. Let $C = C_1 + C_2$ be an integration path in R with C_1 and C_2 encompassing $[a, b]$ and $[a, b]^{-1}$ in the clockwise direction, respectively, such that from Cauchy's Theorem and the identities (5.18) we have

$$\begin{aligned} g_2(z) &= \frac{1}{2\pi i} \oint_C \frac{g_2(\zeta)\, d\zeta}{\zeta - z} = \frac{1}{2\pi i} \left[\int_a^b + \int_{1/b}^{1/a}\right] \frac{g_2^+(x) - g_2^-(x)}{x - z}\, dx \\ &= \frac{1}{\pi} \int_a^b \operatorname{Im} g_2^+(x) \left[\frac{1}{x-z} + \frac{1}{1/x - z}\right] dx \\ &= \frac{1}{\pi} \int_a^b \operatorname{Im} g_2^+(x)\, \frac{1 - 2xz + x^2}{(x-z)(1-xz)}\, dx \end{aligned} \tag{5.19}$$

for $z \in R$. From (5.12), (5.16), and (5.17) it follows that

$$\operatorname{Im} g_2^+(x) = \frac{\log(\dot{\mu}(x)/D(\mu))}{2\sqrt{(x-a)(b-x)(1-ax)(1-bx)}}, \qquad \text{for} \quad x \in [a, b], \tag{5.20}$$

with the square root assumed to be positive for $x \in [a, b]$. From (5.19), (5.20), and (2.39) in Definition 2.38 we deduce that

$$D(\mu; z) = \exp g_1(z) = \exp(u(z) + iv(z)), \qquad \text{for} \quad z \in R, \tag{5.21}$$

and further

$$|D(\mu; z)|^2 = \exp(2u(z)), \qquad \text{for} \quad z \in R. \tag{5.22}$$

Since g_1 is analytic in R, assertion (i) of Lemma 2.40 follows immediately from (5.21). Because of (5.21) we further have

$$\arg D(\mu; z) = v(z), \qquad \text{for} \quad z \in R, \tag{5.23}$$

and assertion (ii) therefore is a consequence of (5.15), since the inner normal derivative times arc length is the differential of the conjugate function. Assertion (iii) follows from (5.21), (5.22), and (5.14) together with (5.12). At last, identity (2.42) is a consequence of (5.13). ∎

Proposition 3.25, i.e., Markov's Theorem for multipoint Padé approximants will be a major tool in the proof of the Theorems 7 and 8. In this connection a comparison between the two different Szegő functions $D(\mu; \cdot)$ and $G(\mu; \cdot)$ introduced in Definitions 2.38 and 3.19, respectively, may be helpful. It will be sufficient to do this comparison on $\overline{\mathbb{D}}$.

LEMMA 5.24. *Let the function F be defined by*

$$F(z) := G(\mu; z)/D(\mu; z) \qquad \textit{for} \quad z \in \overline{\mathbb{D}}. \tag{5.25}$$

Then we have

(i) *The function F is analytic and different from zero on $\overline{\mathbb{D}} \backslash [a, b]$, and it has non-tangential boundary values from both sides of $[a, b]$, almost everywhere.*

(ii) *For the increment of* $\arg F$ *along* $\mathbb{T}$ *we have* $\Delta_{t=0}^{2\pi} \arg F(e^{it}) = 0$.

(iii) *The function $|F|$ is continuous in $\overline{\mathbb{D}}$, and we have*

$$|F(z)|^2 = D(\mu) \qquad \textit{for} \quad z \in [a, b]. \tag{5.26}$$

(iv) *We have*

$$|F(z)|^2 = |G(\mu; z)|^2 \qquad \textit{for} \quad z \in \mathbb{T}. \tag{5.27}$$

Proof. Assertion (i) follows from the assertions (i) and (ii) in Lemma 2.40 together with assertion (i) in Lemma 3.22.

Since $G(\mu; \cdot)$ is analytic and different from zero in $\overline{\mathbb{C}} \backslash \mathbb{D}$, the argument principle implies that $\Delta_{t=0}^{2\pi} \arg G(\mu; e^{it}) = 0$, and therefore assertion (ii) follows from assertion (ii) of Lemma 2.40.

From (2.41) and (3.23) we deduce that

$$|F(x \pm i0)|^2 = D(\mu) \qquad \text{for almost every} \quad x \in [a, b], \tag{5.28}$$

which is less than stated in (5.26). From the definition of $D(\mu; z)$, using (5.3) and (5.12), we get

$$\log |D(\mu; z)| = \int \tfrac{1}{2} \log(\dot{\mu}(x))/D(\mu)\, d\omega_{R,z}|_{[a,b]}, \qquad z \in R.$$

Plugging this and (5.10) in (5.25), the terms involving $\log(\dot{\mu})$ cancel, yielding

$$\log |F(z)| = \tfrac{1}{2} \log D(\mu) \int d\omega_{R,z}|_{[a,b]} + \int \log |G|\, d\omega_{R,z}|_{\mathbb{T}}.$$

We now recognize that $\log |F|$ is the solution to the Dirichlet problem with continuous boundaries data $\frac{1}{2} \log D(\mu)$ on $[a, b]$ and $\log |G|$ on $\mathbb{T}$. Since $[a, b]$ is regular for the Dirichlet problem, (5.26) holds.

Assertion (iv) is an immediate consequence of assertion (iv) in Lemma 2.40. ∎

Proof of the Theorem 7 *and* 8. In order to satisfy the demands of Theorem 9, it would be necessary to consider stationary approximants $\hat{R}_n$ instead of best approximants R_n^*. However, in [8, Theorem 1.3], it has been proved that under the assumption of the Theorems 7 and 8 there exists only one stationary approximant $\hat{R}_n$ for each $n \in \mathbb{N}$ sufficiently large, which is then identical with the uniquely existing best rational approximant R_n^*. In the proof of the Theorems 1–4 it has been shown before (4.23) that the best rational approximants $R_n^* = p_n/q_n$ meet interpolation conditions, and that from the error formula (4.24) and the definition of c_n in (4.25), it follows that

$$e_n(z) := (f - R_n^*)(z) = \frac{\widetilde{q}_n(z)^2}{q_n(z)^2}\, c_n I_n(z) \tag{5.29}$$

with the function I_n defined by

$$I_n(z) := c_n^{-1} \int \frac{q_n(t)^2}{\widetilde{q}_n(t)^2} \frac{d\mu(t)}{t - z}. \tag{5.30}$$

The zeros of $\widetilde{q}_n{}^2$ are interpolation points for R_n^*. All these points are contained in $[a, b]^{-1}$, which implies that they stay away from $[a, b]$ by a positive distance and therefore condition (3.26) in Proposition 3.25 is satisfied, and of course the same is true for (3.8) and (3.9). Since in the Theorems 7 and 8 it also has been assumed that the defining measure μ

belongs to the Szegő class, all assumptions of the Propositions 3.25 and 3.28 are satisfied. From limit (3.29) in Proposition 3.28 we deduce that

$$\lim_{n\to\infty} I_n(z) = \int \frac{d\omega_{[a,b]}(t)}{t-z} = \frac{-1}{\sqrt{(z-a)(z-b)}} =: f_0(z) \tag{5.31}$$

holds locally uniformly for $z \in \overline{\mathbb{C}} \backslash [a, b]$. The sign of the square root in (5.31) is chosen so that $f_0(z) = -1/z + O(z^{-2})$ as $z \to \infty$.

Since $|\widetilde{q}_n(z)/q_n(z)| = 1$ for $z \in \mathbb{T}$, we deduce from the error formula (5.29) that

$$|e_n(z)| = (1 + o(1))\, c_n\, |f_0(z)| \qquad \text{for} \quad z \in \mathbb{T} \tag{5.32}$$

with the Landau symbol $o(1)$ holding uniformly on $\mathbb{T}$ as $n \to \infty$. With the help of Proposition 3.25 we shall extend the asymptotic error estimate (5.32) to $\overline{\mathbb{D}} \backslash [a, b]$. Using the conformal map ψ: $\overline{\mathbb{C}} \backslash [a, b] \to \mathbb{D}$, $\psi(\infty) = 0$, $\psi'(\infty) > 0$, defined in (3.24) we introduce the generalized Blaschke product

$$b_{2n}(z) := \prod_{j=1}^{n} \frac{(\psi(z) - \psi(1/z_{jn}))^2}{(1 - \overline{\psi(1/z_{jn})}\, \psi(z))^2}, \tag{5.33}$$

where z_{jn}, $j = 1, ..., n$, are the n zeros of denominator polynomial q_n. With the notation introduced in (5.29), (5.31), and (5.33), it follows from Proposition 3.25 that

$$e_n(z) = (2 + o(1))\, G(\mu; z)^2\, f_0(z)\, b_{2n}(z) \qquad \text{as} \quad n \to \infty, \tag{5.34}$$

where the Landau symbol $o(1)$ holds locally uniformly for $z \in \overline{\mathbb{C}} \backslash [a, b]$.

Let $1 > \rho = \rho_{[a,b]} > 0$ and φ: $\overline{\mathbb{C}} \backslash ([a, b] \cup [a, b]^{-1}) \to \{\rho < |z| < 1/\rho\}$, $\varphi(b) = \rho$, be those defined in (2.43). Let us consider the sequence of functions

$$g_n(z) := b_{2n}(z)\, \varphi(z)^{2n}\, \rho^{-2n} F(z)^2\, D(\mu)^{-1}, \qquad n \in \mathbb{N}, \tag{5.35}$$

where the function F is defined by (5.25) in Lemma 5.24. It follows from the definitions of the functions b_{2n} and φ and from assertion (iii) of Lemma 5.24 that g_n is analytic in $\overline{\mathbb{D}} \backslash [a, b]$ and $|g_n|$ is continuous in $\overline{\mathbb{D}}$. With (5.26) it further follows that

$$|g_n(z)| = 1 \qquad \text{for} \quad z \in [a, b], \quad n \in \mathbb{N}. \tag{3.36}$$

We now prove that the limit

$$\lim_{n\to\infty} g_n(z) = 1 \tag{5.37}$$

holds locally uniformly for $z \in \mathbb{D} \backslash [a, b]$. From the argument principle, it is clear that

$$\Delta_{t=0}^{2\pi} \arg (b_{2n}(e^{it})) = -2n.$$

In another connection, it follows from the expression of φ as a Green potential (cf. [32, Chapter VIII.6]) and Gauss' theorem, or alternatively from a direct computation using (2.7) (cf. [27, Chapter VI]) that

$$\Delta_{t=0}^{2\pi} \arg (\varphi(e^{it})) = 1.$$

This together with Lemma 5.24 (ii) implies that

$$\Delta_{t=0}^{2\pi} \arg g_n(e^{it}) = 0 \qquad \text{for} \quad n \in \mathbb{N}. \tag{5.38}$$

The definition of g_n together with (5.38) implies that $\log |g_n|$ satisfies condition (4.8) in Lemma 4.7 with $D = \mathbb{D} \backslash [a, b]$. The two components K_1 and K_2 of ∂D are now $[a, b]$ and $\mathbb{T}$. Using Lemma 4.7 in both directions it then follows from (5.36) that

$$\inf_{z \in \mathbb{T}} |g_n(z)| \leqslant 1 \leqslant \sup_{z \in \mathbb{T}} |g_n(z)|, \qquad \text{for} \quad n \in \mathbb{N}. \tag{5.39}$$

On the other hand it follows from (5.32) and (5.34) that

$$c_n = 2(1 + o(1)) \, |G(\mu; z)|^2 \, |b_{2n}(z)| \qquad \text{for} \quad z \in \mathbb{T} \tag{5.40}$$

with the Landau symbol $o(1)$ holding uniformly on $\mathbb{T}$ as $n \to \infty$. With (5.35), (5.27) in Lemma 5.24, and $|\varphi(z)| = 1$ for $z \in \mathbb{T}$, we deduce from (5.40) that

$$|g_n(z)| = (1 + o(1)) \, \tfrac{1}{2} \, c_n \rho^{-2n} \, D(\mu)^{-1} \qquad \text{for} \quad z \in \mathbb{T} \quad \text{as } n \to \infty \tag{5.41}$$

with $o(1)$ again holding uniformly on $\mathbb{T}$. Except for $(1 + o(1))$, we have a constant on the right hand side of (5.41). We therefore deduce from (5.39) and (5.41) that

$$\lim_{n \to \infty} \left[\sup_{z \in \mathbb{T}} |g_n(z)| - \inf_{z \in \mathbb{T}} |g_n(z)| \right] = 0, \tag{5.42}$$

$$\lim_{n \to \infty} \tfrac{1}{2} c_n \rho^{-2n} \, D(\mu)^{-1} = 1. \tag{5.43}$$

From (5.36), (5.41), (5.43) and the maximum principle applied to the harmonic functions $\log |g_n(z)|$, it follows that $\lim_{n \to \infty} |g_n(z)| = 1$ uniformly in $\overline{\mathbb{D}}$. Since g_n is a normal family of analytic functions in $\mathbb{D} \backslash [a, b]$, and since the only possible limit function is 1 as $g_n(0) > 0$, we get (5.37).

From (5.25) we know that $G(\mu;\,\cdot\,) = FD(\mu;\,\cdot\,)$. Inserting this identity into (5.34) we deduce with (5.36) and (5.37) that

$$\begin{aligned} e_n(z) &= (2+o(1))\, f_0(z)\, D(\mu; z)^2\, F(z)^2\, b_{2n}(z) \\ &= (2+o(1))\, f_0(z)\, D(\mu; z)^2\, \varphi(z)^{-2n}\, \rho^{2n}\, D(\mu)\, g_n(z) \\ &= (2+o(1))\, f_0(z)\, D(\mu; z)^2\, \varphi(z)^{-2n}\, \rho^{2n}\, D(\mu) \end{aligned} \tag{5.44}$$

with $o(1)$ holding locally uniformly for $z \in \mathbb{D}\setminus[a, b]$ as $n \to \infty$.

In order to extend (5.44) to $\overline{\mathbb{C}}\setminus(\overline{\mathbb{D}} \cup [a, b]^{-1})$, we use reflection on $\mathbb{T}$. Using the error formula (5.29) twice we deduce that

$$\frac{e_n(1/z)}{I_n(1/z)} = \frac{\widetilde{q}_n(1/z)^2}{q_n(1/z)^2}\, c_n = \frac{q_n(z)^2}{\widetilde{q}_n(z)^2}\, c_n = c_n^2\, \frac{I_n(z)}{e_n(z)} \qquad \text{for} \quad z \in \overline{\mathbb{D}}\setminus[a, b]. \tag{5.45}$$

From (5.43) we know that $c_n = (2+o(1))\, \rho^{2n}\, D(\mu)$ as $n \to \infty$, and from (2.41) and the definition of $D(\mu;\,\cdot\,)$ in (2.39) it follows that

$$D\left(\mu; \frac{1}{z}\right) = \frac{1}{D(\mu; z)} \qquad \text{for} \quad z \in \overline{\mathbb{C}}\setminus([a, b] \cup [a, b]^{-1}). \tag{5.46}$$

Further, it is rather immediate that $\varphi(1/z) = 1/\varphi(z)$. Let now $z \in \overline{\mathbb{C}}\setminus(\overline{\mathbb{D}} \cup [a, b]^{-1})$, then it follows from (5.45) and (5.44) that

$$\begin{aligned} \frac{e_n(z)}{I_n(z)} = c_n^2\, \frac{I_n(1/z)}{e_n(1/z)} &= \frac{(4+o(1))\, \rho^{4n}\, D(\mu)^2}{(2+o(1))\, D(\mu; 1/z)^2\, \varphi(1/z)^{-2n}\, \rho^{2n}\, D(\mu)} \\ &= (2+o(1))\, \rho^{2n}\, \varphi(z)^{-2n}\, D(\mu; z)^2\, D(\mu), \end{aligned} \tag{5.47}$$

where now the Landau symbol $o(1)$ holds locally uniformly in $\overline{\mathbb{C}}\setminus(\overline{\mathbb{D}} \cup [a, b]^{-1})$. With the limit (5.31) the estimates (5.44) and (5.47) imply that

$$\begin{aligned} e_n(z) &= (f - R_n^*)(z) \\ &= (1+o(1))\, \varphi(z)^{-2n}\, \rho^{2n}\, D(\mu; z)^2\, D(\mu)\, \frac{-2}{\sqrt{(z-a)(z-b)}} \end{aligned} \tag{5.48}$$

with $o(1)$ holding locally uniformly in $\overline{\mathbb{C}}\setminus([a, b] \cup [a, b]^{-1} \cup \mathbb{T})$. As

$$e_n(z) \Big/ \varphi(z)^{-2n}\, \rho^{2n}\, D(\mu; z)^2\, D(\mu)\, \frac{-2}{\sqrt{(z-a)(z-b)}}$$

is analytic across $\mathbb{T}$, (5.48) holds locally uniformly in $\overline{\mathbb{C}}\setminus([a, b] \cup [a, b]^{-1})$.

Recalling that the measure $d\omega_{\mathbb{T},[a,b]}(\zeta)$ given in (2.27) is of mass 1, we deduce that

$$\left\| \frac{-2}{\sqrt{(z-a)(z-b)}} \right\| = \sqrt{\frac{8K}{\pi(1-ab)}}.$$

The limit (2.44) in Theorem 7 follows. It is immediate that the estimate (2.45) in Theorem 8 also follows from (5.48).This completes the proof of both theorems. ▮

REFERENCES

1. T. Bagby, The modulus of a plane condenser, *J. Math. Mech.* **17** (1967), 315–329.
2. L. Baratchart, “Sur l’approximation rationelle L^2 pour les systèmes dynamiques linéaires,” Thèse de doctorat, Université de Nice, 1987.
3. L. Baratchart, Rational and meromorphic approximation in L^p of the circle: System-theoretic motivations, critical points and error rates, *in* “Computational Methods and Function Theory” (N. Papamichael, St. Ruscheweyh, and E. B. Saff, Eds.), pp. 45–78, World Scientific, Singapore, 1999.
4. L. Baratchart and F. Mandrea, “Non Hermitian Orthogonality and Poles of Rational Approximants,” INRIA research report 1999.
5. L. Baratchart, V. A. Prokhorov, and E. B. Saff, On meromorphic approximation to Markov functions, manuscript 1999.
6. L. Baratchart, E. B. Saff, and F. Wielonsky, A criterion of uniqueness of a critical point in H_2 rational approximation, *J. Anal. Math.* **70** (1996), 225–266.
7. L. Baratchart, H. Stahl, and F. Wielonsky, Non-uniqueness of rational best approximants, *J. Comp. Appl. Math.* **105** (1999), 141–154.
8. L. Baratchart, H. Stahl, and F. Wielonsky, Asymptotic uniqueness of best rational approximants of given degree to Markov functions in $\mathscr{L}^2$ of the circle, *Constr. Approx.*, in press.
9. L. Baratchart and F. Wielonsky, Rational approximation in the real Hardy space H_2 and Stieltjes integrals, a uniqueness theorem, *Constr. Approx.* **9** (1993), 1–21.
10. S. Basu and N. K. Bose, Matrix Stieltjes series and network models, *SIAM J. Math. Anal.* **14** (1983), 209–222.
11. B. de la Calle Ysern and G. López Lagomasino, Strong asymptotics of orthogonal polynomials with respect to varying measures and Hermite-Padé approximants, *J. Comp. Appl. Math.* **99** (1998), 91–103.
12. V. D. Erohin, On the best approximation of analytic functions by rational functions, *Dokl. Akad Nauk SSSR* **128** (1959), 29–32. [In Russian]
13. J. B. Garnett, “Bounded Analytic Functions,” Academic Press, San Diego, 1981.
14. G. M. Golusin, “Geomerische Funktionentheorie,” Deutscher Verlag der Wissenschaften, Berlin, 1957.
15. A. A. Gonchar, On the speed of rational approximation of some analytic functions, *Math. USSR Sb.* **34** (1978), 131–145.
16. A. A. Gonchar and G. López Lagomasino, On Markov’s theorem for multipoint Padé approximants, *Math. USSR Sb* **34** (1978), 449–459.

17. E. J. Hannan and M. Deistler, "The Statistical Theory of Linear Systems," Wiley, New York, 1988.
18. P. Henrici, "Applied and Computational Complex Analysis," Vol. I, Wiley Interscience, New York, 1977.
19. K. Hoffman, "Banach Spaces of Analytic Functions," Dover, New York, 1988.
20. N. S. Landkof, "Foundations of Modern Potential Theory," Grundlehren der Mathematischen Wissenschaften, Vol. 190, Springer-Verlag, Berlin, 1997.
21. A. L. Levin, The distribution of poles of rational functions of best approximation and related questions, *Math. USSR Sbornik* **9** (1969), 267–274.
22. A. L. Levin and E. B. Saff, Szegő type asymptotics for minimal Blaschke products, *in* "Progress in Approximation Theory" (A. A. Gonchar and E. B. Saff, Eds.), pp. 105–126, Springer-Verlag, Berlin/New York, 1992.
23. L. Ljung, "System Identification, Theory for the User," Prentice–Hall, Englewood Cliffs, NJ, 1987.
24. G. López Lagomasino, On the asymptotics of the ratio of orthonormal polynomials and the convergence of multipoint Padé approximants, *Math. USSR Sbornik* **56** (1987), 216–229.
25. G. López Lagomasino, Asymptotics of polynomials orthogonal with respect to varying measures, *Constr. Approx.* **5** (1989), 199–219.
26. N. I. Muskhelishvili, "Singular Integral Equations," Noordhoff International, Leyden, 1977.
27. Z. Nehari, "Conformal Mapping," McGraw–Hill, New York, 1952.
28. R. Ober, System theoretic aspects of completely symmetric systems, *Oper. Theory, Adv. Appl.* **87** (1996), 233–262.
29. O. G. Parfenov, Estimates of the singular numbers of a Carleson operator, *Mat. Sb.* **131**(173), (1986), 501–518; English transl. in *Math. USSR Sb* **59** (1988).
30. V. A. Prokhorov, Rational approximation of analytic functions, *Mat. Sb.* **184** (1993), 3–32; Engl. transl. in *Russian Acad. Sci. Sb. Math.* **78** (1994).
31. T. Ransford, "Potential Theory in the Complex Plane," London Math. Soc. Student Texts, Vol. 28, Cambridge Univ. Press, Cambridge, UK, 1995.
32. E. B. Saff and V. Totik, "Logarithmic Potentials with External Fields," Grundlehren der Mathematischen Wissenschaften, Vol. 316, Springer-Verlag, Berlin, 1997.
33. D. Sarason, The H^p spaces of an annulus, *Mem. Amer. Math. Soc.*, No. 56 (1965).
34. H. Stahl, Strong asymptotics for orthogonal polynomials with varying weights, *Acta Sci. Math. (Szeged)* **66** (2000), 147–192.
35. H. Stahl and V. Totik, "General Orthogonal Polynomials," Encycl. Math., Vol. 43, Cambridge University Press, Cambridge, UK, 1992.
36. S. Sternberg, "Lectures on Differential Geometry," Chelsea, New York, 1983.
37. G. Szegő, "Orthogonal Polynomials," Vol. 23, Amer. Math. Soc. Colloquium Publ., Providence, RI, 1975.
38. V. Totik, "Weighted Approximation with Varying Weights," Lecture Notes in Math., Vol. 1569, Springer-Verlag, Berlin, 1994.
39. J. L. Walsh, "Interpolation and Approximation by Rational Functions in the Complex Domain," Vol. 20, Amer. Math. Soc., Providence, RI, 1969.
40. J. C. Willems, Dissipative dynamical systems, Part I: general theory, Part II: linear systems with quadratic supply rates, *Arch. Rat. Mech. Anal.* **45** (1972), 321–351, 352–392.